CAP788:DATA SCIENCE TOOLBOX-LABORATORY

L:0 T:0 P:2 Credits:1

Course Outcomes: Through this course students should be able to

CO1 :: understand data and data science in real-time analysis

CO2 :: perform analysis-focused practical's using R and Python

CO3 :: practice data visualization using R and Python

CO4 :: develop model and perform model evaluation using real time data sets

List of Practicals / Experiments:

Practicing basic programs in R

- Write programs using basic loops in R.
- Write Programs using vectors
- Write Programs using Matrices
- Write Programs using Factors
- Write Programs using Data Frames

Practicing Basic of Pyhton

- Write programs using Lists,
- Write programs using Tuples,
- Write programs using Dictionaries
- Write programs using Sets
- Write programs using Scientific Libraries NumPy, SciPy
- Programs using Matplotlib, Pandas

Practicing Data Analysis using R

- Apply Data Preprocessing on imported Data.
- Apply Basic statistical functions in R on Imported Data.
- Apply Visual Presentation and draw various plots.
- Interpretation of results.
- Data Analysis and its interpretation on real-time data set

Practicing Data Analysis using Python

- Practice Data Analysis and apply Correlation Analysis,
- Implementing Regression, and ANOVA on imported data.
- Develop and evaluate the Model.
- Practice Data Visualisation using plots such as Histograms, BoxPlot, GG Plot

Installation of R studio and Python

- Installing R Studio, Loading Data in R,
- Exporting and Importing data using various online platforms,
- R Package, Projects in R and R Markdown
- Installation of Python

Working with GitHub

- Version Control
- GitHub and Git
- Linking Github with R studio
- Linking Github with Pyhton
- Projects Under Github

Text Books:

1. PYTHON FOR DATA ANALYSIS by WES MCKENNY, O'REILLY

References:

1. GITHUB ESSENTIALS by ACHILLEAS PIPINELLIS, PACKT PUBLISHING

2. R FOR DATA SCIENCE: IMPORT, TIDY, TRANSFORM, VISUALIZE, AND MODEL DATA by HADLEY WICKHAM, O'REILLY