DARTMOUTH

Building a Practice of Data-Driven Decision-Making:

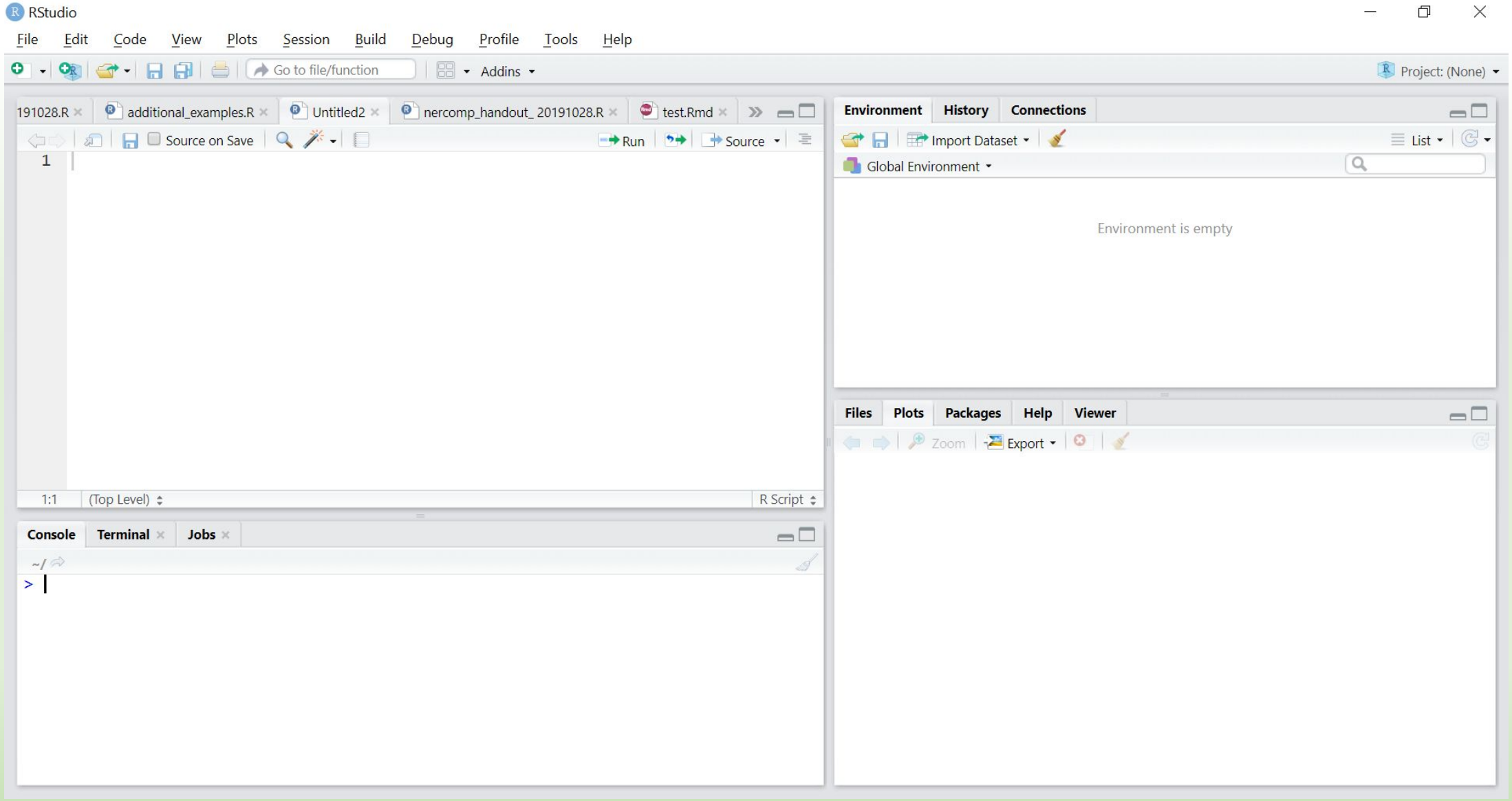
Harvesting the Data

Supplemental Slides

Steve Gaughan
Jing Qi
Dartmouth College

Resources

- The "R" project home page: https://www.r-project.org/
  - Download R https://cran.cnr.berkeley.edu/
- R Studio: https://rstudio.com/
- Software Carpentry https://software-carpentry.org/lessons and http://swcarpentry.github.io/r-novice-inflammation/
- Datasets for today: dartgo.org/workshopsg
- R Documentation https://www.rdocumentation.org/
- R Studio "cheatsheets": https://rstudio.com/resources/cheatsheets/
- Stack overflow site, R: https://stackoverflow.com/questions/tagged/r


RStudio
File Edit Code View Plots Session Build Debug Profile Tools Help
Go to file/function
Addins
Project: (None)
191028.R
additional_examples.R
Untitled2
nercomp_handout_ 20191028.R
test.Rmd
Source on Save
Run
Source
1
1:1
(Top Level)
R Script
Console
Terminal
Jobs
~/
>
Environment
History
Connections
Import Dataset
List
Global Environment
Environment is empty
Files
Plots
Packages
Help
Viewer
Zoom
Export

RStudio

File Edit Code View Plots Session Build Debug Profile Tools Help

Go to file/function | Addins | Project: (None)

191028.R | additional_examples.R | Untitled2 | nercomp_handout_ 20191028.R | test.Rmd

Source on Save | Run | Source

```
117 library(plyr)
118
119 count(sesame_street_characters, 'Name')
120
121 # Simple Pie Chart
122 slices <- c(10, 12,4, 16)
123 lbls <- c("US", "UK", "Australia", "Germany")
124 pie(slices, labels = lbls, main="Pie Chart-Country output")
125
126 # -------------
127 # exercise steps: extract,transform,join,priority list,count,sort,select
128 # read data in from the web
129 df1_web <- read.csv("https://rcweb.dartmouth.edu/homes/f002d69/workshops/da
130 df2_web <- read.csv("https://rcweb.dartmouth.edu/homes/f002d69/workshops/da
131
132 # view data -  View(df1_web)  View(df2_web)
133
```

126:1 (Untitled) R Script

Console | Terminal | Jobs

```
~/
> # vector
> squares <- c(1,4,9,16)
> # make an intentional error, explain error text and how to use it to problem-so
lve
> mean(myvar)
Error in mean(myvar) : object 'myvar' not found
> # Simple Pie Chart
> slices <- c(10, 12,4, 16)
> lbls <- c("US", "UK", "Australia", "Germany")
> pie(slices, labels = lbls, main="Pie Chart-Country output")
>
```

Environment | History | Connections

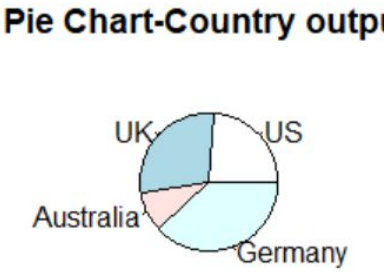
Import Dataset | List

Global Environment

Data	
names	6 obs. of 2 variables
primarycolors	6 obs. of 2 variables
Values	
bekind	TRUE
educational_mons…	"Elmo"
favoritedec	3.1416
favoritenum	25
lbls	chr [1:4] "US" "UK" "Australia" "Germany"

Files | Plots | Packages | Help | Viewer

Zoom | Export | Publish



RStudio

File Edit Code View Plots Session Build Debug Profile Tools Help

Go to file/function | Addins | Project: (None)

additional_examples.R | Untitled2 | nercomp_handout_ 20191028.R | my_mtcars | test.Rmd

Filter

	mpg	cyl	disp	hp	drat	wt	qsec	vs	am	gear	carb
Mazda RX4	21.0	6	160.0	110	3.90	2.620	16.46	0	1	4	4
Mazda RX4 Wag	21.0	6	160.0	110	3.90	2.875	17.02	0	1	4	4
Datsun 710	22.8	4	108.0	93	3.85	2.320	18.61	1	1	4	1
Hornet 4 Drive	21.4	6	258.0	110	3.08	3.215	19.44	1	0	3	1
Hornet Sportabout	18.7	8	360.0	175	3.15	3.440	17.02	0	0	3	2
Valiant	18.1	6	225.0	105	2.76	3.460	20.22	1	0	3	1
Duster 360	14.3	8	360.0	245	3.21	3.570	15.84	0	0	3	4
Merc 240D	24.4	4	146.7	62	3.69	3.190	20.00	1	0	4	2
Merc 230	22.8	4	140.8	95	3.92	3.150	22.90	1	0	4	2
Merc 280	19.2	6	167.6	123	3.92	3.440	18.30	1	0	4	4

Showing 1 to 11 of 32 entries, 11 total columns

Environment | History | Connections

Import Dataset | List | Global Environment

Data

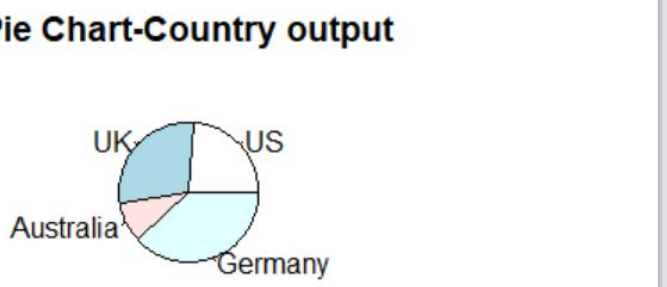
my_mtcars	32 obs. of 11 variables
names	6 obs. of 2 variables
primarycolors	6 obs. of 2 variables

Values

bekind	TRUE
educational_mons…	"Elmo"
favoritedec	3.1416
favoritenum	25

Files | Plots | Packages | Help | Viewer

Zoom | Export | Publish



Console | Terminal | Jobs

```
> # make an intentional error, explain error text and how to use it to problem-so
lve
> mean(myvar)
Error in mean(myvar) : object 'myvar' not found
> # Simple Pie Chart
> slices <- c(10, 12,4, 16)
> lbls <- c("US", "UK", "Australia", "Germany")
> pie(slices, labels = lbls, main="Pie Chart-Country output")
> my_mtcars <- mtcars
> View(my_mtcars)
>
```

RStudio

File Edit Code View Plots Session Build Debug Profile Tools Help

Go to file/function | Addins | Project: (None)

amples.R | nercomp_handout_ 20191028.R | test.Rmd | Untitled2 | r_sample_code.R

Source on Save | Run | Source

```
218 ggplot(mtcars,aes(x=mpg)) + geom_histogram(binwidth=5)
219
220 my_boxplot <- ggplot(mtcars,aes(x=cyl,y=mpg)) + geom_boxplot() + xlab('Cyli
221 my_boxplot
222
223 my_scatplot <- ggplot(mtcars,aes(x=wt,y=mpg)) + geom_point()
224 my_scatplot + xlab('Weight (x 1000lbs)') + ylab('Miles per Gallon') + geom_
225
226
227 my_scatplot <- ggplot(mtcars,aes(x=wt,y=mpg,col=cyl)) + geom_point()
228 my_scatplot + labs(x='Weight (x1000lbs)',y='Miles per Gallon',colour='Numbe
229
230 # faceting facet data - small multiples - views of data
231 my_scatplot <- ggplot(mtcars,aes(x=wt,y=mpg,col=cyl)) + geom_point()
232 my_scatplot + facet_grid(~cyl)
233
234
```

230:1 (Untitled) R Script

Console | Terminal | Jobs

```
> my_boxplot <- ggplot(mtcars,aes(x=cyl,y=mpg)) + geom_boxplot() + xlab('Cylinder
s') + ylab('Miles per Gallon')
> my_boxplot
> my_scatplot <- ggplot(mtcars,aes(x=wt,y=mpg)) + geom_point()
> my_scatplot + xlab('Weight (x 1000lbs)') + ylab('Miles per Gallon') + geom_smoo
th()
`geom_smooth()` using method = 'loess' and formula 'y ~ x'
> my_scatplot <- ggplot(mtcars,aes(x=wt,y=mpg,col=cyl)) + geom_point()
> my_scatplot + labs(x='Weight (x1000lbs)',y='Miles per Gallon',colour='Number of
\n Cylinders')
>
```

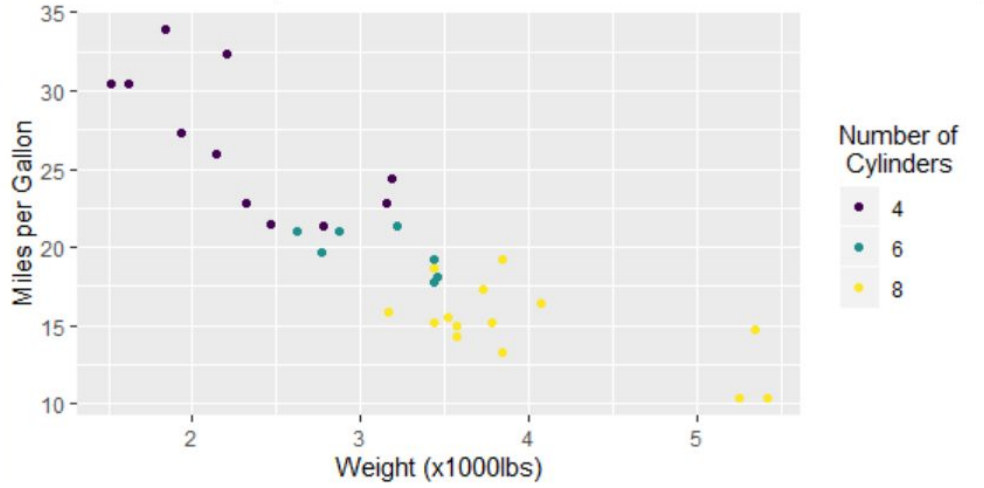
Environment | History | Connections

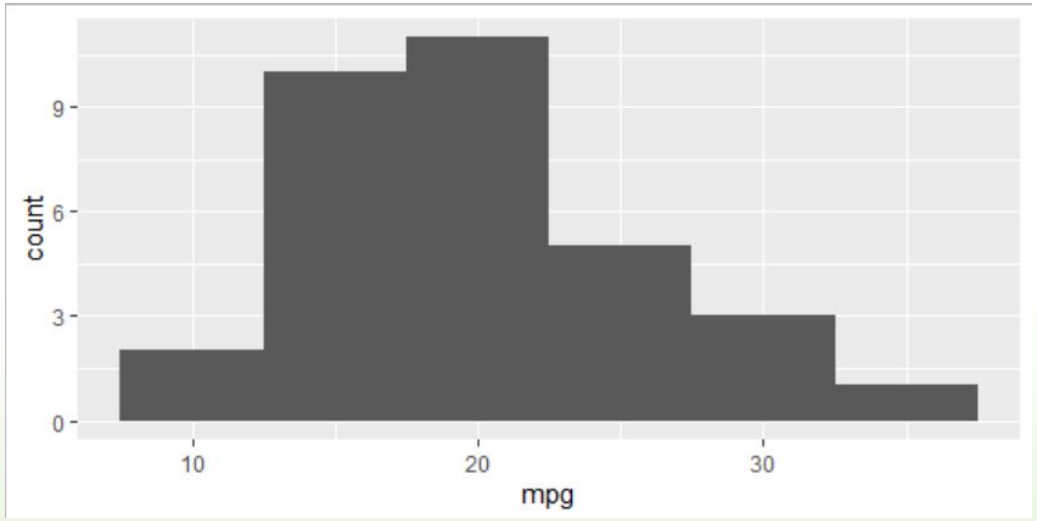
Import Dataset | List | Global Environment

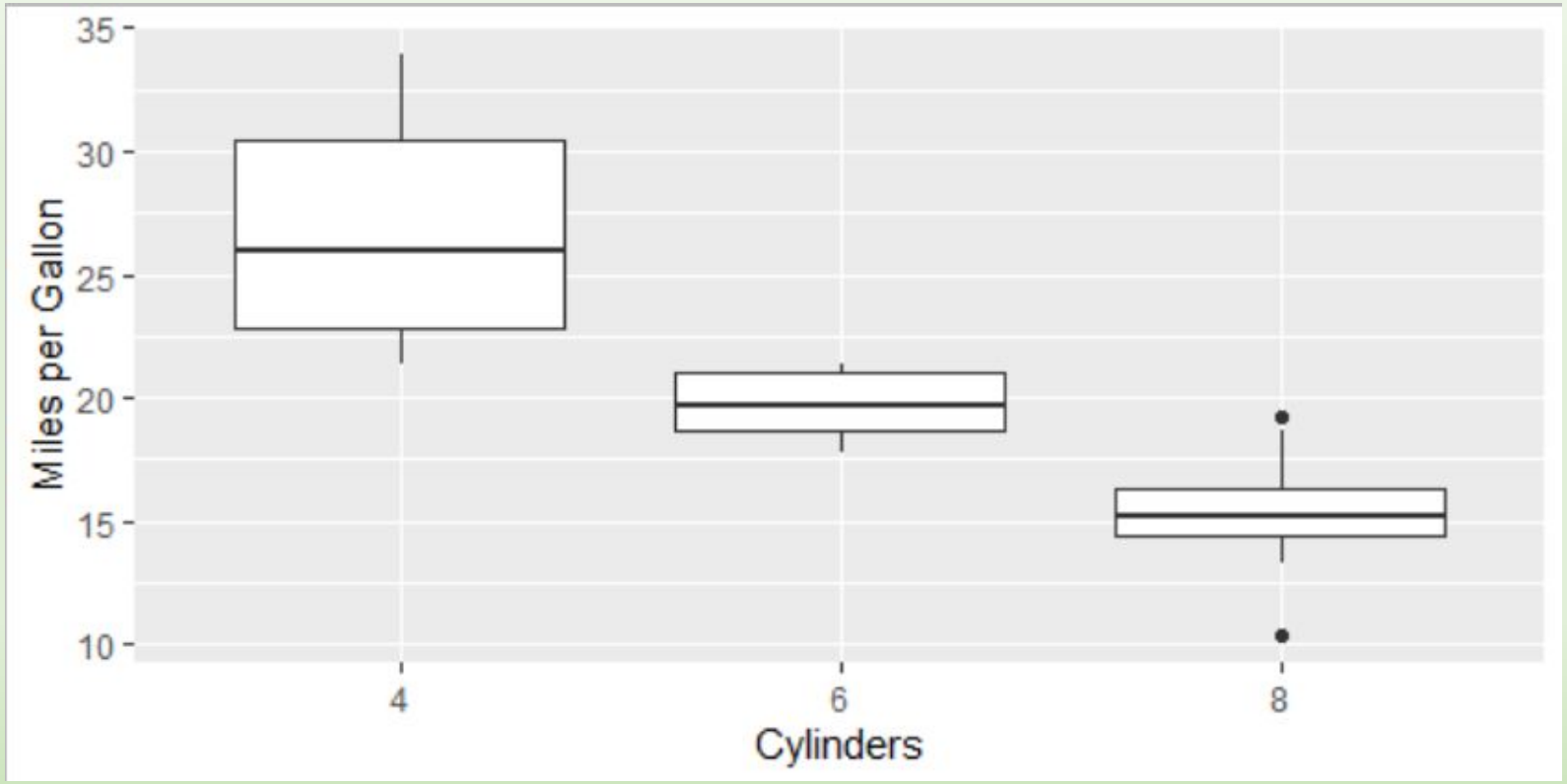
Data	
mtcars	32 obs. of 11 variables
my_boxplot	List of 9
my_scatplot	List of 9

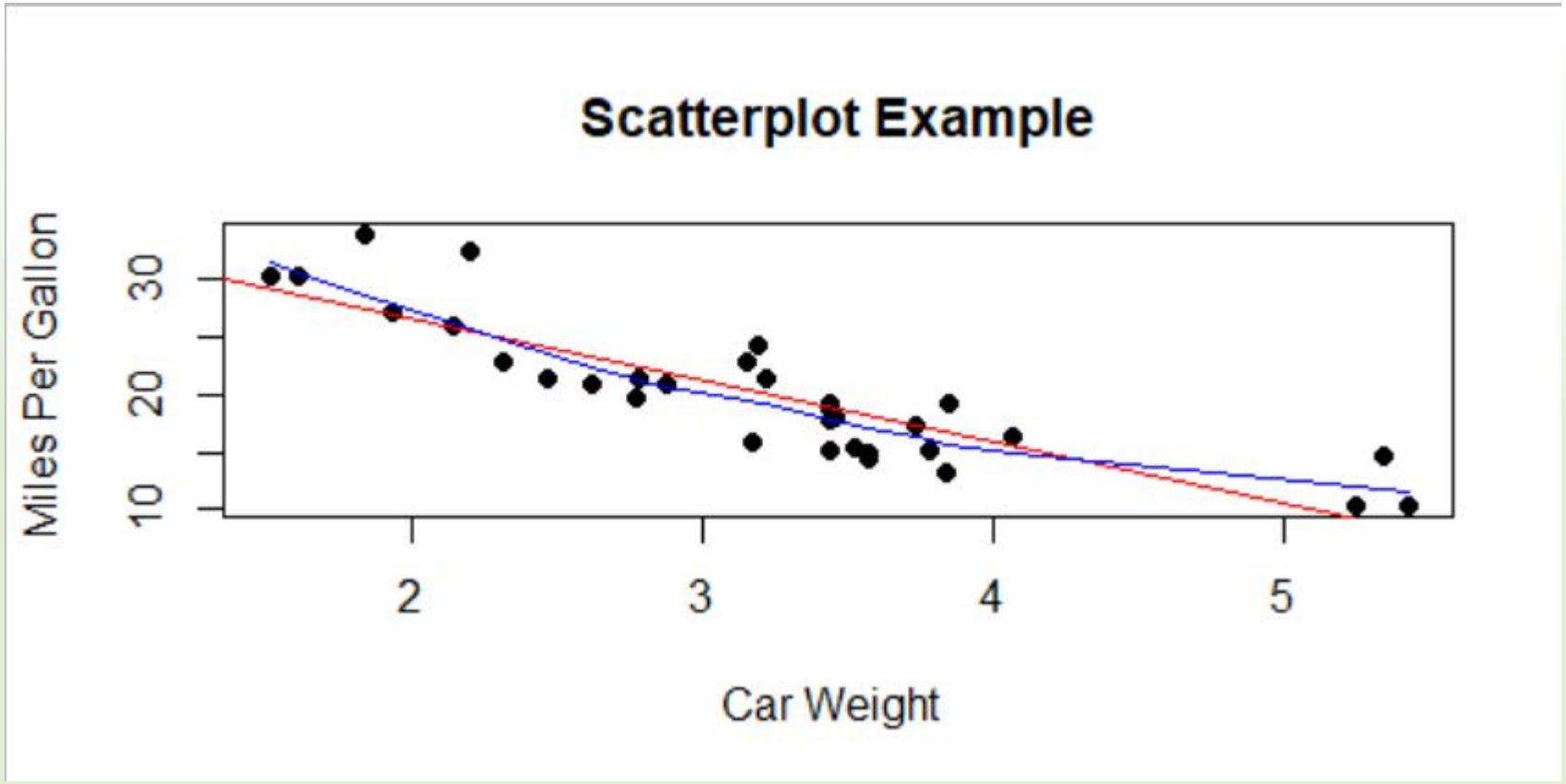
Files | Plots | Packages | Help | Viewer

Zoom | Export | Publish




count
9
6
3
0
10
20
30
mpg


Miles per Gallon
35
30
25
20
15
10
4
6
8
Cylinders


Scatterplot Example
Miles Per Gallon
10
20
30
2
3
4
5
Car Weight

Thank you!

Questions