MANOVA study of Size Effect and Value Effect on Stock Returns and Risks

Presented By

Anish Chourasiya. - 19927
Gaurav Goyal - 199278097
Hariprasad Chintala - 199278009
Nimit Jain - 199278002
Prithvi Mohanty - 199278095
Rohitashva Kumar - 199278076
Sai Harish - 199278028

Size Effect and Value Effect

Size Effect — Small Cap stocks outperform the Large Cap stocks

Value Effect — Low P/B ratio stocks (Value stocks) outperform the High P/B ratio stocks (Growth Stocks)

Small Cap	< 5,000 Cr
Mid Cap	5,000 Cr - 20,000 Cr
Large Cap	> 20,000 Cr

- Market Cap = (No of Shares) * (Share Price)
- P/B Ratio = (Share Price) / (Book Value per Share)
- Book Value = Total Assets – Total Liabilities

MANOVA Model

Factors

- P/B Ratio: Low, High
- Market Cap: Small, Mid, Large

Variables

- 1 Year Return
- 3 Years Return
- 5 Years Return
- 10 Years Return
- 1 Year Risk
- 3 Years Risk
- 5 Years Risk
- 10 Years Risk

Hypothesis

H_0: $\mu_{return\_small\ cap} = \mu_{return\_mid\ cap} = \mu_{return\_large\ cap}$
H_1: $\mu_{return\_small\ cap} \neq \mu_{return\_mid\ cap} \neq \mu_{return\_large\ cap}$

H_0: $\mu_{return\_Low\ P/B} = \mu_{return\_High\ P/B}$
H_1: $\mu_{return\_Low\ P/B} \neq \mu_{return\_High\ P/B}$

H_0: $\mu_{risk\_small\ cap} = \mu_{risk\_mid\ cap} = \mu_{risk\_large\ cap}$
H_1: $\mu_{risk\_small\ cap} \neq \mu_{risk\_mid\ cap} \neq \mu_{risk\_large\ cap}$

H_0: $\mu_{risk\_Low\ P/B} = \mu_{risk\_High\ P/B}$
H_1: $\mu_{risk\_Low\ P/B} \neq \mu_{risk\_High\ P/B}$

H_0 : There is zero interaction effect between the two factors : Market Capitalization and P/B Ratio
H_1: Interaction effect exists in between the factors : Market Capitalization and P/B Ratio

About Data

Factors		Market Cap		
		Small Cap	Mid Cap	Large Cap
P/B Ratio	Low P/B	30 Observations	30 Observations	30 Observations
	High P/B	30 Observations	30 Observations	30 Observations

Sources for Data Collection

THOMSON REUTERS

yahoo! finance

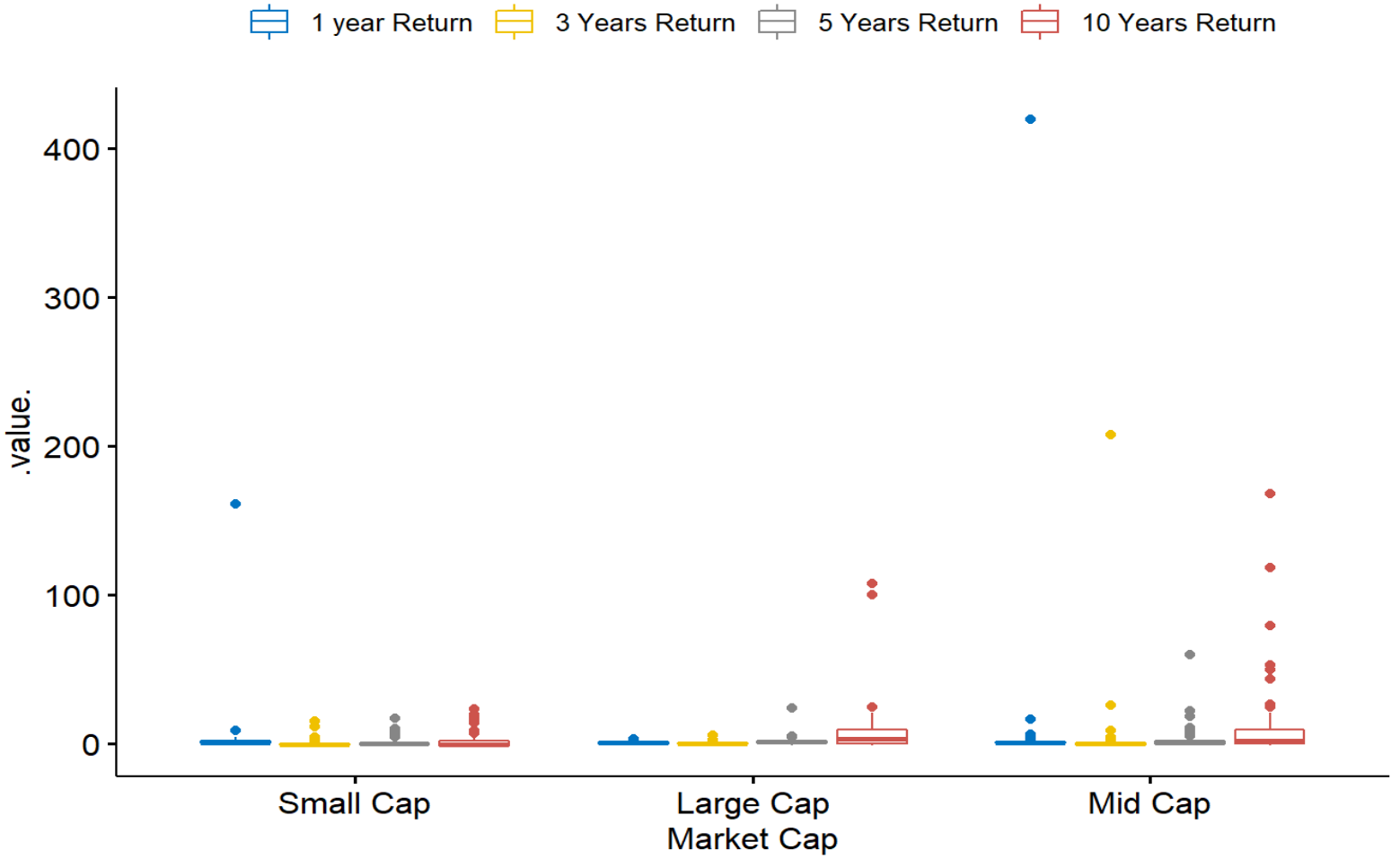
Check for Outliers (Market Caps.)

1 year Return	
Min.	-0.3454
1st Qu.	0.5167
Median	0.9215
Mean	4.5866
3rd Qu.	1.6837
Max.	419.7248

3 Years Return	
Min.	-0.9557
1st Qu.	-0.3514
Median	0.1078
Mean	1.7926
3rd Qu.	0.6192
Max.	207.4500

5 Years Return	
Min.	-0.9397
1st Qu.	-0.1565
Median	0.7093
Mean	1.9558
3rd Qu.	1.8909
Max.	59.7404

10 Years Return	
Min.	-0.9940
1st Qu.	-0.5595
Median	1.4177
Mean	7.8997
3rd Qu.	7.1673
Max.	167.8394



Max values of Returns are very far from the mean which indicate that there are outliers for Returns

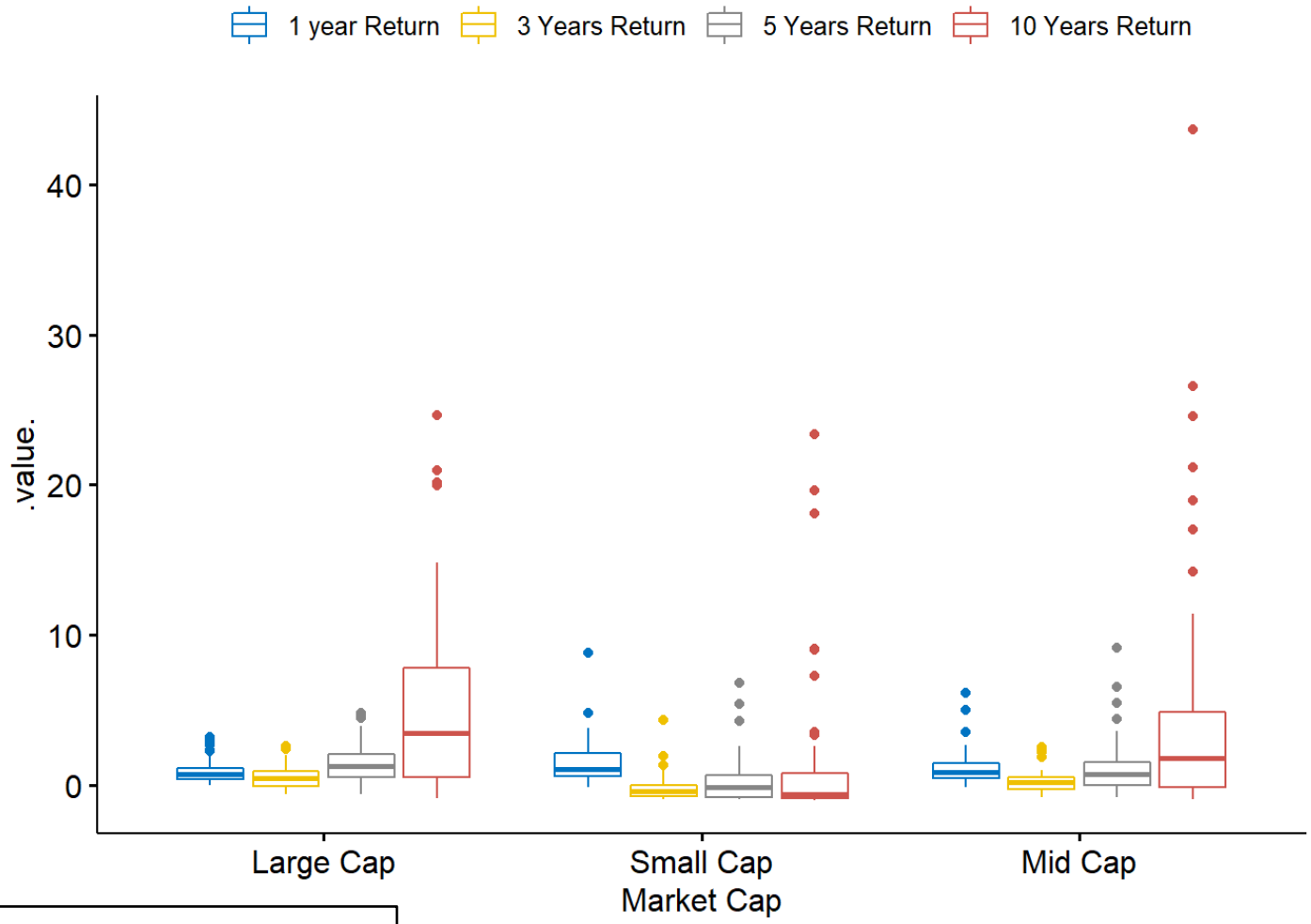
Check for Outliers (P/B Ratio)

```
#wrt to Price to Book Ratio (High/Low)

ggboxplot(

  M, x = c("Price to Book Ratio (High/Low)"), y = c("1 year Return", "3 Years Return",
"5 Years Return","10 Years Return"),

  merge = TRUE, palette = "jco"

)
```

From the box plot you can see most of the outliers are for high PB ratio

Outliers Removed (Market Caps.)

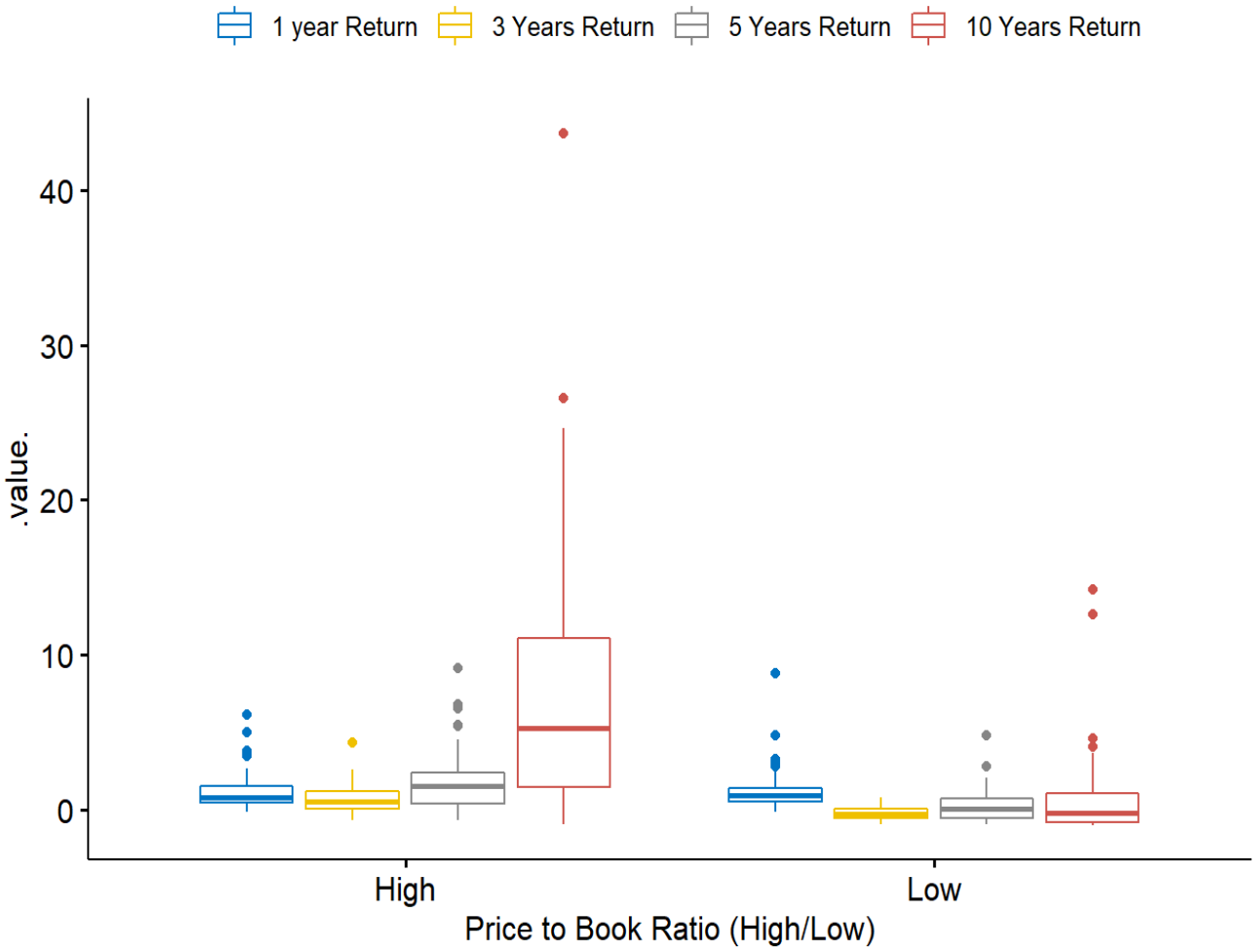
```
# Removing outliers using mahalanobis function:
cutoff=qchisq(1-0.05,4)
cutoff
## [1] 9.487729
mahal=mahalanobis(M[,-c(1,2)],colMeans(M[,-c(1,2)]),cov(M[,-c(1,2)]))
summary(mahal<cutoff)
##    Mode   FALSE    TRUE
## logical      21     159
# 21 outliers
noout=subset(M,mahal<cutoff)
Mo=noout
Mo
## # A tibble: 159 x 10
```



Box plot to verify if the outliers are removed

Outliers Removed (P/B Ratio)

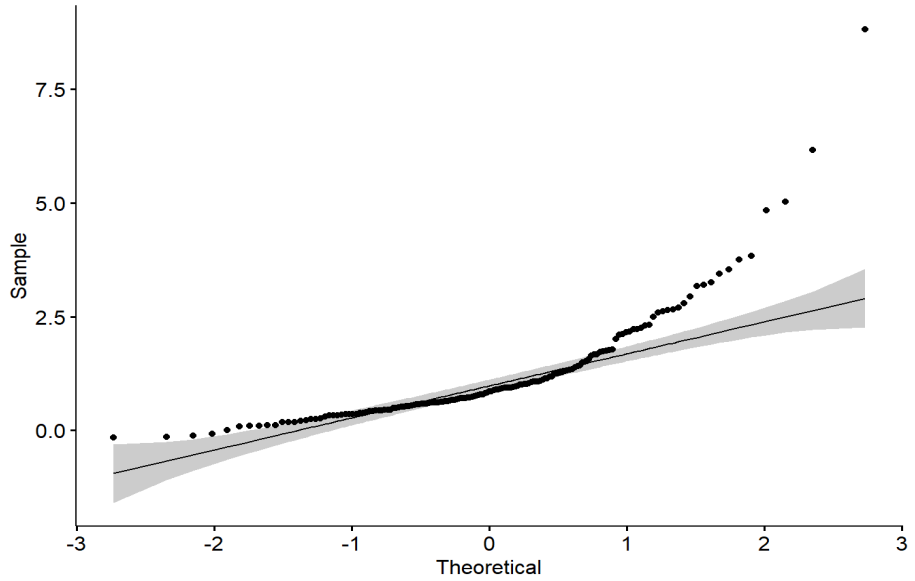
```
ggboxplot(

  Mrto, x = c("Price to Book Ratio (High/Low)"), y = c("1 year Return", "3 Years Retur
n","5 Years Return","10 Years Return"),

  merge = TRUE, palette = "jco"

)
```

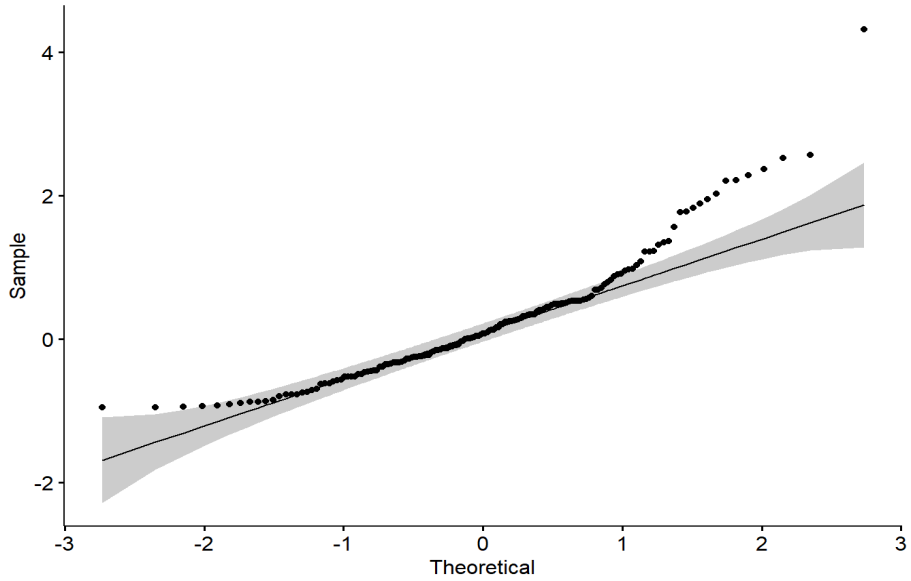


Check for Normality

```
shapiro.test(R[,1])
##
##  Shapiro-Wilk normality test
##
## data:  R[, 1]
## W = 0.75103, p-value = 3.974e-15
```

```
shapiro.test(R[,2])
##
##  Shapiro-Wilk normality test
##
## data:  R[, 2]
## W = 0.9046, p-value = 1.155e-08
```



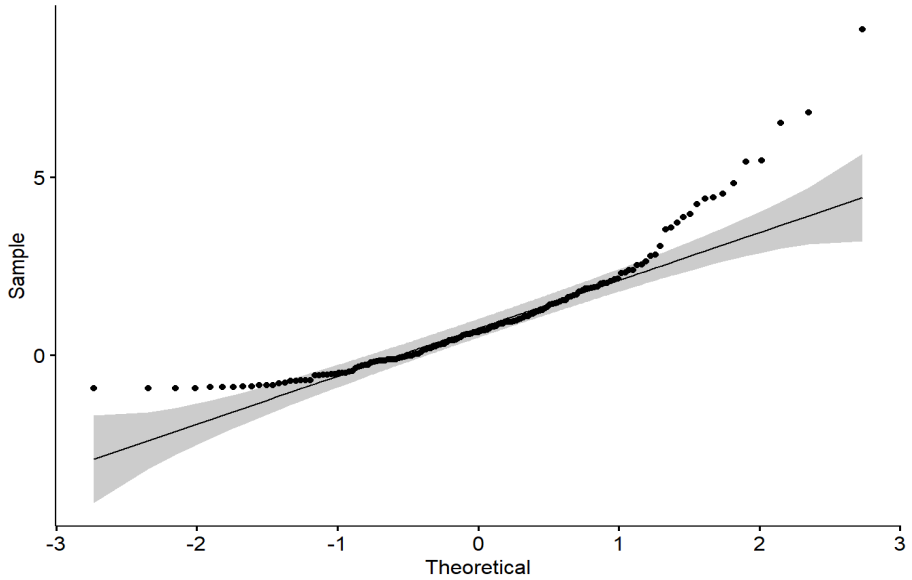


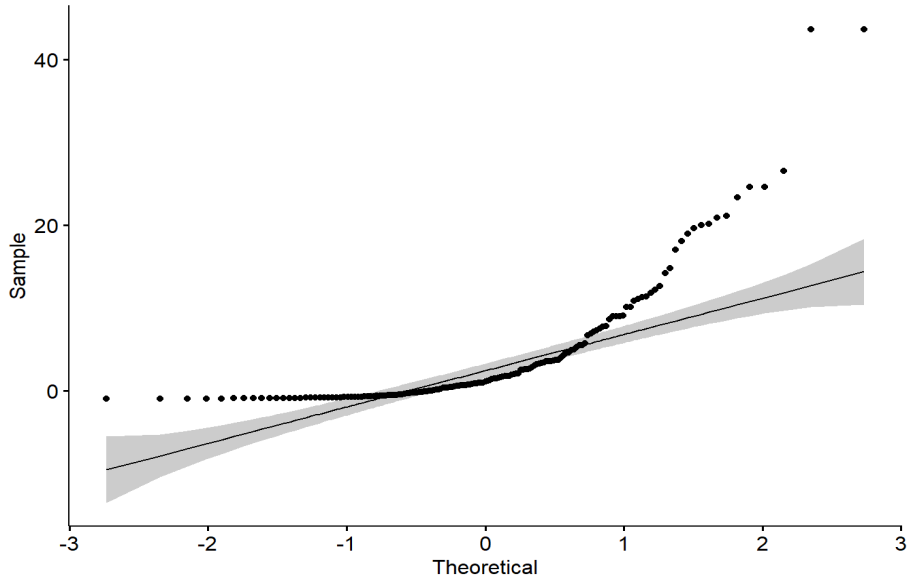
The data is not normal as p-value is less than 0.05 for both the cases

Check for Normality Contd.

```
shapiro.test(R[,3])
##
##  Shapiro-Wilk normality test
##
## data:  R[, 3]
## W = 0.8595, p-value = 4.987e-11
```

```
shapiro.test(R[,4])
##
##  Shapiro-Wilk normality test
##
## data:  R[, 4]
## W = 0.67818, p-value < 2.2e-16
```



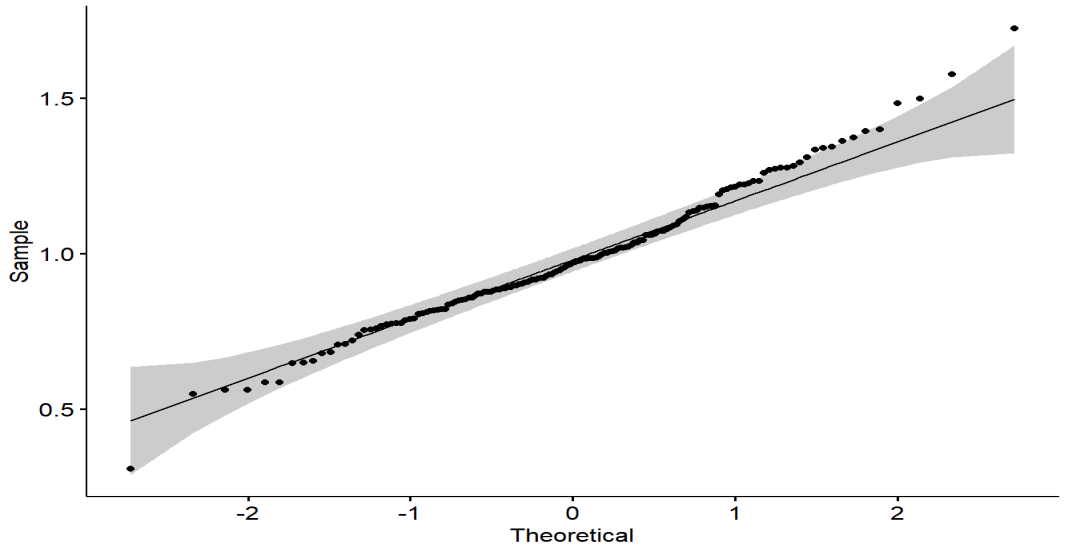


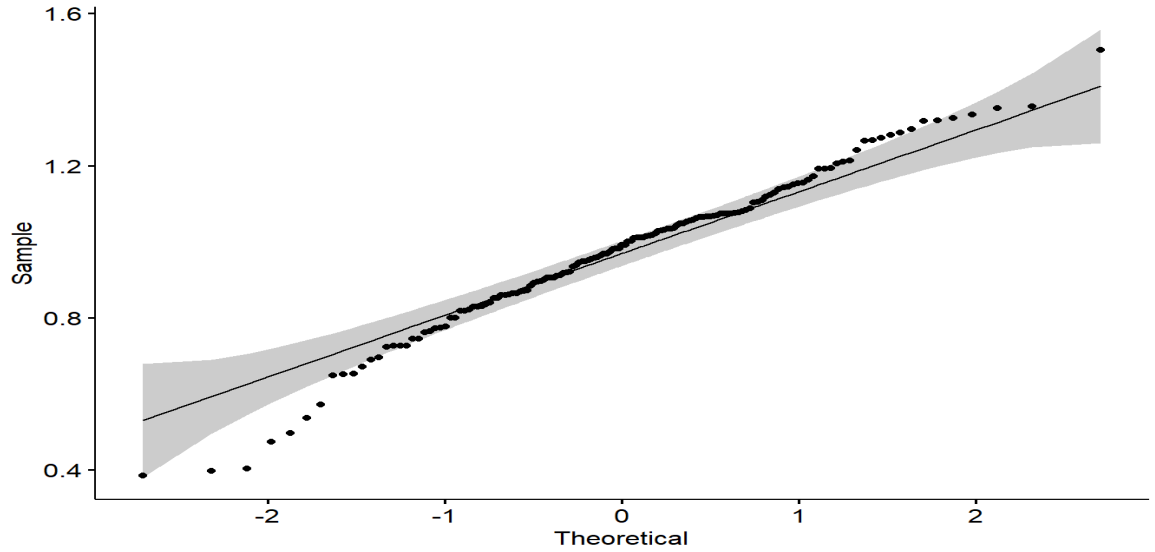
The data is not normal as p-value is less than 0.05 for both the cases

Power Transformation to make it Normal

```
#power Transformation:
# and replacing the coulmn values with particular transformation value
R1=R[,1]
R1n=(R1)^(1/4)
shapiro.test(R1n)
##
##  Shapiro-Wilk normality test
##
## data:  R1n
## W = 0.9852, p-value = 0.09716
```

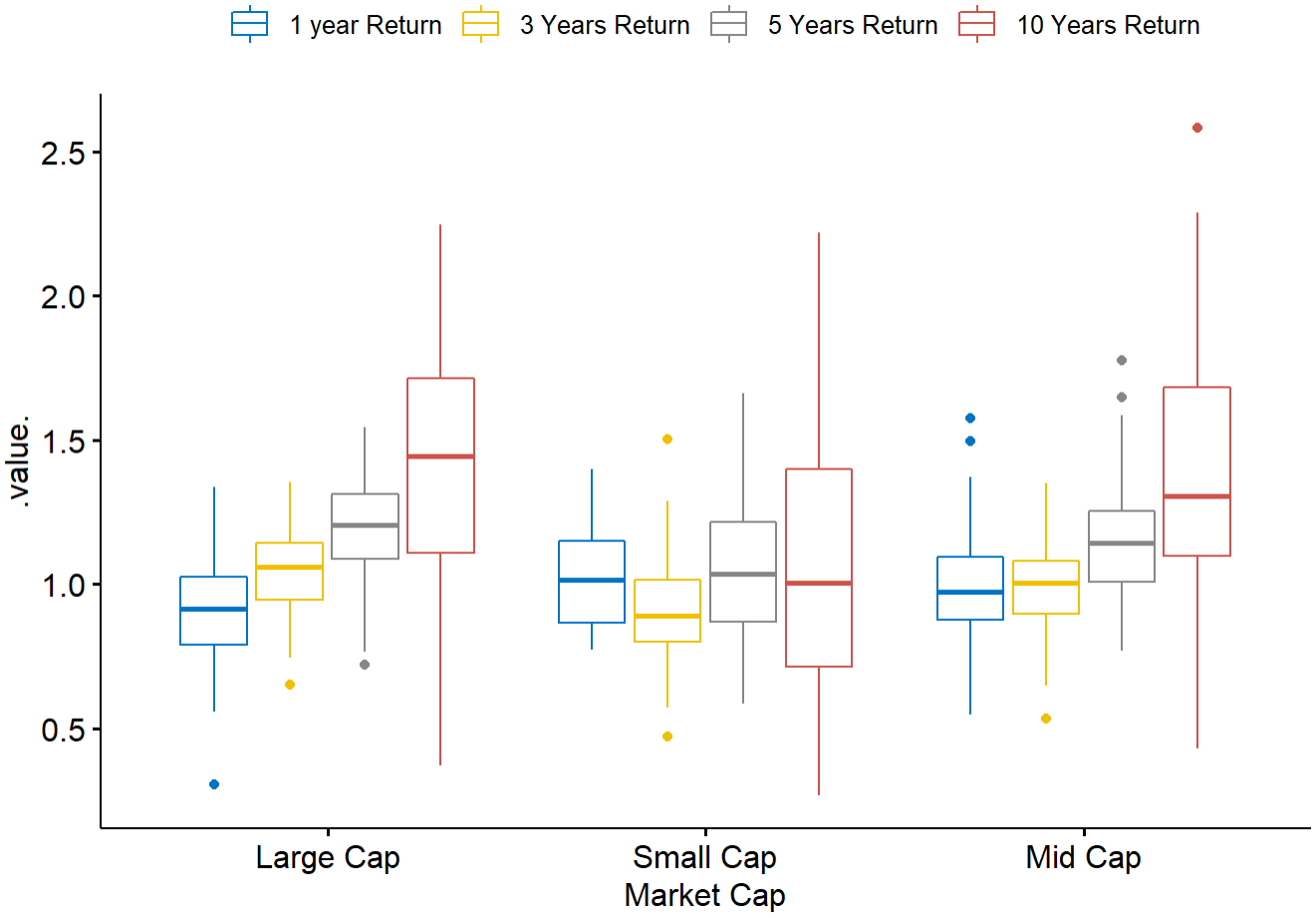
```
R2=R[,2]
R2n=(R2+0.8)^(1/4)
shapiro.test(R2n)
##
##  Shapiro-Wilk normality test
##
## data:  R2n
## W = 0.98189, p-value = 0.04956
```





Final Check for Outliers & Normality (Market Caps.)

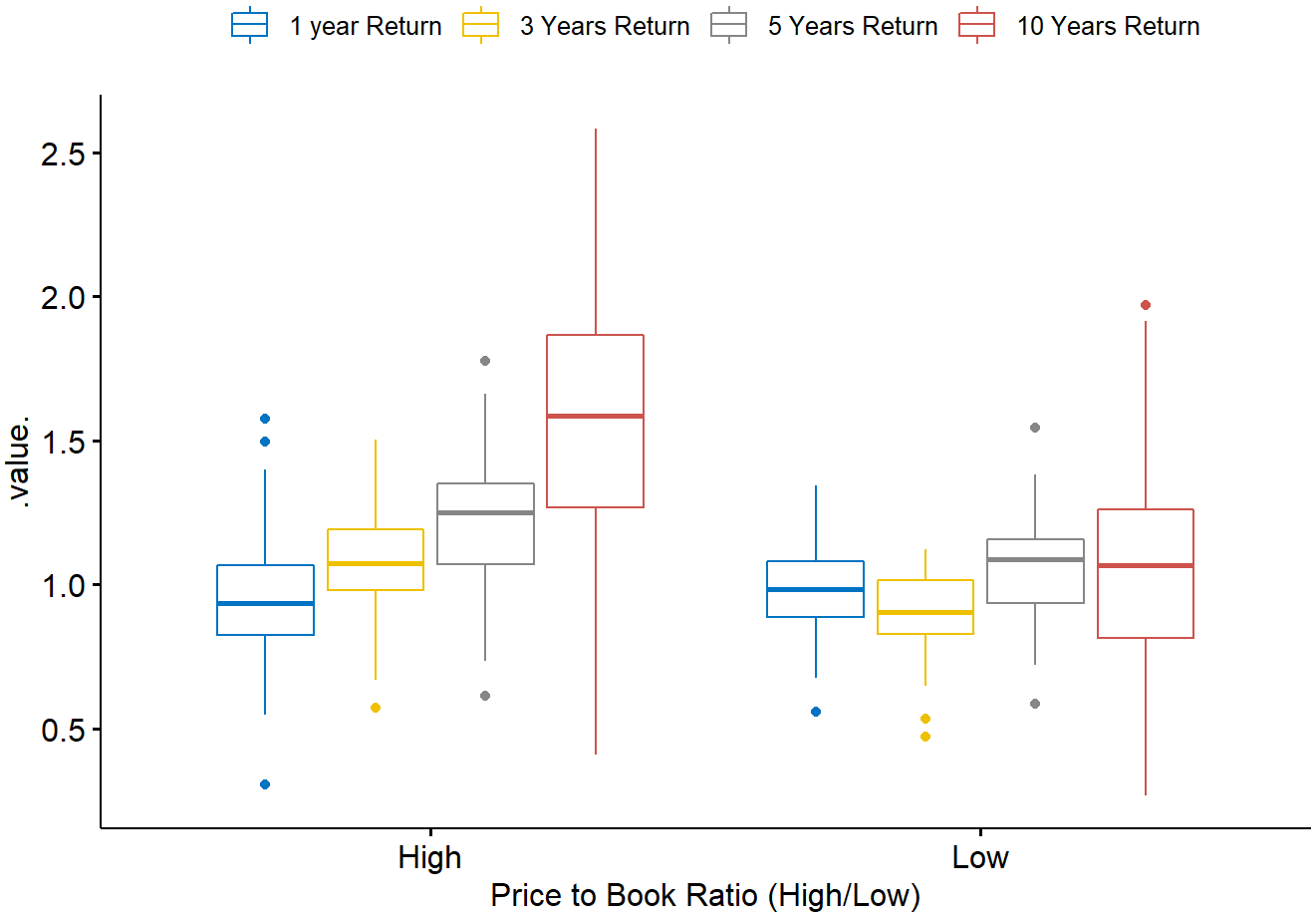
```
ggboxplot(

  Mrton, x = c("Market Cap"), y = c("1 year Return", "3 Years Return","5 Years Return"
,"10 Years Return"),

  merge = TRUE, palette = "jco"

)
```



The box plot again to check the normality and outliers if any

Final Check for Outliers & Normality (P/B Ratio)

```
ggboxplot(

  Mrton, x = c("Price to Book Ratio (High/Low)"), y = c("1 year Return", "3 Years Retu
rn","5 Years Return","10 Years Return"),

  merge = TRUE, palette = "jco"

)
```



The box plot again to check the normality and outliers if any

Check for Additivity

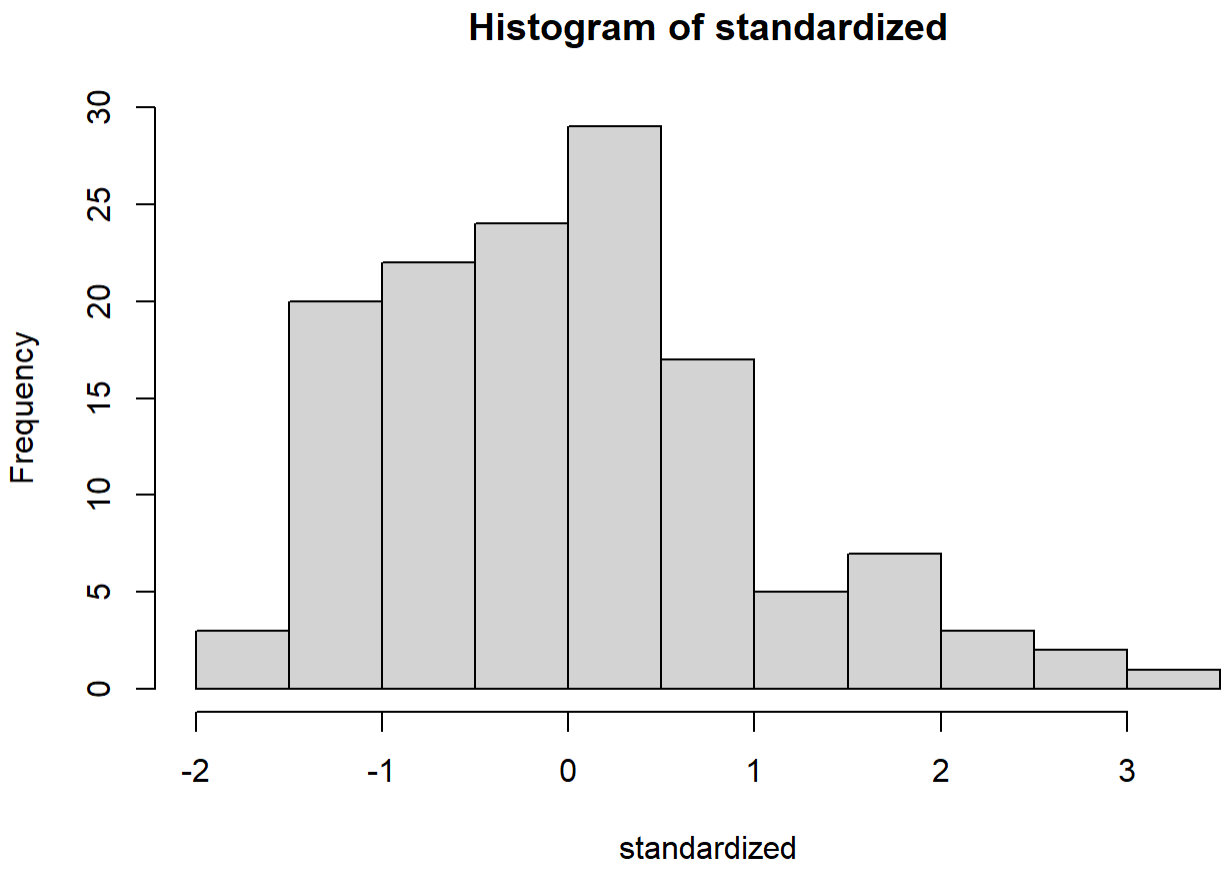
- Checking correlation should have some correlation but coefficients should be less than 0.99
- Correlation close to 1 makes Manova unstable

```
correl=cor(Mrton[,-c(1,2)],use="pairwise.complete.obs")
symnum(correl)
##                 1yR 3YR 5YR 1YR
## 1 year Return   1
## 3 Years Return      1
## 5 Years Return      ,   1
## 10 Years Return     ,   ,   1
```

```
correl
##                 1 year Return 3 Years Return 5 Years Return 10 Years Return
## 1 year Return      1.000000000      0.1983257      0.1882161      0.002850151
## 3 Years Return     0.198325655      1.0000000      0.7846231      0.655625983
## 5 Years Return     0.188216091      0.7846231      1.0000000      0.796460521
## 10 Years Return    0.002850151      0.6556260      0.7964605      1.000000000
# all good
```

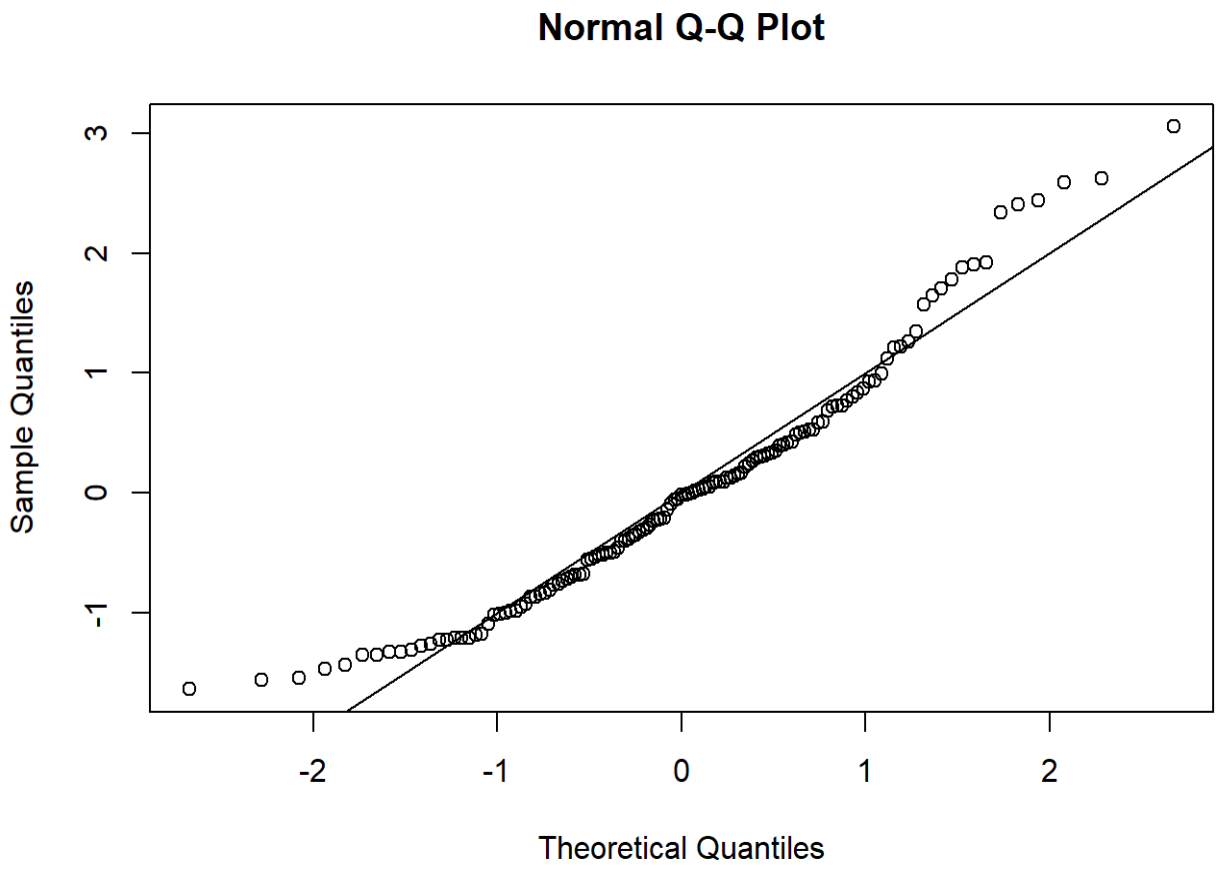
All correlation coefficient are less than 0.99, So we are we good to go

Histogram (Standardized)



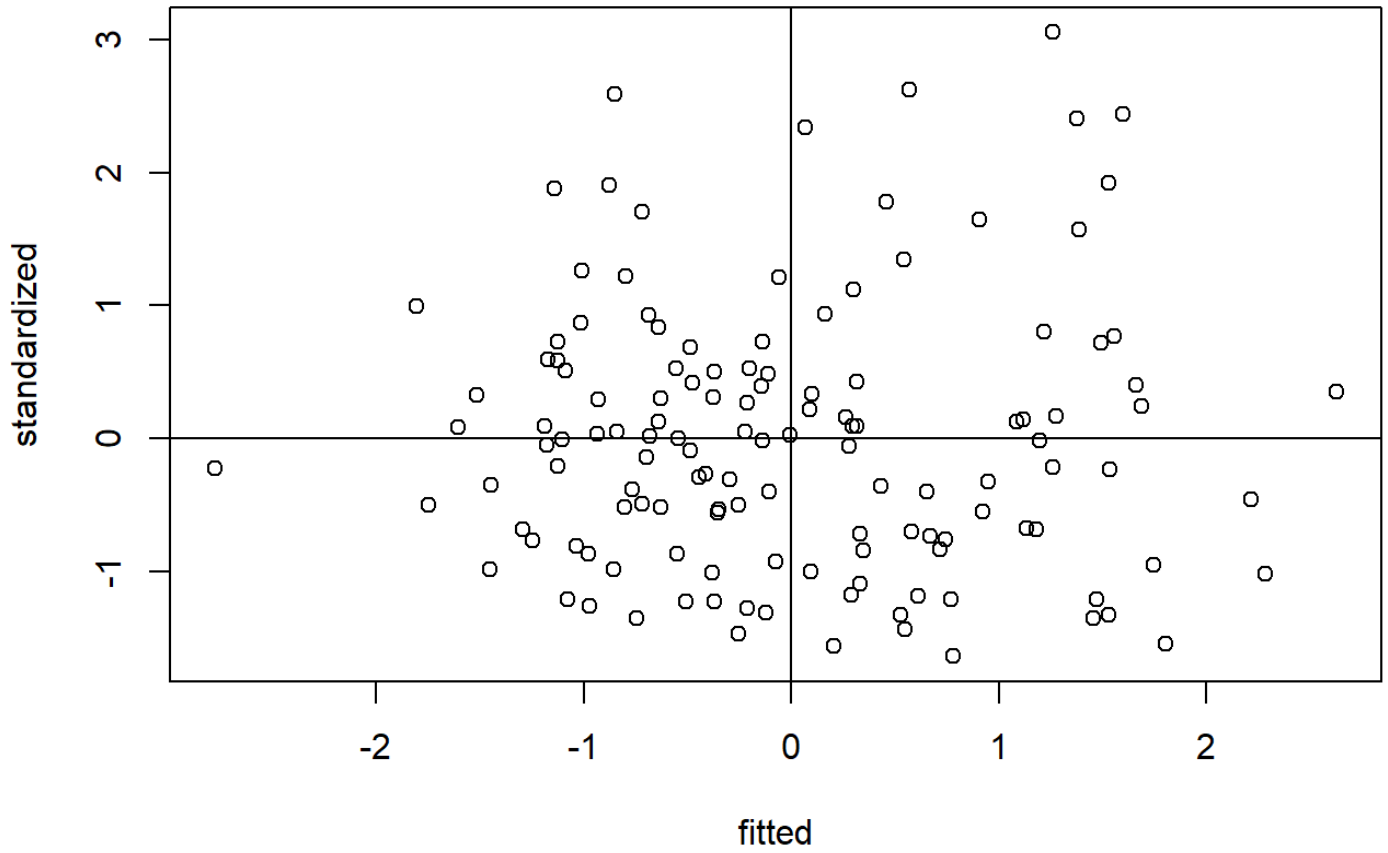
Linearity (Standardized)

```
qqnorm(standardized)
abline(0,1)
```



Homogenity (Standardized)

```
##homogenity
plot(fitted,standardized)
abline(0,0)
abline(v=0)
```



Levene's Test (Market Caps.)

- Levene's Test of Equality of Variance: Used to examine whether the variance between Independent variable groups are equal; also known as homogeneity of variance (Assumption for Manova)
- Since the p value is not significant or greater than 0.05, we cannot reject the hypothesis of i.e., homogeneity in variance

```
leveneTest(Mrton$`1 year Return` ~ Mrton$`Market Cap`*Mrton$`Price to Book Ratio (High
/Low)`,
           data= Mrton, center= mean)
## Levene's Test for Homogeneity of Variance (center = mean)
##        Df F value  Pr(>F)
```

```
leveneTest(Mrton$`3 Years Return` ~ Mrton$`Market Cap`*Mrton$`Price to Book Ratio (Hig
h/Low)`,
           data=Mrton, center= mean)
## Levene's Test for Homogeneity of Variance (center = mean)
##        Df F value Pr(>F)
```

```
leveneTest(Mrton$`5 Years Return` ~ Mrton$`Market Cap`*Mrton$`Price to Book Ratio (Hig
h/Low)`,
           data= Mrton, center= mean)
## Levene's Test for Homogeneity of Variance (center = mean)
##        Df F value  Pr(>F)
## group   5  3.0313 0.01281 *
```

```
leveneTest(Mrton$`10 Years Return` ~ Mrton$`Market Cap`*Mrton$`Price to Book Ratio (Hi
gh/Low)`,
           data= Mrton, center= mean)
## Levene's Test for Homogeneity of Variance (center = mean)
##        Df F value  Pr(>F)
## group   5  2.9263 0.01555 *
```

For 5 year and 10 years return Levene test has a significant p value to reject the hypothesis. It violates the assumption of Manova for this specific data

Multivariate Tests

```
## Multivariate Tests: (Intercept)
##                   Df test stat approx F num Df den Df     Pr(>F)
## Pillai             1  0.932025 425.0513      4    124 < 2.22e-16 ***
## Wilks              1  0.067975 425.0513      4    124 < 2.22e-16 ***
## Hotelling-Lawley   1 13.711332 425.0513      4    124 < 2.22e-16 ***
## Roy                1 13.711332 425.0513      4    124 < 2.22e-16 ***
```

```
## Multivariate Tests: Mrton$`Market Cap`
##                   Df test stat  approx F num Df den Df     Pr(>F)
## Pillai             2 0.3499297  6.627175      8    250 7.4253e-08 ***
## Wilks              2 0.6603257  7.148952      8    248 1.6297e-08 ***
## Hotelling-Lawley   2 0.4988735  7.670180      8    246 3.6322e-09 ***
## Roy                2 0.4655106 14.547207      4    125 8.8075e-10 ***
```

```
## Multivariate Tests: Mrton$`Price to Book Ratio (High/Low)`
##                   Df test stat approx. F num Df den Df     Pr(>F)
## Pillai             1 0.3728771 18.4321      4    124 6.5798e-12 ***
## Wilks              1 0.6271229 18.4321      4    124 6.5798e-12 ***
## Hotelling-Lawley   1  0.5945837  18.4321      4    124 6.5798e-12 ***
## Roy                1 0.5945837 18.4321      4    124 6.5798e-12 ***
```

```
## Multivariate Tests: Mrton$`Market Cap`:Mrton$`Price to Book Ratio (High/Low)
##                   Df test stat  approx F num Df den Df  Pr(>F)
## Pillai             2 0.0483577 0.7743115      8    250 0.62573
## Wilks              2 0.9518073 0.7751185      8    248 0.62502
## Hotelling-Lawley   2 0.0504595 0.7758141      8    246 0.62441
## Roy                2 0.0467508 1.4609635      4    125 0.21807
```

From the summary of Manova we can see that for our factor 1 which is Market Cap, the wilks coefficient section the p value is less than 0.05. we can easily reject the null hypothesis that the Mean of return vectors for all the Market Caps are equal. Similarly for factor 2 which Price to book ratio, since for interaction we have high P value we cannot reject the null hypothesis that there is zero interaction which means since there is no interaction, we must perform Manova for factor 1 and factor 2 separately.

Manova (Factor 1 – Market Cap)

```
summary(R1.av)
##                        Df Sum Sq Mean Sq F value Pr(>F)
## Mrton$`Market Cap`      2  0.299 0.14935   3.677  0.028 *
## Residuals             130  5.280 0.04062
```

```
summary(R2.av)
##                        Df Sum Sq Mean Sq F value Pr(>F)
## Mrton$`Market Cap`      2  0.306 0.15314   4.874 0.0091 **
## Residuals             130  4.085 0.03142
```

```
summary(R3.av)
##                        Df Sum Sq Mean Sq F value Pr(>F)
## Mrton$`Market Cap`      2  0.375 0.18744   3.841 0.0239 *
## Residuals             130  6.343 0.04879
```

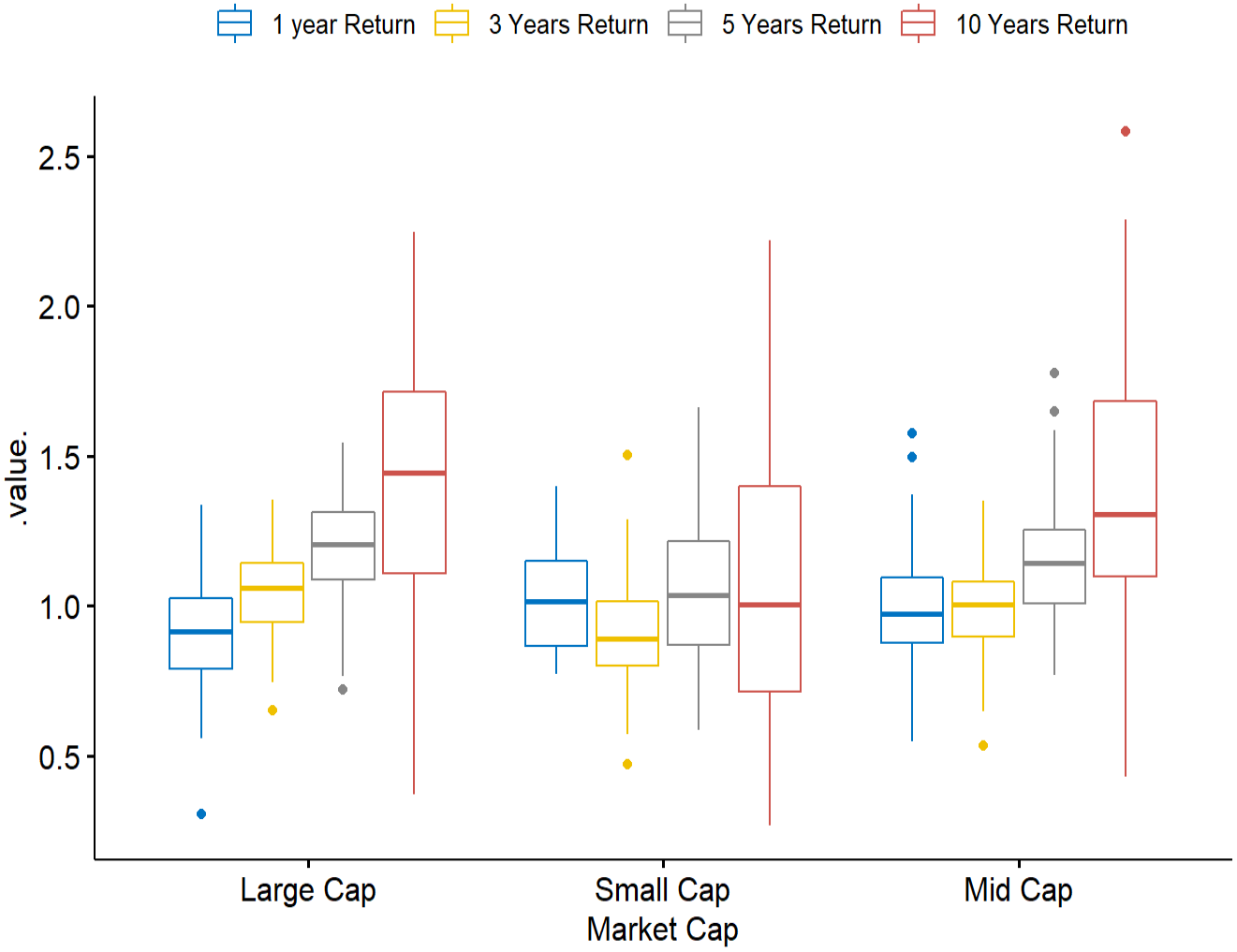
P < 0.5, Reject the null hypothesis. We must see from Bonferroni intervals to know from where the difference is coming from

```
tukey.test <- TukeyHSD(R4.av)
tukey.test
##   Tukey multiple comparisons of means
##     95% family-wise confidence level
##
## Fit: aov(formula = R4.lm)
##
## $`Mrton$`Market Cap``
##                                diff        lwr         upr     p adj
## Mid Cap-Large Cap    -0.01745386 -0.2454210  0.21051325 0.9820017
## Small Cap-Large Cap -0.29802644 -0.5569797 -0.03907322 0.0196893
## Small Cap-Mid Cap    -0.28057259 -0.5453626 -0.01578253 0.0350709
```

For small cap - Large cap interval is always negative, and for small cap - Mid cap interval is always negative as well

Result (Market Caps.)

- There's is no significant difference between the returns of Mid cap and Large Cap companies irrespective of no. of years of return chosen
- For 10 years of return there's a significant difference between small cap and mid cap companies returns was higher for Mid cap compared to small cap companies
- Except for 10 years of returns small and mid cap companies don't have significantly different return
- Small cap companies gave higher 1 years of returns as compared to large cap companies which is little unobvious
- Except for 1 years of return Large cap companies gave higher returns as compared to Small cap companies



Manova (Factor 2 – P/B Ratio)

```
R1_f2.av <- aov(R1_f2.lm)
summary(R1_f2.av)
##                                          Df Sum Sq Mean Sq F value Pr(>F)
## Mrton$`Price to Book Ratio (High/Low)`    1  0.043 0.04328   1.024  0.313
## Residuals                               131  5.535 0.04225
```

P > 0.5, cannot reject the null hypothesis . There's no significant difference between 1 year returns of company with high p/b ratio as compared to company with low p/b ratio. No need to look at Bonferroni interval as only two levels and there's no difference seen

```
R2_f2.av <- aov(R2_f2.lm)
summary(R2_f2.av)
##                                          Df Sum Sq Mean Sq F value   Pr(>F)
## Mrton$`Price to Book Ratio (High/Low)`    1  1.004  1.0042   38.84 5.82e-09 ***
## Residuals                               131  3.387  0.0259
```

P << 0.5, Reject the null hypothesis. There's indeed a significant difference between the 3 years return for company with high pb ratio and company with low pb ratio

```
summary(R3_f2.av)
##                                          Df Sum Sq Mean Sq F value   Pr(>F)
## Mrton$`Price to Book Ratio (High/Low)`    1  1.017  1.0171   23.37 3.68e-06 ***
## Residuals                               131  5.701  0.0435
```

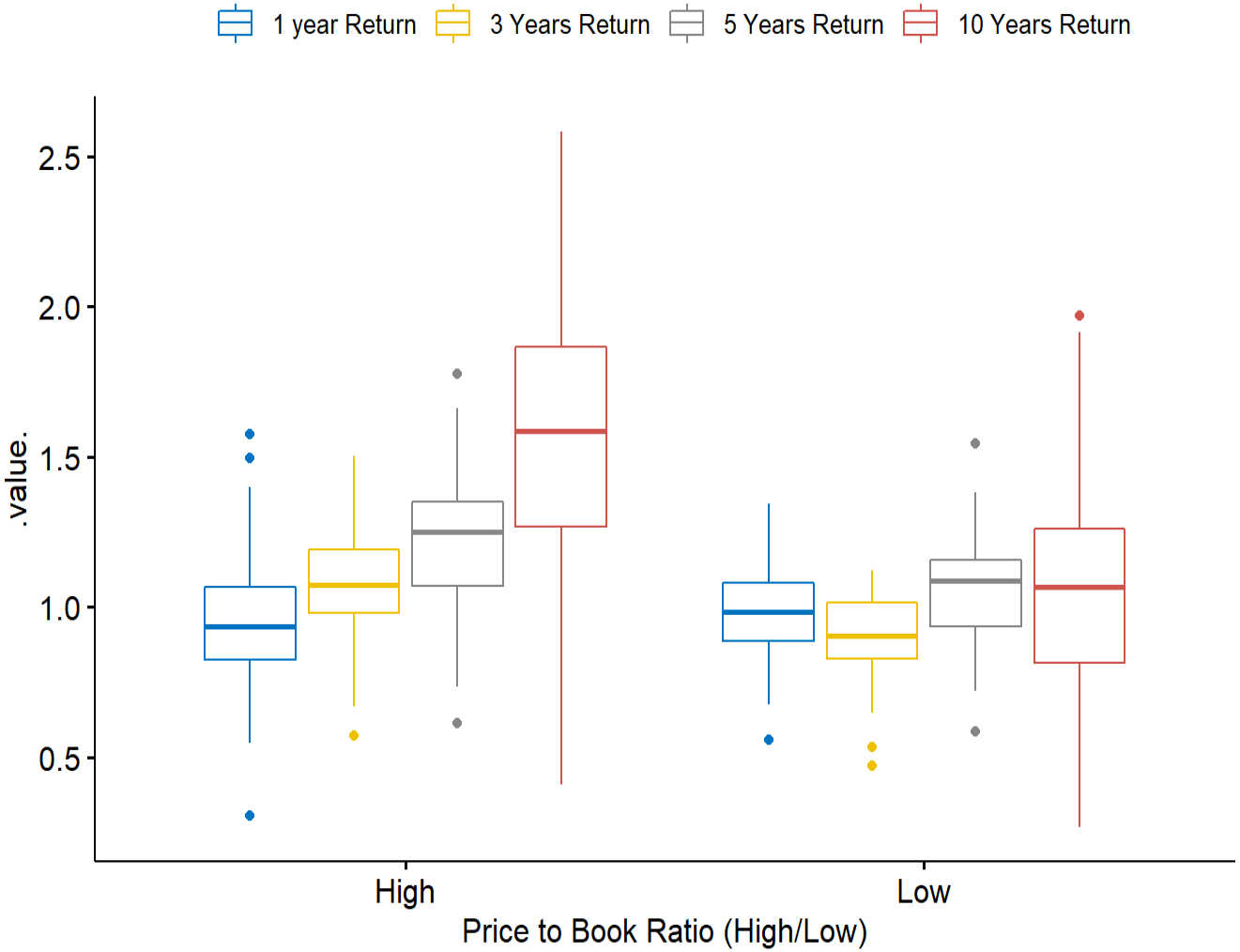
P << 0.5, Reject the null hypothesis. There's indeed a significant difference between the 5 years return for company with high pb ratio and company with low pb ratio

```
summary(R4_f2.av)
##                                          Df Sum Sq Mean Sq F value   Pr(>F)
## Mrton$`Price to Book Ratio (High/Low)`    1  8.881   8.881   49.16 1.13e-10 ***
## Residuals                               131 23.666   0.181
```

P << 0.5, Reject the null hypothesis. There's indeed a significant difference between10 years return for company with high pb ratio and company with low pb ratio

Result (P/B Ratio)

- Except for the 1 years return all other had significant difference between the company with high and low pb ratio
- As predicted from box plot the company with high pb ratio had higher 3 , 5 and 10 years returns
- For 1 years return there was no difference between them



Inferences/Conclusion

Thank You

For Your Attention