**Hypothesis**

1. **Project Description:** This visualization project proposal focuses on elucidating the environmental benefits of reduced human activities on air quality during the COVID-19 pandemic. It aims to create an immersive narrative through data, highlighting the decline in NO2, PM2.5, CO, and SO2 levels across different urban and industrial settings, and the consequences of diminished transportation and shipping activities. The proposal identifies excel and pandas for data cleansing and Tableau for visualization, emphasizing the importance of these tools in handling and presenting complex datasets effectively. This approach aims to highlight the correlation between pandemic-related reductions in human activities and improvements in air quality, offering insights into the environmental impacts of reduced vehicular traffic, industrial activities, air travel, and shipping.
2. **Objective:** The objective of our visualization project is to provide a clear and engaging way to understand the impact of the COVID-19 pandemic on air quality. By examining and presenting the changes in levels of key pollutants (NO2, PM2.5, CO, and SO2) across various urban and industrial environments, we aim to visually demonstrate the direct link between reduced human activities—such as vehicular traffic, industrial operations, air travel, and shipping—and improvements in air quality. This visual analysis will help highlight the potential for adopting sustainable practices that could lead to lasting environmental benefits, encouraging viewers to consider the relationship between human activity and environmental health.

# Hypothesis:

During 2021, Louisiana, alongside other regions like Utah, Ohio, Texas, and Michigan, observed changes in COVID-19 recovery numbers from the prior year. Specifically, Louisiana experienced a notable decrease in recovery numbers by approximately 19.61%, from 28,544,278 in 2020 to 22,947,650 in 2021. This suggests a possible decrease in new COVID-19 cases or improved mitigation efforts. The other states reflected varied trends, with Ohio and Utah also showing reductions in recoveries, while Texas and Michigan saw increases, reflecting the diverse impacts and responses to the pandemic across different states.

In 2021, California and New York saw substantial drops in COVID-19 cases by 45% and 94.74%, respectively, while Ohio's cases decreased by 34.22%, potentially due to effective health measures. Montana's cases, on the other hand, surged by 151.69%, suggesting varied pandemic responses. Texas had a minimal rise in cases at 2.64%, indicating a relatively stable situation. These shifts highlight the complex dynamics of pandemic trends across the U.S.

The visual comparison of death rates across Louisiana, Montana, New York, Ohio, and Texas for the years 2020 and 2021 indicates diverse impacts of health crises, possibly COVID-19, in these regions. The fluctuations in the data points from one year to the next could signify the relative success or challenges of each state's response to the pandemic, including public health interventions and healthcare system resilience.

In the states Arizona, Illinois, Indiana, Michigan, Ohio during the peak of the winter months from December to February, there will be a higher incidence of COVID-19 related deaths among the elderly population compared to other age groups. This increase can be attributed to the combined effects of lower temperatures, which are conducive to viral transmission and respiratory complications.

From Q1 2019 to Q4 2020 in cities Albuquerque, Allentown Bethlehem, Laston, Atlanta Sandy, Springs-Roswell, Bakersfield, Baton Rouge and 7 other cities the NO2 content in atmosphere has reduced in 2020 by an average of 7 points when compared with 2019 corresponding to lesser air travel during covid.

1. **Data Sources:**

<https://data.cdc.gov/NCHS/Provisional-COVID-19-Deaths-by-Sex-and-Age/9bhg-hcku/about_data>

https://[www.epa.gov/air-trends/air-quality-cities-and-counties](http://www.epa.gov/air-trends/air-quality-cities-and-counties)

The dataset for the other three hypothesis has been added to the box.

1. **Number of Records:** 21348
2. **Data Cleansing Tool:** Using Excel and Pandas is an effective way to clean data. Excel is user-friendly and great for quick checks and basic fixes. It's handy for small edits and visual checks. Pandas, on the other hand, is a powerhouse for big datasets and complex cleaning tasks. It's efficient and can handle repetitive jobs well. The two work seamlessly together - you can start in Excel for initial exploration, then switch to Pandas for more advanced tasks and automation. This combination gives you the best of both worlds, making data cleaning easier and more effective.
3. **Visualisation Tool:** We opted for Tableau due to its user-friendly interface and robust visualization capabilities. It efficiently manages large datasets, enabling quick data analysis. The diverse visualization options and seamless connectivity to various data sources meet our requirements. The drag-and-drop functionality enhances efficiency, and Tableau's learning resources support ongoing skill improvement.