

Code: BCA-3005T (For theory) and BCA-3005P (For practical)	DSEC-II	Group-B: Elective-I Basics of Data Analytics using Spreadsheet	1L+T:4P	3 Credits (15 hours theory and 60 hours practical)
Max Marks; Theory: 100 (Int: 25; Ext: 75); Practical: 100				
<p>Course Outcomes: Upon completion of the course, students will be able to</p> <p>CO1: Understand the basics of data analytics and its applications.</p> <p>CO2: Develop proficiency in using spreadsheet software for data manipulation and analysis.</p> <p>CO3: Build and use spreadsheet models for decision making & Communicate data insights effectively</p> <p>Course Outcomes after Lab Programs:</p> <p>CO1: Implement and evaluate supervised learning techniques, including K-Nearest Neighbors, linear regression, and logistic regression, and measure model performance using accuracy, precision, recall, and F1 score.</p> <p>CO2: Apply and visualize clustering algorithms such as K-Means, hierarchical clustering, and DBSCAN on datasets. This practical application helps you understand their real-world use.</p> <p>CO3: Perform dimensionality reduction using Principal Component Analysis (PCA) and interpret the results.</p> <p>CO4: Develop and assess classification models using random forests, support vector machines, and neural networks.</p> <p>CO5: Demonstrate ensemble learning concepts through bagging with random forests and boosting with the AdaBoost algorithm.</p>				
Unit	Topic s		Proposed Lecture	
I	Introduction to Data Analytics Understanding Data and its Types (Structured, Unstructured, Semi-Structured)-What is Data Analytics, Types of Data Analytics, Importance of Data Analytics, Applications of Data Analytics, Introduction to Spreadsheet Tools (Excel/Google Sheets)		7	
II	Data, Ethics, and Industry : Case Studies Data Collection Methods – Different Data Sources & Format – Data Cleaning and Transformation – Handling Missing Data and Outliers, Removing Duplicates – Ethical Considerations in Data Analytics.- Real-world Applications of Data Analytics- Industry-specific Applications (Finance, Marketing, Operations, Healthcare, Manufacturing/Supply Chain) – Case Study Note: Case study is for discussion not to be considered for evaluation.		8	
Lab Programs:	PART-A: Understanding and Describing the Data Introduction to Excel and its Basic Functions Getting started with Excel: Workbook, Worksheet, Cells, and Ranges Data entry and basic formatting techniques Using basic arithmetic functions: SUM, AVERAGE, COUNT, MIN, MAX, ROUND, CEILING, FLOOR Introduction to cell referencing: relative, absolute, and mixed Data Importing and Pre-processing Importing data from various sources(CSV, text files, web data) Data cleaning: removing duplicates, handling missing data, and standardizing formats Data transformation: text-to-columns, data validation techniques Using the "Find & Replace" and "Text Functions"(LEFT,RIGHT, MID,CONCATENATE), Sorting and Filtering Data			

Descriptive Statistics Using Excel

9. Calculating measures of central tendency: mean, median, mode
10. Computing measures of dispersion: range, variance, standard deviation, Coefficient of Variation (CV)
11. Creating and interpreting frequency distributions and histograms
12. Using Excel's "Data Analysis Toolpak" for basic statistical analysis

PART-B: Beyond the Basics: Visualizing and Communicating Data

Advanced Spreadsheet Functions

1. Use logical functions: IF, AND, OR, IFERROR, ISNA
2. Lookup and reference functions: VLOOKUP, HLOOKUP, INDEX, MATCH
3. Data aggregation techniques: SUMIFS, COUNTIFS, AVERAGEIFS
4. Text functions for data manipulation: TRIM, CLEAN, TEXT, RIGHT, LEFT, MID

Data Visualization Techniques

5. Creating various chart types: bar, line, pie, scatter
6. Advanced charting techniques: combo charts, dual-axis charts
7. Data visualization best practices: choosing the right chart, formatting, and styling
8. Creating and customizing Pivot Tables and Pivot Charts

Dashboard Creation

9. Introduction to dashboards: concepts and components
10. Use Pivot Tables and Pivot Charts for dashboard elements
11. Apply conditional formatting for dynamic visual cues
12. Create interactive dashboards with slicers and timeline

Text Books:

1. Mitchell, Tom M. *Machine Learning*. 1st ed., McGraw-Hill, 1997.
2. Kalita, J. K., D. K. Bhattacharyya, and S. Roy. *Fundamentals of Data Science: Theory and Practice*. Elsevier, 2023.
3. Jose, Jeeva. *Beginner's Guide for Data Analysis using R Programming*. Khanna Publishing House, 2024.
4. Nelson, Stephen L., and E. C. Nelson. *Excel Data Analysis for Dummies*. 3rd ed., John Wiley & Sons, 2016.
5. Middleton, Michael R. *Data Analysis Using Microsoft Excel*. 3rd ed., Thomson Brooks/Cole, 2004.

Reference Book:

6. Flach, Peter A. *Machine Learning: The Art and Science of Algorithms that Make Sense of Data*. Cambridge University Press, 2012.
7. Duda, Richard O., Peter E. Hart, and David G. Stork. *Pattern Classification*. 2nd ed., John Wiley & Sons, 2007.
8. Haykin, Simon. *Neural Networks and Learning Machines*. 3rd ed., PHI Learning, 2009.
9. Chollet, François. *Deep Learning with Python*. Manning Publications, 2018.
10. Bishop, Christopher M. *Pattern Recognition and Machine Learning*. Springer, 2006.
11. Goodfellow, Ian, Yoshua Bengio, and Aaron Courville. *Deep Learning*. MIT Press, 2016.
12. Géron, Aurélien. *Hands-On Machine Learning with Scikit-Learn and TensorFlow: Concepts, Tools, and Techniques to Build Intelligent Systems*. 1st ed., O'Reilly Media, 2017.
13. Alexander, Michael, Richard Kusleika, and John Walkenbach. *Excel 2019 Bible*. John Wiley & Sons, 25 Sept. 2018.
14. Ragsdale, Cliff T. *Spreadsheet Modeling and Decision Analysis: A Practical Introduction to Business Analytics*. Cengage Learning Asia, 2015.

SL
M
J
H