



Top Exploratory Data Analysis Tech Related Problems

This document goes through some of the problems that students encounter in this course.

[Installing Packages / Loading Packages](#)

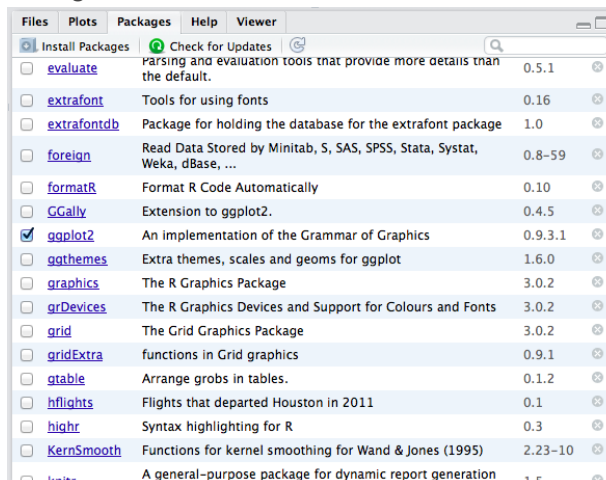
[Setting the working directory:](#)

[Incompatible Version\(s\) of R / Rstudio](#)

[Working with Gapminder Data](#)

Installing Packages / Loading Packages

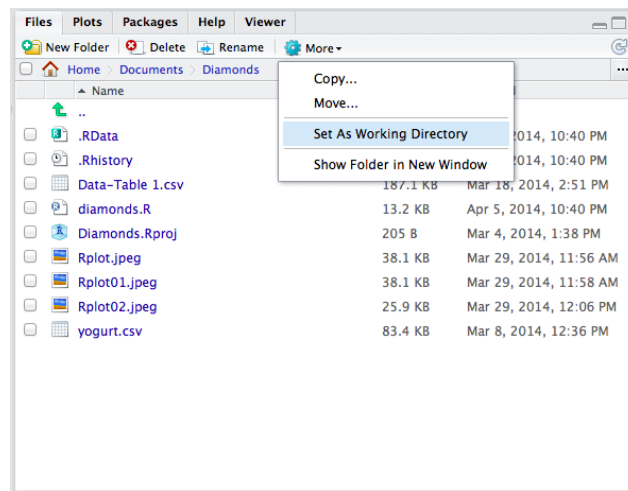
- Programmatically:
 - `install.packages('ggplot2')` OR
`install.packages('ggplot2', dependencies = TRUE)`
 - `library('ggplot2')`
- Interactively:
 - Tools -> Install Packages, type name of package and click install
 - Click on the Packages tab, check the box next to the package



- Packages must be reloaded each time Rstudio is opened
- Packages used in this course: *ggplot2*, *knitr*, *gthemes*, *gridExtra*, *plyr*, *dplyr*, *ahr3*, *energy*, *reshape2*, *GGally*, *scales*, *memisc*, *lattice*, *MASS*, *car*, *reshape*, *RColorBrewer*, *RCurl*

Setting the working directory:

- <http://www.rstudio.com/ide/docs/using/workspaces>
- Check working directory: `getwd()`
- Programmatically: `setwd("filepath of working directory")`
 - R supports tab completion
 - note: in Windows, use slash or double backslash in the file path. R will not recognize backslash
- Interactively: Navigate to working directory in the Files pane, select the More dropdown menu and click Set As Working Directory



Incompatible Version(s) of R / Rstudio

- Check Rstudio version: Rstudio -> About Rstudio
- Check R version: In R, R -> About R
- Update Rstudio: Help -> Check for Updates
- Update R: <http://cran.r-project.org/>

Working with Gapminder Data

- Gapminder datasets often have lots of missing values that require data munging ie: <https://spreadsheets.google.com/a/knowlabs.com/pub?key=rEMA-cbNPaoTpdYxTcwugnw&gid=0>
- Try to find a gapminder dataset that is tidy
- Most of datasets are quite small so data munging can be done manually using Excel