Kaggle's New York Stock Exchange S&P 500 dataset

Introduction

In this project we will analyze real life data from the New York Stock Exchange. You will be drawing a subset of a large dataset provided by Kaggle that contains historical financial data from S&P 500 companies.
any spreadsheet application you like. This includes Google Sheets, Microsoft Excel, etc.

Why this Project?

This project will introduce you to the data analysis process that you will be using throughout the rest of the Nanodegree program. In this project, you will go through the process of calculating summary statistics, drawing an inference from the statistics, calculating business metrics and using models to forecast future growth prospects for the companies. The goal is for you to perform an analysis and also create visual tools to communicate the results in informative ways.

We have provided a clean data set for this project. Although in real life scenarios, data sets often need to be cleaned and processed before analysis can proceed. This project allows you to see what a clean data set should look like.

Background:

We used the *Fundamentals.csv* and *Securities.csv* files provided by Kaggle. The Fundamentals file provides the fundamental financial data gathered from SEC 10K annual filings from 448 companies listed on the S&P 500 index. The Securities file provided the industry or sector information the companies are categorized under on the S&P 500 index.

What skills will I use?

The main goal of this project is for you to demonstrate your ability to:

- interpret the measures of central tendency and spread (mean, median, standard deviation, range)
- use a combination of Excel or Google Sheets functions (e.g., IF statements, INDEX and MATCH, calculating descriptive statistics with the IF statement, drop downs, data validation, VLOOKUP).
- analyze and forecast financial business metrics using Excel or Google Sheets.
- create visualizations of a business metric and use Excel or Google Sheets to create a financial forecast model.