ETL PROJECT

# *FINAL REPORT*

**Data Source:**

1). Yahoo finance using from pandas\_datareader import data.

2). CSV file from [www.nasdaq.com](http://www.nasdaq.com)

**Process:**

* Created a function to get data of FAANG stocks from yahoo.
* Used Pandas to read CSV file.
* Cleaned up data using Pandas.
* Created a connection between MySQL and Python
* Uploaded loaded Data

**What data sources you chose, and why?**

I choose the data based on availability. The reason why a chose this data is to back test a day trading strategy.

**Explication why you have performed the types of transformations you did?**

To back test my strategy, I needed at least 5 years of data with open, close, high, and low price of each day. Luckily, there is a lot of data available related to financial markets. Using python, I got rid off adj close price column since my strategy is focused on day trading. Original data came in very clean which is rare in data science.

**Why you chose the type of the final database**

My data was sequential and that’s why I went with MySQL.

**Schema of the tables/collections in the final database**

I created a single database with 5 different tables for each stock in FAANG.

**Hypothetical use cases for your database**

Due to change in technology, online trading is becoming very popular. I wanted to created a day trading strategy that can beat the market and generate consistent profit. But before any of that, I had to test it and that’s why I chose this data set.