Tutorials Of R Using Titanic Data

Josemari Feliciano

7/10/2017

Tutorial Notes:

The Beauty of R Studio is the ability to have notebooks for data analysis. It allows us to run r-code inline. Think of it as a fancy chemistry notebook where you can run the the experiments in the notebook itself – but for data analysis! For this notebook, I will use the titanic data for analysis. This assumes that you have R and RStudio loaded.

Slight Detour:

Note : mosaic package is required for this if you want to run data yourself and not to simply view them. If you do not have this, run this code in your r-console independently:

```
install.packages("ggplot2")
```

Getting started:

We probably want to load the titanic data first. We probably want to load the mosaic library out of the way. So we can accomplish both with the following script:

```
load("Titanic.Rdata")
library("mosaic")
```

Now that the data has been loaded, we probably want to see which data variables we will deal with! Below, we will use `names()` to print the variable names in our data to have them handy.

```
names(Titanic)
```

```
## [1] "Gender"   "Age"      "Name"     "Fare"     "Class"    "Survived"
```

The output above clearly shows we have:$\ \ $

```
summary(Titanic)
```

```
##     Gender          Age                                        Name
##  Female:388   Min.   : 0.00   Connolly, Miss. Kate              :   2
##  Male  :657   1st Qu.:21.00   Kelly, Mr. James                  :   2
##               Median :28.00   Abbing, Mr. Anthony               :   1
##               Mean   :29.84   Abbott, Master. Eugene Joseph     :   1
##               3rd Qu.:39.00   Abbott, Mr. Rossmore Edward       :   1
##               Max.   :80.00   Abbott, Mrs. Stanton (Rosa Hunt):   1
##                               (Other)                           :1037
##       Fare           Class     Survived
##  Min.   :  0.00   Lower :500   No :618
##  1st Qu.:  8.05   Middle:261   Yes:427
##  Median : 15.75   Upper :284
##  Mean   : 36.69
##  3rd Qu.: 35.50
##  Max.   :512.33
```

##

```
bargraph( ~ Class, data =Titanic)
```