Introduction to R

Drew Schmidt

April 2, 2015

Contents

- What is R?
- R Extensions
- R Interfaces
- R Resources

What is R?

What is R?

- Powerful, free, open source.
- Part programming language, part data analysis package.
- An implementation of the S language.
- Written by Ross Ihaka & Robert Gentleman.
- SAS/SPSS have <15 add-ons, R has over 6400!
- Most popular environment for developing new analysis and vis methods.

The SAS/SPSS Way

1. Data input and management
2. Stats and graphics procedures
3. Output management systems
4. Macro language
5. Matrix language

The R Way

All of these concepts are built into the (single) core language.

Comparing R to Other Tools

- R can be harder to "get going" than other analysis tools.
- But once you do. . .
- Hastily becoming the "lingua franca".

R Extensions

Extensions

- Community and commercial addons.

- Everything from sophisticated analyses to making your computer beep.
- Written in R, Fortran, C, C++ mostly.
- Performance can sometimes be an issue, but generally accurate.
- Package download trends: https://dgrtwo.shinyapps.io/cranview/

The CRAN

- Comprehensive R Archive Network.
- Open source only.
- Over 6400 packages.
- High standard of quality is vigorously enforced.
- Closest thing to a journal for stats software.

```r
install.packages("NameOfPackage")
```

GitHub

- Many authors eschewing CRAN for GitHub.
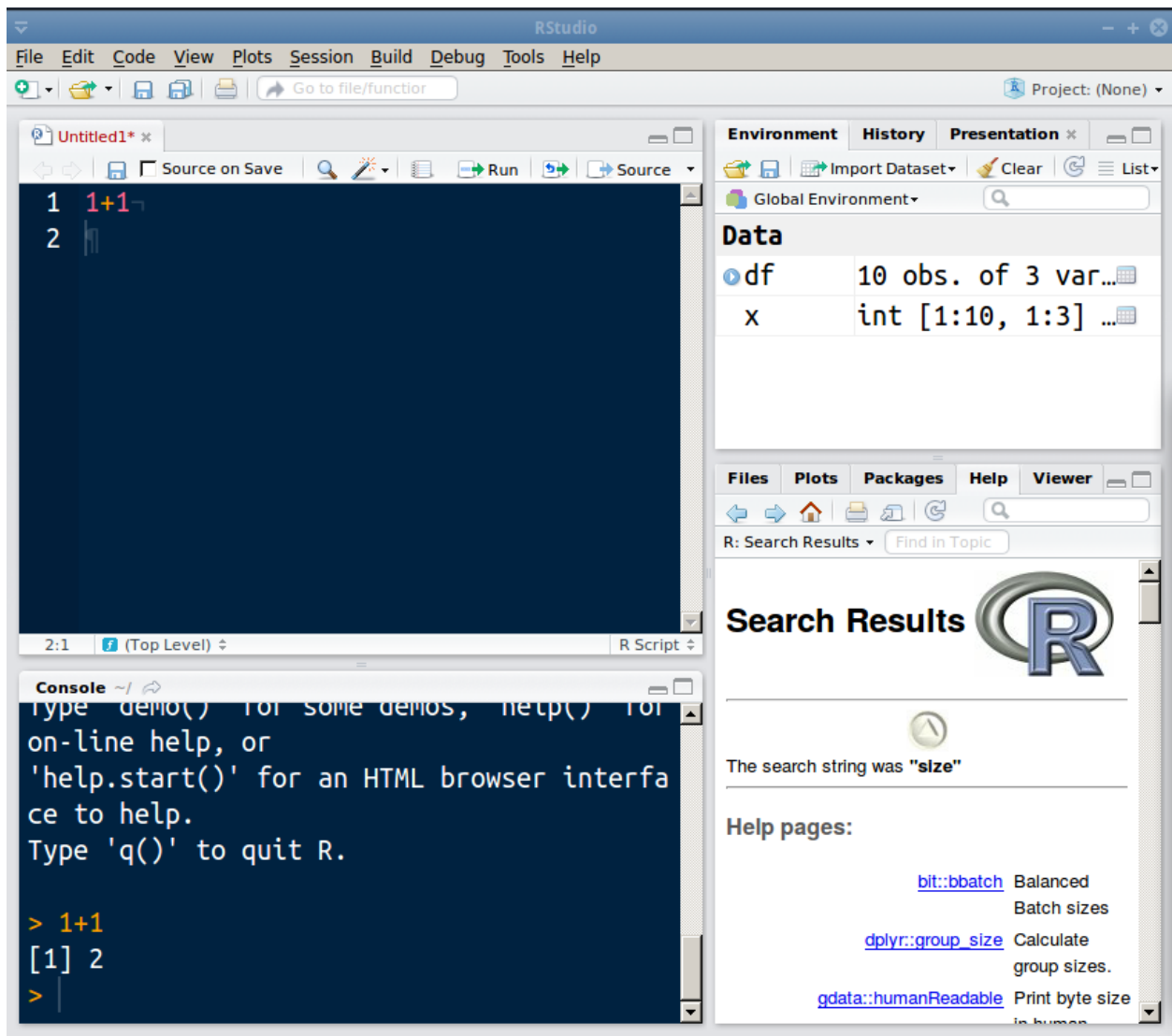- Faster bug fixes, faster bugs (maybe).

```r
install.packages("devtools")
library(devtools)
install_github("GitHubUserName/RepositoryName")
```

R Interfaces

R Interfaces

1. RStudio
2. R Commander
3. Deducer
4. Rattle
5. R from Excel
6. Shiny

RStudio

R Commander

```
╔══════════════════════════════════════════════════════════════╗
  7ん R Commander                          [ _ ] [ □ ] [ × ]

  File   Edit   Data   Statistics   Graphs   Models   Distributions   Tools   Help

  R     Data set:    myd         Summaries          ▶   data set    Model:   <No active model>
                                 Contingency tables ▶
  Script Window                  Means              ▶   Single-sample t-test...
  showData(mydata,                                      Independent samples t-test...   80, max
  summary(mydata)                Proportions        ▶   Paired t-test...
                                 Variances          ▶   One-way ANOVA...
                                 Nonparametric tests▶   Multi-way ANOVA...
                                 Dimensional analysis▶
                                 Fit models         ▶

                                                                                    [ Submit ]
  Output Window

  > summary(mydata)
      workshop      gender        q1              q2              q3
   Min.   :1.0    f   :3    Min.   :1.00    Min.   :1.00    Min.   :2.000
   1st Qu.:1.0    m   :4    1st Qu.:2.00    1st Qu.:1.00    1st Qu.:4.000
   Median :1.5    NA's:1    Median :3.50    Median :2.50    Median :4.000
   Mean   :1.5             Mean   :3.25    Mean   :2.75    Mean   :4.143
   3rd Qu.:2.0             3rd Qu.:4.25    3rd Qu.:4.25    3rd Qu.:5.000
   Max.   :2.0             Max.   :5.00    Max.   :5.00    Max.   :5.000
                                                           NA's   :1.000

  Messages

  NOTE: The dataset mydata has 8 rows and 6 columns.
```
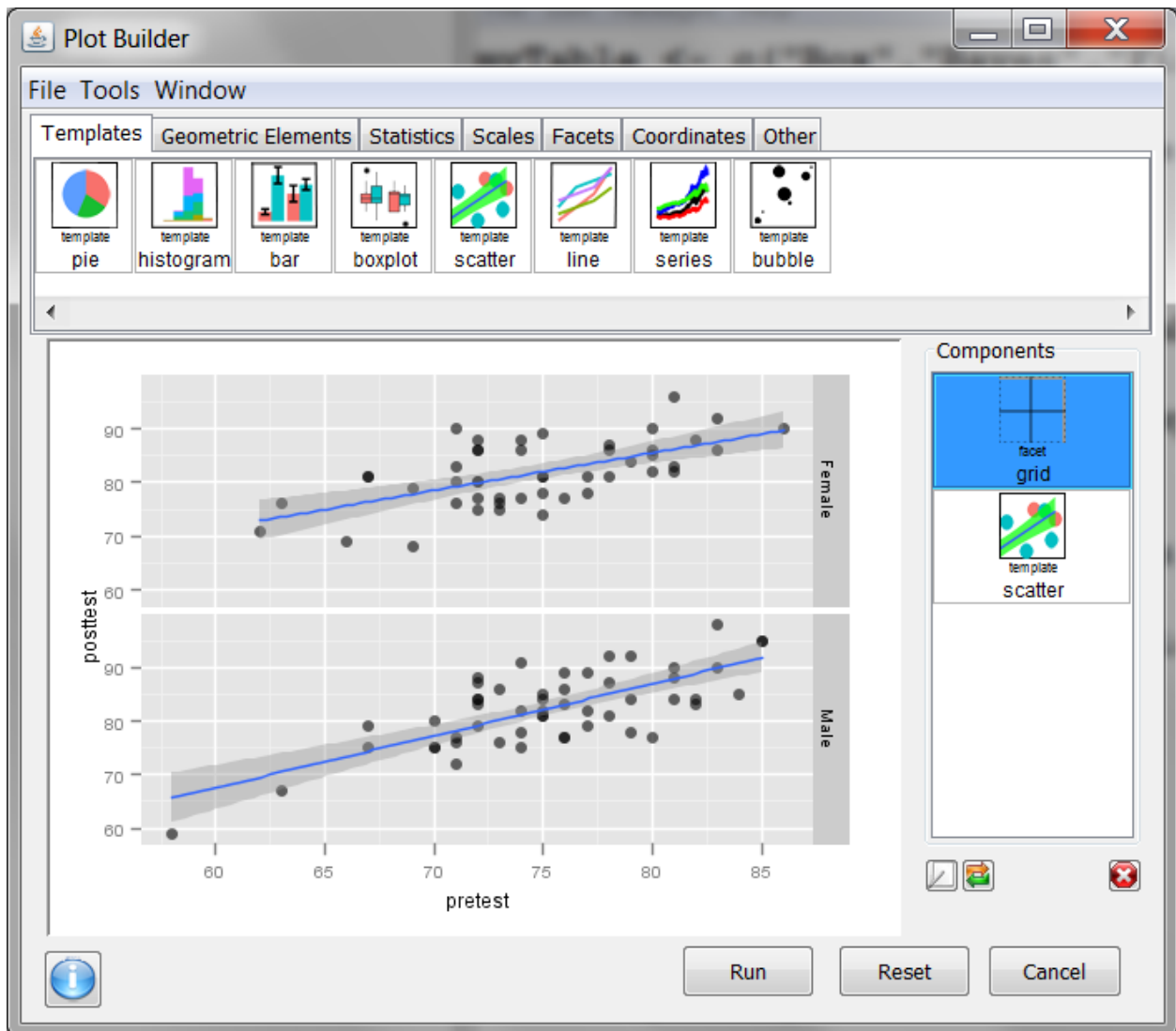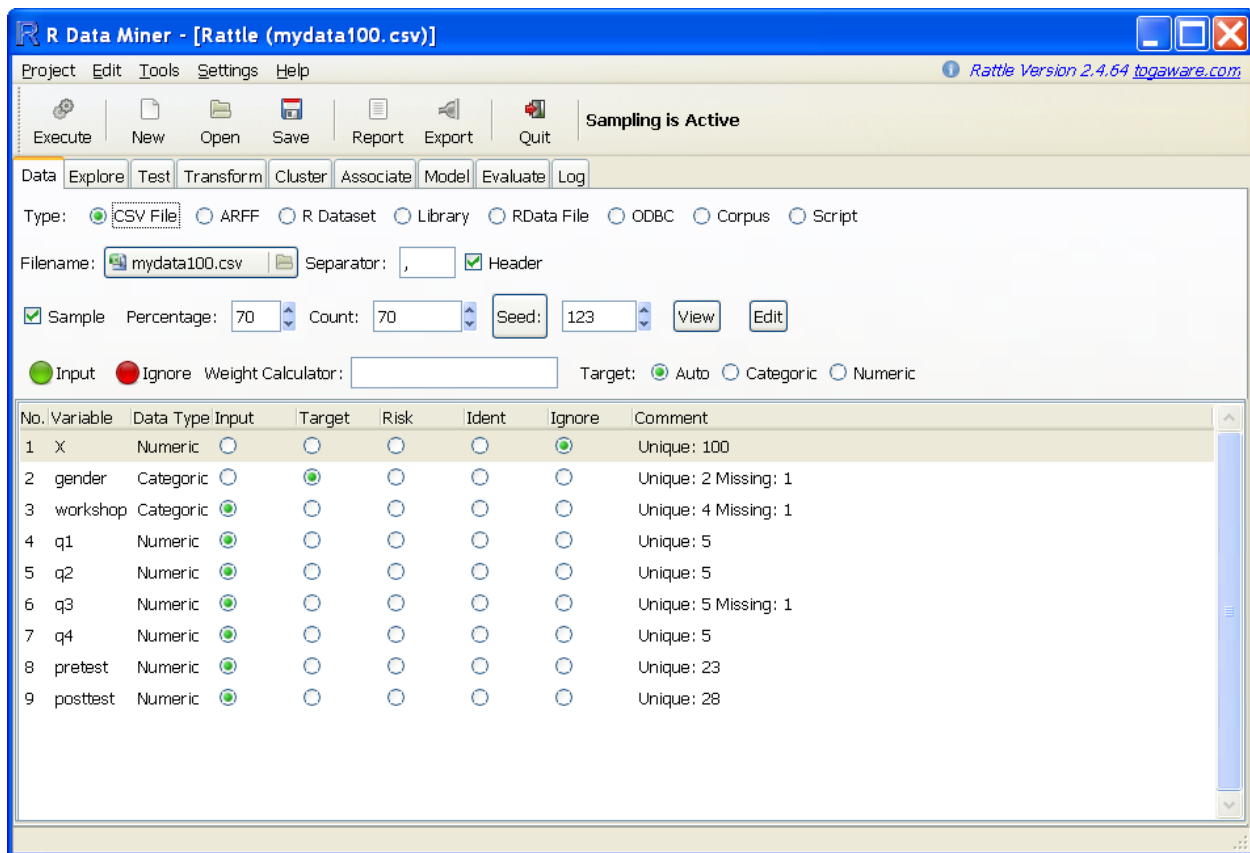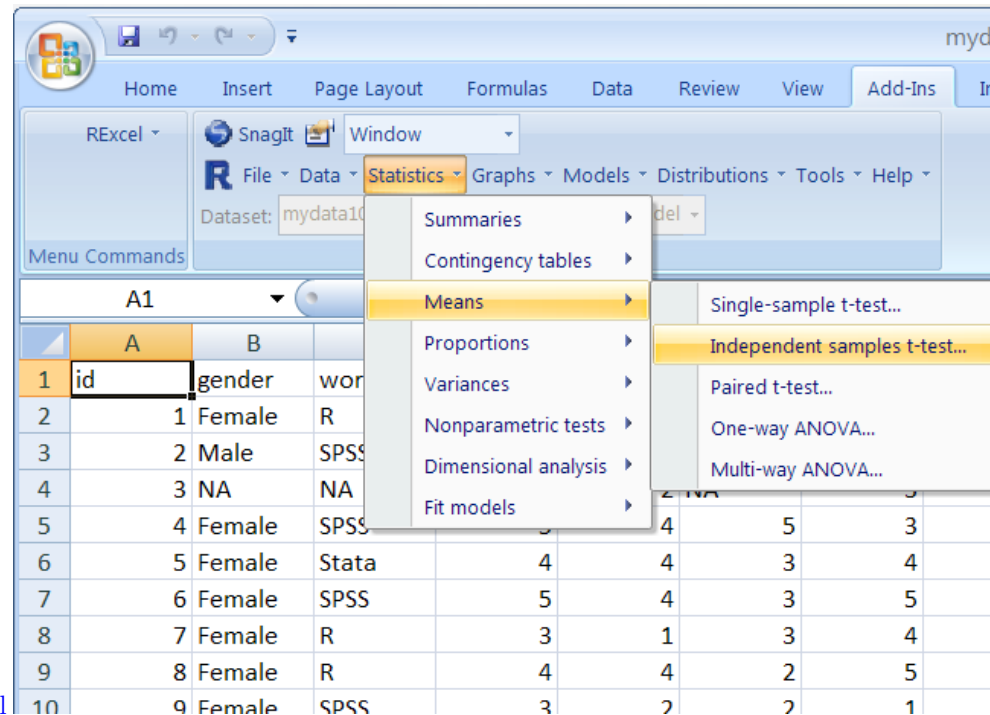
Deducer

Rattle

R from Excel

Shiny

- System for creating your own gui's (webapp)
- Code is 100% R (but can use html/css/javascript/. . .)
- Community quickly getting flooded with webapps

Shiny: Plot Your Data

Live Source code

Plot Your Data

Hello User!

Select Inputs | How To

Choose file to upload
Browse... iris.csv
Upload complete

☐ Use Example Data
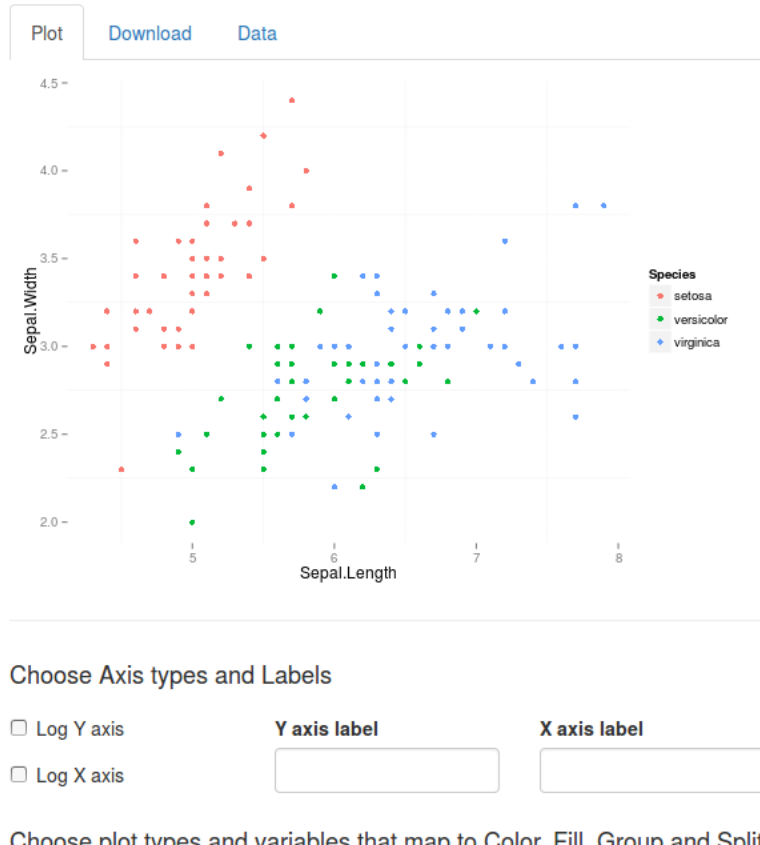
x variable:
Sepal.Length ▼

Include x range:
4.3 ———————————— 7.9
4.3   7.9

y variable:
Sepal.Width ▼

Include y range:
2 ———————————— 4.4
2   4.4

Max number of unique values for filter variables:
250 ▲▼

Plot | Download | Data

Species
• setosa
• versicolor
• virginica

Sepal.Width
Sepal.Length

Choose Axis types and Labels

☐ Log Y axis   Y axis label   X axis label
☐ Log X axis

Choose plot types and variables that map to Color, Fill, Group and Split

Shiny: Radiant

Live Source code

R Resources

Books

- *The Art of R Programming*, Norm Matloff
- *R for SAS and SPSS Users*, Muenchen
- *R for Stata Users*, Muenchen & Hilbe
- *An Introduction to R*
- *The R Inferno*

Comparisons to Other Languages

- aRrgh: a newcomer's (angry) guide to R
- Mathesaurus
- R programming for those coming from other languages

Online Help

- [r] tag on stackoverflow
- The mailing lists (r-help)
- #rstats on twitter

Questions?

This presentation is available at github.com/wrathematics/2015SFSURworkshop

Exercises are also available there.