Report on Faculty Development Programme on Statistical Methods and Techniques Using R

Program Title: Faculty Development Programme on Statistical Methods and Techniques Using R

Duration: One Week

Dates: 2nd -6th January 2024

Venue: CSE 150 LAB

Organized by: DEPARTMENT OF CSE

Program Coordinator: Mr.Dastagiri, Asst.Professor
Mr.N.Sivakumar, Asst.Professor
Ms.R.Sujatha, Asst.Professor

Objective: To provide faculty members with a comprehensive understanding of statistical methods and techniques using R, enhancing their ability to apply these methods in research, teaching, and data analysis.

Resource Persons:

Mr.Pawn, Web Developer, Path creators

Mr. Dattatreya Goud, Decode Soft Tech Solutions Pvt Ltd

Dr. Mohana Sundaram, VIT University, Vellore

Day 1: Introduction to R and Statistical Fundamentals

Morning Session:

The FDP started with formal inauguration where Dr.S.Shathi, HOD of CSE, welcomes the resource person, Dr.S.Srinivas Rao , Principal of MRCET , Dr.T.Venugopal, Dean, and students. Dr.S.Shathi, HOD of CSE, insisted faculty to utilize the opportunity in developing their skills and knowledge on R Programming as it will help to teach Data Science effectively.

- **Introduction to R:**
  - Overview of R programming language and its importance in statistical analysis.
  - Installation and setup of R and RStudio.
  - Basic R syntax and data structures (vectors, lists, data frames).

Afternoon Session:

- **Statistical Fundamentals:**
  - Introduction to basic statistical concepts (e.g., descriptive statistics, probability distributions).
  - Understanding and visualizing data with R (e.g., histograms, scatter plots).
- **Hands-On Activity:**
  - Simple data manipulation and visualization tasks using R.

Day 2: Data Exploration and Visualization

Morning Session:

- **Data Exploration Techniques:**
  - Techniques for exploring and summarizing data (e.g., summary statistics, correlation analysis).
  - Introduction to data cleaning and preprocessing.
- **Visualization with ggplot2:**
  - Overview of ggplot2 for data visualization in R.
  - Creating advanced plots (e.g., box plots, bar charts, line graphs).

Afternoon Session:

- **Hands-On Labs:**
  - Data cleaning and preprocessing exercises.
  - Creating and customizing visualizations using ggplot2.
- **Case Study:**
  - Analyzing a real dataset and presenting findings through various visualizations.

Day 3: Inferential Statistics

Morning Session:

- **Hypothesis Testing:**
  - Introduction to hypothesis testing concepts (e.g., null and alternative hypotheses, p-values).
  - Conducting t-tests, chi-square tests, and ANOVA in R.
- **Regression Analysis:**
  - Basics of simple linear regression and multiple regression.
  - Interpretation of regression outputs.

Afternoon Session:

- **Hands-On Labs:**
  - Performing hypothesis tests and regression analysis on sample datasets.
  - Interpreting results and drawing conclusions.
- **Case Study:**
  - Applying hypothesis testing and regression analysis to a real-world problem.

Day 4: Advanced Statistical Techniques

Morning Session:

- **Multivariate Analysis:**
  - Introduction to principal component analysis (PCA) and factor analysis.
  - Techniques for exploring multivariate data.
- **Time Series Analysis:**
  - Basics of time series data and forecasting methods.
  - Introduction to ARIMA models and their implementation in R.

Afternoon Session:

- **Hands-On Labs:**
  - Conducting PCA and factor analysis on sample datasets.
  - Performing time series analysis and forecasting using R.
- **Case Study:**
  - Applying advanced statistical techniques to complex datasets.

Day 5: Statistical Modeling and Machine Learning

Morning Session:

- **Statistical Modeling:**
  - Overview of various statistical models (e.g., logistic regression, generalized linear models).
  - Model selection and evaluation.
- **Introduction to Machine Learning:**
  - Basics of machine learning concepts and algorithms (e.g., decision trees, k-nearest neighbors).
  - Implementation of machine learning models in R.

Afternoon Session:

- **Hands-On Labs:**
  - Building and evaluating statistical and machine learning models using R.
  - Practical exercises on model tuning and performance assessment.
- **Case Study:**

- Developing and applying statistical and machine learning models to a real dataset.

- **Feedback and Q&A:**
  - Collection of participant feedback on the program.
  - Addressing any remaining questions or concerns.
- **Certificate Distribution:**
  - Presentation of certificates of participation.
  - Closing remarks and program wrap-up.

Participant Feedback:

- Positive feedback on the practical, hands-on approach and the use of real datasets for exercises.
- Appreciation for the comprehensive coverage of statistical methods and R techniques.
- Suggestions for additional topics such as advanced machine learning algorithms and interactive visualizations.

Outcome of the Programme:

- Participants acquired foundational knowledge of R programming and basic statistical concepts, along with practical experience in data manipulation and visualization
- Participants gained skills in data exploration, cleaning, and advanced visualization techniques using ggplot2.
- Participants learned to perform and interpret various inferential statistical tests and regression analyses using R.
- Participants gained knowledge and practical experience with advanced statistical methods, including multivariate analysis and time series forecasting.
- Participants acquired practical skills in statistical modeling and machine learning techniques using R.

Conclusion:

The program successfully met its objectives, providing participants with a robust understanding of statistical methods and techniques using R. The skills and knowledge gained will enhance participants' ability to conduct data analysis and incorporate statistical techniques into their teaching and research. Finally Dr.S.Shanthi, HOD-CSE has thanked Director and Principal of MRCET for providing this wonderful opportunity.

Photos Gallery