

Introduction to R

Steven Tseng

17 September 2020

What is R?

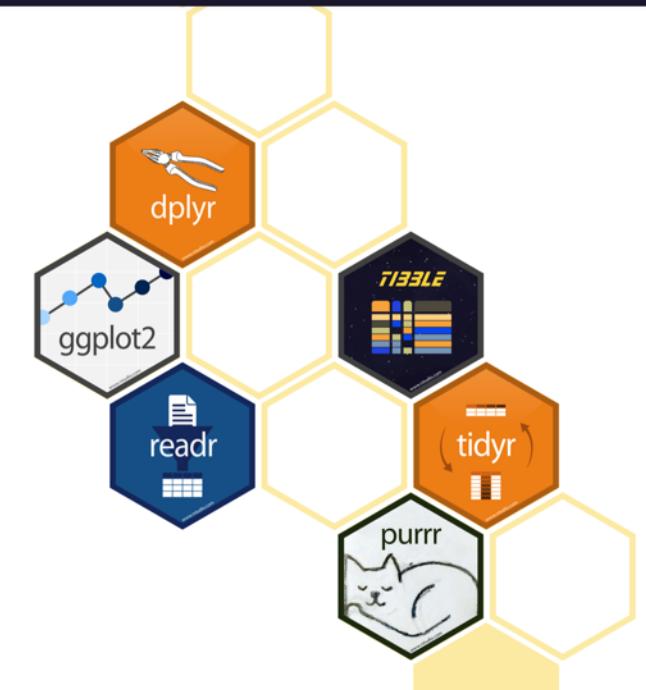
- Statistical Software
- Programming Language



Why R? Lite

- Free
- Open source
- In demand
- Large and growing library of *packages*
- Huge online support network

Tidyverse



Never ask if R can do what you want. It can.

The only limitation to R is the user's programming sophistication. The better you get at statistical programming the further you can explore your data and take your research. This holds for any statistical endeavor whether using R or not.

Michael Clark, Ph.D.
Advanced Research Computing
University of Michigan

Why R? +

- Mode of communication with statisticians, computer scientists, and data scientists
- Open science, reproducible research
- State-of-the-art data visualization
- Introduction to programming
- Forces you to really understand your data
- Empowering

Overview of Series

Three Days of Fun

Session 1

Data Visualization

February 14th ❤️

10am – 1pm

Session 2

Data Wrangling

Data Analysis

March 13th

10am – 1pm

Session 3

Projects

Notebooks

April 10th

10am – 1pm

Goals of Workshop Series

- Provide foundation to allow autonomous exploration
- Maximal hands-on experience



Programming sucks
@UserInputSucks

you call me a developer
i call myself a professional googler

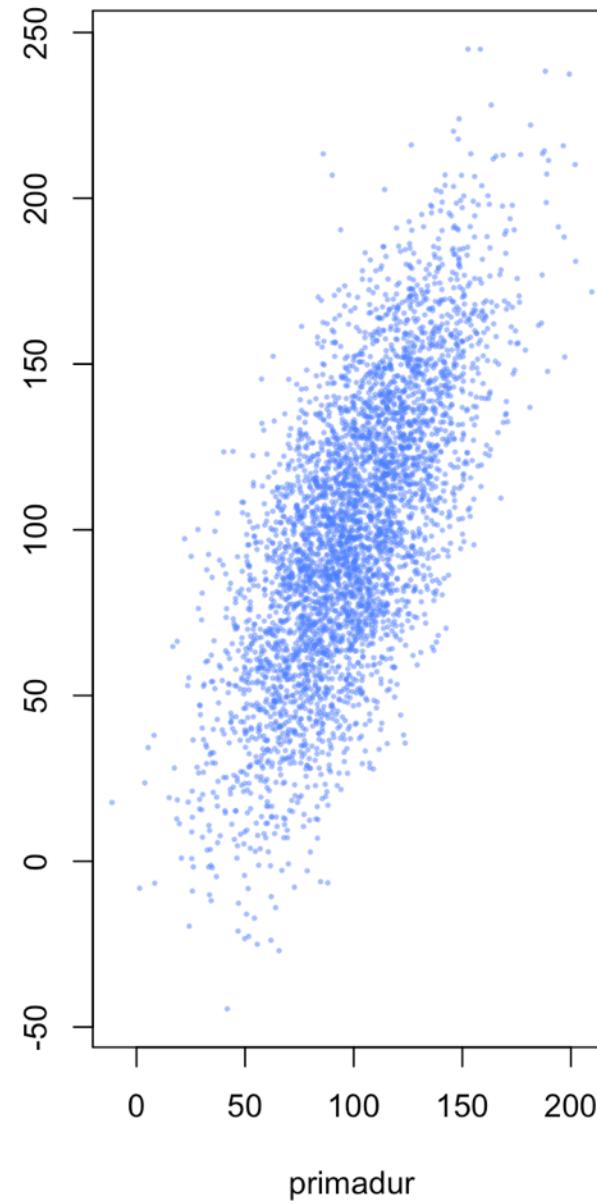
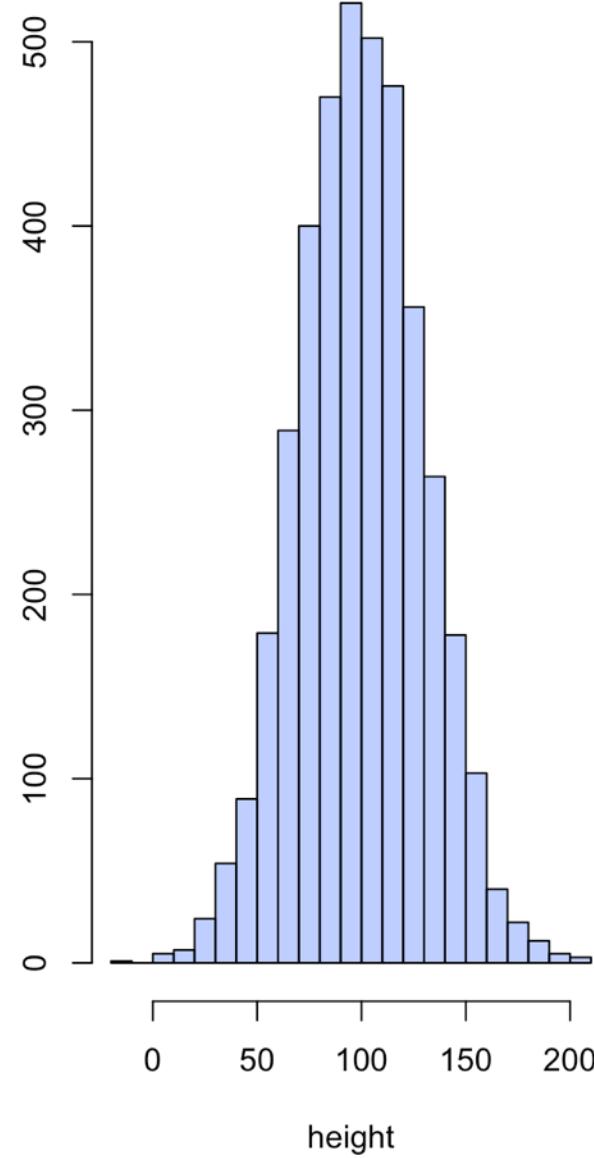
Overview of Session 1

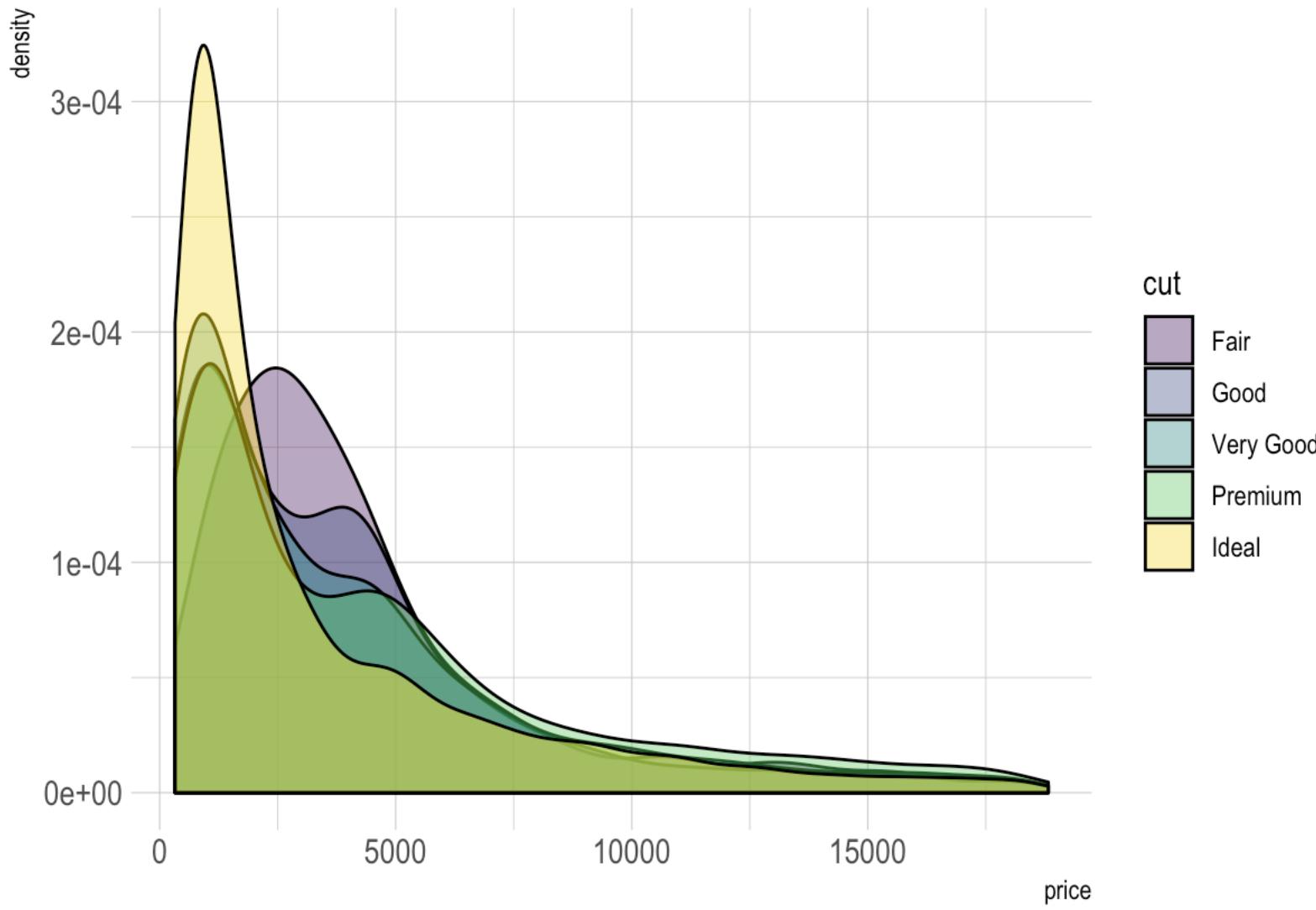
1. Brief introduction to basics of R language
2. Make pretty graphs
3. Make them even prettier

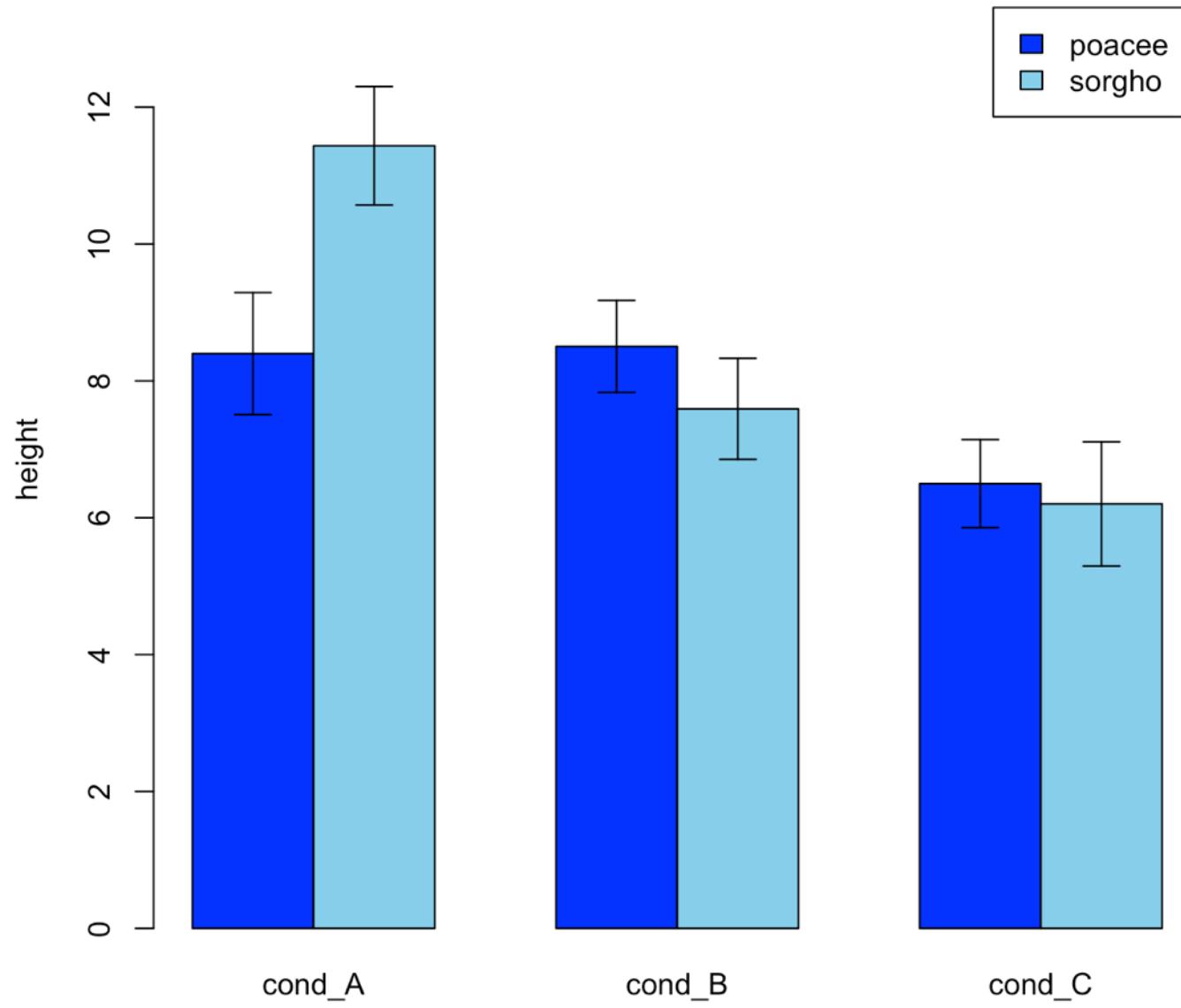
Why Data Viz?

- Universally important skill
- SIOP Top 10 Workplace Trends
- It's really, really cool

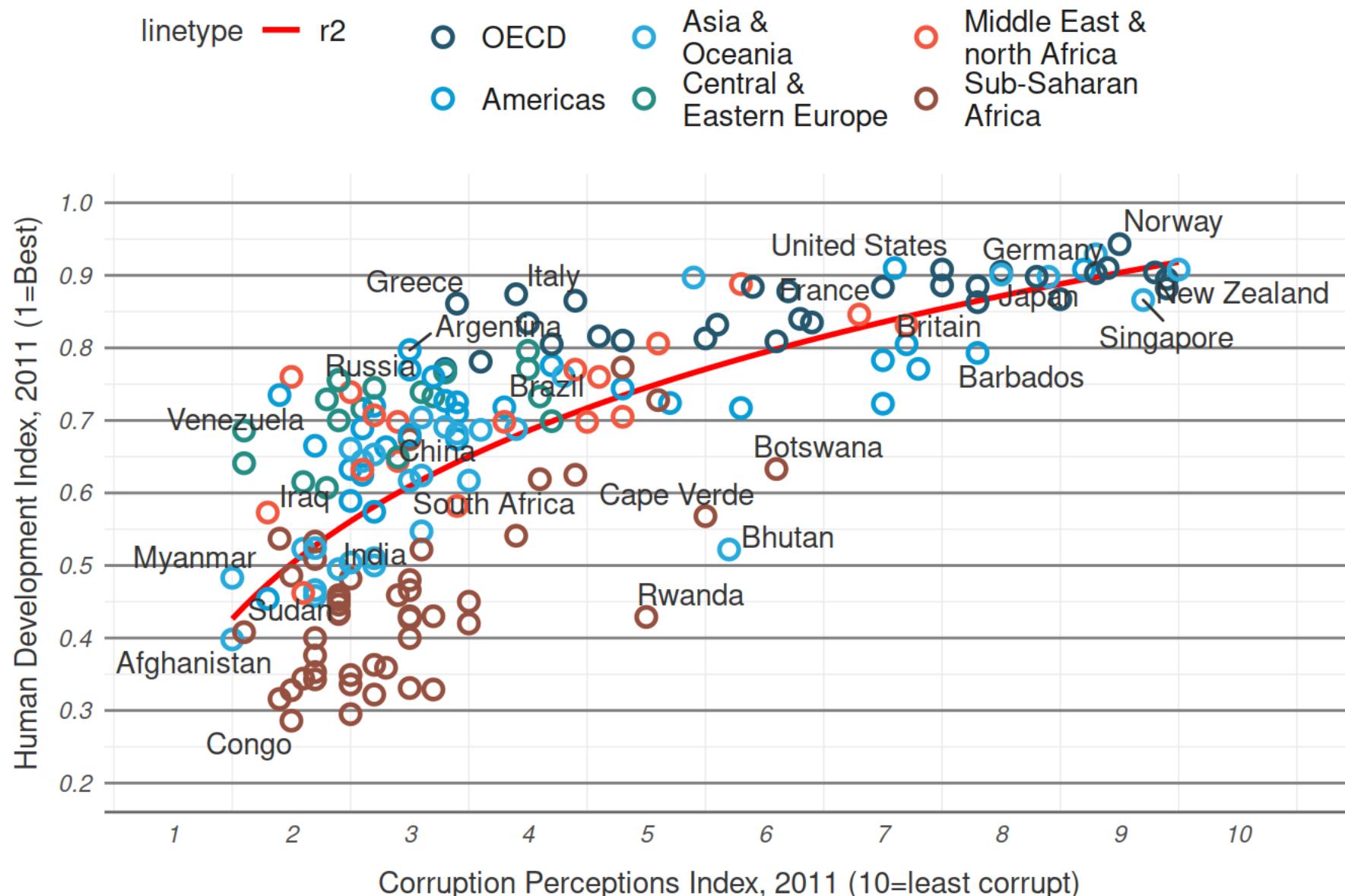
Primadur : Distribution and correlation with Ixos



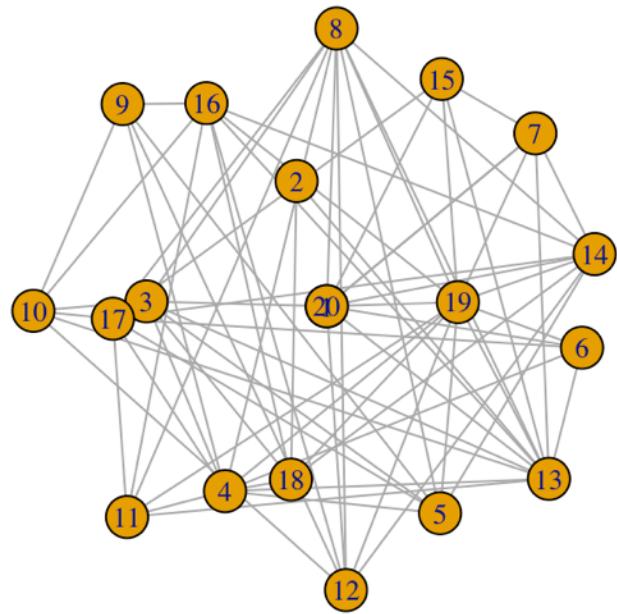




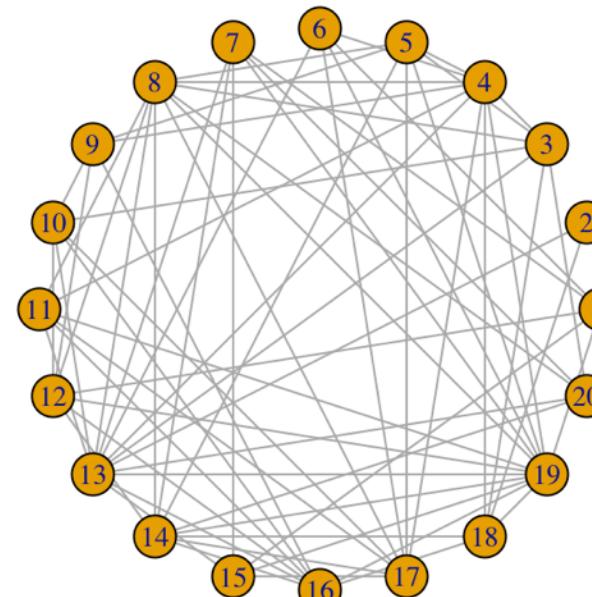
Corruption and Human development



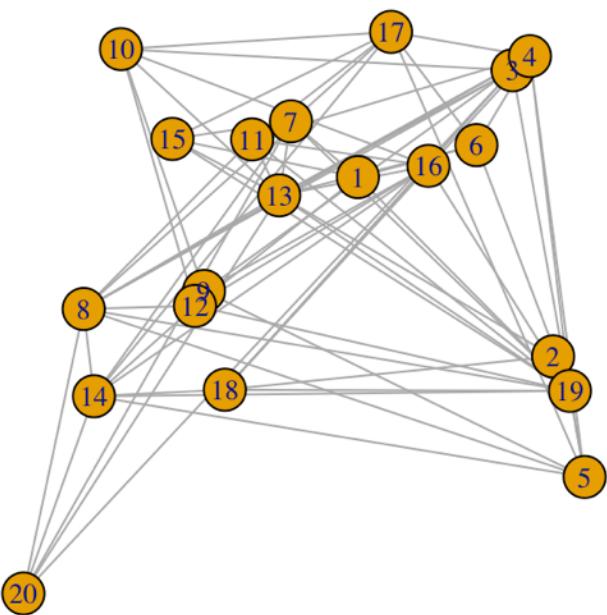
sphere



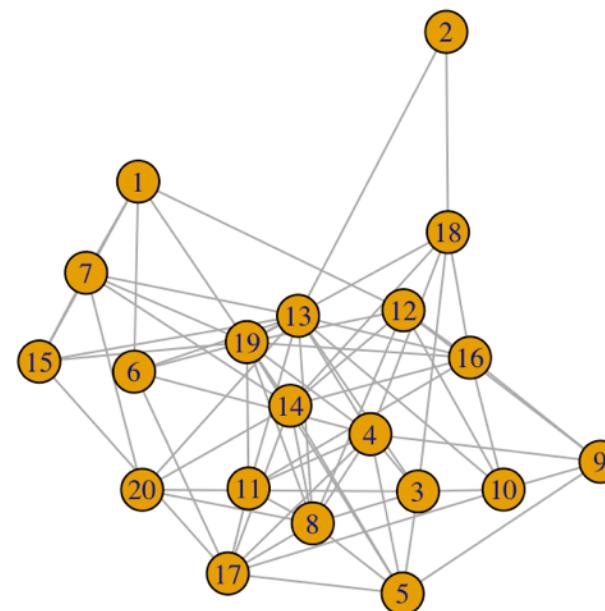
circle

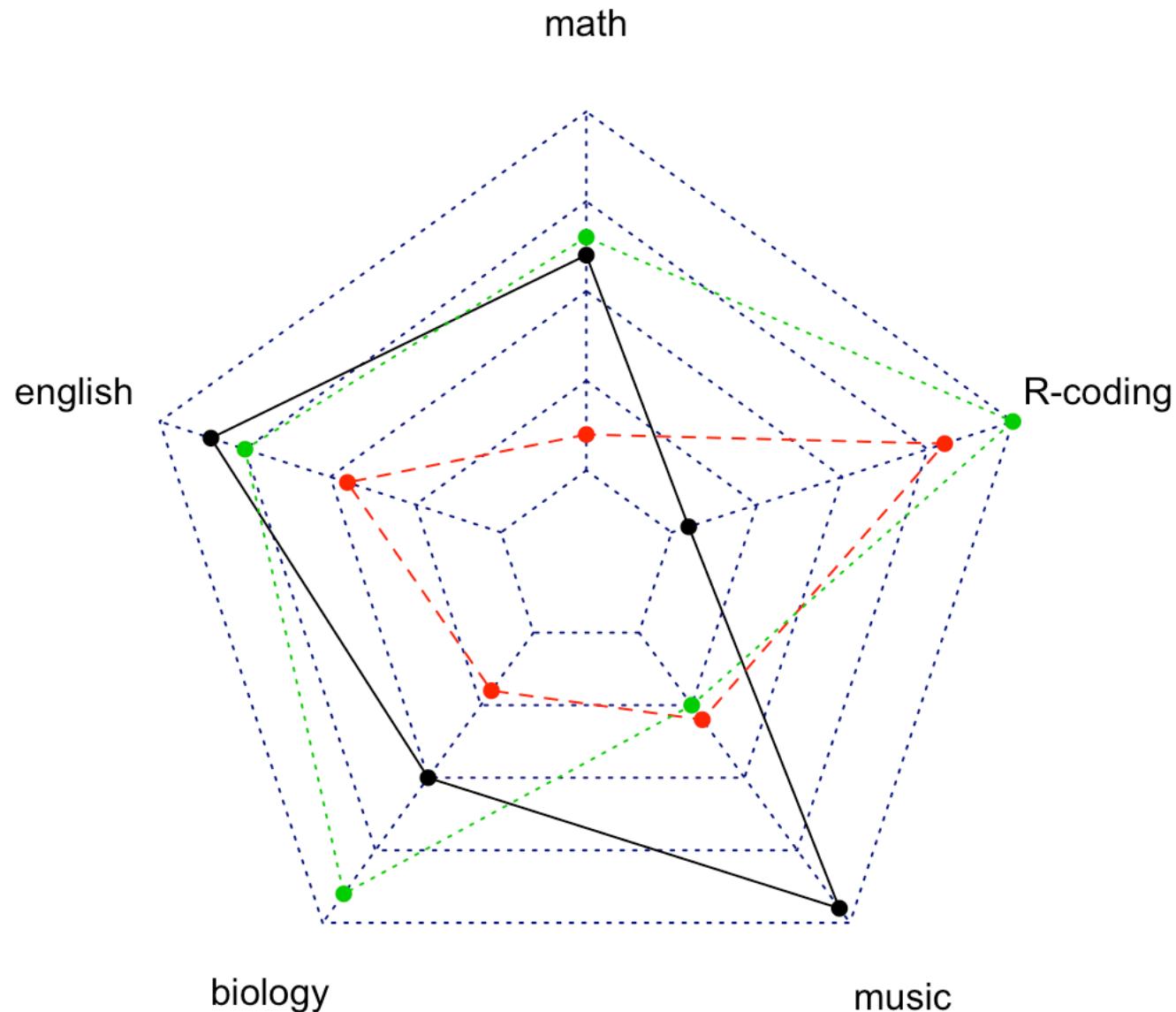


random



fruchterman.reingold





A word cloud visualization of the Gettysburg Address, showing the frequency of words in various colors.

The most prominent words are:

- freedom
- will
- dream
- ring
- day
- one
- together
- shall
- men
- black

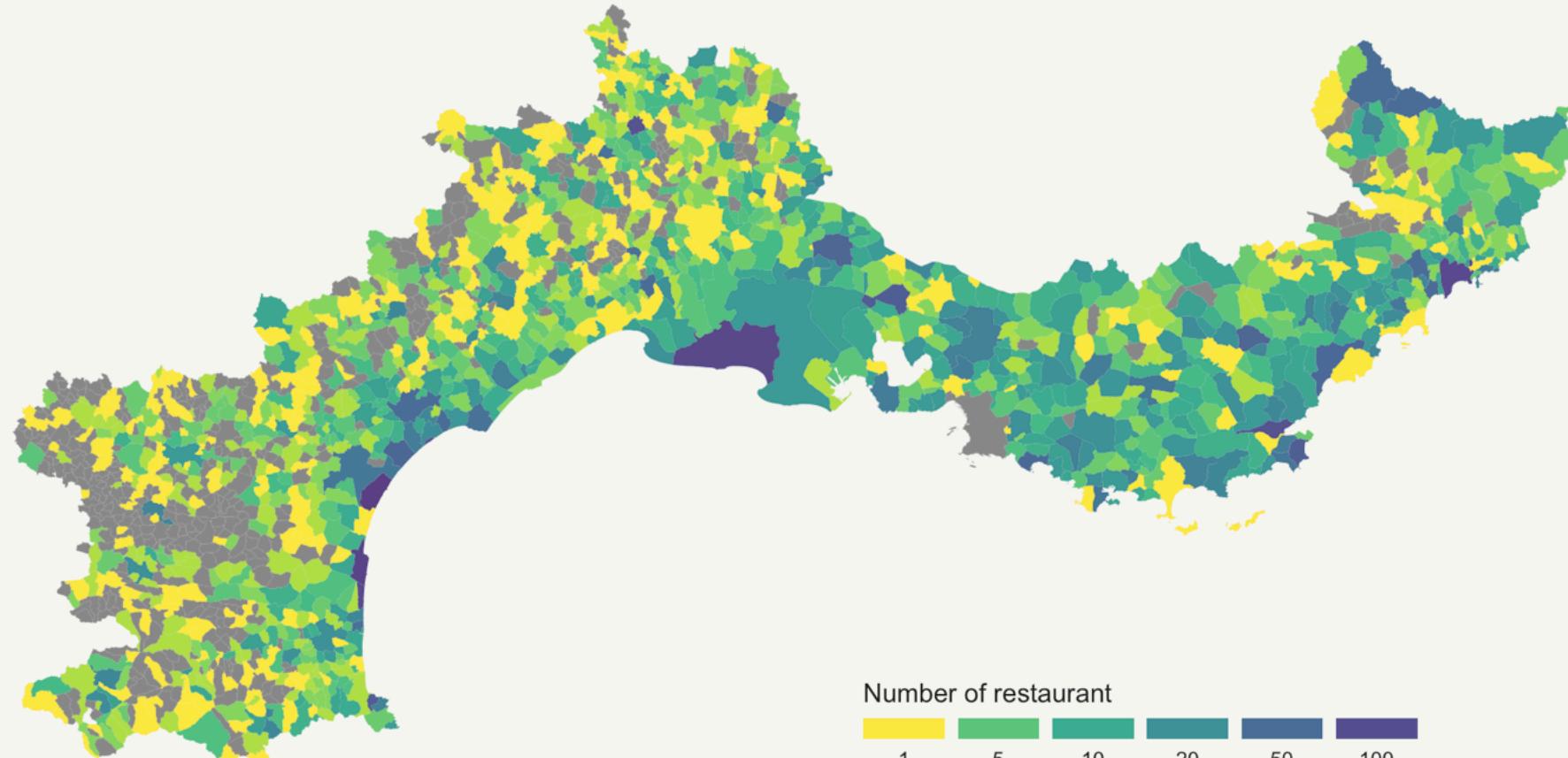
Other significant words include:

- nation
- state
- government
- of
- the
- people
- and
- we
- are
- now
- under
- God's
- protection
- in
- this
- country
- with
- all
- its
- glory
- and
- misery
- here
- before
- us
- which
- shall
- not
- be
- forgotten
- nor
- any
- living
- hereafter
- forget
- what
- they
- did
- here
- which
- shall
- not
- be
- forgotten
- nor
- any
- living
- hereafter
- forget
- what
- they
- did
- here

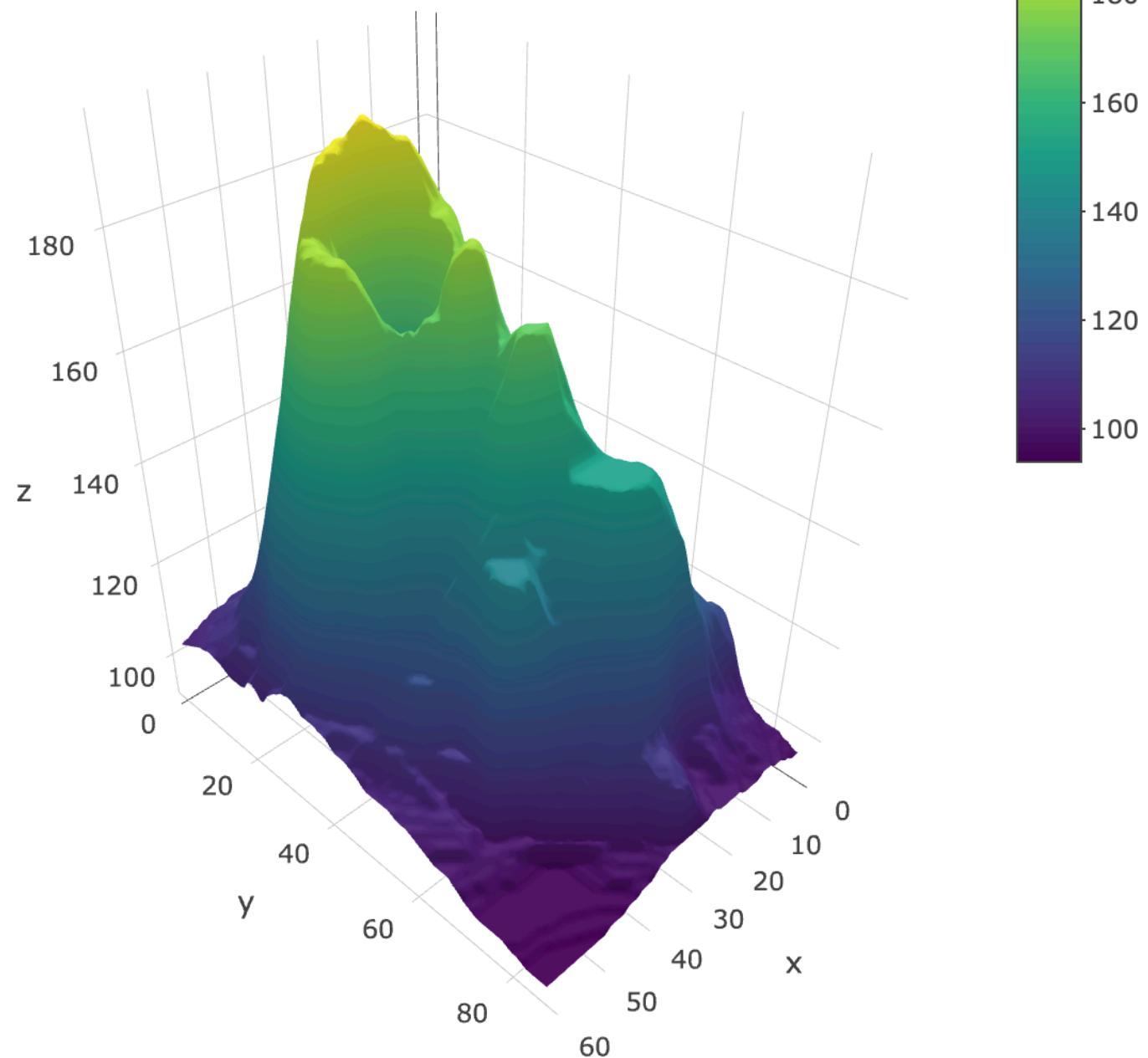


South of France Restaurant concentration

Number of restaurant per city district



Data: INSEE | Creation: Yan Holtz | r-graph-gallery.com



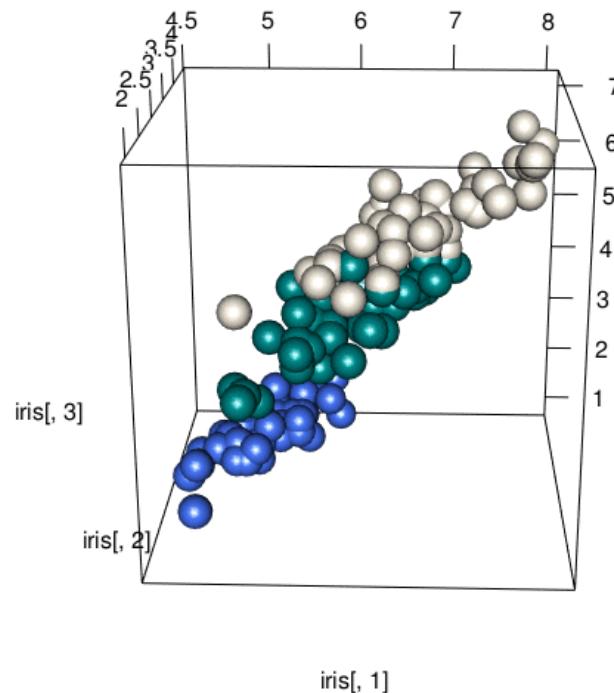
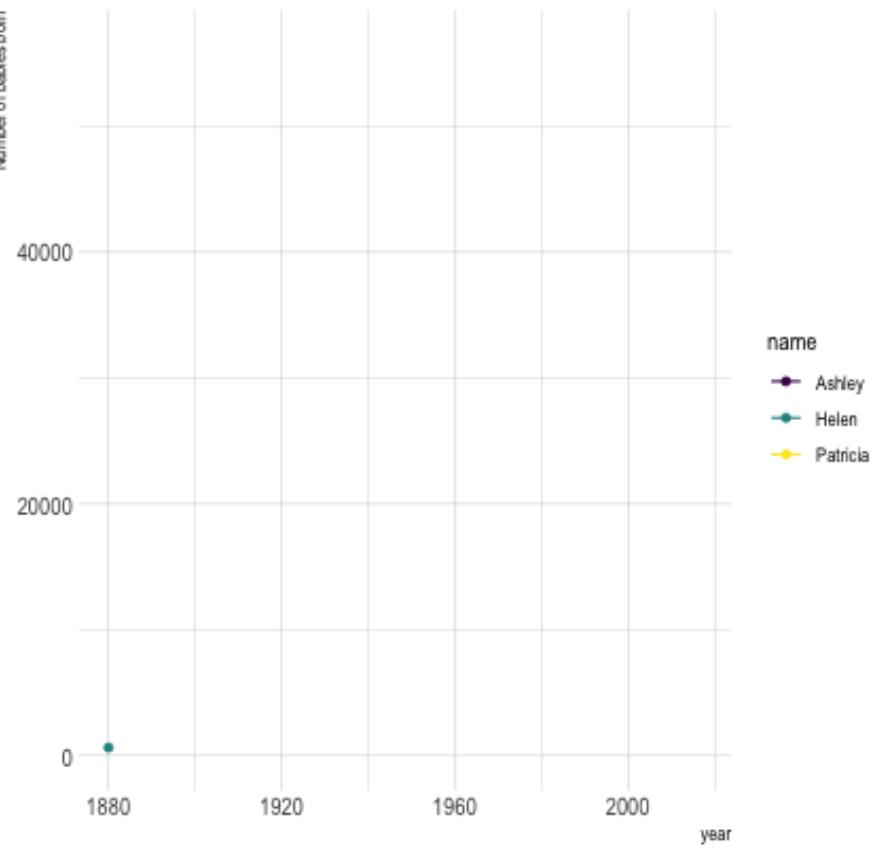


facebook

December 2010

Number of babies born

Popularity of American names in the previous 30 years



Preview of Sessions 2 & 3

Dissertation Primary Analysis

Code ▾

Steven T. Tseng

First created on Dec 24, 2019. Updated on 2020-02-03

- [1 Correlation Matrix \(Full\)](#)
- [2 Spearman Correlations for Some Demographic Variables \(Ordinal\)](#)
- [3 Means and Standard Deviations](#)
- [4 Correlation Matrix \(Partial\)](#)
- [5 Spearman Correlation Matrix \(Partial\)](#)
- [6 Hypothesis Tests](#)
 - [6.1 Hypothesis 1](#)
 - [6.2 Hypothesis 2](#)
 - [6.3 Hypothesis 3](#)
 - [6.4 Hypothesis 4](#)
 - [6.5 Hypothesis 5](#)
 - [6.5.1 Plot Interaction \(Factor Predictor\)](#)
 - [6.5.2 Plot Interaction \(Continuous Predictor\)](#)
 - [6.5.3 Obtaining MSE for Reporting](#)
 - [6.6 Hypothesis 6](#)
 - [6.6.1 Organization-Centric HR Attributions](#)
 - [6.6.2 Employee-Centric HR Attributions](#)
 - [6.7 Hypothesis 7](#)
 - [6.7.1 Using psych::mediate](#)
 - [6.7.2 Using lavaan::sem](#)

Getting Started

Step 1

- Launch RStudio

RStudio

Project: (None)

Console ~/

R version 3.6.2 (2019-12-12) -- "Dark and Stormy Night"
Copyright (C) 2019 The R Foundation for Statistical Computing
Platform: x86_64-apple-darwin15.6.0 (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.

> hello world
Error: unexpected symbol in "hello world"
> print("hello world")
[1] "hello world"

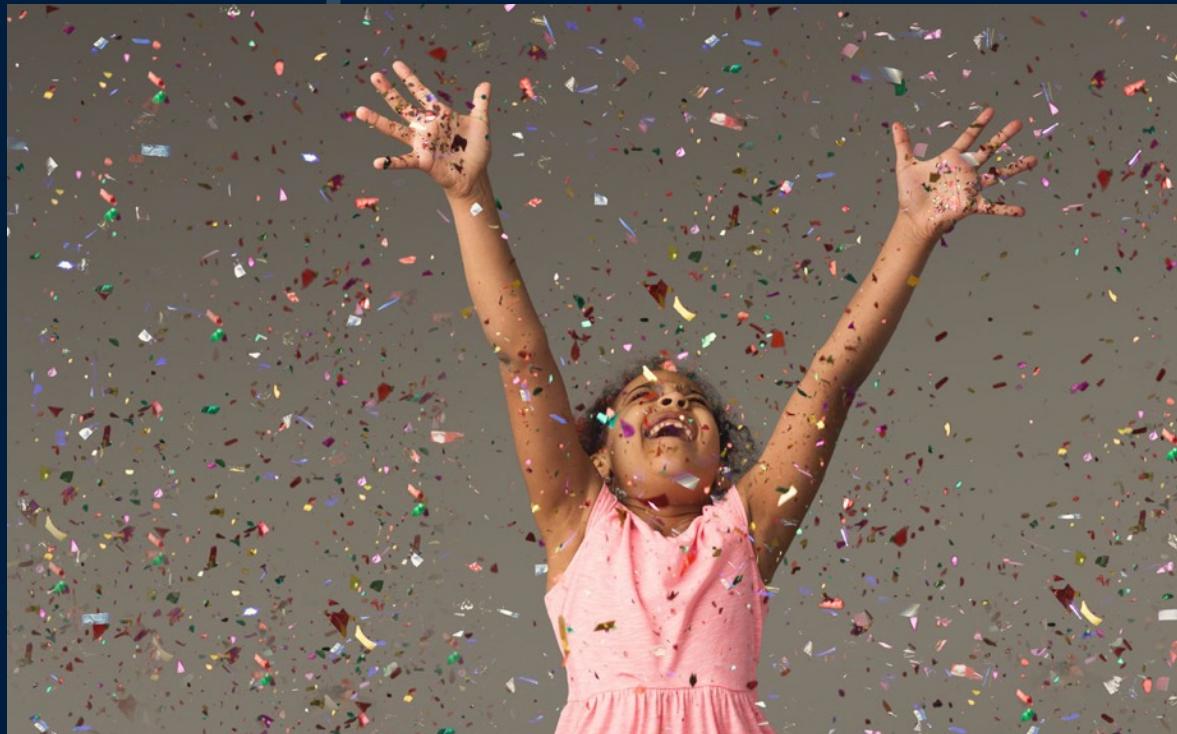
Environment History Connections

Import Dataset Global Environment

Environment is empty

Environment is empty

16, 2016, 11:38 AM
6, 2020, 7:57 PM



Desktop Documents Downloads Dropbox Google Drive

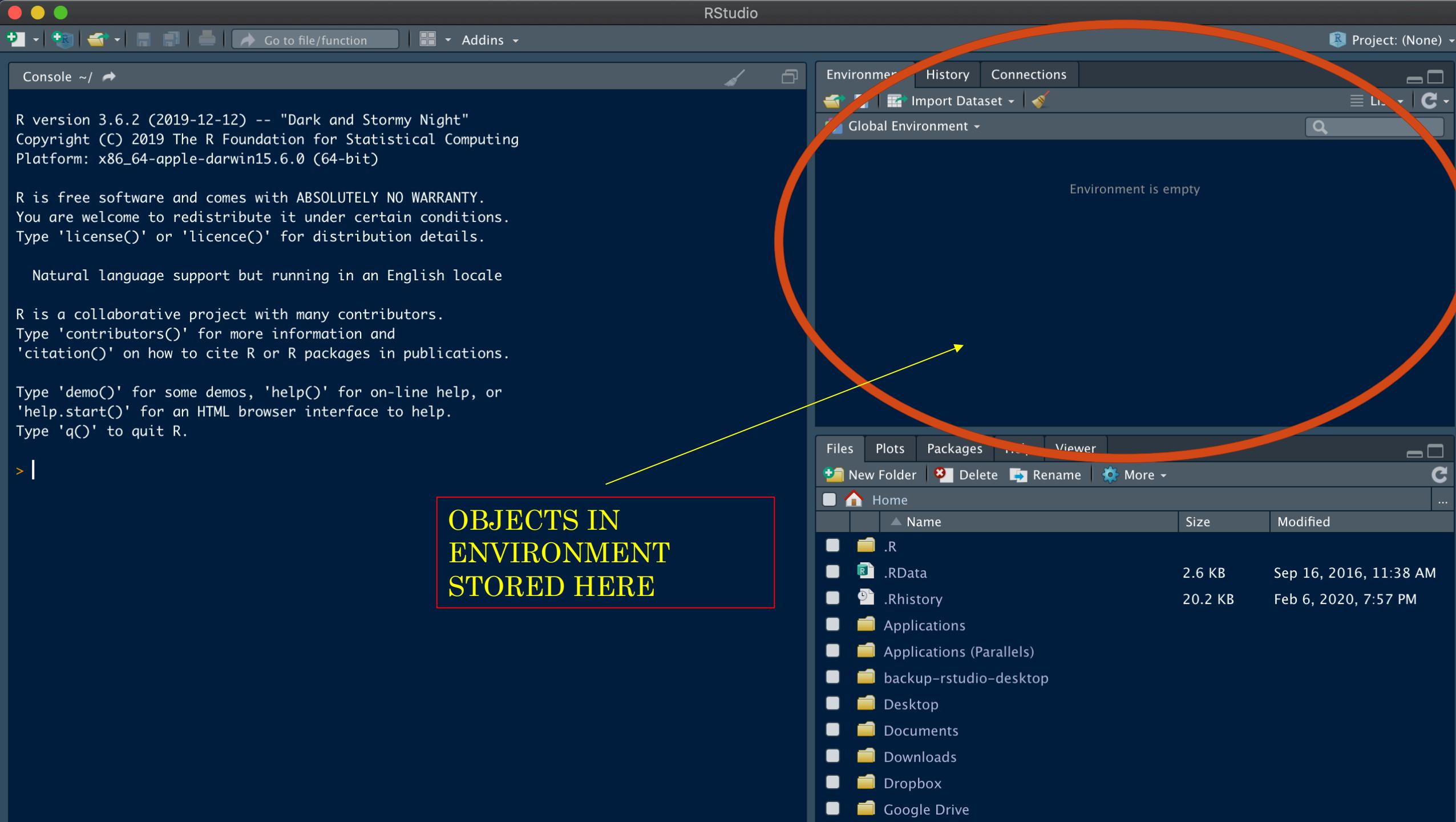
Step 2

- Further explore the console

- $1 + 1$

- 2^2

- $2/2$



Step 3

- Understand the basics of the R language

➤ Object-Oriented

➤ Command to store objects:

<-

Hotkey: **alt + -**

RStudio

Console ~/

```
> x <- 2
> y <- 3
> x
> y
> x*y
```

Environment History Connections

Import Dataset | Global Environment | List | C

Values

x	2
y	3

Files Plots Packages Help Viewer

New Folder | Delete | Rename | More

Home

	Name	Size	Modified
📁	.R	2.6 KB	Sep 16, 2016, 11:38 AM
📄	.RData	20.1 KB	Feb 10, 2020, 1:52 PM
⌚	.Rhistory		
📁	Applications		
📁	Applications (Parallels)		
📁	backup-rstudio-desktop		
📁	Desktop		
📁	Documents		
📁	Downloads		
📁	Dropbox		
📁	Google Drive		

Step 3

- Try:

- `x <- c(1, 2, 3, 4, 5)`



You just ran a FUNCTION!

- `x + 1`

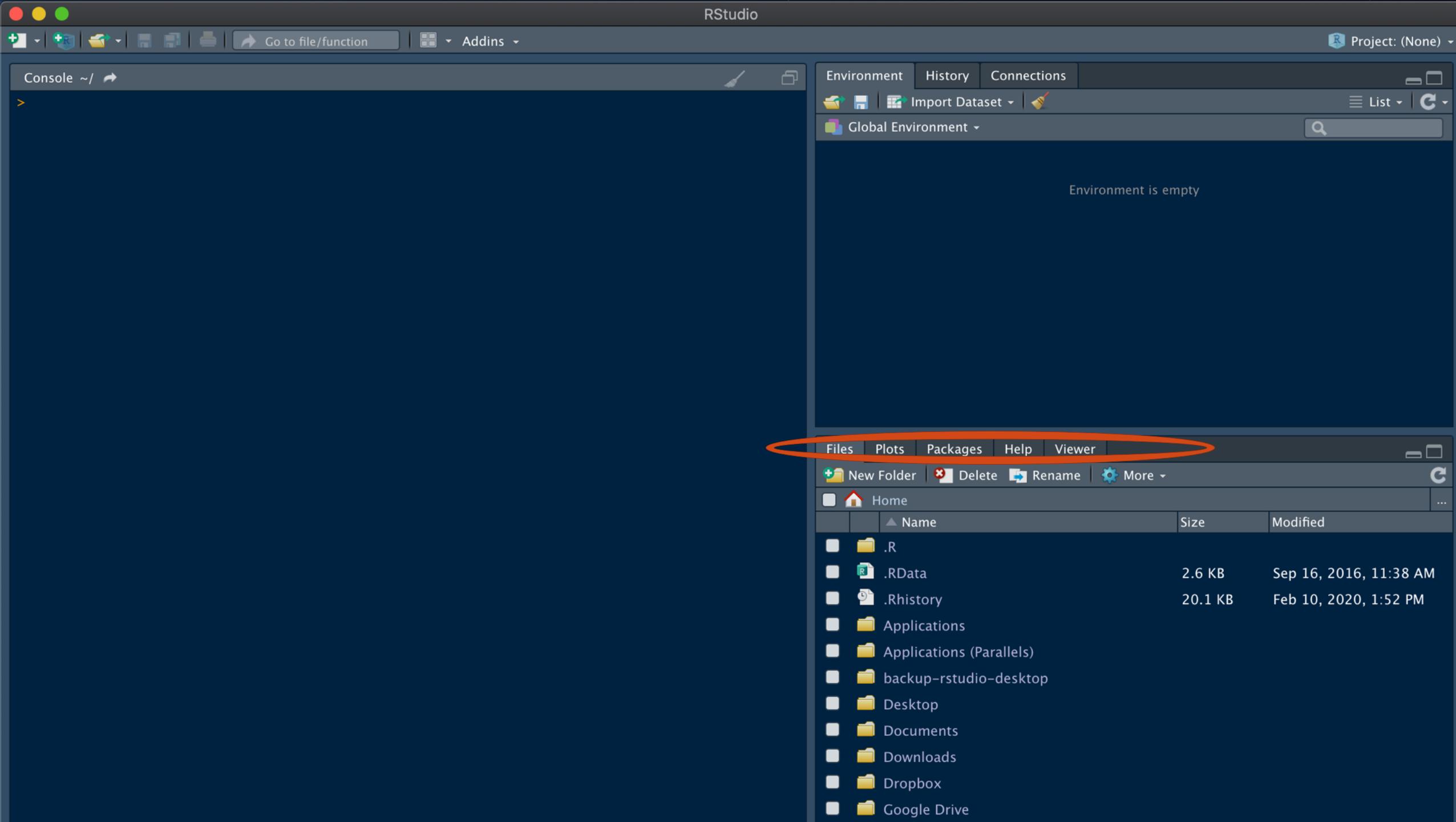
General Form:
function(argument1, argument2, ...)
Ex. `print()`

- `x <- matrix(c(1, 2, 3, 4, 5, 6), 3, 2)`



Step 3

- `data.frame(x)`
- `x <- data.frame(x)`



Step 4

- Learn to get help

➤ ?c()

➤ ?matrix()

matrix {base} R Documentation

Matrices

Description

`matrix` creates a matrix from the given set of values.
`as.matrix` attempts to turn its argument into a matrix.
`is.matrix` tests if its argument is a (strict) matrix.

Usage

```
matrix(data = NA, nrow = 1, ncol = 1, byrow = FALSE,  
       dimnames = NULL)
```

Step 5

- Store code in R scripts

