

Report/Summary

People, especially in mainstream media, have long argued for human contribution to the rapid climate change and its overall impact on the environment. With the power of data science, I decided to discover the truth through data to dispel the myth revolving around climate change and its impact on our society, especially on human health. Before creating any visualization, the data was cleaned to ensure consistency and accuracy while accounting for the ethical implication of the data used during the analysis. Since many businesses and individual readers rely on insights from our analysis to make important decisions; therefore, it is imperative to remove duplicates, misspellings, missing information, and inconsistent data points to avoid making bad decisions leading to hazardous consequences. Additionally, data cleaning cements confidence in our analysis and provides a new perspective to the readers (i.e., doctors, nurses, etc.) that they can trust to develop best practices to streamline operations.

To gain valuable insight from our data surrounding air pollution to make a compelling argument, charts/graphs were generated to visualize the data collected from multiple data sources (websites, databases, APIs, etc.). As shown in Fig. 1, the bubble graph illustrates the emission rate of each country over two decades (1990 -2017), giving a general estimate as to which country contains the most significant emission rate. The bubble graph suggests that Greenland (17,167%) has the highest emission rate, followed by Cambodia (2,495%) and Equatorial (2,445%); therefore, we expect an increase in death percentage in the following countries due to poor air quality. Our hypothesis can be confirmed in Fig. 3 & 4, which shows that countries with high air index (poor air quality) tend to have higher death percentage differences. A similar observation can be made from Fig. 5 which suggests that increasing air index tends to result in a higher death rate compared to a low air index overall. In spite of what the data may suggest, additional data is needed to make definitive claims as other visualization led to inconclusive evidence. For instance, despite being in the top 3 for the overall emission rate, Cambodia (index 2), Greenland (index 1), and Equatorial Guinea (index 1) had the death percentage decreased by 25 percent, 30 percent, and 51 percent, respectively (shown in Fig. 2). Additionally, a side-by-side comparison of emission rate and death percentage of each country shows a sporadic death percentage trend as the emission rate increase; thus, further indicating an inconclusive relationship between air pollution and human health (Fig. 6). In conclusion, the data may suggest that the number of deaths has risen in countries due to increasing emission rate, further research is required that may analysis other variables such as the medical system of the countries or century-old data to conclude. Hopefully, the analysis from this study can be used by future data science enthusiasts to settle the air pollution debate.

<Bubble chart of emission difference percentage across each country>

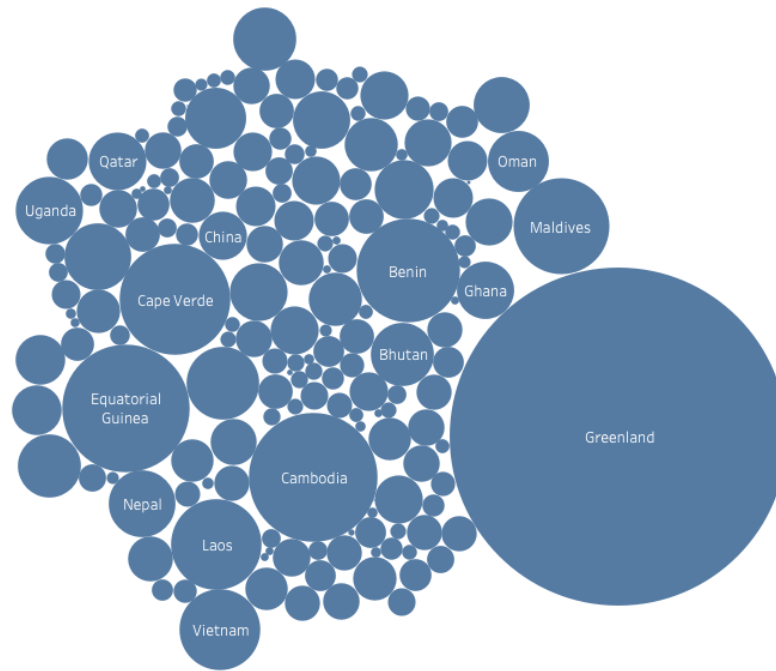


Fig. 1: A bubble chart depicting the emission rate of each country by their relative bubble size.

<Scatter plot of Air Quality vs Death Percentage Difference>

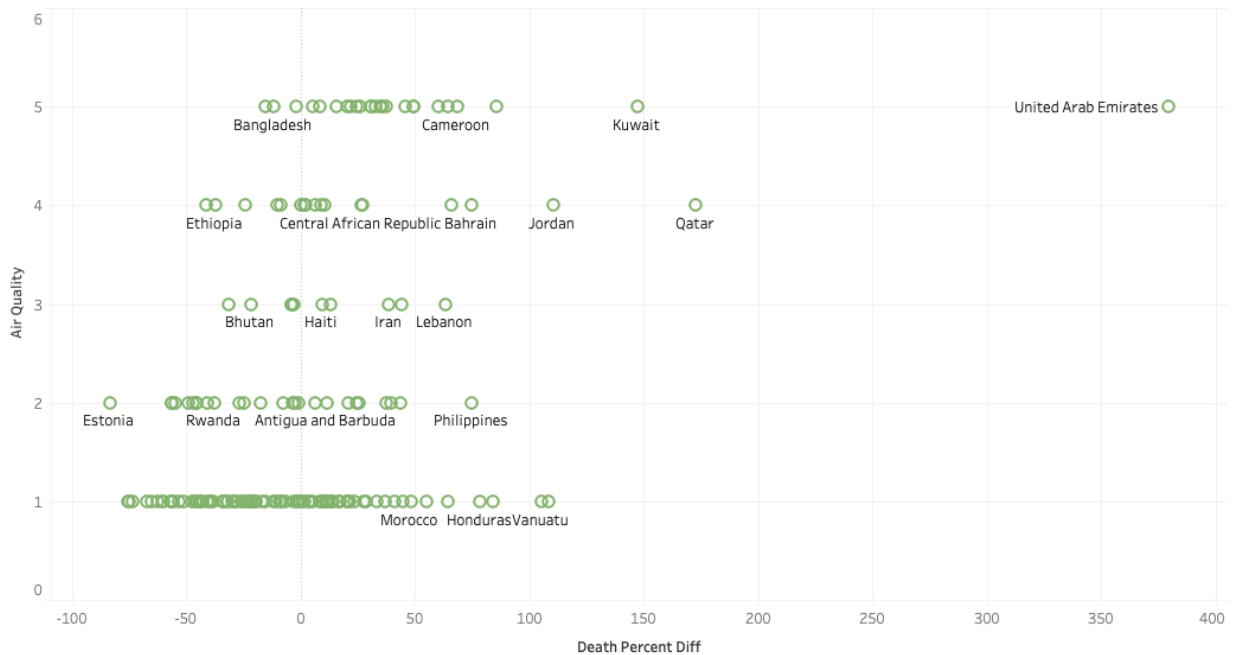


Fig. 2: An Air Quality vs Death Percentage Difference scatterplot showing how the death percentage varies for each air quality index.

<World Map: Air Quality>

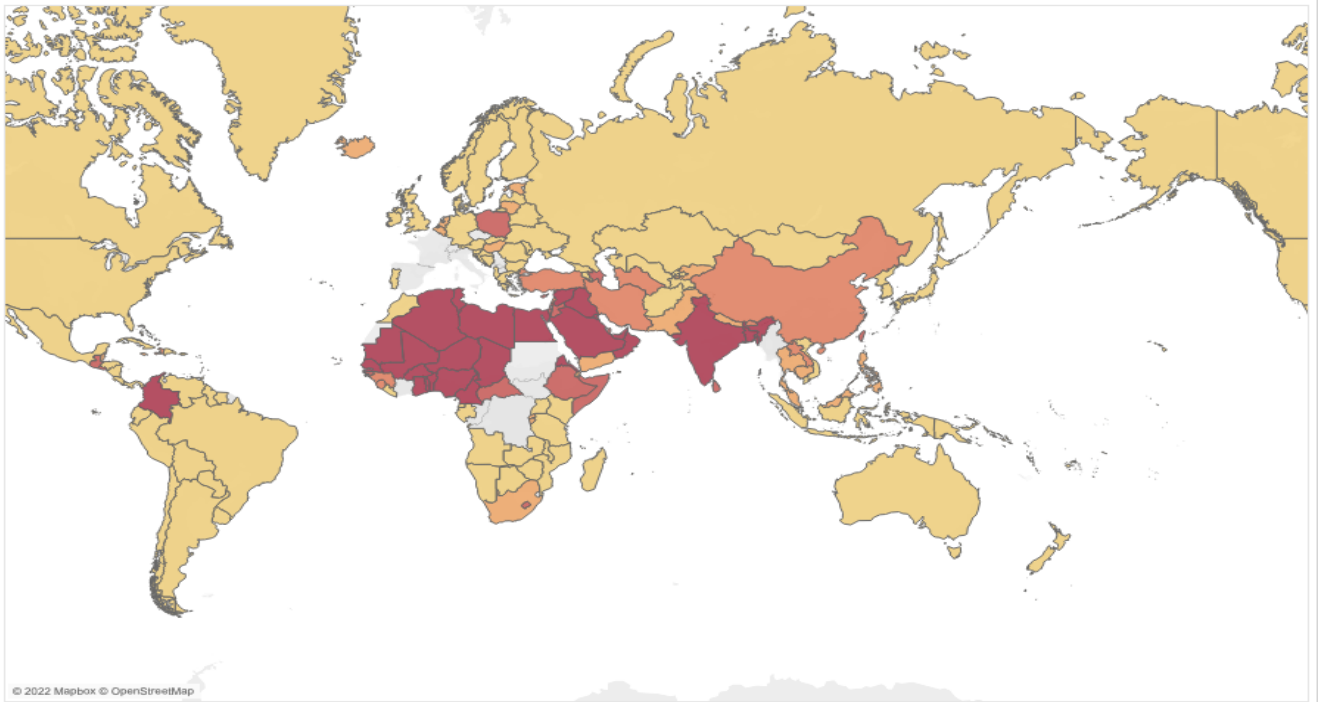


Fig. 3: World map of air quality from the year 1990 to 2017. Countries with darker shades have poor air quality while lighter colors have good air quality.

<World Map: Death Percentage Difference>

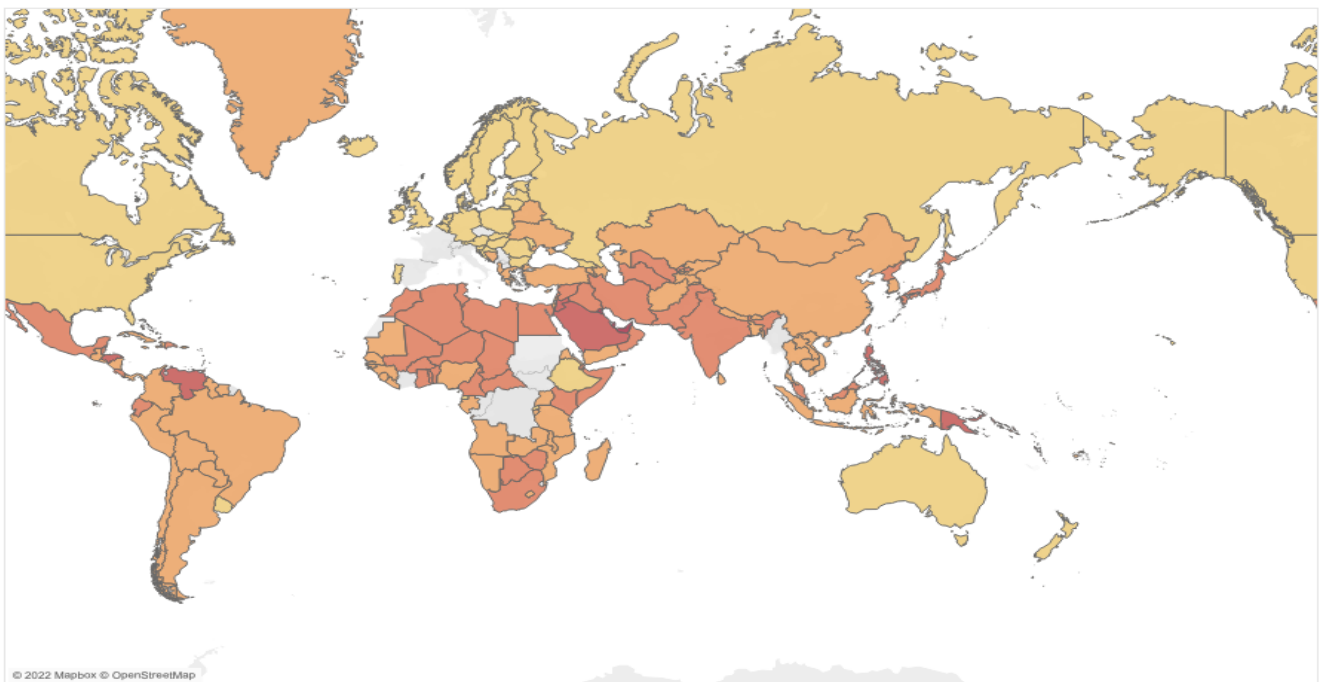


Fig. 4: World map of death percentage difference from the year 1990 to 2017. The percentage value increases as the color become darker.

<Horizontal Bar chart of Death Percentage (Air Quality labeled) >

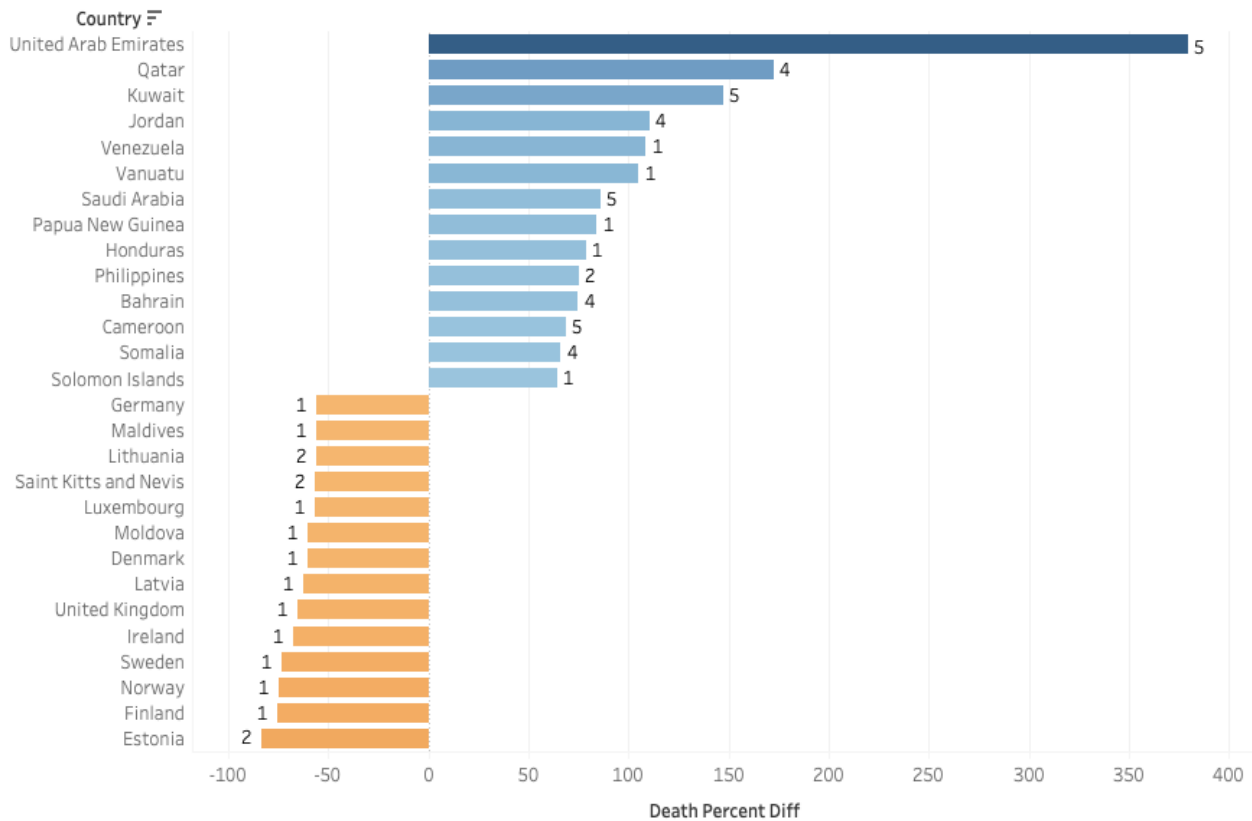


Fig. 5: A horizontal bar chart of the death percentage difference of each country. Each bar is labeled with the air quality index to show the air quality of the region.

<Horizontal Bar chart of Emission Rate & Death Percentage (Emission Rate labeled) >

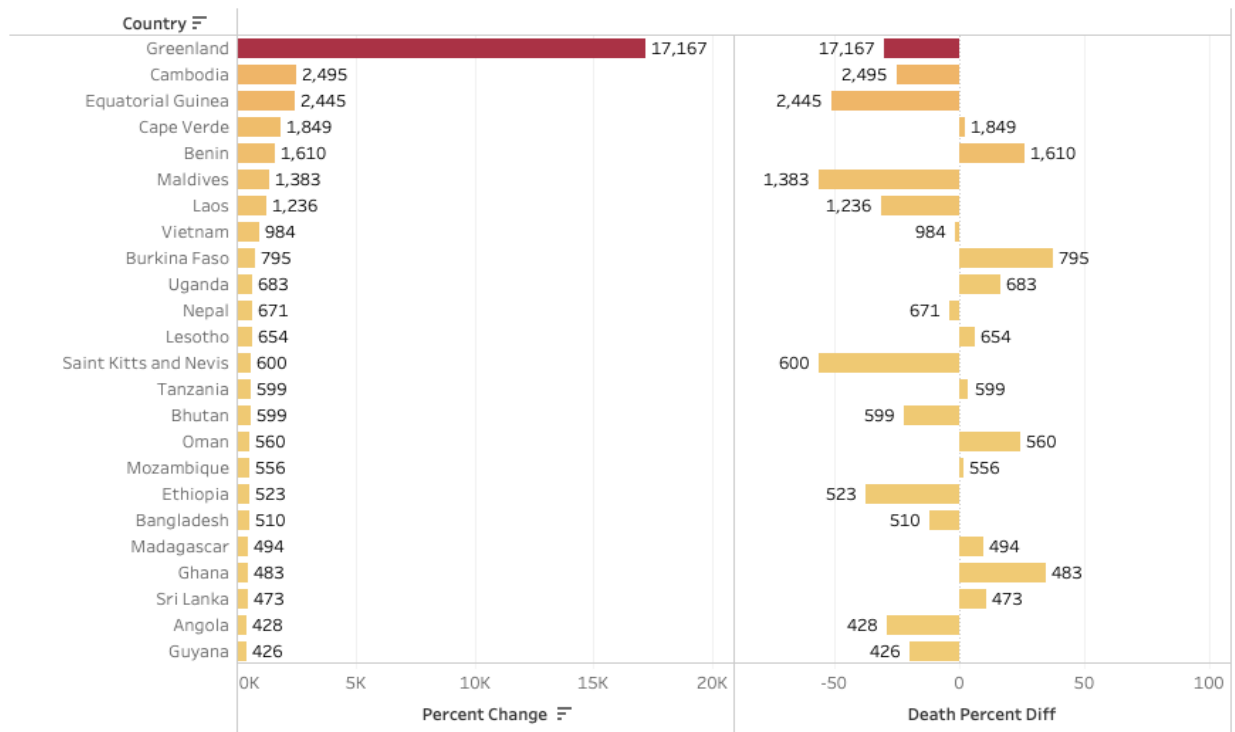


Fig. 6: Horizontal bar chart of emission rate and death percentage difference to understand the relationship between emission rate and death percentage.