1. In this project, we will be focussing on answering following set of questions:
- How does rumors (speculations) and news affect stock prices of companies?
- How do changes in stock price of one firm affect stock prices of other firms in the same industry?
- Is there any relationship between different industries? For example, how real-estate industry impacts manufacturing firms?

2. We are using following data sources to collect data:
- **Intrino (https://intrinio.com)** - Intrino offers APIs to extract historical stock prices and other metrics for different companies.
- **Yahoo! Finance(http://finance.yahoo.com)** - We are using yahoo-finance python module to extract historical stock prices and other metrics from this site.
- **News API(https://newsapi.org)** - We are using this API to extract headlines published on a range of news resources and blogs.

The extracted data from Intrino and Yahoo! Finance is in JSON format. Matching entity between two data sources is the name of the companies. Data extracted from News API is stored in text documents.

3. We created scripts to make API calls to the data sources and collect data. Following is the description of the scripts created:
- **getCompanyData.py** - Makes calls to Intrino API.
- **getCompanyDataYahoo2.py** - Makes calls to Yahoo Finance API and extracts stock prices data of last 10/ 5 /3 /2 /1 years (whichever is greater).
- **getTickers.py** - Makes calls to Intrino API and extracts NASDAQ and NYSE symbols for companies.

4. We will be using text documents for sentiment analysis.

5. Open Source Tools:
We used open source python module "yahoo-finance". Other than that we are just making API calls to get historical data.