

# hw1

Huanyu Liu

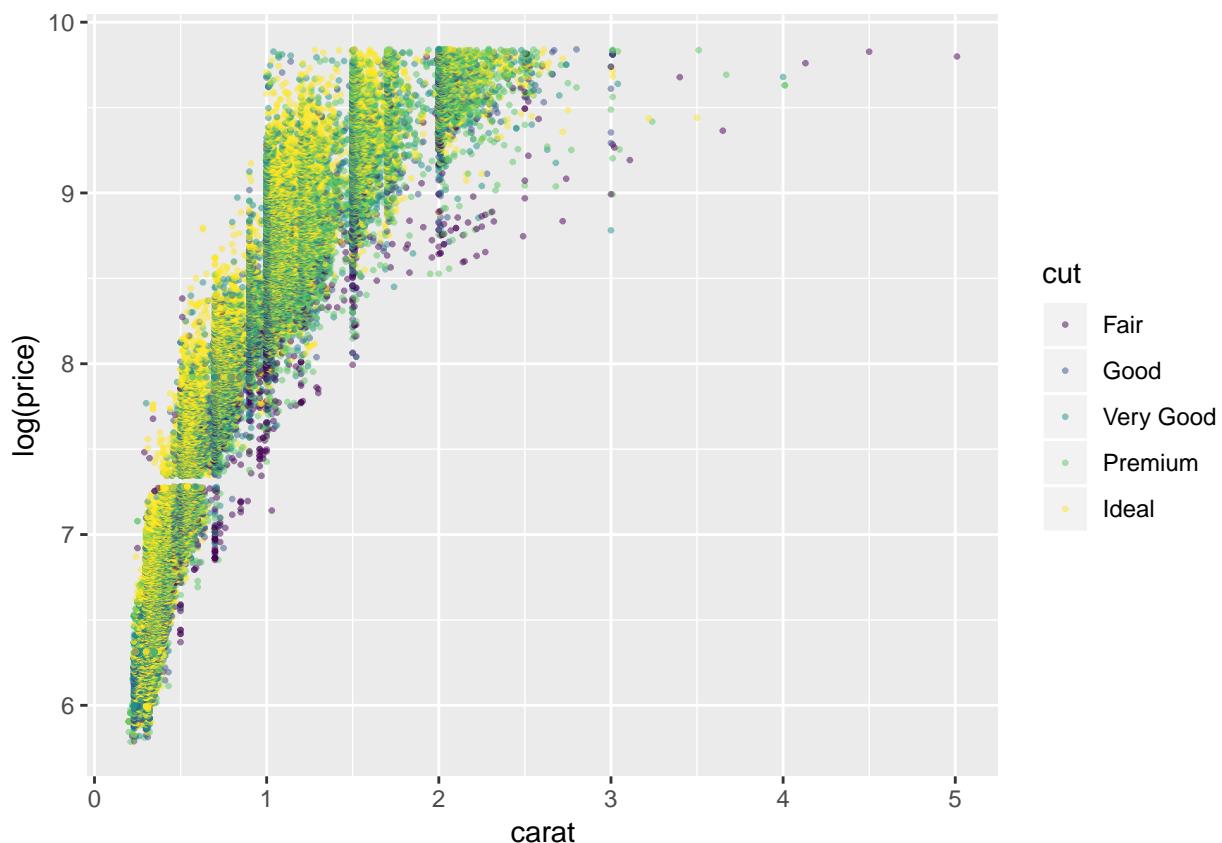
4/7/2019

## Problem 1

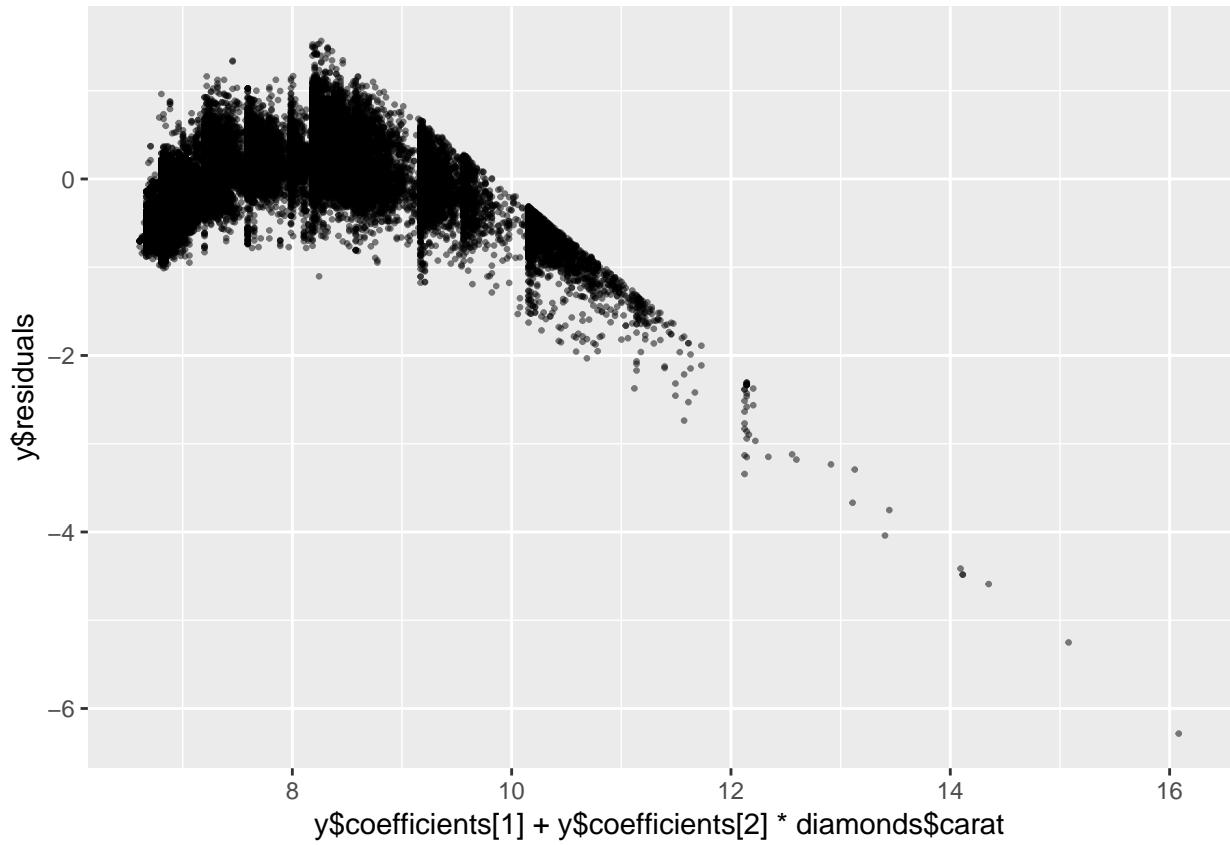
```
library('ggplot2')
library('foreign')
library('data.table')

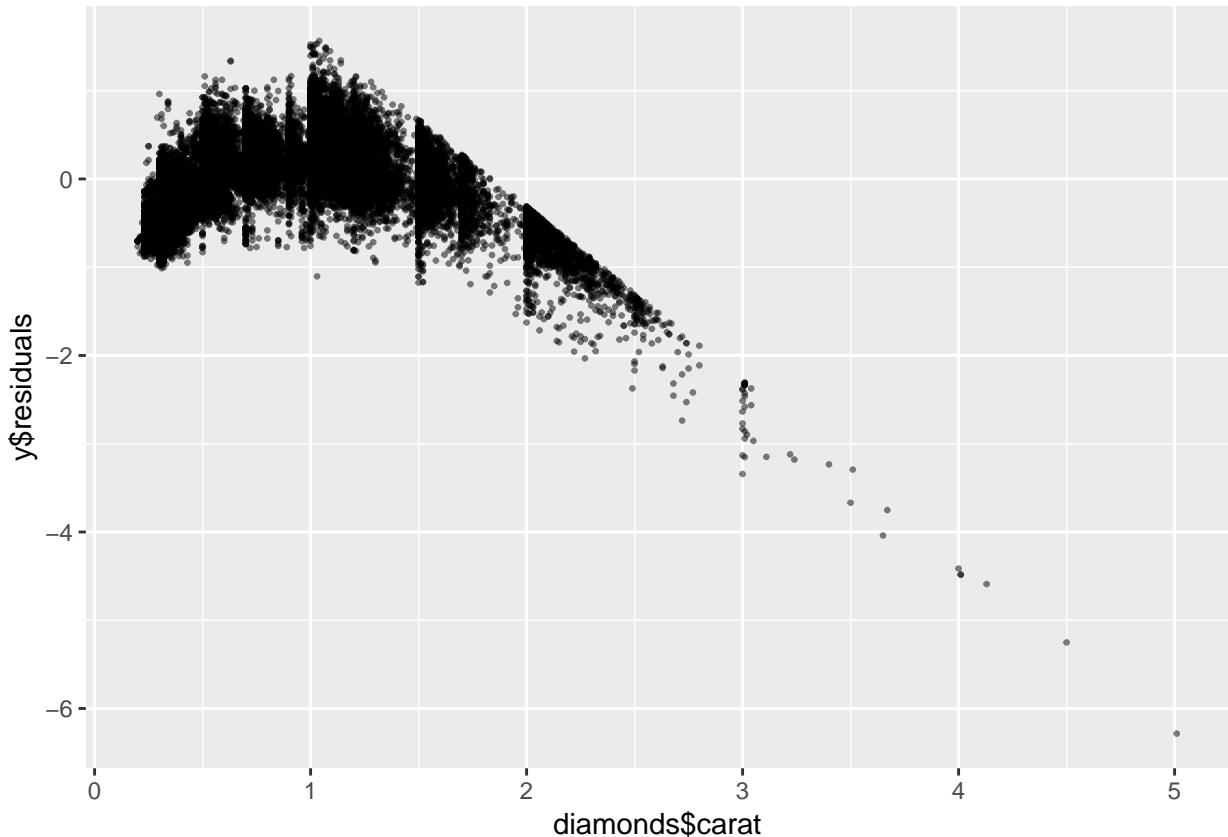
# 1

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))
```



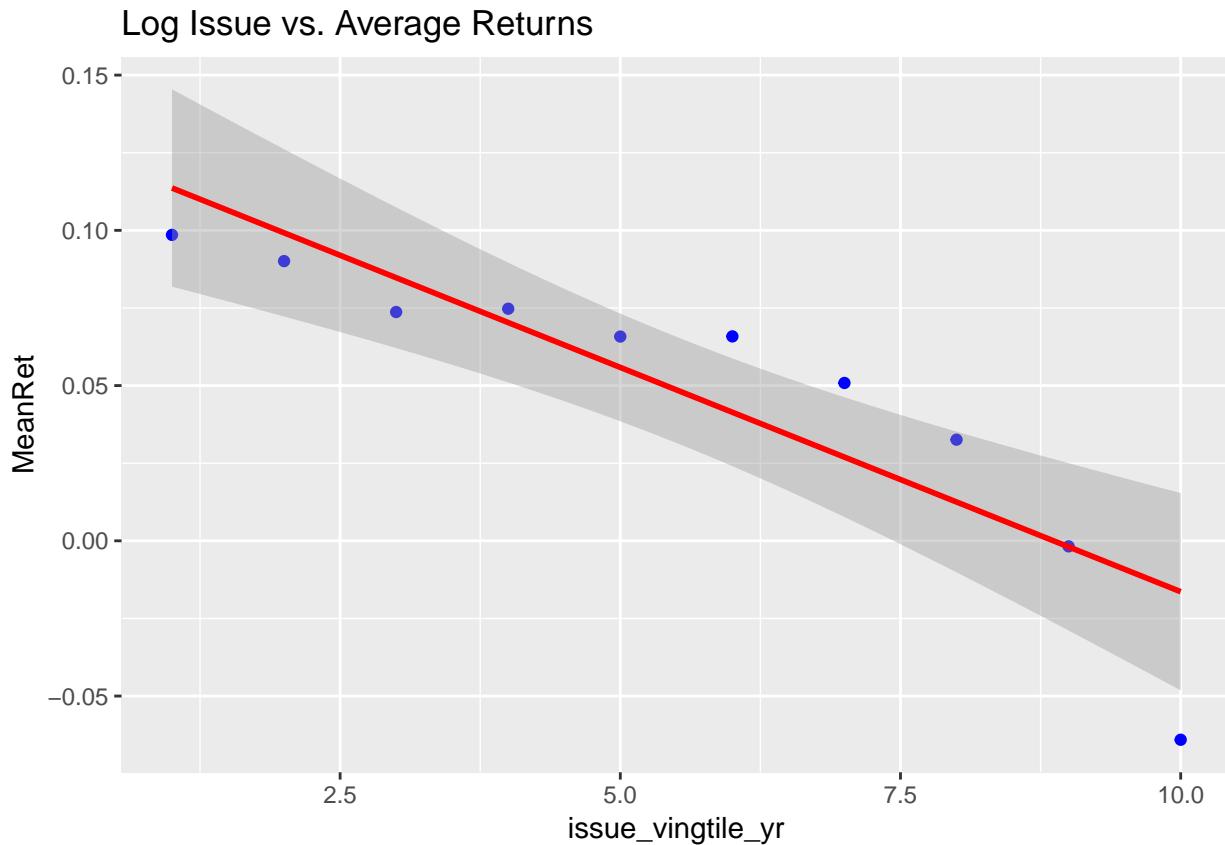


# The residuals are not random. They are correlated with carat. Therefore, the relationship between pri

## Problem 2

```
StockRetAcct_DT = as.data.table(read.dta('/Users/huanyu/Desktop/MachineLearning431/hw1/StockRetAcct_insample.dta'))
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=quantile(StockRetAcct_DT[year == i,]$lnIssue,probs=seq(0,1,by=0.05)))]
}

EW_Issue_MutualFunds = StockRetAcct_DT[,list(MeanRet = mean(lnAnnRet)), by = issue_vingtile_yr]
EW_Issue_MutualFunds = EW_Issue_MutualFunds[!is.na(EW_Issue_MutualFunds$issue_vingtile_yr),]
qplot(issue_vingtile_yr, MeanRet,data = EW_Issue_MutualFunds, col=I("blue"), na.rm = TRUE, main = "Log Return vs Vingtile Year")
```



```
# 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
```

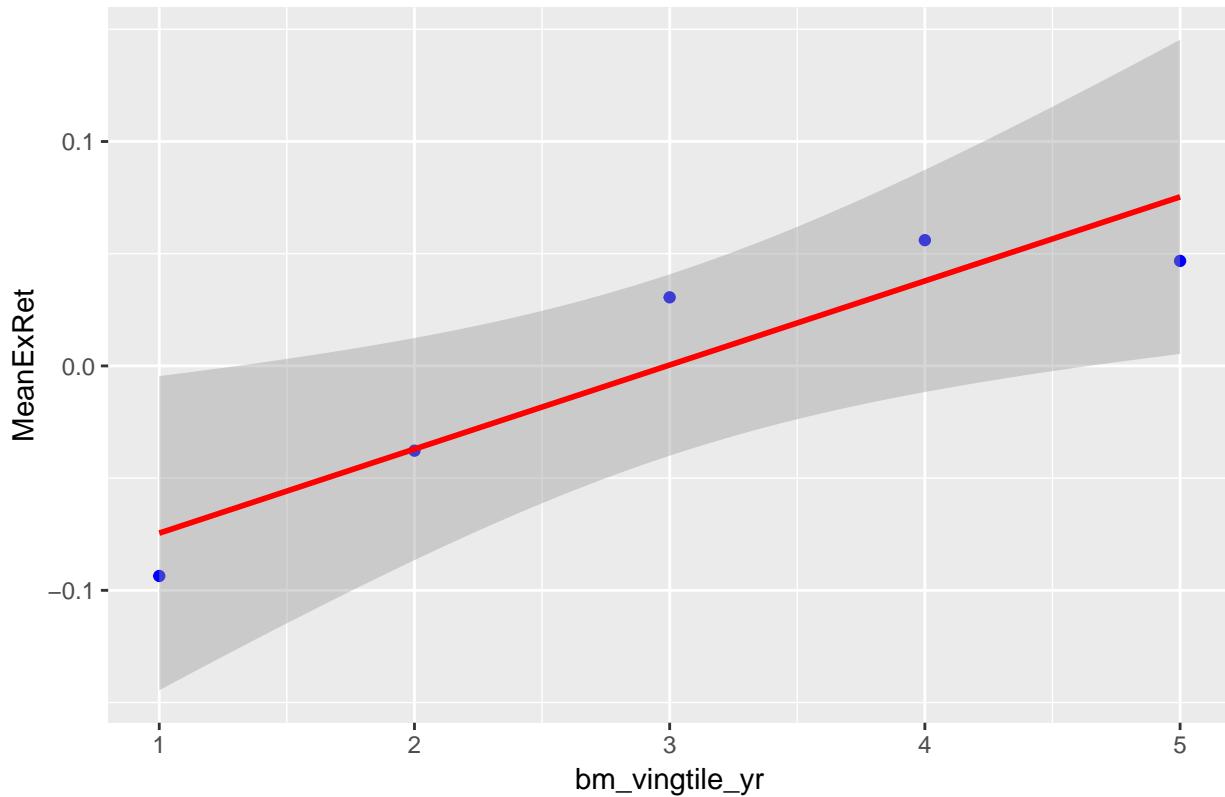
## Problem 3

```
StockRetAcct_DT = as.data.table(read.dta('/Users/huanyu/Desktop/MachineLearning431/hw1/StockRetAcct_ins'))
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,]$lnBM,breaks=quantile(StockRetAcct_DT$lnBM,probs=c(.25,.75)),labels=c('L','H')),]
}
for (i in 1980:2014){
  StockRetAcct_DT[year == i,size_vingtile_yr:=cut(StockRetAcct_DT[year == i,]$lnME,breaks=quantile(StockRetAcct_DT$lnME,probs=c(.25,.75)),labels=c('L','H')),]
}

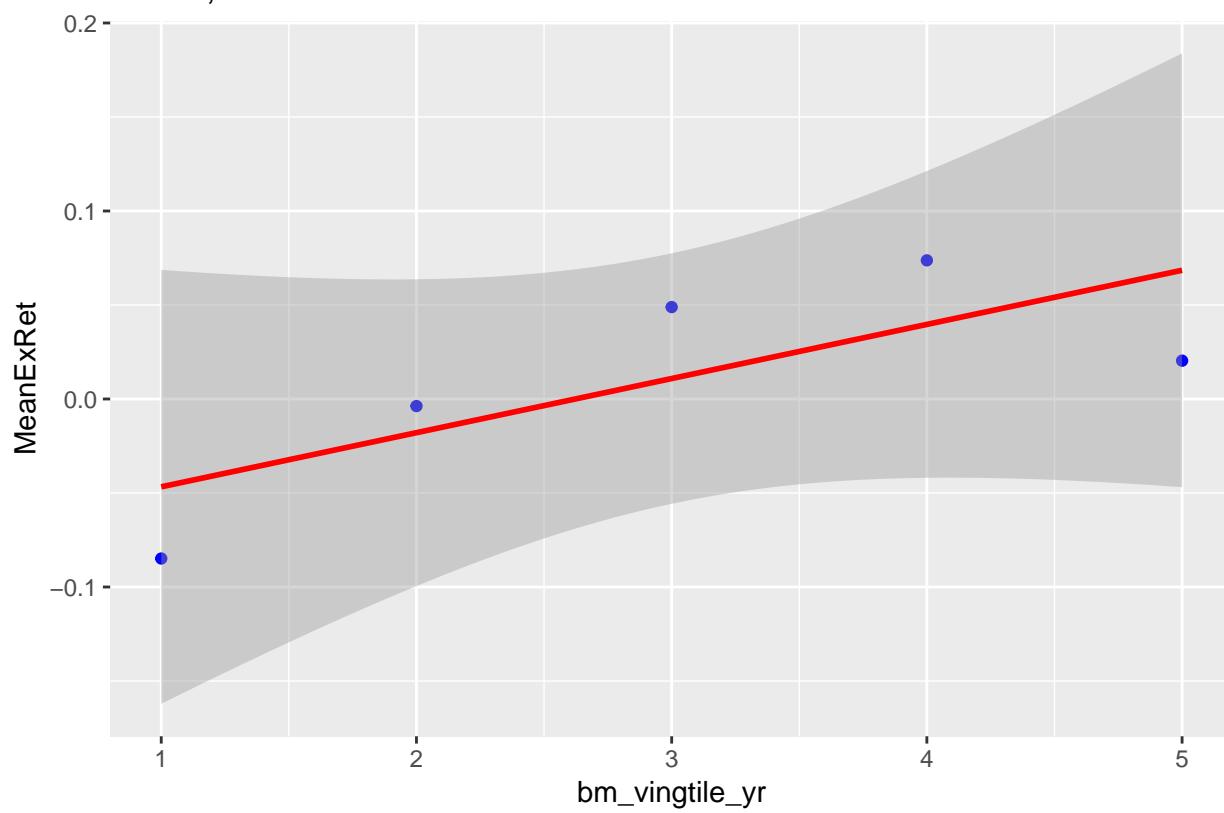
EW_BM_MutualFunds = StockRetAcct_DT[,list(MeanExRet = mean(lnAnnRet)), by = list(bm_vingtile_yr, size_vingtile_yr)]
```

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

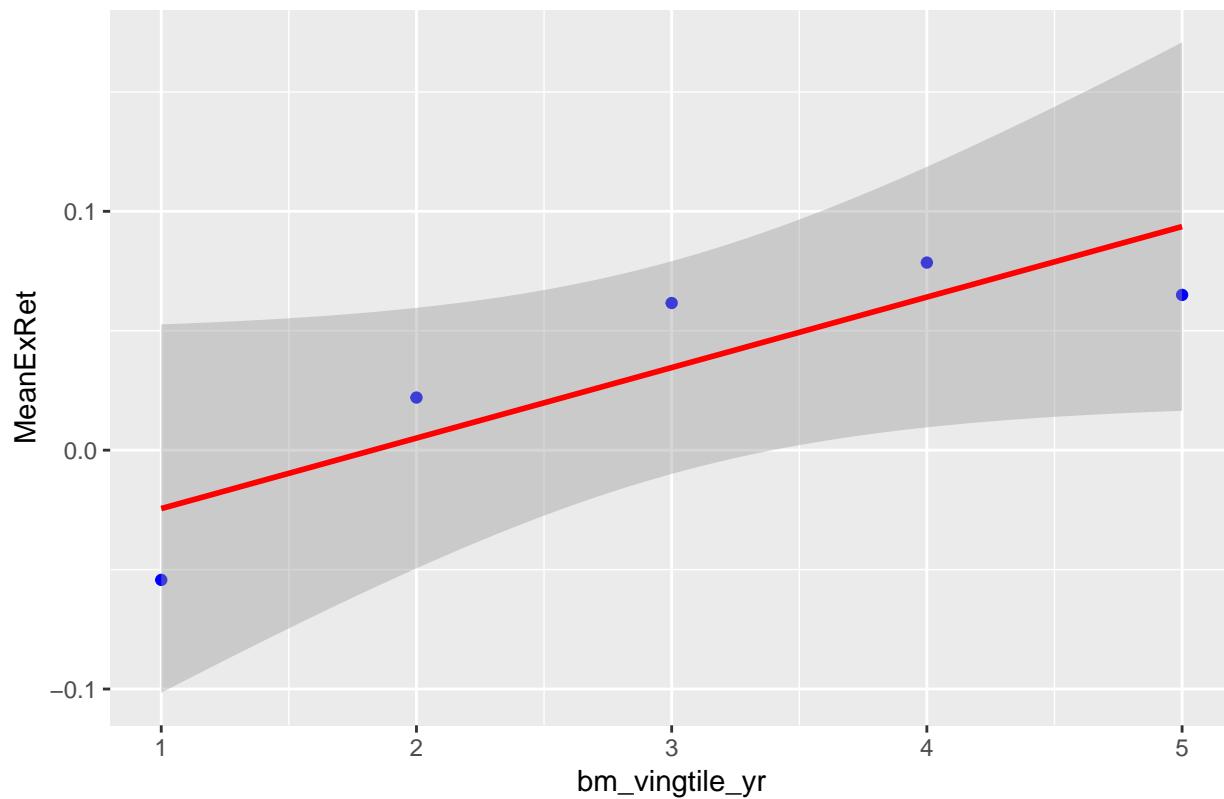
Size 1, Return vs Book-to-Market



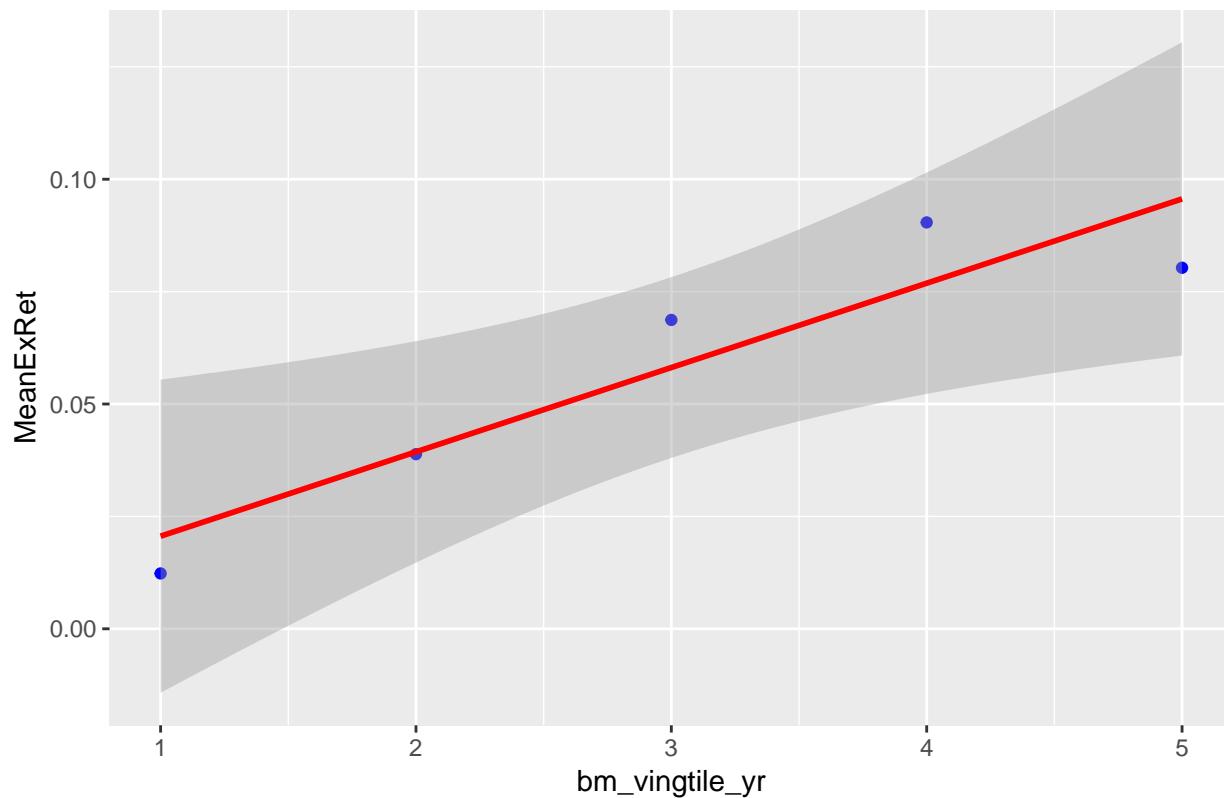
Size 2, Return vs Book-to-Market



Size 3, Return vs Book-to-Market



Size 4, Return vs Book-to-Market



Size 5, Return vs Book-to-Market

