PM 566 Final Project: Analysing Vaccine Preventable Diseases and Vaccine Hesitancy in California

University of Southern California, Fall 2024

Kendall Dimson

Introduction

Long since the rollout of the polio vaccine, there have been a variety of vaccines created in the 1900s to present-day that have been influential in the epidemiological shift, with vaccine efficiency as one of the main modes of primary prevention in infectious disease epidemiology. However, despite continuous efforts to vaccinate, many Vaccine-Preventable diseases are prevalent to this day that threaten the public health system. As a result, in the modern era there have been several vaccine-preventable epidemics that have occurred. According to the Center for Disease Control and Prevention, epidemics are defined as the occurrence of more cases of disease than expected in a given area or among a specific group of people over a particular period of time. There are various factors that can influence the occurrence of these epidemics, including vaccination rates and vaccine hesitancy. This project serves to conduct a comprehensive exploratory analysis into vaccine-preventable diseases in California, with specific focus on fout: Whooping cough (Pertussis), Hepatitis A, Invasive Meningococcal Disease, and COVID-19.

Various research questions include:

Which county had the highest incidence of Pertussis, and when did the peak of the epidemic(s) occur?

Were there any particular epidemics of Hepatitis A outside of Los Angeles county?

What was the statewide trend for Invasive Meningococcal Disease incidence? Was it the same, or different?

How did COVID-19 era vaccine hesitancy impact California counties? Is there an association between vaccine hesitancy and previous Vaccine-Preventable Disease incidence?

Methods

The dataset, vaccine, was extracted through the California Health and Human Services Agency, with data collected from the California Department of Public Health. The data includes the variables: ID, Disease (the disease reported), County (the location the disease was reported from), Year (the year that the disease was reported to CDPH), and count (the cumulative number of cases reported of a given disease in a given year, from the specific county). The dataset includes fifteen Disease categories:

Diphtheria, Invasive Meningococcal Disease, Measles, Pertussis, Tetanus, Mumps, Rubella, Hepatitis A, Acute Hepatitis B, Acute Hepatitis C, Varicella Hospitalizations, Varicella Hospitalization/Death, Chronic Hepatitis B, Haemophilus influenzae, Perinatal Hepatitis B, and Rubella (Congenital Syndrome).

In the data cleaning and wrangling step, counties that recorded 0 cases of a disease were removed. Variety of subsetted datasets were created, subsetting into "california" and "county" datasets since there are some categories in the "County" variable that have a summation number of diseases in California. For each, the filter() function subset. The kable() function was used to print a table of the total number of vaccine preventable diseases reported in the entirety of the dataset. The arrange() function was used to print the top five vaccine-reportable diseases with highest cumulative incidence, as well as to print the top five counties of each disease. Ggplot barplots were used to visualize trends in incidence counts. To analyze the distribution of pertussis cases by counties, I used the leaflet package to code an interactive bubble map, normalized by population density. To visually analyze vaccine hesitancy association, leaflet heat maps were utilized. In the analysis on Pertussis, the data was stratified to compare annual incidence in each California county. In the analysis on Invasive Meningococcal Disease, the data was stratified to visualize bar graphs. All data was analyzed using R software Version 2024.04.2+764.

Results

```
The downloaded binary packages are in
    /var/folders/l4/_n22jzd54wdfdflv_h0vc5sm0000gn/T//RtmpOyot7s/downloaded_packages
```

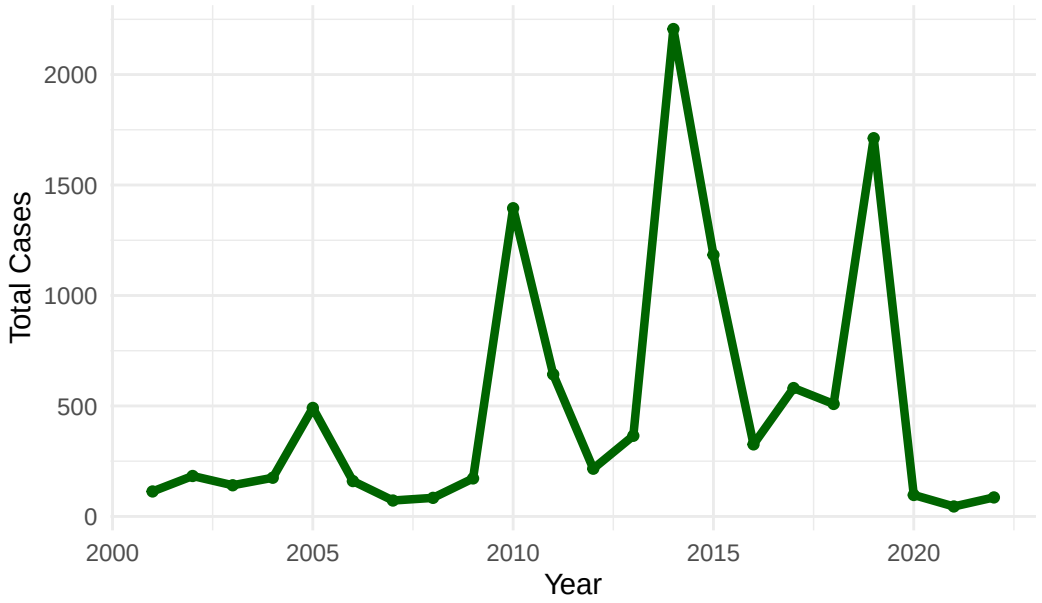
```
The downloaded binary packages are in
    /var/folders/l4/_n22jzd54wdfdflv_h0vc5sm0000gn/T//RtmpOyot7s/downloaded_packages
```

Table 1: Table 1. Total Cases of Vaccine-Preventable Diseases Reported in California, 2001-2022

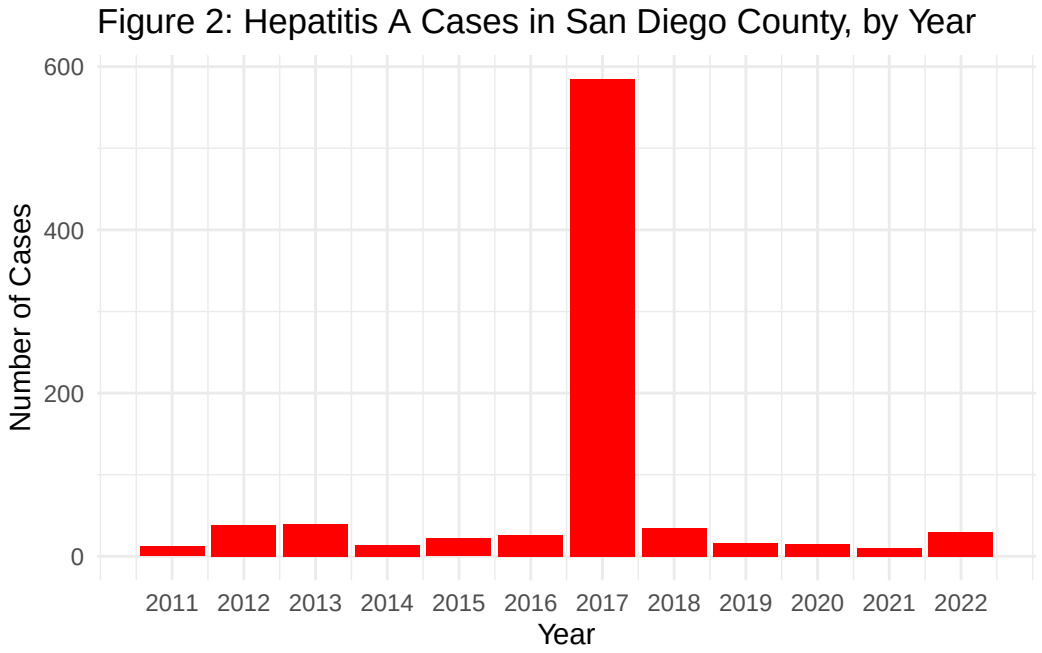
Disease	Total Cases
Pertussis	57219
Hepatitis B, Chronic	4462
Hepatitis A	3035
Invasive Meningococcal Disease	2718
Hepatitis B, Acute	1349
Mumps	1108
Hepatitis C, Acute	1050
Measles	521
Varicella Hospitalizations	406
Varicella Hospitalization/Death	93
Tetanus	86
Rubella	9
Diphtheria	2

There were a total of 72,058 incident vaccine preventable cases recorded. The top five in total cases reported were Pertussis (n=57,219), Hepatitis B, chronic (n=4462) and acute (n=1349), Invasive Meningococcal Disease (n=2718), and Hepatitis A (n=3035).

Figure 1: Trend of Pertussis Cases in Los Angeles per Year



In Figure 1, we investigated the trend of Pertussis cases in Los Angeles each year, the county who had the most cumulative incidence for Pertussis. It was found that Los Angeles county had peaks of case incidence in 2010, 2014, and 2019. After 2020 there was a lower decline of Pertussis cases, but this is likely due to the occurrence of the COVID-19 pandemic.



In Figure 2, it was identified of a Hepatitis A epidemic in San Diego, which greatly stood out from its trend of lower case incidence in recent years. After further research, it was found that the epidemic lasted only one year due to San Diego county's response efforts in vaccination, distribution of hygeine kits, and education.

Figure 3: Invasive Menginococcal Cases in Los Angeles County

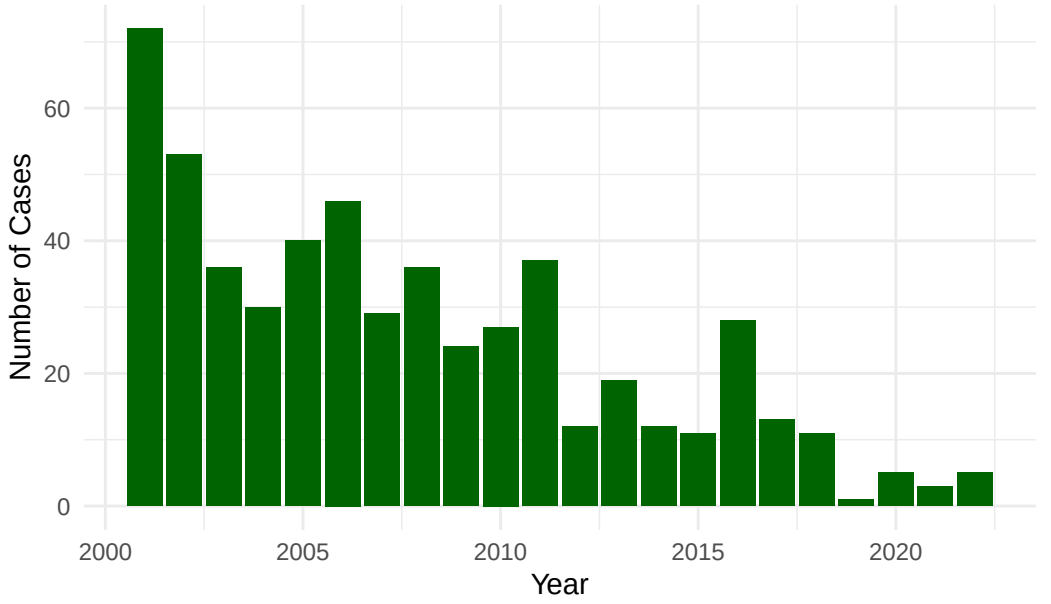
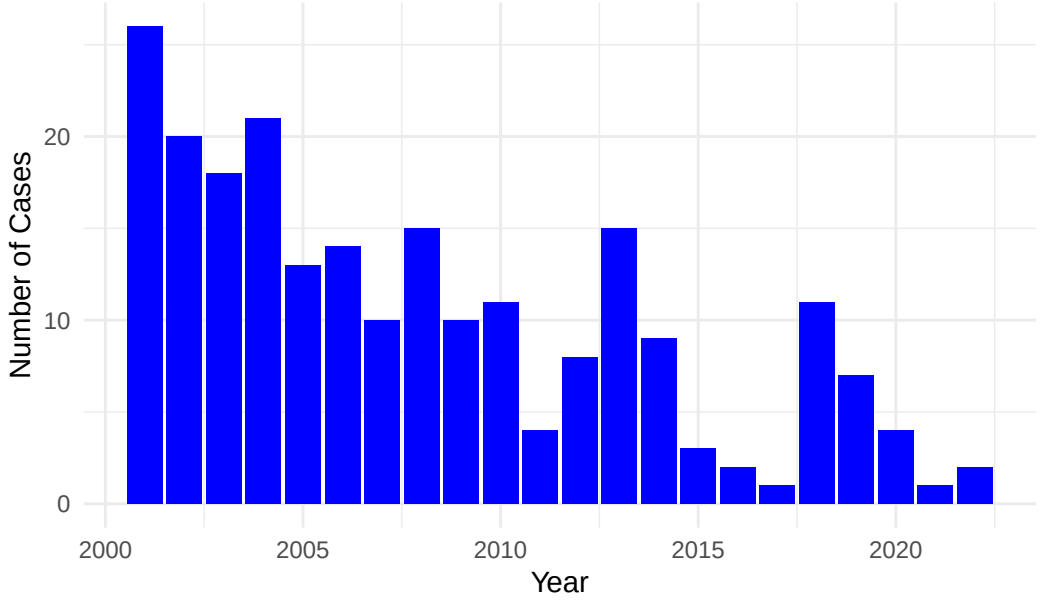
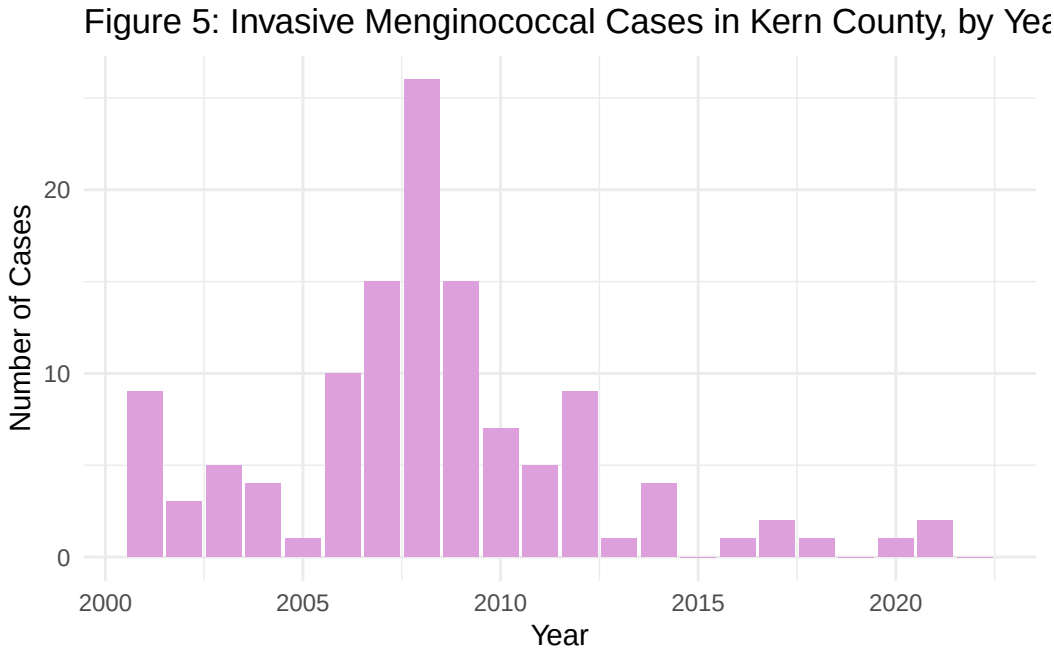


Figure 4: Invasive Menginococcal Cases in San Diego County, I





In exploration of Invasive Meningococcal Disease, a rare but serious bacterial infection, it was first found that the highest incidence cases occurred in Los Angeles (Figure 3), which aligns with the notion that Los Angeles has the largest population size. Los Angeles county reveals a negative trend in Invasive Meningococcal Disease over the years, and this trend also is similar in San Diego County (Figure 4). Interestingly, it was found that Kern County had an epidemic of Invasive Meningococcal Disease, following a unimodal epidemic curve peaking in 2008 with over 25 cases (Figure 5). Likewise to Los Angeles and San Diego counties, Kern has also seen a decrease in case incidence in recent years.

Figure 6. Top 10 Strongly Hesitant Counties − Vaccine

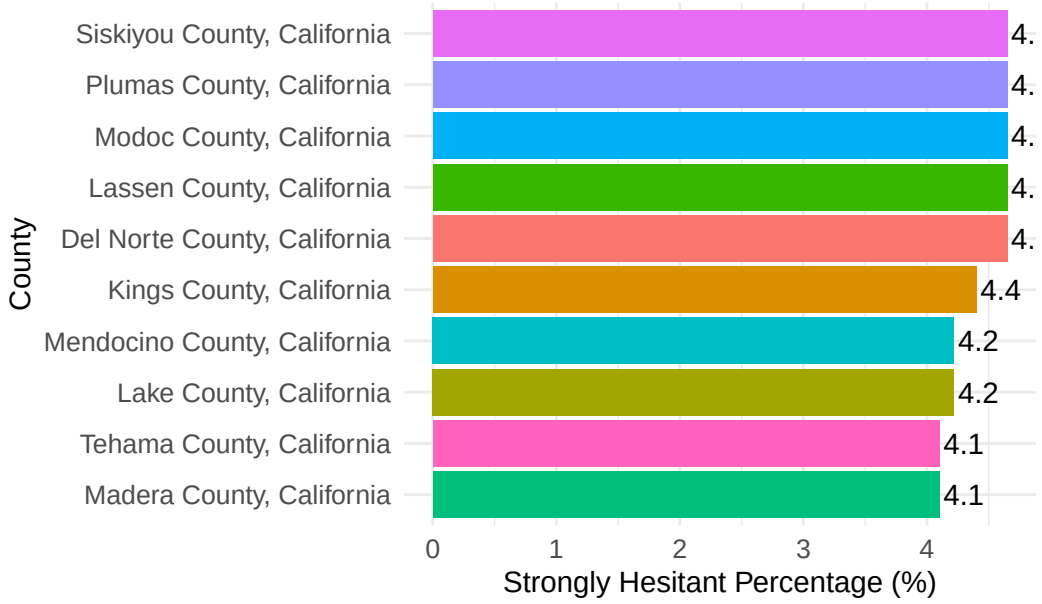
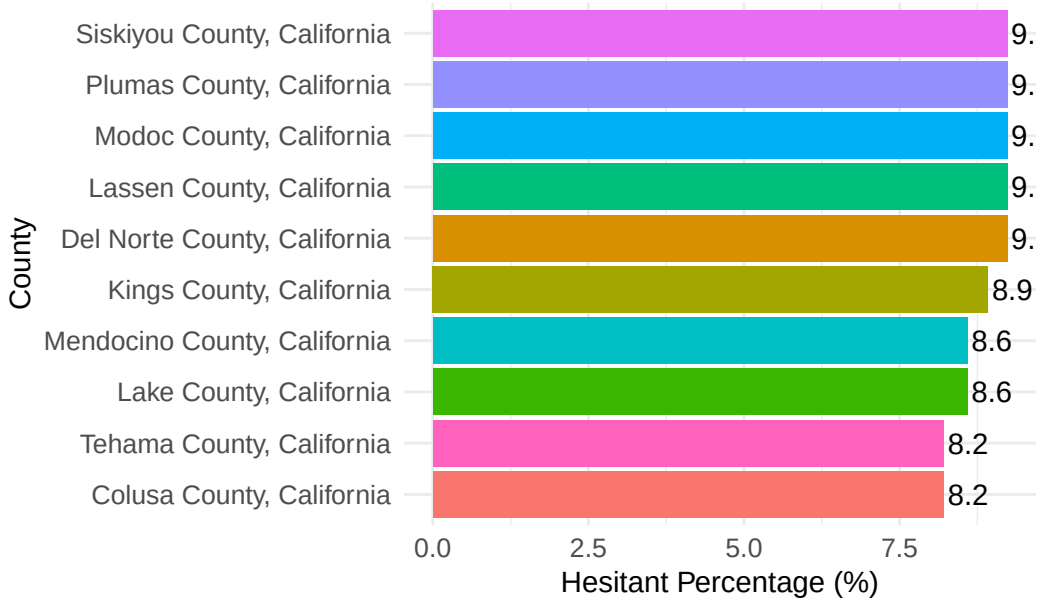
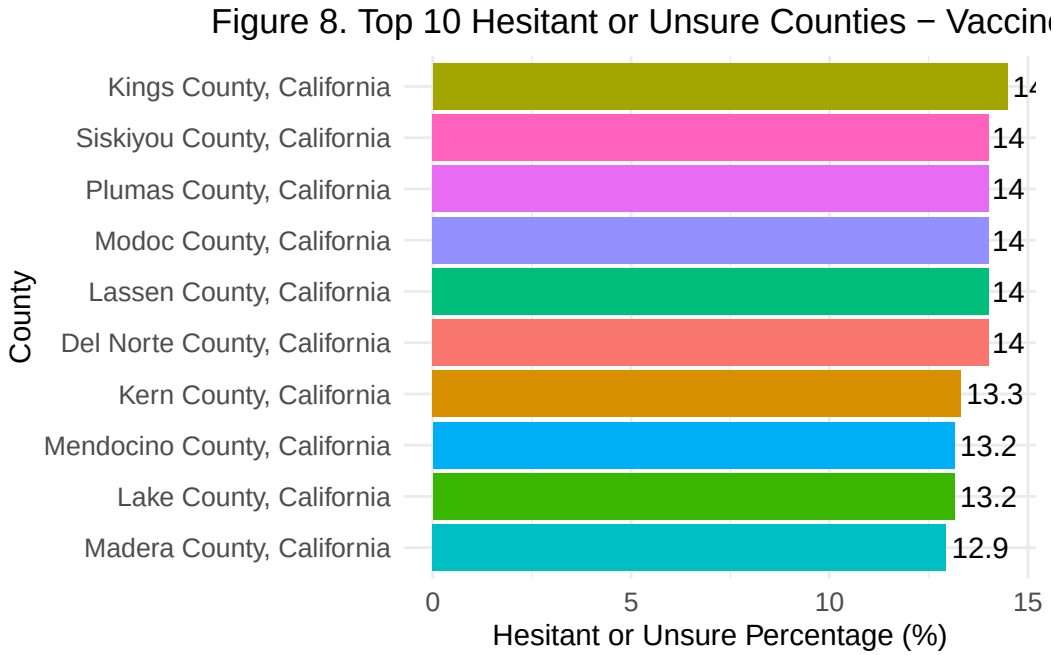


Figure 7. Top 10 Hesitant Counties − Vaccine Hes





Interestingly, in Figures 3, 4 and 5, there was overlap in the California counties that were surveyed with highest percentage of vaccine hesitancy in the state. This included: Siskiyou County, Plumas County, Kings County, and Modoc County as some of the few that stood out in the overlap. However, one of the limitations that was found in this approach is that the percentages of vaccine hesitancy may be too low to capture an appropriate trend nor association. For the survey answer of "strongly hesitant," the highest percentage was in Siskiyou County with 4.7%, which is still a lower percentage. The highest percentage of "hesitant or unsure" in hesitancy was 14% in Kings, Siskiyou, Plumas, Modoc, Lassen, and Del Norte County.

Additionally, another limitation found was in the skewness in the data. Almost none of the counties in the vaccine hesitancy dataset overlap with the Vaccine Preventable Disease incidence dataset. A possibility for this may have been because the statistics of the vaccine preventable diseases focused more in urban, densely-populated counties in California with higher susceptibility of disease transmission. Whereas, the rates of vaccine hesitancy were grouped in rural, less-densely population regions of California.

Conclusions and Summary

In conclusion, it was determined that the highest case incidence of Pertussis occurred in Los Angeles county, with peak epidemics occurring in 2010, 2014, and 2019. This is likely due to the population density of Los Angeles county, but it still provided interesting insight of recent trends in the county.

For Invasive Meningococcal Disease, San Diego and Los Angeles counties were the two with most cumulative case incidence, both having a skewwed right distribution of case incidence decreasing. In contrast, in Kern County, a more rural region of California, had a unimodal distribution of Invasive Meningococcal Disease epidemic occuring in 2008, with the epidemic at the tailend.

COVID-19 vaccine hesitancy was explored, visualizing the counties with the highest percentage of hesitancy responses. However, overall California had a lower percentages of vaccine hesitancy, and the research question is not as particularly relevant to vaccine-preventable diseases since the incident infections occur more in densely populated regions. In future studies, it may be interesting to do a cross-comparison of vaccine hesitancy between California and another dense-populated state, comparing rates with other infectious diseases such as Influenza.

How did COVID-19 era vaccine hesitancy impact California counties? Is there an association between vaccine hesitancy and previous Vaccine-Preventable Disease incidence?