EXPLORATORY DATA ANALYSIS

(Program Elective-V)

Course Code	19IT4702F	Year	IV	Semester	I
Course Category	PE	Branch	IT	CourseType	Theory
Credits	3	L-T-P	3-0-0	Prerequisites	Data Mining
Continuous Internal Evaluation:	30	Semester End Evaluation:	70	TotalMarks:	100

Course Outcomes	Blooms Taxonomy Level
Upon successful completion of the course, the student will be able to	
CO1 Understand the basic concepts of EDA	L2
CO2 Determine various graphical techniques of EDA	L3
CO3 Determine various quantitative techniques of EDA	L3

Contribution of Course Outcomes towards achievement of Program Outcomes & Strength of correlations (H:High, M:Medium, L:Low)													
	PO1	**PO2**	**PO3**	**PO4**	**PO5**	**PO6**	**PO7**	**PO8**	**PO9**	**PO10**	**PO11**	**PO12**	**PSO1**	**PSO2**
CO1	3												1	
CO2	2	3		2									1	
CO3	2	3		2									1	

Syllabus		
Unit No	**Contents**	**Mapped CO**
I	**EDA Introduction**: What is EDA? How Does Exploratory Data Analysis differ from Classical Data Analysis? How Does Exploratory Data Analysis Differ from Summary Analysis? What are the EDA Goals? The Role of Graphics, An EDA/Graphics Example and General Problem Categories	**CO1**
II	**EDA Techniques-I**: Introduction, Analysis Questions, Graphical Techniques: Alphabetic- Autocorrelation Plot, Bihistogram, Block Plot, Bootstrap Plot, Box-Cox Linearity Plot, Box-Cox Normality Plot, Box Plot, Complex Demodulation Amplitude Plot, Complex Demodulation Phase Plot, Contour Plot, DOE Scatter Plot, DOE Mean Plot and DOE Standard Deviation Plot.	**CO1 CO2**
III	**EDA Techniques–II**: Histogram, Lag Plot, Linear Correlation Plot, Linear Intercept Plot, Linear Slope Plot, Linear Residual Standard Deviation Plot, Mean Plot, Normal Probability Plot, Probability Plot, Probability Plot Correlation Coefficient Plot, Quantile- Quantile Plot, Run-Sequence Plot	**CO1 CO2**
IV	**EDA Techniques – III**: Scatter Plot, Spectral Plot, Standard Deviation Plot, Star Plot, Weibull Plot, Youden Plot, 4-Plot, 6-Plot. Graphical Techniques: By Problem Category	**CO1 CO2**
V	**Quantitative Techniques**: Measures of Location, Confidence Limits for the Mean, Two-Sample t-Test for Equal Means, One-Factor ANOVA, Multi-factor Analysis of Variance, Measures of Scale, Bartlett's Test, Chi-Square Test for the Standard Deviation, F-Test for Equality of Two Standard Deviations, Levene Test for Equality of Variances, Measures of Skewness and Kurtosis, Auto correlation, Runs Test for Detecting Non-randomness, Anderson-Darling Test, Chi-Square Goodness-of-Fit Test, Kolmo gorov-Smirnov Goodness-of-Fit Test, Grubbs' Test for Outliers and Yates Analysis	**CO1 CO3**

Learning Resources
Text Books
1.Exploratory Data Analysis Handbook
References
1.Exploratory Data Analysis with Python by Suresh kumar Mukhiya, Usman Ahmed