The proposal should include a description of the question and the broad approach in which the specific data will be used to answer the question (you can choose the datasets provided by one of the four data partners, or identify a different data set and if so, please vet it with us ahead of time)

1. SimilarWeb - Can we predict how page views per streaming giant based on other website traffic? We would try a multitude of different approaches to solve this question with web page data. First, we would try to model each streaming giant's page views based on lagged page views as well as their competitor's lagged views. Then we would try to mix in other entertainment data to see if there is a relationship there (e.g., Twitter searches, movie blogs, Walmart searches for streaming devices like TVs). Finally, we would add in economic data and other forms of consumer spending to see if there is an interesting relationship between that and streaming giant page views (e.g., Netflix, Hulu, etc.).

2. Thinknum – Shareholder Data: With this dataset, we want to get our hands on data on the number of shares held by different types of investors, such as retail traders, large institutional funds, and insiders. We want to draw conclusions about the impact of shareholder behavior on the market performance of individual companies. First of all, we will relate this dataset to the actual market performance of individual companies, such as their stock prices, trading volumes, and earnings to explore how the behavior of different types of shareholders affects the market performance for individua stocks and eventually draw conclusions on this 'causal' relationship. For instance, retail traders tend to buy and sell based on short-term trends, while large institutional funds tend to take a longer-term approach for higher returns. As a result, we might find that companies with a high proportion of retail traders might tend to be more volatile and prone to sudden price swings, while those with a high proportion of institutional funds are more stable and predictable. We could also use this dataset to investigate various theories in behavioral finance.

3. Reveliolabs – Glassdoor company reviews: With this dataset, we can align the Glassdoor company review data with the market performance data by matching the dates of the Glassdoor reviews to the corresponding dates of the market performance data in order to explore how such an 'insider' perspective might affect market performance. For instance, we might be able to examine whether there is a correlation between changes in employee satisfaction ratings and changes in stock prices over time. If such data were available in the dataset, we could also relate company retention rate to its market performance.