Lab : CL-VII, ML

Assignment 1
Study of platform for Implementation - R

AIM

Download the open source software of your interest. Document the distinct features and functionality of the software platform. You may choose WEKA or R or Rjava. I choose R.

Introduction

What is R ?

R is a programming language and free software environment for statistical computing and graphics supported by the R Foundation for Statistical Computing. The R language is widely used among statisticians and data miners for developing statistical software and data analysis

It was created by Ross Ihaka and Robert Gentleman (Hence the name 'R') at the University of Auckland, New Zealand, and is currently developed by the R Development Core Team, of which Chambers is a member. R is named partly after the first names of the first two R authors and partly as a play on the name of S.R is a GNU project.

Its Features

As stated earlier, R is a programming language and software environment for statistical analysis, graphics representation and reporting. The following are the important features of R –

- R is a well-developed, simple and effective programming language which includes conditionals, loops, user defined recursive functions and input and output facilities.
- R has an effective data handling and storage facility,
- R provides a suite of operators for calculations on arrays, lists, vectors and matrices.
- R provides a large, coherent and integrated collection of tools for data analysis.

- R provides graphical facilities for data analysis and display either directly at the computer or printing at the papers.

Installation

1. Install r-base : Write this command in Command Prompt : sudo apt-get install r-base

```
(base) ayan_gadpal@AyanGadpal:~$ sudo apt-get install r-base
[sudo] password for ayan_gadpal:
Reading package lists... Done
Building dependency tree
Reading state information... Done
The following additional packages will be installed:
```

2. Type R on terminal/Command line to get command line for R programming.

```
(base) ayan_gadpal@AyanGadpal:~$ R

R version 3.6.3 (2020-02-29) -- "Holding the Windsock"
Copyright (C) 2020 The R Foundation for Statistical Computing
Platform: x86_64-pc-linux-gnu (64-bit)

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.

>
```

IDE:

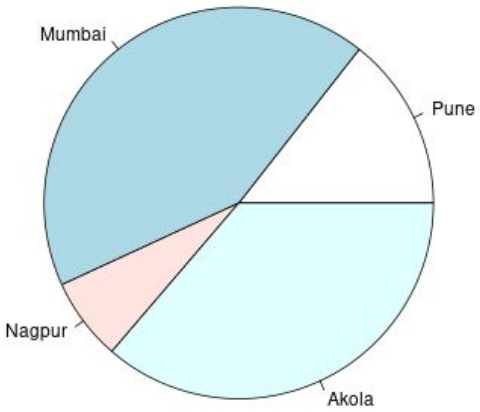
R Studio : RStudio provides free and open source tools for R and enterprise-ready professional software for data science teams to develop and share their work at scale.

- Access RStudio locally.
- Syntax highlighting, code completion, and smart indentation.
- Execute R code directly from the source editor.
- Quickly jump to function definitions.
- Easily manage multiple working directories using projects.
- Integrated R help and documentation.

Demo

A very simple pie-chart is created using just the input vector and labels. The below script will create and save the pie chart in the current R working directory, Which is root in my case.

```
> # Demo By 43308
> # Create data for the graph.
> x <- c(21, 62, 10, 53)
> labels <- c("Pune", "Mumbai", "Nagpur", "Akola")
>
> # Give the chart file a name.
> png(file = "city.png")
>
> # Plot the chart.
> pie(x,labels)
>
> # Save the file.
> dev.off()
null device
```



Conclusion

As a conclusion, R is world’s most widely used statistics programming language. It's the # 1 choice of data scientists and supported by a vibrant and talented community of contributors. R is taught in universities and deployed in mission critical business applications.