Basic R

Learn by doing (90% hands on)

About Me

- Name: Shiv

- BSc and MSc in Mathematics from IIT Kharagpur

- Ph.D in Analytics

- Experience: 19 years (9 Years in Analytics and 10 years in IT)

- Current Role: Chief Data Scientist
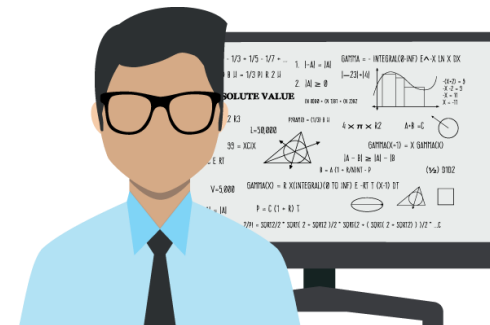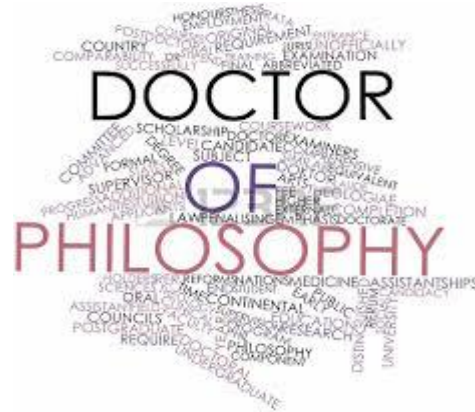
Prerequisites

- Passion for learning
- Good Knowledge of any programming technology

Contents

Introduction

- Installation
- Environment Walkthrough
- How to get help
- Console and Editor
- History
- Working directory

Data structure

- Various types Variables
- Tensors (Vectors)
- Vectors
- List
- Matrices
- Data frames
- Using Cbind, Rbind
- Regular Expressions
- Handling dates

Importing data

- Reading/Saving Data files
- Reading/Saving data objects

Manipulating Data

- Selecting rows/observations
- Selecting columns/fields
- Merging data
- Relabeling the column names
- Converting variable types
- Data sorting
- Data aggregation
- Pivot Tables

Using functions

- Mathematical Functions
- Summary Functions
- String Functions
- User defined functions
- Apply functions

Programming operations

- Truth testing
- Branching
- Looping
- Jump commands
- Arithmetic operations

Charts and Plots

- Box plot
- Histogram
- Line chart
- Scatter plot
- Save images

Miscellaneous

- Logging the various types of messages
- Exceptional Handling in details
- Error messages
- How to get going in open source
- How to learn new module in open source

Basic Statistics

- Measures of Central Tendency
- Measures of Dispersion
- Measure of Shape
- Sampling
- Distributions
- Descriptive Statistics
- Inferential Statistics

All are Hands on topics

Methodology

Content

Theory 10%

Hands on, 90%
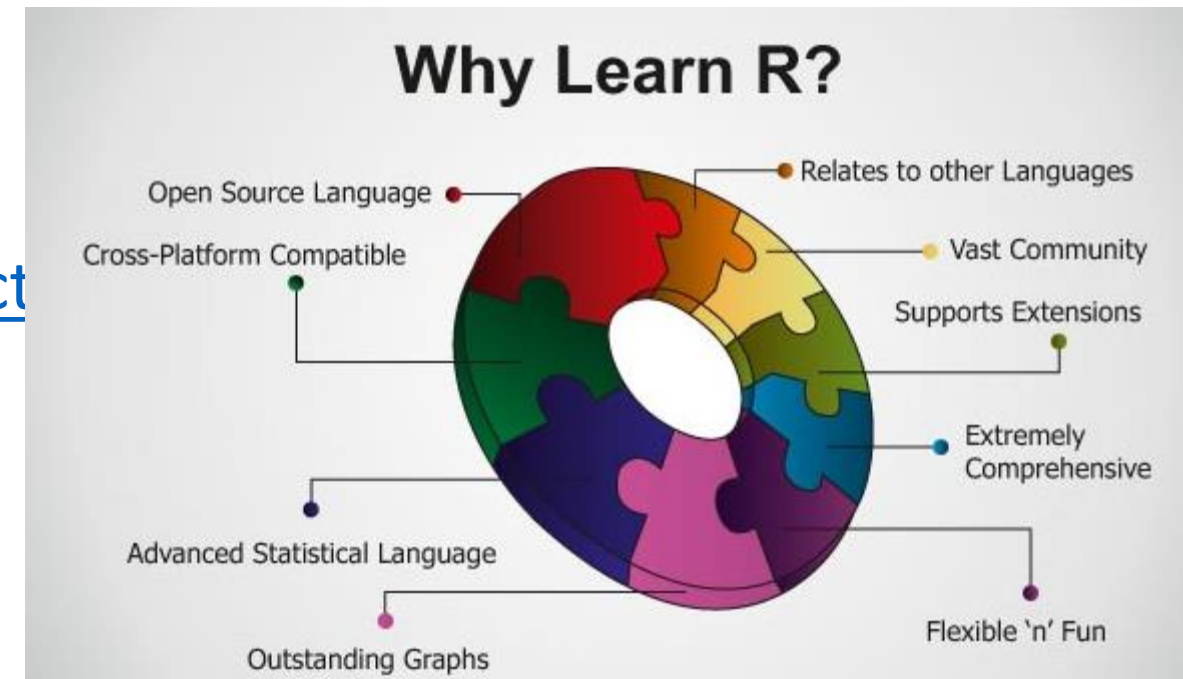
Analytics

Functional, 20%

Technical, 80%
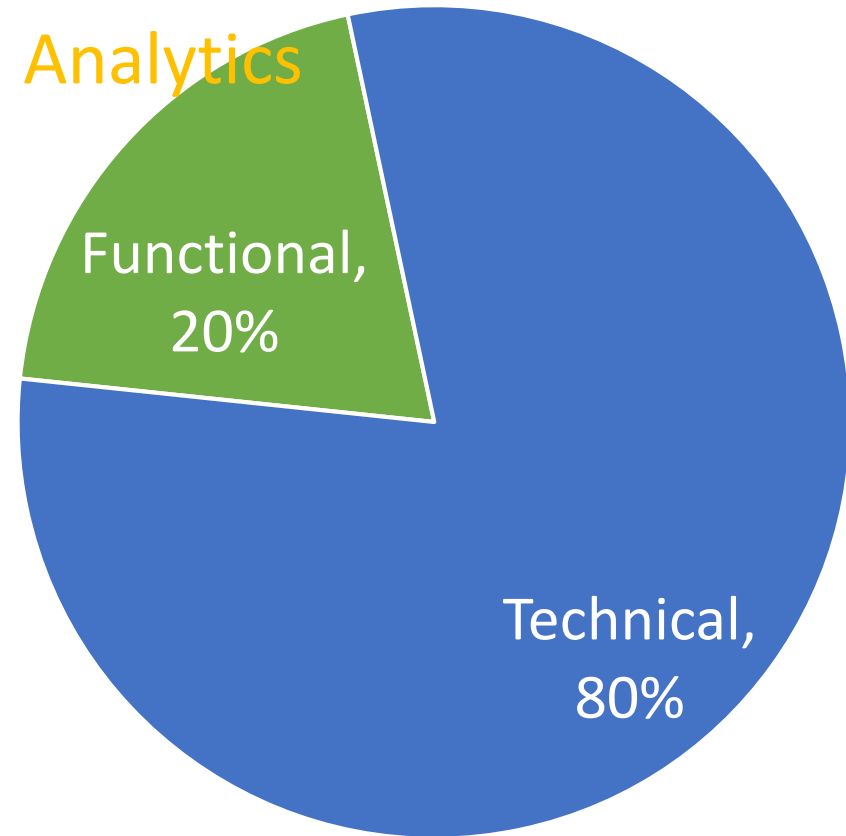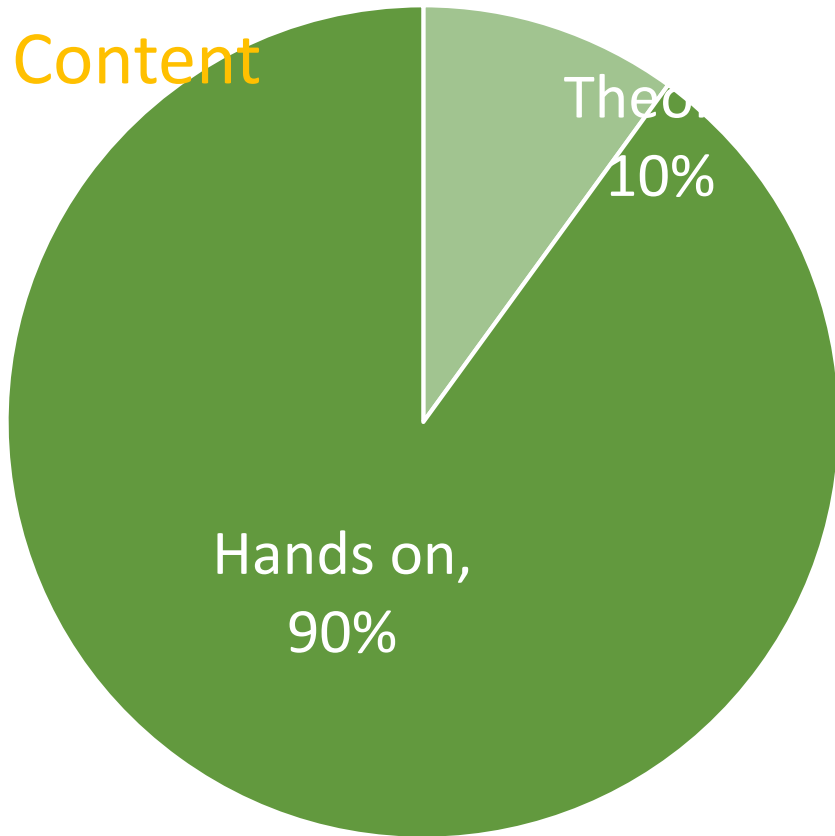
History of R

Birthplace

- Designed by: Ross Ihaka and Robert Gentleman

- Developer: R Core Team and …

- First appeared: August 1993

- Stable release: See at https://cran.r-project

- License: GNU GPL v2[4]

- Source code: C, Fortran, and R

Why Learn R?

- Open Source Language
- Cross-Platform Compatible
- Relates to other Languages
- Vast Community
- Supports Extensions
- Extremely Comprehensive
- Flexible 'n' Fun
- Outstanding Graphs
- Advanced Statistical Language

- Website
  - www.r-project.org
  - https://cran.r-project.org/

- Rstudio as IDE
  - https://www.rstudio.com/

- Rstudio Server

Introduction

- Installation
  - Software: R_Software.txt
  - Libraries: Libraries.R

- Environment Walkthrough
  - R
    - https://cran.r-project.org/
  - Rstudio as IDE
    - https://www.rstudio.com/
  - Rstudio Server
  - Works in RAM

- How to get help
  - Website
    - www.r-project.org
    - https://cran.r-project.org/

- Console and Editor

- Working directory

Various data types

Basic concepts of linear algebra
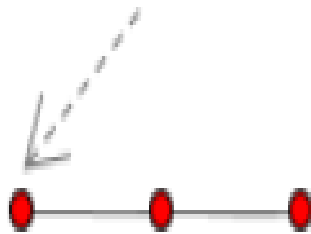
element (number)
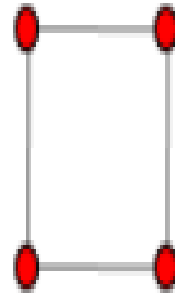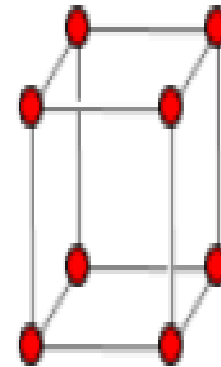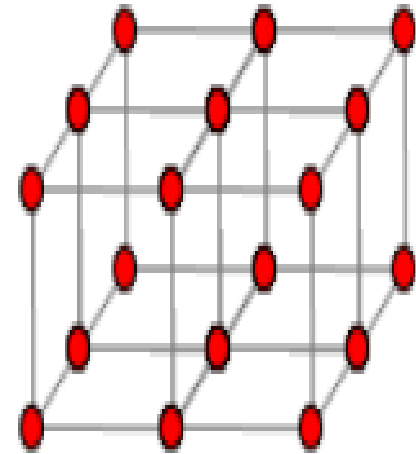
Order 0	Order 1	Order 2	Order 3	Order N
Scalar	Vector	Matrix	Tensor	Tensor

Move to R Studio

Move to R Studio for Hands on

References

- https://regex101.com/

- https://www.google.com/imgres?imgurl=https%3A%2F%2Fwww.nervanasys.com%2Fwp-content%2Fuploads%2F2016%2F04%2FScreen-Shot-2016-04-27-at-10.59.50-AM.png&imgrefurl=https%3A%2F%2Fwww.nervanasys.com%2Fopenai%2F&docid=GggjYZ3JzfRoVM&tbnid=teq-_Mt_keBQiM%3A&vet=10ahUKEwiDqYXb-7_UAhWLs48KHd2dD00QMwh2KC8wLw..i&w=758&h=423&bih=632&biw=1366&q=Reinforced%20learning&ved=0ahUKEwiDqYXb-7_UAhWLs48KHd2dD00QMwh2KC8wLw&iact=mrc&uact=8

- https://images.google.com/

- https://www.google.com/imgres?imgurl=https%3A%2F%2Fcdn.edureka.co%2Fblog%2Fwp-content%2Fuploads%2F2013%2F06%2FData-Scientist.jpg&imgrefurl=https%3A%2F%2Fwww.edureka.co%2Fblog%2Fwho-is-a-data-scientist%2F&docid=q3ij004thhLBlM&tbnid=6_eIIKVKkQtWlM%3A&vet=10ahUKEwi74J_L7sbUAhWHRY8KHYcUBbUQMwi8AShFMEU..i&w=601&h=351&bih=632&biw=1366&q=Data%20Science&ved=0ahUKEwi74J_L7sbUAhWHRY8KHYcUBbUQMwi8AShFMEU&iact=mrc&uact=8

- https://en.wikipedia.org/wiki/R_(programming_language)

- http://www.chioka.in/differences-between-roc-auc-and-pr-auc/

- https://www.google.com/imgres?imgurl=http%3A%2F%2Fstanford.edu%2F~cpiech%2Fcs221%2Fimg%2FkmeansViz.png&imgrefurl=http%3A%2F%2Fstanford.edu%2F~cpiech%2Fcs221%2Fhandouts%2Fkmeans.html&docid=xBG90gMlnM_nKM&tbnid=0dV6dbzMcO1mgM%3A&vet=10ahUKEwiv8PWX49HUAhXKtI8KHVk3BigQMwhSKAUwBQ..i&w=501&h=338&bih=632&biw=1366&q=K-means&ved=0ahUKEwiv8PWX49HUAhXKtI8KHVk3BigQMwhSKAUwBQ&iact=mrc&uact=8

- https://www.otexts.org/

- robjhyndman.com/hyndsight/forecasting/

- OTexts.org/fpp

- http://www.techjini.com/wp-content/uploads/2017/02/mc-learning.jpg

- https://www.r-bloggers.com/six-reasons-to-learn-r-for-business/?utm_source=feedburner&utm_medium=email&utm_campaign=Feed%3A+RBloggers+%28R+bloggers%29

- https://www.datacamp.com/community/blog/r-xts-cheat-sheet

- https://machinelearningmastery.com/improve-deep-learning-performance/

In case you want to learn ML using R online

- Course Name: Data science for AI and Machine learning using R
- Course Link: https://www.udemy.com/data-science-for-ai-and-machine-learning-using-r/?couponCode=AIMLR10
- Why this course: Targeted to corporate requirements (different from traditional academic) so that people have the basic understanding of AI and ML using R
- Cost: $10 (93% discount with Coupon AIMLR10)
- Price validity: Few days as per TnC of Udemy