ACTU PS5841 Data Science in Finance & Insurance - Autumn 2022 (Y. Wang)
Assignment - 11
Assigned 11/5/22, Due 11/13/22

---

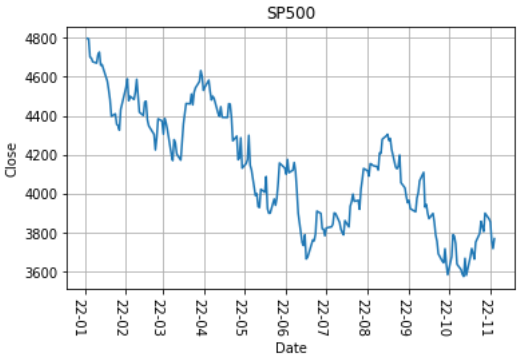
On Wall Street, there was an old adage: "Don't sell stocks on Monday". See
`https://finance.yahoo.com/news/don-t-sell-stocks-monday-140035996.html`
`https://www.amazon.com/Dont-Sell-Stocks-Monday-Watchers/dp/0816010447`
`https://www.amazon.com/Dont-Sell-Stocks-Monday-Hirsch/dp/0140103759`

Let's use a simple logistic regression to investigate the day-of-week effect on the SP500 index.

[*a*] Use `pandas_datareader` to download the year-to-date (from 12/31/21 to 11/4/22) SP500 index levels from Yahoo Finance, into a dataframe.
Note that when you install the package in your virtual environment, the package file name is `pandas-datareader`. When you import the module, the module name is `pandas_datareader`

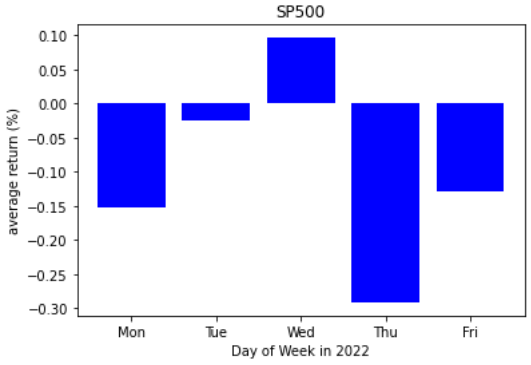
```
df = pandas_datareader.DataReader(name="^GSPC", data_source="yahoo", start="2021-12-31", end="2022-11-4")
```

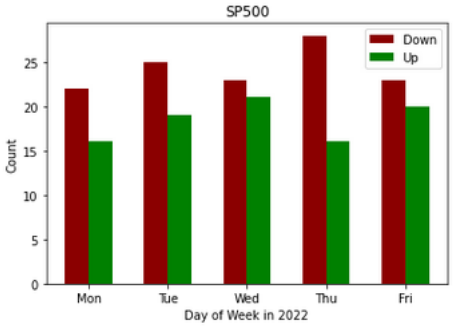
[*b*] Plot the year-to-date SP500 index




[*c*] Let your code calculate the average return (in %) for each day of the week and plot the results.

[*d*] Let your code count the number of up days and the number of down days for each day of the week and plot the results. An up day is one whose close is greater than or equal to the close of the previous trading day.



[*e*] Fit a simple logistic regression with day of the week as the only predictor to estimate the probability of SP500 going up or down given the day of the week. Let your code output the fitted coefficients (including the intercept) and report the log-likelihood achieved.

[*f*] Let your code report the predicted probability of SP500 going up or down for each day of the week.

| | Proba_Up | Proba_Down |
|---|---|---|
| **Mon** | 0.421053 | 0.578947 |
| **Tue** | 0.431818 | 0.568182 |
| **Wed** | 0.477273 | 0.522727 |
| **Thu** | 0.363636 | 0.636364 |
| **Fri** | 0.465116 | 0.534884 |

Please submit your work as hw11.ipynb and hw11.html to `Canvas`.