

EX1markdown

Zezhun Chen

10/13/2021

R Markdown

This is an R Markdown document. Markdown is a simple formatting syntax for authoring HTML, PDF, and MS Word documents. For more details on using R Markdown see <http://rmarkdown.rstudio.com>.

When you click the **Knit** button a document will be generated that includes both content as well as the output of any embedded R code chunks within the document. You can embed an R code chunk like this:

```
summary(cars)
```

```
##      speed      dist
##  Min.   : 4.0    Min.   :  2.00
##  1st Qu.:12.0    1st Qu.: 26.00
##  Median :15.0    Median : 36.00
##  Mean   :15.4    Mean   : 42.98
##  3rd Qu.:19.0    3rd Qu.: 56.00
##  Max.   :25.0    Max.   :120.00
```

Including Plots

You can also embed plots, for example:



Note that the `echo = FALSE` parameter was added to the code chunk to prevent printing of the R code that generated the plot.

Setting up

Remove all the objects in the environment read the data skip, the number of lines of the data file to skip before beginning to read data. `row.names`: giving the column of the table which contains the row names,

```
rm(list = ls())
forbes2000 = read.csv("Forbes2000.csv", row.names = 1, skip=3)
head(forbes2000)
```

##	rank	name	country	category	sales	profits
## 1	1	Citigroup	United States	Banking	94.71	17.85
## 2	2	General Electric	United States	Conglomerates	134.19	15.59
## 3	3	American Intl Group	United States	Insurance	76.66	6.46
## 4	4	ExxonMobil	United States	Oil & gas operations	222.88	20.96
## 5	5	BP	United Kingdom	Oil & gas operations	232.57	10.27
## 6	6	Bank of America	United States	Banking	49.01	10.81

##	assets	marketvalue
## 1	1264.03	255.30
## 2	626.93	328.54
## 3	647.66	194.87
## 4	166.99	277.02
## 5	177.57	173.54
## 6	736.45	117.55

(a)

Find the top 10 ranked companies, the companies with the 10 maximum sales, and the companies with the market value greater than 174 billion USD

```
attach(forbes2000)
```

```
name[rank <= 10]
```

```
## [1] "Citigroup"          "General Electric"  "American Intl Group"
## [4] "ExxonMobil"         "BP"               "Bank of America"
## [7] "HSBC Group"         "Toyota Motor"     "Fannie Mae"
## [10] "Wal-Mart Stores"
```

```
name[marketvalue >= 174]
```

```
## [1] "Citigroup"          "General Electric"  "American Intl Group"
## [4] "ExxonMobil"         "HSBC Group"       "Wal-Mart Stores"
## [7] "Pfizer"            "Microsoft"        "Intel"
## [10] "Vodafone"
```

```
name[sales > sort(sales,decreasing = TRUE)[10]]
```

```
## [1] "General Electric"  "ExxonMobil"
## [3] "BP"               "Toyota Motor"
## [5] "Wal-Mart Stores"  "Royal Dutch/Shell Group"
## [7] "DaimlerChrysler"  "General Motors"
## [9] "Ford Motor"
```

(b)

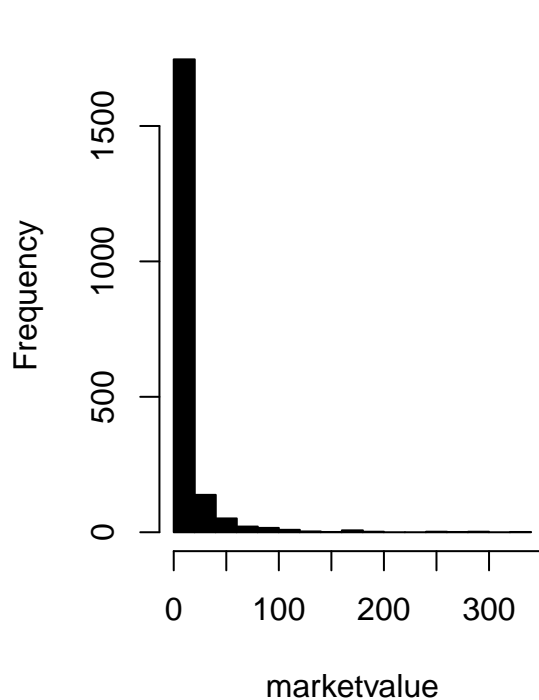
Plot the histograms for marketvalue and log(marketvalue)

```
par(mfrow=c(1,2))
```

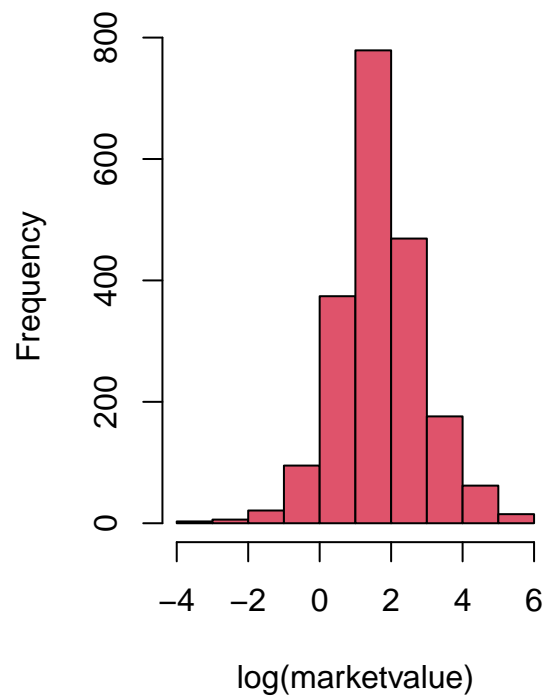
```
hist(marketvalue,main='histogram of market values',col=1)
```

```
hist(log(marketvalue),main='histogram of log market values',col=2)
```

histogram of market values



histogram of log market values



(c)

Compare the outcomes from `mean(profits)` and `mean(profits, na.rm=T)`

```
mean(profits)
```

```
## [1] NA
```

```
mean(profits, na.rm = TRUE)
```

```
## [1] 0.3811328
```

(d)

Median profit for the companies in US and UK separately

```
median(profits[country=='United States'], na.rm = TRUE)
```

```
## [1] 0.24
```

```
median(profits[country=='United Kingdom'], na.rm=TRUE)
```

```
## [1] 0.205
```

(e)

Find all German companies with negative profit

```
name[country == 'Germany' & profits < 0]
```

```
## [1] "Allianz Worldwide"      "Deutsche Telekom"
## [3] "E.ON"                  "HVB-HypoVereinsbank"
## [5] "Commerzbank"           "Infineon Technologies"
## [7] "BHW Holding"           "Bankgesellschaft Berlin"
## [9] "W&W-Wustenrot"         "mg technologies"
## [11] "Nurnberger Beteiligungs" "SPAR Handels"
## [13] "Mobilcom"
```

(f)

To which business category do most of the Bermuda island companies belong

```
table(category[country == 'Bermuda'])
```

```
##
##           Banking           Capital goods           Conglomerates
##           1             1             2
## Food drink & tobacco           Food markets           Insurance
##           1             1             10
##           Media Oil & gas operations Software & services
##           1             2             1
```

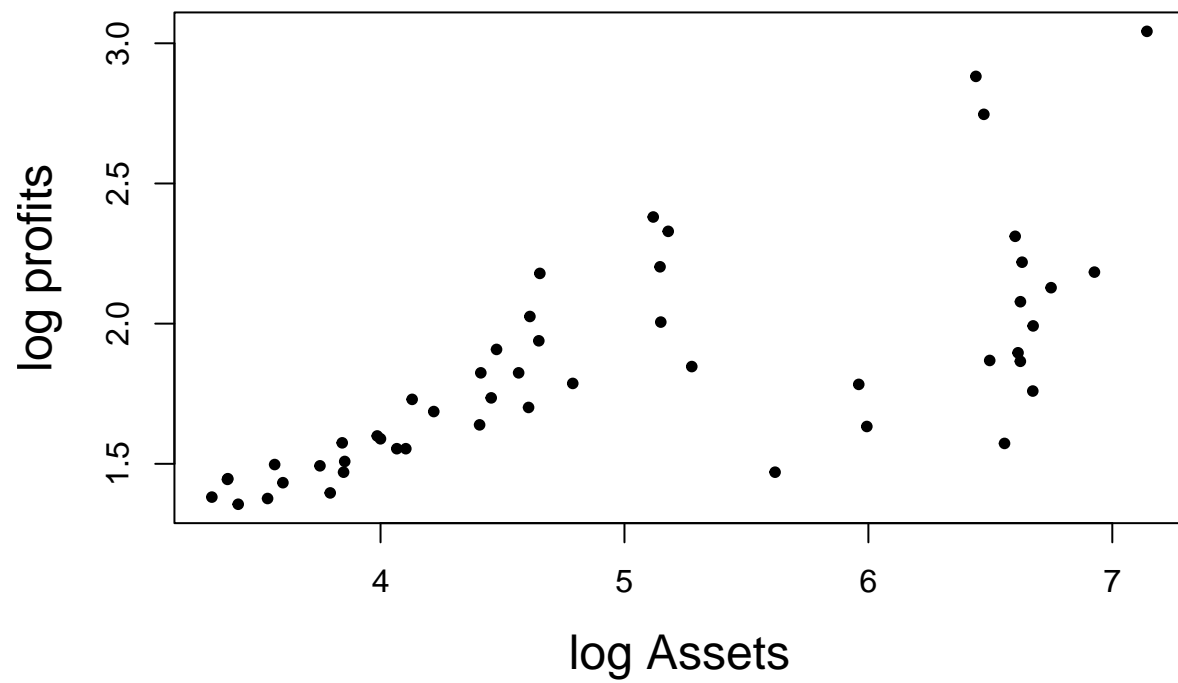
(g)

For the 50 companies with the highest profits, plot profits against assets, using some suitable transformation for each variable if appropriate.

```
profits_sort <- sort(profits, decreasing = TRUE)[1:50]
asset_sort <- assets[profits >= profits_sort[50] & !is.na(profits)]
```

cex the size of point cex.lab the size of label pch the style of points

```
plot(log(asset_sort), log(profits_sort),
     xlab='log Assets', ylab = 'log profits',
     cex.lab=1.5, pch=20)
```



remember to detach when you finish analysis with the data set

```
detach()
```