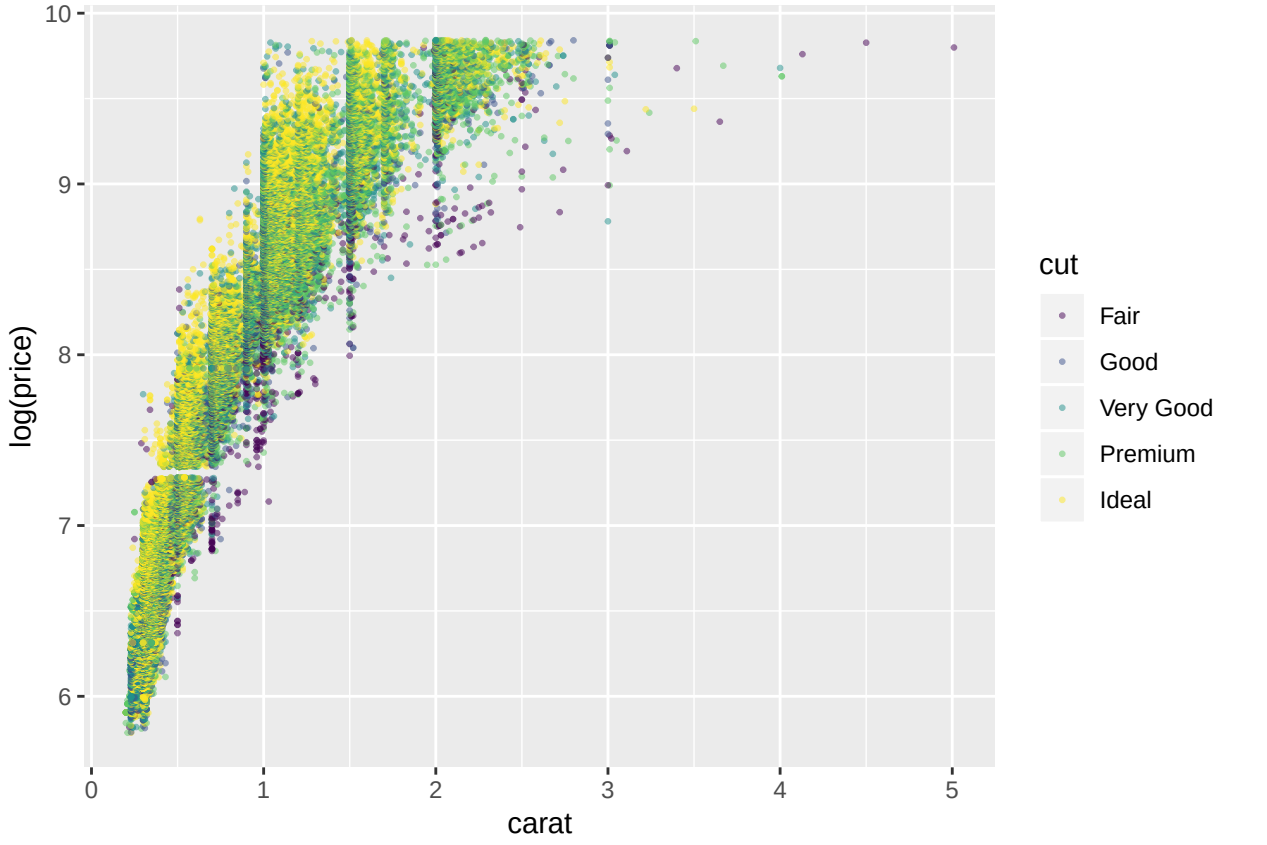
hw1

Huanyu Liu

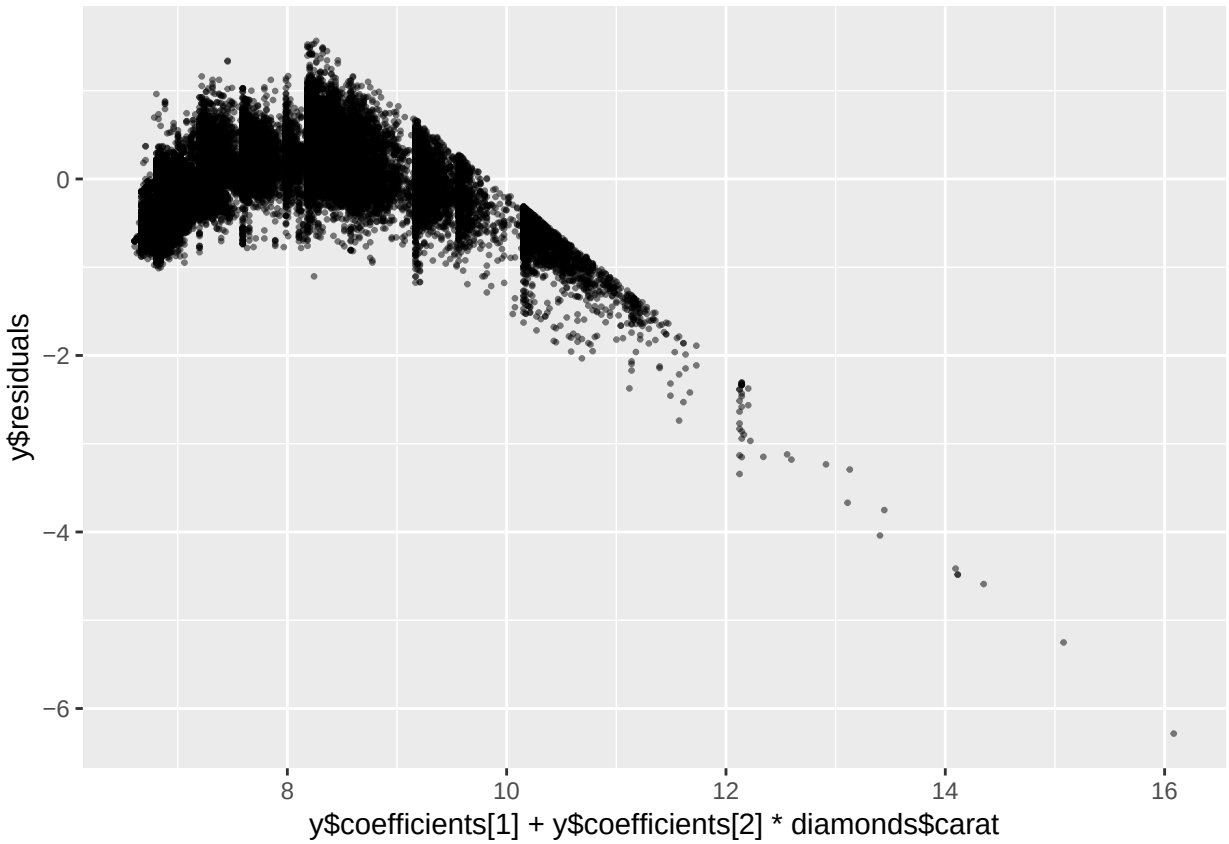
4/7/2019

Problem 1

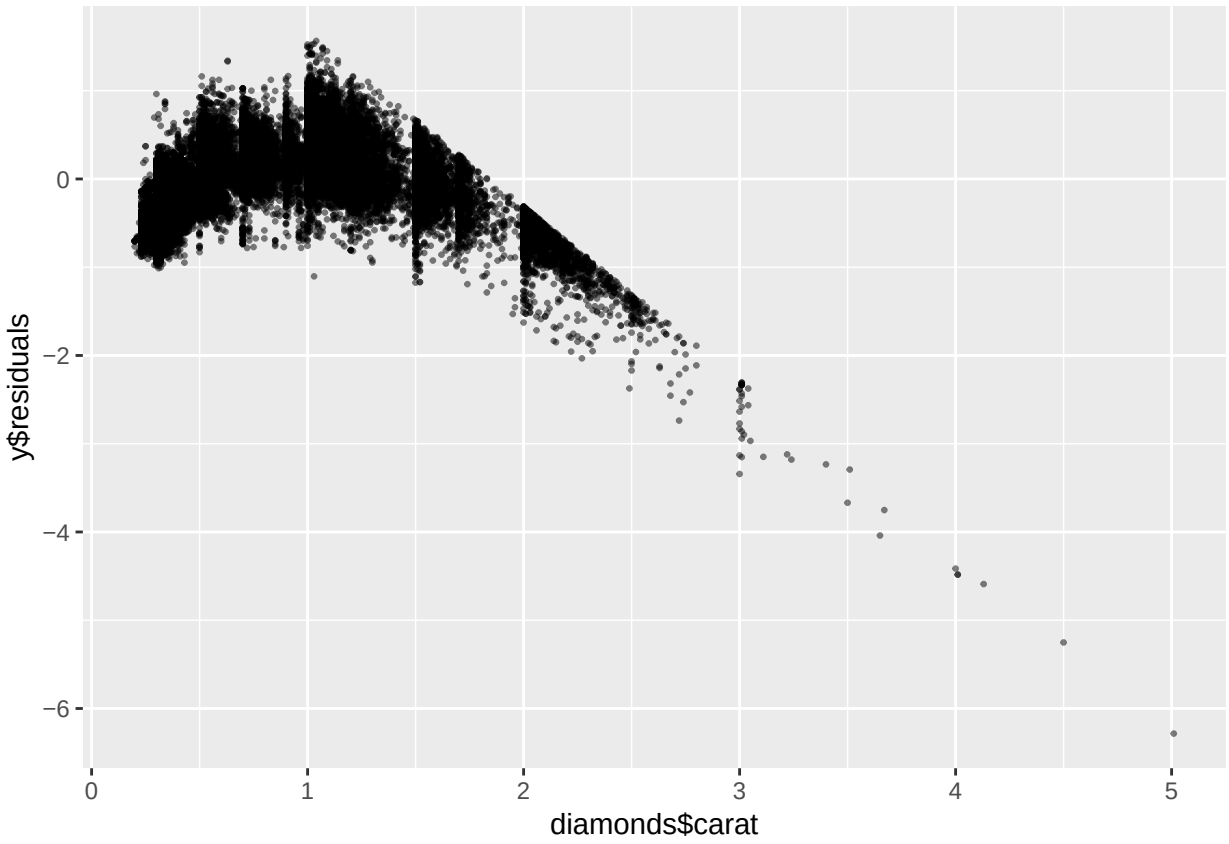
```
library('ggplot2')
library('foreign')
library('data.table')

# 1

data(diamonds)
qplot(carat, log(price), data = diamonds, col = cut, size = I(0.5), alpha = I(0.5))
```




```
# 2
y = lm(log(price)~carat, data = diamonds)
qplot(y$coefficients[1] + y$coefficients[2] * diamonds$carat, y$residuals, size = I(0.5), alpha = I(0.5)
```



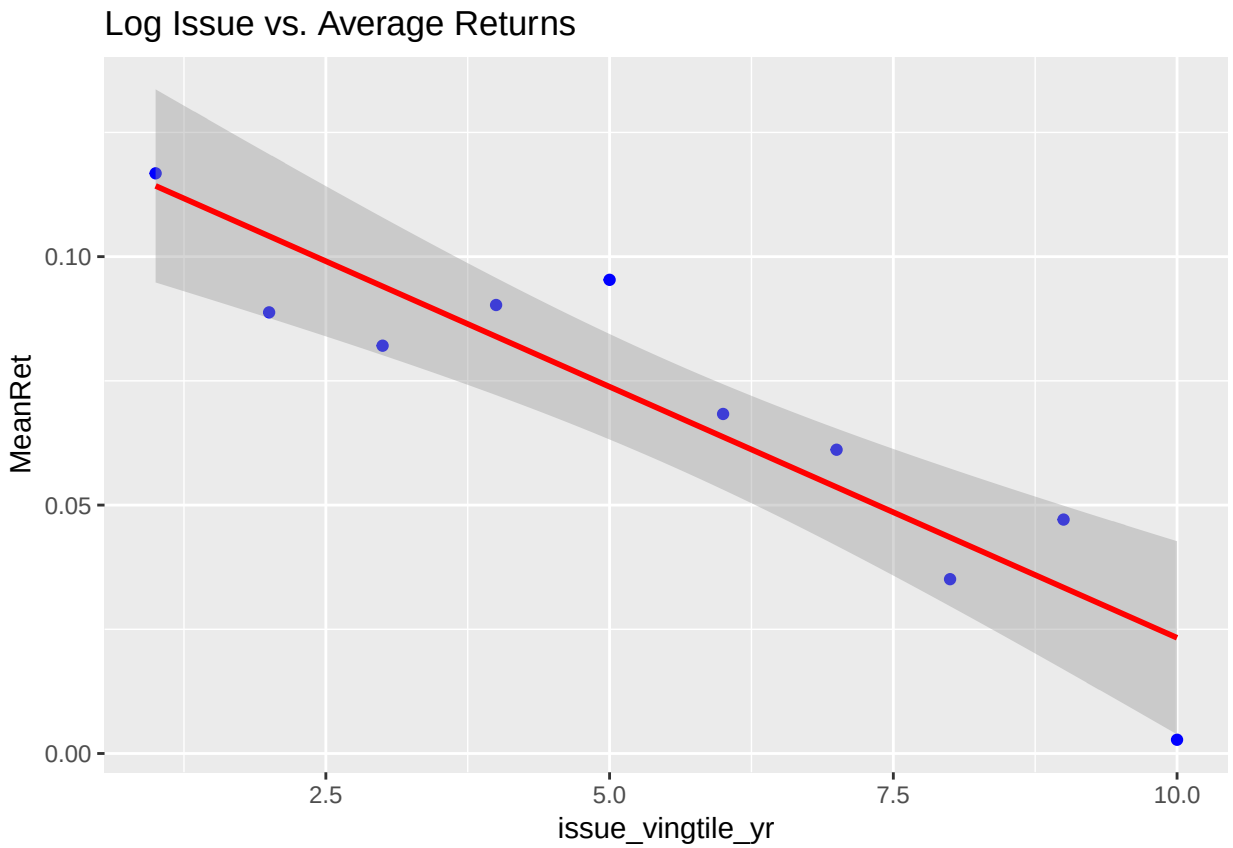

```
qplot(diamonds$carat, y$residuals, size = I(0.5), alpha = I(0.5))
```



The residuals are not random. They are correlated with carat. Therefore, the relationship between price and carat are not linear.

Problem 2

```
StockRetAcct_DT = as.data.table(read.dta('/Users/huanyu/Desktop/MachineLearning431/hw1/StockRetAcct_insa
StockRetAcct_DT = StockRetAcct_DT[!is.na((StockRetAcct_DT$lnIssue)),]
setkey(StockRetAcct_DT, FirmID, year)
StockRetAcct_DT[,ExRet:=exp(lnAnnRet) - exp(lnRf)]
for (i in 1980:2014){
    StockRetAcct_DT[year == i,issue_vingtile_yr:=cut(StockRetAcct_DT[year == i,]$lnIssue,breaks=quantil
}

VW_Issue_MutualFund_yr = StockRetAcct_DT[,list(MeanRetYear = weighted.mean(lnAnnRet, MEwt)), by = list(
VW_Issue_MutualFund = VW_Issue_MutualFund_yr[,list(MeanRet = mean(MeanRetYear)), by = issue_vingtile_yr


qplot(issue_vingtile_yr, MeanRet,data = VW_Issue_MutualFund, col=I("blue"), na.rm = TRUE, main = "Log Is
```



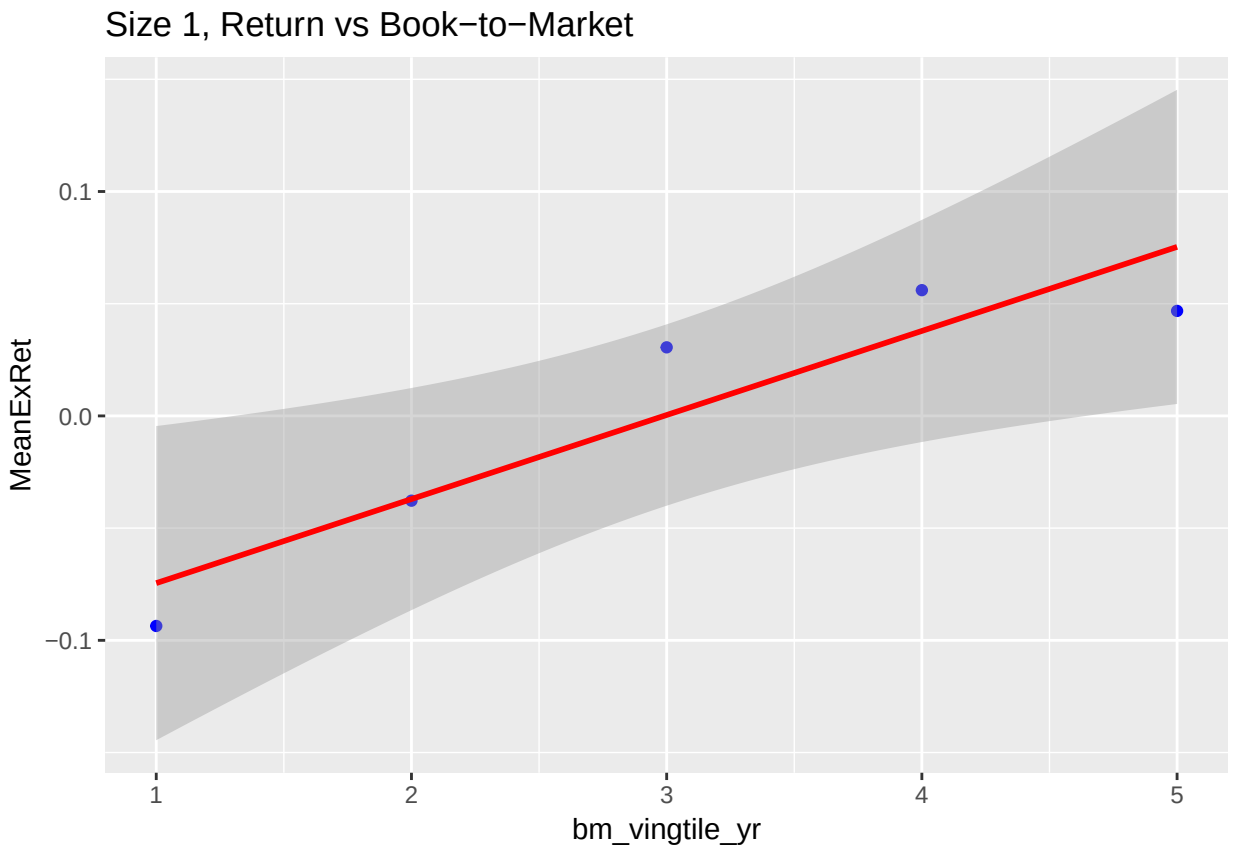

```
# According to the plot, the pattern is not linear.

StockRetAcct_DT[,transIns:= 0][issue_vingtile_yr==1,transIns:= -1][issue_vingtile_yr==10,transIns:= 1][
betas = c(1980:2014)
for (i in 1980:2014){
    betas[i - 1979] = lm(lnAnnRet~transIns,StockRetAcct_DT[year==i][!is.na(transIns)])$coefficients[2]
}
print(mean(betas))
```

```
## [1] -0.07584684
```

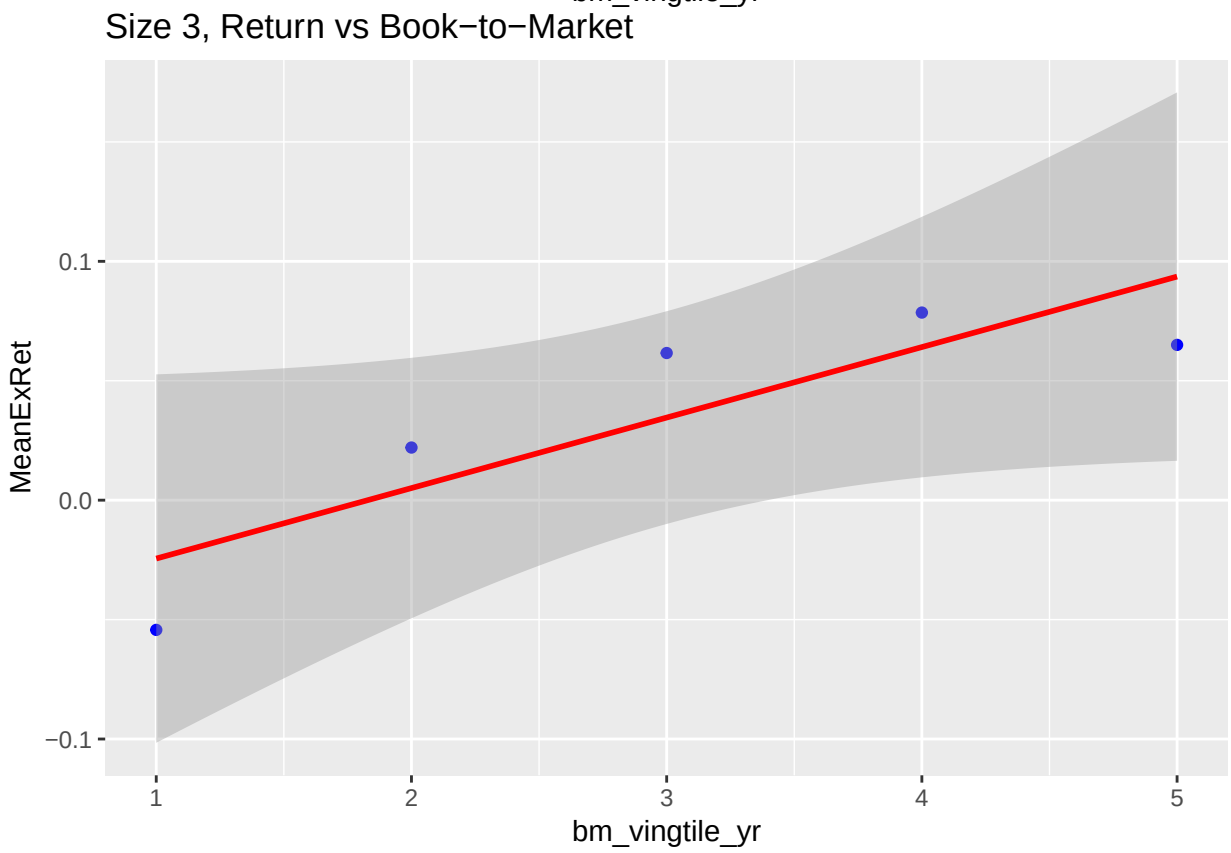
Problem 3

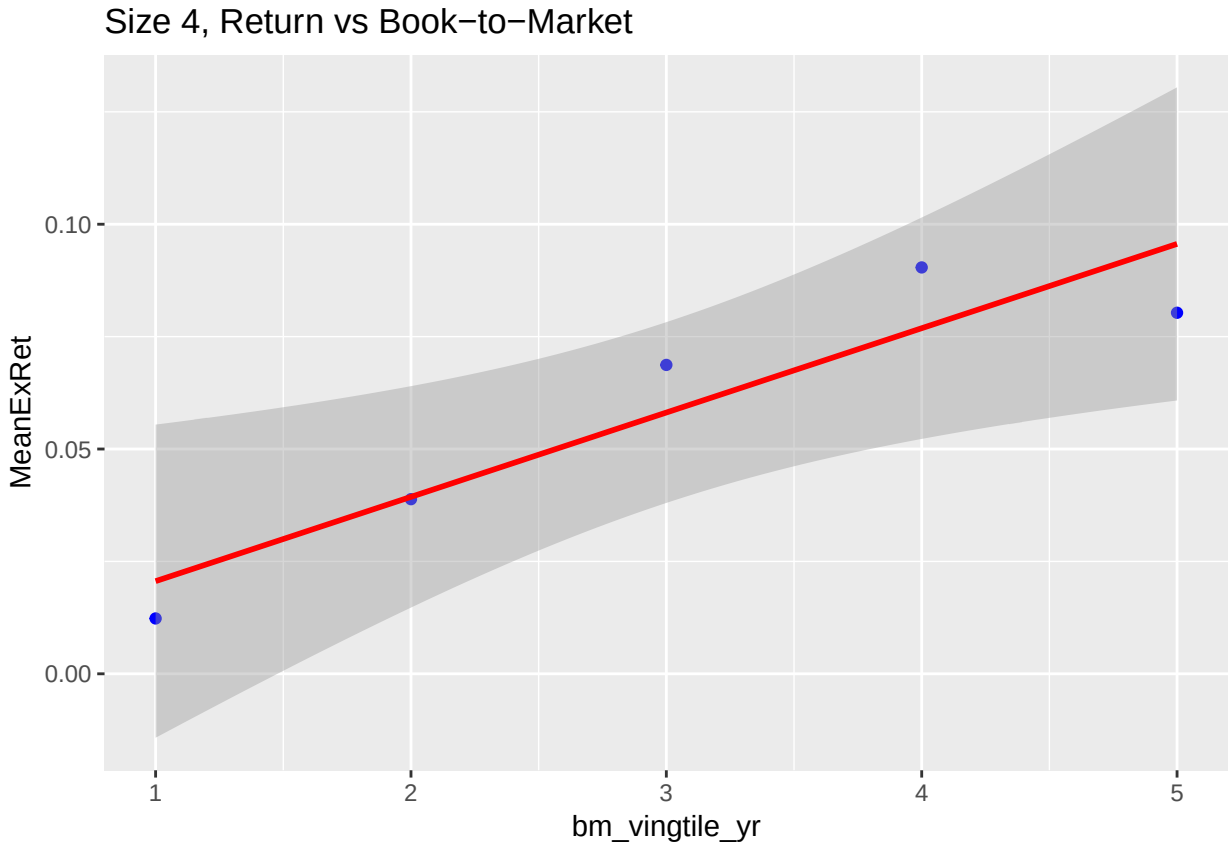
```
StockRetAcct_DT = as.data.table(read.dta('/Users/huanyu/Desktop/MachineLearning431/hw1/StockRetAcct_insa
StockRetAcct_DT = StockRetAcct_DT[!is.na((StockRetAcct_DT$lnBM)),]
StockRetAcct_DT = StockRetAcct_DT[!is.na((StockRetAcct_DT$lnME)),]
for (i in 1980:2014){
    StockRetAcct_DT[year == i,bm_vingtile_yr:=cut(StockRetAcct_DT[year == i,]$lnBM,breaks=quantile(Stock
}
for (i in 1980:2014){
    StockRetAcct_DT[year == i,size_vingtile_yr:=cut(StockRetAcct_DT[year == i,]$lnME,breaks=quantile(St
}

EW_BM_MutualFunds = StockRetAcct_DT[,list(MeanExRet = mean(lnAnnRet)), by = list(bm_vingtile_yr, size_vi
```

```
for (i in 1:5){
    plot = qplot(bm_vingtile_yr, MeanExRet, data = EW_BM_MutualFunds[size_vingtile_yr == i], col=I("blu
    print(plot)
}
```




Size 2, Return vs Book−to−Market
0.2
0.1
0.0
−0.1
MeanExRet
1
2
3
4
5
bm_vingtile_yr


Size 3, Return vs Book−to−Market
0.1
0.0
−0.1
MeanExRet
1
2
3
4
5
bm_vingtile_yr


Size 4, Return vs Book−to−Market
MeanExRet
0.10
0.05
0.00
1
2
3
4
5
bm_vingtile_yr


Size 5, Return vs Book−to−Market
MeanExRet
0.15
0.10
0.05
0.00
1
2
3
4
5
bm_vingtile_yr