Data Immersion
Exercise 2.6
Jordan Novelli

<u>Statistical Visualizations: Scatterplots & Bubble Charts</u>

In the Module Data Immersion Exercise 1.8 there was an excel spreadsheet made for variances, correlation and distribution.

The Correlation Coefficient created for Variable 1 (integrated death and population) was 1. The Correlation Coefficient for Variable 2 (Influenza Visits) was 0.738091643.

The r-squared given by the Tableau trend line is 0.251219. When that number is square rooted the number becomes 0.5012175177.

Because this is a scatterplot of total population and deaths the numerical value from Variable 1 and Tableau are only slightly close but not the same.

_____

To reflect on the scatterplot there is something to notice. As the population increases there is an increase in deaths. Now the population to death ratio is a very small percentage. For example, there are 512 recorded for a population of 38,6952,954. This would mean the death toll is 0.001% of that population. The trend line also increases from the lowest population to the highest population. As the data values cluster around the line there are many that also move away from the line. It's safe to say that there's a moderate relationship but it's not the strongest.

_____

<u>Scatterplot Style Guide Assessment</u>

<u>Text:</u> Count of deaths were added to indicate where the deaths lie per the population. This is overworking the plot. I couldn't figure out how to get the highest deaths per population to show only. I did leave a deaths scale available for interaction. When interacting with that button the scatter points become a lot closer to the trendline.

<u>Color:</u> I put color in the death's variable. This simply makes it easier to interpret some of the deaths from lowest to highest and make the plot a little more interesting.

<u>Other:</u> Be able to interact with the slider to manipulate the death variable is very interesting.

<u>Bubble Chart Style Guide Assessment</u>

<u>Text:</u> No text is added to the chart. The legends to the right-hand side give valuable information. You can interpret that smaller bubbles are used for smaller numbers of deaths while large bubbles are used for larger numbers of deaths recorded.

<u>Color:</u> The color palette used was "Color Blind" to allow all viewers to understand the depiction. The age groups are colored to help indicate which age groups are more affected by the influenza virus.

<u>Other:</u> Not all age groups are included as some of the values are so low that they aren't shown on the chart.

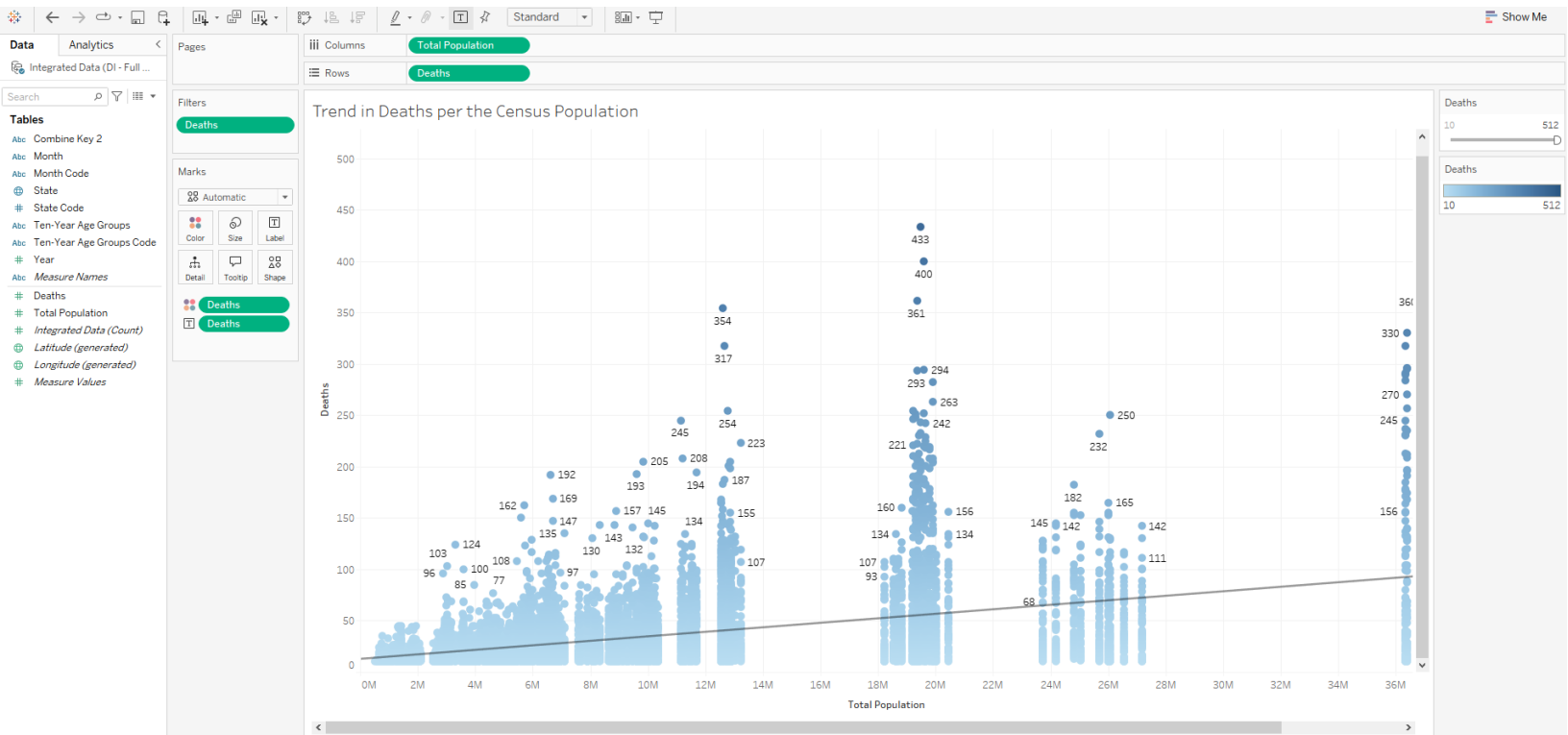Trend in Deaths per the Census Population

Death by Age and Census Population

Tableau Link: Data Immersion 2.6 Scatterplot