

Data Analysis with R

Follow along at: <http://bit.ly/data-analysis-r>

See the code at: <http://bit.ly/data-analysis-r-code>

Open your browser to:

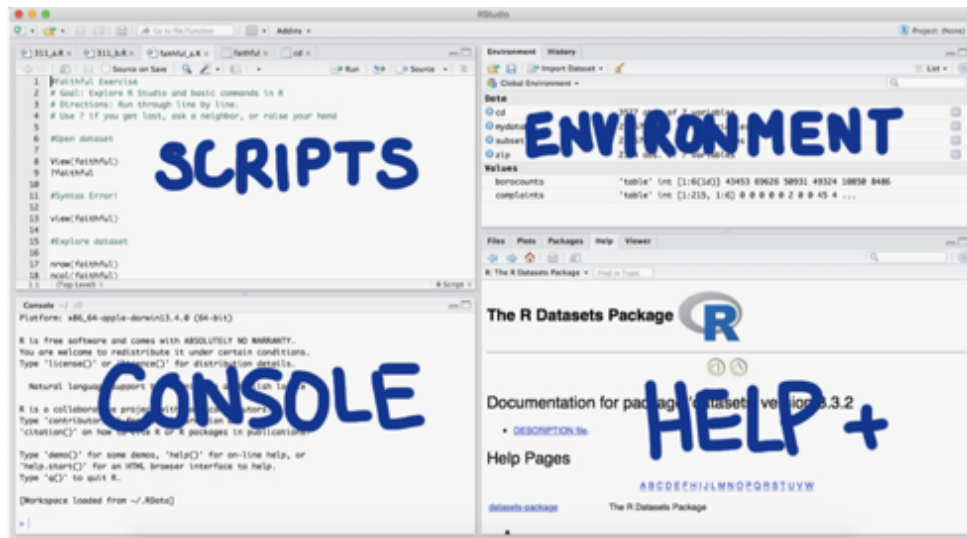
http://student___.datapolitan.com

Username: rstudio

Password: rstudio

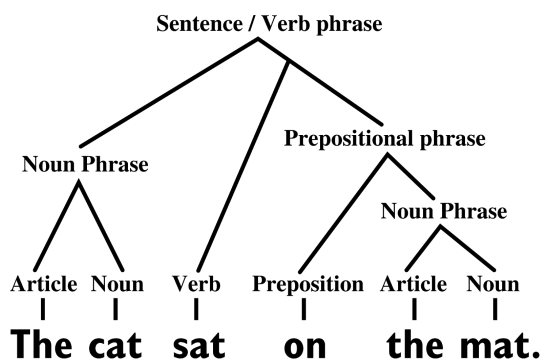
Key Questions for the Morning

What is R?	How is R useful in analysis?
------------	------------------------------



What is Syntax?

Basic constituent structure analysis of a sentence:



R Syntax

```
# basic command
```

```
command(dataset)
View(faithful)
```

```
# select a column
```

```
command(dataset$column)
mean(faithful$waiting)
```

```
# get help
```

```
?help
?faithful
```

Exploring Data

```
View()
# show dataset as spreadsheet in Viewer
```

```
str()
# identify data type and structure
```

```
nrow()
# identify the number of rows
```

```
ncol()
# identify the number of columns
```

```
colnames()
# list the name of every column
```

Visualizing Data

```
hist()
# make a chart with numeric data
```

```
plot()
# plot two numeric variables along an x-y axis
```

```
abline()
# add a trendline to a plot
```

```
table()
# make a table with factor data
```

```
prop.table()
# make a table with percentages
```

```
barplot()
# make a chart with factor data
```

Manipulate Data

```
sort()  
# sort the values in a column
```

```
data.frame()  
# structure data into a matrix
```

```
subset()  
# extract data from a dataframe
```

Calculating Summary statistics

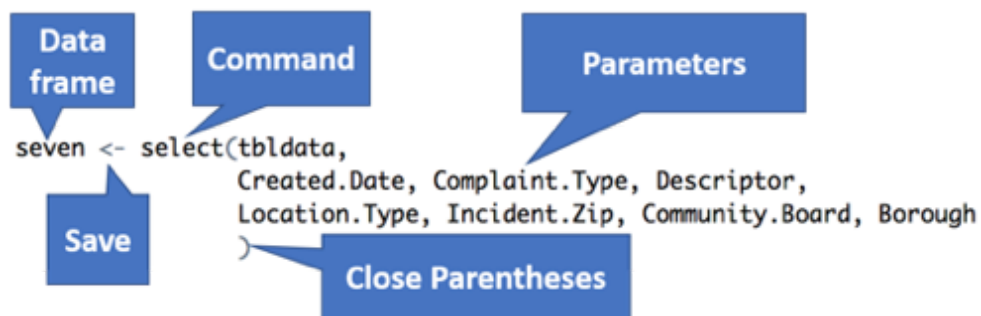
```
min()  
# identify minimum value
```

```
max()  
# identify maximum value
```

```
median()  
# calculate median value
```

```
mean()  
# calculate mean value
```

Important Parts of R Commands We're Using



```
date <- mdy_hms(as.character(cbpkg$Created.Date))
```

Nested Functions



Key Questions for the Afternoon

- What are Algorithms?

- What is Debugging?

How is R/RStudio helpful for me?	What key tasks do I need to learn/practice to best use R/RStudio?
----------------------------------	---

Links

- Download R: <https://cloud.r-project.org/>
- Download RStudio: <https://www.rstudio.com/products/rstudio/download/>
- Download exercise files from this class: <http://bit.ly/data-analysis-r-code>

Learning More

- **Stat Methods** (<http://statmethods.net>) - Great documentation for doing data analysis in R
- **UCLA Stats** (<http://www.ats.ucla.edu/stat/>) - Many examples of statistical analysis with comparisons between R, Stata, SPSS, etc.
- **Swirl** (<http://swirlstats.com>) - install.packages("swirl")
- **R Programming Coursera Course** (<https://www.coursera.org/learn/r-programming>) - free course in R that runs regularly
- **R for Data Science** (<https://r4ds.had.co.nz/>) - Free online book covering many basic topics in data science
- **R Cookbook** (<http://www.cookbook-r.com>) - Free online walkthrough of the basics

Other Resources

- **NYC Open Statistical Programming Meetup** (<https://www.meetup.com/nyhackr>) - Monthly talks about R (also sponsor the NYC R Conference)
- **UCLA Stats Data Analysis Examples** (<https://stats.idre.ucla.edu/other/dae>) - Examples of statistical analysis in different platforms (R, Stata, SPSS, etc.)
- **Tidyverse** (<https://www.tidyverse.org>) - R packages for Data Science
- **Stack Overflow** (<http://stackoverflow.com/>) - One of the best Q&A sites for various technical questions
- **Datapolitan training classes** <http://training.datapolitan.com/>

Contact Us



training@datapolitan.com



[@datapolitan](https://twitter.com/datapolitan)



<http://www.datapolitan.com>

About Datapolitan

For over 5 years, Datapolitan has worked to empower public sector clients with the skills, techniques, concepts, and mindsets necessary to make data meaningful and actionable to effectively manage their resources and realize operational value from the information they collect. We do this through a range of consulting and training services for local government agencies and nonprofit organizations, customized to their strategic vision and operational needs.

Your Notes