Homework 4

Hanao Li hl3202

10/22/2018

```
library(plyr)
debt <- read.csv('debt.csv', as.is = TRUE)
dim(debt)
```

```
## [1] 1171    4
```

```
head(debt)
```

```
##     Country Year    growth     ratio
## 1 Australia 1946 -3.557951 190.41908
## 2 Australia 1947  2.459475 177.32137
## 3 Australia 1948  6.437534 148.92981
## 4 Australia 1949  6.611994 125.82870
## 5 Australia 1950  6.920201 109.80940
## 6 Australia 1951  4.272612  87.09448
```

1.

a.

```
mean.growth <- function(debt){
  return(mean(debt$growth))
}
```

b.

```
library(plyr)
country <- signif(daply(debt, .(Country), mean.growth), 3)
country[names(country) == "Australia"]
```

```
## Australia
##      3.72
```

```
country[names(country) == "Netherlands"]
```

```
## Netherlands
##        3.03
```

2.

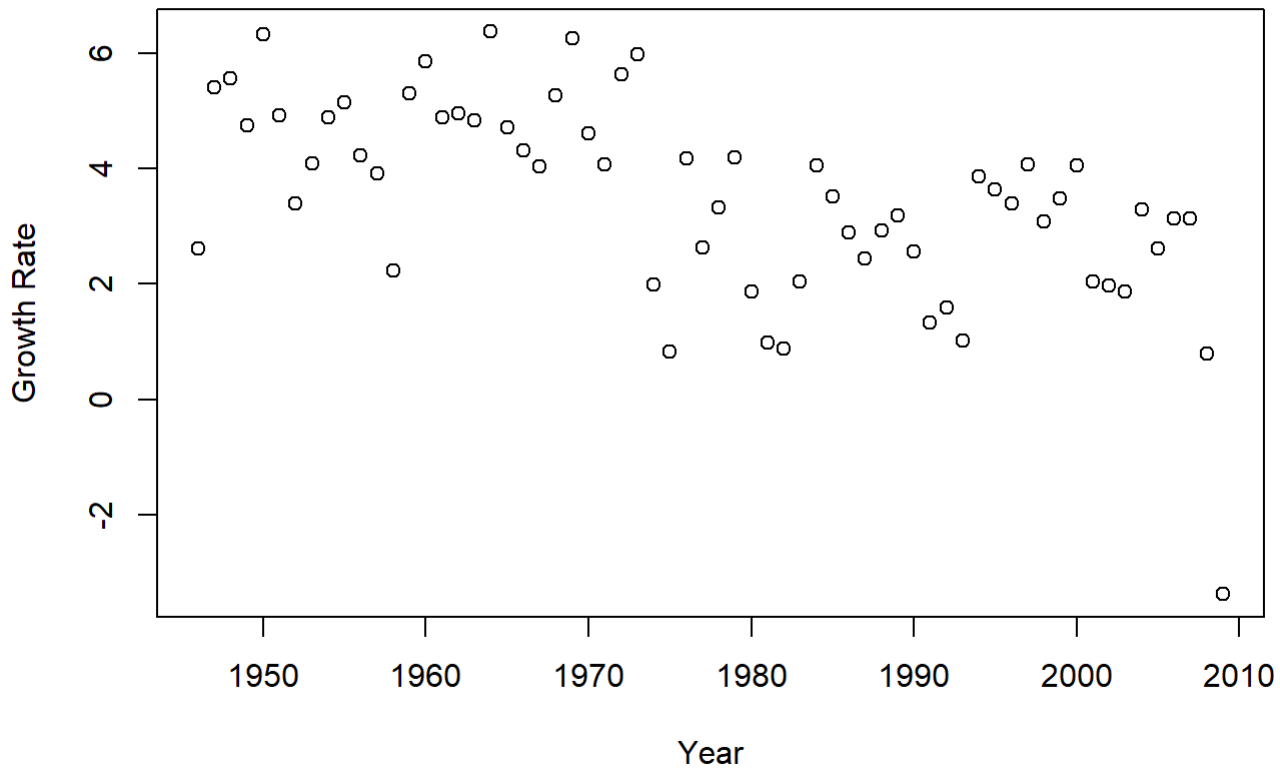
```
rate <- signif(daply(debt, .(Year), mean.growth), 3)
rate[names(rate) == 1972]
```

```
## 1972
## 5.63
```

```
rate[names(rate) == 1989]
```

```
## 1989
## 3.19
```

```
  plot(names(rate), rate, ylab = 'Growth Rate', xlab = 'Year')
```



3.

a.

```
signif(cor(debt$growth, debt$ratio), 4)
```

```
## [1] -0.1995
```

b.

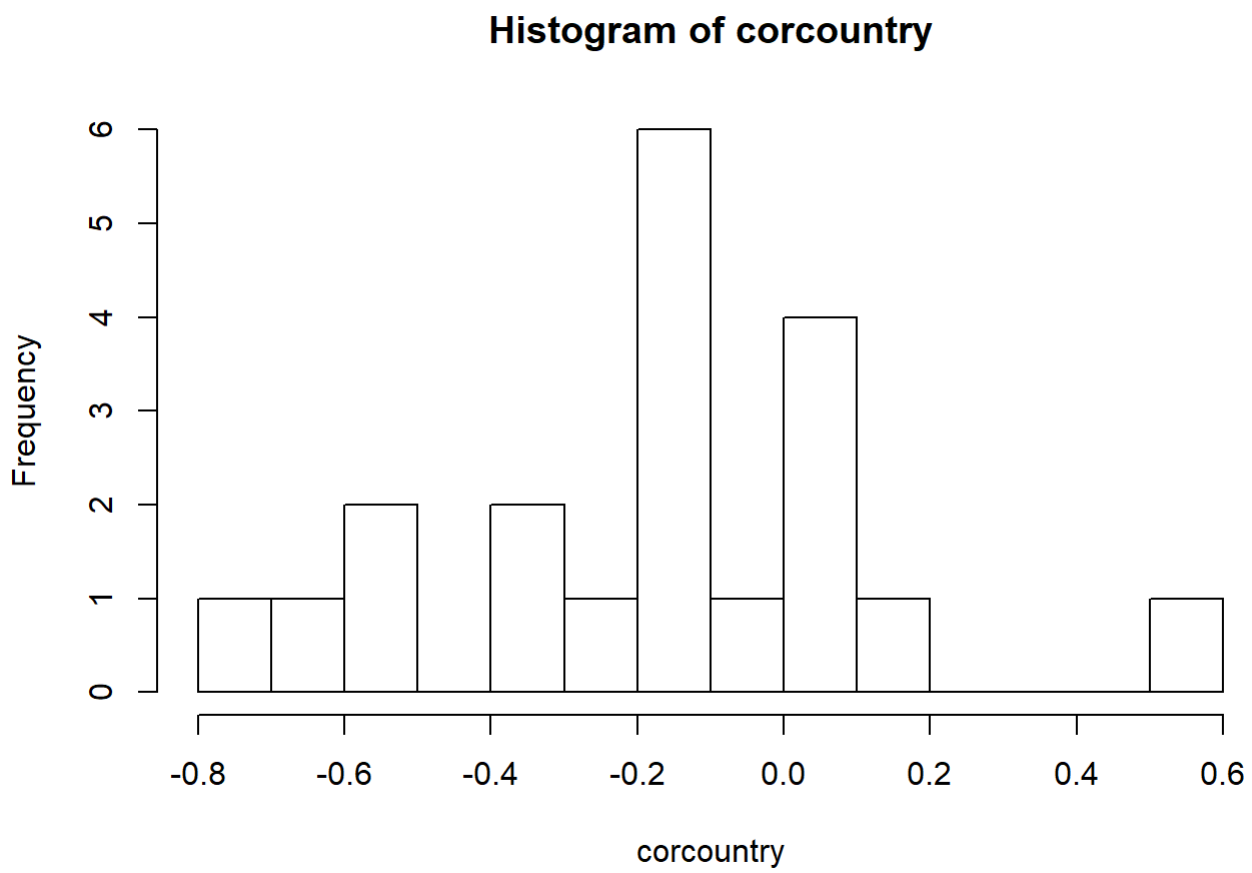
```
corfun <- function(debt){
  return(cor(debt$growth, debt$ratio))
}
signif(corcountry <- daply(debt, .(Country), corfun), 3)
```

```
##   Australia     Austria     Belgium      Canada     Denmark     Finland
##    0.025200   -0.253000   -0.192000    0.075000   -0.168000    0.000581
##      France     Germany      Greece     Ireland       Italy       Japan
##   -0.502000   -0.576000   -0.093500   -0.140000   -0.645000   -0.702000
## Netherlands New Zealand      Norway    Portugal       Spain      Sweden
##   -0.199000    0.161000    0.563000   -0.352000    0.081400   -0.161000
##          UK          US
##   -0.137000   -0.341000
```

```
signif(mean(corcountry), 4)
```

```
## [1] -0.1778
```

```
hist(corcountry, breaks = 10)
```



c.

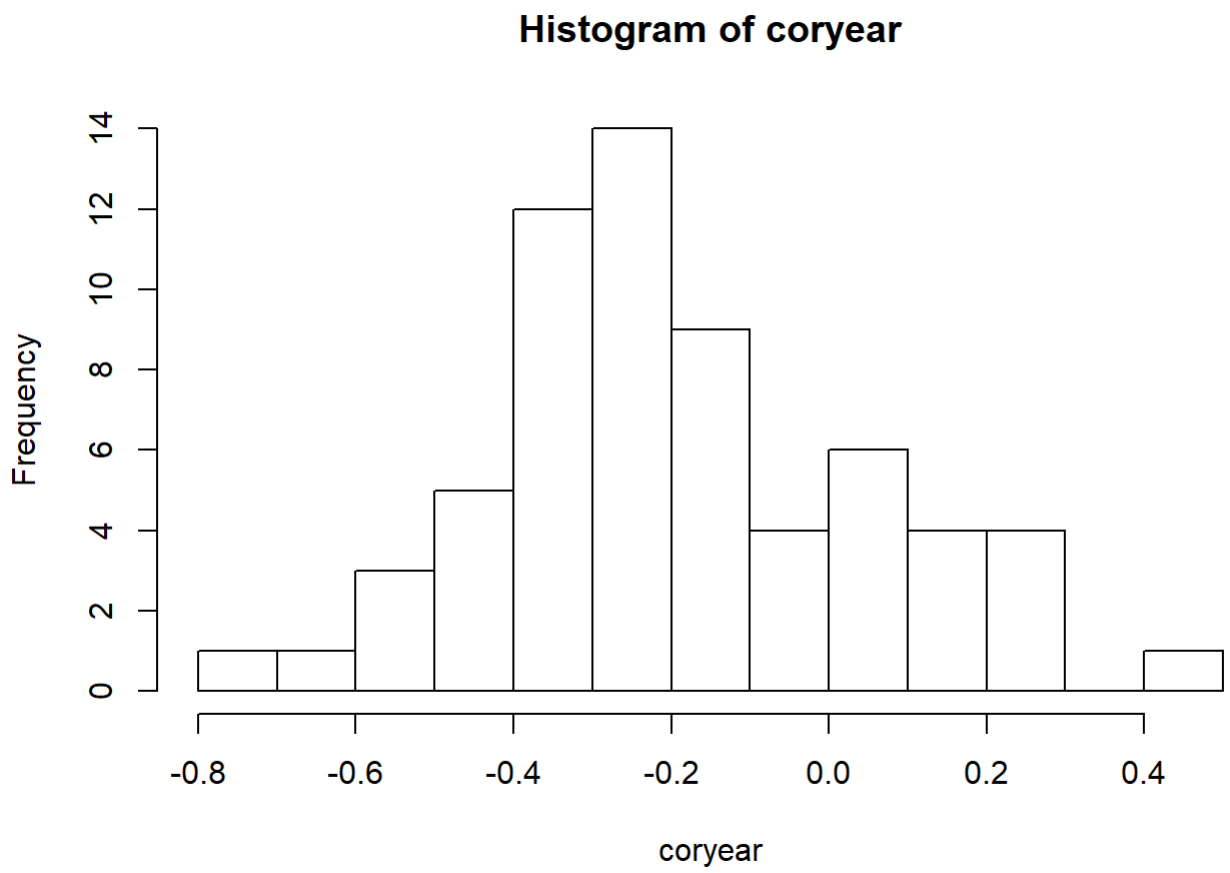
```
signif((coryear <- daply(debt, .(Year), corfun)), 3)
```

```
##     1946     1947     1948     1949     1950     1951     1952     1953
## -0.62000 -0.27400 -0.34000 -0.20000  0.03980 -0.41600 -0.27700 -0.20500
##     1954     1955     1956     1957     1958     1959     1960     1961
## -0.27500 -0.22700 -0.45800 -0.75500 -0.45400 -0.28500 -0.50400 -0.53900
##     1962     1963     1964     1965     1966     1967     1968     1969
## -0.38300  0.12800 -0.36100 -0.31100 -0.31100 -0.27800 -0.18100 -0.25000
##     1970     1971     1972     1973     1974     1975     1976     1977
## -0.51200  0.00872 -0.19600  0.11400  0.26000  0.27100 -0.17100  0.16400
##     1978     1979     1980     1981     1982     1983     1984     1985
##  0.43100 -0.42900 -0.12700  0.03040  0.23900 -0.36200 -0.15600 -0.44900
##     1986     1987     1988     1989     1990     1991     1992     1993
## -0.35800 -0.06890  0.07970  0.06640  0.15600  0.20200 -0.00222 -0.37200
##     1994     1995     1996     1997     1998     1999     2000     2001
## -0.22400  0.05190 -0.35700 -0.11100 -0.26500 -0.25800 -0.13400 -0.23800
##     2002     2003     2004     2005     2006     2007     2008     2009
## -0.34900 -0.06790 -0.17100 -0.31400 -0.19600 -0.34400 -0.09450 -0.20500
```

```
signif(mean(coryear), 4)
```

```
## [1] -0.1906
```

```
hist(coryear, breaks = 10)
```



d.

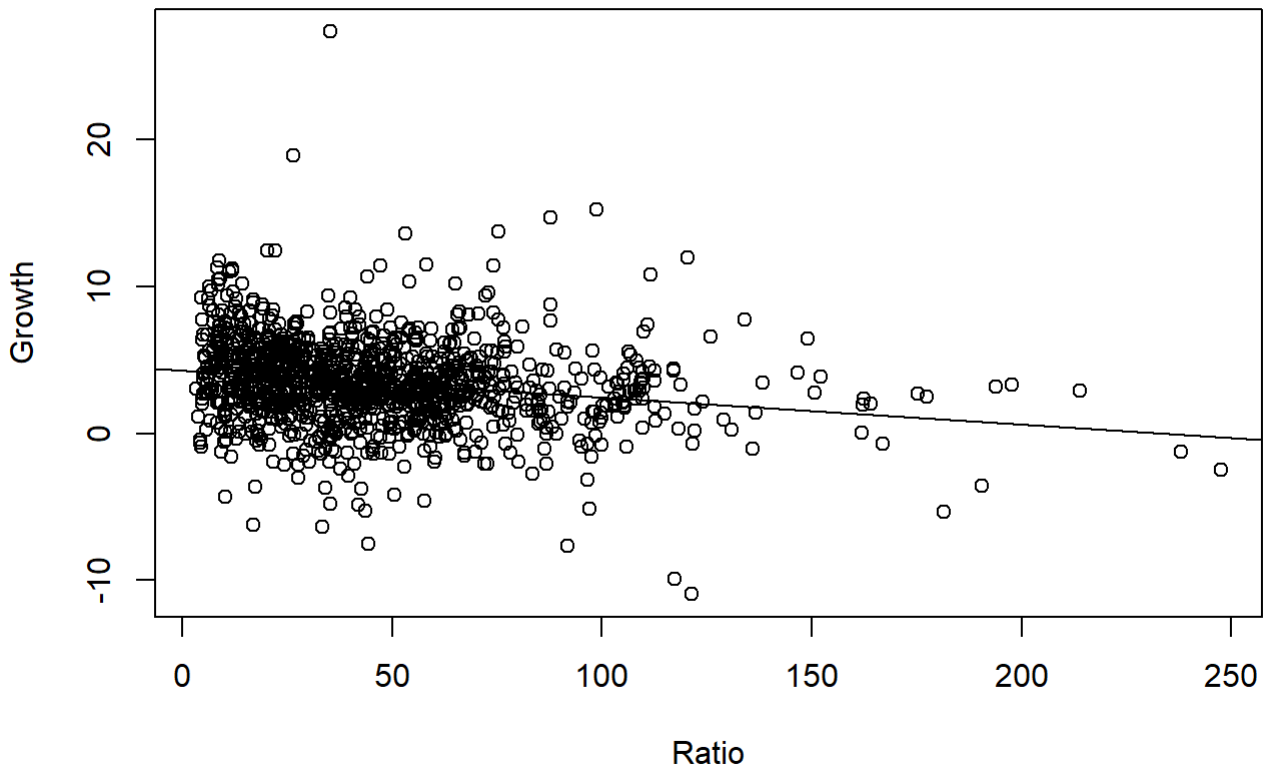
```
# France, Germany, Japan, Italy these four countries have a correlation less than -0.5 and Norwa
y has a correlation larger than 0.5.
# Year 1946, 1957, 1960, 1961, 1970 has a correlation less than -0.5 and Year 1978 has a correla
tion larger than 0.4.
```

4.

```
fit <- lm(debt$growth ~ debt$ratio)
signif(fit$coefficients, 3)
```

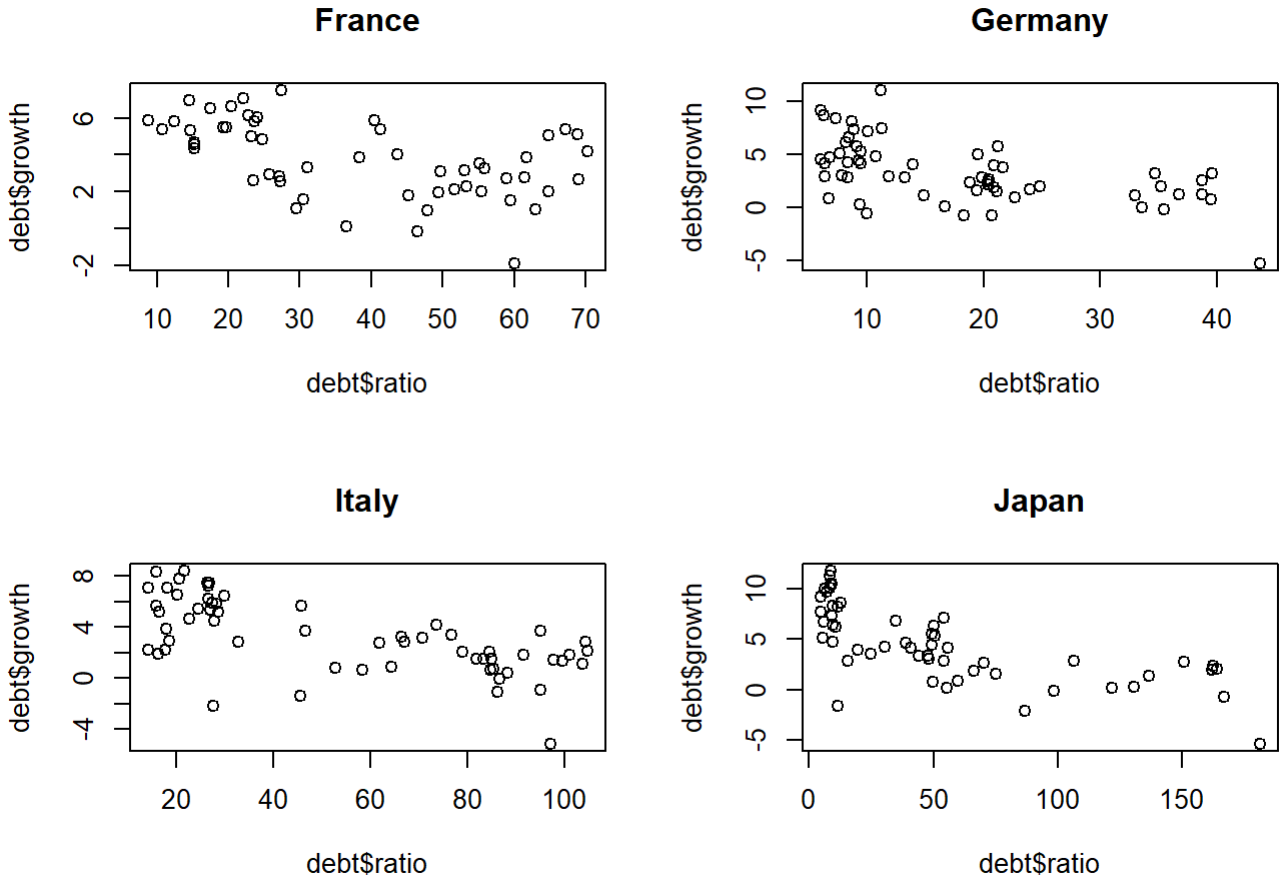
```
## (Intercept)  debt$ratio
##      4.2800     -0.0184
```

```
plot(debt$ratio, debt$growth, xlab = 'Ratio', ylab = 'Growth')
abline(fit)
```



5.

```
plotcor <- function(debt){
  if (debt$Country %in% names(corcountry[corcountry <= -0.5])){
  plot(debt$ratio, debt$growth)
  title(main = debt$Country[1])
  }
}
par(mfrow = c(2,2))
suppressWarnings(d_ply(debt, .(Country), plotcor))
```



6.

a.

```
france <- debt[debt$Country == 'France',]
dim(france)
```

```
## [1] 54  4
```

b.

```
suppressMessages(library(dplyr))
france$next.growth <- ifelse((france$Year + 1) %in% france$Year,  round(lead(france$growth), 3),
NA)
print(france[(france$Year == 1971 | france$Year == 1972),])
```

```
##     Country Year   growth     ratio next.growth
## 392  France 1971 5.372329 10.770552       5.886
## 393  France 1972 5.885827  8.757901          NA
```

7.

```
debt$next.growth <- ifelse((debt$Year + 1) == lead(debt$Year), round(lead(debt$growth), 3), NA)
print(debt[(debt$Country == 'France' & debt$Year == 2009),])
```

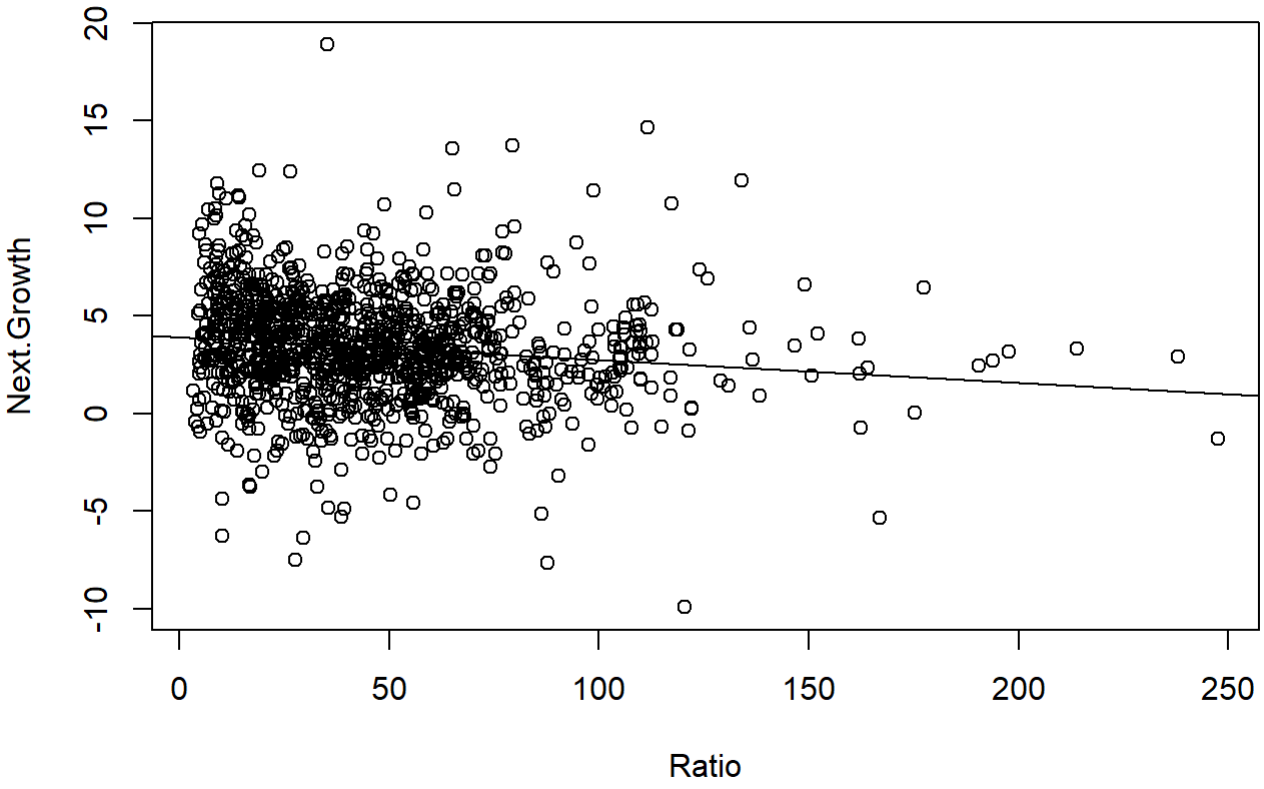
```
##     Country Year    growth    ratio next.growth
## 424  France 2009 -1.906676 60.00151          NA
```

8.

```
fit2 <- lm(debt$next.growth ~ debt$ratio)
signif(fit2$coefficients, 3)
```

```
## (Intercept)  debt$ratio
##      3.9200     -0.0116
```

```
plot(debt$ratio, debt$next.growth, xlab = 'Ratio', ylab = 'Next.Growth')
abline(fit2)
```




```
summary(fit)
```

```
##
## Call:
## lm(formula = debt$growth ~ debt$ratio)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -12.9958  -1.5200  -0.0774   1.5707  23.6960
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)  4.279290   0.148970   28.73  < 2e-16 ***
## debt$ratio  -0.018355   0.002637   -6.96 5.67e-12 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 2.922 on 1169 degrees of freedom
## Multiple R-squared:  0.03979,    Adjusted R-squared:  0.03897
## F-statistic: 48.44 on 1 and 1169 DF,  p-value: 5.666e-12
```

```
summary(fit2)
```

```
##
## Call:
## lm(formula = debt$next.growth ~ debt$ratio)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -12.4487  -1.4566  -0.0376   1.6331  15.3859
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)  3.924729   0.143851  27.283  < 2e-16 ***
## debt$ratio  -0.011608   0.002555  -4.544  6.1e-06 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 2.789 on 1145 degrees of freedom
##   (24 observations deleted due to missingness)
## Multiple R-squared:  0.01771,    Adjusted R-squared:  0.01686
## F-statistic: 20.65 on 1 and 1145 DF,  p-value: 6.104e-06
```

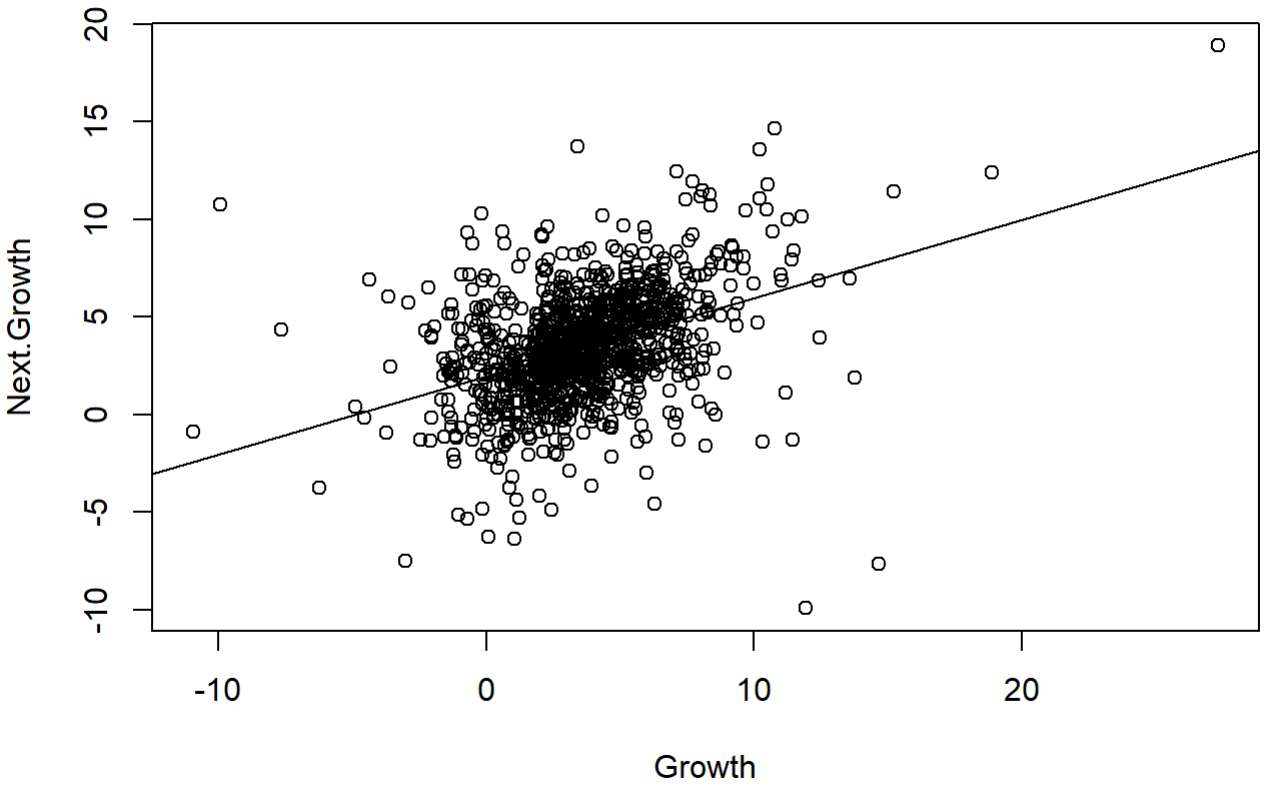
```
# From the summary,we could see that the previous model has a larger R squared value compared to this model, although both of their R squared values are very small. But this model has a smaller residual standard error. From the plots, we could tell that both of them are not doing very wel l.
```

9.

```
fit3 <- lm(debt$next.growth ~ debt$growth)
signif(fit3$coefficients, 3)
```

```
## (Intercept) debt$growth
##       1.970       0.401
```

```
plot(debt$growth, debt$next.growth, xlab = 'Growth', ylab = 'Next.Growth')
abline(fit3)
```




```
summary(fit2)
```

```
##
## Call:
## lm(formula = debt$next.growth ~ debt$ratio)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -12.4487  -1.4566  -0.0376   1.6331  15.3859
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)  3.924729   0.143851  27.283  < 2e-16 ***
## debt$ratio  -0.011608   0.002555  -4.544  6.1e-06 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 2.789 on 1145 degrees of freedom
##   (24 observations deleted due to missingness)
## Multiple R-squared:  0.01771,    Adjusted R-squared:  0.01686
## F-statistic: 20.65 on 1 and 1145 DF,  p-value: 6.104e-06
```

```
summary(fit3)
```

```
##
## Call:
## lm(formula = debt$next.growth ~ debt$growth)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -16.6737  -1.3568   0.0398   1.3995  12.7912
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)  1.97107    0.12040   16.37   <2e-16 ***
## debt$growth  0.40065    0.02643   15.16   <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 2.568 on 1145 degrees of freedom
##   (24 observations deleted due to missingness)
## Multiple R-squared:  0.1671, Adjusted R-squared:  0.1664
## F-statistic: 229.8 on 1 and 1145 DF,  p-value: < 2.2e-16
```

```
# From the summary we could see that this model has a larger R squared value and a smaller resid
ual standard error and from the graph we could see that this model is doing pretty good compared
to the previous model. So we can conclude that current growth is a better predictor of future gr
owth.
```