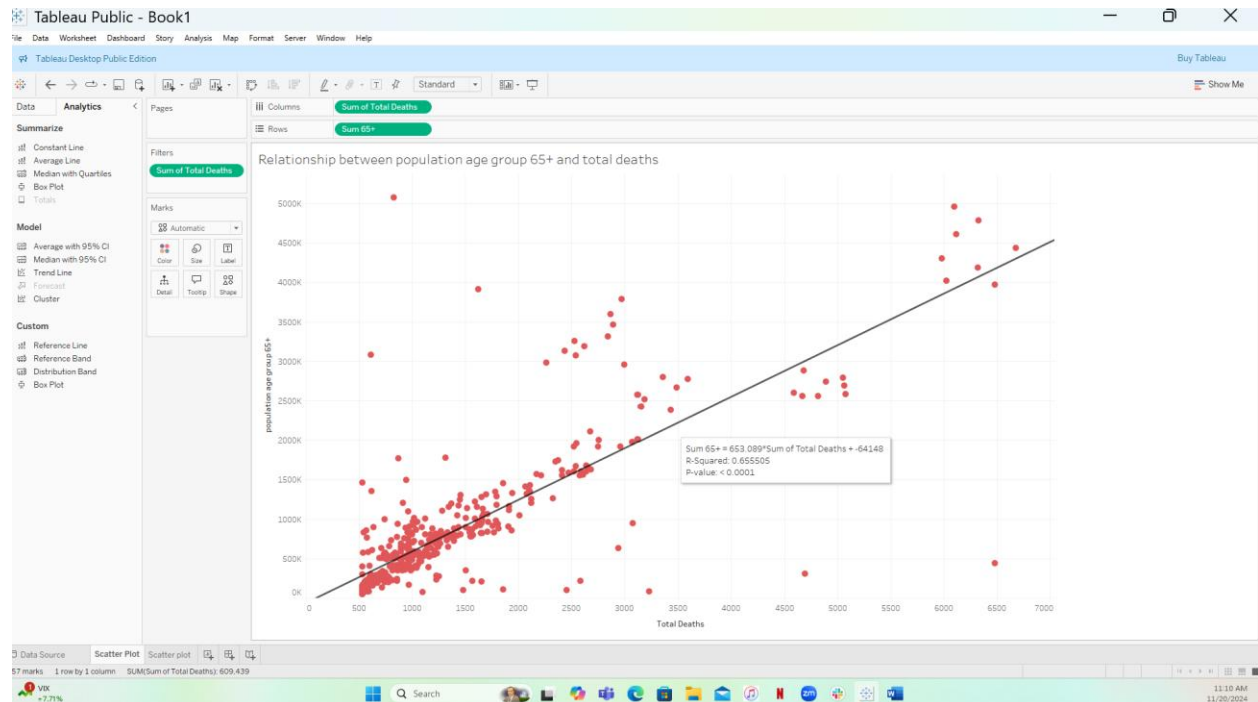


## Data Visualization Exercise 2.6

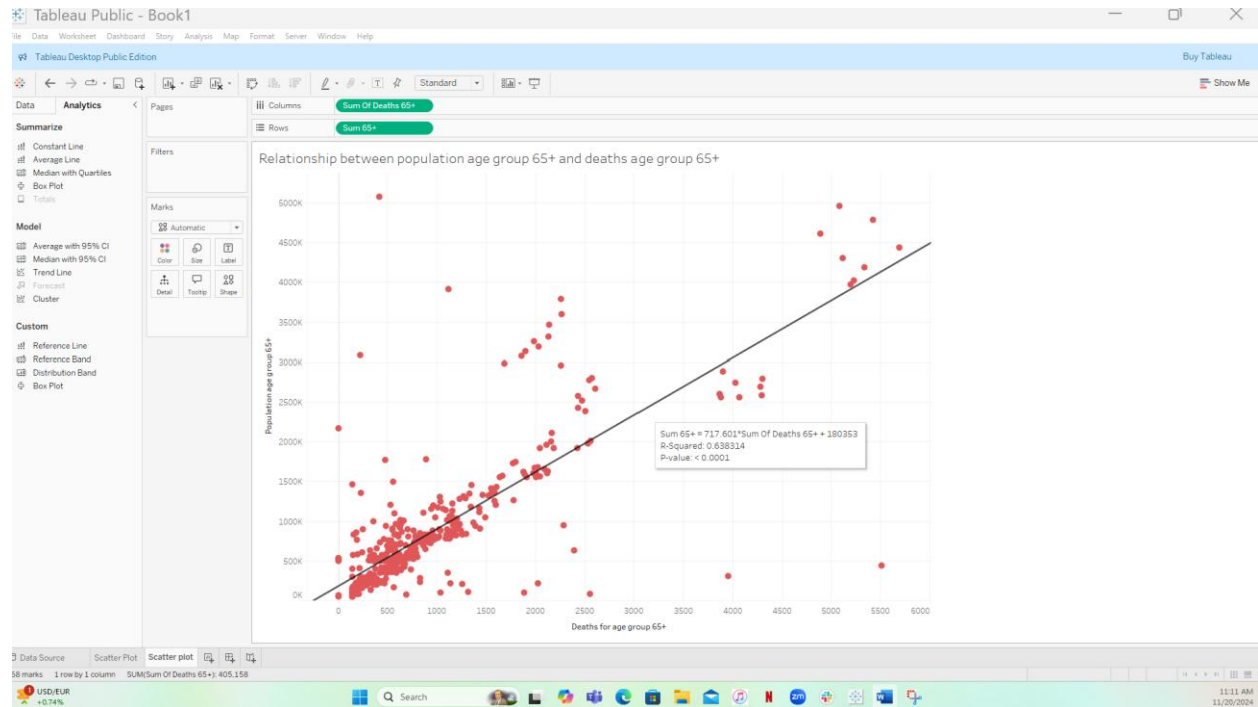
[https://public.tableau.com/app/profile/aaron.gilbert/viz/DataVisualizationExercise2\\_6/Regionbubblechart?publish=yes](https://public.tableau.com/app/profile/aaron.gilbert/viz/DataVisualizationExercise2_6/Regionbubblechart?publish=yes)

### Scatter Plot

**Chart 1:** Relationship between population age group 65+ and total deaths



**Chart 2:** Relationship between population age group 65+ and deaths age group 65+

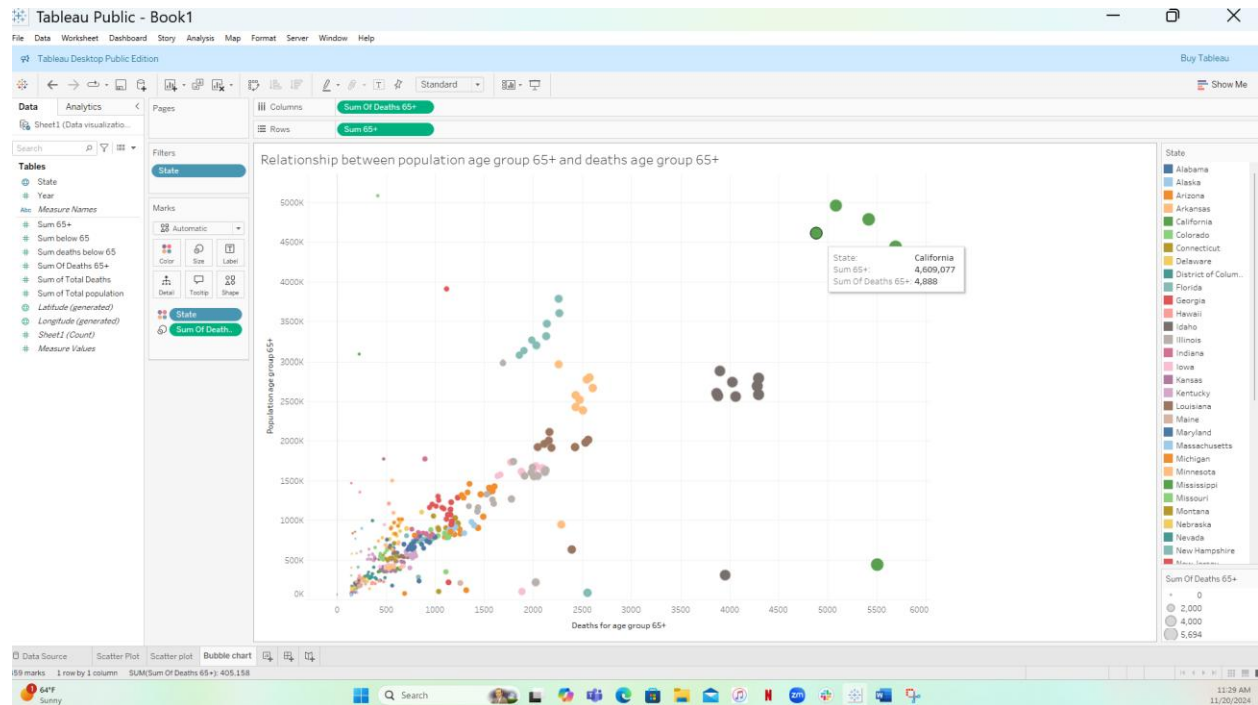


I created two charts; Chart 1 is the relationship between population age group 65+ and total deaths; Chart 2 is the relationship between population age group 65+ and deaths age group 65+

- The trendline's R squared value of chart 1 is 0.655505, which means that the correlation coefficient is 0.8096 (Strong relationship)
- The trendline's R squared value of chart 2 is 0.638314, which means that the correlation coefficient is 0.7989 (Strong relationship)

A scatter plot gives a visual representation of the trend across the whole data. I can see the outliers, the linear relationship or another type like exponential or curve, the strength of correlation of the data across the data. For the charts above, I can see the high correlation between the vulnerable group and death. All the data values are not tightly clustered around the trend line, but there are a lot that are. There are a few extreme values.

## Bubble Chart



For the bubble chart, I added the factor of deaths for age group 65+ for the size of the bubble, and State of the colors category to showcase each state by a different color. We can see that deaths for age group 65+ are predominantly larger, in the states of (1) California, (2) New York, (3) Texas, (4) Pennsylvania, and (5) Florida.

I have simplified the above bubble chart below to represent 4 regions so that there weren't so many colors. This made the visualization easier to identify regions with the highest death rates.

# Aaron Gilbert

