TIEE

Teaching Issues and Experiments in Ecology - Volume 13, March 2018

ISSUES : DATA SET

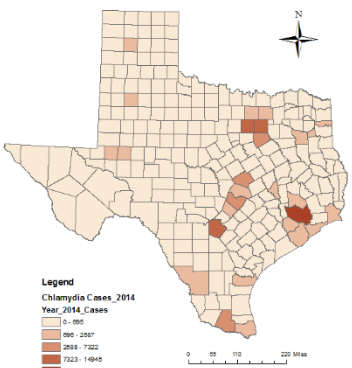
Investigating Sexually Transmitted Disease (STD) Ecologies Using Geographic Information Systems (GIS)

Maruthi Sridhar Balaji Bhaskar [1,*], Jason A. Rosenzweig [2] and Shishir Shishodia[2]

[1]Department of Environmental and Interdisciplinary Sciences, Texas Southern University, 3100, Cleburne Street, Houston, Texas, 77004

[2]Department of Biology, Texas Southern University, 3100, Cleburne Street, Houston, Texas, 77004

*Corresponding author: Maruthi Sridhar Balaji Bhaskar (bhaskarm@TSU.EDU)



Geographic Epidemiology of Chlamydia in Texas

THE ECOLOGICAL QUESTION:

What are the magnitudes and spatial prevalences of Sexually Transmitted Diseases (STDs), such as chlamydia and gonorrhea in Texas? How do the social and economic factors explain the spatial and temporal trends of STD prevalence in Texas?

ECOLOGICAL CONTENT:

Disease ecology, human ecology and behavior, bacterial infections, spatial and temporal spread of disease, Geographic Information Systems (GIS)

WHAT STUDENTS DO:

The students explore and analyze data on two Sexually Transmitted Diseases (STD) - chlamydia and gonorrhea - at the county level. The students download the spatial and temporal datasets of the diseases for the period of over 15 years, from 2000 to 2014, at the county level for the entire state of Texas. The students then process the datasets and import them into Geographic Information Systems (GIS) using ArcGIS software. Spatial and temporal maps of the disease distribution over the entire geographical region are developed for each year. The students use the datasets to develop disease maps for chlamydia and gonorrhea occurrence in Texas for the time period of 2000 to 2014. All the maps developed by the students are put together and discussed at the end of the project.

STUDENT-ACTIVE APPROACHES:

Guided-inquiry or open-ended inquiry, cooperative learning

SKILLS:

The students become adept in the use of Geographic Information Systems (GIS) to analyze and explore health science and biological problems. They analyze the long-term datasets using Microsoft Excel and learn ArcGIS Desktop (Version 10.5) software, including importing tabular data into GIS and create effective maps.

The students explore available disease data and relevant statistics, identify the appropriate and reliable datasets on the websites, download data into excel files, and analyze the data by making graphs. Further, students download the datasets into the GIS, link the data to the spatial datasets, prepare a spatial database of the disease, and download the base layers for making the GIS maps. Students then will interpret, manipulate and analyze the GIS data. Finally, students improve their communication skills through PowerPoint or infographic presentations and write a research report.

Through this module, students learn critical thinking skills and engage in empirical and quantitative learning.

ASSESSABLE OUTCOMES:

- the tables, graphs, GIS maps, and research reports to assess the critical thinking and visual interpretation skills of the students
- written assignments with graphs and GIS maps to assess the empirical and quantitative skills of the students
- class presentations to assess students' understanding, synthesizing capabilities, and communication of scientific data

SOURCE:

Atlas Plus CDC (Center for Disease Control and Prevention) Data:
https://gis.cdc.gov/GRASP/NCHHSTPAtlas/main.html

GIS Data – HGAC (Houston Galveston Area Council):

http://www.h-gac.com/rds/gis-data/gis-datasets.aspx

ACKNOWLEDGEMENTS:

This work was supported by the National Science Foundation HBCU-UP Targeted Infusion (HRD 1622993) and Research Initiation (HRD-1400962) grants. We would also like to thank the Quantitative Undergraduate Biological Education and Synthesis (QUBES) DIG Data Faculty Mentoring Network facilitators and participants for advice, support, and valuable feedback.

OVERVIEW OF THE ECOLOGICAL BACKGROUND

Natural and human-induced disruptions, such as habitat loss, introduction of invasive species, pollution, climate change, and disease spread, are all threats to wildlife health and biodiversity. GIS is an effective tool for managing, analyzing, and visualizing biotic and abiotic data to target areas where interventional management practices are needed and to monitor their effectiveness. In wildlife ecology, GIS can be used to map the habitat requirements, population patches and linkages, and disease levels within populations and the effect of land use and land cover changes on wildlife densities.

GIS facilitates the incorporation of spatial relationships into epidemiological investigations of wildlife diseases. Investigating and mapping diseases in wild animals is a special challenge as the locations of individual animals within the populations at risk are much less predictable than in the case of domestic animals. Several studies have shown the wider applications of GIS for disease surveillance in wildlife, such as estimating the raccoon abundance to administer rabies vaccination (Beasley et al. 2012), estimation of chronic wasting disease in mule deer population (Conner and Miller 2004), tick borne diseases in coyotes (Foley et al. 2005), outbreak of leptospirosis in sea lions (Norman et al. 2008), viral disease incidence in squirrels (Rushton et al. 2000), and West Nile virus incidence in birds (Shuai et al. 2006).

In addition to studying wildlife diseases, GIS can be used to understand the spatial dynamics of human diseases. The number of sexually-transmitted disease (STD) infections in the United States has increased at an alarming rate over the past few decades. According to recent statistics released by the Centers of Disease Control and Prevention (CDC), an estimated 19 million new STD infections occur each year in this country. Moreover, the estimated direct medical cost associated with STD infection is $15.5 billion in one year (Chesson et al. 2004). Chlamydia, a bacterial infection caused by *Chlamydia trachomatis*, is the most prevalent STD infection reported in the United States and this disease is widely spread among sexually active women. *Chlamydia trachomatis* infection is the most commonly reported nationally notifiable disease in the United States, with more than 1,526,658 cases reported to state and local health departments in 2015 (CDC 2015). Chlamydia infections are usually asymptomatic, and this disease can cause pelvic inflammatory disease (PID) if left untreated. PID is an infection affecting the uterus, fallopian tubes, and other reproductive organs, which often results in ectopic pregnancy (CDC 2014).

The second most commonly reported STD is gonorrhea, caused by *Neisseria gonorrhoeae*, with more than 395,216 cases reported in 2015 (CDC 2015). In 2005, the infection rate in the United States from gonorrhea was reported to be 115.6/100,000 (CDC 2006; Raychowdhury et al. 2008). Like chlamydia, gonorrhea is also considered a major cause of PID. According to the CDC, gonorrhea infection is most prevalent among adolescents and young adults.

STDs have long been known for their great impact on health. The most common risk factors for spread of STDs are having more than one sex partner, having a new partner, use of hormonal contraception, not using barrier contraception, and having gonorrhea, a previous chlamydial infection, or cervical ectopy (CDC 2015).

Infectious disease surveillance is a core public health function. Geographic information systems (GIS) have added an exciting new dimension to the epidemiology of infectious diseases. GIS for disease incidence is particularly attractive for surveillance of acute infectious diseases that are related to point source outbreak events. Because of this characteristic, GIS has been applied to, and has greatly enhanced understanding of epidemiology of STDs (Zenilman et al. 2002).

Maps of the geographical distribution of an infectious disease can be used for a variety of reasons. Mapping is a primary goal in spatial epidemiology. Maps of disease distribution and intensity allow an immediate visualization of the status of public health. When based on empirical evidence, maps can support carefully weighted assessments by decision makers on the advantages and disadvantages of alternative courses of action. These may range from helping to plan national or regional scale intervention strategies to offering advice for individuals. These maps can also document a baseline from which intervention success or failure can be monitored.

LITERATURE CITED

Beasley, J.C., W. S. Beatty, T.C. Atwood, S. R. Johnson, and O. E. Rhodes. 2012. A comparison of methods for estimating raccoon abundance: Implications for disease vaccination programs. The Journal of Wildlife Management 76:1290–1297.

CDC. 2006. Sexually Transmitted Disease Surveillance 2005, Centers for Disease Control and Prevention. Atlanta: U.S. Department of Health and Human Services.

CDC. 2014. Division of STD Prevention of the Centers for Disease Control and Prevention. Sexually transmitted disease surveillance 2014. Atlanta: U.S. Department of Health and Human Services.

CDC. 2015. Division of STD Prevention of the Centers for Disease Control and Prevention. Sexually transmitted disease surveillance 2015. Atlanta: U.S.

Department of Health and Human Services https://www.cdc.gov/std/stats15/STD-Surveillance-2015-print.pdf

Chesson, H.W., J.M. Blandford, T.L. Gift, G. Tao, and K.L. Irwin. 2004. The estimated direct medical cost of STDs among American youth 2000. Perspectives on Sexual and Reproductive Health 36:11-19.

Conner, M.M. and M.W. Miller. 2004. Movement patterns and spatial epidemiology of a prion disease in mule deer population units. Ecological Applications 14(6): 1870–1881.

Foley, J.E., E.V. Queen, B. Sacks, and P. Foley. 2005. GIS-facilitated spatial epidemiology of tick-borne diseases in coyotes (*Canis latrans*) in northern and coastal California. Comparative Immunology, Microbiology and Infectious Diseases 28(3): 197-212.

Norman, S.A., R.F. DiGiacomo, F.M.D. Gulland, J. S. Meschke, and M.S. Lowry. 2008. Risk factors for an outbreak of leptospirosis in California sea lions (*Zalophus californianus*) in California. Journal of Wildlife Diseases 44 (4): 837-844.

Raychowdhury, S., S.H. Tedders, and S.K. Jones. 2008. Impact of chlamydia and gonorrhea in Georgia: An urban/rural comparison (2000-2004). Journal of Georgia Public Health Association 1:1-9.

Rushton, S.P., P.W.W. Lurz, J Gurnell, and R. Fuller. 2000. Modeling the spatial dynamics of parapoxvirus disease in red and grey squirrels: a possible cause of the decline in the red squirrel in the UK? Journal of Applied Ecology 37: 997–1012.

Shuai, J., P. Buck, P. Sockett, J. Aramini, and F. Pollari. 2006. A GIS-driven integrated real-time surveillance pilot system for national West Nile virus dead bird surveillance in Canada. International Journal of Health Geographics 5 (1): 17.

Zenilman, J.M., G. Glass, T. Shields, P.R. Jenkins, J.C. Gaydos, and K.T. McKee, Jr. 2002. Geographic epidemiology of gonorrhea and chlamydia on a large military installation: Application of a GIS system. Sexually Transmitted Infections 78: 40-44.

DATA SETS

1. Part 1: Downloading, analyzing, building and Interpreting the STD datasets

Part1_Data_STD_Texas.xlsx

2. Part 2: Developing the geospatial database and mapping the disease incidence

Part2_Chlamydia_TX_Cases_2014.csv
Part2_Chlamydia_TX_Rate_2014.csv
Part2_Gonorrhea_TX_Cases_2014.csv
Part2_Gonorrhea_TX_Rate_2014.csv

3. Part 3: Examining the correlation between the different social factors and the disease incidence

Part3_SocialandOtherfactors_TX_2014.xlsx

The spatial data showing the county-level geographical boundaries of Texas are downloaded from the Houston Galveston Area Council (HGAC) GIS datasets at http://www.h-gac.com/rds/gis-data/gis-datasets.aspx. These data are publicly available through the internet. The HGAC-GIS datasets are the data repository that provides the user with access to wide variety of spatial data for Texas.

The disease incidence data showing the chlamydia and gonorrhea incidence at the county-level for the entire United States is available from the Atlas Plus website of the CDC at https://gis.cdc.gov/GRASP/NCHHSTPAtlas/main.html. These data are publicly available through the internet. The Atlas Plus datasets are the data repository that provides the user with access to the data on HIV, Hepatitis, STD, and TB disease for multiple years at the county-level for the entire nation.

STUDENT INSTRUCTIONS

Background

Ecology is the study of the relationship of organisms in an environment. Disease ecology focuses on the interactions between pathogens and their hosts. In wildlife ecology, Geographical Information Systems (GIS) can be used to map the habitat requirements, population patches and linkages, and disease levels within populations and the effect of land use and land cover changes on wildlife densities.

GIS facilitates the incorporation of spatial relationships into epidemiological investigations of wildlife diseases. Investigating and mapping diseases in wild animals is a special challenge as the locations of individual animals within the populations at risk are much less predictable than in the case of domestic animals. Several studies have shown the wider applications of GIS for disease surveillance in wildlife, such as estimating the raccoon abundance to administer rabies vaccination (Beasley et al. 2012), estimation of chronic wasting disease in mule deer population (Conner and

Miller 2004), tick borne diseases in coyotes (Foley et al. 2005), outbreak of leptospirosis in sea lions (Norman et al. 2008), viral disease incidence in squirrels (Rushton et al. 2000), and West Nile virus incidence in birds (Shuai et al. 2006).

Human societies co-inhabit their environment with many other living organisms, including those that produce the disease. Diseases are most often viewed as serious threats to human and animal health (Brown and Inhorn 1996). An ecological approach to human health and illness emphasizes the fact that the environment and its health risks are, to an extent, created by the various biotic and abiotic interactions (Brown and Inhorn 1996, CDC 2015).

Disease is an inevitable part of life, and coping with disease is a universal aspect of the human experience. During the course of life, all humans harbor infections by disease organisms and suffer the consequences of those infections. The distribution of disease in a population is neither constant nor random. Both the type and severity of diseases that characteristically afflict members of a population vary significantly among societies as a result of differences in culture, ecological setting, and historical period. More important, within a single society, there may be striking variations in terms of the nature and severity of diseases that afflict individuals of different ages, sexes, social classes, and ethnic groups (Brown and Inhorn 1996, CDC 2015). Understanding these epidemiological patterns presents a challenge, because disease distributions are often influenced by various social, environmental and ecological factors.

There are few infectious diseases that are as severe as the sexually-transmitted diseases (STDs). STDs are defined as diseases in which sexual behavior plays an epidemiologically significant role in the transmission of the causative pathogen (CDC 2006). Changes in sexual behaviors have been one of the primary engines driving changing patterns of STDs. As a result, changes in human ecology and behavior related to socioeconomic, demographic, geographic, political, and epidemiologic factors might influence disease patterns.

Chlamydia continues to be the most commonly reported notifiable disease in the US, with 1,526,658 cases reported in 2015 and with increasing rates of reported cases over each of the last two years (CDC 2015). Trends in rates of reported cases of chlamydia are influenced by changes in the incidence of infection, as well as changes in diagnostic, screening, and reporting practices. The national level chlamydia rate is 478.8 cases per 100,000 people (CDC 2015). In 2015, rates of reported cases of chlamydia were highest in the South (520.5 cases per 100,000 people, 7% increase from 2014), followed by the Midwest, West, and Northeast (CDC 2015). Reported case rates and prevalence of chlamydia among individuals of African-American descent continue to be substantially higher than among all other racial and ethnic groups.

Gonorrhea is the second most commonly reported disease in the United States. In 2015, a total of 395,216 cases of gonorrhea were reported in the United States, yielding a rate of 123.9 gonorrhea cases per 100,000 populations (CDC 2015). In 2015, as in previous years, the South had the highest rate of reported gonorrhea cases (146.3 cases per 100,000 people) among the four regions of the United States, followed by the West, the Midwest, and the Northeast (CDC 2015). This increase was largely attributable to an increased rate among men. High gonorrhea rates persist in certain geographic areas, among adolescents and young adults, and in some racial/ethnic groups.

Chlamydia are non-motile, obligatory intracellular organisms. Three species may cause human disease, they are *Chlamydia trachomatis*, *Chlamydia pneumoniae* and *Chlamydia psittaci* (Creatsas and Deligeoroglou 2012). The risk factors for chlamydial infection include having unprotected sex among sexually active young people, new or multiple sex partners, and a sex partner with history of chlamydial infection (CDC 2015). Symptoms of chlamydial infection are usually in the form of burning sensation when urinating, abnormal vaginal discharge in women, pain and swollen testicles and abnormal discharge from penis in men (CDC 2015). *Neisseria gonorrhoeae* is a Gram-negative coccobacillus that invades columnar and transitional epithelial cells, becoming intracellular. Infection with *Neisseria gonorrhoeae* is frequently asymptomatic, while risk factors for gonococcal infection include: age under 25 years, past history of gonococcal infection, co-infection with other STDs, new or multiple sexual partners, lack of barrier protection and drug use (Marrazzo and Martin 2007, Creatsas and Deligeoroglou 2012). Symptoms of gonorrhea may present as vaginitis or cervicitis. Cervicitis is usually in the form of profuse odorless, white-to-yellow vaginal discharge, with no signs of local irritation (Creatsas and Deligeoroglou 2012). In the absence of modern interventions, such as antibiotics and latex condoms, the STDs such as chlamydia, gonorrhea, and syphilis can cause very high rates of infertility (Brunham et al. 1993; Cassels and Singer 2010; Caldwell and Caldwell 1983; Bauch and McElreath 2016).

Chlamydia, gonorrhea, and syphilis are curable with antibiotics. Widespread access to screening and treatment would reduce their spread. Most STD cases continue to go undiagnosed and untreated, putting individuals at risk for severe and often irreversible health consequences, including infertility, chronic pain and increased risk for HIV. STDs also impose a substantial economic burden. CDC estimates that the cost of STD cases to the U.S. healthcare system is about $16 billion each year (CDC, 2015).

Literature Cited:

Bauch, C.T. and R. McElreath. 2016. Disease dynamics and costly punishment can foster socially imposed monogamy. Nature Communications 7:11219.

Beasley, J.C., W.S. Beatty, T.C. Atwood, S.R. Johnson, and O.E. Rhodes. 2012. A comparison of methods for estimating raccoon abundance: Implications for disease vaccination programs. The Journal of Wildlife Management 76:1290–1297.

Brown, P.J. and M.C. Inhorn. 1996. Health issues in human sections. Pp: 187-214. In Medical Anthropology: Contemporary theory and method, Editors: T.M. Johnson and C.E. Sargent. Anthropology of infectious disease: International Health Perspectives. ICC Publication, Book 4, Praeger Press, Pp: 584.

Brunham, R.C., M. Cheang, J. McMaster, G. Garnett, and R. Anderson. 1993. Chlamydia trachomatis, infertility, and population growth in sub-Saharan Africa. Sexually Transmitted Diseases 20:168–173.

Caldwell, J.C. and P. Caldwell. 1983. The demographic evidence for the incidence and cause of abnormally low fertility in tropical Africa. World Health Statistics 36:2–34.

Cassels, S. and B.H. Singer. 2010. Population decline induced by gonorrhea and tuberculosis transmission: Micronesia during the Japanese occupation, 1919–45. Journal of Population Research 27:293–313.

CDC. 2006. Sexually Transmitted Disease Surveillance 2005, Centers for Disease Control and Prevention. Atlanta: U.S. Department of Health and Human Services.

CDC. 2015. Division of STD Prevention of the Centers for Disease Control and Prevention. Sexually transmitted disease surveillance 2015. Atlanta: U.S. Department of Health and Human Services https://www.cdc.gov/std/stats15/STD-Surveillance-2015-print.pdf

Conner, M.M. and M.W. Miller. 2004. Movement patterns and spatial epidemiology of a prion disease in mule deer population units. Ecological Applications 14:1870–1881.

Creatsas, G. and E. Deligeroglou. 2012. Microbial ecology of the lower genital tract in women with sexually transmitted diseases. Journal of Medical Microbiology 61: 1347-1351.

Foley J.E., E.V. Queen, B. Sacks, and P. Foley. 2005. GIS-facilitated spatial epidemiology of tick-borne diseases in coyotes (*Canis latrans*) in northern and coastal

California. Comparative Immunology, Microbiology and Infectious Diseases 28:197-212.

Marrazzo, J.M. and D.H. Martin. 2007. Management of women with cervicitis. Clinical Infectious Diseases 44: S102–S110.

Norman, S.A., R.F. DiGiacomo, F.M.D. Gulland, J.S. Meschke, and M.S. Lowry. 2008. Risk factors for an outbreak of leptospirosis in California sea lions (*Zalophus californianus*) in California. Journal of Wildlife Diseases 44:837-844.

Rushton, S.P., P.W.W. Lurz, J. Gurnell, and R. Fuller. 2000. Modeling the spatial dynamics of parapoxvirus disease in red and grey squirrels: a possible cause of the decline in the red squirrel in the UK? Journal of Applied Ecology 37:997–1012.

Shuai, J., P. Buck, P. Sockett, J. Aramini, and F. Pollari. 2006. A GIS-driven integrated real-time surveillance pilot system for national West Nile virus dead bird surveillance in Canada. International Journal of Health Geographics 5:17.

Websites

CDC fact sheet on reported STDs in the United States
https://www.cdc.gov/nchhstp/newsroom/docs/factsheets/STD-Trends-508.pdf

CDC 2015 Sexually Transmitted Disease Surveillance Data Report
https://www.cdc.gov/std/stats15/default.htm

Information on Sexually Transmitted Diseases
https://www.cdc.gov/std/

STD Terminology and Definitions
https://www.cdc.gov/std/stats/casedefinitions.htm

STD Data and Statistics
https://www.cdc.gov/std/stats/default.htm

Background on chlamydia and the disease characteristics
https://www2a.cdc.gov/stdtraining/ready-to-use/chlamydia.htm

Background on gonorrhea and the disease characteristics
https://www2a.cdc.gov/stdtraining/ready-to-use/gonorrhea.htm

GIS and Public health
https://www.cdc.gov/gis/index.htm

GIS Training modules from CDC
https://www.cdc.gov/gis/gis-training.htm

Cartographic guidelines for public health
https://www.cdc.gov/dhdsp/maps/gisx/resources/cartographic_guidelines.pdf

Causes, symptoms, treatments and cure for STDs
https://www.nichd.nih.gov/health/topics/stds/conditioninfo/Pages/causes.aspx

Applications of GIS in different disciplines
http://gisgeography.com/gis-applications-uses/

Texas STD Surveillance Report
https://www.dshs.texas.gov/hivstd/reports/

Approach

The learning outcomes that will be accomplished through the implementation of these exercises are:

1. Understanding information about Sexually Transmitted Diseases (STDs), facts, causes, symptoms, transmission, and treatment of STDs.
2. Understanding national and state-level surveillance and geographical trends in the spread of STDs.
3. Interpreting and understanding the differences between the chlamydia and gonorrhea infections, their symptoms, treatments and the annual number of cases reported within the US.
4. Understanding the impact of gender, racial, social and behavioral risk factors on the prevalence of STDs.
5. Understanding the principles and applications of Geographic Information Systems (GIS) in disease mapping.
6. Constructing STD datasets of the state or other geographical location of their choice.
7. Preparing the tables and graphs from the downloaded STD datasets and interpreting the tables and graphs created from the excel file.
8. Downloading and linking datasets of STDs into Arc GIS software and making spatial maps of STD incidence for the chosen geographical region.
9. Preparing spatial and temporal maps of disease spread.
10. Explaining the results through answering the questions within the exercises or presenting the results in powerpoint format and discuss.
11. Writing a final research report.

Student worksheets

Part 1: Downloading, analyzing, building and Interpreting the STD datasets

Purpose

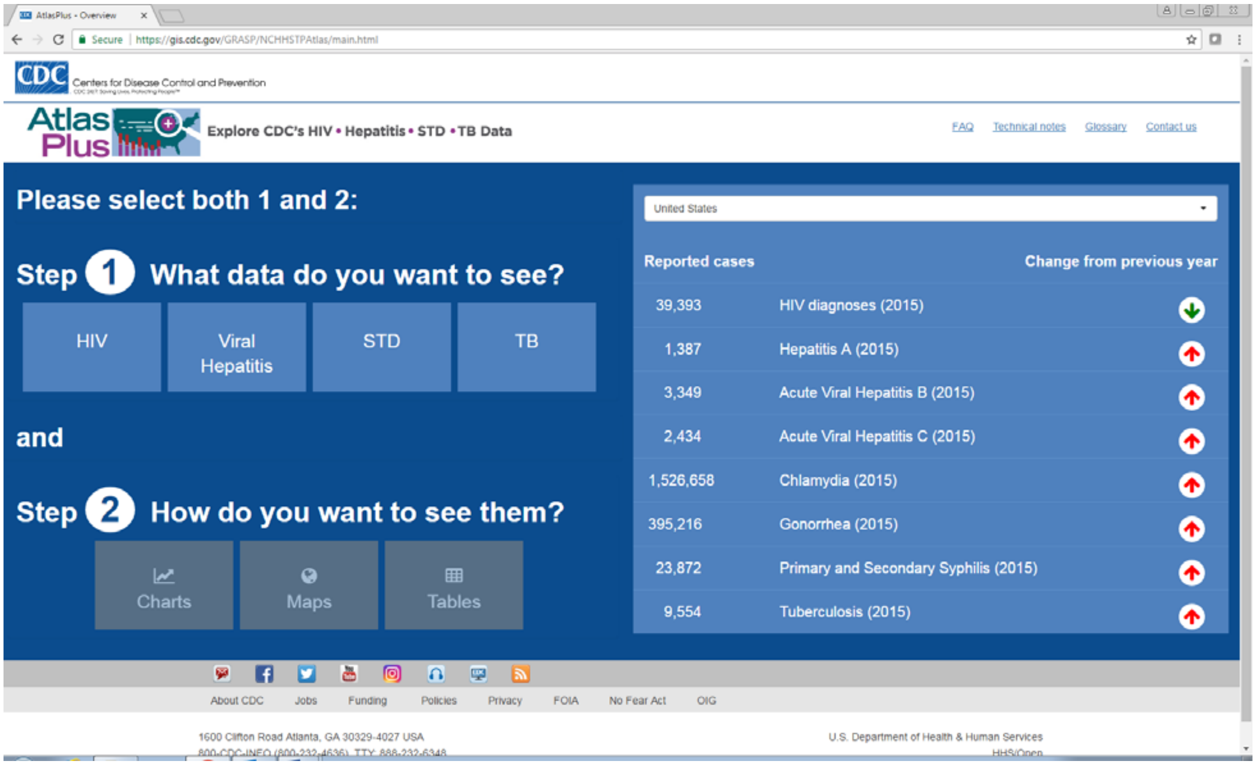
In this exercise, you will download the county-level data for chlamydia and gonorrhea disease counts for the state of Texas for each of the years 2000 to 2014 and then construct an Excel sheet with all the assembled data.

Introduction and background

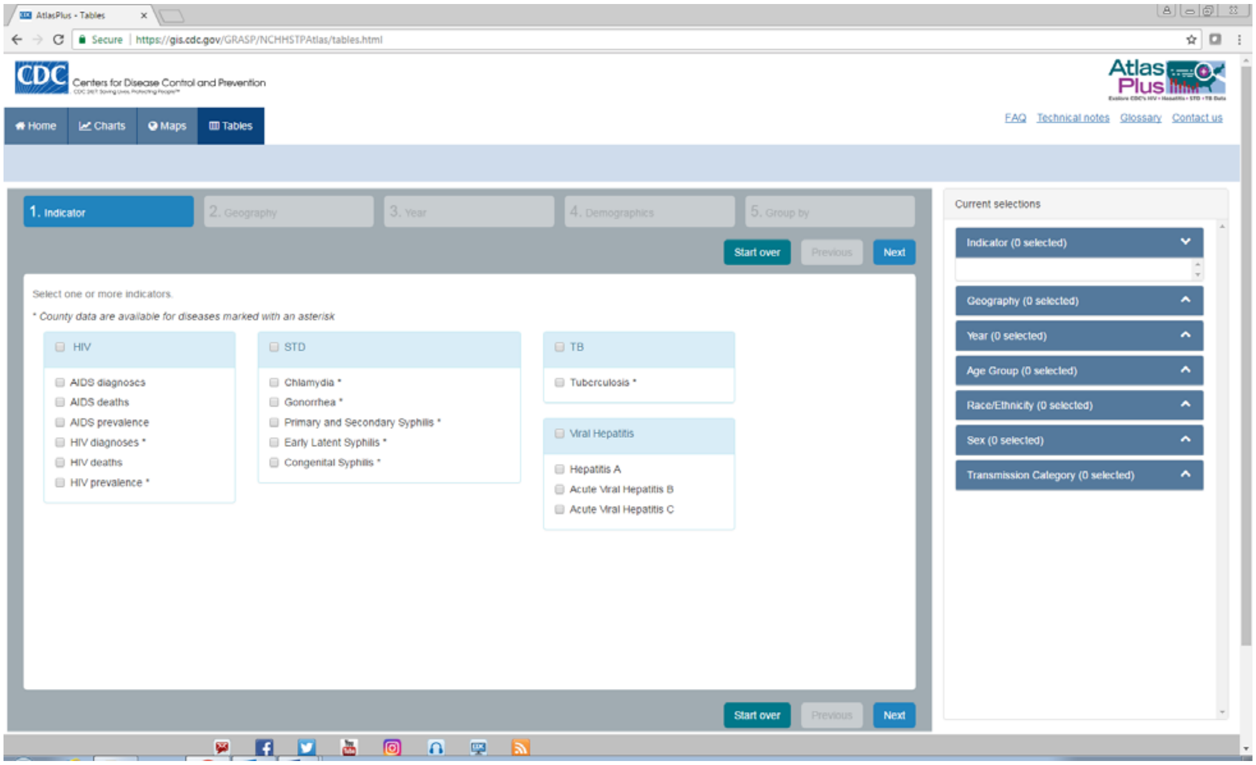
The disease incidence data showing the chlamydia and gonorrhea incidence at the county level for the entire United States is available from the Atlas Plus website of CDC at https://gis.cdc.gov/GRASP/NCHHSTPAtlas/main.html. These data are publicly available through the internet. The Atlas Plus datasets are the data repository that provides the user with access to the data on HIV, Hepatitis, STD and TB disease for multiple years at the county level for the entire nation.

Accessing and Downloading the CDC Data

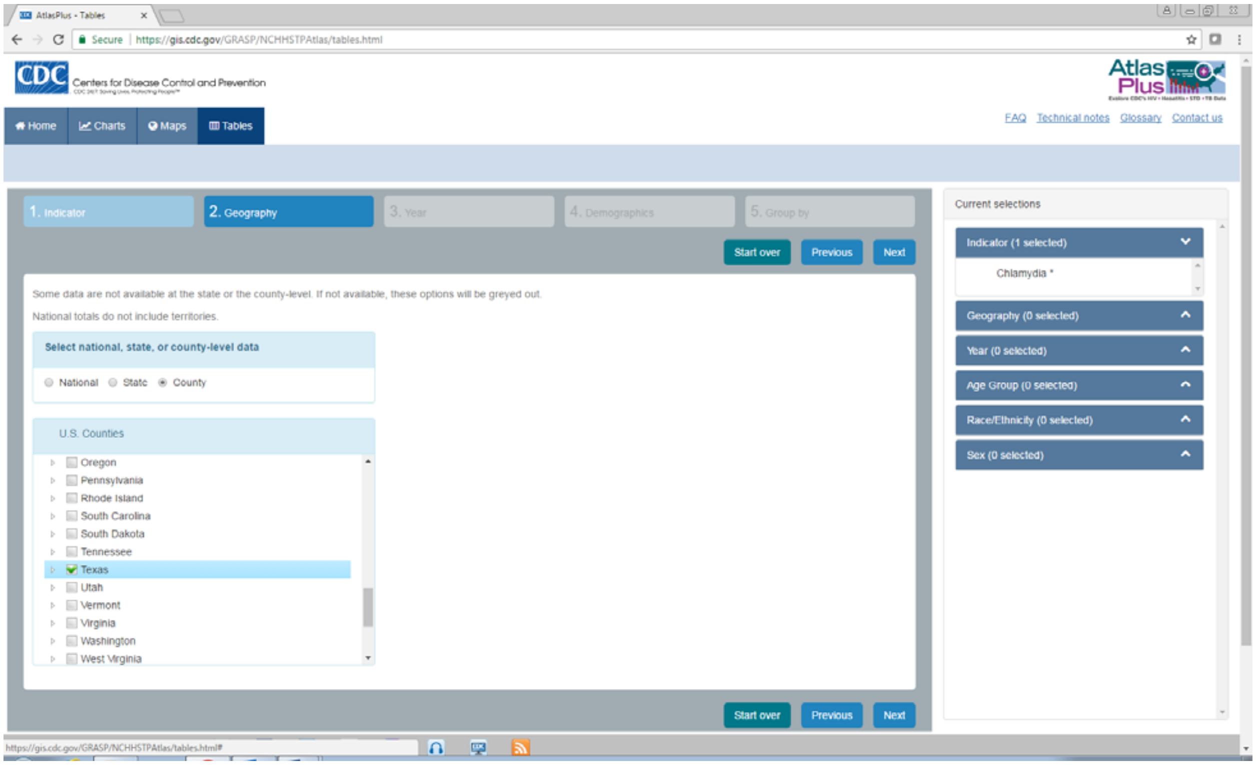
Navigate to the website https://gis.cdc.gov/GRASP/NCHHSTPAtlas/main.html and then select the “STD” disease you want to map in the step 1 and then select “Tables” in step 2.



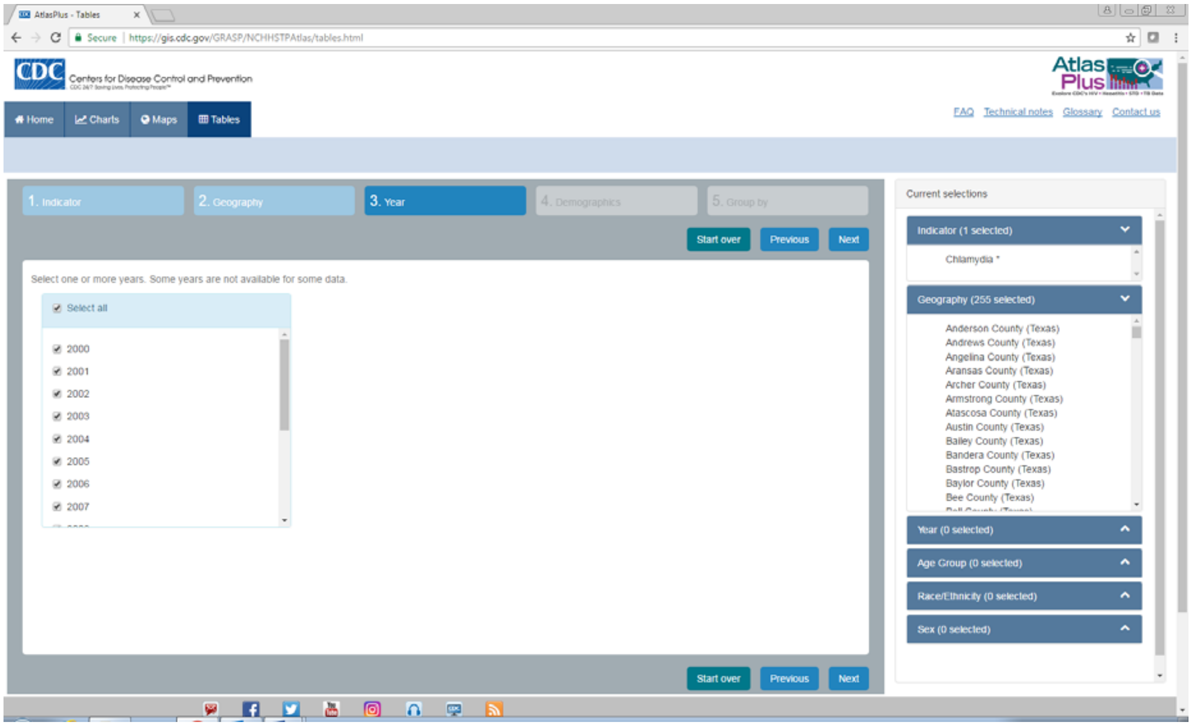
This will result in transfer to the next window, where you will be given the option to choose what type of STD data you want to download in the "option1: Indicator". The rest of the options are greyed out and you will be prompted to select in each tab and then complete all the five tabs given on the website. In this window, select "Chlamydia" and then hit next.



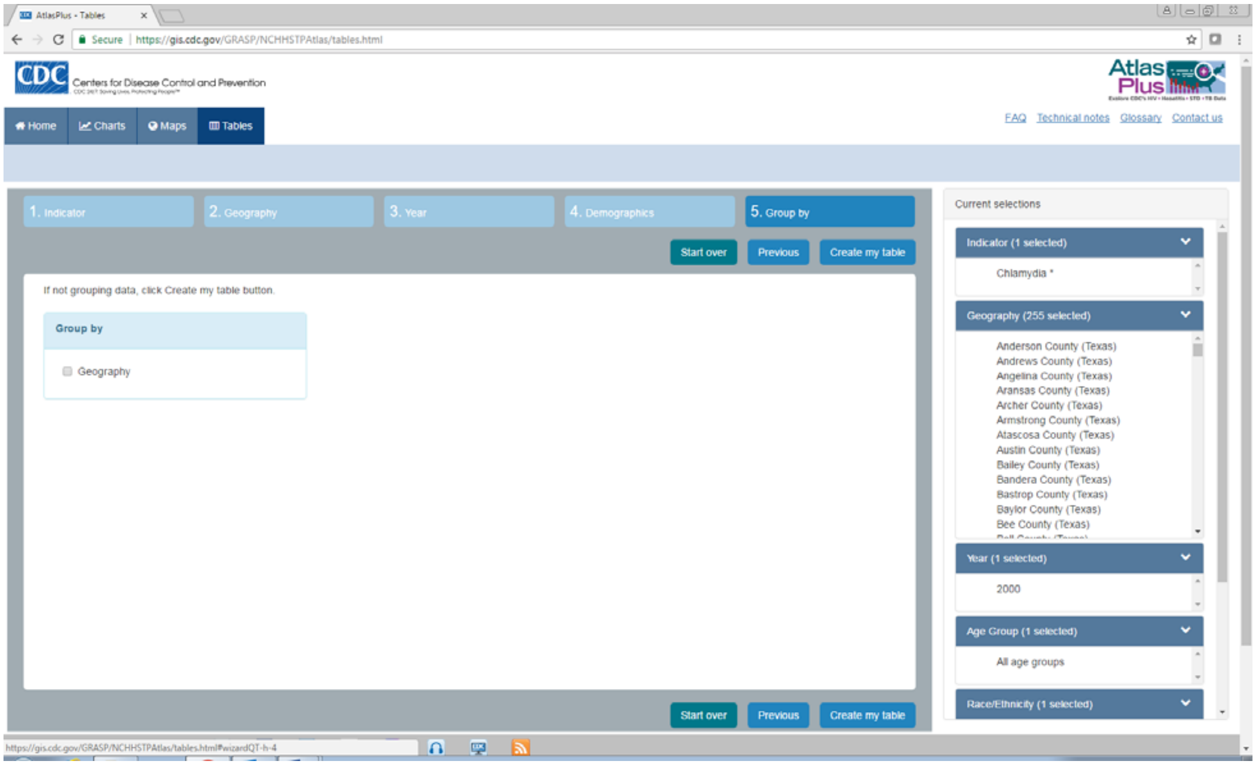
Then, you will be navigated to the other window where you need to select the geographical region in the tab. Select ‘County’ and then ‘Texas’ as State and then click next.



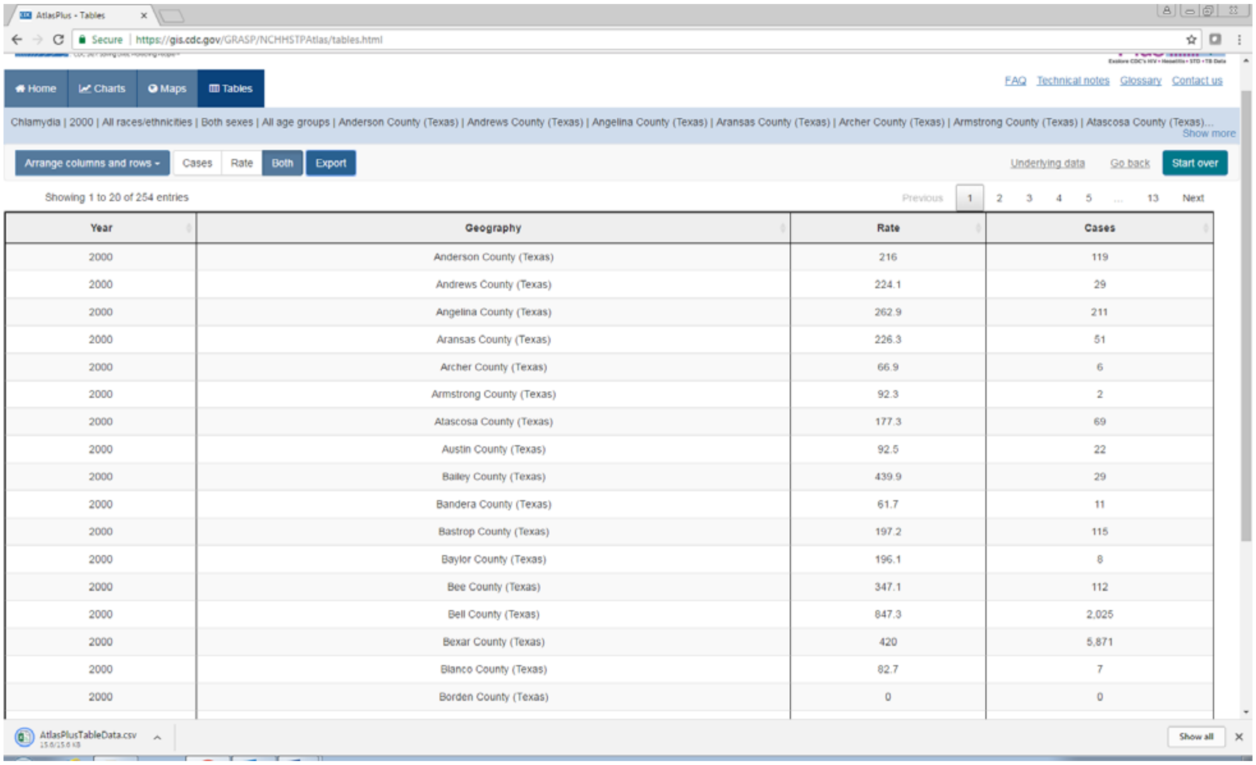
In the next window, you are prompted to select the year for which the data should be downloaded. If you have a good and fast internet connection, then click “select all”; otherwise, select 2-3 years at a time and download the data. You must download the data for all the given years. Then, click next.



Keep the default option for the "demographic" and then keep the radio button of geography unchecked in the "Group by" tab and then click "Create my Table".

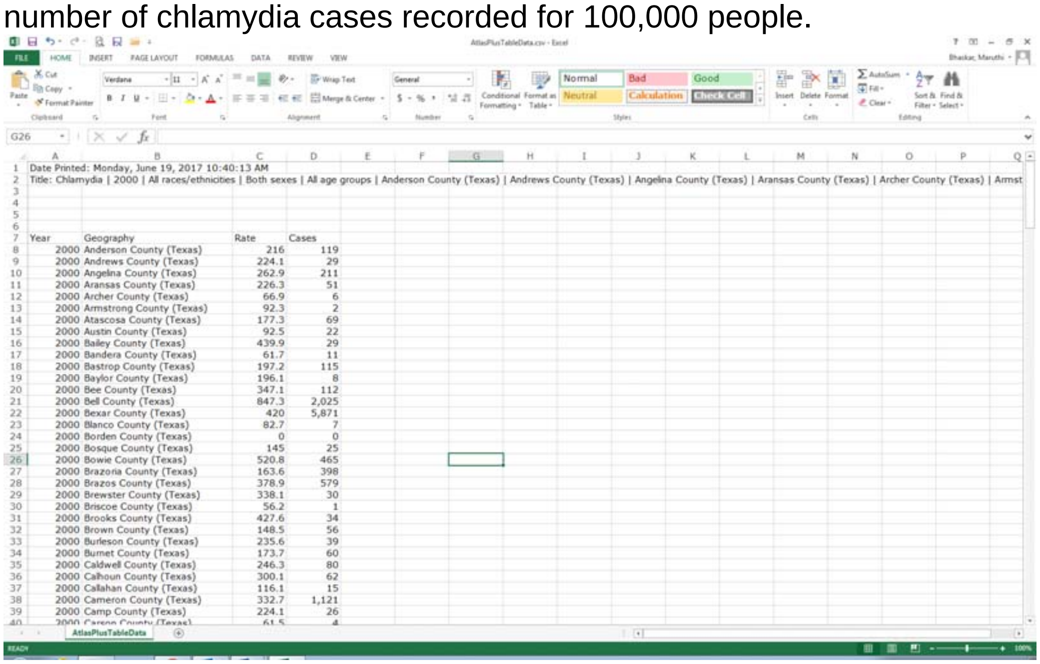


You will see a table where all your data was created and then click “Export” on the top and your data will be downloaded as a .csv file.



Year	Geography	Rate	Cases
2000	Anderson County (Texas)	216	119
2000	Andrews County (Texas)	224.1	29
2000	Angelina County (Texas)	262.9	211
2000	Aransas County (Texas)	226.3	51
2000	Archer County (Texas)	66.9	6
2000	Armstrong County (Texas)	92.3	2
2000	Atascosa County (Texas)	177.3	69
2000	Austin County (Texas)	92.5	22
2000	Bailey County (Texas)	439.9	29
2000	Bandera County (Texas)	61.7	11
2000	Bastrop County (Texas)	197.2	115
2000	Baylor County (Texas)	196.1	8
2000	Bee County (Texas)	347.1	112
2000	Bell County (Texas)	847.3	2,025
2000	Bexar County (Texas)	420	5,871
2000	Blanco County (Texas)	82.7	7
2000	Borden County (Texas)	0	0

The downloaded data file will be seen as follows. You will see two columns "Cases" and "Rate" for each year you are downloading. 'Case' refers to the total number of chlamydia cases reported in the county for the year and the 'Rate' refers to the number of chlamydia cases recorded for 100,000 people.

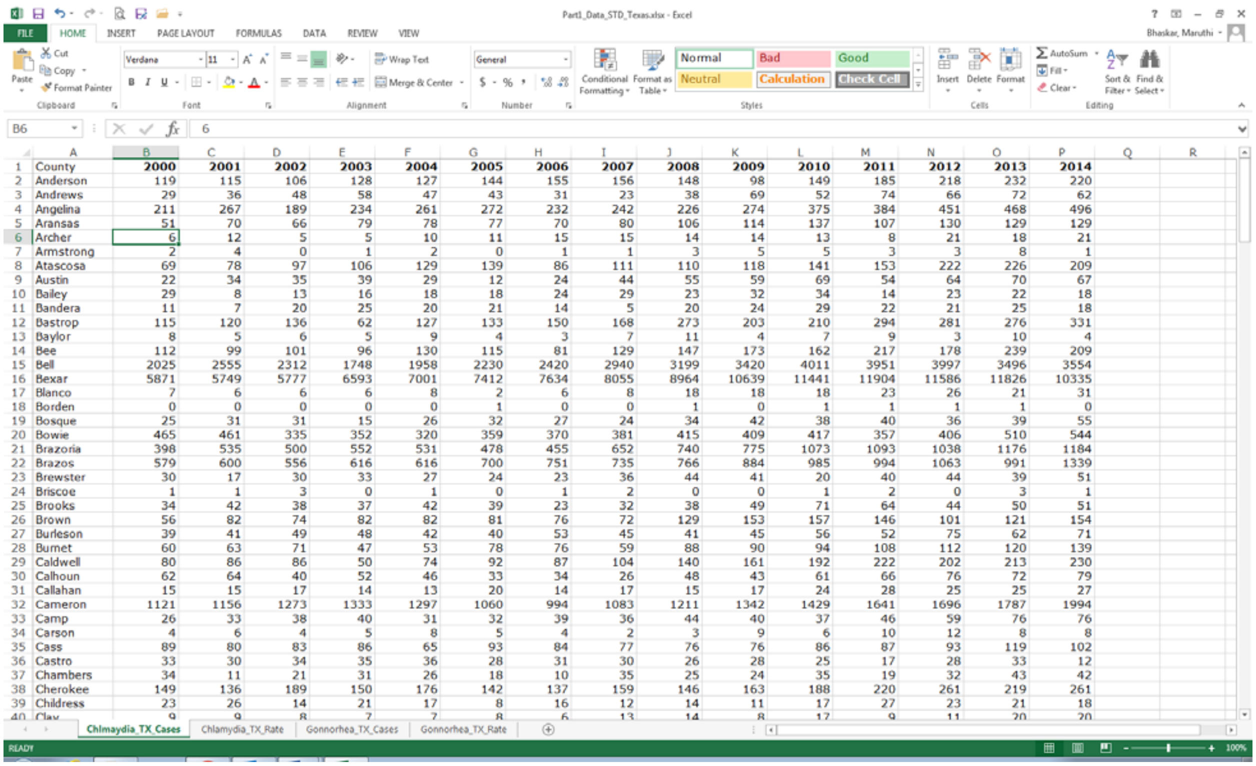


Date Printed: Monday, June 19, 2017 10:40:13 AM

Title: Chlamydia | 2000 | All races/ethnicities | Both sexes | All age groups | Anderson County (Texas) | Andrews County (Texas) | Angelina County (Texas) | Aransas County (Texas) | Archer County (Texas) | Armst

Year	Geography	Rate	Cases
2000	Anderson County (Texas)	216	119
2000	Andrews County (Texas)	224.1	29
2000	Angelina County (Texas)	262.9	211
2000	Aransas County (Texas)	226.3	51
2000	Archer County (Texas)	66.9	6
2000	Armstrong County (Texas)	92.3	2
2000	Atascosa County (Texas)	177.3	69
2000	Austin County (Texas)	92.5	22
2000	Bailey County (Texas)	439.9	29
2000	Bandera County (Texas)	61.7	11
2000	Bastrop County (Texas)	197.2	115
2000	Baylor County (Texas)	196.1	8
2000	Bee County (Texas)	347.1	112
2000	Bell County (Texas)	847.3	2,025
2000	Bexar County (Texas)	420	5,871
2000	Blanco County (Texas)	82.7	7
2000	Borden County (Texas)	0	0
2000	Bosque County (Texas)	145	25
2000	Bowie County (Texas)	520.8	465
2000	Brazoria County (Texas)	163.6	398
2000	Brazos County (Texas)	378.9	579
2000	Brewster County (Texas)	338.1	30
2000	Briscoe County (Texas)	56.2	1
2000	Brooks County (Texas)	427.6	34
2000	Brown County (Texas)	148.5	56
2000	Burleson County (Texas)	235.6	39
2000	Burnet County (Texas)	173.7	60
2000	Caldwell County (Texas)	246.3	80
2000	Calhoun County (Texas)	300.1	62
2000	Callahan County (Texas)	116.1	15
2000	Cameron County (Texas)	332.7	1,121
2000	Camp County (Texas)	224.1	26

Arrange all the data and make a single excel file with all the Chlamydia data for all the 'case' and all the 'rate' data. Your excel file should be seen as follows.



County	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014
Anderson	119	115	106	128	127	144	155	156	148	98	149	185	218	232	220
Andrews	29	36	48	58	47	43	31	23	38	69	52	74	66	72	62
Angelina	211	267	189	234	261	272	232	242	226	274	375	384	451	468	496
Aransas	51	70	66	79	78	77	70	80	106	114	137	107	130	129	129
Archer	6	12	5	5	10	11	15	15	14	14	13	8	21	18	21
Armstrong	2	4	0	1	2	0	1	1	3	5	5	3	3	8	1
Atascosa	69	78	97	106	129	139	86	111	110	118	141	153	222	226	209
Austin	22	34	35	39	29	12	24	44	55	59	69	54	64	70	67
Bailey	29	8	13	16	18	18	24	29	23	32	34	14	23	22	18
Bandera	11	7	20	25	20	21	14	5	20	24	29	22	21	25	18
Bastrop	115	120	136	62	127	133	150	168	273	203	210	294	281	276	331
Baylor	8	5	6	5	9	4	3	7	11	4	7	9	3	10	4
Bee	112	99	101	96	130	115	81	129	147	173	162	217	178	239	209
Bell	2025	2555	2312	1748	1958	2230	2420	2940	3199	3420	4011	3951	3997	3496	3554
Bexar	5871	5749	5777	6593	7001	7412	7634	8055	8964	10639	11441	11904	11586	11826	10335
Blanco	7	6	6	6	8	2	6	8	18	18	18	23	26	21	31
Borden	0	0	0	0	0	1	0	0	1	0	1	1	1	1	0
Bosque	25	31	31	15	26	32	27	24	34	42	38	40	36	39	55
Bowie	465	461	335	352	320	359	370	381	415	409	417	357	406	510	544
Brazoria	398	535	500	552	531	478	455	652	740	775	1073	1093	1038	1176	1184
Brazos	579	600	556	616	616	700	751	735	766	884	985	994	1063	991	1339
Brewster	30	17	30	33	27	24	23	36	44	41	20	40	44	39	51
Briscoe	1	1	3	0	1	0	1	2	0	0	1	2	0	3	1
Brooks	34	42	38	37	42	39	23	32	38	49	71	64	44	50	51
Brown	56	82	74	82	82	81	76	72	129	153	157	146	101	121	154
Burleson	39	41	49	48	42	40	53	45	41	45	56	52	75	62	71
Burnet	60	63	71	47	53	78	76	59	88	90	94	108	112	120	139
Caldwell	80	86	86	50	74	92	87	104	140	161	192	222	202	213	230
Calhoun	62	64	40	52	46	33	34	26	48	43	61	66	76	72	79
Callahan	15	15	17	14	13	20	14	17	15	17	24	28	25	25	27
Cameron	1121	1156	1273	1333	1297	1060	994	1083	1211	1342	1429	1641	1696	1787	1994
Camp	26	33	38	40	31	32	39	36	44	40	37	46	59	76	76
Carson	4	6	4	5	8	5	4	2	3	9	6	10	12	8	8
Cass	89	80	83	86	65	93	84	77	76	76	86	87	93	119	102
Castro	33	30	34	35	36	28	31	30	26	28	25	17	28	33	12
Chambers	34	11	21	31	26	18	10	35	25	24	35	19	32	43	42
Cherokee	149	136	189	150	176	142	137	159	146	163	188	220	261	219	261
Childress	23	26	14	21	17	8	16	12	14	11	17	27	23	21	18
Clay	9	9	8	7	7	8	6	13	14	8	17	9	11	20	20

Prepare another dataset for gonorrhea too, following the same steps as described above. Once you prepare your datasets, you will be analyzing the data and making the line graphs and time series graphs.
Using the data in your Excel sheets, answer the following questions:

1. Chlamydia disease incidence data
    a) Graph the time series data of chlamydia disease counts for five counties with highest incidence.
    b) Graph the time series data of chlamydia disease rate for five counties with highest incidence.
    c) Describe the trend of disease incidence and disease rate in these five counties over the last 15 years.
    d) Why are the disease counts always higher in particular counties over the others?

2. Gonorrhea disease incidence data

a) Graph the time series data of gonorrhea disease counts for five counties with highest incidence.
b) Graph the time series data of gonorrhea disease rate for five counties with highest incidence.
c) Describe the trend of disease incidence and disease rate in these five counties over the last 15 years.
d) Why are the disease counts always higher in particular counties over the others?

Part 2: Developing the geospatial database and mapping the disease incidence

Mapping and Exploring the Sexually Transmitted Diseases (STDs) Using Geographic Information Systems (GIS)

Purpose

In this exercise, you will map the distribution and the geographical heterogeneity of chlamydia and gonorrhea incidence across Texas. You will also analyze the spatial and temporal changes in the distribution of the total cases of disease and the rate of disease incidence in Texas.

Mapping the spatial and temporal heterogeneity of chlamydia and gonorrhea in Texas will identify the regions (counties in Texas in this exercise) with higher incidence of disease, which can help to identify the STD prevention strategies for core regions instead of targeting uniformly across the state of Texas. Overlap of core areas among STDs will help to plan the intervention and prevention strategies that can be combined to target multiple STDs effectively. Geostatistical techniques can also be used to visualize disease patterns and identify emerging outbreaks. Evaluations of the spatial distribution of STDs can guide both prevention and intervention efforts, as well as provide additional clues to the local transmission of disease.

Also, the purpose of the exercise is to familiarize you to the ESRI's ArcGIS software, displaying data, searching data, navigating the ArcGIS files, and improving your knowledge of the application of GIS in the health and biological sciences.

Introduction and background

In this lab, you will be working with ESRI's ArcGIS 10.3 or 10.5 software. ArcGIS is widely used in industry by professional GIS users. Working in ArcGIS typically uses three main modules: Catalog, Arc Toolbox, and Arc Map. These modules represent the three basic necessities of GIS: data management, data analysis, and data output/mapping. In this lab, we will cover the Arc Map module in greater depth, as well as discover some of its key functionality.

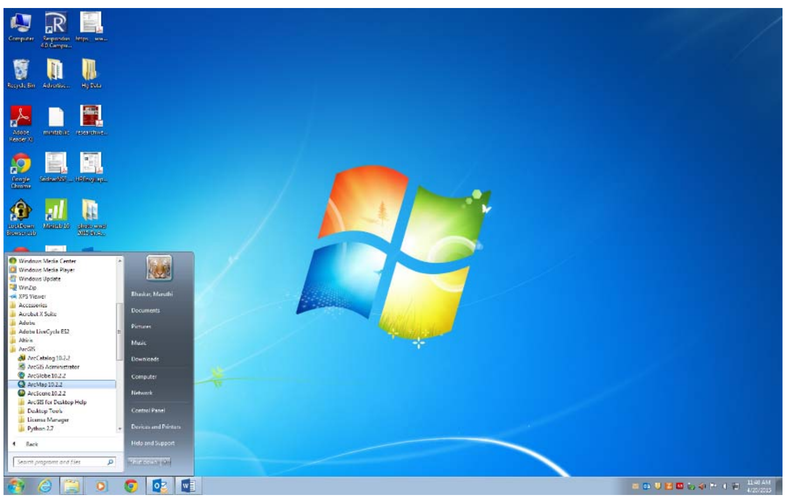
Data for This Class Exercise

All the files required for this class are posted under a specified folder.
The folder has many subfolders. They are:

1. Texas Counties Political
2. Chlamydia Data
3. Gonorrhea Data

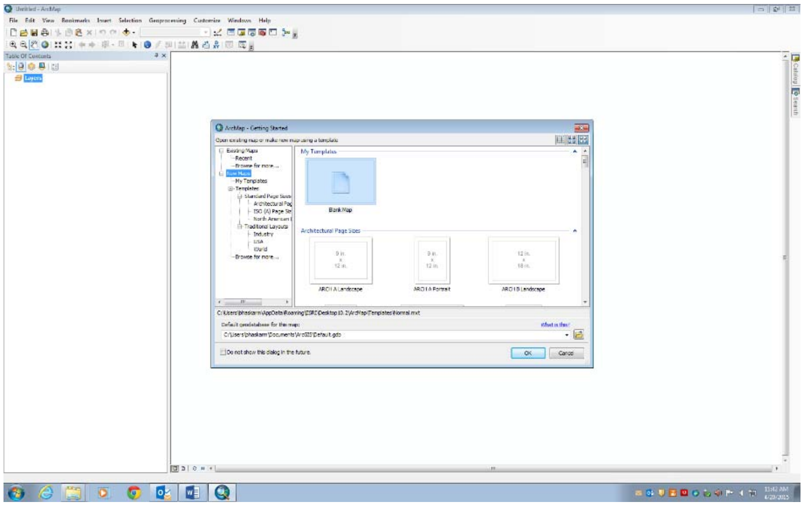
Introduction to ArcMap
ArcMap is the central module of ArcGIS used for displaying, organizing, browsing, managing and processing your data, and working with map documents. The other two main modules, ArcCatalog and ArcToolbox are accessible from within the ArcMap application.

To start ArcMap, click on the Start button and navigate to All Programs -> ArcGIS -> Arc Map 10.5.



When ArcMap starts, it asks you what map document you want to create or work with.

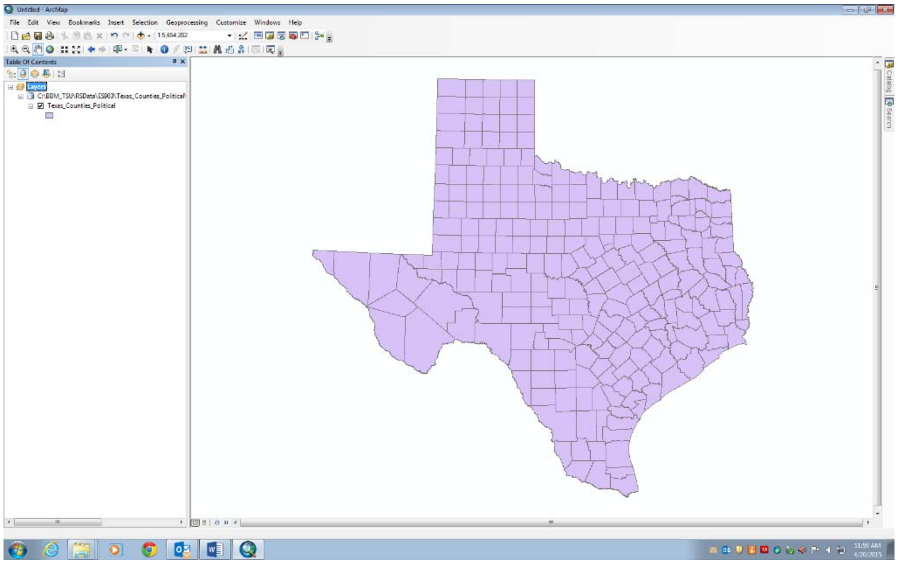
At this point, just click on OK, to have ArcMap start with a blank map document. Your ArcMap application will look very much like this:



ArcMap looks similar to many Windows applications. There is a menu across the top of the window – with items such as “File”, “Edit”, “View”, and “Help”. Under the menu, there are one or more rows of small icons representing convenient shortcuts to often used actions. On the left side of the application is a window with the subtitle “Table of Contents”. The large window in the rest of the application is the map area (it has no subtitle) and is where the map data are displayed. On the right edge of the window are two small sideways icons, one says “Catalog” and the other says “Search”.

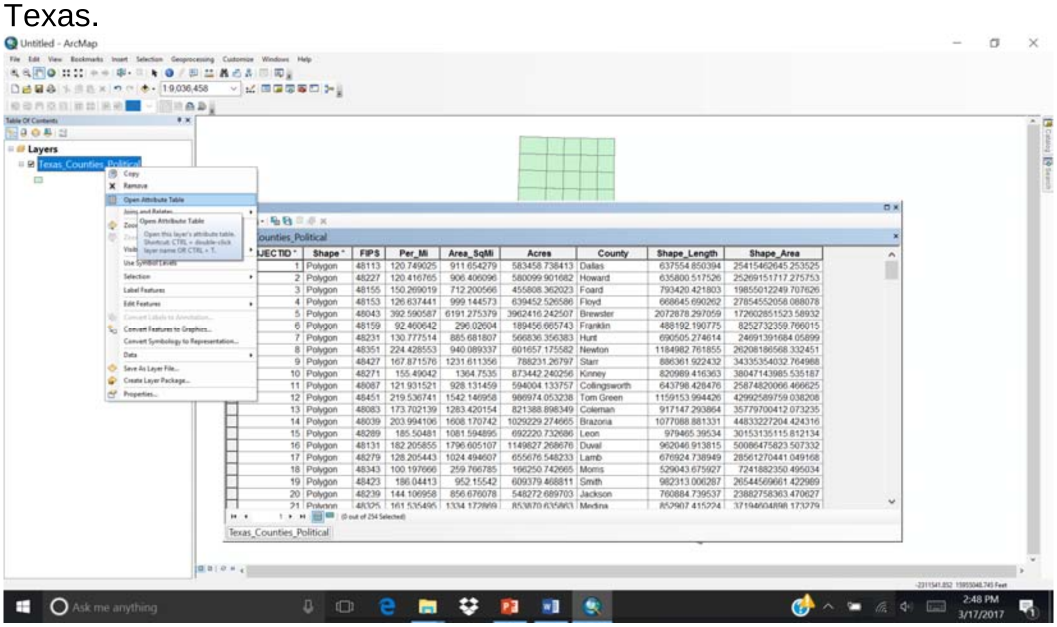
Connecting to your lab data

Now, let’s get the data in the “C:\\ClassName” (Create a folder with the name of your class, for easy tracking of files). In the ArcMap window, in the second row, click on the “+” icon called “Add Data” and then navigate to the folder ‘ClassName’ and then click on the folder “Texas_counties_Political” and then click add. The image will be displayed as shown below.

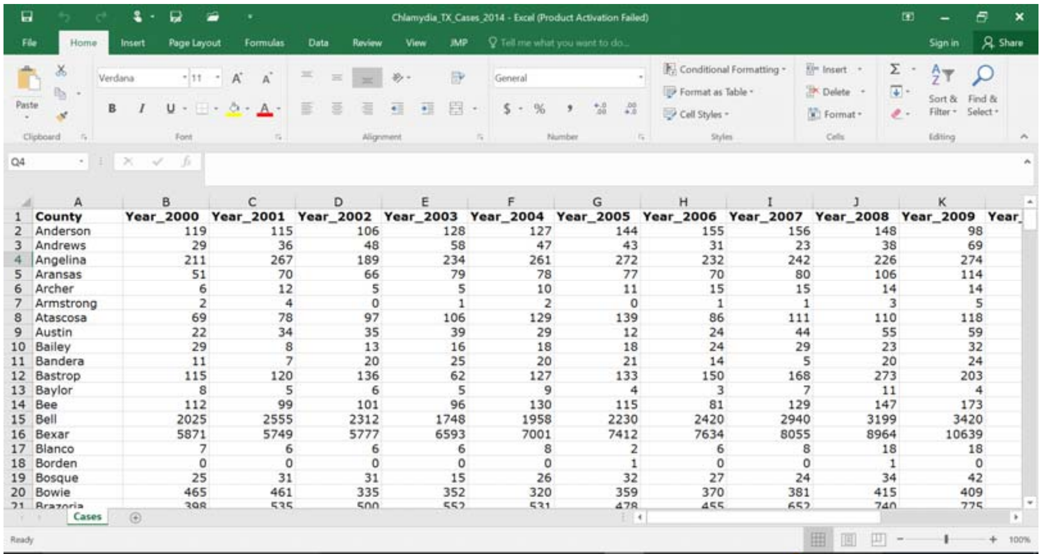


Now to look at the tabular data behind the displayed image, right click on the "Texas_Counties_Political" layer in the table of contents. In the pop-up menu, click on the attribute table and then you can see the table displayed with all the data. The attribute table can be resized by holding along the edges.

In the attribute table, you can see the column called 'County' displaying the names of all the counties in the shape file. Now to this table you should link the Chlamydia disease data, so that you can display the spatial distribution of Chlamydia incidence in Texas.



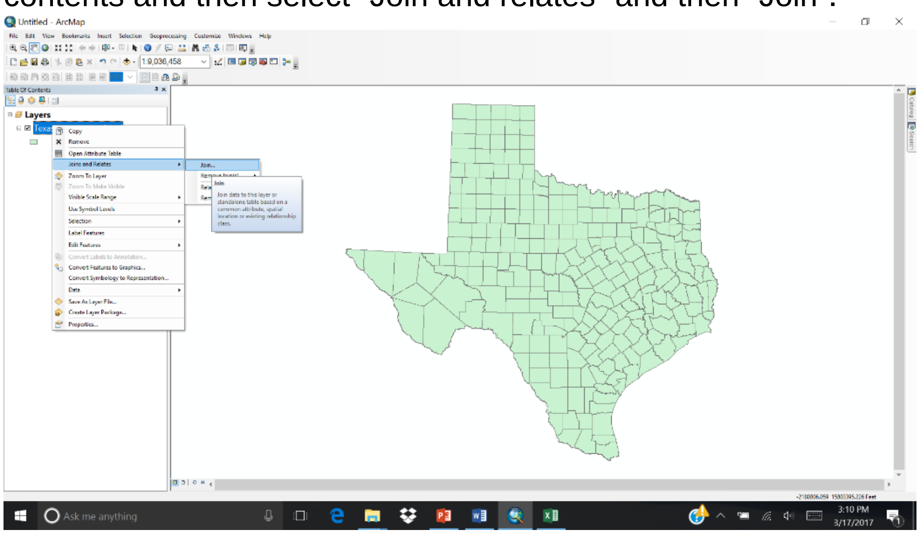
Now open the Excel file “Part2_Chlamydia_TX_Cases_2014.csv”. The excel file shows the Chlamydia cases recorded in Texas by county for the years from 2000 to 2014. Now we need to import this data into the GIS, so that we can make the geographical distribution of the data and to conduct further geostatistical analysis.



	County	Year_2000	Year_2001	Year_2002	Year_2003	Year_2004	Year_2005	Year_2006	Year_2007	Year_2008	Year_2009
2	Anderson	119	115	106	128	127	144	155	156	148	98
3	Andrews	29	36	48	58	47	43	31	23	38	69
4	Angelina	211	267	189	234	261	272	232	242	226	274
5	Aransas	51	70	66	79	78	77	70	80	106	114
6	Archer	6	12	5	5	10	11	15	15	14	14
7	Armstrong	2	4	0	1	2	0	1	1	3	5
8	Atascosa	69	78	97	106	129	139	86	111	110	118
9	Austin	22	34	35	39	29	12	24	44	55	59
10	Bailey	29	8	13	16	18	18	24	29	23	32
11	Bandera	11	7	20	25	20	21	14	5	20	24
12	Bastrop	115	120	136	62	127	133	150	168	273	203
13	Baylor	8	5	6	5	9	4	3	7	11	4
14	Bee	112	99	101	96	130	115	81	129	147	173
15	Bell	2025	2555	2312	1748	1958	2230	2420	2940	3199	3420
16	Bexar	5871	5749	5777	6593	7001	7412	7634	8055	8964	10639
17	Blanco	7	6	6	6	8	2	6	8	18	18
18	Borden	0	0	0	0	0	1	0	0	1	0
19	Bosque	25	31	31	15	26	32	27	24	34	42
20	Bowie	465	461	335	352	320	359	370	381	415	409
21	Brazoria	398	535	500	552	531	478	455	652	740	775

Linking the Excel Data to the Attribute Table

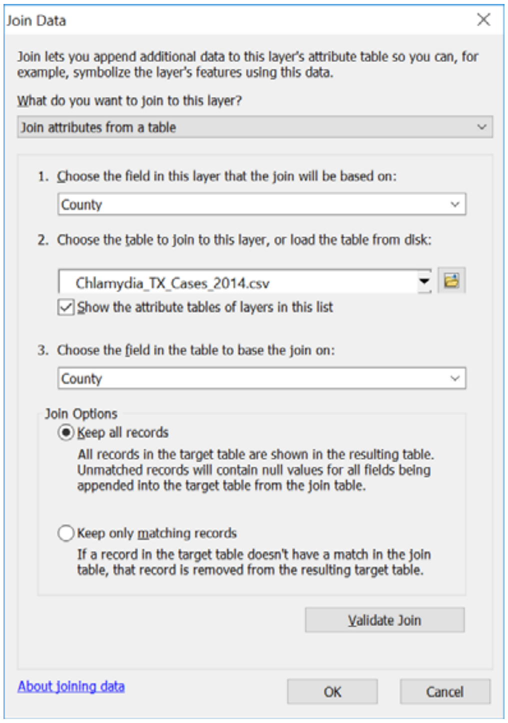
In order to import and display the data in the excel file, one way is to link the data to the existing attribute table field in the geospatial data. You can link these two tables using the ‘Common Key Element’ in both the tables, which is the “County” column. For this again right click on the “Texas_Counties_Political” layer in the table of contents and then select “Join and relates” and then “Join”.



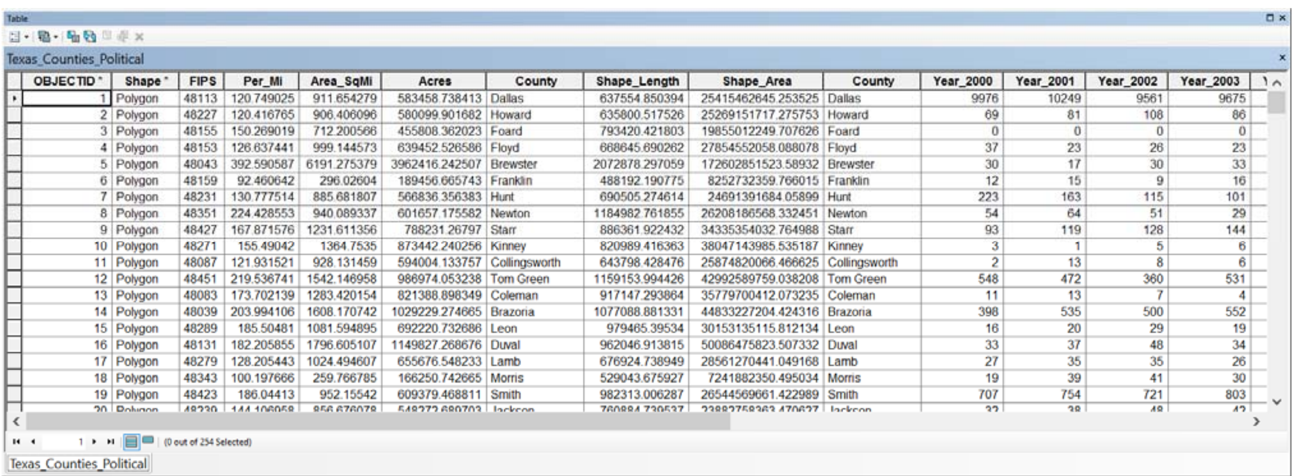
TIEE

Teaching Issues and Experiments in Ecology - Volume 13, March 2018

In the pop-up ‘Join Data’ window make sure to select “county” in the first field and then in the second field navigate to the “Part2_Chlamydia_TX_Cases_2014.csv” and in the third field choose “County” column within the excel file and then leave all default and then click “Validate Join” and then click “OK”.



Now when you open your attribute table, you can see all the data is being appended to the existing attribute table. Now, you can display and do any sort of analysis on these data in the GIS.



Table

Texas_Counties_Political

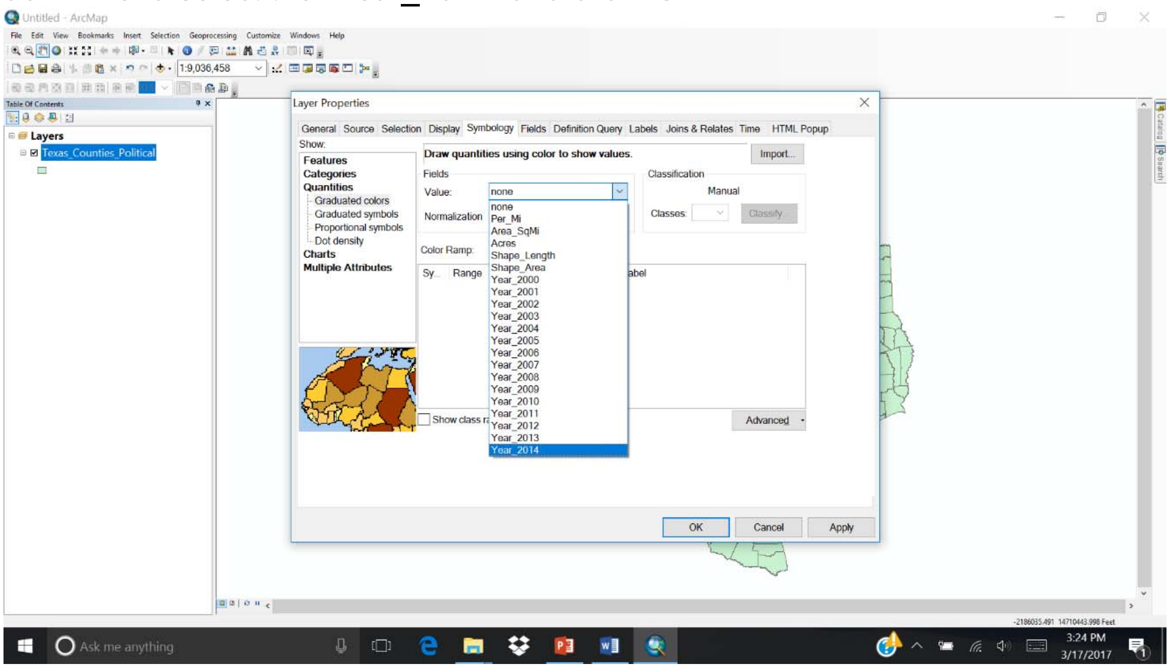
OBJECTID *	Shape *	FIPS	Per_Mi	Area_SqMi	Acres	County	Shape_Length	Shape_Area	County	Year_2000	Year_2001	Year_2002	Year_2003
1	Polygon	48113	120.749025	911.654279	583458.738413	Dallas	637554.850394	25415462645.253525	Dallas	9976	10249	9561	9675
2	Polygon	48227	120.416765	906.406096	580099.901682	Howard	635800.517526	25269151717.275753	Howard	69	81	108	86
3	Polygon	48155	150.269019	712.200566	455808.362023	Foard	793420.421803	19855012249.707626	Foard	0	0	0	0
4	Polygon	48153	126.637441	999.144573	639452.526586	Floyd	668645.690262	27854552058.088078	Floyd	37	23	26	23
5	Polygon	48043	392.590587	6191.275379	3962416.242507	Brewster	2072878.297059	172602851523.58932	Brewster	30	17	30	33
6	Polygon	48159	92.460642	296.02604	189456.665743	Franklin	488192.190775	8252732359.766015	Franklin	12	15	9	16
7	Polygon	48231	130.777514	885.681807	566836.356383	Hunt	690505.274614	24691391684.05899	Hunt	223	163	115	101
8	Polygon	48351	224.428553	940.089337	601657.175582	Newton	1184982.761855	26208186568.332451	Newton	54	64	51	29
9	Polygon	48427	167.871576	1231.611356	788231.26797	Starr	886361.922432	34335354032.764988	Starr	93	119	128	144
10	Polygon	48271	155.49042	1364.7535	873442.240256	Kinney	820989.416363	38047143985.535187	Kinney	3	1	5	6
11	Polygon	48087	121.931521	928.131459	594004.133757	Collingsworth	643798.428476	25874820066.466625	Collingsworth	2	13	8	6
12	Polygon	48451	219.536741	1542.146958	986974.053238	Tom Green	1159153.994426	42992589759.038208	Tom Green	548	472	360	531
13	Polygon	48083	173.702139	1283.420154	821388.898349	Coleman	917147.293864	35779700412.073235	Coleman	11	13	7	4
14	Polygon	48039	203.994106	1608.170742	1029229.274665	Brazoria	1077088.881331	44833227204.424316	Brazoria	398	535	500	552
15	Polygon	48289	185.50481	1081.594895	692220.732686	Leon	979465.39534	30153135115.812134	Leon	16	20	29	19
16	Polygon	48131	182.205855	1796.605107	1149827.268676	Duval	962046.913815	50086475823.507332	Duval	33	37	48	34
17	Polygon	48279	128.205443	1024.494607	655676.548233	Lamb	676924.738949	28561270441.049168	Lamb	27	35	35	26
18	Polygon	48343	100.197666	259.766785	166250.742665	Morris	529043.675927	7241882350.495034	Morris	19	39	41	30
19	Polygon	48423	186.04413	952.15542	609379.468811	Smith	982313.006287	26544569661.422989	Smith	707	754	721	803

(0 out of 254 Selected)

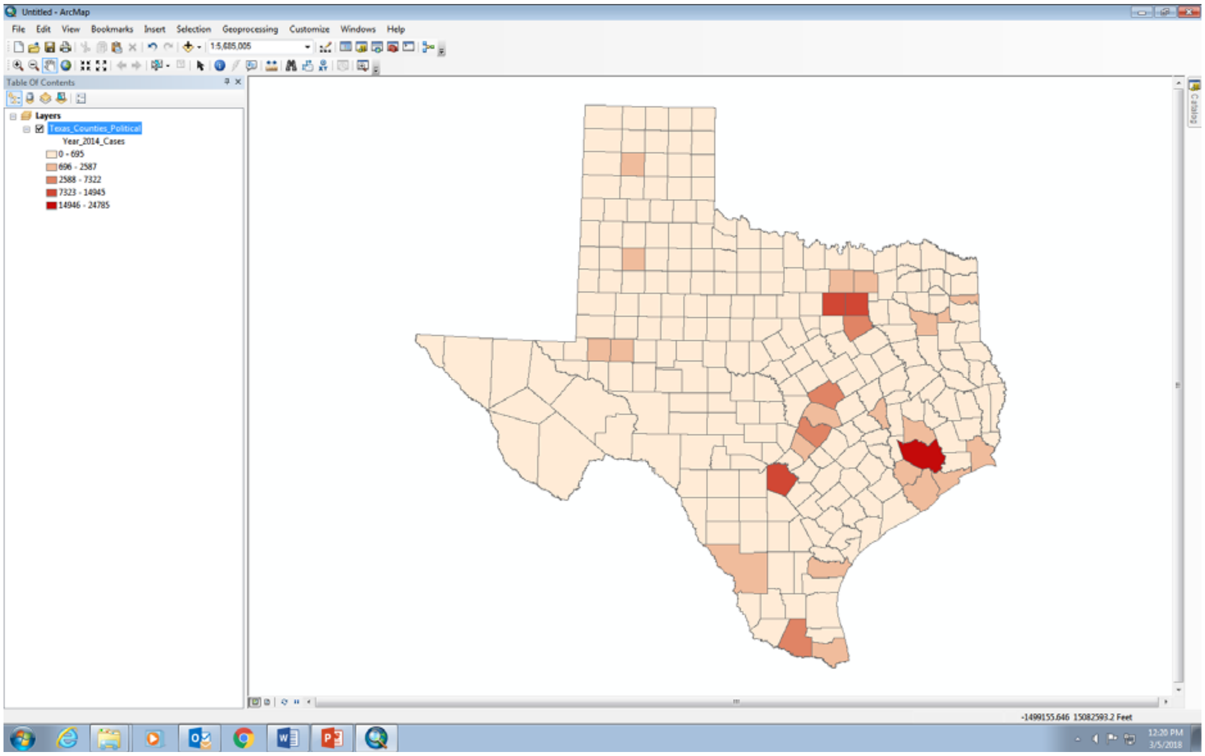
Texas_Counties_Political

Creating the Chlamydia Spatial Distribution Maps

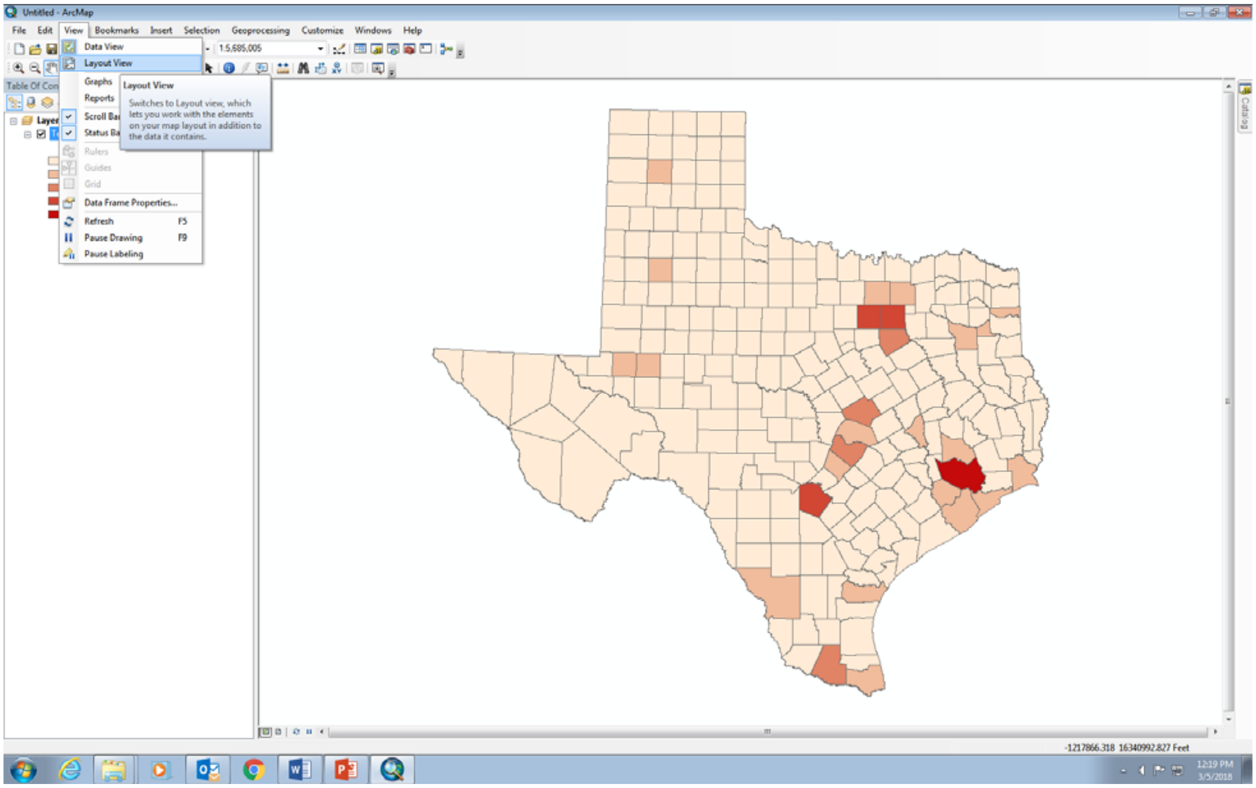
Now double click on the “Texas_Counties_Political” layer in the Table of Contents, a layer properties window will open up. In the “Layer Properties” window, click on the “Quantities” in the left pane and then select ‘Graduated Colors’ and in the Values pull down menu select the “Year_2014” and click “OK”.



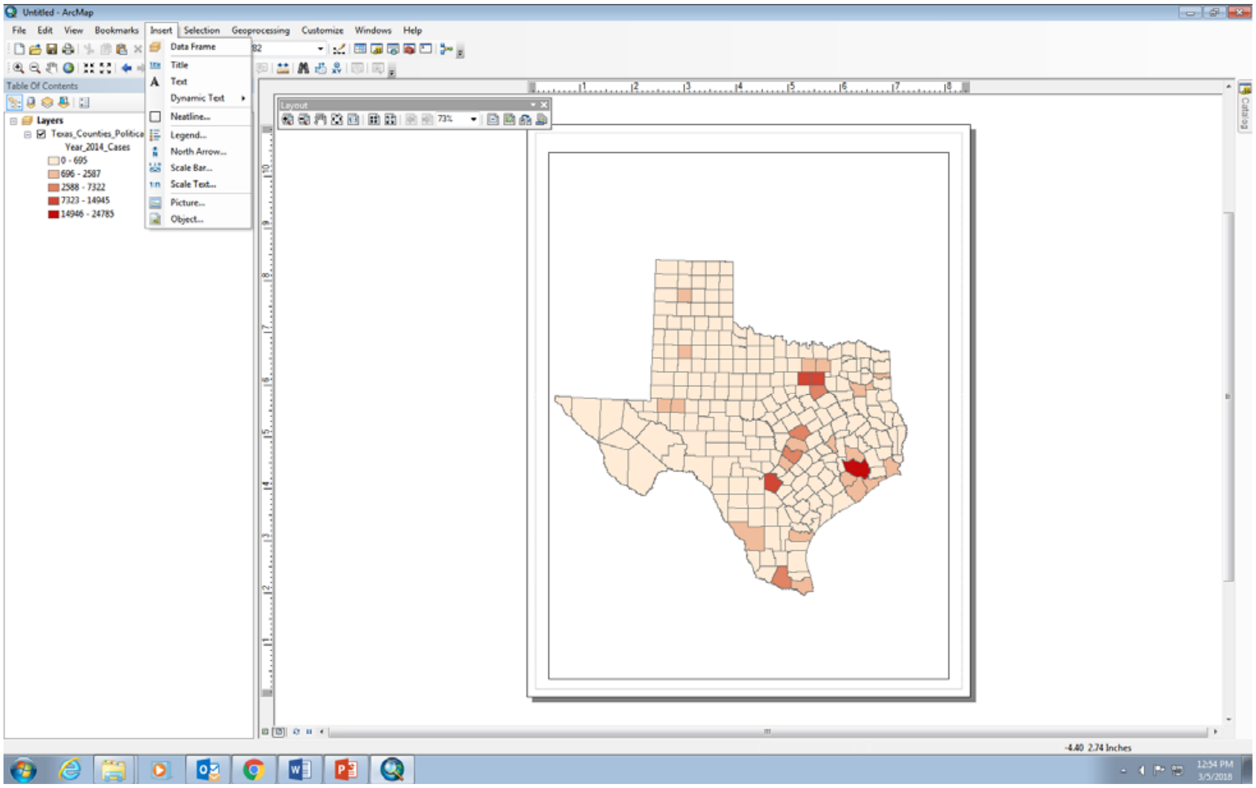
Now you can see the Texas map showing all the reported Chlamydia cases in 2014. You can choose the different color palettes and the one that is shown here assigns all the ‘dark’ colors to the counties with high incidences and ‘light’ colors to the counties with low incidence. The data will be displayed in default classification type ‘Natural Breaks (Jenks)’ with five classes. Natural breaks are data-specific classifications and are inherent to the data. The five class breaks that are identified are the best group similar values, that maximize the differences between classes. This classification is based on the Jenks Natural Breaks algorithm.



To get your map ready for being exported as an image, click on the “View” in the top menu bar and then select “Layout view”.



Once the image is displayed in the layout view, you can click "Insert" in the menu bar and then insert a, 1) Title, 2) Legend, 3) Scale bar.

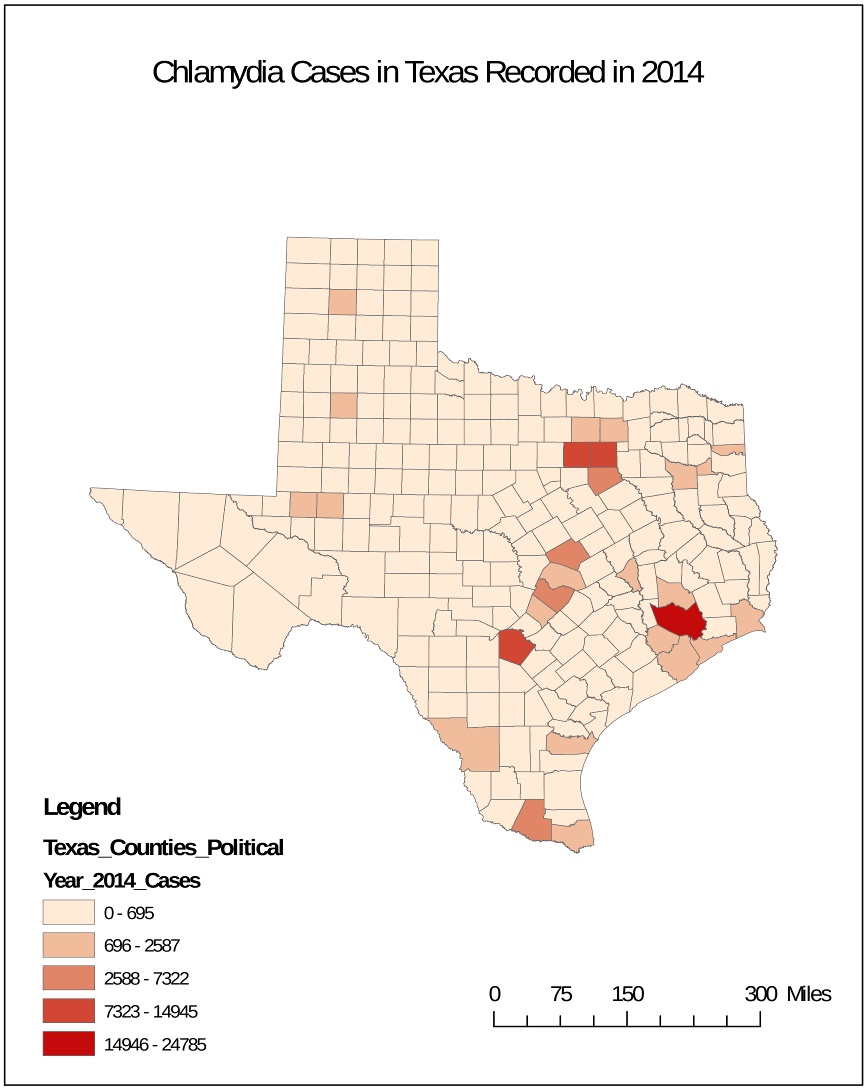


For inserting the title, click “Insert” in the menu bar and then ‘Title’, a pop-up window ‘Insert Title’ will appear where you can give the title for the map and click “ok” for the title to appear on the map.

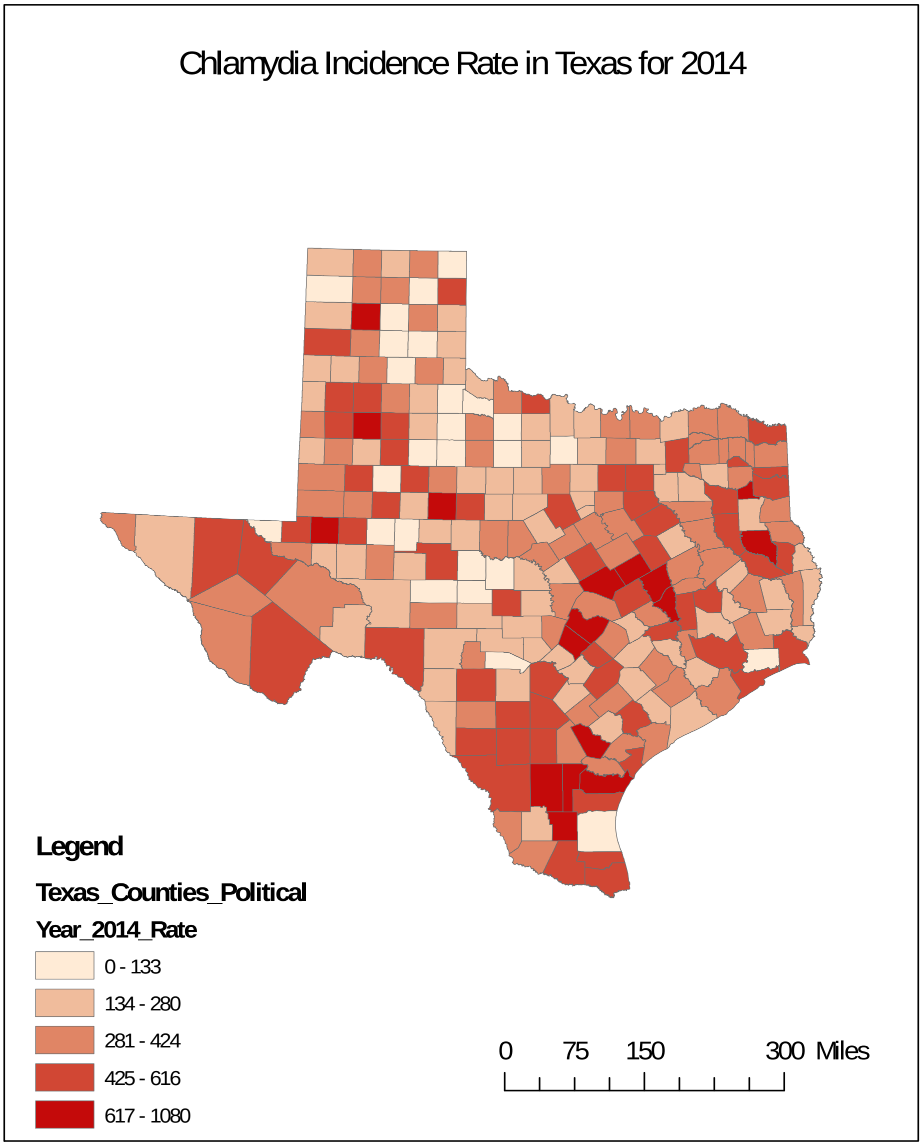
For inserting the legend, click “Insert” in the menu bar and then ‘Legend’, a pop-up window ‘Legend Wizard’ will appear and then click ‘next’ by keeping all the defaults in the next four windows and then click finish. The legend that appear can be moved and placed at the bottom left corner of the map.

For inserting the scale, click “Insert” in the menu bar and then ‘Scale Bar’, a pop-up window ‘Scale Bar Selector’ will appear and then click on the scale type of your choice and click ‘ok’. The scale bar that appears can be moved and placed at the bottom right corner of the map.

Once you have finished your map, then click on “Edit” in the menu bar and then select “Copy Map to Clipboard” and then paste it either in a Word document or a Powerpoint. Your map should look like the map shown here. This is the map showing the total number of Chlamydia incidence cases in 2014 for each county.



Following the same steps that you used earlier, use the "Part2_Chlamydia_TX_Rate_2014.csv" data to create a map of the Rate of incidence. Chlamydia incidence rate is defined as the number of chlamydia cases per 100,000 people.



Each student will be assigned to map the disease data for one selected year following the procedures above. Based on your map, answer the following questions:

1. Map the Chlamydia disease incidence data

a) Which geographical regions of the Texas shows highest counts for chlamydia cases for a given year?

b) Which geographical regions of the Texas shows highest chlamydia rate for a given year?

c) Explain why the chlamydia incidence is mostly concentrated in certain geographical region of Texas.

2. Map the gonorrhea disease incidence data
   a) Which geographical regions of the Texas shows highest counts for gonorrhea cases for a given year?
   b) Which geographical regions of the Texas shows highest gonorrhea rate for a given year?
   c) Explain why the gonorrhea incidence is mostly concentrated in certain geographical region of Texas.

All the answers will be compiled and discussed in the following class.
Then, you will develop a poster, an info graph, or a presentation based on the geospatial analysis and submit a research report.

Part 3: Examining the correlation between the different social factors and the disease incidence

Purpose
In this exercise, you will use the excel data provided to explore the correlation between the different social factors, such as teen birth rate, annual violent crimes, children in poverty, primary care physicians, number of uninsured people, number of households with severe problems with the number of chlamydia cases and gonorrhea cases in each county of Texas.

Analyze the data and explain why certain factors are more correlated while others are less correlated to the disease data. You can do this by conducting a correlation analysis.

The county level health ranking data along with a wide variety of social factors can be downloaded from http://www.countyhealthrankings.org/rankings/data/tx.

You will also explore the relationship between the number of chlamydia and the number of gonorrhea cases during each year among the counties within state of Texas. Based on your analysis, answer the following questions:

1. Correlate the Chlamydia disease incidence data to the different social factors
   a) Which factors among the given social factors are more correlated to the reported cases of chlamydia incidence?

2. Correlate the gonorrhea disease incidence data to the different social factors

a) Which factors among the given social factors are more correlated to the reported cases of gonorrhea incidence?

Part 4: Application of spatial disease ecology to monitor diseases in wildlife

Purpose
In this exercise, you will explore and apply the broader knowledge of disease ecology and spread by applying the knowledge you gained to a non-human ecological system. Apply the knowledge of the GIS and spatial epidemiology that you learned for the surveillance of one of the selected wildlife disease of your choice.

Unlike the human diseases, there are many unique challenges involved in studying diseases in wildlife populations, such as poorly delineated population boundaries and difficulties finding animals to sample. However, spatial epidemiology and GIS have proven to be valuable tools for mapping wildlife populations, analyzing disease clusters and identifying environmental predictors of disease.

1. Apply the knowledge of the GIS and spatial epidemiology that you learned in the above exercises to the surveillance of one of the selected wildlife diseases of your choice. Choose the wildlife disease of your choice by going through the websites given below and explain how you will monitor and map the spread of the disease in a given study area?

Websites

Wildlife Diseases in Texas
https://tpwd.texas.gov/huntwild/wild/diseases/

US National Wildlife Health Center
https://www.nwhc.usgs.gov/

Fish and Wildlife Disease
https://www.usgs.gov/science/mission-areas/ecosystems/fish-wildlife-disease

NOTES TO FACULTY

The goal of this teaching exercise is to introduce the students to a project-based learning approach of mapping the disease dynamics using the Geographic Information Systems (GIS). Project Based Learning (PBL) adopts a multidisciplinary approach using real-world problems, such as actual disease incidence as in the case of this exercise, in bringing together knowledge and skills. The faculty can design the appropriate course materials providing the flexibility for a move away from the traditional teaching in large classes to a more student-centered teaching and learning environment.

PBL is an individual or group activity that goes on over a period of time, resulting in a product, presentation, or performance. It typically has a timeline and milestones, and other aspects of formative evaluation as the project proceeds. PBL for mapping the disease incidence can be introduced as a lecture and then as hands-on lab activity. The lecture consists of an introduction to the characteristics, causes, and symptoms of STD prevalence, disease ecology, and influence of different social factors affecting the geographical spread of the disease. The laboratory curriculum consists of carefully selected and designed problems that demand from the learner acquisition of critical knowledge, problem-solving proficiency, self-directed learning strategies, and team participation skills.

Students are given open-ended projects or problems with more than one approach or answer, intended to simulate professional situations. The PBL learning approaches are defined as student-centered and include the teacher in the role of facilitator or coach. Advantages of PBL include the encouragement of student initiative, self-directedness, inventiveness, and independence. However, the PBL demands from students a heightened level of self confidence, motivation, and ability to organize their own work plans.

The dataset and GIS exercise are introduced to the students after the completion of the lectures on the bacterial and microbial infections, Sexual Transmitted Diseases (STDs) and Epidemiology. Also, the students are introduced to the concept of GIS and the applications of GIS in various fields of study including the biological sciences. By the end of the course, students will develop a comprehensive understanding of the health data and its interpretation at the local and regional scales during multiple years.
The exercise modules given are divided into Parts 1 to 4. Depending on the available class time, class size, class type, student level, availability of the computers, software, and the speed of the internet connection, the faculty can decide which parts of the exercise to use or whether to use all the exercise modules in sequence. The faculty can use the ESRI ArcGIS software as shown in the student worksheets or can also use any other open source GIS software such as Quantum GIS (QGIS) to complete these exercises.

Part 1: Downloading, analyzing, building and Interpreting the STD datasets

The "Part 1: Downloading, analyzing, building and interpreting the STD datasets" will help the students explore the different disease datasets available on the CDC website. This module is suitable for one class period. This includes chlamydia, gonorrhea, HIV, TB, viral hepatitis and different stages of syphilis. The instructor can make the students download the datasets at the national, state or county level within each state. The students then can download the data based on multiple years, sex, age, and geography.

The students can also analyze the different online interactive maps of the data at https://gis.cdc.gov/GRASP/NCHHSTPAtlas/charts.html or interactive geographic maps of the data at https://gis.cdc.gov/GRASP/NCHHSTPAtlas/maps.html or can download the tabular data at https://gis.cdc.gov/GRASP/NCHHSTPAtlas/tables.html. The students will gain knowledge of the different graph formats, understand the spatial distribution of the data and the different downloadable dataset formats. The choice of this particular topic was guided by the desire to show the students in the biology major, the importance of the CDC datasets in public health, and also to show them the relative importance of STD diseases compared to other diseases.

The instructor can either assign each student to observe, analyze and interpret the disease data of a certain geographical area by assigning the student a particular set of counties or states to analyze and report, so that the students can compare and present their results at the end of the class. If the time does not permit, the instructor can make the entire class to do the same exercise. The part 1 exercise introduces the students to the CDC interactive website with all the disease data. The students can explore and build knowledge based on their interest of a specific disease or geographical region of choice, such as researching and finding the information about their home states or home counties. Students can also form hypotheses about which demographic and social factors influence the disease incidence in their local geographical regions.

The instructor can choose a variety of questions based on the online datasets using the interactive website https://gis.cdc.gov/GRASP/NCHHSTPAtlas/charts.html such as

1. What is the rate of chlamydia incidence among the age group of 20-24 during the year of 2015 in the United States?
2. Which age group among the females has the highest number of chlamydia cases reported in Texas for the year 2014?
3. Describe the trend of gonorrhea disease incidence in Texas during the period of 2000 to 2014?
4. Describe the trend of chlamydia disease incidence in Harris County of Texas during the period of 2000 to 2014?

5. What is the rate of chlamydia incidence among the age group of 20-24 among Black African women during the year of 2015 in the United States?

The instructor can choose a variety of questions based on the online datasets using the interactive website https://gis.cdc.gov/GRASP/NCHHSTPAtlas/maps.html such as

1. Describe the geographical spread of chlamydia incidence among the age group of 20-24 during the year of 2015 in the United States?
2. Describe the geographical spread of gonorrhea incidence among the age group of 20-24 among white during the year of 2015 in the state of Texas?
3. Submit and explain the map of the geographical spread of gonorrhea incidence among all the age groups during the year of 2014 in the United States?
4. Submit the change over time graph of the gonorrhea incidence in the United States and explain which states are having high incidence rate over time?

Part 2: Developing the geospatial database and mapping the disease incidence

The "Part 2: Developing the geospatial database and mapping the disease incidence" will help the students download the text data files of different years for the chosen geographical region. This module is suitable for one to two class periods. The students will gain knowledge on the different dataset formats and in preparing the geospatial datasets from the given excel data. This exercise will help the students realize the missing data points within the years and understand the importance of the quality of the data. The choice of this topic was guided by the desire to show biology students the importance of the geospatial science in public health, and also to show them the relative importance of knowledge pertaining to the geographical spread of the STD diseases.

The structured exercise can be completed in a one-hour lab period, if the students are already familiar with GIS software, where the students can download the excel dataset of chlamydia and gonorrhea incidence in Texas and develop spatial maps of the data. The exercise can be broken down into two lab sessions, if the students are being introduced to the GIS software or the students can do chlamydia disease in one session and gonorrhea data in the other. Anything not completed within the lab period can be completed as homework. The instructor can assign small groups of students or to a project to extract and prepare geospatial datasets for different geographical regions, or may be different counties within a single state or even different time periods for the same region. The instructor can ask the students to synthesize the data at the end of the assigned period of study, which can extend from single week to multiple weeks, where the students can present their data to the class as a poster, info graph, or an in class presentation.

The instructor can choose a variety of questions based on the GIS exercise, such as

1. Provide the geospatial map of the chlamydia count in Texas for the Year 2014?
2. Provide the geospatial map of the rate of incidence of gonorrhea in Texas for the Year 2014 and describe which counties show higher rate of incidence and which counties shows the lowest rate of incidence?
3. How does the map of chlamydia count and chlamydia rate of disease incidence vary in Texas for the year of 2013?
4. Overlay the map of chlamydia incidence on the top of gonorrhea incidence? And explain how the disease incidence of both these vary?

Part 3: Examining the correlation between the different social factors and the disease incidence

The "Part 3: Examining the correlation between the different social factors and the disease incidence" will help the students to compare and correlate the disease data with several other social factors. This module is suitable for one class period. The students will learn about the impact of different social factors on the STD diseases. Also the students will map the data and calculate the correlation coefficients between the different parameters.

The county-level health ranking data along with a wide variety of social factors can be downloaded from http://www.countyhealthrankings.org/rankings/data/tx. The instructor can suggest the students explore the influence of teen birth rate, annual violent crimes, children in poverty, primary care physicians, and the number of uninsured people on the STD incidence. The data can be downloaded yearly for each state of your choice. Specific questions can be developed where the students can download the data from their home states or states assigned by the instructor in the follow up labs.
The instructor can choose a variety of questions based on the exercise, such as

1. What is the correlation between the number of primary care physicians available and the chlamydia cases recorded in each county in Texas for the year of 2014? Show the correlation in the graph.

2. What is the correlation between the number of children in poverty in the county and the chlamydia cases recorded in each county in Texas for the year of 2014? Show the correlation in the graph.
3. How does the correlation hold when you compare the rate of chlamydia incidence with rate of poverty? Explain?

The PBL approach requires extra involvement and time commitment from faculty or instructors. The faculty role is very important at the design stage of the problem. Strong guidance is needed on how to tackle project work in order to reduce the likelihood of students attempting to undertake overly ambitious projects or an overly simple project. Project specifications should be more detailed for the students. Careful piloting and testing of proposed projects should be undertaken in advance of the presentation of the research results to establish a reasonable estimate of time required for successful completion of the project by the student or the group of students. Sample projects should be provided in the lab, like the chlamydia and gonorrhea disease data mapping of this exercise, to show students the scope of project and what they should achieve at the end of the project. It should be recognized that extra demands are made upon faculty both in terms of personal involvement and of time commitment in evaluating or assessing projects.

Part 4: Application of spatial disease ecology to monitor diseases in wildlife

The “Part 4: Application of spatial disease ecology to monitor diseases in wildlife” will help the students explore and apply the broader knowledge of disease ecology to a non-human ecological system. This module is suitable for one class period. The students will apply the knowledge gained in the disease data mining, spatial mapping of the data, and correlation between different data parameters in exercises of parts 1-3 to monitor the wildlife disease patterns and spread. The instructor can suggest and direct the students to explore the different wildlife diseases such as Avian Influenza, Chronic Wasting Disease, White Nose Syndrome, Anthrax and others.

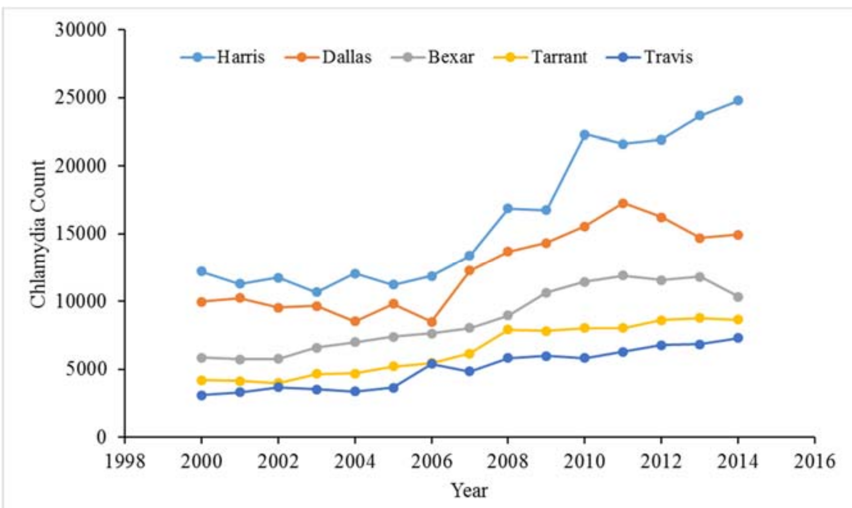
Assessment

The lab reports, answers to the questions embedded in the lab exercises and in the examinations, and answers to the quizzes will serve as assessable outcomes for each of the exercise modules. A research proposal, research report, info graph or poster presentation, or an oral presentation at the end of the project will serve as an assessable outcome for the project, if a PBL approach is implemented wherein all the exercise modules are utilized and an individual or group project is assigned for the students. The students will be learning and

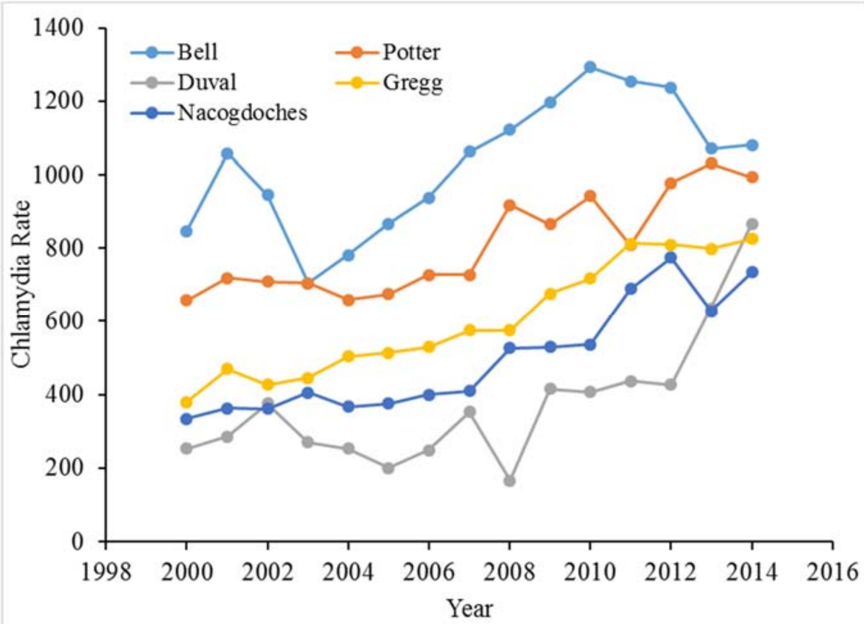
demonstrating the data analysis skills using excel and other databases, map making and spatial querying using the GIS software, communication skills using Powerpoint, poster or info graph making.
Answers to embedded lab questions:

Part 1: Downloading and Accessing the Sexually Transmitted Diseases (STDs) Data

1. Chlamydia disease incidence Data
   a) Graph the time series data of chlamydia disease counts for five counties with the highest incidence.



   b) Graph the time series data of chlamydia disease rate for five counties with the highest incidence.



c) Describe the trend of disease incidence and disease rate in these five counties over the last 15 years?

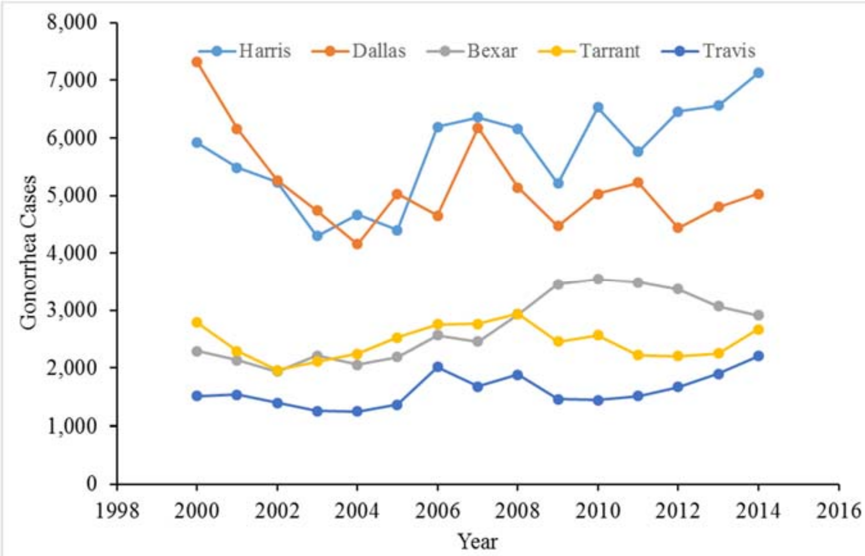
The disease counts in Harris County are showing an increasing trend over the last decade compared to all other counties.

Bell, Potter and Gregg Counties are showing an increasing rate of chlamydia. Bell County rate is slowly decreasing in the last couple of years.

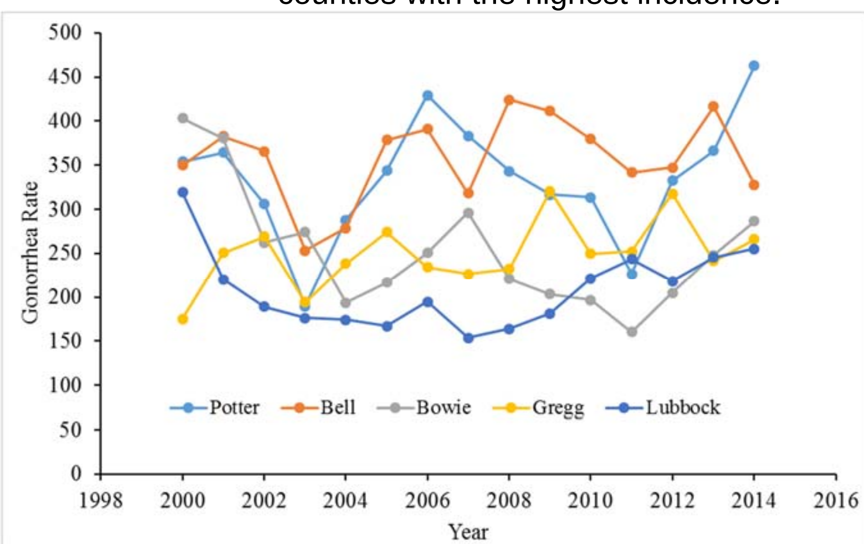
d) Why are the disease counts always higher in particular counties over the others?

The disease counts are higher in counties with a larger population, as you are looking at a bigger population size. Most of the counties within a big metropolitan region show higher count, but when you calculate the disease rate, it will be lower because of the higher population size.

2. Gonorrhea disease incidence Data
   a) Graph the time series data of gonorrhea disease counts for five counties with the highest incidence.



b) Graph the time series data of gonorrhea disease rate for five counties with the highest incidence.



c) Describe the trend of disease incidence and disease rate in these five counties over the last 15 years?

The disease counts in Harris County is showing an increasing trend over the last decade compared to other counties. Potter and Lubbock Counties are showing an increasing rate of gonorrhea.

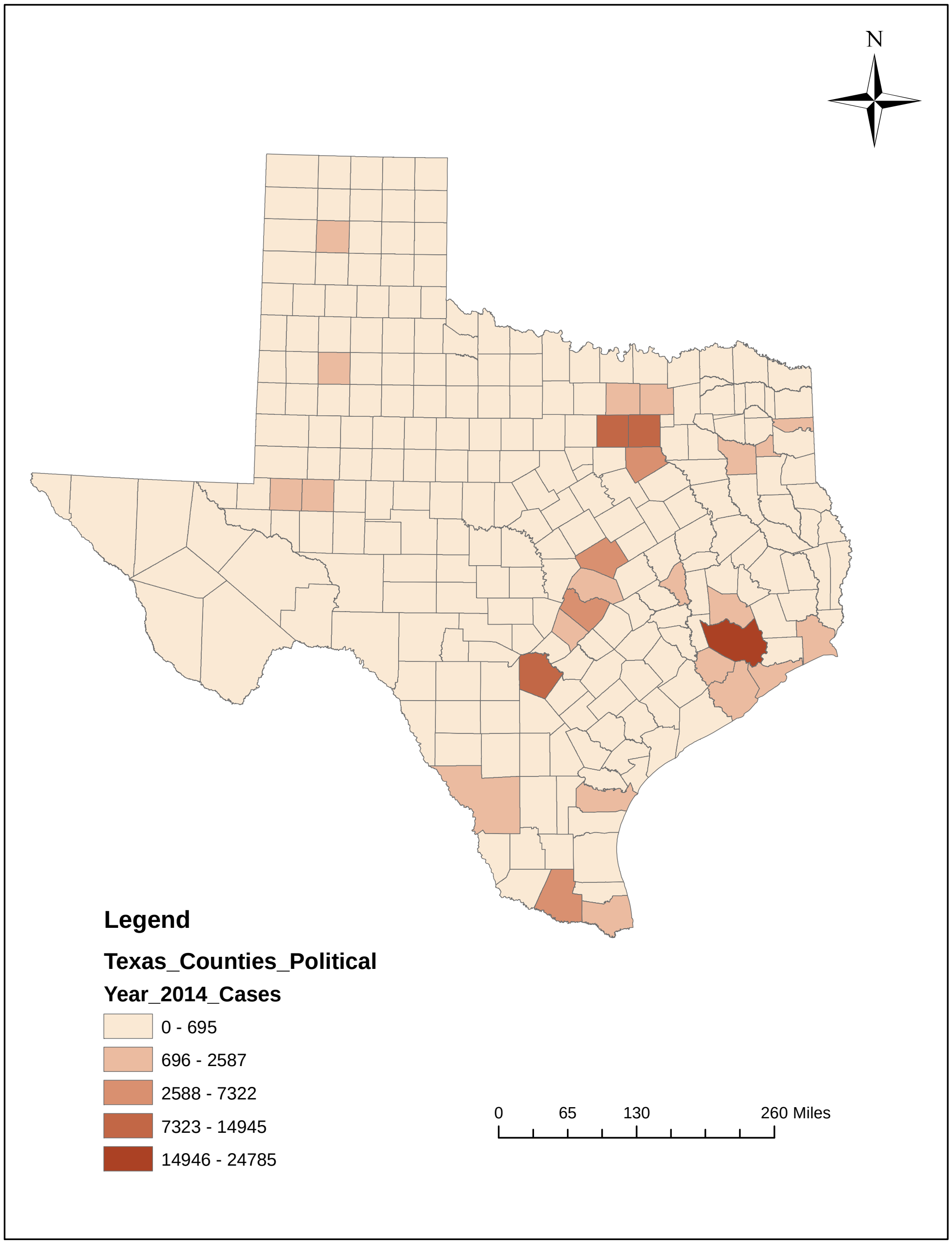
d) Why are the disease counts always higher in particular Counties over the others?

The Harris County is the home to the Houston Metropolitan area, which is the largest city in Texas with the largest population. Hence, the disease counts are higher because of higher population.

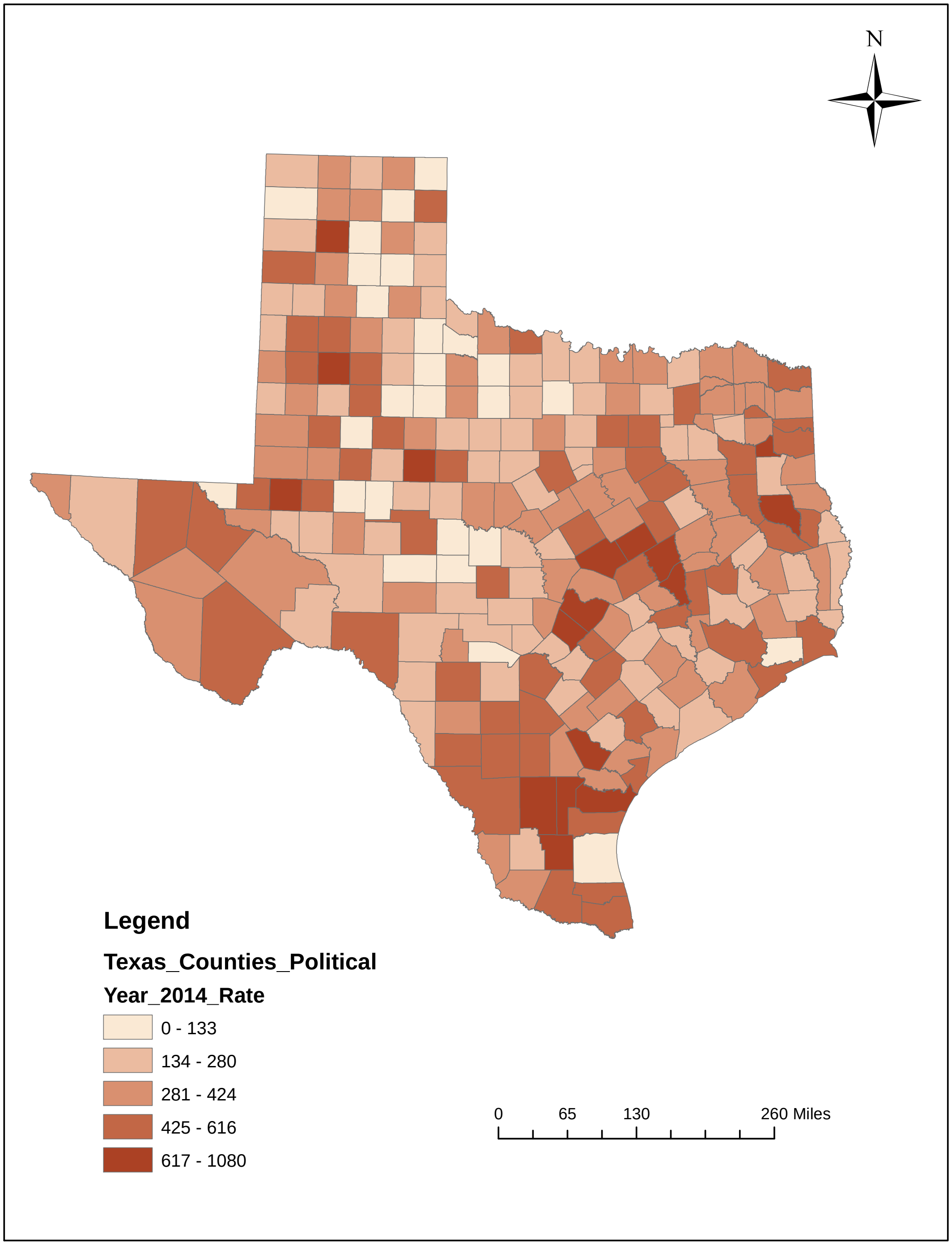
Part 2: Developing the geospatial database and mapping the disease incidence

1) Map the Chlamydia disease incidence Data

a) Which geographical regions of the Texas shows highest Chlamydia cases for (a given year) 2014.



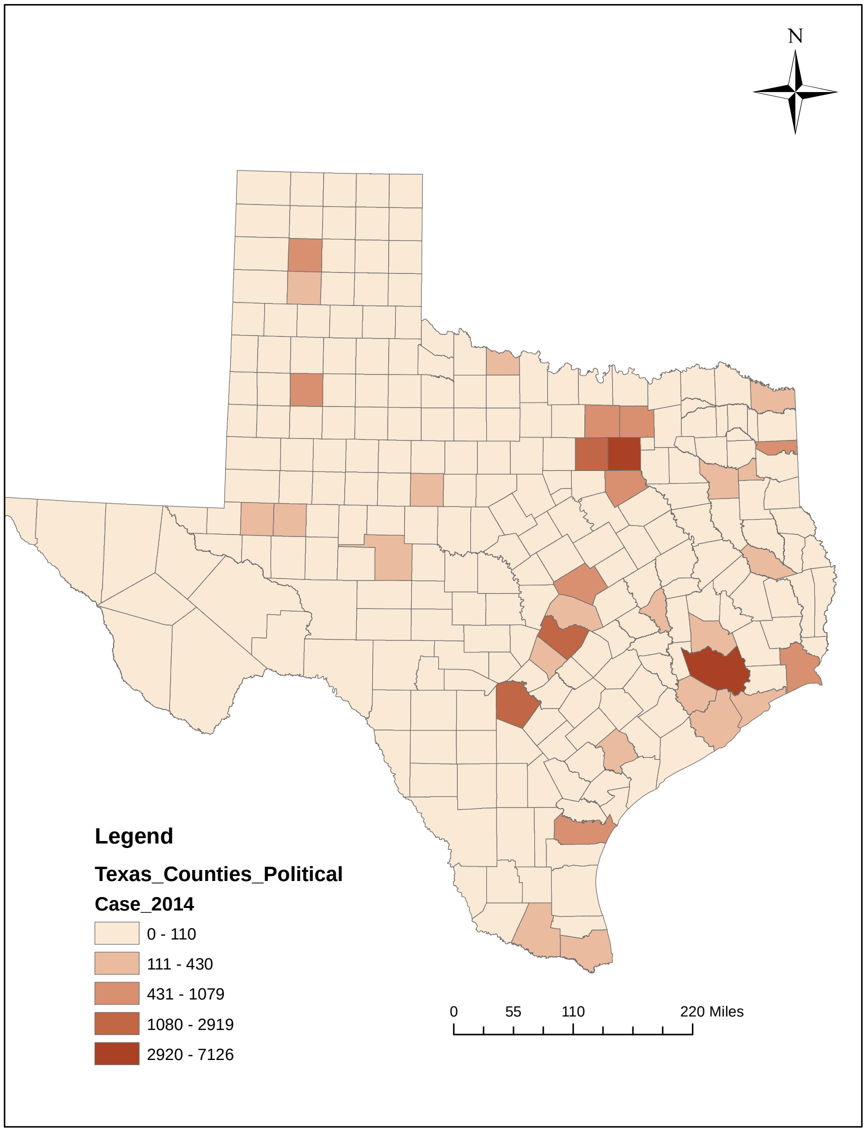
b) Which geographical regions of the Texas shows highest Chlamydia rate for (a given year) 2014.


N
Legend
Texas_Counties_Political
Year_2014_Rate
0 - 133
134 - 280
281 - 424
425 - 616
617 - 1080
0
65
130
260 Miles

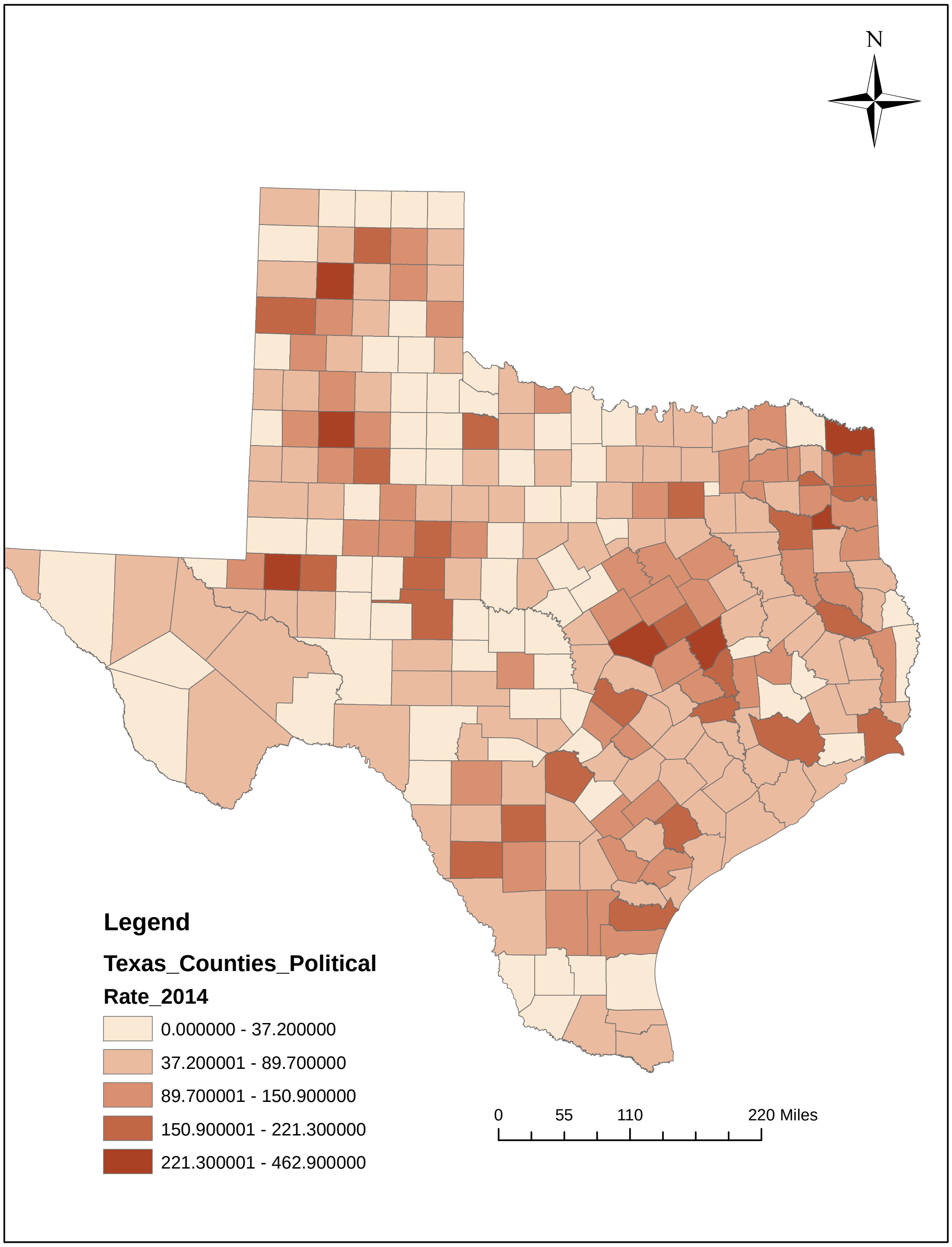
c) Describe why the chlamydia incidence is mostly concentrated in certain geographical region of Texas.

Most of the chlamydia incidence is located in the south and south east counties of the Texas compared to the north and western parts of the state.

2) Map the gonorrhea disease incidence Data
   a) Which geographical regions of the Texas shows highest gonorrhea cases for (a given year) 2014.



b) Which geographical regions of the Texas shows highest gonorrhea rate for a for (a given year) 2014.

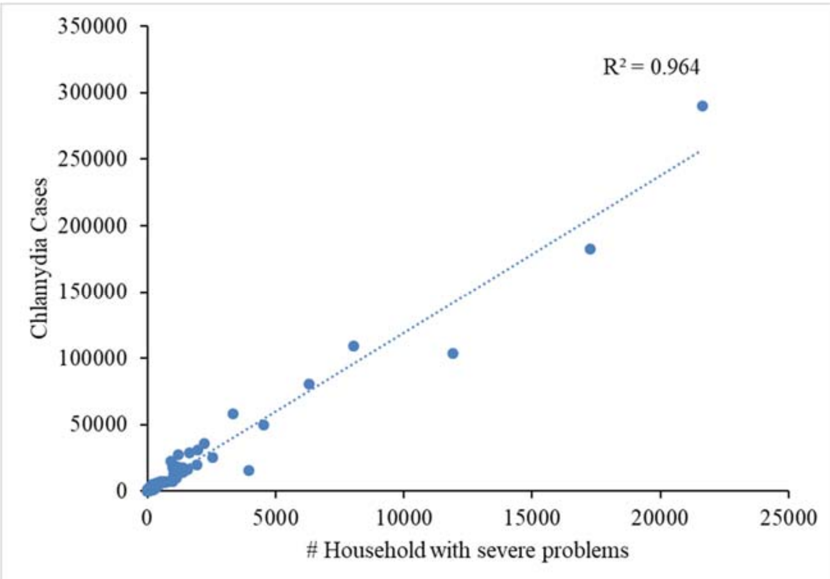


c) Explain why the gonorrhea incidence is mostly concentrated in certain geographical region of Texas.

The gonorrhea cases are higher in the Eastern and the southeastern part of the Texas and the cases are low in the west and southwest parts of the state. The possible reasons for the higher incidence are that the eastern region is the home for the big cities, with high job growth and with younger population and all those may be contributing for the disease spread in this region.

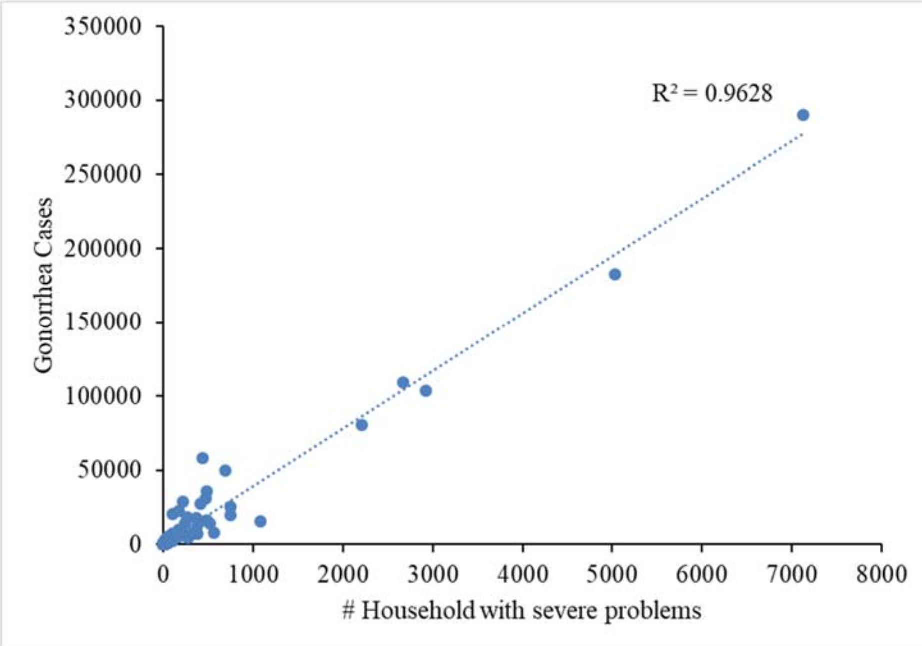
Part 3: Examining the correlation between the different social factors and the disease incidence

1) Correlate the Chlamydia disease incidence data to the different social factors
   a) Which factors among the given social factors are more correlated to the reported cases of chlamydia incidence?



The correlation between the numbers of household with severe problems to the number of chlamydia cases is stronger for the year 2014 compared to all the other social factors.

2) Correlate the gonorrhea disease incidence data to the different social factors
   a) Which factors among the given social factors are more correlated to the reported cases of gonorrhea incidence?



The correlation between the numbers of household with severe problems to the number of gonorrhea cases is stronger for the year 2014 compared to all the other social factors.

Part 4: Application of spatial disease ecology to monitor diseases in wildlife

1) Apply the knowledge of the GIS and spatial epidemiology that you learned in the above exercises for the surveillance of one of the selected wildlife disease of your choice. Choose the wildlife disease of your choice by going through the websites given below and explain how you will monitor and map the spread of the disease in a given study area?

The answers will vary depending on the choice of the disease the student will select to monitor and study.

Guidelines for Research Report:

Format for a scientific research report

1. Title:

- A concise and informative title of not more than 10-15 words.
- The title should be self-explanatory and easy to understand for the general audience.
- The title should be specific enough to describe the contents of the paper, but not so technical that only specialists can understand.

2. Keywords:

Provide 5-6 keywords

3. Abstract:

The abstract should be broken down into different components. It should include the following:

1. ***Introduction and Purpose*****:** One to two sentences explaining why you studied this particular topic and what is significant about it. How does this research add to existing knowledge?
2. ***Materials and Methods*****:** One to two sentences outlining the methods you used to conduct the research. How was the data collected and processed?
3. ***Results:*** Three to four sentences about what you found in the research.
4. ***Interpretation:*** Three to four sentences discussing the results and what does they mean?
5. ***Conclusion:*** One sentence summarizing what was learned and why it is significant

- Your abstract should be one paragraph, of 200-250 words, which summarizes the purpose, methods, results and conclusions of the paper.
- Don't use abbreviations or citations in the abstract.

4. Introduction:

- What question did you ask in your experiment? Why is it interesting?
- The introduction should summarize the relevant literature so that the reader can understand the back ground and why the research is needed to be done. Three to four paragraphs should be enough.
- Introduction should conclude with a sentence explaining the specific research question that needs to be answered.
- It should then lead into the hypothesis that will be tested and the objectives that are framed to accomplish the research goals.

5. Materials and Methods:

- This section chronologically describes the instruments, software and the process that you undertook to complete the research.
- Describe the study area along with a map or figure
- Describe the datasets that were used, from where they were downloaded, how they are processed along with the data analysis software that was used.
- Describe the laboratory methods and the statistical methods that were utilized.

6. Results:

- The Results section describes but does not interpret the major findings of your study.
- Present the data using graphs and tables to reveal any trends that were found. Describe these trends to the reader. The presentation of data should be chronological, to correspond with the Materials and Methods that was explained.
- Organize logically and use headers to emphasize the ordered sections.
- Illustrate and summarize the findings by organizing the data and emphasizing the trends and patterns with appropriate visuals. Integrate visuals with text so that the reader can better understand.
- Decide on a logical order that tells a clear story, that makes it easy to understand. Generally, this will be in the same order as presented in the methods section.
- Do not include any references in this section as you are presenting your results, so you cannot refer to others here.

7. Discussion:

- Explain the main findings of the study. Evaluate, analyze, and explain the significance and implications of the study, principles and concepts that was supported or rejected.
- Discuss the strengths and limitations of the study. Explain key limitations such as questions that left unanswered, major experimental constraints, lack of correlation, and negative results, if any.
- Discuss the main results with reference to previous research reported. Discuss agreement or contrast with any previously published work.
- Discuss the policy and practical implications of the results.
- Recommend areas for future study and explain your choices.

8. Conclusions:

- Do not reiterate the data or discussion.
- Provide clear scientific justification for the research study.

9. Acknowledgements:

- Provide the names of people who contributed to the work, but did not contribute sufficiently to earn authorship.
- Acknowledge the funding sources supported to conduct the study.

10. References:

Provide complete citations for any articles or websites or reports or book chapters or any other materials referenced in the text.

Research Report Grading Rubric

Component	**Outstanding**	**Good**	**Average**	**Below Average**	**Score**
Title	Descriptive of the question and work performed. Clear, concise and informative title.	Gives a general description but is partially clear and concise.	Gives a general description but is neither clear nor concise.	Lack of general description of the work performed.	5 points
Keywords	Clearly relevant keywords	Partially relevant keywords	Slightly relevant keywords	Poorly relevant keywords	5 points
Abstract	States clearly the question being asked and the hypothesis being tested. States the methodology, major findings and conclusions.	Missing at least one component of good abstract.	Missing at least two components of good abstract.	Missing at least three component of good abstract.	15 points
Introduction	Strong introduction of research topic, questions and relevant terms. Specific thesis statement.	Convey research topic, questions and relevant terms. General thesis statement.	Convey research topic, but not questions and relevant terms. Vague thesis statement.	Does not adequately convey research topic, questions. Poor thesis statement.	10 points
Materials and Methods	Describes clearly how the study was performed. Gives logical,	Most steps are understandable but some are confusing.	Most steps are understandable but some lack details and are confusing.	Most steps are not clearly understandable and lack details and are confusing.	10 points

	detailed and easy to follow steps.				
Results	Results described clearly in narrative as well as in graph or table formats along with captions. Data presented in logical manner to draw conclusions.	Missing at least one component of good results. Data is described clearly but have errors in graphs and tables.	Missing at least two components of good results. Data is presented in a confusing manner.	Missing at least three component of good results. Raw and unprocessed data with no clear presentation	15 points
Discussion	Well organized discussion with clear reference to data and current knowledge.	Discussion can be understood but not clear to follow.	Discussion can be understood but difficult to follow.	Lacking several characters of good discussion.	15 points
Conclusion	Strong review of conclusions and strong integration with the thesis statement.	Good review of conclusions and integration with the thesis statement.	Average review of conclusions and integration with the thesis statement.	No clear review of conclusions and no proper integration.	10 points
References	All references and citations are correctly written.	At least one references or citations missing.	At least two references or citations missing.	A number of references or citations missing.	15 points

COPYRIGHT STATEMENT

GENERIC DISCLAIMER

Adult supervision is recommended when performing this lab activity. We also recommend that common sense and proper safety precautions be followed by all participants. No responsibility is implied or taken by the contributing author, the editors of this Volume, nor anyone associated with maintaining the TIEE web site, nor by their academic employers, nor by the Ecological Society of America for anyone who sustains injuries as a result of using the materials or ideas, or performing the procedures put forth at the TIEE web site, or in any printed materials that derive therefrom.