Kadi Sarva Vishwavidyalaya, Gandhinagar

25CASE212: Data Analytical Tools

Program: B.E. Computer Engineering / Computer Science and Engineering / Information Technology (Year - 2)

Teaching Scheme: 0L - 0T - 4P (2 Credits)

Prerequisite: Students should possess basic programming knowledge (preferably in Python), along with an understanding of statistics and probability fundamentals. They should also be familiar with database concepts and have hands-on experience using tools such as Excel for basic data handling and visualization.

Rationale: The Data Analytics Tools course is designed to equip students with essential analytical, statistical, and computational skills required in the data-driven world. With the growing demand for insights from massive datasets, this course bridges theory and practical applications using Python, Power BI, and R.

1. Course Objectives:

The Purpose of learning this course is

- **To explain** the fundamental concepts, processes, and importance of data analytics in real-world applications.
- **To apply** data analytical tools such as Python, R, Power BI, and Tableau for data processing and visualization tasks.
- **To perform** data cleaning, transformation, and exploratory data analysis on structured and unstructured datasets.
- **To analyze** datasets using statistical and analytical techniques to extract meaningful patterns and insights.
- **To design and develop** interactive dashboards and visual reports for effective data-driven decision making. Data-driven sustainability insights

2. Teaching and Evaluation Scheme:

Teaching Scheme					Examination Scheme				
L	**T**	**P**	**C**	**Hrs / Week**	**IE**	**EE**	**CIA**	**Practical**	**Total Marks**
-	-	4	2	4	00	00	30	20	50

- **IE**: Internal Exam
- **EE**: External Exam (End Semester)
- **CIA**: Continuous Internal Assessment
- **Practical**: Practical Exam (End Semester)

3. LIST OF PRACTICALS:

Lab	Details	Hours	CO
1	**Python – Data Loading and Preprocessing** To load and preprocess a dataset using Python libraries.	02	CO1
2	**Python – Exploratory Data Analysis (EDA)** To perform exploratory data analysis using Python.	02	CO1
3	**Python – Visualization using Matplotlib/Seaborn** To create data visualizations for analysis.	02	CO1, CO2
4	**Power BI – Importing Datasets** To import data into Power BI from multiple sources.	02	CO2
5	**Power BI – Data Transformation using Power Query** To clean and transform data using Power Query Editor.	04	CO2
6	**Power BI – Creating Dashboards** To build an interactive dashboard using Power BI.	04	CO2, CO5
7	**Power BI – DAX Formulas for Calculated Fields** To create calculated fields using DAX.	04	CO2
8	**R – Data Import and Manipulation** To import and manipulate data using R.	04	CO3
9	**R – Data Visualization with ggplot2** To visualize data using ggplot2 library.	04	CO3
10	**R – Correlation and Regression Analysis** To perform correlation and regression analysis in R.	04	CO3
11	**R – Hypothesis Testing** To test statistical hypotheses using R.	04	CO3
12	**R – Classification or Clustering Model** To implement classification or clustering algorithms.	04	CO3
13	**Integration of Python Results with Power BI** To integrate Python analytics results into Power BI.	04	CO4
14	**Forecasting using R or Python** To perform predictive analysis using forecasting models.	04	CO4
15	**Mini Project: Data Analytics Case Study** To apply data analytics concepts to a real-world case.	04	CO4, CO5
16	**Capstone Dashboard Development** To develop a comprehensive analytical dashboard integrating all tools.	04	CO5
17	**Presentation and Evaluation of Analytical Findings** To present and communicate analytical insights effectively.	04	CO5

4. Course Outcomes (COs)

After completion of the course, the student will be able to:

CO No.	Course Outcome (CO)	RBT Level (Cognitive Domain)
CO1	**To explain** the fundamental concepts, processes, and importance of data analytics in real-world applications.	Understand
CO2	**To apply** data analytical tools such as Python, R, Power BI, and Tableau for data processing and visualization tasks.	Apply
CO3	**To perform** data cleaning, transformation, and exploratory data analysis on structured and unstructured datasets.	Apply, Create
CO4	**To analyse** datasets using statistical and analytical techniques to extract meaningful patterns and insights.	Apply , Analyse
CO5	**To design and develop** interactive dashboards and visual reports for effective data-driven decision-making. Data-driven sustainability insights	Apply, Create

5. Mapping of COs with POs & PSOs:

CO	PO1	PO2	PO3	PO4	PO5	PO6	PO7	PO8	PO9	PO10	PO11	PSO1	PSO2
CO1	3	2	–	–	2	–	–	–	1	–	1	3	2
CO2	3	3	2	1	3	–	–	–	1	–	1	3	2
CO3	2	2	3	2	3	–	1	–	1	–	1	3	3
CO4	2	3	3	2	3	–	1	2	1	–	1	3	3
CO5	2	2	3	3	3	–	1	2	2	1	2	3	3
Avg.	**2.4**	**2.4**	**2.2**	**1.6**	**2.8**	**0**	**0.6**	**0.8**	**1.2**	**0.2**	**1.2**	**3**	**2.6**

6. Reference Reading:

6.1 Text Books:

1. Wes McKinney – Python for Data Analysis, O'Reilly.
2. Alberto Ferrari & Marco Russo – Introducing Microsoft Power BI.
3. Hadley Wickham – R for Data Science, O'Reilly.
4. Provost & Fawcett – Data Science for Business.

6.2 Reference Books:

1. Wes McKinney – Python for Data Analysis, O'Reilly.
2. Alberto Ferrari & Marco Russo – Introducing Microsoft Power BI.

6.3 LIST OF OPEN-SOURCE SOFTWARE/LEARNING WEBSITE:

1. **Anaconda Distribution –** Open-source platform for Python, R, and data science tools. (https://www.anaconda.com/)
2. **Jupyter Notebook / JupyterLab –** Interactive environment for writing and executing Python code. (https://jupyter.org/)
3. **Microsoft Power BI Desktop (Free Version)** – Used for creating dashboards and reports. (https://www.microsoft.com/en-us/power-platform/products/power-bi/)

7. Course Outcomes Mapping Table

CO No.	Course Outcome (CO)	POs/PSOs Map
CO1	**To explain** the fundamental concepts, processes, and importance of data analytics in real-world applications.	PO1, PO2, PO5, PO9, PO11, PSO1, PSO2
CO2	**To apply** data analytical tools such as Python, R, Power BI, and Tableau for data processing and visualization tasks.	PO1, PO2, PO3, PO4, PO5, PO9, PO11, PSO1, PSO2
CO3	**To perform** data cleaning, transformation, and exploratory data analysis on structured and unstructured datasets.	PO1, PO2, PO3, PO4, PO5, PO7, PO9, PO11, PSO1, PSO2
CO4	**To analyse** datasets using statistical and analytical techniques to extract meaningful patterns and insights.	PO1, PO2, PO3, PO4, PO5, PO7, PO8, PO9, PO11, PSO1, PSO2
CO5	**To design and develop** interactive dashboards and visual reports for effective data-driven decision-making.	PO1, PO2, PO3, PO4, PO5, PO7, PO8, PO9, PO10, PO11, PSO1, PSO2

8. Evaluation Scheme:

- The course evaluation shall Continuous Internal Assessment (CIA) and External Practical Examination.
- The Continuous Internal Assessment (CIA) (30 marks) shall be considered based on assignments given during the term and practical implementation carried out by the student.
- The Practical Examination (70 marks) shall be conducted externally and will include implementation, performance, and viva-voce.
- Students are required to secure minimum qualifying marks in each component separately as per university rules.

9. Pedagogical Framework:

- The course adopts a blended teaching approach coding demonstrations, and interactive problem-solving sessions.
- Emphasis is placed on learning by doing through extensive laboratory work, assignments, and mini-projects.
- Conceptual understanding is reinforced with real-world examples and algorithm visualization techniques.
- Students are encouraged to engage in collaborative learning, discussions, and peer problem-solving. Data-driven sustainability insights
- Continuous feedback through assessments and lab performance supports progressive skill development and mastery.

10. Innovative Practices:

- **Problem-Based Learning (PBL):** Each practical is designed as a real-world problem statement (e.g., sales analysis, customer segmentation), encouraging students to think critically and apply analytical tools to derive meaningful insights.
- **Multi-Tool Integration Approach:** Students are exposed to multiple industry-relevant tools such as Python, R, Power BI, and Tableau, and are encouraged to integrate outputs across platforms (e.g., Python results in Power BI dashboards).
- **Use of Real-World Datasets:** Instead of predefined datasets, students work with datasets from platforms like Kaggle, government portals, and industry case studies to simulate real-life analytics scenarios.
- **Mini Project & Capstone-Based Learning:** Students undertake a domain-specific mini project and a capstone dashboard, promoting independent learning, creativity, and end-to-end implementation of analytics workflows.
- **Data Storytelling & Presentation Practice:** Emphasis is given not only on analysis but also on communicating insights through storytelling, dashboards, and presentations, enhancing both technical and soft skills.

11. Self-Learning Component:

- Students are encouraged to engage in self-paced learning through online coding platforms (NPTEL/SWAYAM), tutorials, and technical resources to strengthen their understanding of algorithms and data structures. (E.g. Programming, *Data Structures And Algorithms Using Python* by Prof. Madhavan Mukund (IIT Madras) - https://nptel.ac.in/courses/106106145)
- Regular practice of coding problems beyond the syllabus helps in improving problem-solving skills and logical thinking.
- Learners can explore advanced topics such as dynamic programming, graph algorithms, and optimization techniques independently.
- Participation in coding competitions, hackathons, and open-source contributions promotes continuous self-improvement.
- This approach fosters lifelong learning habits, enabling students to keep pace with evolving technologies and industry demands.

Sustainable Development Goals (SDG) Integration

The course is aligned with **SDG 9 (Industry, Innovation and Infrastructure)** by equipping students with modern data analytics and visualization skills required for building intelligent, data-driven systems. It also supports **SDG 13 (Climate Action)** by enabling analysis and interpretation of environmental and climate-related data for informed decision-making. Overall, the course promotes the use of data for sustainable development and responsible innovation.

8. SDG Mapping with Course Outcomes

CO	Course Outcome	Relevant SDGs	Justification
CO1	Understand fundamentals of data analytics and preprocessing using Python	SDG 9	Builds foundational skills required for modern data-driven infrastructure and digital innovation.
CO2	Perform data visualization and reporting using Power BI	SDG 9	Enables development of industry-relevant dashboards and reporting systems for decision-making.
CO3	Apply statistical and analytical techniques using R	SDG 9, SDG 13	Supports analytical modeling useful in industrial optimization and environmental data analysis such as climate trends.
CO4	Integrate tools for advanced analytics and interpret data insights	SDG 9, SDG 13	Encourages development of integrated systems for smart infrastructure and data-driven climate monitoring solutions.
CO5	Develop dashboards and present analytical findings	SDG 9, SDG 13	Helps in communicating insights for sustainable industrial growth and environmental awareness through data storytelling.