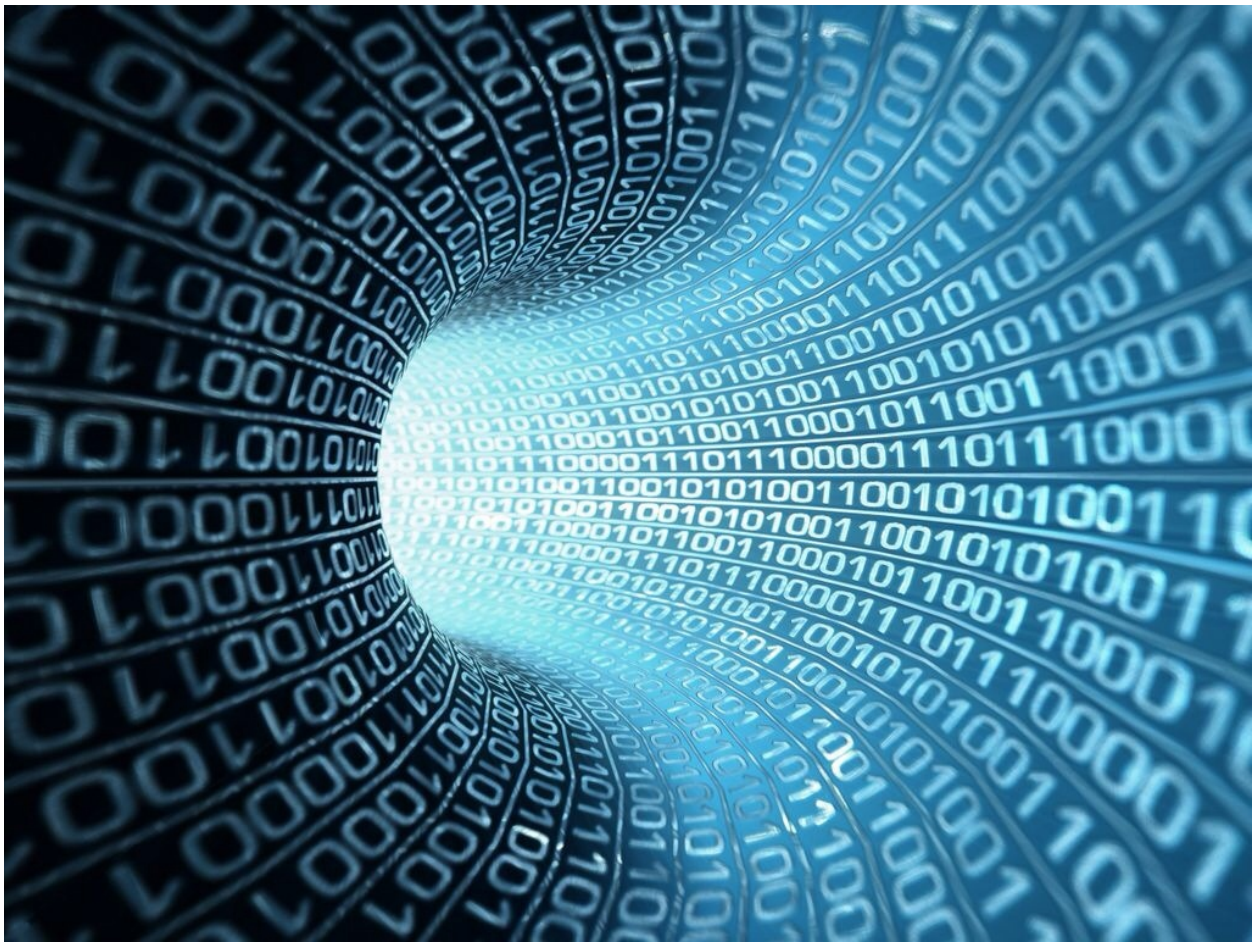
Report: DIAML Assignment 1



I, the undersigned, have read the entire contents of the syllabus for course 18-785 (Data

Inference and Applied Machine Learning) and agree with the terms and conditions of

participating in this course, including adherence to CMU's AIV policy.

Signature: Tunga Tessema

Andrew ID: tchamiss

Full Name: Tunga Tessema

Librairies

I have used the **math**, **matplotlib.pyplot** and **pandas** librairies. I have used the math library to use the exp() function. I have used matplotlib.pyplot to plot the graphs and I have used pandas to read excel and csv files, convert them to dataframes and perform operations on them.

Question 1

We fold the paper until its thickness is greater than the height of mount everest. For this I have used a while loop that terminates once the thickness of the paper is greater than the value. Everytime the paper is folded its thickness doubles. I have found that the paper's thickness is greater after 24 folds

Question 2

I used the formula given to calculate the final volume of the water by incrementing the time t. For this I used a while loop that terminates when the final volume is less than half of the original volume. I have found that after around 6.9 units of time the water in the reservoir has less than half of its initial volume.

Question 3

We wanted to calculate how much money we have after a certain number of years. I used a for loop that runs based on the number of years given. Each time it adds the annual interest to the previous amount of money. I have found the following:
After 1 year, I have: 105
After 2 years, I have: 110
After 3 years, I have: 116
After 4 years, I have: 122
After 5 years, I have: 128

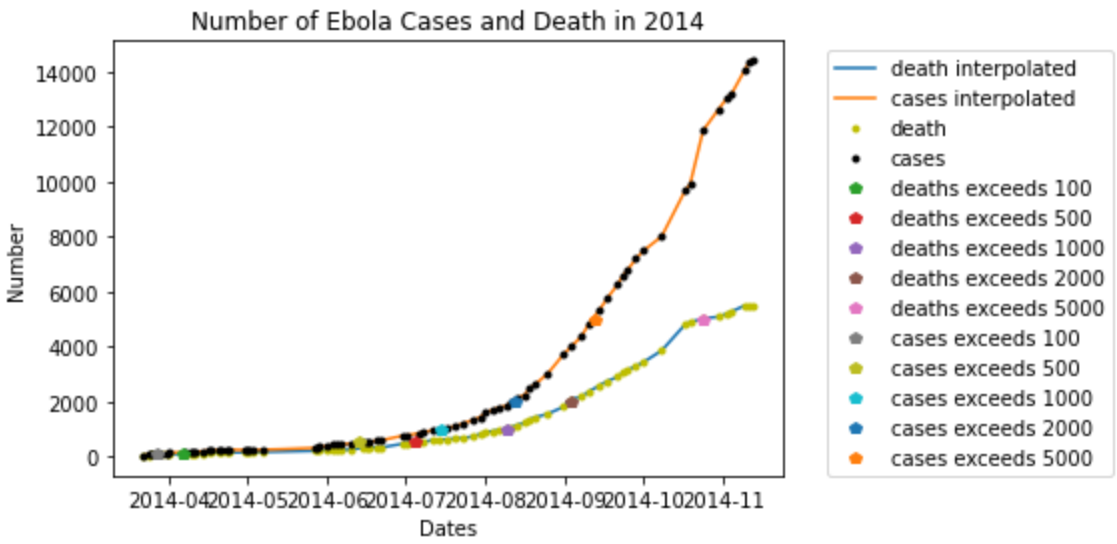
Question 4

We wanted to calculate the monthly payment for a 20 000 loan with a monthly interest rate of 1%. For this I used a function that receives the number of months and uses the loan calculation formula to calculate the monthly payment. I have found the following:
Monthly payment for 1 year: 1777
Monthly payment for 2 years: 941
Monthly payment for 3 years: 664

Question 5

We wanted to calculate the number of days where we will be able to return our initial investment. For this I used a while loop that loops until the profit is greater or equal to the investment. It also stores each value in an array to use it later for the graph plot. I used the matplotlib.plot library to plot the graph. An important thing to remember was that the number of customers should be rounded to the nearest whole number because there is no such thing as a zero point something human.
I have found that 74 is the breakeven day.

Question 6

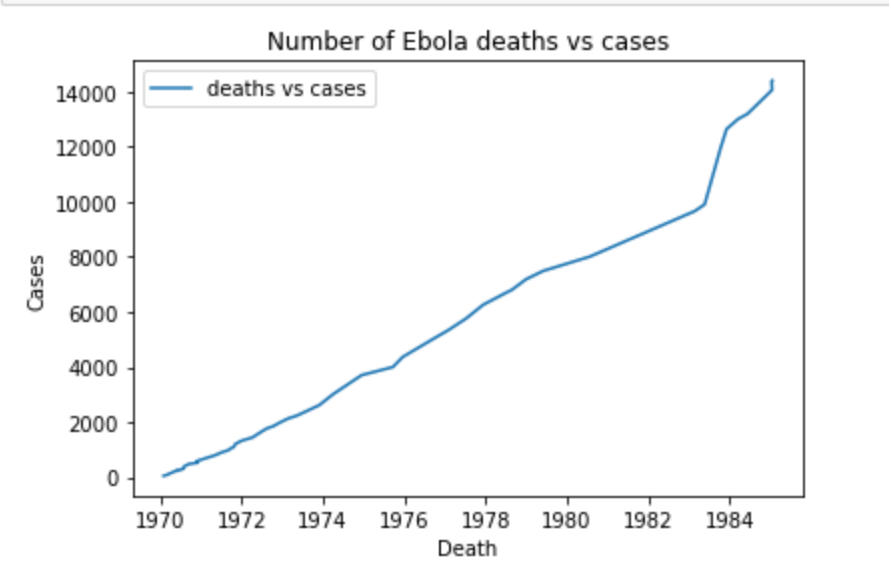


As you can see on the above graph, we have periods of days where the data for ebola cases and deaths is not collected. The black and yellow dots are interrupted. What interpolation does is it fills these gaps with data by calculating the estimates from given data points. By using the interpolation function I was able to get data for everyday and estimate the dates where the number of cases and deaths exceeded 100, 500, 1000, 2000 and 5000.
By plotting the graph we can see that at the beginning the number of cases and deaths were very close but as time went on, only about half of the people who have ebola die. This shows that the way ebola was handled has improved overtime.

Question 7

We wanted to calculate the average growth rate of ebola cases and deaths. First, I calculated the growth rates using the pct_change function of pandas and stored it in new columns called growth_rate_cases and growth_rate_death. I then got their average by using the mean function of pandas.
We saw that on average ebola cases increased by approximately 2.51% everyday and ebola deaths increased by 2.33% everyday.

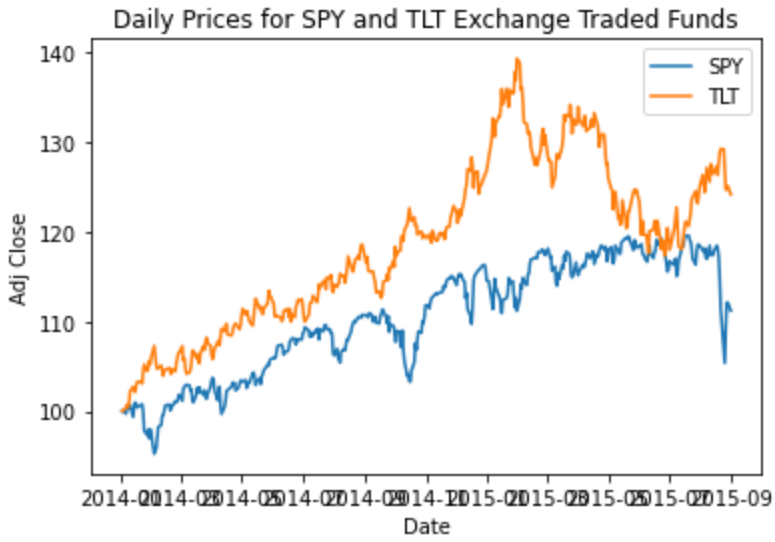
Question 8



We were asked to estimate the average death to case ratio. I first created another column by dividing the number of deaths with the number of cases. I then used the mean function to get the average . I have found a death ratio of approximately 0.558. This shows that about 56% of the people who have been diagnosed with ebola die.

The above graph shows us the number of deaths vs the number of cases, we can see that as the number of cases increase the number of deaths increase but not at the same rate. The number of cases is way bigger than the number of deaths.

Question 9



We were asked to plot the daily price of SPY and TLT exchange traded funds. Both ETFs had different starting prices but I have normalized them to start with the same price of 100. The normalization is done by dividing every value by the first value in the column and multiplying it by 100. This helps us to compare the prices of the two ETFs. As we can see on the graph, overall both prices are going up. But we notice something interesting, when SPY's price is decreasing rapidly, TLT's price is increasing rapidly. Also, their lowest and highest prices are very different. SPY's peak seems to be around 120, while TLT's peak is around 140. If someone wanted to buy among the two ETFs, buying the TLT one would be recommended.

Question 10

```
SPY average daily return in %: 0.028612224182557648
SPY min daily return in %: -4.210697162689259
SPY max daily return in %: 3.8394124544050445

TLT average daily return in %: 0.05537842887416558
TLT min daily return in %: -2.4324931167024166
TLT max daily return in %: 2.6468904269034477
```

We were asked to calculate the average, min and max daily return of the two ETFs. The result once again shows that TLT is a better choice. TLT's average daily return is almost double that of SPY. The daily return measures the dollar change in a stock's price as a percentage of the previous day's closing price. This means that TLT's daily price increase is twice that of SPY's.

In addition, when prices decrease SPY's daily price has decreased as low as 4.2% while TLT's had only decreased at 2.4%.
While prices are increasing, SPY's maximum price increase is 3.8% while TLT's while TLT's maximum increase is by 2.6%. This shows that it happens that SPY increases at a higher rate than TLT but since SPY starts from a lower point than TLT, its final price after the increase might still be lower than that of TLT.