**Slide 1- Rachel to format - DONE**

Title Page

**Slide 2- Rachel to format - DONE**

Mission Statement - Goals of the analysis

**Slide 3- Rachel to format - DONE**

Our data

* Data sets we started with
* How we condensed
  + Reduced Diseases: Measles, Mumps, Pertussis, Tetanus
  + Reduced Counties: Sonoma, Yuba, LA, Alameda
* How we chose the ones we did (scattered ideally to represent similar counties)

**Slide 4- 2 charts total - Rachel - DONE**

General Visualizations

* Graph #1 - Bar chart showing population per county per year
  + X axis: year
  + Y axis: population count
  + Bars: one bar per county for each year
  + Graph #2 - Bar chart showing disease counts per county per year
  + X axis: year
  + Y axis: disease count
  + Bars: same idea, but instead of county it is per disease

**Slide 5 through 8 - 4 graphs total - Clay - DONE**

Analysis - population against disease count over time

* Our guess of what the information will show:
  + With cities that the population increases over time, the disease count increases over time. Is this consistent over small, medium or large counties.
* Limitations: We are using these counties as a representation of other similar sized counties
* Graph: Scatter plot with dots and lines - One graph per disease
  + X axis: year
  + Y axis: number of people
  + Lines:
* Total population of LA
* Disease #1 count in LA
* Repeat for each county & disease

**Slide 9- 4 charts total - Gabbie - DONE**

Analysis- Total Disease count over time- broken down by disease

Question we are asking: is the total disease count consistent over time or does it increase in this county?

* Graph per county - can probably be on the same slide
  + X axis: year
  + Y axis: total disease count
  + Stacked bars - each disease stacks to see total, and also visualize the breakdown of which disease is more common
* Should we show the total population trend above these bars? So that we can see the population trend too?

Is this data too similar to the previous data? The difference is that we can see which disease makes up the majority of the total disease count

**Slide 10- 1 chart - Gabbie**

Question we are asking: Is our analysis representative of other similar sized counties in the state?

Limitation/depth of information: we did not analyze the square milage or population density of each county (making the counties rural vs urban)

* Graphing ratio of disease count per population
  + (ex. Number of disease counts per ten thousand people)
  + X axis: county name (also show categorization of city)
  + Y axis: disease count/population count
  + Stacked bar chart (bar per disease)

**Slide 11- 2 maps - Elise - Combine into beginning of presentation**

Showing the counties visually in map view (the 15 counties)

1st map: highlight county, darker color for higher population (the 15 counties)

2nd map: highlight county, darker color for higher disease count (the 15 counties)

Indicate the 4 that we chose to move forward with

Once we make this slide should this slide be before the others? To show how we narrowed down our data set?

Other maps could be added if we find something could be useful

**Slide 12- Rachel to format - DONE**

List other things we could have considered if we expanded the dataframe:

* Ethnicity
* All counties
* Other similar sized counties in other states
* Age
* Gender
* Poverty level
* Spread of disease
* Availability of vaccines/medical clinics
* Social media influence
* Immigration
* Predictive analysis