Understanding and Comparing Stock Markets and Exchange Rates

By Harsha Kumar, hk425

Summary

This report explores the relationship between exchange rates and stock price by analysing and comparing the datasets provide. The data has been analysed by producing a times series plots of each dataset to find any similarities between the datasets or and common effects in the graphs. The report also includes a correlation analysis to find the pairs with the best correlation and to be able to calculate a hypothesis test. The results of the hypothesis test show that most pairs were able to accept the alternative hypothesis. Lastly the report covers the simple regression analysis to find the best models that fit each of the dataset through the R squared value.

1.Introduction

This report explores the relationship between different exchange rates and stock exchanges by identifying patterns in the data along with analysing the correlations of each relationship. The datasets consist of data collected monthly across a period of January 1976 to December 2023.

Three of the datasets represent the exchange rates of the British pound, the Australian dollar, the Japanese yen from the US dollar. The other three datasets represent the stock market indexes from different regions. The Dow Jones from the US, FTSE 100 from the UK and the Nikkei 225 from Japan.

Dow Jones tracks the prices of 30 of most traded stocks on the New York Stock Exchange which allows people to determine the overall direction of stock prices.

FTSE 100 is an index made of shares from the 100 top companies on the London Stock Exchange which is determined by the price movement of the stocks.

Nikkei 225 is a price-weighted equity index consisting of 225 stocks selected from the domestic common stocks in the Prime Market of the Tokyo Stock Exchange. Price-weighted equity index means that each component of the index has been weighed separately according to their current share prices.

The software used to analyse the datasets and create visualisations is R studio and R markdown. The report explore the software and methods use in section 2.

This report will cover a time series plot of the monthly prices for each data set between the years of 01January 1982 to 01 December 2023 in section 3 .A correlation analysis for all the data sets in section 4 and a simple linear regression analysis for every pair of datasets in section 5.

2.Methods

The main software used in this project is RStudio to analyse the datasets using R and R markdown to format the code and comments to a HTML webpage for easy access. To be able to access the datasets in RStudio you must import each of the csv files that you want access in this case I used 6 different csv files of the currency exchanges and the given stock exchanges.

To be able to work with this data I had to import many libraries to access certain functions like ggplot or the correlation functions. Next to make sure the data was useable I had to reformat the data. For example, I had to change the date column of all the datasets to make sure the format was right else I wouldn't be able to plot the data in a time series graph due to the way the date was saved.

To plot a time series graph of each data frame I had to filter each of the data to only cover the correct period which was between 1982 to 2023 and select only the price column in the new dataset. However, the Dow Jones variable didn't have the same length of rows as the other filtered sets, this was due to the original Dow Jones dataset having multiple collections of data per month. The analysis of the data only required the first of each month, so I filtered the dataset again to make sure it only included the correct values. The next issue was that the FTSE_100 dataset didn't have all the years required since the data collected only begins in the year 2001. Due to this the axis of the plotted graph for that filtered dataset didn't begin at the correct year and only began in the year the data was available. I wanted to keep all the axis of each graph the same to make it easier to compare each to each other, so I extended the original dataset to 1982 and filled each of the rows with N/a which allowed RStudio to plot the correct period and create a graph with the desired axis.

To calculate the correlation analysis for all these datasets and then a pairwise correlation estimation and hypothesis I used a correlation matrix. This creates a matrix of all the possible comparisons from a dataset and out puts all the correlation data like r value and p value for hypothesis testing. To use the matrix, I had to create a new data frame of all the data I required from each original data set, and the inputted data must be the same length as each other. Since FTSE_100 was the smallest dataset, I filtered the remaining datasets to only contain the same dates as the FTSE_100 and had to make sure the filtered Dow Jones dataset only had the data for the first of each month like the other data's. From the new datasets I selected on the price column which I concatenated into a new data frame that I could use when creating the matrix.

To calculate the linear regression analysis, I created a summary of the linear model which provides the coefficients of the regression formula along with the intercepts which I used to construct and equations. The summary also provides the R squared value which can be used to determine how well the data fits the model. I did this for all the datasets to get a total of 30 comparison and used the comparisons to find the best R value for each set of data.

3.Time series plot of the data.

In this section we analyse the time series graphs for each data set and how the price changes across each month.

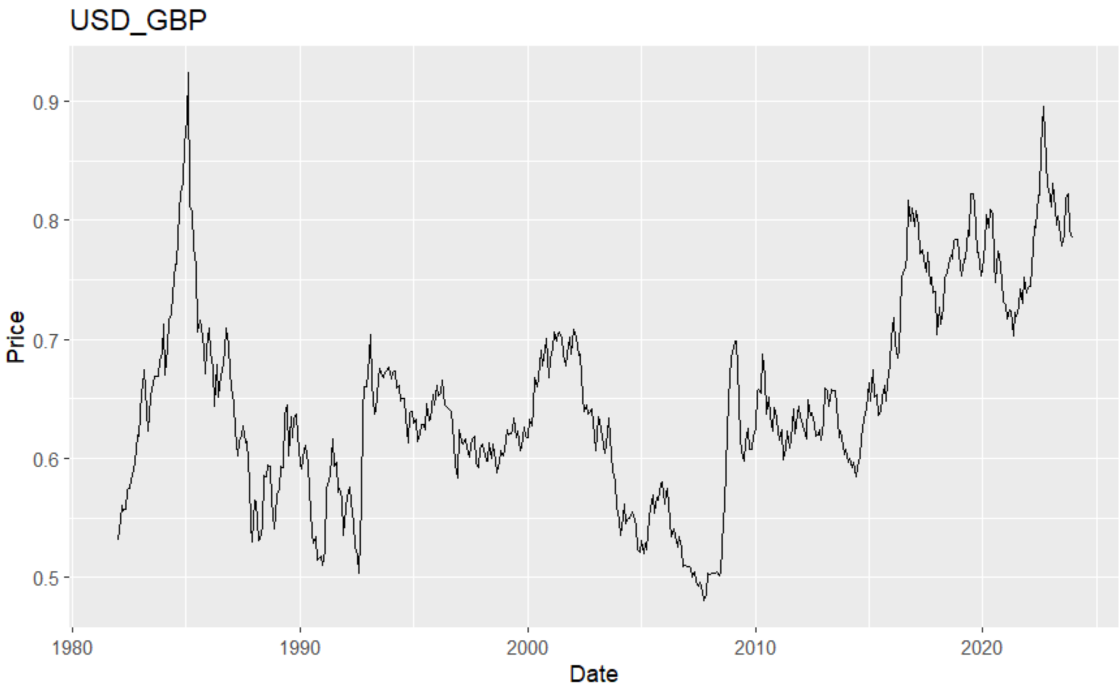


Figure 1

Figure 1 shows the plot of the currency exchange between USD and the British pound. The largest peak in the graph occurs in the year 1985 where the currency exchange reached 0.92, due to the economic troubles in the UK, including the miner strikes along with the Federal reserve increasing interest rates in the US, prompting more investors in the dollar.

The next peak happens in 1992 because of the 1992 sterling crisis. This crisis was due to the collapse of the pound forcing the UK to withdraw from the European Exchange Rate Mechanism (ERM). The ERM was an adjustable exchange rate agreement that most of the nations in Europe joined to link their currencies to prevent large fluctuation in relative value and keep the rate stable. However, The UK couldn't keep the exchange above the lower limit and had to exit the agreement caused by mass selling of the pound by traders.

From 2000's and onwards the rates begin to continuously fluctuate with the greatest decline in the year 2007 caused by the US having the subprime housing crisis, involving a nationwide drop in housing prices and increased default rates for certain mortgages which went on to cause the 2008 global financial crisis plotted at a steep incline in 2008.

The Brexit vote in 2016 and the COVID-19 pandemic also had noticeable impacts on the exchange rate shown by the graph. Brexit led to uncertainty because of ongoing negotiations with the European Union and while the pandemic disrupted global economies and worldwide trade.

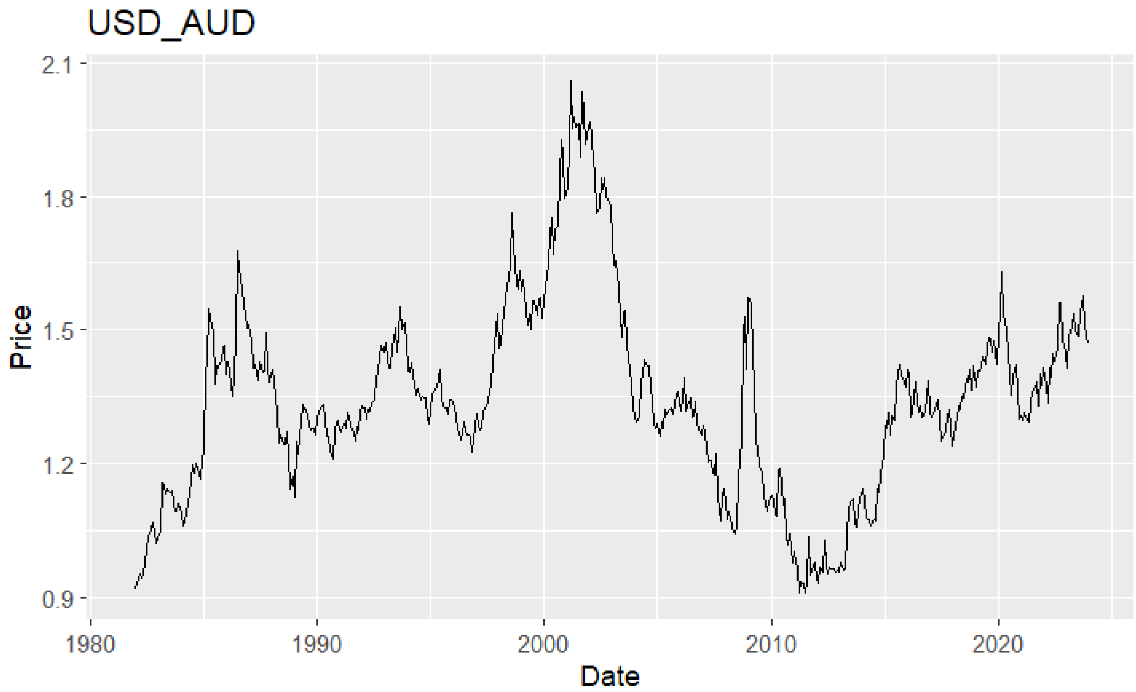


Figure 2

Figure 2 is a graph of the currency exchange between the US dollar and the Australian dollar. The graph shows that the first significant peak occurs around 1987, this could be due to the US increasing interest rates appreciating the US dollars value compared the Australian dollar. In the next decade after many fluctuations in price the price decreases suddenly in 1997 cause by the Asian financial crisis where exports and trade came to a halt reducing import prices and reducing the exchange rate.

There is a notable peak in 2001 raising the exchange rate to 2.06 the highest value the exchange rate increases to where Australian economy recorded its highest rate of inflation which was an aftermath of the Goods and Services Tax (GST) on 1 July 2000, leading on to a sudden decline in the next five years as the value of the Australian dollar begins to increase.

in 2008 the global financial crisis has a significant impact on economies worldwide including Australia which contributes to the increase in exchange rate as well. After this major price increases the graph shows many fluctuations along the years with the graph hitting a minimum value in the 2011 with a price of 0.9 before increasing again. This increase around 2013 could, be due to Australia decrease in trade which would have cause the depreciation in the Australian dollar. Lastly the graph shows large changes in trend in 2020 which was due to the pandemic which disrupted many economies and trade.

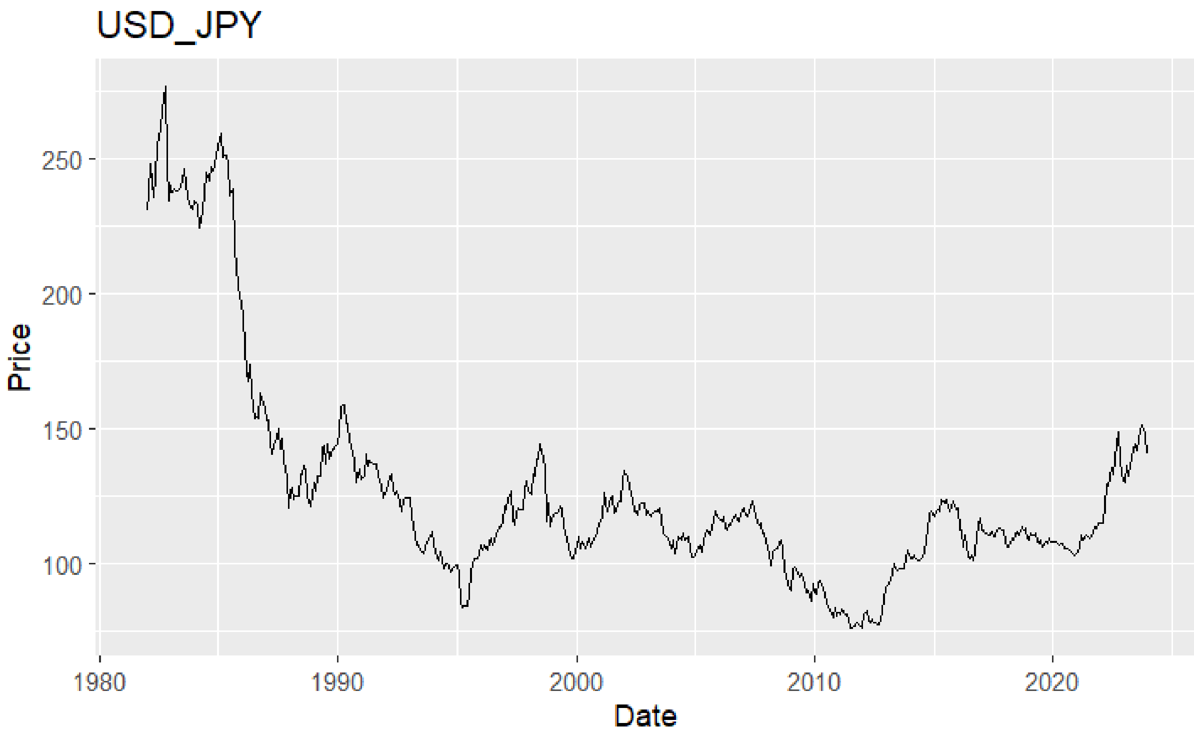


Figure 3

Figure 3 shows the currency exchange between USD and the Yen. The graph begins with the exchange rate being the highest throughout the timeline. The graph hits the max peak in 1982 where the price is 277 yen per dollar.

In 1985 the graph peaks again due to the plaza accord established by the G-5 nations to manipulate exchange rates by depreciating the US dollar before leading to a large decline in the years to come. This decline was because of Japan's rapid economic growth and the increased exportation by the country increasing the yens values and decreasing the exchange rate price.

However, due to excessive investment in land and stocks during the 1980's resulted in forming an asset price bubble due to the high expectations in Japan's future growth. This bubble collapsed in 1991 shown in the graph by the sharp peak just after 1990. This led to a slow economic growth in the later years.

In the next decades there are many fluctuations in the rate, but it stays in the same range before dipping significantly towards 2010. At this point the world was undergoing the 2008 financial crisis resulting in many currencies being depreciated and large changes in exchange rate. However, Japan was immune to this event and the value of the yen only continued to increase while compared to the US dollar showing in the graph as the low number of yen the dollar could buy. Later in the years the exchange gets affected by covid 19 which impacts on the exchange rate shown by the graph due to affect global trade.

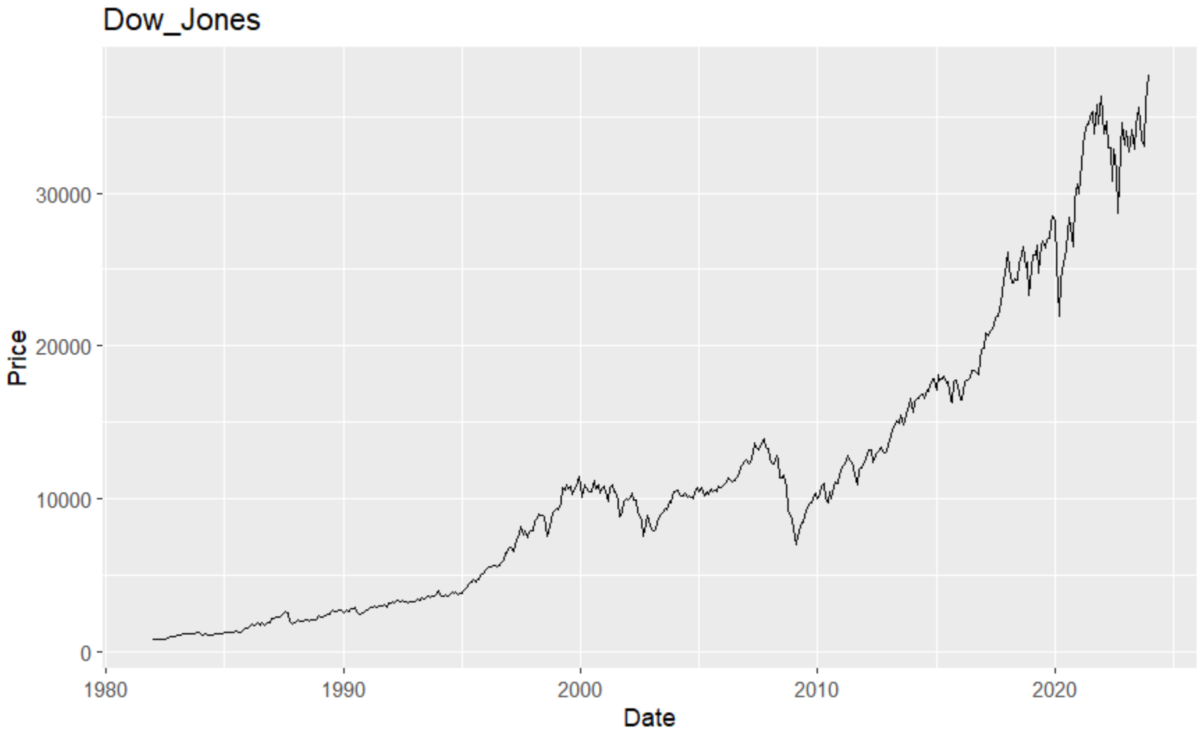


Figure 4

Figure 4 shows the prices of 30 of the most traded stocks on the New York Stock Exchange grouped by the Dow Jones Industrial Average. The graph begins with the stock being low. There are many fluctuations along the years but overall, there is a steady increase in the price from 1995. The graph also shows many sudden declines the most noticeable around 2008 with a large drop in stock value, most likely due to the housing crisis along with the 2008 crisis causing a stock market crash forcing the trade to stay below 10,000 for the first time in 5 years show by the graph. The market seems to stabilise by 2010 and the price continues to increase until 2020 where the market crashes again caused by Covid 19 and the instability it caused. The graph shows many fluctuations in price along with the lack of increase for the next two years due to the concern and the impact the pandemic had on the country.

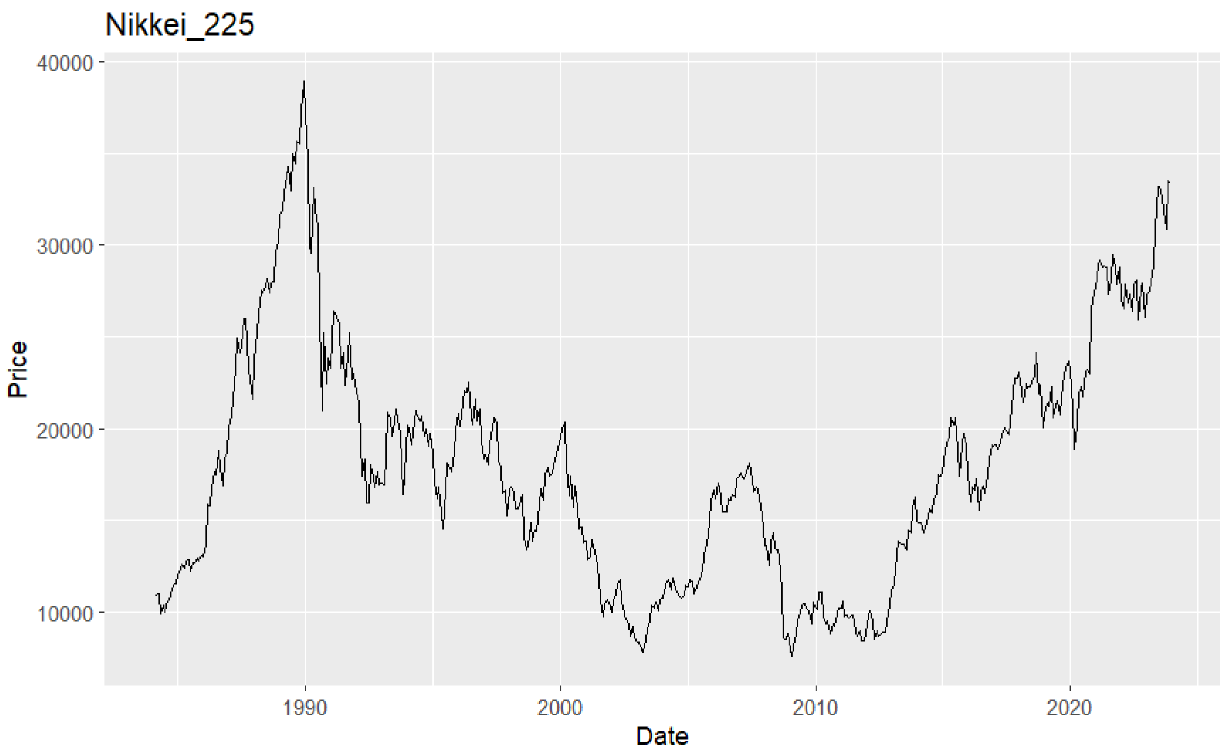
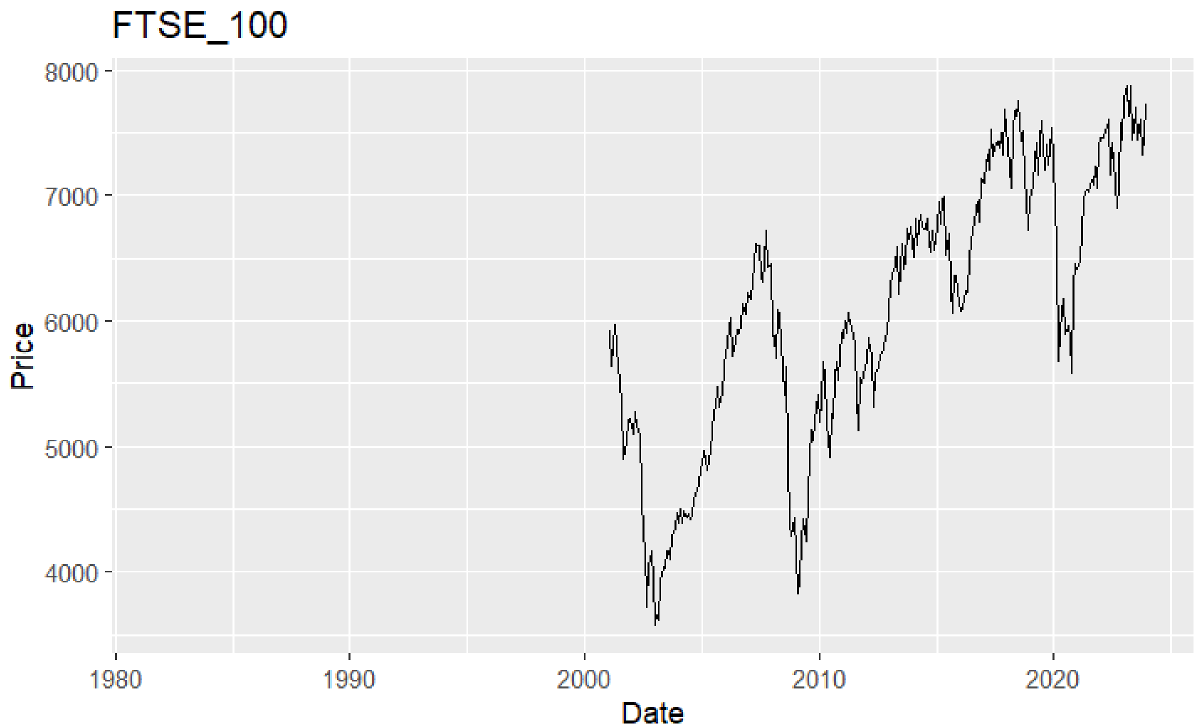


Figure 5

Figure 5 shows the prices of Nikkei 225 which consists of 225 stocks in the Prime Market of the Tokyo Stock Exchange over years. The graph shows a huge incline during the late 1980's which correlates with Japan's rapid economic growth. The incline peaks in 1989 with a stock price of 38916. This massive peak also declines at a rapid rate cause by the collapse of Japan's asset bubble in 1991 cause by over investment and high asset prices. This decline led to fluctuations in the coming years till the stock market dropped to a low in 2003 before recovering slightly until 2008. In 2008 due to the financial crisis the stock prices decreased again to an all time low of 7568 in 2009. Around 2010 the market begin to rise again and continued to rise with many minor fluctuations. Covid 19 seems to have affected the prices slightly with the graph showing a decrease slightly however the market is able to recover, and the stock value is able to continue in gaining value.



This graph represents a time-series plot of the stock prices of FTSE 100 which is index made up of shares from the 100 biggest companies by market capitalisation on the London Stock Exchange. The graph begins in 2001, it immediately starts to decrease in value until it hits the minimum of 3567 in 2003. This massive decrease could be due to the economic uncertainty especially after the sudden burst of the dot-com bubble in 2001 which may have decreased investments, another reason could be due to weak corporate earnings, since many companies listed in the FTSE_100 reported weak earnings in that period. In the next 5 year the graph shows a large incline in price producing a peak in 2007, this is short lived with the stock markets crashing in 2008 due to the global financial crisis as shown in the graph. Over the next few year the index recovers and begins to increase steadily until the pandemic around 2020 which threw the prices into an unbalance, but the prices recover and continues inclining at the steady where it reaches a high of 7876 in 2023.

Overall, all the graphs have a few similar changes across the year with most graphs showing effects of the global crisis cause a drop in stock prices and increased exchange rates, the only rate immune to the disaster was the USD_JPY dataset. This dataset once plotted clearly shows the price continues to decrease. All the graphs show the effects of covid too and how the values change. One thing I noticed when plotting was that it wasn't possible to analyse the graphs for seasonal effects that occur often, due to the range of the data and the scale of the x axis the graphs went up in years instead of months, so when looking at the graphs and the significant changes that occurred they match with economical events and not seasonal event.

4.correlation analysis

```
# Correlation Matrix (auto-method)

Parameter1        |       Parameter2 |     r |         95% CI | t(273) |         p
-----------------------------------------------------------------------------------
USD_GBP.Price     |    USD_AUD.Price |  0.37 | [ 0.26,  0.47] |   6.61 | < .001***
USD_GBP.Price     |    USD_JPY.Price |  0.32 | [ 0.21,  0.42] |   5.52 | < .001***
USD_GBP.Price     |  Dow_Jones.Price |  0.74 | [ 0.68,  0.79] |  18.07 | < .001***
USD_GBP.Price     | Nikkei_225.Price |  0.63 | [ 0.55,  0.70] |  13.41 | < .001***
USD_GBP.Price     |   FTSE_100.Price |  0.56 | [ 0.47,  0.63] |  11.11 | < .001***
USD_AUD.Price     |    USD_JPY.Price |  0.69 | [ 0.62,  0.75] |  15.77 | < .001***
USD_AUD.Price     |  Dow_Jones.Price |  0.09 | [-0.03,  0.20] |   1.44 | 0.150
USD_AUD.Price     | Nikkei_225.Price |  0.16 | [ 0.04,  0.27] |   2.69 | 0.023*
USD_AUD.Price     |   FTSE_100.Price | -0.16 | [-0.27, -0.04] |  -2.60 | 0.023*
USD_JPY.Price     |  Dow_Jones.Price |  0.40 | [ 0.30,  0.50] |   7.32 | < .001***
USD_JPY.Price     | Nikkei_225.Price |  0.58 | [ 0.49,  0.65] |  11.64 | < .001***
USD_JPY.Price     |   FTSE_100.Price |  0.29 | [ 0.17,  0.39] |   4.92 | < .001***
Dow_Jones.Price   | Nikkei_225.Price |  0.95 | [ 0.94,  0.96] |  50.57 | < .001***
Dow_Jones.Price   |   FTSE_100.Price |  0.81 | [ 0.76,  0.84] |  22.60 | < .001***
Nikkei_225.Price  |   FTSE_100.Price |  0.81 | [ 0.77,  0.85] |  23.00 | < .001***

p-value adjustment method: Holm (1979)
Observations: 275
```

Figure 6

In this analysis I conducted a correlation analysis using all the datasets. I used the correlation::correlation() function available with the correlation package in r for the analysis. If using the auto method, the function automatically finds the best method for the correlation estimation and outputs the calculate results in a matrix form that includes the correlation coefficient labelled as r in the results shown in figure 6, along with the confidence interval, the t-value and the p-value for each comparison performed. Figure 6 shows the output of the correlation matrix after comparing all the possible pairs of the datasets and using this data set I can conduct a hypothesis test where the null hypothesis is if there is no correlation between the pair and the alternative hypothesis being there is some correlation present.

Based on the correlation matrix here are some findings that show the three possible results:

The comparison of the USD to GBP exchange rate compared to the USD to AUD exchange rate shows that there is a correlation of 0.37 between the two suggesting that if the USD to GBP rates increase then the USD to AUD rates may also increase as well. The matrix also shows that p value calculated is less than 0.001. since the p value is less than the 1% significance level it means that there is strong evidence to shows that there is a correlation between the pair and that I can reject the null hypothesis.

Another good comparison would be the USD to AUD rates compared to the Dow Jones stock market data. This pair produces a r value of 0.09 suggesting that there is not much correlation between the two sets of data this matches with the p value of 0.15. this p value is a lot greater than the 5% significance level showing that there is evidence to suggest no correlation and in this case I must reject the alternative hypothesis.

The third example would be the pair USD to AUD rates against the Nikkei 225 stock data, the pair has a r value of 0.16 showing that there is correlation just between the data just not much of it. The p value is 0.023 which is small enough to reject the null hypothesis at a significance level of 5% but still able to reject the alternate hypothesis at a 1% level, showing that less strong evidence against the null hypothesis but its still significant at a 5% level.

Out of all the 15 comparisons only 1 had evidence to reject the alternate hypothesis at a 5% significance level and 2 had evidence to reject the alternate hypothesis at a 1% significance level all which were comparisons to the stock exchange data, the remaining pairs all had evidence to reject the null hypothesis.

5. simple linear regression analysis

I have conducted a linear regression analysis on all the datasets which produced 30 possible combinations with results, I decided to only analyses the best fits for each dataset.

USD_GBP:

Figure 7

```
Call:
lm(formula = GBP$Price ~ DowJones$Price)

Residuals:
      Min        1Q    Median        3Q       Max
-0.154222 -0.047577  0.005657  0.045058  0.146691

Coefficients:
                Estimate Std. Error t value Pr(>|t|)
(Intercept)    5.179e-01  8.937e-03   57.95   <2e-16 ***
DowJones$Price 8.377e-06  4.636e-07   18.07   <2e-16 ***
---
Signif. codes:
0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.06465 on 273 degrees of freedom
Multiple R-squared:  0.5446,     Adjusted R-squared:  0.543
F-statistic: 326.5 on 1 and 273 DF,  p-value: < 2.2e-16
```

The best comparison for this dataset was GBP against Dow Jones. Figure 7 shows the summary of the linear regression analysis which contains data for the coefficients which can be used to create a regression line along with the R squared value which shows the goodness of the fit and the p value which I later used for the hypothesis test.

Gradient: 8.38e-0.6

Intercept: 5.18e-01

This creates the regression formula of y = 8.38e-06x + 5.18e-01.

The R squared value for this comparison is 0.54 showing that the model has a goodness fit of 54%. This R squared value is greater than 50% which shows that it is a good model for the dataset. Next I conducted a hypothesis test where the null hypothesis is no correlation, and the alternative hypothesis is that there is a correlation. The p value for this analysis is less than 2.2e-16 and since this is smaller than the significance level of 5% I can reject the null hypothesis.

USD_AUD:

Figure 8

```
Call:
lm(formula = AUD$Price ~ JPY$Price)

Residuals:
     Min       1Q   Median       3Q      Max
-0.32554 -0.12292 -0.02673  0.06277  0.59267

Coefficients:
             Estimate Std. Error t value Pr(>|t|)
(Intercept) 0.1344446  0.0765880   1.755   0.0803 .
JPY$Price   0.0109487  0.0006942  15.772   <2e-16 ***
---
Signif. codes:
0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.1763 on 273 degrees of freedom
Multiple R-squared:  0.4768,     Adjusted R-squared:
0.4749
F-statistic: 248.8 on 1 and 273 DF,  p-value: < 2.2e-16
```

The best comparison for USD_AUD was with the dataset USD_JPY. Figure 8 shows the data produced for the regression analysis.

Gradient:0.01

Intercept: 0.13

These coefficients produce a regression line y = 0.01x + 0.13.

The R squared value for these datasets in 0.48 which means that the goodness fit of the model is 48%. However, even though this comparison had the highest R squared value the percentage is below 50% suggesting that it could be changed to be a better fit for the data. For the hypothesis test the p value for this analysis is less than 2.2e-16 and since this is smaller than the significance level of 5% I can reject the null hypothesis.

USD_JPY:

Figure 9

```
Call:
lm(formula = JPY$Price ~ AUD$Price)

Residuals:
    Min      1Q  Median      3Q     Max 
-29.895  -7.877  -0.792   5.805  31.650 

Coefficients:
            Estimate Std. Error t value Pr(>|t|)    
(Intercept)   51.314      3.735   13.74   <2e-16 ***
AUD$Price     43.546      2.761   15.77   <2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 11.12 on 273 degrees of freedom
Multiple R-squared:  0.4768,	Adjusted R-squared:  0.4749 
F-statistic: 248.8 on 1 and 273 DF,  p-value: < 2.2e-16
```

This comparison is between USD_JPY and USD_AUD since it was the pair with the best R squared value. Figure 9 shows the data produced for the calculation.

Gradient:43.54

Intercept:51.31

These coefficients produce a regression line of y = 43.55x +51.31

The R squared value is 0.48 showing that the model has a goodness fit of 48%. However, even though this comparison had the highest R squared value the percentage is below 50% suggesting that it could be changed to be a better fit for the data. For the hypothesis test the p value for this analysis is less than 2.2e-16 and since this is smaller than the significance level of 5% I can reject the null hypothesis.

Dow Jones:

Figure 10

```
Call:
lm(formula = DowJones$Price ~ Nikkei$Price)

Residuals:
    Min      1Q  Median      3Q     Max 
-6995.3 -1234.8   144.5  1733.6  5225.7 

Coefficients:
              Estimate Std. Error t value Pr(>|t|)    
(Intercept)  -2.504e+03  4.232e+02  -5.918 9.76e-09 ***
Nikkei$Price  1.208e+00  2.389e-02  50.574  < 2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 2621 on 273 degrees of freedom
Multiple R-squared:  0.9036,	Adjusted R-squared:  0.9032 
F-statistic:  2558 on 1 and 273 DF,  p-value: < 2.2e-16
```

Figure 10 shows the data calculated by the comparison of Dow Jones to Nikkei 225 since they had the highest R squared value.

Gradient:1.21

Intercept: -2.50e+03

These values create a regression line of y= 1.21x - 2.50e+03

The R squared value is 0.90 meaning that comparison has a high goodness fit of 90% since the value is extremely close to 1 this shows that the model is the best fit for the data. %. For the hypothesis test the p value for this analysis is less than 2.2e-16 and since this is smaller than the significance level of 5% I can reject the null hypothesis.

Nikkei:

Figure 11

```
Call:
lm(formula = Nikkei$Price ~ DowJones$Price)

Residuals:
    Min      1Q  Median      3Q     Max
-4638.0 -1284.0  -208.3   909.7  5292.6

Coefficients:
               Estimate Std. Error t value Pr(>|t|)
(Intercept)    3.458e+03  2.851e+02   12.13   <2e-16 ***
DowJones$Price 7.480e-01  1.479e-02   50.57   <2e-16 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 2063 on 273 degrees of freedom
Multiple R-squared:  0.9036,	Adjusted R-squared:  0.9032
F-statistic:  2558 on 1 and 273 DF,  p-value: < 2.2e-16
```

Figure 11 shows the data calculated by the comparison of Nikkei 225 to Dow Jones since there were the best comparison.
Gradient:7.48e-01
Intercept: 3.46e+03
These values create a regression line y = 7.48e-01x +3.46e+03
The R squared value is 0.90 meaning that comparison has a high goodness fit of 90% since the value is extremely close to 1 this shows that the model is the best fit for the data. For the hypothesis test the p value for this analysis is less than 2.2e-16 and since this is smaller than the significance level of 5% I can reject the null hypothesis.

FTSE_100:

Figure 12

```
Call:
lm(formula = FTSE$Price ~ Nikkei$Price)

Residuals:
     Min       1Q   Median       3Q      Max
-1426.32  -495.74    65.71   553.47  1034.37

Coefficients:
             Estimate Std. Error t value Pr(>|t|)
(Intercept)  3.894e+03  1.016e+02   38.33   <2e-16 ***
Nikkei$Price 1.319e-01  5.734e-03   23.00   <2e-16 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 629.2 on 273 degrees of freedom
Multiple R-squared:  0.6597,	Adjusted R-squared:  0.6584
F-statistic: 529.1 on 1 and 273 DF,  p-value: < 2.2e-16
```

This calculation showing in figure 12 is a comparison of the datasets for FTSE and Nikkei.
Gradient:1.32e-01
Intercept:3.89e+03
The regression line for this pair is y = 1.32e-01x + 3.89e+03.
The R squared is calculated to be 0.66 meaning that this model has a fit of 60% since the value is about 50% it shows that the model has a good fit for the FTSE dataset. For the hypothesis test the p value for this analysis is less than 2.2e-16 and since this is smaller than the significance level of 5% I can reject the null hypothesis.

All the comparisons were the best models for each of the 6 datasets since they had the highest R squared values. However, only one of the comparisons had a percentage that was the closest to 1 the rest of the comparison remained around the 50% area all those models whether they are good fit or not can be improved a lot more to create a better model for the datasets. Another thing I noticed is that a lot of the R squared values for the remaining comparisons are not about 50% and since a majority of the pairs show this result it could suggest that a linear model is not the right way to produce a fit for the data, A way to improve the models would be to check for non-linearity since most the graphs created in task 1 don't really follow a linear pattern. We could also use the residual data to look for any other patters that could show if a linear model would be the best fit for the data.

6.Conclusion

To conclude this report analyses the correlation between each data sets, the graphs show the time series analysis of the data and after an analysis where most the graph had similar changes. Most of the graphs had significant changes in 2008 due to the global financial crisis resulting in many increases in exchange rates and the decreasing of stock prices. However, the exchange rate between US and Japan managed to stay stable and didn't have any changes the price continued to decrease. Another common occurrence shown in the graphs is how Covid 19 affected the rates and stock, during that period all the exchange rates showed fluctuations in exchange rates and stock prices decreasing. Out of all the graphs the UK was affected the most shown by the increase in exchange rate as well the huge drop in stock prices.

The correlation analysis shows that most of the comparisons have a correlation with each other. Most of the hypothesis tests prove that there is a correlation by providing evidence that rejects the null hypothesis. However out of all the 15 comparisons, 2 couldn't provide evidence to reject the null hypothesis at the 1% level but did at the 5% level and 1 comparison had a p-value greater than 5% meaning the alternate hypothesis was rejected. These all happened to be comparisons with USD_AUD and the other stock price data sets.

Lastly the regression analysis managed to find the best datasets for comparison, however after analysis the R squared values is seems like there have been better ways to analyse the datasets. Out of all the 30 comparisons only 2 were able to produce a high R squared value, these comparisons were both between Dow Jones and Nikkei. The high R squared value produced clearly shows that the analysis is linear and that the produced models show a good fit for the data. But since the rest of the data sets were only able to produce value around 50% and lower this shows that the model could also be non-linear and that should take other calculated information into account for example the residual data could also be used to show any other patterns hidden in the data.

Appendix

Appendix A.1 R code for loading data

```
library(tidy verse)
library(seasonal)
library(fpp2)

USD_GBP <- read_csv('USD_GBP Historical Data monthly.CSV')
USD_AUD <- read_csv('USD_AUD Historical Data monthly.CSV')
USD_JPY <- read_csv('USD_JPY Historical Data monthly.CSV')
Dow_Jones<- read_csv('Dow Jones Industrial Average Historical Price Data m
onthly.CSV')
Nikkei_225 <-read_csv('Nikkei 225 Historical Price Data monthly.CSV')
FTSE_100 <- read_csv('FTSE 100 Historical Price Data monthly.CSV')
```

Appendix A.2 R code to change the dates in the dataset to the correct format.

```
USD_GBP[['Date']]<- as.Date(USD_GBP[['Date']],format = "%d/%m/%Y")
USD_AUD[['Date']]<- as.Date(USD_AUD[['Date']],format = "%d/%m/%Y")
USD_JPY[['Date']]<- as.Date(USD_JPY[['Date']],format = "%d/%m/%Y")
Dow_Jones[['Date']]<- as.Date(Dow_Jones[['Date']],format = "%d/%m/%Y")
Nikkei_225[['Date']]<- as.Date(Nikkei_225[['Date']],format = "%d/%m/%Y")
FTSE_100[['Date']]<- as.Date(FTSE_100[['Date']],format = "%d/%m/%Y")
```

Appendix A.3 R code that filters the data and selects only the price column.

```
a = USD_GBP %>% filter(Date >= '1982-01-01')
b = USD_AUD %>% filter(Date >= '1982-01-01')
c = USD_JPY %>% filter(Date >= '1982-01-01')
d = Dow_Jones %>% filter(Date >= '1982-01-01')
d = d %>% filter(day(Date) ==1)
e = Nikkei_225 %>% filter(Date >= '1982-01-01')
f = FTSE_100 %>% filter(Date >= '1982-01-01')
new_f = seq(as.Date('1982-01-01'),as.Date('2001-01-01'),by = 'month')
df = data.frame(Date = new_f)
f = bind_rows(f,df)
a = select(a,Price,Date)
b = select(b,Price,Date)
c = select(c,Price,Date)
d = select(d,Price,Date)
e = select(e,Price,Date)
f = select(f,Price,Date)
```

Appendix A.4 R code to plot the filtered datasets.

```r
ggplot()+
  geom_line(data = a, aes(x = Date, y=Price))+ggtitle('USD_GBP')

ggplot()+
  geom_line(data = b, aes(x = Date, y=Price))+ggtitle('USD_AUD')

ggplot()+
  geom_line(data = c, aes(x = Date, y=Price))+ggtitle('USD_JPY')

ggplot()+
  geom_line(data = d, aes(x = Date, y=Price))+ggtitle('Dow_Jones')

ggplot()+
  geom_line(data = e, aes(x = Date, y=Price))+ggtitle('Nikkei_225')

ggplot()+
  geom_line(data = f, aes(x = Date, y=Price))+ggtitle('FTSE_100')
```

Appendix A.5 R code filters the original data to a new set of dates and selects only the price column.

```r
GBP = USD_GBP %>% filter(Date >= '2001-02-01')
AUD= USD_AUD %>% filter(Date >= '2001-02-01')
JPY= USD_JPY %>% filter(Date >= '2001-02-01')
DowJones = Dow_Jones%>% filter(Date >= '2001-02-01')
DowJones =  DowJones%>% filter(day(Date) ==1)
Nikkei = Nikkei_225%>% filter(Date >= '2001-02-01')
FTSE = FTSE_100 %>% filter(Date >= '2001-02-01')
a = select(GBP,Price)
b = select(AUD,Price)
c = select(JPY,Price)
d = select(DowJones,Price)
e = select(Nikkei,Price)
f = select(FTSE,Price)
```

Appendix A.5 R code that creates a data frame combining the filtered column and then uses the data frame to conduct a correlation analysis and create a matrix of comparisons.

```r
test1 <- data.frame(c('USD_GBP'=a,'USD_AUD'=b,'USD_JPY'=c,'Dow_Jones'=d,'N
ikkei_225'=e,'FTSE_100'=f))


correlation::correlation(test1,include_factors = TRUE, method = "auto")
```

Appendix A.6 R code that creates a summary of the linear regressions for all the GBP comparisons

```r
#GBP comparisons:
summary(lm(GBP$Price~AUD$Price)) #Regression equation y = 0.15x + 0.47

summary(lm(GBP$Price~JPY$Price))#Regression equation  y = 0.002x+0.45

summary(lm(GBP$Price~DowJones$Price))#Regression equation  y = 8.38e-06x +
5.18e-01

summary(lm(GBP$Price~Nikkei$Price))#Regression equation  y = 9.09e-06x +
5.14e-01
```

```
summary(lm(GBP$Price~FTSE$Price))#Regression equation  y = 4.96e-05x + 3.6
3e-01

 (AUD$Price~FTSE$Price))#Regression equation y = -3.51e-05x +1.54
```

Appendix A.7 R code that creates a summary of the linear regressions for all the AUD comparisons

```
#AUD comparisons:
summary(lm(AUD$Price~GBP$Price))#Regression equation y = 0.94x +0.70

summary(lm(AUD$Price~JPY$Price))#Regression equation  y = 0.01x + 0.13

summary(lm(AUD$Price~DowJones$Price))#Regression equation  y = 2.51e-06x +
1.29

summary(lm(AUD$Price~Nikkei$Price))#Regression equation y = 5.90e-06x +1.2
3

summary(lm(AUD$Price~FTSE$Price))#Regression equation - y = -3.51e-05x +1.
54
```

Appendix A.8 R code that creates a summary of the linear regressions for all the JPY comparisons

```
#JPY comparisons:
summary(lm(JPY$Price~GBP$Price))#Regression equation  y = 50.86x + 75.53

summary(lm(JPY$Price~AUD$Price))#Regression equation  y = 43.55x +51.31

summary(lm(JPY$Price~DowJones$Price))#Regression equation  y = 7.37e-04 +
9.65e+01

summary(lm(JPY$Price~Nikkei$Price))#Regression equation y = 1.33e-03x+ 8.7
4e+01

summary(lm(JPY$Price~FTSE$Price))#Regression equation y = 4.07e-03x + 8.46
e+01
```

Appendix A.9 R code that creates a summary of the linear regressions for all the Dow Jones comparisons

```
#Dow_Jones comparisons:
summary(lm(DowJones$Price~GBP$Price))#Regression equation  y = 65018x -257
75

summary(lm(DowJones$Price~AUD$Price))#Regression equation  y = 3016x+13334

summary(lm(DowJones$Price~JPY$Price))#Regression equation  y = 222.41-6953
.54

summary(lm(DowJones$Price~Nikkei$Price))#Regression equation  y= 1.21x - 2
.50e+03
```

```
summary(lm(DowJones$Price~FTSE$Price))#Regression equation y = 6.32x -2.09
e+04
```

Appendix A.10 R code that creates a summary of the linear regressions for all the Nikkei 225 comparisons

```
#Nikkei_225 comparisons:
summary(lm(Nikkei$Price~GBP$Price))#Regression equation  y = 43680x -12538

summary(lm(Nikkei$Price~AUD$Price))#Regression equation  y = 4328x +10601

summary(lm(Nikkei$Price~JPY$Price))#Regression equation  y=248.85x -10756.
58

summary(lm(Nikkei$Price~DowJones$Price))#Regression equation y = 7.48e-01x
+3.46e+03

summary(lm(Nikkei$Price~FTSE$Price))#Regression equation  y = 5.00x - 1.39
e+04
```

Appendix A.11 R code that creates a summary of the linear regressions for all the FTSE 100 comparisons

```
#FTSE_100 comparisons:
summary(lm(FTSE$Price~GBP$Price))#Regression equation  y = 6281.1x + 1895.
2

summary(lm(FTSE$Price~AUD$Price))#Regression equation  y = -687.8x + 6976.
4

summary(lm(FTSE$Price~JPY$Price))#Regression equation  y = 20.04x + 3871.1
1

summary(lm(FTSE$Price~DowJones$Price))#Regression equation  y = 1.03e-01x
+ 4.27e+03

summary(lm(FTSE$Price~Nikkei$Price))# Regression equation y = 1.32e-01x +
3.89e+03
```