Getting Data in and Out of R

1). Working with Excel Files

R needs excel special packages to read from Microsoft Excel. Some of the packages are XLConnect, xlsx, gdata. Xlsx package would be used here. To use it, we need to install it.

install.packages("xlsx")

To verify whether excel package has been installed

Verify the package is installed.
any(grepl("xlsx",installed.packages()))

If installed (TRUE) then load library into the workspace
Load the library into R workspace.
library("xlsx")

To read the excel workbook
d = read.xlsx("city.xlsx", sheetName = 1)

Data is read and stored as a Data frame.

2). Working with JSON Files
JSON stands for JavaScript Object Notation. JSON is a text format for storing and transporting data. JSON is "self-describing" and easy to understand. JSON is primarily used to transmit data between server and application.

To read JSON file, we need to install rjsonlite package.

install.packages("jsonlite")

To load
library("jsonlite")

To Read
jsonresult = fromJSON("sample.json")
Jsonresult

To Convert to JSON Data Frame
json_data_frame= as.data.frame(jsonresult)
Json_data_frame

Selection/Subsetting of Rows based on Column Value

1. ***data[data$dept == 'IT',]*** To select rows who work in IT department
2. ***subset(data,dept=='IT',)*** // another way of doing the same operation
3. To select rows whose salary is greater than 600 and working in IT department
   subset(data,salary>=600 & dept =='IT')
   data[data$dept == 'IT'& data$salary>=600]

4. ***data$name[data$dept == 'IT'& data$salary>=600]*** To retrieve only the names (leaving out other columns) who are getting salary > 600