HW 4 PREVIEW

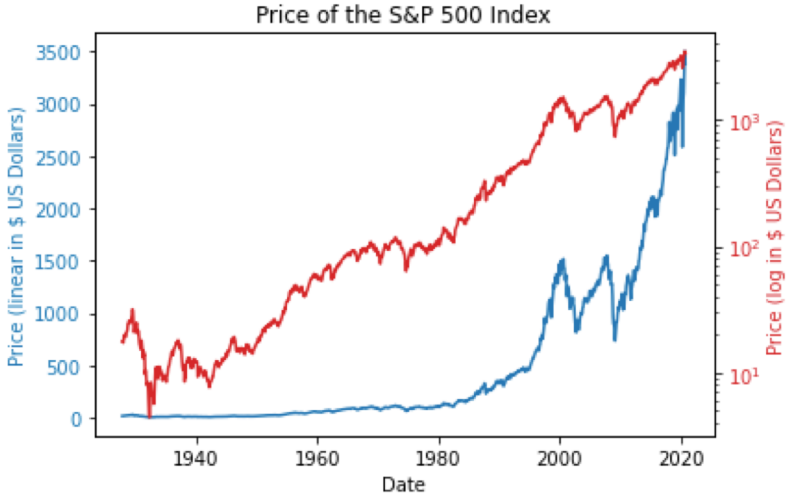
Data-X GSI Team

BEFORE WE START

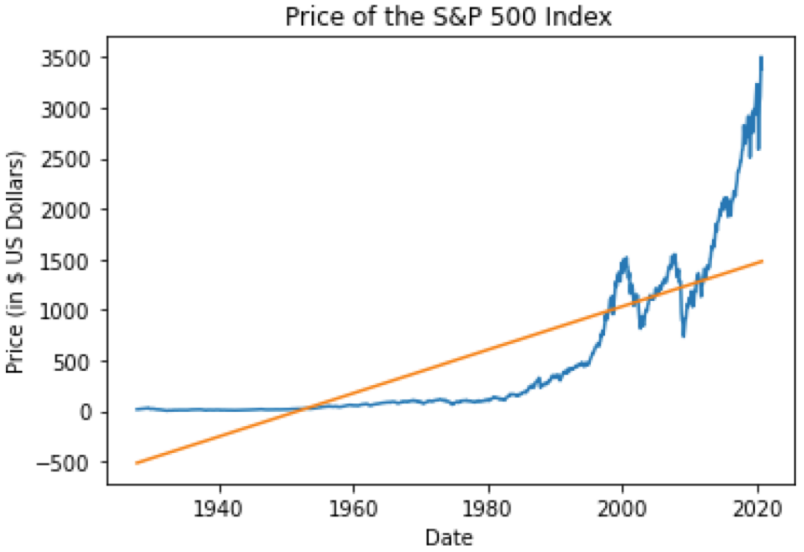
- NumPy:
  - Array
  - Data-X NumPy Module
  - NumPy Manual
- Pandas:
  - DataFrame
  - Series
  - Data-X Pandas Module
  - Pandas Documentation
- Matplotlib:
  - Matplotlib Tutorial

HW 4 STRUCTURE

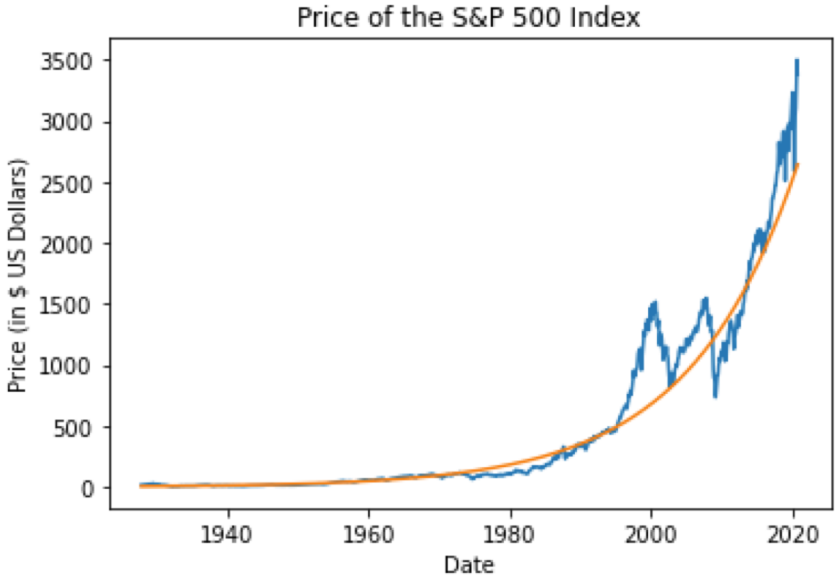
- Part 1: Data Pre-processing
  - 1a) Load data from csv to Pandas DataFrame
  - 1b) Drop unrelated column in the DataFrame
  - 1c) Drop some rows in the DataFrame and make the date consistent among three DF
- Part 2: Data Visualization & Linear Regression
  - 2a) Plot the data with y-left and y-right.
  - 2b) Linear regression of S&P 500 Index.
  - 2c) Exponential regression of S&P 500 Index.
  - 2d) Scatter plot of month-to-month return of S&P 500 Index and 30 Year Treasury Bond.
  - 2e) Plot the histogram


Price of the S&P 500 Index
Price (linear in $ US Dollars)
Price (log in $ US Dollars)
Date

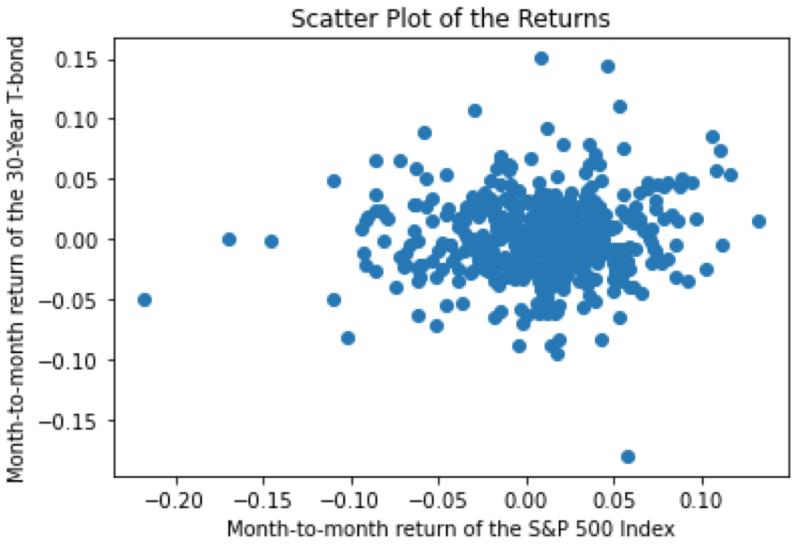
(a)


Price of the S&P 500 Index
Price (in $ US Dollars)
Date

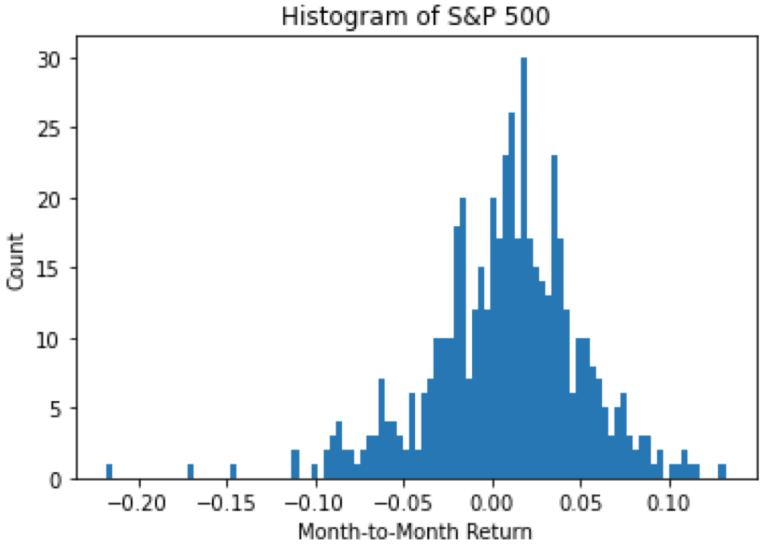
(b)


Price of the S&P 500 Index
Price (in $ US Dollars)
Date

(c)


Scatter Plot of the Returns
Month-to-month return of the 30-Year T-bond
Month-to-month return of the S&P 500 Index

(d)


Histogram of S&P 500
Count
Month-to-Month Return

(e)

SOME USEFUL FUNCTIONS

- **Pandas:**
  - Series.to_numpy(), df.drop_duplicates(), pd.to_dateime(s), and series.astype()
  - Series.str.split(), series.str.join, pd.read_csv(), series.subtract() and df.drop()
- **NumPy:**
  - Np.mean(), np.var(), np.square(), np.exp()
- **Matplotlib:**
  - axis.set_xlabel(), axis.set_ylabel(), axis.tick_params(), and axis.twinx()
  - Plot.yscale(), plot.title(), plot.show(), and plot.subplot()
  - Plot.plot(), plot.scatter(), plot.hist()

SUBMISSION

1. You will be able to run otter grader on your local this time (pip install otter-grader)
2. Run the whole Jupyter Notebook (Cell->Run all)
3. Make sure all the plots are shown correctly
4. Submit the Jupyter Notebook (.ipynb) to Gradescope