



# **CAPSTONE PROJECT**

# ON THE AGENDA

**01**

INTRODUCTION, BENEFITS, AND FINDINGS

**02**

TOP 5 DISEASES BY TOTAL CASES

**03**

TOTAL DISEASE CASES BY YEAR

**04**

DISEASE CONCENTRATION BY COUNTY

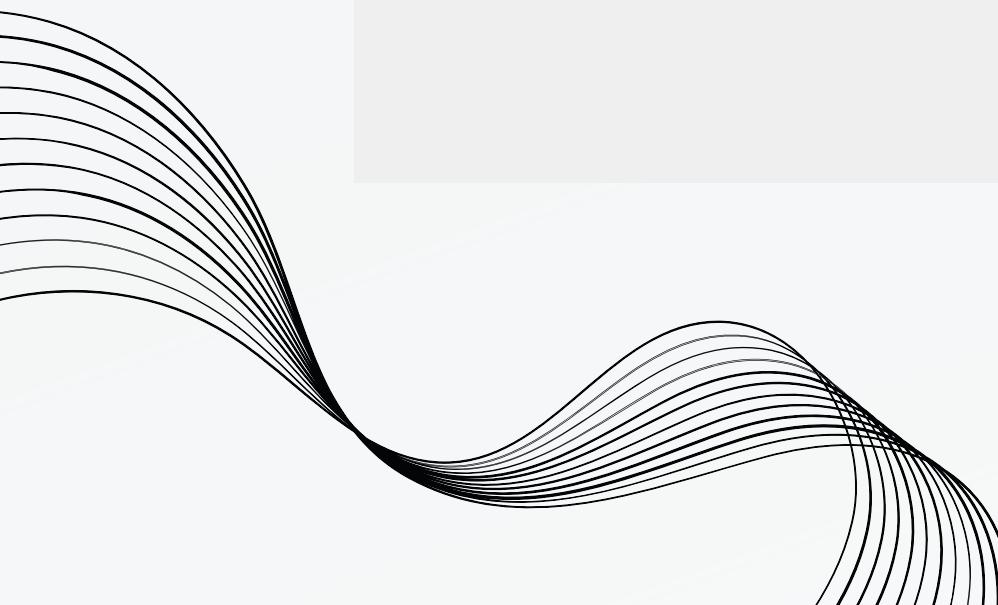
**05**

VALLEY FEVER CASES BY COUNTY

# INTRODUCTION

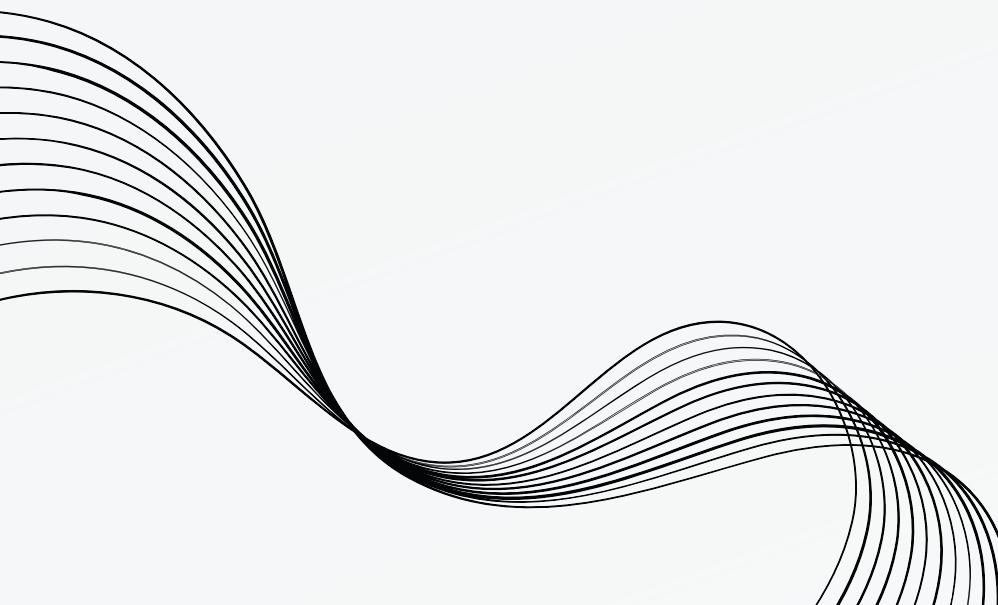
The main objective is to analyze **Infectious Disease Cases by County in California** from 2001 to 2022 and train an MLM to assess the historical trends and try to predict future cases based on County demographics.

Data from <https://data.gov/> and  
<https://www.census.gov/>



# WHO BENEFITS FROM THIS?

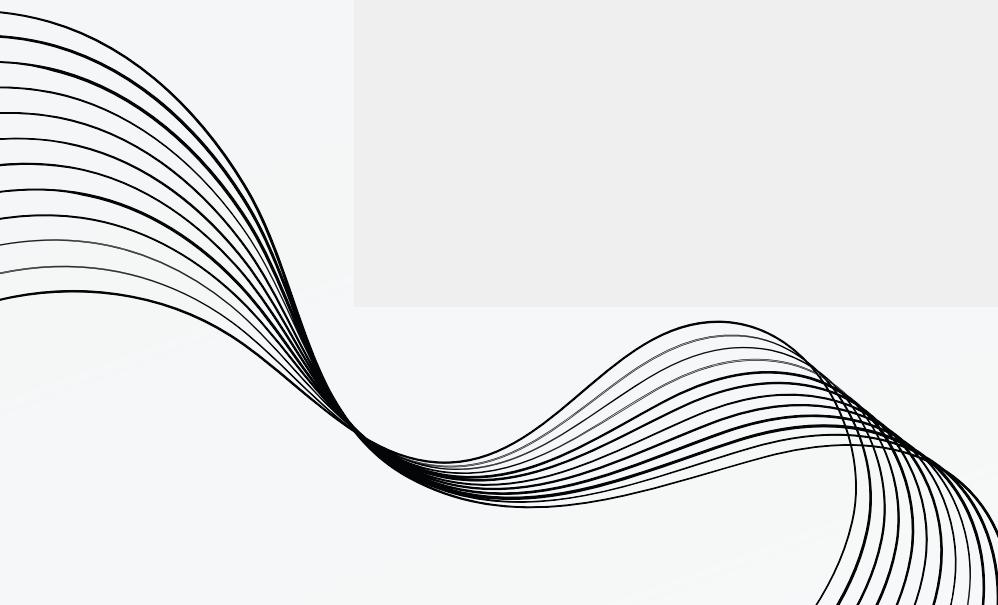
Infectious disease specialists, Doctors, and other medical professionals can utilize a predictive MLM to forecast future cases and use it as a study tool. The visuals can help non-medical professionals to follow along as well.



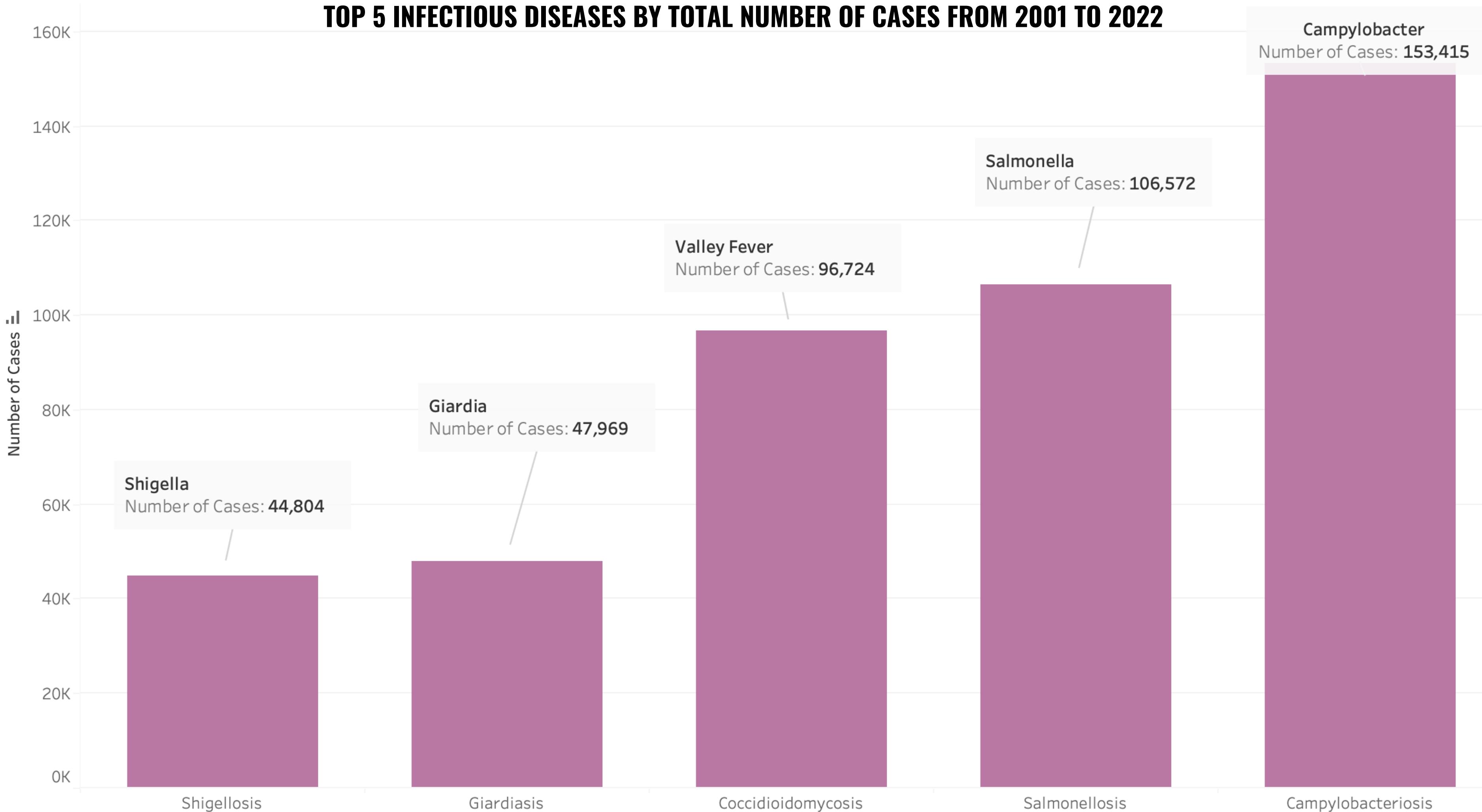
# INITIAL DATA FINDINGS

The disease dataset was well over 180,000 rows but after separating columns and merging with the demographic table, the shape is now 62234 rows x 78 columns.

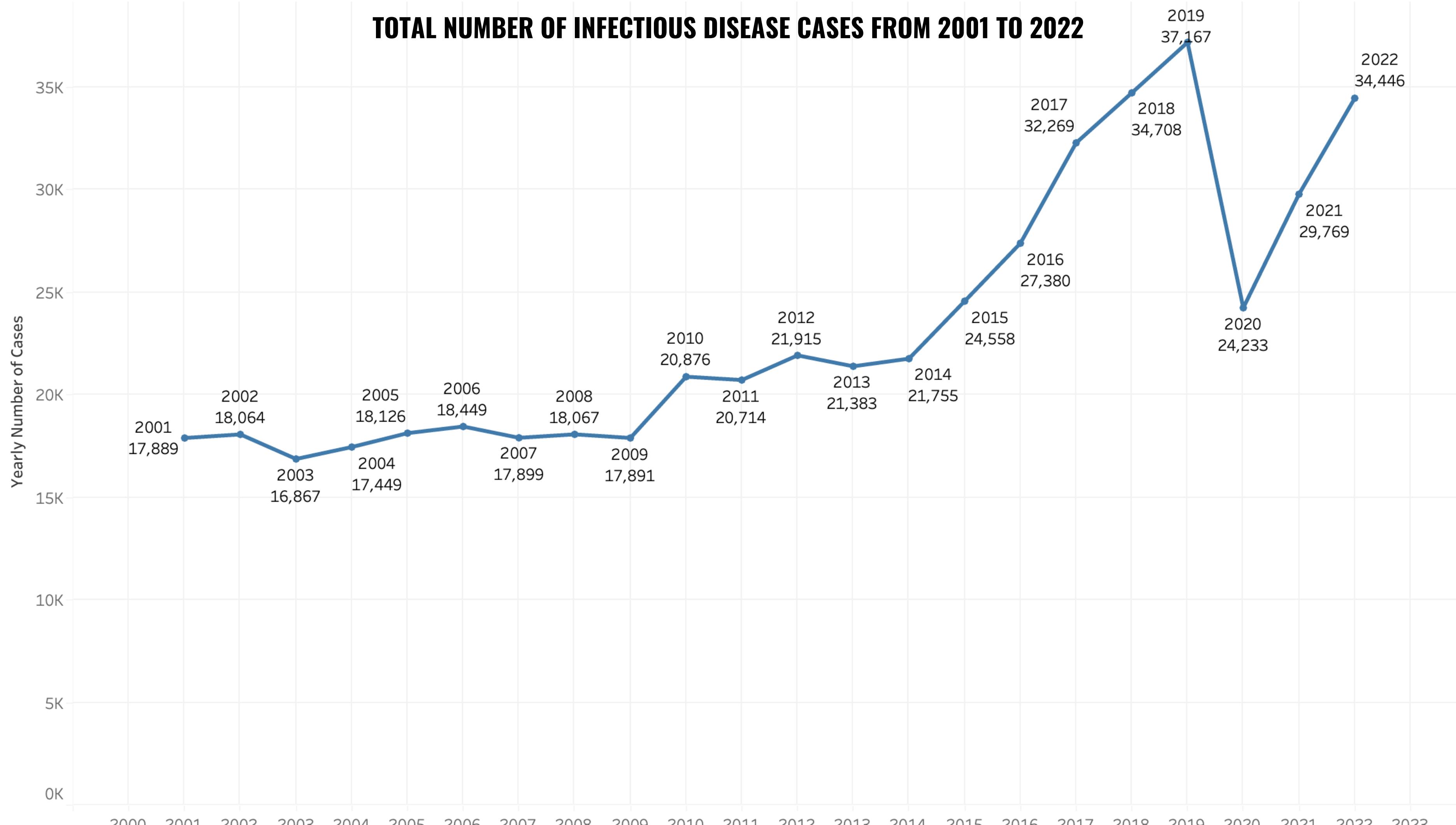
The most challenging part was getting the demographic table to match the size before merging. This was due to the diseases having different testing ranges.



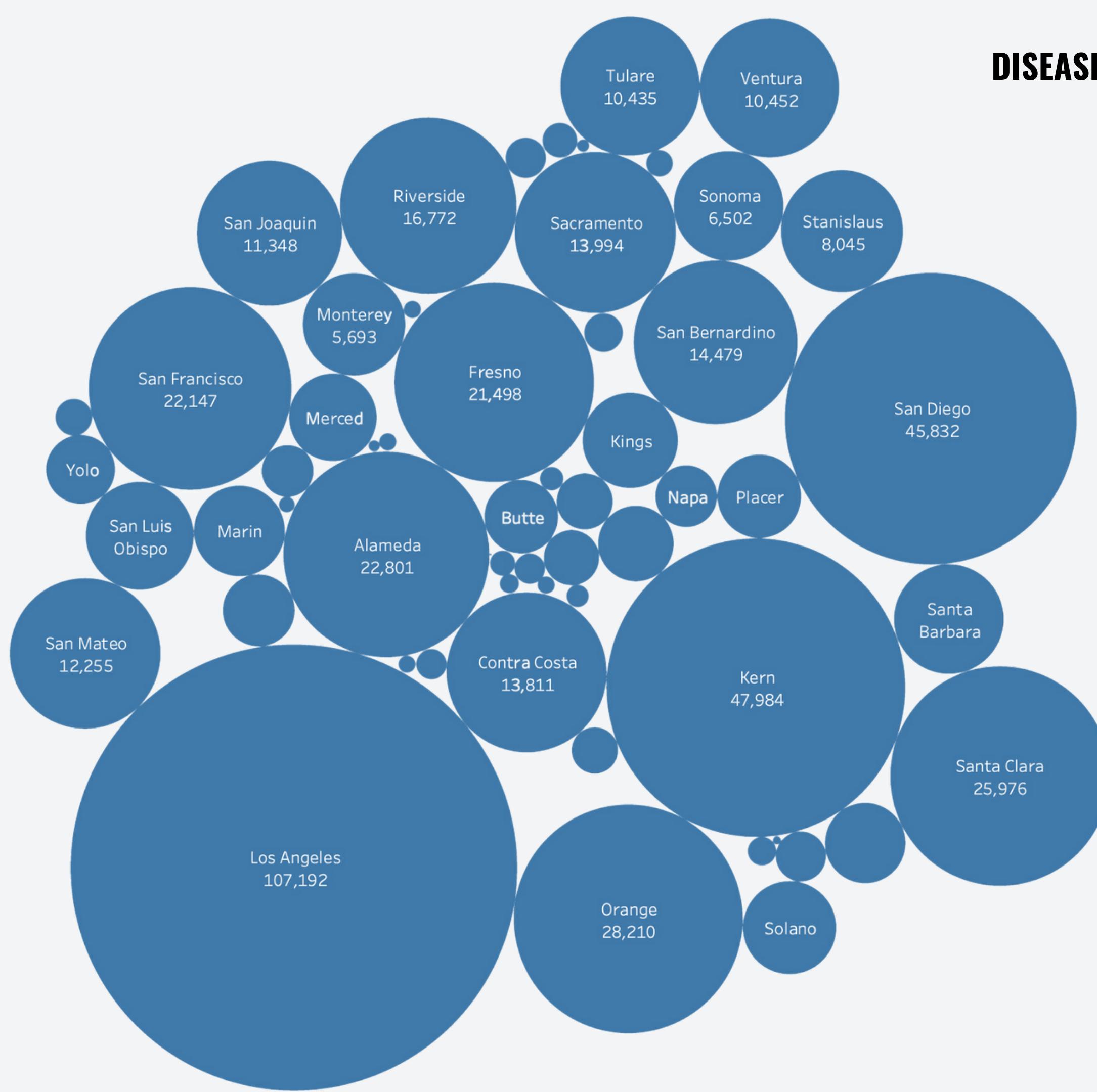
# TOP 5 INFECTIOUS DISEASES BY TOTAL NUMBER OF CASES FROM 2001 TO 2022



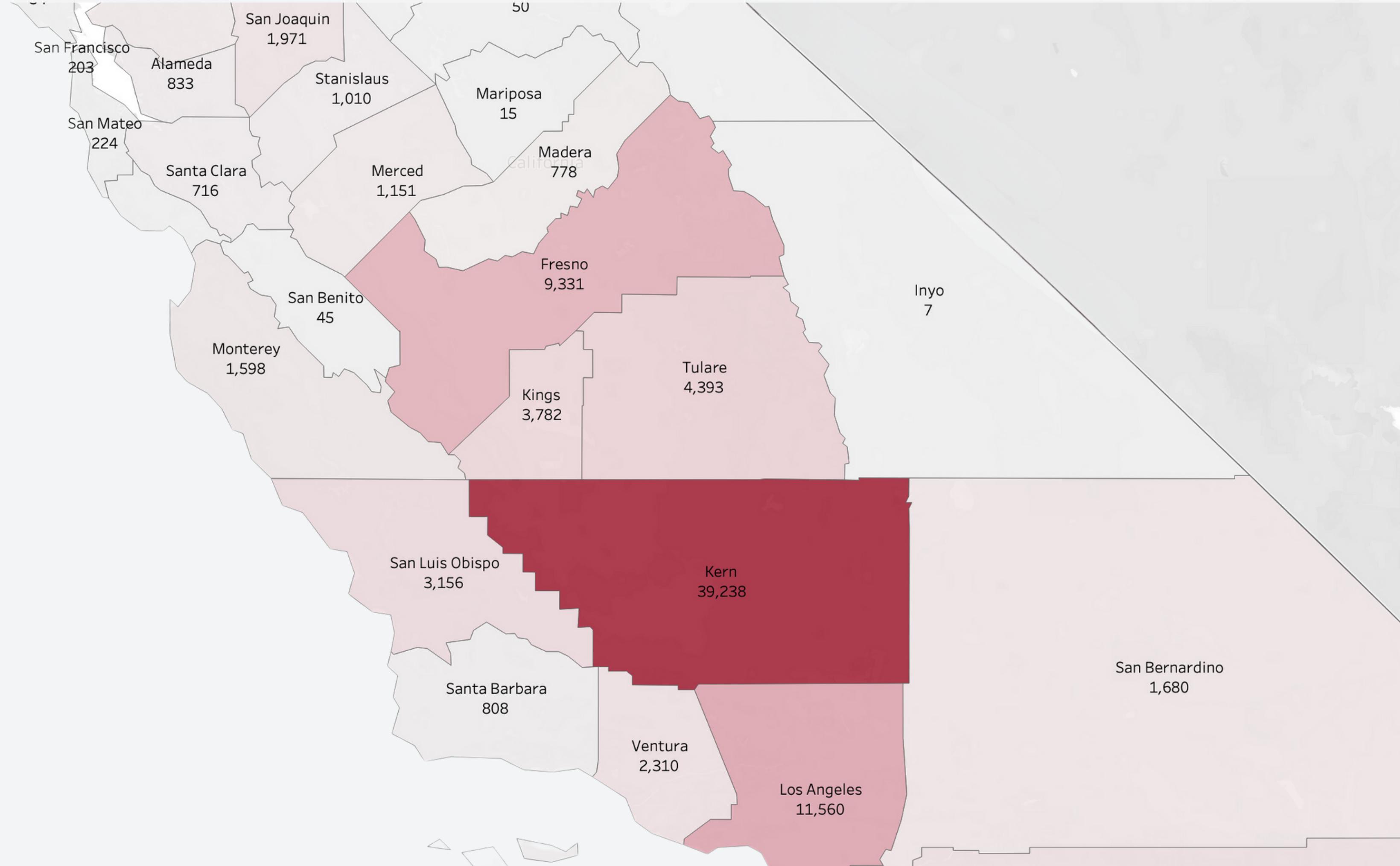
# TOTAL NUMBER OF INFECTIOUS DISEASE CASES FROM 2001 TO 2022



# DISEASE CONCENTRATION BY COUNTY FROM 2001 TO 2022



# VALLEY FEVER CONCENTRATION BY COUNTY FROM 2001 TO 2022



Thanks  
for watching!  
Questions?

