HW 3

Zheng Li

5/7/2020

QF 104 HW 104 - Zheng Li

#Importing everything I need for the code

```
library("readxl")
library(fBasics)
library(reshape)
```

Question 1

#Merging the two database

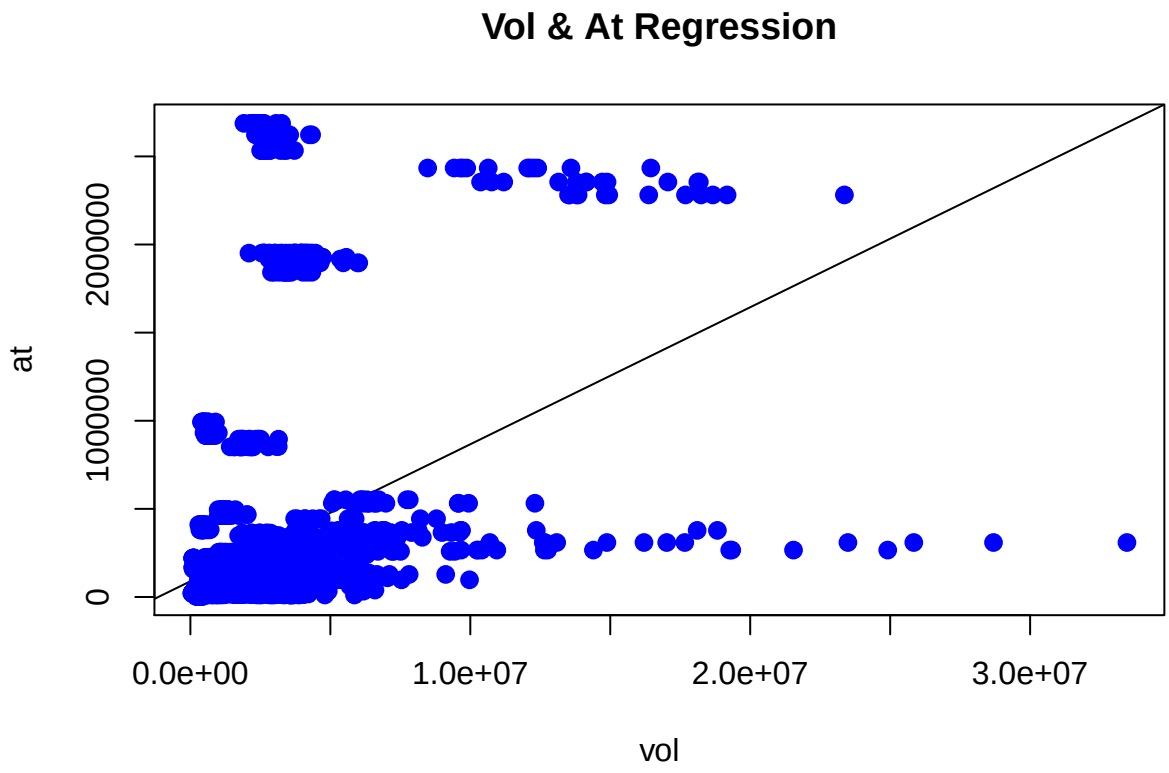
```
compustat <- read_excel("COMPUSTAT.xlsx")
crsp <- read_excel("CRSP.xlsx")

merged <- merge(compustat, crsp, by = c("fyear","cusip8"))
```

#Report and graph the linear regression model of data “vol”(x) and “at”(y).

```
x1 <- merged[["vol"]]
y1 <- merged[["at"]]
relation1 <- lm(y1~x1)
print(summary(relation1))
```

```
##
## Call:
## lm(formula = y1 ~ x1)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -2381395  -146521   -90266   -22832  2449581
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept) 8.907e+04  1.025e+04   8.691   <2e-16 ***
## x1          7.775e-02  3.132e-03  24.823   <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 441900 on 2830 degrees of freedom
## Multiple R-squared:  0.1788, Adjusted R-squared:  0.1785
## F-statistic: 616.2 on 1 and 2830 DF,  p-value: < 2.2e-16
```

```
plot(x1,y1,col = "blue",main = "Vol & At Regression",
     abline(relation1),cex = 1.3,pch = 16,xlab = "vol",ylab = "at")
```



Question 2

#Change the merged data set from monthly to yearly by using the aggregate function.

```
yearMerged1 <- aggregate(merged$vol, by=list(merged$fyear, merged$cusip8, merged$at), FUN=sum, na.rm=TRU
```

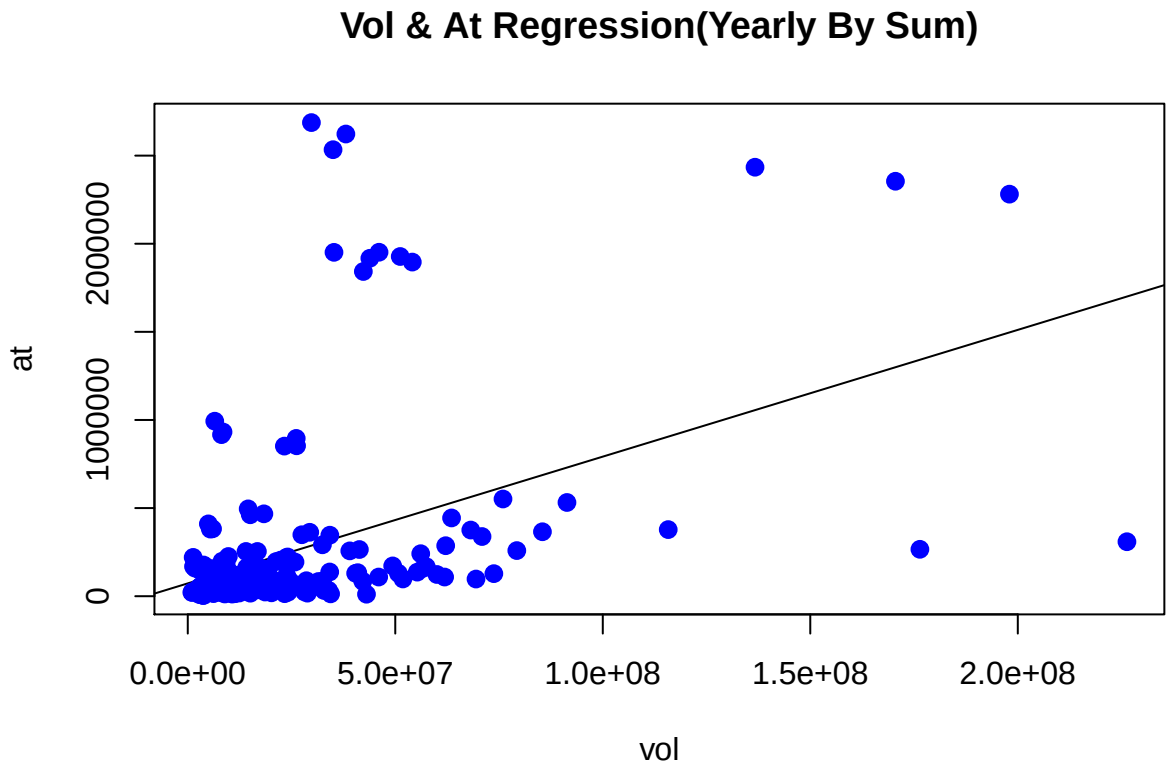
#Then just repeat the steps in question 1 to report and graph the new regression model.

```
x2 <- yearMerged1[["x"]]
y2 <- yearMerged1[["Group.3"]]
relation2 <- lm(y2~x2)
print(summary(relation2))
```

```
##
## Call:
## lm(formula = y2 ~ x2)
##
## Residuals:
##       Min        1Q    Median        3Q       Max
## -1391253  -144044   -83215   -24333  2400430
##
## Coefficients:
```

```
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept) 7.265e+04  3.587e+04   2.025    0.044 *
## x2          7.193e-03  9.449e-04   7.612 6.58e-13 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 438300 on 234 degrees of freedom
## Multiple R-squared:  0.1985, Adjusted R-squared:  0.1951
## F-statistic: 57.95 on 1 and 234 DF,  p-value: 6.581e-13
```

```
plot(x2,y2,col = "blue",main = "Vol & At Regression(Yearly By Sum)",
     abline(relation2),cex = 1.3,pch = 16,xlab = "vol",ylab = "at")
```



Question 3

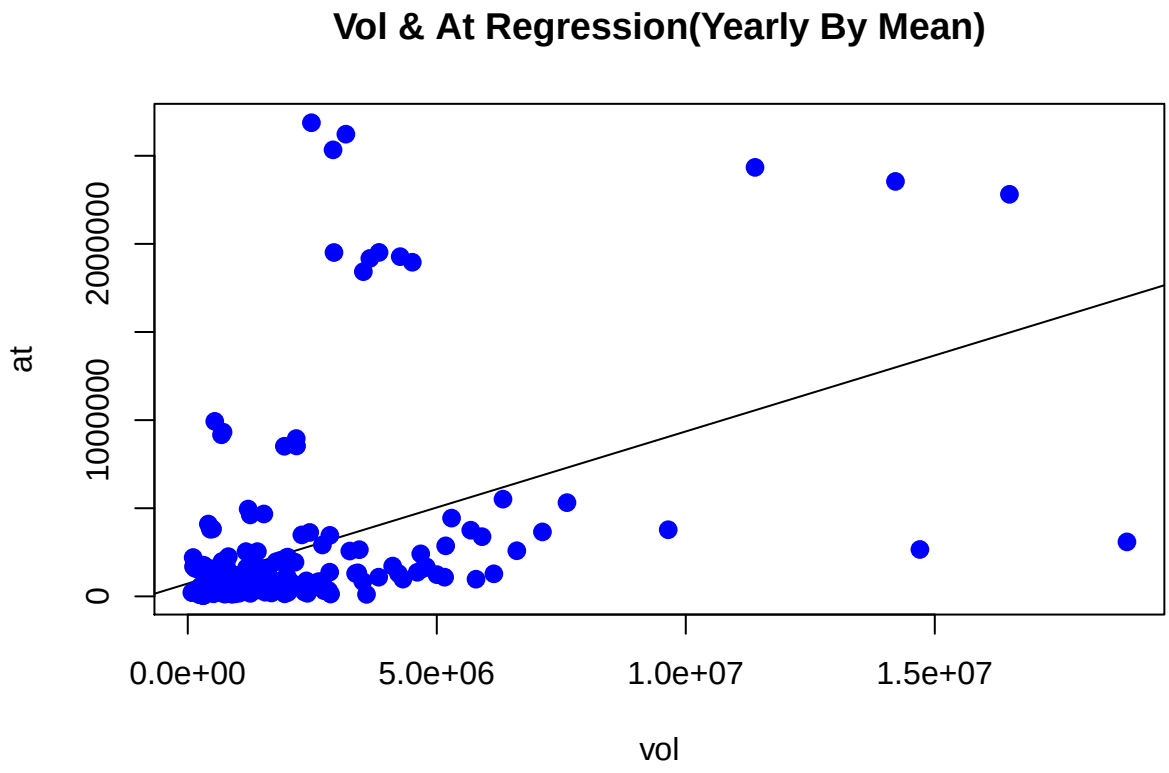
#Repeat what I did in question 2 but change "sum" to "mean" for aggregation.

```
yearMerged2 <- aggregate(merged$vol, by=list(merged$fyear, merged$cusip8, merged$at), FUN=mean, na.rm=T
x3 <- yearMerged2[["x"]]
y3 <- yearMerged2[["Group.3"]]
relation3 <- lm(y3~x3)
print(summary(relation3))
```

```
##
## Call:
## lm(formula = y3 ~ x3)
```

```
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1391253  -144044   -83215   -24333  2400430
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept) 7.265e+04  3.587e+04   2.025    0.044 *
## x3          8.631e-02  1.134e-02   7.612 6.58e-13 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 438300 on 234 degrees of freedom
## Multiple R-squared:  0.1985, Adjusted R-squared:  0.1951
## F-statistic: 57.95 on 1 and 234 DF,  p-value: 6.581e-13
```

```
plot(x3,y3,col = "blue",main = "Vol & At Regression(Yearly By Mean)",
     abline(relation3),cex = 1.3,pch = 16,xlab = "vol",ylab = "at")
```



Question 4

#Tabulate the coefficients from your three regressions.

```
Table <- matrix(c(relation1$coefficients, relation2$coefficients, relation3$coefficients),ncol = 2, byr
colnames(Table) <- c("Intercept","x")
rownames(Table) <- c("Montly(Question 1)","Yearly By Sum(Question 2)","Yearly By Mean(Question 3)")
print(Table)
```

```
##                              Intercept           x
## Montly(Question 1)            89068.21 0.077749490
## Yearly By Sum(Question 2)     72645.18 0.007192748
## Yearly By Mean(Question 3)    72645.18 0.086312975
```