COURSE TITLE	**DATA DRIVEN ANALYTICS**
COURSE CODE	**05MF0205**
COURSE CREDITS	**4**

Objective:

1. Students will be able to master data handling techniques encompassing loading, cleaning, and preprocessing across diverse formats including Excel, CSV, JSON, and SQLite, using Pandas.
2. Students will be able to develop proficient visualization skills through Matplotlib and Seaborn, enabling the creation of customized plots for effective data exploration and communication.
3. Students will be able to acquire web scraping proficiency with Beautiful Soup, Selenium, and Scrapy, enabling the extraction of structured data from websites for analysis.
4. Students will be able to apply acquired skills in a mini-project focusing on technical and fundamental analysis of financial and stock data, integrating web scraping and data visualization.
5. Students will be able to engage in guided financial projects tailored for professionals like Chartered Financial Analysts (CFA) and Chartered Accountants (CA), applying data analytics techniques to real-world financial scenarios.

Course Outcomes: After completion of this course, student will be able to:

1. Students will demonstrate proficiency in loading, cleaning, and pre- processing data from various sources, including Excel, CSV, JSON, and SQLite, using Pandas, ensuring data integrity and usability for analysis.
2. Students will be able to effectively visualize data using Matplotlib and Seaborn, creating insightful and visually appealing plots to communicate patterns, trends, and insights derived from the data.
3. Students will gain competency in web scraping techniques using BeautifulSoup, Selenium, and Scrapy, allowing them to extract structured data from websites efficiently and ethically for analysis purposes.
4. Students will apply their acquired skills in a mini-project focused on technical and fundamental analysis of financial and stock data, integrating web scraping and data visualization techniques to derive actionable insights.
5. Students will demonstrate proficiency in loading, cleaning, and preprocessing data from various sources, including Excel, CSV, JSON, and SQLite, using Pandas, ensuring data integrity and usability for analysis. They can effectively visualize data using Matplotlib and Seaborn, creating insightful plots, and gain competency in web scraping with Beautiful Soup, Selenium, and Scrapy, applying these skills in a mini-project focused on the technical and fundamental analysis of financial and stock data

Pre-requisite of course: Basic Knowledge about Python and Finance

DR. SUNIL LALCHAND BAJEJA
Digitally signed by (Name of HOD)

DR. R SRIDARAN
Digitally signed by (Name of Dean/ Principal)

Teaching and Examination Scheme

Theory Hours	Tutorial Hours	Practical Hours	ESE	IA	CSE	Viva	Term Work
3	0	2	50	30	20	25	25

Contents : Unit	Topics	Contact Hours
1	**Data Loading** - Loading Data with Pandas (Excel - CSV), Creating Json File and reading Json file, Reading Data from SQLite , Cleaning and Preparing Data , , Irrelevant Columns Name, Outliers , Duplicates Value handling , Missing Data Handling , Columns that have to be processed, Unexpected Data Values.	11
2	**Processing with Data Visualization** - Basic plotting techniques: line plots, scatter plots, bar plots, and histograms, Customizing plots with titles, labels, Colors, and styles, , Multi-panel plots: subplots, Plotting categorical data: bar charts, pie charts, and box plots, Customizing and Styling Plots with Seaborn.	11
3	**Web Scrapping** - Web Scrapping Introduction and libraries introduction – Beautiful Soup - Selenium – Scrapy HTML for Web Scrapping , Tags & Element - Tree Structure , Working with Beautiful Soup , Working with Selenium , Working with Scrapy.	11
4	**Mini Project on Finance** - Technical and Fundamental Analysis on the choice of your topic	12
Total Hours		**45**

DR. SUNIL LALCHAND BAJEJA
Digitally signed by (Name of HOD)

DR. R SRIDARAN
Digitally signed by (Name of Dean/ Principal)

Suggested List of Experiments:

Contents : Unit	Topics	Contact Hours
1	**Unit 1** • How would you load a CSV file named "data.csv" into a Pandas DataFrame using one example? • Explain the steps involved in handling missing data in a dataset using Pandas using one example. • What are the common techniques to identify and handle outliers in a dataset u sing one example? • Describe how you would rename specific columns in a DataFrame to make them more descriptive using one example. • Discuss the process of removing duplicate rows from a DataFrame using Pandas using one example. • How can you create a scatter plot using Matplotlib to visualize the relationship between two variables using one example? • Explain the purpose of using histograms in data visualization and provide an example where histograms are useful using one example. • How would you customize the appearance of a plot in Matplotlib, including adding titles, labels, and legends using one example? • Compare and contrast line plots and bar plots, and provide scenarios where each type of plot is appropriate using one example. • Describe the steps to create a box plot using Seaborn and interpret the insights it provides about the data distribution using one example.	15
2	**Unit 2** • What is web scraping, and how does it differ from web crawling, explain with suitable examples? • Explain the role of HTML tags and attributes in web scraping using Beautiful Soup using examples. • Discuss the advantages and limitations of using Selenium for web scraping compared to Beautiful Soup using examples. • How can you handle dynamic content and JavaScript-rendered pages when scraping using Selenium using examples? • Describe the structure of a Scrapy spider and explain how it facilitates web scraping of multiple pages on a website using examples. • Outline the steps involved in collecting financial data from a company's annual report for analysis. • How would you calculate financial ratios such as debt-to-equity ratio and return on investment (ROI) using Python? • Explain the concept of technical analysis in stock market forecasting and provide examples of technical indicators. • Discuss the significance of fundamental analysis in evaluating the financial health and performance of a company. • Describe how you would use Python libraries to visualize trends in stock prices and trading volumes for technical analysis.	15
Total Hours		**30**

Textbook :

1 Python for Data Analysis: Data Wrangling with Pandas, NumPy, and IPython,, Wes McKinney, O'Reilly Media, 2017

References:

1 Python for Data Analysis, Python for Data Analysis, Wes McKinney, O'Reilly Media, 2017

DR. SUNIL LALCHAND BAJEJA
Digitally signed by (Name of HOD)

DR. R SRIDARAN
Digitally signed by (Name of Dean/ Principal)

References:

2 Data Visualization with Python and Matplotlib, Data Visualization with Python and Matplotlib, Benjamin Walter Keller, Independently published, 2018

3 Web Scraping with Python: Collecting More Data from the Modern Web, Web Scraping with Python: Collecting More Data from the Modern Web, Ryan Mitchell, O'Reilly Media, 2018

Suggested Theory Distribution:

The suggested theory distribution as per Bloom's taxonomy is as follows. This distribution serves as guidelines for teachers and students to achieve effective teaching-learning process

Distribution of Theory for course delivery					
Remember / Knowledge	**Understand**	**Apply**	**Analyze**	**Evaluate**	**Higher order Thinking / Creative**
10.00	20.00	25.00	25.00	10.00	10.00

Instructional Method:

1 Board Work

2 PPT

3 Demo

Supplementary Resources:

1 https://prolearn.mit.edu/data-science-and-big-data-analytics-making-data-driven-decisions

2 https://towardsdatascience.com/

DR. SUNIL LALCHAND BAJEJA

Digitally signed by (Name of HOD)

DR. R SRIDARAN

Digitally signed by (Name of Dean/ Principal)