The
Data
Driven Manager

Working With Data

In RStudio and ROIStat

Learning Objectives

- Install R and RStudio
- Install ROIStat
- Access R packages
- Create and import data files into RStudio and ROIStat
- Work with data frames in RStudio
- Work with data in ROIStat

Getting Started

With R and RStudio

Introduction to R and RStudio

- R is the programming language
- RStudio is the user interface

1 Install R

- Download R at http://www.r-project.org
- Once downloaded, double click to install R

2 Install RStudio

- Download RStudio at http://www.rstudio.com or

  https://posit.co

- Once downloaded, double click to install RStudio

3 Open RStudio

4 Install lolcat

```
install.packages("devtools")
require(devtools)
install_github("burrm/lolcat")
require(lolcat)
```

Packages

Once you have installed a package, it needs to be loaded any time you open RStudio

```
require( )
library( )
```

R Syntax

Case sensitive

Assignment operator: <-

Import Data into
RStudio

Text Data

Excel Data

Copy / Paste

```
df<-read.table("clipboard", sep="\t",
header = TRUE)
View(df)
```

Import Data into RStudio

The datasets we will be using in our course will be imported as a data frame, with different variable names.

To select a variable: Dataframe$variable

Data Types

Character   Character or string values ("abc" "length")

Numeric   Real numbers, with / without decimals (5, 16.7)

Integer   Whole number (123L) designated with "L"

Logical   TRUE, FALSE (T, F)

Data Structures

Vectors   Numeric, Character, Logical, Integer

1 x n

Matrices   All columns in a matrix must have the
same data type (numeric, character, etc.)
and the same length.
m x n

Data **Structures**

Arrays   Same as matrices but can have more
than two dimensions

Lists   An ordered collection of objects
(components). A list allows you to
gather a variety of (possibly unrelated)
objects under one name.

Data **Structures**

Data Frames A data frame similar to a matrix, but different columns (vectors) can have different classes (numeric, character, factor, etc.)

Data Structures

Factors   Tell R that a variable is nominal by making it a factor.

`factor( )`

An ordered factor is used to represent an ordinal variable.

`ordered( )`

Useful **Functions**

length(object)   # number of elements or components
str(object)   # structure of an object
class(object)   # class or type of an object
names(object)   # names

Useful Functions

```
c(object,object,...)       # combine objects into a vector
cbind(object, object, ...)   # combine objects as columns
rbind(object, object, ...)   # combine objects as rows
```

Useful Functions

object   # prints the object

ls()   # list current objects

rm(object)   # delete an object

Useful Functions

newobject <- **edit**(object)   # edit copy and save as newobject

fix(object)   # edit in place

Variable Labels

```
names(Sample)[3] <- "Level"

colnames(Sample)<-c("Samp", "Date", "Level", "Wt","Ht")
rownames ( )
```

Value Labels

```
Sample$Factor<- factor(Sample$Factor,
levels = c(1,2,3),
labels = c("Low", "Medium", "High"))
```

Missing Values

Testing for Missing Values

is.na(Sample)   # returns TRUE when value missing

Missing Values

Excluding Missing Values from Analyses

Arithmetic functions on missing values yield missing values.

mean(Sample$Height)

[1] NA

> mean(Sample$Height, na.rm=T)

Date Values

Character to Date

use as.Date( ) to convert strings to dates

Sample$Time<-as.Date(Sample$Time)

Getting Started

In ROIStat

Introducing ROIstat!

- A **free** graphic user interface for R

ROIStat

- Steve Ouellette, MS, CMC®
- Former prof at EMP
- Applied stats geek
- Consultant since 1996
- Author of Galileo's Telescope (2021)

ROIstat Use Case

- Professors
  - FREE
  - Concentrate on teaching how and when to use statistics
  - When students are ready for R, it is already installed
- Students
  - FREE
  - Concentrate on learning statistics, rather than struggling to get the output from your script
- Professionals
  - Most of what you will need with analytics can be done with ROIstat
  - It is fast to get in, get an answer, and get on with your life
  - Oh yeah, FREE

Go to ROI-Ally.com, hover over Services...

Home   About   Services ⌄   Case Studies ⌄   the ROI alliance   Blogs ⌄   📖 Books

HOW WE CAN HELP

📈 Management Consulting

📊 Industrial Statistics

💬 On-Site Training   ›

🌐 Online Training

🖥 ROIstat Software   **Click here**

Achieve   lence

Retu   stment from

Lear

Learn More →

Register with the site

Introducing ROIstat

The **free and easy** graphic user interface harnessing the power of R for professors, students, and working professionals alike.

Register first, then...

⬇ Download and Install

Click here

Professor / Teacher

Students don't need to know how to use R to learn the power of statistics.

Do you spend more time instructing introductory students in how to script in R instead of how and when to use statistics? This GUI still runs in R (and is free!) but allows you to **focus on helping the students understand statistics**. When they are ready for more advanced statistics and need to learn R, they will already have it installed and working.

Student

This graphical user interface gives you a tool that allows you to **focus on learning how to use statistics to solve your problems**, not on how to write a command in R. The packages that are used in ROIstat have been shown to give the correct answers, so you don't have to worry about how some package you downloaded is going to do an analysis.

Oh, and its FREE!

Professional Engineer, Six Sigma Belt, Data Analyst

Are you frustrated when you know R can do something you need, but you don't remember the command or the syntax? Do you spend a lot of time going back to your notes rather than getting the answer?

Once you learn how statistics can be used to solve practical, everyday problems, you want to **get in, get the answer, and get on with your life**. ROIstat puts the power of R into a graphic user interface so you don't have to remember what to type to get results from your data quickly and accurately.

Oh, and it is FREE!

The Usual Stuff

User Registration

* Required field

Name *

Username *

Confirm Password *

Email Address *

Confirm Email Address *

Captcha *

Web Site Privacy

*

By signing up to this web site and agreeing to its Privacy Policy, you agree to this web site storing your email address. We do not sell contact information - we only use your information to conduct our business and to contact you about new content and events. You can unsubscribe at any time.

Go back to the ROIstat page
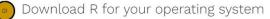
Introducing **ROIstat**

The **free and easy** graphic user interface harnessing the power of R for professors, students, and working professionals alike.

Register first, then...

⬇ Download and Install

Click here

Professor / Teacher

Students don't need to know how to use R to learn the power of statistics.

Do you spend more time instructing introductory students in how to script in R instead of how and when to use statistics? This GUI still runs in R (and is free!) but allows you to **focus on helping the students understand statistics**. When they are ready for more advanced statistics and need to learn R, they will already have it installed and working.

Student

This graphical user interface gives you a tool that allows you to **focus on learning how to use statistics to solve your problems**, not on how to write a command in R. The packages that are used in ROIstat have been shown to give the correct answers, so you don't have to worry about how some package you downloaded is going to do an analysis.

Oh, and its FREE!

Professional Engineer, Six Sigma Belt, Data Analyst

Are you frustrated when you know R can do something you need, but you don't remember the command or the syntax? Do you spend a lot of time going back to your notes rather than getting the answer?

Once you learn how statistics can be used to solve practical, everyday problems, you want to **get in, get the answer, and get on with your life**. ROIstat puts the power of R into a graphic user interface so you don't have to remember what to type to get results from your data quickly and accurately.

Oh, and it is FREE!

Follow the Instructions

01 Download R for your operating system

- https://www.r-project.org/
- Click on "download R"
- Select a geographic area near you
- Select the version of R for your operating system
- Double-click to install

02 Download RStudio

- https://www.rstudio.com/products/rstudio/
- Download the Open Source Edition
- Double-click to install

03 Open RStudio

04 Install Necessary Packages

- Copy and paste the following into the RStudio console, then hit the Enter key

```
install.packages("devtools")
require(devtools)
install_github("burrm/lolcat")
install_github("dreamRs/datamods")
install_github("dreamRs/shinyWidgets")
install_github('rstudio/DT')
install_github("tidyverse/dplyr")
```

- Wait for a bit as the various packages and their dependencies are installed.

05 Download ROIstat

- DISCLAIMER: ROIstat is GUI that presents an easy way to access a few R packages created and administered by other developers. We have little to no control over these packages, therefore we can offer no warranty as to ROIstat's fitness for your data analysis. Every effort has been made to check ROIstat's results, but if a developer were to change a calculation or format without our knowledge, inaccurate results could occur.
- By clicking on the download button below, you acknowledge this as a precondition to using the GUI.

☁ Download ROIstat

06 Start the ROIstat app

- Unzip the file you downloaded to someplace convenient for you
- Open RStudio
- Select File | Open
- Navigate to where you unzipped the ROIstat file
- Open app.R
- You will see a bunch of code load. This is what makes the UI and the server code for ROIstat work
- To start the app, click on the green arrow above and to the right of the code window
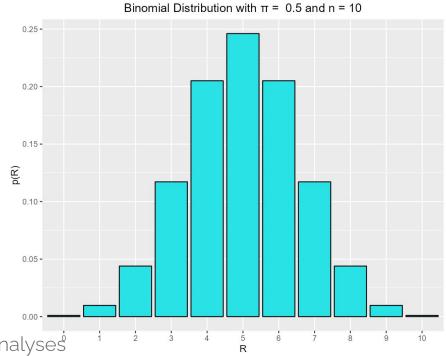
07 From Now On, Run the App from RStudio

- Just open RStudio and run the app from now on. You don't need to know a line of code to use R!

How to...

- Start the app after installation
- Import data
  - Load file
  - Paste data
- Select, modify, and filter data
- Generate a distribution
- Calculate a sample size or power
- Test for normality
- Perform one- and two-sample tests
- Perform correlation and association analyses

Binomial Distribution with $\pi = 0.5$ and $n = 10$

Other Features

- Reference column-based tests
- ANOVA
- SPC
- MSA
- Crosstabs
- Enhancement of test output, such as additional tests and graphical output
- Save and Load functionality