**Methodology**

**1. Data Collection**

The data collection will focus on 10 years’ timeline from 2011 to 2020 focusing on the UK market index in the London Stock Exchange (LSE) comparing with the data on the UK EPU index and the trading volume in the LSE. The data collection will be on secondary data approach where the data is obtained through the available primary data through public website including Yahoo Finance for London Stock Exchange (LSE) index, Global EPU Index platform for UK EPU index and financial website for investor sentiment. In addition, in order to control the potential impact of various macro indicators on the model, using the Consumer Price Index (CPI), and Treasury bond yield () as control variables.

**2. Data Pre-processing**

We use the consumer confidence index (CCI) of the United Kingdom (<https://data.oecd.org/leadind/consumer-confidence-index-cci.htm>) as a measure of investor sentiment based on previous studies. Based on responses to questions about their anticipated financial condition, their sentiment about the general economic situation, unemployment, and savings capability, the consumer trust indicator predicts future trends in household consumption and saving. An indicator above 100 indicates that consumers are more optimistic about the potential economic situation, and as a result, they are less likely to save and more likely to spend money on big purchases in the coming year. Indicator below 100 reflect a negative outlook on potential economic developments, likely leading to a desire to save more and spend less.

Meanwhile, the official website (<http://www.policyuncertainty.com>) will be used to collect the EPU index. The index is focused on articles in newspapers about policy uncertainty. The Times, The FT, The Telegraph, Sunday Times, The Guardian, The Daily Mail, The Mirror, The Daily Express, The Sun, The Evening Standard, and The Northern Echo are among the 11 UK newspapers represented. The phrases uncertainty, economic or market, and policy related terms are used in the news articles. ' tax,' ' regulation,"spending,"policy,' are words that are applicable to policy.

The FTSE 100 index (<https://www.londonstockexchange.com/indices/ftse-100>) is an index of the 100 largest companies listed on the London Stock Exchange as the representative of the return of stock market in UK. The index is a barometer of the British economy and one of the most important stock indexes in Europe. The return of stock market is measured by the increase or decrease of the index in a monthly period.

As for the collection of control variables, CPI index will be collected as one of them, a macroeconomic indicator that reflects changes in the price level of consumer goods and services generally purchased by households, to represent the influence of social economic influencing factor to stock return. Besides, according to the Capital Asset Pricing Model (CAPM), relationship exerts between systematic risk and expected return for assets, and this is particularly common when it comes to the measurement for the return of stocks. Therefore, the return of a risk-free bond (systematic risk) should also be considered as a control variable and I obtain the annual yield of the UK treasury bond to act as the systematic risk in the research.

The regression test will be including in the research methodology as this is crucial in explaining the phenomena of the significant relationship of the independent variable in EPU index and investor sentiment against the dependent variable in the market return (Demir & Ersan, 2018). The stepwise regression model will be used to predict the value of the dependent variable and estimate the effect of some explanatory variable on the dependent variable. The proposed research methodology will emphasize on the use of Rstudio software in generating the statistical data that will help in interpretation and hypothesis testing to draw the finding for this research.

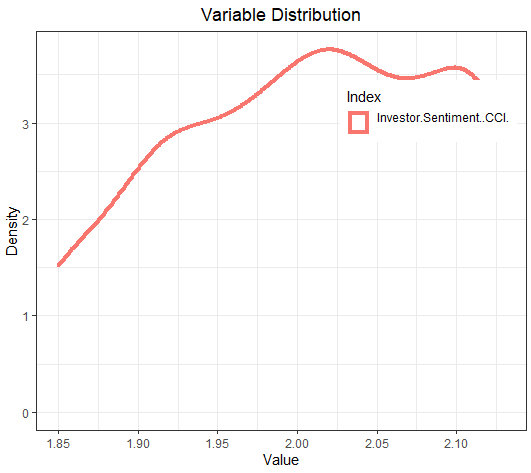
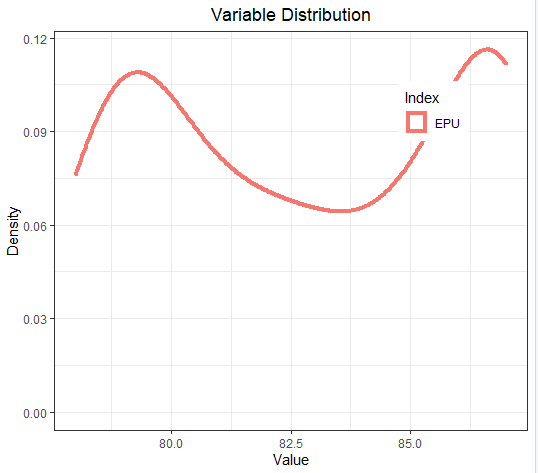
**3. Data Analysis**

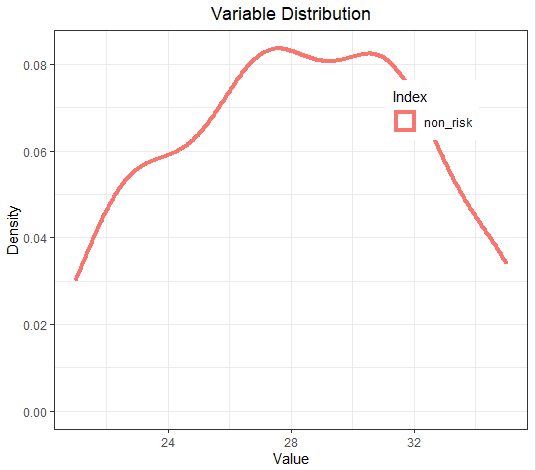
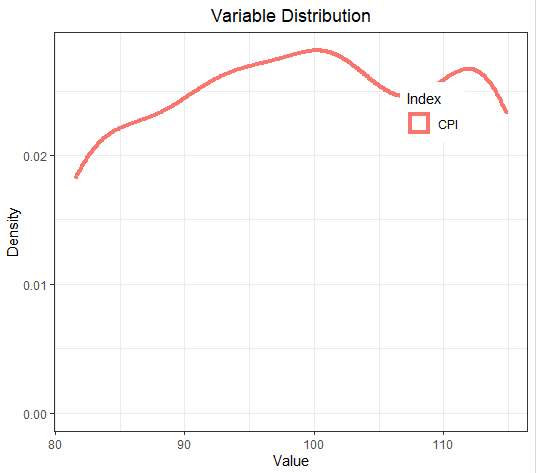
According to the Table 1. Out of five indexes, CCI have the most steadily fluctuations with a 0.0864 standard deviation. CPI and Rf are close with 2.099 and 3.675 standard deviation . We can see that Rt went through tremendous fluctuation during the ten years, with a mean of 0.46 and a standard deviation of 27.09. Maximum Market Return rate is 89.48 while Minimum Market return rate is 3.75. CPI is the only index that has a negative mean while all others have positive ones. CPI and Rf are the only two indexes that have negative minimum.

**Table 1**

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Variable | Min | Max | Mean | Std | 1th Qu | 3rd Qu |
|  | 78.138% | 112.517% | 82.16% | 76.384% | 81.461% | 92.494% |
|  | -22.568% | 1.016% | -0.030% | 2.099% | 0.01% | 0.38% |
|  | 1.85% | 2.13% | 2.07% | 0.0864% | 1.93% | 2.08% |
|  | -13.808% | 12.352% | 0.143% | 3.675% | 31.68% | 57.4% |
|  | 3.75% | 89.48% | 46.35% | 27.09% | 20.65% | 67.78% |

**Figure 1**





According to Table1, The distribution of EPU growth rate shows a binomial trend with a first peak around 75 and a second peak around 90. The distribution of CCI growth rate indicates a right-skewed pattern. The distribution of CPI also has a binomial trend with two peaks located around 100 and 115.

**Figure 3**

图表

描述已自动生成

Figure 3 indicates that the return of the FTSE 100 index fluctuated slightly around 0. However, the annual yield of the UK Treasury Bond behaves the opposite. It dropped from 0.9 in 2011 to 0.2 around mid 2012, then fluctuated around 0.4 between 2013 and 2016. Afterwards it went from 0.2 in 2017 to 0.8 in 2018 in a sudden. Finally it had a second major drop during late 2019 to early 2020.

**Pearson**

|  |  |  |  |
| --- | --- | --- | --- |
|  | t | p-value | Correlation |
| EPU | 7.0954 | 1.004e-10 | 0.5452451 |
| CCI | -1.039 | 0.3009 | -0.094814 |
| CPI | -0.68989 | 0.4916 | -0.06311563 |
|  | 6.3202 | 4.719e-09 | 0.5013136 |

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  |  | EPU | CCI | CPI |  |
|  | 1 |  |  |  |  |
| EPU | 0.54524513 | 1 |  |  |  |
| CCI | -0.09481488 | -0.04781042 | 1 |  |  |
| CPI | -0.06311563 | 0.05956421 | 0.79072931 | 1 |  |
|  | 0.50131358 | 0.09650454 | -0.18550843 | -0.04696916 | 1 |

Through covariance analysis, it can be concluded that the correlation coefficients of EPU, CCI, CPI and Rf are 0.54524514, -0.09481488, -0.06311563, 0.50131358. EPU and Rf have positive correlation with Rt while CI and CPI have negative correlations. CCI has negative correlation with EPU while CPI and Rf have positive correlations. R have negative correlation with both CCI and CPI.

**Unit Root Test**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | EPU | CCI | CPI |  |
| **ADF (1st difference)** | -3.2282 | -0.2992 | -0.613 | -0.5693 |
| **PP (level)** | 0.2145 | 0.6109 | 0.544 | 0.5529 |
| **PP (1st difference)** | 0.001438 | 0.5761 | 0.4499 | 0.4687 |

The Augmented Dickey Fuller Test(ADF) is unit root test for stationarity. Unit roots can cause unpredictable results in your time series analysis.

|  |  |
| --- | --- |
| > adf.test(CPI)  Augmented Dickey-Fuller Test  data: CPI  Dickey-Fuller = -3.403, Lag order = 4, p-value = 0.05739  alternative hypothesis: stationary | > adf.test(EPU)  Augmented Dickey-Fuller Test  data: EPU  Dickey-Fuller = -2.2438, Lag order = 4, p-value = 0.4753  alternative hypothesis: stationary |
| > adf.test(CCI)  Augmented Dickey-Fuller Test  data: CCI  Dickey-Fuller = -4.1766, Lag order = 4, p-value = 0.01  alternative hypothesis: stationary | > adf.test(Rf)  Augmented Dickey-Fuller Test  data: Rf  Dickey-Fuller = -4.8326, Lag order = 4, p-value = 0.01  alternative hypothesis: stationary |
| > adf.test(Rt)  Augmented Dickey-Fuller Test  data: Rt  Dickey-Fuller = -4.71, Lag order = 4, p-value = 0.01  alternative hypothesis: stationary | |

|  |  |
| --- | --- |
| Null Hypothesis | Alternative Hypothesis |
| H0 *– There is a unit root* | H1 *–* The time series is stationary |

Since only the p value for EPU and CPI are larger than 0.05 and the other three variable is smaller than 0.05, we are going to say that we fail to reject the null hypothesis for EPU and CPI. We reject the null hypothesis for Rt/Rf/CCI.

**4. Regression Model**

The establishment of the regression model will refer to the research of Rehman and Apergis (2019).

. (Rehman and Apergis, 2019)

*= Monthly return rate of the FTSE 100 index;*

*CCI = Consumer confidence index (investor sentiment);*

*EPU = Economic policy uncertainty;*

*CPI = Consumer Price Index*

*=* *Treasury bond yield.*

**5. Hypothesis Testing**

Three Hypothesis Tests are performed in this research.

|  |  |
| --- | --- |
| *H1* | *Whether there is significant relationship between EPU and the return of stock market* |
| *H2* | *Whether there is significant relationship between investor sentiment and return on stock market* |
| *H3* | *Whether there is a positive correlation between EPU, investor sentiment and return of stock market* |

Dependent Variable

Independent Variable

Economy Policy Uncertainty Index

Return on stock market

Investor Sentiment

**5.1 Testing of Hypothesis 1**

Simple Linear Regression Model between Rt and EPU was performed. The equation was stated as below:

|  |  |
| --- | --- |
| Null Hypothesis | Alternative Hypothesis |
| H0 *- There is no significant relationship between EPU and return of stock market*. | H1 *- There is a significant relationship between EPU and return of stock market*. |

The regression results are shown as following:

Call:

lm(formula = Rt ~ EPU, data = ER\_Final\_Data)

Residuals:

Min 1Q Median 3Q Max

-0.57387 -0.13868 0.03098 0.17142 0.49304

Coefficients:

Estimate Std. Error t value Pr(>|t|)

(Intercept) -3.241855 0.520334 -6.230 7.28e-09 \*\*\*

EPU 0.044517 0.006274 7.095 1.00e-10 \*\*\*

---

Signif. codes: 0 ‘\*\*\*’ 0.001 ‘\*\*’ 0.01 ‘\*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 0.2281 on 119 degrees of freedom

Multiple R-squared: 0.2973, Adjusted R-squared: 0.2914

F-statistic: 50.34 on 1 and 119 DF, p-value: 1.004e-10

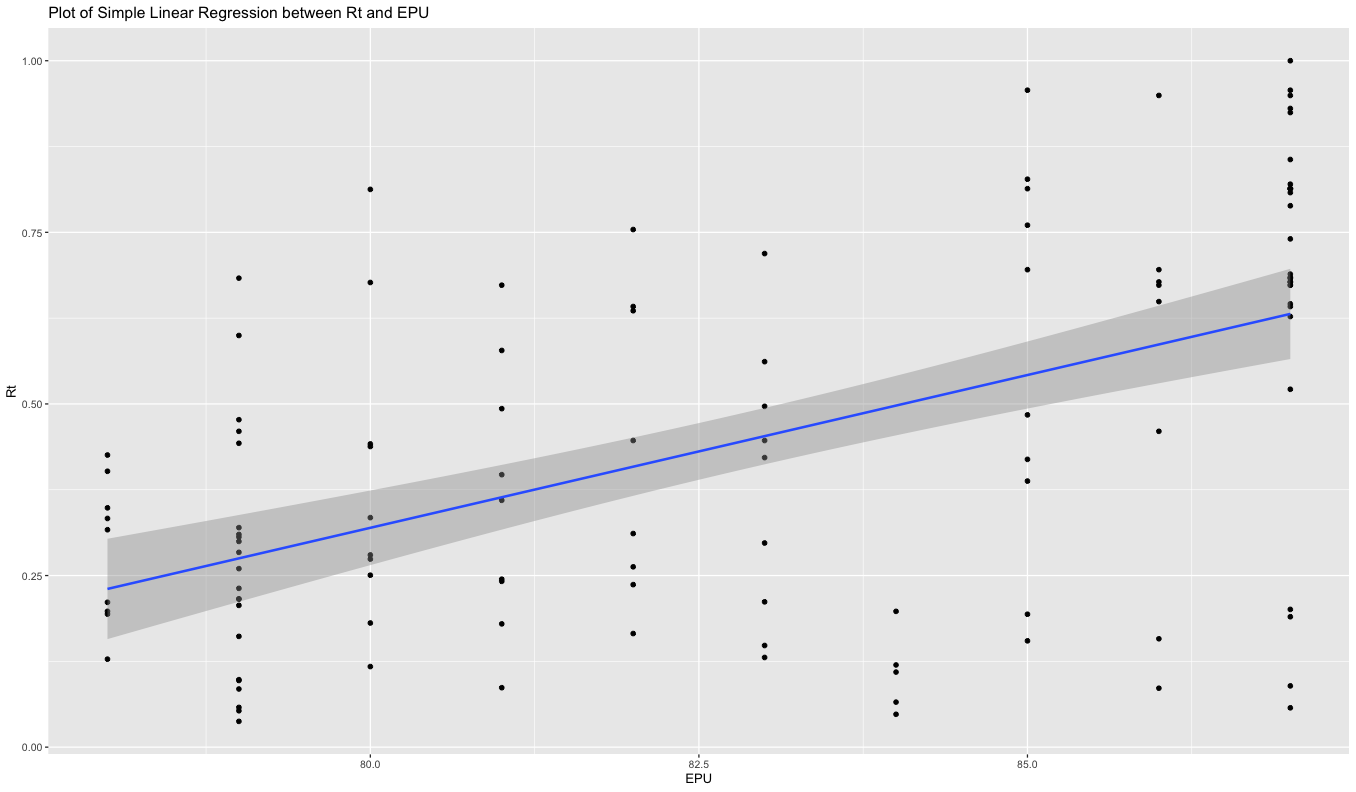
The p-value for the F-statistics is 1.004e-10, which means the model is highly statistically significant. This means that explanatory variables are significantly related to the responsive variable. Since the p value is smaller than 0.05, we reject the null hypothesis. Multiple R-squared and Adjusted R-squared are 0.2973 and 0.2914 respectively, which indicate a bad fit of the data and model.

> ggplot(model1, aes(x=EPU, y=Rt)) +

+ geom\_point() +

+ geom\_smooth(method=lm)

+ ggtitle(“Plot of Simple Linear Regression between Rt and EPU”)

****

**5.2 Testing of Hypothesis 2**

Multiple Linear Regression Model of Rt and CCI/CPI was performed. The equation was stated as below:

|  |  |
| --- | --- |
| Null Hypothesis | Alternative Hypothesis |
| H0 *- There is no significant relationship between investor sentiment and return on stock market.* | H1 *- There is significant relationship between investor sentiment and return on stock market.* |

The regression results are shown as following:

Call:

lm(formula = Rt ~ CCI + CPI, data = ER\_Final\_Data)

Residuals:

Min 1Q Median 3Q Max

-0.41424 -0.24149 -0.01593 0.24571 0.52829

Coefficients:

Estimate Std. Error t value Pr(>|t|)

(Intercept) 1.1239208 0.6904294 1.628 0.106

CCI -0.3757739 0.4693455 -0.801 0.425

CPI 0.0007809 0.0036941 0.211 0.833

Residual standard error: 0.272 on 118 degrees of freedom

Multiple R-squared: 0.009365, Adjusted R-squared: -0.007425

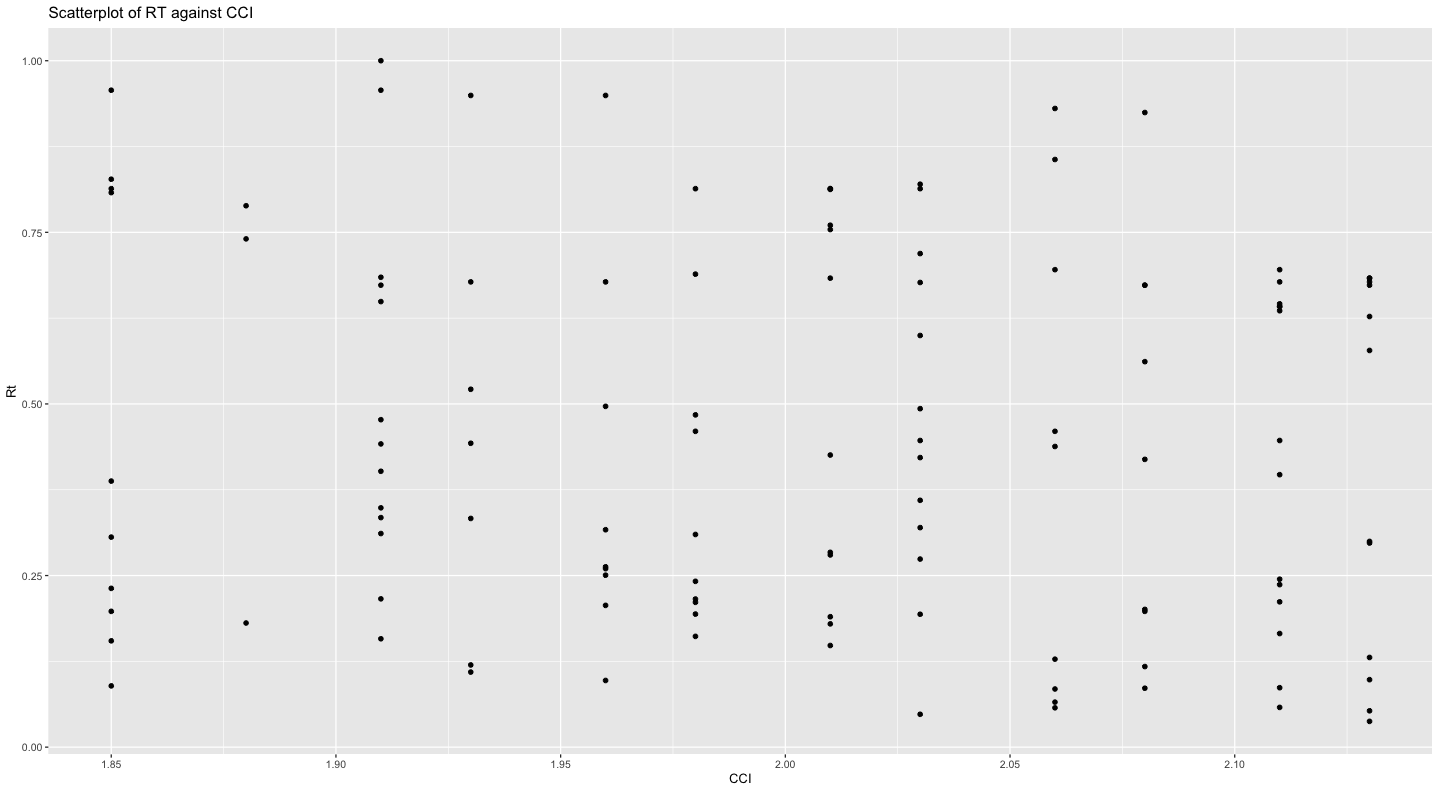
F-statistic: 0.5578 on 2 and 118 DF, p-value: 0.574

The p-value for the F-statistics is 0.574, which means the model is not statistically significant. Since the p value is larger than 0.05, we fail to reject the null hypothesis. There is no significant relationship between Rt and CCI and CPI. Multiple R-squared and Adjusted R-squared are 0.009365 and -0.007425 respectively, which indicate almost no fitting pattern between the data and model.

> ggplot(model2,aes(x=CCI, y=Rt))

+ geom\_point()

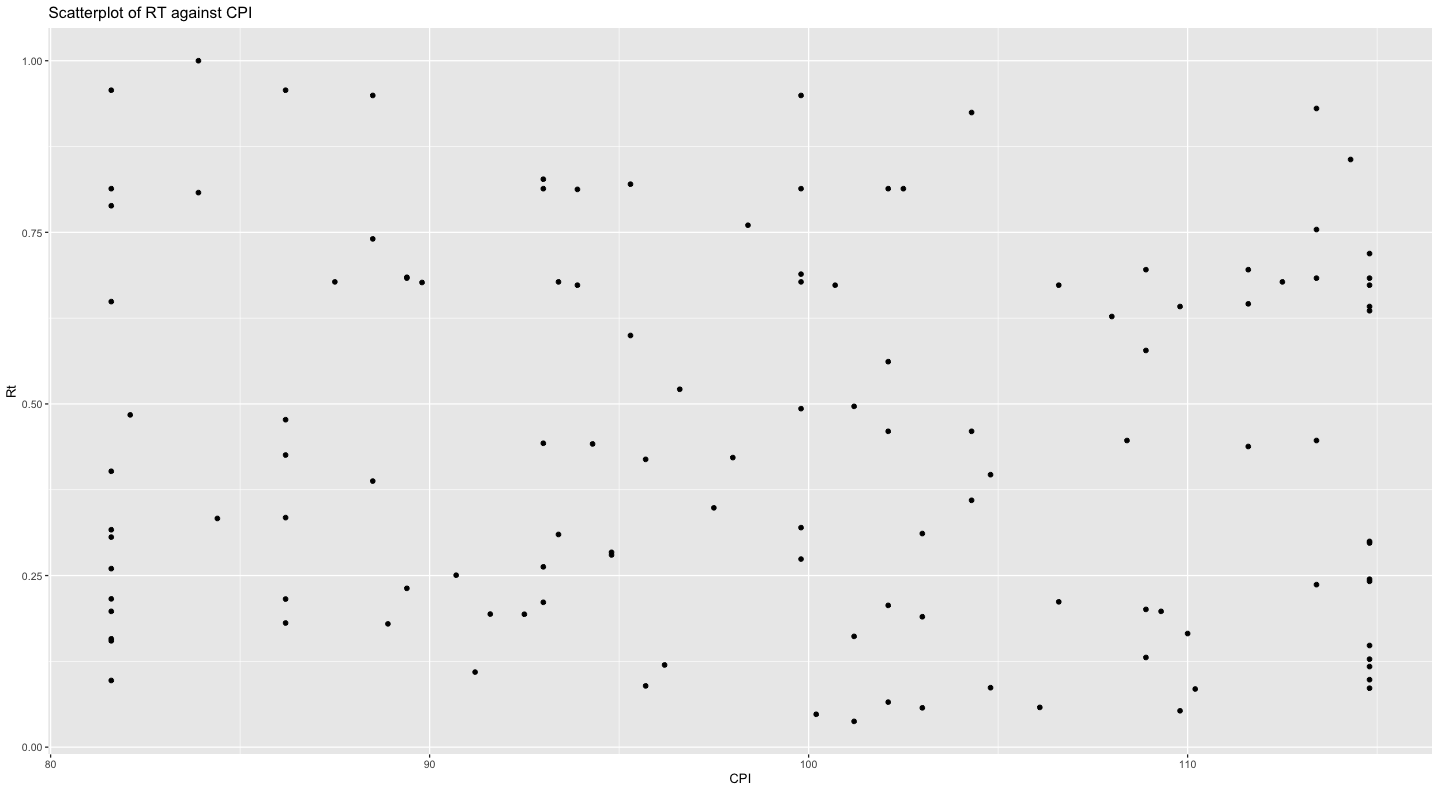
+ ggtitle("Scatterplot of RT against CCI")



> ggplot(model2,aes(x=CPI, y=Rt))

+ geom\_point()

+ ggtitle("Scatterplot of RT against CPI")



**5.3 Testing of Hypothesis 3**

Multiple Linear Regression analysis was performed between Rt and EPU/CPI/CCI/Rf. The equation was stated as below:

The hypothesis test was stated as following:

|  |  |
| --- | --- |
| Null Hypothesis | Alternative Hypothesis |
| H0*: There is no correlation between EPU, investor sentiment and return of stock market*. | H1: *There is a positive correlation between EPU, investor sentiment and return of stock market*. |

The regression results are shown as following:

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | Estimate | Std. Error | **t value** | **Pr(>|t|)** |
| (Intercept) | -4.722845 | 0.733061 | -6.443 | 2.78e-09 |
| EPU | 0.042667 | 0.005339 | 7.992 | 1.09e-12 |
| CCI | 0.631954 | 0.342929 | 1.843 | 0.0679 |
| CPI | -0.005706 | 0.002660 | -2.145 | 0.0340 |
|  | 0.032966 | 0.004595 | 7.175 | 7.34e-11 |

|  |  |
| --- | --- |
| Residual standard error | 0.191 on 116 degrees of freedom |
| Multiple R-squared | 0.5197 |
| Adjusted R-squared | 0.5032 |
| F-statistic | 31.38 on 4 and 116 DF |
| p-value | < 2.2e-16 |

The p-value for the F-statistics is < 2.2e-16, which means the model is highly statistically significant. This means that at least one of the four explanatory variables is significantly related to the responsive variable. Since the p value is smaller than 0.05, we reject the null hypothesis.

Multiple R-squared and Adjusted R-squared are 0.5197 and 0.5032 respectively, which indicate a mediocre fit of the data and model. 3 explanatory variables have less than 0.05 p values while CCI have a p value that is slightly over 0.05. The p value of EPU(1.09e-12) and Rf(7.34e-11) is significantly smaller than CCI(0.0679) and CPI(0.0340). 3 explanatory variables have positive estimate values while CPI have an estimate that is slightly under 0.

In summary, changing in EPU and Rf are more significantly associated with changing in Rt while changing in CCI and CPI are slightly significantly associated with changing in Rt. EPU, Rf  and Rf arepositively related with Rt while CPI is negatively related with Rt.

**6. Conclusion**

The relationship between EPU and Rt is statistically significant since the p value is way smaller than 0.05. However, a Multiple R-squared of 0.2973 and an Adjusted R-squared of 0.2914 indicate that the relationship isn’t necessarily linear. Further regression model can be performed between EPU and Rt.

The relationship between CCI, CPI and Rt is not statistically significant since the p value is above 0.05. A Multiple R-squared of 0.009365 and an Adjusted R-squared of -0.007425 indicate that the relationship did not exist between the model and data.

Finally, a multiple regression model is performed between Rt and EPU/CCI/CPI/Rf, from the result it can be analysed that EPU and Rf are more significantly associated with changing in Rt when compared to CCI and CPI.

In summary, Investor sentiment plays a minor to minimal role in Return on Stock Market. Economic Policy Uncertainty and Treasury bond yield have way more significant relationship with Return on Stock Market when compared to investor sentiment. Nevertheless, from the result of our model, the relationship between them is not necessarily linear. Further model including logistic, polynomial or other type of regression can be performed to determine the exact type relationship between Return on Stock Market and Economic Policy Uncertainty as well as Treasury bond yield.