**Title:** Healthcare Cost Analysis Across the U.S.

**Team Members**: Silvio Zabala, Lynn Soors, Justin Kim, Connor Casey, Chandara In, Daniel Simonson

**Project Description/Outline**:

This project aims to analyze and compare healthcare rates across different states in the U.S. and the types of insurance coverage. By examining the relationship between types of insurance coverage and healthcare expenditure, the project seeks to uncover patterns and insights that may inform healthcare policy and economic planning.

**Research Questions:**

1. What is the difference in healthcare costs between private insurance and public or state-funded healthcare?
2. How does the amount of coverage of public vs private insurance impact how much the average person has in healthcare costs a year?
3. How has the proportion of insured versus uninsured people changed yearly across states from 2011 to 2020?
4. What are the trends in private insurance enrollment by state from 2011 to 2020?
5. How do the numbers of people with direct-purchase insurance versus employer-based insurance change yearly by state from 2011 to 2020?
6. What is the relationship between the number of insured people in each state and the total healthcare expenditure by state for each year from 2011 to 2020?
7. How has healthcare expenditure changed from 2011 to 2020 across different states?
8. What are the trends in healthcare insurance rates from 2011 to 2020, and how do they vary by state?

**Hypothesis**:

* Higher insurance coverage rates correlate with reduced total healthcare expenditure per state, as insured individuals are more likely to engage in preventive care, reducing long-term healthcare costs.
* Public insurance coverage rates correlate with higher total healthcare expenditure per state, as insured individuals are more likely to engage in preventive care, reducing long-term healthcare costs.
* States with a higher proportion of private insurance coverage may experience higher healthcare expenditures due to potentially higher costs associated with private healthcare services.
* Healthcare expenditures have risen consistently over the years.
* Health insurance coverage rates have shown an upward trend over time.

**Datasets**:

* Government Health Expenditure Data: To analyze healthcare costs across states.
* Insurance Coverage Data: To evaluate the changes in insurance coverage rates across different states.
* Types of insurance rates by state.
* Example dataset: <https://1drv.ms/x/c/f12522726470ea74/ETVwB4ygY6lKpRLz97K8Sd4B_rz2xYm6ipL16JqNuJu3Yg>

**Rough Breakdown of Tasks**: We recommend focusing your analysis on multiple techniques, such as aggregation, correlation, comparison, summary statistics, sentiment analysis, and time-series analysis. Don’t forget to include plots during both the exploration and analysis phases.

* Data Collection: Gather datasets from reliable government sources.
* Data Cleaning and Manipulation: Prepare the data for analysis, ensuring consistency and accuracy across variables and timeframes.
* Data Visualization: Create visualizations to illustrate healthcare cost disparities, trends in insurance coverage across states.
* Statistical Analysis: Conduct statistical analyses to test hypotheses and identify significant trends.
* Reporting: Summarize findings and insights in a clear, concise report, supported by visualizations.