Assignment 5
Study of Open Source Analytical Software

AIM
Study of platform for Implementation of Assignments. Download and install Python, and the open source software - WEKA and R. Document the distinct features and functionality of the software platform.

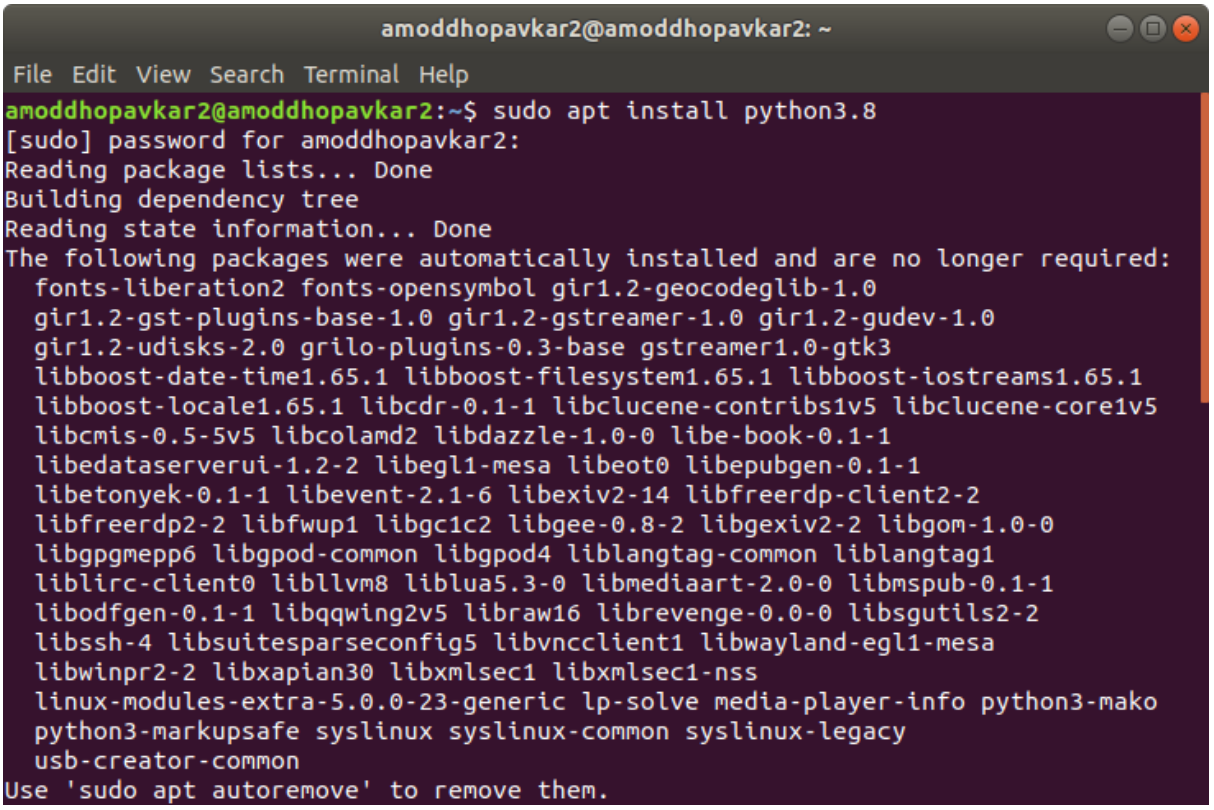
OBJECTIVE
To study -
- Concept of open source analytical software(WEKA and R).
- Concept of statistical analysis.
- Distinct features and functionality of open source software.

THEORY
INTRODUCTION TO PYTHON
Python is an interpreted, high-level and general-purpose programming language. Created by Guido van Rossum and first released in 1991, Python's design philosophy emphasizes code readability with its notable use of significant whitespace.

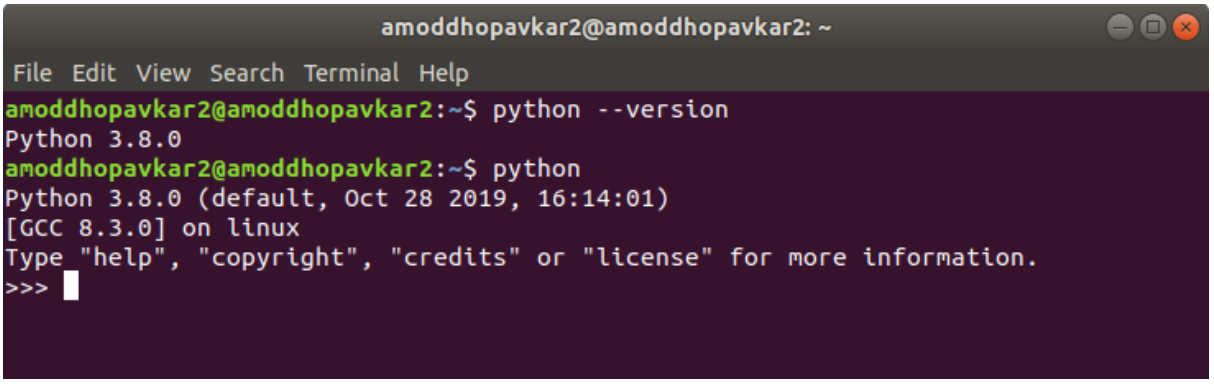
Features -
- Easy to Code. Python is a very developer-friendly language which means that anyone and everyone can learn to code it in a couple of hours or days.
- Open source and free.
- GUI support.
- Object Oriented approach.
- High-level language.
- Highly portable.
- Highly dynamic.
- Integrated by nature.

Installation -

- Install python with the command “sudo apt install python3.8”.



- Check the installed version and run the python shell using commands “python --version” and “python”.



INTRODUCTION TO 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.

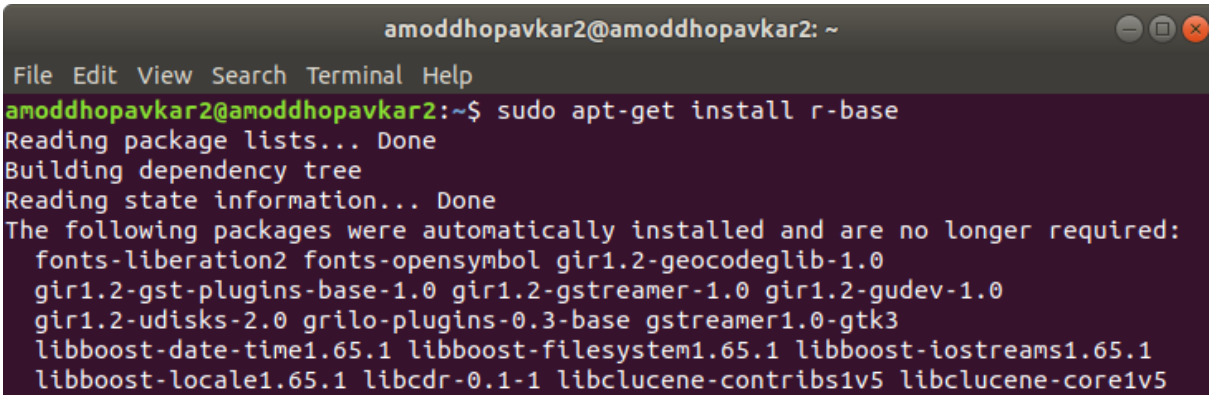
It was created by Ross Ihaka and Robert Gentleman (Hence the name 'R') at the University of Auckland, New Zealand, and is currently developed by the R Development Core Team, of which Chambers is a member. R is named partly after the first names of the first two R authors and partly as a play on the name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

Features -

- 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 -

- Install R with the command "sudo apt-get install r-base".

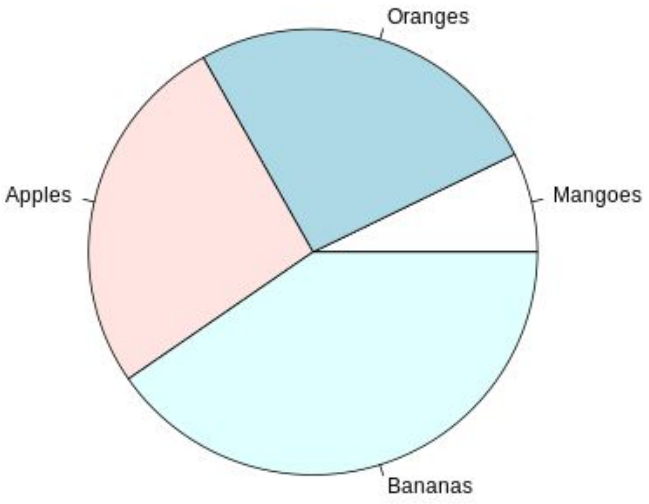


- Type “R” on terminal/Command line to get command line for R programming.

```
amoddhopavkar2@amoddhopavkar2: ~
File  Edit  View  Search  Terminal  Help
amoddhopavkar2@amoddhopavkar2:~$ R

R version 3.4.4 (2018-03-15) -- "Someone to Lean On"
Copyright (C) 2018 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.
```

- Demo - Create a simple pie chart in R for the sale of various fruits.

```
> # demo by amoddhopavkar2
> # A basic pie chart
> x <- c(12, 43, 44, 67)
> labels <- c("Mangoes", "Oranges", "Apples", "Bananas")
>
> # Name the file
> png(file = "fruits.png")
>
> # Plot the chart
> pie(x,labels)
>
> # Save the file
> dev.off()
null device
```



INTRODUCTION TO WEKA

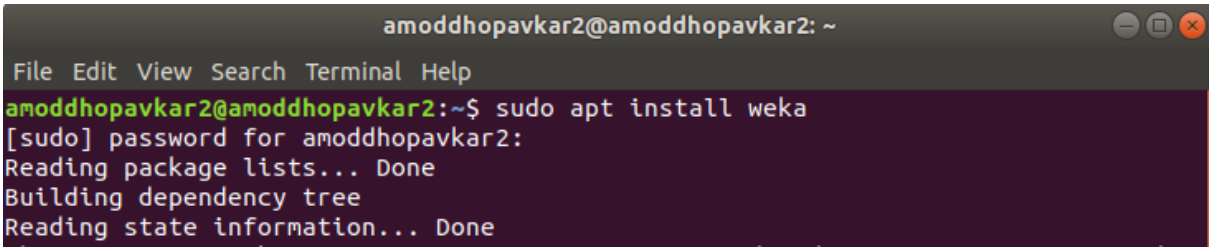
WEKA is open source software under the GNU General Public License.The system is written using object oriented language Java. There are several different levels at which WEKA can be used. Weka provides implementations of state-of-the-art data mining and machine learning algorithms. WEKA contains modules for data preprocessing, classification, clustering and association rule extraction.

Features -

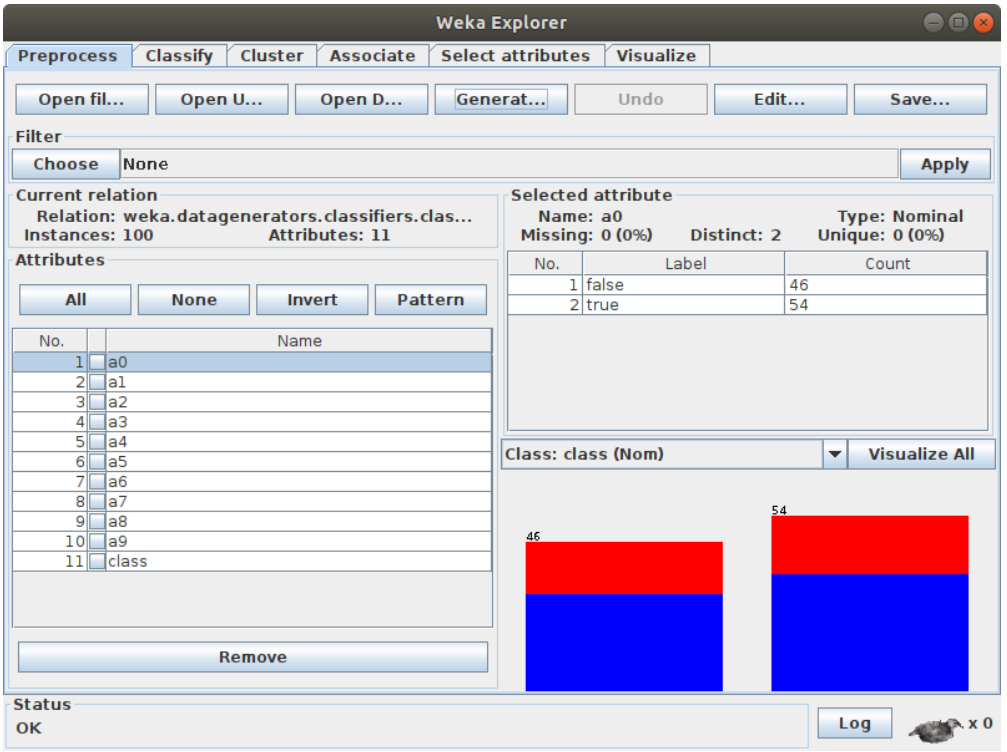
- Preprocess
- Classify
- Cluster
- Associate
- Select attributes
- Visualize

Installation -

- Install using the command “sudo apt install weka”



- Demo



CONCLUSION

Downloaded Python and the open source softwares R-base, RStudio and WEKA, and studied the distinct features and functionality of these software platforms.