TRAINING MANUAL ON WILDLIFE HEALTH INFORMATION MANAGEMENT

Fifth Cycle

Workshop for OIE National Focal Points for Wildlife

WORLD ORGANISATION FOR ANIMAL HEALTH
Protecting animals, preserving our future

TRAINING MANUAL ON WILDLIFE HEALTH INFORMATION MANAGEMENT

Fifth Cycle

Workshop for OIE National Focal Points for Wildlife

World Organisation for Animal Health
12, rue de Prony, 7501 Paris, France
Tel.: 33(0) 1 44 15 18 88
Fax: 33(0)1 42 67 09 87
http://www.oie.int/en/

Cover Photo: LeAnn White, USGS National Wildlife Health Center
François Diaz, World Organisation for Animal Health

Table of Contents

FOREWORD

OIE Collaborating Centre for Research, Diagnosis and Surveillance of Wildlife Pathogens

This workbook has been developed for the 5th Cycle Training for OIE Focal Points for Wildlife. During the trainings associated with the previous cycles, focal points were surveyed to determine the topics they were most interested in receiving future training. From that survey, it was evident that there was a need to provide training in wildlife health information management. Despites its practical importance, this topic is not well represented in the wildlife health literature, and our goal was to help to fill this gap with the material presented within this workbook.

The workbook is structured to follow a logical progression associated with wildlife health information management. We begin by detailing why it is important to share wildlife health information. We examine common challenges associated with information sharing, and explore potential opportunities to overcome these challenges. Next, we examine general principles to establish a wildlife health information network, and provide some case studies of real-world examples of successful networks. This is intended to address who should be included in a network, and how to establish the network structure for acquiring wildlife health information. We then progress to discussing what type of data should be collected, and describe some best practices for data management and curation. This section explores what to collect and share with regards to wildlife health information. Finally, we conclude with a discussion on data dissemination with a particular focus on using Geographic Information Systems (GIS) as a tool for disseminating wildlife health information. This last section describes how to share wildlife information, which is generally the ultimate goal of wildlife health information management.

We intend this workbook to provide the basis for developing a general understanding of wildlife health information management; however, this topic is broad in scope, and application of the presented concepts is often highly dependent on local conditions, challenges and governmental structures. Therefore, we present the material at a general level drawing upon the literature, expert opinion and personal experience. Throughout the material, we provide references and electronic links to additional resources for readers who desire to develop a more in-depth understanding of the topics presented.

Wildlife health information management is becoming increasingly important component of surveillance and management of wildlife diseases. We hope this workbook will help focal points in maintaining and/or improving the health of wildlife in their countries for the benefit of wildlife, domestic animal and human health.

Opportunities and Challenges of Sharing Wildlife Health Data

The Case for Sharing Wildlife Health Information

Surveillance for wild animal pathogens is a critical component of a national wildlife health program. Surveillance is defined as the systematic ongoing collection, collation, and analysis of information related to animal health and the timely dissemination of information to those who need to know, so action can be taken (OIE *Terrestrial Animal Health Code*). As described in the OIE 2nd and 4th Cycle Workshops for OIE National Focal Points for Wildlife, surveillance is essential to all of the other components of a program (prevention, detection, risk analysis, and management). It is through wildlife pathogen surveillance that a country identifies what pathogens are present in its wild animal populations, and in which geographic areas and host species they occur. Surveillance is the primary mechanism for detecting new, emerging diseases, and is also used to measure the proportion of animals in a population that are infected (the prevalence of disease within the population). All of this information is required to assess health risks associated with international trade or internal movement of wild animals, to meet international obligations for disease reporting, to develop effective management actions, and maintain wildlife health.

Integral to the concept of effective wildlife health surveillance is that the information needed for decision making and action is made accessible to the person(s) responsible for taking action. Indeed, the need for rapid exchange of information about animal diseases was a major objective for the establishment of the World Organisation for Animal Health (OIE), and the deployment of the OIE's World Animal Health Information System (WAHIS) to facilitate the transfer of information among member countries about diseases with the potential to impact international trade (OIE 2018). Submission of information on notifiable terrestrial and aquatic animal diseases to OIE is among the formal obligations of OIE member countries. However, there are many other diseases affecting wildlife that have not met the OIE criteria for listing. The OIE Working Group on Wildlife Diseases selected 53 of these infectious and non-infectious diseases to be monitored because of their importance to wildlife, to maintain biodiversity, and to serve as early warning to protect animal and human health. Reporting on these diseases is performed by OIE Delegates and the nominated focal points through a complimentary reporting system known as WAHIS-Wild. Although the reporting of non-listed diseases is voluntary it is highly encouraged as sharing this information through WAHIS-Wild improves understanding of the wildlife health situation worldwide.

In order for wildlife health information to be provided to WAHIS-Wild it must first be collected from individuals at the local, regional, and national levels. The focus of this workshop is to provide strategies for collecting, curating, and sharing wildlife health information at multiple scales. Finding ways to share information is necessary since even within a single country, generally no individual agency has the resources or, in some cases, the authority, to perform primary data collection on the health of wild animals. Since data must be shared to get broad coverage, agencies and individuals must work together to determine which data fields are essential and when the information should be shared for timely decision-making. By learning to work together in this way, entities at multiple scales can combine efforts to form a collaborative global health surveillance network. The World Health Organization and OIE both function under the premise that combining health datasets across geographic barriers is the most efficient mechanism for detecting epidemics and monitoring the status and trends of diseases in humans and animals throughout the world.

The value of surveillance is realized through effective and efficient delivery of useful information (Hall *et al.* 2012). Delays in sharing information can result in lost opportunities for interventions. Failure to share information resulting in incomplete estimates of the magnitude of a health issue can also affect allocation of resources and decisions about research priorities and disease control. Therefore, in the next sections we discuss some of the major challenges associated with sharing wildlife health information as well as potential solutions for overcoming these challenges. The preponderance of published information on the topic of scientific data sharing currently focuses on academia. Nevertheless, many of the issues at the heart of these examples are universal to data sharing and can be applied to the wildlife health field to improve information exchange within and among countries. *NOTE:* The term **"data steward"** used throughout this section refers to anyone (an individual, group of individuals, or agency) that has data that could potentially be shared.

The Challenge of Sharing Wildlife Health Information

From the standpoint of individual scientists, if data are well organized, documented, preserved, verified, and accessible, they can be more readily analyzed, used, and shared. Sharing data benefits individuals through increased transparency and opportunities for scientific collaboration, discovery, and innovation (van Panhuis *et al.* 2014). Nevertheless, data sharing has numerous challenges.

> "*Despite a global commitment to the use and sharing of public health data, this can be challenging in reality. No systematic framework or global operational guidelines have been created for data sharing in public health. Barriers at different levels have limited data sharing but have only been anecdotally discussed or in the context of specific case studies"* (van Panhuis *et al.* 2014).

Furthermore, even though most scientists would acknowledge the benefits of sharing information from the standpoint of the wildlife health community, when it comes to sharing their own data there are frequently delays in sharing information, or information may not be shared at all.

> *"Researchers are often acting in ways that make sense for them, that is, not to share data. For example, they don't waste precious research time on making data sharable, and they don't risk getting scooped by exposing valuable data too early. Although everyone is acting sensibly according to their own interests, the collective result is that far less data is shared than everyone agrees would be optimal"* (Van den Eynden and Bishop 2014).

Legal, regulatory, and ethical limitations on data sharing

Unless the disease has agricultural or human health importance, there are often few formal rules, regulations, and legislative authorities governing the sharing of wildlife health information, which results in the decision to share being voluntary. For example, in Australia there is legal requirement that any organization that suspects or diagnoses a notifiable disease in wildlife to immediately report it to the government authority. However, Wildlife Health Australia reports that it more challenging to obtain accurate data on non-notifiable diseases, as there are no mandates for reporting.

Some countries may have legal restrictions that do not allow some types of wildlife health information to be shared. Often these restrictions are caused by a fear that releasing the information could have potential trade or other economic implications. There may also be reluctance to share data due to fear that information will be released to the public before the agency responsible for managing the disease has been given an opportunity to begin intervention strategies or biocontrol. This could result in delays in interventions as agencies divert resources (e.g., personnel time) to responding to public and press inquiries.

Some data stewards may be reluctant to share data due to concerns that they will be unable to ensure that the data is fully understood by external users, leading to potential misuse or misrepresentation of the data. For example, data regarding on-going and historical wildlife disease events occurring in the United States of America is publicly served on USGS-National Wildlife Health Center's (NWHC) Wildlife Health Information Sharing Partnership event reporting system (WHISPers). While these data can be used to monitor status and trends of wildlife disease in the United States, the data are passively collected and voluntarily submitted to NWHC. The absence of data records for some states may arise because there are fewer disease events in those states *or* that those states have not reported them to NWHC (i.e., reporting bias). This detail becomes important when scientists attempt to draw conclusions not supported by the data. For example, if the data are being used to follow the spread of a pathogen across the country, there may be a large gap between two detections of a pathogen that exceeds the distance moved by hosts. If the person or agency conducting the analysis does not understand that there is a gap in detections due to reporting, they may conclude that the jump in transmission must be due to mechanical spread rather than spread by the hosts themselves.

Data stewards may also be unwilling to share for personal or political reasons. They may be afraid that the importance of their program may be diminished if someone else has access to their program's data (Bernstein and Sweeney 2012). They may also be afraid that a better-funded entity may be able to more quickly analyze and publish the data, thereby taking credit for the research idea that led to the data collection. In cases when the wildlife health information does not pertain to agriculture or human health, the specifics of what, when, and how to share are largely left up to individual scientists. Regardless of how often data use without proper acknowledgement occurs, if scientists are concerned about this possibility, then it is a real impediment to data sharing and acknowledging it will help ensure that the concern is properly addressed.

Administrative barriers

Some agencies may not be bound by laws that regulate wildlife health information sharing, but may not have a culture of sharing because they are concerned that releasing a finding (e.g., the detection highly pathogenic avian influenza in wild birds) could have trade implications. The lack of guidance on data sharing and dissemination can also be an impediment. In some cases groups may be unaware of what data is relevant to others and the necessary timeframe in which to share that data for effective wildlife health interventions. In other cases, individuals may be willing to share but need guidance on what level of information their own agency deems acceptable for sharing. For example, in a 2018 survey of partners conducted by the Canadian Wildlife Health Cooperative, individuals expressed their personal desire to share wildlife health information but noted that major impediments that prevented sharing were was institutional or agency policies that discourage it or the lack the of clear guidance on what was permissible.

Competition for limited resources can also be seen as an administrative barrier as the resources required to manipulate, code, and transmit data for sharing (Bernstein and Sweeney 2012) may be needed for other priorities established by local, state, or regional government entities. Given that resources are often more restricted in the wildlife health field than in the agricultural or public health fields, and that resources vary widely among regions and countries, sharing data about diseases that only affect wildlife is not always viewed as a top priority. As described below, data need to be thoroughly described and curated so that they can be interpreted correctly by users, which is also resource-intensive. Performing these tasks can also affect the time between disease detection and data release which can in turn affect the value of information if a rapid response is needed (Bernstein and Sweeney 2012).

Finally, the appropriate infrastructure, such as a centralized data repository to facilitate data sharing, may not be available. In a system where each agency collects, interprets, and uses their own data, it may be unclear: who else is interested in the data, who else is collecting similar or complementary data, and how these parties can work together to share data in a time- and cost-effective way.

Technical barriers

The biggest technical challenge facing health information sharing is "finding efficient and effective ways of combining multiple sources of complex data and information into meaningful and actionable knowledge" (Savel and Foldy 2012). For example, some Wildlife Health Australia partners have their own databases and may not use the national wildlife health surveillance system for primary data entry. Therefore, to provide data to the national system requires duplication of effort entering the data in both places.

Data incompatibilities are also problematic, and can be caused by differences in hardware and software, coding, formatting, terminology definitions, storage formats (Bernstein and Sweeney 2012), or language barriers. Technical barriers are a prominent issue for wildlife health as many commonly used data fields are not standardized and diagnostic test procedures are not uniform. In some cases, important data may not have been collected (e.g., total number of animals affected) or, if data are collected for short-term purposes such as outbreak detection, data preservation and archiving may not have been prioritized (van Panhuis *et al.* 2014). Data stewards may also lack the technical skills and resources for curating large datasets that are needed for locating, retrieving, and disseminating information.

Some technical barriers may have straightforward solutions, but time and creativity are required to make the combined datasets useful. For example, there is no standard coordinate reference system for recording locations so units must frequently be converted. Another common inconsistency is the use of both text and numbers to code the same field (e.g., lesion severity may be classified as mild, moderate, and severe or 1, 2, and 3). In some cases, standardizing terminology may require more time-intensive solutions. For example, an agency may define a wildlife disease "event" as all of the mortality caused by a disease within a given time period regardless of location (e.g., all parks in a country with the same disease). Another agency may define an "event" as having specific geographic boundaries (e.g., each park is considered a single event). A third agency may label each submission of samples to a laboratory as an individual "event". This type of data inconsistency may be inconsequential if the purpose of combining data is to see geographical extent of disease; however, if the events reported by these three agencies occurred over three separate years and the intention was to examine the size of the events over time, the conclusion would be inaccurate. Still more complicated is the categorization of disease fields as test methods may vary and case definitions do not exist for most diseases that primarily affect wildlife. For example, one agency may classify a frog as having chytridiomycosis based on histopathological findings. Another may require molecular tests in addition to histopathology before diagnosing an individual with chytridiomycosis. Still another may consider an animal to have the disease based on epidemiological ties to an outbreak. These different classifications would make it difficult to perform even simple summaries such as the number of frogs with chytridiomycosis.

Overcoming Barriers to Wildlife Health Data Sharing

Recent technological advances have allowed large quantities of data to be recorded, stored, and disseminated electronically at a rate that was not previously possible with paper or physical samples. While many of the published discussions on barriers to sharing data have focused on how to increase data sharing among agencies and academic institutions, many of these data sharing challenges are universal. Therefore, once a barrier in wildlife health data sharing has been identified, the techniques outlined below may be useful for increasing sharing of wildlife health information within and among countries.

Overcoming legal, regulatory, and ethical limitations to wildlife health data sharing

Understanding the specific obstacles or concerns of data stewards is important for developing the right strategy to overcome barriers for data sharing. For example, no amount of incentives will be sufficient if the steward does not have control over the data or the legal authority to share is unclear. In these cases it may be necessary to develop transparent legal agreements. In many cases though, non-binding documents such as memorandums of understanding or letters of intent that include language to allow release of the data to specific partners while preventing public sharing will be sufficient. If some but not all aspects of the data are restricted, data sharing agreements may include de-identifying some data fields (e.g., names of farms, geographic locations, specific dates), aggregating data (e.g., to larger geographic location), or applying time delays for release to other groups or the public.

Outlining data use policies can help overcome some of the barriers associated with partner concerns about privacy, data ownership, and how collected data will be used. A thorough example of this type of policy can be found on Wildlife Health Australia (WHA) Data Management documentation located here:

https://wildlifehealthaustralia.com.au/DataManagementWildlifeDiseaseSurveillance.aspx.

In Australia, ownership of wildlife health information provided to WHA remains with the submitter, confidentiality of information is controlled by the data submitter, and the release of information by WHA is primarily summary or aggregated data that includes confidentiality statements where appropriate. WHA also assists with reporting barriers in their country by having the recognition and authority of their government to administer Australia's national wildlife health information system. By maintaining the national system they can ensure that proper protocols for reporting are followed so that information provided by partners on notifiable diseases events is communicated to relevant government authorities before sharing the data.

If the data steward is concerned that data may be misused or misinterpreted, the data requester could provide explicit explanations describing how the data will be used. The data steward can then assess whether the data is appropriate for that use. This may alleviate concerns and reduce the amount of time that the data steward might spend detailing where it might not be appropriate to use their dataset. Another proposed solution to increase data sharing while recognizing that data misinterpretation is possible, is the creation of a standardized citation designation for "data authors" (Bierer *et al.* 2017) who would only be held responsible for integrity of the data set, not the analyses or associated conclusions made from subsequent use of the data. Alternatively data stewards may want to consider providing a waiver or other documentation describing the data's "fitness of purpose" if they are concerned about interpretation issues. For example, to assist with secondary data use quality control in Canada, the Canadian Wildlife Health Cooperative reserves the right to review final documents with conclusions based on their data.

Overcoming administrative barriers

Data sharing requires time and resources that are chronically lacking in the majority of wildlife programs. Therefore, one of the most common administrative barriers is that data sharing has not been made a priority, and there are no dedicated resources for data sharing (e.g., gathering, cleaning, and transforming data). In these cases, beginning requests to share data or to request others' data with a formal recognition of the time commitment that this task requires can be useful. It is also crucial to communicate the practical value of shared data. For example, both the National Wildlife Health Center (NWHC) in the United States of America and the Canadian Wildlife Health Cooperative (CWHC) in Canada serve as a repository for wildlife health information in their respective countries. Each organization directly generates some portion of their country's wildlife health information through investigation of wildlife disease events. NWHC employs diagnostic staff that work directly for the Center and CWHC utilizes a network of university-based diagnostic staff to conduct investigations. But both organizations also collect and manage wildlife health data provided by other partners throughout their countries. The primary incentive that they have used when soliciting this data is active communication with an emphasis on the value of centralized wildlife health data for providing their countries with situational awareness and a national perspective on wildlife health, which is required for timely and accurate decision-making. CWHC partners also reportedly value the centralized communications, reporting, and data management capabilities available through CWHC. By serving as a central repositories for Canadian wildlife health information, CWHC is also able to recover some of their costs through fee for service data requests from industry or other stakeholders needing access to specific sets of data.

Wildlife Health Australia provides numerous mechanisms to overcome administrative barriers. They allow data contributors to provide regular feedback on how Australia's national wildlife health information system can be improved. They also provide a forum for pro-actively sharing information through regular teleconferences for partners to share and discuss wildlife health events being investigated around the country. They also report that partners contribute data because they have been made aware of the national benefits of a centralized wildlife health information system and want to contribute because they sharing as mutually beneficial.

NWHC's development and hosting of WHISPers, the wildlife disease reporting system described above, is an example of the successful use of novel infrastructure to make data sharing and reporting easier. Ultimately, WHISPers is envisioned as a centralized curated repository of wildlife health data within the United States with customizable levels of data accessibility. Furthermore, it is a way for stewards, partners and stakeholders to connect and work together to optimize use of resources and maximize the utility of captured data.

While the benefits to the wildlife health community or society as a whole may be sufficient to convince some data stewards to provide data, others may require more direct benefits or additional incentives. Another way to incentivize partners to share data is to communicate the potential for increased visibility of the data collector or their program. The increased visibility of shared data has been demonstrated in academia through findings that studies with shared data were more often cited in published scientific literature (Piwowar *et al.* 2007, Piwowar and Vision 2013). The shared data can be mutually beneficial to both the data steward and the data requestor. For example, partners contributing data to CWHC's data repository leverage the national scope of the CWHC partner network and their communications to enhance the profile of their individual programs. CWHC, in turn, uses the shared information to provide a detailed national perspective on wildlife health issues through reports and other communication tools.

Shared data can also promote new collaborations and the ability to pursue analyses that were not possible with an individual dataset. For example, researchers examining factors contributing to the spread of amphibian chytridiomycosis may be able to combine datasets to understand factors occurring on a regional or national scale. Similarly, researchers may find collaborators from fields other than their own, allowing for analyses or perspectives the initial researcher would not have considered. Scientists benefitting from shared data can also be strong advocates for recruiting new partners in the future.

However, it also important to note that despite a shift towards open data, increased data sharing, and many opinion articles on strategies for addressing barriers, there are few evidence-based reports demonstrating the relative effectiveness of various techniques (Rowhani-Farid *et al.* 2017). One of the few tested incentives is the use of open science "badges" which are awarded to authors who voluntarily make all digitally shareable data and materials available in an open repository. These symbols are then linked to publications, typically in one of the two formats below (from *https://cos.io/our-services/open-science-badges-details/*):

Variation One:

Large, color badges are presented with accompanying text and hyperlinks as a Figure:

Figure 1. Badges earned for open practices: Open Data, Open Materials, and Preregistered. Experiment materials and data are available at http://dvn.iq.harvard.edu/dvn/study?globalId=hdl:1902.1/18708, and the preregistered design and analysis plan is accessible at http://opensciencframework.org/project/TVyXZ/

Variation Two:

Small, gray-scale badges are presented in the article header and accompanying text and hyperlinks appear in a footer at the bottom of the first page:

PERSPECTIVES ON SOME KIND OF SCIENCE

Research Report

Writing a Sample Article

And Writing Other Things Too

Alfred R. Wallace[1] and Charles R. Darwin[2]

[1]Collegiate School, Leicester, England, and [2]University of Edinburgh Medical School, Edinburgh, Scotland

Although these badges may seem trivial, they provide the incentive of public recognition, and are one of the few evidence-based practices shown to increase data sharing rates (Kidwell *et al.* 2016, Rowhani-Farid *et al.* 2017). Development of badges or some other standardized way to visually provide recognition of a partners contributions could be relatively easily developed within the wildlife health community. These symbols could be displayed along with partner emblems on regional or national websites, reports, or publications to recognize partners that regularly contribute data and information to the wildlife health community.

Many of the techniques described above are directed at increasing sharing of information that has already been collected. If you need to find data or encourage data collection, developing strong partner networks is a key mechanism for locating, accessing, and soliciting wildlife health information. Partner network formation takes time and is covered in detail in a separate section of this workbook. However, it is worth noting here that incentives can also be used to encourage data collection as well as sharing. Communicating with partners and the public regarding the value of the data can be achieved by a number of mechanisms including newsletters, social media, or a centralized website that summarizes the status of the project, ways the data has been used, and measurable outcomes. Gamification, the use of game elements such as leaderboards or prizes (Deterding *et al.* 2011), has been successfully employed in large-scale citizen science monitoring efforts (Fritz *et al.* 2017). Data collectors may also appreciate opportunities to improve their data collection skills through tutorials or workshops. The use of training as an incentive has the added benefit to the data requester of improving the quality of the data. In a study examining the use of crowdsourced data for remote sensing and visual interpretation, the authors found that the quality of data provided by non-experts increased over time. Tracking changes in data quality can be used to target training materials to those areas where increase accuracy is most needed as well to determine the effectiveness of the training materials.

A final technique to overcome administrative barriers is monetary incentives. Wildlife Health Australia provides some support for administrative work required to contribute data to their national wildlife surveillance system. Partners are also able to apply for funding to support significant disease investigations.

Overcoming technical barriers

Solutions to technical challenges will be necessary throughout the data sharing process. Well-defined data can be more efficiently and effectively shared and the availability of metadata, or data about the data, can significantly increasing the usefulness of a dataset. Information described in metadata often includes where the data was collected, why the data was collected, and, importantly, who is responsible for the dataset in case there are questions about the data. Metadata often follow a standard format and details about how to create effective metadata are provided in a later section of this workbook.

The creation of definitions for essential data fields can assist the data steward and data requester in combining datasets. If these definitions can be established prior to the collection of data it can significantly increase the efficiency of data sharing as it minimizes the need to translate from one format or definition to another. For example, the development of widely-accepted case definitions for wildlife diseases would significantly increase the ability to share information on common or emerging diseases of interest. Case definitions typically include field, gross, histopathologic, laboratory, and epidemiologic criteria for assigning an individual to a specific disease category. Where one exists, the case definition in the specific chapter of the OIE Terrestrial or Aquatic Code should be used (see *http://www.oie.int/en/international-standard-setting/terrestrial-code/access-online/*). Although case definitions for diseases that are restricted to wildlife are less common than for those that can also affect humans or domestic animals, they have been developed for several important wildlife diseases (see White *et al.* 2016, NWHC 2014).

The creation of case definitions for all wildlife diseases would ensure that one of the most important fields being shared among partners, the diagnosis, was being consistently applied. In fact, in a recent workshop assessing the essential functions of national wildlife health programs that involved 11 countries (Australia, Canada, China, the Netherlands, New Zealand, Republic of Korea, Sweden, Switzerland, Thailand, Great Britain, and the United States) participants expressed a need for standardized case definitions specific to wildlife health and disease (Nguyen *et al.* 2017). These definitions could be made broadly available via publication in journals or through a specific organization charged with curating them. For example, the Centers for Disease Control maintains case definitions for human diseases tracked in the United States through their National Notifiable Diseases Surveillance System website (*https://wwwn.cdc.gov/nndss/conditions/*). Similarly, SNOMED CT (Systematized Nomenclature of Medicine-Clinical Terms) is a standardized set of clinical health terminology use to support the effective exchange of human health information worldwide. The terms are maintained by SNOMED International, an organization that develops and maintains the terminology, thereby ensuring consistency among all user groups. Although there is not yet a wildlife extension for SNOMED CT, there is a veterinary extension that includes codes for use in agriculture and domestic animals (*http://vtsl.vetmed.vt.edu/*).

Data requesters can also increase the likelihood that data stewards will comply with voluntary data requests by requesting only the information required for the task they are performing. Minimizing requested information reduces the burden on the data steward for data entry or data manipulation required to transmit and combine the data with other datasets. Minimizing requests to preselected or commonly used fields also minimizes misinterpretation of data. User-friendly data entry platforms and options to bulk upload data through APIs (application programming interface) on the user interface of large shared databases will also likely improve compliance with requests for data.

When data are highly diverse and combining datasets is not feasible it may be worthwhile to maintain a central repository for data on a particular topic. Canadian Wildlife Health Cooperative utilizes this technique for white-nose syndrome (WNS), which affects North American bat species, whereby they maintain a repository of which individuals and organizations are collecting information, and what type of data they collect so they can facilitate data users and data stewards working on this disease.

Assessing Country-Specific Barriers and Solutions for Data Sharing

Although the topic of data sharing in the wildlife health field is not new, it has received renewed attention due to the widespread availability of technology that allows for large quantities of data to be recorded, stored, and disseminated electronically. The information we have provided is intended as an overview of the potential challenges and solutions that others have proposed or found useful for overcoming barriers to data sharing. An important note from Wildlife Health Australia's experience in building a national wildlife health information system is that building trust your organization can appropriately manage sensitive and confidential data is crucial, but takes time, often years. The questions below can be used to develop an understanding of the particular barriers and potential solutions for your region, country, agency, or partners.

1. Which partners in your country collect wildlife health data? What types of information does each partner collect? Are there any obvious gaps in wildlife health data collections (i.e., data needed for decisions not currently being collected)?

   ....

   ....

   ....

2. Assuming you can obtain all of the data you need from partners, how will it be used (situational awareness, generation of regional/national summaries on wildlife health, publication, management actions, etc)? Generate a list of how the data will be used and, if possible, link it with who is collecting the data so that you can directly link partner data with products or outcomes during communications with them.

   ....

   ....

   ....

3. Are there any incentives for sharing data with you that partners should be aware of (e.g., complementary resources or services, data management support, access to analytical tools, or standardized terminology)?

   ....

   ....

   ....

4. What are the specific barriers your partners have to sharing data? *Comments:* Determining barriers to data sharing can accomplished through informal or formal surveys (phone, in-person, email) of partners. Do not assume all partners have the same barriers to data sharing. Be aware that some partners may have more than one barrier to data sharing.

   ....

   ....

   ....

5. Which of the barriers could you currently address? Are there areas where a compromise can be reached (e.g., skewing exact location information for privacy)? Are there common barriers among key partners that could be addressed first? Are there barriers that may require additional resources (e.g., technology development such as bulk data upload)? If you are having trouble addressing a particular barrier, consider contacting other OIE National Focal Points for Wildlife or Collaborating Centres to see if they have found solutions for similar problems.

   ....

   ....

   ....

Literature cited

Bierer BE, Merce MD, Pierce HH. 2017. Data authorship as an incentive to data sharing. New England Journal of Medicine. 376: 1684-1687.DOI: 10.1056/NEJMsb1616595

Deterding S, Sicart M, Nacke, L, O'Hara K, Dixon D. 2011. Gamification. Using game-design elements in non-gaming contexts. CHI Extended Abstracts on Human Factors in Computing Systems 2425–2428. DOI: 10.1145/1979742.1979575

Fritz S, See L, Perger C, McCallum I, Schill C, Schepaschenko D, Duerauer M, Karner M, Dresel C, Laso-Baya JC, Lesiv M, Moorthy I, Salik CF, Danylo O., Sturn T, Albrecht F, You L, Kraxner F, Obersteiner M. 2017. A global dataset of crowdsourced land cover and land use reference data. Scientific Data 4: 170075. doi:10.1038/sdata.2017.75

Hall HI, Correa A, Yoon PW, Braden CR. 2012. Lexicon, definitions, and conceptual framework for public health surveillance. Morbidity and Mortality Weekly Report 61: 10-14.

Kidwell MC, Lazarević LB, Baranski E, Hardwicke TE, Piechowski S, Falkenberg L-S,Kennett C, Agnieszka S, Sonnleitner C, Hess-Holden C, Errington TM, Fiedler S, Nosek BA. 2016. Badges to Acknowledge Open Practices: A Simple, Low-Cost, Effective Method for Increasing Transparency. PLoS Biol 14(5): e1002456*. https://doi.org/10.1371/journal.pbio.1002456*

National Wildlife Health Center (NWHC). 2014. Diagnostic Categories for Reporting Cases of Bat White-Nose Syndrome (WNS) Including a Summary of revisions to WNS case definitions for the 2014/2015 season. Available at *https://www.nwhc.usgs.gov/disease_information/white-nose_syndrome/Case%20Defintions%20for%20WNS.pdf* Accessed 15 March 2018

Nguyen NT, Duff JP, Gavier-Widén D, Grillo T, He H, Lee H, Ratanakorn P, Rijks JM, Ryser-Degiorgis M-P, Sleeman JM, Stephen C, Tana T, Uhart M, Zimmer P. 2017. Report of the workshop on evidence-based design of national wildlife health programs: U.S. Geological Survey Open-File Report 2017–1038, 18 pp. *https://doi.org/10.3133/ofr20171038.*

OIE (World Organisation for Animal Health). 2018. Animal health in the world - Overview. http://www.oie.int/en/animal-health-in-the-world/oie-listed-diseases-2017/. Accessed May 2018.

Piwowar HA, Day RS, Fridsma DB. 2007. Sharing Detailed Research Data Is Associated with Increased Citation Rate. PLoS ONE 2(3): e308. *https://doi.org/10.1371/journal.pone.0000308*

Piwowar HA, Vision TJ. 2013. Data reuse and the open data citation advantage. PeerJ 1:e175 *http://dx.doi.org/10.7717/peerj.175*

Rowhani-Farid A, Allen M, Barnett AG. 2017. What incentives increase data sharing in health and medical research? A systematic review. Research Integrity and Peer Review 2:4. *https://doi.org/10.1186/s41073-017-0028-9*

Savel TG, Foldy S. 2012. Role of public health informatics in enhancing public health surveillance. Morbidity and Mortality Weekly Report Supplements 61:20-24.

Van den Eynden V, Bishop L. 2014. Incentives and motivations for sharing research data, a researcher's perspective. 48 pp. *http://www.data-archive.ac.uk/media/492924/ke_report-incentives-for-sharing-research-data.pdf knowledge-exchange.info/Default.aspx?ID=733*. Accessed March 2018.

Van Panhuis W G, Paul P, Emerson C, Grefenstette J, Wilder R, Herbst A J, Heymann D, Burke D S. 2014. A systematic review of barriers to data sharing in public health. BMC Public Health 14:1–9.

White, CL, Forzan, MJ, Pessier, AP, Allender, MC, Ballard, JR, Catenazzi, A, Fenton, H, Martel, A, Pasmans, F, Miller, DL, Ossiboff, RJ, Richgels, KLD, Kerby. JL. 2016. Amphibian: A case definition and diagnostic criteria for *Batrachochytrium salamandrivorans* chytridiomycosis. Herpetological Review. 47: 207-209.

Key Components Required to Establish and Maintain Partner and Stakeholder Networks

Successful wildlife health networks have the following characteristics:

1. *Customized.* There is no one size fits all approach to building and maintaining partner networks, instead network approaches and structures should be tailored to address specific needs.

2. *Right people*. The success of the network relies on getting the correct people - those with access and permission to share (or use) wildlife health data - to actively participate in the network.

3. *Partner buy-in.* When building a network, make sure to include partners in the planning and implementation phases to ensure support and cooperation from those critical to the networks success.

4. *Mutually beneficial.* Networks that have been successfully maintained ensure transparency in decisions and communications, equity in the way partners are treated, and provide a mutual benefit for all parties involved.

Introduction

This section will cover general components required to establish and maintain a partner and/or stakeholder network, followed by specific case examples of successful wildlife health networks. Given the significant differences among and within nations across the globe regarding the specific pieces and steps required to create a partner network, a definitive guide on the subject is not possible. However, we will provide some general ideas and suggestions that are widely applicable and may help to establish successful wildlife health networks.

We begin by defining a *network* as a communications arrangement linking people who are engaged in similar activities. A *wildlife health partner network* establishes communication guidelines for stakeholders and partners engaged in wildlife health. For example, a network, established to address diseases that have major impacts on all three health sectors, may create communication plans and oversight committees for interagency coordination of public health, agricultural health, and wildlife health.

For the purpose of this chapter, a *stakeholder* is defined as a person or entity that has an interest or concern with regards to wildlife health. A *partner* is defined as a person or entity who actively participates in the regulatory, management, identification, or social aspects of wildlife health. In general, partners are also stakeholders, but many stakeholders are not necessarily considered partners. For example, a wildlife manager, collaborating with a diagnostic or research laboratory to better understand patterns of disease outbreaks in wildlife that he/she actively manages, would be considered a *partner* in the disease outbreak investigation. In contrast, if a major disease outbreak occurs in wild birds at a national park, individuals inhabiting neighborhoods that border the park would not typically be considered partners unless citizen science is being used to help identify the outbreak (see Case Example 1).

Networking is often taught as a foundational principle for persons working in the business field (see *https://hbr.org/topic/networking* for introductory materials). In brief, *business networking* is the practice of purposefully connecting with other professionals in your field to both expand your social network and develop mutually beneficial business relationships. It is often described as a good method for expanding a business, finding financial support for innovative ideas, and for finding job opportunities. All of these outcomes are thought to increase success and professional development of the person networking. This type of networking is often aimed at meeting the professional goals of an individual, but could also be exploited to form an informal wildlife health partner network.

Beyond *business networking* is the practice of developing *partnerships* - that is bringing together multiple organizations from different sectors to work together to complete a common objective (Partnering Toolbook, 2003). A wildlife health partner network is, by definition, a partnership. Thus, the majority of this chapter will focus on the practice and principles involved in developing partnerships, which are directly applicable to building and maintaining a partner network for wildlife health. Partner networks in wildlife health can have a variety of structures depending on the goals and objectives of the wildlife health partner network. For example, they can vary in scale from regional, national, or international; in scope by being disease specific, sector specific (such as public health, agricultural health), or covering One Health; and in types of structure from informal to formal (Table 1). *Table 1* provides examples of wildlife health partner networks from around the globe. Later we will describe different types of structures, and the advantages and disadvantages of each as you progress from informal to formal partner networks (Partnering Toolbook, 2003).

Partner networks can be formal or informal. For example, a person may meet others who work in the wildlife health field at a meeting and maintain relationships with these people as a way to share information about general wildlife health - this is an example of an informal wildlife health partner network, established and maintained through business networking. Informal networks are also often used during wildlife disease events to gather information about the event and provide situational awareness to partners and stakeholders. For reportable diseases with zoonotic or agricultural implications, the network of partners may be more structured and include a steering committee or working group with specific guidelines on the roles and responsibilities of partners and the frequency and type of communications. This would exemplify a semi-formal partner network. A formal partner network in wildlife health is often developed by agencies or non-profits that are supported by legislation and/or sufficient resources. For example, the U.S. Department of Agriculture, which has a legal mandate to protect agricultural resources, runs the National Agricultural Health Laboratory Network (NAHLN) that partners with diagnostic laboratories around the United States to ensure standardized diagnostics for notifiable diseases. NAHLN laboratories receive funding to test for these diseases and maintain their certification. The formal structure of NAHLN allows the organization to provide aid, through personnel or funding, to partners within the network. For more in-depth examples of different network structures, see the case examples at the end of this chapter.

Table 1: Examples of different partner networks around the world, categorized by their scale, scope, and structure.

Scale	Scope	Structure	Lead Organisation	Network Name	Link
Regional	Disease specific	Informal	USGS National Wildlife Health Center	Avian Monitoring for Botulism Lakeshore Events (AMBLE)	https://www.nwhc.usgs.gov/our_research/amble/
National	Wildlife Health	Informal	USGS National Wildlife Health Center	Wildlife Health Information Sharing Partnership - event reporting system (WHISPers)	https://www.nwhc.usgs.gov/whispers/
National	Disease specific	Semi-formal	U.S. Department of Agriculture	Interagency Avian Influenza Steering Committee on Surveillance in Wild Birds	https://www.aphis.usda.gov/aphis/newsroom/news/sa_by_date/sa_2015/sa_july/migratory-birds
International	Animal Health	Semi-formal	World Organisation for Animal Health	Collaborating Centres and Reference Laboratory Network	http://www.oie.int/en/scientific-expertise/overview/
International	Wildlife Health	Semi-formal	European Wildlife Disease Association	EWDA Network for Wildlife Health Surveillance in Europe	http://ewda.org/ewda-network/
National	Wildlife Health	Semi-formal	Canadian Wildlife Health Cooperative	Canadian Wildlife Health Cooperative	http://www.cwhc-rcsf.ca/who_we_are.php
National	Wildlife Health	Semi-formal	Wildlife Rehabiltiators Network of New Zealand	Wildlife Rehabiltiator's Network of New Zealand	https://www.wrennz.org.nz/
National	Wildlife Health	Formal	Wildlife Health Australia	Australia Wildlife Health Network	https://www.wildlifehealthaustralia.com.au/AboutUs.aspx
National	Animal Health	Formal	U.S. Department of Agriculture	National Animal Health Laboratory Network	https://www.aphis.usda.gov/aphis/ourfocus/animalhealth/lab-info-services/nahln/ct_national_animal_health_laboratory_network

How Do You Build a Partner Network?

The first step in building a partner network is to identify stakeholders and partners that will be important contributors to the planning process and resulting network. Tied closely to this task, is defining the need that this network is addressing. Often the impetus for starting a wildlife health network is to gather and collate information collected independently by local groups across a large area to improve situational awareness of diseases affecting humans, agriculture, or wildlife. Additional reasons for initiating networks may include exchanging knowledge and building technical capacity that follows standards set by the lead organization. One of the main benefits of building partner networks is that diagnostic results and other response measures are implemented by a trusted partner. This is especially true if partners in the network provide surge capacity during disease outbreaks. For example, OIE Reference Laboratories or Collaborating Centres are designated after careful review of their qualifications and quality. These laboratories and centres can be valuable components of a network, and often can be leveraged during critical situations by providing reliable diagnostic tests, expert advice to responding agencies, and/or providing trained staff to aide in response efforts. Regardless of the purpose of the network, it is critical that the purpose be explicitly defined because it will be the foundation upon which the network will be built. Once the purpose of the network is defined, then the lead organization can start to identify partners whose participation is important for the success of the network. Important questions to consider when defining partners include:

- What partners already collect information that would be valuable to the network?
- What partners have the capacity to, but are not currently, collecting information that would be valuable to the network?
- What partners will use the collected information?

If the list of potential partner groups or organizations is long (>10), it may be beneficial to rank partners as *primary partners* - those that already focus on similar work and may already be involved in the initial planning phases; *secondary partners* - those that could do this work but are not currently a major player; and *key partners* - those partners who may have a large influence on the success of the network. Key partners may be regulatory authorities, funding agencies, or those with high esteem in the field of interest or are politically important. Initial network development will be most successful if it is focused on key and primary partners first, and subsequently incorporates secondary partners.

Once the list of partners is created, it will be important to also identify a point-of-contact for each partner group. A point-of-contact is a person within the group that will actively participate in the network, and be an advocate for it within their own group or organization. There may be more than one point-of-contact for a partner group depending on the complexity of the group's mission and the importance of their participation in the network. It will also be critical that key decision makers within each partner group are aware and supportive of the efforts, if they are not the point-of-contact. At this stage, professional relationships (i.e. informal networks) can play a helpful role in initiating contacts. Even if the personal contact within an organization is not the correct person needed by the network, they may be able to identify (and even potentially introduce) the correct person(s). *The success of the network relies on getting the correct people - those with access and permission to share (or use) wildlife health data - to actively participate in the network.*

Table 2 shows an example exercise that could help categorize potential partners (adapted from the NACCHO strategic planning guide: *https://www.naccho.org/uploads/downloadable-resources/Programs/Public-Health-Infrastructure/StrategicPlanningGuideFinal.pdf)*.

Table 2: An example exercise that could be used to identify potential network partners, their ranking (key, primary, secondary), and their point-of-contact.

Partner Organization	Partner Ranking	Point-of-contact

After the list of potential partners has been established, it is often useful at this point to map out the current state of partner relationships to identify those already well-aligned to participate in a network and highlight those relationships that need additional effort. Below is a network mapping exercise that can be used to establish the current state of the partner network, and through the follow up assessment questions, identify connections that can be strengthened to improve partner relationships and the functioning and viability of the network. This exercise is adapted from Harvard ManageMentor Network Mapping Tool: *http://www.harvardbusiness.org/sites/default/files/PDF/network_mapping_tool.pdf*.

Wildlife Health Network Mapping Exercise

Use this network mapping exercise to identify and assess the relative strengths and weaknesses of your existing or planned partner networks. The directions to map your partner network are below, followed by the exercise and some assessment questions to help identify action items to improve the network.

Map your partner network: Follow the directions below to fill in your table of partners and map your partner network.

1. You may complete more than one network mapping tool, so name your network in a way that is relevant to this specific network.

2. Describe the network's purpose/objectives. Refer to common reasons to establish wildlife health networks in the text above. If making multiple network maps (e.g. for different sectors or level or partners), please be specific for the objective of each network map.

3. List the partners and/or point-of-contacts who are currently members of your network.

Title	
Purpose	
Partner Group	**Point-of-contact**

4. Map the partners listed in your table onto the blank circular diagram:

a. Consider yourself as the center of the network.

b. Position each partner on the map:

i. Use the rings around the center to position your network partners according to how frequently you interact with them:

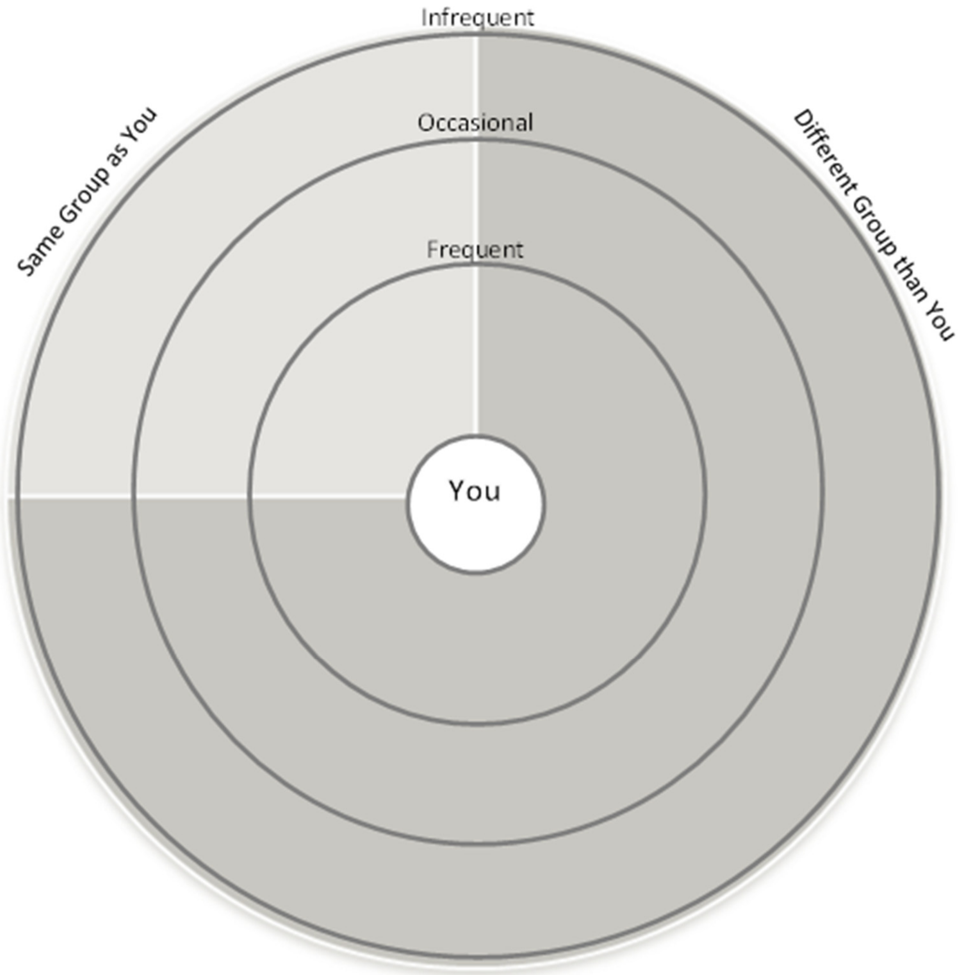
- Frequently (e.g., at least monthly)
- Occasionally (e.g., quarterly to annually)
- Infrequently (e.g., less than annually)

ii. Position the partners on the relevant section of the diagram:

- Same organization or group as you
- Different organization or group than you

iii. Draw lines connecting you to the partners of the network whom you have direct contact with.

iv. Draw lines connecting partners of your network if the members know one another. If you anticipate having many lines, it may be helpful to use a different color or a dotted line for these connections.



Assess your partner network: Think about the following questions to help identify opportunities to improve your network.

1. *Partners*. Who are you really dependent on to achieve your network's purpose? Are all of those groups and individuals included? Are there groups or individuals who will be critical in the future who are not part of your network?

2. *Structure***.** Is your network cohesive—i.e., do most people in your network know one another? Or is it a "bridging" type of network—where you're connected to people who aren't connected to one another? Do you see opportunities to connect groups or individuals who would benefit from knowing one another (thus providing some benefit back to your network partners)?

3. *Composition and breadth***.** Does your network include both strong and weak ties? Are most of your network partners clustered within your group or organization? Given your network's purpose, are there risks associated with any clustering in a particular health sector?

4. *Closeness***.** Are there important groups or individuals in your network with whom you should interact more frequently (e.g., someone whom you underutilize)? What initial reflections do you have on the quality of relationships within the network?

5. *Patterns***.** What other patterns on your map provide insight?

Improve your partner network: In the space below, list two or three specific opportunities to improve your partner network (e.g., add key relationships that are missing, spend more time with some partners, increase the diversity of your network).

1.

2.

3.

Even longstanding wildlife health organizations may find this above exercise useful. For example when the U.S. Geological Survey National Wildlife Health Center completed this exercise they identified the need to improve their relationship with state wildlife health departments because they are underrepresented in the network. In response, several actions were initiated including increasing communications and training opportunities for these departments to improve their participation in the wildlife health partner network.

During the planning process, in addition to the above mapping exercise, it may be beneficial to hold a workshop that includes primary and key partners to gather input on the scale, scope, and structure of the network being built. This would also be an ideal time to discuss the objectives of the partner network and requirements for participation. The most effective partner networks have buy-in, active participation and strong belief in the network's mission. A proven way to garner support is by providing these partners input in the network planning process. A common phrase used during team building exercises is "weigh-in leads to buy-in" - meaning that if stakeholders and partners feel like their contributions and concerns have been considered during the design and development of the network, they are much more likely to be supportive and participate in the resulting network structure. Thus, the planning workshop should be face-to-face if possible so the members of the partner network can build rapport with the lead organization's network coordination staff and with each other. If possible, a hired facilitator can help make sure there is participation from partner groups. The result of the planning workshop should be a clear set of guidelines and requirements that balance the abilities of partners with the needs of the lead organization.

To develop the structure of the network, as part of the aforementioned workshop or outside of that forum, the group of partners should discuss the risks of the network to their own organization's function and mission, as well as the benefits of the collaboration. The lead organization should also solicit information on the obstacles, and potentially resources needed by each partner, to meet the network's stated objectives. The group can brainstorm effective solutions for the most common obstacles. The lead organization should then establish the network structure, which can range from informal coordination calls to more formal membership requirements or funding agreements. Below three different types of partner relationships are described with advantages and disadvantages associated with each. These descriptions are adapted from Box 3 of The Partnering Toolbook by Ros Tennyson (p.17), 2003, The International Business Leaders Forum (IBLF) and the Global Alliance for Improved Nutrition (GAIN) *https://thepartneringinitiative.org/publications/toolbook-series/the-partnering-toolbook/*.

1) **Informal** networks can include personal contacts developed through business-style networking (see definition above), but can also include working, focus, or task groups. These types of networks generally do not have a signed agreement among partners, though they may have some kind of guiding document. Informal networks are probably the most common form of wildlife health network.

Advantages	***Disadvantages***
Greater freedom to explore ideas and intentions and to build new relationships	Not being taken seriously enough by external agencies or other key players
Cheaper – the major resource demand is time rather than cash	Too easily neglected when those involved are diverted by their other priorities
Non-bureaucratic	Not structured enough for the coordination and management of resources

2) Semi-Formal networks generally have written guidelines and protocols for participation that may be publicly served through a website or other media. Semi-formal networks also generally have signed agreements among the partners, such as project charters, memorandums of understanding, or collaborative agreements (see table 3). Generally these agreements are signed by all partners and partners agree to provide the resources needed to meet stated objectives (generally this is staff time).

Advantages	***Disadvantages***
Develops a greater profile	Needs greater coordination
Providing an 'umbrella' for a wide range of loosely linked activities	Requires more agreement on policies and operational principles
Building commitment from a wider constituency	More complex decision-making processes

3) Formal networks have a recognized organization, such as a non-profit association or foundation or a governmental agency that is charged with leading and maintaining the network. Formal networks are often supported by government mandates or regulations, or have resources to contribute to network operations, either through dedicated staff at the lead organization or through funding for partners to participate.

Advantages	***Disadvantages***
Increased authority and capacity to exert influence	Subject to legislative restrictions on action
More focused activities and greater likelihood of sustainability	Tendency to become over-bureaucratic and impersonal
Enhanced ability to mobilize and manage large-scale resources	Increasingly high administrative costs as opposed to project costs

Some common types of agreements used by these types of networks are defined below in *Table 3*. Regardless of the structure or type of a network established, the requirements and expectations for participants in the network should be clearly stated and outlined in any guiding documents or signed agreements.

Table 3: Common types of guiding documents or signed agreements used in partner relationships, on a scale from semi-formal (project charter) to formal (funding agreement).

Name	Type of document	Definition	Available Templates
Project Charter	Guidance	A document that defines the objectives, roles and responsibilities, and participating partners.	https://www.acf.hhs.gov/sites/default/files/fysb/mou_508.pdf
Memorandum of Understanding	Signed Agreement	A document that includes the information in a project charter, but is also signed by participating partners to show alignment and commitment to the stated objectives.	https://www2a.cdc.gov/cdcup/library/templates/CDC_UP_Project_Charter_Template.doc
Cooperative Agreement	Signed Agreement	Includes the information above, but also includes resource commitments from participating partners (generally represented by salary to cover staff time spent on the network).	https://www.loc.gov/programs/static/national-recording-preservation-plan/tools-and-resources/documents/Cooperative-Agreement-MBRS-project.pdf
Funding Agreement	Signed Agreement	A legal contract to transfer funds or other resources (materials, personnel, etc.) between partners to complete a stated objective.	https://www2.usgs.gov/usgs-manual/500/SM%20500.20_Technology%20Transfer_Figure%20500.20-3.doc

How do you maintain a partner network?

Once the initial structure and guiding documents are established and the network becomes operational, maintaining the network needs a continued commitment from the lead organization in terms of dedicated personnel and resources. The network coordination staff need to continue to work on relationships with partners and as the network grows they can establish participation by additional partners (secondary partners as described above). In addition, network coordination staff should re-evaluate the program regularly (at least every 5 years, but ideally annually) to make sure the network is meeting its objectives, and all primary and key partners are still engaged in the network. If the network is not meeting objectives or lacks the correct membership, then careful consideration should be given to how to engage new partners, disengage partners that are no longer needed, and remove obstacles to participation (see Overcoming Barriers to Wildlife Health Data Sharing section of this workbook).

The most effective partner networks will strive to incorporate three key principles - *equity*, *transparency*, and *mutual benefit* (see Figure 1). *Equity* acknowledges the right of each partner to be "at the table", that is that each partner is a valuable part of the network and has an equal right to participate and provide input. For example, suppose a province has dedicated wildlife health staff and actively participates in the partner network. A second province does not have dedicated staff for wildlife health and while officially part of the network, generally does not actively contribute. Both provinces request aid for an ongoing wildlife health event from the lead organization. Equity is providing the same aid to both partners, even though one has a better relationship with the lead entity.

Transparency is defined as being open and honest about decisions, processes, and relationships that affect the partner network. For example, in the above response, the lead organization should report how much aid it provides both provinces and the key deciding factors that led to that decision. Best practices in transparency suggest common decisions be clearly articulated in guidance documents so that partners can understand decisions for most situations. For example, developing communication plans can help foster transparency and reassure partners that important wildlife disease findings will be communicated in a timely and appropriate manner. However, the guidelines should be flexible enough to adapt to changing situations, new technology, and new emerging wildlife diseases. The lead organization can still make exceptions for extenuating circumstances, but those exceptions should be well communicated and documented.

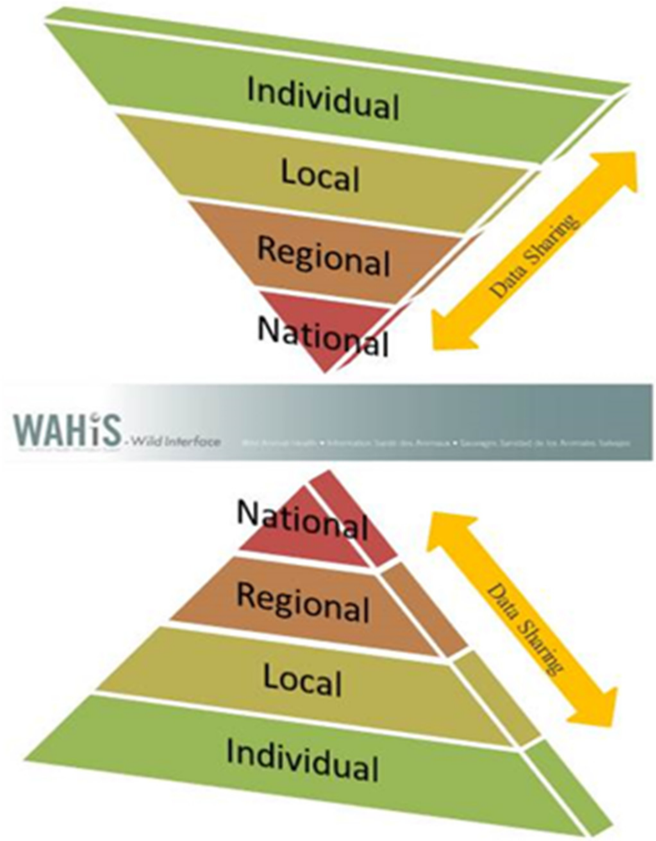


Figure 1. Example of information flow within and between two countries' partner networks for reporting wildlife health information.

Lastly, and perhaps most importantly, the relationship with the network needs to be *mutually beneficial* to partners and the lead organization. Ongoing discussions are needed during planning, implementation, and management of a partner network with partner groups to ensure that participation in the network benefits all parties. For example, there is often a hierarchy of increasing scale of data sharing within a partner network (see Figure 1 below) with data collection beginning with individuals that report wildlife disease observations to local authorities. These local authorities report this information to regional partners, and regional partners report to national agencies. Finally, national agencies may send this informational to international bodies such as OIE. Throughout this process and at multiple levels of this hierarchy, data are collated, and may be analyzed for drivers or trends. One way to assure partner networks are mutually beneficial is to ensure that collated data and associated findings travel back through the hierarchy all the way down to the individual level. This information can then be used in disease management activities, promotion of situational awareness, and provide the partners within the network with an understanding of how their data fits into the national trends (see Data Dissemination section of this workbook), providing mutual benefits to the national network and the local data generator.

It is important to maintain participation in wildlife health partner networks, and often long-term efforts are supported by either regulations or resources. Regulations can be used to enforce rules whereas resources are often used to encourage partners in their efforts to share information with the network. An example of a regulatory requirement is a legal mandate that diagnosis of certain notifiable diseases be reported to a lead organization. In contrast, resources can be provided to partner organizations to conduct routine testing for important diseases contingent upon following designated testing protocols and completing mandatory reporting of results. In the absence of these two motivating factors, partner networks can be maintained by strong partner relationships sustained through good communication and the efforts of network coordination staff that strive to ensure continuing equity, transparency, and mutual benefits.

Partner Networks in Practice

Now that we have covered the basic principles involved in developing and maintaining wildlife health partner networks, we present specific examples of application of these principles that cover the spectrum from informal to formal structures, regional to national scale, and disease-specific to broad in scope. We used a combination of interviews with the network coordinator, information from publicly accessible websites, and published literature to answer the following questions about each network:

1. What are the objectives of the partner network?
2. Who are the partners in the network?
3. How is engagement of partners maintained in the network?
4. How is the network structured?
5. How are data managed in your network?
6. How is information shared or disseminated back to partners?

The first case example is from a beach monitoring program, AMBLE, run by the U.S. Geological Survey National Wildlife Health Center. This program involved both partner organizations who helped plan and coordinate the network as well as citizen scientists who made and recorded mortality observations. AMBLE is a good example of an informal network with a regional focus (the Great Lakes Basin of the US) on a specific disease (avian botulism). The second case example highlights the Canadian Wildlife Health Cooperative, which is a semi-formally structured network among veterinary universities and the lead organization that monitors and maintains data about wildlife health mortality events throughout Canada. The last case example focuses on Wildlife Health Australia, which is a formal network among many different partners across several health sectors (including Australian zoos, the Australian agricultural department, and others).

Case Example 1 - Informal Partners Network (AMBLE)

Answers were provided by Jenny Chipault, lead technician on the AMBLE project at U.S. Geological Survey National Wildlife Health Center.

1. What is the purpose/objectives of the partner network?

The goal of AMBLE (Avian Monitoring for Botulism Lakeshore Events) was to empower concerned citizens to monitor bird health and beach conditions along miles of Lake Michigan shoreline, thus increasing knowledge of avian botulism trends.

2. Who are the partners in the network?

AMBLE had two related and connected networks: partner organizations (those that helped with planning and implementation) and citizen scientists (those that collected the data). The lead organization was several hours away from the field collection sites, so we actively recruited local natural resource agencies to participate. The initial contact with potential partner organizations was via email, phone, and in-person meetings, resulting in 15 organizations who agreed to participate. The partners included management agencies or non-profits with an interest in wildlife health issues; these agencies were motived to participate by wanting to better understand botulism outbreaks to improve management and/or to be able to answer questions about the outbreaks for the public. To do the actual beach monitoring, we actively recruited citizen scientist groups, which was ideal in locations with higher population densities. We connected with volunteers through our partner organizations, which used bulletin boards, websites, ads in the local newspapers or radio stations to recruit citizen scientists. Those that participated were willing to monitor shoreline near their house for free, in exchange for being a part of the collective effort to solve a local environmental problem. In addition, in areas where human populations were sparse, we contracted local biologists to monitor the beaches. Data from all groups and areas were collected following the same shared protocol so that results were comparable across space and time.

3. How is engagement of partners maintained in the network?

We maintained engagement through active and consistent feedback. We issued easily-digestible newsletters to both the citizen scientists and partner agencies. We regularly communicated through the network about publications, presentations, and media coverage that stemmed from data collected by the network. An emphasis of the lead organization was to provide information and benefits back to the partner organizations, often presenting at chapter meetings of citizen science groups. We held annual in-person workshops with partners to show our commitment to the project and to ensure partners had summaries of the results-to-date. Lastly, we recognized citizen scientist efforts by giving out awards at the end of each season, such as the person who reported the rarest bird during their monitoring or the person who collected the most carcasses. Citizen scientists were given the opportunity to provide feedback on an evaluation sheet distributed at the end of every season and the program was adjusted based on their suggestions.

4. How is the network structured?

The results of the network needed to address the research and management needs of our key partners, which included an inter-agency group of researchers funded by a grant from the Great Lakes Restoration Initiative (GLRI). To effectively do this, we needed a network of agencies and citizen scientists (see #2) to cover the entire Great Lakes region. Determining which partner organizations would be primary and which would be secondary happened organically; some organizations volunteered resources, such as locations for citizen science training or staff time, to get the program initiated. The network coordinator had a good relationship with the points of contact at those primary organizations, and made an effort to build and maintain personal relationships through extracurricular activities after network trainings or workshops.

5. How are data managed in the network?

Because many of the many people (>100) collecting data were not professional scientists, all data fields were explained very clearly and there was a standardized datasheet and written protocol. Data collectors also received accompanying training and their training participation was tracked. Early in the project, we developed an online data portal for citizen scientists to enter their data [Aside: this helps with quality assurance, as those who collect data are more likely to catch mistakes when entered]. The portal was designed to be user-friendly and to mirror the standardized datasheet. User access to data was controlled through roles and permissions built into the data portal, which fed into a centralized database housed at the lead organization. Network coordination staff at the lead organization did quality checks on the data and followed up with the volunteers regarding any unusual values. The data portal is still active and can be found here: *https://www.nwhc.usgs.gov/amble/.*

6. How is information shared or disseminated back to partners?

Network coordination staff emailed newsletters and notices about presentations, publications, or media coverage to the group and held annual workshops for partner organizations and citizen scientists. Additionally, data and information was posted to a network website. Collected data were used in four peer-reviewed manuscripts, as well as at least a dozen presentations.

Case Example 2 - Semi-formal Partner Network (CWHC)

Answers were summarized from the CWHC website (http://www.cwhc-rcsf.ca/) as well as from responses to requests for information sent via email to Craig Stephens, the Chief Executive Officer of CWHC, and Patrick Zimmer, Chief Operating Officer of CWHC.

1. What is the purpose/objectives of the partner network?

The CWHC is dedicated to generating the knowledge needed to assess and manage wildlife health and working with others to ensure that knowledge is applied in a timely fashion.

2. Who are the partners in the network?

At the core of the CWHC is a partnership linking Canada's five veterinary colleges and the British Columbia Animal Health Centre. These represent the primary partners that help to plan and operate the network. Branching from that core are secondary partners that include both the public and private sectors, which allows us to access critical expertise needed to detect and assess wildlife health issues and make sure our results find their way to people who need to make decisions on wildlife management, wildlife use, public health and agriculture.

3. How is engagement of partners maintained in the network?

Because CWHC primary partners generate the bulk of the data, secondary partners within the distributed network (provinces/territories) look to the CWHC to provide either the primary wildlife health surveillance program for their jurisdiction or to augment existing surveillance programs. In general the CWHC is a data provider, responding to numerous requests for data from government partners and other researchers. As the designated OIE Focal Point for Wildlife in Canada, the CWHC is obligated to provide summary data on selected diseases every 6 months to the Canadian Food Inspection Agency (CFIA). In those instances where we collect/report on data we are not generating ourselves, the data provider provides their data on the basis of believing in broader regional and national awareness and health intelligence. The primary incentive to do so is the additional knowledge one receives from being part of the broader program, the central communications, reporting and data management that does not need replicating (hence cost savings) and assistance in decision making brought by enhanced awareness.

4. How is the network structured?

The CWHC has no legislated mandate, yet we have provided a national program for Canada based more or less on handshakes and good will for the past 25 years. Based on its position and partnerships, the CWHC serves as a focal point to coordinate and link federal, provincial and territorial programs, priorities and information. This includes our role as a spearhead, a secretariat and facilitator of national strategies and plans to address wildlife health challenges.

5. How are data managed in the network?

CWHC has a centralized database to collect and curate data, primarily the data generated by veterinary university partners.

6. How was information shared or disseminated back to partners?

The CWHC has quarterly reports, annual reports, newsletters, fact sheets, and other communication documents. It also provides data, surveillance, and technical expertise as requested by partners and user groups. The central office provides communications and social media support. The University-based CWHC members publish data and results in peer reviewed publications.

Case Example 3 - Formal Partners Network (WHA)

Answers were summarized from the WHA website (https://www.wildlifehealthaustralia.com.au/Home.aspx) as well as from documents and email responses to requests for information from Tiggy Grillo, National Coordinator and Australian OIE Focal Point for Wildlife Health.

1. What is the purpose/objectives of the partner network?

Wildlife Health Australia (WHA) is the peak body for wildlife health in Australia. Our principal objectives are the protection and enhancement of the natural environment. We link, inform and support people and organisations who work with or have an interest in wildlife health and provide leadership, coordination, technical advice, facilitation, communications and professional support. Through our activities we also help link the environment, animal health and public health sectors. We provide a framework that helps us protect and enhance our natural environment by better identifying, assessing, articulating and managing risks posed by wildlife diseases to Australia. The framework assists Australia in better identifying, assessing, articulating and managing these risks. We provide the framework for Australia's general wildlife health surveillance system.

2. Who are the partners in the network?

WHA has a network of more than 700 wildlife health professionals, members of the public and those with an interest in wildlife health and the protection and enhancement of Australia's natural environment. Our network includes representatives from federal, state, and territory conservation, agriculture and human health agencies and industries, universities, zoos, wildlife clinics, private practitioners, wildlife care groups, hunters and fishers, diagnostic pathology services and members of the public. Primary surveillance partners include every state or territory, plus the Australian Antarctic Division, who each has a WHA coordinator, appointed by the state or territory chief veterinary officer, based in the agriculture agency. There is also a corresponding point-of-contact within the environment agency of each state or territory. Additional complementary partners were added in stages, with the first a collaboration with the Zoo and Aquarium Association and includes 10 participating zoo based wildlife hospitals. Additional partners that were added more recently include 8 wildlife hospitals and clinics and 7 university veterinary clinics and/or pathology departments.

3. How was engagement of partners maintained in the network?

WHA encourages participation of partner groups by building trust-based relationships and networks that focus on listening, two-way communication (transparency), co-investment and confidence in the system, facilitating understanding of different perspectives (political, social, ethical, and/or commercial), ensuring partners get what they need in return (mutual benefit), and proactively providing information to partners. Building trust takes time. A proven track record of appropriately managing sensitive and confidential data is crucial. The time required to build the trust necessary to obtain data from different stakeholder groups may take time (which may be years). In addition, WHA staff make a point of thanking partners, respecting confidentiality and intellectual property of partners, and acknowledging partner contributions. Lastly, WHA provides training and administrative support for partners contributing data and has access to funds that can support significant disease investigations.

4. How is the network structured?

WHA is a for purpose (not-for-profit), incorporated association registered under the Associations Incorporation Act (2009) in New South Wales, Australia but operates nationally. WHA is a registered charity. WHA extends the work of the Australian Wildlife Health Network (AWHN) and is administered under good organisational governance principles. An elected management committee (Board) oversees the activities of WHA.

5. How are data managed in the network?

WHA administers the national electronic Wildlife Health Information System (eWHIS) database, a web-enabled, secure database capturing information relating to wildlife health surveillance and disease investigation in Australia. WHA receives wildlife health data from sources both within governments (including state and territory agricultural, environment and health agencies) and from sources outside of governments (such as university researchers, zoo wildlife hospitals and private veterinary practitioners). Other sources include the Australian Registry of Wildlife Health, specific targeted projects such as the avian influenza wild bird surveillance program, and national datasets such as Australian Bat Lyssavirus (ABLV) testing. WHA collates and moderates the data in eWHIS to ensure that it is as accurate as possible.

6. How is information shared or disseminated back to partners?

Quarterly teleconferences, quarterly reports, fact sheets, and weekly digests are developed to communicate with partners about on-going activities and wildlife health events. WHA emphasizes good communication and maintaining partner trust and confidentiality for submitted data. WHA also provides outbreak response support through expert consultation, assistance with dissemination of agreed messaging and potential funding opportunities.

Conclusion

Even though each case example above had different structures, scopes, and objectives, there were some similarities in the way they answered the six questions. In particular, all three networks emphasized the importance of communication and information sharing on a regular and real-time basis with partners - a main way that the network is providing *mutual benefit* to partner groups. In addition, all three networks had strong partner engagement by building trusting relationships and being responsive to partner needs - an important way to get *partner buy-in* and maintain participation in the network. In the first example, the organizers of AMBLE spent considerable time and effort recruiting citizen scientists - that is ensuring they had the *right partners* - for the network to be successful. The same can be said on a larger scale for the third example, where secondary partners are continually being added to expand and strengthen the existing network. Lastly, based on the objectives of the partner network, all three examples were *customized* to successfully meet the stated purpose. These four traits, customized, right partners, partner buy-in, and mutually beneficial, are critical components of successful wildlife health partner networks.

Collection and Curation of Wildlife Health Information

Although most wildlife health professionals receive training in how to collect data for a specific research project or purpose, few have formal training in data management. In this section, we provide a broad overview of many of the basic concepts we have found useful for curating wildlife health information. Our intent is to emphasize the importance of considering how data is managed and stored, so that it can be more easily used and shared for wildlife health initiatives. More detailed information on any of these topics can be found in textbooks (e.g., Briney 2015) or on-line resources dedicated to data management. The specific objectives for this section include:

- Provide an overview of the data management lifecycle as a useful outline for many of the broad concepts needed to curate wildlife health information;
- Review the need to establish common data fields and define them during wildlife disease surveillance;
- Describe data dictionaries and common methods used to construct metadata;
- Provide an overview of quality assurance and quality control techniques commonly used during data management;
- Describe best practices to consider during data storage including the importance of storing data in machine readable formats and the benefits of storing data in databases compared to spreadsheets;
- Provide list of open source freeware that can be utilized for storing and managing wildlife health information.

Data Management Life Cycle

Data management has a number of steps that are often referred to as the data management life cycle. The six primary elements of the data lifecycle are depicted in the Figure 2 below with the shaded arrows below the boxes representing aspects of data management that should take place throughout the life cycle (from Faundeen *et al.* 2013).

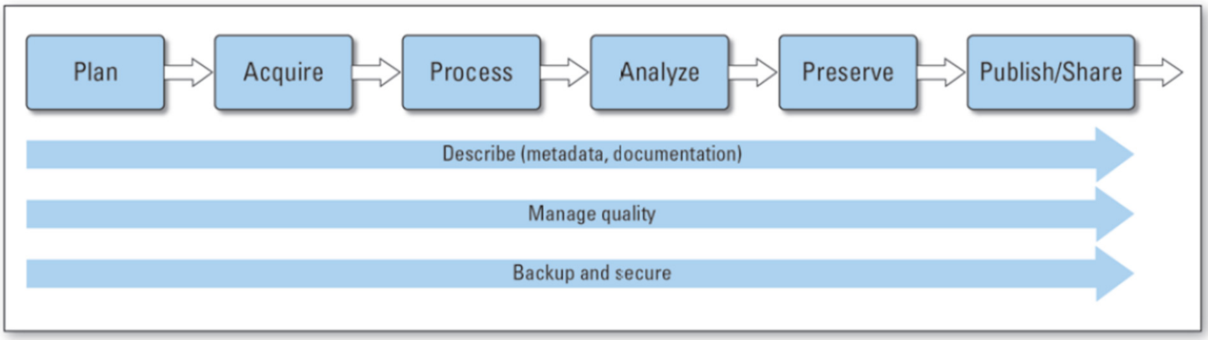


Figure 2. Six elements of the data life cycle.

The first phase, the *plan* phase, includes determining why data collection is needed, what data need to be collected, how they will be used to answer questions, and what standards should be applied to data (e.g., terminology, units of measure, spatial scale). In our experience, the *plan* phase is not always given full consideration, as scientists are often most interested in starting the *acquire* phase where the data points of interest are actually collected. During the *process* phase, data should be verified, organized, and transformed into an appropriate output for subsequent use. The *analyze* phase is used to detect patterns, develop explanations, and test hypotheses. The most commonly used action during the *preserve* phase is the storage of data on individual computers or external hard drives, but may also include data archiving. In the *publish/share* phase, data are made available to the public or other agencies directly (e.g., sending data to OIE) or indirectly (e.g., making data available in a data repository). *Describing*, *managing quality*, *backing up, and securing* are also extremely important aspects of data management and as indicated by the long shaded arrows in Figure 2 above; these processes should be considered during the planning phases and performed throughout the entire life cycle.

Although all of the steps in the data life cycle are important, in this section we will be focusing primarily on methods that can be used to *plan, process, preserve, describe,* and *manage the quality* of wildlife health information to ensure data collected by numerous partners can be combined and used for regional and national wildlife health surveillance. Multiple methods and activities used to *acquire* and *analyze* wildlife health data were discussed in the 4th cycle OIE workshop: Wildlife Disease Outbreak Investigations. Activities associated with Outbreak Investigation are often performed by multiple partners from a variety of backgrounds (e.g., hunters, veterinarians, the public, and wildlife rehabilitators) within a country. The development of wildlife health partner networks is the topic of the previous section. Some basic approaches for *publishing/sharing* data will be included in the last section of this workbook entitled "Data Dissemination".

Wildlife Health Data Planning

Establishing primary data fields

A basic premise when combining multiple datasets is that similar information was collected. As described in 4th cycle OIE workshop, most wildlife disease outbreak investigations seek to determine answers to questions about the who (what population is affected), what (disease etiology), where (geographic extent of the problem), and when (time frame of the event) of a wildlife disease event. Typical primary data fields used to capture the answers to these questions include:

- Location of the event
- Onset and end date of observed morbidity and mortality
- Species affected
- Number of affected animals
- Clinical signs observed in affected animals
- Diagnoses
- Laboratory where diagnostics were performed
- Contact: person(s) reporting the event and their contact information in case additional information is needed

Defining primary data fields

For optimal global wildlife health surveillance, there would ideally be a set of universally accepted definitions for these primary data fields to assure a common understanding of terminology and methods of measurement. However, since global standards do not currently exist for most data fields, it becomes critically important for each data collector to document how they defined and measured the information in each field. This, ideally would be completed *before* the data are collected. Although some data standards can be applied after the data are collected, or even once they have been shared, it is more complicated and time consuming than applying them at the time of collection. In some cases it may be impossible to acquire needed information from the data if the standard was not applied at the time of collection. For example, if exact location data is needed (e.g., latitude and longitude) but only the county, province, or general description of the location was collected by the partner, it may limit the usefulness of the data. Note, that some standards, such as date, time, and country codes, may already be available (see International Organization for Standardization *https://www.iso.org/home.html*). However, allowing partners to provide feedback on terminology can promote buy-in, acceptance, and use of common terminology standards. Some options and considerations for defining the most common fields are described below.

Location of the event: Although the definition of location may seem straightforward, it is important to clarify the scale and accuracy of location information collected. Latitude and longitude provide the most precise estimate of location and can be accessed through most smartphones or GPS- enabled mobile devices. Specifying preferences for formats (e.g., decimal degrees, degrees minutes seconds) and geodetic datums (e.g., WGS 84, NAD 83, ETRS 89) will reduce time needed for conversions after data are received. Requesting additional information about the location such as the name of the park, body of water, town, or other landmarks may also useful, particularly if the distribution of affected animals spans large spatial scales. If there are concerns about proprietary location information, conservation concerns (e.g., releasing exact locations of species at risk may be undesirable) or issues accessing private land, exact locations can be expanded to larger spatial scales (e.g., county, province, state, country) or rounded to one or two decimals (if using decimal degrees), before data are distributed more widely.

Onset and end date of observed morbidity and mortality: In most wildlife health datasets, the onset and end dates of the disease event will be an estimate because it is unlikely that the first or last incidence of disease was directly observed. Nevertheless, these dates are still critical for establishing temporal cycles of the events and understanding whether the events change temporally. The preferred format for the date should be specified (see detailed example of various date formats later in this section).

Species affected: Species affected typically refers to all species observed dead or sick at a given location during an event. Since many species have common names that vary by region or language, utilizing scientific names as the standard for this field can help ensure consistency in reporting. However, even scientific names are subject to change as new discoveries in systematics are made, so utilizing a system such as the Integrated Taxonomic Information System Taxonomic Serial Number (ITIS TSN*; http://www.itis.gov*) to code species may be helpful. For each scientific name, ITIS includes the authority, taxonomic rank, associated vernacular names, a unique taxonomic serial number, data source information (publications, experts, etc.) and data quality indicators. Another benefit of using an established system such as ITIS is that changes to taxonomic information in the ITIS database are tracked, so that if a scientific name changes you can see its history and determine whether similar changes need to be applied to your database or to queries involving that species. Although the initial emphasis for ITIS is on categorizing and tracking North American taxa, the intention is to expand geographic coverage worldwide.

Number of affected animals: The number of affected individuals typically refers to a count of the number of dead and sick animals for a particular species at a given location during an event. There could be separate fields indicating number sick and number dead for each species if it important to distinguish between the two. A total count of all affected animals can be calculated from numbers reported from individual species. For many locations and vegetation types, actual counts of affected animals may be difficult to perform, so estimates of numbers affected may be more easily obtained. However, since individual researchers vary in their likelihood to over or underestimate, it is important to note in the database and metadata that these numbers are estimates. Estimates can also be generated using population dynamic models for unmarked individuals such as double-observer counts or distance sampling. The benefit of these approaches are that the probability of detecting a sick/dead individuals is estimated directly, which permits a correction for missing individuals and allows for the estimated number of individuals to be accompanied by a measure of precision.

Clinical signs: Variables such as clinical signs can be used in conjunction with other variables such as species to suggest a potential cause for the event prior to performing diagnostics. However, clinical signs are not always disease-specific and may be difficult to standardize so should be used with caution during reporting and for analyses.

Collecting information on clinical signs may also be useful for reporting to OIE on non-listed diseases. For example, persons entering data into WAHIS-Wild are requested to "enter for each disease, infection or infestation and for each category of species (domestic species, wildlife) one of the following occurrence codes. If no information is available on an OIE-listed disease, please select the occurrence code "no information" on the WAHIS online notification application or indicate "…" on the paper form." Example of OIE occurrence codes for WAHIS-Wild are shown below.

OIE occurrence codes indicating disease presence in domestic species or in wildlife

Disease present	+	The disease is present with clinical signs in the whole country (in domestic species or wildlife
Disease limited to one or more zones	+0	The disease is present with clinical signs, and limited to one or more zones/compartments (in domestic species or wildlife)

OIE occurrence codes indicating the presence of the infection or infestation in domestic species or in wildlife without clinical signs

Infection/infestation	+?	Confirmed infestation or infection using diagnostic tests, but no clinical signs observed (in domestic species or wildlife).
Infection/infestation limited to one or more zones	+?0	Confirmed infestation or infection using diagnostic tests, but no clinical signs observed and limited to one or more zones/compartments (in domestic species or wildlife).

Diagnoses: Providing a list of diagnoses can help standardize the preferred diagnostic terminology (e.g., infection with chytrid fungus vs chytridiomycosis) and make partners aware of which diagnoses are of interest for reporting purposes. The OIE maintains a list of wild animal diseases that are not OIE-listed but that have been selected for monitoring by the OIE's Working Group on Wildlife Diseases. Disease are included on this list because of their importance for wild animals, and for early warning purposes to protect human and wildlife health. This list can be found here: *http://www.oie.int/wahis_2/public/wahidwild.php/Diseaseinformation/popup/diseaselist*.

If you are soliciting data from partners you should specify whether you are interested in only confirmed cases or both suspect and confirmed cases for a particular disease. It is also extremely important to define whether partners should report the cause of death for all individual examined animals or only the primary diagnosis for the majority of animals affected in an event (i.e., the *event* diagnosis). See the USGS-National Wildlife Health Center's Wildlife Health Information Sharing Partnership event reporting system (WHISPers) for an example of a system that reports *event* diagnoses: https://www.nwhc.usgs.gov/whispers/.

Lastly, it is critically important to ensure that the criteria used to make the diagnoses were consistently applied among animals within the same event and among events over time and space. Case definitions as defined by the OIE are "a set of criteria used to distinguish a case animal or an epidemiological unit from a non-case". Case definitions typically include a scientifically accepted and clearly defined set of field, gross, histopathologic, laboratory, and epidemiologic criteria used to assign an individual to a specific disease category for reporting purposes, and are important for counting and classifying cases of a disease consistently across jurisdictions. Case definitions have been previously established for several diseases affecting wild species that are also of agricultural or human health importance and may impact international trade of animals or their products. Where one exists, the case definition in the specific chapter of the OIE Terrestrial or Aquatic Code should be used (see *http://www.oie.int/en/international-standard-setting/terrestrial-code/access-online/* and *http://www.oie.int/en/international-standard-setting/aquatic-code/access-online/*). If the Terrestrial or Aquatic Code does not provide case definition for the disease of interest, the criteria used to classify a case may be based on other sources or generated internally. Examples of case definitions from OIE and United States Department of Agriculture-Animal Plant Health Inspection Service-Veterinary Services, that outline criteria for classifying individuals as confirmed positive or suspect cases can be found here:

- Middle East Respiratory Syndrome Coronavirus: http://www.oie.int/en/scientific-expertise/specific-information-and-recommendations/mers-cov/
- Ranavirus: http://www.oie.int/fileadmin/Home/eng/Health_standards/aahm/current/chapitre_ranavirus.pdf
- Novel Swine Enteric Coronavirus: https://www.aphis.usda.gov/animal_health/animal_dis_spec/swine/downloads/secd_case_definition.pdf
- Highly Pathogenic Avian Influenza: https://www.aphis.usda.gov/animal_health/emergency_management/downloads/hpai/case_definition.pdf
- West Nile Virus: https://www.aphis.usda.gov/vs/nahss/equine/wnv/case_definition_west_nile_virus_09_15_11.pdf

The pathogens in the case definition examples above affect wildlife but also have significant human or agricultural animal health implications. Established case definitions for diseases that are restricted to wildlife are less common, but have been developed for several recent emerging wildlife diseases in the United States, including *Batrachochytrium salamandrivorans* chytridiomycosis (White *et al.* 2016) and white-nose syndrome (*https://www.nwhc.usgs.gov/disease_information/white-nose_syndrome/Case%20Defintions%20for%20WNS.pdf*). Utilizing common definitions as well as the continued development of case definitions for diseases of importance to wildlife would promote more consistent reporting of wildlife diseases internationally. Typically, case definitions include the following fields:

- Individual, place, and time: common species, age groups, locations, and time of year typically associated with the disease;
- Diagnostic description: clinical, gross necropsy, and histopathology observations diagnostically compatible with the disease;

- Laboratory criteria for diagnosis: diagnostic findings or laboratory test results sufficient for suspicion or confirmation of diagnosis;
- Epidemiologic linkage criteria: description of epidemiologic linkages, such as geographic scale or time period, that define linkage to other cases;
- Case classification categories: common classifications include suspect, probable, and confirmed, with the diagnostic and epidemiologic information used as criteria for classification in each category;
- References: may include scientific literature or agency websites used to construct the definition.

The USGS-National Wildlife Health Center is currently developing case definitions for several other common and emerging wildlife diseases in the United States (e.g., avian botulism, avian pox, avian salmonellosis, avian cholera, and snake fungal disease) to increase consistency in reporting these diseases within the wildlife health community. These case definitions will be shared online as they are developed.

Laboratory where diagnostics were performed: Collecting contact information (name, address, email, phone number) for the laboratory where a diagnosis was made is important, particularly for new, emerging, or reportable diseases, as follow-up on test results or methods may be warranted.

Collector: The name and contact information for person(s) reporting the event should be collected for each event in case additional information is needed. This field is also useful to acknowledge a partner's role in providing the information.

Other possible data fields: Each piece of information requires a partner's time to collect and share, so the decision to collect additional data fields should be considered carefully, and its use and value should be weighed against the potential reduction of information sharing due to the additional effort for partners. For example, recording the age class of affected animals may be useful for providing a timely suspect diagnosis since some disease agents preferentially affect juveniles or adults. Recording sex of affected animals may also be important as behavioral differences (e.g., foraging, nesting, and territorial defense) may result in differences in mortality rates. However, it is also important to recognize that sex and age can be difficult to determine for some species. Therefore, information in these fields may be unreliable or difficult to standardize so should be used with caution during reporting and for analyses.

Constructing a data dictionary

If you intend to combine wildlife health datasets from multiple partners, providing partners with a data dictionary that includes field definitions and units of measurement before they collect data can help ensure consistency among datasets. Data dictionaries allow you to define variables and provide contextual information that cannot be captured in the dataset itself. This information should be stored with the data so that it can be properly interpreted by current and future users of the dataset. The following list of information that is often captured in a data dictionary was compiled by Briney (2015):

- Variable name
- Variable definition
- How the variable was measured
- Data units
- Data format
- Minimum and maximum values
- Coded values and their meaning

- Representation of null values
- Precision of measurement
- Known issues with the data (missing values, biases, etc)
- Relationship to other variables
- Other important notes about the data

Table 4 below is an example of a data dictionary for several of the key fields described in this section. The units and data type for each field has been specified in the "Data Type" column. A detailed description of each field has been provided in the "Data Definition" column.

Table 4: Example data dictionary.

	A	B	C
1	**Field Name**	**Data Type**	**Data Definition**
2	Event_Type	Integer	1 = Morbidity/Mortality: Sick or dead animals linked spatially and temporally. Occurrence of single animals is included if there is special interest in the species, the suspected agent, the location, or the time of year (e.g., a solitary species, an endangered species, a possible new pathogen, or a range or temporal expansion for an existing pathogen) 2 = Surveillance: positive detections of a pathogen during active surveillance of healthy live, hunter-killed, or euthanized animals (that were not sick before euthanizing).
3	Start_Date	ISO 8601 YYYYMMDD	Beginning date for event (considering all locations).
4	End_Date	ISO 8601 YYYYMMDD	Ending date for event (considering all locations).
5	Affected	Integer	Total number of individuals affected in event. A count of sick plus dead for a morbidity/mortality event and a count of positives for a surveillance event.
6	Dianosis_id	Integer	ID for event diagnosis. Foreign key link to diagnosis look up table.
7	Species	Integer	Species ID. Foreign key link to species lookup table
8	Population	Integer	Estimate of the total population of this species at this location (population at risk). Use the peak number during the course of the event.
9	Sick	Integer	Actual count of the number of sick or injured animals of this species at this location. Include euthanized animals, if any. Use 0 if known to be 0 (instead of leaving blank). Leave blank if there is no count. Avoid re-count of animals, especially if there are repeated visits to a location to assess wildlife health. Consider whether animals initially observed sick were later counted as dead; if so, only count them as dead. Numbers reported should reflect either a snapshot of morbidity/mortality as observed during a one-time site visit, or a synopsis of the numbers affected over the course of an event (e.g., dead = cumulative dead during multiple site visits and sick = number remaining sick or recovered from being sick at the end of the event).
10	Dead	Integer	Actual count of the number of dead animals of this species at this location. Do NOT include euthanized animals. Use 0 if known to be 0 (instead of leaving blank). Leave blank if there is no count. Avoid re-count of animals, especially if there are repeated visits to a location to assess wildlife health. Consider whether animals initially observed sick were later counted as dead; if so, only count them as dead. Numbers reported should reflect either a snapshot of morbidity/mortality as observed during a one-time site visit, or a synopsis of the numbers affected over the course of an event (e.g., dead = cumulative dead during multiple site visits and sick = number remaining sick or recovered from being sick at the end of the event).

Processing Wildlife Health Data

Quality assurance and quality control

The two main types of data errors are: errors of commission and errors of omission. Errors of commission refer to the entry of incorrect or inaccurate data. They can be caused by a malfunctioning instrument or human error such as mistyped data. Errors of omission result from data failing to be recorded (DataOne 2012a). The most common errors of omission are forgetting to collect a measurement or accidentally skipping a line or field on a spreadsheet during data entry. Malfunctioning equipment, such as a GPS running out of battery power, can also cause errors of omission.

Activities used to prevent or correct errors in a dataset are typically referred to as quality assurance and quality control. Quality assurance (QA) refers to the processes and methods used to prevent data errors whereas quality control (QC) refers to detection and repair of data quality issues (US Geological Survey 2018). Basic QA procedures to consider *before* data entry include providing data field definitions (see examples above) and specifying measurement units. Quality assurance plans also can include step-by-step instructions for entering data which can be communicated through online tutorials, in-person trainings, or written manuals.

Methods used to eliminate errors *during* data entry often involve the design of the database itself. Whenever possible, atomized data should be used whereby only one piece of information is in each cell of the spreadsheet or database. For example, if the location of mortality event is Pond A, Park B, Region C, Country D, each segment of this location should be its own field. Multiple pieces of information in a single cell will make data analysis more challenging. Restricting the type of information that can be entered into a cell is another common QA procedure. For example, lookup tables or drop-down lists can be used to standardize acceptable values or terminology for data fields. You can also set up a field to accept only text or numerical values, choose range of values that a field will accept, or set a field to accept only unique values (more details on this below in the section on creation of spreadsheets and databases).

Double data entry is another possible QA procedure that can be used during data entry. Double data entry involves two people independently keying in data and then using a computer program to examine the two entries for discrepancies. This process is time consuming, but may be worthwhile if the data is extremely valuable and other methods are not possible (DataOne 2012a).

A final but important step in QA during data entry is the documentation of any changes made to the data after they have been entered. This may involve a simple process such as creating a text file to accompany the data set or using a scripted program that records data changes, the date, and the person who made the changes (DataOne 2012a).

After data are entered they should undergo QC procedures, including evaluation for missing or impossible values, which can be assessed through basic statistical summaries such as means and standard errors. Outliers or extreme values can sometimes be assessed by sorting and visually inspecting data or by graphing the data using normal probability plots, scatter plots, or subtracting the mean from individual values for a particular data field (DataOne 2012a). Locations can be mapped to ensure the latitude and longitude are correct.

Online resources describing additional data quality management practices can be found at:

- Food and Agriculture Organization of the United Nations: *http://www.fao.org/docrep/005/ac665e/ac665e07.htm*
- Distributed Active Archive Center for Biochemical Dynamics: *https://daac.ornl.gov/datamanagement/*
- A. D. Chapman. 2004. Principles of Data Quality: Report for the Global Biodiversity Information Facility. Global Biodiversity Information Facility, Copenhagen. *https://www.gbif.org/document/80509/principles-of-data-quality*

Preserving and Describing Wildlife Health Data

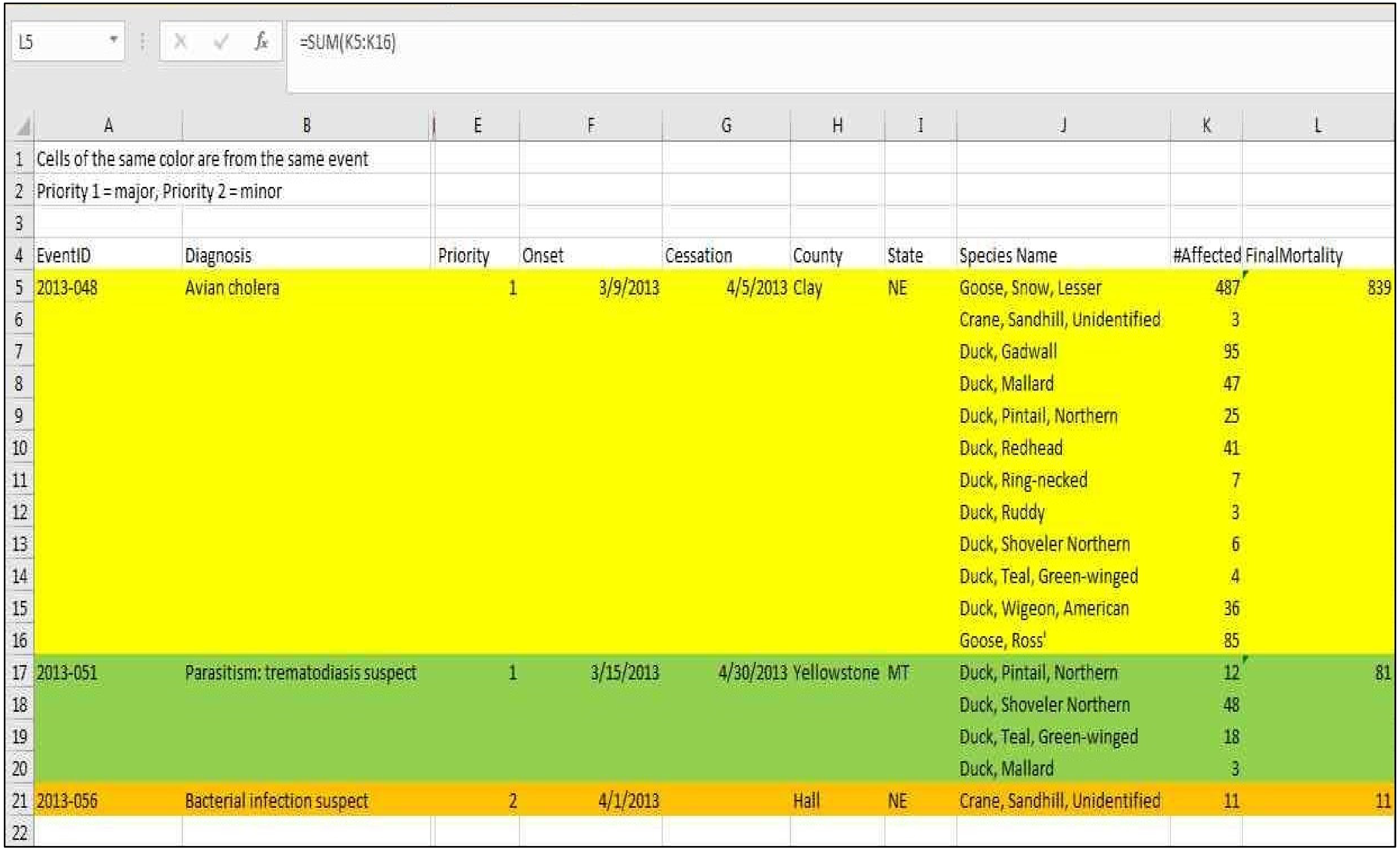
Storing wildlife health data

Although handwritten field notes or datasheets are still very common for initial data collection, data also ultimately need to be stored in a "machine readable" format regardless of whether you are storing wildlife health data for an individual site or for an entire country. "Machine readable" is defined as information that can be read and understood by a computer, without human aid, due to both its format and structure. Appropriate machine readable formats for various types of data include:

- Tabular data: CSV (comma-separated values)
- Textual data: HTML, plain text (.TXT), accessible PDFs (text is accessible by a computer), XML
- Geographic data: GeoJSON, CSV, KML, GML

Data saved in electronic format is not the same as machine readable. Photographs stored electronically are not machine readable and neither are some data entry shortcuts used in spreadsheets.

To illustrate the difference between machine- and human-readable formats. Several common spreadsheet conventions that scientists use to make the data human readable, but which are not machine-readable are illustrated on the spreadsheet on the next page. First, highlighting is used in the example spreadsheet to denote which data are part of the same event. While this may make it easy for a person to visualize this information, it is not captured in a way that the computer could read. In the example below, columns C and D have also been hidden to signify that this information should not be considered in a summary, but the computer would not be able to distinguish the hidden columns from other columns or capture the fact that they are should not be included in a summary. Although formulas (e.g., column L is the sum of associated highlighted rows in column K) are a great shortcut available in most spreadsheets, calculated values are not maintained when spreadsheets are saved as CSV files; however, spreadsheets may offer a means to convert formulas to calculated values. A final example in the Figure 3 below is the insertion of documentation or notes about the data in the data file (rows 1 and 2). This information should be included in a metadata or "readme" file (see section on metadata below for more information), but not in the data file.



L5 | =SUM(K5:K16)

	A	B	E	F	G	H	I	J	K	L
1	Cells of the same color are from the same event									
2	Priority 1 = major, Priority 2 = minor									
3										
4	EventID	Diagnosis	Priority	Onset	Cessation	County	State	Species Name	#Affected	FinalMortality
5	2013-048	Avian cholera	1	3/9/2013	4/5/2013	Clay	NE	Goose, Snow, Lesser	487	839
6								Crane, Sandhill, Unidentified	3	
7								Duck, Gadwall	95	
8								Duck, Mallard	47	
9								Duck, Pintail, Northern	25	
10								Duck, Redhead	41	
11								Duck, Ring-necked	7	
12								Duck, Ruddy	3	
13								Duck, Shoveler Northern	6	
14								Duck, Teal, Green-winged	4	
15								Duck, Wigeon, American	36	
16								Goose, Ross'	85	
17	2013-051	Parasitism: trematodiasis suspect	1	3/15/2013	4/30/2013	Yellowstone	MT	Duck, Pintail, Northern	12	81
18								Duck, Shoveler Northern	48	
19								Duck, Teal, Green-winged	18	
20								Duck, Mallard	3	
21	2013-056	Bacterial infection suspect	2	4/1/2013		Hall	NE	Crane, Sandhill, Unidentified	11	11
22										

Figure 3. An example of a spreadsheet displaying information in a non-machine readable format.

Below we list some best practices that should be used whenever storing data.

- Use one header row;
- Limit column headings to 12 characters or less as truncating can occur when exported to various file formats;
- Create descriptive field or column headings (e.g., AvianData_18). Use only letters, numbers, and underscores in column headings. Do not use spaces, hyphens, or other character symbols such as + - & * as they may be interpreted as mathematical operators by some statistical programs;
- Each field or column should contain only one type of information (e.g., text or "string" data, integer numbers);
- Use consistent formats for each field (e.g., all dates should be in the same format);
- Use consistent codes (e.g., if abbreviations are used for species names they should be consistent each time a species is entered);
- Specify what notation is used for missing data. In numerical fields "9999" is often used to indicate a missing value. In text fields "NA" is often used to reflect "Not Applicable" or "Not Available";
- If using a spreadsheet, each line should be complete. All cells in each row and column should be filled in so that loss of information does not occur if data in a single column are sorted;
- Do not use highlighting, text boxes, shapes, or charts in the spreadsheet containing the data.

Spreadsheets

Spreadsheets are one of most common and simplest ways to enter and store data. They use rows and columns to compose a table that more or less simulate a paper worksheet. Examples of open-source freely available spreadsheets include:

- LibreOffice Calc: https://www.libreoffice.org/
- Googledoc Sheets: https://docs.google.com
- Apache Open Office Sheets: https://www.openoffice.org/download/index.html
- Gnumeric Spreadsheet http://www.gnumeric.org/

The benefits of storing data in spreadsheets include:

- Ability to perform simple calculations, add mathematical formulas to cells, and easily create summary figures such as charts and graphs;
- Easy to use for tracking simple data;
- Allow for flexible cell content type (e.g., numbers and text can be used in the same column).

Some data validation rules can be enforced in spreadsheets through restricting the allowable type of data for a field (text versus numeric), size of information that can be entered into the cell (e.g., specifying the number of decimal places required for a numeric field), or format (e.g., dates, times). For example, the date March 14, 2012 can be represented as:

- 2012-03-14
- 14 March 2012

- 12-03-14
- 14-03-12
- 14/03/12
- 14 March, 2012

Without controlling the format of the values entered into the date field it would be much harder to integrate multiple datasets, even for this field alone. As mentioned above, whenever possible, try to use data standards for each of the fields you are collecting. For example, specifying ISO 8601 format for date (https://www.iso.org/iso-8601-date-and-time-format.html) would let users know that the first option (2012-03-14) was the desired format for date.

Providing users with a list of choices through mechanisms such as pick lists or drop-down menus, is another mechanism to assure data quality in spreadsheets. Dropdown menus reduce entry errors such as misspellings, and demonstrate to users what data you are interested in capturing and how it should be displayed. For example, the list of of non-OIE listed diseases (*http://www.oie.int/wahis_2/public/wahidwild.php/Diseaseinformation/popup/diseaselist*) affecting wild animals could be used to create a diagnosis list in a Google Sheet as illustrated in Figure 4.

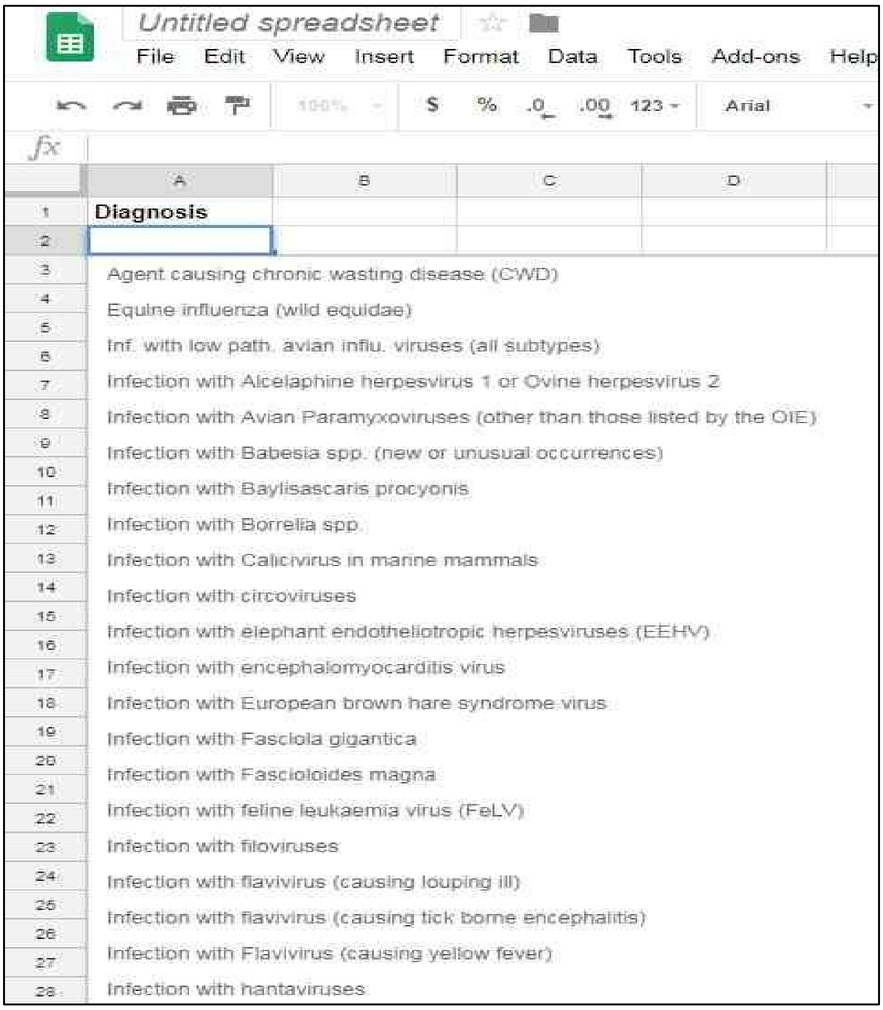


Figure 4. Diagnosis list in Google Sheets.

Instructions on how to create in-cell drop-down lists in various spreadsheets can typically be found online. Below are step-by-step instructions for creating dropdown lists for Google Sheets (from *https://support.google.com/docs/answer/186103?co=GENIE.Platform%3DDesktop&hl=en*):

1. Open a spreadsheet in Google Sheets and select a cell or cells where the drop-down list will be applied.
2. Click **Data** > **Data validation**.
3. Next to "Criteria," choose an option:

   List from a range: allows you to create a list within the spreadsheet and then choose the cells that contain the list to be applied to those cells.

   List of items: allows you to manually enter the items for the list. They should be separated by commas and no spaces.

4. The cells where the list has been applied will have a Down arrow ▾ .
5. If you enter data in a cell that doesn't match an item on the list, you'll see a warning. "Reject input" next to "On invalid data" to prevent users from entering items not on the list.
6. Click **Save**.

Databases

Databases provide a number of benefits over spreadsheets including allowing multiple-user access, and the ability to handle large datasets and difficult data queries. Databases also have many more options for automatic data integrity checks.

Benefits of storing data in a database include:

- Ability to store large quantities of data;
- Ability to perform complex queries and quickly generate summary reports;
- Allowance for multiple users to enter and manipulate data while maintaining data integrity;
- Ability to create user-friendly front ends that provide a clear, uniform way for users to enter data;
- Explicit control over data types (e.g., the type of data to be entered into a field is considered for all data fields during database design) and ability to enforce business rules and validity checks (e.g., counts of affected animals must be positive);
- Ability to define multiple types of relationships between data fields and sets of data. See below for examples of wildlife health data fields that may have different types of relationships within a relational database;
- Ability to create user-friendly forms to make data entry simple for users;
- Ability to incorporate other types of data such as images or document attachments (in some cases).

Examples of free open source database software include:

- CUBRID: *https://www.cubrid.org/*
- Firebird: *http://www.firebirdsql.org/en/start/*
- MariaDB: *https://mariadb.org/*
- MongoDB: *https://www.mongodb.com/*
- MySQL: *https://www.mysql.com/*
- PostGreSQL: *https://www.postgresql.org/*
- SQLite: *https://www.sqlite.org/index.html*

Similar to spreadsheets, relational databases include tables comprised of rows and columns but also allow the user to define the relationships between the tables. An example of the types of relationships between variables that might exist in wildlife health database can be seen in Figure 5 below.

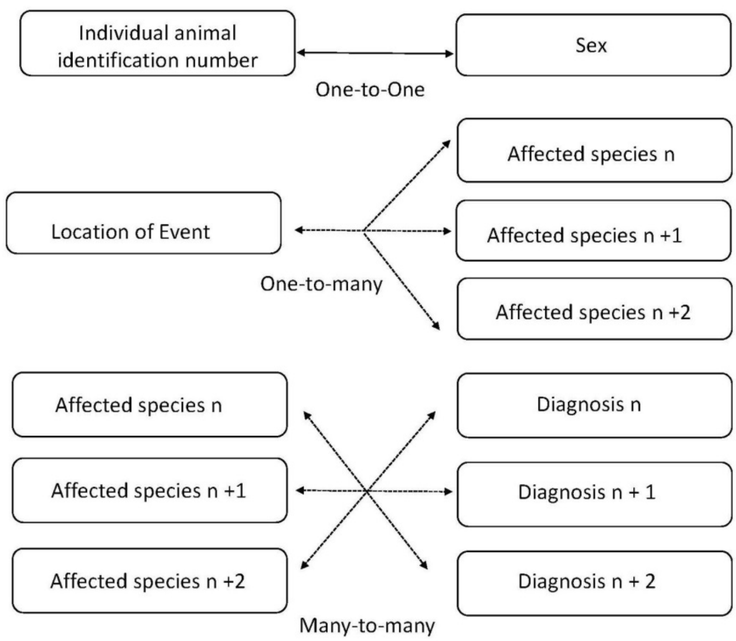


Figure 5. Example of different variable relationships.

Metadata

For data to be useful to others beyond its original collector, they must be accessible, discoverable, and interpretable. All datasets should be stored with metadata, or documentation that describes the data. Metadata help the originator of the data track important details about the data, describe the data to new potential users, and facilitate search and retrieval of the data if deposited in a data repository.

Metadata can be as simple as free text "readme" style files that are stored with the spreadsheet or database. To be most useful, readme files should include the following minimum information (adapted from Cornell Research Data Management Service Group 2018):

- Title of dataset;
- Contact information for person(s) responsible for data collection;
- Date or range of dates on which data were collected;
- Information describing geographic location(s) of data;
- If the metadata describe multiple files or spreadsheets, provide a short description of what data are contained within each filename and the date it was created;
- Description of any restrictions or licenses associated with the data;
- Description of methods used for data collection;
- Description of methods used for data analysis whenever processed data are given;
- List of variables, including spelled-out abbreviations and definitions of column headings;
- Units of measurement;
- Definitions for codes or symbols used to record missing data

There are also a variety of standardized, structured, machine readable formats for metadata. A standardized metadata format may be required for data that will be deposited into a data repository. A metadata standard describes both the content of the information to be included in the metadata as well as the format in which this information should be stored.

- A catalog of metadata standards by scientific discipline can be found at: *http://www.dcc.ac.uk/resources/metadata-standards*
- An example of ecology metadata standards can be found at: *https://knb.ecoinformatics.org/#external//emlparser/docs/index.html*

In conclusion, all steps of the data management life cycle can be made easier and more effective by deliberate consideration of the suggestions outlined in this chapter. This in turn will allow more optimal use of the data and increase the quality and utility of wildlife health research both nationally and internationally. Close attention to such considerations as data standards and definitions, QA/QC, and data storage, while often not the primary interest of the investigators, is critical for the generation and maintenance of high quality, sharable, and actionable wildlife health data.

Literature cited

Briney K. 2015. Data Management for Researchers: Organize, Maintain, and Share your data for Research Success. Exeter: Pelagic Publishing, UK.

Cornell Research Data Management Service Group. 2018. Metadata and describing data. From *https://data.research.cornell.edu/content/writing-metadata*.Accessed April 2018.

DataOne. 2012. DataONE Education Module: Data Quality Control and Assurance. DataONE. From *http://www.dataone.org/sites/all/documents/L05_DataQualityControlAssurance.pptx*. Accessed April 2018.

Faundeen JL, Burley TE, Carlino JA, Govon, DL, Henkel HS, Holl SL, Hutchison VB, Martín E, Montgomery ET, Ladino CC, Tessler S, Zolly LS. 2013. The United States Geological Survey Science Data Lifecycle Model: U.S. Geological Survey Open-File Report 2013–1265, 4 pp.

US Geological Survey. 2018. Data management: data quality. Retrieved April 6, 2018. From *https://www2.usgs.gov/datamanagement/qaqc.php*

White, CL, Forzan, MJ, Pessier, AP, Allender, MC, Ballard, JR, Catenazzi, A, Fenton, H, Martel, A, Pasmans, F, Miller, DL, Ossiboff, RJ, Richgels, KLD, Kerby. JL. 2016. Amphibian: A case definition and diagnostic criteria for *Batrachochytrium salamandrivorans* chytridiomycosis. Herpetological Review. 47: 207-209.

Data Dissemination

In the previous sections, we have discussed motivations for sharing wildlife health information, associated challenges and potential solutions, the importance of proper data curation, and ways to build partner networks to facilitate the collection of this information. In this section, we will provide some tools and guidelines for disseminating wildlife health information to various stakeholder groups. In particular, we will focus on the use of geographic information systems (GIS) and how to use GIS-created maps to effectively communicate intended messages about wildlife health.

Raw Data

One of the most common forms of data dissemination is the sharing of the raw data or wildlife health information. In this context, we define raw data as the information directly collected during an outbreak or during wildlife disease surveillance, and lacks any additional summary or statistical analyses. Often raw data is stored in a table, which may consist of rows of individual animals sampled, and columns of the locations of disease cases or sampled individuals, diagnoses or disease state, animal attributes (e.g., sex, age, etc.), date and time of discovery, and any other additional information of interest to the investigator. *Table* 5 displays an example format of raw wildlife health data.

Table 5: West Nile cases reported to the United States Geological Survey's National Wildlife Health Center (*https://www.nwhc.usgs.gov/whispers/*).

Event ID	Start Date	End Date	Affected	States	Species	Event Diagnosis
160485	8/3/2017	8/15/2017	11	WI	American Crow, Great Horned Owl, Cooper's Hawk, Red-tailed Hawk	West Nile Virus
160521	8/1/2017	10/31/2017	5	MI	Ruffed Grouse	West Nile Virus

The raw data may be disseminated to interested parties directly in this table format. But, as previously discussed, to ensure the end user can successfully access the information, the data should be maintained in a machine-readable format such .csv files with appropriate metadata. Additionally, often data is shared to facilitate statistical summaries and analyses, and maintaining the data in machine-readable formats makes the data readily available to the statistical programs (e.g., *https://www.r-project.org/*) for accomplishing these tasks.

The raw data may also need to be disseminated and transferred manually into other databases, in which it is standardized, organized and made available to interested parties. This may be the case when sharing data between agencies within or between jurisdictions, or when sharing with international animal health organizations such as the OIE. We will discuss how to share wildlife health information with the OIE via the WAHIS Wildlife interface (http://www.oie.int/animal-health-in-the-world/wahis-wild-interface/) later in this course, and provide hands-on opportunities in this regard. Sharing wildlife health information as raw data, is the simplest way to disseminate data because it requires no additional effort beyond that expended for data collection and curation; however, beyond scientists or agencies that want access to the raw data for conducting analyses, generally sharing wildlife health information in this format is undesirable.

Other Formats

There are a variety of different formats, in addition to sharing the raw data, that can be used to disseminate wildlife health information. During the 4th Cycle Training for National Focal Points for Wildlife Training, different formats were discussed in terms of outbreak investigations. It was mentioned that the most appropriate format is highly dependent on the situation and audience. For example, when disseminating information on particular diseases to partners within a network or other agency scientists, a scientific report detailing the relevant background, methods, findings and implications may be useful. Often these reports will contain graphs, charts and tables that succinctly summarize the information. An example of a simple report that was disseminated to partners within a network can be found at https://pubs.er.usgs.gov/publication/sir20125271.

In contrast, to effectively communicate wildlife health information, to stakeholders without scientific training, such as the general public or decision-makers, it is often more effective to distill the information within a scientific report into a short document or "Fact Sheet", which captures the most salient features of the surveillance efforts or the outbreak investigation in a simple and direct manner. These documents should be worded for non-scientific audience, and be visually attractive with interesting pictures, charts and graphs largely conveying the intended message rather than text. These documents should focus less on the methods, and mainly on the findings and implications. Examples of this format can be found at https://www.nwhc.usgs.gov/publications/fact_sheets/.

These are just a few examples of formats that can be used to disseminate data. There are numerous other methods that could also be implemented; however, detailing each of these potential formats and their strengths and weaknesses is beyond the scope of this workbook. Rather, we will focus the remainder of this workbook on describing one of the most useful tools for presenting wildlife health information: a map. Maps allow the creator to present to the viewer, the spatial and temporal context of the wildlife health information along with other important features on the landscape that may be directly impacted. Maps also provide useful tools for planning disease response efforts, developing hypotheses regarding causative agents, factors affecting growth and spread rates, and identifying important stakeholders. The remaining sections will focus on the use of maps as tools for communicating wildlife health information.

Mapping Wildlife Health Information

In the 4th Cycle Training of OIE National Focal Points for Wildlife, the use of Google Earth Pro (*https://earth.google.com/download-earth.html*) was presented as a simple and easy to use mapping tool for wildlife outbreak data. This remains an easy-to-use and effective option for quick mapping; however, we will focus on the use of GIS as a more sophisticated, scientific platform for generating maps of wildlife health information. We will provide a general introduction to GIS mapping in this training; however, to become proficient in GIS we recommend taking additional trainings, exploiting on-line resources, and/or acquiring one of numerous texts available on the subject.

There are a number of software platforms that provide the tools and framework for creating a GIS. ESRI (*www.esri.com*) is a leading company that produces many GIS products that are used globally, and provide a wide range of functionality; however, licensing expenses can be high. There are also open-source software for GIS construction. QGIS (*www.qgis.org/*) is one example of these types of software that is widely used and provides much of the capability of the more expensive ESRI products; however, there are many open-source GIS applications with some excelling for specific applications. An Internet search is a good place to begin to evaluate the suite of potential tools available, and obtain instructional materials.

What is a GIS? A GIS is a system designed to capture, store, manipulate, analyze, manage, and present all types of spatial or geographical datafor the purpose of understanding patterns and relationships (*https://en.wikipedia.org/wiki/Geographic_information_system*). A simple heuristic description of a GIS is a layering of useful information that when viewed in concert provides new insights not readily distinguishable before. Figure 6 below provides a graphic describing a simple GIS constructed to map wildlife disease outbreak data. This figure highlights the real strength of a GIS is the linking useful spatial or spatio-temporal data layers together. There is an ever-increasing amount of georeferenced data becoming readily available for use in capturing the spatial and temporal context of wildlife health information. A GIS provides the means to harness this wealth of information in an efficient and rigorous way.

Installing QGIS

For this training, we will demonstrate the use of QGIS. QGIS can be downloaded from the following url: *https://www.qgis.org/*, and is available for Linux, Mac and Windows operating systems. Follow the installation instructions to install the GIS platform on your local machine. We will use the Windows version for this training. Readers using other operating systems are referred to the QGIS documentation at *https://www.qgis.org/en/docs/index.html* for the relevant instructions for their operating systems.

We will only minimally cover the capabilities of QGIS in this training, and readers interested in continued learning can access the documentation for QGIS at *https://www.qgis.org/en/docs/index.html*. Additionally, there are numerous resources available on the internet that can be easily found using common internet search engines. Readers who are already well versed in GIS applications can skip to the "Better Map Design" portion of this section.

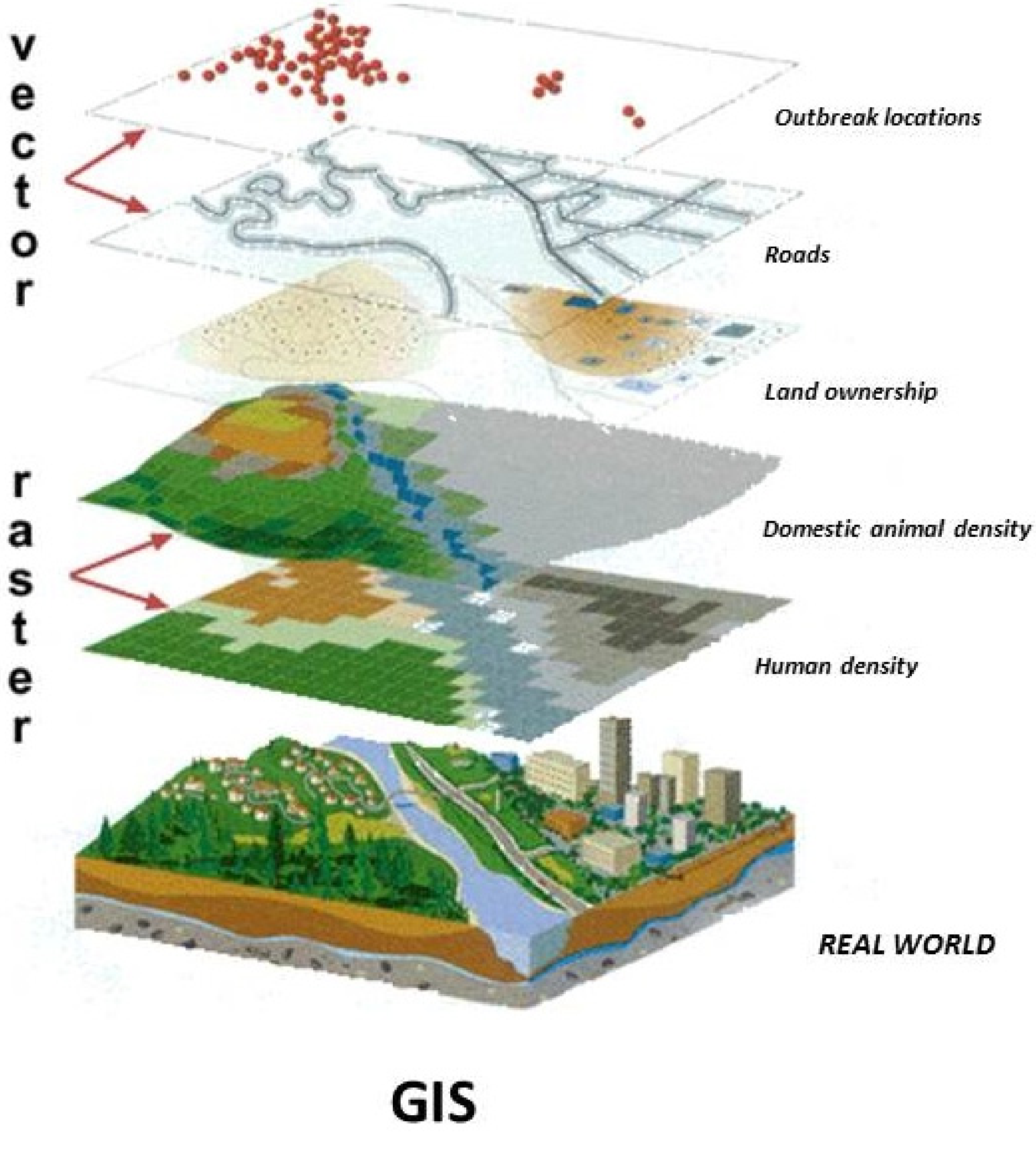
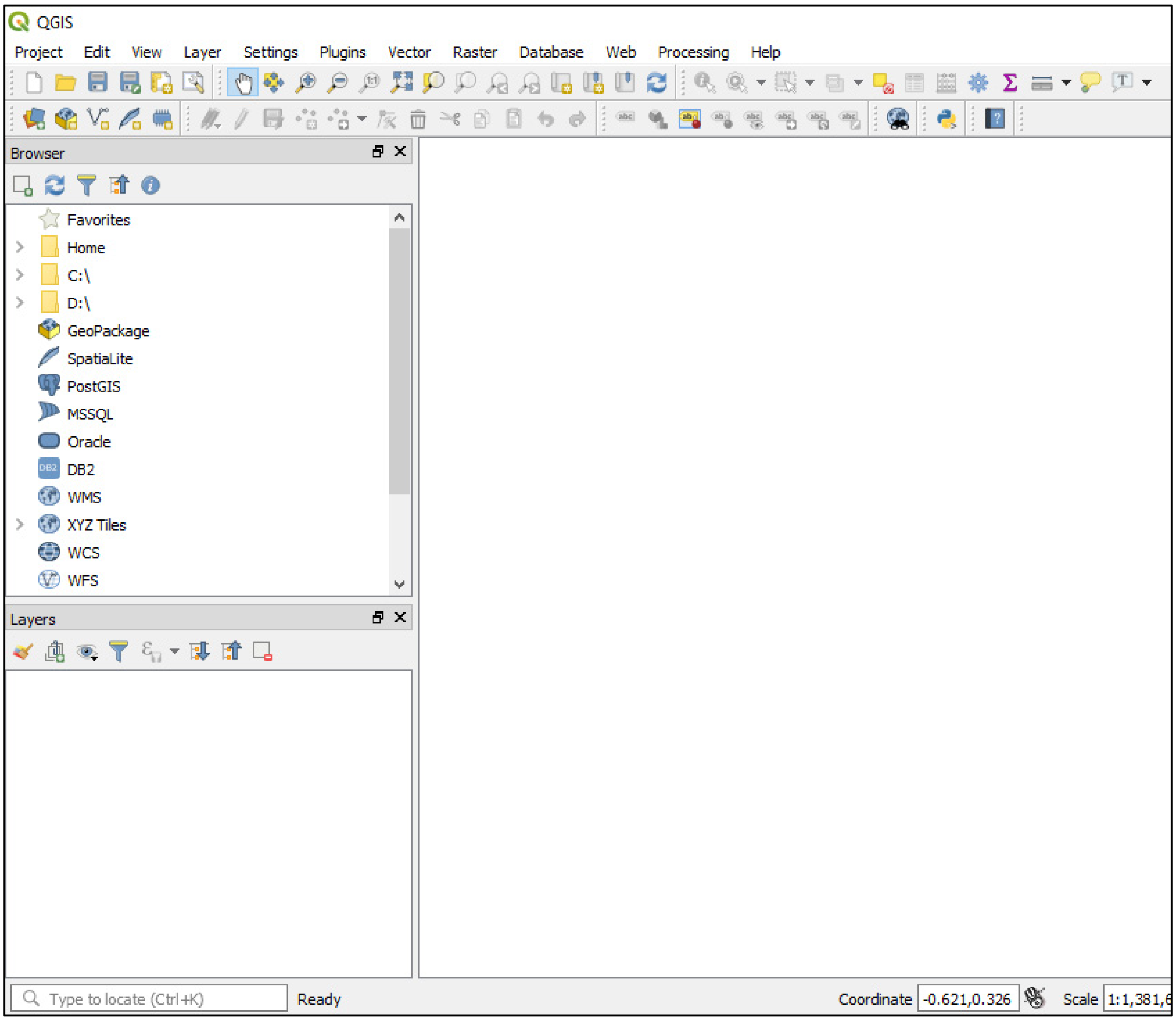


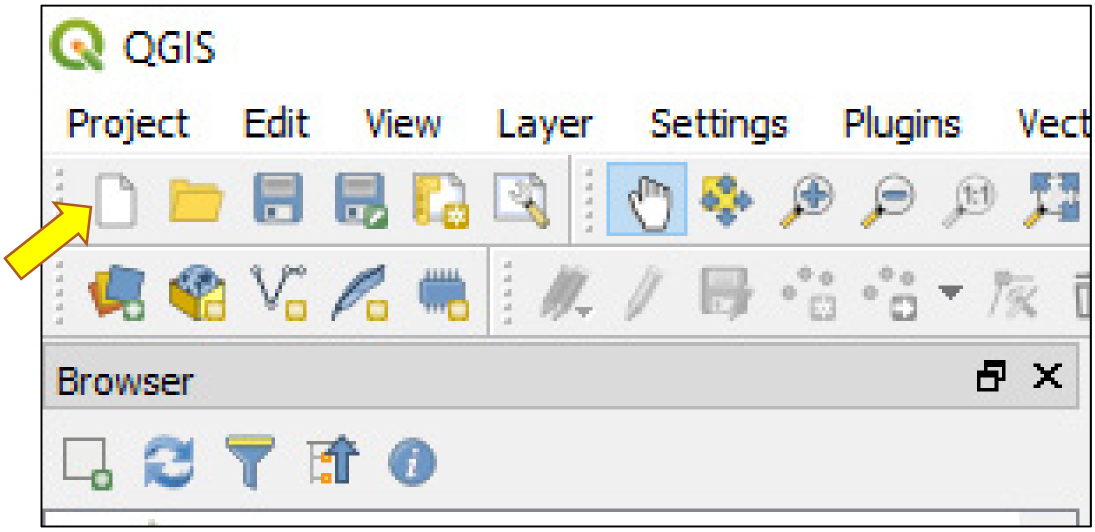
Figure 6. Example of a GIS for a wildlife disease outbreak
(Source: http://doc.arcgis.com/en/arcgis-online/reference/geo-info.htm).

Creating a GIS

To begin, open the QGIS application from where it is installed on your local computer. Once open the following screen should appear.



To start a new GIS project click on the "blank page" icon or left click on the "Project" menu and select "new".

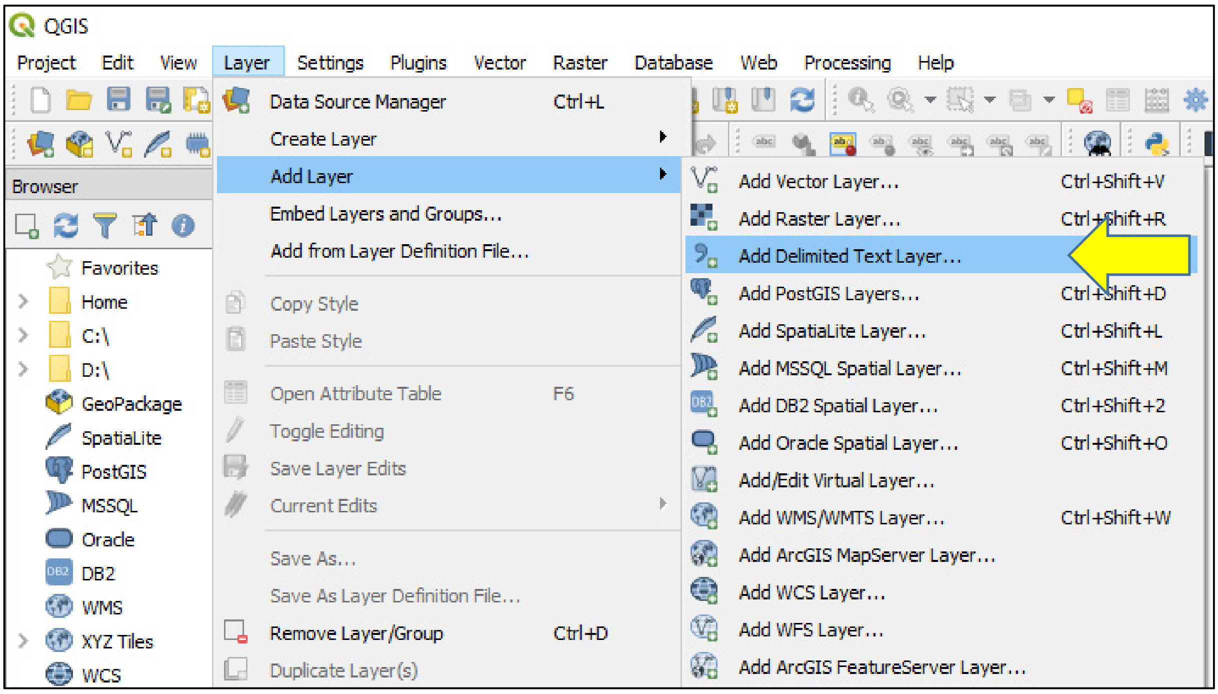


This will open a blank project. To illustrate the creation of a GIS for wildlife health information commonly encountered during wildlife disease outbreaks or surveillance, we will assume that we have our raw wildlife health information in a .csv file. Specifically, we will be examining an example *Table 6* containing wildlife cases in the state of Illinois in Midwestern United States. This dataset consists of 13 cases of striped skunk (*Mephitis mephitis*) or raccoons (*Procyon lotor*) reported to the wildlife health agency as part of a suspected rabies outbreak in the region.

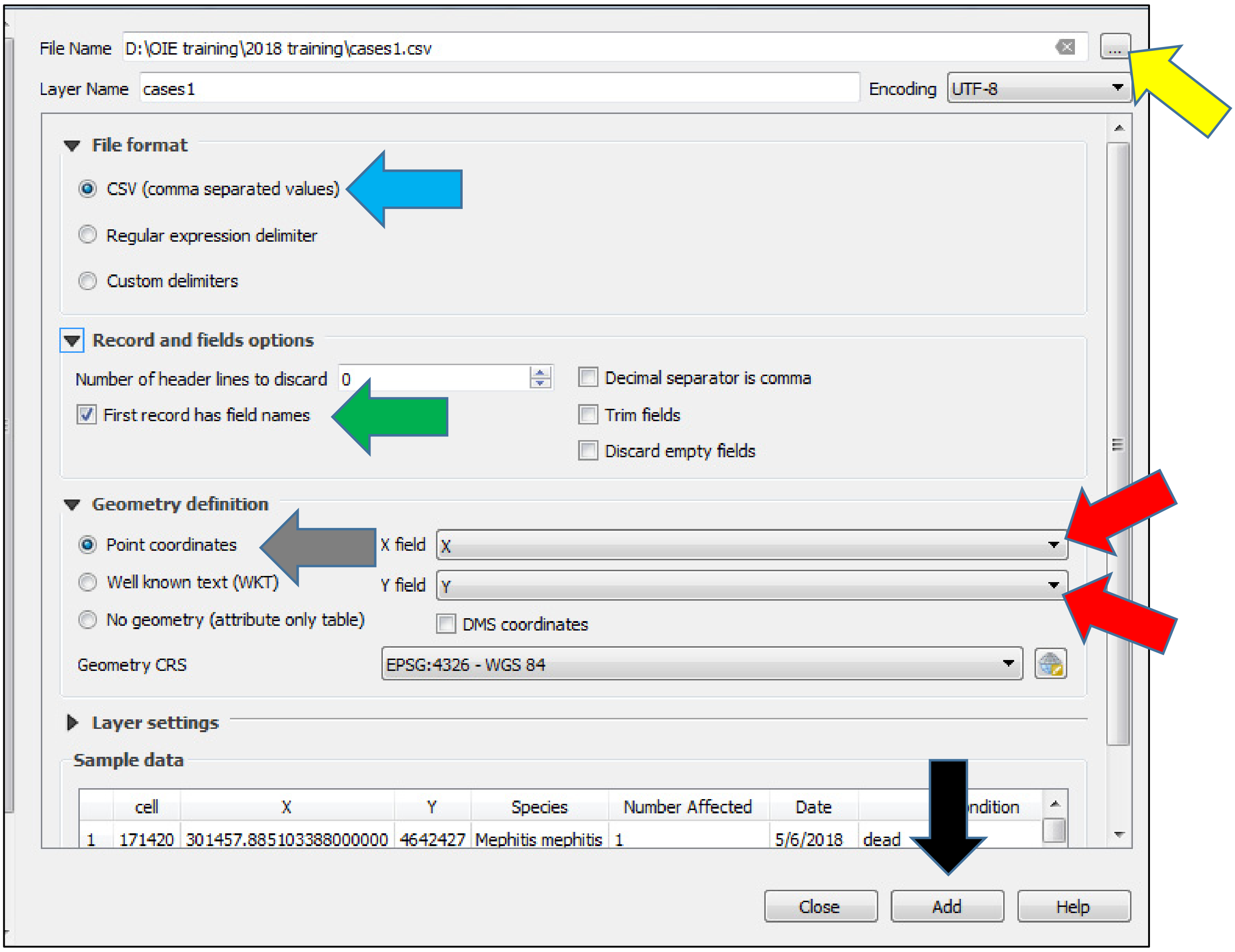
Table 6: Example dataset of reported to the wildlife health agency.

X	Y	Species	Number Affected	Date	Condition
301458	4642427	*Mephitis mephitis*	1	5/6/2018	dead
300956	4640438	*Procyon lotor*	1	5/30/2018	neurologic signs
300564	4640578	*Mephitis mephitis*	1	4/30/2018	aggressive behavior/euthanized
303228	4639932	*Mephitis mephitis*	1	4/12/2018	neurologic signs
300289	4640485	*Mephitis mephitis*	1	4/23/2018	road-kill
296153	4638906	*Mephitis mephitis*	2	4/13/2018	dead
298286	4640684	*Mephitis mephitis*	1	4/15/2018	dead
300390	4640618	*Mephitis mephitis*	1	4/18/2018	dead
300244	4641372	*Procyon lotor*	1	4/26/2018	neurologic signs
299496	4640230	*Mephitis mephitis*	1	4/20/2018	dead
300936	4641733	*Procyon lotor*	1	5/2/2018	dead
302551	4644136	*Mephitis mephitis*	1	5/4/2018	aggressive behavior/euthanized
297002	4643303	*Mephitis mephitis*	1	5/15/2018	neurologic signs

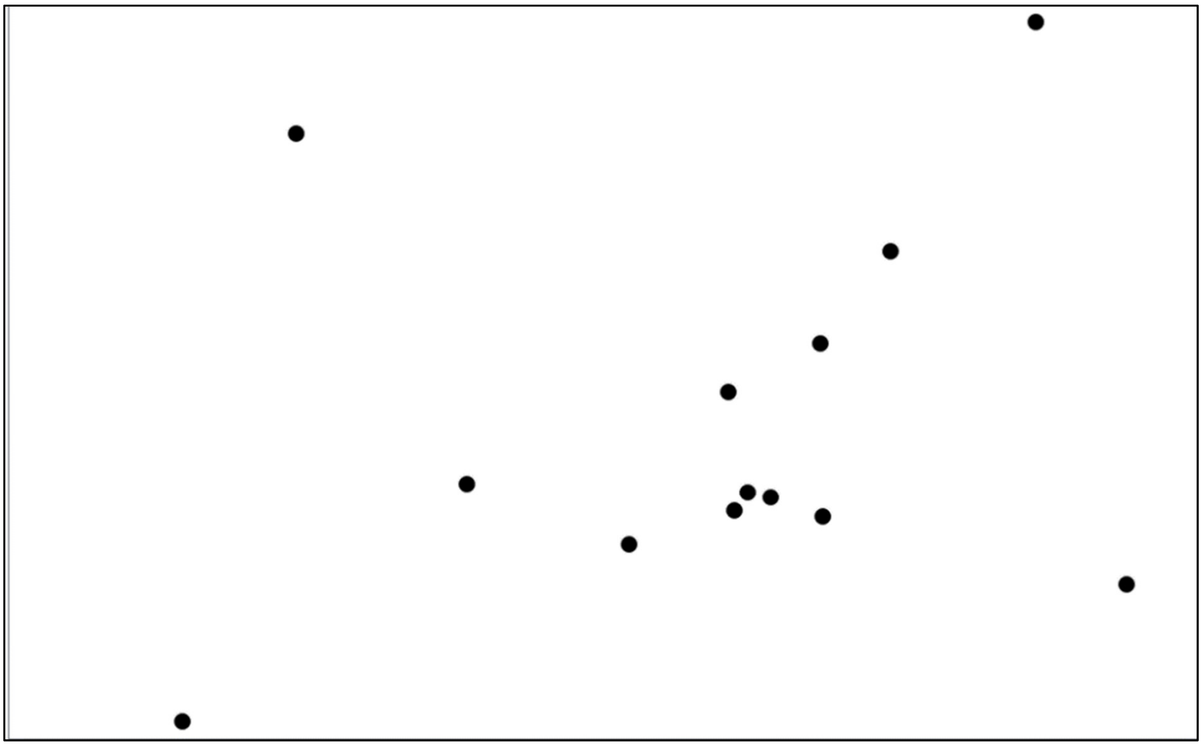
We will import this text file into QGIS. To accomplish this task, we click on the "Layer" menu, then on the "Add Layer" submenu, and finally on the "Add Delimited Text Layer…" (yellow arrow).



This will open the following dialog window. To import your file, left click on the "..." button in top right corner (yellow arrow). This will open a standard file explorer window that you can use to navigate to the folder of interest and select the appropriate file (e.g., cases1.csv in this example). Next, you can check that the file format is correct, by clicking on the black triangle next to "File Format". We can see that QGIS recognized the file extension and assigned it as a .csv file (blue arrow). Next we click the black triangle next to "Record and field options", and make sure that QGIS recognizes we have a header row (i.e., the box "First record has field names is checked; green arrow"), or in other words the first row of our .csv file contains the column headings. Then we click the black triangle next to geometry definition, and we check the appropriate type of geometry is selected. In this case we want to create a point file (gray arrow). We also check that QGIS recognizes the correct fields from our table for the X and Y coordinates (red arrows). In our example we have fields named "X" and "Y", which QGIS recognizes automatically, but if we have a different name for these fields in our .csv file, we would have to specify the appropriate field names for the X and Y location values. Finally, we can click the "Add" button (black arrow), and QGIS will create a new point layer using our .csv file.



The result of this operation is shown below; the locations of the animals reported in this outbreak are now displayed.

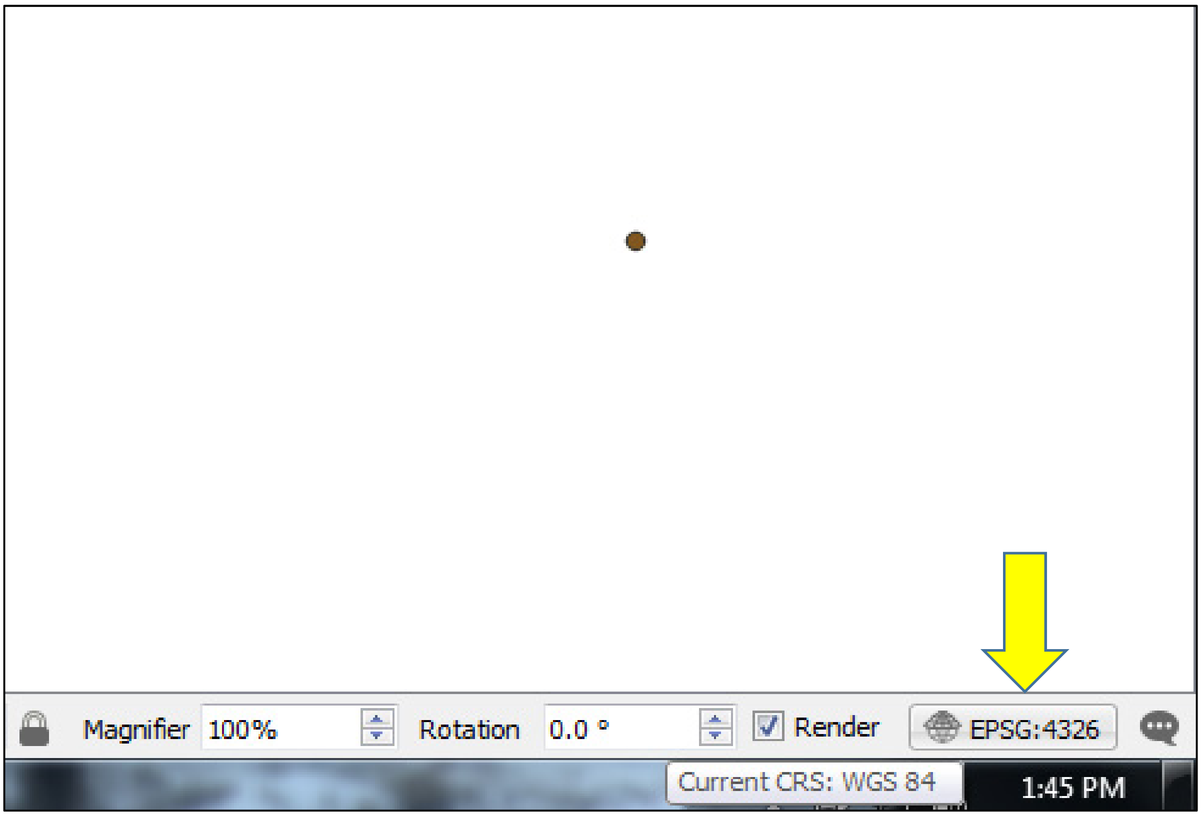


Coordinate reference systems

The power of GIS is the ability to link different data layers together; however, to achieve this linkage it is critical that the geographic coordinate reference system of each layer is known. A geographic coordinate reference system (CRS) is a mathematical representation of the Earth that allows for any location to be designated by a set of numbers (*https://en.wikipedia.org/wiki/Geographic_coordinate_system*). This is essentially how a GIS program knows where to place objects on the map. An example of a commonly used coordinate reference systems is the geographic system, which uses latitude and longitude to specify locations.

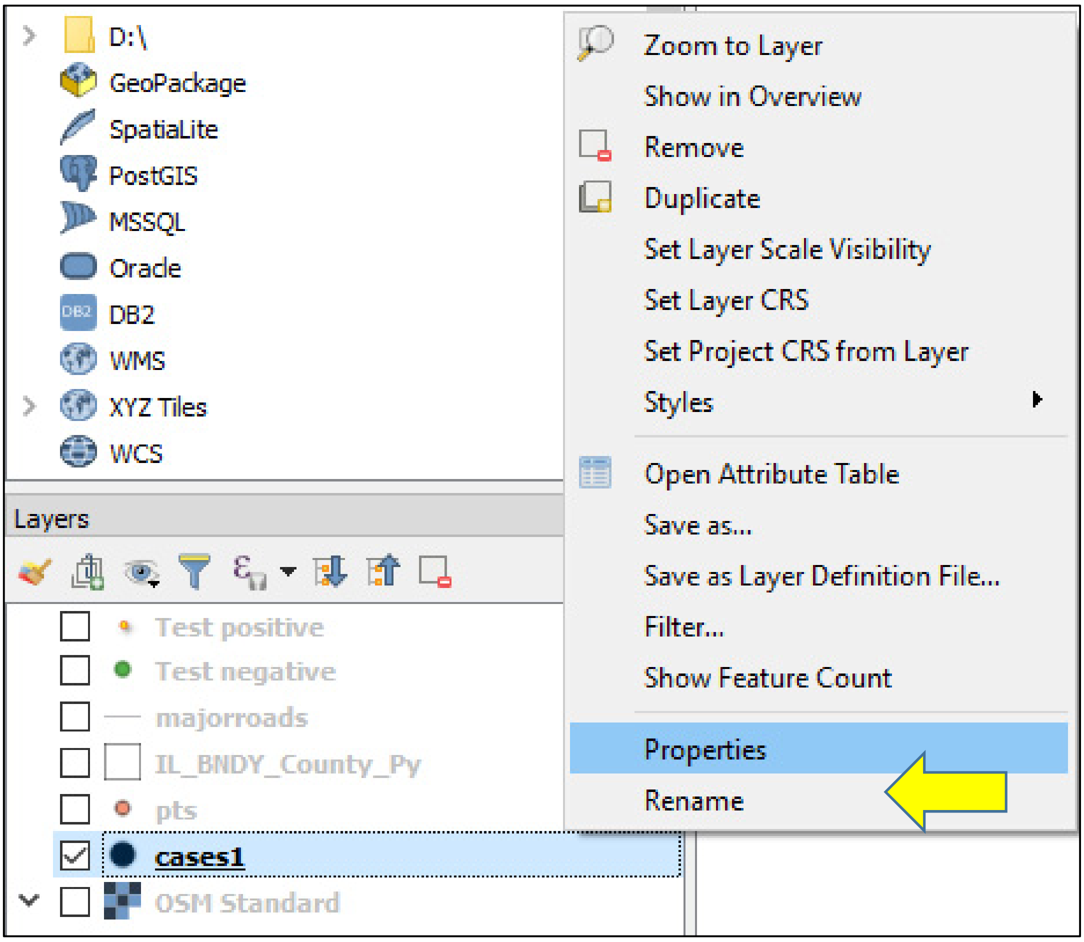
Another important component of the CRS is whether the GIS layer is projected or not. Map projections are mathematical models that transform the 3-dimensional surface of the Earth into a flat 2-dimensional map. Map projections are necessary because it is impossible to project a 3-dimensional surface (i.e., the Earth) onto a 2-dimensional surface (i.e., a map) and maintain the area, angles and distances accurately. All map projections create some level of distortion, and understanding the type of distortion for the area that is being mapped is critical. For example, some projections such as "equal-area" projections maintain the local areal relationships, whereas "conformal" projections maintain local angles, and "equidistant" projections maintain distance and directional relationships. The map creator needs to decide which projection best suits their needs for the locations of interest. There are numerous map projections available, and a list of potential projections are provided at the following url: (http://www.epsg.org/). This list is maintained by the European Petroleum Survey Group (EPSG), and each projection/coordinate reference system is given a unique number. This EPSG number is often used in GIS systems to assign or track the CRS associated with each layer. In short, the quality of a map projection for a particular task varies by location on the Earth and the cartographer's goals.

In general, if you are downloading GIS data layers from a reputable source, the CRS is often assigned to the layer, and is detailed in the metadata that accompanies the layer. However, if you are creating a new GIS layer, as we are in our example, we need to manually specify the CRS because QGIS will automatically set it to the project's CRS if we do not specify it when we create the layer. We did not set a CRS when we created our layer. Thus, we must determine the current project CRS by using the mouse to hover the point over the EPSG text in the lower right portion of the screen.

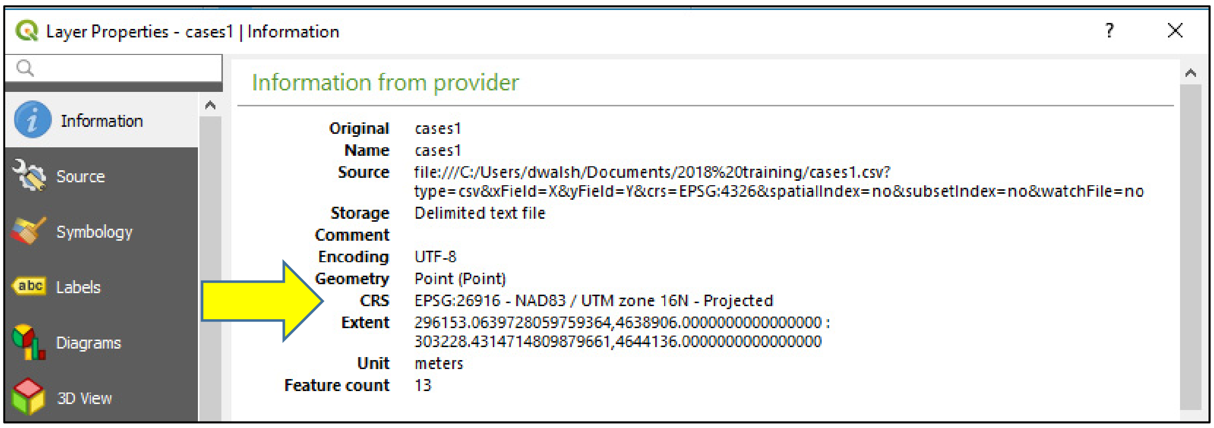


This tells us that the currently the project in our example is displayed in EPSG: 4326, or in other words the WGS 84 CRS (*https://en.wikipedia.org/wiki/World_Geodetic_System*). The project CRS is the display CRS, and all layers that are in this GIS project will be projected into this project CRS. It is necessary to have a single project CRS so that QGIS can align the various spatial data layers, each with a potentially different CRS, into one map. The project CRS is also the CRS that each layer will be assigned when it is created, if it is not specified automatically or by the user during layer creation. Left clicking on the "EPSG:4326" button allows the user to set their project CRS, it can also be set based on the CRS of a layer, which we describe later.

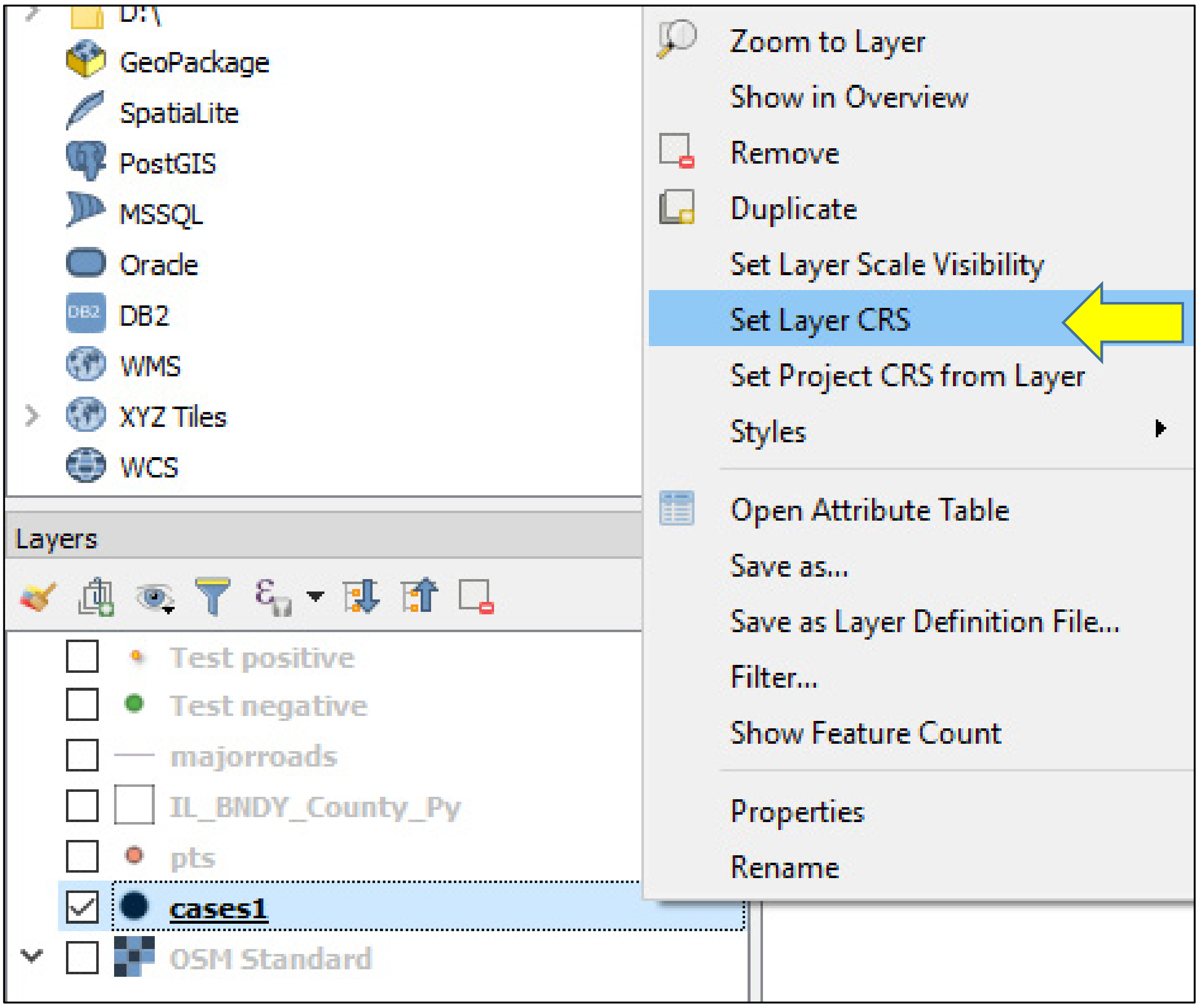
Now let's determine the CRS for the layer we just created, knowing ahead of time the data was collected in the Universal Transverse Mercator projection (UTM; *https://en.wikipedia.org/wiki/Universal_Transverse_Mercator_coordinate_system*). We can do that by going to the "Layers" pane in the bottom left of the screen, and right clicking on our layer, which is entitled "cases1". This brings up a pop-up menu, and we scroll down to "Properties", and left click.



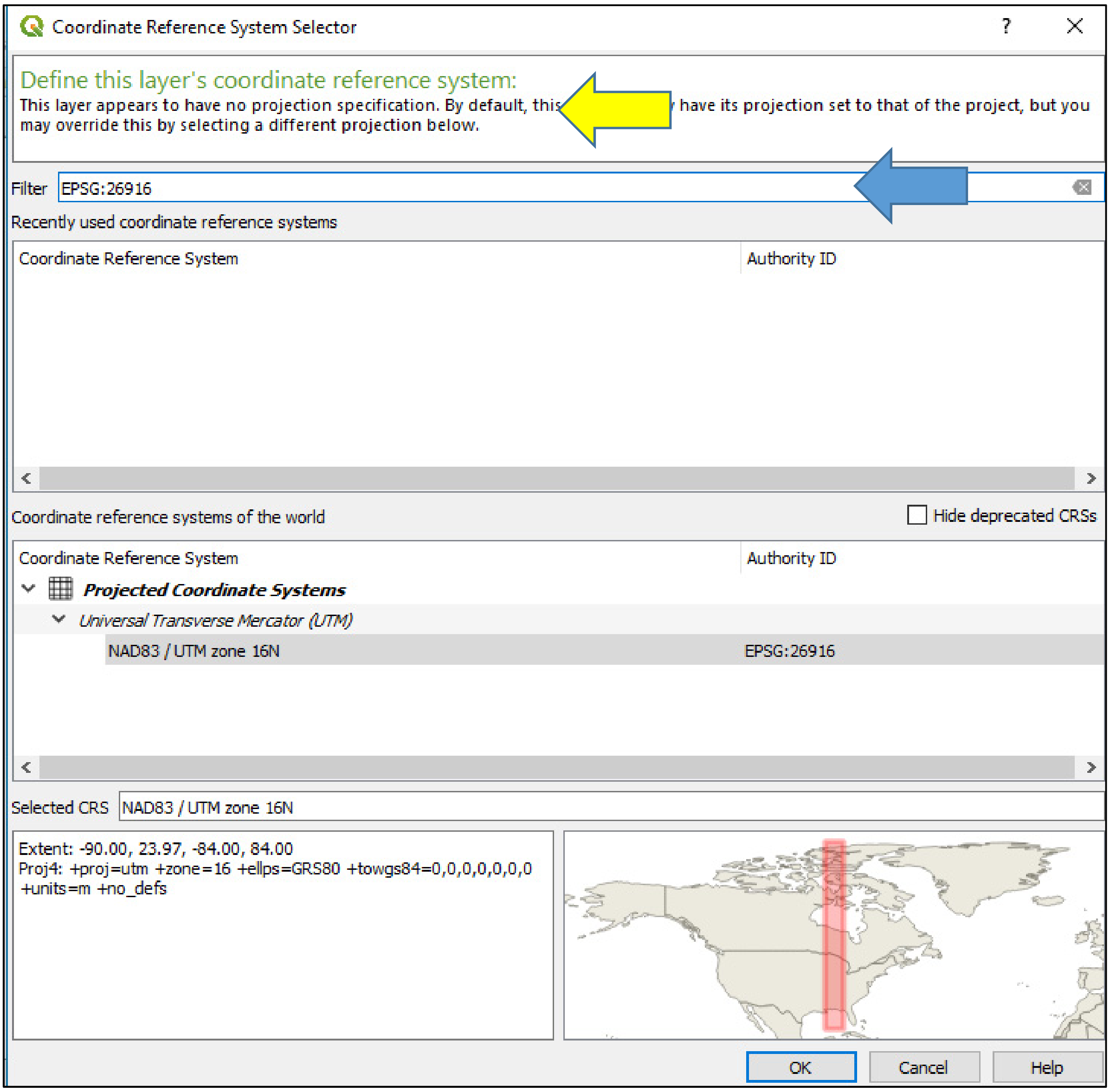
We then left click on "Information", and look to see what the current CRS is for the "cases1" layer (yellow arrow). It is clear that is indeed specified as the project CRS, EPSG:4326.



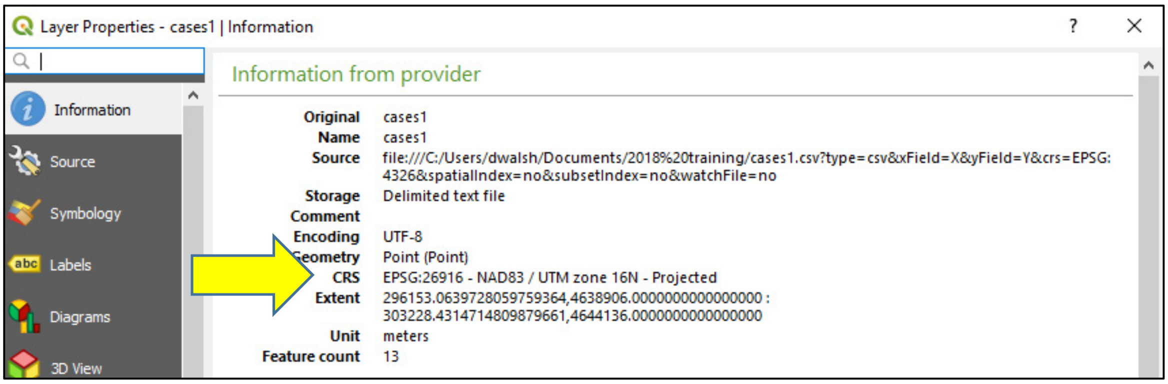
Thus, we need to reset the CRS for this layer to the correct CRS, NAD83/ UTM 16N, otherwise the location data will not be placed in its proper spatial location. To do this, we once again right click on the "cases1" layer in the "Layers" pane in the lower left corner of the screen. This brings up a pop-up menu, and we scroll down to "Set Layer CRS", and click it.



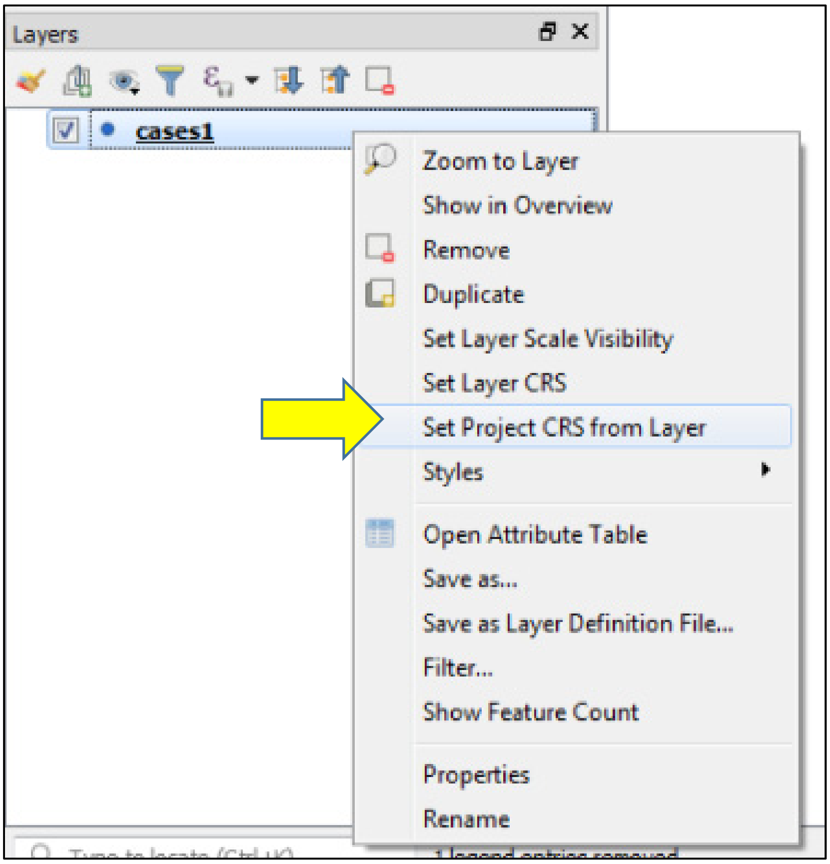
The following dialog box pops up, which tells us that there is no coordinate reference system currently set for this layer (yellow arrow). Therefore, we need to set the layer's coordinate system.



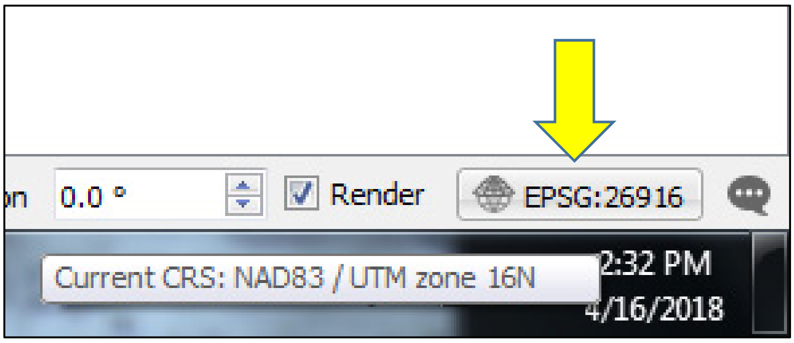
We set the CRS by typing in the name or EPSG number, in this case EPSG:26916 (i.e., NAD83/UTM zone 16N) into the "Filter" line (blue arrow; above), and then click ok. This has now set the CRS for this layer. We can check this is the case by right clicking on our layer, "cases1", in the lower left corner "Layers" pane, and then clicking on "Properties". This displays all of the properties of the "cases1" layer. If we click on "information", we can then see all the metadata associated with this layer, including the CRS. It has now been changed to the correct CRS.



QGIS also allows you to use a layer's CRS to set the CRS for the entire project. This is accomplished by right clicking on the name of the layer you want to use to set the project's CRS in the "Layers" pane in the lower left corner of the screen. This brings up a dialog box. Next click on "Set Project CRS from Layer CRS". This will take the CRS from the layer and assign it as the project's CRS.



Now the project CRS is set to NAD83/UTM 16N, and all layers that created or added will be displayed in this CRS. We can confirm that this is the project CRS by looking at the CRS button down in the bottom right corner of the screen (yellow arrow).

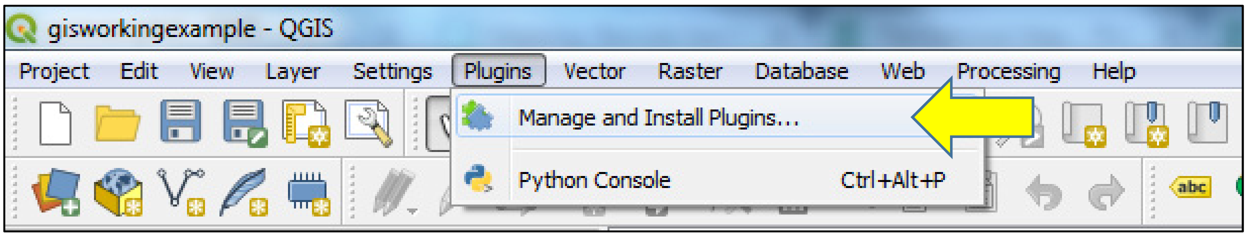


It is indeed been changed to EPSG:26916 as we intended.

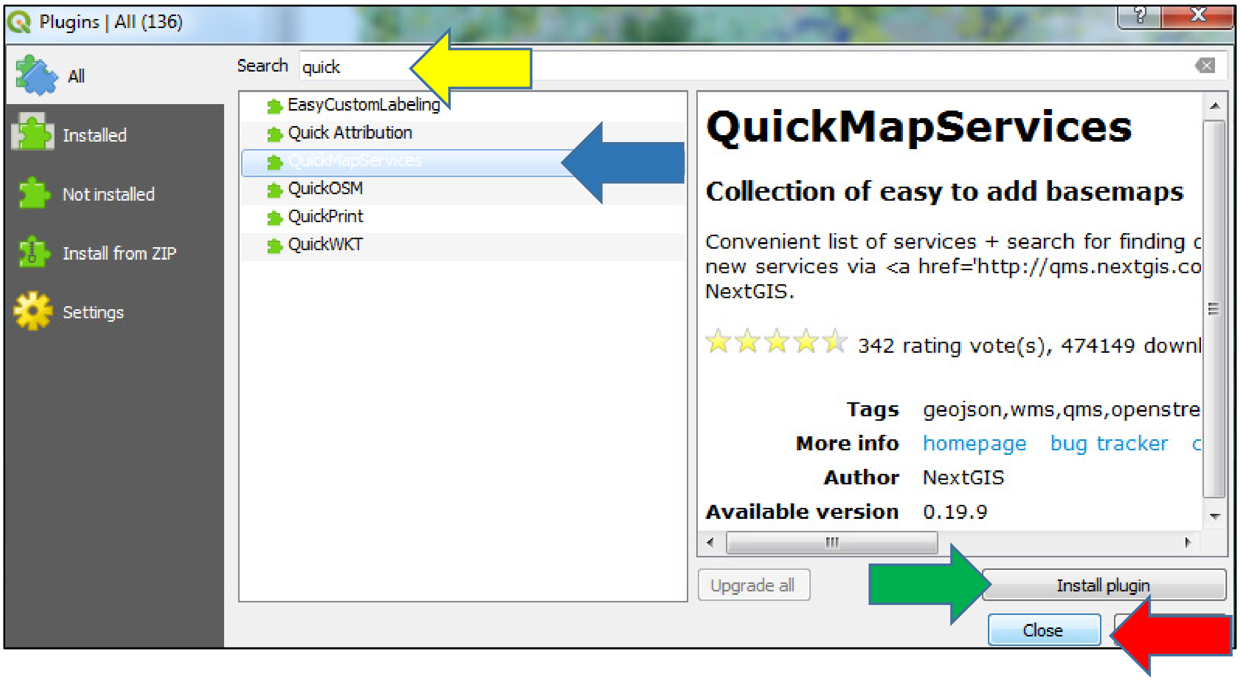
Adding Existing Layers

Now that we have assigned the coordinate system, we can start to explore the real power of a GIS…the ability to link different spatial datasets. These datasets may be other layers that we create manually, but most often they are created by other GIS users. A general internet search for GIS data from reputable sources (e.g., often government agencies) for your area of interest is useful to understand what layers are available for your project. For our example, we need to add some visual context to our wildlife health information because right now our map provides the spatial layout of our disease cases, but does not give us any other information. In GIS, we often start by adding a "base layer". This layer provides a background that helps the map viewer develop a better understanding of spatial context of the data pictured. Depending on the scale of your map, this may be an aerial photo, a topographic map or some other common spatial data layer that helps the viewer visualize the context of the wildlife health information. We will add a layer to our map of the infrastructure characteristics of our region of interest. To add a new layer to our existing map project, we will demonstrate how to install and use a "plugin" in QGIS. Specifically, we will use the "QuickMapServices" plugin. Plugins in QGIS are additional programs or add-on features for QGIS that a user can enable to extend the capabilities of QGIS. The library of plugins can be found at the following url: *https://plugins.qgis.org/*.

Note the following procedure is applicable for installing any plugin; however, we will demonstrate the method using the QuickMapServices plugin. We start by left clicking on the "Plugins" Menu at the top of the screen. Then we select "Manage and Install Plugins…".



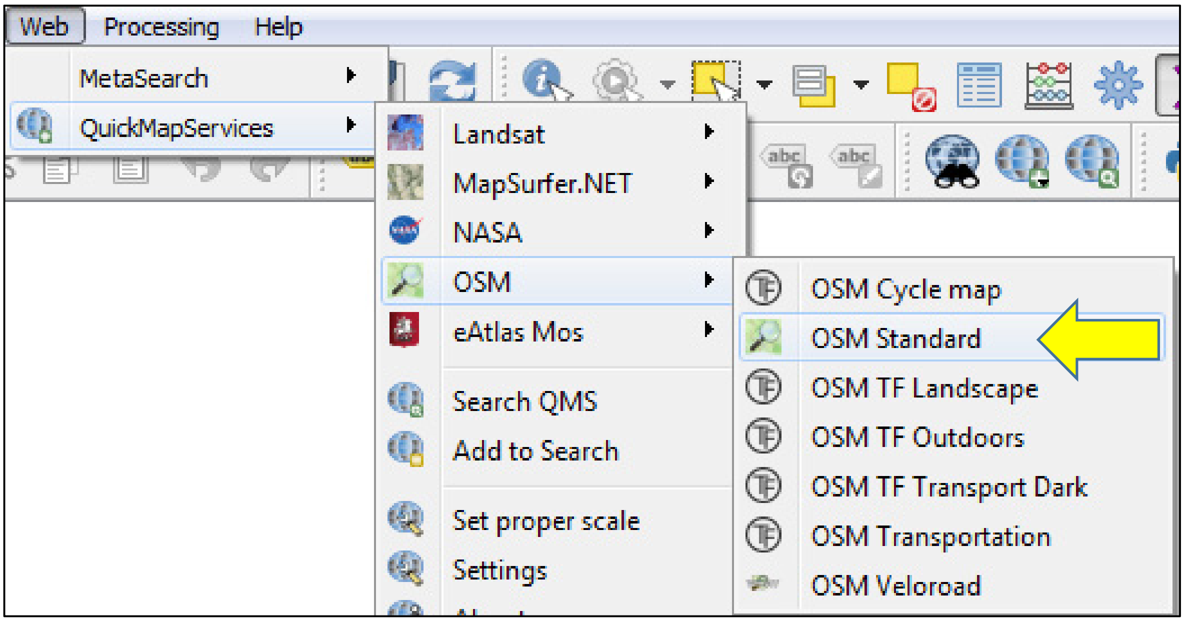
This opens the following menu.



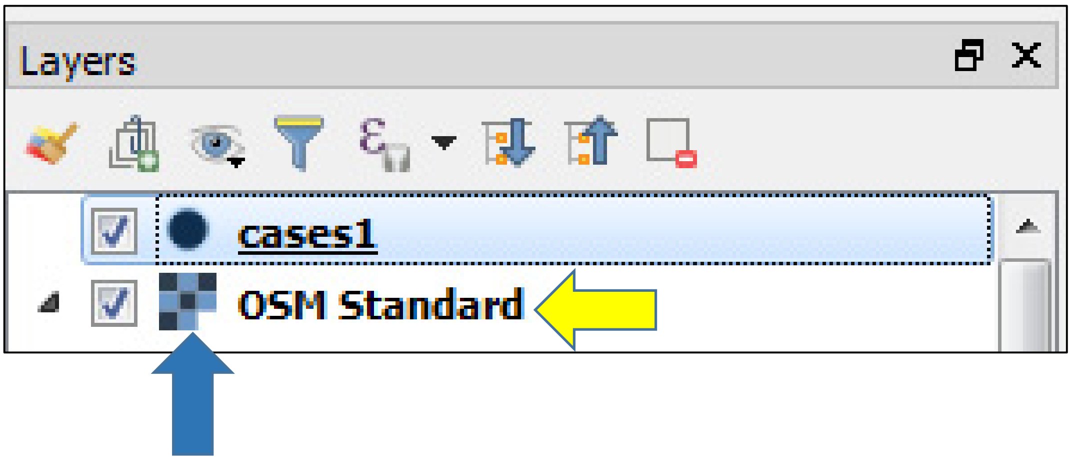
To find the plugin, type "quick" into the search box at the top (yellow arrow). This brings up a list of plugins with "quick" in the name. We scroll down and left click on QuickMapServices (blue arrow). This brings up a description, help links, etc. associated with the plugin. This is indeed the plugin we would like to install, and so we click on "Install plugin" (green arrow). This will install the plugin. Once it has installed we can click the close button (red arrow).

We can now use the QuickMapServices to add a base map layer to our project from a number of different GIS web services (*Note: This plugin requires an internet connection*). To add the base layer, we click on the "Web" menu at the top of the screen. We then scroll down to QuickMapServices, and left click. Next we scroll down to OSM (i.e., Open Street Map; *https://www.openstreetmap.org/*), which provides street maps of the world under an open license agreement, and left click. Finally, we left click on OSM standard (yellow arrow). This will add the "standard" open street map to our project. It is worth noting there a number of different types of base maps that could be added using this process. It may be helpful to experiment with adding various base maps to determine which base map is most suitable for a particular project.

Note, we are demonstrating the base services associated with the QuickMapServices plugin. Additional, base maps are available via contributed services. Information on how to access these services can be found at the following url: http://nextgis.com/blog/quickmapservices-with-contributed-services/. We focus solely on the base map services because they are validated by the plugin authors.



We can see on the map displayed on the following page (note: displayed in landscape view) that our disease cases run along the Rock River, and next to Illinois Highway 2. Also, note that we can now see that a second layer named "OSM Standard" (yellow arrow) has been added in the "Layers" pane on the bottom left of the screen. It also tell us that it is a Raster file as designated by the "checker board" of blue squares next to the layer name (e.g., blue arrow).

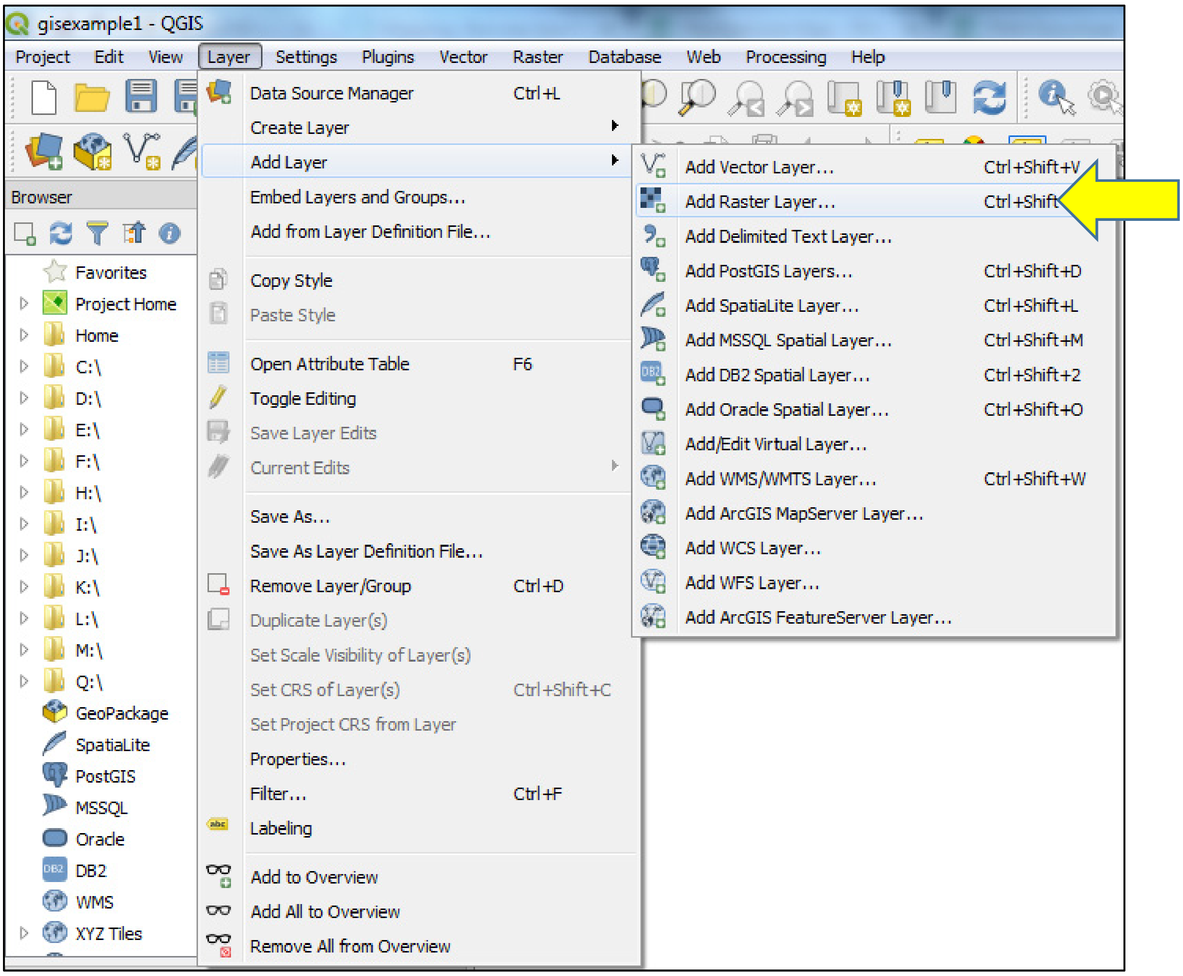



Flagg Road
Lost Nation Road
Lost Nation Road
IL 2
Rock River
Grand Detour
IL 2
Rock River
Rock River
White Oak Lane

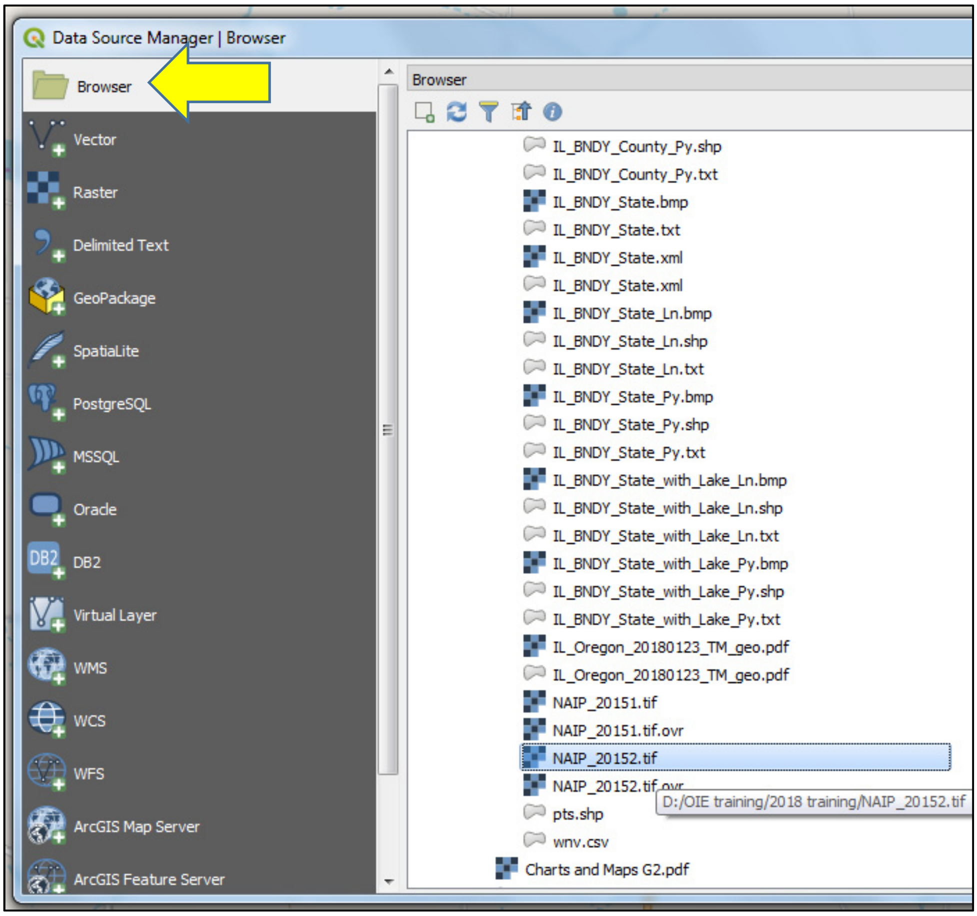
A raster is a type of image that is composed of rows and columns of square blocks or pixels that each are given a specific value(s). When you zoom in close on a raster image, you can see these pixels and the image becomes "pixelated". Many GIS datasets are stored in raster format, particularly large images (e.g., satellite imagery). The resolution of a raster image is the length of a side of one of the pixels composing the image in the units of the map projection (e.g., meters). We will show how to determine the resolution below.

Now that we have added a base map layer, we can create and readily share a map of our wildlife health information that an average viewer could readily interpret. Our data has been given spatial context by the named roads and water features provided in the OSM base map. We now truly have a GIS project because we have linked multiple spatial data layers.

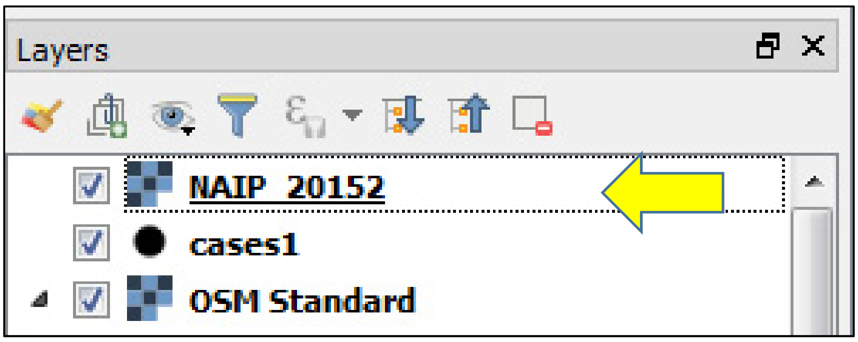
Perhaps, however, we would like to have a more in-depth understanding of the characteristics of the area surrounding our disease cases. Another useful layer we could add would be an orthophoto of the site. An orthophoto is an aerial image that has been geometrically corrected to remove distortions associated with taking a picture from aerial platform flying above the Earth. I have previously downloaded an orthophoto image, which is a raster layer, and I will now add it to our project. To do this, we will left click on the "Layer" menu at the top of the screen. We will then scroll down to "Raster" layer, and then left click on "Add Raster Layer…".

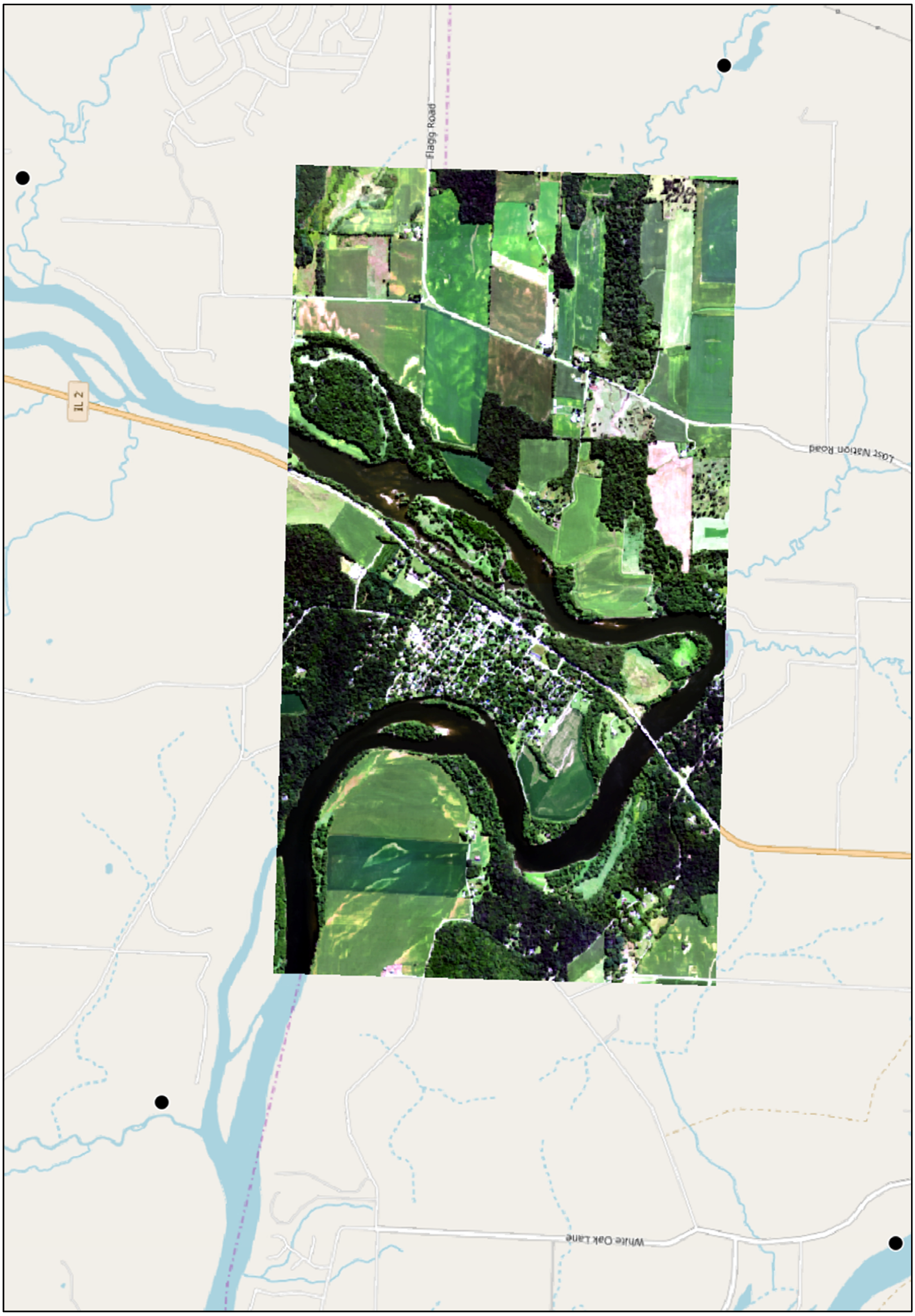


This will open the window below. Left click on "Browser" (yellow arrow), and then navigate to the folder containing the file of interest. Double left click on the file. It will be added to the project. Then close the window.



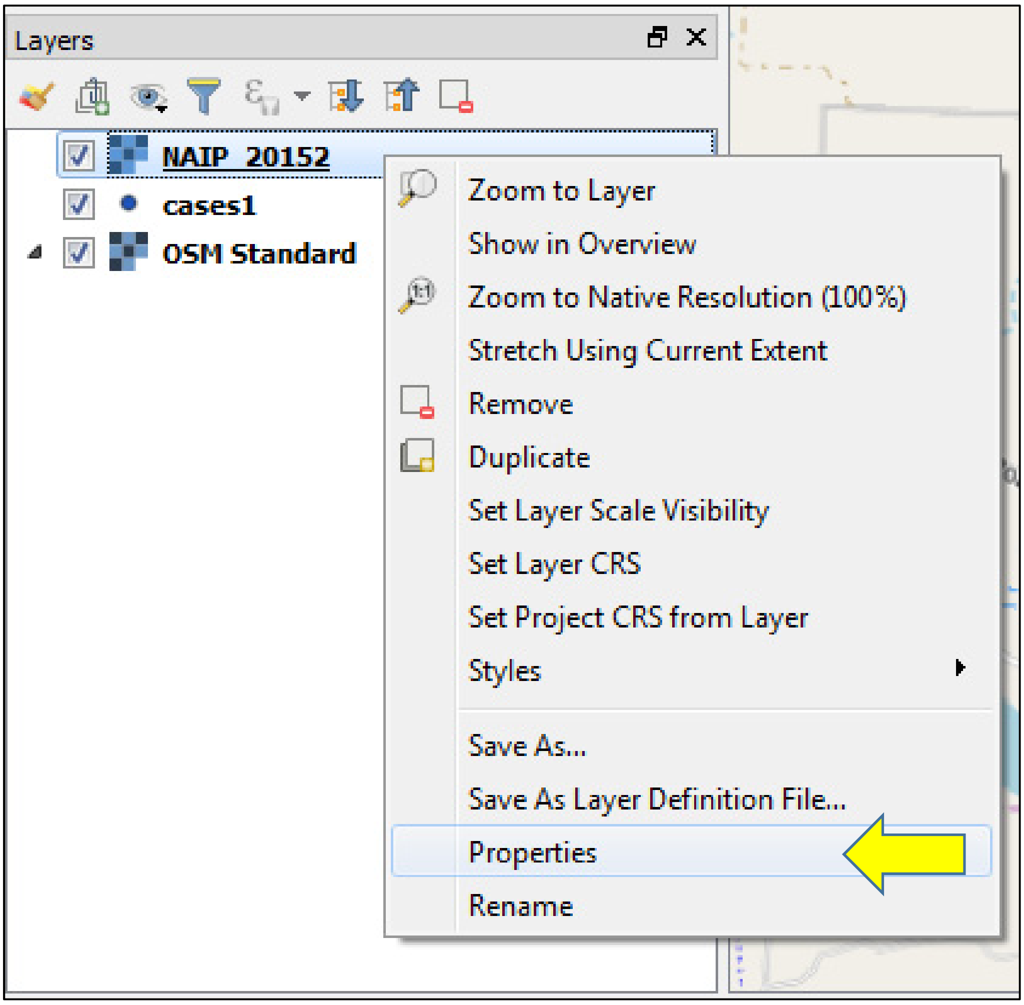
If we now look at our screen, we can see the orthophoto has been added as a raster layer in the "Layers" pane at the bottom left of the screen (yellow arrow). This layer provides us the ability to examine the area surrounding our disease cases in much more detail when compared to the simple base map layer that we added previously. The map with the orthophoto is displayed on the following page.



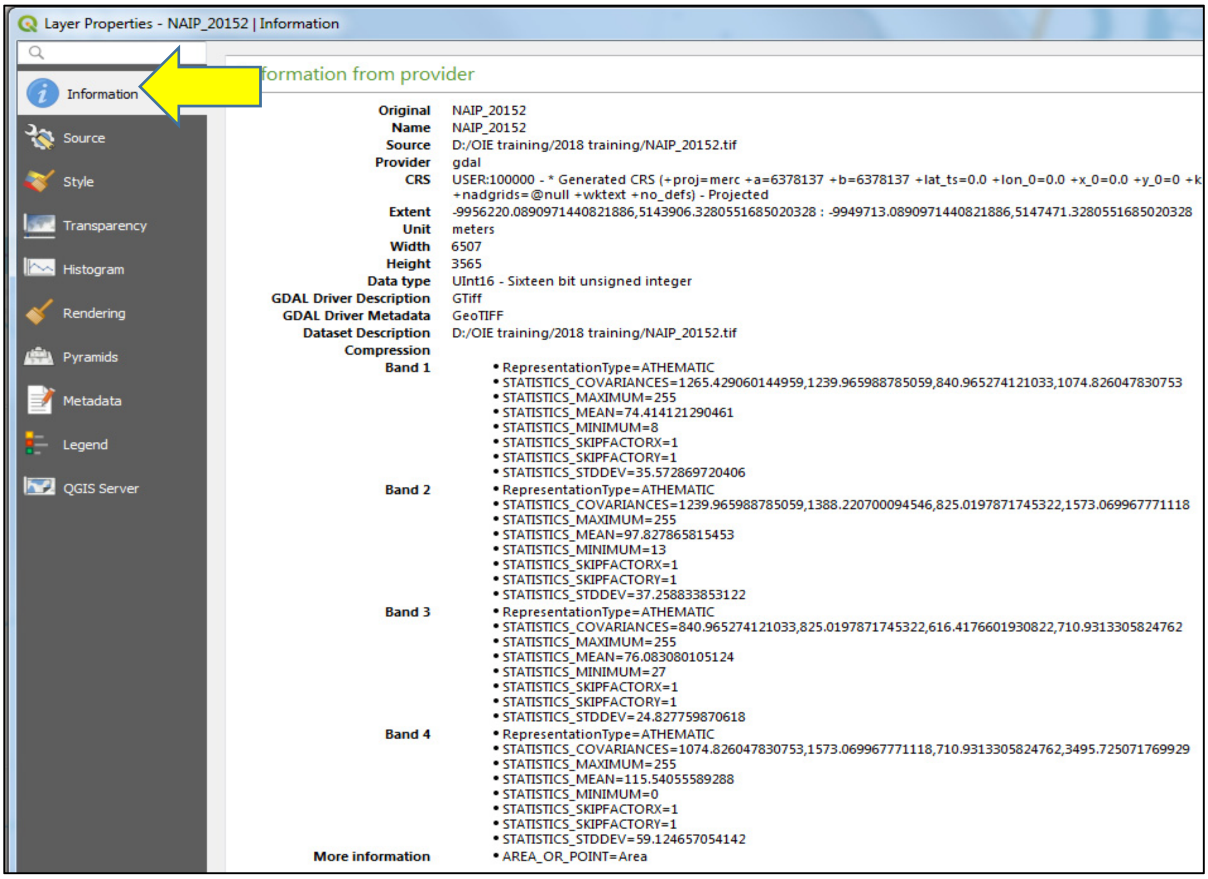

Flagg Road
IL 2
Lost Nation Road
White Oak Lane

It is worth noting that in the map above that most of the disease cases are no longer visible. We will address this problem shortly.

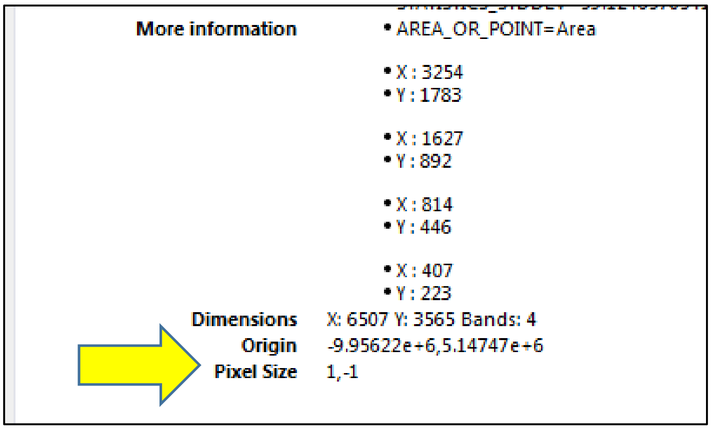
If we would like to know properties of this raster layer, for example the resolution of the raster image, we can right click on the name of the layer in the "Layers" pane, and select "Properties". Then click on "Information".



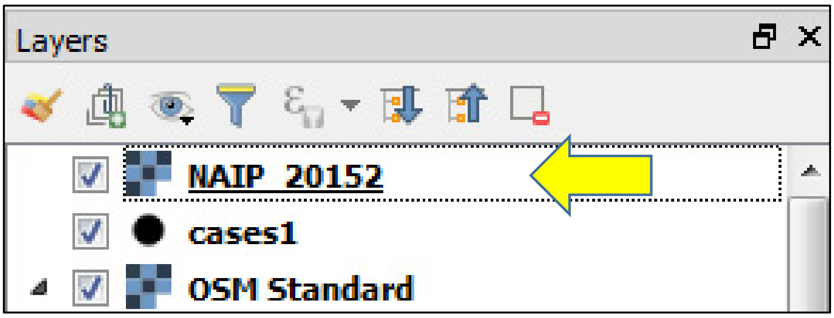
This brings up the following window, which provides a wide variety of information regarding the layer. It provides the CRS (as shown previously); the measurement units (i.e., meters); the size of the layer, which is specified in the number of pixels in the width and height; type of raster (i.e., GeoTiff); and the number and characteristics of the values associated with the layer (i.e., Band1 – Band 4). This information is useful for understanding the properties of the raster file, and often complements a read-me file, which usually accompany raster files that are acquired from reputable sources, and further explain the raster file's structure and properties.



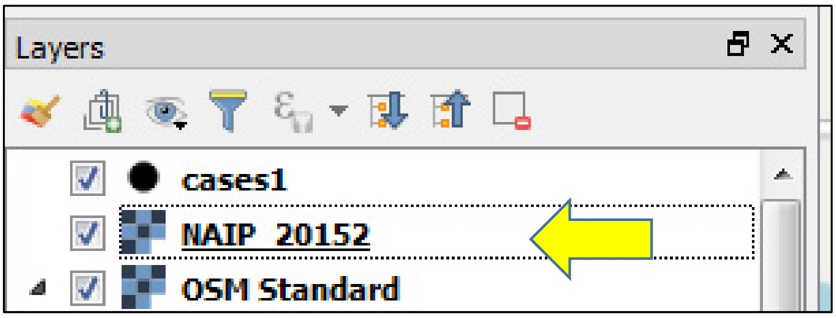
To find the resolution of the image we scroll this window down further, and find "Pixel Size", which for our example is 1 meter because the units are in meters.



Now, let's go back and look more closely at our project. As previously mentioned, we see that the points representing our disease cases are no longer visible because the raster layer we just added is "on top" of or above the cases1 layer. We can examine the order of the layers in our project by looking in the "Layers" pane. The layer on the top of the list in this cases is our recently added raster image (yellow arrow). The second layer is the "cases1" layer, and the third is our "OSM Standard" base layer. This is not what we want.

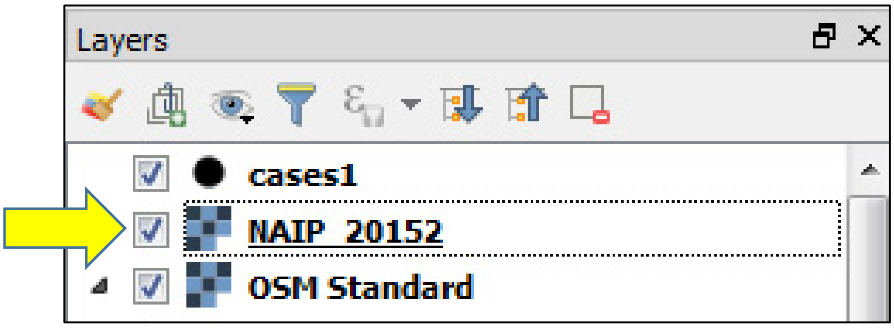


To move the cases1 layer to the top of the layer stack, we hold down the left mouse button and drag the cases 1 layer to the top. We can do this with any layer to achieve the correct ordering of the layers for our particular problem.

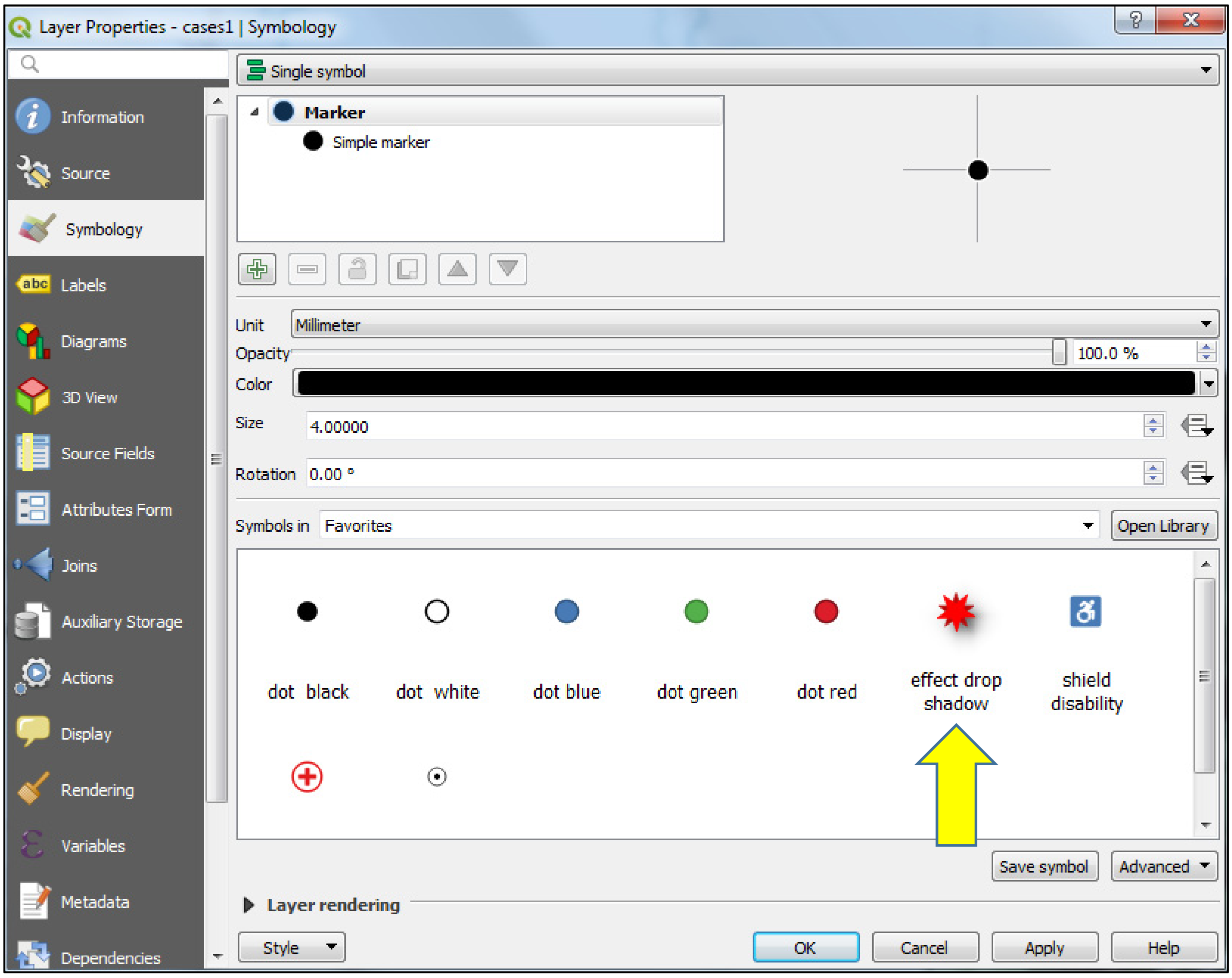


The cases now are on top of the orthophoto (i.e., the black dots shown on the map below).

We can also turn layers "on/off" by simply clicking on the box next to the name of each layer in the "Layers" pane (yellow arrow). If the box next to the layer's name is checked it means that it is "on" or visible to the user. In the figure below, I turn off the OSM base layer, and it is no longer displayed.

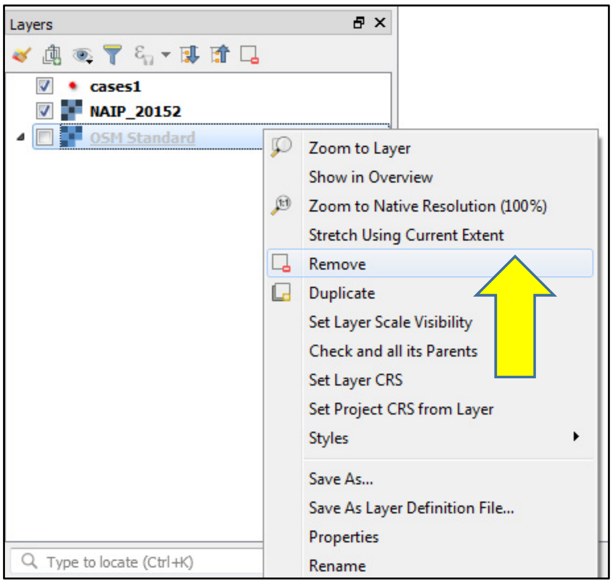


Another feature of a GIS is the ability to change the style, color, etc. of the features. In our example, the locations of our disease cases are represented by a blue dot, but we can change this to something that is more visible. To change the appearance of our symbol for our disease cases, we right click on the cases1 layer. We select "Properties", which opens a new window. We then select "Symbology".

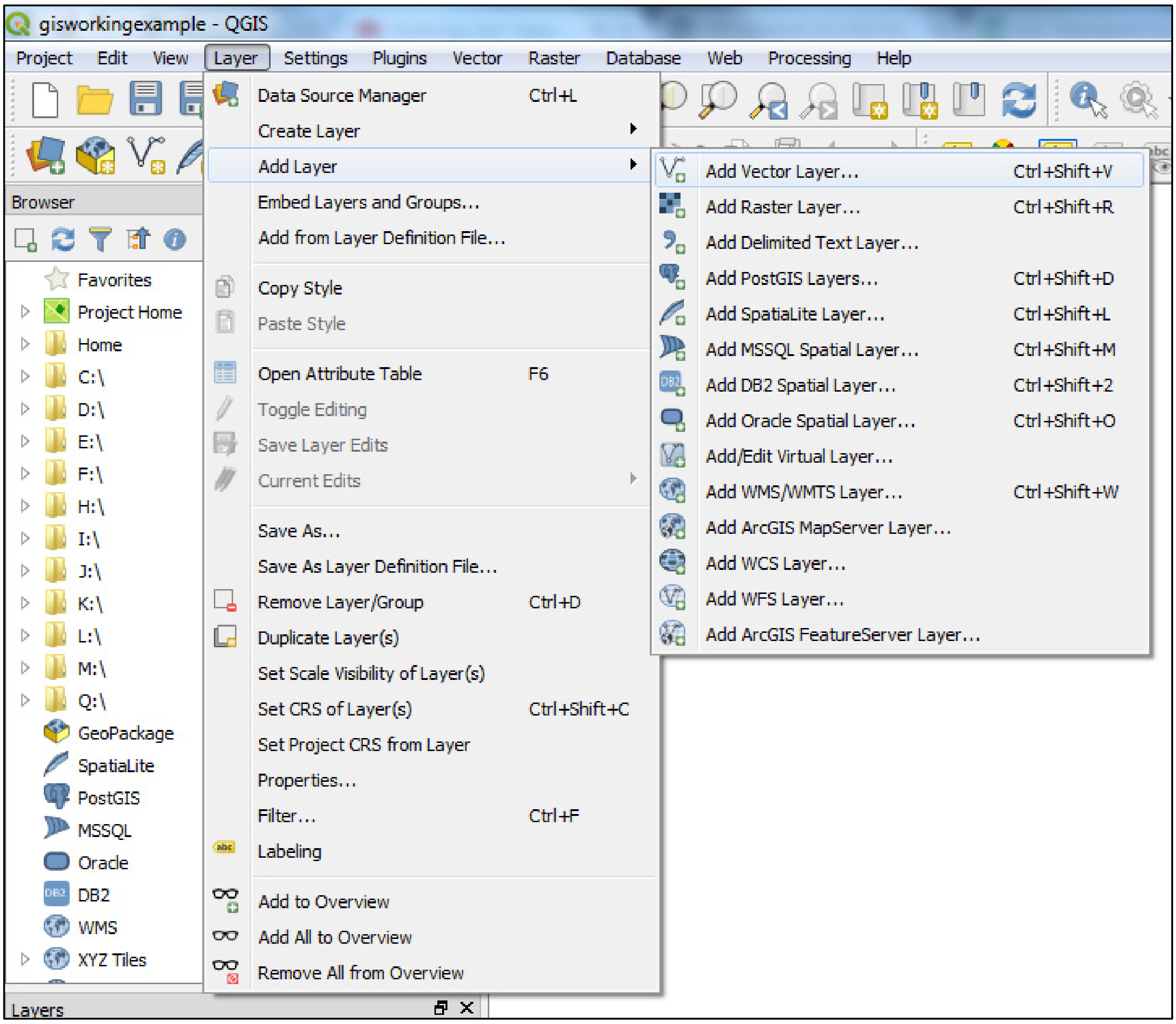


There are numerous options to change the appearance of our symbols. I will change the appearance to "effect drop shadow" by clicking on that new symbol type (yellow arrow), but there are numerous other possibilities based on the GIS user's preference. I then click ok. Now we can see the symbols are much more prominent.

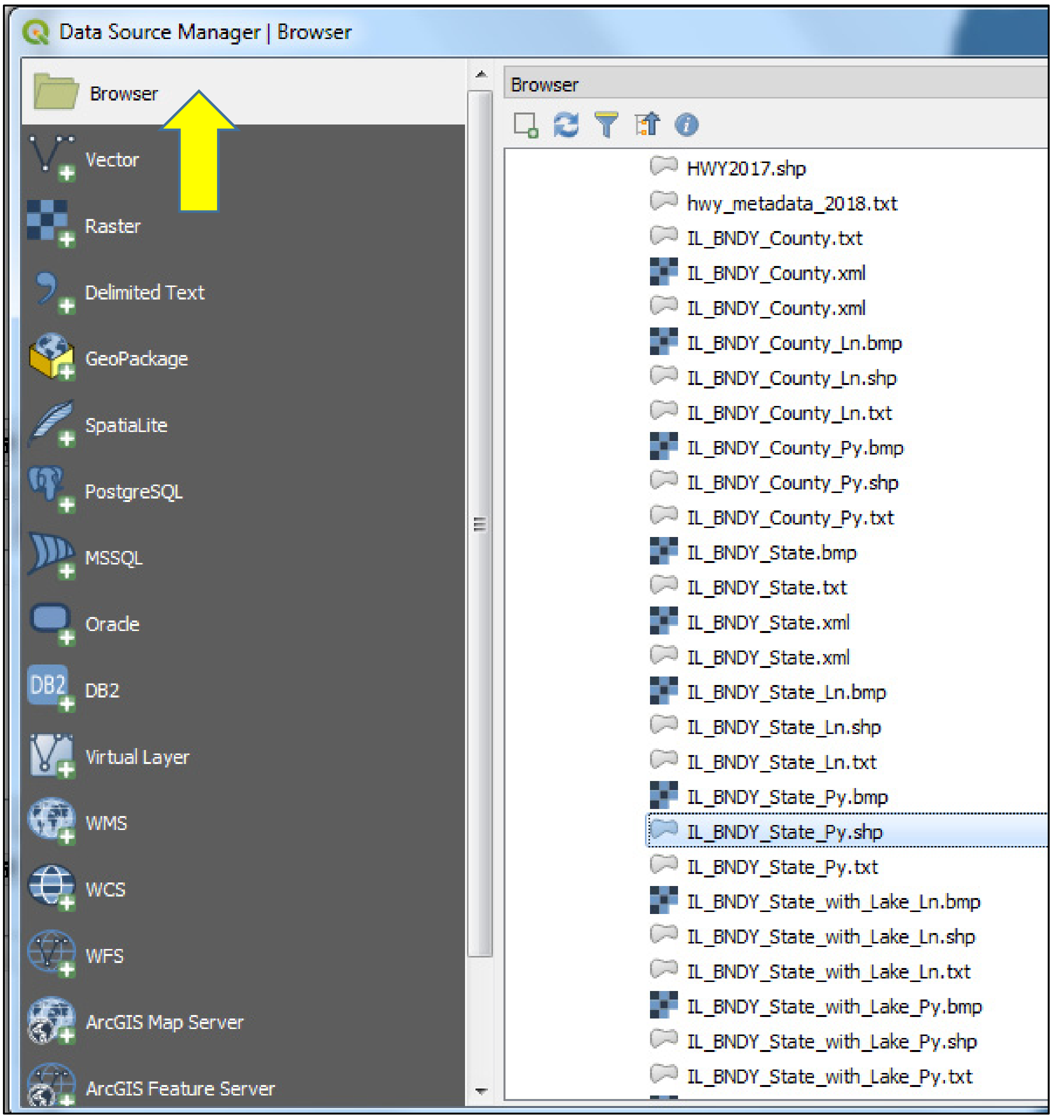
It is also possible to remove layers that are no longer of interest. We do this by simply right clicking on the name of the layer that is no longer of interest in the "Layers" pane, and selecting "Remove". This will remove the layer from the project.



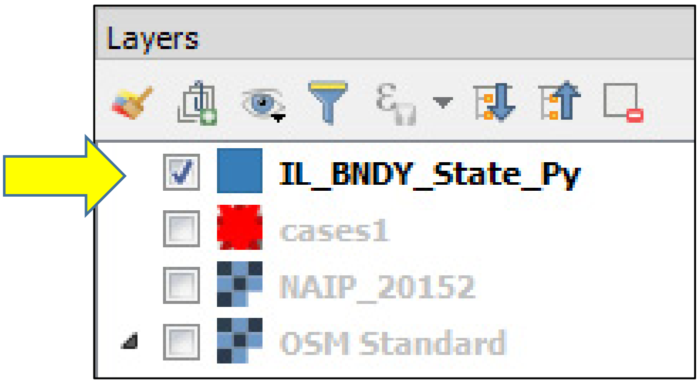
To this point we have talked mainly about raster data layers; however, there are also vector data layers. Unlike raster data layers that are created by pixels, vector data layers are created using mathematical calculations that join series of points to form lines or polygons. They may also just consist of a series of points such as our cases1 layer. Vector data is scalable, which means that its resolution doesn't change when you zoom in or out on the layer, whereas raster data can become pixelated when you zoom in close. An interesting discussion of the strengths and weaknesses of raster and vector data can be found at the following url: *https://gisgeography.com/spatial-data-types-vector-raster/*, or by conducting an internet search. To demonstrate how to add a vector layers to a GIS project, we will add the outline of the state of Illinois, which is the state that contains our disease cases. The layer was downloaded previously, and is a polygon vector layer. To add this vector layer, select the Layer menu at the top of the screen. Next click on "Add Layer", and then click on "Add Vector Layer".



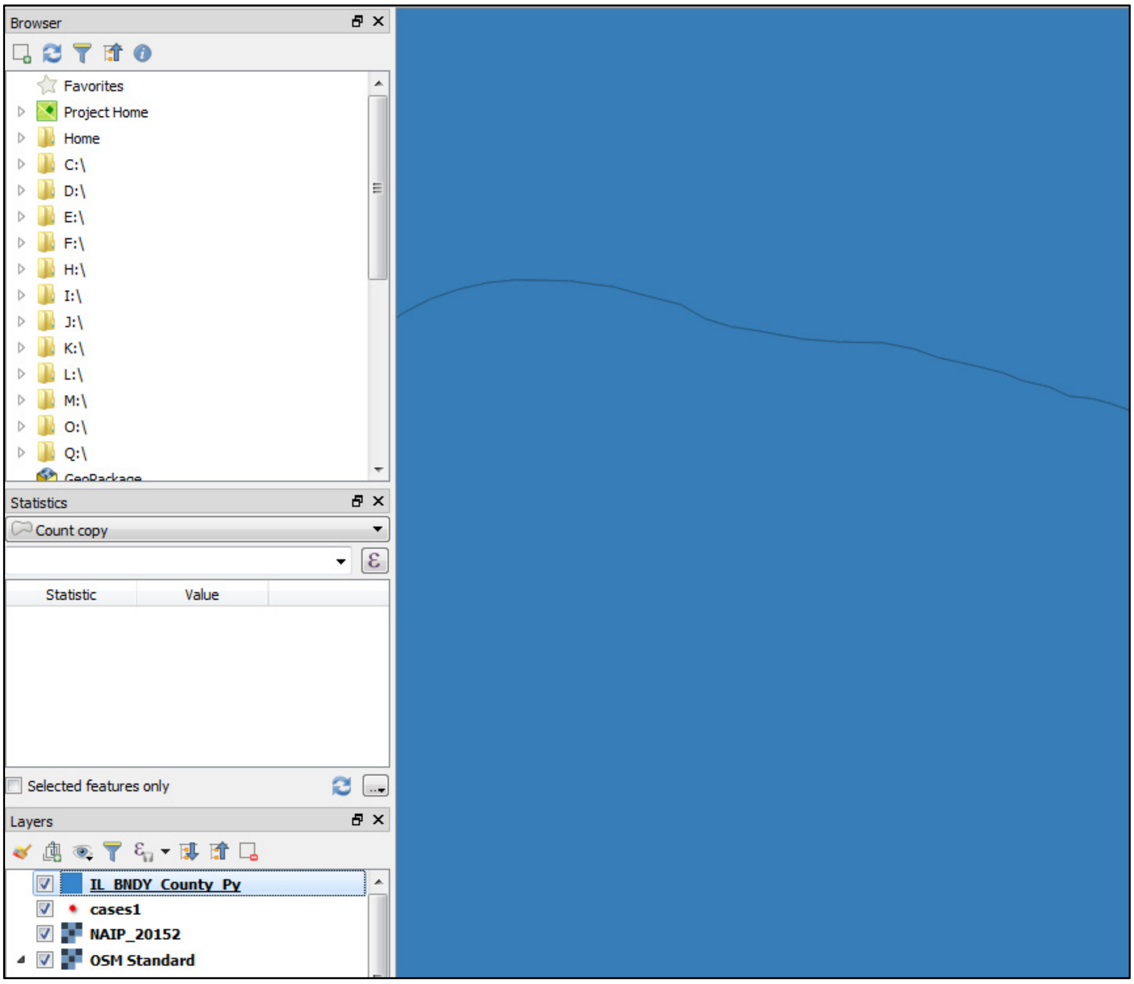
Next click on the “Browser” to navigate to the layer’s location on your local computer. Select the vector layer you want to add to the project and double click on its name (i.e., IL_BNDY_STATE_Py.shp). Then close the Browser window.



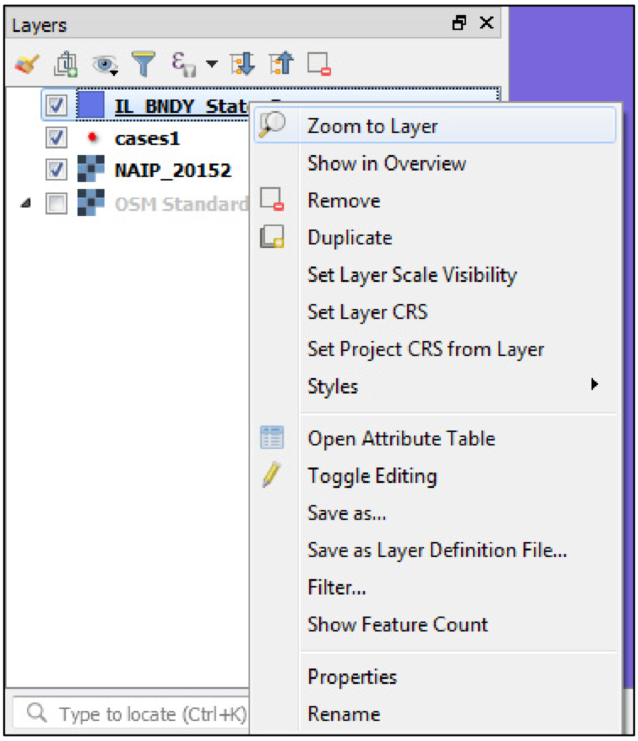
This will add the layer to your project, and its name will show in the “Layers” pane in the lower left corner of the screen (yellow arrow).



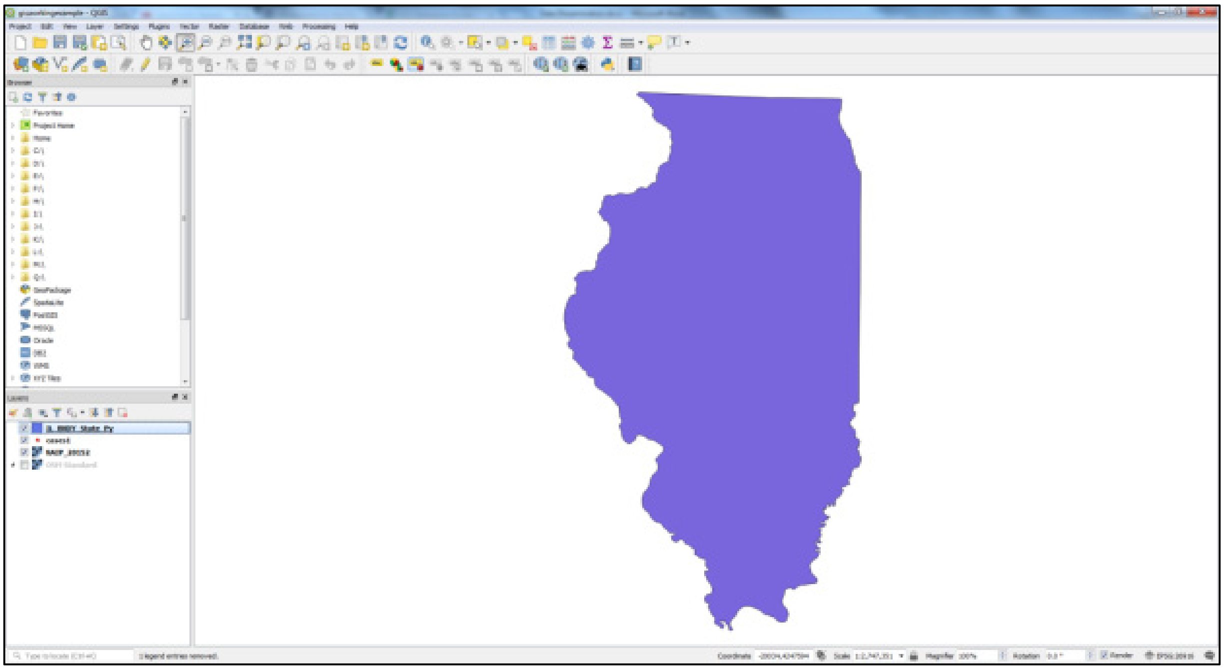
Unfortunately, because we are zoomed in to the location of our cases, we are unable to see the boundary of the state. However, because it is a polygon, and it is a "filled" shape, all we can see the purple color associate with the polygon. To see the full extent of this layer, we need to zoom to the layer's extent.



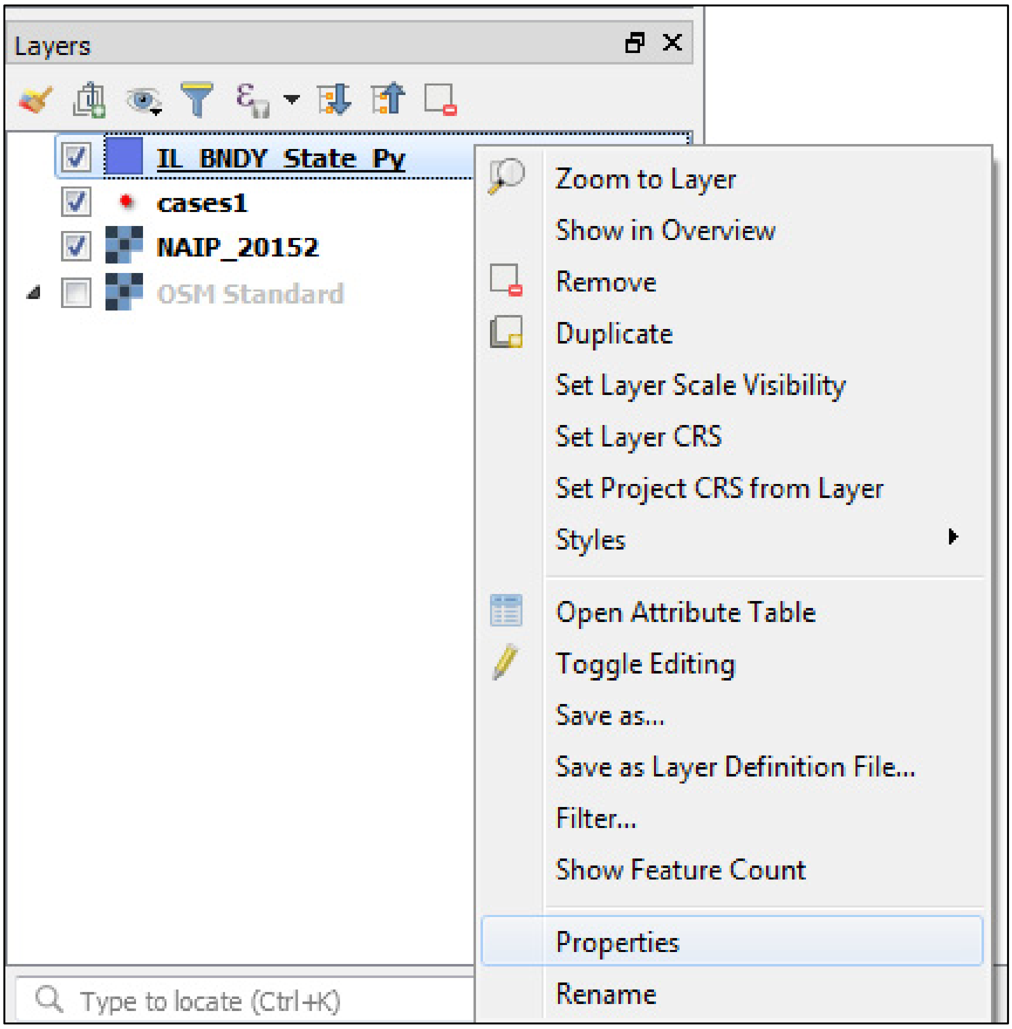
To zoom out to see the entire layer, we right click on the layer's name and select "Zoom to Layer".



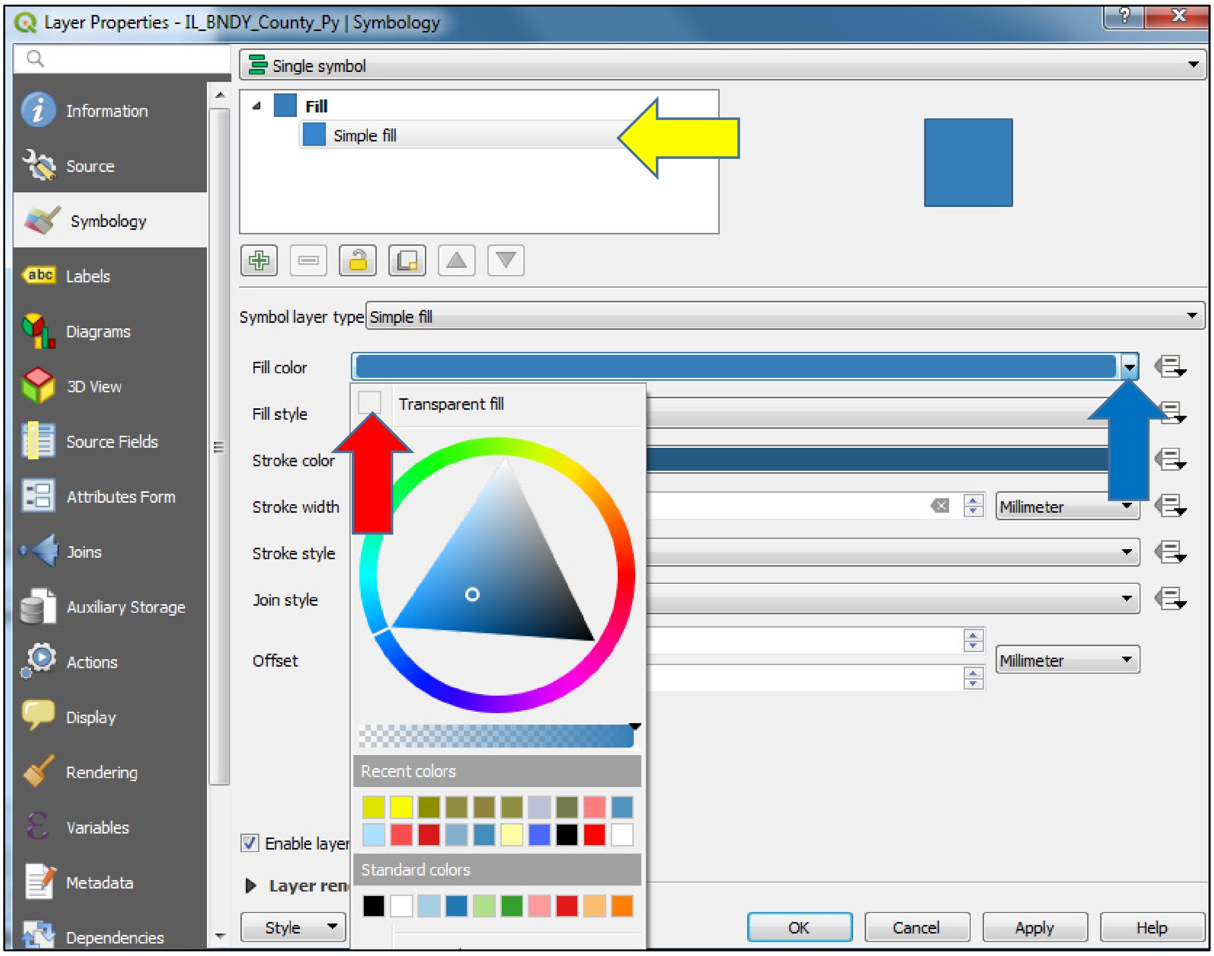
This zooms our GIS project out so that the entire layer is visible. We can now see the outline of the state, but we can no longer see our disease cases because they are below the state boundary in the stack of layers, and the polygon is filled.



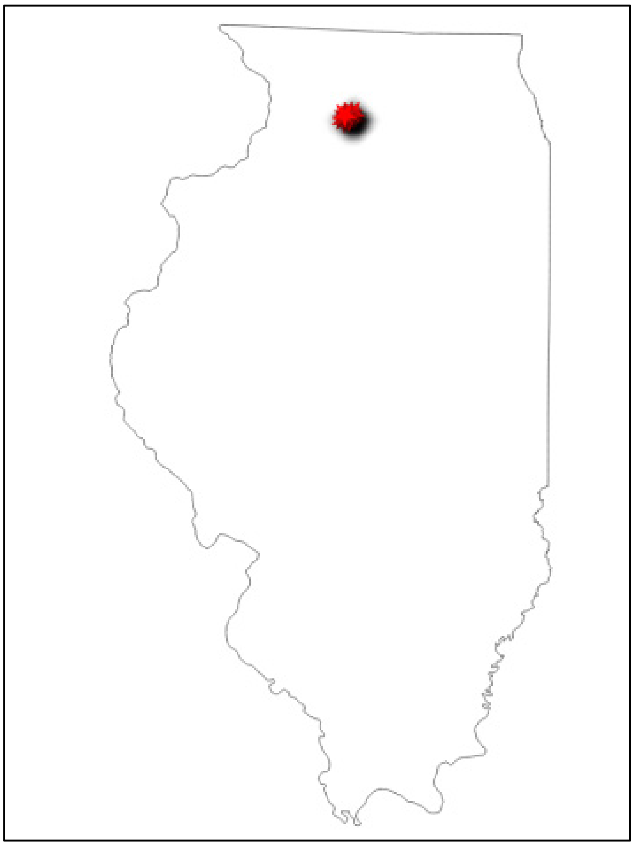
We can either move cases up in the layer stack, as previously described, or we can choose to only display the outline of the polygon (i.e., not filled). To accomplish the latter task, we will right click on the polygon layer's name in the "Layers" pane in the bottom left of the screen, and select "Properties". This will create a new window from which we will select "Symbology".



In this window, there many different options to change how the polygon is displayed in our GIS project. We would like the polygon not to be filled. Therefore, we first click on "Simple fill" (yellow arrow), and click the black triangle next to "Fill color" (blue arrow). This open a second window with options for specifying a color with which to fill our polygon. To make our polygon unfilled, we check the "Transparent fill" box (red arrow). Then click ok.



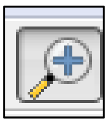
This creates an outline of the state where our disease cases occur, and we can see where within the state our disease event occurred. This provides a much larger spatial context to our wildlife health information. However, it is worth noting that this may not be best the way to visualize this information because at this scale all the disease cases appear as just a series of symbols stacked on top of each other. It is not possible to know how many cases occur and whether they occur only at one location or if they are spread over a greater spatial extent.



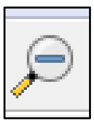
Interactive Map Viewing

One of the benefits of capturing wildlife health information in a GIS is the ability to view the data interactively which is impossible when using static maps. For example, we can easily zoom in or out on specific elements within the map to understand their spatial context at different scales. One method is to use the mouse wheel to zoom in and out. There are also a series of tools for zooming in and out of a GIS project. These are located at the top of the screen.

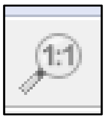
There are eight different zoom tools that can be used. Starting from the left, the first tool is used to manually select a rectangular region to zoom *IN*.



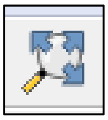
The second tool is used to manually select a rectangular region to zoom *OUT*.



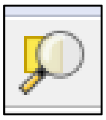
The third tool is used to zoom to "native resolution". This tool is used to zoom into a raster layer, but only to a resolution before the image becomes pixelated.



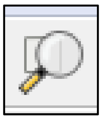
The fourth tool is used to zoom to the project extent, which means the screen is zoomed out to the extent of the largest layer.



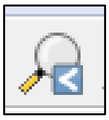
The fifth tool is used to zoom to selected features. We will discuss selecting features shortly.



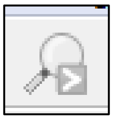
The sixth tool is used to zoom to the extent of the active/ live layer. The active/ live layer is the layer that is highlighted in the "Layers" pane, and a layer can become the active/ live layer by left clicking on the name of that layer in the "Layers" pane.



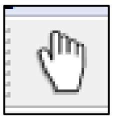
The seventh tool is used to go back to the previous extent of your GIS project. It is essentially an undo button for the zoom tools.



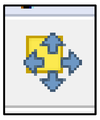
Finally, the last tool is the next extent zoom tool, and serves as a redo button for the zoom tools.



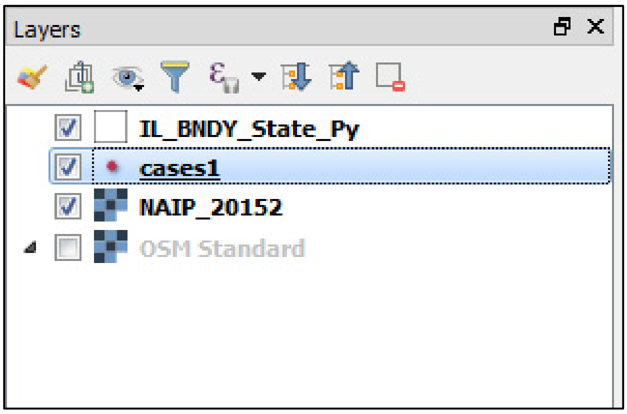
Another useful tool, located left of the zoom tools on the menu bar is the "Pan map" tool, which allows the user to keep the GIS project at the same resolution (i.e., not zoomed in or out), and "drag" or pan the viewing screen to a new area of the project.



The user can also use the "Pan Map to Selection Tool", which is right of the "Pan Map" tool, to pan the map to selected features within the active/ live layer. We will discuss selected features shortly. It is worth noting that the purpose of any of the tools can be identified by hovering the mouse over them in the menu bar.

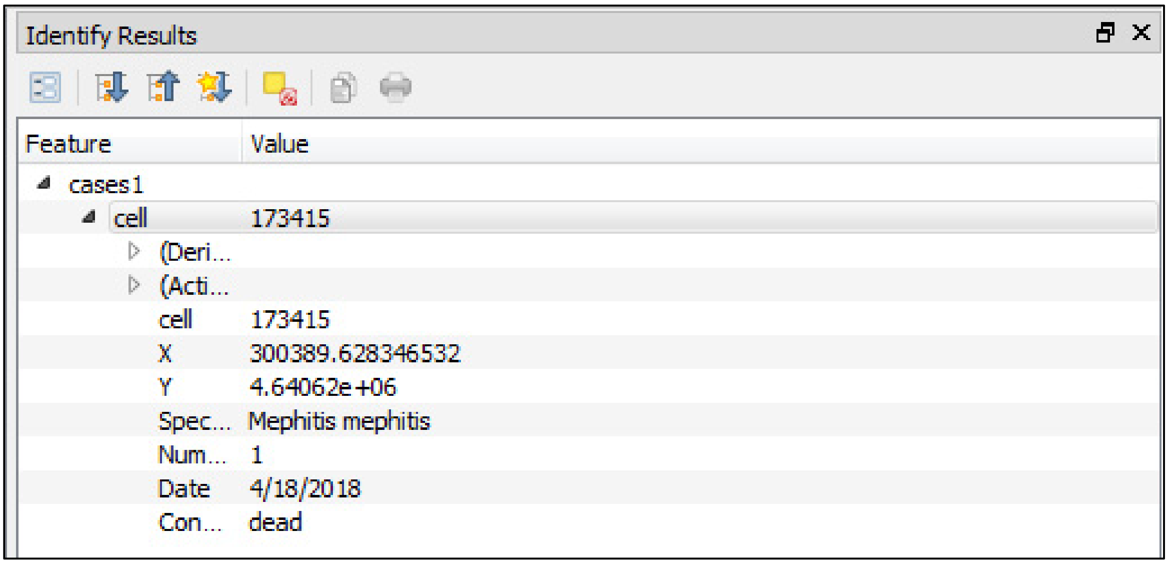


Another useful tool is the "identify tool". With this tool we can left click on elements within our map, and view the associated attribute information. To use the identify tool first left click on the layer of interest in the "Layers" pane. This will make the layer of interest the active/ live layer.

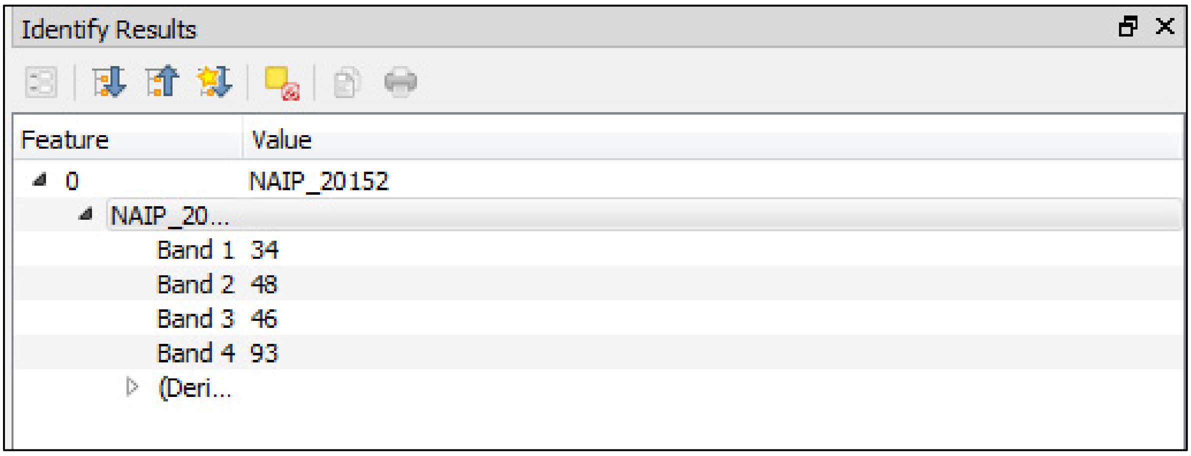


Next, left click on the identify tool which is located in the top tool bar. Now, we left click on any feature

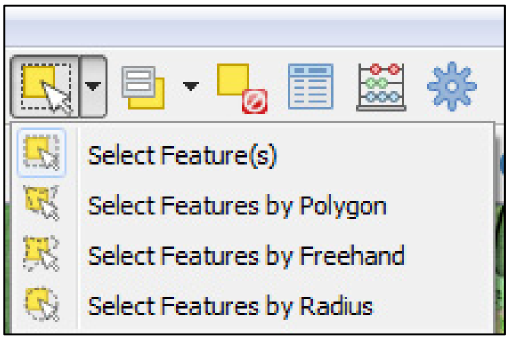
in our layer of interest, and a window opens on the right of the screen that provides all the attribute information associate with that feature. In our example, after clicking on the identify tool icon, I click on one of the disease cases (yellow arrow), and we can see all the information associated with that case in the identify results pane.



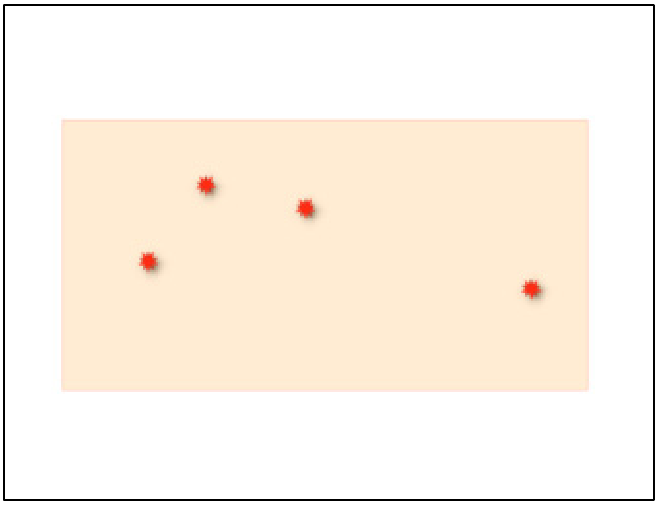
This is the information for that case from our original .csv table that we used to create the layer. The identify tool can be used either on vector data layers, as just shown, or on raster data layers. The latter provides information on individual pixel values. The following shows the identify results when a raster layer is active, which in this case is the orthophoto.



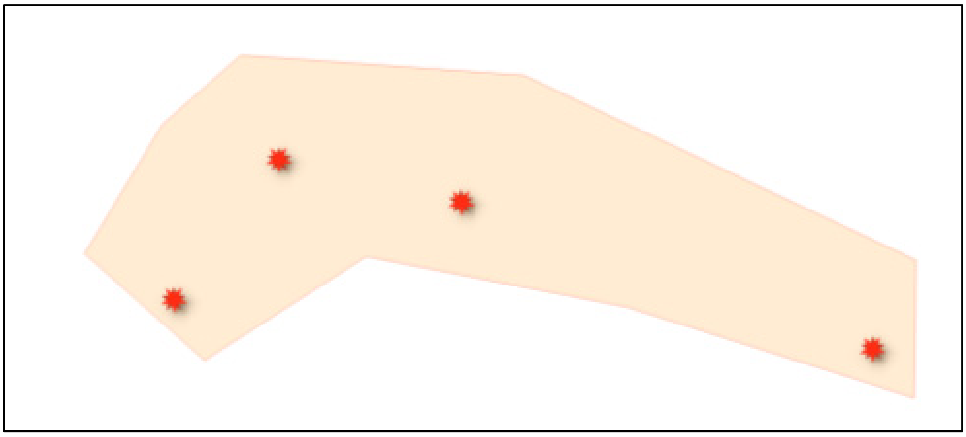
We can also select features over an area using the "Select Features by area or single click tool".



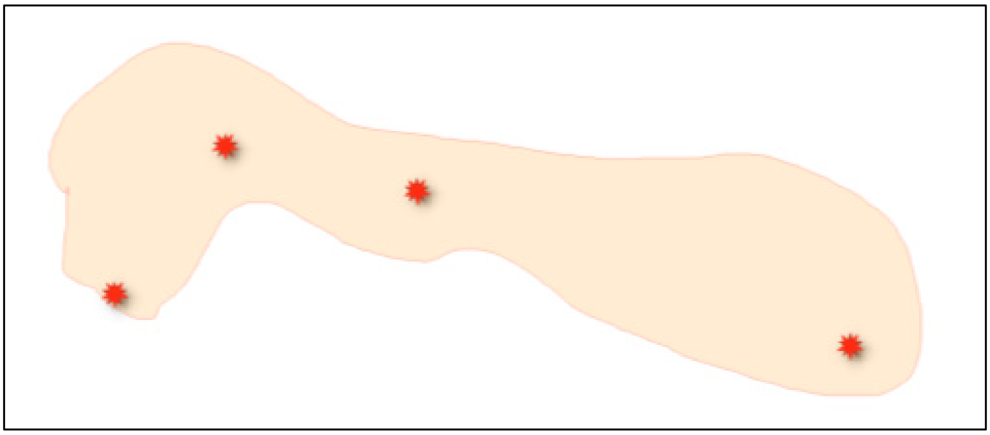
There are four options available with this tool. The first allows you select features from an area using a rectangle, which you specify by clicking and dragging over the features of interest as shown below.



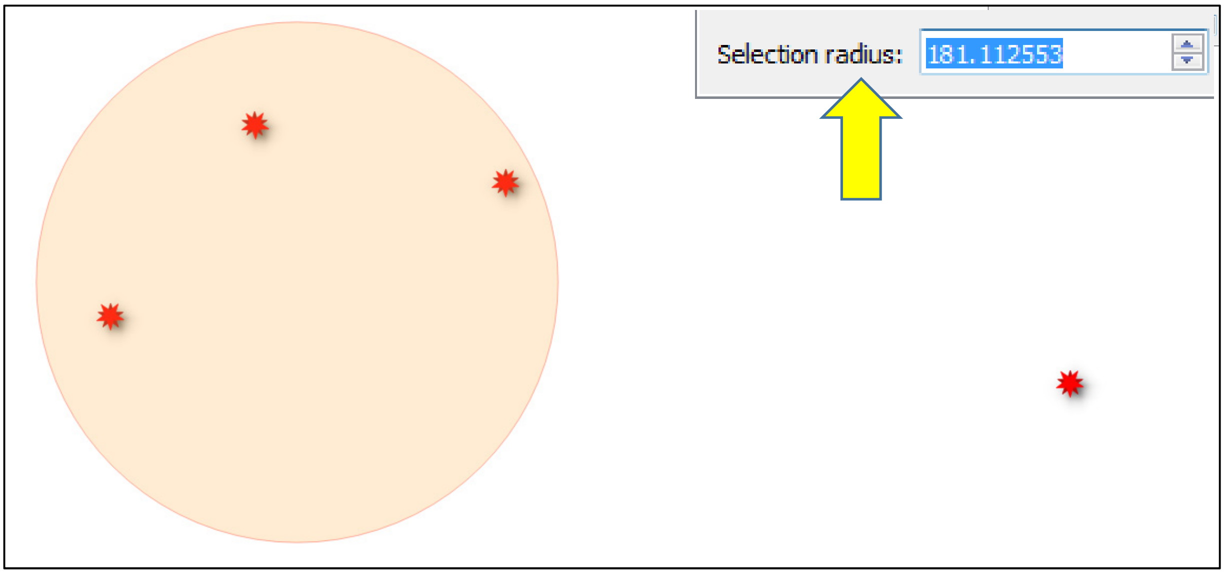
The second option allows you to create a non-rectangular polygon to select features. This is accomplished by left-clicking the mouse to create the vertices of the polygon, and then right clicking to close the polygon.



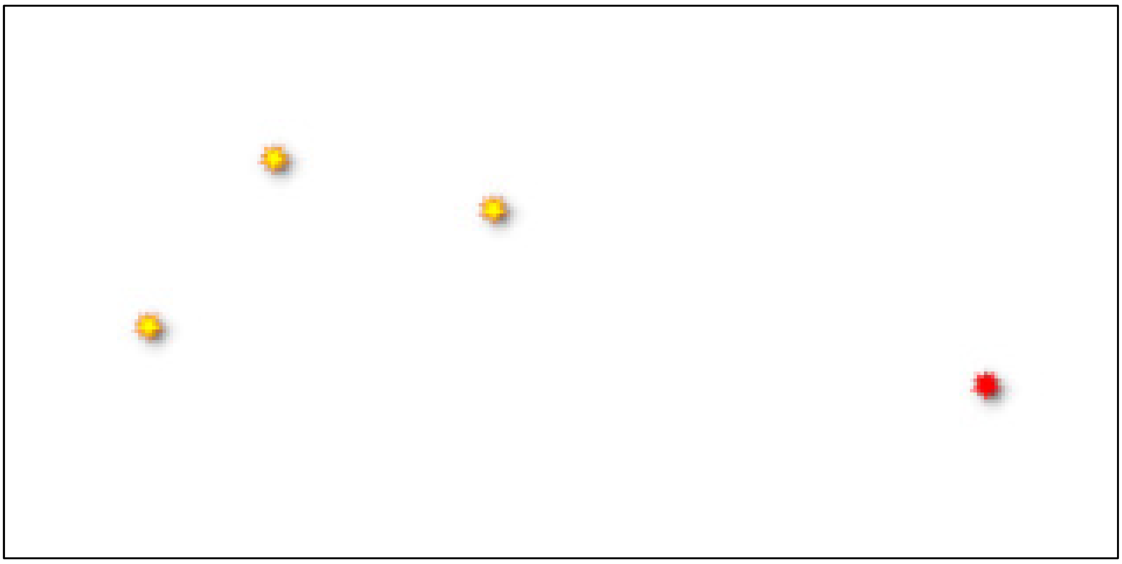
The third option, permits the user to select the features by "free hand". This means that you can just draw around the features that you would like to select. When your drawing encompasses the features you would like to select, left click the mouse.



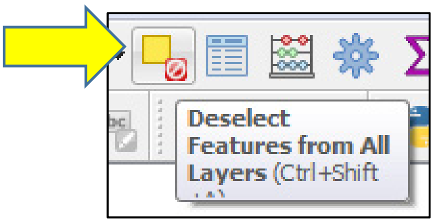
The final option is to select feature "by radius". This involves drawing a circle around the features of interest. To use this tool you first need to left click where you would like the center of the circle to be placed. Then use your mouse to draw the circle, or specify the radius manually in the dialog box (yellow arrow). When your circle encompasses the features you would like to select, left click the mouse. Also it is important to note when using this option, you can right click the mouse which allows you to remove the current selection area, and to redraw the selection area of interest.



When features are selected using the select features tool, they will be highlighted in a different color. Below the features were selected using the circle in the image above, and they are displayed in yellow rather than red to indicate they have been selected.



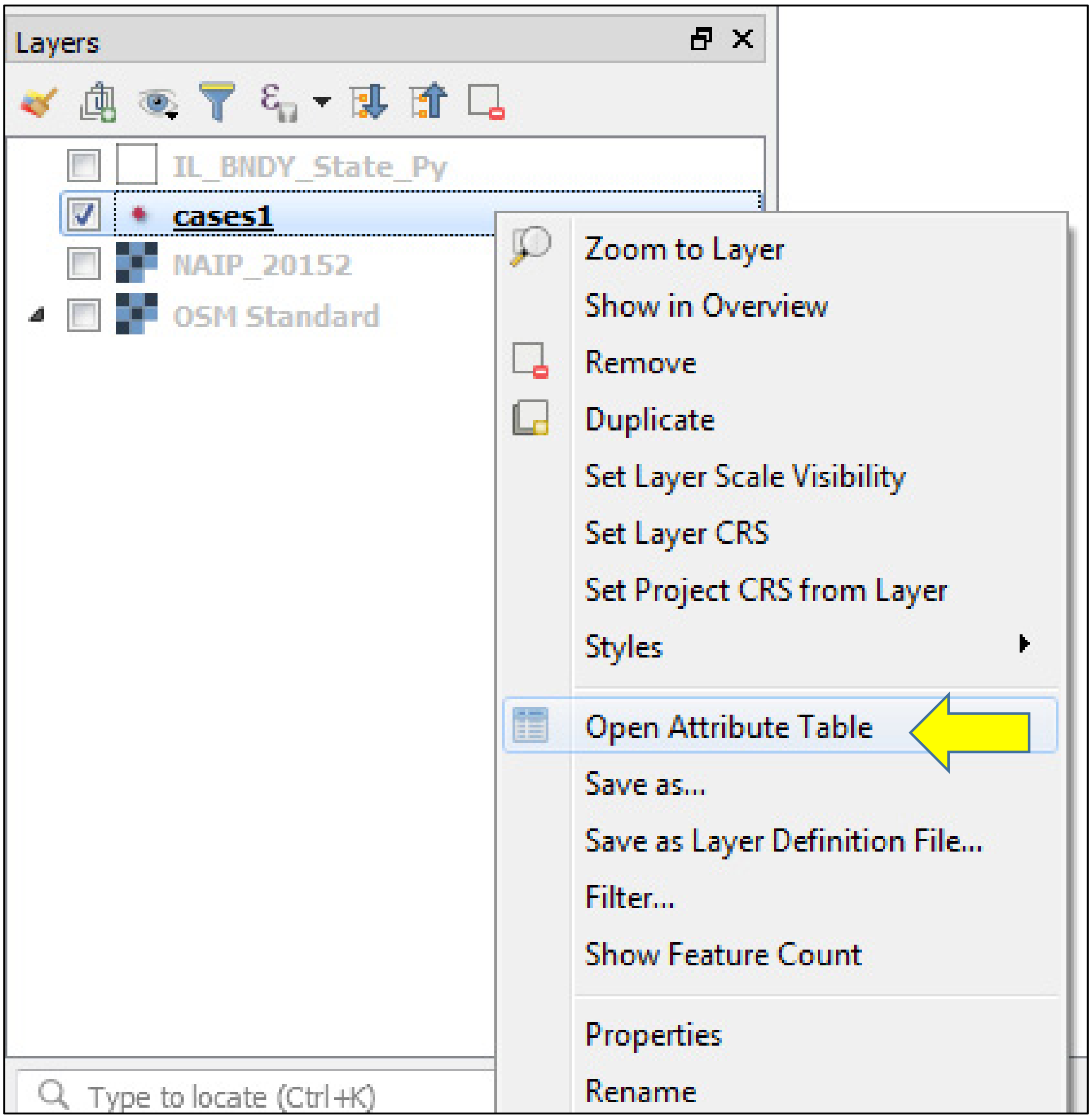
To unselect features, simply click on the “Deselect Features tool”.



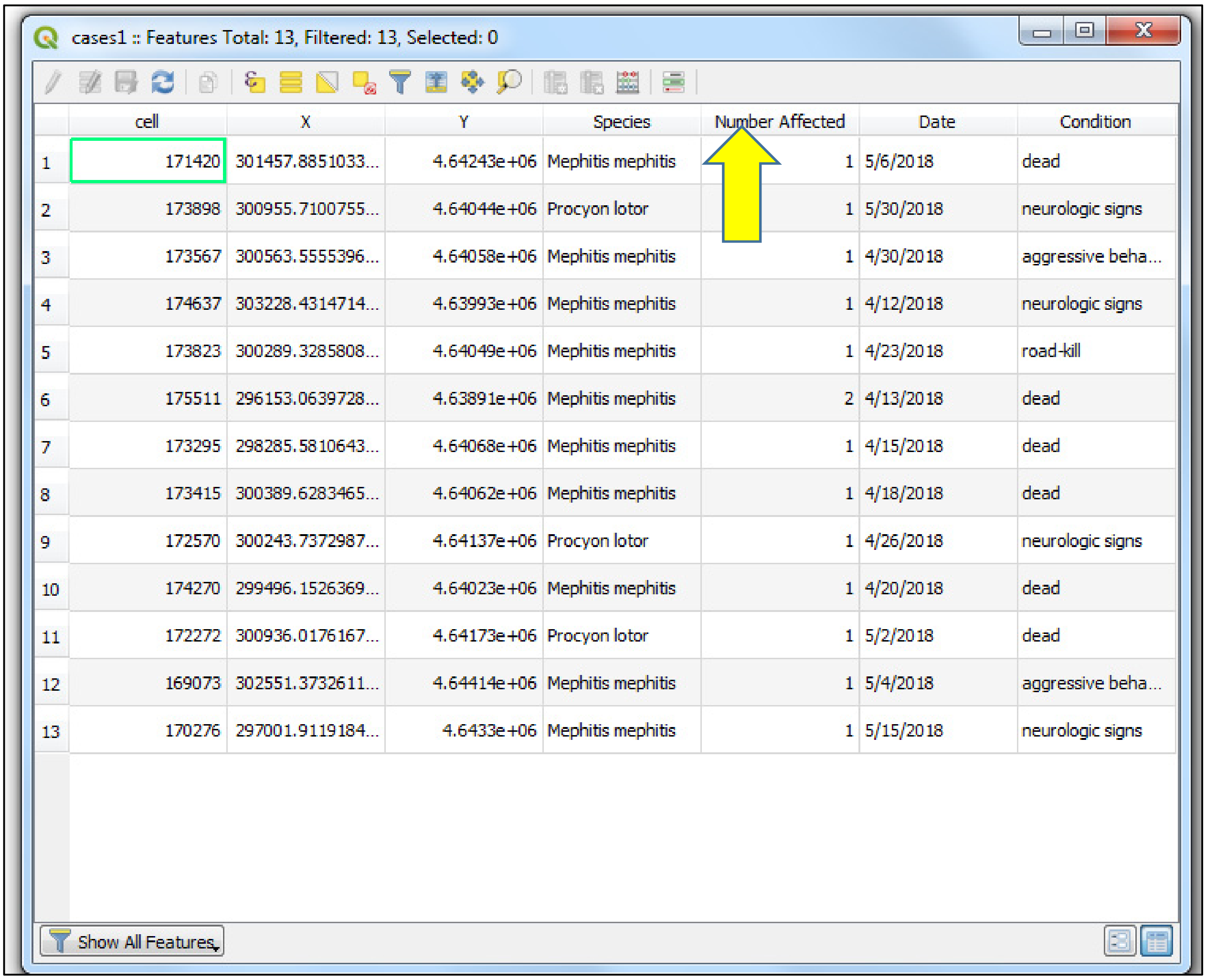
There are several other tools available at the top of the screen, which we will not discuss. However, we encourage you to explore these additional capabilities of QGIS. The QGIS “Help” menu at the top of the screen and the QGIS documentation (*https://qgis.org/en/docs/index.html*)provide good places to start when exploring the functionality of these various tools.

Attribute Information

We have mentioned previously that one of strengths of GIS is the ability to interweave or connect spatial data. Up to this point, we have focused on describing how to bring these spatial datasets together, and to view them interactively within QGIS. However, generally we are not only interested in the features (i.e., point, lines, polygons, etc.) in each layer, but also the attribute information associated with each of these features. Attribute information is the data that is tied to the feature. In some instances, this may just be the type of feature and its spatial location. In other situations, such as in our example (i.e., cases1 layer), there is other relevant data associated with each location. For example, the attribute information for the cases1 layer includes the species, number affected, and condition of the animal upon discovery. Being able to access the attribute information associated with a layer is simple. First, right click on the layer of interest in the “Layers” pane, and select “Open Attribute Table” (yellow arrow).



This will open the attribute table and display all the information for each of the features in the layer. It is worth noting that in our example this table is the same as the .csv file we originally used to create the layer. In this table, each row corresponds to a unique feature in the attribute table. Thus, our layer contains thirteen unique features. The table can be sorted by any column by left clicking on the column heading (yellow arrow). It is worth mentioning that sorting in the attribute table will not have an effect on the map.



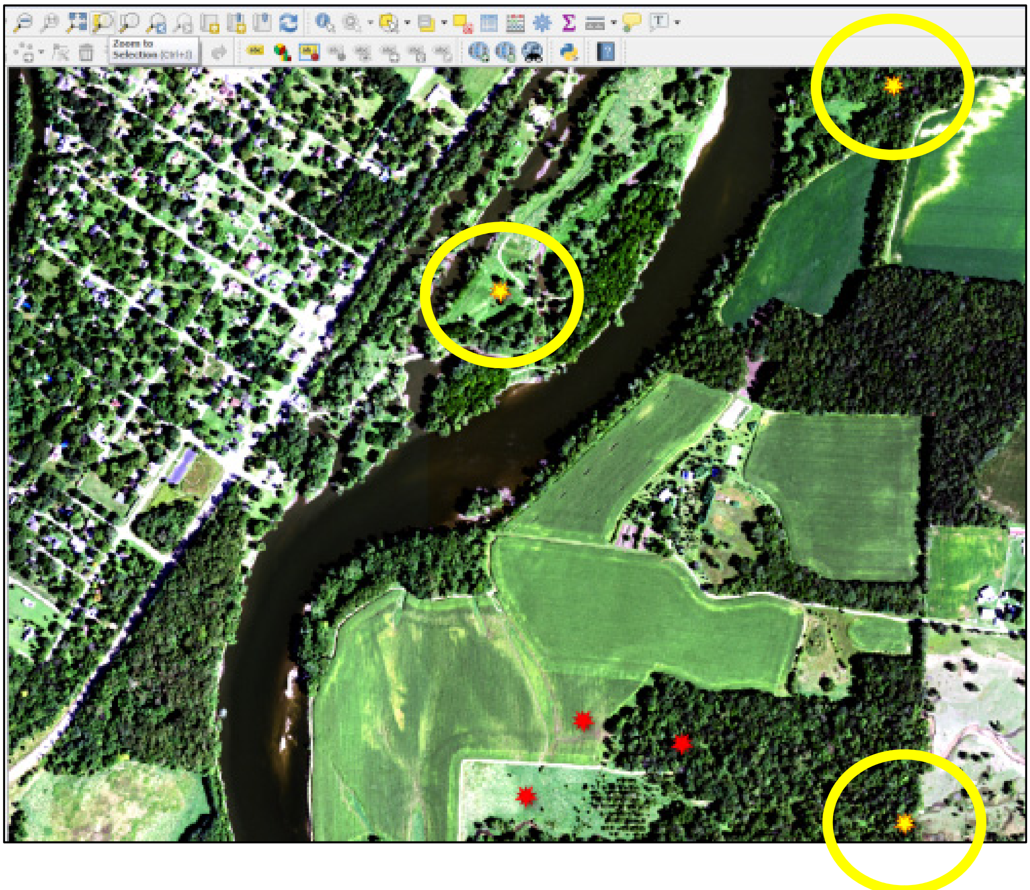
	cell	X	Y	Species	Number Affected	Date	Condition
1	171420	301457.8851033...	4.64243e+06	Mephitis mephitis	1	5/6/2018	dead
2	173898	300955.7100755...	4.64044e+06	Procyon lotor	1	5/30/2018	neurologic signs
3	173567	300563.5555396...	4.64058e+06	Mephitis mephitis	1	4/30/2018	aggressive beha...
4	174637	303228.4314714...	4.63993e+06	Mephitis mephitis	1	4/12/2018	neurologic signs
5	173823	300289.3285808...	4.64049e+06	Mephitis mephitis	1	4/23/2018	road-kill
6	175511	296153.0639728...	4.63891e+06	Mephitis mephitis	2	4/13/2018	dead
7	173295	298285.5810643...	4.64068e+06	Mephitis mephitis	1	4/15/2018	dead
8	173415	300389.6283465...	4.64062e+06	Mephitis mephitis	1	4/18/2018	dead
9	172570	300243.7372987...	4.64137e+06	Procyon lotor	1	4/26/2018	neurologic signs
10	174270	299496.1526369...	4.64023e+06	Mephitis mephitis	1	4/20/2018	dead
11	172272	300936.0176167...	4.64173e+06	Procyon lotor	1	5/2/2018	dead
12	169073	302551.3732611...	4.64414e+06	Mephitis mephitis	1	5/4/2018	aggressive beha...
13	170276	297001.9119184...	4.6433e+06	Mephitis mephitis	1	5/15/2018	neurologic signs

We have demonstrated how to select features manually in the map viewer window. The attribute table can also be used to select features by left clicking on the row number, which will highlight the row. This is useful to determine where features with certain attributes are located in the viewing window. For example, suppose we would like to know where all the raccoon (*Procyon lotor; pictured*) cases are located. We left click on all the rows where the Species is designated as *Procyon lotor*.

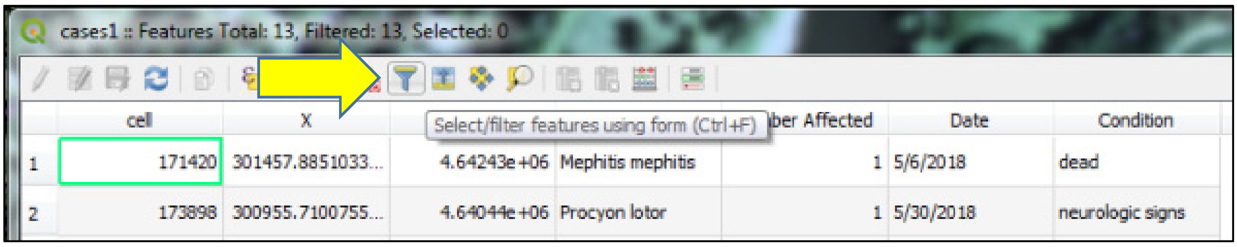
cases1 :: Features Total: 13, Filtered: 13, Selected: 3

	cell	X	Y	Species	Number Affected	Date	Condition
1	171420	301457.8851033...	4.64243e+06	Mephitis mephitis	1	5/6/2018	dead
2	173898	300955.7100755...	4.64044e+06	Procyon lotor	1	5/30/2018	neurologic signs
3	173567	300563.5555396...	4.64058e+06	Mephitis mephitis	1	4/30/2018	aggressive beha...
4	174637	303228.4314714...	4.63993e+06	Mephitis mephitis	1	4/12/2018	neurologic signs
5	173823	300289.3285808...	4.64049e+06	Mephitis mephitis	1	4/23/2018	road-kill
6	175511	296153.0639728...	4.63891e+06	Mephitis mephitis	2	4/13/2018	dead
7	173295	298285.5810643...	4.64068e+06	Mephitis mephitis	1	4/15/2018	dead
8	173415	300389.6283465...	4.64062e+06	Mephitis mephitis	1	4/18/2018	dead
9	172570	300243.7372987...	4.64137e+06	Procyon lotor	1	4/26/2018	neurologic signs
10	174270	299496.1526369...	4.64023e+06	Mephitis mephitis	1	4/20/2018	dead
11	172272	300936.0176167...	4.64173e+06	Procyon lotor	1	5/2/2018	dead
12	169073	302551.3732611...	4.64414e+06	Mephitis mephitis	1	5/4/2018	aggressive beha...
13	170276	297001.9119184...	4.6433e+06	Mephitis mephitis	1	5/15/2018	neurologic signs

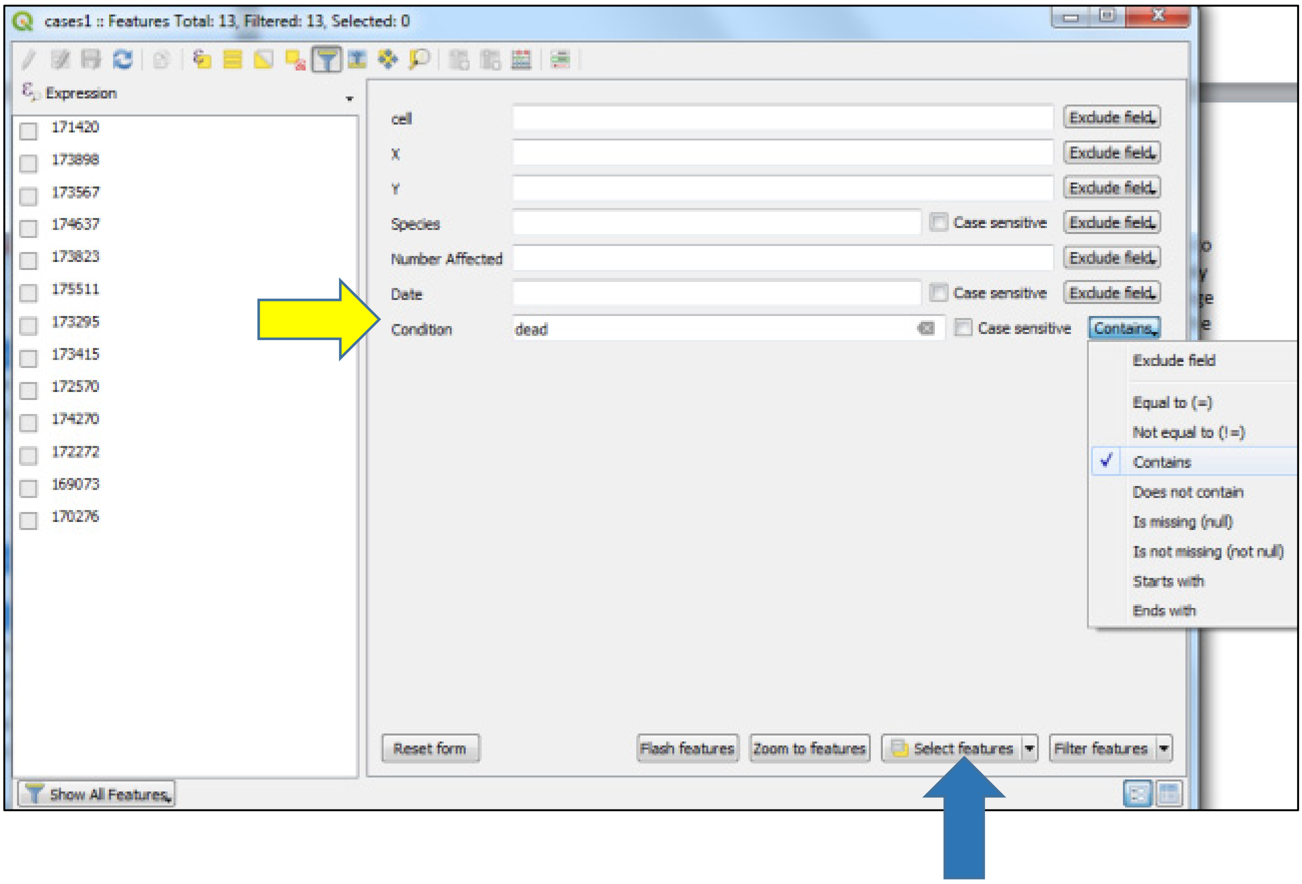
Then we click on the “Zoom to selected features” tool. This will zoom our viewing window so all features associated with raccoon cases are in the extent, and they will be highlighted yellow (yellow circles) whereas other species’ features will be red.



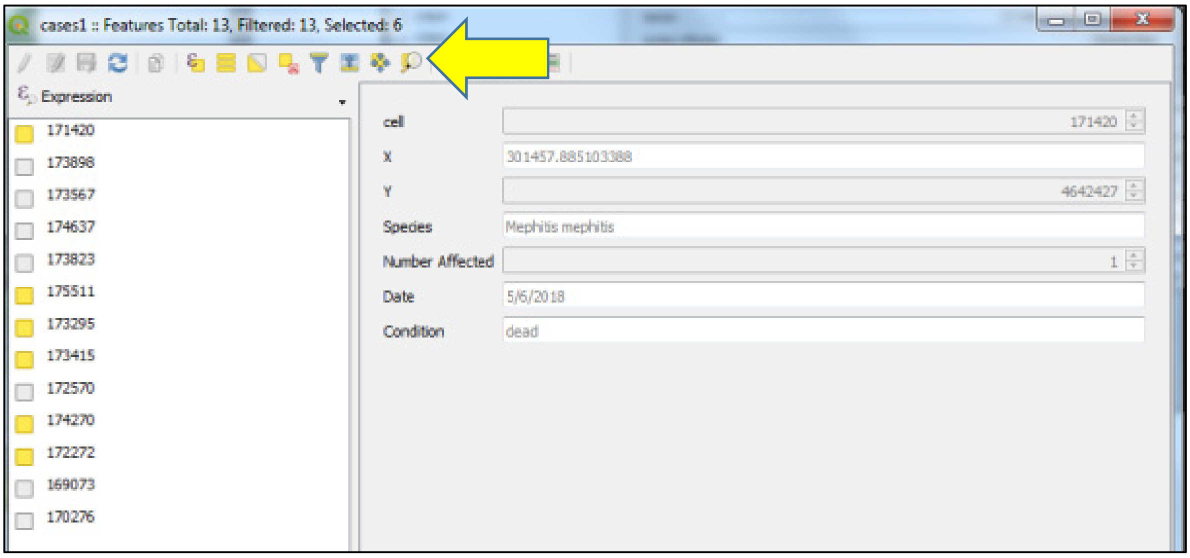
This provides a useful way to view our attribute data under different conditions. QGIS also provides tools to select features based on their attributes using the “Select features using form” tool, which essentially allows us to query our attribute information. This is particularly useful when the layer contains a large number of features, which would make manual selection both tedious and error prone. To access the tool open the attribute table for the layer of interest, as just describe, and then left click on the icon (yellow arrow) as shown below.



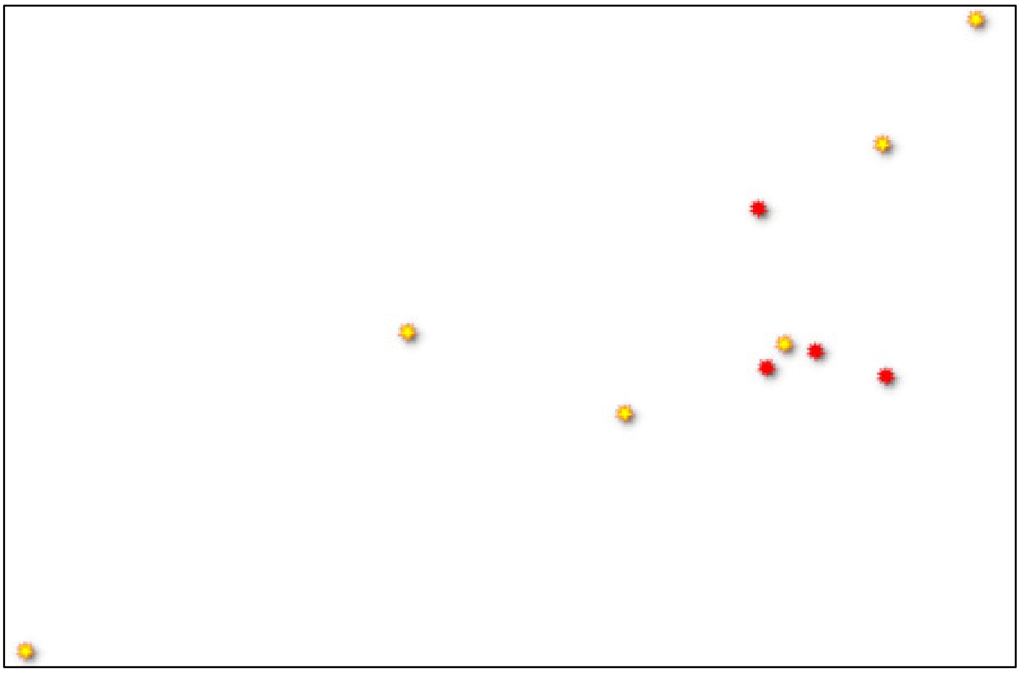
This opens the following dialog window that allows us to create a query of the attribute information. For this example, we will select all features where the condition of the disease case was listed as “dead”. Thus, we write dead in the “Condition” row (yellow arrow), and left click on the tab at the far right of the row. This opens a menu from which we select “Contains”. This means that our query is: select from the attribute table all features whose Condition contains dead. We then collect on the “Select features” tab (blue arrow) at the bottom of the window to run our query.



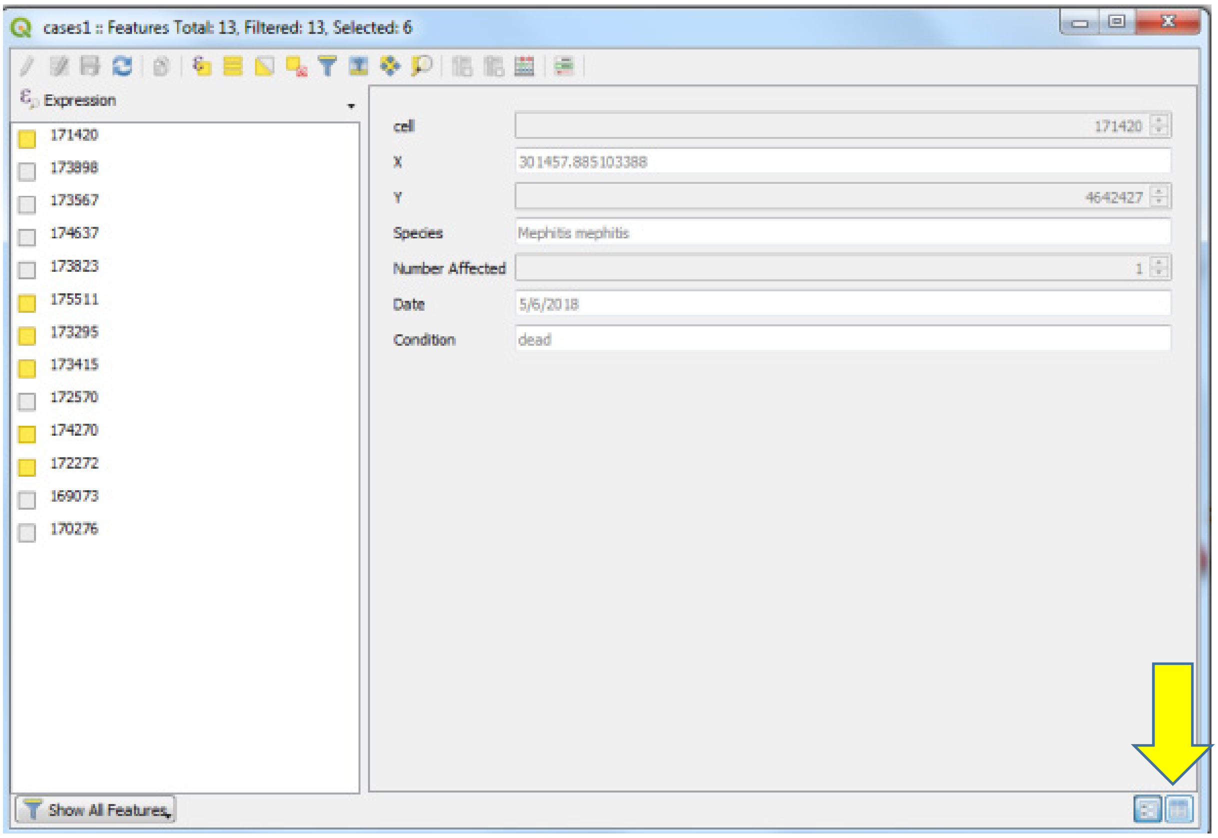
This query determines that there are six features that the condition of the disease case was listed as dead. They are highlighted in yellow on the left side of the screen.



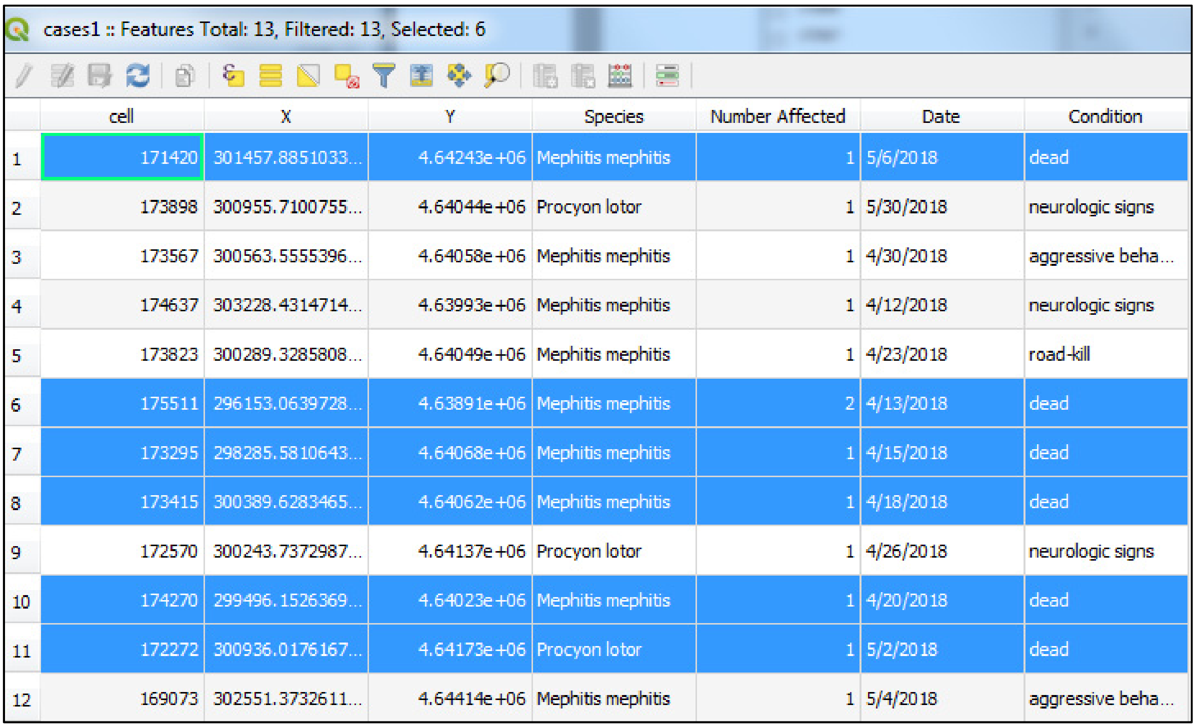
To view these selected features, we click on the “Zoom to selected features” tool (yellow arrow-above). We can see all six selected features are displayed as yellow icons, and non-selected features appear in red.



To switch out of the query/filter view and return to the table view, left click on the bottom right icon (yellow arrow).

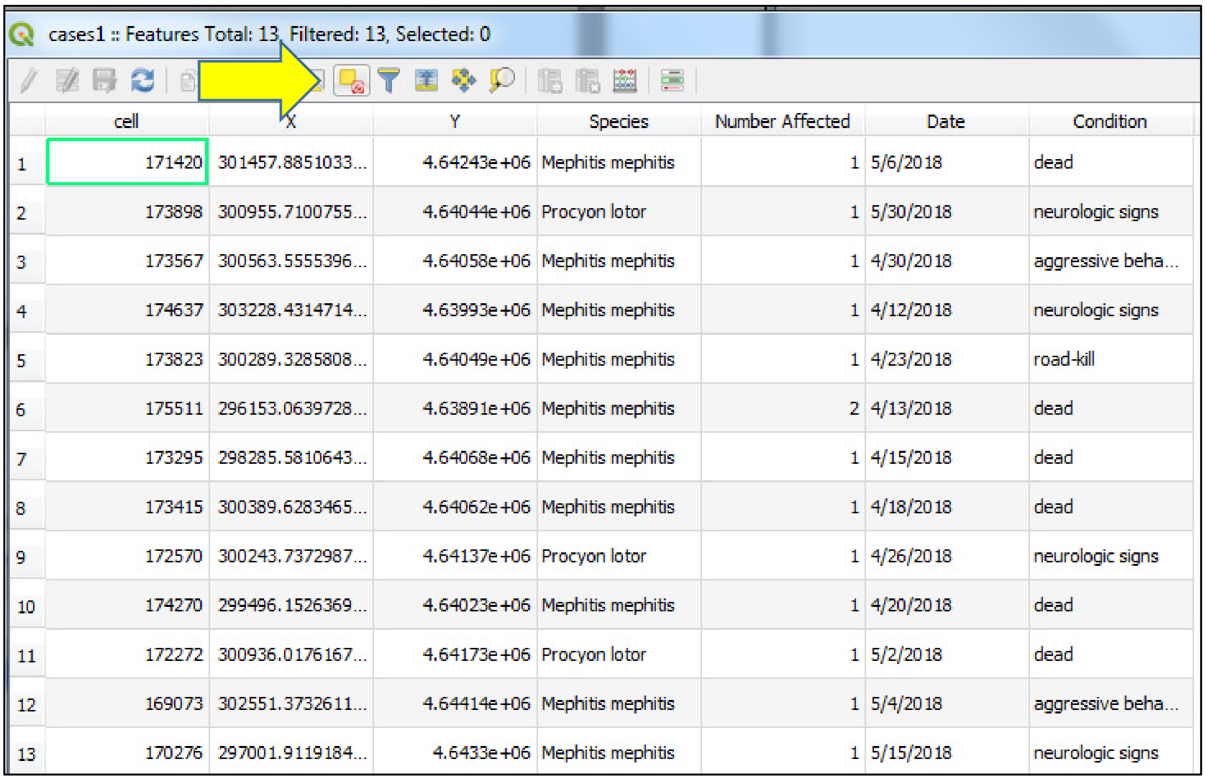


This will return the screen to the table view, and we can see the highlighted rows correspond to our six selected features.



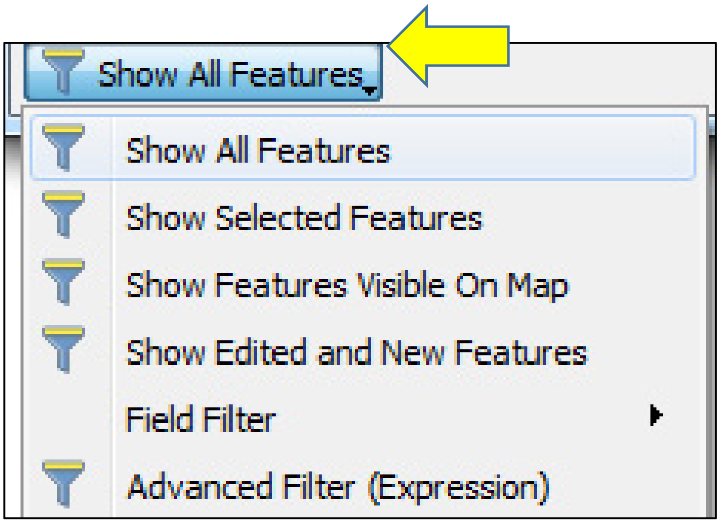
	cell	X	Y	Species	Number Affected	Date	Condition
1	171420	301457.8851033...	4.64243e+06	Mephitis mephitis	1	5/6/2018	dead
2	173898	300955.7100755...	4.64044e+06	Procyon lotor	1	5/30/2018	neurologic signs
3	173567	300563.5555396...	4.64058e+06	Mephitis mephitis	1	4/30/2018	aggressive beha...
4	174637	303228.4314714...	4.63993e+06	Mephitis mephitis	1	4/12/2018	neurologic signs
5	173823	300289.3285808...	4.64049e+06	Mephitis mephitis	1	4/23/2018	road-kill
6	175511	296153.0639728...	4.63891e+06	Mephitis mephitis	2	4/13/2018	dead
7	173295	298285.5810643...	4.64068e+06	Mephitis mephitis	1	4/15/2018	dead
8	173415	300389.6283465...	4.64062e+06	Mephitis mephitis	1	4/18/2018	dead
9	172570	300243.7372987...	4.64137e+06	Procyon lotor	1	4/26/2018	neurologic signs
10	174270	299496.1526369...	4.64023e+06	Mephitis mephitis	1	4/20/2018	dead
11	172272	300936.0176167...	4.64173e+06	Procyon lotor	1	5/2/2018	dead
12	169073	302551.3732611...	4.64414e+06	Mephitis mephitis	1	5/4/2018	aggressive beha...

To deselect features, we simply click on the “Deselect features” tool (yellow arrow) at the top of the screen.

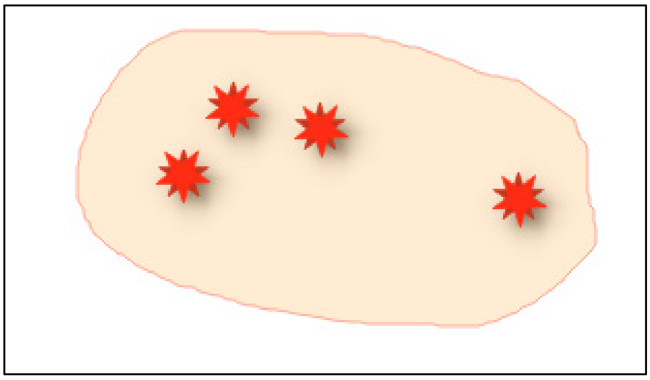


	cell	X	Y	Species	Number Affected	Date	Condition
1	171420	301457.8851033...	4.64243e+06	Mephitis mephitis	1	5/6/2018	dead
2	173898	300955.7100755...	4.64044e+06	Procyon lotor	1	5/30/2018	neurologic signs
3	173567	300563.5555396...	4.64058e+06	Mephitis mephitis	1	4/30/2018	aggressive beha...
4	174637	303228.4314714...	4.63993e+06	Mephitis mephitis	1	4/12/2018	neurologic signs
5	173823	300289.3285808...	4.64049e+06	Mephitis mephitis	1	4/23/2018	road-kill
6	175511	296153.0639728...	4.63891e+06	Mephitis mephitis	2	4/13/2018	dead
7	173295	298285.5810643...	4.64068e+06	Mephitis mephitis	1	4/15/2018	dead
8	173415	300389.6283465...	4.64062e+06	Mephitis mephitis	1	4/18/2018	dead
9	172570	300243.7372987...	4.64137e+06	Procyon lotor	1	4/26/2018	neurologic signs
10	174270	299496.1526369...	4.64023e+06	Mephitis mephitis	1	4/20/2018	dead
11	172272	300936.0176167...	4.64173e+06	Procyon lotor	1	5/2/2018	dead
12	169073	302551.3732611...	4.64414e+06	Mephitis mephitis	1	5/4/2018	aggressive beha...
13	170276	297001.9119184...	4.6433e+06	Mephitis mephitis	1	5/15/2018	neurologic signs

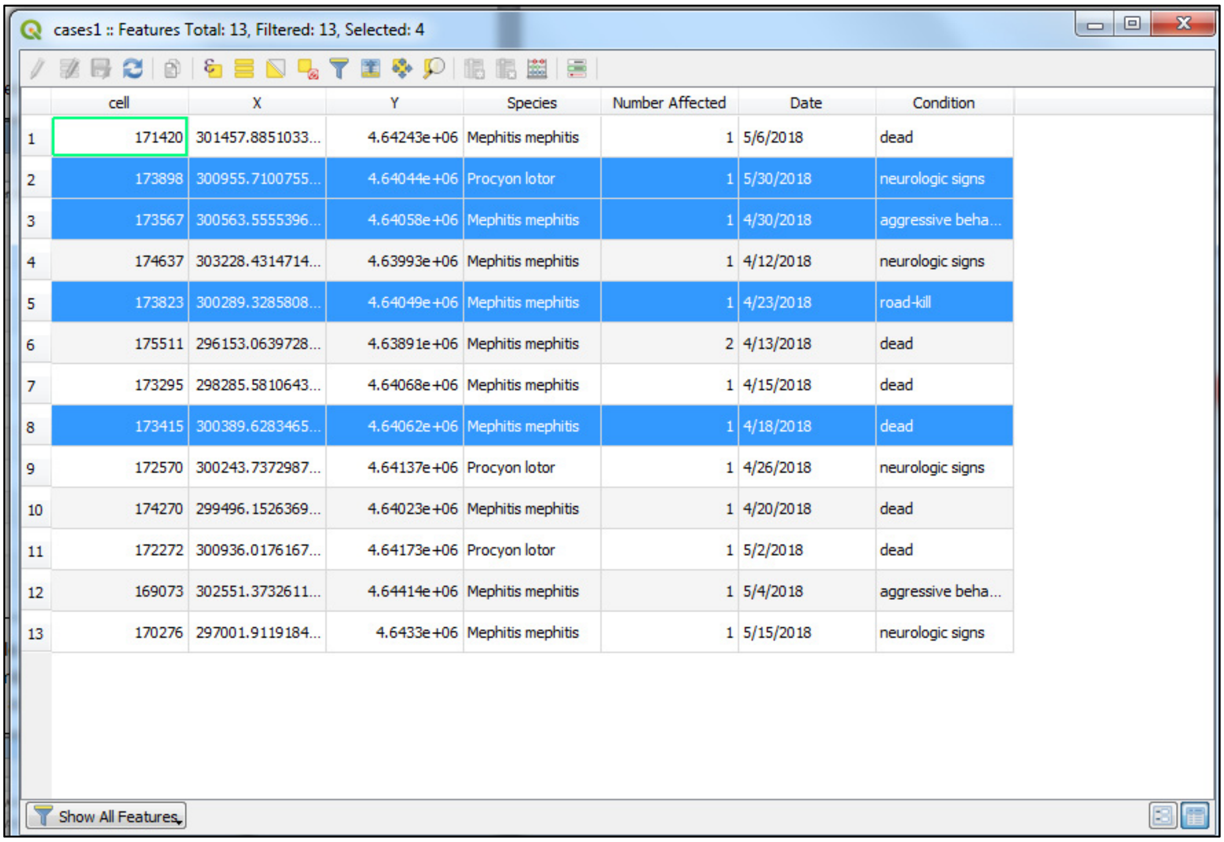
QGIS also provides the ability to alter what features are visible in the attribute table by left clicking on the button on the bottom left of the screen (yellow arrow). There are a number of different options available, and these options are useful particularly when the layer contains a large number of features.



In summary, QGIS provides two different methods for selecting features: 1) feature selection can be accomplished by opening the attribute table and selecting features as just described or 2) by selecting features using the "Select Features by area or single click tool" and selecting the features from the viewing window as we explained earlier. To demonstrate the second method selects the features in the attribute table, we will select features using the select features by "free hand" tool, and then open the attribute table and confirm these features are highlighted.



They are indeed highlighted as shown below.


cases1 :: Features Total: 13, Filtered: 13, Selected: 4

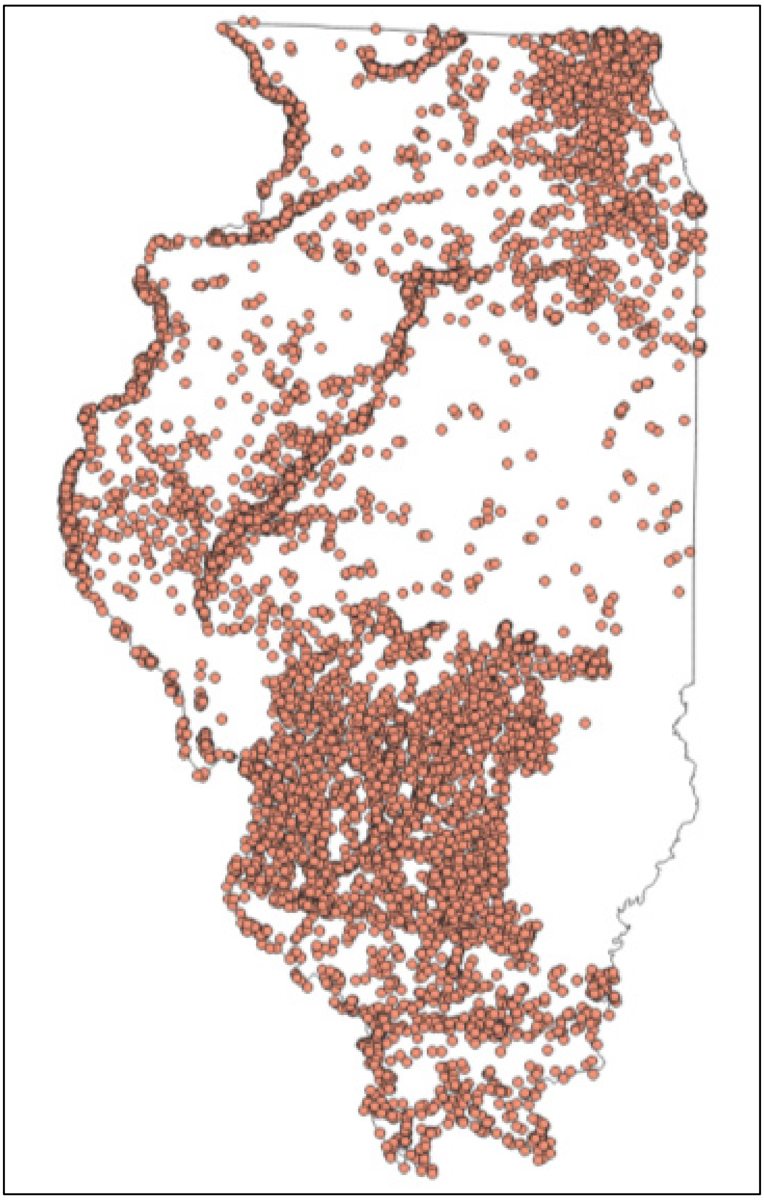
	cell	X	Y	Species	Number Affected	Date	Condition
1	171420	301457.8851033...	4.64243e+06	Mephitis mephitis	1	5/6/2018	dead
2	173898	300955.7100755...	4.64044e+06	Procyon lotor	1	5/30/2018	neurologic signs
3	173567	300563.5555396...	4.64058e+06	Mephitis mephitis	1	4/30/2018	aggressive beha...
4	174637	303228.4314714...	4.63993e+06	Mephitis mephitis	1	4/12/2018	neurologic signs
5	173823	300289.3285808...	4.64049e+06	Mephitis mephitis	1	4/23/2018	road-kill
6	175511	296153.0639728...	4.63891e+06	Mephitis mephitis	2	4/13/2018	dead
7	173295	298285.5810643...	4.64068e+06	Mephitis mephitis	1	4/15/2018	dead
8	173415	300389.6283465...	4.64062e+06	Mephitis mephitis	1	4/18/2018	dead
9	172570	300243.7372987...	4.64137e+06	Procyon lotor	1	4/26/2018	neurologic signs
10	174270	299496.1526369...	4.64023e+06	Mephitis mephitis	1	4/20/2018	dead
11	172272	300936.0176167...	4.64173e+06	Procyon lotor	1	5/2/2018	dead
12	169073	302551.3732611...	4.64414e+06	Mephitis mephitis	1	5/4/2018	aggressive beha...
13	170276	297001.9119184...	4.6433e+06	Mephitis mephitis	1	5/15/2018	neurologic signs

Show All Features

As with most of the QGIS capabilities we have shown thus far, there are numerous additional options and tools for working with attribute data that we are not able to detail here. But, there are many tutorials, help files and documentation for those who would like a more in-depth understanding of QGIS.

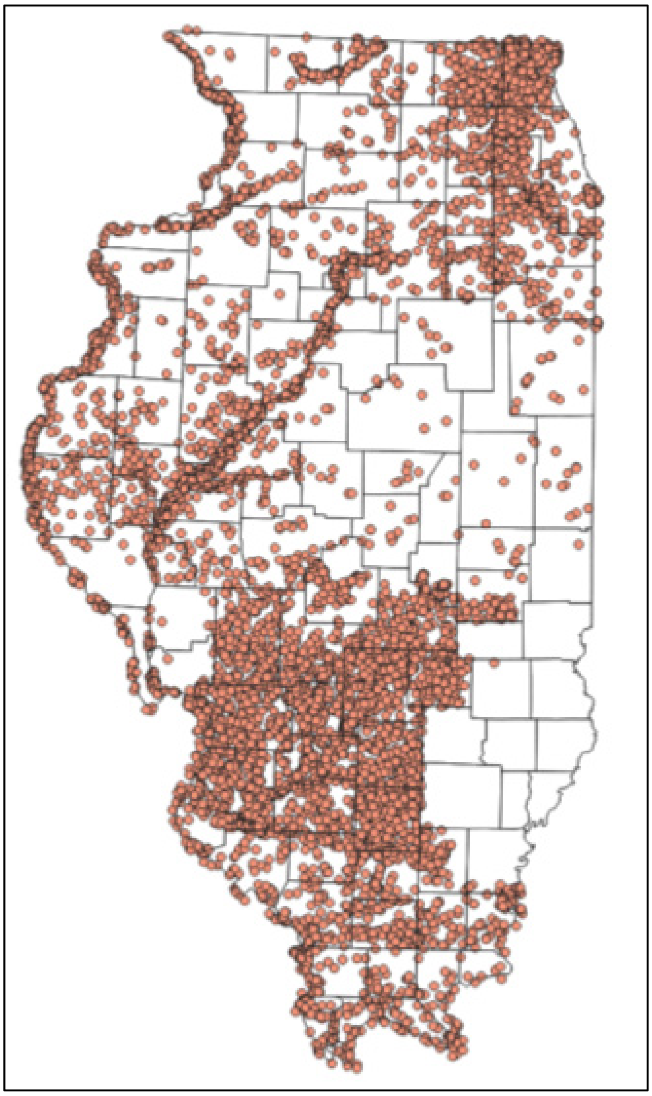
Statistical Mapping

It is common when mapping wildlife health information to display summaries or statistics of the data rather than the raw data. In our example thus far, we have a small dataset consisting of only 13 records; however, let's suppose now that we have a much larger dataset consisting of 10,000 records. If we try to display all the locations of each of these cases, the map would be cluttered and difficult to interpret. To illustrate our point, we will create a new layer from a .csv file named, "pts". We will add the layer as described above using the "Add Delimited Text Layer…" tool. We also set the CRS to NAD 83/UTM using the instructions provided earlier. We then display the information.

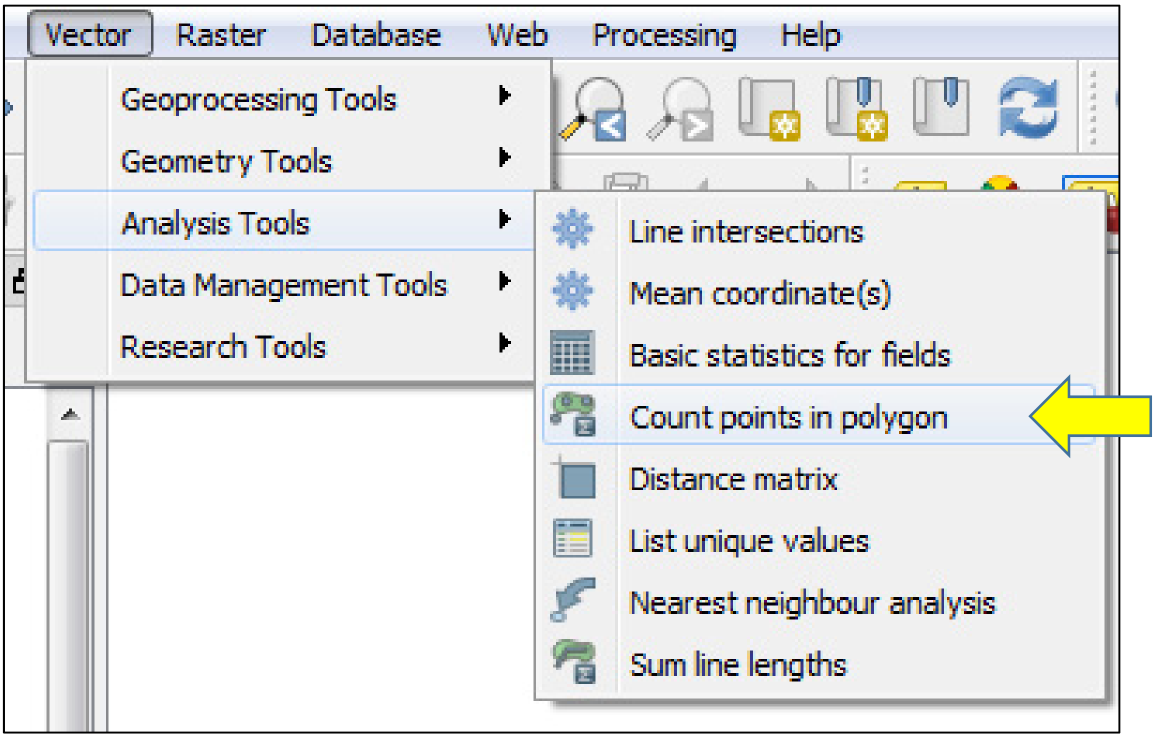


Displaying all 10,000 records in this manner, provides some general perceptions of the data, but it is difficult to comprehend specific or local characteristics of the disease outbreak. In this case, it would be more useful to provide some statistical summary of the data that would allow the map viewer to quickly understand important aspects (e.g., areas with the greatest number of cases or highest intensity).

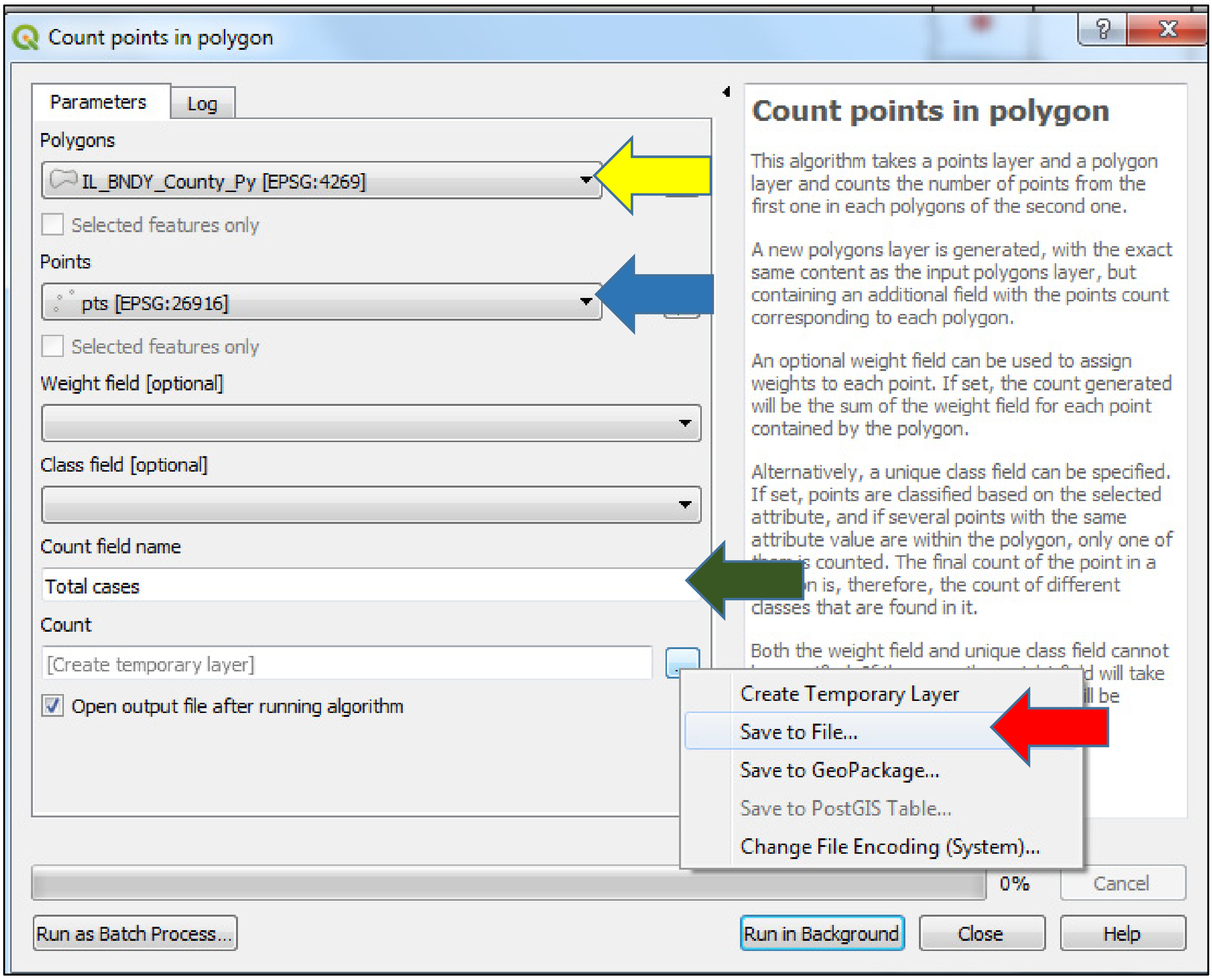
One common technique for creating statistical maps is to discretize the region into aerial units. In other words break up the region of interest into a number of non-overlapping sections. Then create summaries for each unit. This discretization permits the comparison of these units. In our example, we discretize or break up the state of Illinois into polygons of various shapes and sizes. These polygons will correspond to the counties, which are administrative districts within the state. To do this, we will first add a polygon vector layer of the counties using the “Add Vector Layer…” tool described previously. The result of this addition is shown below. We can now see the aerial units that we will use to create statistical maps.



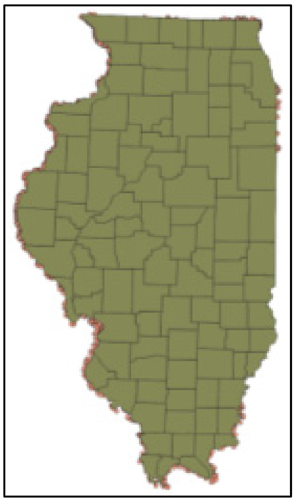
Let’s first create a map that displays the number of cases in each county. The first step is to tabulate or count the number of points within each county polygon. This would be a tedious, error-prone task if done manually. Luckily, QGIS has a simple tool that will complete the counts automatically. Open the “Vector” menu at the top of the screen, select the “Analysis Tools” option, and then choose the “Count points in polygon” tool (yellow arrow).



This will open a dialog window that gives us a number of options. First, we will choose the name of the polygon layer from which you would like to count points (yellow arrow), and then we will select the points layer that you would like to summarize in each polygon (blue arrow). Next we will change the name of the field that will be displayed in the attribute table from “Numpoints” to “Total cases” (green arrow). We also would like to save this file rather than just creating a temporary file that would not be available for future work after closing the project. To save the file, we click on the “…” button (red arrow) under the “Count” heading. This will open a dialog box where we will select “Save to File”. This will open a browser window from which we can navigate to the directory on our computer we would like to save the file, and will allow us to assign a name to it. We can also select the type of file to save. Generally, this will be a shapefile, so we select the .shp extension. Next we left click on the “Save” button. Then we select “Run in Background”, and close the Window. There are additional tool options that can be explored. This will create a new layer in our “Layers” pane named “Count”. The layer will be a vector data layer composed of polygons, and will automatically have the same CRS as the polygons file not the points file. We can check this by examining the layer’s properties, and if the CRS is not correct it can be changed using the techniques already detailed above to match the project’s or any other layer’s CRS.

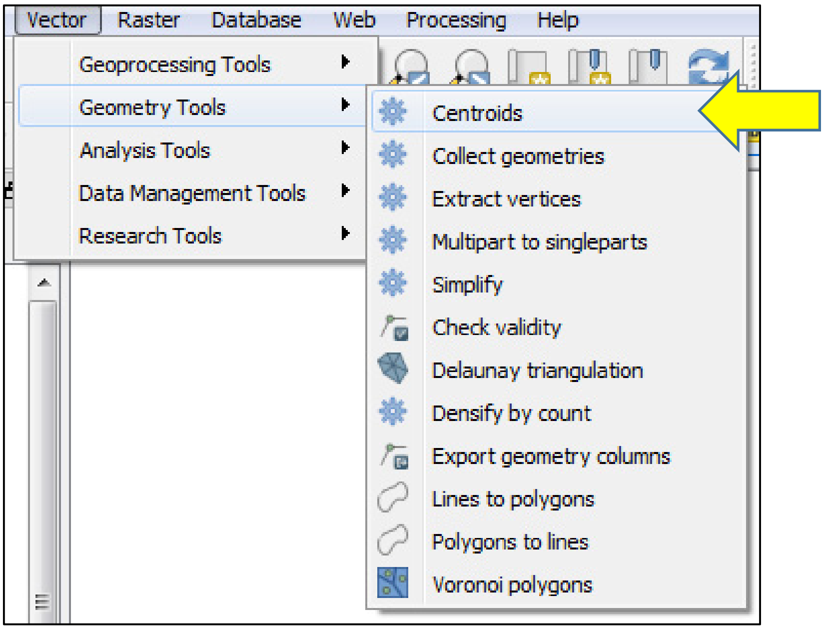


The count layer is displayed as a series of filled polygons, which is not very useful for displaying differences in case numbers by county.

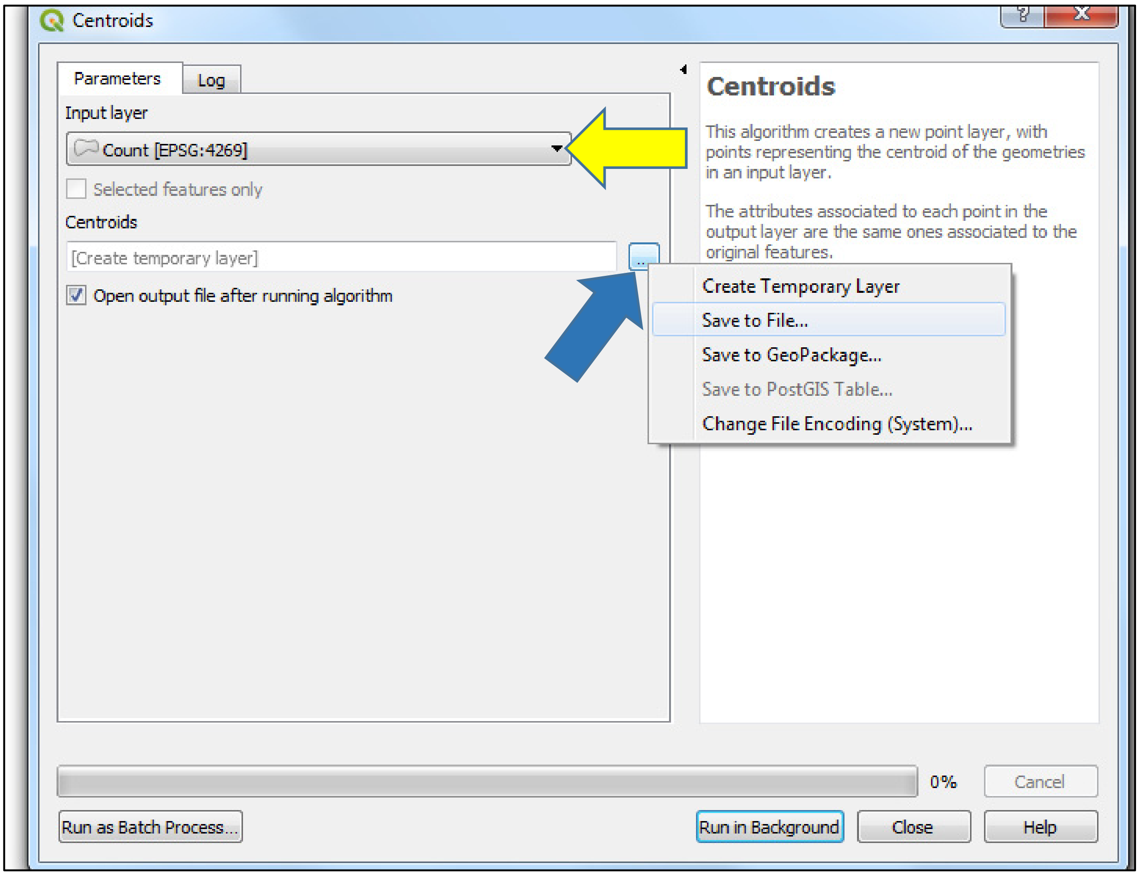


Graduated Symbols Map

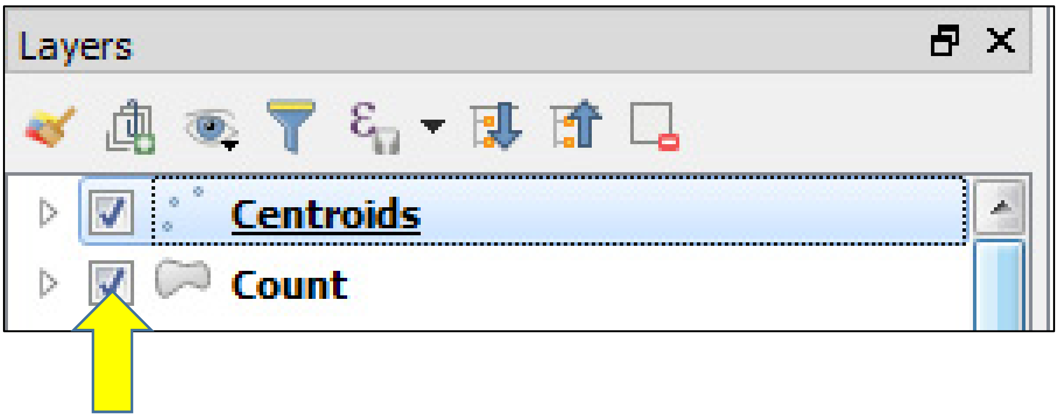
To improve our map, we will create a new point layer from this polygon layer. Note, we switch to a point layer because it is easier to specify graduated symbols for point layers compared to polygon layers. Specifically, generating the legend when finalizing the map is problematic when using polygon layers. To create the new point layer, open the Vector menu at the top of the screen, select “Geometry Tools”, and choose “Centroids”.

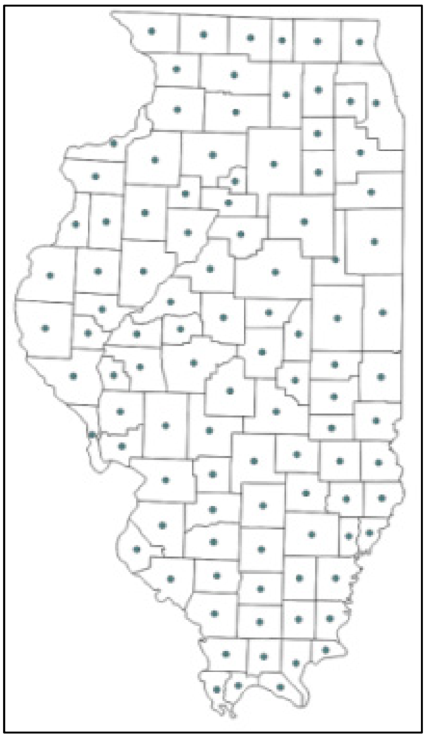


The following dialog box allows you to select the polygon file to use in creation (yellow arrow). In our example that is the “Count” layer we just created. We will want to save the file for future use, so we will save it as a shapefile (.shp). If we do not save it, QGIS will only create a temporary file for the centroids, which will not be available for future use or when we reopen our project in the future after closing it. To save the file, we left click on the “…” button (blue arrow) under the “Centroids” heading. This will open a dialog box where we will select “Save to File”. This will open a browser window from which we can navigate to the directory on our computer we would like to save the file, and will allow us to assign a name to it. We can also select the type of file to save. Generally, this will be a shapefile, so we select the .shp extension. Next we left click on the “Save” button. Lastly, we left click the “Run in Background” button, and close the window. This tool will create a layer containing a point, located at the centroid of each polygon.

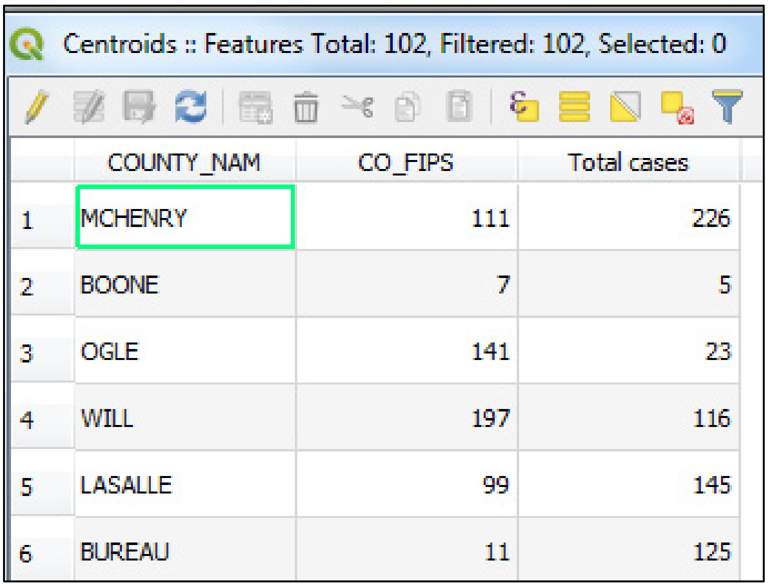


This produces a point layer of the centroids of each of our polygons, which is displayed below, after turning off the "Count" layer by unclicking the checkbox in the "Layers" pane (yellow arrow).





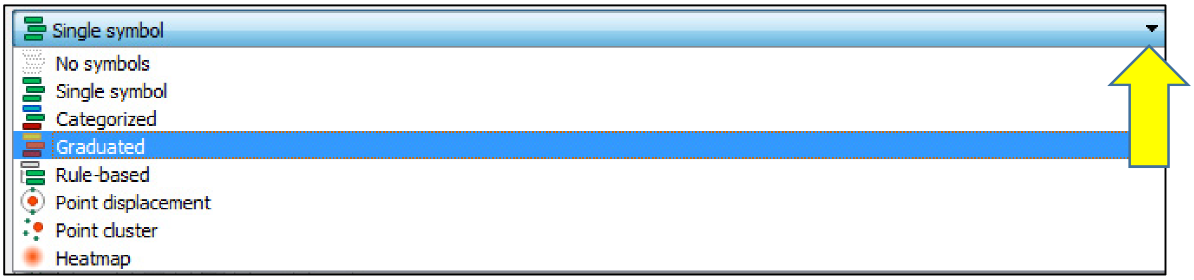
Additionally, if we open the attribute table we can see the attribute information including the county name, county id and total number of cases from the polygon layer associated with each point.



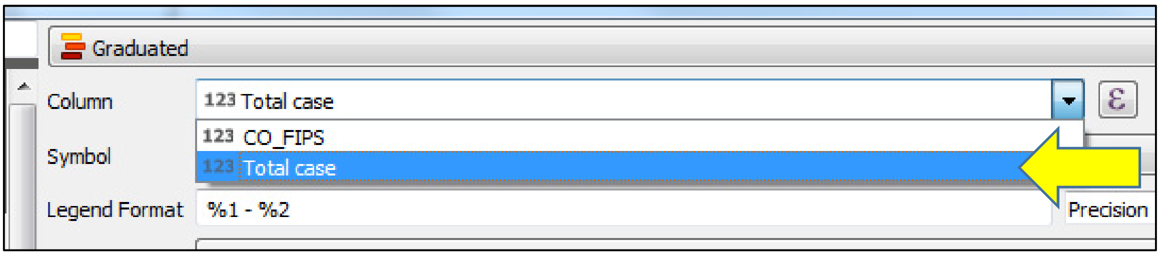
	COUNTY_NAM	CO_FIPS	Total cases
1	MCHENRY	111	226
2	BOONE	7	5
3	OGLE	141	23
4	WILL	197	116
5	LASALLE	99	145
6	BUREAU	11	125

To make our map more informative, we would like to change the symbology of this new point layer, which is called "Centroids", to have a graduated symbols map based on the number of cases in each county. Graduated symbology is a technique of displaying quantitative data such that the size of the symbol corresponds to the class associated with the magnitude of the point's data value (e.g., Total cases). This technique bins or groups the quantitative data into a number of classes to permit ease of viewing. For example, we might create a class that contains all values for the total number of cases that are from 100-200. This class would then receive some size of symbol. Any classes consisting of total number of cases less than this 100 would have smaller symbols, and any classes greater than 200 would have larger symbols. Thus, when viewed, the relative magnitude of the number of cases reported in each county is quickly seen and understood.

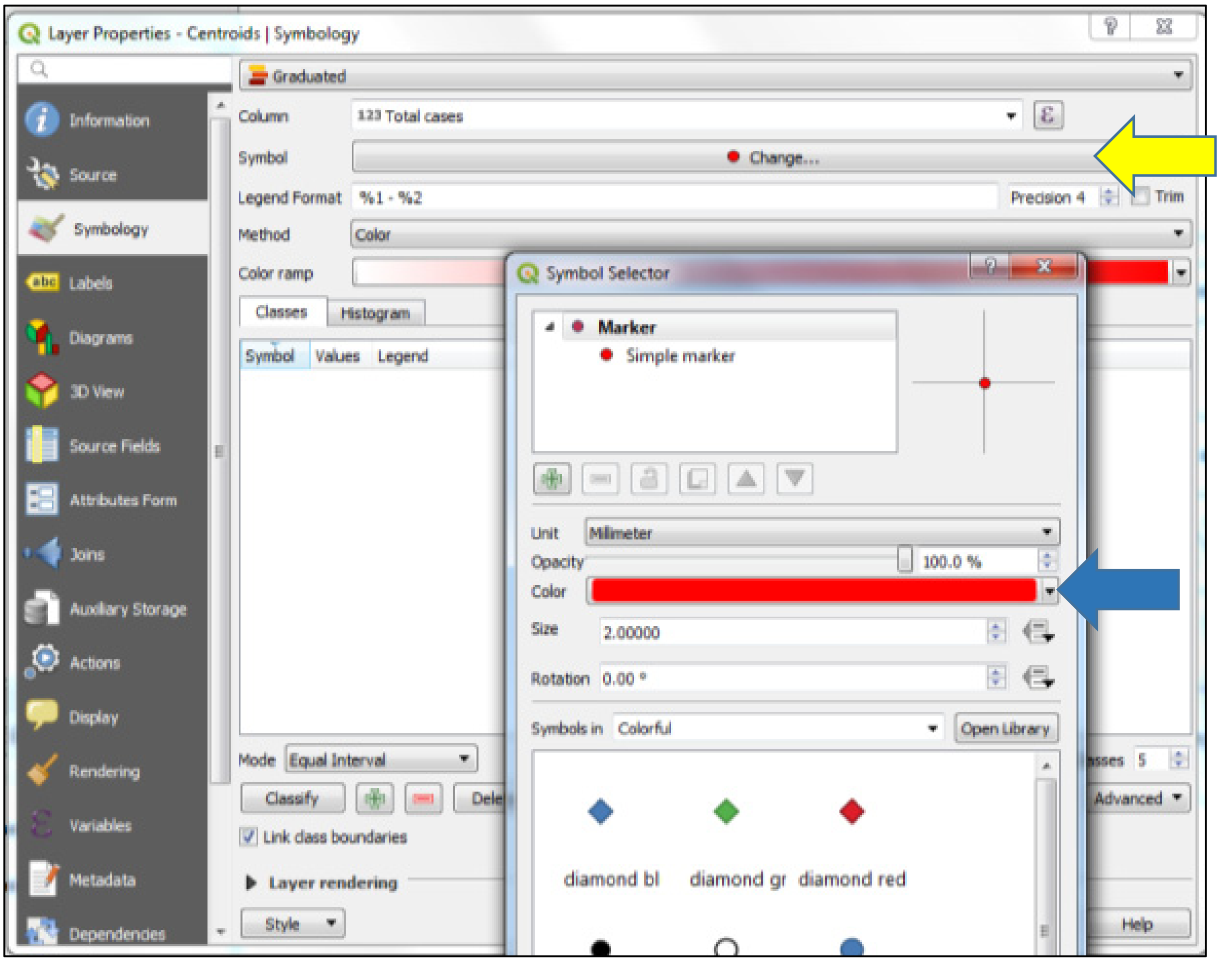
To create a graduate symbols maps, we right click on the "Centroids" layer in the "Layers" pane, and select "Properties" and then "Symbology". This will open the window that will allow us to change how our statistical results are depicted on the map. First, we will change from "Single symbol" to "Graduated" by clicking on the black triangle at the top of the window (yellow arrow), and selecting "Graduated" from the drop down list.



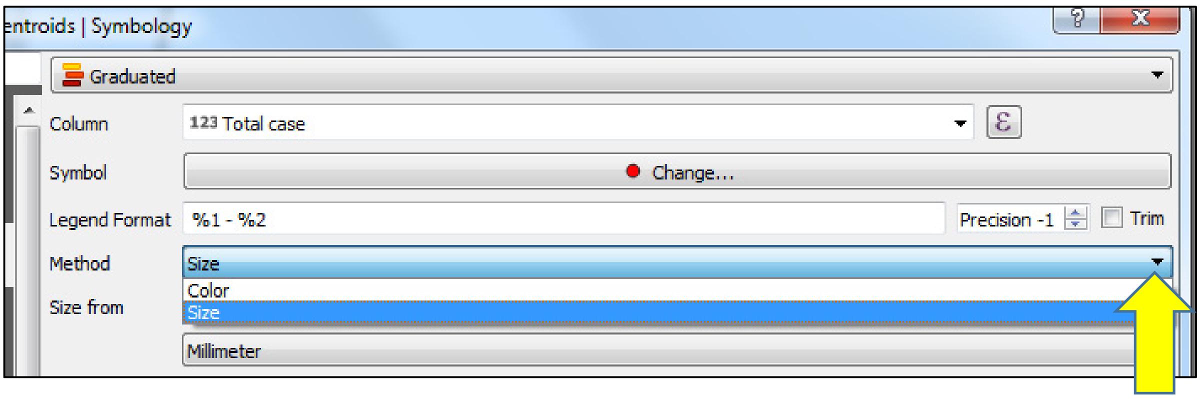
This will change the window and give us options for creating graduated symbol maps. In this new window, we first choose which field from our attribute table that we will use to create the differences in sizes of the symbols. For our example, we will select "Total cases".



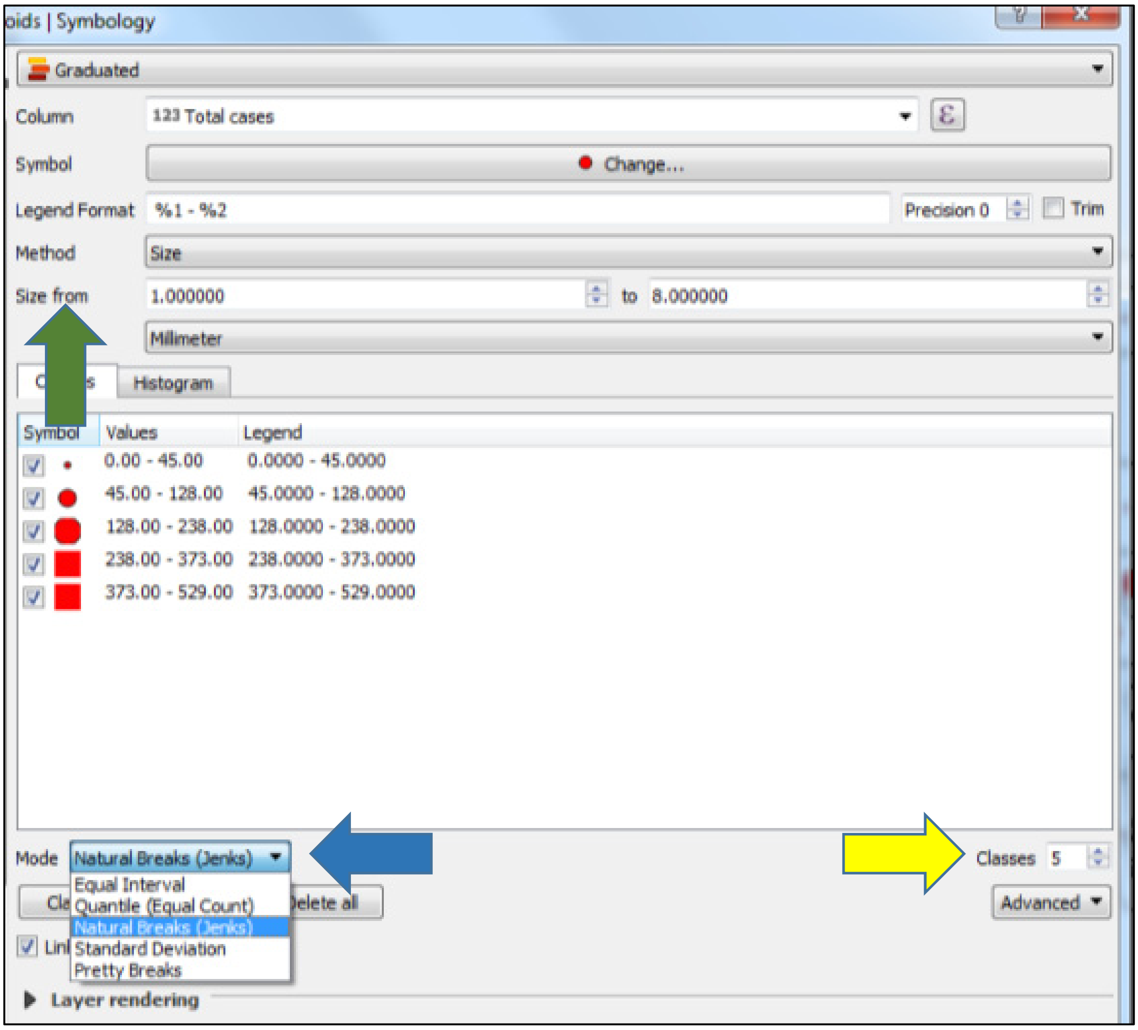
We will then change the color of the symbol to red. We select the "Change" symbol option (yellow arrow). This generates a new window from which we select the color of our choice (blue arrow). For our example, we select red. Then select "OK" to close the window.



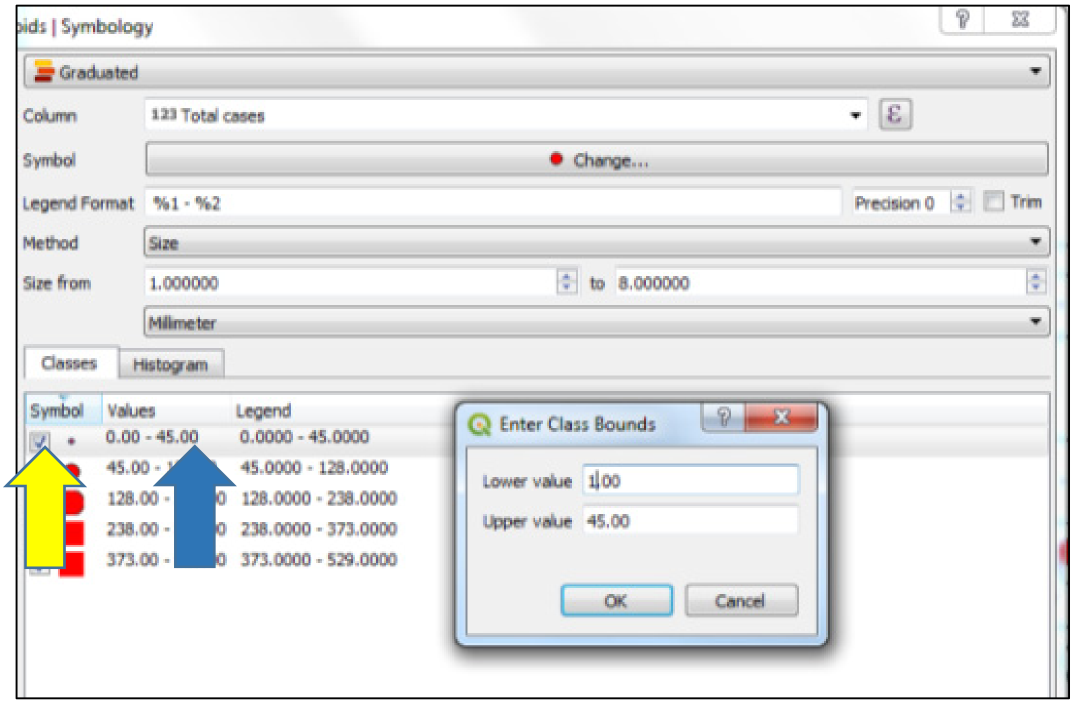
Next, we change the method (yellow arrow) from "Color" to "Size". This indicates that we want our symbols to be different sizes based on the total number of cases for each county, rather than having one symbol with different colors for the varying total number of cases.



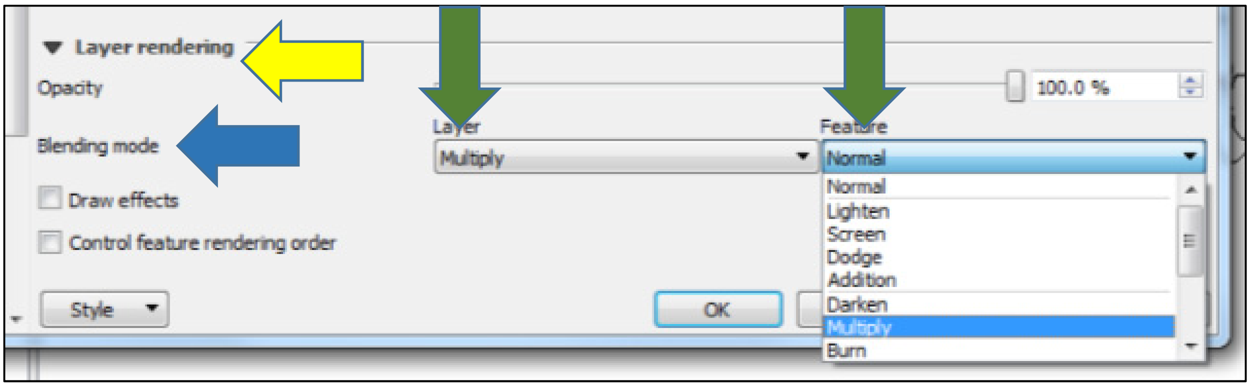
We now will place each point into a specific class. We will leave the default as 5 classes (yellow arrow), but this could be increased/decreased. There are a number of QGIS options (blue arrow) that we can use to create the bins/classes to which we will assign each point. For this example, we will select the "Natural breaks (Jenks)" option that uses the empirical distribution of the data to determine how we should break up our data. This creates 5 different classes and shows the boundaries for each class, and the size of the symbol increases as the boundary values increase. Each point will be placed in the class for which it is greater than or equal to the lower boundary and is less than the upper boundary. For this example, we will leave the range of sizes that the classes will take as the default range (i.e., 1-8; green arrow).



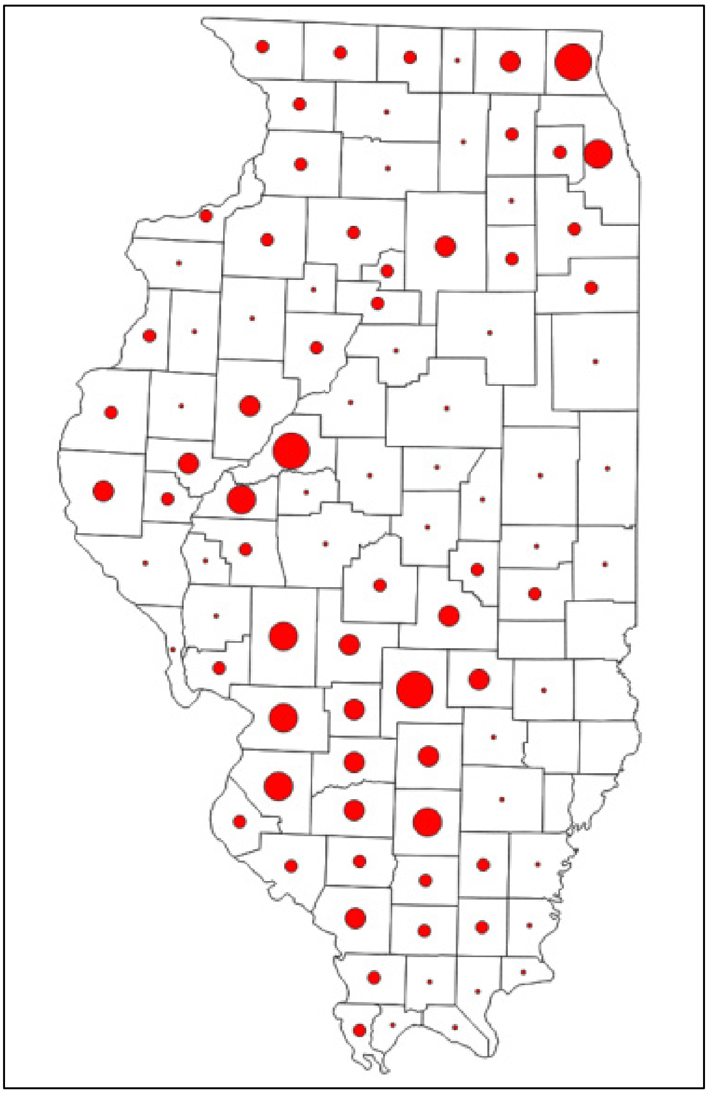
Looking at the boundary values for each class above, the first class starts at zero. We want this class to start at one rather than zero. This will allows us to readily distinguish counties with no cases because they will lack any point symbol. To make this boundary change, double left click on the values (yellow arrow) for the symbol that needs to have the boundaries values changed. This creates a dialog box in which we change the lower bound from zero to one. This may need to be repeated for the "Legend" bounds (blue arrow) as well. Then left click "OK".



Lastly, we want to make sure that overlapping symbols will be visible, and any layers with spatial data below the point symbols for this layer can be seen. This is particularly important if the large symbols might obscure smaller symbols or key features of underlying layers. Thus, we will adjust the options for "Layer rendering". We click on the small black triangle (yellow arrow). Then under the "Blending mode" (blue arrow), we select "Multiply" for the "Layer" and the "Feature" options (green arrows). This will complete our symbology changes, and therefore we click "OK".



The final map is displayed below. The counties with a large number of cases are quite evident and easily distinguished. Additionally, the counties without any cases are distinct because their centroid is not visible. These types of maps provide useful summaries of a large amount of data, and can be readily shared with the public or decision makers without the need for significant explanation.

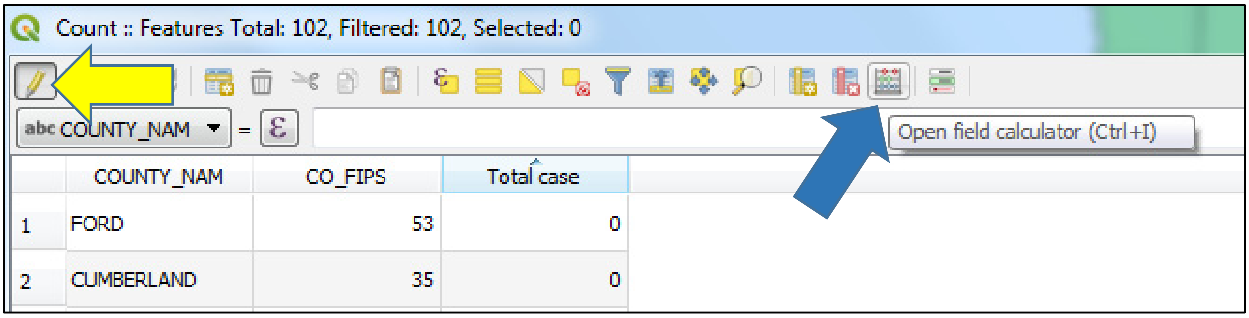


Choropleth Map

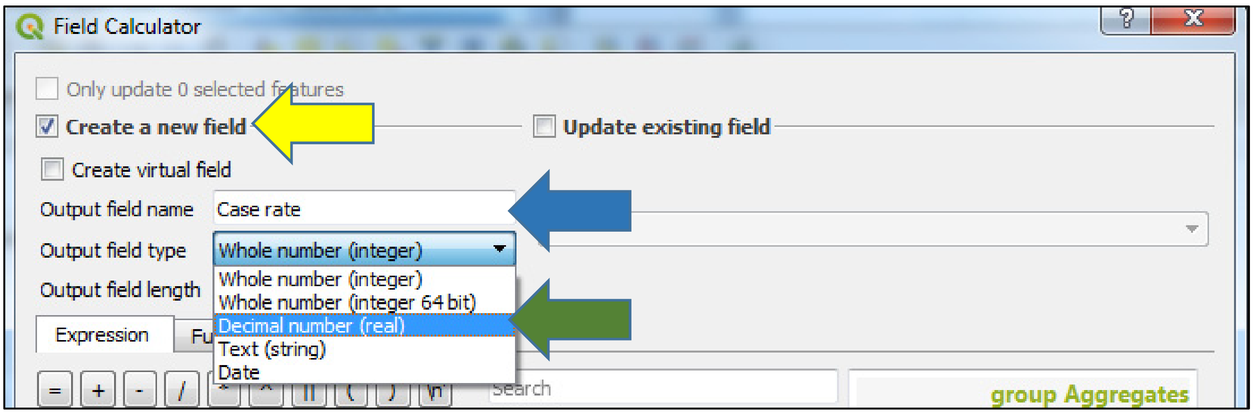
Choropleth maps are one of the most common types of maps for displaying wildlife health information. They provide an effective means of displaying rate or intensity information. Similar to the graduated symbols map above, they discretize the region of interest and assign each discrete unit to a specific class. However, rather than using the size of symbol to describe the magnitude of the attribute data, they fill each discrete unit with different colors that represent the varying classes of intensities or rates (e.g., prevalence, cases/km^2, etc.).

To demonstrate how to construct a choropleth map in QGIS, we will use the same dataset that we used to create the graduated symbols map. However, instead of displaying the raw number of cases, we will create a choropleth map to show the number of cases/km^2. In essence, we are standardizing the number of cases in each discrete unit or county by the amount of area encompassed in the boundaries of the county. This is useful to determine if the county has a large number of cases because it is a large county, or if the intensity of the county is actually higher than other counties in the state of Illinois.

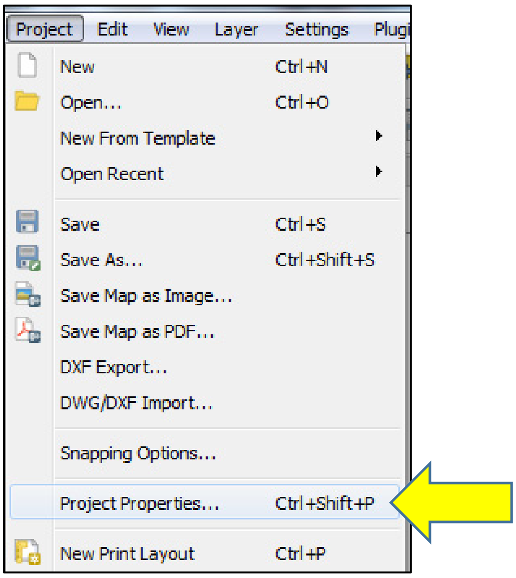
We begin by determining the cases/km² for each county. We open the attribute table for the "Count" layer we created earlier by right clicking on the layer and selecting "Open Attribute Table". We need to add a new column to this table, which requires us to turn on the "Edit" mode for the table. By left clicking on the "Edit" tool (yellow arrow). This allows us to edit the attribute table. Next we will left click on the "Field calculator" tool (blue arrow). This will open a new dialog window.



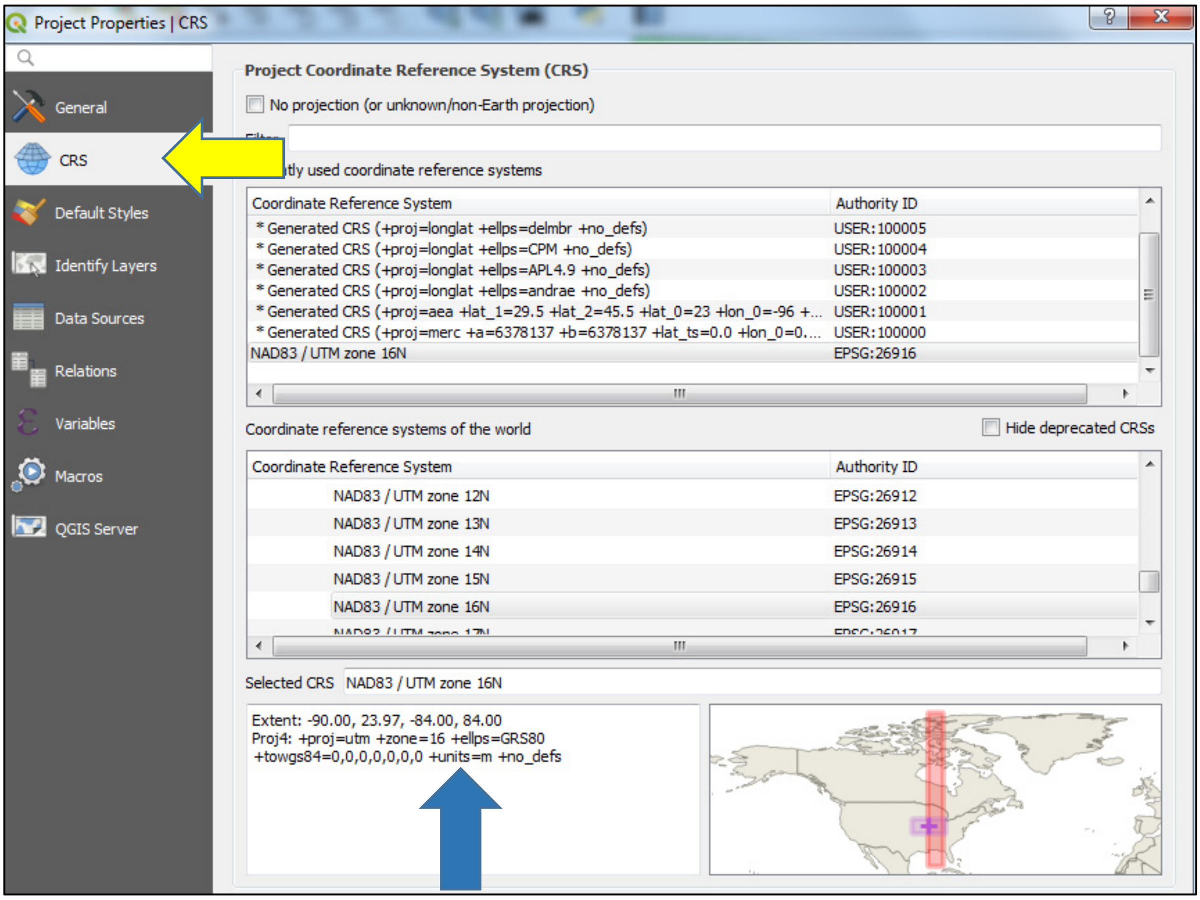
In this new window, we begin by checking the "Create a new field" box (yellow arrow). We name our "Output field name" as "Case rate" (blue arrow). We then choose the "Output field type", which we select as "Decimal number (real)" (green arrow) because we are dealing with a rate.



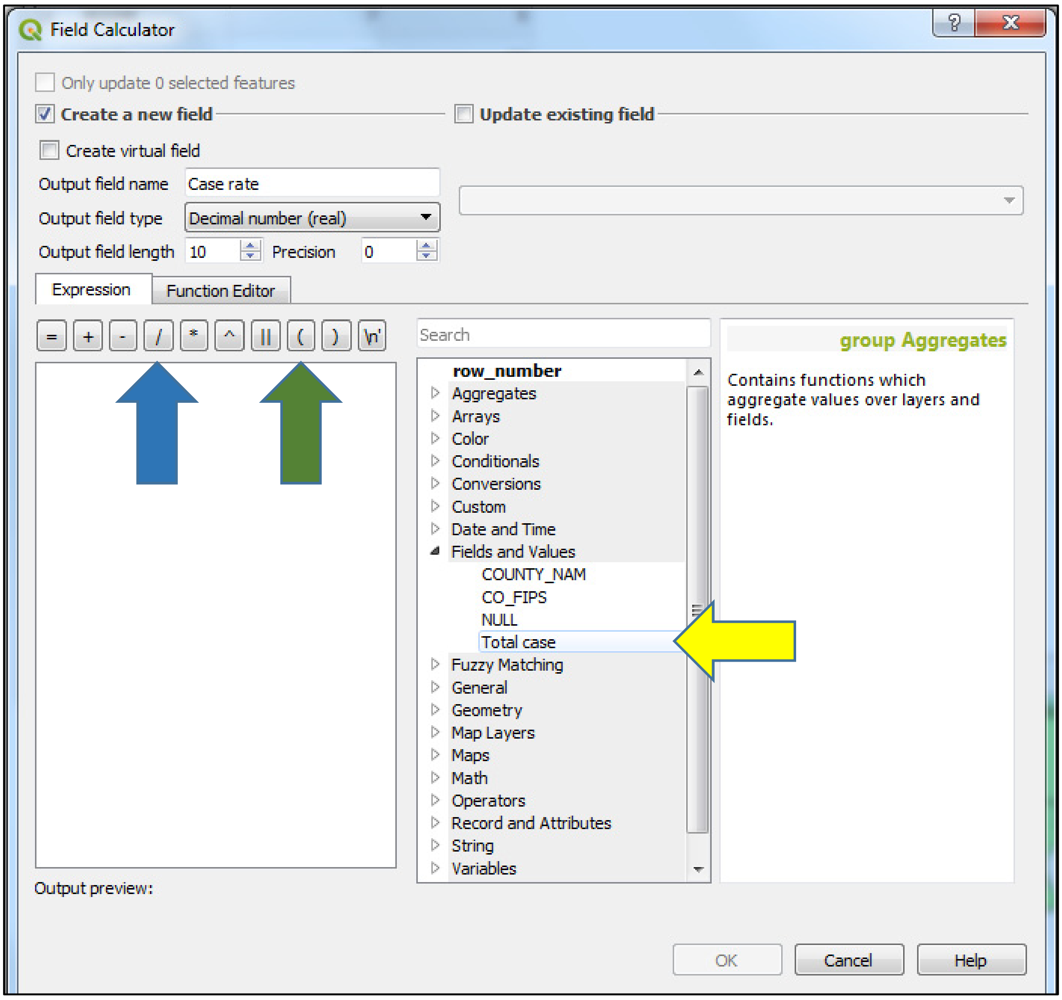
Now, we need to specify the expression for calculating the cases/km² in each county. There are many useful functions available in the "Field Calculator" to help the user conduct a wide array of analyses based on the attribute information of a layer. In our example, we will create a simple expression: "Total case"/(area$/(1000*1000)). This expression asks QGIS to take the number of cases and divide it by the area of each county. However, because the area units are in meters², we divide by 1000*1000 to convert them into km². If you are unsure of the units of your project, you can determine the units by clicking on the "Project" menu at the top of the screen, and then selecting "Project Properties".



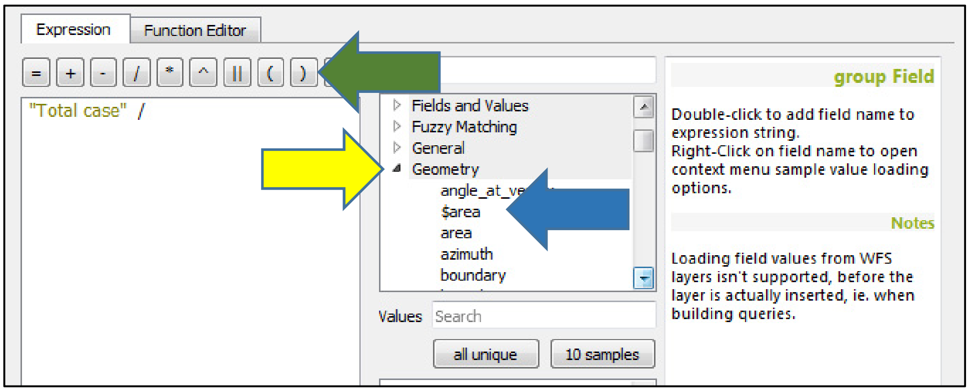
This will open the information window for the project. Select CRS (yellow arrow), and then look to see what units are associated with the project's CRS (blue arrow). In this case, we see the units are "m" indicating meters.



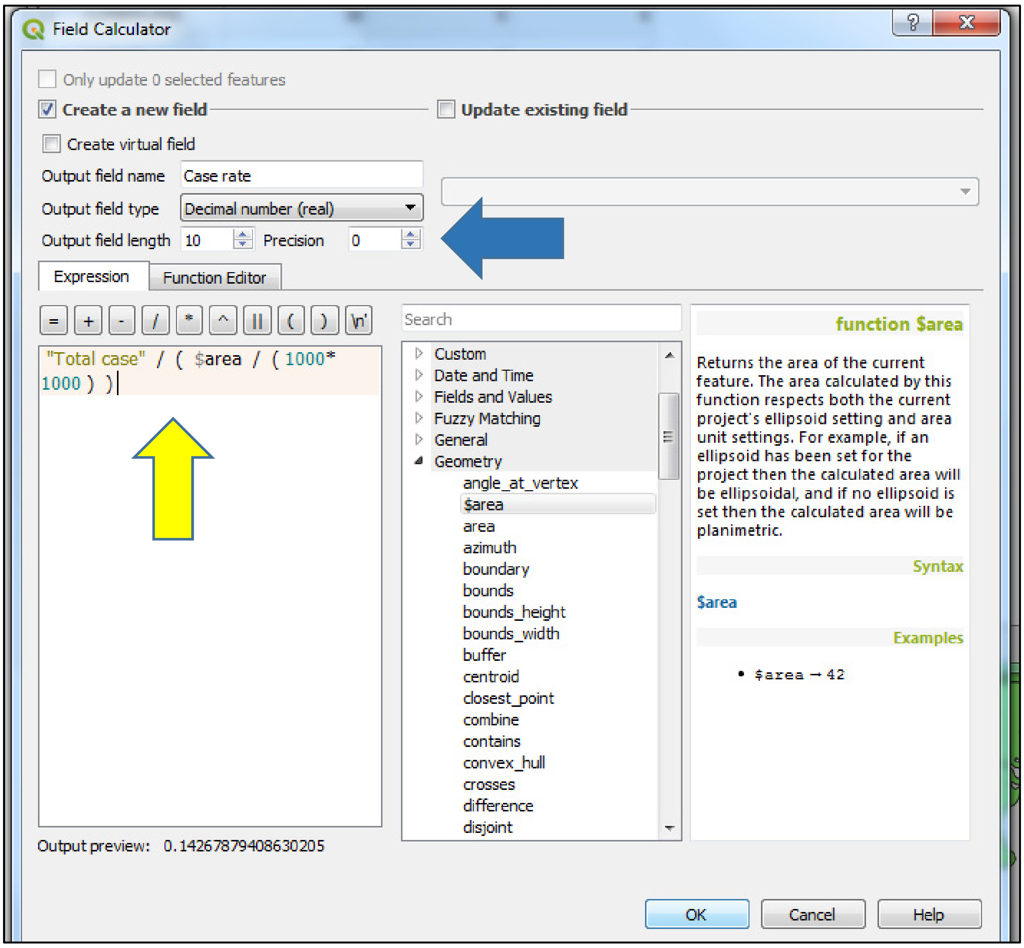
Now that we have established our units are indeed in meters, we need to create our expression we defined above in QGIS. In the "Field Calculator" window, we left click on the arrow next to the "Field and Values" group, and double left click on the "Total case" field.



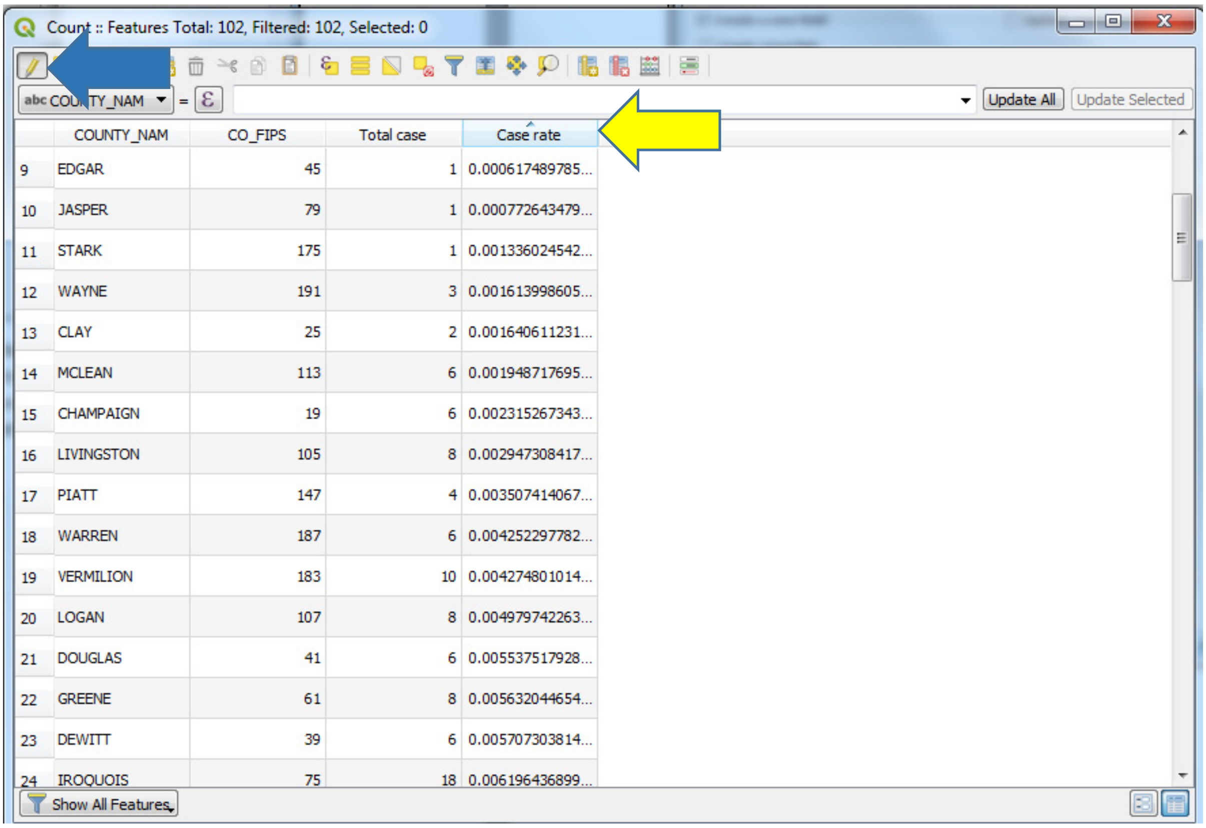
This will add the field name, "Total case" to our expression. Next, we left click on the "/" or divide by button (blue arrow, above), and on the left parentheses button (green arrow). Then we left click on the arrow next to the "Geometry" group (yellow arrow, below) and double left click on "$area" (blue arrow, below). This tells QGIS that our expression requires the area of each county to be calculated.



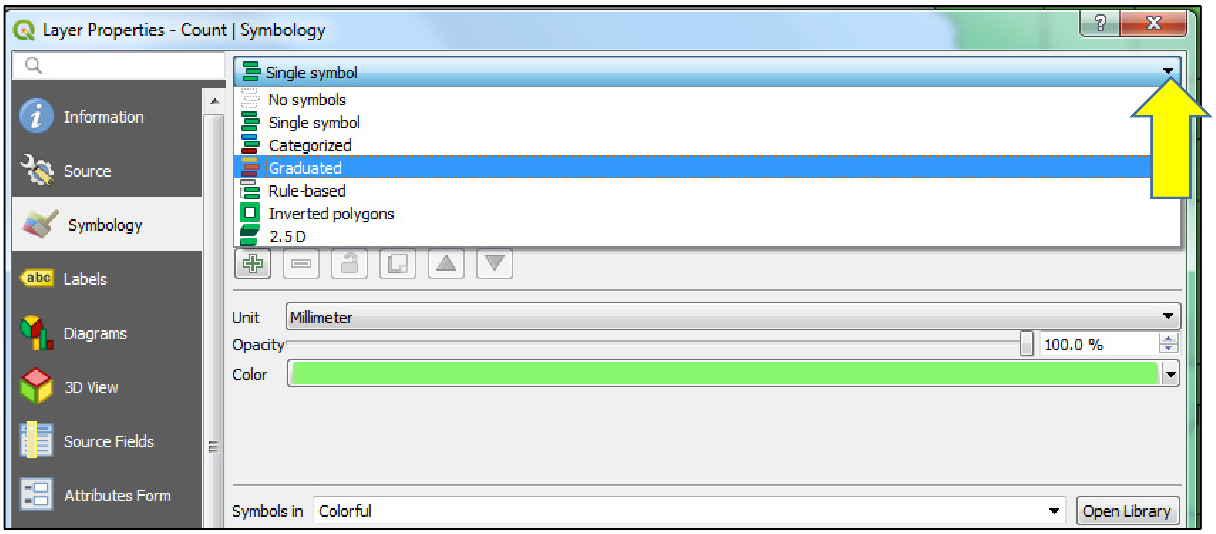
We conclude our expression, by left clicking on the divide by button, followed by the left parentheses button, typing “1000*1000”, and concluding by left clicking on the right parentheses button twice (green arrow, above). We can see that our expression now matches the one we described earlier. We also need to set the “Precision” of our new field from the default of zero to five (i.e., five digits after the decimal point; blue arrow). If we don’t set the precision, when we save the table QGIS will round to the nearest whole number (i.e., precision=0). Finally, we click the “OK” button at the bottom of the window.



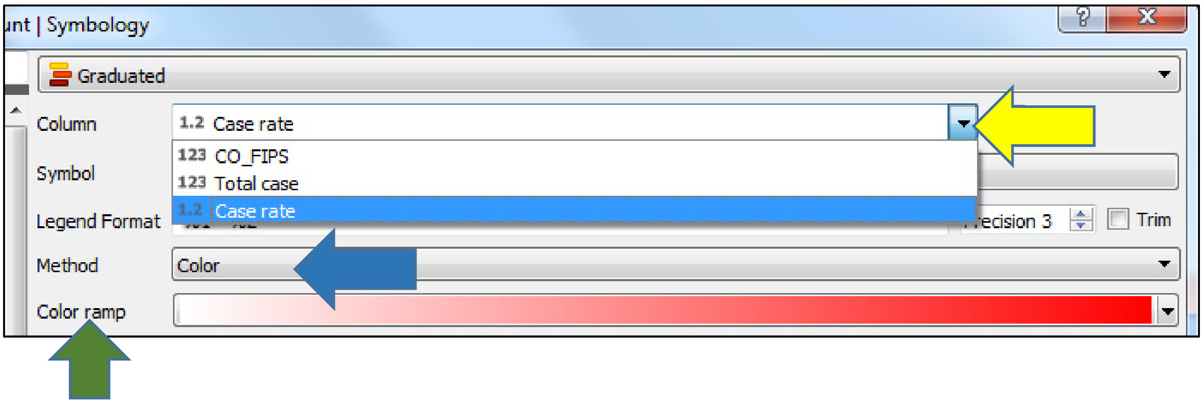
Now, when we open our attribute table, we can see that we have a new column called, "Case rate" that contains the number of cases/km² for every county. To stop editing the table we click on the "Edit" tool (blue arrow) and save our changes. This provides us the requisite data from which we will create our choropleth map.



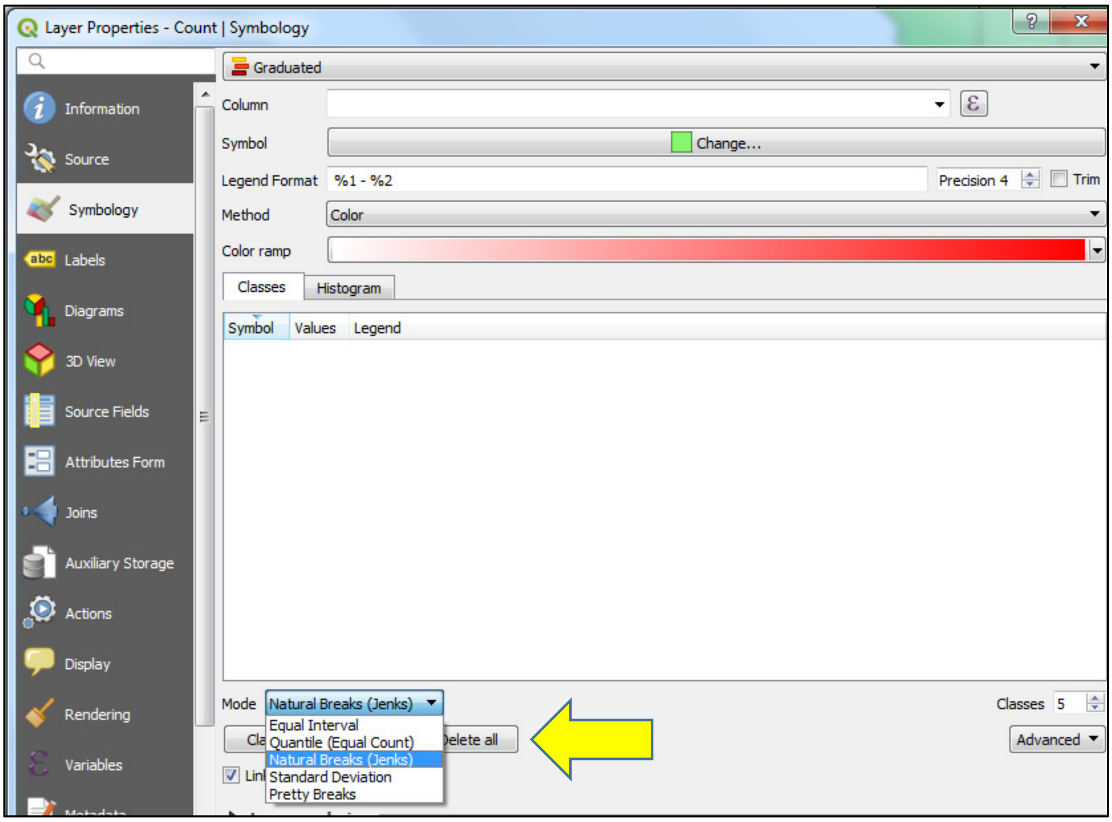
To build our map, we right click on the "County" layer in the "Layers" pane, and select "Properties". We then select "Symbology", and change the option at the top of the window from "Single symbol" to "Graduated" by left clicking on the black triangle (yellow arrow), and selecting "Graduated".



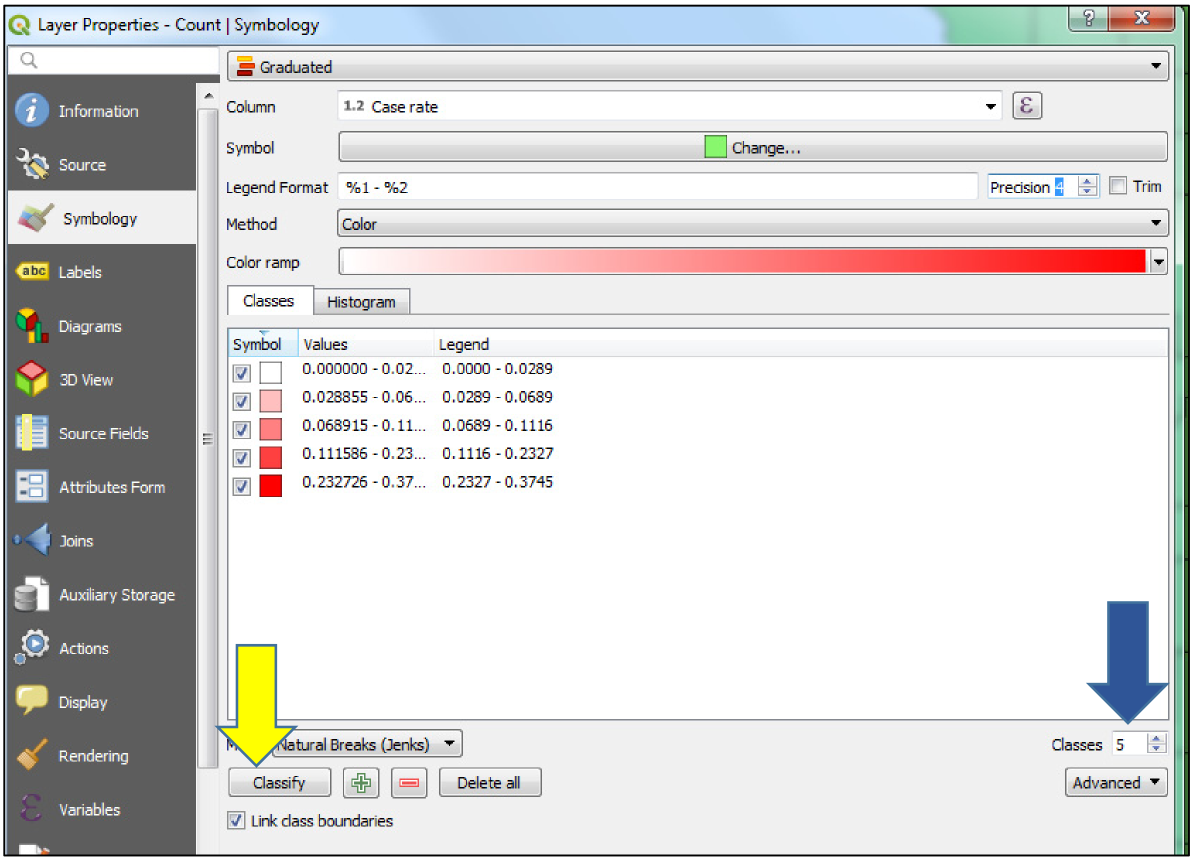
This opens a new window that is identical to the one we saw previously with the graduated symbols map. We first select the field from the attribute table we would like to be depicted as varying color classes on our map (i.e., “Case rate”)” by left clicking on the black triangle next to the Column row (yellow arrow). We will leave the Method (blue arrow) as “Color”, and use the default “Color ramp” (green arrow); however, you can choose from a large number of available color ramps, or create your own.



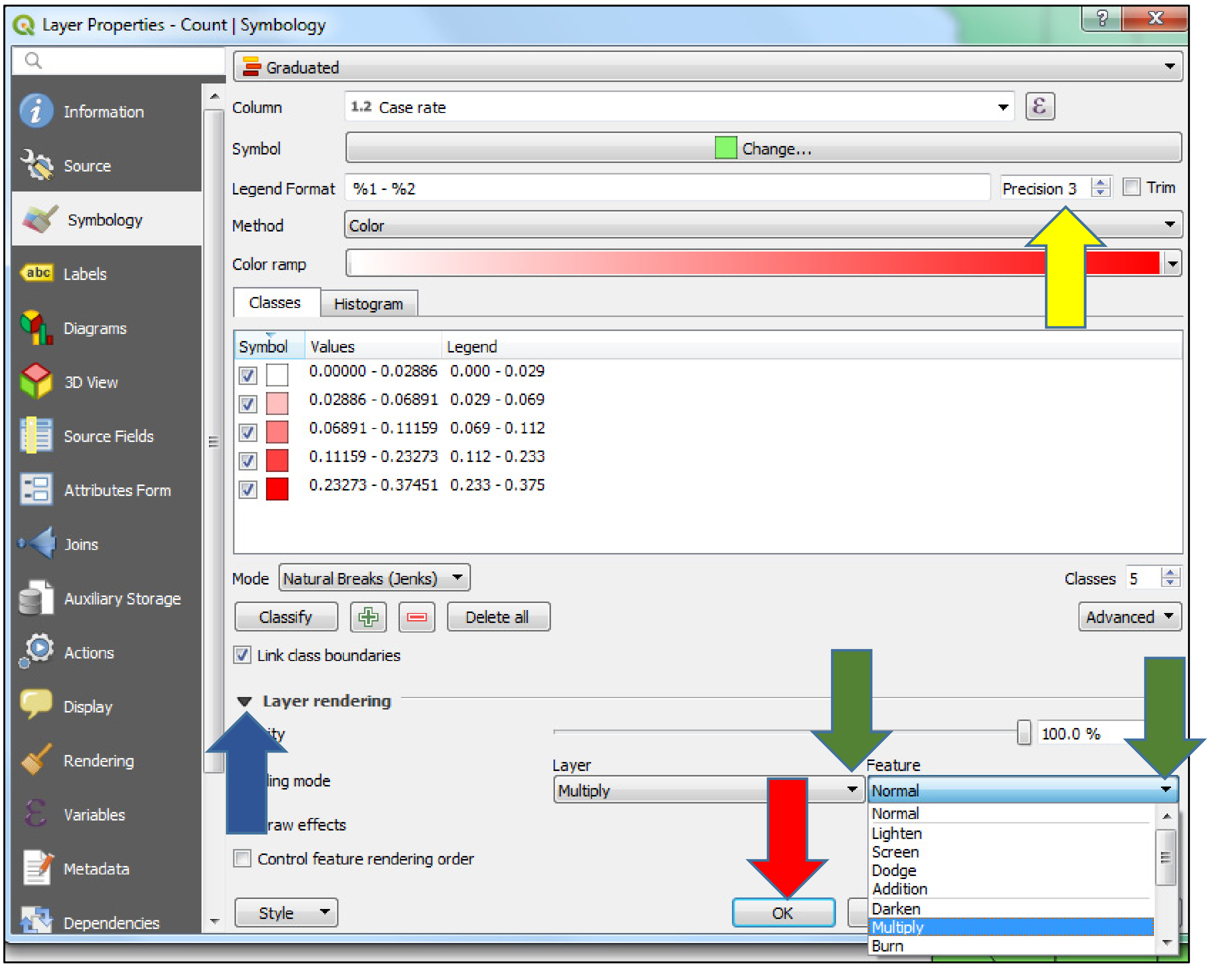
We now specify how QGIS will break up our values for “Case rate” into discrete classes. We left click on the black triangle (yellow arrow) associated with the “Mode”, and choose “Natural breaks (Jenks)”. There are other options that you can experiment with if desired.



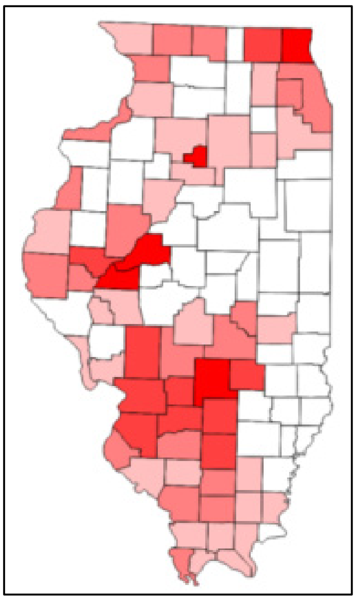
We then left click on classify (yellow arrow) to have QGIS create the classes. Note we will leave the number of classes as 5 (blue arrow), but this can be increased or decreased as desired. We can now see the 5 classes, associated boundary values, and the colors with which they will be displayed.



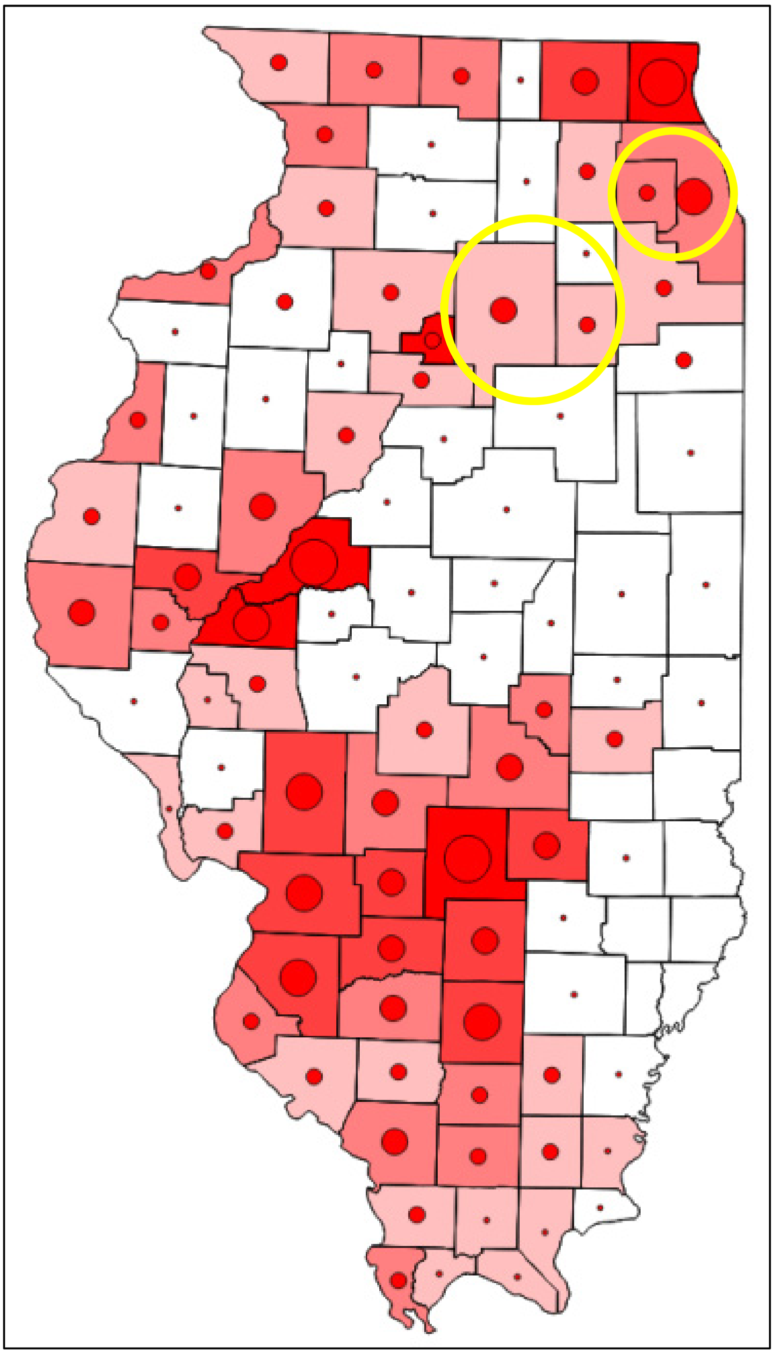
For our "Legend" values we would like to only have 3 decimal places displayed. So, we change the "Precision" (yellow arrow) from 4 to 3. The legend values are now rounded to 3 decimal places. We also would like underlying map layers to be visible below our colors. So, we click on the black triangle next to "Layer rendering" (blue arrow). We choose "Multiply" for the "Layer" and "Feature" options. We then click "OK" (red arrow) to create our choropleth map.



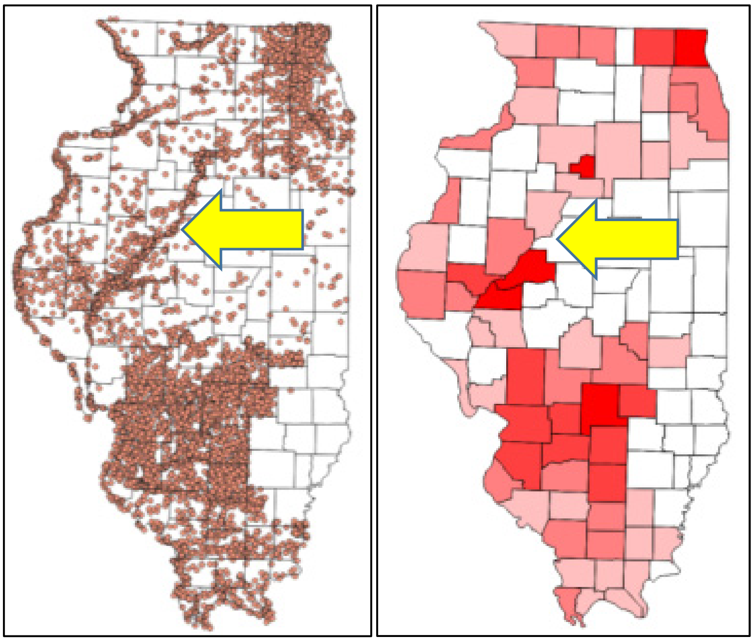
Here is our final choropleth map.



We can see that in general this is very similar representation of our wildlife health information as displayed in the graduated symbols map, which depicted the total number of cases per county. However, there are a few counties where the choropleth map shows similarities between counties that the graduated symbols map depicts as considerably different. To see this, we will turn on our point layer called "Centroids" by left clicking on the check box next to the name of the "Centroids" layer in the "Layers" pane. Also assure that the "Centroids" layer is at the top of the layer stack. The result is a map that displays both our choropleth as well as graduated symbols maps. Examples of differences between the two maps are circled in yellow.



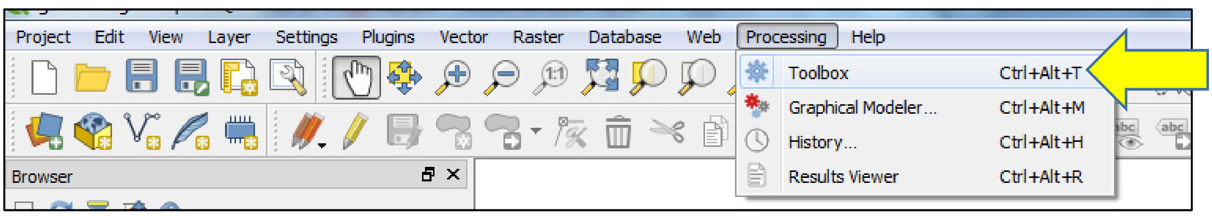
It is worth noting, for ease of quickly understanding large-scale patterns, discretization and statistical summaries can mask fine-scale spatial patterns. In our example, when we looked at the map of the raw data, it is clear that the cases have some interesting spatial patterns. For example, there seems to be a linear string of cases through the central portion of the state (yellow arrow); however, this information is much less evident when we examine the choropleth maps.



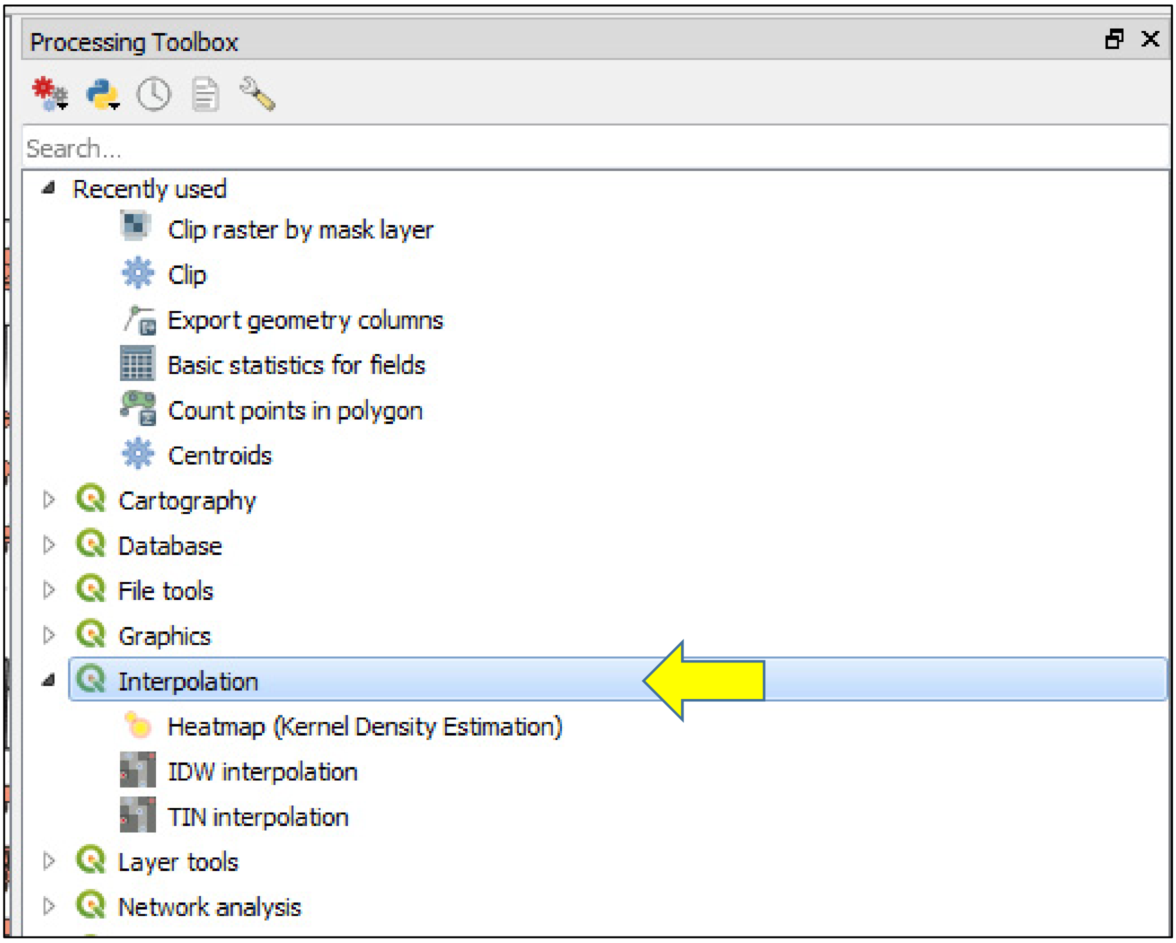
Heat map

The final type of GIS map we are going to show how to create in this training is a heat map or kernel density map. This type of map uses a non-parametric statistical model based on the point locations of our disease cases to estimate a probability density function for the number of cases in our regions of interest (e.g., Illinois). The result is a continuous surface or heat map that displays the relative level of number of cases in Illinois. This type of map doesn't require us to discretize the space into aerial/discrete units like the graduated symbols or choropleth maps. Note: heat maps can only be created from point layers.

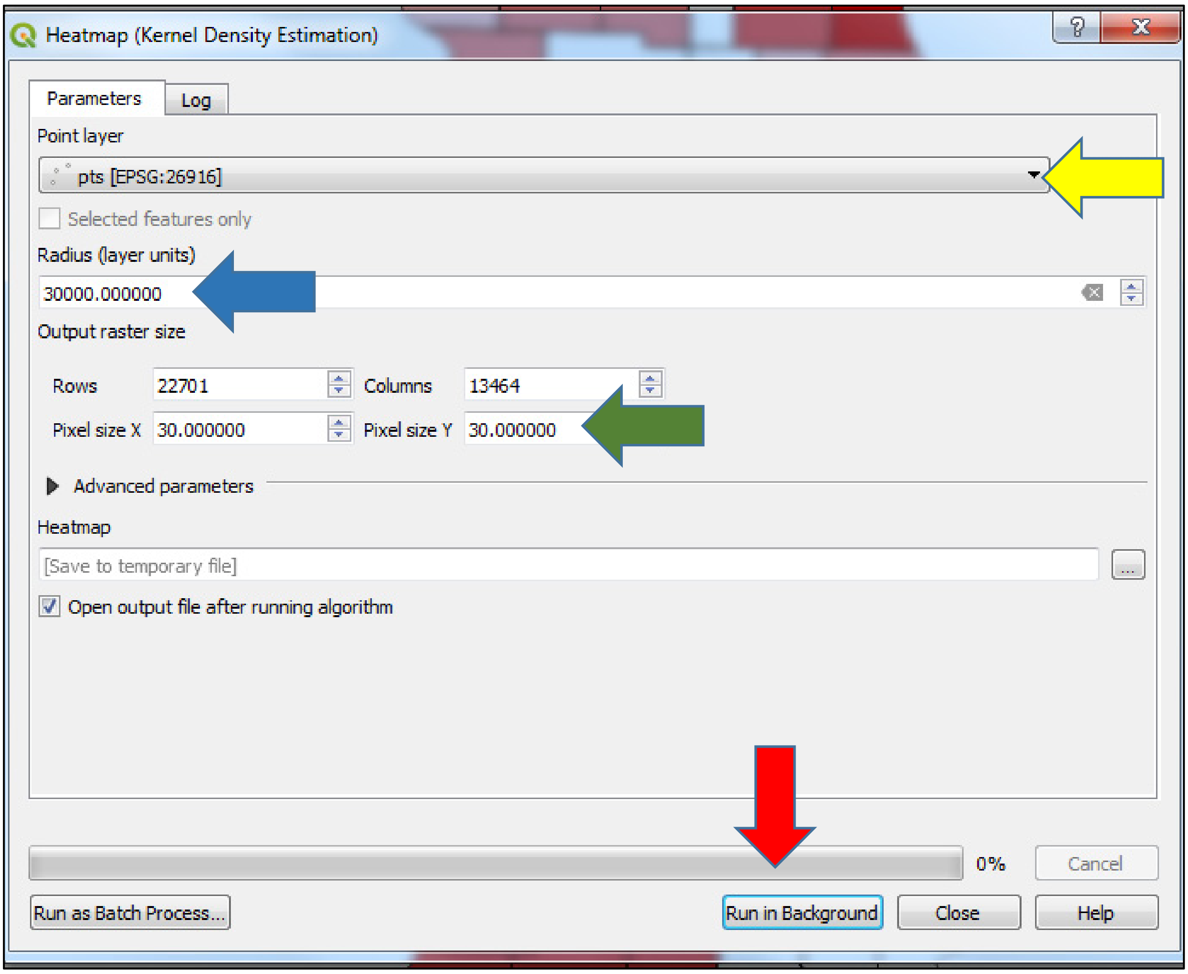
To create a heat map in QGIS, we open the "Processing" menu at the top of the QGIS window and select "Toolbox" to display geoprocessing tools available in QGIS.



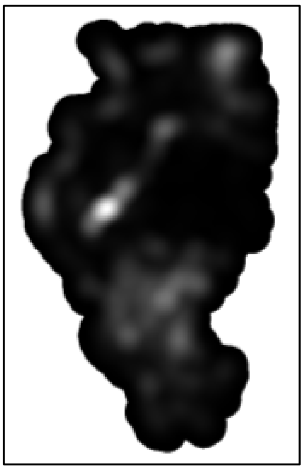
This opens the toolbox window. The tool we are interested in is under the "Interpolation" group, which we access by left clicking on the arrow next to "Interpolation". We then double left click on the "Heatmap" tool (yellow arrow).



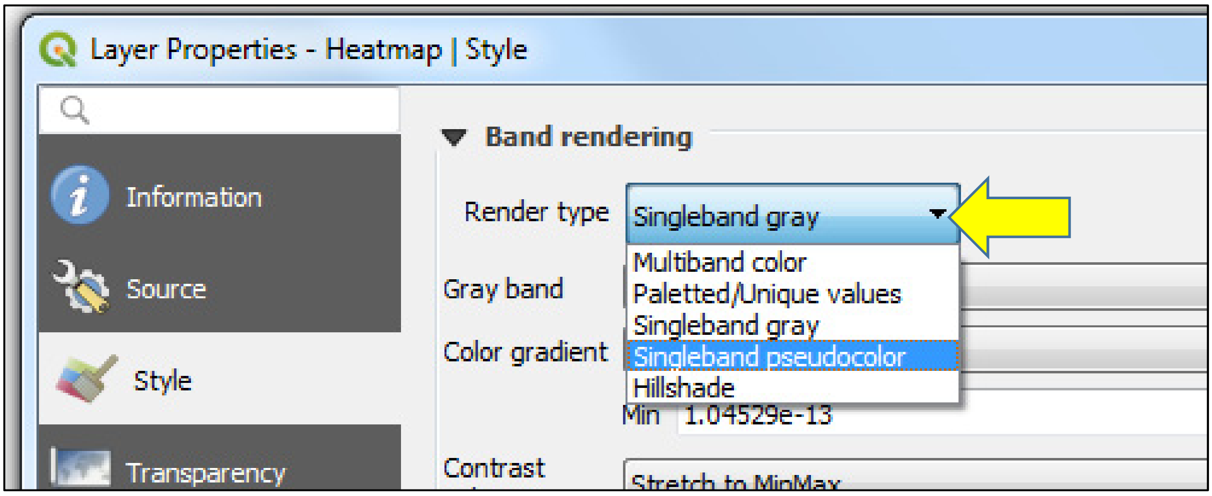
This opens the dialog window for the Heatmap tool. We select our points layer from which we would like to create a heat map (i.e., "pts"; yellow arrow). We then need to specify our "Radius" (blue arrow). The radius is the parameter that determines the amount of smoothing that will be done across the region of interest. A heuristic explanation is that given a specific location on the map, all points within the radius of that location will contribute to the kernel density estimate at that location. Any points outside the radius will not affect the kernel density estimate at that location. From this description, it is clear that a larger radius will make a smoother heat map, whereas a smaller radius will make a less smooth kernel density surface (i.e., radius of zero ≈ returns a map of the points). For our example, we will set our radius to 30,000 meters, which will generate a smooth surface. Note that as the radius increases so does the computing time. Lastly, because the area covered by our map is the entire state of Illinois, we will set the "Output raster size" to have "Pixel size X and Pixel Size Y" = 30 meters (green arrow). If you set the pixel sizes to be too small (e.g., 1 meter), QGIS may not be able to generate the heat map likely due to memory restrictions. It is worth emphasizing that the heat maps that this tool produces are raster images, which we discussed previously, and that is why we need to specify the resolution/pixel size of our heat map. We will just save this map to a temporary file because we will do further manipulations with it, and we will save these manipulations to file for future use. We conclude by clicking "Run in Background" (red arrow).



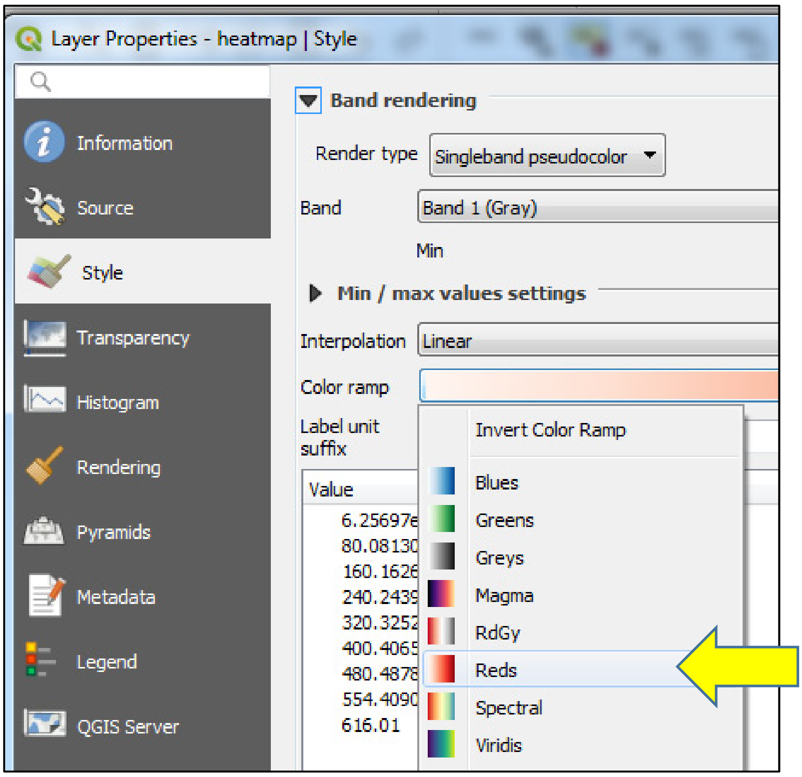
The kernel density calculations may take some time depending on the size of the "Radius" you specified, the pixel size you selected, and the number of points in the points layer that was used. Once QGIS has finished running a new map is produced that shows the relative density of disease cases throughout Illinois in gray-scale with black representing low density and white representing high density.



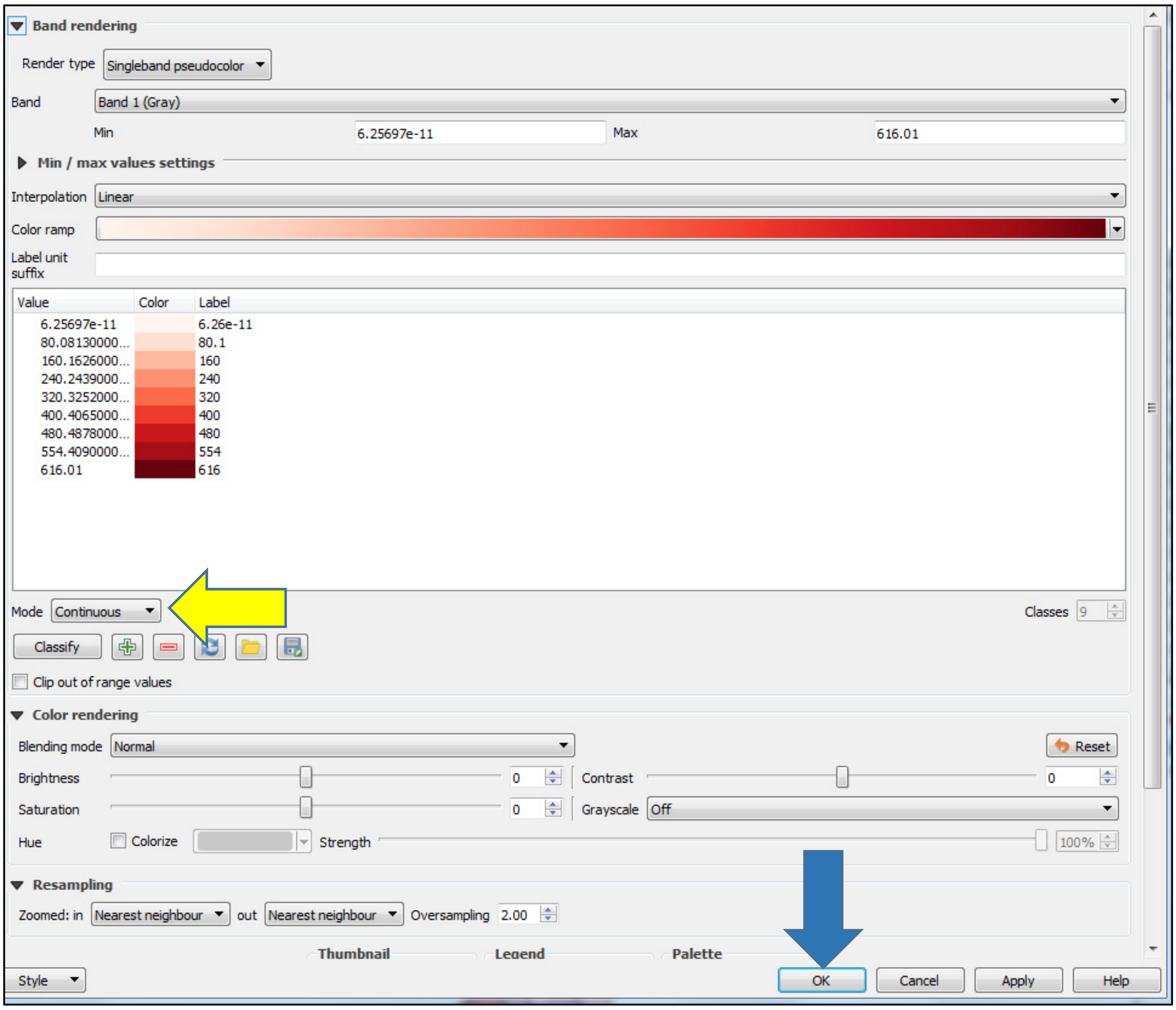
This gray-scale representation of the heat map is not visually appealing for our problem. So we will change the color ramp that is used to display the raster. We right click on the raster layer's name (e.g., Heatmap) in the "Layers" pane, and select properties. We then choose "Style" in the dialog window that opens. Next we click on the black triangle next to "Render type" (yellow arrow), and select the "Singleband pseudocolor" option.



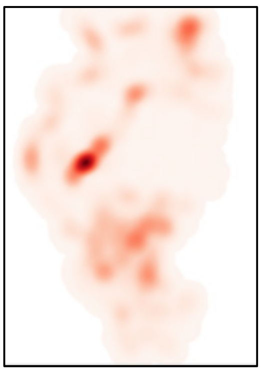
This will allow us to choose a color depiction of our raster image. We left click on the black triangle for "Color ramp", and select "Reds" (yellow arrow) to match the previous maps we have constructed thus far.



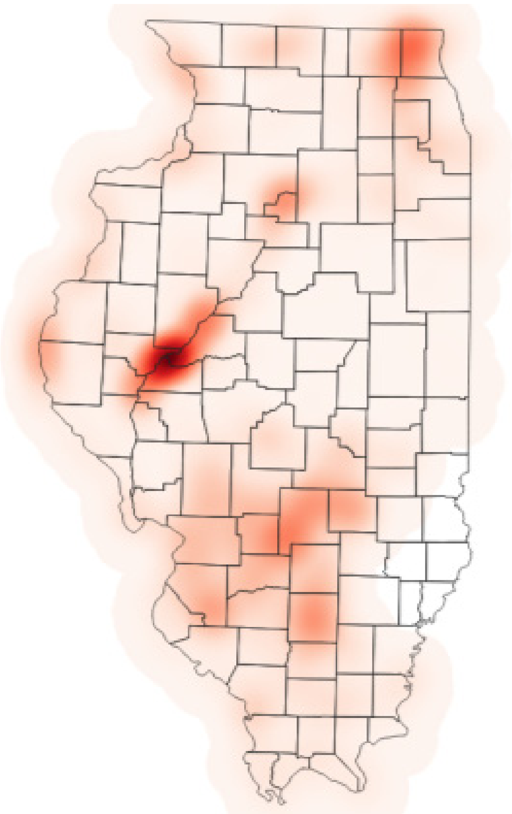
We leave the “Mode” as “Continuous” (yellow arrow) for our example, but you could create classes/bins if desired. We finish by clicking “OK” (blue arrow).



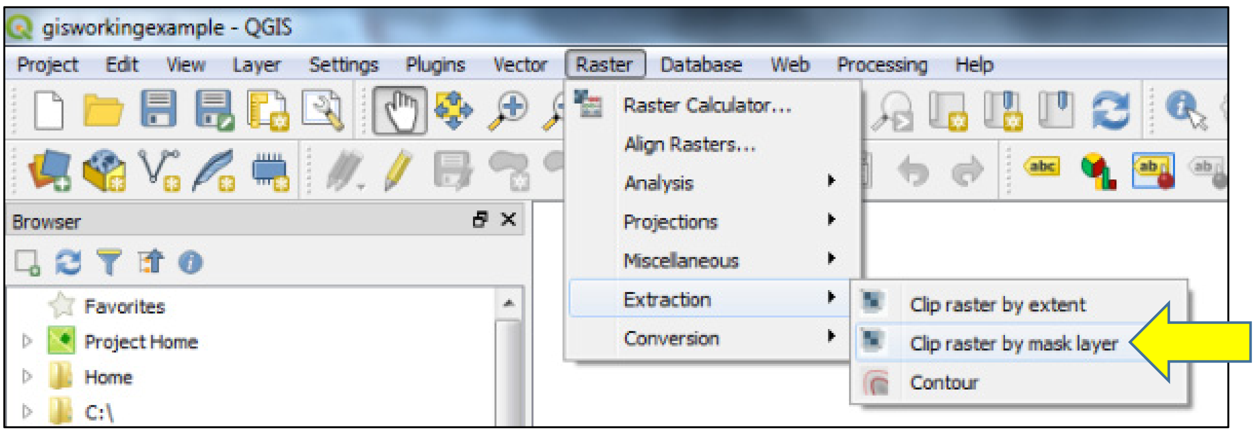
The result is now shown below. We can see that this map is easier to interpret compared to the original gray-scale raster image above.



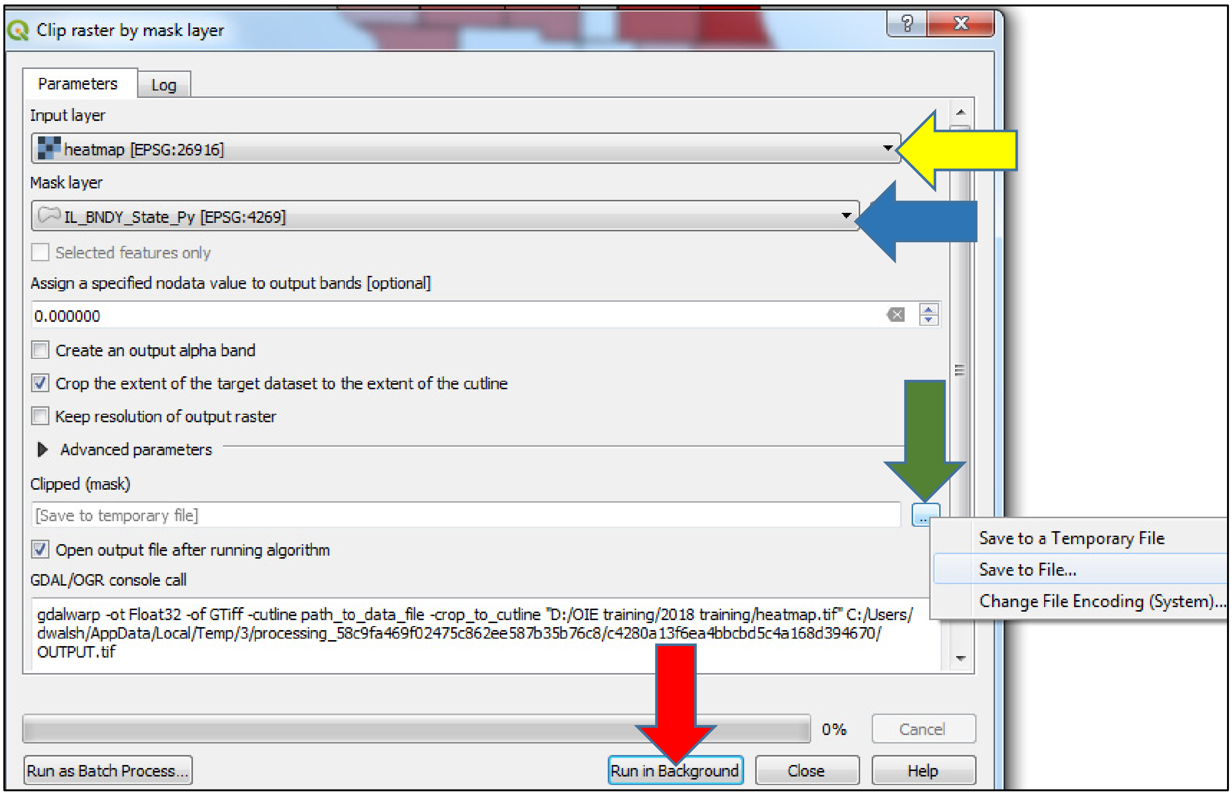
If we overlay the county polygon layer on this heat map, by dragging the county polygon layer to the top of the layer stack in the "Layers" pane, we see that the heat map extends beyond the state of Illinois. This is due to the nature of the smoothing algorithm. This is undesirable for our purposes.



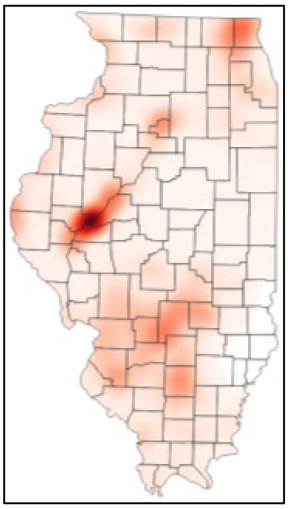
To correct this problem, we will clip the heat map raster using the county polygons layer. We open the "Raster" menu at the top of the screen, select the "Extraction" toolbox, and then choose "Clip raster by mask layer" (yellow arrow).



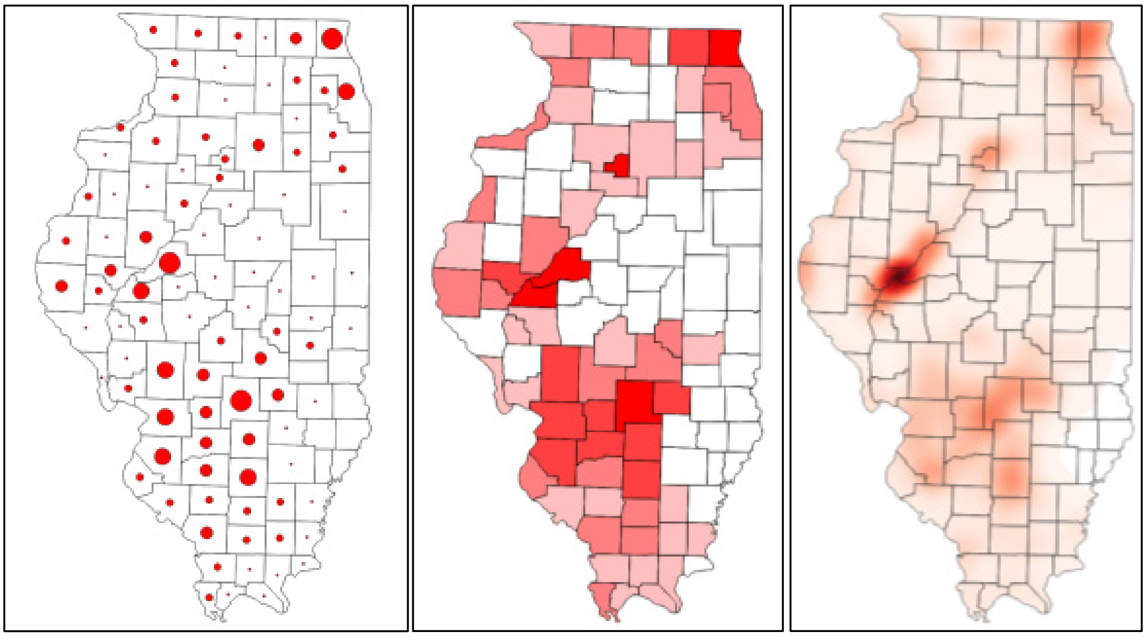
This will open a new dialog window in which we will choose the input raster layer that we intend to clip (yellow arrow). We also choose the vector layer that we will use for clipping the raster layer, which is the county polygon layer in our example (blue arrow). Lastly, we will choose "Save to File" by left clicking on the "…" button (green arrow). This will open a browser window that we can use to navigate to the directory on our computer where we would like to save our file as a GeoTiff, and allows us to specify a file name (e.g., "heatmap"). We then run the tool by clicking the "Run in Background" button (red arrow).



The result is the final heat map that we can use for disseminating our wildlife health information. This image preserves some of the more fine-scale spatial information, and clearly shows that the rates of cases within counties are heterogeneous. This is the type of the information that is lost with choropleth or graduated symbols maps.



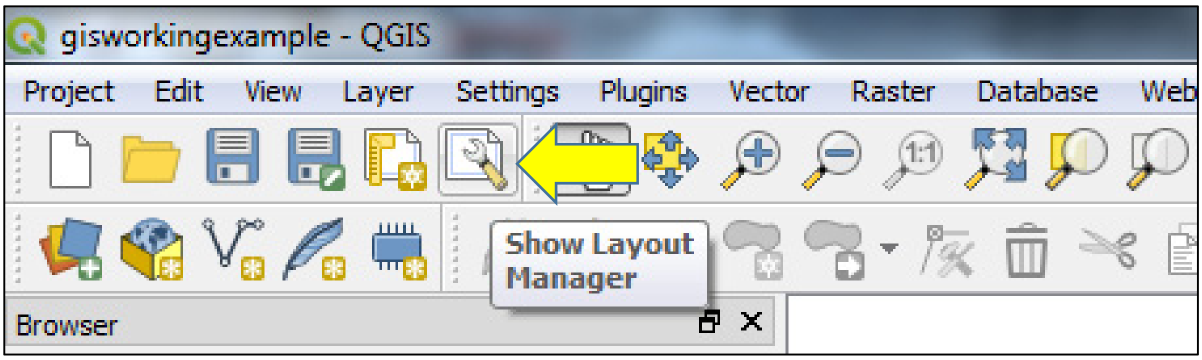
We conclude by placing all three maps side by side so it is apparent how they display the same information in different ways, and can be used to convey different messages.



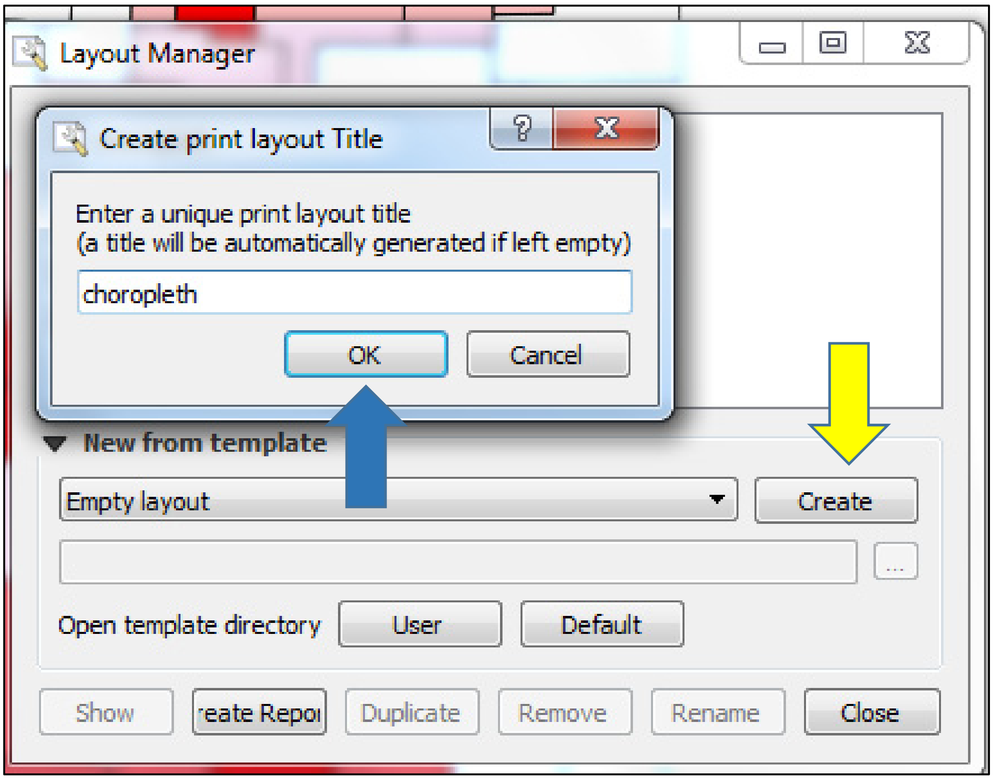
We will conclude our discussion of map creation in QGIS by describing how to make "polished" maps that can be shared.

Creating Map Layouts

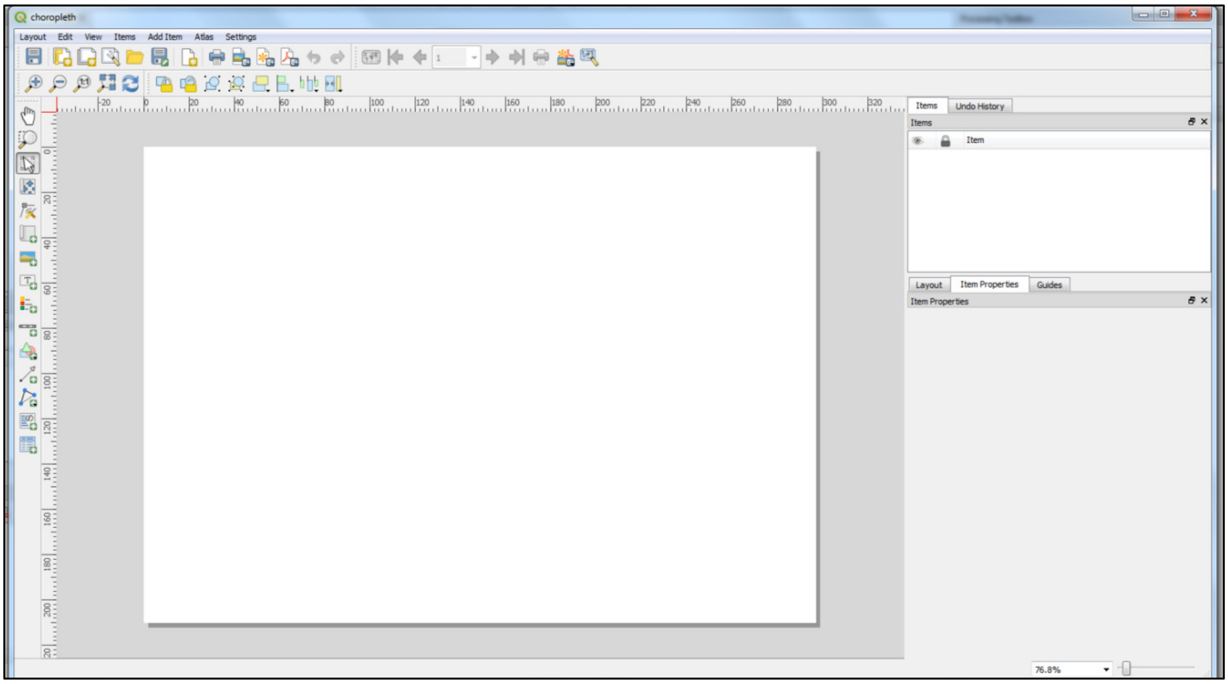
We have demonstrated some of the capabilities of QGIS and detailed how to combine spatial data layers to create a GIS for visualizing wildlife health information. We will now describe how to create a final map that can be disseminated to interested parties. This will involve creating a map layout. For this we will use the choropleth map we previously created. If that layer is not the top layer in the layer stack, in the "Layers" pane, turn off all the layers but "Count" by unchecking the boxes next to all of the layers' names. We then begin by left clicking on the layout manager tool (yellow arrow).



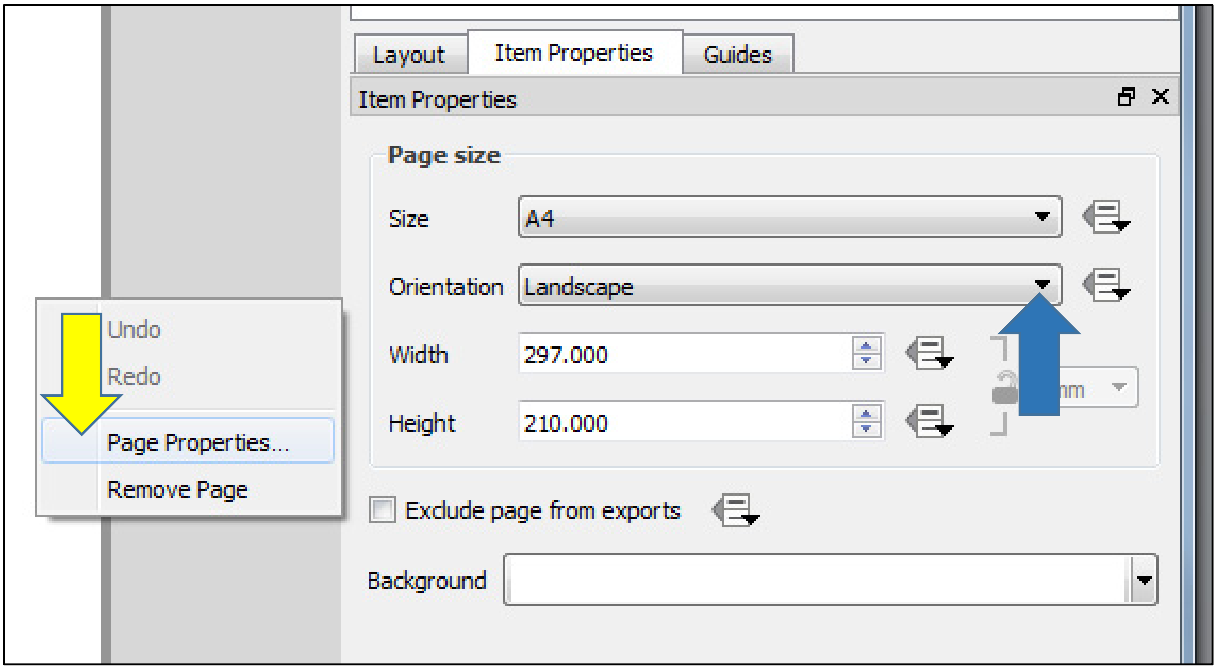
This opens a prompt where we can specify the name of our new layout, or open an existing layout. We will create a new layout by clicking "Create" (yellow arrow), and will specify the name of our layout as "choropleth" because we will use our choropleth map in this example. We left click "OK" (blue arrow).



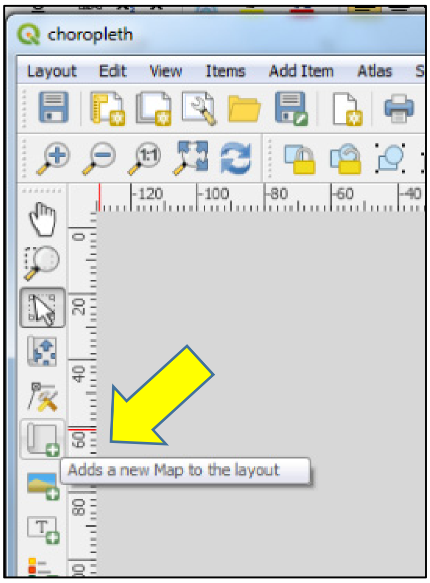
We now will see the layout manager window.



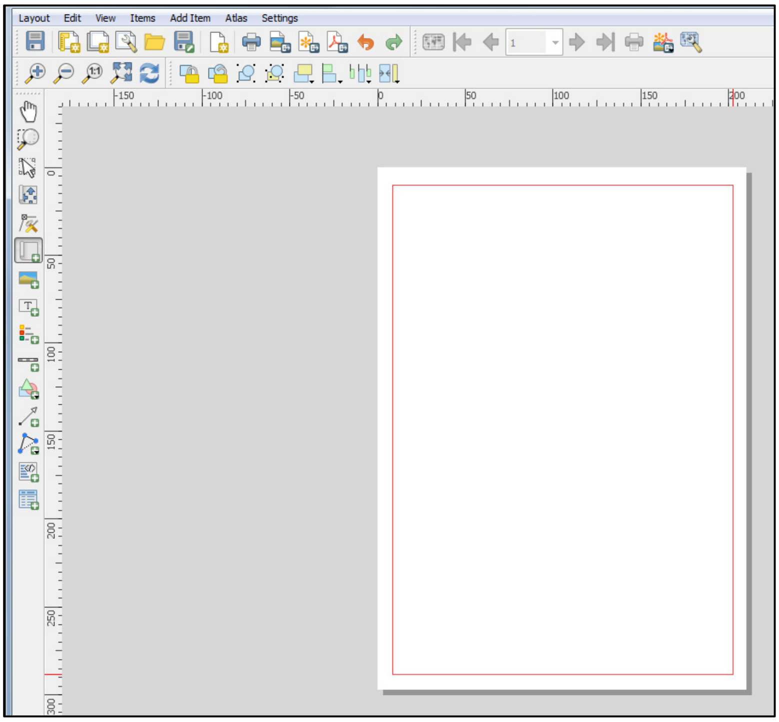
First, we need to change our layout template from "landscape" to "portrait" because our map has a larger height than width. We do this by right clicking on the center of the empty map template, and choosing "Page properties" (yellow arrow). We can then go over to the "Item Properties" tab, and change the orientation from "Landscape" to "Portrait" by clicking on the black triangle (blue arrow) in the "Orientation" row and selecting the "Portrait" option.



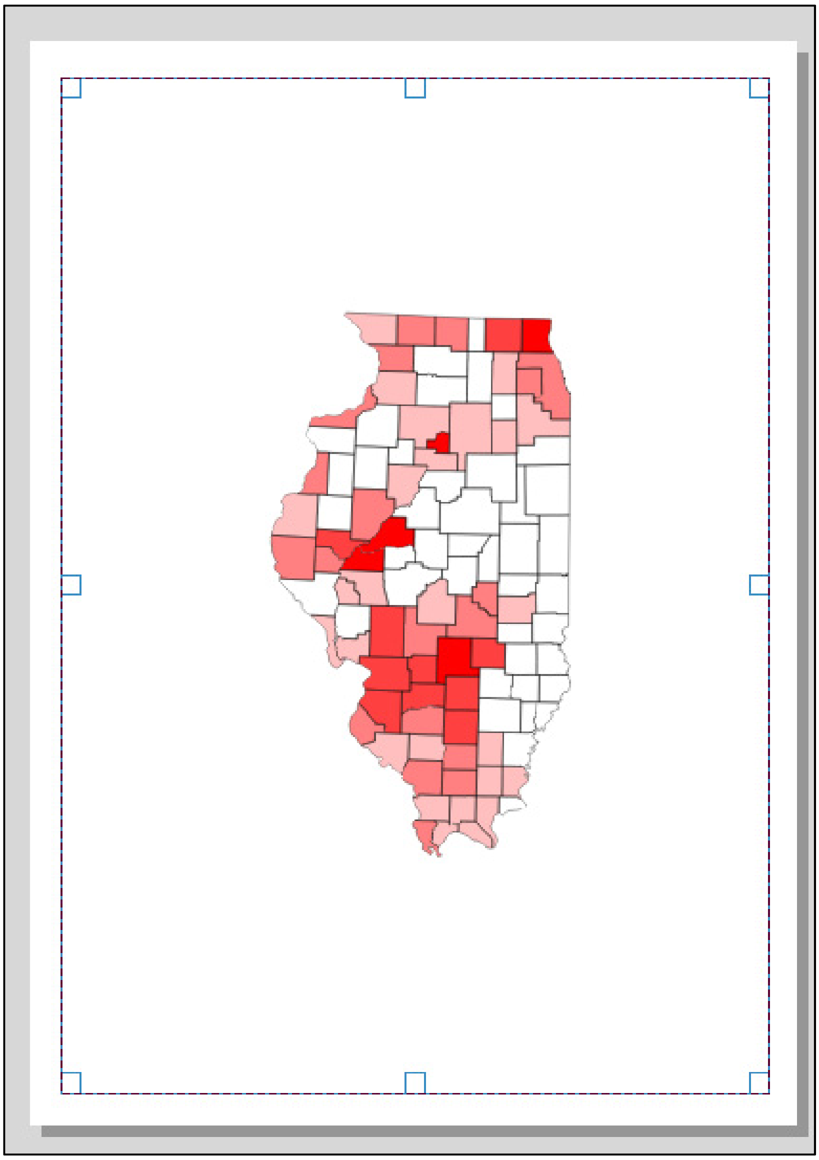
The first element we are going to add to our layout is our choropleth map. We left click on the "Add new map" tool (yellow arrow).



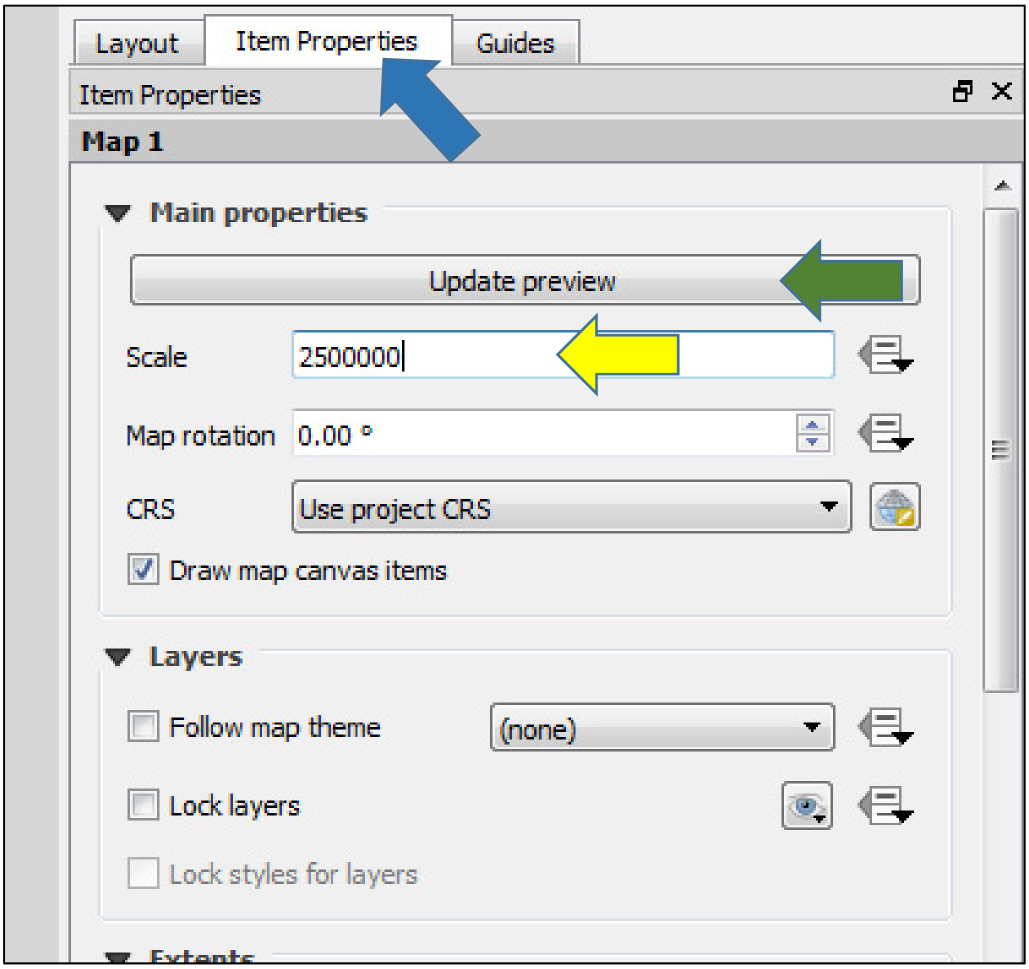
A “black cross” appears that we use to create a red rectangle on the template that will designate the region on the template that our map will occupy. First, we move our mouse to the location we would like to start our rectangle. We then hold down our left mouse button and drag the mouse to create the rectangle of the desired size.



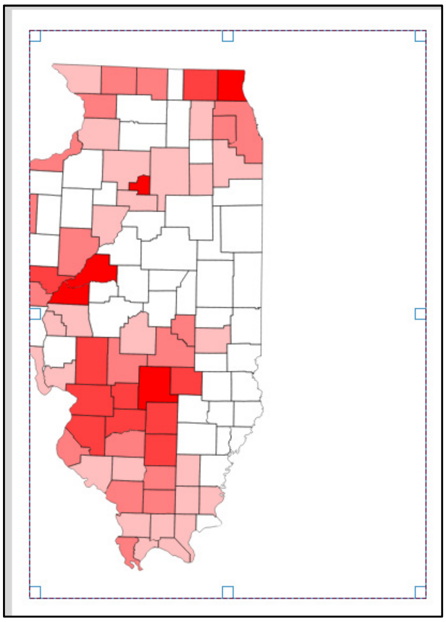
We now see our choropleth map displayed on the template. However, it is too small relative to size of our layout.



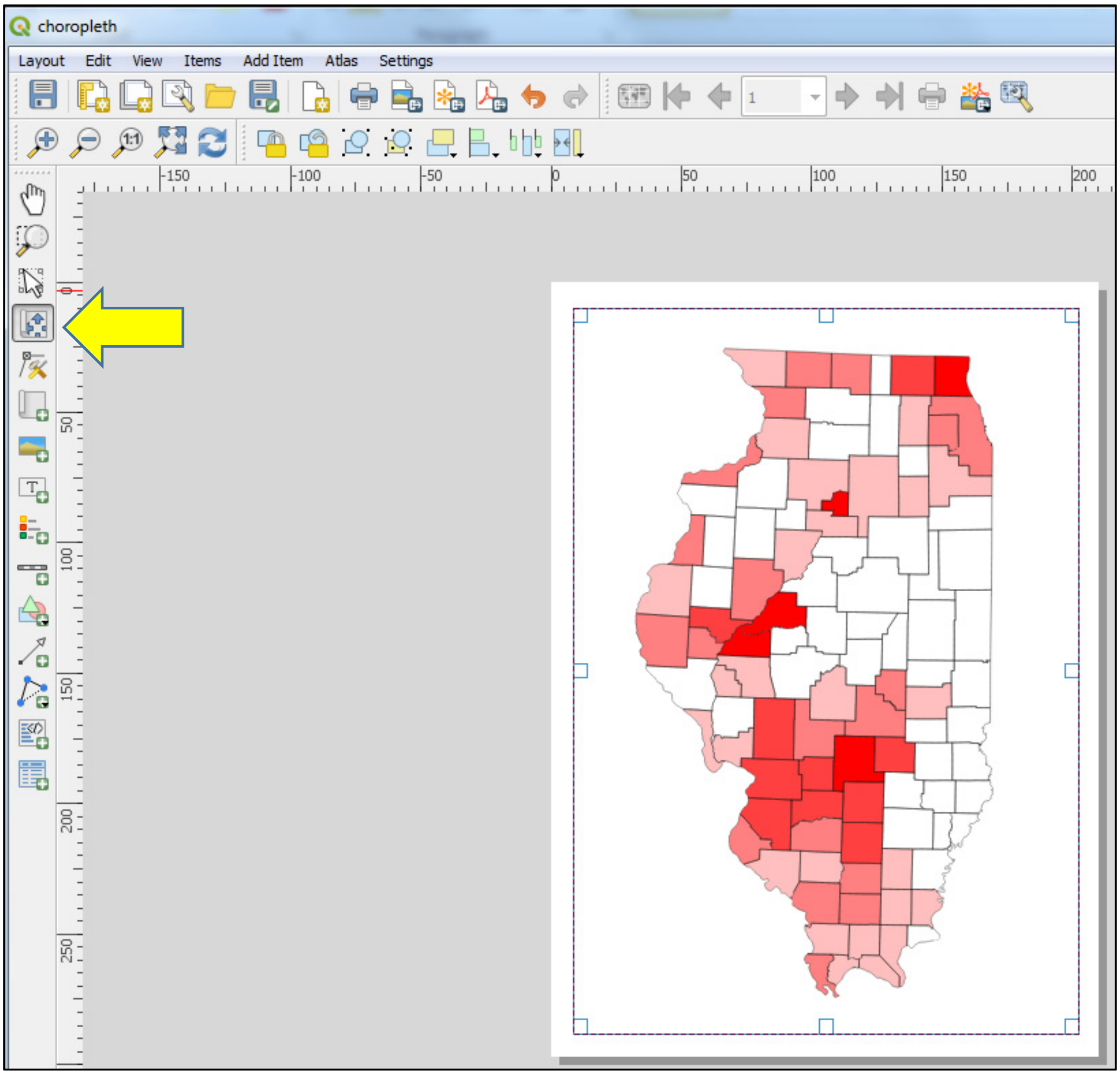
So, we will change the scale, by specifying 2500000 for the "Scale" (yellow arrow) under the "Item Properties" tab (blue arrow). You may have to play with different scale values to determine the one most appropriate for your map. For reference, reducing the number entered in the scale value results in a "zoomed in map" whereas increasing the number "zooms the map out". Once we have specified our new scale we left click "Update preview" (green arrow) to see what effect changing the scale has had on the appearance of our map. We repeat until we have the size we desire.



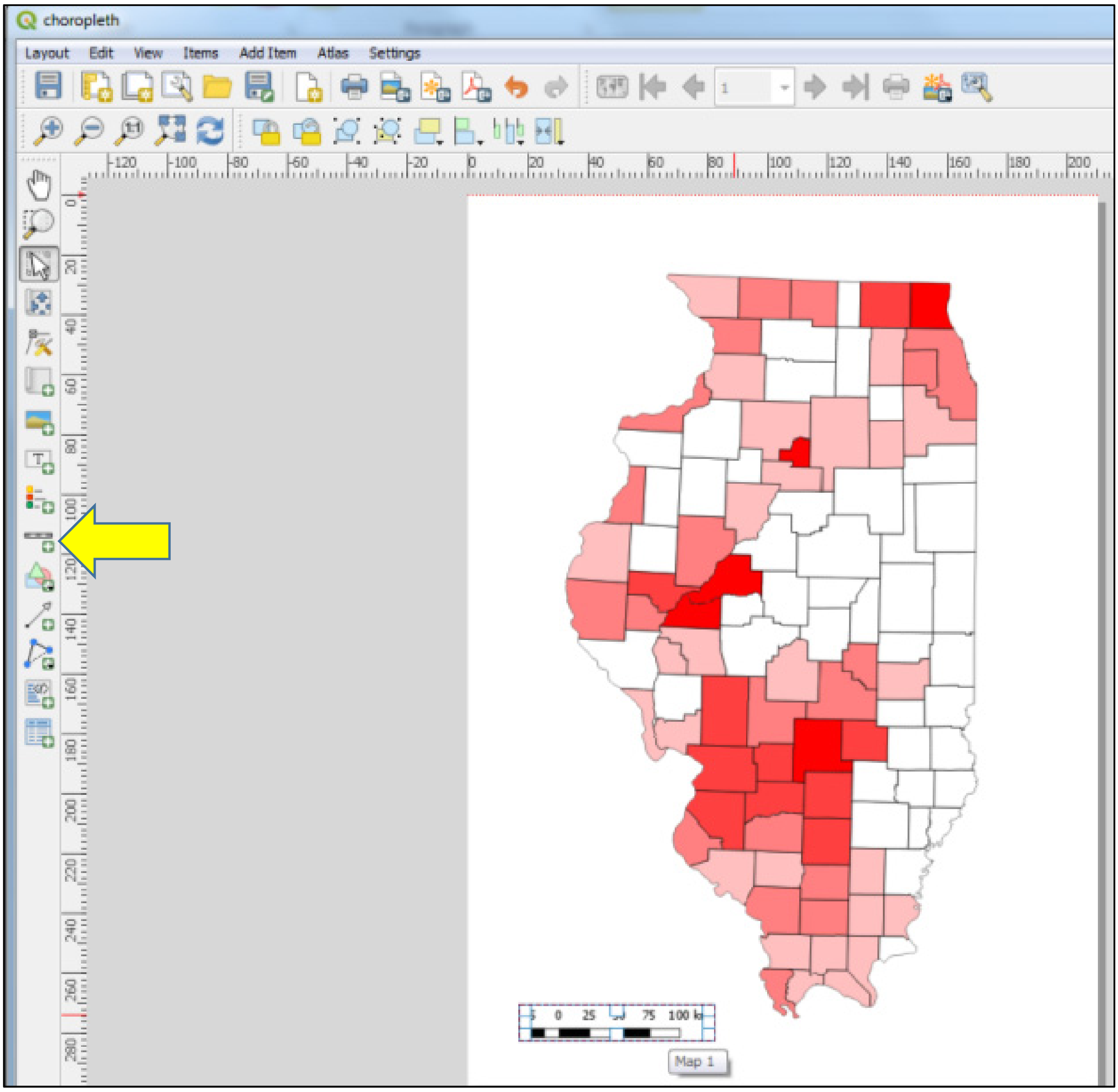
We can see that now the choropleth map fills the template.



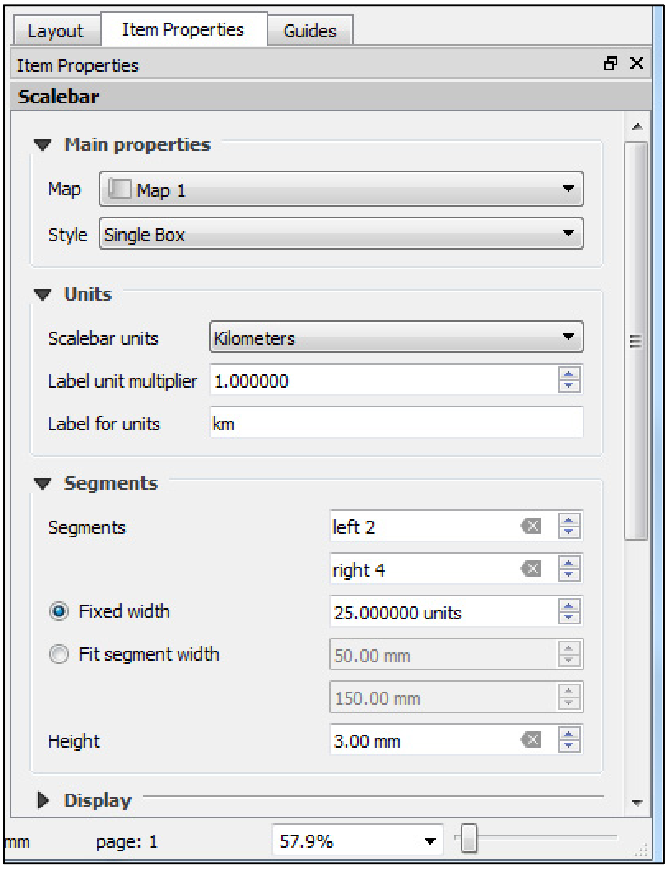
However, it is clearly off-center in our template. So, we click on the "Move Item" tool (yellow arrow). We then move our mouse over our map, and while holding the left mouse button down, we drag the map to the desired location on our template. It is now centered.



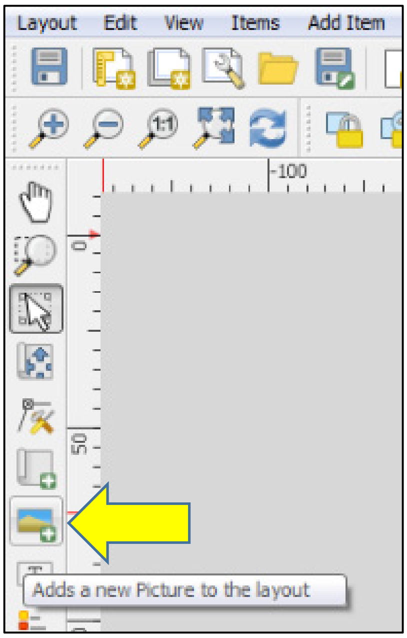
The next element we will add is a scale bar. Scale bars are important map elements because they allow the map user to determine distances on the map. To add the scale bar, we left click on the "Add scale bar tool" (yellow arrow), and then move our cursor to where we would our scale bar to be placed. We then hold the left mouse button down, and drag the cursor to create the scale bar of the desired height and length.



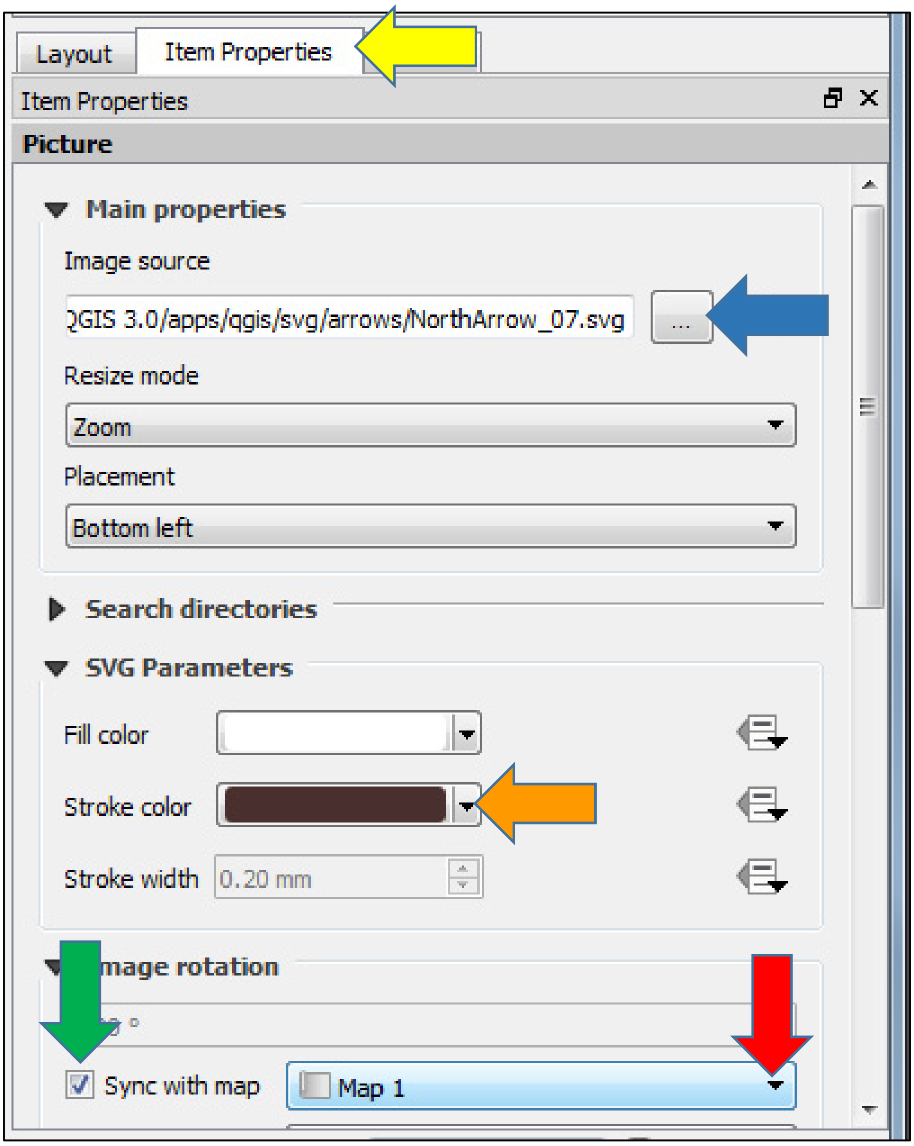
We can change the properties of the scale bar by left clicking on the newly created scale bar so it is selected, and then going to the "Item Properties" tab. There are many options to choose from. For example, you can change the units of the scale bar (e.g., meters to kilometers, etc.), the font, color, etc. We will just use the default settings for this example.



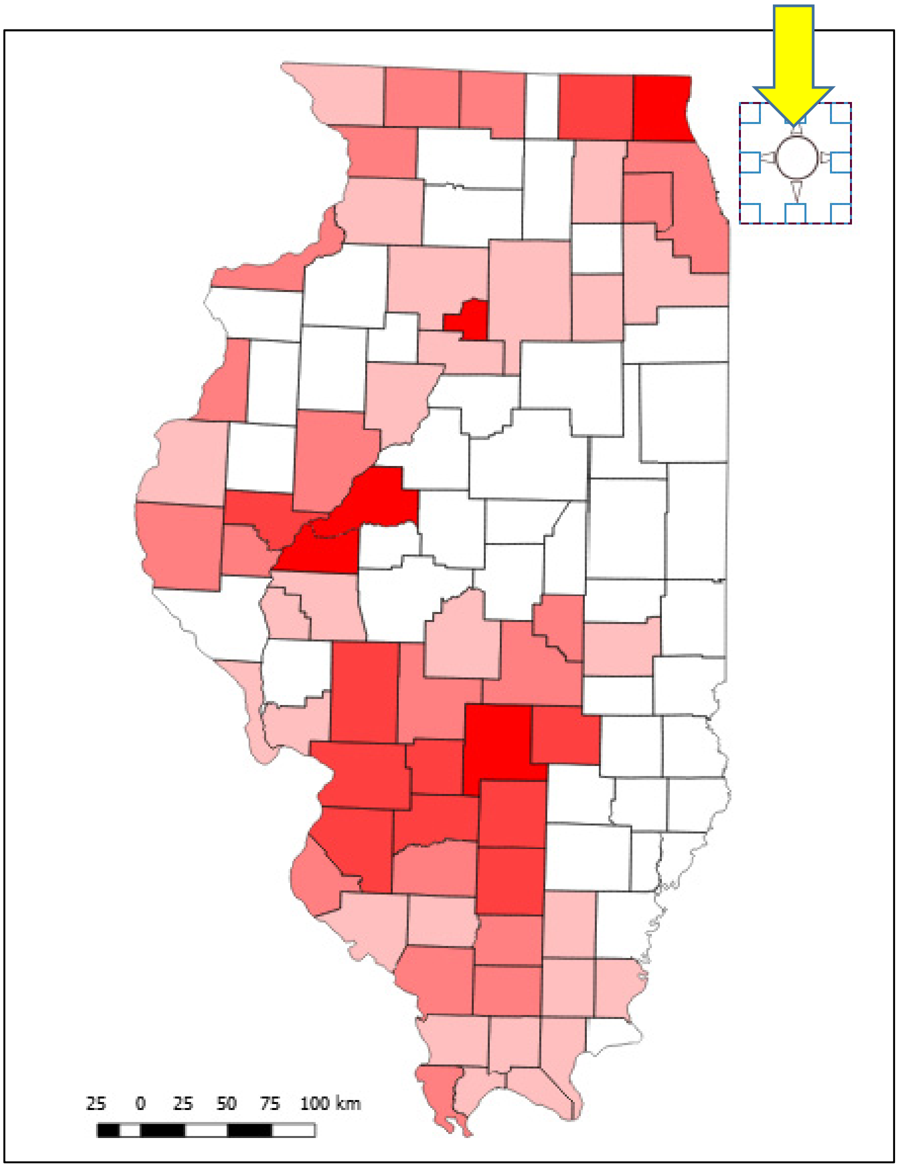
The next element we will add is a North arrow. This element helps the map user know the orientation direction of the map, which is often critical for actually using the map. The North arrow is added by left clicking the "Add picture tool" (yellow arrow).



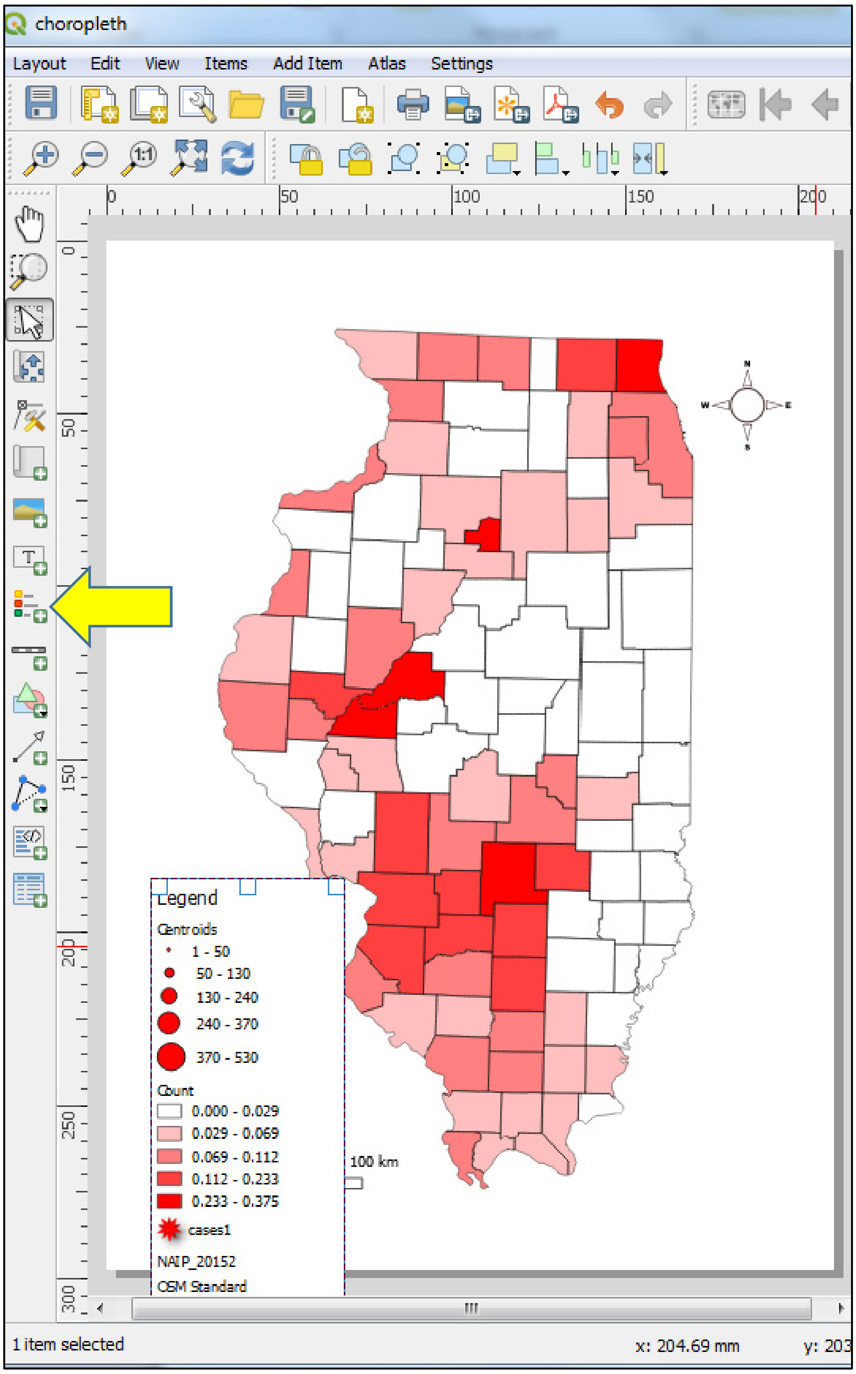
This will turn the mouse cursor into a black cross. Move the cursor to the location where you would like to place the North arrow, and while holding down the left mouse button "drag" a rectangle to the size of North arrow you would like to create, and then release the mouse button. This will create a frame into which the North arrow will be placed. Next go to the "Item properties" tab (yellow arrow), and click on the "…" button (blue arrow) to open a browser window to navigate to the folder that contains the image you would like to use. QGIS comes with a number of North arrow images that can be selected. They are found in the following directory: "C:\Program Files\QGIS 3.0\apps\qgis\svg\arrows". We will choose "NorthArrow_07.svg" for our map. We want to sync the North arrow with our map, so we will check the "Sync with map" (green arrow) option, and choose "Map1" as the map with which to sync by clicking on the black triangle (red arrow) and selecting "Map1". Finally, we want to de-emphasize the North arrow so that is not so prominent. Thus, we change the "Stroke color" (orange arrow) from black to dark gray.



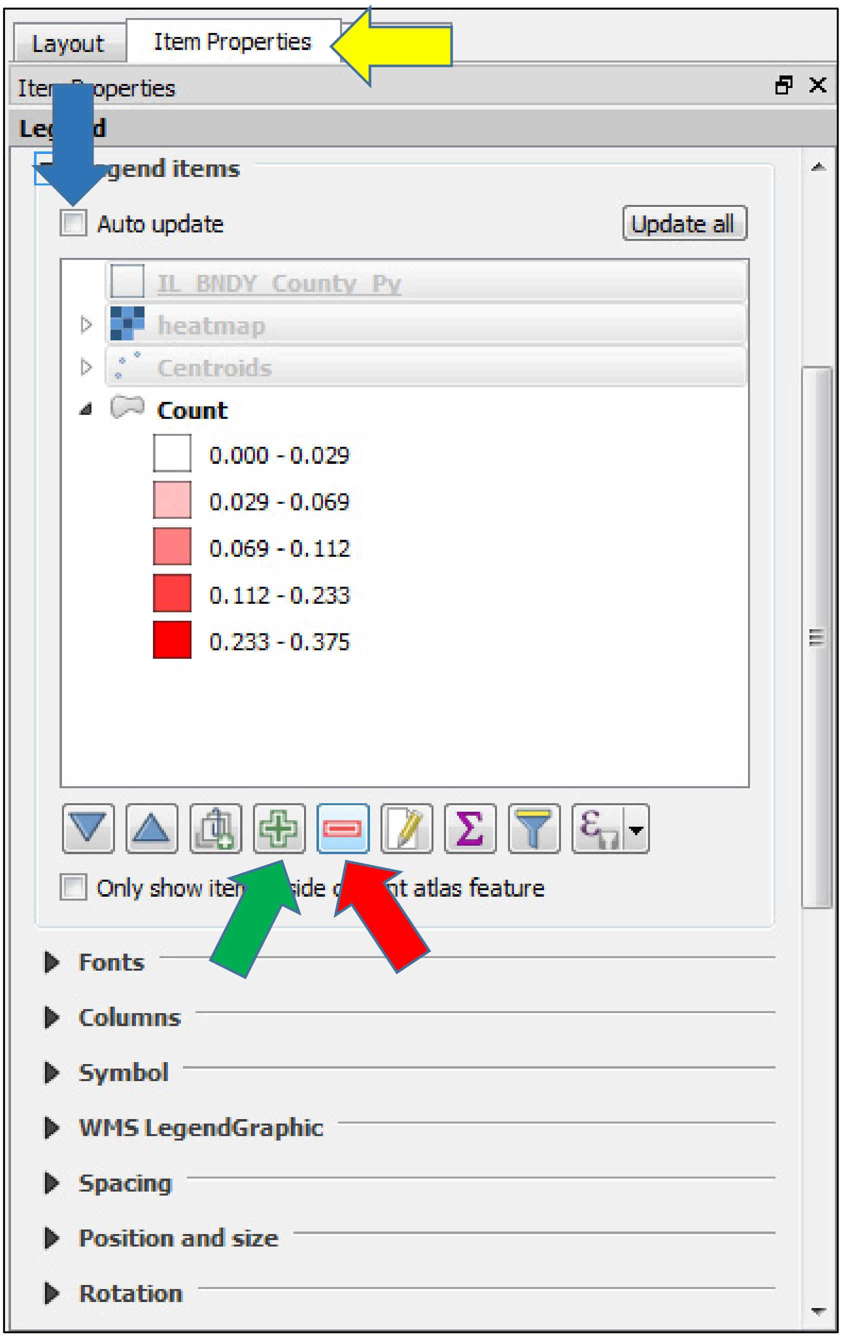
The map with the North arrow is now displayed. If we would like to resize the North arrow, we left click the North arrow, and then we hover the mouse cursor over any of the "boxes" surrounding the North arrow (yellow arrow) until the cursor turns into a bi-directional arrow. We then hold the left mouse button down and drag until the arrow is the desired size. If the North arrow needs to be repositioned, we simply move the mouse cursor to the center of the rectangle containing the North arrow. The cursor will become a cross composed of two arrows, we then hold the left mouse button down and drag the North arrow to the desired location. This procedure to resize and reposition works with all map elements.



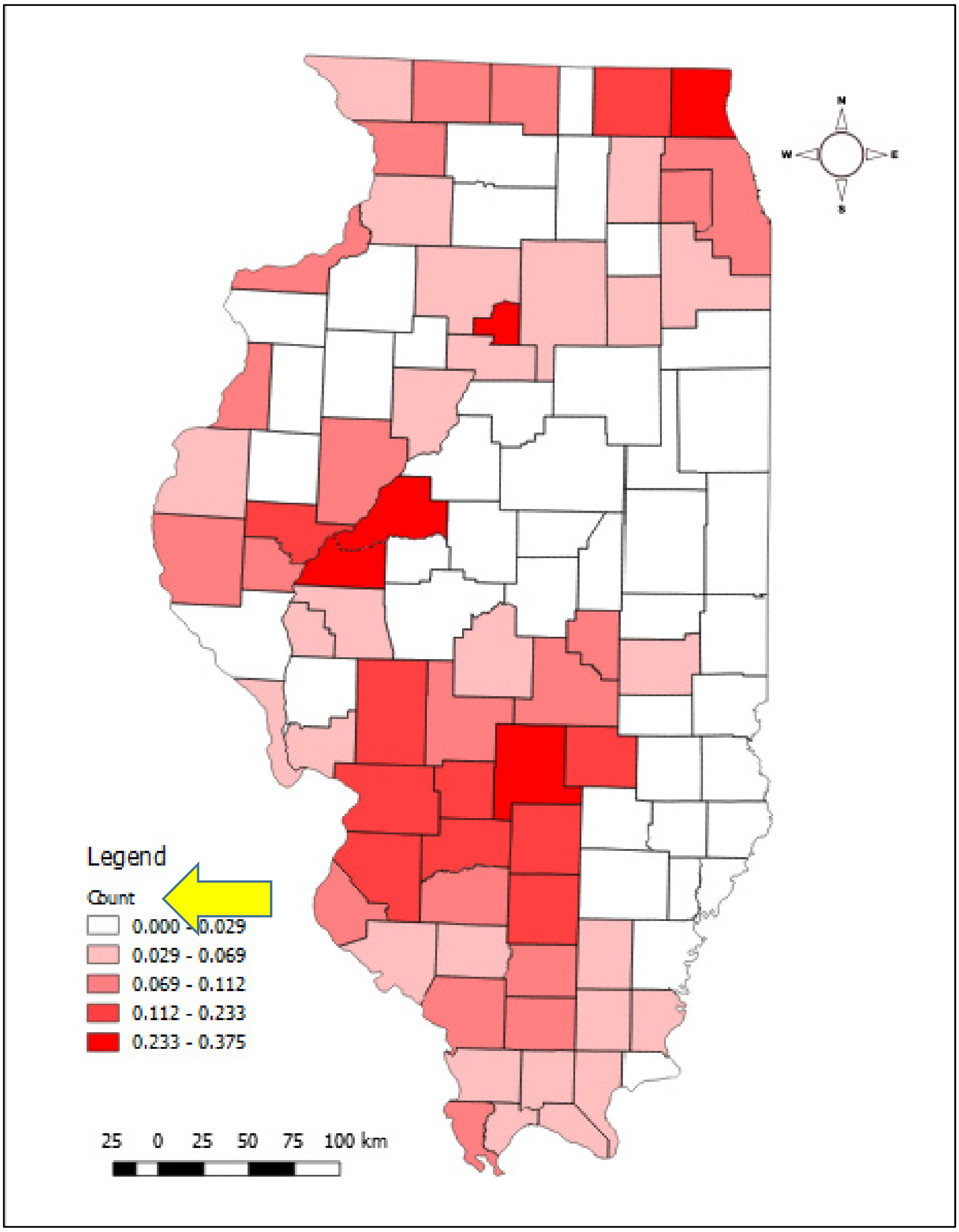
For statistical maps such as choropleth maps, it is important to provide a legend that will allow the map user to understand the meaning of the various color classes that are displayed. Therefore, the next map element we will add is a legend. We click on the add legend tool (yellow arrow). This will turn the mouse cursor into a black cross. Move the cursor to the location where you would like to place the legend, and while holding down the left mouse button "drag" a rectangle to the size of legend you would like to create and then release the mouse button. This will create a frame containing the legend. Note, the newly displayed legend contains an item for each layer in the "Layers" pane of our GIS project. However, we would like the legend to only contain information for the choropleth map.



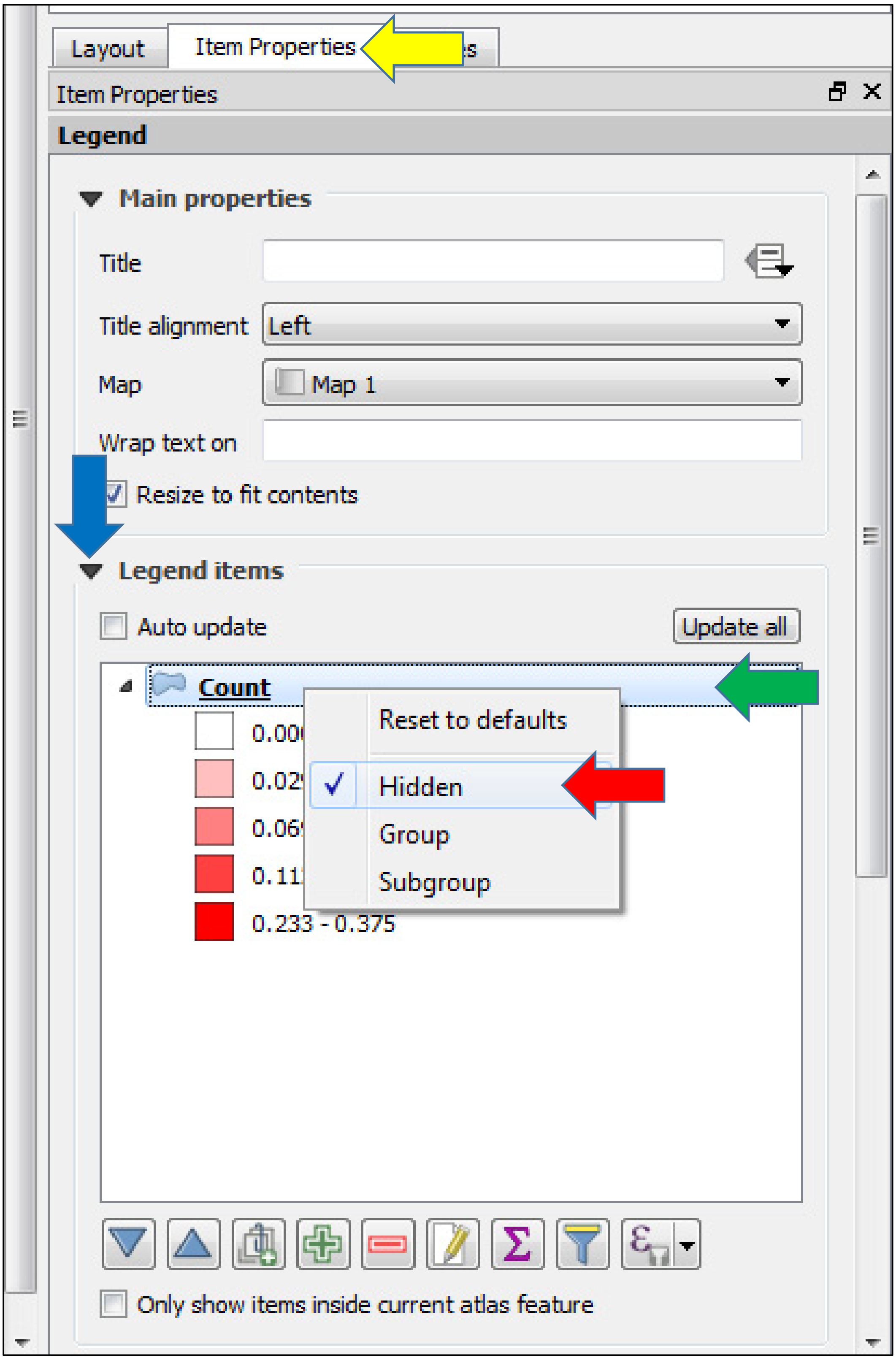
To remove the unwanted legend items, we go to the "Item properties" tab (yellow arrow). We uncheck the "Auto update" box (blue arrow). We then left click on each layer we want to remove under the "Legend items" group, and click on the red "–" button (red arrow). We then repeat until only the choropleth layer (i.e., Count) remains. If we would like to add back in a layer, we click on the green "+" button (green arrow), and select the layer to add.



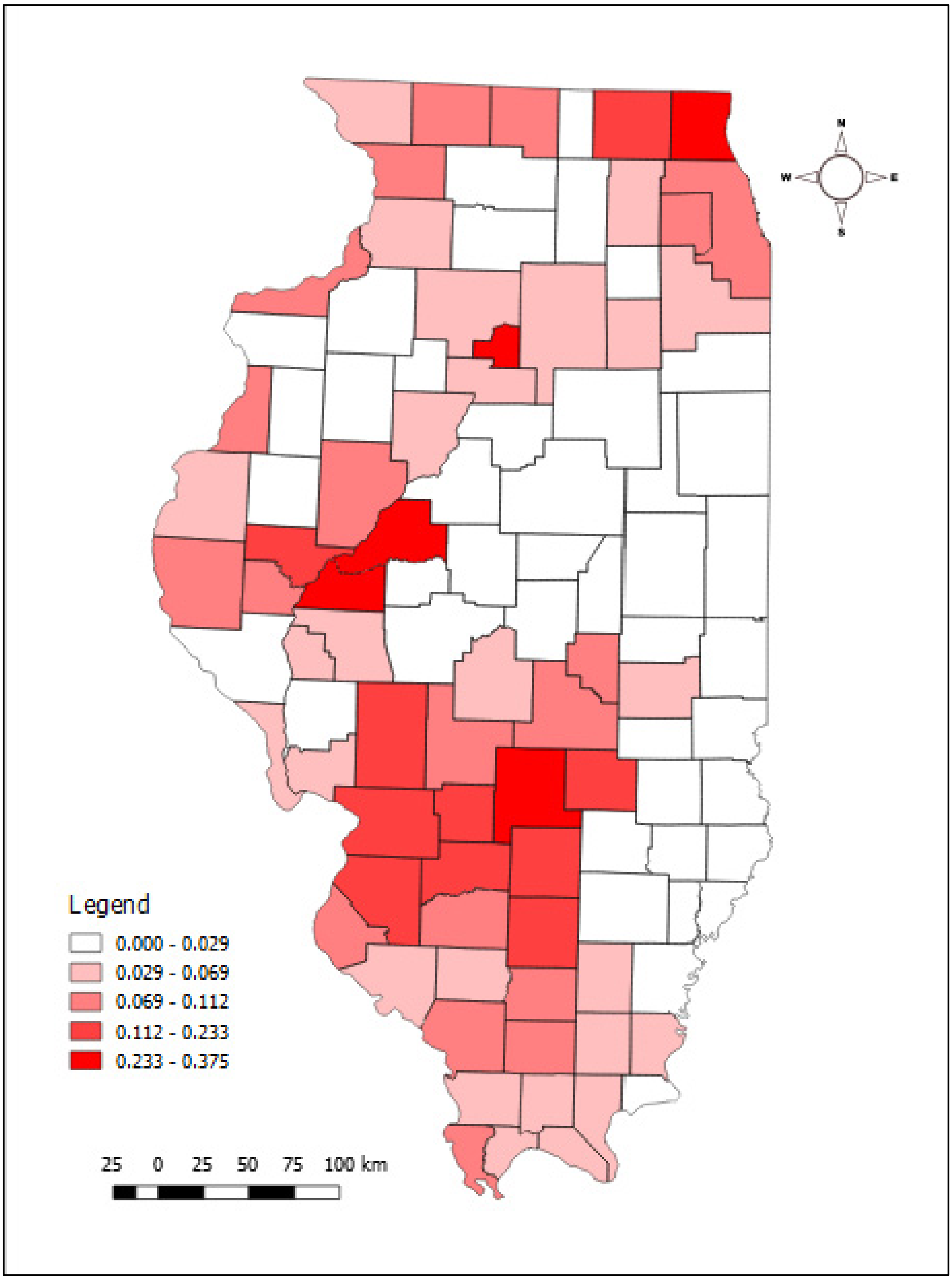
The legend now contains only the relevant information that we want the map user to be able to view. We can now reposition the legend as needed using the techniques described above for repositioning the North arrow.



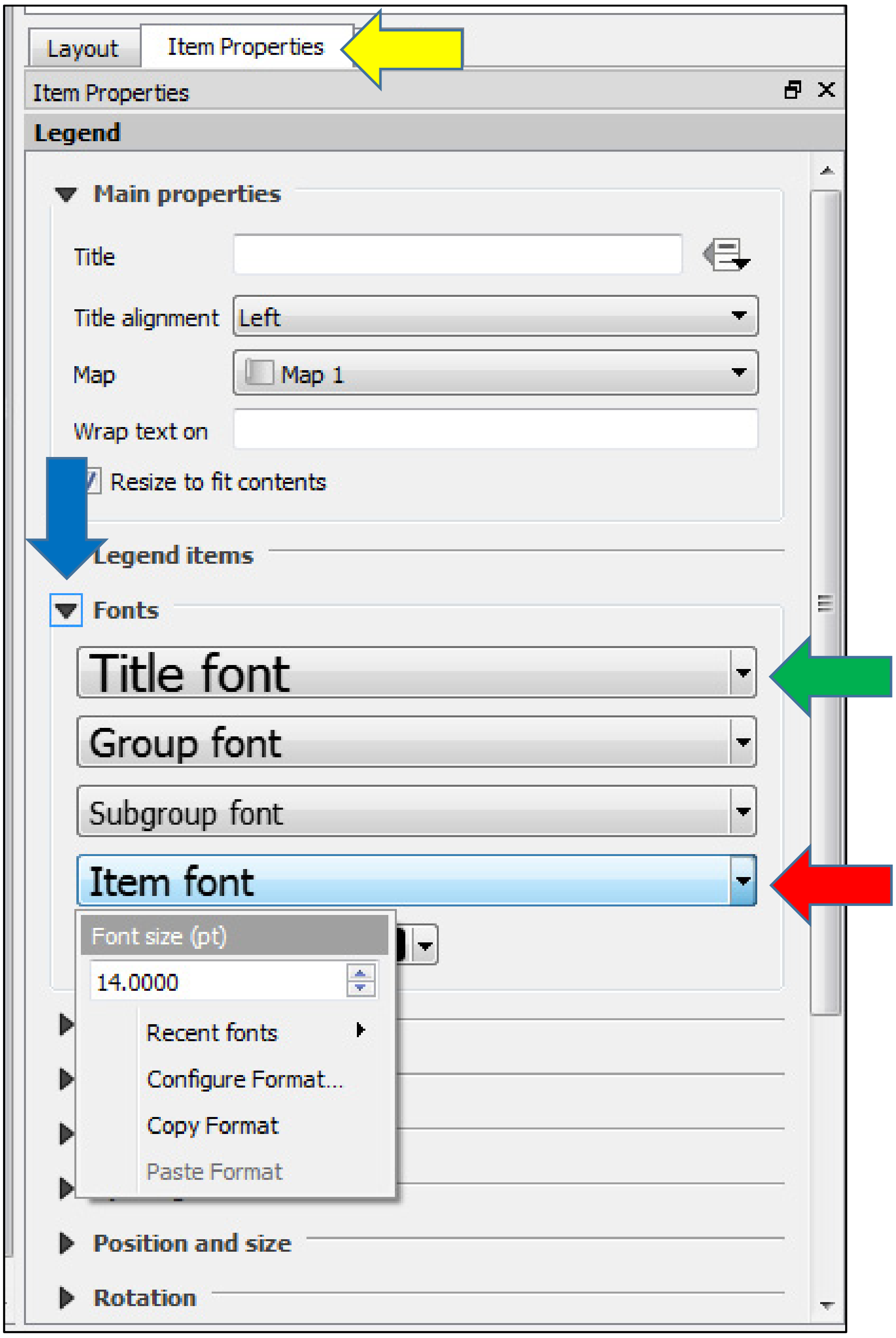
There are many different options for constructing or altering the appearance of the legend which we will not describe, but there are numerous on-line resources available to demonstrate these capabilities. However, we do want to remove the "Count" subgroup label (yellow arrow, above), which is displays the name of our layer, from the legend. To do this, we left click on the black triangle (blue arrow) next to the "Legend items" group under the "Item properties" tab (yellow arrow). Next we right click on "Count" (green arrow) and select "Hidden" (red arrow).


Layout
Item Properties
Item Properties
Legend
Main properties
Title
Title alignment
Left
Map
Map 1
Wrap text on
Resize to fit contents
Legend items
Auto update
Update all
Count
Reset to defaults
Hidden
Group
Subgroup
0.233 - 0.375
Only show items inside current atlas feature

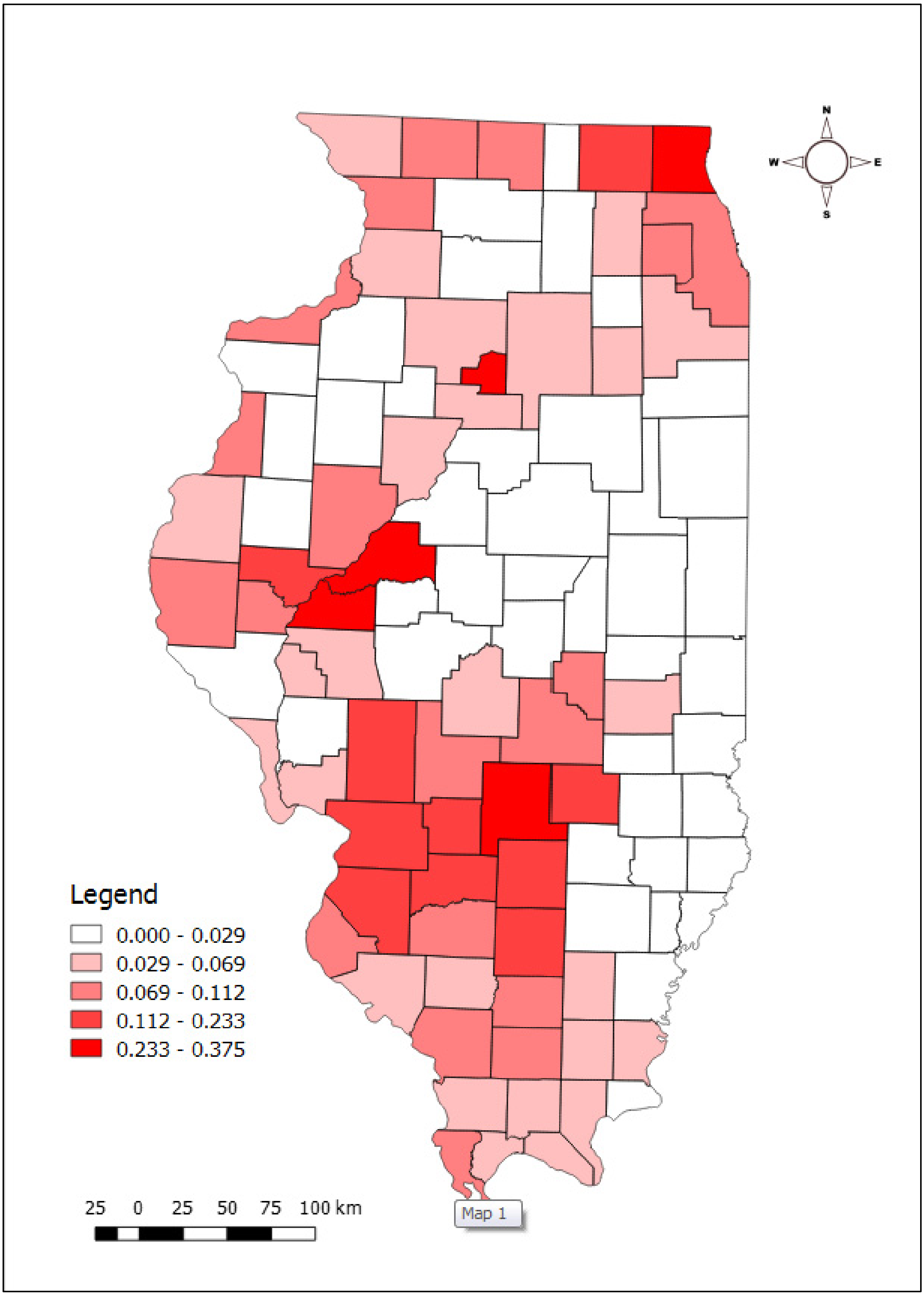
This will remove the "Count" title from the legend as we desired, and just display the colors and associated rates for each class.



Lastly, we would like to increase the size of the font of the title and items in the legend. Thus, we left click on the black triangle (blue arrow) next to "Fonts" group under the "Item Properties" tab (yellow arrow). Then we left click on the black triangle (green arrow) for "Title font" and select 18, and then we left click on the "Item font" (red arrow) and choose 14.

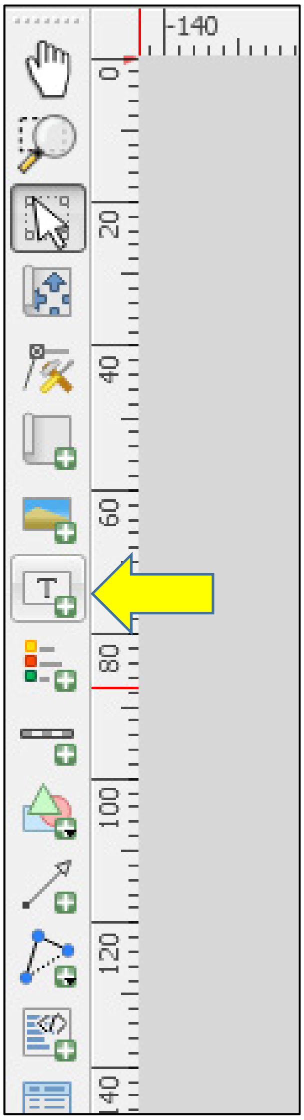


This finalizes the legend of our map. Next we will add a title element.

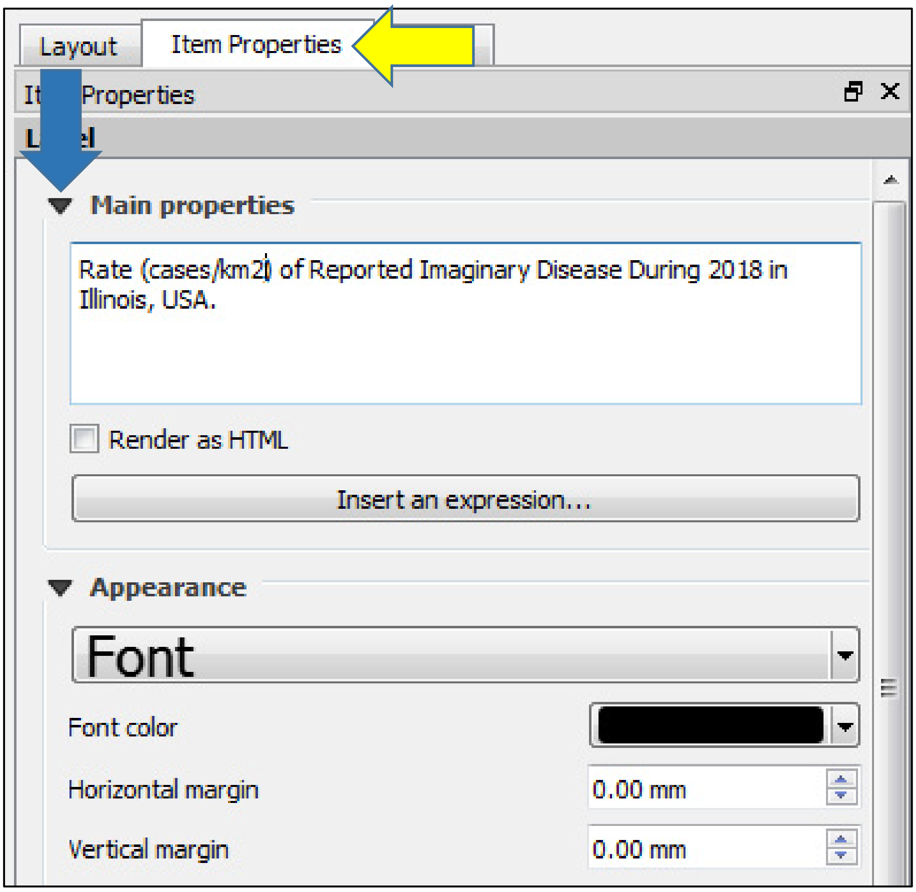


N
W
E
S
Legend
0.000 - 0.029
0.029 - 0.069
0.069 - 0.112
0.112 - 0.233
0.233 - 0.375
25 0 25 50 75 100 km
Map 1

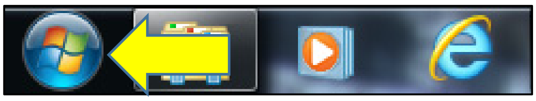
Title text is important to include on a map so that the map user can readily understand the information being presented. To annotate our map layout with text, we click on "Add new label" tool (yellow arrow).



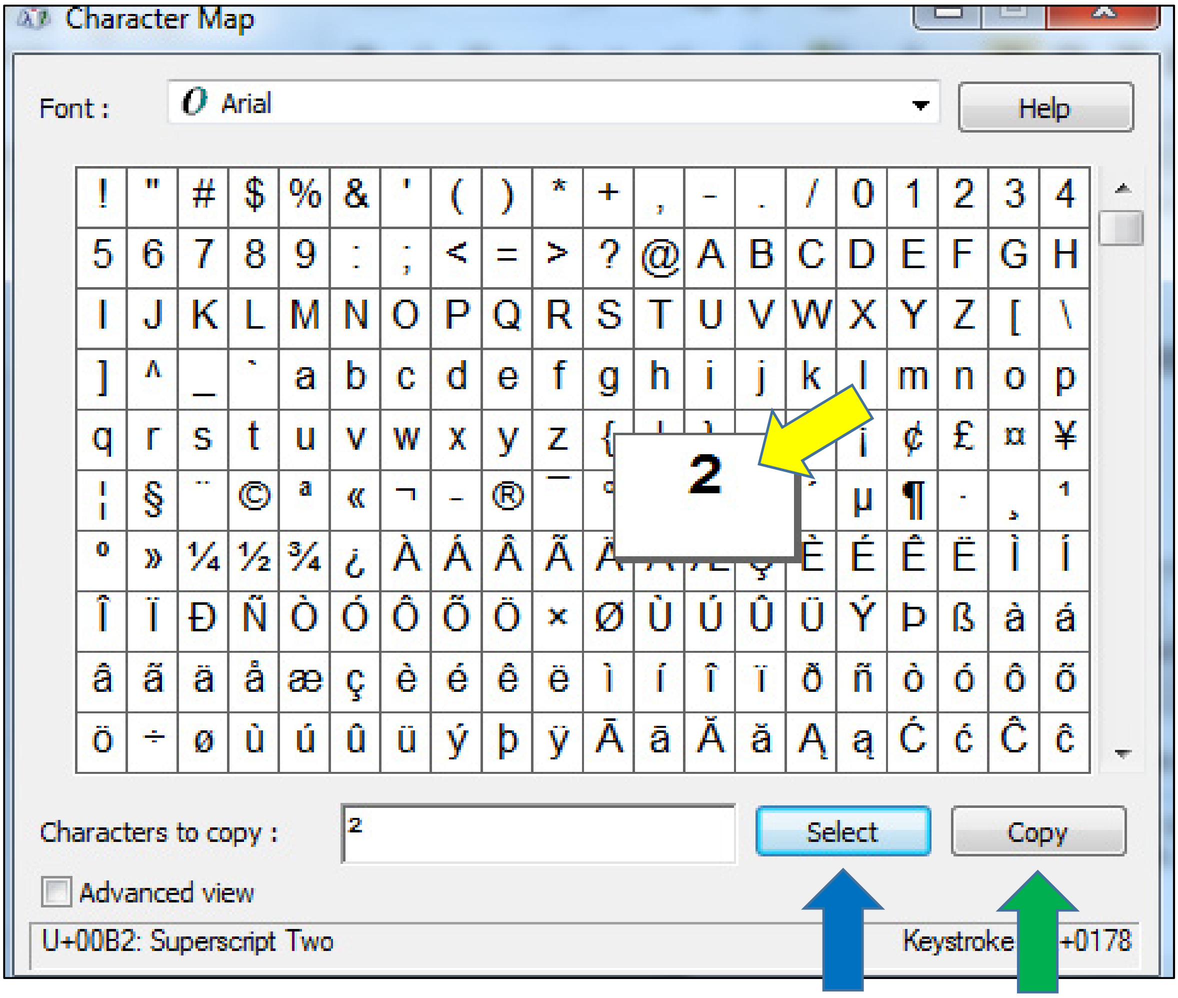
We would like to add the title, "Rate (cases/km^2) of Reported Imaginary Disease During 2018 in Illinois, USA". We add this text in the "Main properties" group (blue arrow) under the "Item Properties" tab (yellow arrow). However, to account for the superscripted "2" after km, we need to complete the following steps.



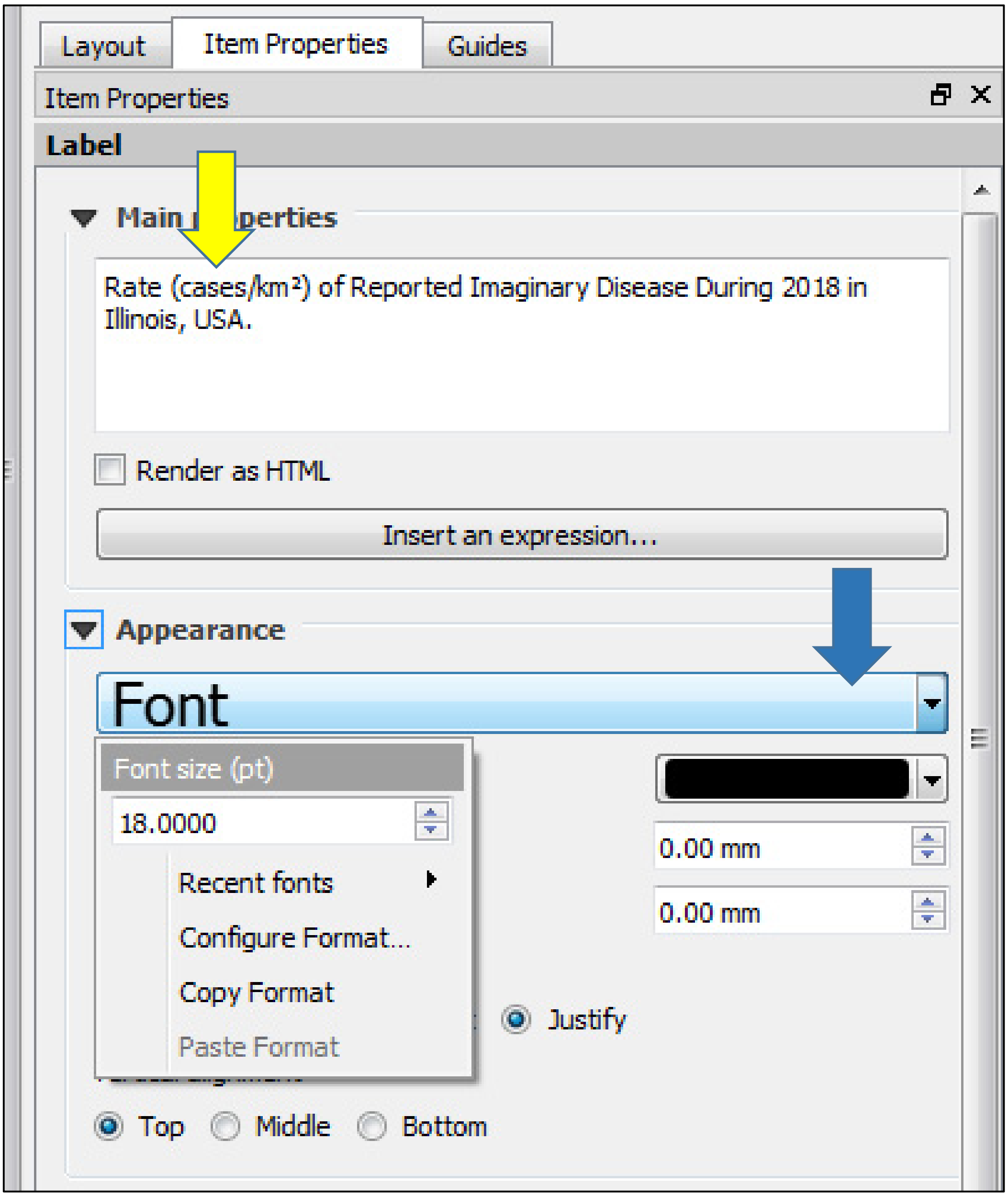
First, ***outside of QGIS***, we open Character Map by left clicking the Windows Start button at the bottom of our screen.



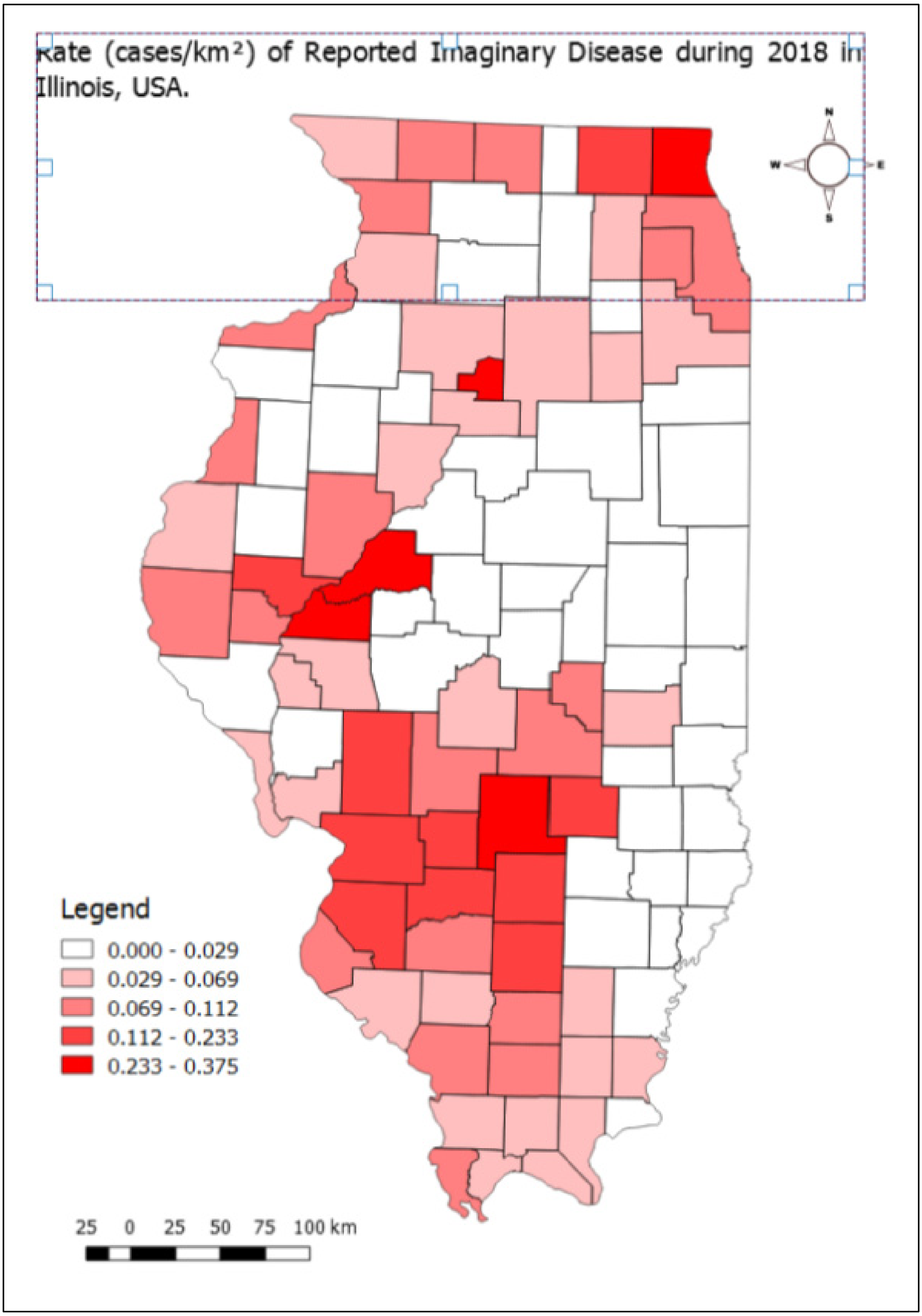
This opens the Windows menu, and we then left click All Programs, then choose Accessories, then choose System Tools, and then left click Character Map. In the Font list, type or select the font you want to use (e.g., Arial). Then scroll down and find the superscripted number 2 that we will use in our legend (yellow arrow). We then left click Select (blue arrow), and then left click Copy (green arrow).



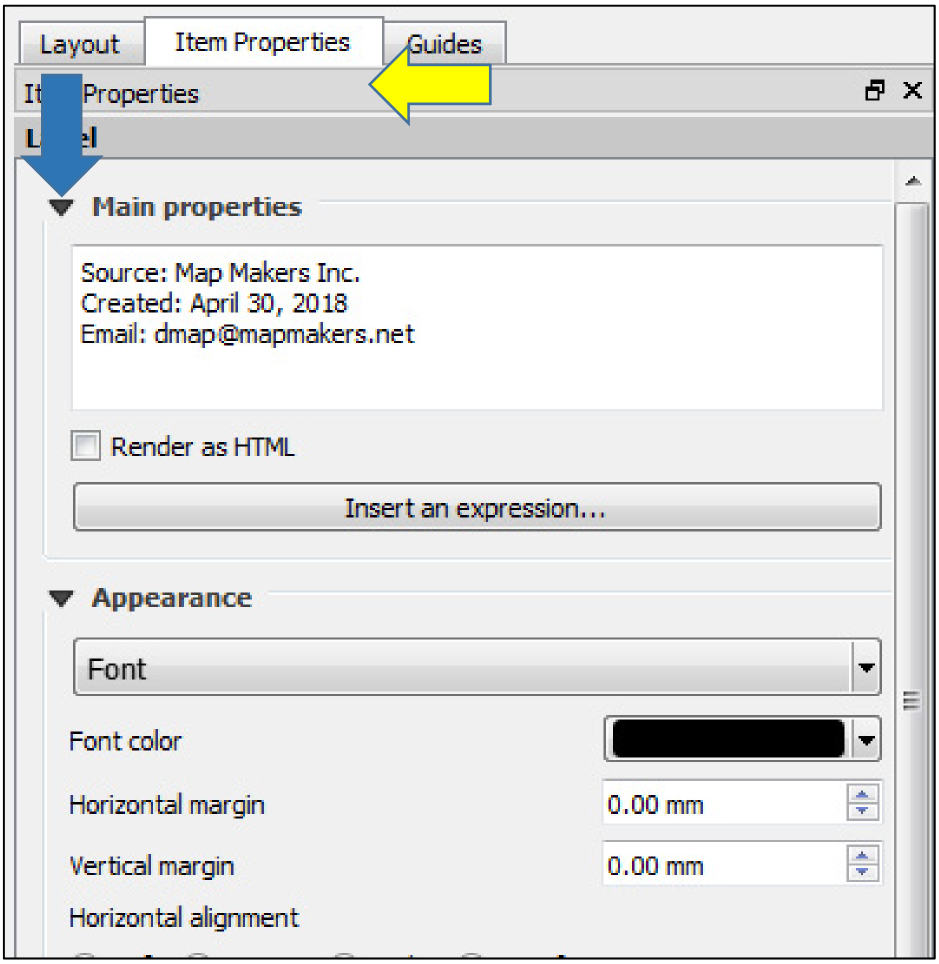
We then return to QGIS and paste the superscripted 2 behind "km" in the text contained in our "Main properties" group (yellow arrow) by left clicking behind "km" and right clicking and selecting "Paste". Finally, we would like to increase the font of our title text. Thus, we left click on the black triangle (blue arrow) for "Font" and specify the size as 18.



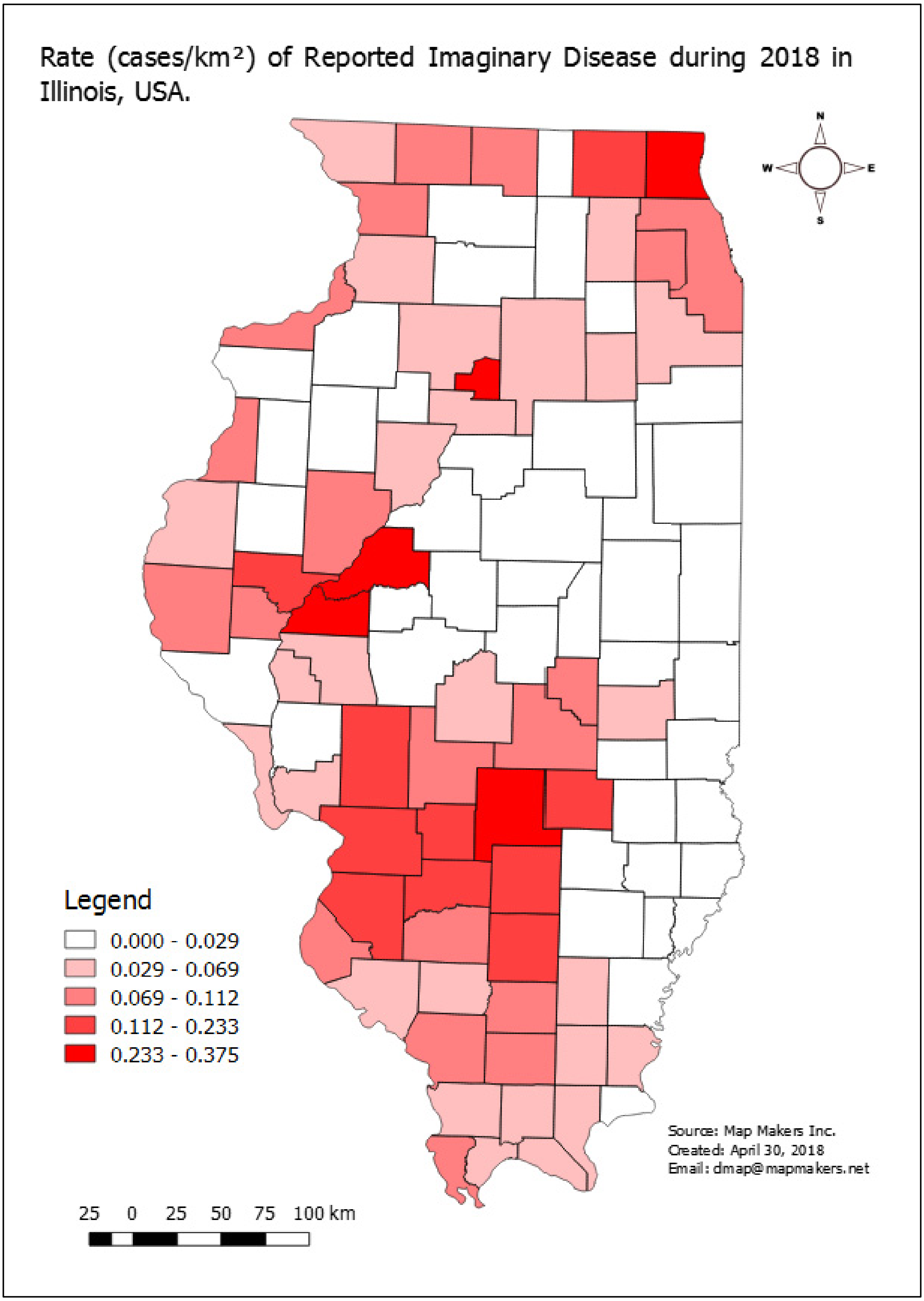
We then resize the text box surrounding our title as necessary to achieve the desired appearance of the title.


Rate (cases/km²) of Reported Imaginary Disease during 2018 in Illinois, USA.
N
W
E
S
Legend
0.000 - 0.029
0.029 - 0.069
0.069 - 0.112
0.112 - 0.233
0.233 - 0.375
25 0 25 50 75 100 km

The final map element we will add is a "source information box", which will include the date the map was published, the source of the map, and contact information for the map author. This element is less critical than previous elements, but it is good practice to include this information so map users can have this basic meta-data easily accessible. Additionally, the date is useful information because wildlife health information is both spatially and temporally dynamic, and the map viewer needs to understand the time-frame to which the map is pertinent. To add this, information we will first add text using the same procedure we just used for adding the map title. We left click on the "Add new label" tool, and then in the "Main properties" group (blue arrow) under the "Item Properties" tab (yellow arrow) we type in the following information: "Source: Map Makers Inc.", "Created: 4/30/2018", and "email:dmap@mapmakers.net". This is just example source material, but we will use it to create the text portion of our source metadata.



We resize the text using the procedures described for resizing the North arrow, and achieve our desired appearance of the "source information box". This is the last major map element that we will add.

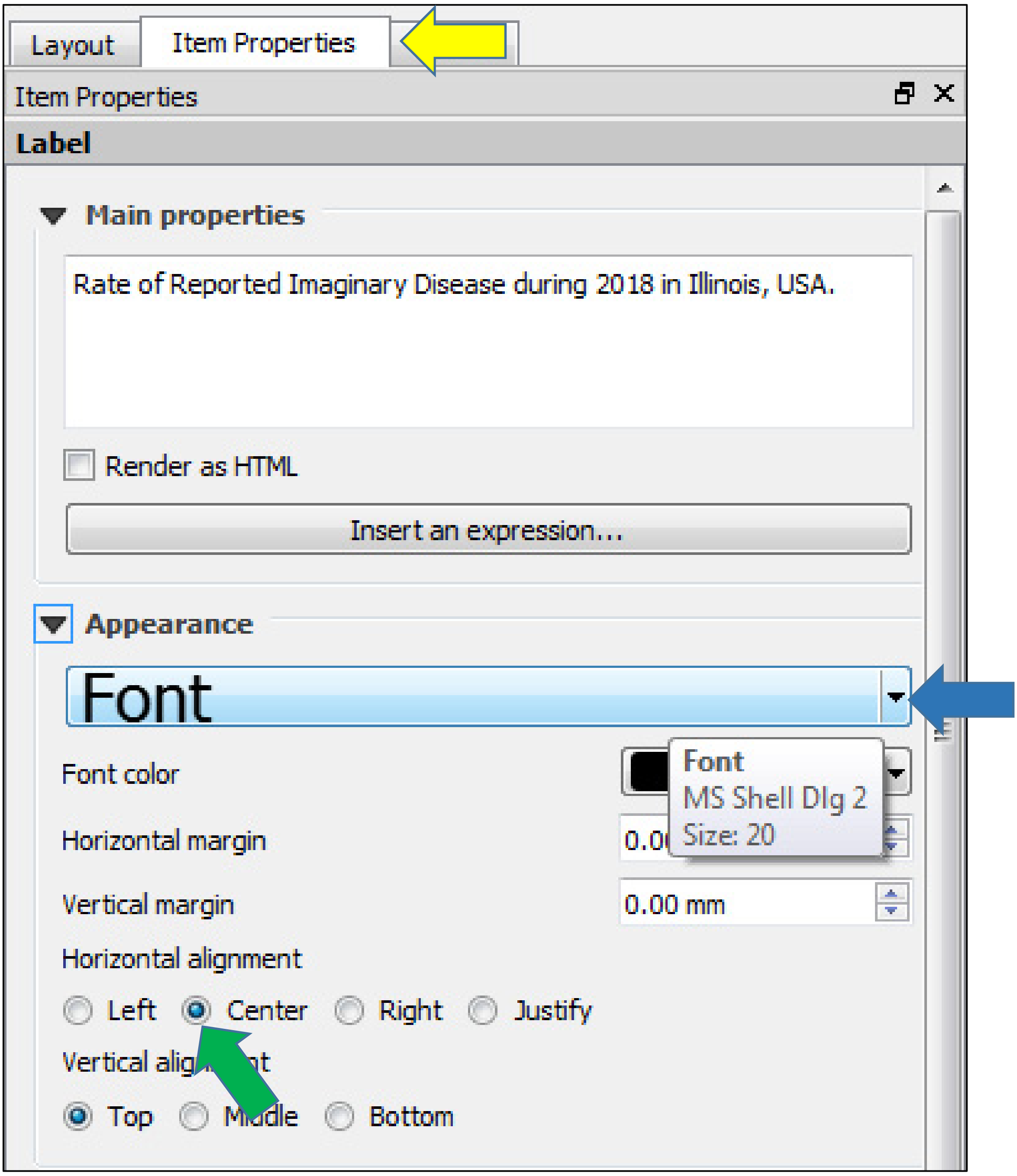

Rate (cases/km²) of Reported Imaginary Disease during 2018 in Illinois, USA.
N
W
E
S
Legend
0.000 - 0.029
0.029 - 0.069
0.069 - 0.112
0.112 - 0.233
0.233 - 0.375
Source: Map Makers Inc.
Created: April 30, 2018
Email: dmap@mapmakers.net
25 0 25 50 75 100 km

Better Map Design

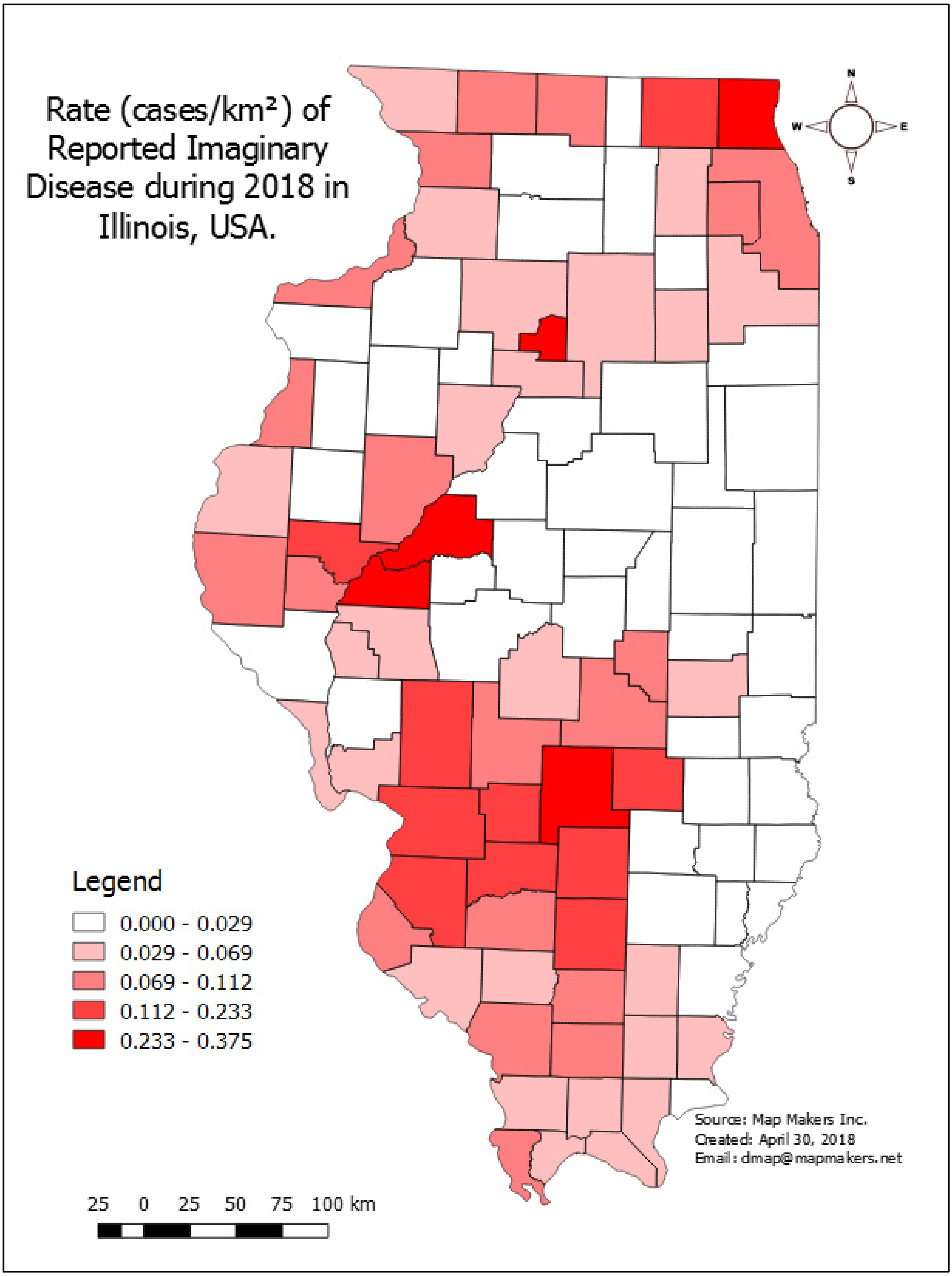
One of the key aspects of a well-designed map is that it needs to be aesthetically pleasing. A map that is attractive to the viewer will more quickly convey the intended message, and will allow the viewer to focus on the most important elements in the map. Conversely, a poorly constructed map will often be distracting to the viewer, and will require much more diligence on the viewers part to extract the relevant information. Thus, it is imperative for us as the map creator to ensure that the layout of our map compliments its intended purpose. Creating a good map layout, is generally an iterative process, and involves extensive trial and error. We will provide some tips and aspects to consider, and demonstrate them with the map layout we just created; however, ultimately, there is no single "best" layout, rather it is at the discretion of the map maker to determine what layout suitably fulfills the purpose of the map. A good potential reference on the subject of map design is the book, "Designing better maps: a guide for GIS Users" by Cynthia Brewer.

One of the major considerations when designing a map is the visual hierarchy. The visual hierarchy is the order in which the map elements are perceived by the user. In simple terms, the visual hierarchy describes how the map viewer's "eye" detects various map elements. The visual hierarchy of elements in a map is derived from their size, color, location in the layout and surrounding empty space. Generally, the title and the main map's features should be the most prominent aspects of the map, and other elements should be lower on the visual hierarchy. This means that when the viewer looks at the map their eye should be immediately drawn to the title and the main map, and as they scan the map they become cognizant of the other elements. This hierarchy is logical because most often the information contained in the title and main map are the data that the map creator is trying to communicate to the viewer. For example, if the North arrow is too large or too dark, this may move the North arrow too far up the map hierarchy. This would cause the map viewer's eye to be drawn to the North arrow rather than the title or the main map. This would create a confusing map layout.

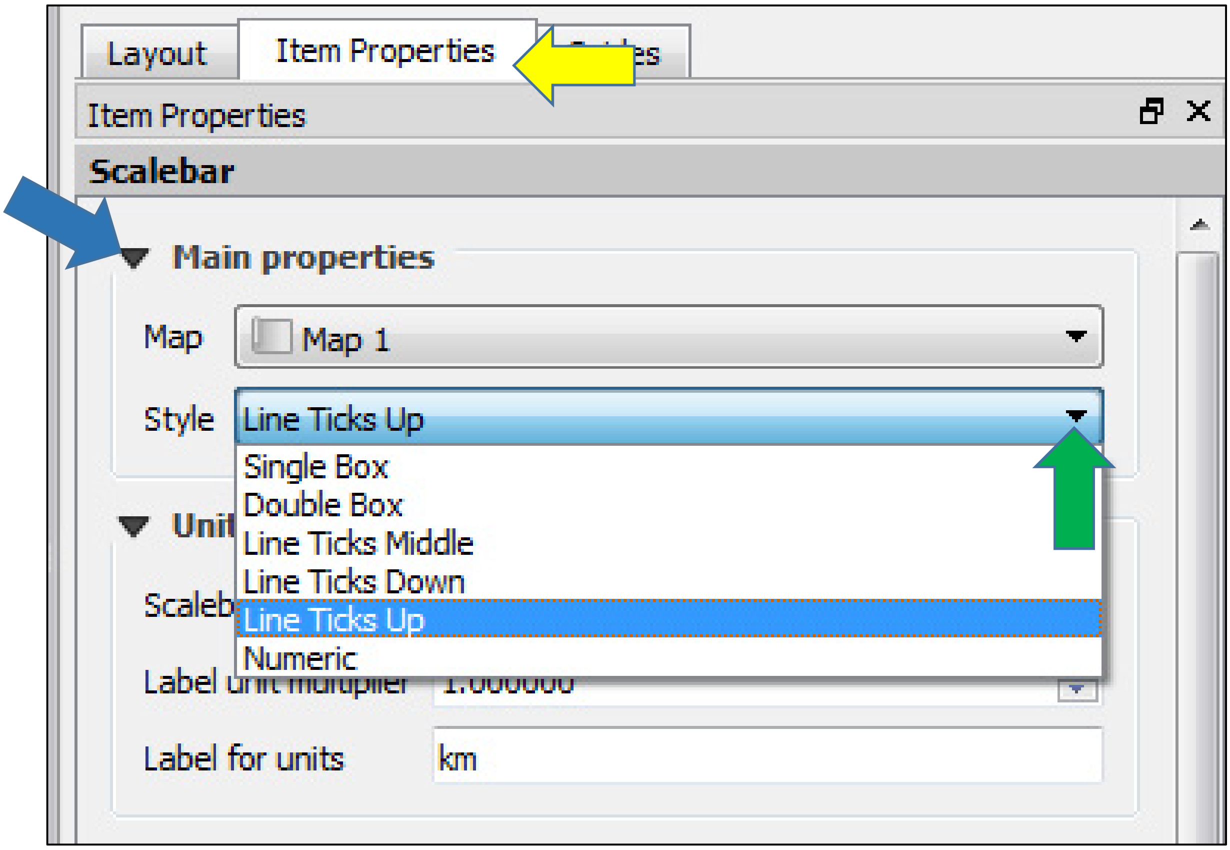
Examining the map we just created, we can see that the main map is indeed featured prominently in our layout. However, the title seems to need more emphasis, and the scale bar seems to be too prominent. So, we will correct these problems. First, we will change the positioning and size of the title. We left click on the title, and then go to the "Item Properties" tab (yellow arrow). We then left click on the black triangle (blue arrow) next to "Font" and change the font size from 18 to 20. Also the "(cases/km^2)" should be removed from the title and moved to the legend where it is more appropriate. We delete the "(cases/km^2)" from the title, and change the alignment of the text from "Left" to "Center" by left clicking on the appropriate radio button under "Horizontal alignment" (green arrow). We then decrease the size of the title's frame using the resizing techniques described previously for the North arrow, and reposition the title left of the main map rather than above it. We also reposition the main map to the right to accommodate the new title location. This will help to better integrate the title and the main map, and emphasize them as the most important elements on the map.



The new layout is shown below. This has now brought the title up the visual hierarchy as desired; however, it has also created additional layout issues. For example, the North arrow now is too large, the source information is too close to the main map, and we need to add "(cases/km²)" to the legend. We also need to bring the scale bar down the map hierarchy. We start by resizing the North arrow to reduce its visual prominence using techniques described previously. Next, we move the source information to allow better spacing from the main map. To move the scale bar down the visual hierarchy, we will change the style of the scale bar.

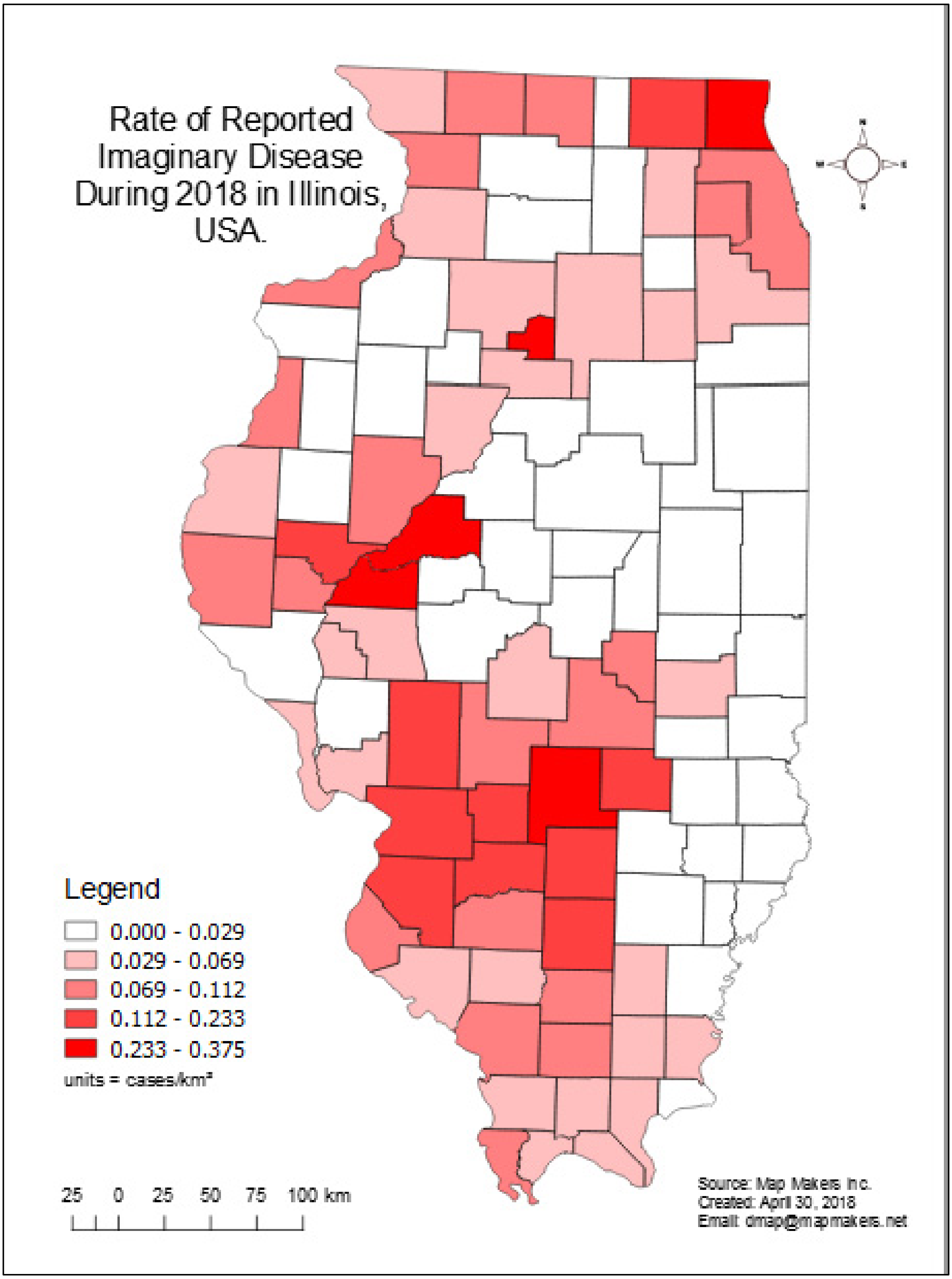

Rate (cases/km²) of Reported Imaginary Disease during 2018 in Illinois, USA.
N
W
E
S
Legend
0.000 - 0.029
0.029 - 0.069
0.069 - 0.112
0.112 - 0.233
0.233 - 0.375
Source: Map Makers Inc.
Created: April 30, 2018
Email: dmap@mapmakers.net
25 0 25 50 75 100 km

To change the scale bar style, we left click on the scale bar. In the "Item Properties" window (yellow arrow), we left click on the black triangle next to "Main properties" (blue arrow). We then left click on the black triangle next to "Style" (green arrow), and select the "Line Ticks Up" option. This will change the style of the scale bar to a less visually dominant style.

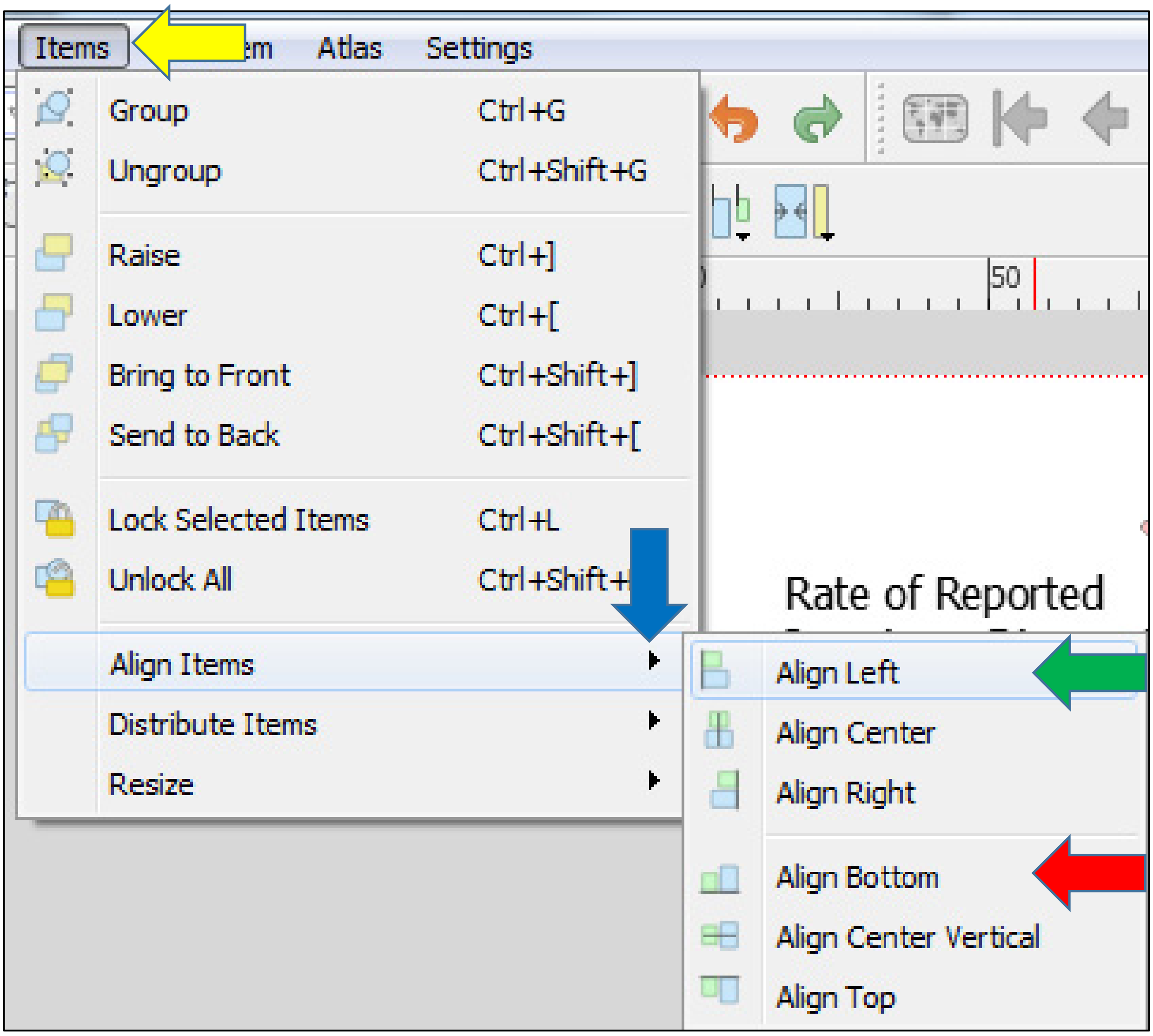


Next, we will add the "(cases/km^2)" to the legend. We simply will add a text annotation below the legend that reads, "units = cases/km^2". This will provide the map user with a quick reference to the units of the boundary values for each class displayed in the legend, but provides this information in the proper position of the visual hierarchy rather than in the title. We add the text by clicking on the "Add a new label" tool, and entering the "units = cases/km^2" in the "Main properties" group under the "Item Properties" tab. We will leave the text annotation's properties at the default setting. We then resize the frame containing this new text annotation, and position it directly below the last class of the legend. The new map layout is displayed below, and now it better captures the visual hierarchy.

There are still several aspects of this layout that could be improved. The first is alignment of map elements. Assuring that map elements are aligned is a simple way to increase the professional look of your map, and avoid the map having a cluttered appearance. In our example, we will align the left side of the legend, the text box containing, "units = cases/km^2", the scale bar and the title. We will then align the bottom of the scale bar and the source information. First left click on the title. While holding down the "Shift" key, left click on the legend, the text box containing, "units = cases/km^2", and the scale bar. All four map element should now be selected. Now go to top of the screen and left click on the "Items" menu (yellow arrow). This will open a drop-down menu in which we will select "Align

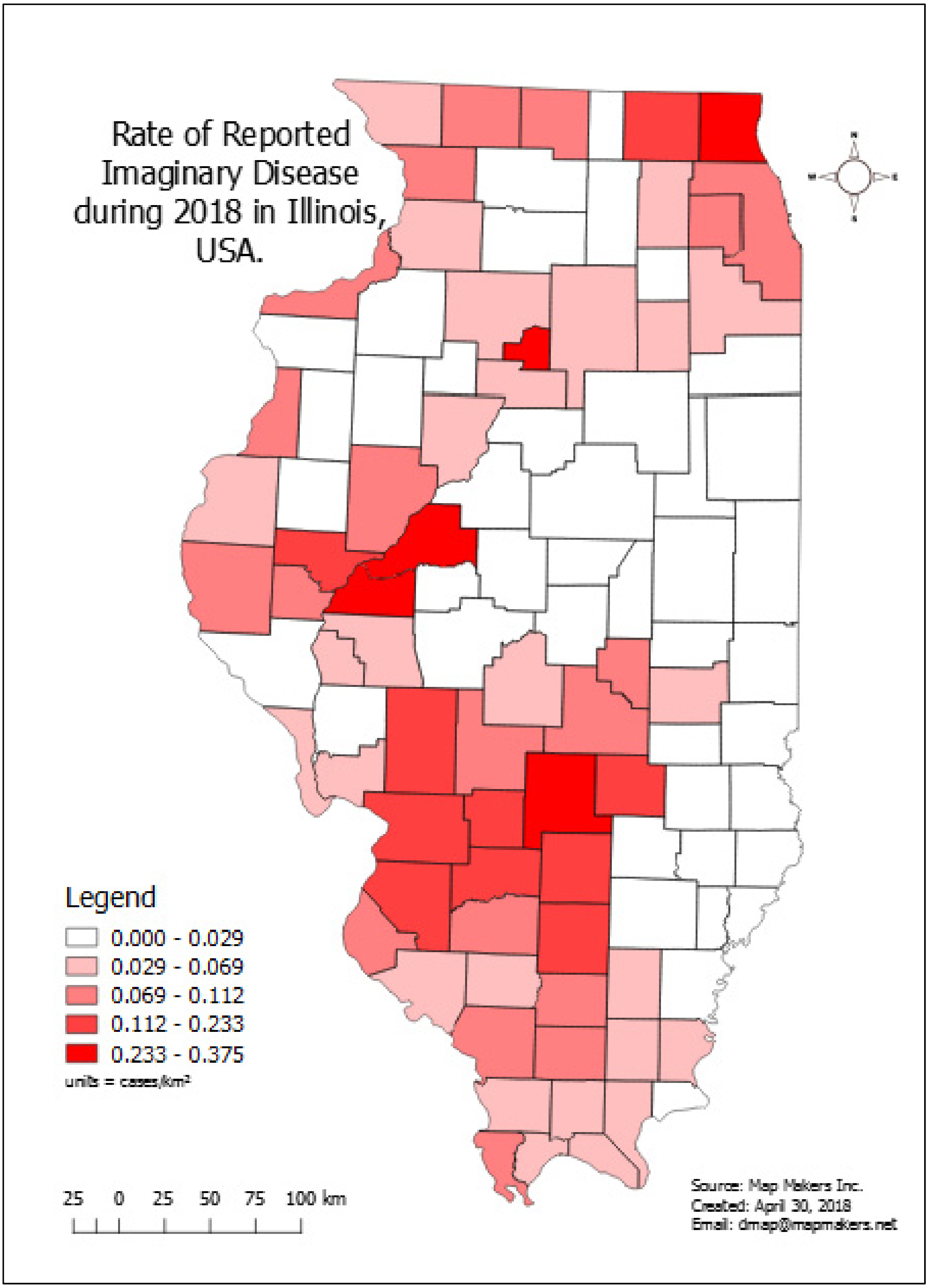

Rate of Reported
Imaginary Disease
During 2018 in Illinois,
USA.
Legend
0.000 - 0.029
0.029 - 0.069
0.069 - 0.112
0.112 - 0.233
0.233 - 0.375
units = cases/km²
25 0 25 50 75 100 km
Source: Map Makers Inc.
Created: April 30, 2018
Email: dmap@mapmakers.net

Items” (blue arrow), and then left click “Align left” (green arrow).



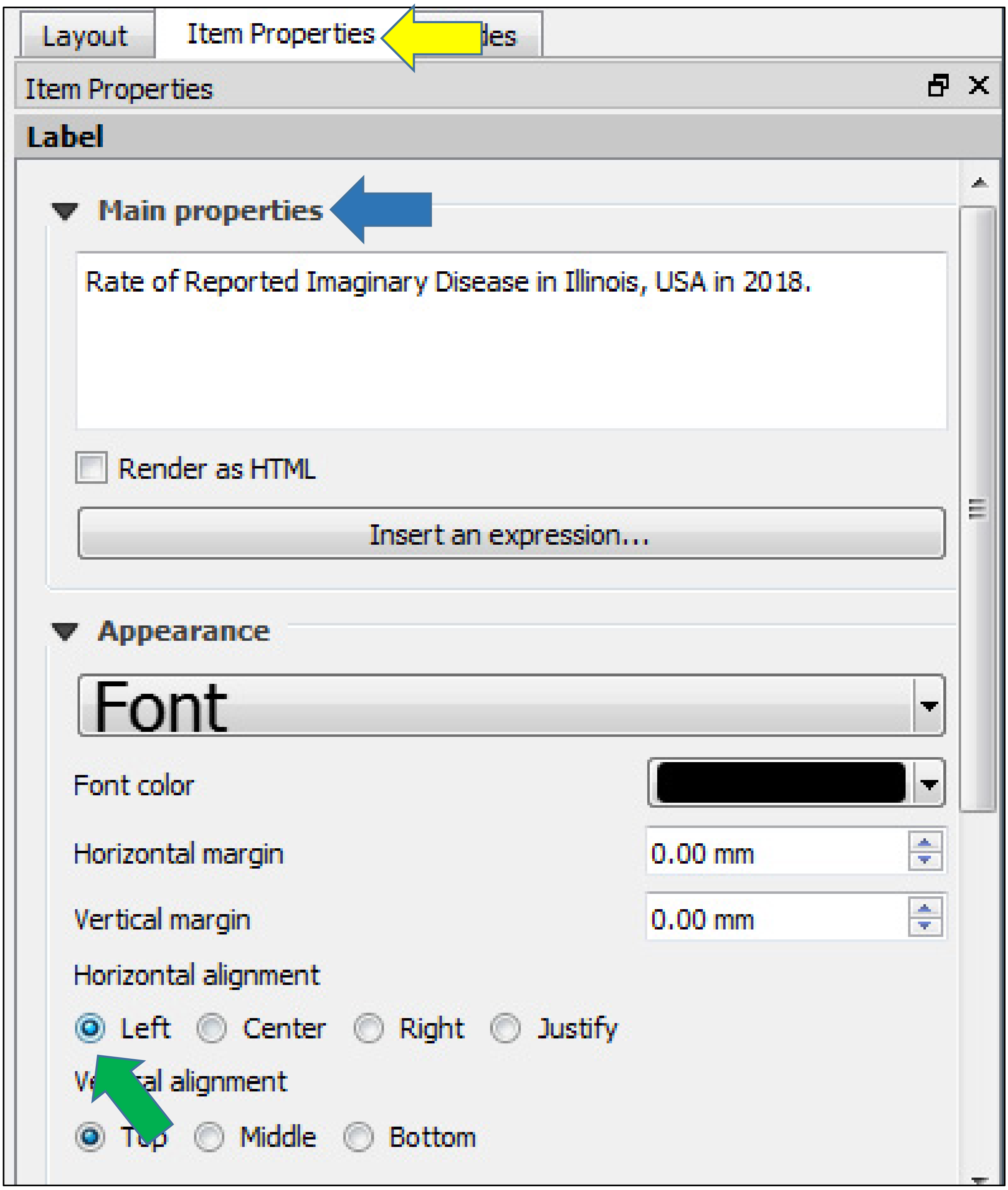
This will align the left side of each of these map elements automatically. We will repeat this procedure by left clicking on the scale bar, and then while holding down the “Shift” key left click on the source information text box. Now both these map elements are selected. We then go back to the “Items” menu, “Align Items”, and then select “Align Bottom” (red arrow above). This will align the bottom of these two map elements. Our map layout is now displayed below.

To emphasize the iterative nature of developing a good map layout, examining the current layout. We determined that the alignment of the title text would look better if it were displayed with left alignment rather than center alignment. Also, we should change the title text form “Rate of Reported Imaginary Disease during 2018 in Illinois, USA” to “Rate of Reported Imaginary Disease in Illinois, USA in 2018”.

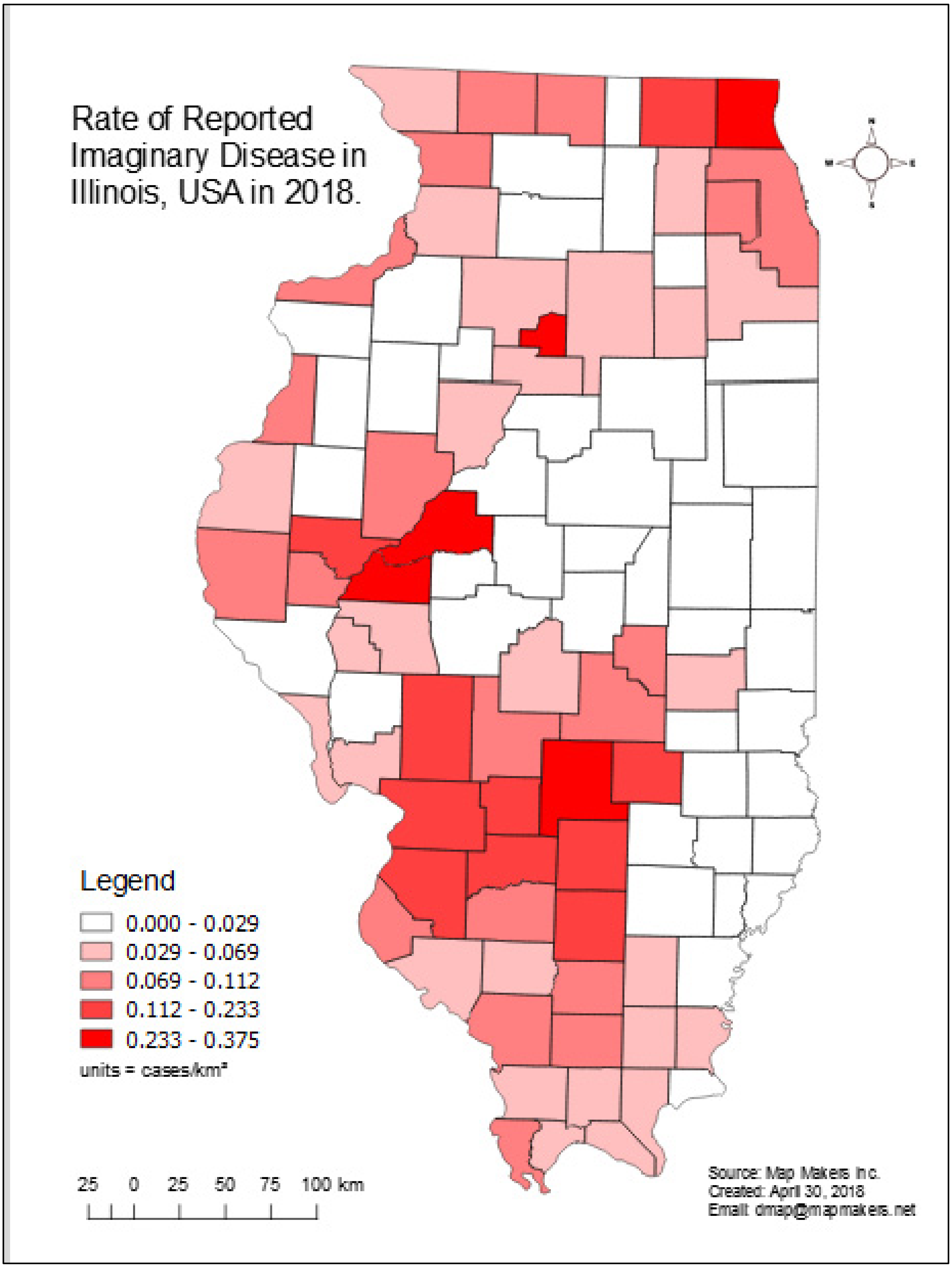


Rate of Reported Imaginary Disease during 2018 in Illinois, USA.
Legend
0.000 - 0.029
0.029 - 0.069
0.069 - 0.112
0.112 - 0.233
0.233 - 0.375
units = cases/km²
25 0 25 50 75 100 km
Source: Map Makers Inc.
Created: April 30, 2018
Email: dmap@mapmakers.net

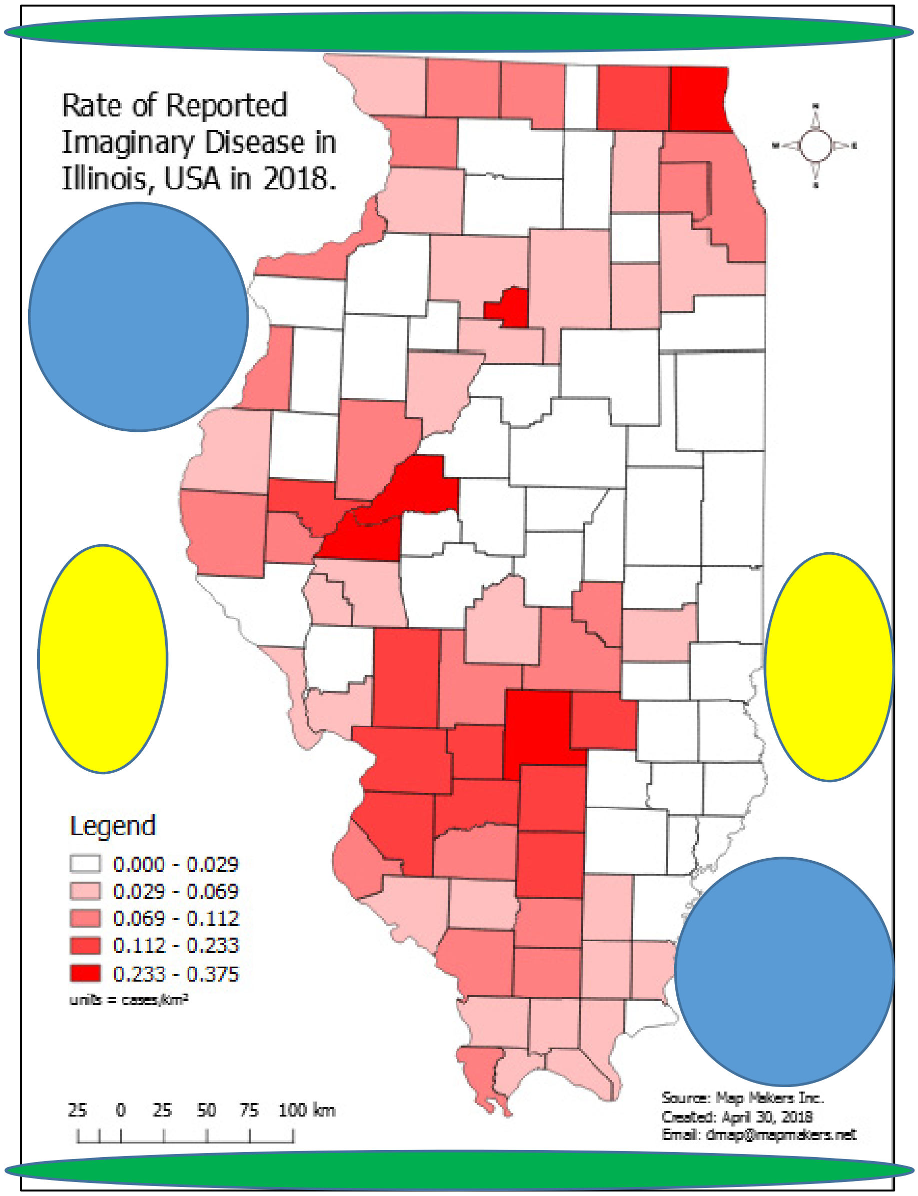
To make these changes, we left click on the title. Under the “Item Properties” tab (yellow arrow), we change the text in the “Main properties” group (blue arrow) to “Rate of Reported Imaginary Disease in Illinois, USA in 2018”. Lastly, we change the horizontal alignment from “Center” to “Left” by left clicking on the appropriate radio button (green arrow).



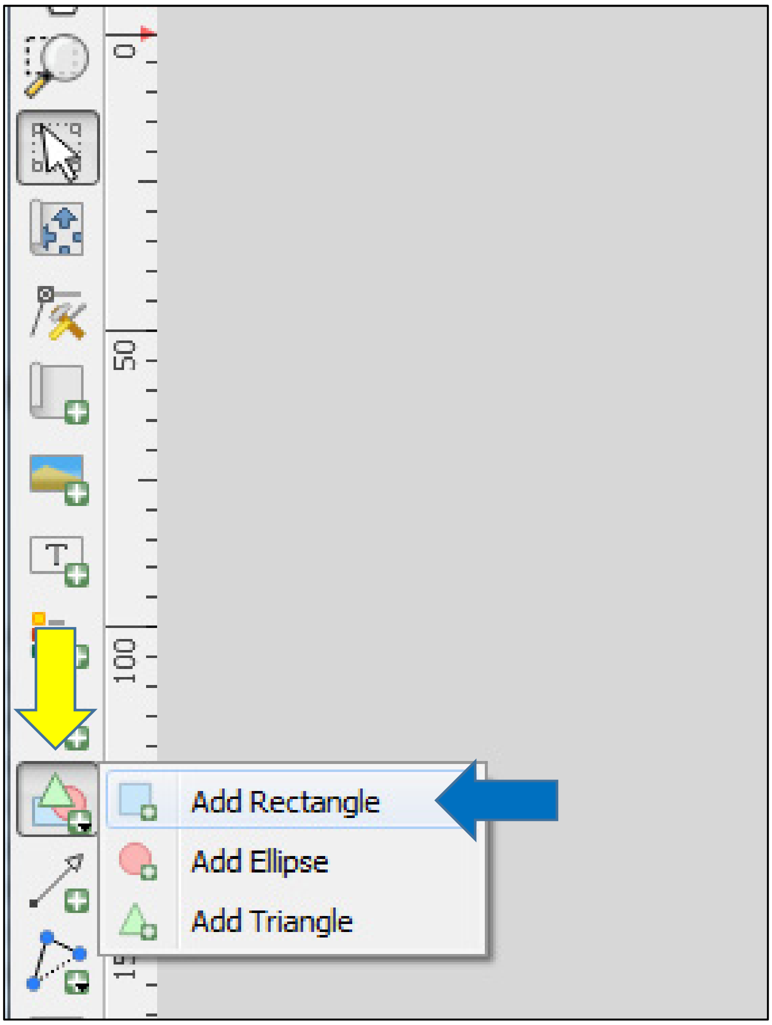
Our final map layout is shown below, which is a more professional and aesthetically pleasing map than our original version.



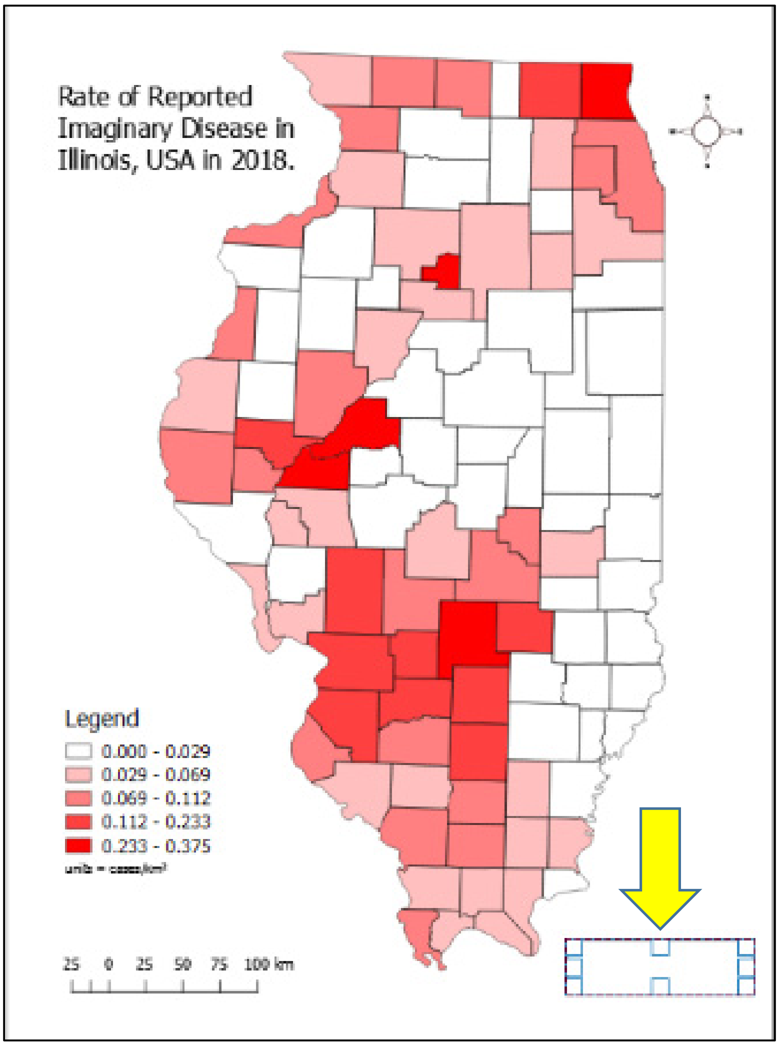
Although we have not explicitly stated it until now, an important component of creating aesthetically pleasing and interpretable maps is using white/empty spaces thoughtfully throughout the map. For example, in a complex map with many elements empty spaces may be used to break up the information to provide visual relief for the viewer, and present the information in a coherent manner. Additionally, it is an important to balance empty spaces. The idea behind balancing white space is that when an empty space is created on portion of the map, the map elements are positioned so that that space is balanced in other regions of the map layout. We have been implementing this idea in the map layout we have been designing by deciding where we place our map elements. If we look at the map layout we just created, we demonstrate where we have balanced our empty spaces (same colored circles). These empty spaces may not balance exactly due to the shapes of map elements, but despite being inexact they provide the desired aesthetic effects. Thus, it is critical to view empty spaces and their placement as an important component of the design of every map layout.



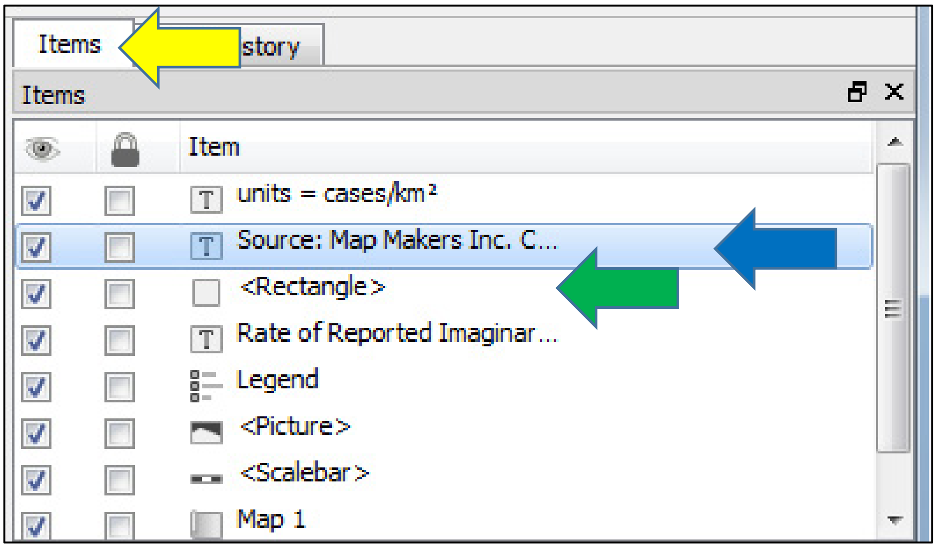
Another tool in creating a map layout is the use of boxes or neat lines. These elements are used to accentuate or separate other map elements from each other into logical groupings. For example, if a map layout requires a second map that denotes the location of the main map element (i.e., the main map is a district within a country, and the second map depicts the locations of that district within the country), a box may be placed around it to clearly separate the two maps. In our example, perhaps we want to put a greater emphasis on the source information. To do this we will draw a box around it, which will bring it up the visual hierarchy. To add a box, we click on the "Add shape" tool (yellow arrow), and select "Rectangle" (blue arrow).



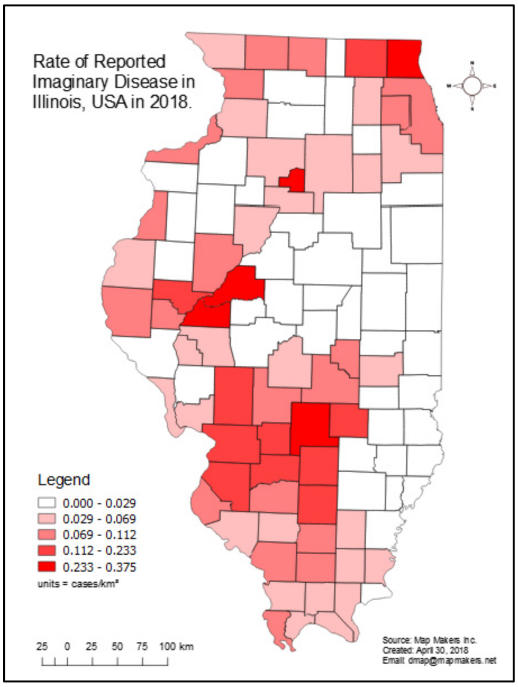
This turns the mouse cursor into a black cross. We move the cursor to where we would like to place our rectangular box, and then while holding down the left mouse button drag out the box to the desired size. The box can be resized and repositioned as needed using the procedures previously described for the North arrow. The box is created, but by default it has a black outline and is filled with the color white. This results in the source information being obscured from view as shown below.

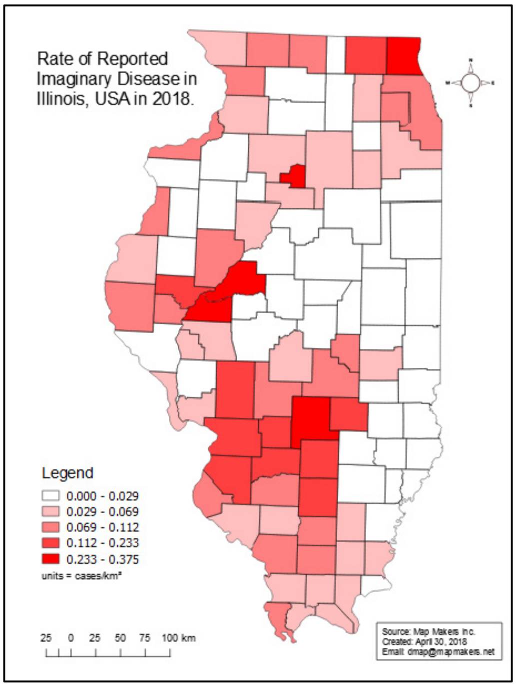


To remedy this problem, we could make the fill color for the box be transparent; however, we will instead demonstrate how to move the map elements up and down in visibility within the layout. The map layout can be thought of as a stack of map elements, and the map elements on the bottom of the stack may be obscured by map elements higher in the stack. This is the problem with which we are currently faced. The box is higher in the stack then the source information. So, we will move the source information above the box. To do this we go to the "Items" tab (yellow arrow) in the top right corner of the screen. Under this tab, is a list of all the map elements in the layout. We can turn on/off each element's visibility by left clicking on the left-most check box next to each element's name. To move the source information so that it is visible, we move the mouse cursor over the "Source:" text element (blue arrow), and then hold down the left mouse button and drag the element above the "<Rectangle>" element (green arrow). The source information is now visible as desired. This moving of map elements is a useful feature of QGIS to create effective map layouts, particularly when there are overlapping map elements.



Below, is a comparison of the map layout with and without the box around the source information. It is clear that the box moves the source information up the visual hierarchy, and more readily draws the map viewer's eye to that portion of the map.

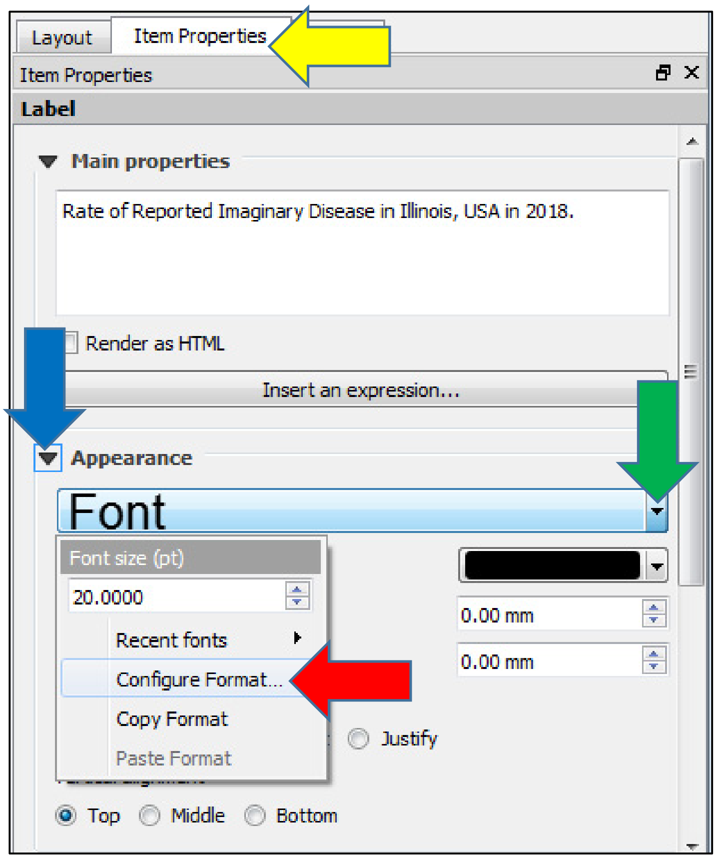




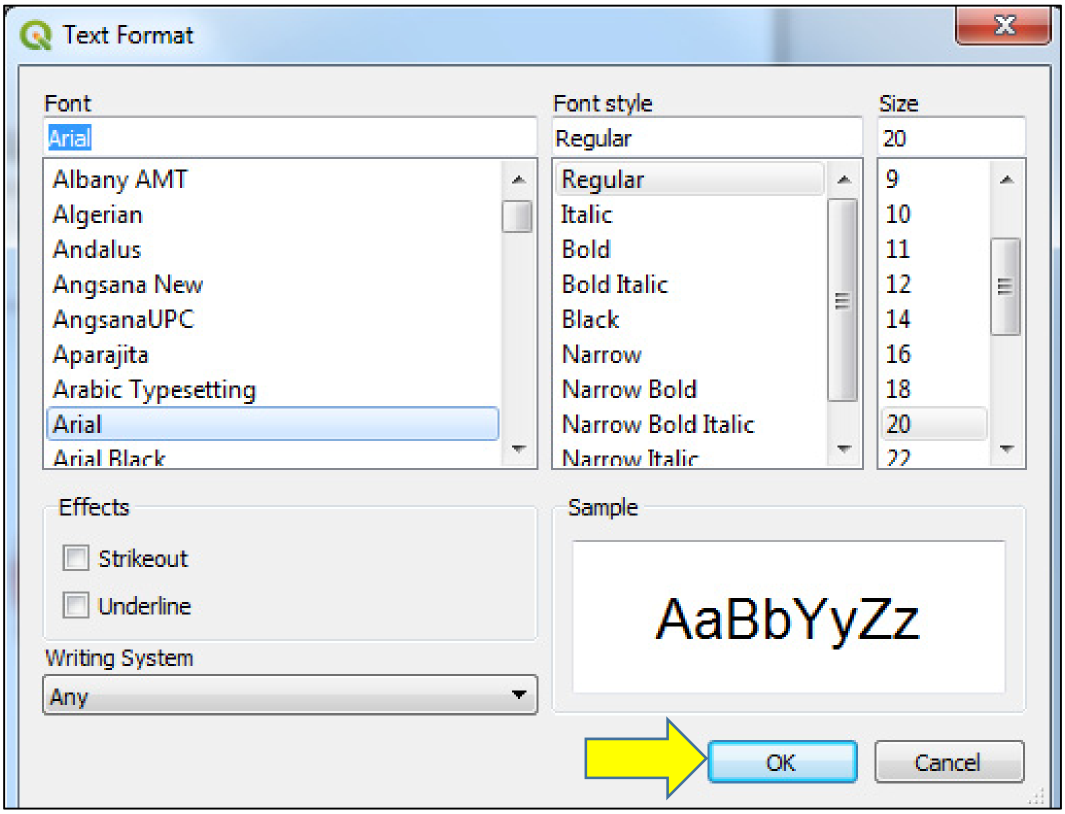
Neat lines and boxes can be important elements of any map layout; however, they should be used sparingly. Maps with many boxes may create a situation where it is difficult to balance the empty spaces effectively lowering the aesthetic value of the map. Additionally, boxes and neat lines by design break-up the "visual flow" of a map, and therefore when creating the layout it is important to decide if such boxes add value and enhance the interpretation of the map's information or do they make the map disjointed and lack continuity.

The last map design component we will briefly discuss is the choice of fonts. There are a wide breadth of font types available to use in your map layout, and many styles of each font are also options (e.g., italics, bold, etc.). However, care must be taken when choosing the font because it is tempting to select a "display" or "decorative" font when creating a map layout to give the map a "sophisticated or artistic feel". But often the simpler fonts are more appropriate to maintain the visual hierarchy as well as clarity of the map information. If a decorative font is chosen, it generally should only be used for the title. Another important consideration when choosing the font type is ensuring legibility of the text at the resolution that the map will be viewed. Experimenting with the font types at the appropriate resolution and size at which the map will be displayed, may help to select the most appropriate font. An additional consideration, particularly when sharing GIS maps in an electronic format, is the availability of the font that was used to create the map on the map viewer's computer. If the appropriate font is not available, the computer may substitute another font that may impact the quality of the map. Therefore, commonly available fonts such as Arial or Times New Roman may be safer to use for maps that will be widely disseminated in electronic format. Lastly, the size of the text is another important consideration that we have mentioned previously. The selection of the appropriate size effects the legibility of the map information, but also can move text up and down the visual hierarchy.

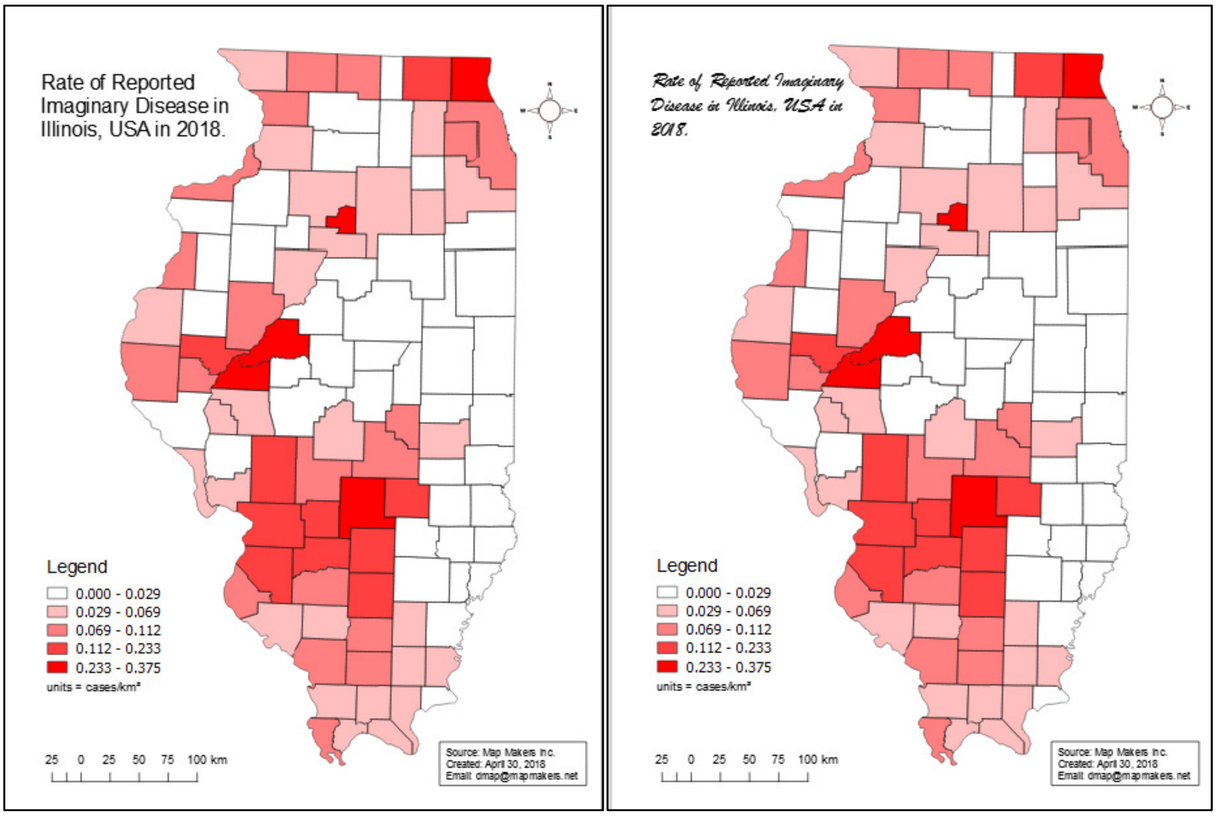
To change the font, left click on the text box containing the text for which you would like to change the font. This will select that map element. Under the "Item Properties" tab (yellow arrow), below the "Appearance" group (blue arrow), left click on the black triangle next to "Font", and choose "Configure Format" (red arrow).



This will open a dialog menu from which you can select and preview the font of your choice, change the style and determine the appropriate size. Once you have chosen the desired font, style and size, left click "OK" (yellow arrow).



Below, we compare a decorative font, Brush Script MT, for the title of our choropleth map to the font we initially chose, Arial. Although the new font provides some artistic tones to the map, it doesn't enhance the presentation of the information in the map. Additionally, it decreases the overall legibility of one of the most important text elements. Therefore, we would prefer the simpler Arial font for this map layout.



In summary, designing a map layout requires significant thought and a thorough understanding of the purpose of the map. In general, it is good practice to attempt to keep the map as simple as possible, which will help to keep the map viewer focused on the important information being displayed and not be distracted by frivolous map elements. Asking the questions: 1) does this map element strengthen or confuse the message of the map, and 2) where should this element belong in the visual hierarchy, when deciding whether to add a map element can be important guides to determine whether the element is necessary and help to establish its "look and feel". Experimentation with the layout and construction of map elements is essential to ensure that the most aesthetically pleasing and informative map is produced. Having others critique a map layout, is an effective way to assure that the message of the map is clear and the visual hierarchy has been constructed in the intended way. Lastly, it should be stressed that there is no single "correct layout". There may be numerous different, well-thought out arrangements that may be equally effective at achieving the purpose for which the map was created. Choosing among them may come down to artistic differences and personal preference. The key message is to think critically about the map layout of the map and how each element interacts and contributes to the "story" it is telling.

Messaging with Maps

The final topic we will discuss regarding wildlife health information dissemination via maps is establishing the message that we intend to convey with our map product. Arguably, we could have started with this topic because we have repeatedly discussed that a map's purpose largely drives the construction of a GIS, and how maps are designed as products from that GIS. However, we have elected to present this information at the end of this section so that we conclude with the most important topic of the section, and we can use the maps that we have constructed above as the basis for illustrating the salient points we are presenting, rather than focusing on the software details.

Key Questions

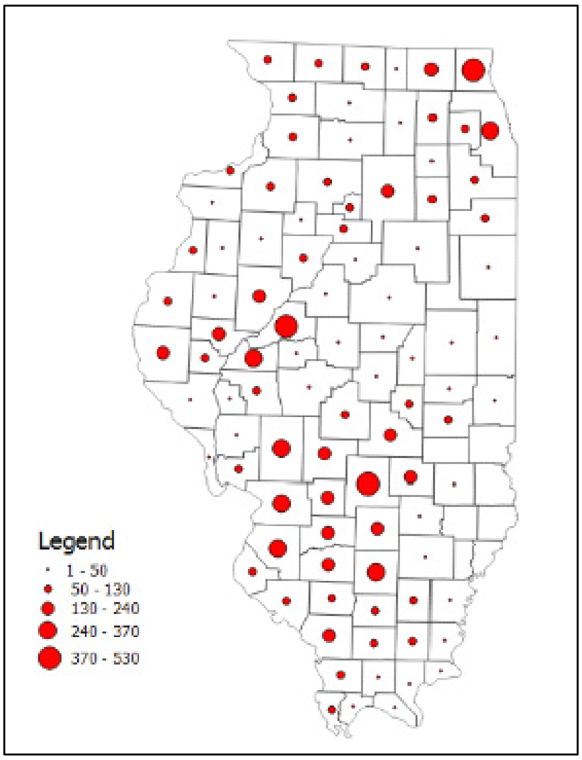
The first key aspect of messaging with maps is to define why is this map being created? Answering this question is the foundation of the creation of every GIS and associated map. Thinking about this question when viewing a map created by someone else, also helps us understand the underlying motive of the map maker, and to identify potential use or misuse of map elements to convey the creator's message. In our case, for example, the response to this question may be to provide situational awareness of an on-going wildlife disease outbreak; describe the amount of effort and/or results of disease surveillance efforts in wildlife; or perhaps communicate risk associated with a wildlife health event that may have implications for human or domestic animal health. Each of these potential example responses to this foundational question, may require different data layers being created/imported into our GIS and definitely demand varying map designs.

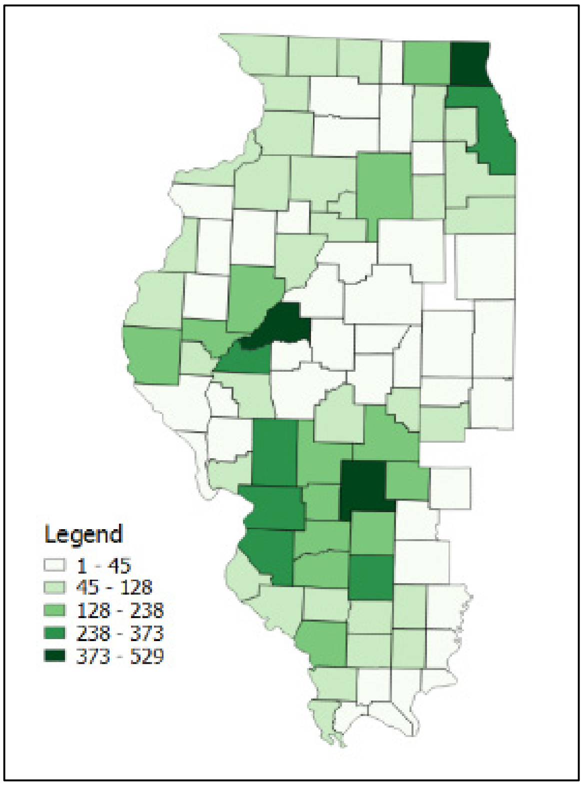
The second key question that should be answered is: who is the audience that will be viewing the map product? This question will determine the type and level of information that can be included in a map. For example, maps that will be shared with other agency scientists may contain detailed spatial information that can be used to develop scientific hypotheses regarding a wildlife disease outbreak such as aerial imagery or detailed infrastructure layers that can be used for planning outbreak responses. Maps that will be given to decision makers may focus more on disease response efforts or threat level to key resources, and avoid the detailed information essential to the scientists. Likewise, maps that will be disseminated to the general public may be quite simple with only a minimal number of essential map elements so that interpretation of the map is clear with minimal explanation. Thus, the "why" and "who" are questions that dictate GIS and map creation.

In the following, we will demonstrate several examples of how the answer to these two questions change the map appearance. We will use the example datasets we have been employing to demonstrate how to create a GIS and map layouts in QGIS. We will not go into the details of how to implement these maps in QGIS in this portion of the training; however, all the techniques we will use have been described above.

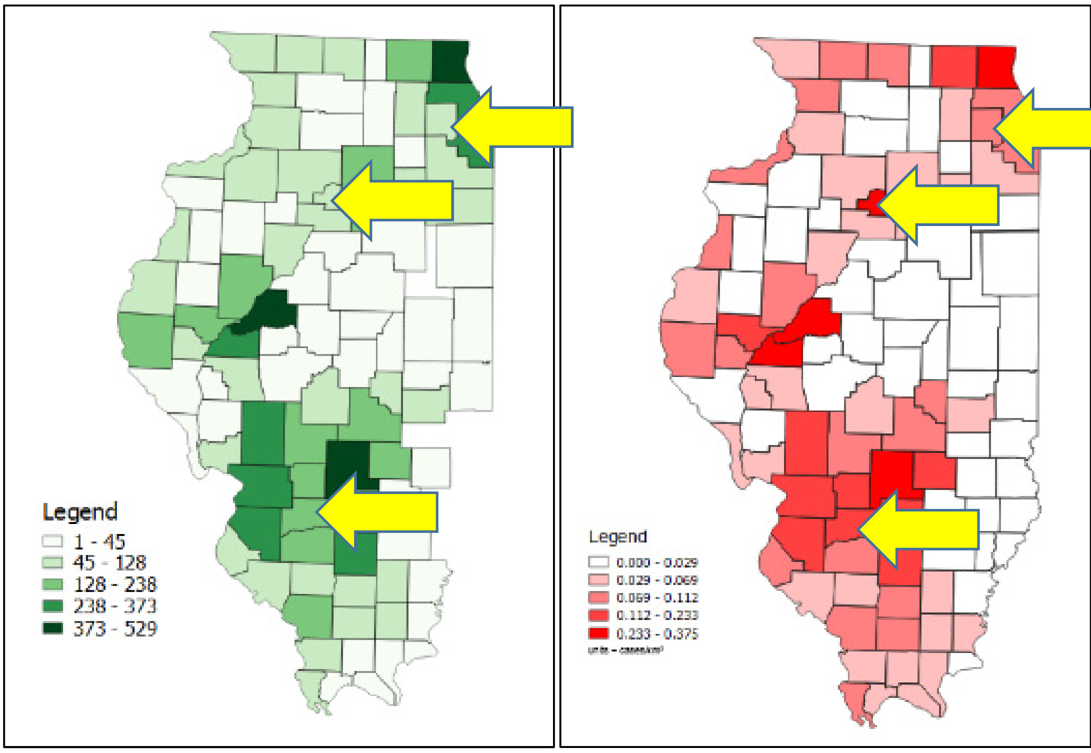
Map Symbol Choice

We will begin by creating two different types of maps, graduated symbols and a choropleth map, to depict total number of disease cases reported in each county in Illinois. Thus, the purpose of each map is to communicate to the map viewer the magnitude of the disease cases reported in each county, and provide a relative comparison across counties. If we examine each map critically, we see that the graduated symbols map is a better visual representation of the total number of disease cases because the large point symbols represent a high total number of cases in both large and small counties, and small symbols represent the opposite. Also, this map makes it easy for the viewer to ascertain the area of each county that contains each point symbol. However, in the choropleth map, the various colors or area symbols are misleading because area symbols in choropleth maps are usually used to display intensity or rate over the entire county. For example, a color may represent either a small county with a dense number of reported cases, or a large county with the same number of cases spread widely across the county. In other words, the choropleth map is not a good way to visualize the relative comparison across counties of the total number of reported cases. In general, it is best to use a point symbol map to display the magnitude of cases whereas the choropleth map is better suited to depict intensities or rates.


Legend
1 - 50
50 - 130
130 - 240
240 - 370
370 - 530


Legend
1 - 45
45 - 128
128 - 238
238 - 373
373 - 529

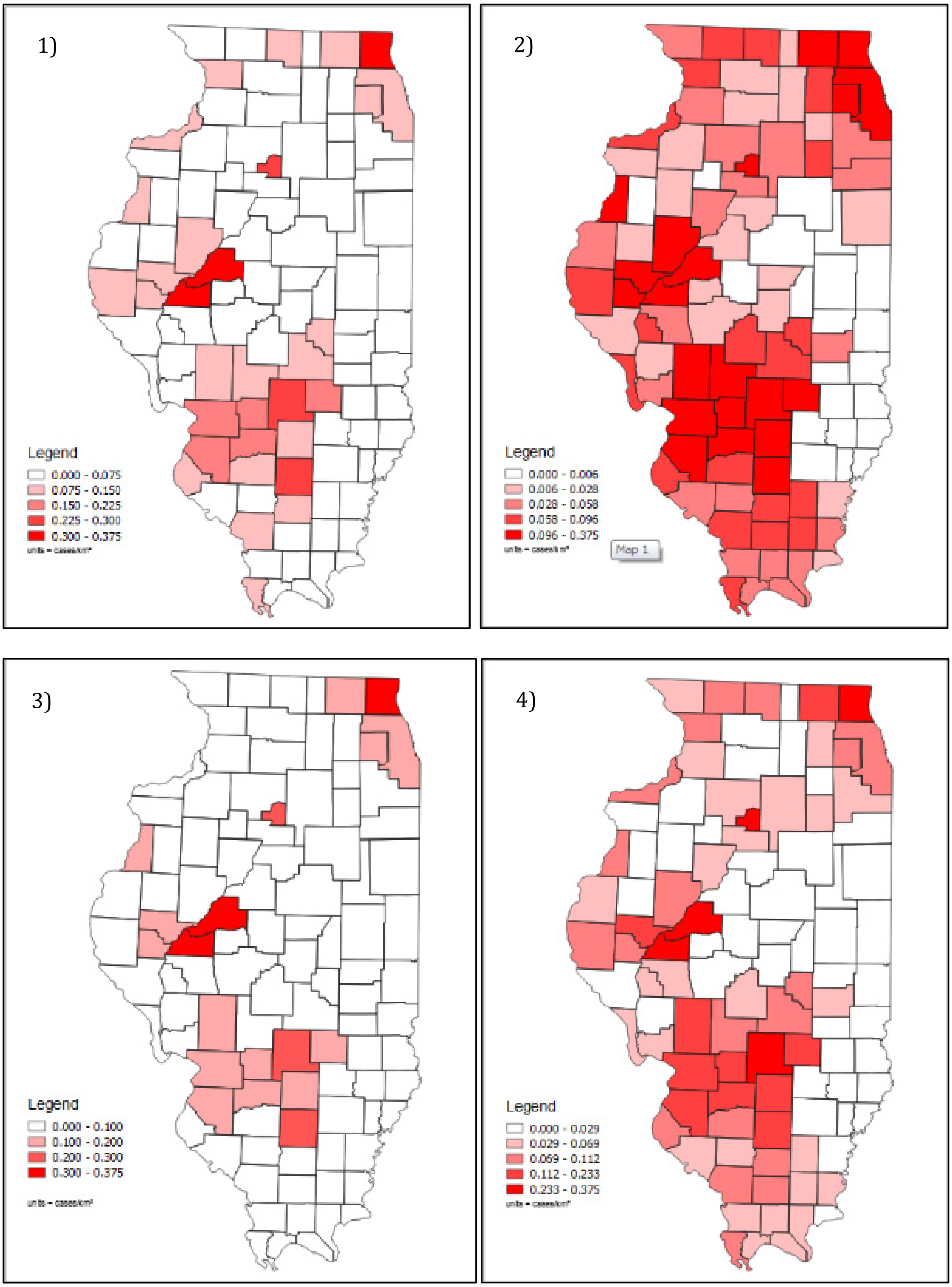
The maps below compare the choropleth maps showing total number vs. intensity of cases. Several counties where magnitude of disease cases vs. rates provide differing perceptions to the map viewer are highlighted (yellow arrows).



By accounting for area in the second map, several counties that seemed less important when looking at the total number choropleth map now rise in importance when we see the relative number of cases is high given their area. Thus, the key message is that the map creator needs to be cognoscente of the importance of their choice of symbols, and should match their symbol choice to the type of statistical measurement that the map is trying to display.

Aggregation Effects

A second important topic to consider when creating maps that consist of statistical summaries is the impact of the how observations are aggregated or broken into classes (i.e., class breaks scheme). The class breaks scheme can significantly impact the message portrayed by the map. We will use the choropleth map depicting the disease case rate to illustrate the potential issues. Note, the data used in each of this map is the same, only the class breaks are changed. Below, we show four different maps with class breaks created using the following breaking rules: 1) equal-interval; 2) equal count; 3) pretty breaks; and 4) natural breaks (i.e., the method that we have used previously). Note the differing boundary values for each class between the maps.


1)
Legend
0.000 - 0.075
0.075 - 0.150
0.150 - 0.225
0.225 - 0.300
0.300 - 0.375
2)
Legend
0.000 - 0.006
0.006 - 0.028
0.028 - 0.058
0.058 - 0.096
0.096 - 0.375
Map 1
3)
Legend
0.000 - 0.100
0.100 - 0.200
0.200 - 0.300
0.300 - 0.375
4)
Legend
0.000 - 0.029
0.029 - 0.069
0.069 - 0.112
0.112 - 0.233
0.233 - 0.375

Comparing each of these maps, demonstrates the impact of the seemingly innocuous choice, of how to aggregate the data into classes. Clearly, this choice can dramatically affect the resulting message portrayed by the map. So, which map is correct? The answer is they all are correct and they all are wrong…it depends on the message that the map maker is trying to convey and for whom the message is intended.

If we examine map 1, we see that the class breaks boundary values increase by 0.075 and the resulting map depicts most of the counties as white indicating low rates of reported disease cases with only a few showing high rates. This class breaks scheme would be useful if the message of the map is intended to characterize the impacts of the disease are mainly limited to a few counties. Examples of potential uses for this might be to justify to decision-makers focusing response efforts on a few counties when resources are limited, or if the intent is to provide the public with the map that emphasizes that most regions have low disease case rates.

In map 2, we used an equal count class breaks scheme, which creates class breaks such that each class has the same number of counties within it. In contrast to map 1, map 2 emphasizes that most counties have been affected by this disease outbreak with few counties being white. Example messaging for which this map would be useful is to demonstrate the wide geographic area that has been affected, or if the purpose of the map is to enhance public awareness and to encourage taking appropriate disease response efforts.

Map 3 breaks the counties into classes whose breaks are evenly spaced based on simple round numbers (i.e., 0.1). Map 3 is similar to map 1, but is more extreme in the sense that more counties are colored white, and fewer are classed as having a high case rate. This map would portray a similar message to map 1, and the simple class breaks would make it a good candidate for public dissemination.

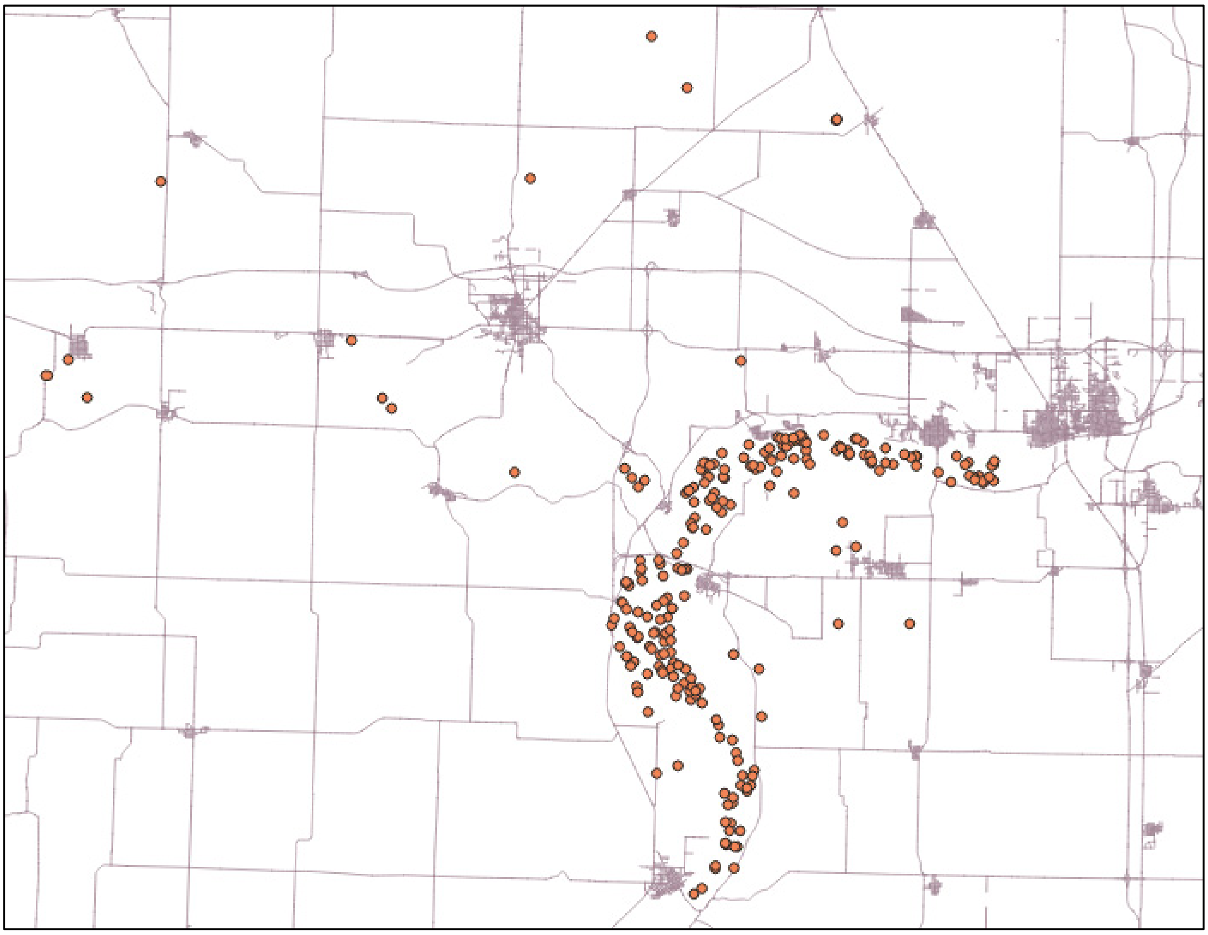
The final map uses the empirical distribution of the case rates across the counties to choose where to create the class breaks. This map provides an intermediate display between maps 1 and 3 and map 2. The potential messaging that for which this map may be useful is to provide a reasonable depiction of the geographic extent of the outbreak, but to also accentuate the most impacted counties, as measured by rate of reported cases.

Thus, the key message is that GIS provides many useful tools including a many built-in options to aggregate data for use in maps. However, a variety of these options or user-defined methods should be explored to ensure that the purpose and intended message of the map is clearly presented in a manner that is easily interpreted by the intended audience. Additionally, given the visual impact of the choice of class breaks, it is important when viewing these types of statistical maps to look at them critically to determine if the assertions depicted in the map or accompanying text are truly supported by the data, or are they an intentional or unintentional artifact of aggregation choices made by the map creator.

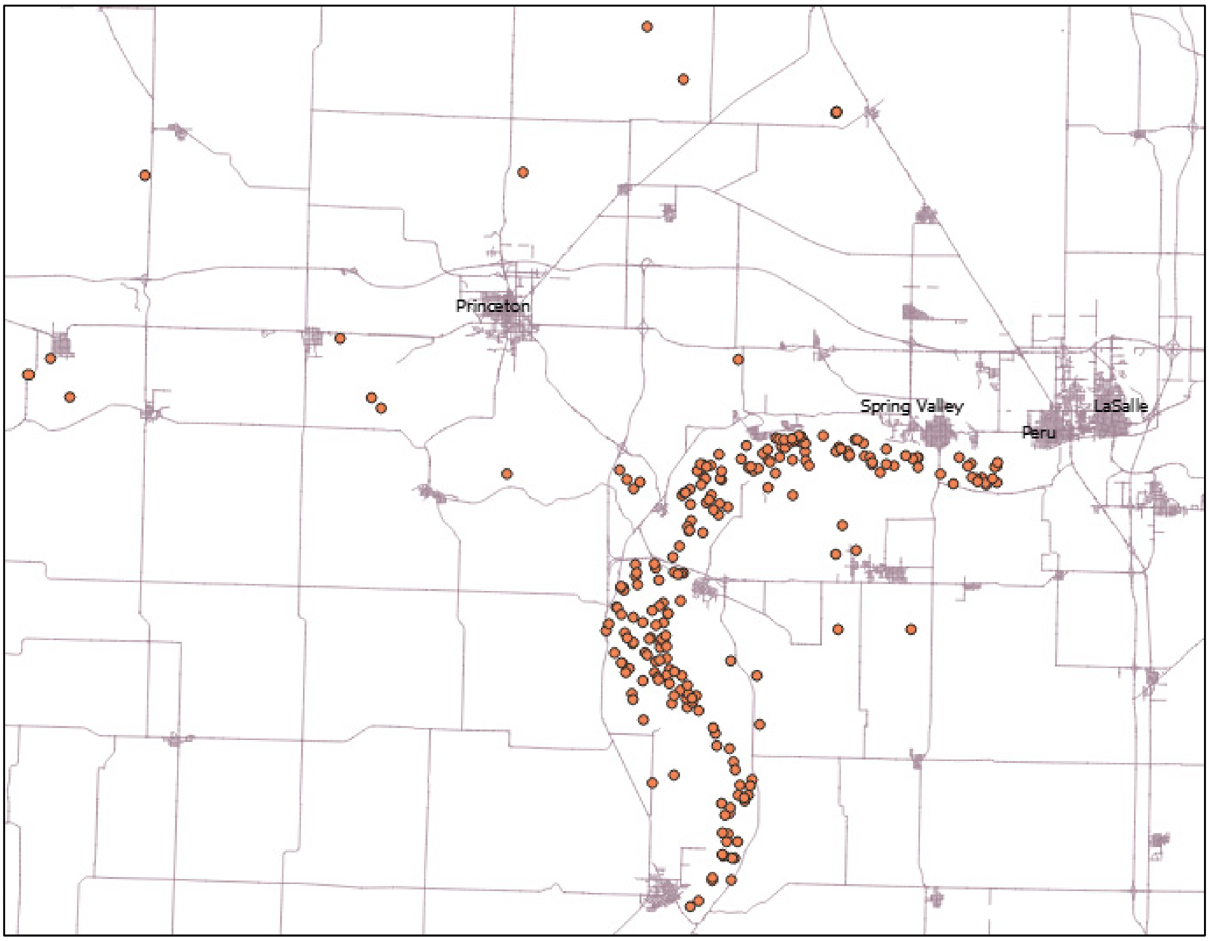
Call outs

We will define call outs as text elements that are inserted in the map layout to improve the aesthetics of the map or more often to enhance the intended message for a specific audience. Our interest lies in the latter case. A potentially useful call out that can be important, particularly when dealing with wildlife diseases that have a zoonotic potential or may impact domestic livestock, is the addition of city or place names of importance. We will demonstrate the use of call outs using the location data for the reported disease cases in a two county region in Illinois, and displaying these locations in a point layer.

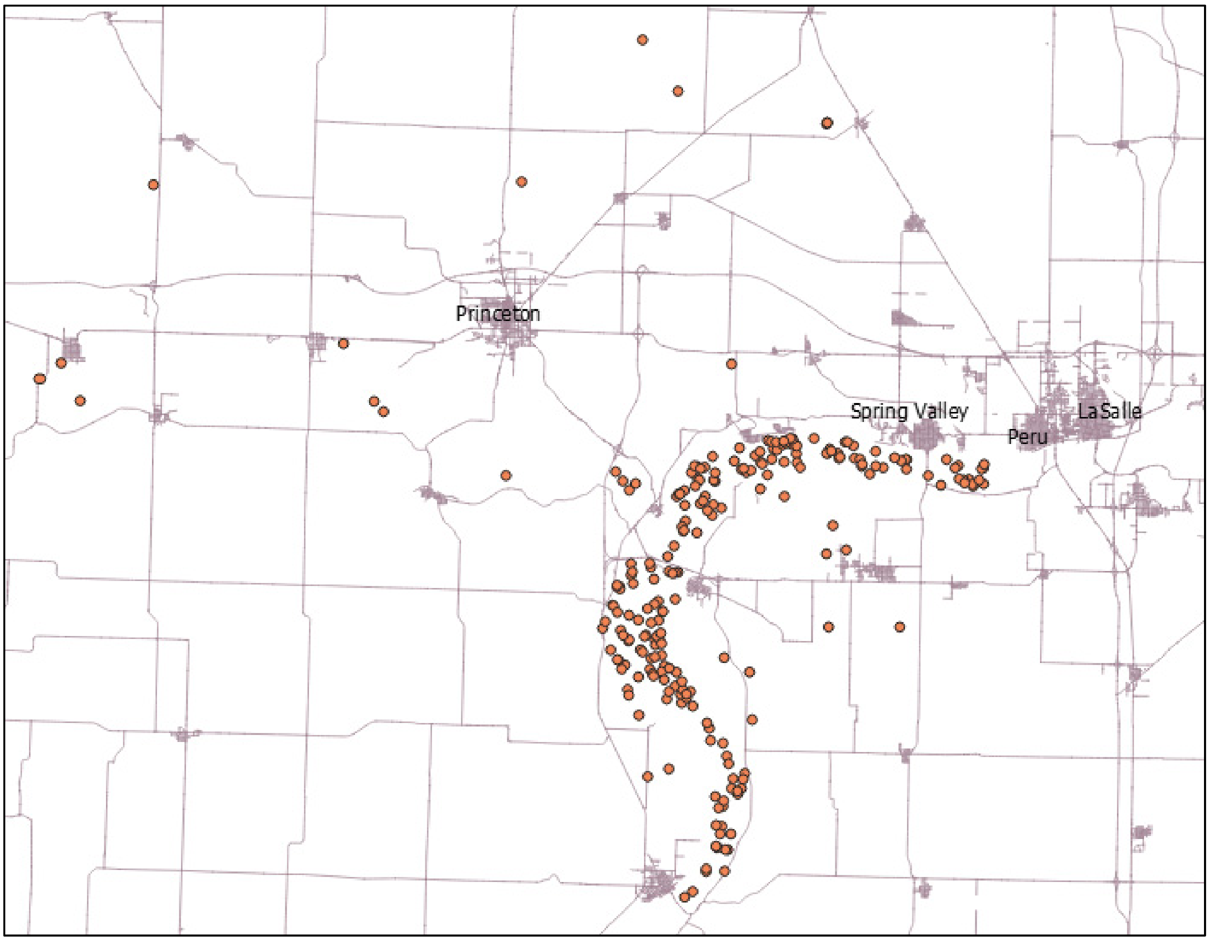
Our first map simply depicts the roads and location of disease cases reported. The map adequately shows the spatial information tied to these cases, but the significance of these cases is not readily apparent. Thus, this map could be improved.



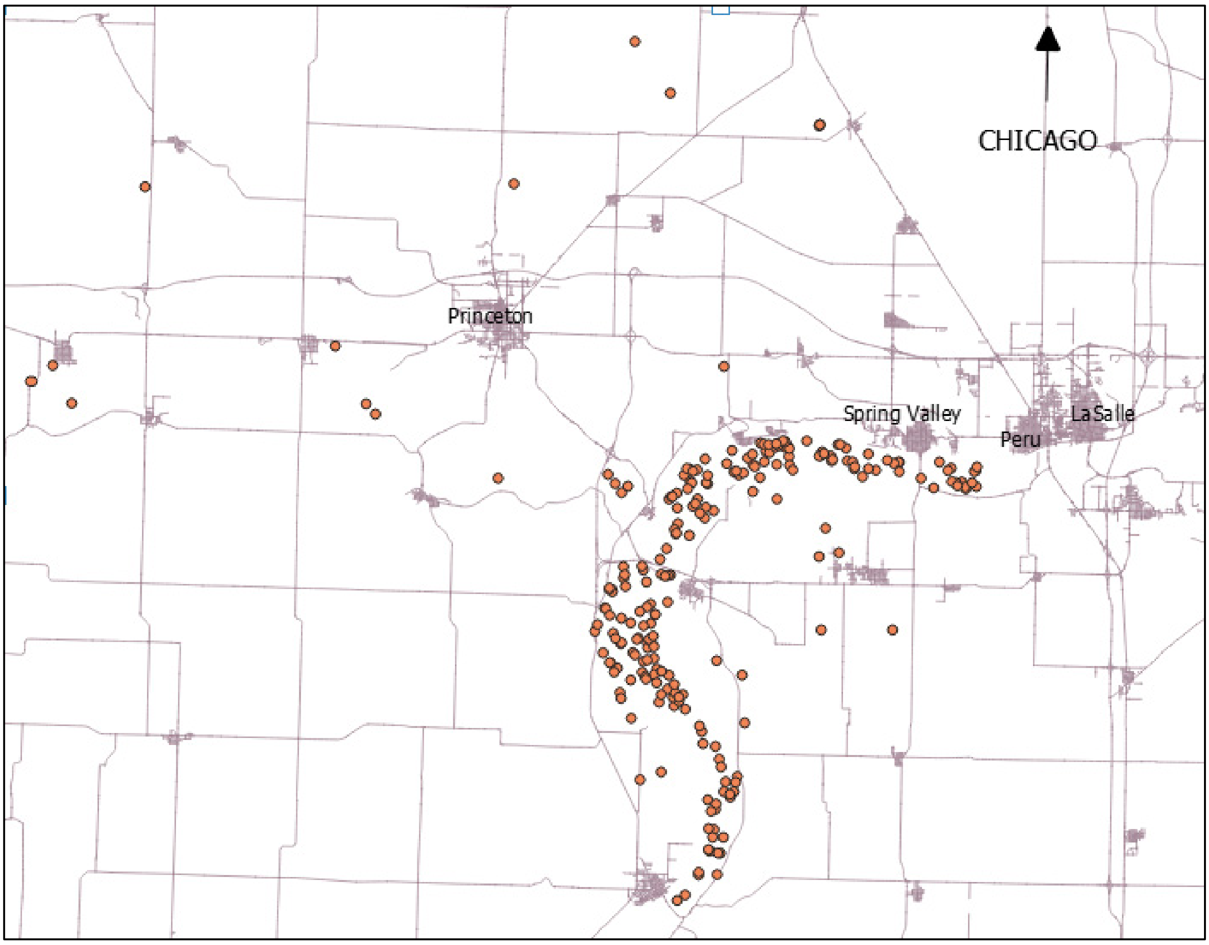
We can add the names of the local towns and villages to allow the map viewer to visualize the human or domestic animal risk associated with this wildlife disease outbreak. The following map includes the callouts for the town names. However, the emphasis of this map is still on the location information of the disease cases because the call outs are down too far on the visual hierarchy.



To bring the town names up the visual hierarchy, we will increase the font size of the call outs in the following map. The greater emphasis on the town names in the map layout helps the viewer to be visually drawn to the call outs and then to the location of the disease cases. Inherently, this should lead to the viewer mentally connecting this spatial information, and then begin to consider the risk posed by this outbreak to the more densely populated regions of the map.



This final map may be suitable as constructed, particularly if the intended audience has the requisite knowledge of the local area and will understand the significance of these towns. But, perhaps the map is intended to seek resources from a state agency for a disease response effort. Decision makers at the state-level may not believe that this outbreak is worthy of resource allocation given this map presents the problem at essentially a local-level. This is problematic because increasing the scale of the map will help put this outbreak in a more appropriate context (i.e., state-level), but it will also result in a loss of clarity of the location information (i.e., points will be overlapping). In these cases, sometimes an additional callout(s) can be particularly useful. In our example, we will add a call out that indicates the direction towards Chicago, the most populous city in Illinois. We will use a large font so that this this call out will be high in the visual hierarchy. Although Chicago is hundreds of kilometers from this site, including its name in call out makes the map viewer associate this outbreak with Chicago. Thus, this map might be a more useful tool for attempting to acquire resources from decision makers at the state-level.

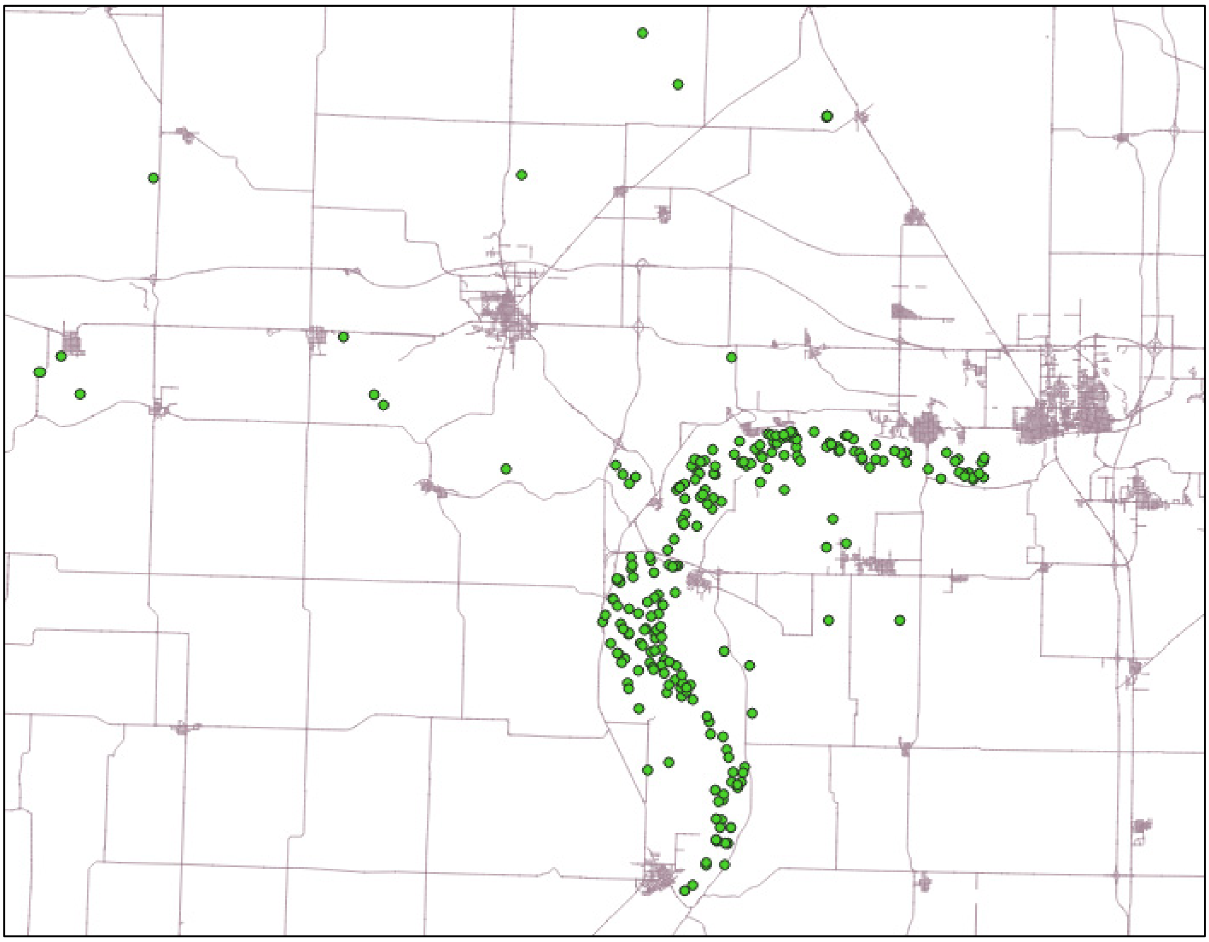


The key message here is when mapping wildlife health information often times the purpose of the map is to communicate risk to viewers. It is critical to think about the audience, and how they will perceive the risk after viewing the map. Using call outs provides a tool to effectively enhance a disease map to assist the viewer in associating the mapped disease information with risk to highly valued places or resources.

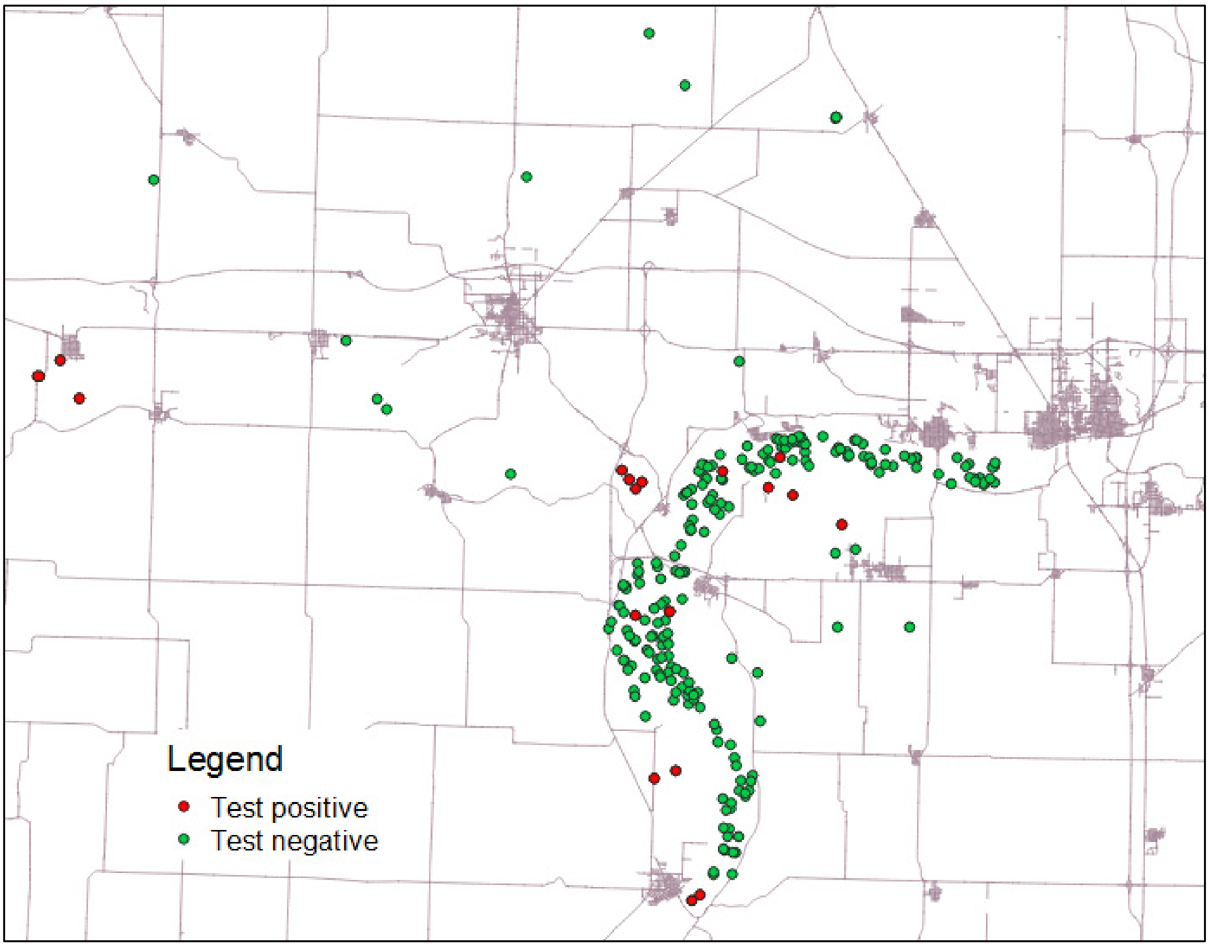
Color Choice

When designing maps, there are many aspects to the choice of the color of map elements that can enhance or detract from the message a map is intended to convey. Every color is made up of three components: hue, lightness/value, and saturation. Hue is associated with the various electromagnetic wavelengths, and is what we normally associate with color names (e.g., red, green, etc.). Lightness or value refers to the how light or dark the hue is that is being displayed, and saturation refers to the vividness of the color. These three components are what make up the colors that map viewers perceive. In general, when designing a map, hue and lightness are generally the components that are manipulated; however, there are instances where the choice of saturation level is important. It is also worth noting that humans can only distinguish a finite number of different color shades, so it is best to try to minimize the number of colors used in maps so that color differences and associated classes can be readily distinguished when viewed.

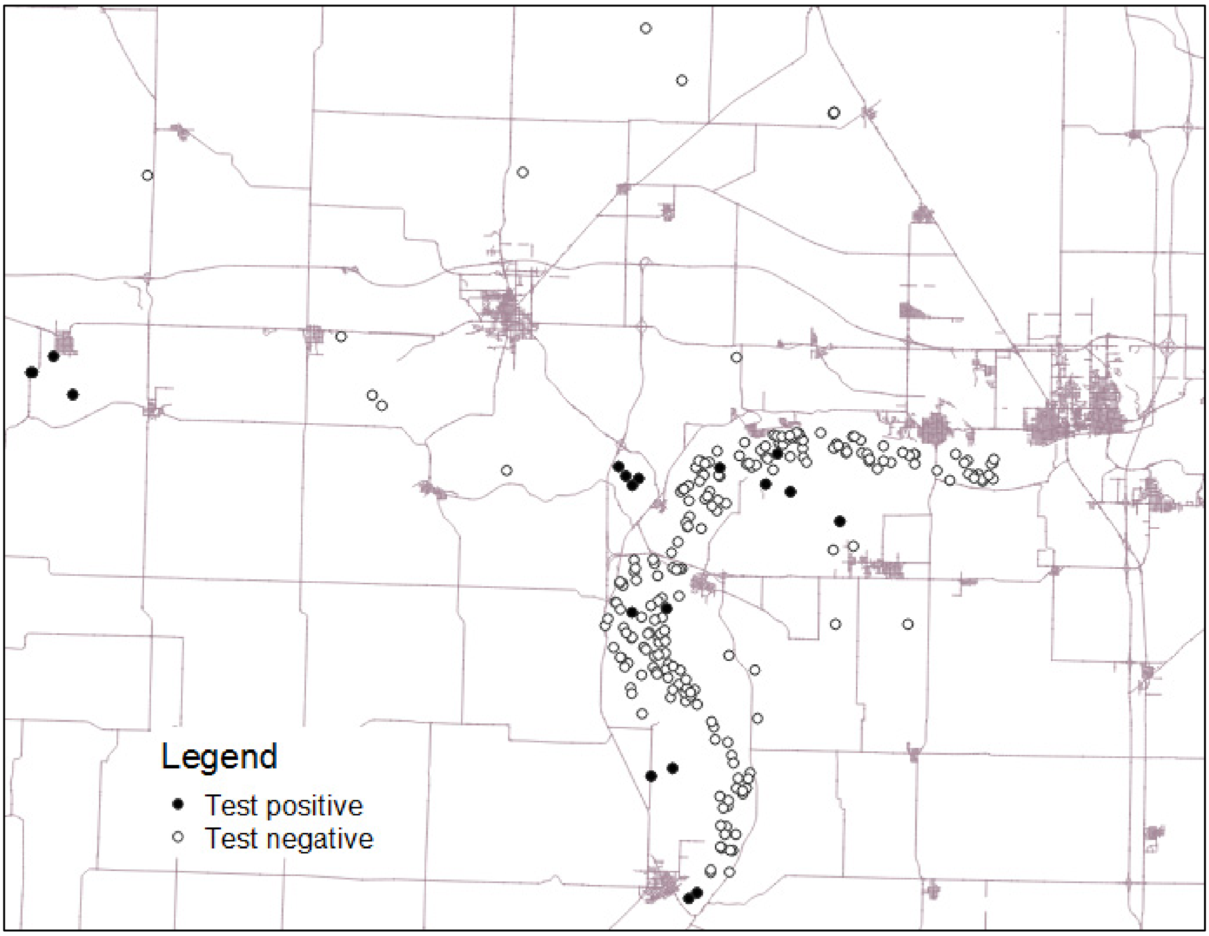
A common mapping task in wildlife health is the depiction of the results of wildlife disease surveillance. Using the same point layer, let's assume that the points depict the locations of animals that were tested for a disease as part of a targeted surveillance effort. The map below depicts solely the locations of animals that were tested for disease, but does not provide any information on the disease status of the individual animals sampled. For example, it may be useful for describing to local residents the sampling effort that has been conducted in the region. It may also be useful for agency personnel to assess gaps in spatial sampling.



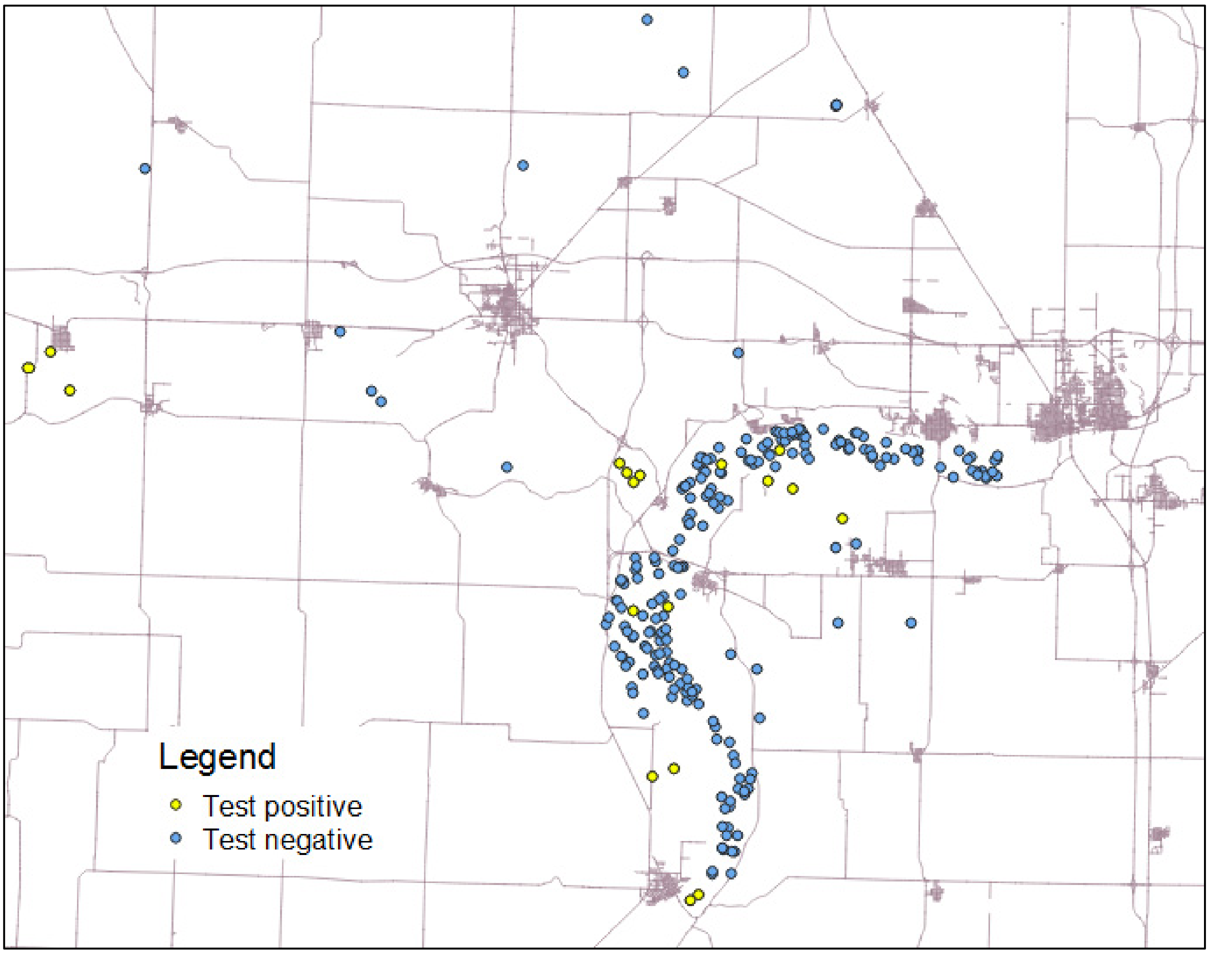
Now suppose we would like to depict the locations of individual animals that tested positive. This can be accomplished in several different ways depending on the message that is desired. In the map below we use different hues to denote the location of test negative and positive animals. Colors were chosen so that the positives (hue = 359° on the "color wheel", value/lightness = 100%, and saturation= 100%) are distinct, but are not significantly higher on the visual hierarchy than the negatives (hue = 141°, value/lightness = 100%, and saturation= 80%). This map would be valuable when the map creator would like the viewer to be able to differentiate between positive and negative animals, but still wants the overall surveillance effort to be the main message of the map without a significant visual fixation on the positive animals.



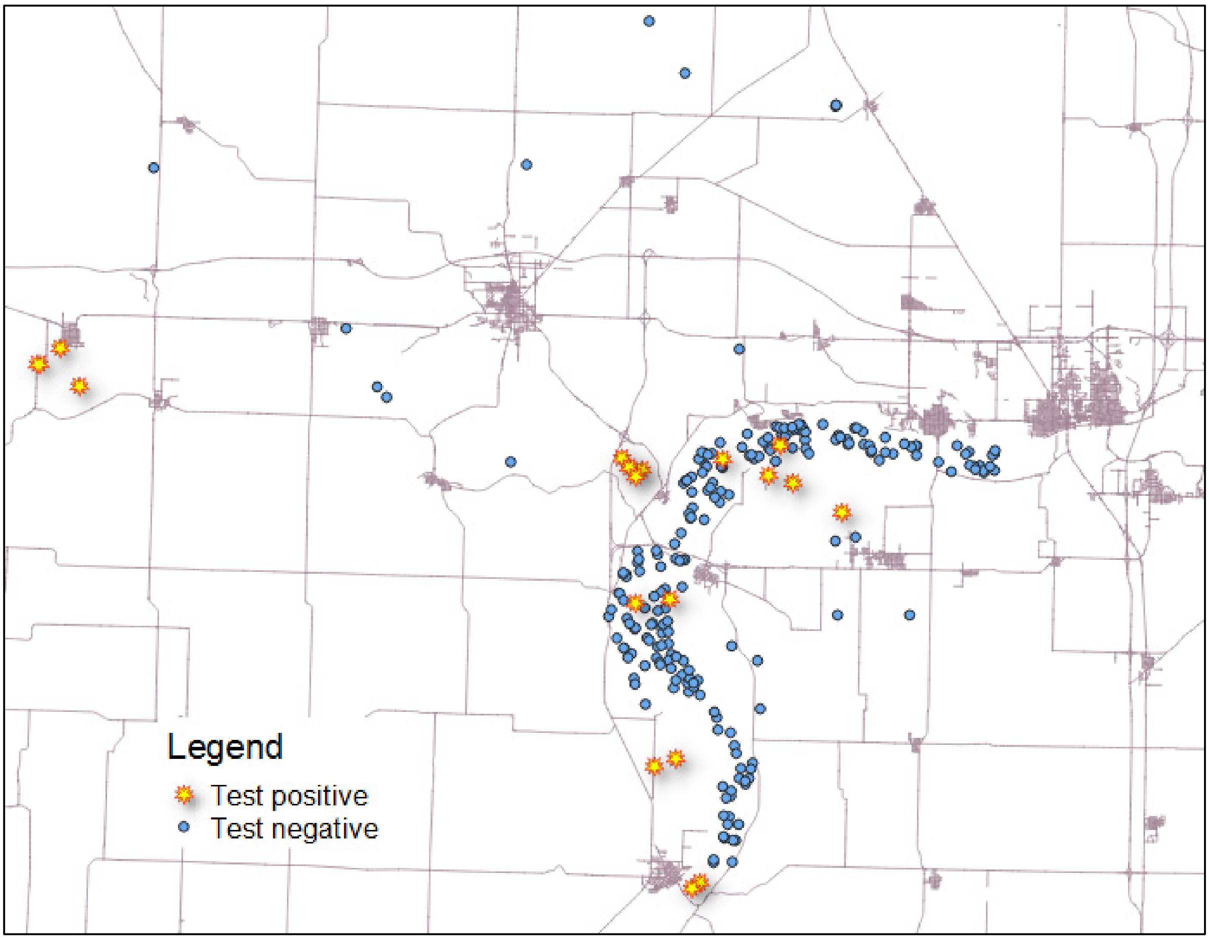
However, if the purpose of the map is to draw attention to the positive cases perhaps for risk communication, then we can change our color selection to bring the location of positive cases to the top of the visual hierarchy. In the map below we achieve this by making the positive case locations black (i.e., hue = 0°, saturation = 0% and value/lightness = 0%), and the negative case locations white (i.e., hue = 0°, saturation = 05 and value/lightness = 100%). This map still depicts the surveillance effort; however, now that is no longer the primary message, and the viewer has to spend a greater amount time studying the map to ascertain this information.



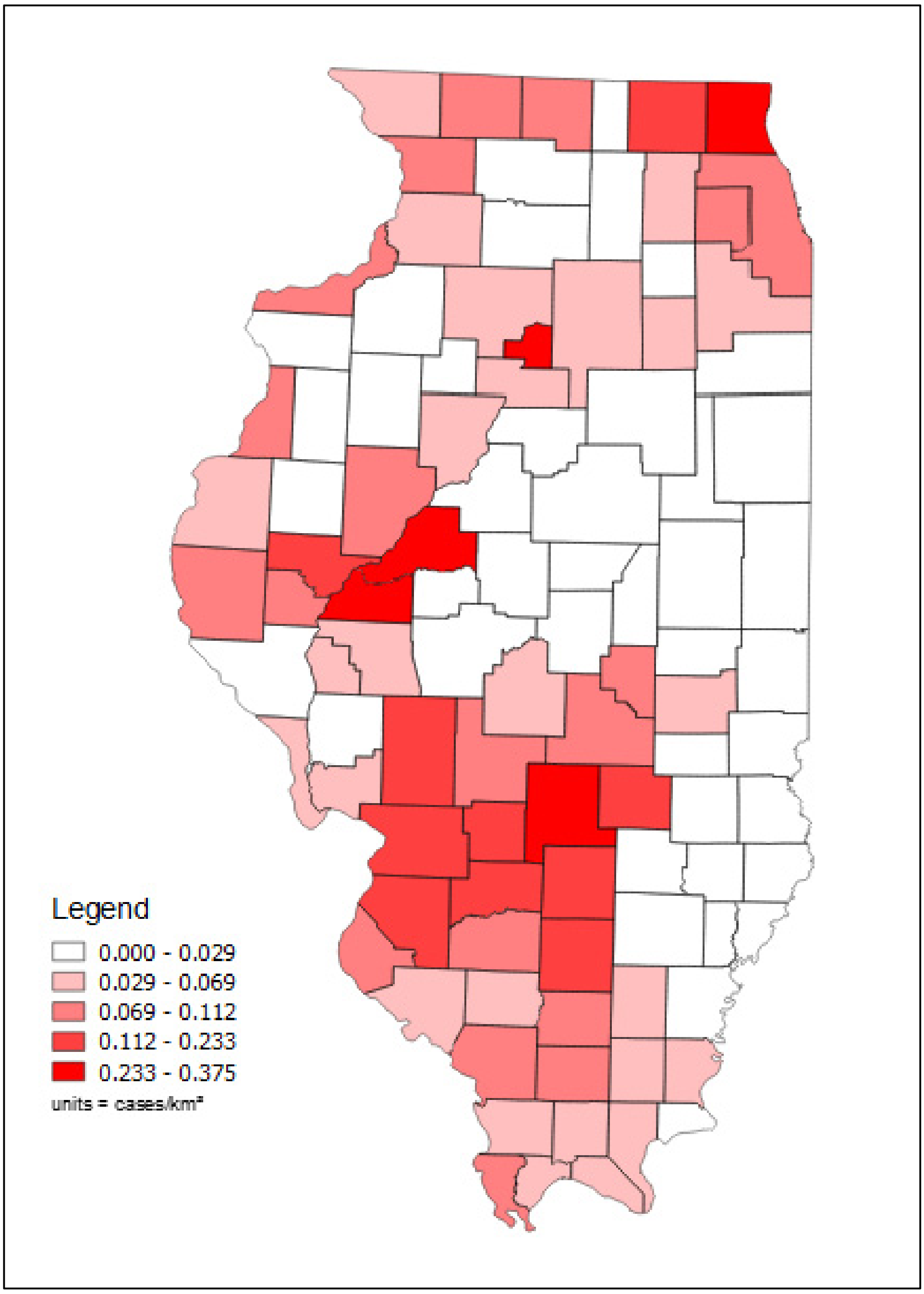
The effect of color can be accentuated by also increasing/decreasing the size of the symbols. Changing symbol size allows for a greater range of colors to be used to differentiate the positive and negative animals, which may increase the aesthetic appeal of the map. The map below uses yellow (i.e., hue = 62°, saturation = 100% and value/lightness = 100%) and blue (i.e., hue = 211°, saturation = 56% and value/lightness = 93%) to depict the locations. The message of the map is still the same as the previous map, and the elements maintain the same visual hierarchy, but this map is more engaging.



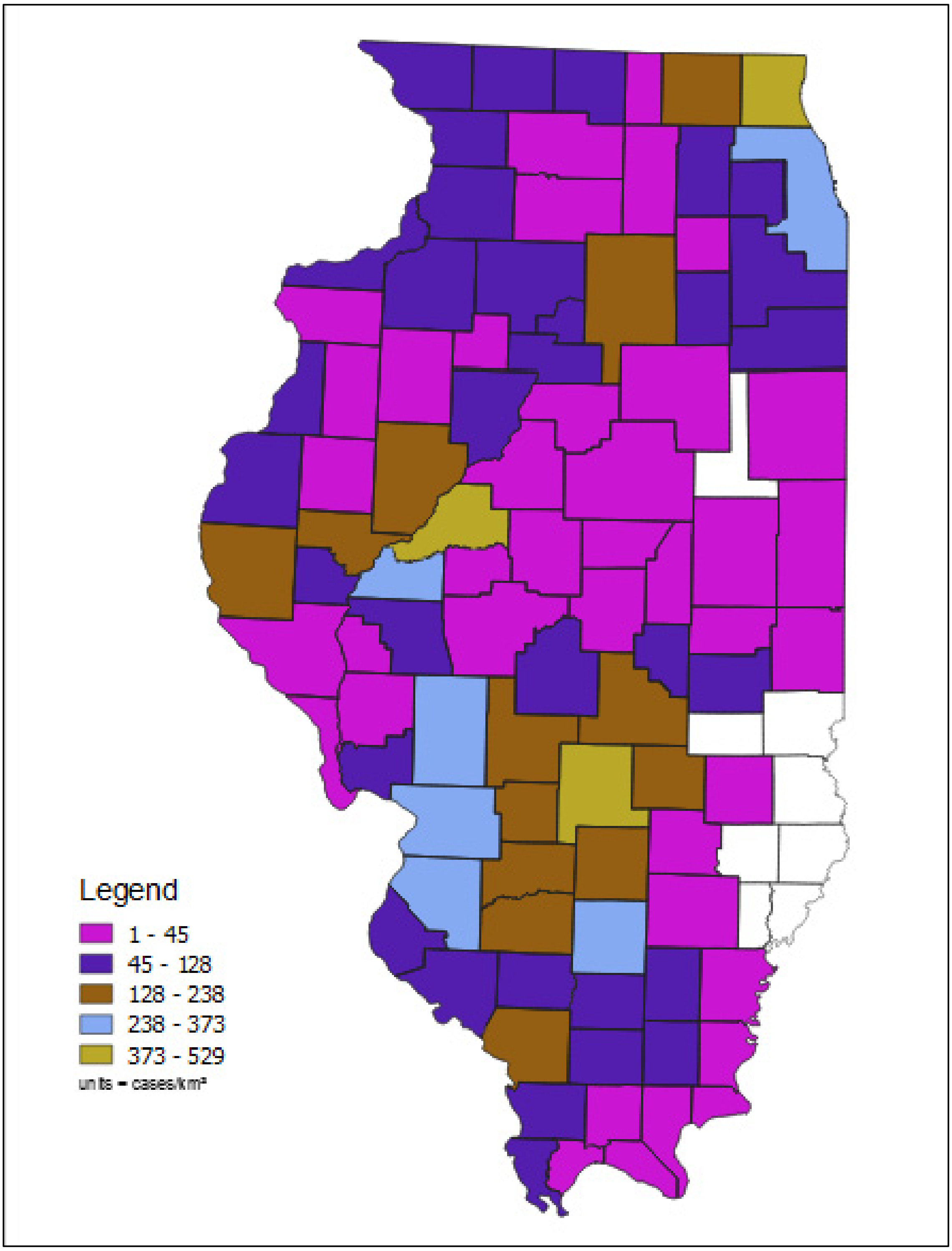
Symbol type can also be changed along with color and symbol size to accentuate the positive cases. The following map provides the same message as the previous two maps, but makes the differences between test positive and negative animals even more dramatic. Thus, it is clear that color choice is a critical decision when creating a map and attempting to ensure that its message is clear and as intended.



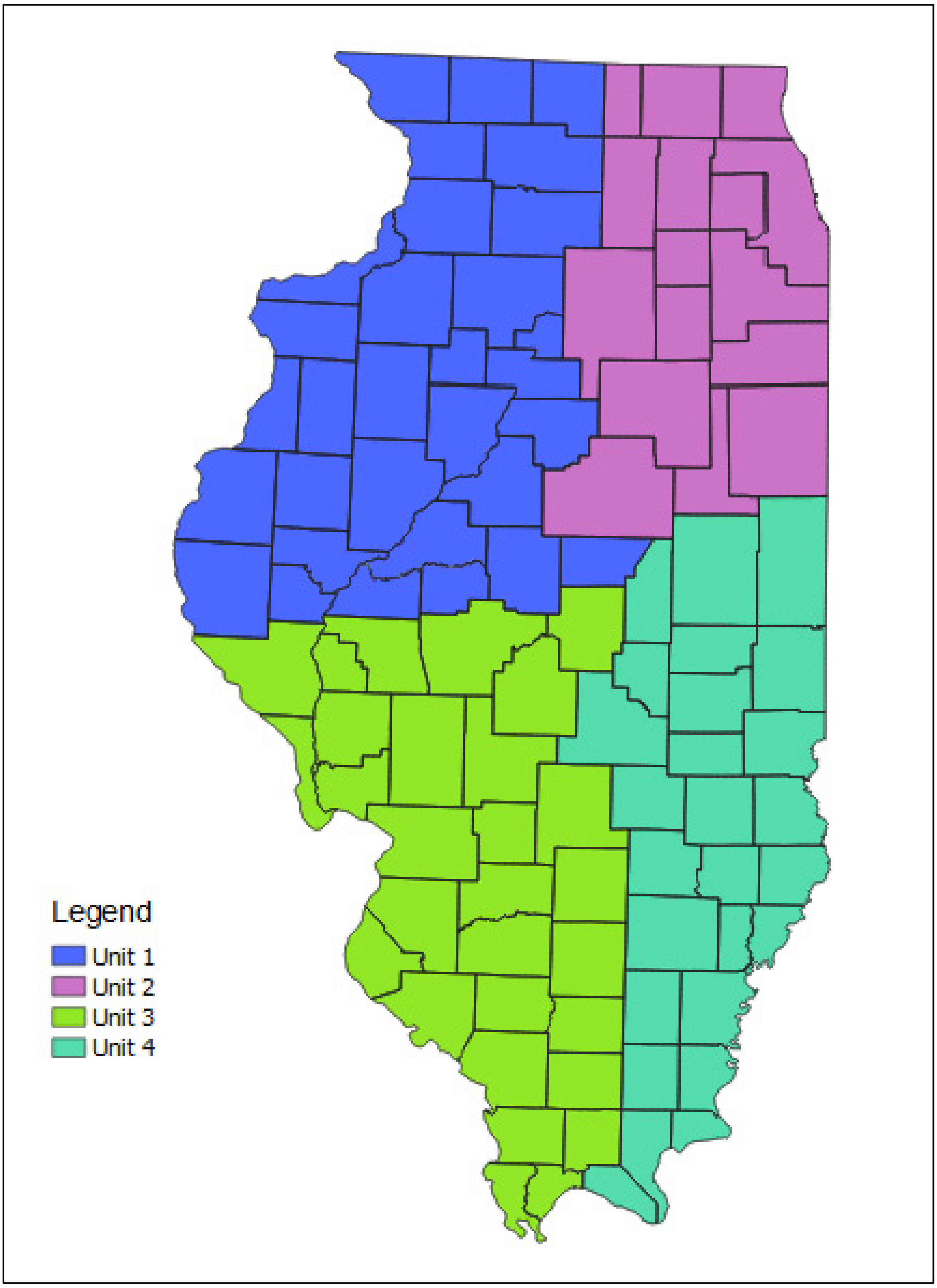
We have already seen the use of color in our discussion of statistical map display. In particular, choropleth maps utilize color to display intensities/rates via a color ramp that assigns the color choices to each class. Care must be taken when selecting a color ramp so the data will be displayed logically. In general, when dealing with data that has a natural order (e.g., increasing diseases case rates), it is recommended to choose a hue and then change the lightness to create different colors for each class, which is called a sequential color scheme. Most often when this scheme is employed the darker colors represent the larger values. We implemented this type of color ramp earlier where selected red as our hue, and then create 5 different colors by changing the lightness with the dark representing the counties with the highest number of cases/km^2. QGIS provides a number of built-in color ramps that are visually appealing, and do not require manual construction.



Although GIS programs provide the ability to use any color in color ramps for ordered data, it is highly recommended to avoid the temptation to try and create a visually stunning map by using many hues without any logical ordering, which is also called a qualitative color scheme. These color ramps will confuse the viewer and often obfuscate the message of the map. The map below is such a map where we replaced the color ramp used in the map above, and depict the class differences via varying hues. Although perhaps artistically interesting, the naturally ordering of the data is lost and the map's message is lost.

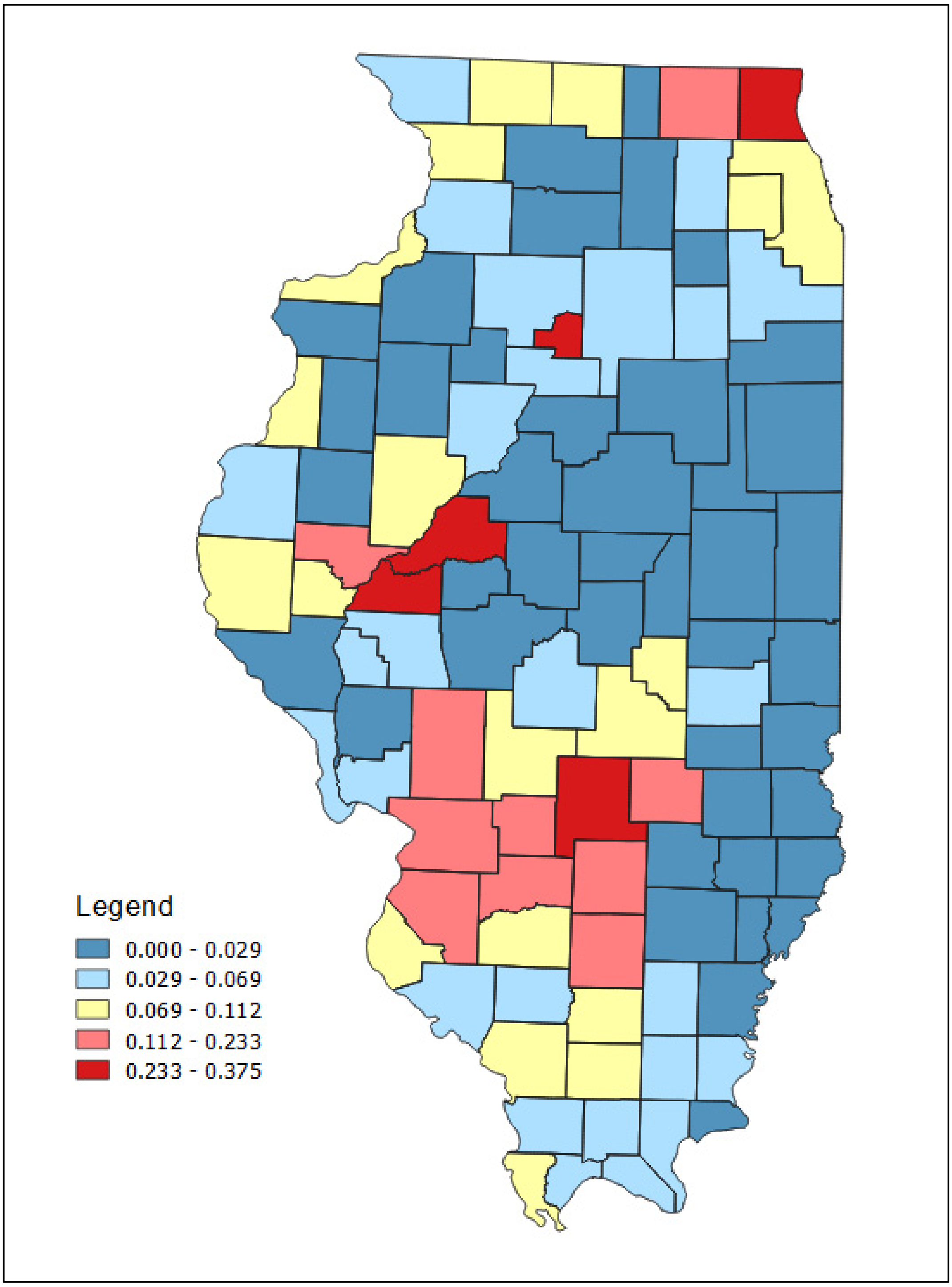


Although the qualitative color scheme used above is inappropriate for data with a natural ordering, it can be quite useful if the map is depicting unordered categorical data (e.g., land cover). In the map below for each county, we display the wildlife health agency's administrative unit to which it belongs. This information may be helpful for coordinating disease response efforts.



The qualitative color scheme works well for this information because although there is a number associated with each administrative unit there is no logical ordering associated with that number. Thus, various hues can be used together, and the map still clearly conveys the necessary message.

An additional color scheme that is useful is a diverging scheme. This scheme is primarily employed to show changes. It leverages two colors on each end of the class values (e.g., blue = lowest class and red = highest class), and then the lightness/value of these colors is reduced as they approach the midrange. The middle or no change class is represented with the lightest color. The map below shows the use of a diverging scheme depicting the cases rates for the counties.



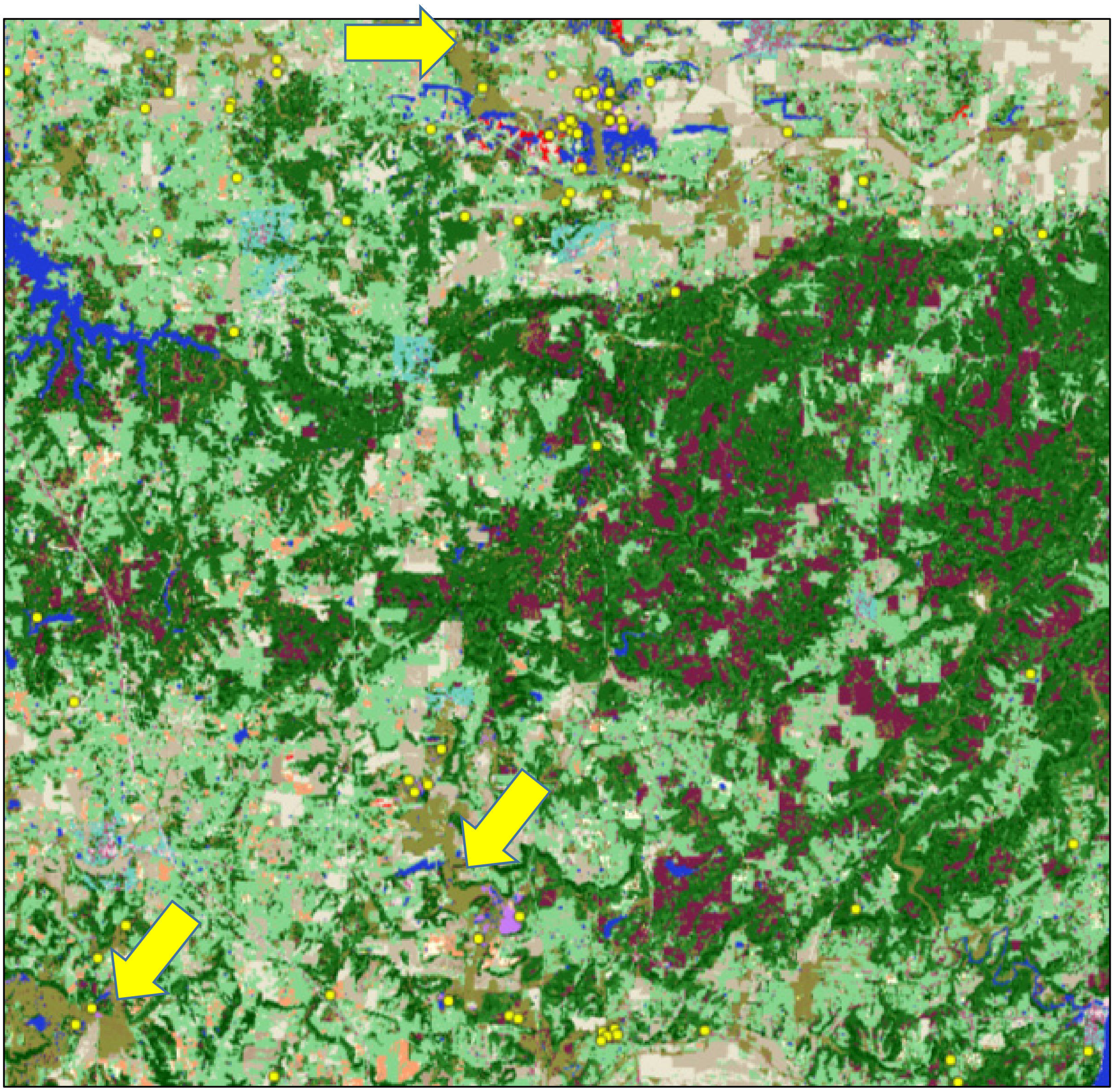
Although the above map uses two different hues, it does not result in confusion for the map viewer like the qualitative scheme because there is logical structure to the color ramp. The lightest color represents the class containing the average case rate. The blues represent counties with below average case rates, and the reds are counties with above-average case rates. Thus, the viewer can quickly scan the map and identify counties that are above or below average, and those that are at the extremes are high in the visual hierarchy (i.e., darker reds and blues). Diverging schemes are useful tools when the purpose of the map is to communicate areas that are different than "normal" or "average", or have changed in some respect.

For all the color considerations we have discussed, it is also important to understand how the color perceived by the viewer will be influenced by the surrounding colors on the map. This phenomenon is known as simultaneous contrast, and is defined as the visual effect that causes the contrast between juxtaposed colors to be exaggerated. A simple example of simultaneous contrast is shown below. Which green rectangles are the same color?

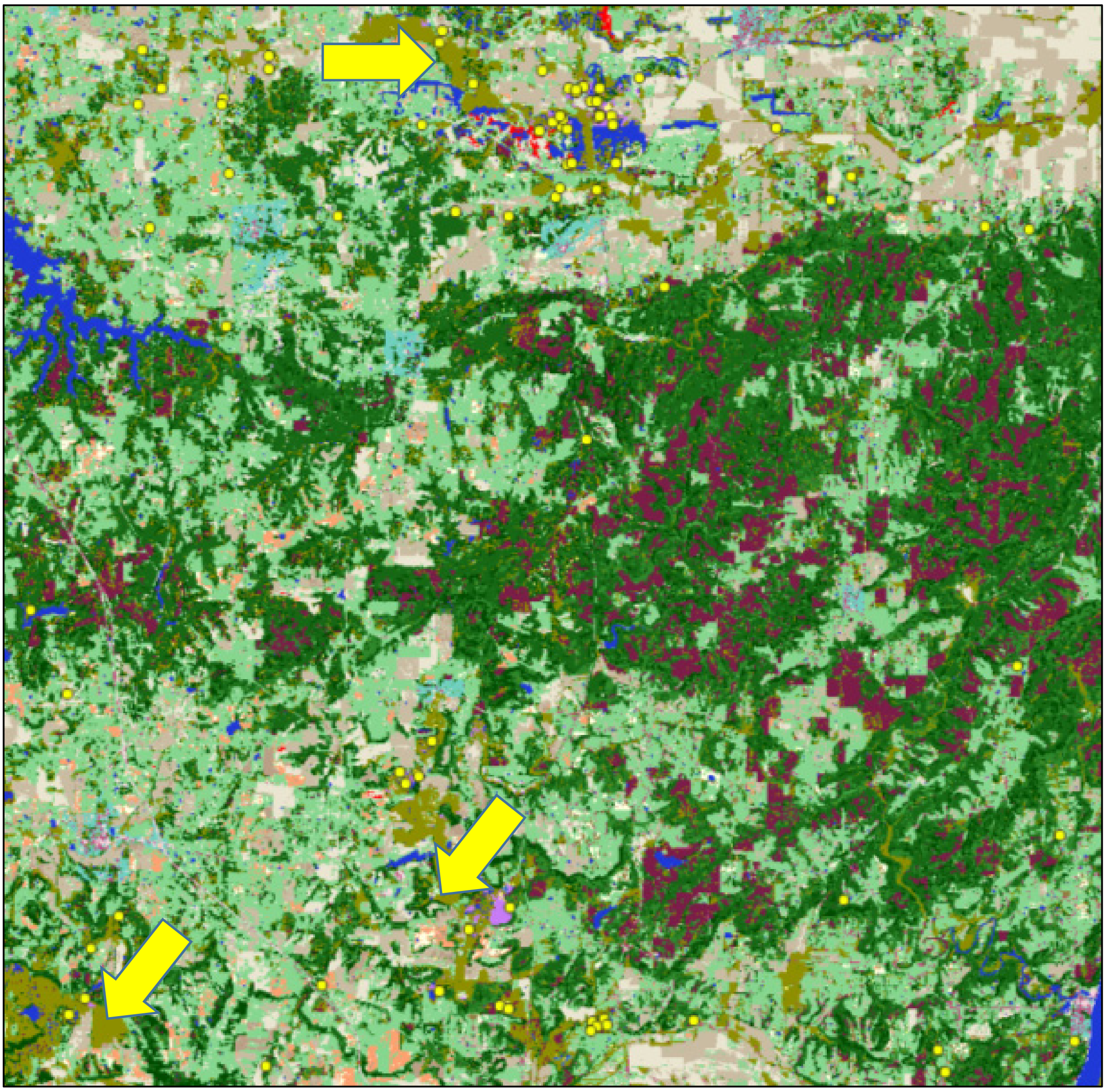
The color of the green rectangles in the first group of blocks are actually the same, but for most readers they appear different (right being more yellow) based solely on the background in which they appear. The green rectangles in the second set of block are actually different, but for most readers they appear the same. Below the rectangle are pictured side by side to demonstrate their differences or lack thereof.

Simultaneous contrast can impact the messaging of a map if the viewers cannot distinguish actual differences, perceive differences in color that do not exist, or if it unintentionally alters the visual hierarchy of the map. This can lead the viewer to assign the wrong class to a region, or the message of the map being muddled. This problem is most problematic when there are regions that are completely surrounded by a contrasting color. Totally avoiding the effects of simultaneous contrast may be impossible because the spatial locations of the data dictate color placement on a map. But, experimenting with a variety of colors on the map will help to identify areas where this phenomenon may be occurring, and can be avoided with a different color choice.

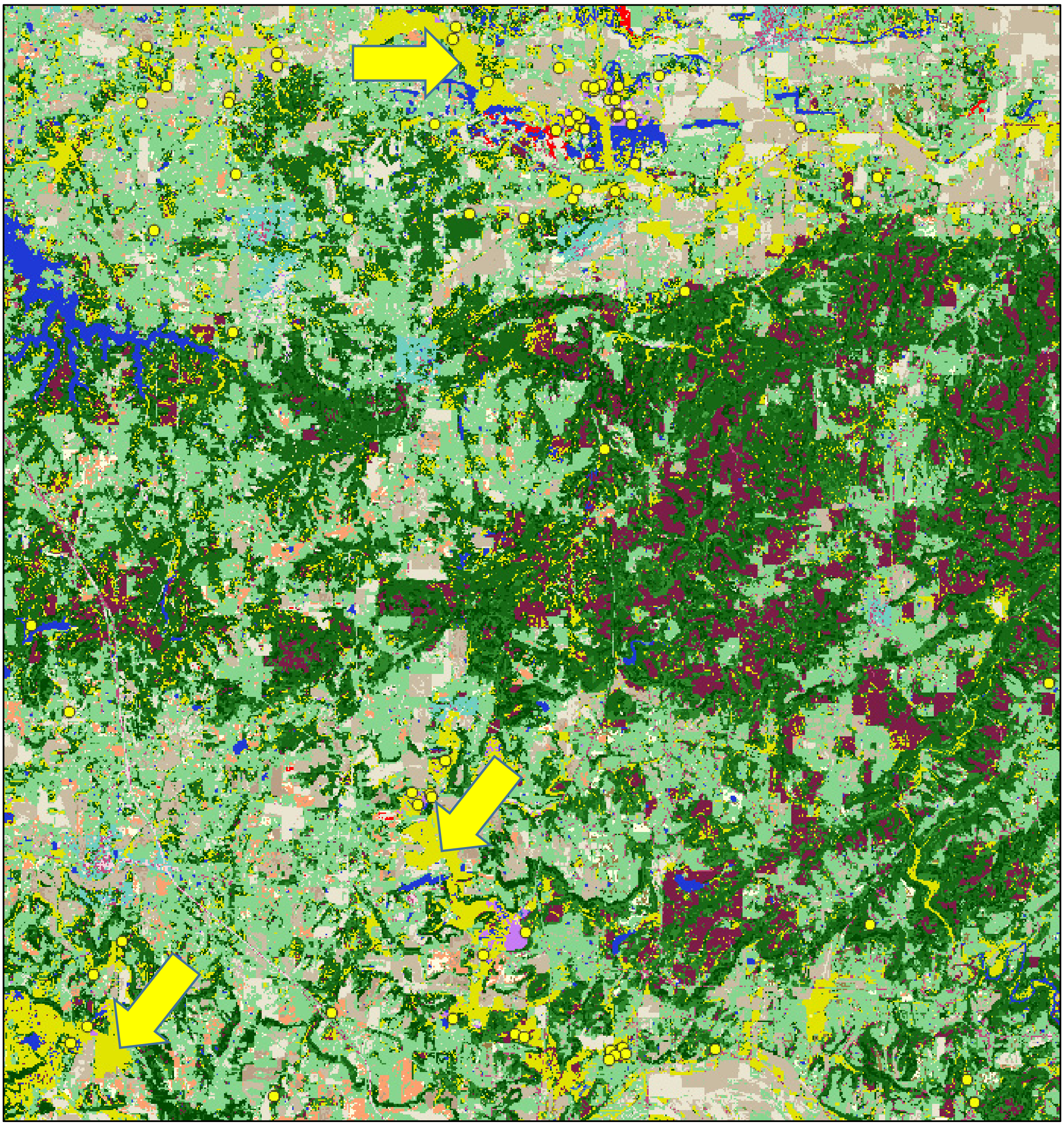
Color choice can also be critical when there are important areas on the map, but they are small relative to the map size. These areas can often times become visually lost, and be overwhelmed by the features that dominate the map. In the following map, we show locations of diseases cases in southern Illinois as yellow points. We also have include a land cover layer based on information sourced from: Illinois Natural History Survey's 1999-2000 1:100 000 Scale Illinois Gap Analysis Land Cover Classification,



Raster Digital Data, Version 2.0, September 2003. The purpose of this map to is to qualitatively evaluate the scientific hypothesis that land cover type may be influencing the distribution of disease cases. Wetland classes are rare, but may be important given the clustering of points near these areas; however, because they are small regions they are hard to distinguish using the current color choice (yellow arrows). To increase the emphasis on this land cover type, we will increase the emphasis by using high-saturation colors. In the following map we use exactly the same hue and lightness, but increase the saturation to 100%. Now, the wetland areas are more readily apparent across the map.



This effect could be further enhanced by also increasing the lightness/value of the color. In the following map, we increase the lightness as well, and now we can clearly see more wetland areas within the dark green, forested regions, that were difficult to distinguish before. Thus, it is good practice if there are small areas of high significance in a map to use high-saturation color value to ensure that they are not lost within all the colors that are depicted. In our example, the wetland areas may be important components of the ecology of the disease we are investigating (e.g., breeding ground for vectors), and being unable to readily visualize these areas in our map may lead to exclusion of relevant hypotheses regarding the impacts of land cover on the etiology of the disease.



We conclude our discussion of color choice by encouraging careful thought about the potential significance a color may have for the intended audience. Often judicious use of color of particular map elements can take advantage of the emotional responses of people to colors. For example, earlier in this training we have made extensive use of the color red when displaying our disease cases. This choice was intentional because in many regions red is associated with warning or fire, and this aligns with the purpose of the maps to communicate disease risk to human, domestic and wild animal populations and emphasize the dire nature of the data. However, peoples' emotional responses to color can vary due to many intrinsic demographic characteristics. Thus, once again knowing the intended audience is essential to ensure the color choices will likely have the desired effect and reinforce the message the maps is trying to communicate.

In summary, to truly take advantage of maps as communication tool for disseminating wildlife health information, it is essential to clearly outline the purpose of the map, the message that it is intended to convey, and the audience it is meant to inform or influence. This information will provide the foundation upon which careful map design choices can be made. The creation of good maps is a process of trial and error, and experimentation to find the best choices is required. It is also helpful to have the map critiqued by others who may be able to determine if the map is conveying the proper message in a manner that is simple to understand by the audience. In the end, a well-constructed map can be one of the most useful scientific tools available for wildlife health to inform and influence: scientists, agency personnel, the general public, and decision makers.

Conclusion

Throughout this section, we have only provided a brief introduction into the concepts and uses of GIS for data dissemination. Our intent in this section was to provide a simple, easy-to-use guide to help interested readers begin using GIS to disseminate and explore their wildlife health information. Cartography is both an art and a science, and we do not expect readers to become experts based solely on this training. However, there are numerous training opportunities available via the internet, text books, and tutorials that can be used to supplement the basic foundation that we have built in this training, and to unleash the power of GIS for wildlife health applications.

We intentionally devoted considerable effort to not only describing how to use available GIS tools, but how also to use them effectively through good design choices. It was also our intent to help educate the readers about the mapping to encourage them to not only be able to use this tool for wildlife health investigations, but also to allow them to more critically evaluate maps that others create. All maps inherently have inaccuarcies resulting from the map projection, the removal of some geographic features for clarity, the choice of aggregation, color choice, etc. An educated map viewer can evaluate if these inaccuracies, either intentional or uintentional, have negatively influenced or falsely portray the information. Given the importance that we place on maps and the fact that they are often considered to be definitive documents, it is essential to view them critically and realize that they are the product of many choices by the map maker, and as such may portray biases, agendas, and beliefs of that person. Additionally, acknowledging the fact that the maps that we create also have the same shortcomings, we can strive to minimize these effects and use maps to effectively disseminate information for the benefit of wildlife health.

Wildlife Disease Outbreak Communication Plan

Introduction

Wildlife disease outbreaks have the potential to cause widespread impacts on public health, agricultural species, and wildlife populations. Emerging infectious diseases (e.g., SARS, avian influenza) can be particularly challenging to manage as the circumstances of disease outbreaks may vary by many factors, including type of agent, scale of exposure, and mode of transmission, etc. Planning and preparing in advance of an infectious disease emergency is critical for an effective response, as events may unfold quickly and unpredictably. Effective communication is an essential element of outbreak management that requires the establishment of a disease outbreak communication plan. Furthermore, because of the multi-sector impact of wildlife diseases, communication and coordination among agencies, stakeholders and sectors during disease outbreaks has become increasingly important. This, and the rapidity by which information can be shared globally requires that competent authorities coordinate and communicate clearly, quickly and credibly about risks to animals, humans, and the environment.

A good plan should be flexible enough to address any type of outbreak and should contain the following information, among others:

- describe and designate roles and responsibilities for different outbreak scenarios;
- designate who is accountable for leading the response;
- designate who is responsible for implementing various actions;
- designate who needs to be consulted during the process;
- designate who needs to be informed about what is taking place;
- designate who will be the lead spokesperson and backup for different scenarios;
- include procedures for public information verification, clearance and approval;
- include procedures for coordinating with important stakeholders and partners (for example, with other government agencies, law enforcement, emergency management staff and elected officials);
- include risk communication messages for the public and stakeholders;
- include a media communication strategy and methods of monitoring media and social media reports, postings, etc.

Developing a Communication Plan

In 2011, the Member Countries of the OIE adopted international standards on the communication of Veterinary Services and Aquatic Animal Health Services. Relevant provisions are included in the section dealing with quality services of the Terrestrial (chapter 3.3, section3), and the Aquatic Animal Health Codes (chapter 3.2, section 3) respectively. Furthermore the OIE has developed, in collaboration with the World Health Organization, a Communication Handbook for Veterinary Services (*www.oie.int/communication_handbook*) to assist Member Countries in strengthening skills and competencies in communications and facilitate networking among sectors. Readers are referred to these and other resources at the end of this chapter that will be useful in developing robust risk communication plans. In addition, *Table* 7 below provides a checklist for an emergency or outbreak communication plan.

TABLE 7: EMERGENCY RISK COMMUNICATION PLAN CHECK LIST*
Signed endorsement from the senior leadership
Designated line and staff responsibilities for the communication teams
Internal information verification and clearance/approval procedures
Agreements on information release authorities (who releases what/when/how)
Media contact lists
Procedures to coordinate with field response teams
Designated spokespersons
Emergency response team after-hours contact numbers
Contact information for other agencies and partners
Agreements/MOUs/procedures to form joint response teams and/or coordinate
Procedures to secure needed resources continuously
Identified methods of information dissemination to all audiences (partners, stakeholders, publics) via email, media, social media, press releases, leaflets and fact sheets, etc.
Adapted from the CDC Crisis Emergency Risk Communications (CERC) checklist:** ***https://emergency.cdc.gov/cerc/resources/templates-tools.asp

Interagency Coordination

As mentioned, during the development of a serious health emergency or disease outbreak, it will be necessary to implement an effective approach to communication and coordination among stakeholders, especially in the public health, veterinary, and environmental/wildlife sectors. This will ensure that all key partners are kept informed of the status of events, and will enable decision-making to occur with all available information. It is important that the communication plan outline the approach to communication among these stakeholders and related agencies, as appropriate. Regular, frequent communication at different phases/stages/levels will be needed, and it will be necessary to also define how ad-hoc messages and coordination may occur at different points. Effective communications and coordination will enable these agencies to work together to anticipate, characterize, contain, and mitigate disease outbreaks as quickly as possible.

The communication plan should clearly describe how communications and coordination among public health, veterinary, and environmental agencies, and other partner agencies, will be handled in the case of an emergency or disease outbreak. It should also identify communications and coordination responsibilities for the various agencies and partners that will be involved in the response. The plan should:

- identify how a coordinated approach to response will be maintained across various agencies;
- establish efficient mechanisms for decision-making during an emergency or disease outbreak that will support coordinated approaches, but allow for sufficient individual agency flexibility; and
- describe how partners will be notified of important events and how they will be kept informed of the current status and related response efforts.

Conclusion

Communication is the foundation upon which a "One-Health" approach to managing emerging infectious diseases rests. The contribution of wildlife health data is one component of the overall information exchange that is required to effectively combat disease outbreaks; however, sharing wildlife health information requires a significant and continued commitment of resources by responsible parties. To ensure that this commitment is justified, establishing detailed communication plans is critical because these plans provide the basis for efficiently communicating the necessary information to affect the crucial decisions surrounding a disease outbreak. Thus, a well-developed communication plan is the framework supporting "One-Health" endeavors.

Resources

OIE Communication Handbook for Veterinary Services (*www.oie.int/communication_handbook*)

CDC Crisis and Emergency Communication Website (*https://emergency.cdc.gov/cerc/index.asp*)

Minnesota Department of Health Public Health and Quality Improvement Resources & Tools (*http://www.health.state.mn.us/divs/opi/qi/toolbox/swimlane.html*)

WORLD ORGANISATION FOR ANIMAL HEALTH • *Protecting animals, preserving our future*
12 rue de Prony 75017 Paris France • tel.: 33(0)1 44 15 18 88 • fax: 33(0)1 42 67 09 87 • www.oie.int • oie@oie.int