Chapter 2: Getting Started

Ryan McWay

## **ECON 427- Applied Econometrics Capstone**

## *Packages*

Installing the read excel package

# install.packages('readxl')

Or, if you already have it, you can update the package to make sure you have all the cool new tricks installed

# update.packages('readxl')

*For the purpose of knitting the log of this file (printing out the code and results), I have to comment out these past two commands (this is done with the # before the line of code). If you want to run them yourself, I suggest doing them in the console/terminal and not in the script itself. And also remember to remove the # before you run the code.*

To utlize this package, I need to call it from my library of packages

library('readxl')

## Warning: package 'readxl' was built under R version 3.5.3

library(knitr)

## Warning: package 'knitr' was built under R version 3.5.3

## *Import Data*

Import in the UKHP.xls using the readxl package, check out a sample of the data. View will show you all the data if you would like to view it. Declaring the type is only useful in terms of ensuring the data format and sometimes when trying to speed up import load time.

Rate1 <- read\_excel("~/Ryan/USF/2020-Spring/intro\_finance\_metrics/data/Chapter\_1/Excel/RATE1.xls",  
 col\_types = c('numeric', 'numeric','numeric', 'numeric'))  
head(Rate1)

## # A tibble: 6 x 4  
## OBS RATING EASE HOT  
## <dbl> <dbl> <dbl> <dbl>  
## 1 1 2.8 3.7 0  
## 2 2 4.3 4.1 1  
## 3 3 4 2.8 1  
## 4 4 3 3 0  
## 5 5 4.3 2.4 0  
## 6 6 2.7 2.7 0

## *Data Description*

Display the summary stats of this dataframe, and the variable names

summary(Rate1)

## OBS RATING EASE HOT   
## Min. : 1 Min. :1.200 Min. :1.600 Min. :0.00   
## 1st Qu.: 7 1st Qu.:2.800 1st Qu.:2.400 1st Qu.:0.00   
## Median :13 Median :3.600 Median :2.800 Median :0.00   
## Mean :13 Mean :3.416 Mean :2.836 Mean :0.28   
## 3rd Qu.:19 3rd Qu.:4.100 3rd Qu.:3.100 3rd Qu.:1.00   
## Max. :25 Max. :5.000 Max. :4.800 Max. :1.00

names(Rate1)

## [1] "OBS" "RATING" "EASE" "HOT"

## *Changing and Creating Data*

Creating a new variable

Rate1$log\_rating <- log(Rate1$RATING)

## *Graphics and Plots*

A quick graphic

par(cex.axis = 1, cex.lab = 1, lwd = 2)  
hist(Rate1$RATING)

