DATA RESEARCH LAB

Welcome to the next step in the DATA Research Lab application process. Based on your background and credentials, we have identified you as a potential fit for the Research Lab. However, to further assess your fit, we would like you to complete a short assignment. You are free to use any solution you want (Excel, R, Python, Tableau, etc). Here is the problem that you must solve:

A publicly traded company is trying to determine various aspects of executive pay and its impact on firm performance, as well as other indicators. This company has compiled a collection of data regarding executive pay across various publicly traded companies. They would like you to analyze this data set, and offer insights and recommendations.

Please select EXACTLY TWO of the following ways to analyze the data:

1. **Data Collection**: Explain what other types of data you think could be paired with the provided data. What sources or methods might you use to gather the data? What additional variables in the new data will complement the provided data? Explain how data you would go out and gather would complement the existing data provided to you. (Note: In this task, you are not required to actually gather the data. We simply want to hear from you your thought process as to where we should find other data that would complement this data well to help *solve* the problem).

2. **Data Wrangling**: Pair the provided data set with other data (weather, stock, economic data, etc). Explain how the addition of your chosen data set will help an analyst provide additional insights into the problem above. You must also merge the dataset you found with the one provided to you in this assignment to create a single dataset.

3. **Data Modeling**: Design a model that can explain executive salary with only the data on hand. Ensure to specify the model, and why you believe the variables included in your model should be there. This should be an econometric model, with assumptions of your model fully tested.

4. **Machine Learning**: Pair this data with stock data. Write code that will pair each row in the data with the company's corresponding high stock price during the specified year. Then, design a model that can predict the stock price using various aspects of the executive pay data.

5. **Descriptive Analysis:** Compute various summary statistics and cross tabulations on the provided data set. Provide us insights about executive from the descriptive analysis.

6. **Data Visualization:** Design a few visualizations that capture the story behind executive pay. These could be any plots of your choosing, or a full on dashboard.

DATA.

Deliverables (DUE Jan 3rd):

- Indicate which TWO of the six tasks you chose to undertake.
- Provide code, additional data files, picture files, and any and all other material that you used or generated to answer the business question at hand.
- A 5-minute recording (a zoom recording link) of yourself that presents on the result. You do not need to use a powerpoint presentation, but it would be helpful in presenting your results.