Assignment1

By fedesoriano in Kaggle

12/09/2021

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 (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:

```
chooseCRANmirror(graphics = getOption("menu.graphics"), ind = 79,
                 local.only = FALSE)
install.packages("vcd")
```

```
## Installing package into 'C:/Users/Srija/Documents/R/win-library/4.1'
## (as 'lib' is unspecified)
```

```
## package 'vcd' successfully unpacked and MD5 sums checked
##
## The downloaded binary packages are in
##  C:\Users\Srija\AppData\Local\Temp\RtmpgRI0H5\downloaded_packages
```

```
airlines <- read.csv("USAirlines.csv")




#summary of the dataset US Airlines

summary(airlines)
```

```
##        X              firm          year          output
##  Min.   : 1.00   Min.   :1.0   Min.   :1970   Min.   :0.03768
##  1st Qu.:23.25   1st Qu.:2.0   1st Qu.:1973   1st Qu.:0.14213
##  Median :45.50   Median :3.5   Median :1977   Median :0.30503
##  Mean   :45.50   Mean   :3.5   Mean   :1977   Mean   :0.54499
##  3rd Qu.:67.75   3rd Qu.:5.0   3rd Qu.:1981   3rd Qu.:0.94528
##  Max.   :90.00   Max.   :6.0   Max.   :1984   Max.   :1.93646
##       cost             price             load
##  Min.   :  68978   Min.   : 103795   Min.   :0.4321
##  1st Qu.: 292046   1st Qu.: 129848   1st Qu.:0.5288
##  Median : 637001   Median : 357434   Median :0.5661
##  Mean   :1122524   Mean   : 471683   Mean   :0.5605
##  3rd Qu.:1345968   3rd Qu.: 849840   3rd Qu.:0.5947
##  Max.   :4748320   Max.   :1015610   Max.   :0.6763
```

```
# mean of the firm from the dataset Airlines
mean(airlines$firm)
```

```
## [1] 3.5
```

```
#median of the firm from the dataset Airlines
median(airlines$firm)
```

```
## [1] 3.5
```

```
#mode of the firm from the dataset Airlines
mode(airlines$firm)
```

```
## [1] "numeric"
```

```
#install the package dplyr
install.packages("dplyr")
```

```
## Installing package into 'C:/Users/Srija/Documents/R/win-library/4.1'
## (as 'lib' is unspecified)
```

```
## package 'dplyr' successfully unpacked and MD5 sums checked
```

```
## Warning: cannot remove prior installation of package 'dplyr'
```

```
## Warning in file.copy(savedcopy, lib, recursive = TRUE): problem copying C:
## \Users\Srija\Documents\R\win-library\4.1\00LOCK\dplyr\libs\x64\dplyr.dll to C:
## \Users\Srija\Documents\R\win-library\4.1\dplyr\libs\x64\dplyr.dll: Permission
## denied
```

```
## Warning: restored 'dplyr'
```

```
##
## The downloaded binary packages are in
##  C:\Users\Srija\AppData\Local\Temp\RtmpgRI0H5\downloaded_packages
```

```
library(dplyr)
```

```
##
## Attaching package: 'dplyr'
```

```
## The following objects are masked from 'package:stats':
##
##     filter, lag
```

```
## The following objects are masked from 'package:base':
##
##     intersect, setdiff, setequal, union
```

```
select(airlines, firm : cost)
```

```
##    firm year   output    cost
## 1     1 1970 0.952757 1140640
## 2     1 1971 0.986757 1215690
## 3     1 1972 1.091980 1309570
## 4     1 1973 1.175780 1511530
## 5     1 1974 1.160170 1676730
## 6     1 1975 1.173760 1823740
## 7     1 1976 1.290510 2022890
## 8     1 1977 1.390670 2314760
## 9     1 1978 1.612730 2639160
## 10    1 1979 1.825440 3247620
## 11    1 1980 1.546040 3787750
## 12    1 1981 1.527900 3867750
## 13    1 1982 1.660200 3996020
## 14    1 1983 1.822310 4282880
## 15    1 1984 1.936460 4748320
## 16    2 1970 0.520635  569292
## 17    2 1971 0.534627  640614
## 18    2 1972 0.655192  777655
## 19    2 1973 0.791575  999294
## 20    2 1974 0.842945 1203970
## 21    2 1975 0.852892 1358100
## 22    2 1976 0.922843 1501350
## 23    2 1977 1.000000 1709270
## 24    2 1978 1.198450 2025400
## 25    2 1979 1.340670 2548370
## 26    2 1980 1.326240 3137740
## 27    2 1981 1.248520 3557700
## 28    2 1982 1.254320 3717740
## 29    2 1983 1.371770 3962370
## 30    2 1984 1.389740 4209390
## 31    3 1970 0.262424  286298
## 32    3 1971 0.266433  309290
## 33    3 1972 0.306043  342056
## 34    3 1973 0.325586  374595
## 35    3 1974 0.345706  450037
## 36    3 1975 0.367517  510412
## 37    3 1976 0.409937  575347
## 38    3 1977 0.448023  669331
## 39    3 1978 0.539595  783799
## 40    3 1979 0.539382  913883
## 41    3 1980 0.467967 1041520
## 42    3 1981 0.450544 1125800
## 43    3 1982 0.468793 1096070
## 44    3 1983 0.494397 1198930
## 45    3 1984 0.493317 1170470
## 46    4 1970 0.086393  145167
## 47    4 1971 0.096740  170192
## 48    4 1972 0.141500  247506
## 49    4 1973 0.169715  309391
## 50    4 1974 0.173805  354338
## 51    4 1975 0.164272  373941
## 52    4 1976 0.170906  420915
```

```
## 53    4 1977 0.177840  474017
## 54    4 1978 0.192248  532590
## 55    4 1979 0.242469  676771
## 56    4 1980 0.256505  880438
## 57    4 1981 0.249657 1052020
## 58    4 1982 0.273923 1193680
## 59    4 1983 0.371131 1303390
## 60    4 1984 0.421411 1436970
## 61    5 1970 0.051028   91361
## 62    5 1971 0.052646   95428
## 63    5 1972 0.056348   98187
## 64    5 1973 0.066953  115967
## 65    5 1974 0.070308  138382
## 66    5 1975 0.073961  156228
## 67    5 1976 0.084946  183169
## 68    5 1977 0.095474  210212
## 69    5 1978 0.119814  274024
## 70    5 1979 0.150046  356915
## 71    5 1980 0.144014  432344
## 72    5 1981 0.169300  524294
## 73    5 1982 0.172761  530924
## 74    5 1983 0.186670  581447
## 75    5 1984 0.213279  610257
## 76    6 1970 0.037682   68978
## 77    6 1971 0.039784   74904
## 78    6 1972 0.044331   83829
## 79    6 1973 0.050245   98148
## 80    6 1974 0.055046  118449
## 81    6 1975 0.052462  133161
## 82    6 1976 0.056977  145062
## 83    6 1977 0.061490  170711
## 84    6 1978 0.069027  199775
## 85    6 1979 0.092749  276797
## 86    6 1980 0.112640  381478
## 87    6 1981 0.154154  506969
## 88    6 1982 0.186461  633388
## 89    6 1983 0.246847  804388
## 90    6 1984 0.304013 1009500
```

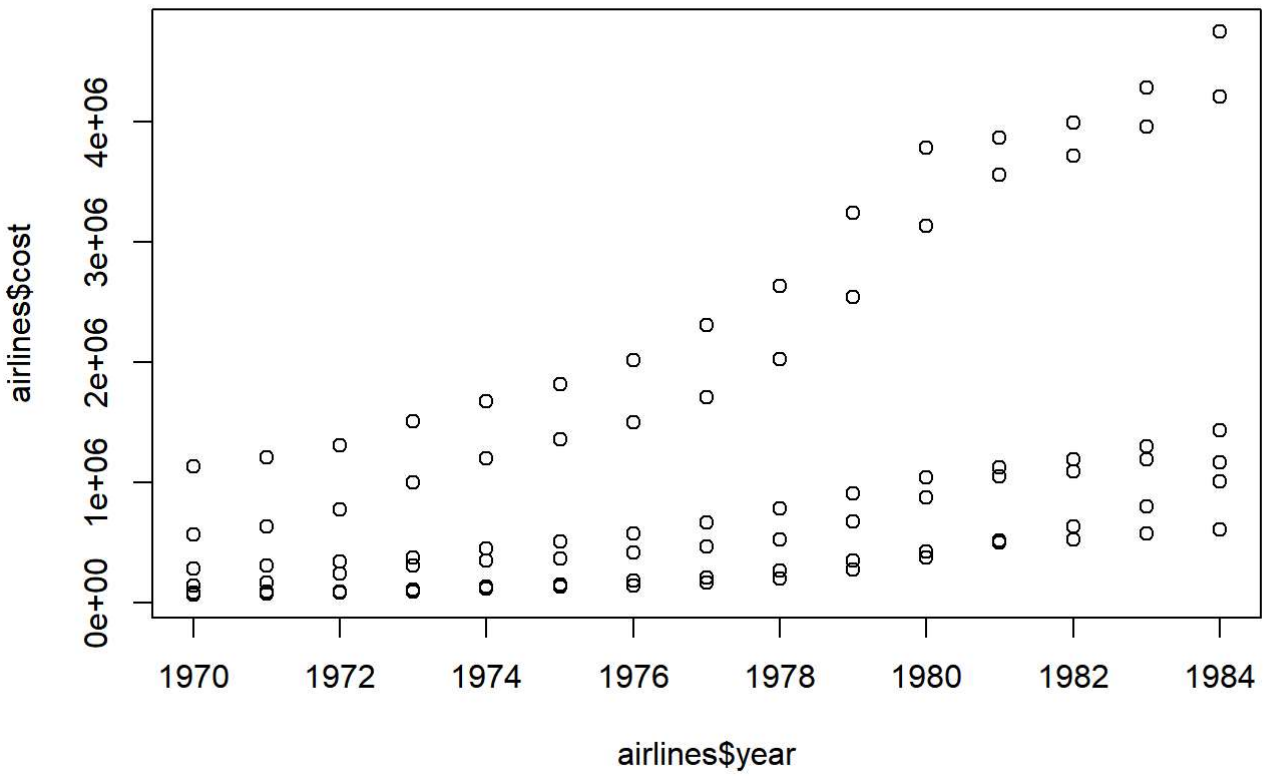
```
filter(airlines, firm ==1)
```

```
##     X firm year   output    cost  price     load
## 1   1    1 1970 0.952757 1140640 106650 0.534487
## 2   2    1 1971 0.986757 1215690 110307 0.532328
## 3   3    1 1972 1.091980 1309570 110574 0.547736
## 4   4    1 1973 1.175780 1511530 121974 0.540846
## 5   5    1 1974 1.160170 1676730 196606 0.591167
## 6   6    1 1975 1.173760 1823740 265609 0.575417
## 7   7    1 1976 1.290510 2022890 263451 0.594495
## 8   8    1 1977 1.390670 2314760 316411 0.597409
## 9   9    1 1978 1.612730 2639160 384110 0.638522
## 10 10    1 1979 1.825440 3247620 569251 0.676287
## 11 11    1 1980 1.546040 3787750 871636 0.605735
## 12 12    1 1981 1.527900 3867750 997239 0.614360
## 13 13    1 1982 1.660200 3996020 938002 0.633366
## 14 14    1 1983 1.822310 4282880 859572 0.650117
## 15 15    1 1984 1.936460 4748320 823411 0.625603
```

```
summary(airlines)
```

```
##        X              firm          year          output
##  Min.   : 1.00   Min.   :1.0   Min.   :1970   Min.   :0.03768
##  1st Qu.:23.25   1st Qu.:2.0   1st Qu.:1973   1st Qu.:0.14213
##  Median :45.50   Median :3.5   Median :1977   Median :0.30503
##  Mean   :45.50   Mean   :3.5   Mean   :1977   Mean   :0.54499
##  3rd Qu.:67.75   3rd Qu.:5.0   3rd Qu.:1981   3rd Qu.:0.94528
##  Max.   :90.00   Max.   :6.0   Max.   :1984   Max.   :1.93646
##       cost             price              load
##  Min.   :  68978   Min.   : 103795   Min.   :0.4321
##  1st Qu.: 292046   1st Qu.: 129848   1st Qu.:0.5288
##  Median : 637001   Median : 357434   Median :0.5661
##  Mean   :1122524   Mean   : 471683   Mean   :0.5605
##  3rd Qu.:1345968   3rd Qu.: 849840   3rd Qu.:0.5947
##  Max.   :4748320   Max.   :1015610   Max.   :0.6763
```

```
plot(airlines$year, airlines$cost)
```



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