

Online Certificate Course on Descriptive Statistics Using R Language

Introduction and Background:

In the digital era every activity related to digital gadgets is leaving digital footprints behind generating humongous set of data .This data can be leveraged to generate a data driven decision systems necessary for smooth working of complex and data hungry systems of the digital world . In other words data is like oxygen in digital era and thus necessary to learn , understand and efficiently handle the data . One should know basics of Statistics to understand data and a Computer programming language to handle that data efficiently. In view of this we propose to start a entry level online course titled as “ Descriptive Statistics using R language”.

R is a programming language and free software environment for statistical computing and graphics supported by the R Foundation for Statistical Computing and fully compatible with Linux, Mac OS and Windows OS . The R language is widely used among statisticians and data miners for developing statistical software and data analysis. R language was developed by Ross Ihaka and Robert Gentleman at Aukland University in around 1992.

Eligibility:

Any student of Statistics or Mathematics (UG/PG) interested in learning data summarization tools and R language environment and its capabilities.

Prerequisites:

Be able to use computers and know basic Statistics and Mathematics concepts.

Duration of the Course:

THREE WEEKS (16th Aug 2020 to 3rd Sept 2021)

Five video lectures will be released and assignments will be given in one week. The lectures can be viewed at any time during that week and assignments shall have to be submitted before given deadline. The questions and discussions facility will be available through Google Classroom.

After successful completion of the course a certificate will be issued by Shri Shivaji Science College, Amravati depending on the performance.

Course Content

- Introduction R language
- R environment Saving ,Storing and Retrieving work
- R datatypes, variables and constants
- Matrix operation in R
- Methods of Data Input
- Importing Data from Excel
- Graphics with R
- Primary , Secondary Data
- Frequency , Cumulative Frequency , Classification of Data, Classification Rules
- Measures of Central Tendency
- Measures of Dispersion
- Skewness , Kurtosis
- Correlation and Regression Techniques

Weekly Plan

Week 1 : R Language

History Of R Language, Why R

Installation, Environment and R Prompt

R as Calculator, Arithmetic Operators , BODMAS rule in Evaluation of expression

Data Types in R

Assignment Operator , adding comments

Variable naming conventions

Evaluation of Arithmetic Expressions

Vectors in R and its arithmetic , Data Frames

Matrix Command , Accessing and Matrix operations , Functions in R

Methods of Data Inputs , Importing Data from Excel , Useful Built in Functions

Week 2 : Diagrammatic Representation of Data using R

Importance of Diagrammatic representation of Data

Drawing of Bar Diagrams

Drawing of Subdivided and Multiple Bar Diagrams

Drawing of Pie Charts

Histograms

Frequency Polygons

Stem and Leaf Plot

Q-Q plots

Week 3 : Data Summarization Measures using R

What is Central Tendency

Types of Central Tendency Measures

Arithmetic Mean and Its Properties and computation

Median and Mode computation

Partition Values

Measures of Dispersion

Range Quartile Deviation Mean Deviation

Standard deviation

Skewness and Kurtosis, Correlation & Regression computation using R