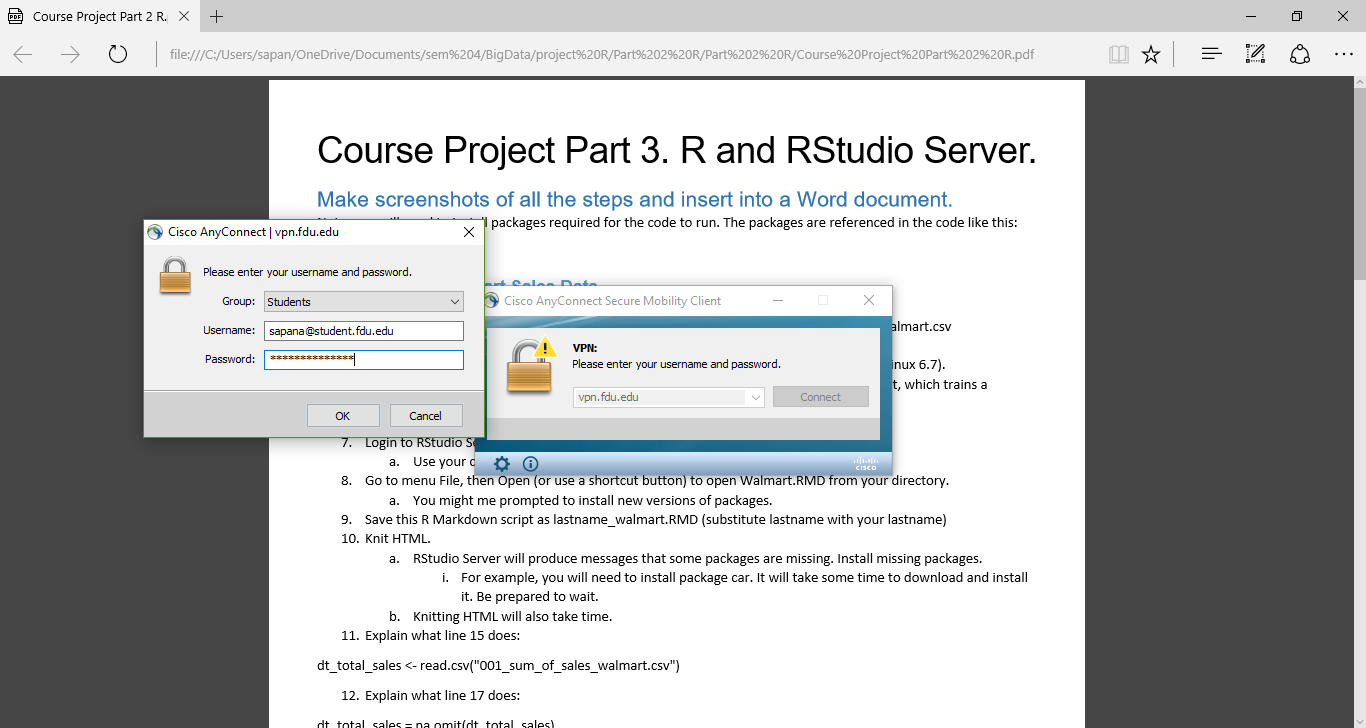
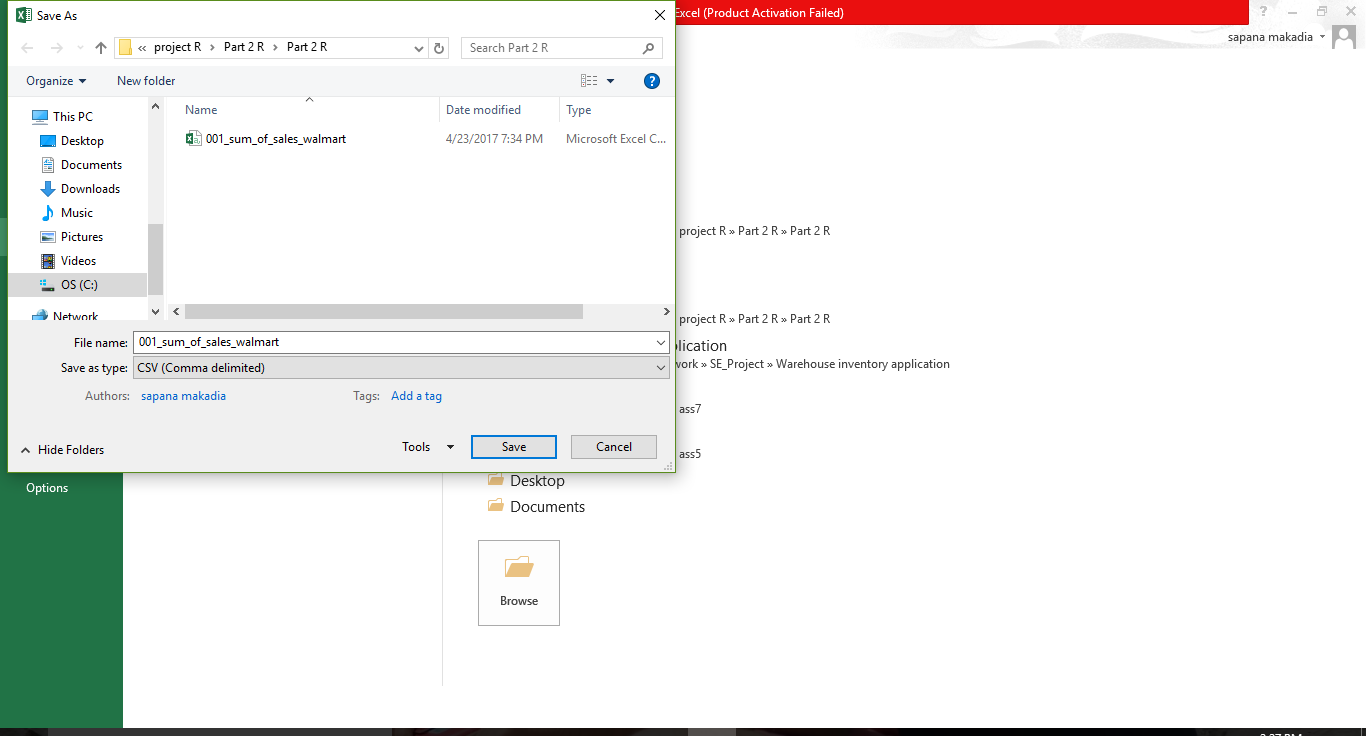
**Part-1**

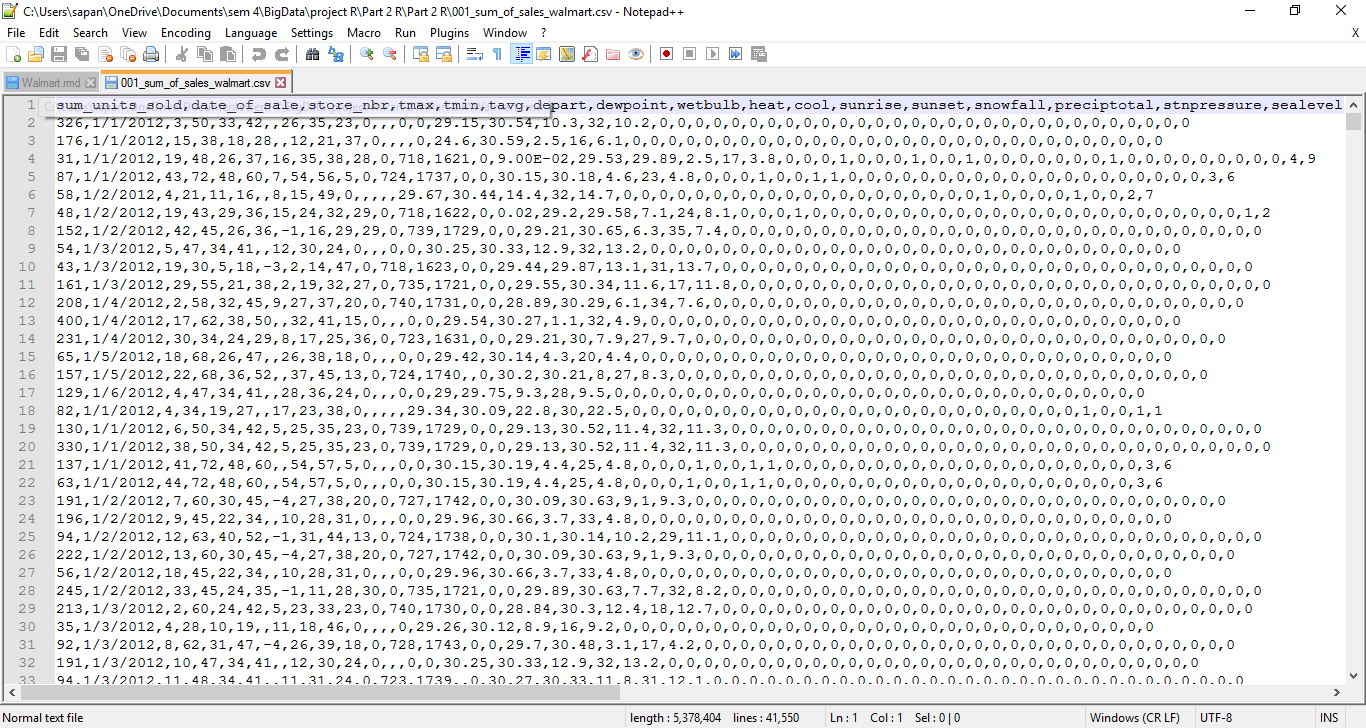
Connect to FDU VPN



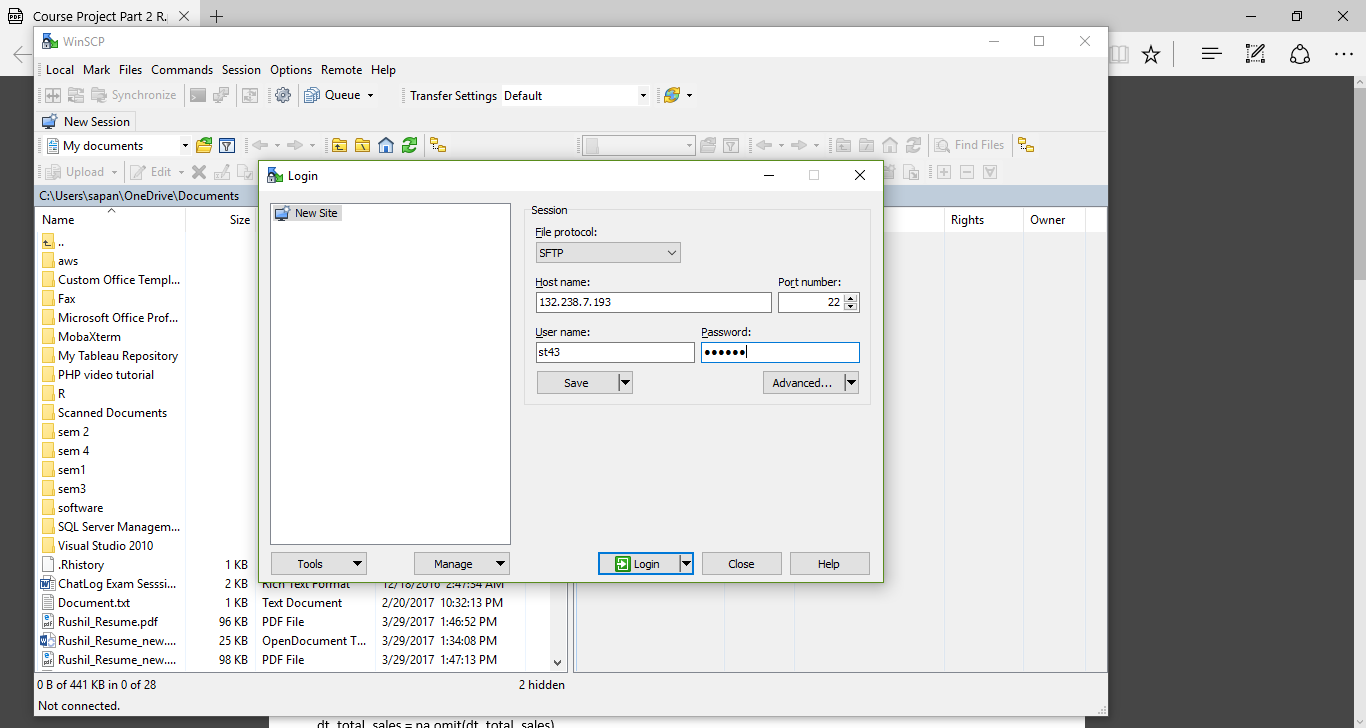
Download a dataset in CSV format from WebCampus. File name is 001\_sum\_of\_sales\_walmart.csv

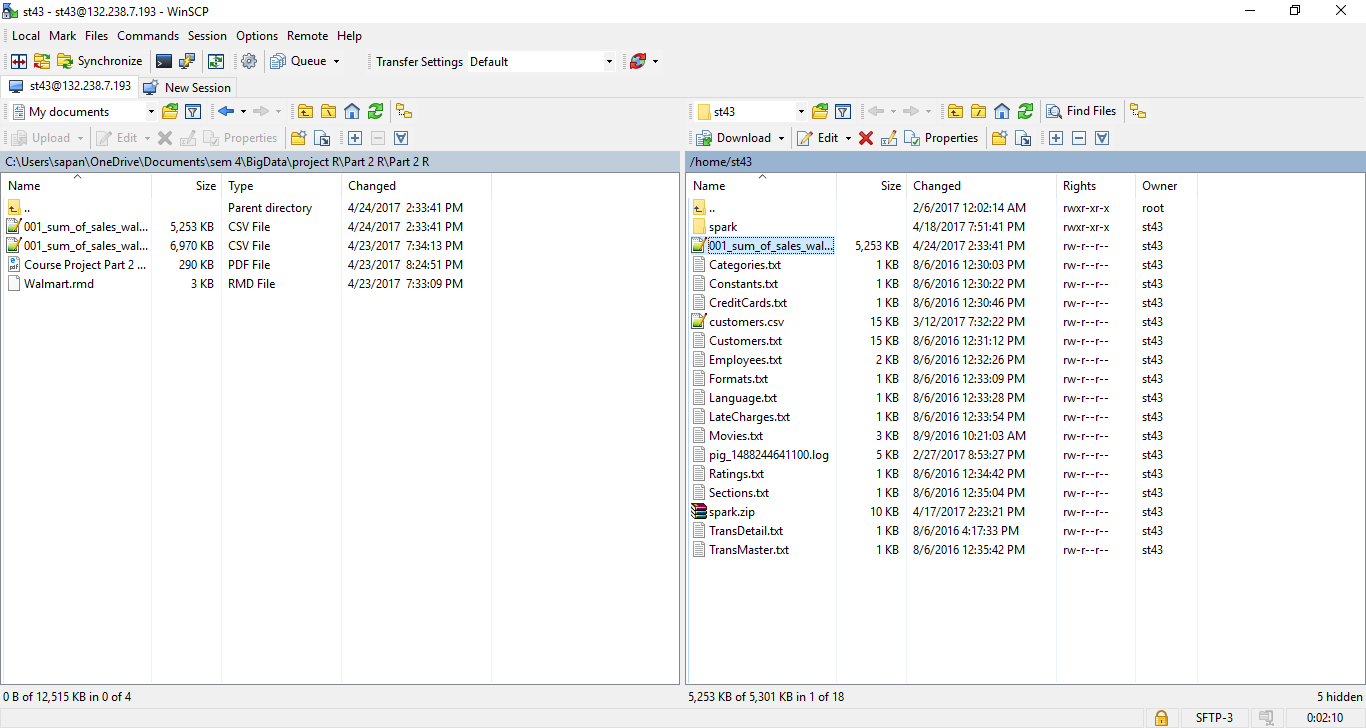


001\_sum\_of\_sales\_walmart.csv file

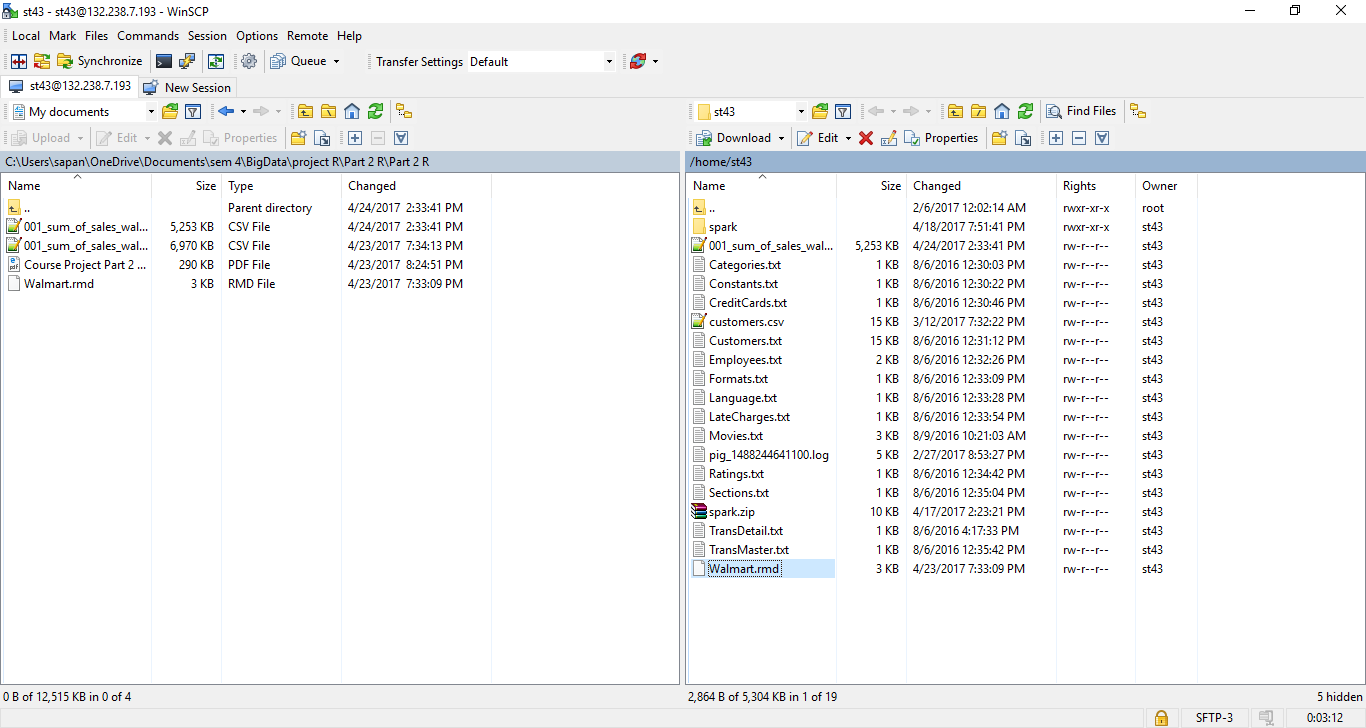


Transfer the downloaded CSV dataset to your home directory in Linux

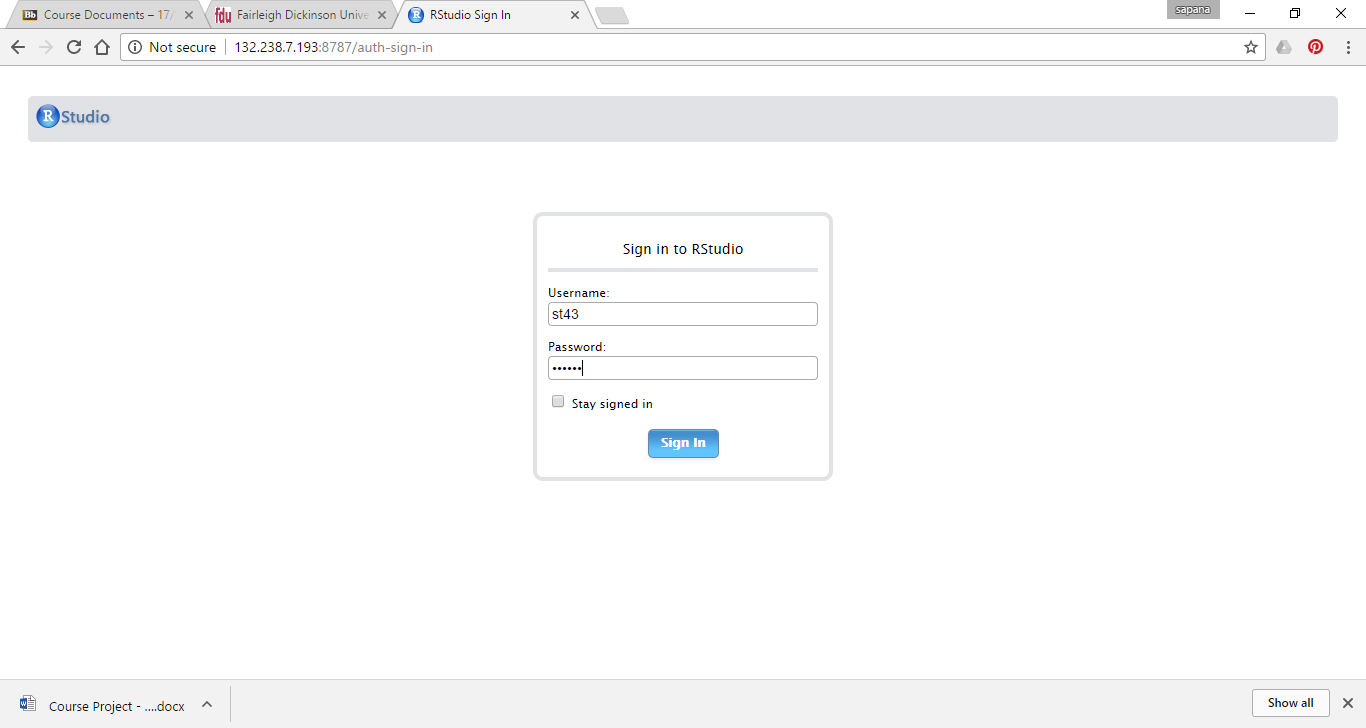




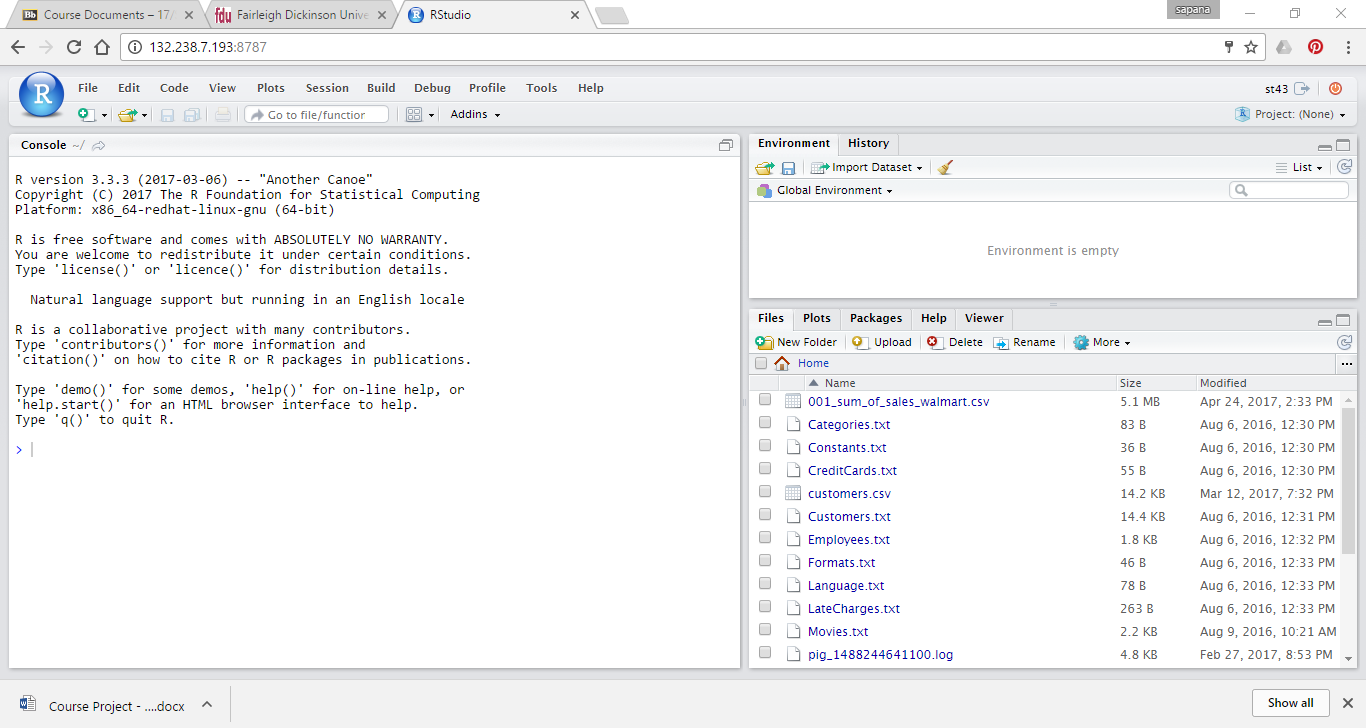
Transfer Walmart.RMD to your home directory in Linux



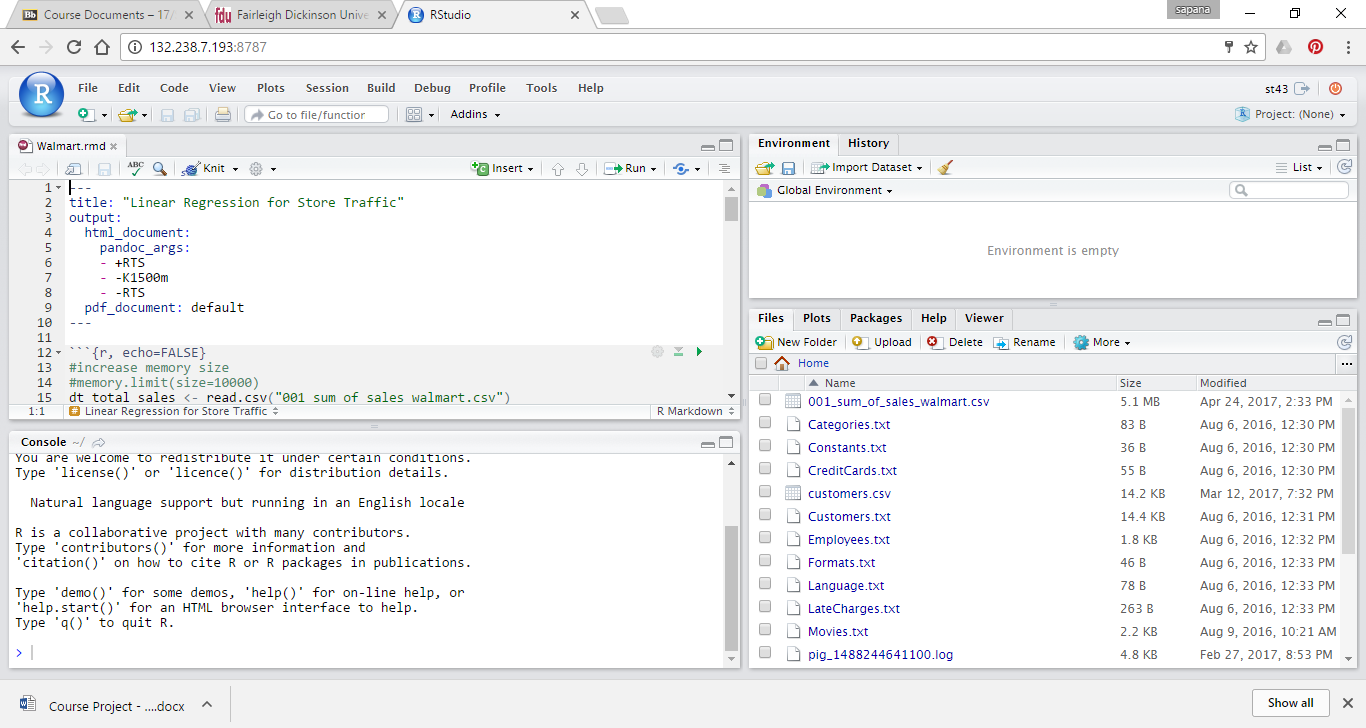
Login to RStudio Server by navigating to http://132.238.7.193:8787/



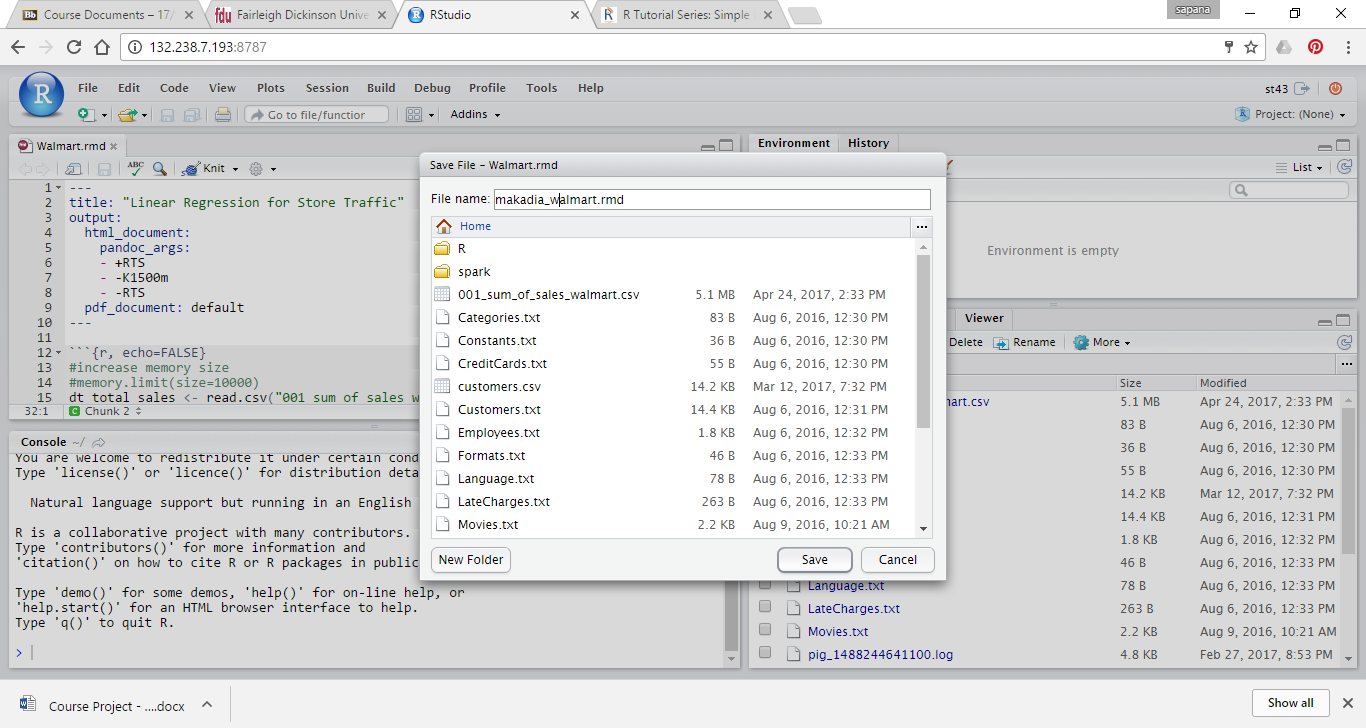
R studio



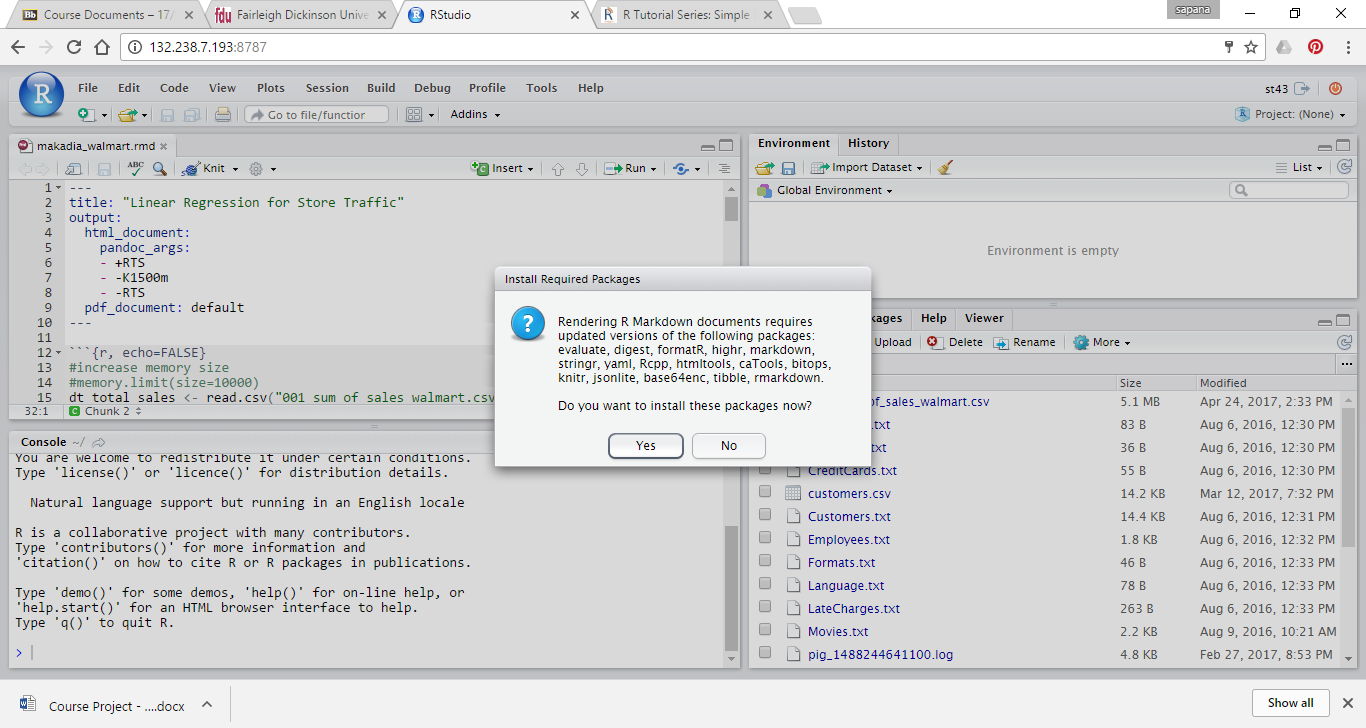
Go to menu File, then open Walmart.RMD from your directory

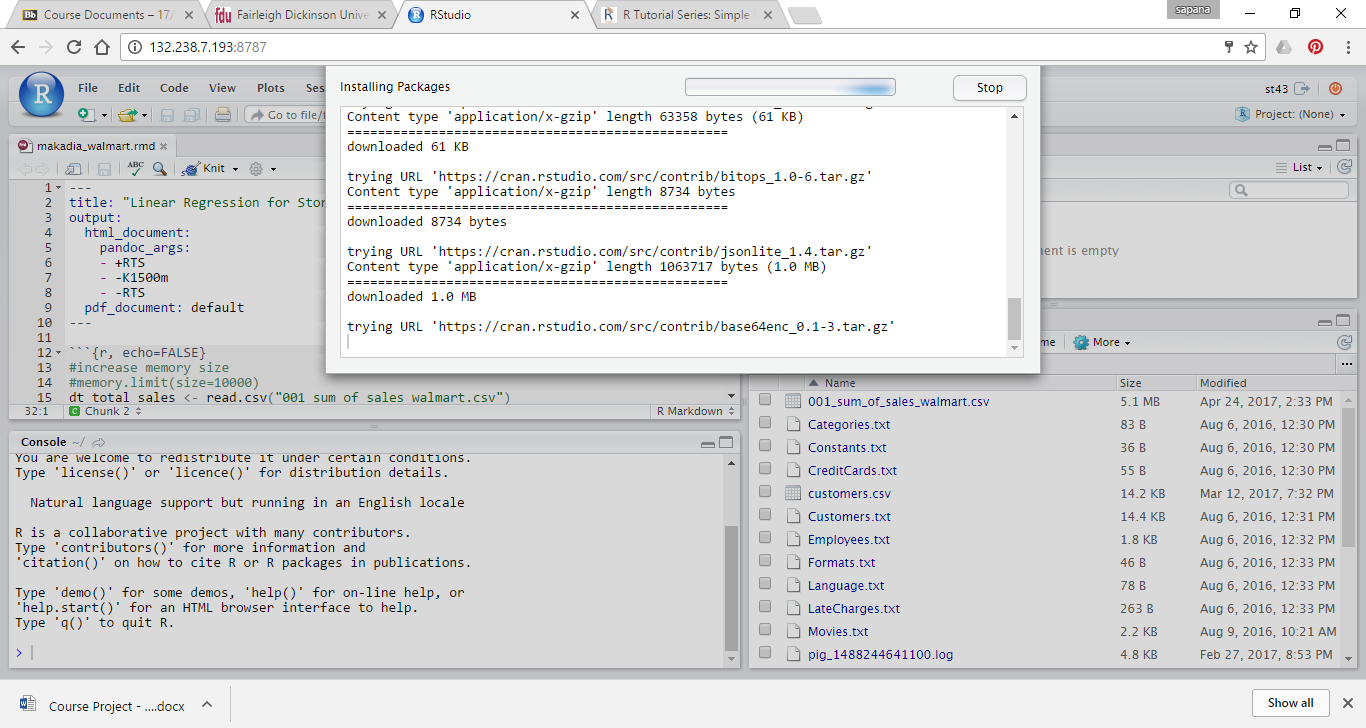


Save this R Markdown script as Makadia\_walmart.RMD

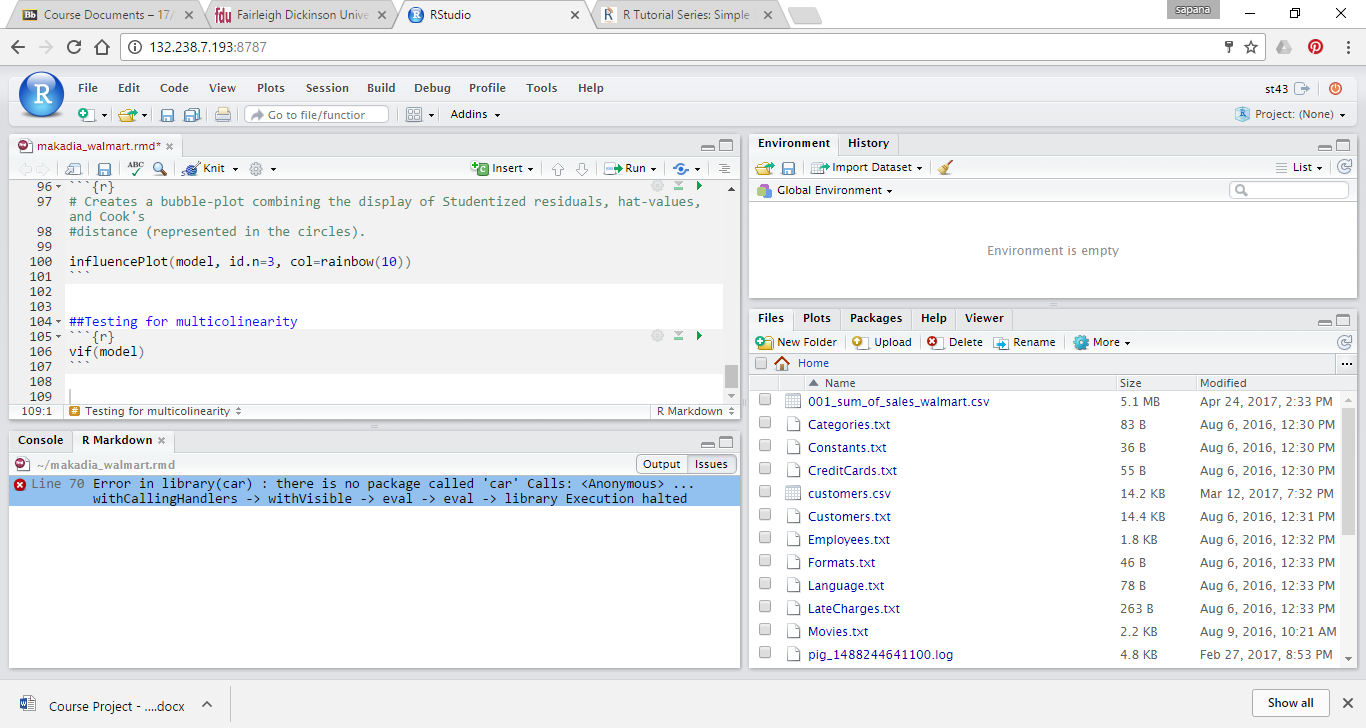


RStudio Server will produce messages that some packages are missing. Install missing packages

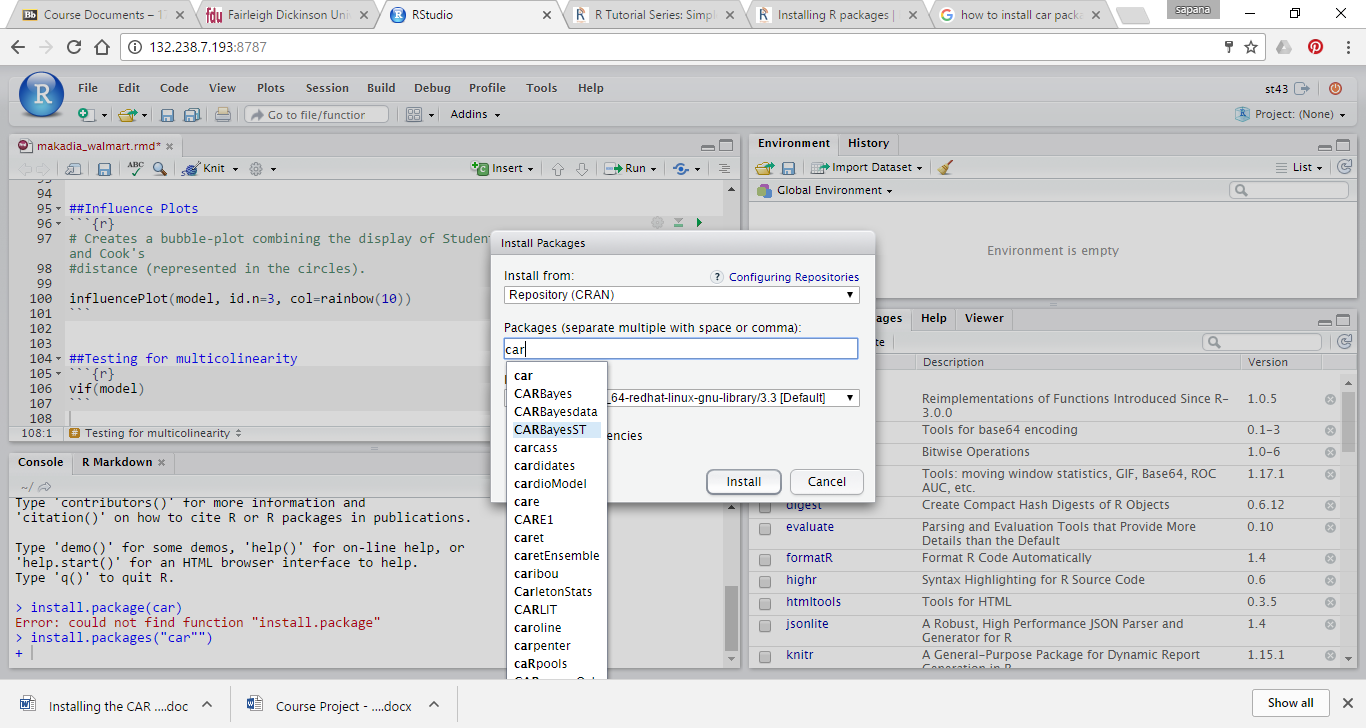


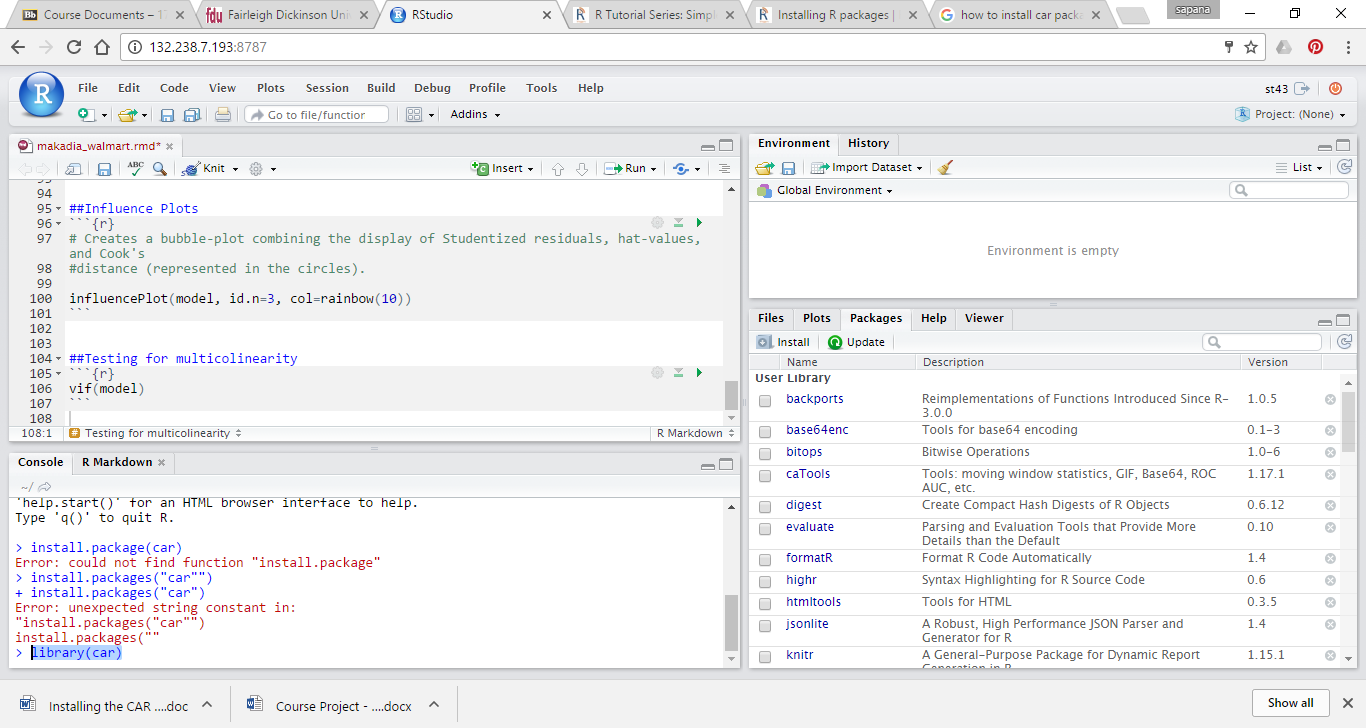


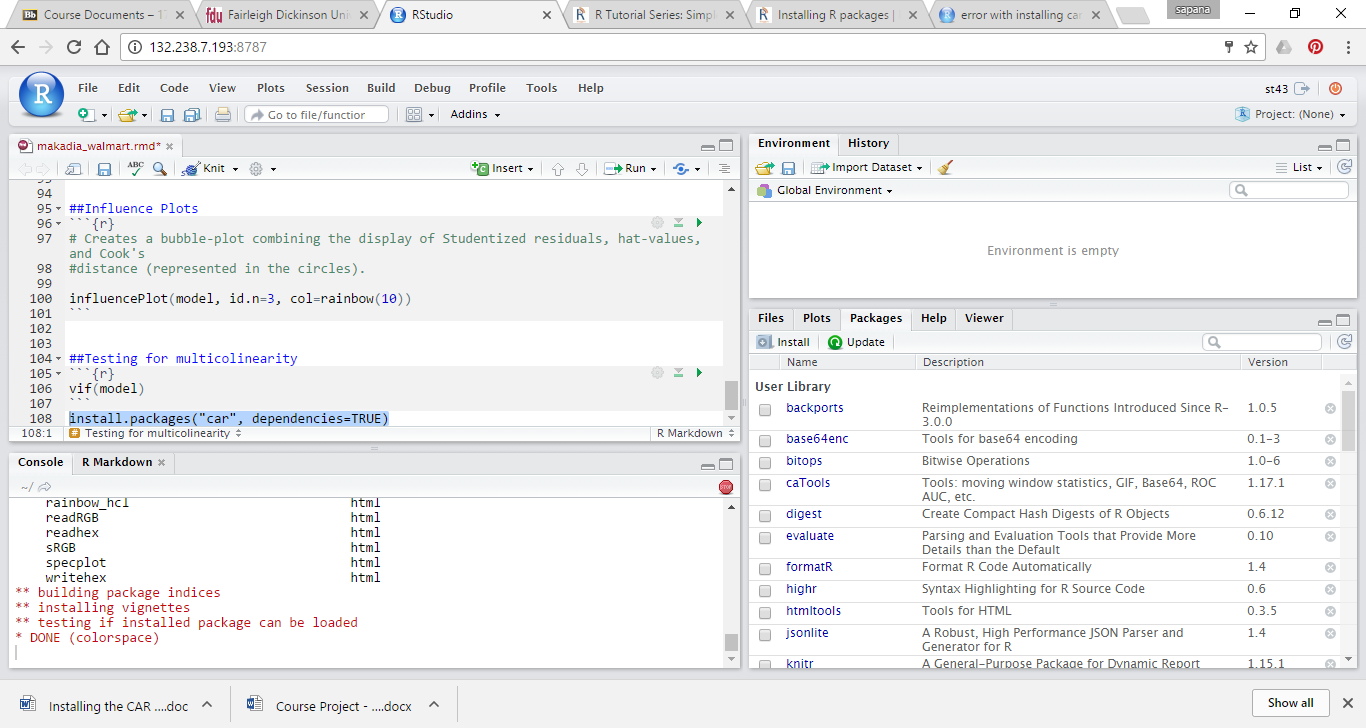
you will need to install package car otherwise it gives you error



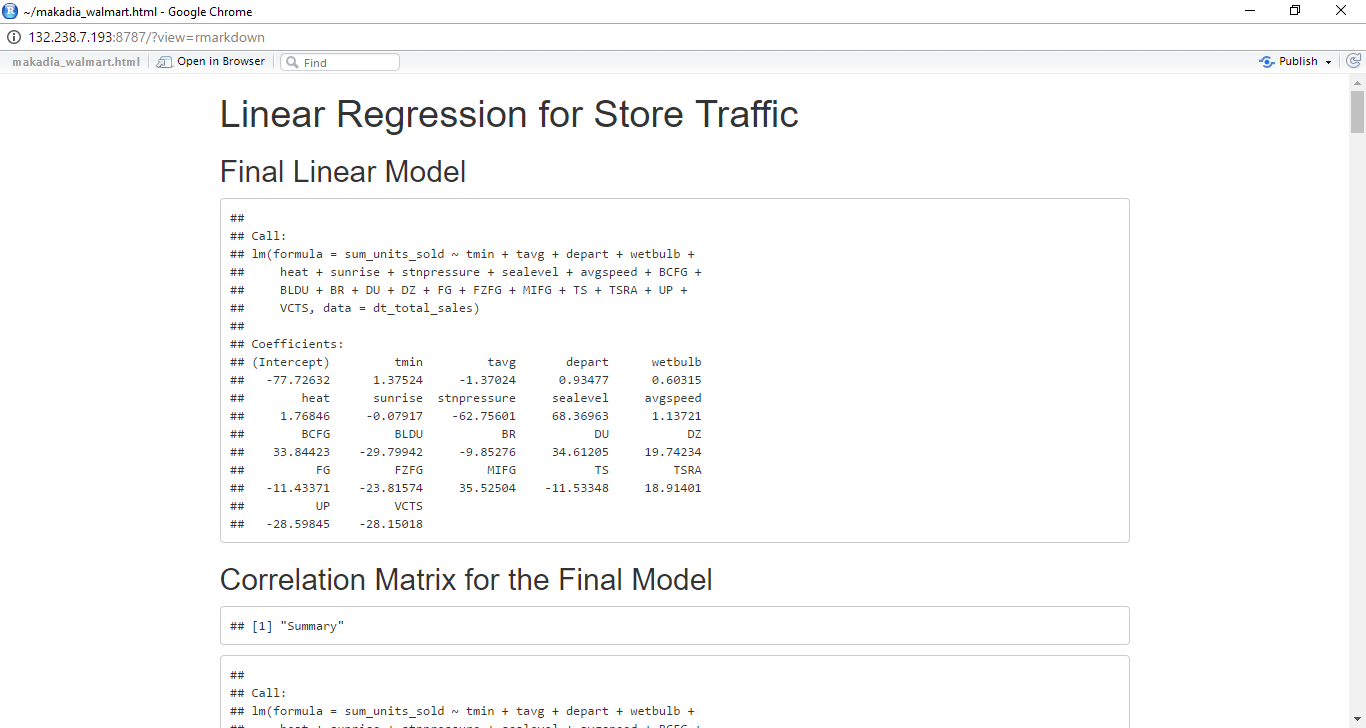
Installing “car” package

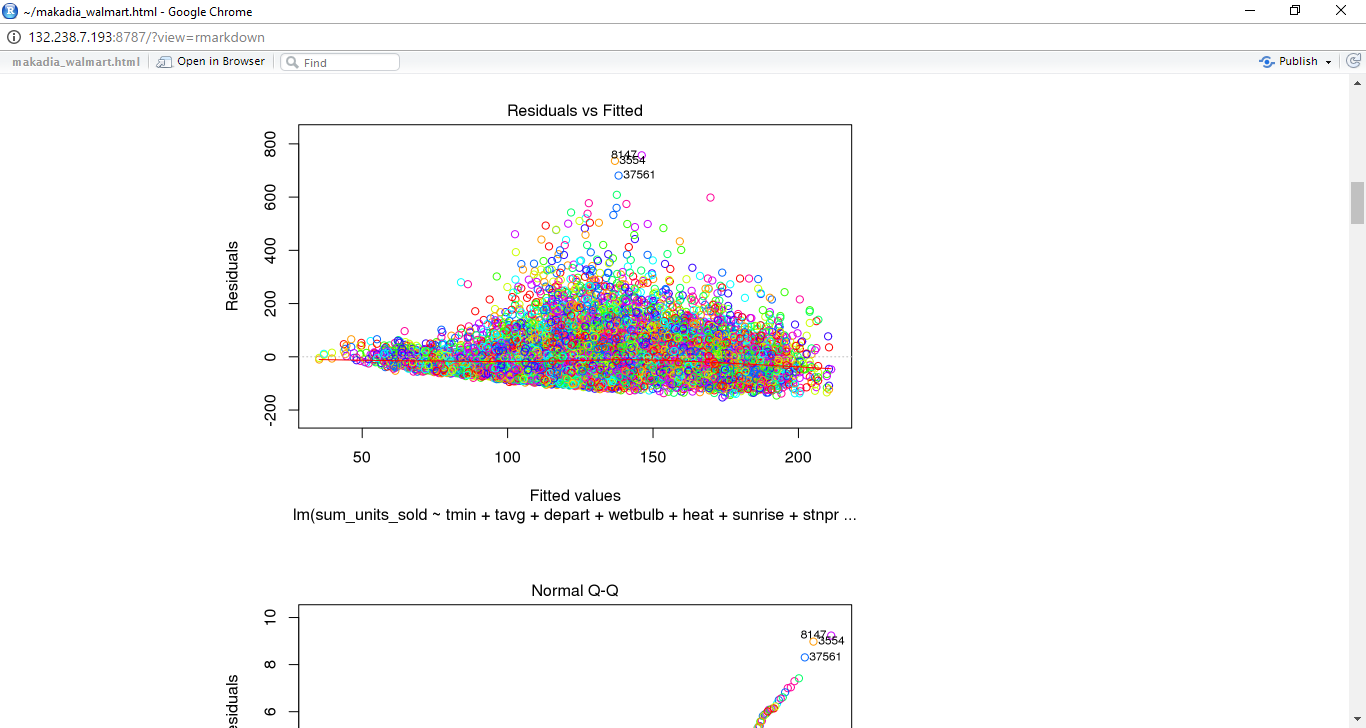






Knit HTML







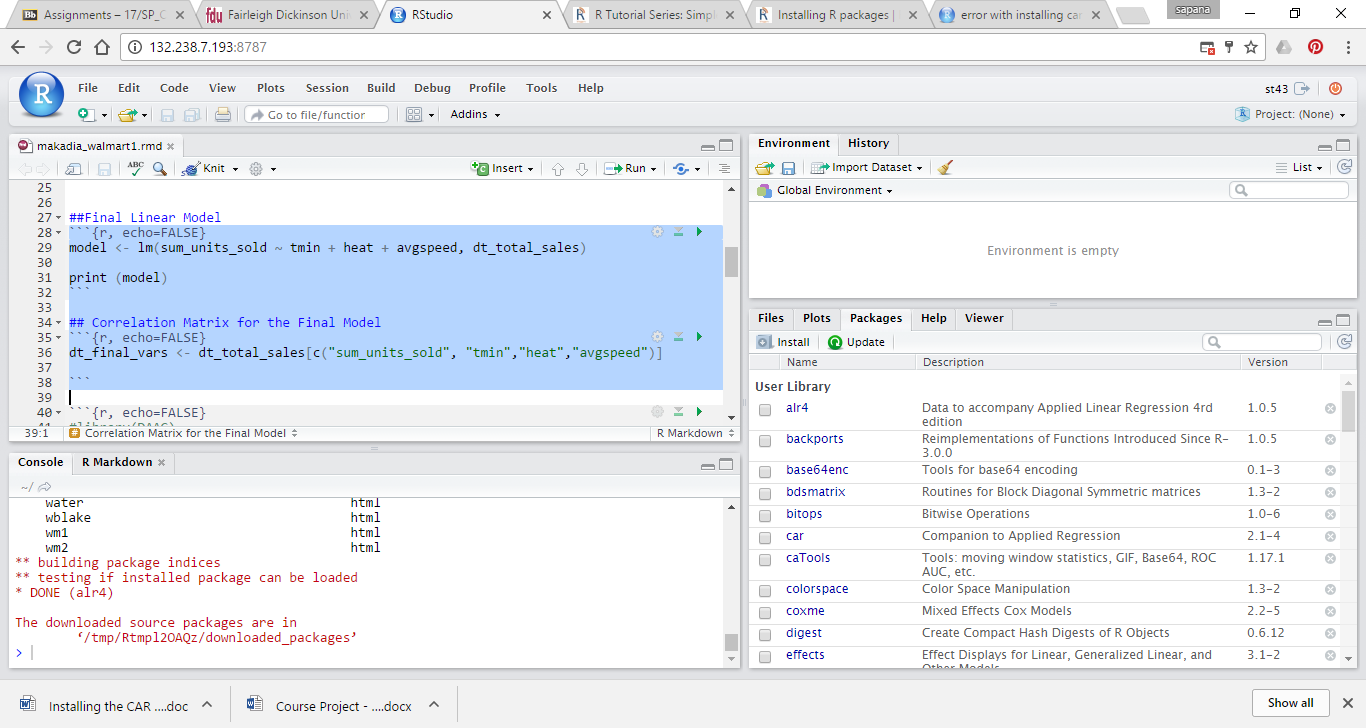
**Explain what this line does: dt\_total\_sales <- read.csv("001\_sum\_of\_sales\_walmart.csv")**

* It will read the csv file 001\_sum\_of\_sales\_walmart.csv and store data into data frame called dt\_total\_sales

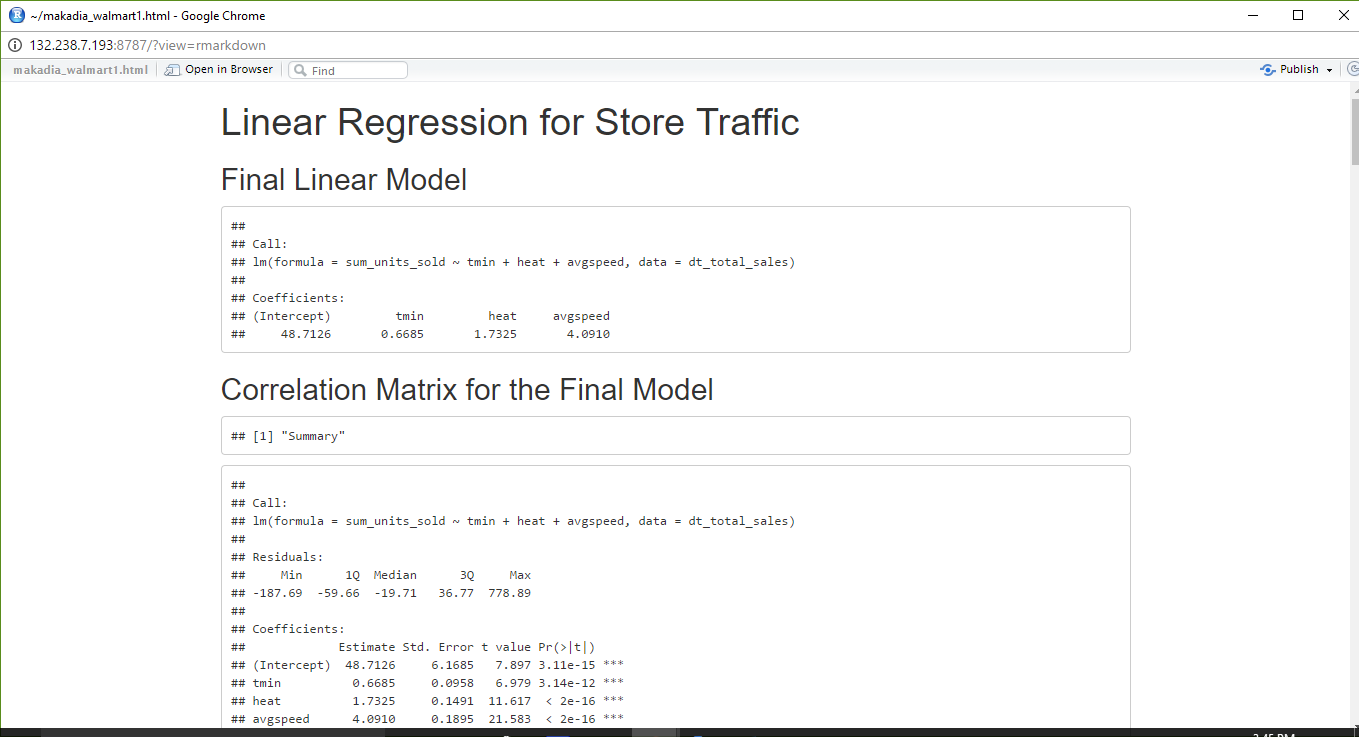
**Explain what this line does: dt\_total\_sales = na.omit(dt\_total\_sales)**

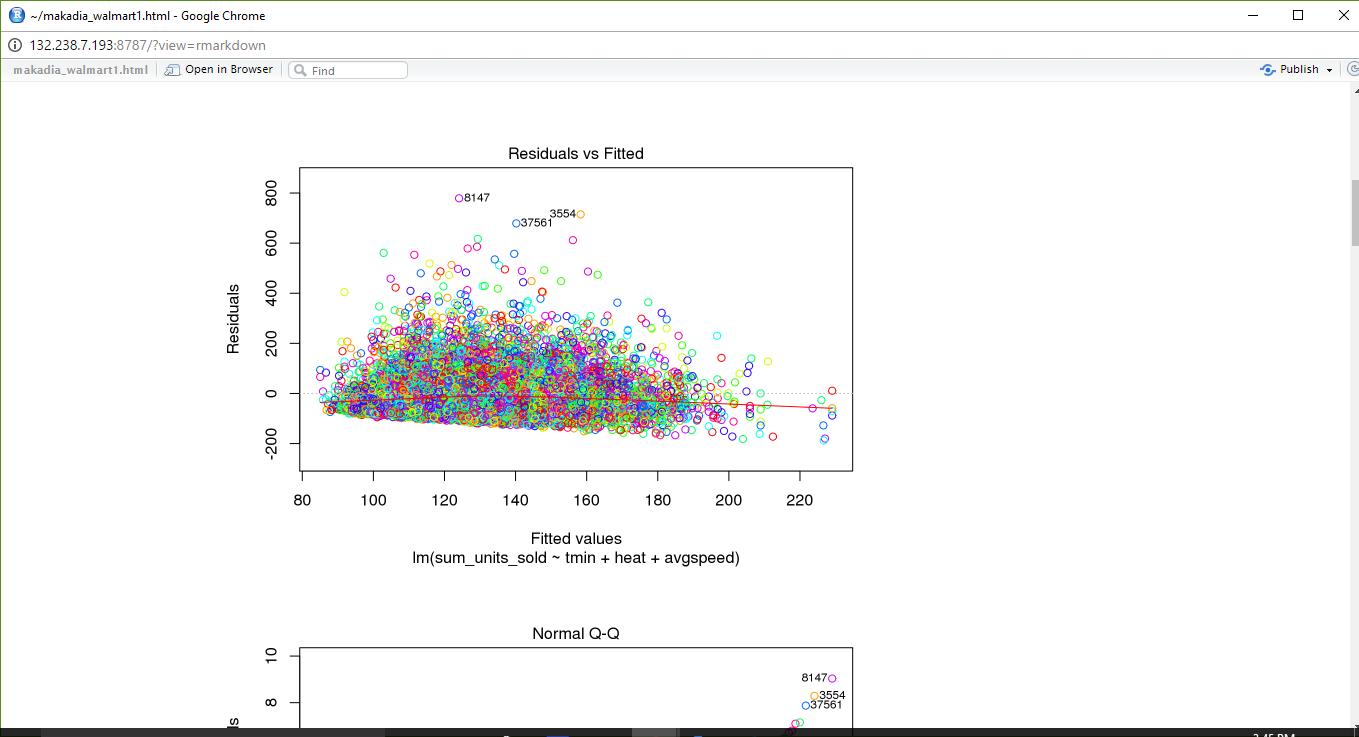
* returns the object with observations removed if they contain any missing values

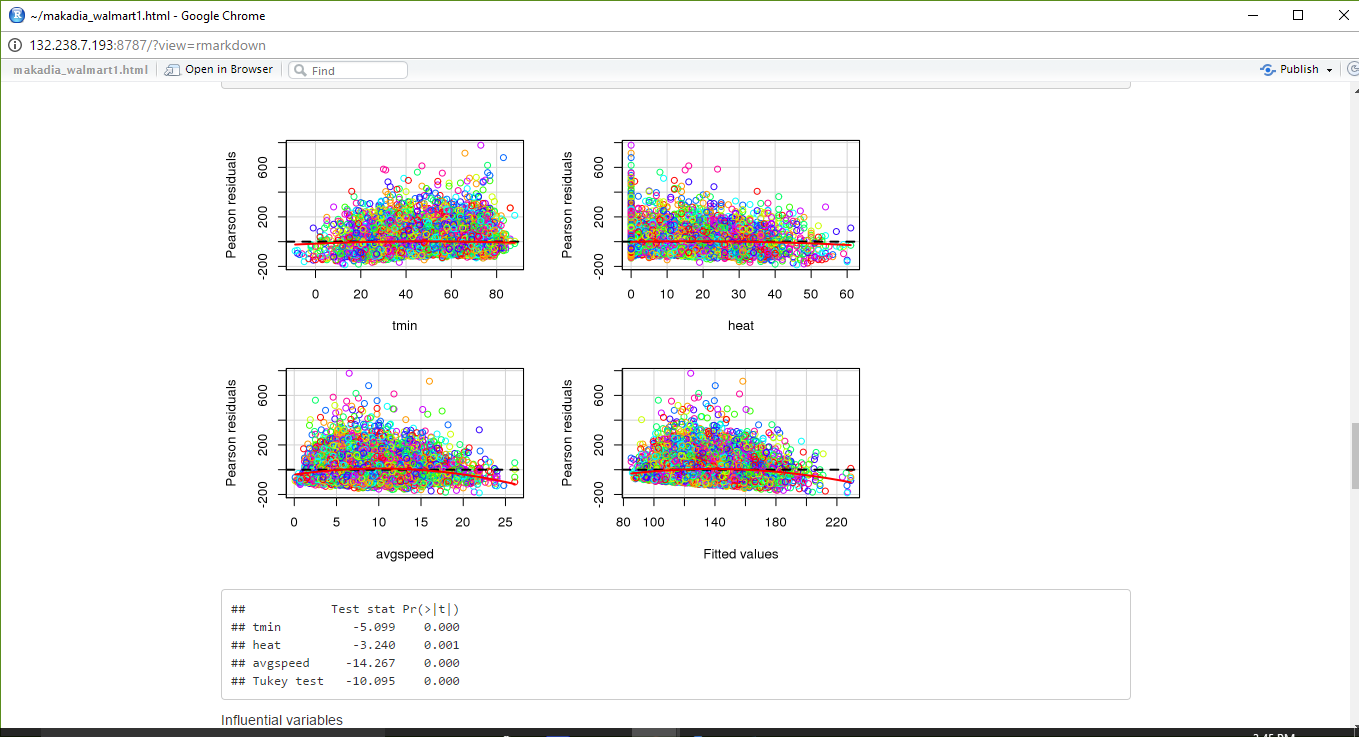
Change the code by simplifying the final model to predict sum\_units\_sold based only on the following variables: tmin, heat and avgspeed.



Knit HTML

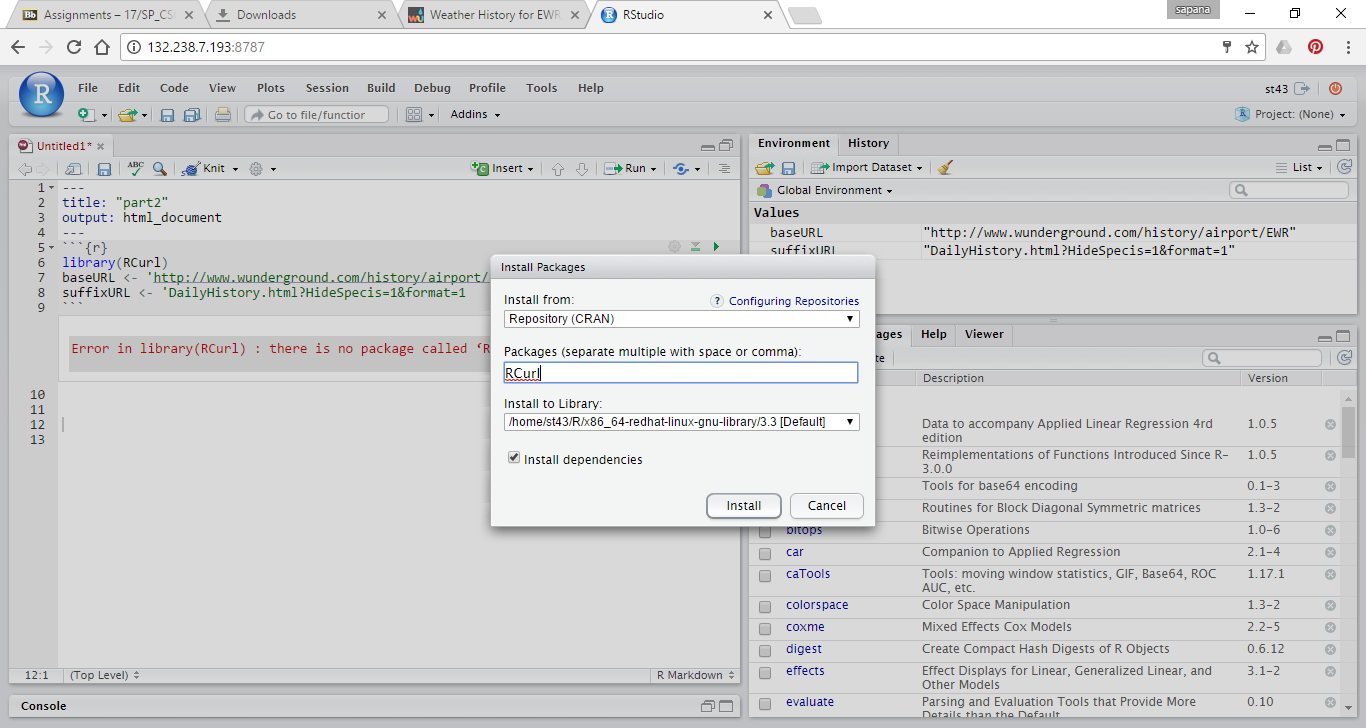




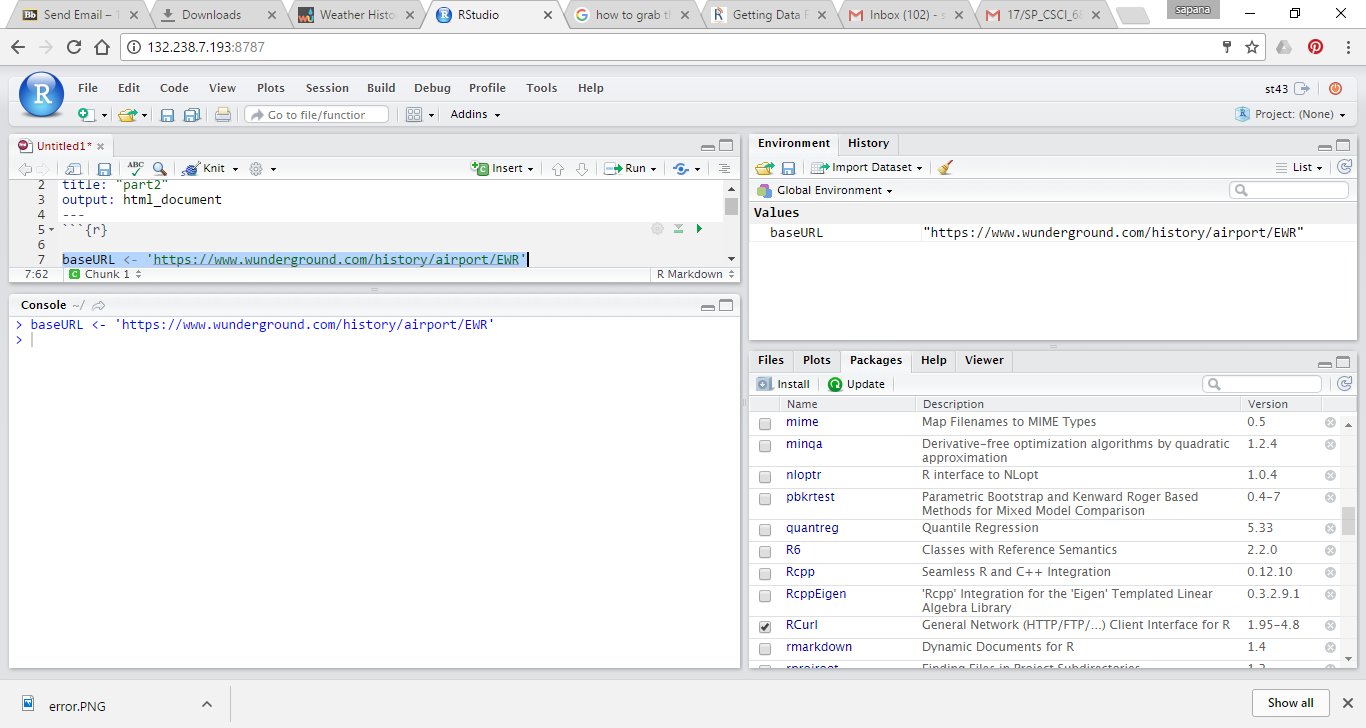


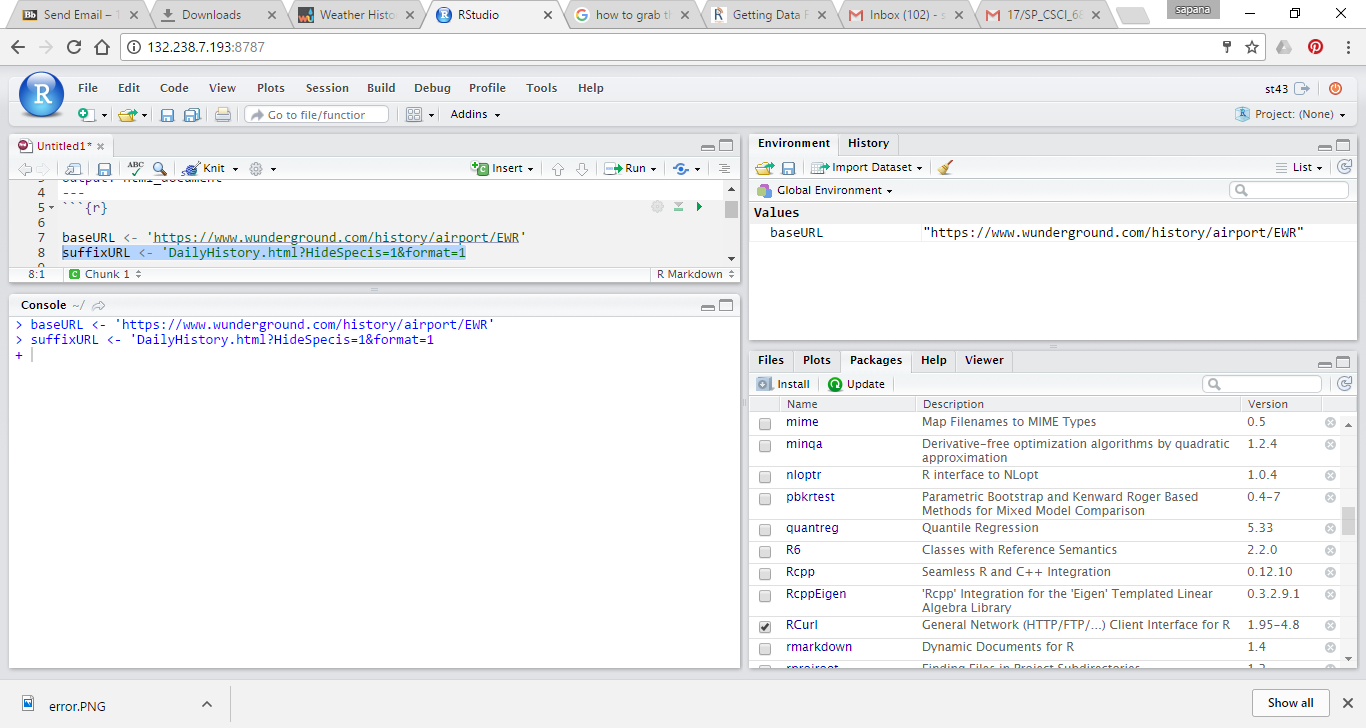
**Part-2**

You need to install RCurl package first.

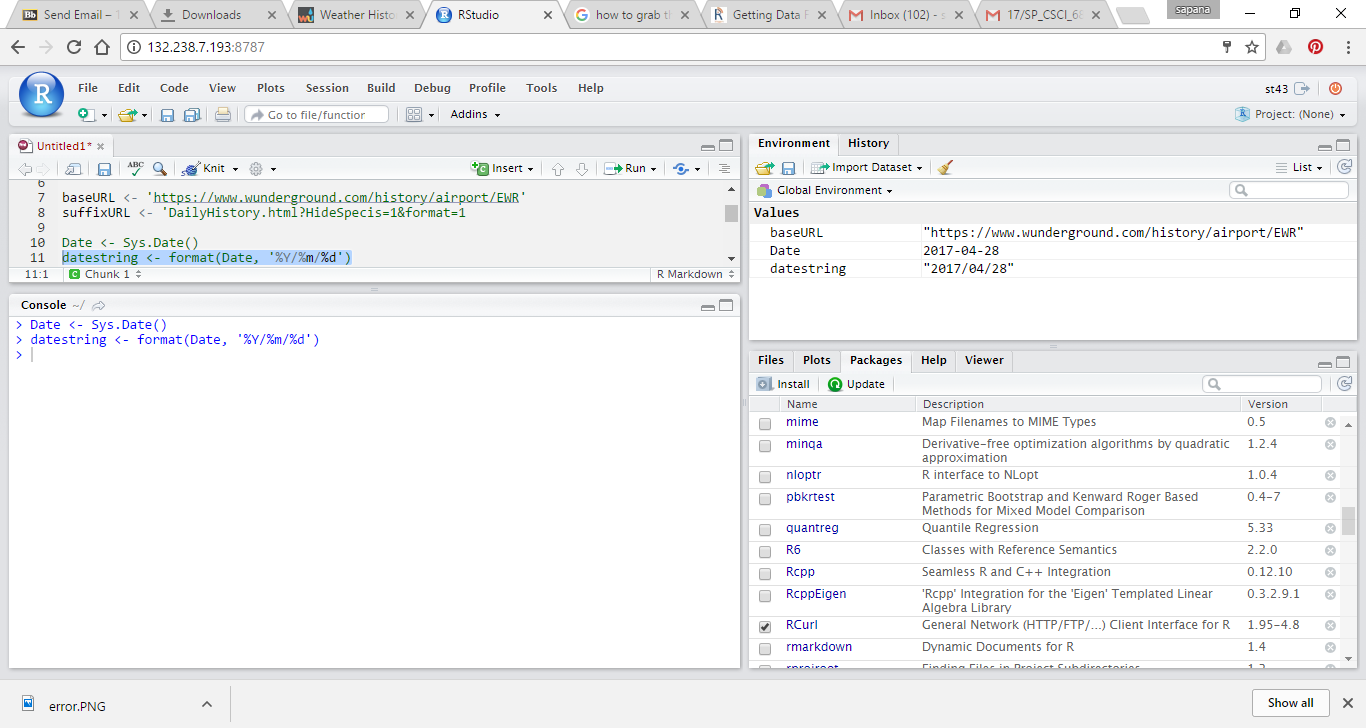
****

First, we define the beginning and end of the URL, which will never change for the Newark Airport

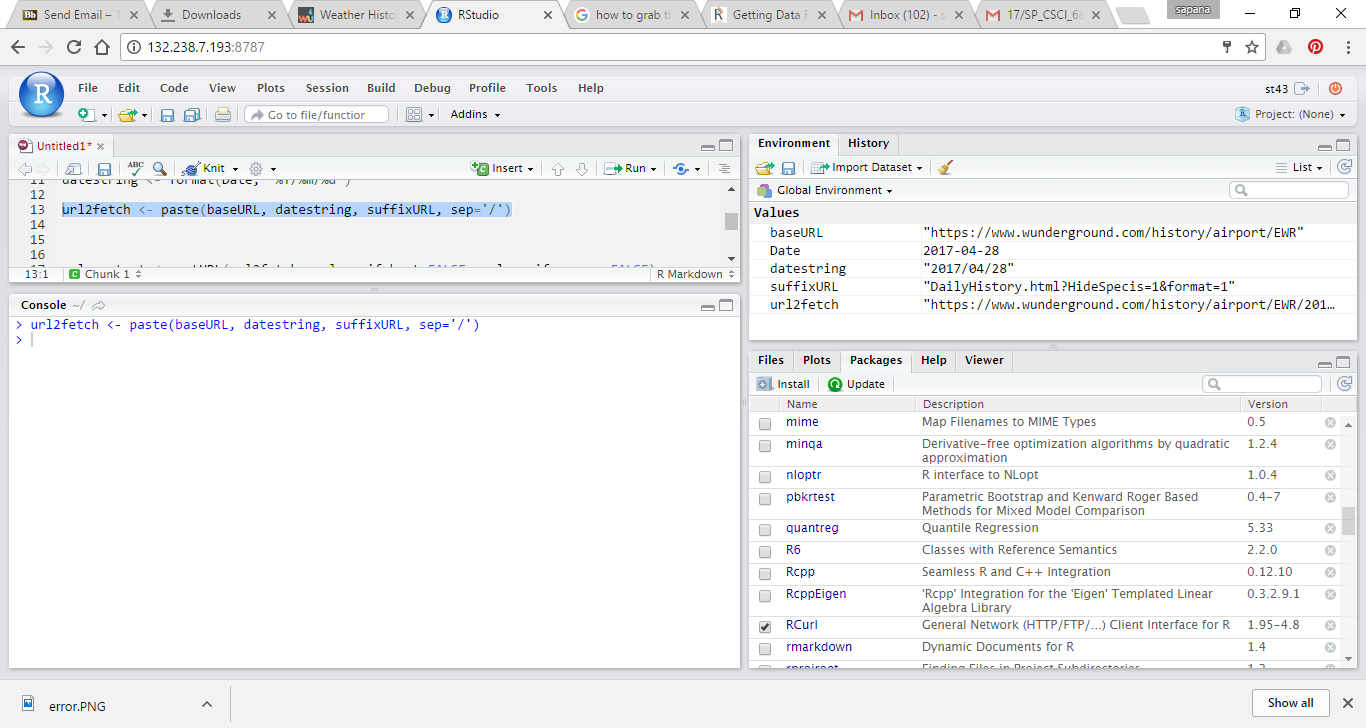




Then we get today’s date and from produce a string in format year/month/day



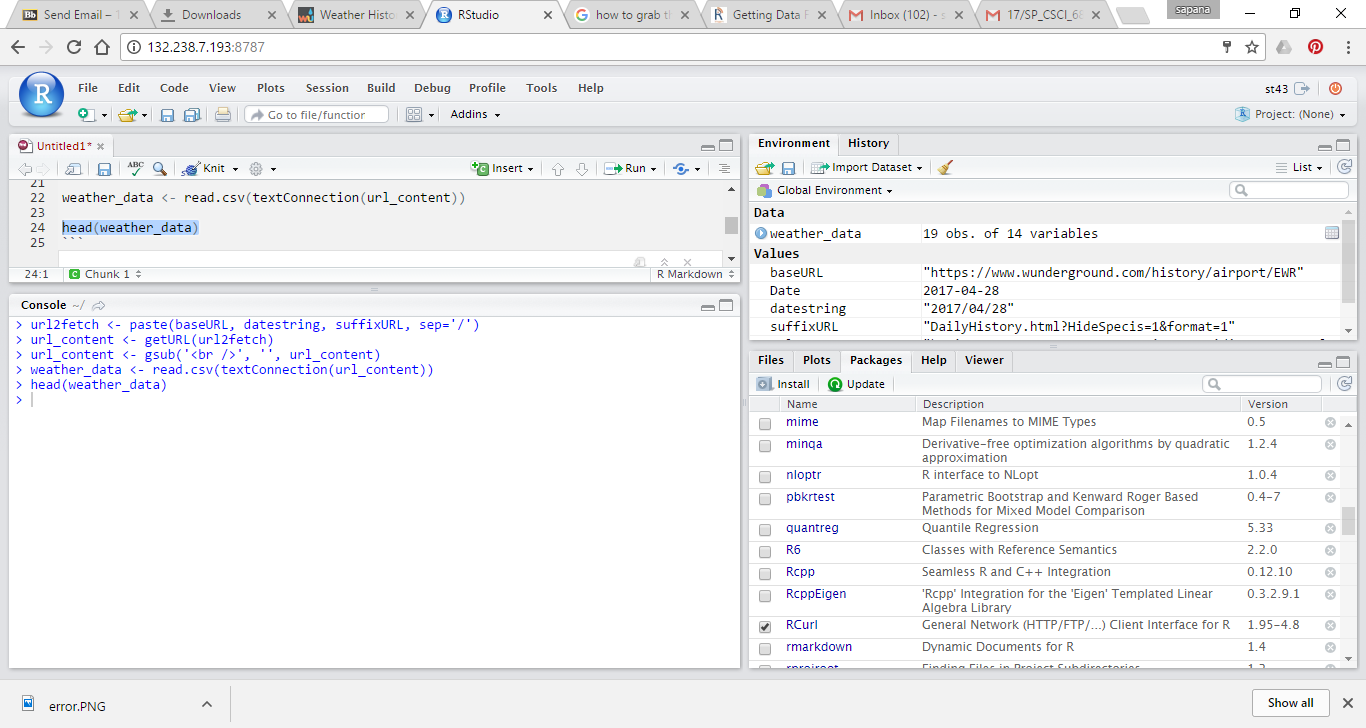
Combine all the parts of the URL and add the date to it.



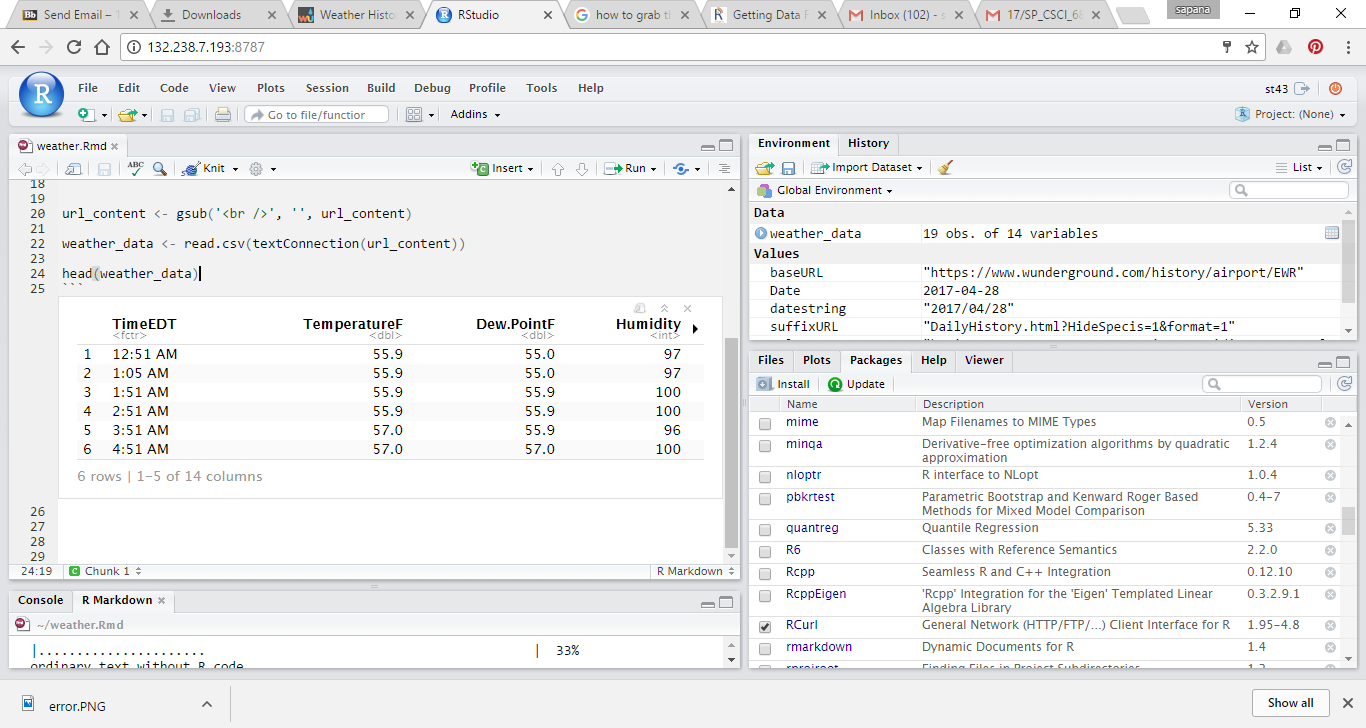
Grab the contents of the webpage at that URL using the RCurl method



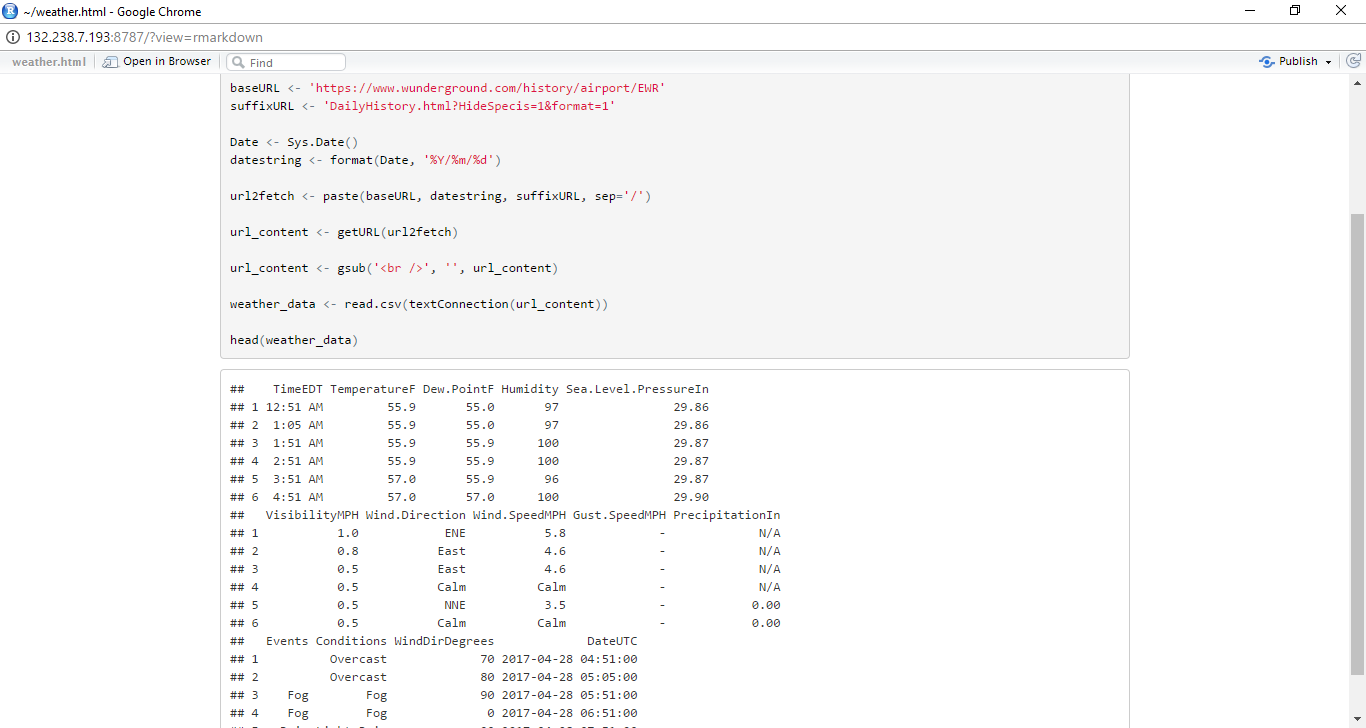
we have the content of the page in a string. We need to read the string into a data.frame object



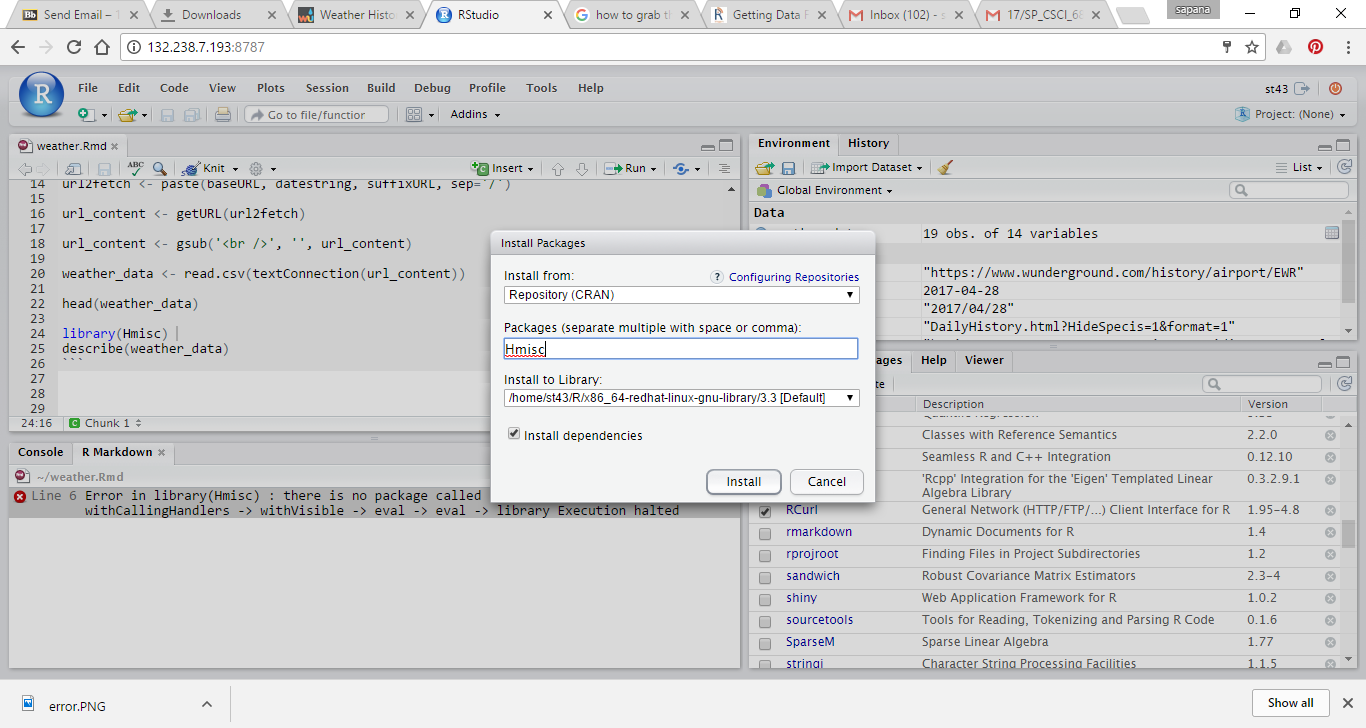
First 6 rows of data when you run head(weather\_data) line



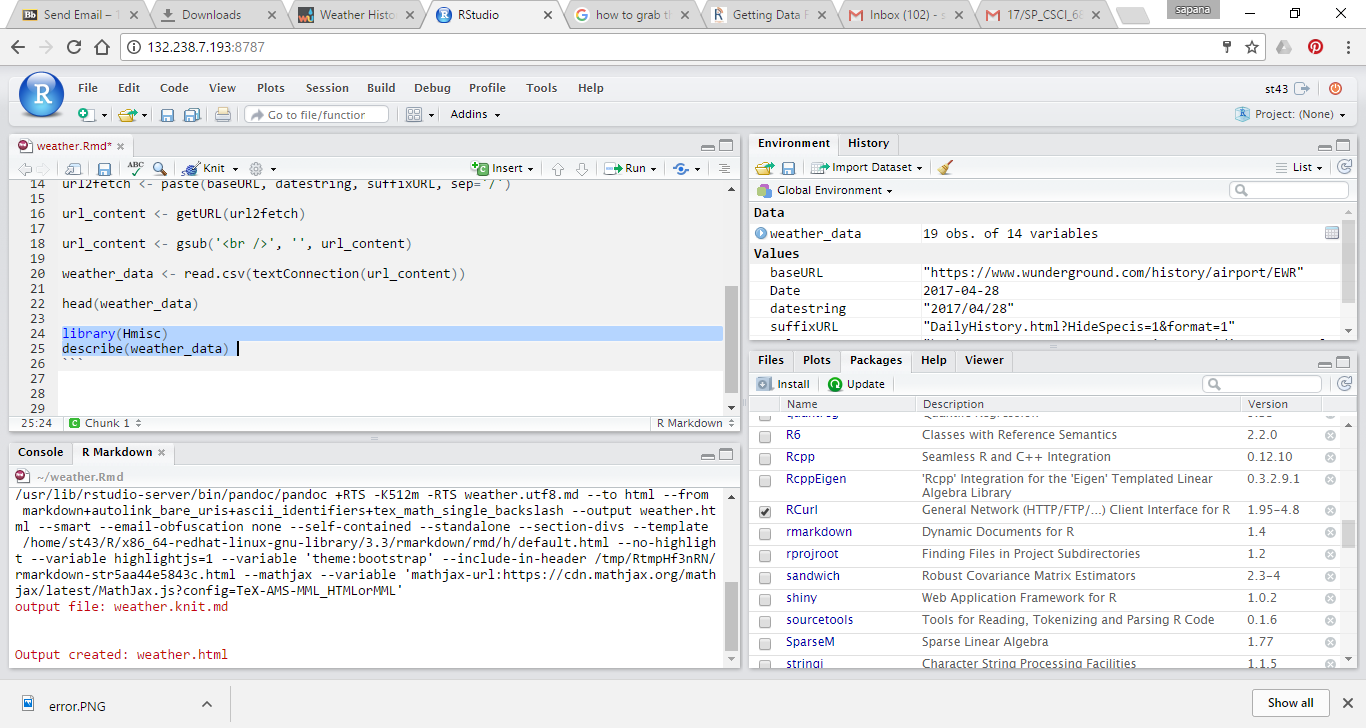
Output when you knit HTML



The describe function belongs to the Hmisc package. You need to install this package.



Let’s get extended descriptive statistics of the dataset by variable



Knit HTML

