Interactive Applications for Modeling and Analysis with Shiny
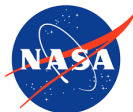
Presenter: Nicole Bishop
Cindy Fryer, Paul Guill

NASA GSFC Code 405

August 26, 2015

Introduction

- RStudio/Shiny offers quick and easy ways to build interactive applications for statistical models.

What is R?

- Programming language for data analysis and statistical computing
- Free and open source
- Widely used in academia, industry and the government
- Powerful language with flexible syntax
- A collection of over 7,000+ libraries
- Steep learning curve but a lot of learning materials available

RStudio

- Integrated Development Environment (IDE) for R
- Free and open source
- Includes R console, a code editor, file browser, help files and graphical display
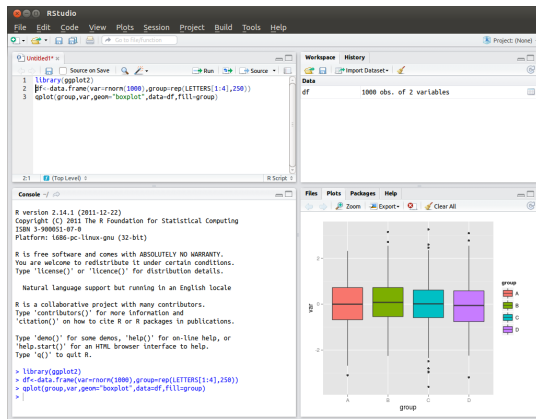
RStudio Advanced Features

- Version control with Git and SVN

- Reproducible reporting tools (knitr, Sweave)
  - Incorporate Markdown or LaTeX markup languages
  - Embed R code and equations into documents
  - Automatically regenerate if assumptions, data or code change
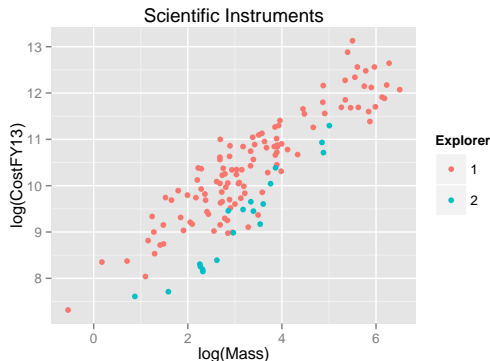 - Output as HTML, PDF or MS Word

Data Analysis

Data Preprocessing

- Read-in, clean and transform data

Exploratory Analysis

- Descriptive statistics
- Graphical displays

Scientific Instruments

```
InstData <- read.csv("Inst_data.csv", header = TRUE, row.names = 1)
summary(InstData)
InstData$Optical <- ifelse(InstData$InstFamily == "Optical", 1, 0)
InstData$Explorer <- as.factor(InstData$Explorer)
InstData$log_mass <- log(InstData$Mass)
ggplot(InstData, aes(log(Mass), log(CostFY13), color = Explorer)) + geom_point()
```

Fit and Evaluate Models

Model Selection and Validation

- Variable selection, dimension reduction methods
- Resampling methods (e.g. cross validation)

Model Fitting

- Model diagnostics

Predict New Observations

```
regfit <- regsubsets(log(CostFY13) ~ log(Mass) + log(Variable1) + log(Variable2) +
    log(Variable3) + Variable4 + Variable5, data = InstData, nbest = 3)
summaryHH(regfit)
lm_model <- lm(log(CostFY13) ~ log(Mass), data = InstData, x = TRUE, y = TRUE)
summary(lm_model)
plot(lm_model)
input_var <- data.frame(Mass = 50)
log_median <- predict.lm(lm_model, input_var, interval = "prediction")[1]
median <- exp(log_median)
```

What is Shiny?

- R package created by RStudio

- Framework for building browser-based applications

- Turns R code into interactive and dynamic displays

- Applications can be developed entirely with R code

- Can incorporate transparency

How to Make Analyses Interactive

```r
ggplot(iris, aes(x = Sepal.Length, fill = as.factor(Sepal.Width),
    group = as.factor(Sepal.Width))) + geom_density(alpha = 0.75) +
    theme(axis.line = element_line(colour = "gray", size = 0.75),
        panel.background = element_blank(), plot.background = element_blank())
```
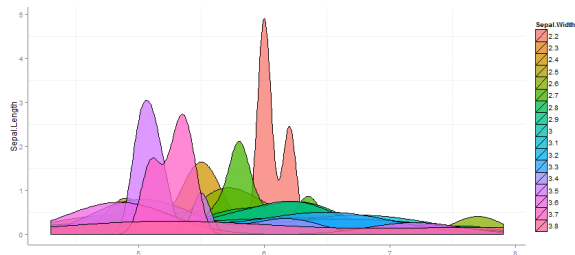

Select Options

Density plot

Motivation for Using Shiny

Transparent

- Not a black box

Reproducible

- Ease of model validation

Flexible

- Users interactively communicate with the data and analysis
- Can very quickly develop multiple models with the same database
- Tons of R packages available

Building a Shiny Application

ui.R : controls look and layout of the app

server.R: instructions used to build the app
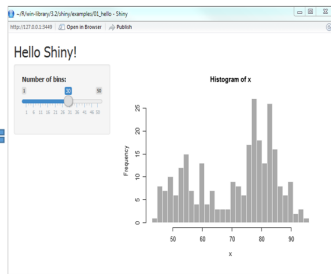
Additional Features of Shiny

- Interactive visualizations, data manipulation, pivot tables, download datasets

- Dynamic report generation using knitr

- Fully extensible and customizable with HTML, JavaScript and CSS

Deploying Shiny Applications

Various Options Available

- Run locally, requires R software to be installed

- Enterprise hosting service (Linux server, RStudio's Shiny server)

Shiny Demo

- Development of instrument cost model using NASA's ONCE database

- RStudio
    1. Data analysis
    2. Model development
    3. Generate report

- Shiny Application

Conclusion

- Shiny is an economical (free)

- Shiny provides an easy way for R users to build interactive applications for statistical modeling

- Easy for non-R users to use

- R/RStudio/Shiny is under rapid development and is gaining popularity, resulting in more and more capabilities being added

Try it yourself!

- Download at
  `https://www.rstudio.com/products/rstudio/download/`

References

http://www.rstudio.com/

http://rmarkdown.rstudio.com/

http://shiny.rstudio.com/