

# **Intro to R**

**An Introduction to the R Statistical  
Programming Language**

# Goals

By the end of the Presentation, we will have learned:

- What R and RStudio are / are not
- How to start a new R project
- How to write and run your first R script
- How to expand the functionality of R with packages and user-defined functions
- How to load, manipulate, and save your data
- How to visualize your data

# What is R



```
Rterm (64-bit)
R version 3.6.2 (2019-12-12) -- "Dark and Stormy Night"
Copyright (C) 2019 The R Foundation for Statistical Computing
Platform: x86_64-mingw32/x64 (64-bit)

R is free software and comes with ABSOLUTELY NO WARRANTY.
You are welcome to redistribute it under certain conditions.
Type 'license()' or 'licence()' for distribution details.

  Natural language support but running in an English locale

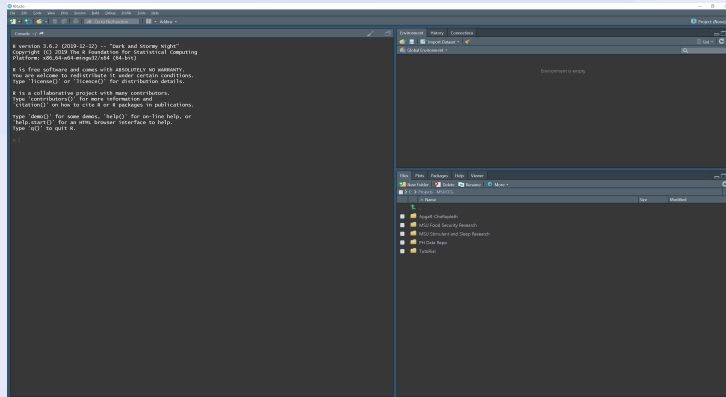
R is a collaborative project with many contributors.
Type 'contributors()' for more information and
'citation()' on how to cite R or R packages in publications.

Type 'demo()' for some demos, 'help()' for on-line help, or
'help.start()' for an HTML browser interface to help.
Type 'q()' to quit R.

> -
```

- Programming Language
- Optimized for data manipulation, data visualization, and statistical tests
- Functionality is expanded by the community (packages) and by the user (user-defined functions)

# What is RStudio



- Integrated Development Environment
- Software that helps users successfully write in R
- Code Completion, Syntax Highlighting, Project Management, Virtual Environment and File Viewer

# What is the Difference?



- Code Interpreter
- Necessary to run any R code
- Interactive Development Environment
- What you will likely write code with

# What R Is and What R Is Not

## What R Is:

- Powerful tool for handling data
- Versatile and customizable
- Can be made to fit your precise needs

## What R Is Not:

- 'Finished'
- Ready to use Software (SPSS, QGIS)

# So What Can R Do?

welch Two Sample t-test

```
data: mtcars$mpg by mtcars$am
t = -3.7671, df = 18.332, p-value = 0.001374
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
 -11.280194 -3.209684
sample estimates:
mean in group 0 mean in group 1
 17.14737      24.39231
```

Terms:

```
as.factor(mtcars$cyl) Residuals
Sum of Squares      824.7846  301.2626
Deg. of Freedom        2        29
```

Residual standard error: 3.223099

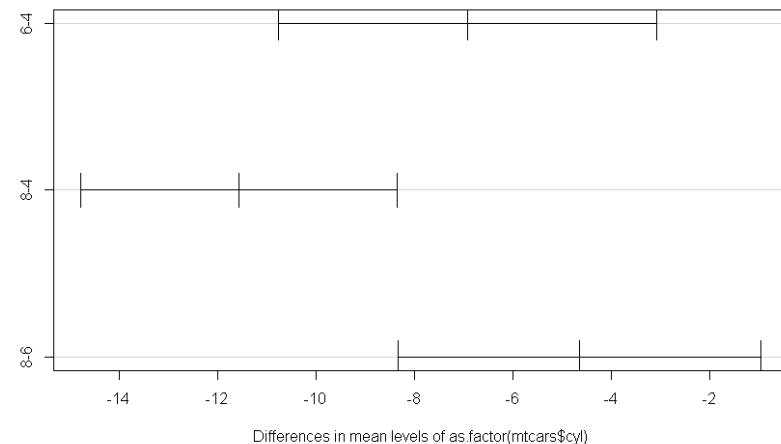
Estimated effects may be unbalanced

```
> summary(aov_examp)
```

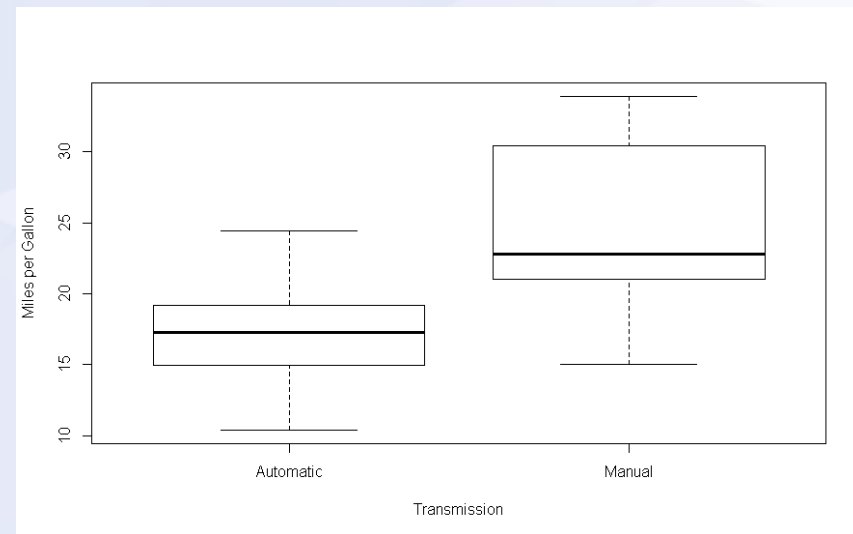
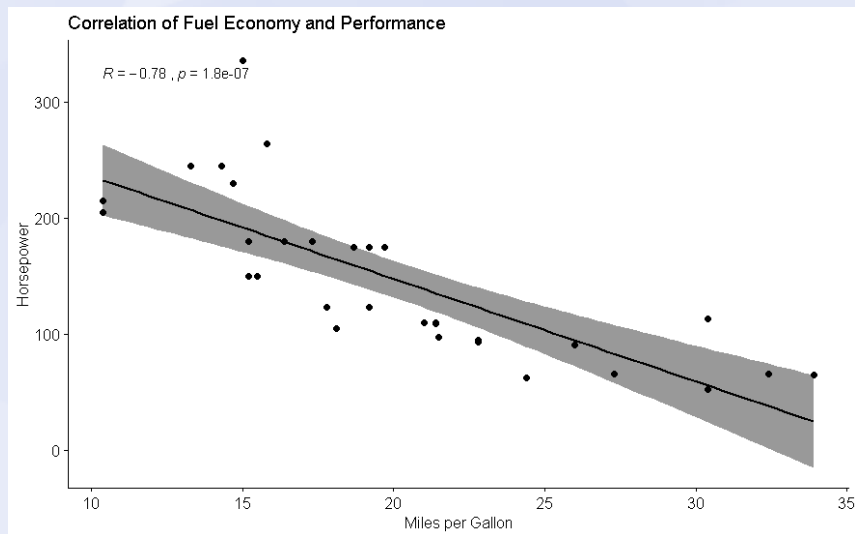
	Df	Sum Sq	Mean Sq	F value	Pr(>F)
as.factor(mtcars\$cyl)	2	824.8	412.4	39.7	4.98e-09 ***
Residuals	29	301.3	10.4		

---  
signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

95% family-wise confidence level



# So What Can R Do?





# So What Can R Do?

## Atlantic County

### Summary of Secondary Data

#### Atlantic County, New Jersey Department of Education 20 Year Summary

Measure	2000	2019	20 Year Change
Student Enrollment	41,906	42,716	810
Students Receiving Free Lunch (N)	13,392	21,363	7,971
Students Receiving Free Lunch (%)	32.0%	50.0%	18.1%
Students Receiving Reduced Price Lunch (N)	3,709	2,866	-843
Students Receiving Reduced Price Lunch (%)	8.9%	6.7%	-2.1%

#### Atlantic County, American Community Survey 8 Year Summary

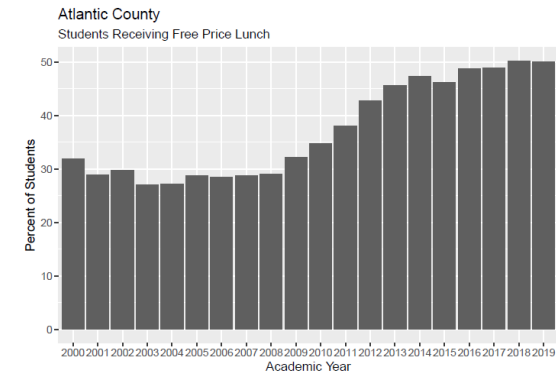
Note: Estimate (N) Data is Suppressed

Measure	2011	2018	8 Year Change
Total Households (Estimate)	101,418	99,874	-1,544
Households receiving Social Security benefits (%)	29.5%	35.2%	5.7%
Households receiving Supplemental Security Income (%)	4.3%	6.4%	2.1%
Households receiving Cash Public Assistance Income (%)	3.3%	2.8%	-.5%
Households receiving SNAP benefits (%)	8.0%	14.0%	6.0%
Households at or below 100% FPL (%)	9.6%	10.9%	1.3%
Households with children under 18 at or below 100% FPL (%)	15.1%	18.3%	3.2%
Households with children under 5 at or below 100% FPL (%)	20.0%	23.3%	3.3%

### Students Receiving Free Lunch

The total number of students receiving free lunches in Atlantic County changed from 13,392 (32.0%) in the Academic Year 2000 to 21,363 (50.0%) in the Academic Year 2019; representing a 20-year change in the number of students receiving free lunches of 7,971 (18.1%). The current rate of students receiving free lunch (50.0%) is greater than the State's current average (31.6%).

The percentage of students each year receiving free lunch is graphed below.



# So What Can R Do?

Atlantic County, New Jersey, 2011



The percentage of households receiving Supplemental Nutrition Assistance Program (SNAP) Benefits, by census tract, is mapped below.

Atlantic County, New Jersey, 2011



Atlantic County, New Jersey, 2018



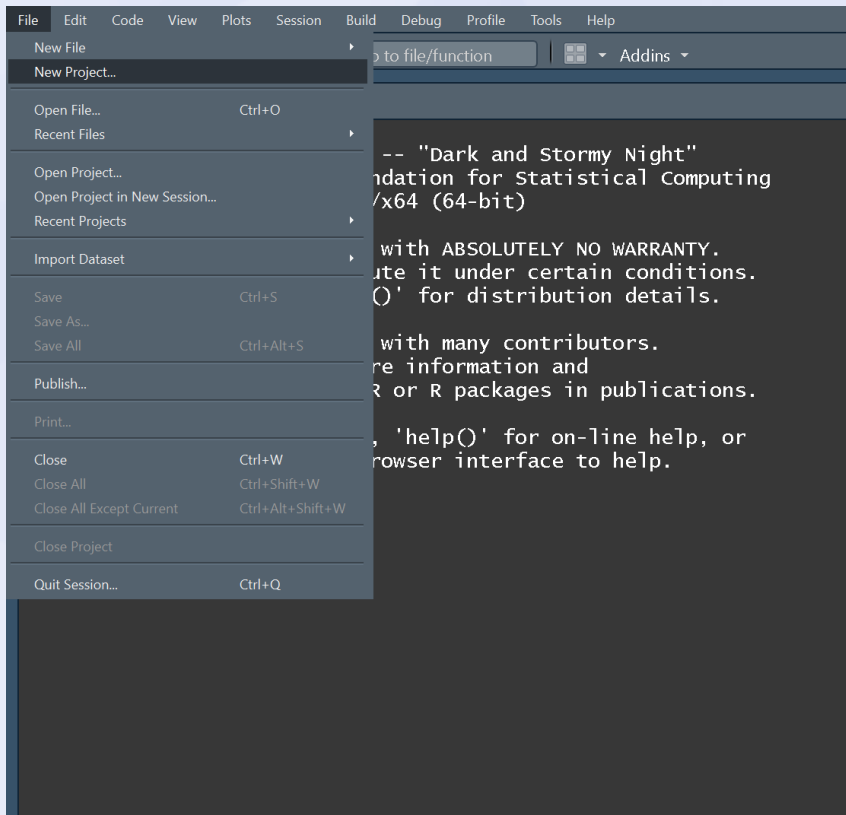
# Getting Started

- ~~What R and RStudio are / are not~~
- How to start a new R project
- How to write and run your first R script
- How to expand the functionality of R with packages and user-defined functions
- How to load, manipulate, and save your data
- How to visualize your data

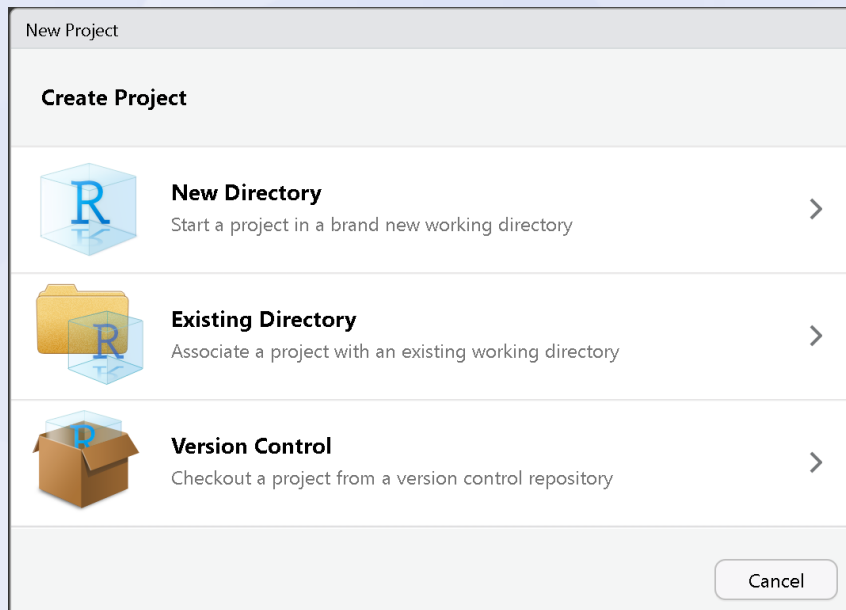
# Getting Started

## Starting a new Project

- Open RStudio
- File > New Project



# Getting Started



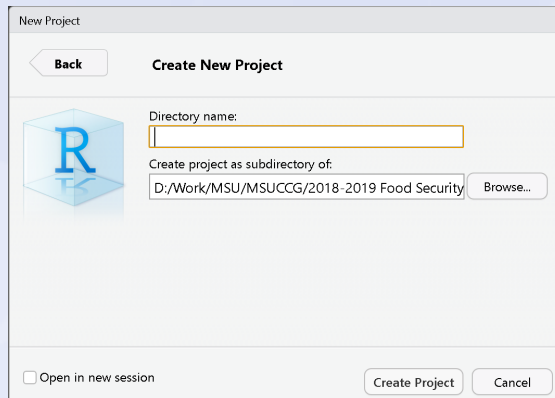
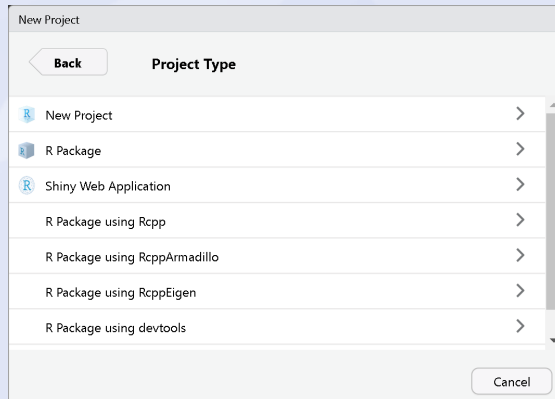
- > New Directory  
or
- > Existing Directory

*Note:*  
*Directory == Folder*

# Getting Started: New Directory

If starting a project in  
a new directory:

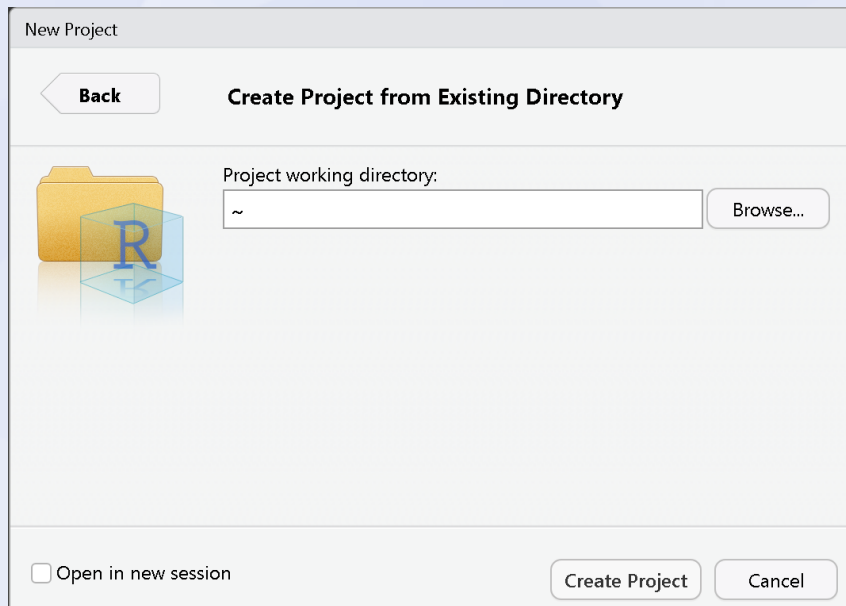
- New Project >
- Browse



# Getting Started: Existing Directory

If starting a project in  
an existing directory:

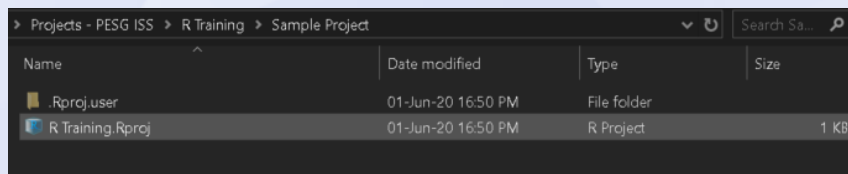
- Browse



# Getting Started: Project Directory

Starts with:

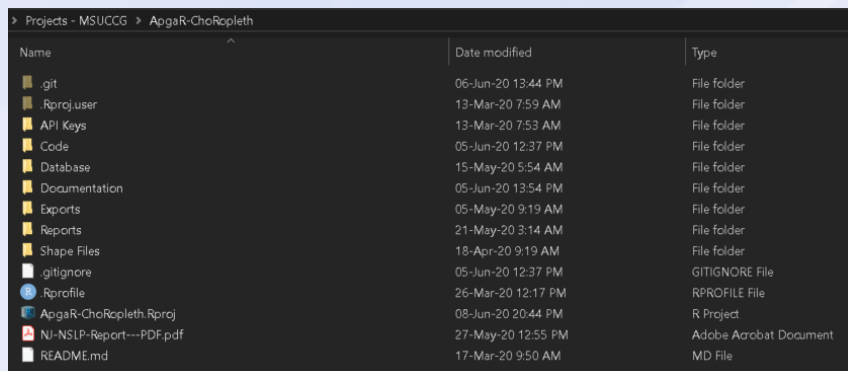
- `.Rproj`



Name	Date modified	Type	Size
.Rproj.user	01-Jun-20 16:50 PM	File folder	
R Training.Rproj	01-Jun-20 16:50 PM	R Project	1 KB

Should Contain:

- User Code
- User Data



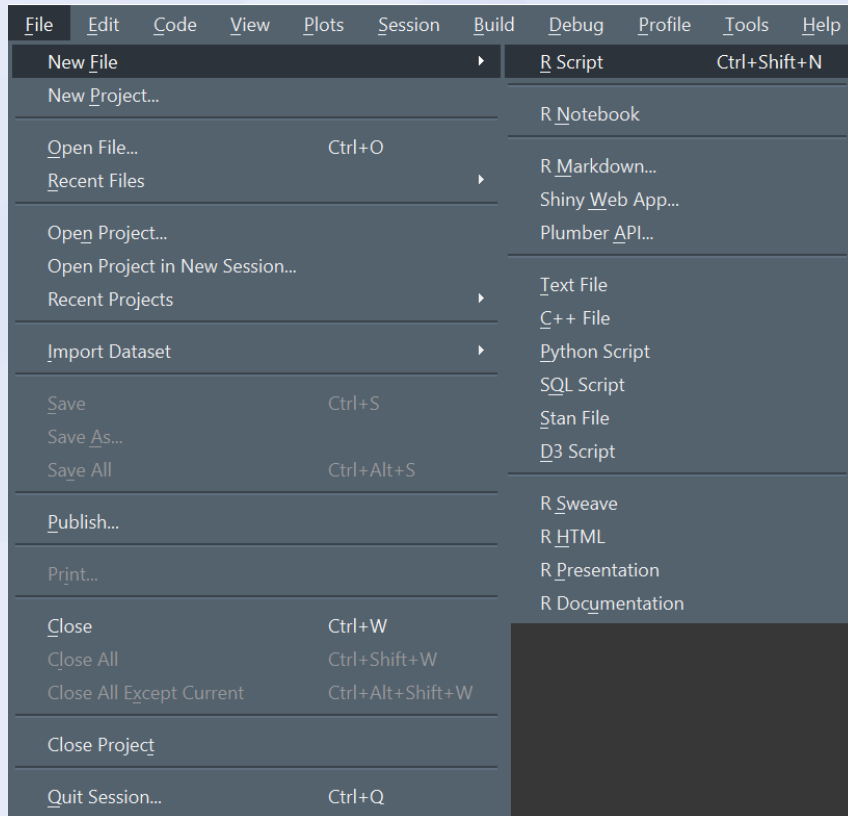
Name	Date modified	Type
.git	06-Jun-20 13:44 PM	File folder
.Rproj.user	13-Mar-20 7:59 AM	File folder
API Keys	13-Mar-20 7:53 AM	File folder
Code	05-Jun-20 12:37 PM	File folder
Database	15-May-20 5:54 AM	File folder
Documentation	05-Jun-20 13:54 PM	File folder
Exports	05-May-20 9:19 AM	File folder
Reports	21-May-20 3:14 AM	File folder
Shape Files	18-Apr-20 9:19 AM	File folder
.gitignore	05-Jun-20 12:37 PM	GITIGNORE File
.Rprofile	26-Mar-20 12:17 PM	RPROFILE File
ApgaR-Choropleth.Rproj	08-Jun-20 20:44 PM	R Project
NJ-NSLP-Report---PDF.pdf	27-May-20 12:55 PM	Adobe Acrobat Document
README.md	17-Mar-20 9:50 AM	MD File



# Writing in R

- ~~What R and RStudio are / are not~~
- ~~How to start a new R project~~
- How to write and run your first R script
- How to expand the functionality of R with packages and user-defined functions
- How to load, manipulate, and save your data
- How to visualize your data

# Writing in R



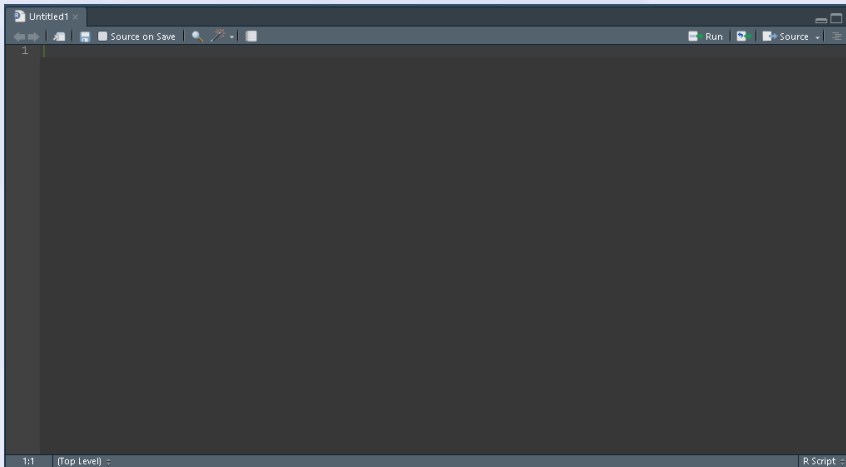
## Starting a New Script

- Open Your New Project
- File > New File > R Script

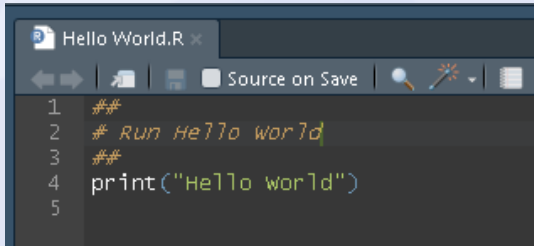
# Writing in R

## Your First .R File

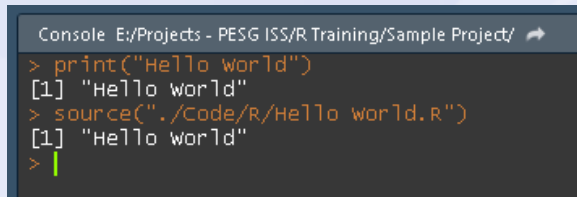
- Opens in Rstudio
- Can edit in any Plain-Text Editor, as needed



# Writing in R



```
Hello World.R x
Source on Save
1 ##
2 # Run Hello world
3 ##
4 print("Hello world")
5
```



```
Console E:/Projects - PESG ISS/R Training/Sample Project/
> print("Hello world")
[1] "Hello world"
> source("../Code/R/Hello world.R")
[1] "Hello world"
> |
```

- Write your instructions for R
- Run or Source your Code
- View Results in Console

*Note:*

*Run == Execute one or more line of code*

*Source == Execute entire contents of a script*

# Writing in R

```
##  
# Prepare the R Project Workspace  
##  
source("../Code/R/Boot/Load Packages.R") # Loads Packages  
source("../Code/R/Boot/Load Functions.R") # Loads User Defined Functions  
source("../Code/R/Boot/Load API Key.R") # Loads API Keys to the global environment  
writeLines(sprintf("%s %s READY:", Sys.time(), getwd()))
```

```
Console E:/Projects - MSUCCG/Apgar-Choropleth/ ↗  
2020-06-08 20:44:24 LOADED: R Packages  
2020-06-08 20:44:24 LOADED: Project Audit Tools  
2020-06-08 20:44:24 LOADED: Codex Query Tools  
2020-06-08 20:44:24 LOADED: dBase Query Tools  
2020-06-08 20:44:24 LOADED: Database - ACS Tools  
2020-06-08 20:44:24 LOADED: Database - NJ DOE Tools  
2020-06-08 20:44:24 LOADED: Database - NJ DOT Tools  
2020-06-08 20:44:24 LOADED: Shape Files - TIGER/Line Tools  
2020-06-08 20:44:24 LOADED: Google Maps API Tools  
2020-06-08 20:44:24 LOADED: Statistical Analysis Tools  
2020-06-08 20:44:24 LOADED: ACS Report Generator Tool  
2020-06-08 20:44:24 LOADED: API Keys  
2020-06-08 20:44:24 E:/Projects - MSUCCG/Apgar-Choropleth READY:
```

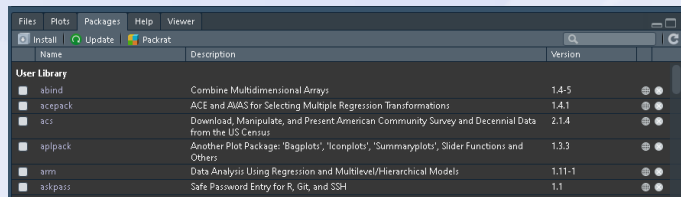
- You can use scripts to make complex lists of instructions
- Scripts can run other scripts
- Automate repetitive processes

# Writing in R

- ~~What R and RStudio are / are not~~
- ~~How to start a new R project~~
- ~~How to write and run your first R script~~
- How to expand the functionality of R with packages and user-defined functions
- How to load, manipulate, and save your data
- How to visualize your data

# Adding to R: Packages

```
##  
# Packages for reading from and writing to .dbf, xls, xlsx, ; shape file, file formats  
##  
library(foreign)  
library(shapefiles)  
  
library(readxl)  
library(openxlsx)  
  
library(sf)  
library(sp)
```



The screenshot shows the R Package Manager window with the 'User Library' tab selected. It displays a table of installed packages with columns for Name, Description, and Version. The packages listed are abind, acepack, acs, aplpack, arm, and askpass.

Name	Description	Version
abind	Combine Multidimensional Arrays	1.4-5
acepack	ACE and AIAS for Selecting Multiple Regression Transformations	1.4.1
acs	Download, Manipulate, and Present American Community Survey and Decennial Data from the US Census	2.1.4
aplpack	Another Plot Package: 'Bagplots', 'Iconplots', 'Summaryplots', 'Slider Functions and Others	1.3.3
arm	Data Analysis Using Regression and Multilevel/Hierarchical Models	1.11-1
askpass	Safe Password Entry for R, Git, and SSH	1.1

- Packages add to the base functionality of R
- Authored by other professionals
- Traditionally open-source

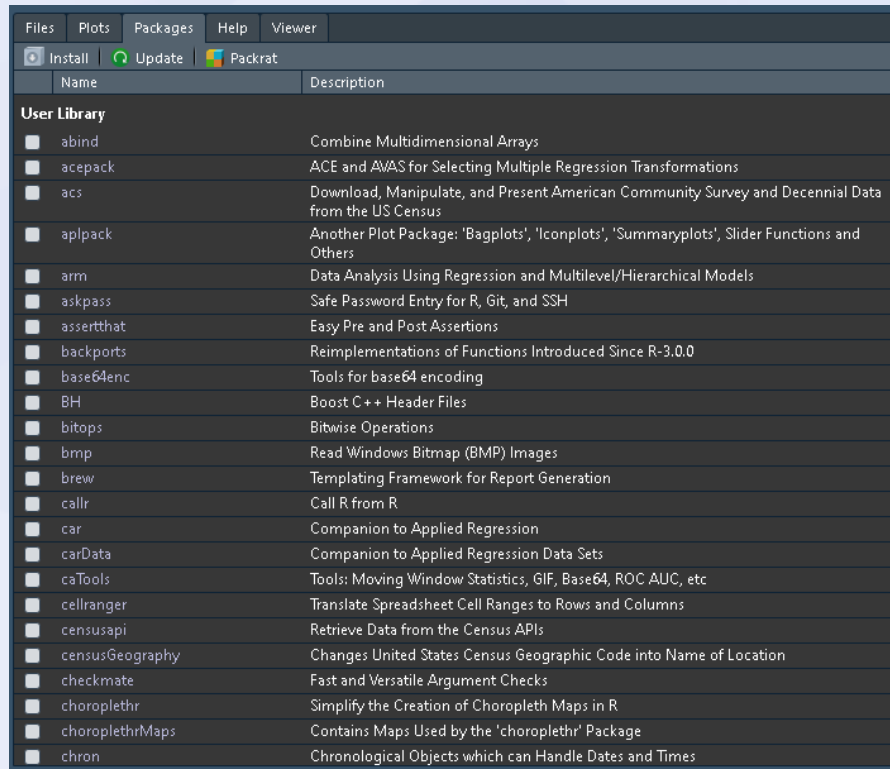
# Adding to R: Packages

```
##  
# Packages for reading from and writing to .dbf; xls, xlsx, ; shape file, file formats  
##  
library(foreign)  
library(shapefiles)  
  
library(readxl)  
library(openxlsx)  
  
library(sf)  
library(sp)  
  
##  
# Packages for sorting, filtering, manipulating, Data and Data Frames  
##  
library(dplyr)  
library(tidyr)  
library(numform)  
library(tidyselect)  
library(data.table)  
  
##  
# Package for creating static and interactive maps  
##  
library(tmap)  
  
##  
# Packages for API and WWW interfaces, web scraping  
##  
library(httr)  
library(jsonlite)  
library(xml2)  
library(rvest)  
  
##  
# Packages for creation of graphics  
##  
library(ggpubr)  
library(ggplot2)  
  
##  
# Packages for report generation  
##  
library(knitr)  
library(rmarkdown)
```

- The library command is used to load any package that you have downloaded
- Recommended packages listed here



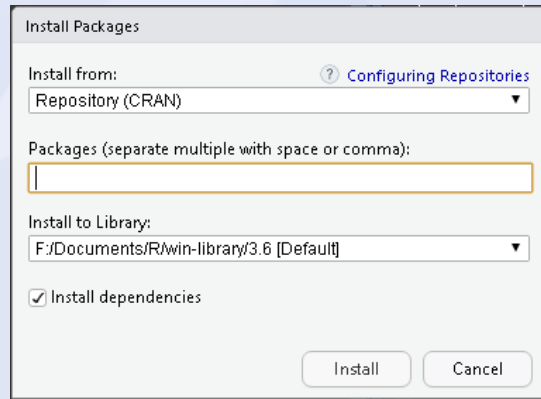
# Adding to R: Packages



	Name	Description
<b>User Library</b>		
<input type="checkbox"/>	abind	Combine Multidimensional Arrays
<input type="checkbox"/>	acepack	ACE and AVAS for Selecting Multiple Regression Transformations
<input type="checkbox"/>	acs	Download, Manipulate, and Present American Community Survey and Decennial Data from the US Census
<input type="checkbox"/>	aplpack	Another Plot Package: 'Bagplots', 'Iconplots', 'Summaryplots', Slider Functions and Others
<input type="checkbox"/>	arm	Data Analysis Using Regression and Multilevel/Hierarchical Models
<input type="checkbox"/>	askpass	Safe Password Entry for R, Git, and SSH
<input type="checkbox"/>	assertthat	Easy Pre and Post Assertions
<input type="checkbox"/>	backports	Reimplementations of Functions Introduced Since R-3.0.0
<input type="checkbox"/>	base64enc	Tools for base64 encoding
<input type="checkbox"/>	BH	Boost C++ Header Files
<input type="checkbox"/>	bitops	Bitwise Operations
<input type="checkbox"/>	bmp	Read Windows Bitmap (BMP) Images
<input type="checkbox"/>	brew	Templating Framework for Report Generation
<input type="checkbox"/>	callr	Call R from R
<input type="checkbox"/>	car	Companion to Applied Regression
<input type="checkbox"/>	carData	Companion to Applied Regression Data Sets
<input type="checkbox"/>	caTools	Tools: Moving Window Statistics, GIF, Base64, ROC AUC, etc
<input type="checkbox"/>	cellranger	Translate Spreadsheet Cell Ranges to Rows and Columns
<input type="checkbox"/>	censusapi	Retrieve Data from the Census APIs
<input type="checkbox"/>	censusGeography	Changes United States Census Geographic Code into Name of Location
<input type="checkbox"/>	checkmate	Fast and Versatile Argument Checks
<input type="checkbox"/>	choroplethr	Simplify the Creation of Choropleth Maps in R
<input type="checkbox"/>	choroplethrMaps	Contains Maps Used by the 'choroplethr' Package
<input type="checkbox"/>	chron	Chronological Objects which can Handle Dates and Times

- A list of downloaded packages is available in the Packages tab

# Adding to R: Packages



```
##  
# Scripts allow you to have  
# more complicated instructions  
# Or instructions in order  
# such as installing and loading  
# a new package  
##  
install.packages("ggplot2")  
library(ggplot2)
```

- New packages can be installed using RStudio's search window
- Packages can be installed manually

# Adding to R: Functions

- Functions are used to create variable and /or repeatable instructions

```
##  
# Correlations Tests  
##  
correlation = function(x, y) {  
  ##  
  # Build Temporary Data Frame  
  # Bypasses the limits of ggscatter  
  # Also allows for secondary data cleaning  
  ##  
  cor_df = data.frame(x, y) # Assign temp data frame from passed vectors  
  names(cor_df) = c("X", "Y") # rename them to 'X' and 'Y', for reference  
  ##  
  # Run Tests  
  ##  
  # Correlation Test using only matched pairs, in case of missing data.  
  cor_test = cor.test(cor_df$x, cor_df$y, use = "complete.obs")  
  cor_plot = ggscatter(cor_df, "X", "Y", add = "reg.line", conf.int = T, cor.coef = T)  
  ##  
  # Return Tests  
  ##  
  print(cor_test)  
  cor_plot  
}
```

*Note:*

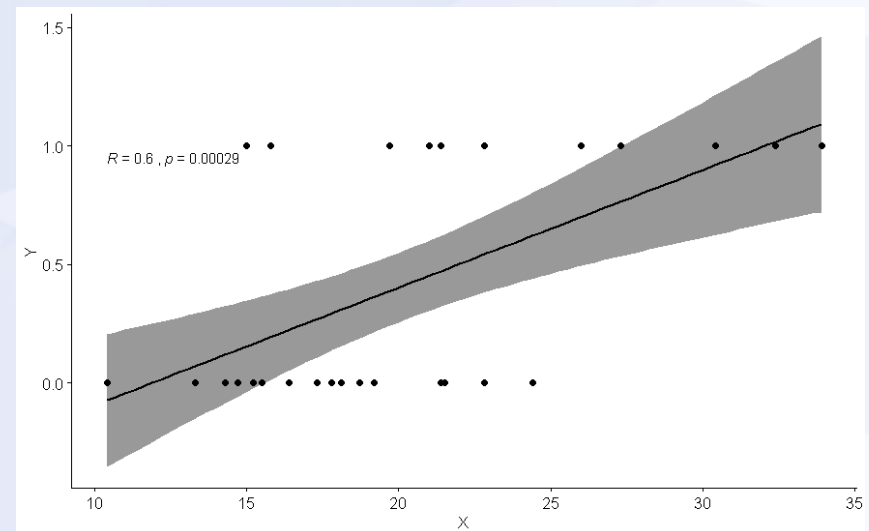
*AKA User-Defined  
Functions*

# Adding to R: Functions

```
> correlation(mtcars$mpg, mtcars$am)

Pearson's product-moment correlation

data: cor_df$X and cor_df$Y
t = 4.1061, df = 30, p-value = 0.000285
alternative hypothesis: true correlation is not equal to 0
95 percent confidence interval:
 0.3175583 0.7844520
sample estimates:
      cor 
0.5998324
```



# Adding to R: Functions

Functions	
animate_acs_tract	function (state, profile, geosub, measure, column, delay = 100, loop = TRU...
batch_acs_tract	function (state, profile)
build_acs_tract	function (state, profile, year, measure)
check_acs_dBase	function (profile, year)
check_acs_tract	function (state, profile, year, measure)
check_tiger_county	function (year)
check_tiger_municipality	function (state, year)
check_tiger_tract	function (state, year)
correlation	function (x, y)
fetch_tiger_county	function (year)
fetch_tiger_municipality	function (state, year)
fetch_tiger_tract	function (state, year)
load_acs_county_dBase	function (state, profile, year)
load_acs_dBase	function (state, profile, year)
load_acs_state_dBase	function (state, profile, year)
load_acs_tract	function (state, profile, year, measure)
load_from_dBase	function (source, data, year)
load_njdoe_dBase	function (report, year)
load_njdot_dBase	function (report, year)
load_njogis_county	function ()
load_njogis_municipality	function ()
load_tiger_county	function (state, year)
load_tiger_tract	function (state, year)
map_acs_tract	function (state, profile, year, geosub = "state", measure = "est", column ...
mean_diff	function (x, y)
query_acs_metadata	function (profile, year, query)
query_county_dBase	function (state, query)
query_gmap_distance	function (origin, destination)
query_gmap_place	function (place)
query_municipality_dBase	function (state, query)
query_state_dBase	function (state, substate)
save_to_dBase	function (df, name)

- All User-Defined Functions that are loaded are visible inside the Project Environment
- All of these functions are your own personal R package

# Adding to R: Functions

- All User-Defined Functions are written as scripts
- To load them, source the script that defines the function

```
##
# Project Management Functions ----
##
# Project Audit Functions
##
source("../Code/R/Boot/Cleanup Project.R")
writeLines(sprintf("%s LOADED: Project Audit Tools", Sys.time()))
##
# Codex Query Functions
# Local Functions and Single web Function for determining FIPS codes and GEOID values
##
source("../Code/R/Database/Codex/Query State dBase.R")
source("../Code/R/Database/Codex/Query County dBase.R")
source("../Code/R/Database/Codex/Query Municipality dBase.R")
writeLines(sprintf("%s LOADED: Codex Query Tools", Sys.time()))
##
# dBase Import/Export Functions
# Local Functions for sending data to and retrieving data from dBase structure
##
source("../Code/R/Database/Save to dBase.R")
source("../Code/R/Database/Load from dBase.R")
writeLines(sprintf("%s LOADED: dBase Query Tools", Sys.time()))
```

# Writing in R

- ~~What R and RStudio are / are not~~
- ~~How to start a new R project~~
- ~~How to write and run your first R script~~
- ~~How to expand the functionality of R with packages and user-defined functions~~
- How to load, manipulate, and save your data
- How to visualize your data

# Loading Your Data

- R can handle data from most common file formats
- Natively supports .csv, .txt
- Packages support dBase, Excel, JSON, SAS, SPSS, and more
- Can load data saved to your computer, or available on a network

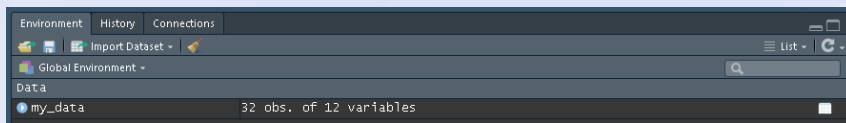


# Loading Your Data

```
my_data = read.csv("../data/mtcars_data.csv")
```

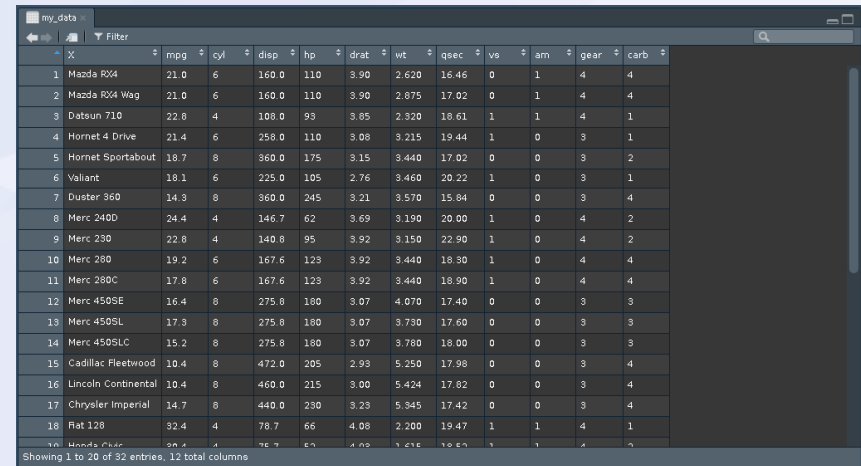
```
load_acs_data = function(state, profile, year) {  
  ##  
  # Generate State Variables from the passed state argument  
  # Passes the state argument to the fetch script  
  # Returns the appropriate ANSI, USPS, and NAME of the queried state  
  ##  
  state_query = query_state_dbase(state)  
  state_fips = state_query[1,1] # 2 Digit ID Number  
  state_usps = state_query[1,2] # 2 Character State Name Abbreviation  
  state_name = state_query[1,3] # Full State Name  
  ##  
  # Read the Census Data Table into the virtual environment  
  ##  
  head_data_csv = read.csv(sprintf("../PH Data Repo/US Census/ACS/%s/%s/ACSDP5Y%s.csv", toupper(profile), year, year),  
    as.is = TRUE, header = TRUE, as.data.frame = FALSE)  
  body_data_csv = read.csv(sprintf("../PH Data Repo/US Census/ACS/%s/%s/ACSDP5Y%s.csv", toupper(profile), year, year),  
    as.is = TRUE, header = TRUE, as.data.frame = FALSE)  
  head_data_csv[1] = "GEO_id" # Renames vector 1 to GEO_id  
  head_data_csv[2] = "GEO_Name" # Renames vector 2 to GEO_Name  
  colnames(body_data_csv) = head_data_csv #Assigns the appropriate Vector Names as the column names of the data  
  acs_data = body_data_csv #Assigns a new object for easier reference  
  setorder(acs_data, GEO_id) #Sets order of the new object by renamed GEO_id vector  
  ##  
  # Data Clean up to Remove unwanted Margin of Error Vectors  
  ##  
  head_acs_data = names(acs_data) # Temp variable for storing names of the state variable, for name manipulation  
  head_acs_data = gsub(sprintf("%s_0", toupper(profile)), "X", head_acs_data) # Removes data profile ID Code from variable  
  head_acs_data = gsub("PE", "PER", head_acs_data, fixed = T) # Renames Percent Estimate Code to 'Per'  
  head_acs_data = gsub("E", "EST", head_acs_data, fixed = T) # Renames Estimate Code to 'Est'  
  head_acs_data[1] = "GEO_id" # Redefine Vector Name, because of overwriting of above gsub function  
  head_acs_data[2] = "GEO_Name" # Redefine Vector Name, because of overwriting of above gsub function  
  names(acs_data) = head_acs_data # Injects names from temp variable into the state variable  
  acs_data = acs_data %>% select(-ends_with("M")) # Removes all Margin of Error and Percent Margin of Error Vectors, w  
  ##  
  # Check to see if a state argument was passed  
  # If a state argument is passed, only the defined state will be loaded  
  # If no argument is passed, loads the entire data table  
  ##  
  if (missing(state) == F) {  
    if (state_usps != "US") {  
      state_acs_data = acs_data[grepl(sprintf("US%s", state_fips), acs_data$GEO_id),] # Uses the substring to sort the da  
      writeLines(sprintf("%s LOADED: Census Data Table %s from the %s ACS for the state of %s.", Sys.time(), profile, year, state_name))  
      return(state_acs_data) # If using a state variable, returns the data frame built with only one state  
    } else {  
      writeLines(sprintf("%s LOADED: Census Data Table %s from the %s ACS.", Sys.time(), profile, year)) # If using only
```

# Loading Your Data



The Environment pane in RStudio displays the following information:

Environment	History	Connections
Global Environment		
Data		
my_data	32 obs. of 12 variables	



The Data Viewer displays the first 20 rows of the dataset. The columns are: X, mpg, cyl, disp, hp, drat, wt, qsec, vs, am, gear, carb.

X	mpg	cyl	disp	hp	drat	wt	qsec	vs	am	gear	carb
1 Mazda RX4	21.0	6	160.0	110	3.90	2.620	16.46	0	1	4	4
2 Mazda RX4 Wag	21.0	6	160.0	110	3.90	2.875	17.02	0	1	4	4
3 Datsun 710	22.8	4	108.0	93	3.85	2.320	16.61	1	1	4	1
4 Hornet 4 Drive	21.4	6	258.0	110	3.08	3.215	19.44	1	0	3	1
5 Hornet Sportabout	18.7	8	360.0	175	3.15	3.440	17.02	0	0	3	2
6 Valiant	18.1	6	225.0	105	2.76	3.460	20.22	1	0	3	1
7 Duster 360	14.3	8	360.0	245	3.21	3.570	15.84	0	0	3	4
8 Merc 240D	24.4	4	146.7	62	3.69	3.190	20.00	1	0	4	2
9 Merc 230	22.8	4	140.8	95	3.92	3.150	22.90	1	0	4	2
10 Merc 280	19.2	6	167.6	123	3.92	3.440	18.30	1	0	4	4
11 Merc 280C	17.8	6	167.6	123	3.92	3.440	18.90	1	0	4	4
12 Merc 450SE	16.4	8	275.8	180	3.07	4.070	17.40	0	0	3	3
13 Merc 450SL	17.3	8	275.8	180	3.07	3.730	17.60	0	0	3	3
14 Merc 450SLC	15.2	8	275.8	180	3.07	3.780	18.00	0	0	3	3
15 Cadillac Fleetwood	10.4	8	472.0	205	2.93	5.250	17.98	0	0	3	4
16 Lincoln Continental	10.4	8	460.0	215	3.00	5.424	17.02	0	0	3	4
17 Chrysler Imperial	14.7	8	440.0	230	3.23	5.345	17.42	0	0	3	4
18 Fiat 128	32.4	4	78.7	66	4.08	2.200	19.47	1	1	4	1
19 Honda Civic	30.4	4	75.7	62	4.08	1.615	16.86	1	1	4	2

Showing 1 to 20 of 32 entries, 12 total columns

# Saving Your Data

- Just like loading data, R can save data to most common file formats
- Natively supports .csv, .txt
- Packages support dBase, Excel, JSON, SAS, SPSS, and more
- Can save data to your computer, or save to a device available on a network

# Saving Your Data

```
##  
# Save your data to a specific place|  
##  
write.csv(mtcars, "./Data/mtcars_data.csv")
```

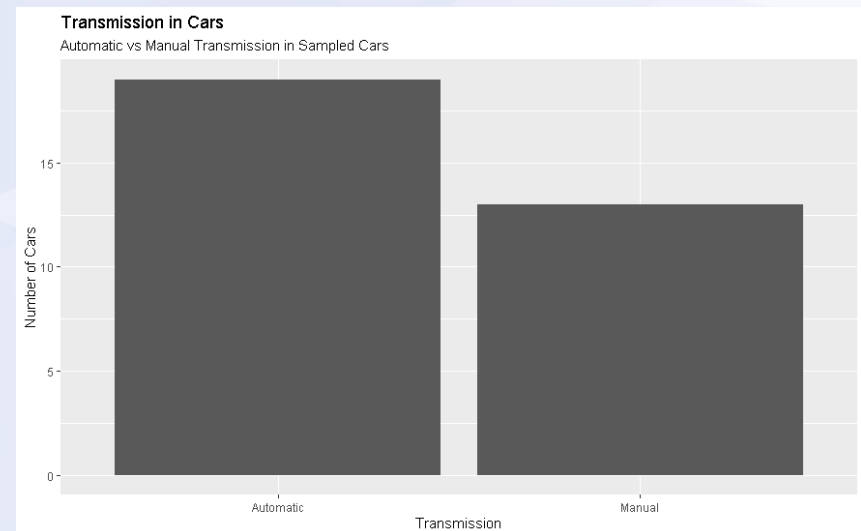
- Almost all packages that read a file format will let you save in that format
- Some file formats are easier to use than others

# Writing in R

- ~~What R and RStudio are / are not~~
- ~~How to start a new R project~~
- ~~How to write and run your first R script~~
- ~~How to expand the functionality of R with packages and user-defined functions~~
- ~~How to load, manipulate, and save your data~~
- ~~How to visualize your data~~

# Visualizing Your Data

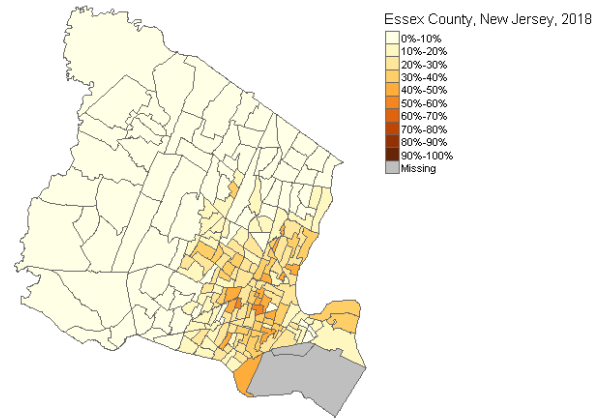
```
mtcars$am_text = gsub(0, "Automatic", mtcars$am)
mtcars$am_text = gsub(1, "Manual", mtcars$am_text)
ggplot(mtcars, aes(am_text)) + geom_bar() +
  labs(x = "Transmission", # The X Axis Label
       y = "Number of Cars", # The Y Axis Label
       title = "Transmission in Cars", # The Bolded Graph Title
       subtitle = "Automatic vs Manual Transmission in Sampled Cars") # The all Graph Title
```



# Visualizing Your Data

```
##  
# Plot Map  
# State Level Estimate  
##  
if (toupper(measure) == "EST" & map_of == "state") {  
  tmap_object = tm_shape(tmap_data) +  
    tm_polygons(col = column,  
               title = sprintf("%s, %s", state_name, year)) +  
    tm_legend(legend.outside = T,  
              legend.outside.position = "right") +  
    tm_layout(frame = F,  
              main.title = acs_name,  
              main.title.size = .9  
              )  
}
```

Percent Estimate,  
INCOME AND BENEFITS (IN 2018 INFLATION-ADJUSTED DOLLARS),  
Total households,  
With Food Stamp/SNAP benefits in the past 12 months



# Visualizing Your Data

```
1 title:
2   '[](./Reports/Logos/MSU_Crest_Black.png){width=3in}'
3
4   New Jersey Student Hunger Audit
5 author:
6   - Chris Snyder, MPH [Montclair State University, snyderc@montclair.edu]
7   - Kelsey Bent, MPH, CHES
8 output: pdf_document
9
10 ----
11
12 [r Prepare workspace, message=FALSE, warning=FALSE, include=FALSE]
13 ##
14 ## Prepare workspace
15 ## Do Not Source Prepare workspace ; It clears the Params Variable
16 ## Rprofile is run at startup
17 tmap_mode("plot") # Set Map View Mode for Report
18
19 [r Load data, message=FALSE, warning=FALSE, include=FALSE]
20 ##
21 ## NJ DOE
22 ## Load the dBase Files
23 ## Assign a year column
24
25 ay2000 = load_njdoe_dbase("Free_Reduced_Lunch", "2000"); ay2000$year = "2000"; setorder(ay2000, count_Lab1)
26 ay2001 = load_njdoe_dbase("Free_Reduced_Lunch", "2001"); ay2001$year = "2001"; setorder(ay2001, count_Lab1)
27 ay2002 = load_njdoe_dbase("Free_Reduced_Lunch", "2002"); ay2002$year = "2002"; setorder(ay2002, count_Lab1)
28 ay2003 = load_njdoe_dbase("Free_Reduced_Lunch", "2003"); ay2003$year = "2003"; setorder(ay2003, count_Lab1)
29 ay2004 = load_njdoe_dbase("Free_Reduced_Lunch", "2004"); ay2004$year = "2004"; setorder(ay2004, count_Lab1)
30 ay2005 = load_njdoe_dbase("Free_Reduced_Lunch", "2005"); ay2005$year = "2005"; setorder(ay2005, count_Lab1)
31 ay2006 = load_njdoe_dbase("Free_Reduced_Lunch", "2006"); ay2006$year = "2006"; setorder(ay2006, count_Lab1)
32 ay2007 = load_njdoe_dbase("Free_Reduced_Lunch", "2007"); ay2007$year = "2007"; setorder(ay2007, count_Lab1)
33 ay2008 = load_njdoe_dbase("Free_Reduced_Lunch", "2008"); ay2008$year = "2008"; setorder(ay2008, count_Lab1)
34 ay2009 = load_njdoe_dbase("Free_Reduced_Lunch", "2009"); ay2009$year = "2009"; setorder(ay2009, count_Lab1)
35 ay2010 = load_njdoe_dbase("Free_Reduced_Lunch", "2010"); ay2010$year = "2010"; setorder(ay2010, count_Lab1)
36 ay2011 = load_njdoe_dbase("Free_Reduced_Lunch", "2011"); ay2011$year = "2011"; setorder(ay2011, count_Lab1)
37 ay2012 = load_njdoe_dbase("Free_Reduced_Lunch", "2012"); ay2012$year = "2012"; setorder(ay2012, count_Lab1)
38 ay2013 = load_njdoe_dbase("Free_Reduced_Lunch", "2013"); ay2013$year = "2013"; setorder(ay2013, count_Lab1)
39 ay2014 = load_njdoe_dbase("Free_Reduced_Lunch", "2014"); ay2014$year = "2014"; setorder(ay2014, count_Lab1)
40 ay2015 = load_njdoe_dbase("Free_Reduced_Lunch", "2015"); ay2015$year = "2015"; setorder(ay2015, count_Lab1)
41 ay2016 = load_njdoe_dbase("Free_Reduced_Lunch", "2016"); ay2016$year = "2016"; setorder(ay2016, count_Lab1)
42 ay2017 = load_njdoe_dbase("Free_Reduced_Lunch", "2017"); ay2017$year = "2017"; setorder(ay2017, count_Lab1)
43 ay2018 = load_njdoe_dbase("Free_Reduced_Lunch", "2018"); ay2018$year = "2018"; setorder(ay2018, count_Lab1)
44 ay2019 = load_njdoe_dbase("Free_Reduced_Lunch", "2019"); ay2019$year = "2019"; setorder(ay2019, count_Lab1)
45
46 ##
47 ## Generate a Total Table
48 ##
49 ayTotal = bind_rows(ay2000, ay2001, ay2002, ay2003, ay2004, ay2005, ay2006, ay2007, ay2008, ay2009,
50                      ay2010, ay2011, ay2012, ay2013, ay2014, ay2015, ay2016, ay2017, ay2018, ay2019)
51
52 ##
53 ## Convert errant data to integers
54 ##
```



## New Jersey Student Hunger Audit

Chris Snyder, MPH\* Kelsey Bent, MPH, CHES

### Background Information

New Jersey presently has over 1.3 million students enrolled in schools during the 2018-2019 academic year; an increase of over 75,000 students in the total student body over the past 20 years. The National School Lunch Program (NSLP) is a federally assisted meal program operating in public and nonprofit private schools, which provides nutritionally balanced, free or reduced-cost lunches to children each school day. Children may be determined categorically eligible for the free meal program through the participation in certain Federal Assistance Programs, such as Supplemental Nutrition Assistance Program (SNAP), formerly known as Food Stamps. In addition, children may also qualify for free or reduced-price lunch program based upon their household income compared to the Federal Poverty Level (FPL). Children from families with incomes at or below 130% FPL are eligible for free meals; those with incomes between 130% and 185% FPL are eligible for reduced price meals.

The New Jersey Department of Education (NJDOE) (link) maintains publicly available enrollment data, which includes the number of students receiving free and reduced-price lunches. Over the last 20 years, we have seen a staggering increase of 10% across the state in use of the free and reduced-price lunch program. This may be indicative of a much larger issue, for as statewide economic factors change, the risk of student food insecurity becomes more prevalent. Further, while students are receiving assistance during K-12 schooling, there is no assistance available as they enter institutions of higher education.

The 2015-2020 Dietary Guidelines for Americans acknowledged that there is a connection between food insecurity and lifelong health outcomes. Other scholars have noted a link between food security and educational attainment and academic performance.

As such, an audit of Student Food Security in New Jersey is required, in order to determine where the areas of highest risk to food security and student well-being, and to suggest the changes necessary to the campus environment to be able to support students from these communities.

\*Montclair State University, snyderc@montclair.edu



# Writing in R

- ~~What R and RStudio are / are not~~
- ~~How to start a new R project~~
- ~~How to write and run your first R script~~
- ~~How to expand the functionality of R with packages and user-defined functions~~
- ~~How to load, manipulate, and save your data~~
- ~~How to visualize your data~~

# Demo

- Logic
- Saving Data
- Loading Data
- Sourcing Scripts
- Loading Packages
- Running a User-Defined Function
- Visualizing Data
- Automation with .bat Files

# Questions?

[snyderc@montclair.edu](mailto:snyderc@montclair.edu)

<https://github.com/MSUCCG>