FIN3080 Assignment2

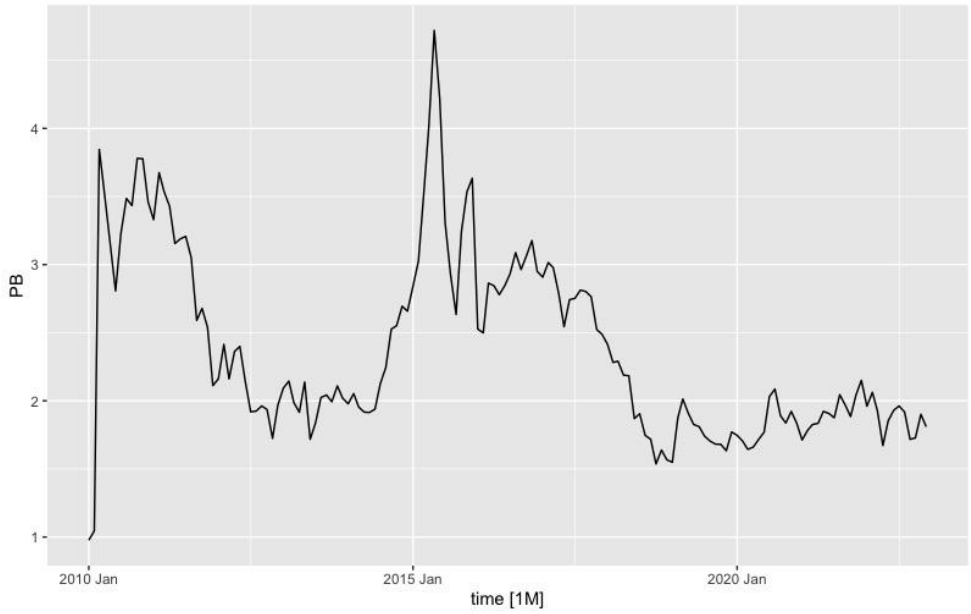
朱元昊 121090885

This assignment contains actually three part. The first part is to calculate the monthly P/B ratio, the second is to do a linear regression, and the third one is the construct portfolios and draw pictures.

For the first part, we need to download the datasets and calculate the P/B ratio. Here, I additionally downloaded the 'Total Market Value' field, to suit the formula of:

P/B = Total Market Value/Book Value = Total Market Value/ (Total Asset - Total Liabilities)

The question asks us to calculate the monthly P/B, however, we only have the quarterly Balance Sheet. We know that in each month, we only can use the already existed Balance Sheet, for example when we calculate April's P/B ratio we should use March's Balance Sheet. Thus, we can expand the month in the Balance Sheet, and use every last quarter's data to fill the next two months' blanks. Therefore we get the monthly Balance Sheet with total asset minus total liabilities to be the monthly Book Value. Then we merge the datasets, by the index of the month, we can calculate the P/B ratio. And we can draw the median P/B as the picture below:



For the second part, it asks us to do the linear regression of P/B ratio on ROE and Stock Volatility. Here, because we are only interested in the data and the relationships in 2010Q4, so I just download the data of ROE and Stock Volatility on 2010.12.31. For Stock Volatility, it represents the risk of the log return, we can divide it by square root of four, i.e. by half, to

be out quarterly volatility. For ROE, I notice that there are four types of ROEs, and from the descriptions we can see that they are all reasonable, just calculated by different Average Balance of Shareholders' Equity's time choices. So here I download them all and try two of them to see the results (Using ROE-A and ROE-TTM as examples, since the other two only have difference of the method of averaging). So now we can mutate the data sets with corresponding month, and merge the datasets of ROE and Stock Volatility and the calculated P/B ratio, use the linear regression function to do the linear regression and we can get the results below:

P/B regress on ROE-A and Stock Volatility:

```
Call:
lm(formula = PB ~ returnA + volatility, data = Q1)

Residuals:
     Min       1Q   Median       3Q      Max
-10.6805  -1.7073  -0.7744   0.8553  11.6291

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)   0.1328     0.4056   0.327    0.743
returnA       0.3139     0.6091   0.515    0.606
volatility    9.1767     0.8915  10.294   <2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 2.635 on 1167 degrees of freedom
  (14 observations deleted due to missingness)
Multiple R-squared:  0.08357,   Adjusted R-squared:  0.082
F-statistic: 53.21 on 2 and 1167 DF,  p-value: < 2.2e-16
```

Where 'returnA' means the ROE-A, and 'volatility' means the Stock Volatility. From the result we see that P/B is positively correlated to the Stock Volatility and ROE, and the R squared has achieved 8%. The result satisfies the formula of : $P/B=P/E * ROE$. However the coefficients' p-value are not so well, it has only a little that can let us beliece.

For regressing P/B on ROE-TTM and the Stock Volatility, we find that:

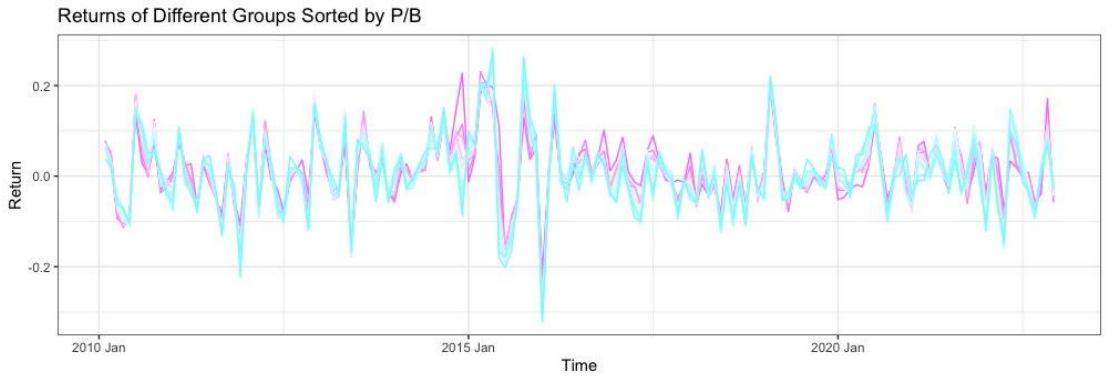
```
Call:
lm(formula = PB ~ returnTTM + volatility, data = Q1)

Residuals:
     Min       1Q   Median       3Q      Max
-10.8093  -1.6882  -0.7593   0.8623  11.7339

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)  0.08208    0.39888   0.206    0.837
returnTTM    0.55877    0.11256   4.964 7.92e-07 ***
volatility   9.19331    0.88219  10.421  < 2e-16 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 2.608 on 1167 degrees of freedom
  (14 observations deleted due to missingness)
Multiple R-squared:  0.1023,    Adjusted R-squared:  0.1008
F-statistic: 66.51 on 2 and 1167 DF,  p-value: < 2.2e-16
```

Still, P/B is positively correlated with ROE and the Stock Volatility. This time, the p-values are satisfied and the R-squared has achieved 10%, clearly, this result is better than using ROE-A. My understanding is that ttm value can apply more accurate and true value for the corresponding month/quarter, while ROE-A just use one time’s balance sheet, it has huge time effect or even bias.

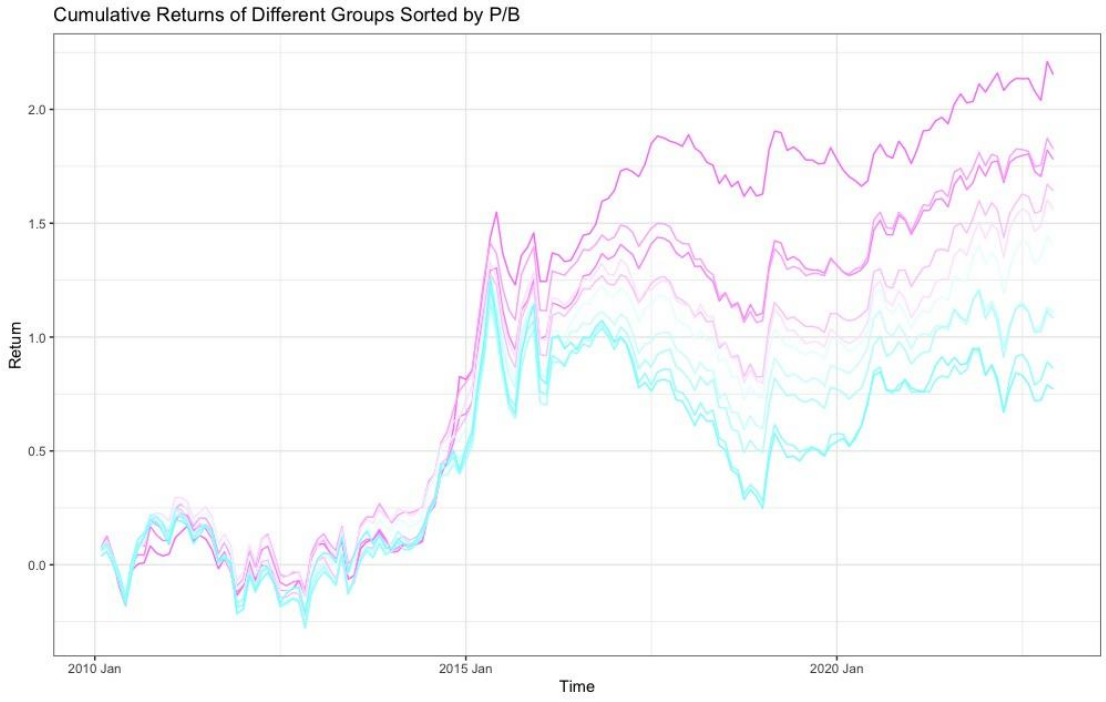
In addition, when dealing with the datasets, I found many calculated P/B ratios that are so ridicules, so I removed some P/B ratios as outliers, only keep P/B that between quantile of 0.25 - 3IQR to quantile of 0.75 + 3IQR to do the linear regression, and the observations reduced about 62.

Now we come to the last part. The question asks us to construct 10 different portfolios, by P/B.

Because we need to sort them by last month’s P/B ratio, so I moved the return part, reduce their month by one, so that the P/B ratio can be corresponded rightly to the return. Now I manually use the function to mutate a column which shows the group of the stocks sorted by P/B. Then because the question says we should let them be equally weighted, I denote this as equally money invested among them, so the total return will be easily the mean of the returns. Now I use a for loop to construct a table that represents 10 time series with their returns, and get the graph as below:



Where the different 10 groups, from less P/B to more P/B, their color changes from purple to blue. We can see that the volatility of the blue curves is more than the purple curves, which means the more P/B ratio, the more the volatility will be, this satisfies our regression result. However, this graph is hard to show the difference of the return, now we consider the cumulative return, which is calculated by returns add their past years return, to show whether there is different between different groups. The graph is shown below:



Also, the different 10 groups, from less P/B to more P/B, their color changes from purple to blue. Now we can easily see that the purple curves has more cumulative returns than the blue curves, which means the low P/B ratio can bring more returns generally. This is true because generally, for high P/B ratio, it may have high total market value, which means its stock price is high, or even be overvalued, at the same time, those with low P/B ratios are often be undervalued by the market, thus if we can construct the portfolios with low P/B ratios, we

may have high returns in the future. And also, with the finding above, low P/B ratios will bring less volatility, which means low risks, it has two sides benefits.