

Aug 14, 2008

R AnalyticFlow

A flowchart-style GUI for R

Ryota Suzuki / Ef-prime, Inc.

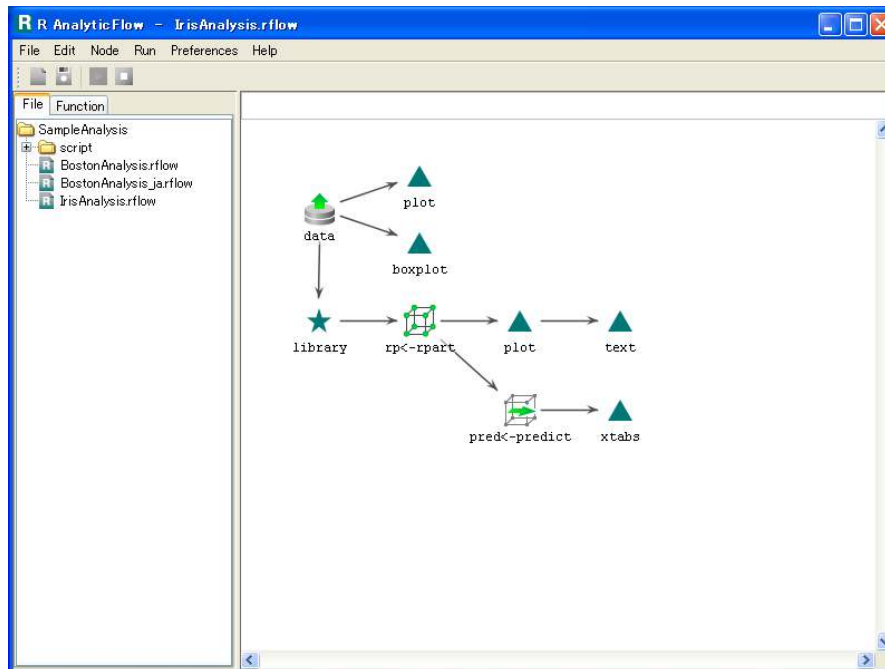
Introduction

- This talk contains...
 - This is “R AnalyticFlow”
 - What’s good?
 - About the software
 - Demo
 - Where are we heading?



This is “R AnalyticFlow”

Main window



Console window

```

R R console
Command
Copyright (c) 2000-2011 R Foundation for Statistical Computing
ISBN 3-900051-07-0

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.

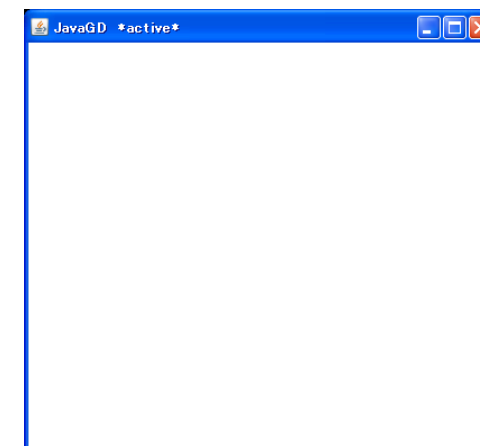
Natural language support but running in an English locale

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.

>
  
```

Graphic window



This is “R AnalyticFlow”



The screenshot displays the R AnalyticFlow interface. The main window, titled "R AnalyticFlow - IrisAnalysis.rflow", shows a project structure on the left and a central workspace with a flow diagram. The flow diagram consists of nodes: "data" (a cylinder icon), "library" (a star icon), "plot" (a triangle icon), "boxplot" (a triangle icon), "rp<-rpart" (a cube icon), "pred<-predict" (a cube icon), "xtabs" (a triangle icon), and another "plot" (a triangle icon). A dashed orange box highlights the "data", "plot", and "boxplot" nodes. A callout box below this diagram shows a simplified version of the "data" node branching into "plot" and "boxplot".

To the right, the "R console" window shows the following text:

```
Command
copyrig (c) 2000-2011 R Foundation for Statistical Computing
ISBN 3-900051-07-0

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.

>
```

Below the console window, a text box contains the text: "This is an “analysis flow”".

This is “R AnalyticFlow”



“Run” on a node...

The screenshot displays the R AnalyticFlow interface. The main window shows a workflow diagram with nodes: 'data' (disk icon), 'library' (star icon), 'plot' (triangle icon), 'boxplot' (triangle icon), 'rp<-rpart' (cube icon), 'pred<-predict' (cube icon), and 'xtabs' (triangle icon). A context menu is open over the 'library' node, listing actions: 'Run' (highlighted), 'Clear and Run', 'Copy Script', 'Remove Node', 'Detach', 'Branch', 'Select Following Subflow', and 'Remove Edge'. The 'Run' option is highlighted. The R console window shows the command `plot(iris[, 1:4], col = as.integer(iris$species) + 1)` and its output, including the R license text. Below the console is a JavaGD window.

This is “R AnalyticFlow”



The screenshot shows the R AnalyticFlow interface. The main window displays a workflow diagram with nodes: data, library, rp<-rpart, plot, boxplot, and text. The R console window shows the following output:

```

Command

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.

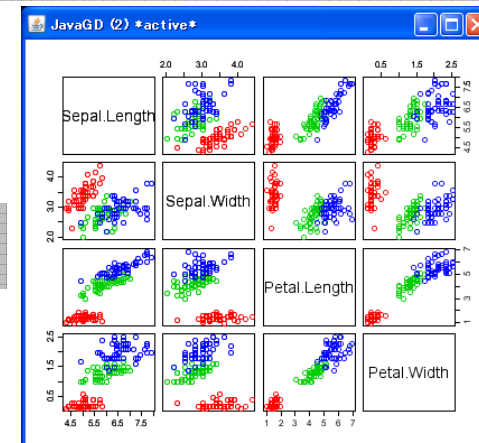
Type 'demo()' for some demos, 'help()' for on-line help, or
'help.start()' for an HTML browser interface to help.
Type 'q()' to quit R.

> data(iris)
> plot(iris[, 1:4], col = as.integer(iris$species) + 1)
>
  
```

```

> data(iris)
> plot(iris[, 1:4], col = as.integer(iris$species) + 1)
  
```

R codes are executed!



What's good?

■ Visual.

- Easier to think, share and remember.
- Once created, anyone can use it.

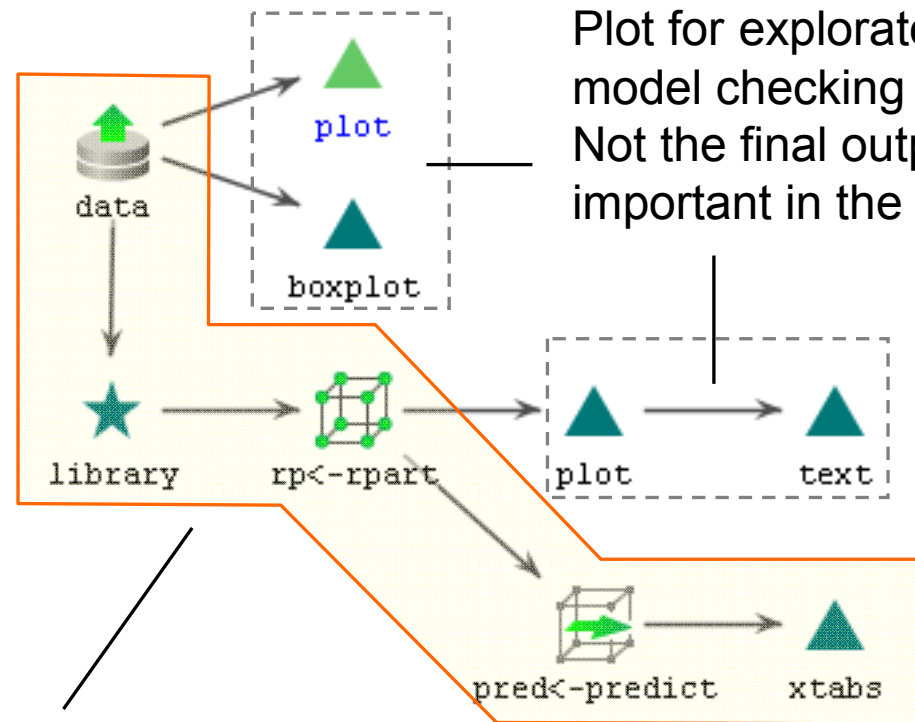
■ Flexible.

- Data analysis is not always straightway.
- It can represent “branching” processes.



▶ Branching?

“Branching”



Branches:

Plot for exploratory analysis, model checking etc.
Not the final outputs, but important in the analysis process.

Mainstream:

Loading data and library, modeling and prediction.
Prediction result is the final output.

About the software

■ System requirements

- OS: Windows, Linux and Mac OS X.
 - Linux version runs on Mac OS X.
- Java ≥ 5
- R $\geq 2.5.0$
 - Required packages: rJava, JavaGD and codetools.
 - For multi-byte Windows environment, modified library pack is required (currently available for R $\leq 2.6.2$).



About the software

■ Development

- Java
- Many open source softwares... thank you!
 - R, JRI, rJava, JavaGD, JUNG, Apache libraries, etc.



About the software

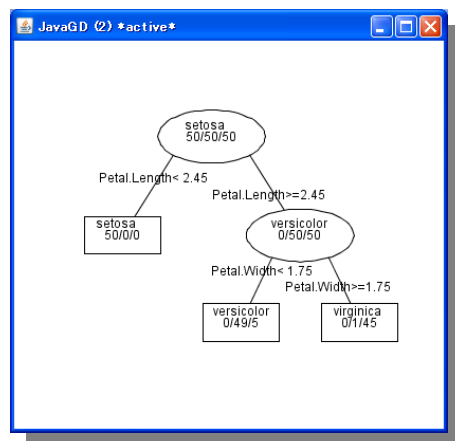
■ Availability

- Freely available!
- Open sourced under the BSD license.
- Download from:



<http://www.ef-prime.com/>

Demo



R console

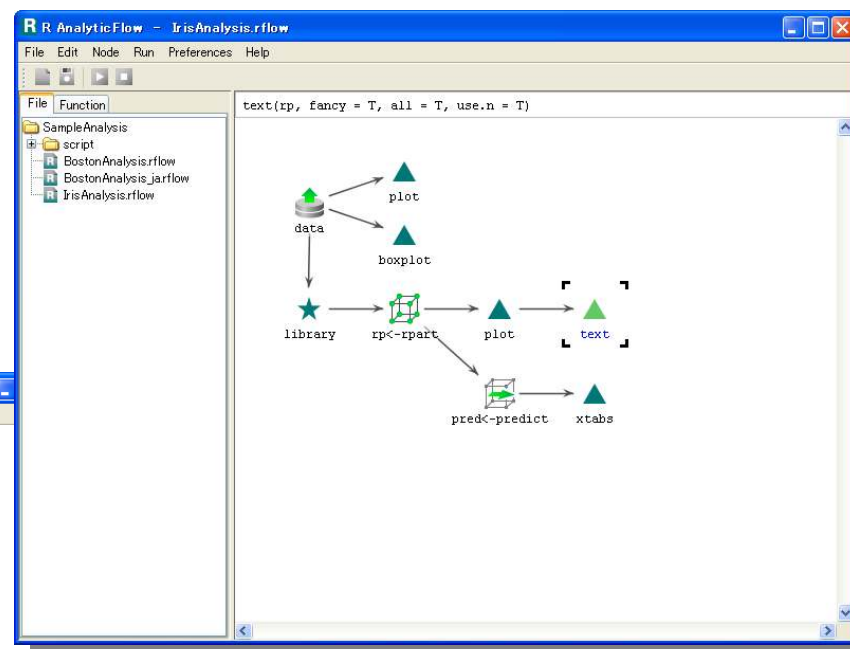
```

Command
> data(iris)
> library(rpart)
> rp <- rpart(Species ~ ., iris)
> plot(rp, margin = 0.2, branch = 0.3)
> text(rp, fancy = T, all = T, use.n = T)
>
  
```

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.



Where are we heading?

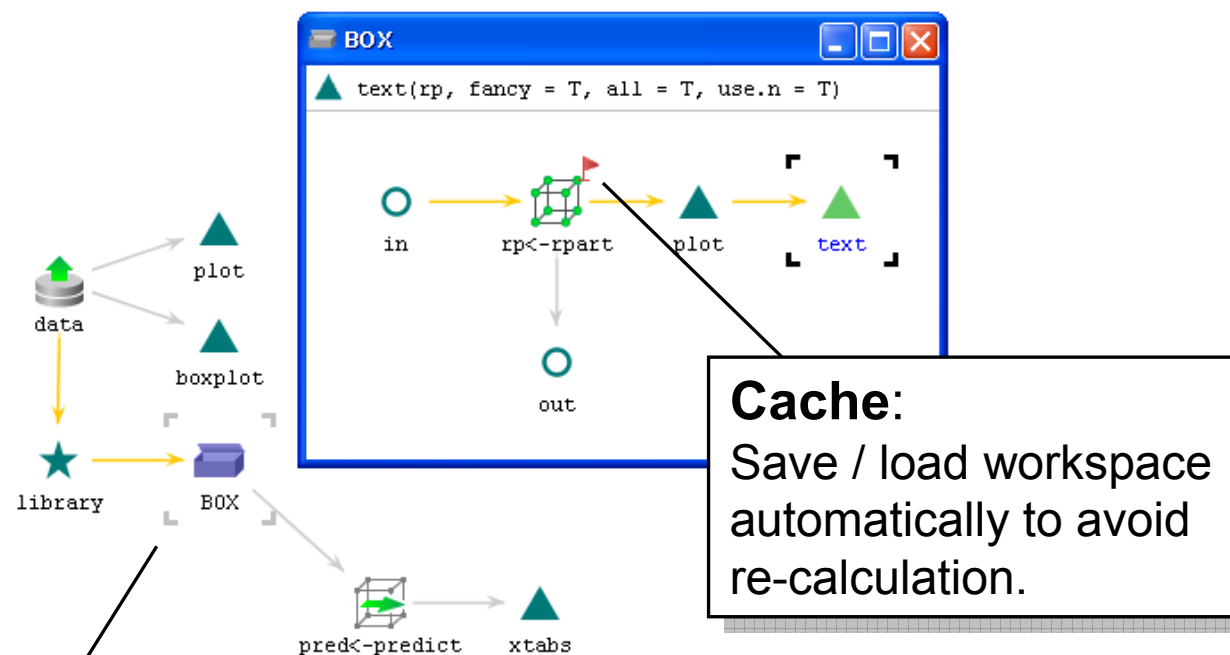
■ Current status

- “Preview edition”
 - Basic functions are well-developed and tested
 - Many advanced functions are under development



Where are we heading?

- Next version
 - A milestone version with important new functions



Box:
Contains a subflow nestedly.

Where are we heading?

■ About us: Ef-prime, Inc.

- A data analysis company
 - Location: Tokyo, Japan
- Services:
 - Data analysis consulting
 - Software development
 - Training



■ About me...

- Ryota Suzuki
 - An author of package “pvclust”
 - Running the company with friends

Where are we heading?

■ Our motivation

- Our needs
 - R is flexible, reliable and open
 - One of the best choice, but we need more
- Not only a useR!
 - We hope to commit to R, with our viewpoint of business data analysis
 - There will be much in common with other fields of application!



Where are we heading?

■ Need your help

- Maintenance is tough.
 - For a small company, especially without charge
- Any kind of help is welcomed!
 - Bug / typo report
 - Comments, requests
 - Request for paid support
 - > Data analysis
 - > Software customization
 - > Documentation, etc.



Thank you!

- Enjoy your analysis with...



<http://www.ef-prime.com/>

mail: contact@ef-prime.com