



# *Data Visualization*

## Why are We Here?



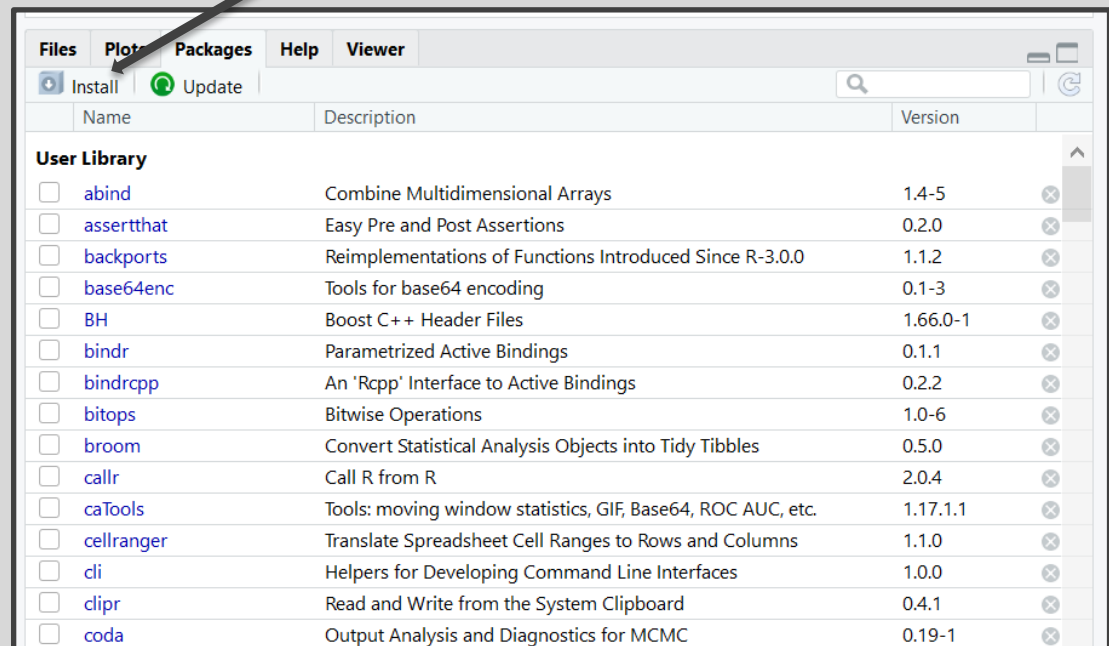
- To Develop a Personal and Intimate Relationship With R
- To Install Some Key R Packages
  - Tidyverse
  - Rmarkdown
- To Practice Coding via R Scripts
- To Learn Elements of ggplot2
- Practice Making Visually Stunning Pictures that Change Lives

# Initial Steps in RStudio



- Install Tidyverse Package
  - > `install.packages("tidyverse")`
- Other Packages To Be Installed
  - RColorBrewer
  - Rmarkdown

Select Install and Search on CRAN



# Initial Steps in RStudio

- To Use the Package

- Code

```
> library("tidyverse")
```

- Check Box for Tidyverse




Files	Plots	Packages	Help	Viewer
Install		Update		
		Name	Description	Version
<input checked="" type="checkbox"/>		tidyverse	Easily Install and Load the 'Tidyverse'	1.2.1
<input type="checkbox"/>		tinytex	Helper Functions to Install and Maintain 'TeX Live', and Compile 'LaTeX' Documents	0.6
<input type="checkbox"/>		utf8	Unicode Text Processing	1.1.4
<input type="checkbox"/>		viridisLite	Default Color Maps from 'matplotlib' (Lite Version)	0.3.0
<input type="checkbox"/>		whisker	{{mustache}} for R, logicless templating	0.3-2
<input type="checkbox"/>		withr	Run Code 'With' Temporarily Modified Global State	2.1.2
<input type="checkbox"/>		xfun	Miscellaneous Functions by 'Yihui Xie'	0.3
<input type="checkbox"/>		XLConnect	Excel Connector for R	0.2-15
<input type="checkbox"/>		XLConnectJars	JAR Dependencies for the XLConnect Package	0.2-15
<input type="checkbox"/>		xml2	Parse XML	1.2.0
<input type="checkbox"/>		yaml	Methods to Convert R Data to YAML and Back	2.2.0
<input type="checkbox"/>		zoo	S3 Infrastructure for Regular and Irregular Time Series (Z's Ordered Observations)	1.8-3
<b>System Library</b>				
<input type="checkbox"/>		boot	Bootstrap Functions (Originally by Angelo Canty for S)	1.3-20
<input type="checkbox"/>		class	Functions for Classification	7.3-14
<input type="checkbox"/>		cluster	"Finding Groups in Data": Cluster Analysis Extended Rousseeuw et al	2.0.7-1

# ggplot2

- Help Page: [Link](#)
- Comes with Preloaded Datasets



 **ggplot2** part of the [tidyverse](#)

Reference

### Data

ggplot2 comes with a selection of built-in datasets that are used in examples to illustrate various visualisation challenges.

<b>diamonds</b>	Prices of 50,000 round cut diamonds
<b>economics</b>	US economic time series
<b>faithfuld</b>	2d density estimate of Old Faithful data
<b>midwest</b>	Midwest demographics
<b>mpg</b>	Fuel economy data from 1999 and 2008 for 38 popular models of car
<b>msleep</b>	An updated and expanded version of the mammals sleep dataset
<b>presidential</b>	Terms of 11 presidents from Eisenhower to Obama
<b>seals</b>	Vector field of seal movements
<b>txhousing</b>	Housing sales in TX
<b>luv_colours</b>	<code>colors()</code> in Luv space

## ggplot2



- Many Useful Plots and Charts Provided
  - See Cheat Sheet: [Link](#)  
(Also on Course Website)
  - Called Geoms (Geometric Objects)
  - The Geom you choose Must Comply with the Type of Variables You are Analyzing
- Organized by Type of Data
  - Univariate
  - Bivariate
  - Mixtures of Categorical and Numeric

ggplot2



- ggplot2 General Form

Fill in Blank With  
Name of Data in R

```
> ggplot(data=_____) +
```

```
  geom_TYPE(mapping=aes(x=____,y=____,  
etc.))
```

Fill in Blanks from  
Variables in Data

## Many Examples



- Locate Supplement on Website
- Download Rmd File
- Open Rmd File on Computer
- Knit the Rmd File to HTML format
- View Graphs with Me



Closing



Disperse  
and Make  
Reasonable  
Decisions