- To acquaint with basic of data analysis and handling of missing observations.

Learning Outcomes:

- Apply the fundamentals of R and its features for data analysis and interpretation.
- Proficiently apply basic R syntax and effectively work with different data types
- Explore data visualization in R and its importance in data analysis
- Interpret the result of bivariate data analysis techniques including cross tabulation and scatter plot.
- Interpreting the result of a fitted linear regression model.

UNIT I

Introduction to R and its features, Installing R and R-Studio, Basic R syntax and data types, Importing data into R from various file formats (e.g., CSV, Excel, text files), Problems faced in case of missing data and data cleaning techniques.

UNIT II

Introduction to data visualization in R, Basic plotting functions in R, Creating and customizing various types of plots, including: Histograms, Pie charts, Frequency polygons, Frequency curves, Bar charts, Box plots. Adding labels, titles, and legends to plots, saving and exporting plots.

UNIT III

Measures of central tendency: Mean, median, and mode, weighted mean, Geometric mean, Measures of dispersion: Range, variance and standard deviation, Coefficient of variation, Interquartile range, Percentiles and quartiles.

UNIT IV

Measures of skewness: Coefficient of skewness and its interpretation, Measures of kurtosis: coefficient of kurtosis. Introduction to bivariate analysis: Cross tabulation, scatter plots. Pearson's and Spearman's correlation coefficient.

UNIT V

Simple linear regression, fitting of regression models in R, Evaluating regression models, Least Squares method, interpreting regression result and prediction using fitted regression model.

SKILL ENHANCEMENT COURSE IN STATISTICS: SEC-151L
(Statistical Data Analysis using R)

Full Marks=30 [End Semester Exam (30)]
Pass Marks =12 [End Semester Exam (12)]

1. Plotting simple graphs in R (Histograms, Bar Diagram, Pie Diagram, Boxplot, Stem-leaf, ogives).
2. Adding labels, titles, and legends to plots, Saving and exporting plots.
3. Tabulation of raw data in R
4. To compute mean, median and mode for a grouped frequency data in R
5. To compute Geometric mean and Harmonic mean.
6. To compute mean, median, variance, covariance, standard deviation in R.
7. Computation of partition values, skewness and kurtosis in R.
8. To compute correlation and lines of regression in R.
9. Random number generation from different distributions in R.
10. Fitting of simple linear regression in R and its interpretation.
11. Fitting of polynomials and exponential curves in R.
12. Fitting of Binomial and Poisson distribution in R.
13. Problems based on selecting random sample in R (with and without replacement).
14. Problems based on plotting normal probability plot in R (P-P plot and Q-Q plot).

SUGGESTEDREADINGS:

1. Gardener,M. (2012). Beginning R:The Statistical Programming Language. Wiley Publications.
2. Braun, W.J.,& Murdoch, D.J. (2007). A First Course in Statistical Programming with R. Cambridge University Press. New York
3. Moore, D.S., McCabe, G.P., &Craig, B.A. (2014). Introduction to the Practice of Statistics. W.H. Freeman
4. Cho,M.J.,Martinez,W.L.(2014). Statistics in MATLAB: A Primer. Chapman and Hall/CRC

*****