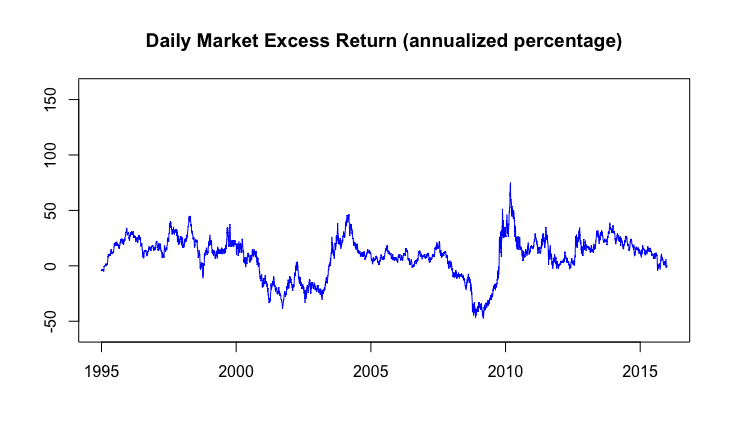
Fin 654

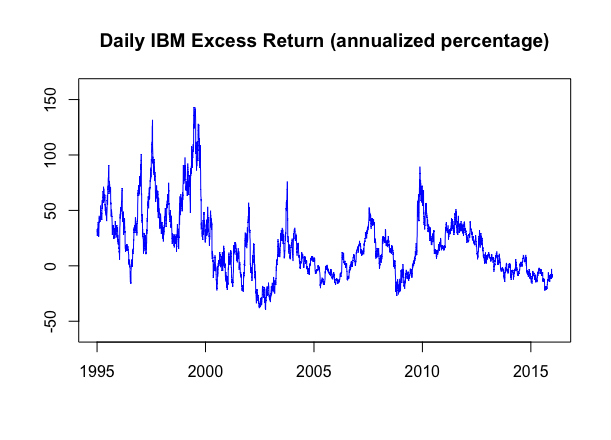
Financial Analytics Tianle Chen, Yankun Wang, Henglong So, Jiarong Yang

Case 1: Capital Asset Pricing Model

**Step 1.3 FIG Time-Series Plot of Yearly Returns**

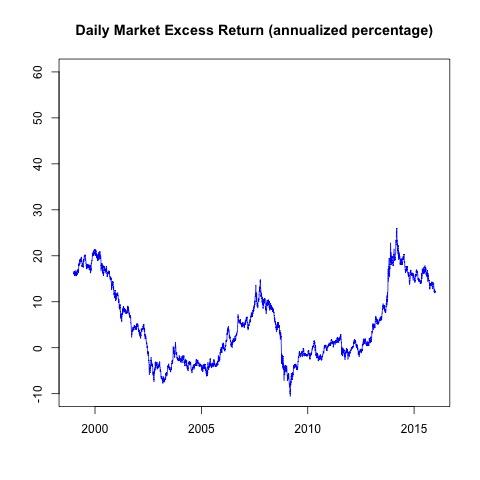


The first plot shows that annualized market excess returns (percentage) at a daily frequency, from 1994-01-03 to 2015-12-31. The horizontal axis present the years of investment and the vertical axis shows the percentage daily return. According to data, the stock price of market took a sharp downturn period in 2000-2003 and 2008-2009, which might due to overall market crisis. Recall during 2001-2002, it was a large bear market due the internet bubble bursting, and Financial Crisis Happened in 2008-2009 with a crisis in the subprime mortgage market in the United States. However the market excess return over the 20 years investment horizon has yield in a range between -47.49% to 75%.

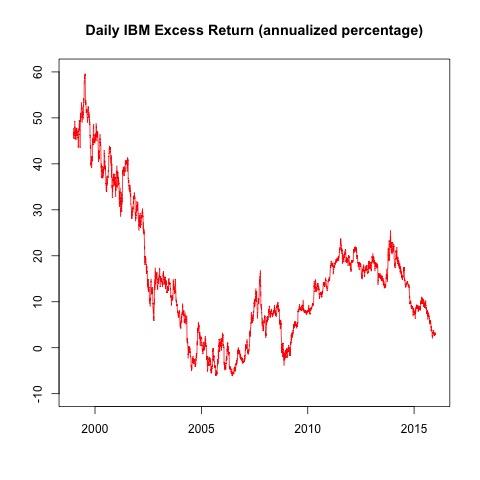


The second plot shows that annualized IBM excess returns ( percentage) at a daily frequency, from 1994-01-03 to 2015-12-31. The horizontal axis present the years and the vertical axis shows the daily excess returns. In the graph, we can see that a decline in the returns in 2002-2004. Which is correlating same trend as market excess returns did. The annualized return of the IBM reached its peak level in 1999-06-21, it was 143.1283% from inception date. After the dip in 2002, we see IBM has a larger bounce back in its excess returns than that of market returns in 2003, where IBM has grew its sale at double digits growth rate. IBM excess return over 20 year investment horizon stayed in range between -39.86% and 143.12% over this 20 years.

**Step 1.5 Time-Series Plot of Five-year of annualized Return**

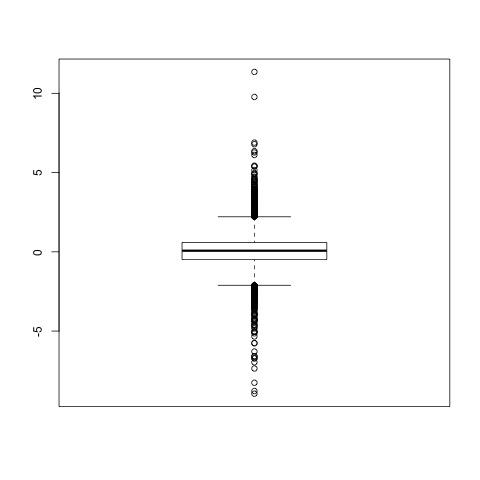


The first plot shows that annualized market excess returns ( percentage) over five-year rolling at a daily frequency, from 1994-01-03 to 2015-12-31. The horizontal axis present the years of investment and the vertical axis shows the percentage daily return. According to data, the stock price of market took a sharp downturn period in 2000-2003 and 2008-2009, which might due to overall market crisis. Recall during 2001-2002, it was a large bear market due the internet bubble bursting, and Financial Crisis Happened in 2008-2009 with a crisis in the subprime mortgage market in the United States. Generally, the market excess return have mostly stayed in range between -10% and 20% over this 20 years.

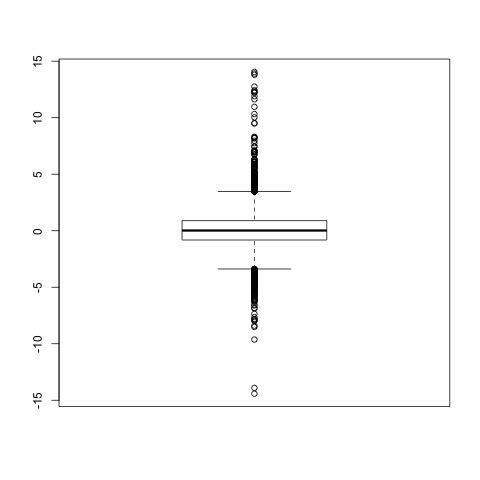


The second plot shows that annualized IBM excess returns ( percentage) over five-year rolling at a daily frequency, from 1994-01-03 to 2015-12-31. The horizontal axis present the years and the vertical axis shows the excess returns. In the graph, we can see that a decline in the returns in 2000-2004 and 2008-2009. The annualized return of the IBM reached the highest rate on 1999-07-15, it was 59.64612%. In the period from 2000 to 2003, IBM has a larger bounce in its excess returns than that of market returns for most of the time. Generally, the IBM excess return have mostly stayed in range between -10% and 60% over this 20 years.

**Step 2.1 FIG Plot**



The “Daily Market Excess returns” box plot illustrates that the median of daily market excess return is also close to zero, which means that the proportions for both negative and positive market excess return is almost equal. The middle 50% of the market excess returns range from around -0.25% as first quartile to 0.25% as third quartile. The maximum and minimum daily Market excess return is close to 11% and -9% respectively, much higher or lower than those of market excess returns.



The “Daily IBM excess return “box plot illustrates that the median of daily market excess return is also close to zero, which means that the proportions for both negative and positive market excess return is almost equal. The middle 50% of the market excess returns range from around -0.4% as first quartile to 0.4% as third quartile. Also, unlike that for market excess return, the data spread wildly in the upper outlier than that in the lower one. The maximum and minimum daily IBM excess return is close to 15% and -15% respectively, much higher or lower than those of market excess returns. Compared with index investors, IBM shareholders more likely to gain higher excess return but in the meanwhile they will take risk to loss more.

**Step 2.3 Numerical Moments**

|  |  |  |
| --- | --- | --- |
|  | IBM | Market |
| Mean | 0.0690 | 0.0313 |
| Std | 1.8165 | 1.1866 |
| Skewness | 0.5800 | -0.1169 |
| Kurtosis | 8.0602 | 7.6395 |
| Correlation | 0.5956 | NA |

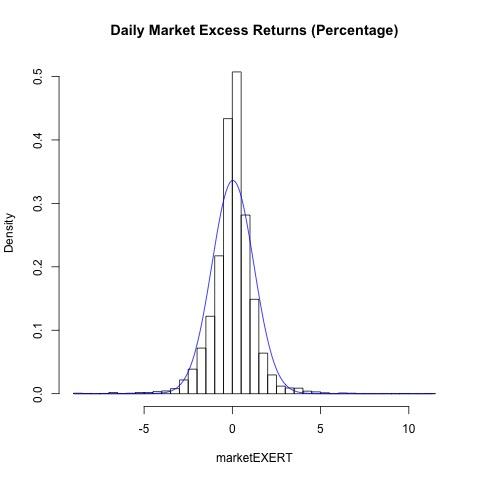
What the table shows above are the mean, standard deviation, skewness, kurtosis and correlation of IBM daily excess returns and market daily excess returns.   
**Comparison:**  
**Mean**: IBM’s mean is 0.0690, market’s mean is 0.0313, which means the IBM has a higher average excess return than the market, i.e. IBM is outperformance relative to the markets.  
**Std**: Standard deviation shows the dispersion of the dataset. IBM’s std is 1.8165 and the market’s std is 1.1866, which means that the IBM stock price is more volatile than the market stock prices,which give more risk to the investor.

**Skewness**: Skewness is a measure of the degree of asymmetry of a distribution. IBM’s skewness is 0.5800, it has a positive and higher value in absolute value, and the market’s skewness is -0.1169, which means that the IBM has a longer tail than the market and it’s on the right. While market has a shorter tail on the left. The longer the tail, the more the data disperse from the mean.

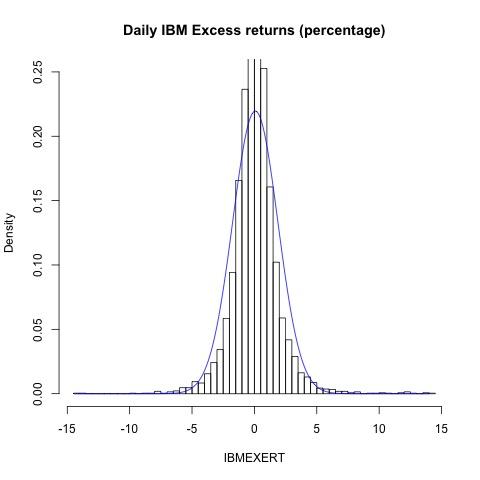
**Kurtosis**: kurtosis is a measure that describes the shape of a distribution's tails in relation to its overall shape. The kurtosis of IBM is 8.0602, the kurtosis of market is 7.6395. A higher kurtosis means that the IBM has a longer tail, i.e. more outliers than the market.

**Correlation**: measures the degree that two variables move in relation to each other. A positive number between IBM and market means that those two has a positive relationship. in our study we found IBM and the market has a positive correlation at 0.5956 which explain us that IBM stock is moving same direction to the market.

**Step 3.1 Histogram**

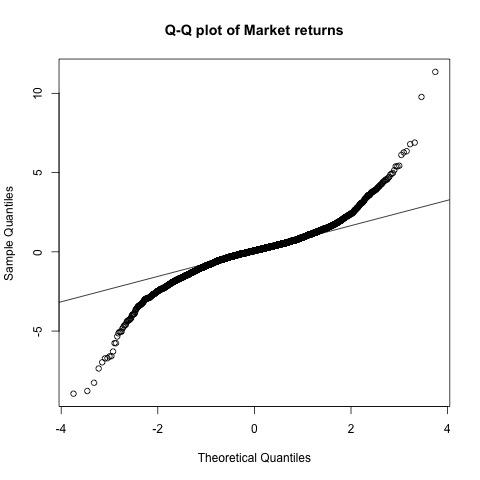


In the “Daily Market excess returns” histogram, the y-axis represents the frequency of the daily market excess returns in each bin and the x-axis represents daily market excess returns with 0.5%(?) as unit. The curve is the shape of standard normal distribution. The range with highest frequency in the histogram is from 0% to 0.5%. It is most likely for market shareholders to earn excess return form -1% to 1% that is more than index investors earn. We can see that the distribution of daily market excess return is more peak than that of normal distribution and has a fatter tail. Except the peak of the distribution, the rest of shape is along with the standard normal distribution. The frequencies on the right tail is higher than those on the left tail. In the extreme circumstance, the probability of gaining a positive excess return is greater than that of negative one. Compared with index investors, market shareholders more likely to gain higher excess return but in the meanwhile they will take risk to loss more.

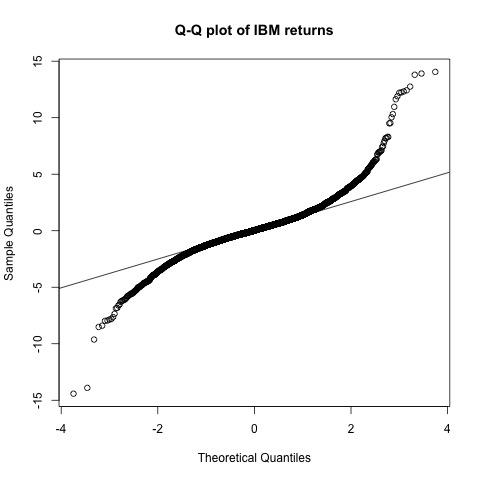


In the “Daily IBM excess returns “histogram, the y-axis represents the frequency of the daily IBM excess returns in each bin and the x-axis represents daily IBM excess returns with 0.5% as unit. The curve is the shape of standard normal distribution. The range with highest frequency in the histogram is from -1% to 1%. It is most likely for IBM shareholders to earn excess return form -1% to 1% that is more than index investors earn. We can see that the distribution of daily market excess return is more peak than that of normal distribution and has a fatter tail. Except the peak of the distribution, the rest of shape is along with the standard normal distribution. The frequencies on the right tail is higher than those on the left tail. In the extreme circumstance, the probability of gaining a positive excess return is greater than that of negative one. Compared with index investors, IBM shareholders more likely to gain higher excess return but in the meanwhile they will take risk to loss more.

**Step 3.2 QQ plot**



In the Q-Q plot of market excess returns, the horizontal axis shows the theoretical quantiles generated from the normal distribution while the vertical axis shows the quantile of market excess returns in our sample dataset. And the line crosses the dots means where the most of dots are if the data were normally distributed. From this plot, we learn that most dots are close to the theoretical quantiles line. The rest dots also draw an upward curve on the higher end and a downward curve on the lower end. In that case, data of market excess returns fit fat-tail distribution instead of normal distribution.



Same in the Q-Q plot of IBM market excess returns, the horizon axis shows the theoretical quantiles generated from the normal distribution while the vertical axis shows the quantile of IBM excess returns in our sample dataset. Most dots are close to the theoretical quantiles line, while the rest ones draw a curvature on the lowerend and higher end. It demonstrates that the data of IBM excess returns are distributed as fat-tail distribution instead of normal distribution, similar with the market excess returns.

However, it is obvious that more dots exist above the thermotical quantiles line in the plot of IBM excess return than the one of market excess return and the curvature is ever steeper. It means the tails of IBM plot is fatter than of market. IBM excess returns will be higher when on the same probability.

**Step 3.3 The Jarque-Bera Test**

Since some normality tests may be not reliable when n is large, Jarque-Bera test is usually run before one of these tests to confirm normality. In this step, we will conduct our null and alternative hypothesis based on the data of market excess returns and IBM excess returns

For market excess returns:

H0: The data of market excess returns comes from normal distribution.

H1: The data of market excess returns doesn’t come from normal distribution.

Decision Rules: Reject if p-value < 5%

Conclusion: After running the codes in R, we found the p-value <2.2e-16, which is obviously lower than 5%. In that case, we decided to reject the hypothesis, which means the data of market excess returns doesn’t come from normal distribution.

For IBM excess returns:

H0: The data of IBM excess returns comes from normal distribution.

H1: The data of IBM excess returns doesn’t come from normal distribution.

Decision Rules: Reject if p-value < 5%

Conclusion: After running the codes in R, we found the p- value <2.2e-16, which is obviously lower than 5%. In that case, we decided to reject the hypothesis, which means the data of IBM excess returns doesn’t come from normal distribution.

**Step 3.4 The Lilliefors test**

The Lilliefors test is also a test for normality. It is an improvement on the Kolomogorov-Smirnov test correcting the K-S for small values at the tails of probability distributions.

For market excess returns:

H0: The data of market excess returns comes from normal distribution.

H1: The data of market excess returns doesn’t come from normal distribution.

Decision Rules: Reject if p-value < 5%

Conclusion: After running the codes in R, we found the p-value <2.2e-16, which is obviously lower than 5%. In that case, we decided to reject the hypothesis, which means the data of market excess returns doesn’t come from normal distribution.

For IBM excess returns:

H0: The data of IBM excess returns comes from normal distribution.

H1: The data of IBM excess returns doesn’t come from normal distribution.

Decision Rules: Reject if p-value < 5%

Conclusion: After running the codes in R, we found the p- value <2.2e-16, which is obviously lower than 5%. In that case, we decided to reject the hypothesis, which means the data of IBM excess returns doesn’t come from normal distribution.

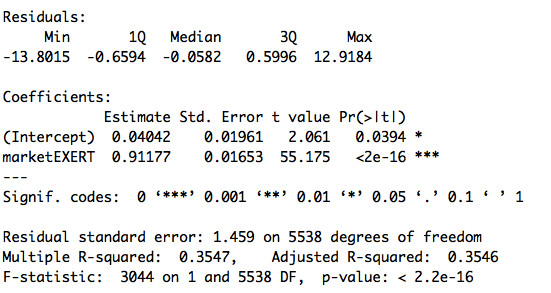
**Step 4 Model Setup**

In order to estimate CAPM, we have to set up a linear regression model following this equation based on case 1:



In this model, alpha is the intercept (representing estimated IBM daily return) and beta is the slope (estimated difference in the IBM excess return if market excess return happened to be 1 unit larger). With this regression, we can estimate the stock’s beta coefficient.

**Step 5 Model Estimation**



Estimated equation:



First of all, residuals are the difference between the actual value and the predicted value of IBM excess return. As shown above, the output of model residual is broken down to 5 summary points, which shows a symmetrical distribution (1Q+3Q close to zero, median close to zero and min+max close to zero) and it implies that predicted values are close to actual points.

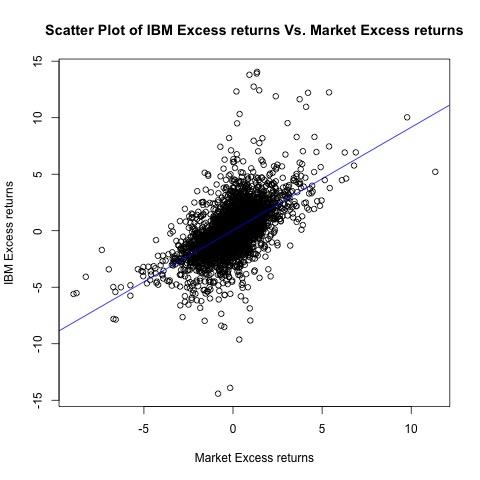
In this model, alpha equals 0.04042, which means the estimated daily average risk adjusted return of IBM is 0.04042%, and this number can vary by 0.0196% because standard error of intercept is 0.0196. In other word, the excess return of IBM would be 0.04042% if the market return is 0.

What’s more, the beta is 0.091177, which implies that for each additional 1% increase in the daily market excess return, we would expect an increase of 0.91177% in the daily excess return of IBM on average, and the increase could vary by 0.0165% as the standard error of slope shown.

The R-squared of a linear model describes the amount of variation in the response that is explained by the least squares line, and we can know how well the model is fitting the actual data known R-squared. In this case, 35.47% of the variance found in the IBM excess returns can be explained by the market excess returns.

Finally ,F-statistic and P-value of the model are good indicators of whether there is a relationship between our predictor and the response variables. The further the F-statistic is from 1 the better it is, and the closer the P-value is from 0 the better it is. In this case, F-statistic is 3044 which is definitely larger than 1 and the P-value is close to 0.

**Step 8 Plot the OLS Line**



This scatter plot in which the values of two variables are plotted along two axes, the pattern of the resulting points revealing any correlation present. In other words, a scatter plot can show if there is any connection between the two variables. This scatterplot has been labeled as illustrating a strong, positive, linear relationship between market excess return and IBM excess return.

**Step 9.1 Is the risk adjusted return zero?**

In this section we discuss uncertainty in the estimates of y-intercept for our regression model. The excess returns of stock i (yi,t+1 = Ri,t+1 − Rf, t) against the market excess returns (Xi, t+1 = Rm,t+1 − Rf, t) with an intercept αi . Since it is a two-side test, we will use t (df, 2.5%) to finish the test.

For the intercept αi generated in previous models:

H0 : αi = 0. The risk adjusted excess return is zero.

HA : αi ≠ 0. The risk adjusted excess return is not zero. It could either be positive or negative.

Decision Rules: Reject if t-value < t (df, 2.5%)

Conclusion: After running the codes in R, we found the t- value of intercept=2.061, which is obviously lower than t (df, 2.5%) = 1.960392. In that case, we decided to reject the hypothesis, which means the risk adjusted excess return is not zero.

**Step 9.2 Is the risk exposure higher than one?**

In this section we will discuss uncertainty in the estimates of the slope for our regression model. Since it is a one-side test, we will use t (df, 5%) to do the test

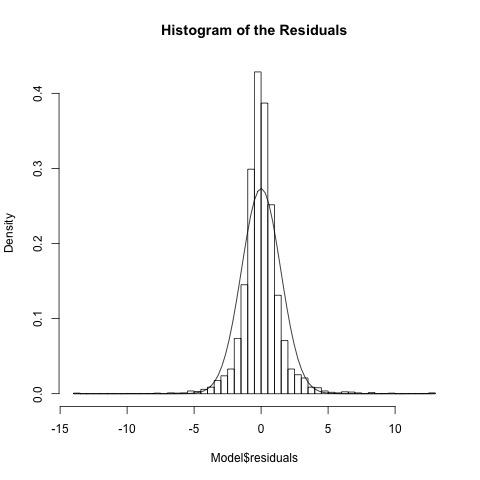
For the slope βi generated in previous models,

H0 : βi ≤ 1. The risk exposure is one.

HA : βi > 1. The risk exposure is higher than one.

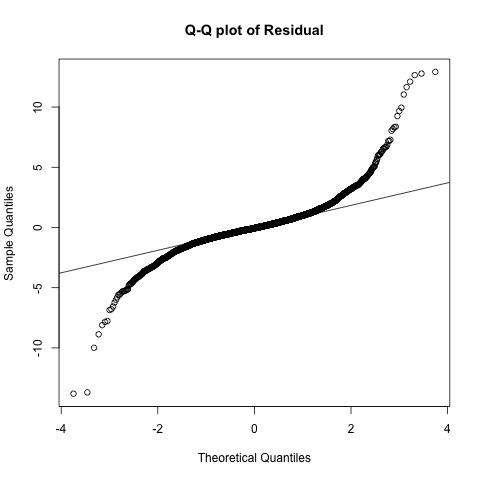
Conclusion: After running the codes in R, we found the t-value of slope =-5.338838, which is lower than t(df, 2.5%) = 1.645129. In that case, we can not reject the null test, which means the risk exposure is one.

**Step 11.1 Histogram of the Residuals**

****

In the “Histogram of the Residuals”, the y-axis represents the frequency of the daily IBM excess returns in each bin and the x-axis represents daily IBM excess returns with 0.5% as unit. The curve is the shape of standard normal distribution. The range with highest frequency in the histogram is from -1% to 1%. It is most likely for IBM shareholders to earn excess return form -1% to 1% that is more than index investors earn. We can see that the distribution of daily market excess return is more peak than that of normal distribution and has a fatter tail. Except the peak of the distribution, the rest of shape is along with the standard normal distribution. The frequencies on the right tail is higher than those on the left tail. In the extreme circumstance, the probability of gaining a positive excess return is greater than that of negative one. Compared with index investors, IBM shareholders more likely to gain higher excess return but in the meanwhile they will take risk to loss more.

**Step 11.2 QQ-plot of the Residuals**

****

In the Q-Q plot of residuals, the horizon axis shows the theoretical quantiles values generated from the normal distribution while the vertical axis shows the quantile of residuals generated from previous linear model. And the line crosses the dots means where the most of dots are if the data were normally distributed. From this plot, we learn that most dots are close to the theoretical quantiles line. The rest dots also draw an upward curve on the higher end and a downward curve on the lower end. In that case, data of residuals generated in previous regression models does not come from normal distribution.

**Step 11.3 The Jarque-Bera Test**

For the Jarque-Bera test for the disturbances term:

H0: The residuals of previous regression model comes from normal distribution.

H1: The residuals of previous regression model doesn’t come from normal distribution.

Decision Rules: Reject if p-value < 5%

Conclusion: After running the codes in R, we found the p- value <2.2e-16, which is obviously lower than 5%. In that case, we decided to reject the hypothesis, which means the residuals of previous regression model doesn’t come from normal distribution.

**Step 11.4 The Lilliefors test**

For the lilliefors test for the disturbances term:

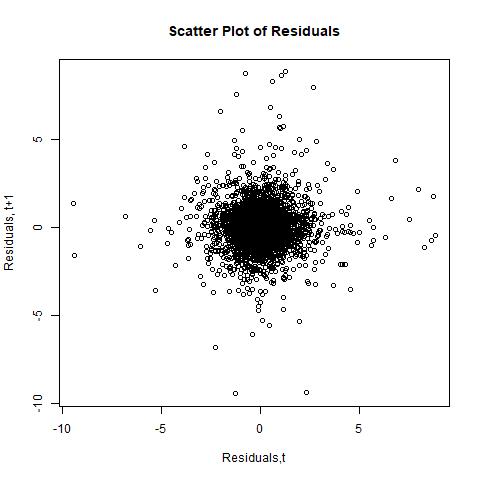
H0: The residuals of previous regression model comes from normal distribution.

H1: The residuals of previous regression model doesn’t come from normal distribution.

Decision Rules: Reject if p-value < 5%

Conclusion: After running the codes in R, we found the p - value <2.2e-16, which is obviously lower than 5%. In that case, we decided to reject the hypothesis, which means the residuals of previous regression model does not come from normal distribution.

**Step 12.1 Scatter plot of ei,t+1 vs ei,t**



This scatter plot in which the values of two variables are plotted along two axes, the pattern of the resulting points revealing any correlation present. In other words, a scatter plot can show if there is any connection between the two variables. base on the graph, we concluded that the residual t+1 and residuals t are not relate. we could say that the two sets of data have no correlation.

**Step 12.2 Durbin-Watson Test**

The Durbin-Watson test is used to test the null hypothesis that linear regression residuals are uncorrelated, against the alternative that autocorrelation exists.

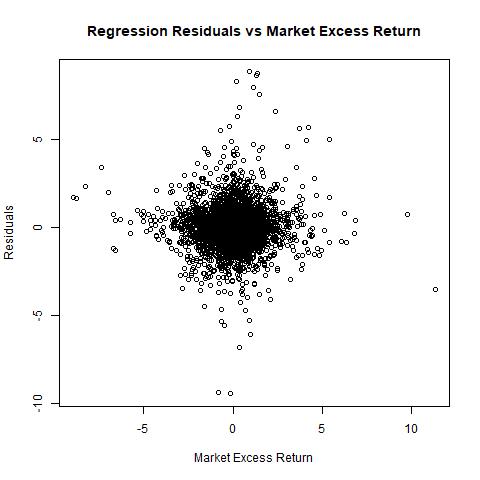
Conclusion:

In our testing we found P-value = 0.2771 which is greater than 0. it shows a significantly large p-value on the validity of the null hypothesis and indicated a high correlation among residuals.

**Step 12.3 Ljung-Box Q-Test**

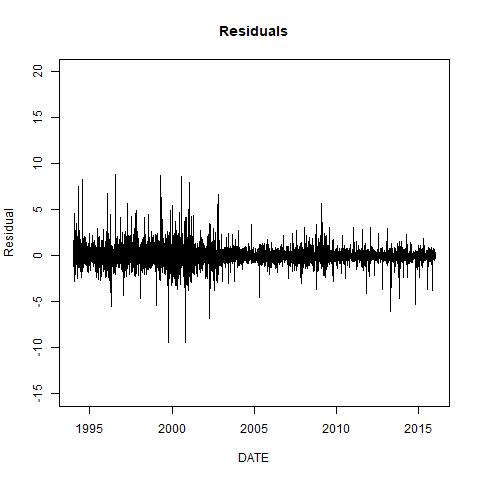
The Ljung-Box Q-test is a portmanteau test that assesses the null hypothesis that a series of residuals exhibits no autocorrelation for a fixed number of lags m, against the alternative that some autocorrelation coefficient.

**Step 13 Non-linearity**



Nonlinearity is a common issue when examining cause-effect relations. Such instances require complex modeling and hypothesis to offer explanations to nonlinear events. in our graph, we concluded that the residual and market excess return are not related. we could say that the two sets of data have no correlation.

**Step 14 Heteroskedasticity**



In our graph, we see the heteroscedasticity produced a distinctive fan shape in residual plot. the vertical range of the residuals decreases as the fitted values decrease.