Course 5: Python Project for Data Science

Week 1:

Project Overview

- Assume the role of a data scientist
- Helping with a new startup investment firm
    - Help customers invest their mondey in stocks
- Job
    - Extract financial data (various sources using python libraries and webscraping)
        - Historical share price
        - Quarterly revenue reporting
    - Visualize this data in a dashboard to identify patterns or trends
    - Stocks to be working with:
        - Tesla
        - Amazon
        - AMD
        - GameStop
- Dashboard Analytics
    - View of key performance indicators in a clear way
    - Analyzing a data set and extracting key performance indicators
- Watson Studio
    - Using Skills Network Labs and Watson studio
        - Skills Network Labs is a sandbox environment for learning and completing labs
        - Watson Studio is a suite of tools and a collaborative environment for data scientists, data analysts, AI and machine learning engineers and domain experts to develop and deploy projects

Stock Shares

- Stock
    - Security that represents the ownership of a fraction of a corporation
    - Entitles the owner of the stock to a proportion of the corporation's assets and profits equal to how much stock they own
    - Units of stock are called "shares"
- Investors can buy a stock and sell it later
- Stock increases, investor profits. Decreases, investor incurs a loss
- Determining stock price
    - Number of outstanding shares
    - Size of the company's future profits
    - Much more
- People trade stocks throughout the day

- Stock ticker
    - Report of the price of a certain stock, updated continuously throughout the trading session by various stock market exchanges
- y-finance API will obtain stock ticker and extract information about the stock

Extracting Stock Data using Python Library

- yfinance as yf
    - Stock info
    - Historical share price data
    - Historical dividends data
- Apple = yf.Ticker(“AAPL”)
    - .info = information as python dictionary
    - .history(period = “value”) = share price of a stock overtime
        - 1d, 5d, 1mo, 3mo, 6mo, 1y, 2y, 5y, 10y, ytd, max
        - Returns dataframe
            - Date = index
            - Open, High, Low, Close, Volume, Stock Splits
            - reset_index = reset index of data frame (inplace = True changes dataframe itself)
    - .dividends = money returned per share an investor owns

Extracting Stock Data Using Web Scraping

- import requests
    - Download webpage
- from bs4 import BeautifulSoup
    - Parse text to html
    - Html then gets converted to dataframe using Python