



Statistical Analysis in R Participant's Guide

September 30th, 2020

Instructors: Brain Ashford
Yvonne Phillips



About BeVera



- Founded in 2013
- HQ in Metro Atlanta
- Veteran-Owned Small Business
- Data Science Training and Consulting
- Technical Staffing
- Dedicated to making Data Science understandable and usable at every level of the organization
- Helping our customers become self-sufficient

BK Ashford, Director of Data Science

- Over 22 years of Data Science solution delivery experience
- Analytical Consultant for SAS Institute for over 11 years
- Successful Data Science solution delivery across multiple industries
 - Insurance Carriers
 - Major Airlines
 - Credit Bureaus
 - Retail
- Education:
 - M.Sc. Applied Mathematics, Lehigh University
 - B.S. Mathematics, Morehouse College



Yvonne Phillips, Instructor

- 18 years of Predictive Analytics work
- Currently Adjunct Professor, Morehouse College – Computer Science/Data Science
- Experienced analytic professional with demonstrated history in:
 - Insurance: Underwriting Auto & Property, Motor Vehicle Reporting Analytics
 - Government: Federal and State/Local government customer solutions needs
 - Financial/Credit Cards
 - Retail
- Education:
 - M.Sc., Decision Science, Georgia State University, Robinson College of Business
 - B.S., Mathematics, Spelman College

Read article:

Introduction to Research Statistical Analysis.pdf - Adobe Acrobat Reader DC

File Edit View Sign Window Help

Home Tools Introduction to Res... x

71 (1 of 5) 116%

Education

Introduction to Research Statistical Analysis: An Overview of the Basics

Christian Vandever¹

Abstract

Description

This article covers many statistical ideas essential to research statistical analysis. Sample size is explained through the concepts of statistical significance level and power. Variable types and definitions are included to clarify necessities for how the analysis will be interpreted. Categorical and quantitative variable types are defined, as well as response and predictor variables. Statistical tests described include t-tests, ANOVA and chi-square tests. Multiple regression is also explored for both logistic and linear regression. Finally, the most common statistics produced by these methods are explored.

Keywords

statistical analysis; sample size; power; t-test; anova; chi-square; regression

Introduction

Statistical analysis is necessary for any research project seeking to make quantitative conclusions. The following is a primer for research-based statistical analysis. It is intended to be a high-level overview of appropriate statistical testing, while not diving too deep into any specific methodology. Some of the

Author affiliations are listed at the end of this article.

Correspondence to:
Christian Vandever
HCA Healthcare Graduate
Medical Education
2000 Health Park Drive
Brentwood TN, 37027
(Christian.Vandever@hca-healthcare.com)

Search 'Measure'

Edit PDF

Export PDF

Adobe Export PDF

Convert PDF Files to Word or Excel Online

Select PDF File

Introductio...nalysis.pdf

Convert to

Microsoft Word (*.docx)

Document Language:
English (U.S.) [Change](#)

Create, edit and sign PDF forms & agreements

[Start Free Trial](#)

Module 1: R Programming Language



<https://cran.r-project.org/doc/manuals/R-intro.pdf>

RStudio is a four pane work-space for 1) creating file containing R script, 2) typing R commands, 3) viewing command histories, 4) viewing plots and more.

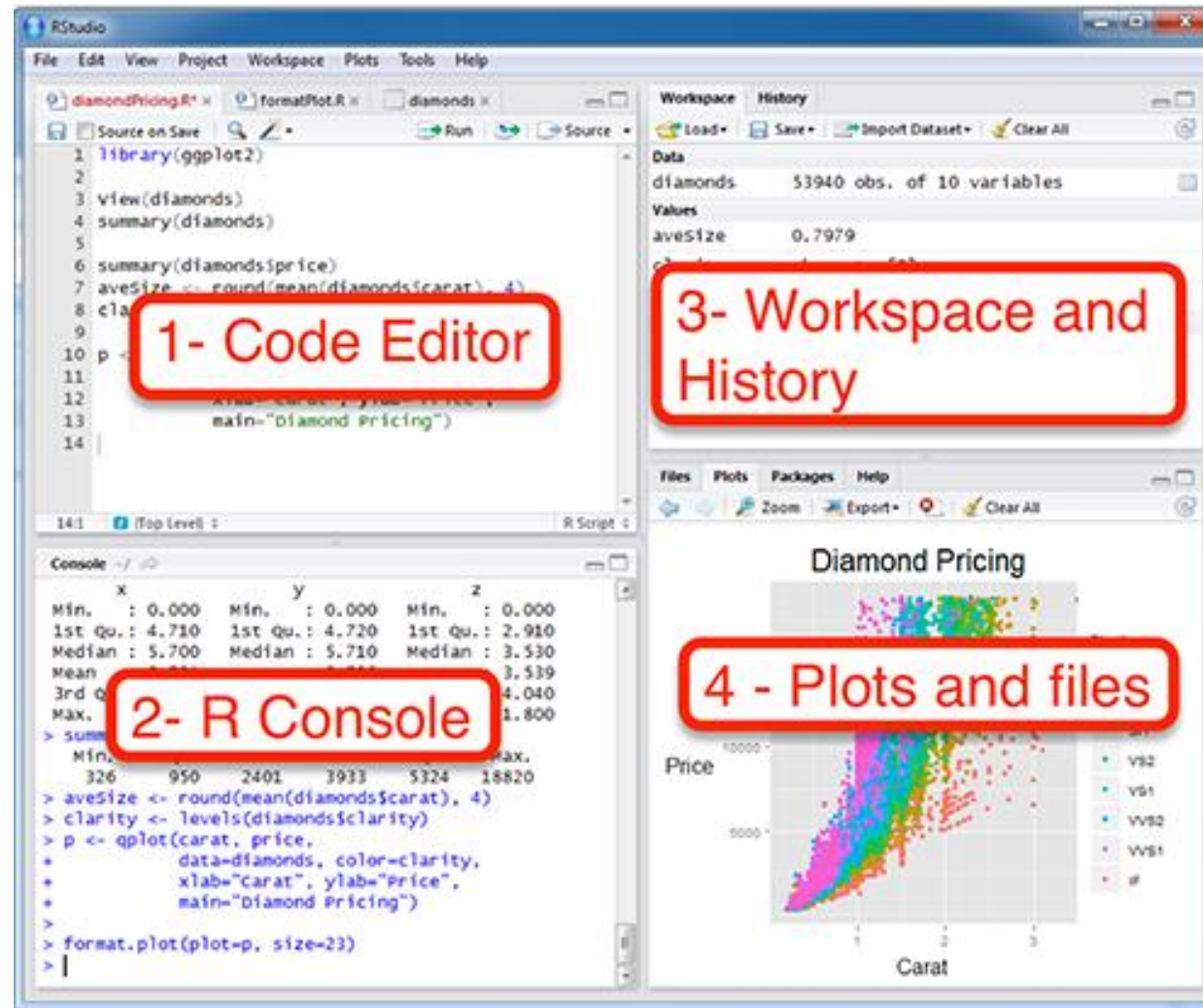
1.Top-left panel:

- Code editor allowing you to create and open a file containing R script.

- The R script is where you keep a record of your work. R script can be created as follow:
File → New → R Script

2.Bottom-left panel:

- R console for typing R commands



3.Top-right panel:

- Workspace tab: shows the list of R objects you created during your R session
- History tab: shows the history of all previous commands

4.Bottom-right panel:

- Files tab: show files in your working directory
- Plots tab: show the history of plots you created. From this tab, you can export a plot to a PDF or an image files
- Packages tab: show external R packages available on your system. If checked, the package is loaded in R.

Importing your data



R allows for the import of different data formats using specific packages that can make your job easier:

- [readr](#): importing flat files into R
- [readxl](#): getting excel files into R
- [haven](#): import SAS, STATA and SPSS data files into R.
- [RMySQL](#) and [Rpostgresql](#): Databases, access and manipulate via DBI
- [rvest](#): webscraping

Example: `mydata=read.csv(file="C:/.....csv", header=TRUE)`

<https://www.datacamp.com/community/tutorials/r-data-import-tutorial>

Manipulating your data



- [tidyr](#) package: tidying the data
- [stringr](#) package: string manipulation
- [dplyr](#) package: for data frame like objects
 - set of verbs: mutate(), select(), filter(), summarise(), arrange()
- [data.table](#) package: heavy data wrangling tasks
- [zoo](#), [xts](#) and [quantmod](#): Performing time series analysis



Using packages

1

```
install.packages("readr")
```

Downloads files to computer

1 x per computer

2

```
library("readr")
```

Loads package

1 x per R Session