Homework 4

UNI: qy2205
Name: Quan Yuan
Email: qy2205@columbia.edu

```
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)
library(plyr)
mean.growth <- function (data){
  return(mean(data$growth))
}
# (b)
country.ave.growth <- signif(daply(debt, .(Country), mean.growth), 3)
country.ave.growth
```

```
##   Australia     Austria     Belgium      Canada     Denmark     Finland
##        3.72        4.44        3.18        3.65        2.66        3.57
##      France     Germany      Greece     Ireland       Italy       Japan
##        3.78        3.31        2.93        3.93        3.25        4.45
## Netherlands New Zealand      Norway    Portugal       Spain      Sweden
##        3.03        3.07        3.83        4.00        3.20        3.07
##          UK          US
##        2.41        3.00
```

```
country.ave.growth[c("Australia","Netherlands")]
```

```
##   Australia Netherlands
##        3.72        3.03
```

Report: the average GDP growth rates is highest in Japan and lowest in UK. Austria, Jpan and Portugal have high average GPA growth rate (above 4). UK, Greece and Denmark have low GPA growth rate (below 3).

2

```
each.year.ave.growth <- signif(daply(debt, .(Year), mean.growth), 3)
each.year.ave.growth
```

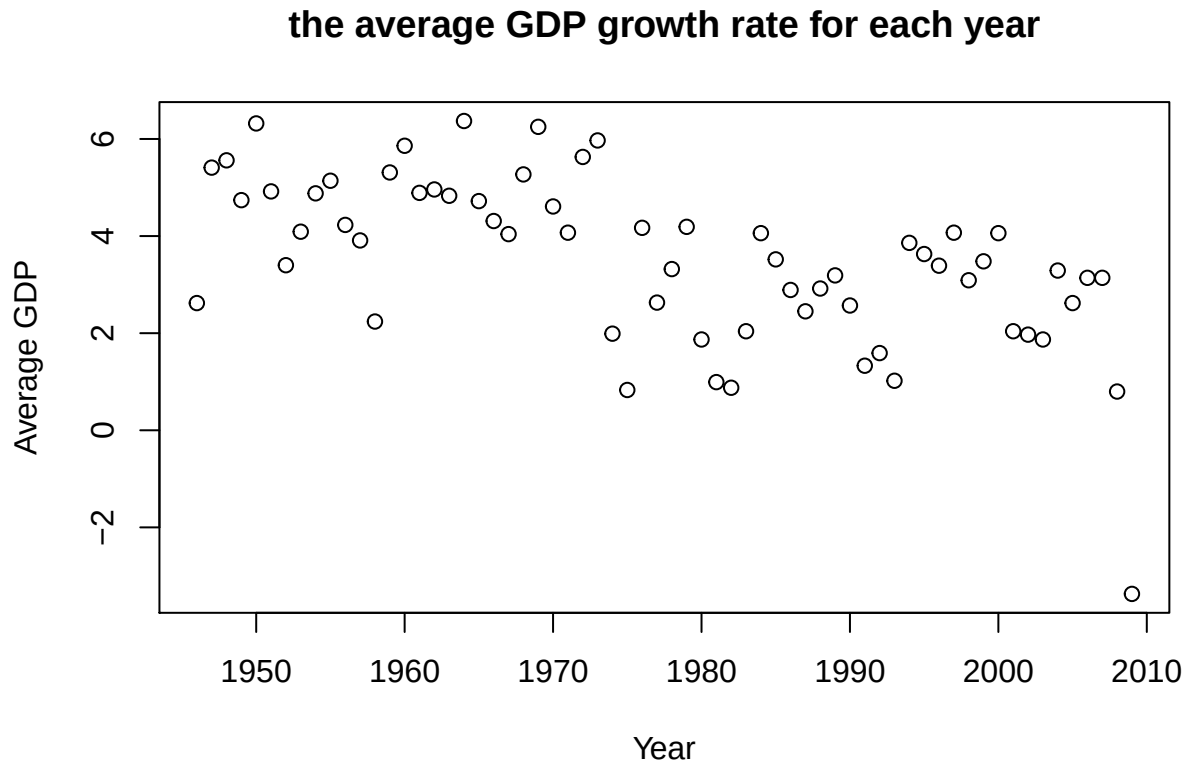
```
##   1946   1947   1948   1949   1950   1951   1952   1953   1954   1955
```

```
##  2.620  5.410  5.560  4.740  6.320  4.920  3.400  4.090  4.880  5.140
##   1956   1957   1958   1959   1960   1961   1962   1963   1964   1965
##  4.230  3.910  2.240  5.310  5.860  4.890  4.960  4.830  6.370  4.720
##   1966   1967   1968   1969   1970   1971   1972   1973   1974   1975
##  4.310  4.040  5.270  6.250  4.610  4.070  5.630  5.970  1.990  0.830
##   1976   1977   1978   1979   1980   1981   1982   1983   1984   1985
##  4.170  2.630  3.320  4.190  1.870  0.992  0.876  2.040  4.060  3.520
##   1986   1987   1988   1989   1990   1991   1992   1993   1994   1995
##  2.890  2.450  2.920  3.190  2.570  1.330  1.590  1.020  3.860  3.630
##   1996   1997   1998   1999   2000   2001   2002   2003   2004   2005
##  3.390  4.070  3.090  3.480  4.060  2.040  1.970  1.870  3.290  2.620
##   2006   2007   2008   2009
##  3.140  3.140  0.798 -3.370
```

```
each.year.ave.growth[c("1972","1989")]
```

```
## 1972 1989
## 5.63 3.19
```

```
year <- as.numeric(names(each.year.ave.growth))
year.ave.gdp <- unname(each.year.ave.growth)
plot(year, year.ave.gdp, xlab = 'Year', ylab = 'Average GDP', main = 'the average GDP growth rate for ea
```



3

```
# (a)
cat('the correlation coeffcient between GDP growth and the debt ratio over the whole data set', signif(
```

```
## the correlation coeffcient between GDP growth and the debt ratio over the whole data set -0.1995
```

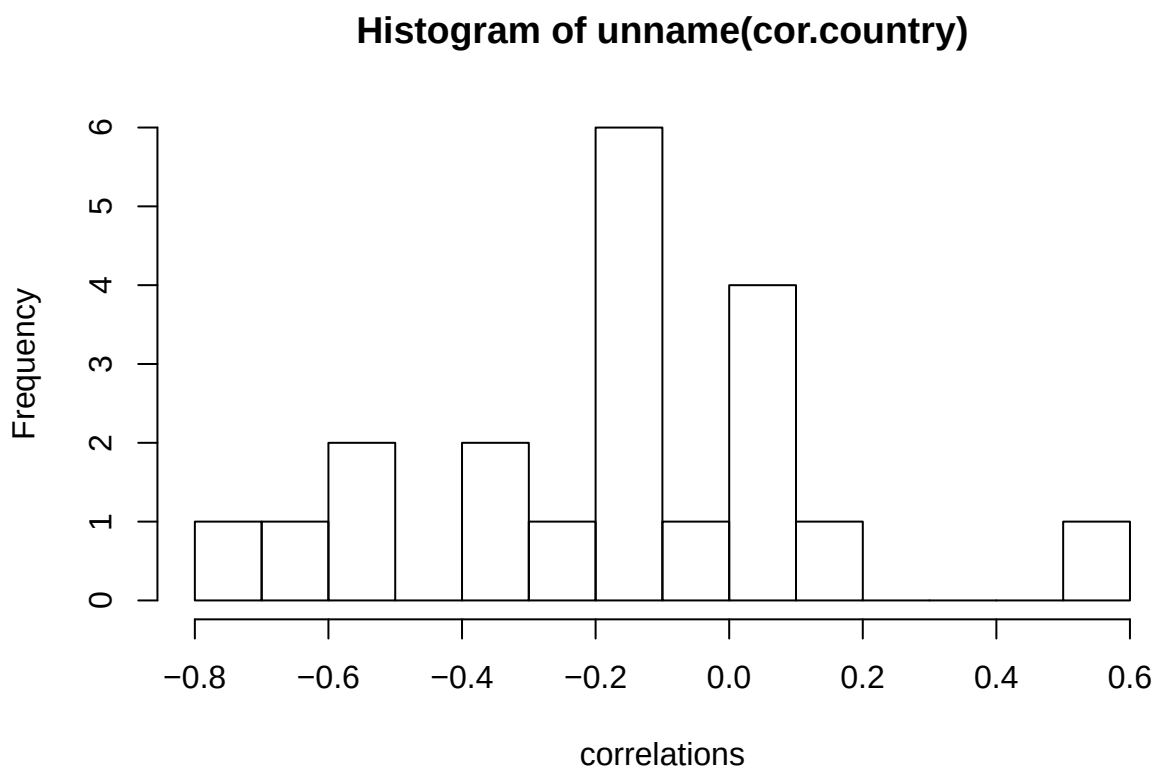
```r
# (b)
debt.gdp.cor <- function (data) {
  return(cor(data$growth, data$ratio))
}
cor.country <- daply(debt, .(Country), debt.gdp.cor)
signif(cor.country, 3)
```

```
##   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
```

```r
cat('The mean of correlations (country):', signif(mean(cor.country),4), '3 digits:', signif(mean(cor.co
```

```
## The mean of correlations (country): -0.1778 3 digits: -0.178
```

```r
hist(unname(cor.country), breaks = 10, xlab = 'correlations')
```




```r
# (c)
cor.year <- daply(debt, .(Year), debt.gdp.cor)
signif(cor.year, 3)
```

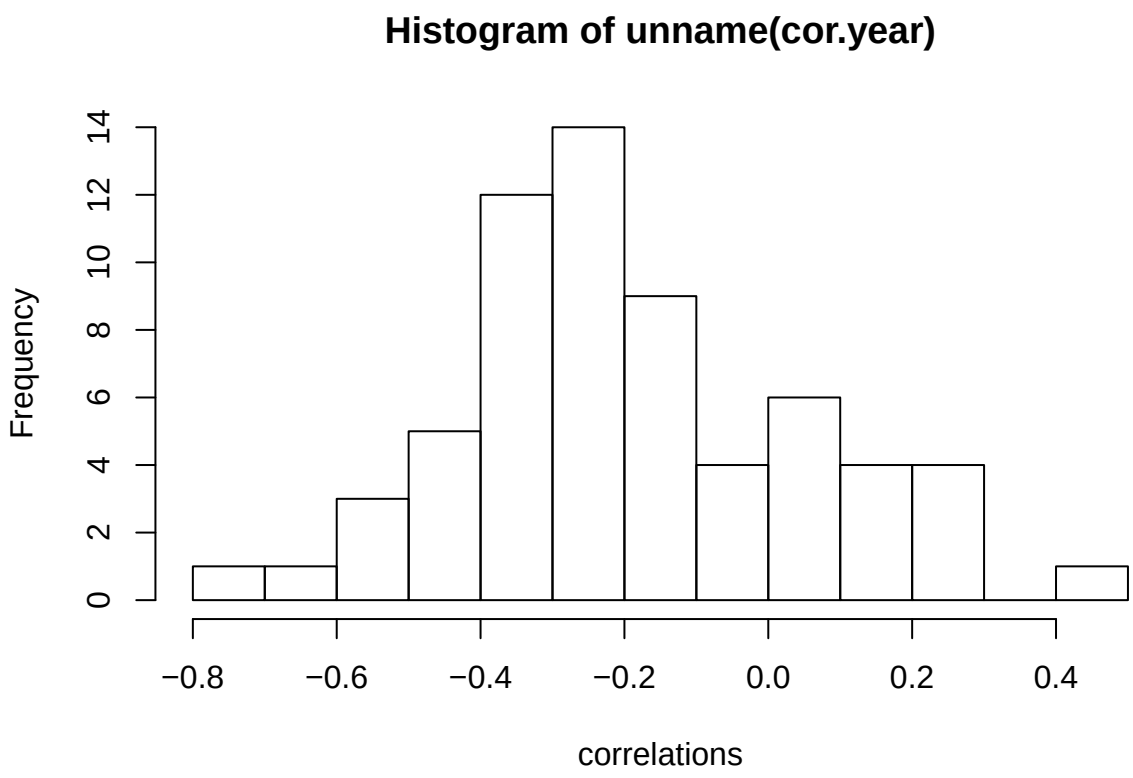
```
##     1946     1947     1948     1949     1950     1951     1952     1953
## -0.62000 -0.27400 -0.34000 -0.20000  0.03980 -0.41600 -0.27700 -0.20500
```

```
##     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
```

```
cat('The mean of correlations (year):', signif(mean(cor.year),4), '3 digits:', signif(mean(cor.year),3))
```

```
## The mean of correlations (year): -0.1906 3 digits: -0.191
```

```
hist(unname(cor.year), breaks = 10, xlab = 'correlations')
```




```
# (d)
cat('Country: Norway has very high positive corr, Jpan and Germany have high negative corr')
```

```
## Country: Norway has very high positive corr, Jpan and Germany have high negative corr
```

```
cat('Year: 1946 and 1957 have high negative corr, 1978 has high positive corr')
```

```
## Year: 1946 and 1957 have high negative corr, 1978 has high positive corr
```

4

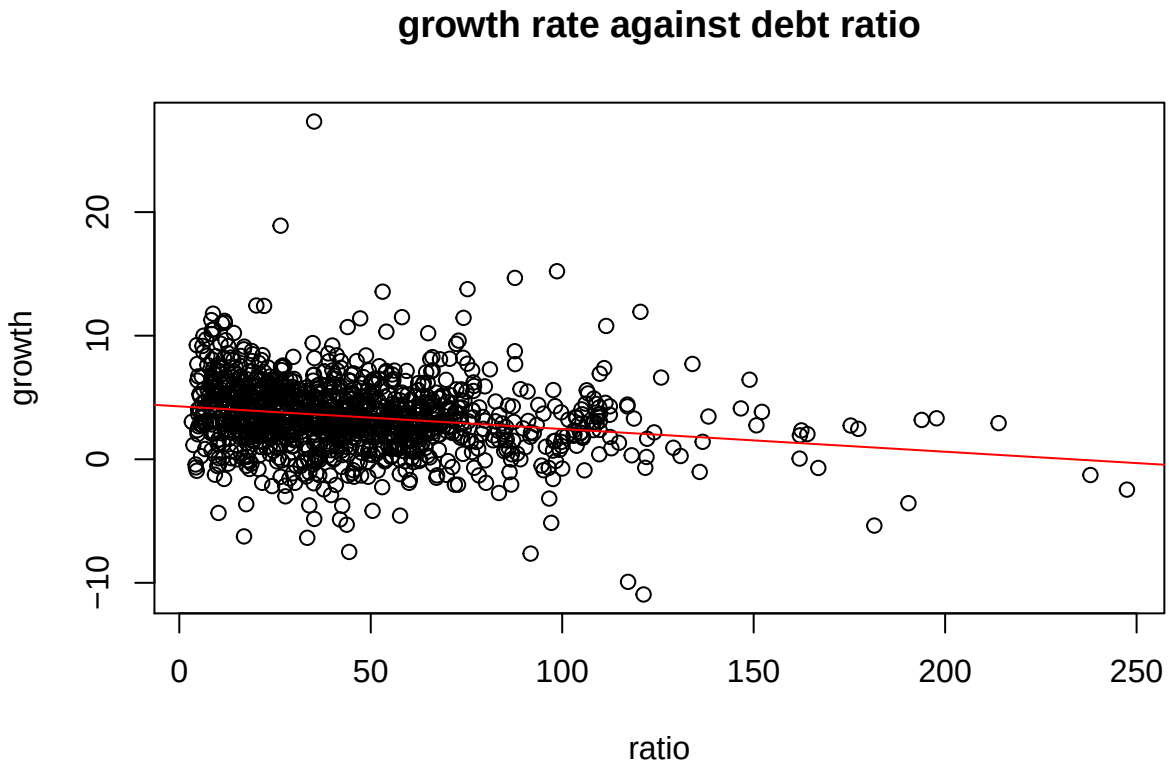
```
lm0 <- lm(debt$growth ~ debt$ratio, data = debt)
cat('slope is ', signif(lm0$coefficients[2], 3))
```

```
## slope is  -0.0184
```

```
cat('intercept is ', signif(lm0$coefficients[1], 3))
```

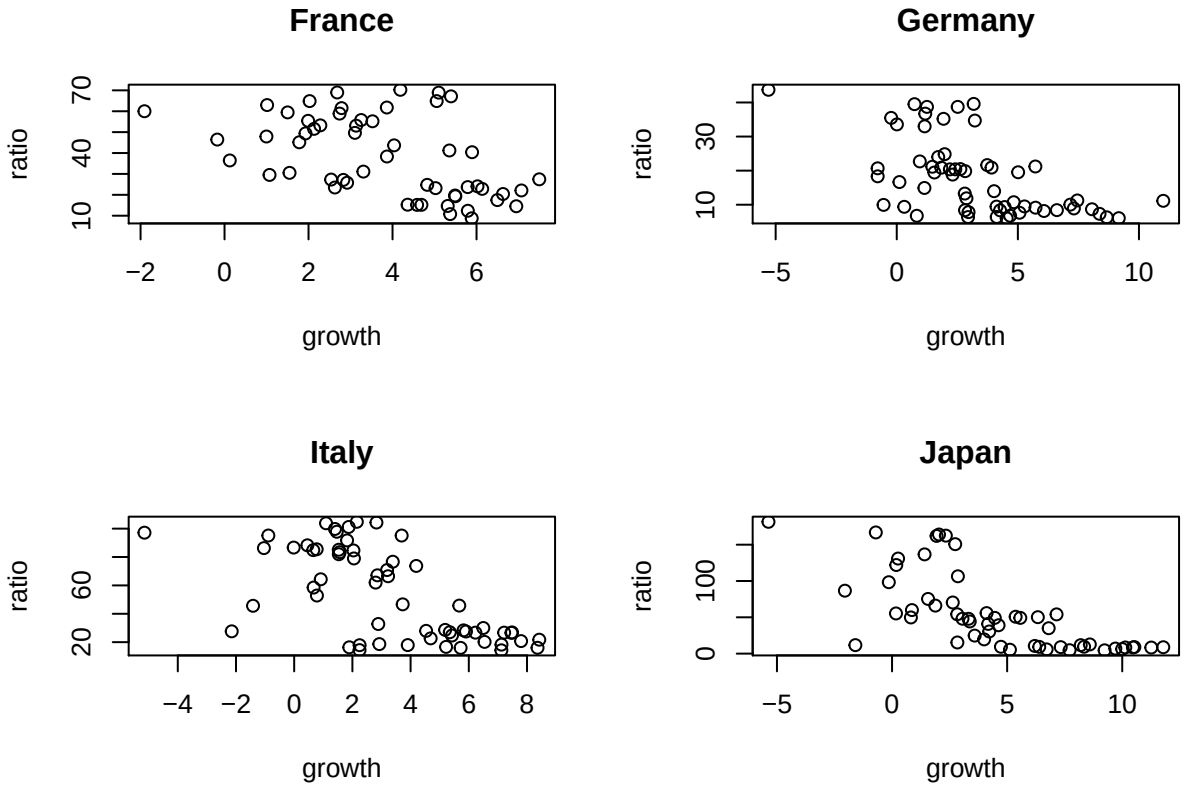
```
## intercept is  4.28
```

```
plot(debt$ratio, debt$growth, xlab = 'ratio', ylab = 'growth', main = 'growth rate against debt ratio')
abline(lm0, col = 'red')
```



5

```
four.countries <- names(cor.country[cor.country < -0.5])
country.plot <- function (data){
  plot(data$growth, data$ratio, main = unique(data$Country), xlab = 'growth', ylab = 'ratio')
  return(NA)
}
par(mfrow = c(2,2))
daply(debt[debt$Country %in% four.countries, ], .(Country), country.plot)
```




```
##  France Germany   Italy   Japan
##      NA      NA      NA      NA
```

6

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

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

```
# (b)
France$judge <- ifelse((c(France$Year[2:nrow(France)] - 1, NA) == c(France$Year)) == TRUE, 0, NA)
France$next.growth <- c(France$growth[2:nrow(France)], NA) + France$judge
France <- subset(France, select = -judge)
cat('Year 1971:', signif(France[France$Year == 1971,]$next.growth, 3))
```

```
## Year 1971: 5.89
```

```
cat('Year 1972:', France[France$Year == 1972,]$next.growth)
```

```
## Year 1972: NA
```

7

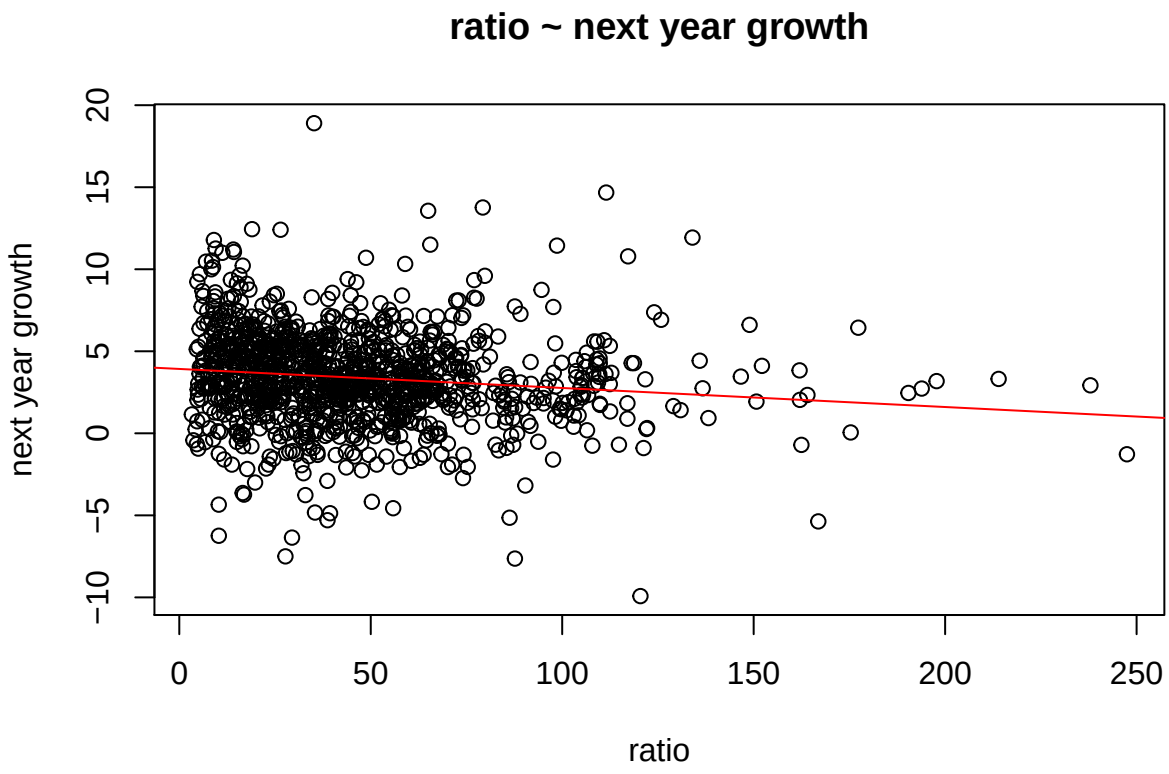
```
get.next.growth <- function (data) {
  data$judge <- ifelse((c(data$Year[2:nrow(data)] - 1, NA) == c(data$Year)) == TRUE, 0, NA)
  data$next.growth <- c(data$growth[2:nrow(data)], NA) + data$judge
```

```
  data <- subset(data, select = -judge)
  return(data)
}
data.list <- ddply(debt, .(Country), get.next.growth)
cat('France Year 2009:', data.list[(data.list$Country == 'France')&(data.list$Year == 2009), ]$next.grow
```

```
## France Year 2009: NA
```

8

```
plot(data.list$ratio, data.list$next.growth, xlab = 'ratio', ylab = 'next year growth', main = 'ratio ~
lm1 <- lm(data.list$next.growth ~ data.list$ratio)
abline(lm1, col = 'red')
```




```
cat('Intercept is:',signif(lm1$coefficients[1], 3))
```

```
## Intercept is: 3.92
```

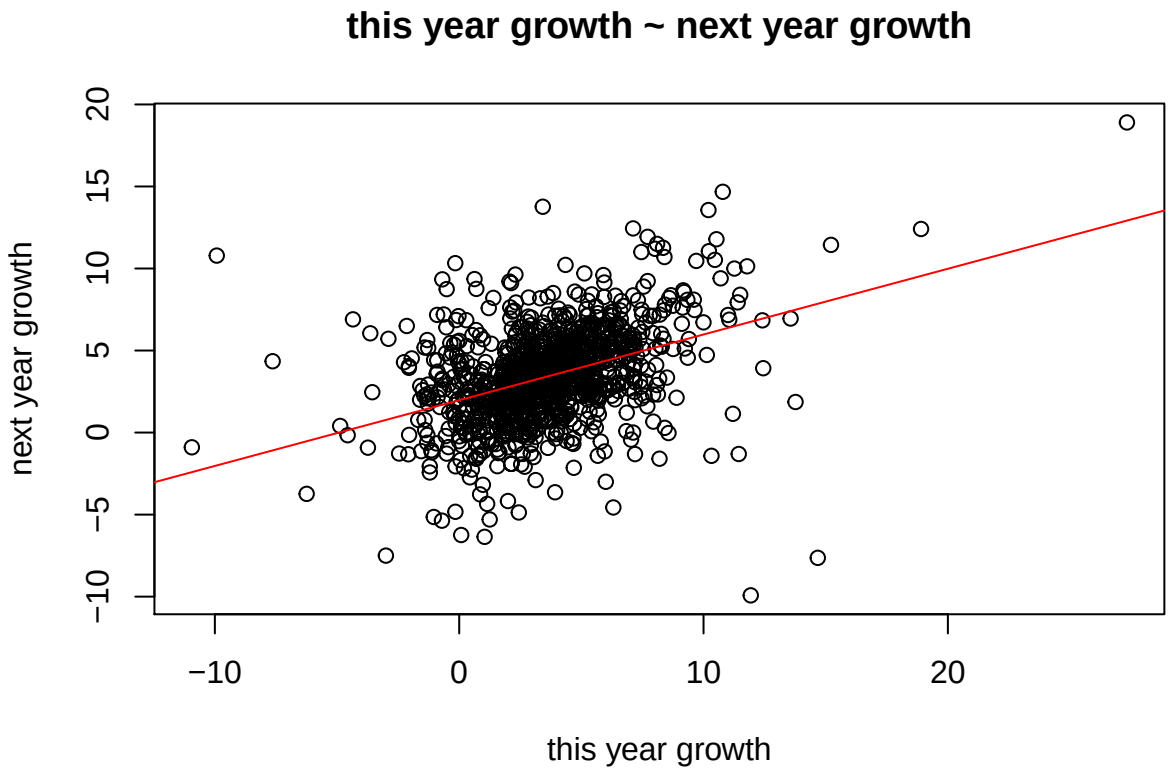
```
cat('Slope is:',signif(lm1$coefficients[2], 3))
```

```
## Slope is: -0.0116
```

the slope becomes closer to zero, which means the the next year GDP growth is less related to the current debt ratio. The intercept becomes samller, which means the next year GDP growth is likely to be lower than this year if debt ratio is 0

9

```
plot(data.list$growth, data.list$next.growth, xlab = 'this year growth', ylab = 'next year growth', mai
lm2 <- lm(data.list$next.growth ~ data.list$growth)
abline(lm2, col = 'red')
```



Answer: current growth is a better predictor of future growth since it has a more significant slope.