SYLLABUS

DATA VISUALISATION USING PYTHON

UNIT-I: INTRODUCTION TO DATA SCIENCE AND PYTHON TOOLS

Introduction: Introduction to Data Science, Exploratory Data Analysis and Data Science Process. Motivation for using Python for Data Analysis, Introduction of Python Jupyter Notebook. Essential Python Libraries: NumPy, pandas, matplotlib, SciPy, scikit-learn, statsmodels, seaborn.

UNIT -II: DATA MANIPULATION AND ANALYSIS WITH PANDAS

Getting Started with Pandas: Arrays and vectorized computation, Introduction to pandas Data Structures, Essential Functionality, Summarizing and Computing Descriptive Statistics. Data Loading, Storage and File Formats. Reading and Writing Data in Text Format, Web Scraping, Binary Data Formats, Interacting with Web APIs,

Interacting with Databases Data Cleaning and Preparation. Handling Missing Data, Data Transformation, String Manipulation.

UNIT-III: ADVANCED DATA PREPARATION AND VISUALIZATION TECHNIQUES

Data Wrangling: Hierarchical Indexing, Combining and Merging Data Sets Reshaping and Pivoting. Data Visualization matplotlib: Basics of matplotlib, plotting with pandas and seaborn, other python visualization tools. Advanced categorical and numeric plots.

UNIT-IV: DATA AGGREGATION, GROUP OPERATIONS, AND TIME SERIES ANALYSIS

Data Aggregation and Group operations: Group by Mechanics, Data aggregation, General split-apply-combine, Pivot tables and cross tabulation

Time Series Data Analysis: Date and Time Data Types and Tools, Time series

Basics, date Ranges, Frequencies and Shifting, Time Zone Handling, Periods and Periods Arithmetic, Resampling and Frequency conversion, Moving Window Functions.

UNIT -V: ADVANCED PANDAS OPERATIONS AND WORKFLOWS

Categorical Data: cleaning data and visualization techniques, Advanced GroupBy methods ,Use Techniques for Method Chaining.