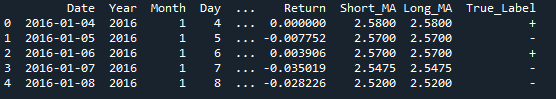
Linda Zhang

Class: CS 677

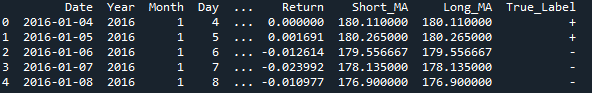
Date: 03/26/2022

**Homework Problem #1.1**(Line 12-Line 21): I created a new column called "True\_Label".I used the "Return" column to determine if it was a '-' or a '+'.I set the training year from 2016,2017,and 2018. The testing year is set at 2019 and 2020.

ZNGA



SPY



**Homework Problem #1.2**(Line 24-Line 36): I took the first training data which contains the

first three year(training data) and determined assumed that all of the days are

independent of each other. I then calculated the probability that the sign

'+' will appear next.

ZNGA

The probability of the next day being ‘+’ is 53.8%.

SPY

The probability the next day will be positive is 55.43%

**Homework Problem #1.3**(Line 39-Line 60):

ZNGA

Probability of a positive day after three negative days is 36.448% over negative day

Probability of a positive day after two negative days is 34.821% over negative day

Probability of a positive day after one negative days is 36.727% over negative day

SPY

Probability of a positive day after three negative days is 38.88%

Probability of a positive day after two negative days is 37.32%

Probability of a positive day after one negative day is 37.31%

**Homework Problem #1.4**(Line 63-Line 82):

ZNGA

Probability of a negative day after three positive days:31.690% over negative day

Probability of a negative day after two positive days:34.090% over negative day

Probability of a negative day after one positive days:49.752% over negative day

SPY

Probability of a negative day after three positive days is 34.73%

Probability of a negative day after two positive days is 33.12%

Probability of a negative day after one positive days is 48.07%

**Homework Problem #2.1**(Line 86-Line 163): I first created a function that would take in a string

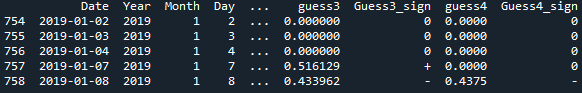
and search for the pattern. In order to be able to search for the value, I had

to change it into a string. I then created compute function which would look

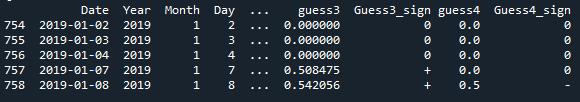
at the W value and predict the values for Year 2019 to 2020 based on the count

of the pattern that was counted in Year 2016,2017 and 2018.

ZNGA



SPY



**Homework Problem #2.2**(Line 167-Line 190):Calculated accuracy of the different W.( I ended up calculating the positive accuracy and the negative accuracy in 3.3 and 3.4 and the code for that was written there instead of in this section of code.

ZNGA

Accuracy of W=2 is 50.7%

Accuracy of W=3 is 50.2%

Accuracy of W=4 is 47.6%

SPY

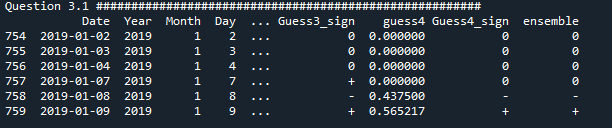
Accuracy of W=2 is 58.56%

Accuracy of W=3 is 58.48%

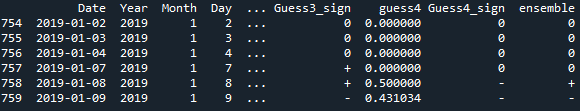
Accuracy of W=4 is 55.2%

**Homework Problem #3.1**(Line 193-Line 219): Created ensemble column that takes the more popular sign of the three W.

ZNGA



SPY



**Homework Problem #3.2**(Line 222-Line 225): looked at the accuracy of the ensemble column.

ZNGA

The accuracy of the ensemble column is 49.2%

SPY

Accuracy of the ensemble column 58.8%

**Homework Problem #3.3**(Line 222-Line 255):

ZNGA

I did not improve my accuracy for predicting ‘-’ for the majority of W values by using the ensemble column. It was better than W=3, but the accuracy was not better than W=2 and W=4

Negative Prediction Accuracy for ensemble is 11.6%

Negative Prediction Accuracy for W=2 is 13.74%

Negative Prediction Accuracy for W=3 is 8.98%

Negative Prediction Accuracy for W=4 is 15.2%

SPY

I did not improve my accuracy for predicting ‘-’ for the majority of W values by using the ensemble column. It was better than W=2, but the accuracy was not better than W=2 and W=4

Negative Prediction Accuracy for ensemble 5%

Negative Prediction Accuracy for W=2 is 0%

Negative Prediction Accuracy for W=3 is 9.58%

Negative Prediction Accuracy for W=4 is 14.2

**Homework Problem #3.4**(Line 258-Line 281):

ZNGA

I did improve my accuracy for predicting ‘+’ for the majority of W values by using the ensemble column. It was better than W=2 and W=4 but worse for W=3.

Positive Prediction Accuracy for ensemble is 37.6%

Positive Prediction Accuracy for W=2 is 37.05%

Positive Prediction Accuracy for W=3 is 41.31%

Positive Prediction Accuracy for W=4 is 32.4%

SPY

I did improve my accuracy for predicting ‘+’ for the majority of W values by using the ensemble column. It was better than W=3 and W=4 but worse for W=2.

Positive Prediction Accuracy for ensemble is 53.8%

Positive Prediction Accuracy for W=2 is 58.56%

Positive Prediction Accuracy for W=3 is 48.90%

Positive Prediction Accuracy for W=4 is 41%

**Homework Problem #4.1**(Line 284-Line 287)

ZNGA

True Positive for W=2 is 184

True Positive for W=3 is 206

True Positive for W=4 is 162

True Positive for ensemble is 188

SPY

True Positive for W=2 is 292

True Positive for W=3 is 244

True Positive for W=4 is 205

True Positive for ensemble is 269

**Homework Problem #4.2**(Line 307-Line 327)

ZNGA

False Positive for W=2 is 149

False Positive for W=3 is 173

False Positive for W=4 is 142

False Positive for ensemble is 160

SPY

False Positive for W=2 is 208

False Positive for W=3 is 160

False Positive for W=4 is 137

False Positive for ensemble is 183

**Homework Problem #4.3**(Line 331-Line 351)

ZNGA

True Negative for W=2 is 69

True Negative for W=3 is 45

True Negative for W=4 is 76

True Negative for ensemble is 58

SPY

True Negative for W=2 is 0

True Negative for W=3 is 48

True Negative for W=4 is 71

True Negative for ensemble is 25

**Homework Problem #4.4**(Line 354-Line 374)

ZNGA

False Negative for W=2 is 98

False Negative for W=3 is 76

False Negative for W=4 is 120

False Negative for ensemble is 94

SPY

False Negative for W=2 is 0

False Negative for W=3 is 48

False Negative for W=4 is 87

False Negative for ensemble is 23

**Homework Problem #4.5**(Line 377-Line 392)

ZNGA

True Positive Rate for W=2 is 65.24%

True Positive Rate for W=3 is 73.04%

True Positive Rate for W=4 is 57.44%

True Positive Rate for ensemble is 66.66%

SPY

True Positive Rate for W=2 is 100%

True Positive Rate for W=3 is 83.56%

True Positive Rate for W=4 is 70.20%

True Positive Rate for ensemble is 92.12%

**Homework Problem #4.6**(Line 396-Line 408)

ZNGA

True Negative Rate for W=2 is 31.65%

True Negative Rate for W=3 is 20.64%

True Negative Rate for W=4 is 34.86%

True Negative Rate for ensemble is 26.60%

SPY

True Negative Rate for W=2 is 0%

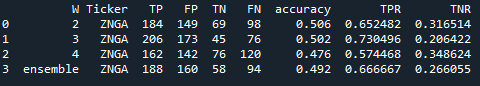
True Negative Rate for W=3 is 0.23076923076923078

True Negative Rate for W=4 is 0.34134615384615385

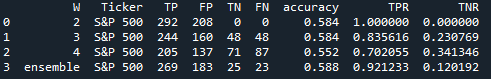
True Negative Rate for ensemble is 0.1201923076923077

**Homework Problem #4.7**(Line 396-Line 408)

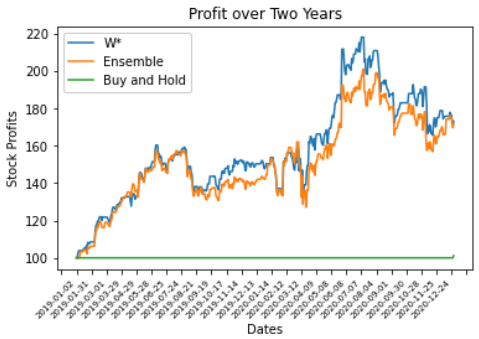
ZNGA



SPY



**Homework Problem #5.1**(Line 432-Line 486)

****

**Homework Problem #5.2**

In year 4 my profits went up and then it went back down again and the same thing happened in year 5.