



B.Com Honours

Semester II

Calicut University

Data Analytics for Business

Course Code: COM2MN109 • Module 4 Notes

1. Data Analytics Process and Visualization Tools

Extracting business value from data requires a structured process and effective visualization tools to present findings to stakeholders. This final module covers the data analytics lifecycle, data mining frameworks, data visualization principles, introductory machine learning concepts, and analytical tools.

The Data Analytics Lifecycle

A standard analytics project progresses through six structured phases:

1. **Discovery:** Defining the business problem, establishing project goals, and identifying data sources.
2. **Data Preparation:** Setting up sandboxes, importing, cleaning, and pre-processing datasets.
3. **Model Planning:** Determining the analytical techniques (e.g., linear regression, decision trees) and selecting key variables.
4. **Model Building:** Writing code (Python/R) to train models and testing their accuracy on validation datasets.
5. **Communicate Results:** Summarizing findings and creating dashboards for business stakeholders.
6. **Operationalize:** Deploying the model into production environments for real-time decision support.

Data Visualization & Data Dictionary

Data Visualization Principles

Presenting numbers visually using appropriate chart types. E.g., Bar charts for categorical comparisons, Line charts for time-series trends, and Scatter plots for analyzing correlation. Dashboarding tools like Tableau and PowerBI provide interactive, real-time visual summaries.

Data Dictionary

A centralized document containing metadata —descriptions, types, formats, and relationships of all database columns. Essential to ensure consistency and clarity for the engineering team.

Introduction to Machine Learning (ML) & SQL Clustering

****Machine Learning**** is a branch of artificial intelligence focused on building applications that learn from data and improve accuracy over time without being explicitly programmed:

- **Supervised Learning:** Training algorithms on labeled datasets where inputs and outputs are known (e.g., predicting house prices based on size).
- **Unsupervised Learning:** Training algorithms on unlabeled datasets to find hidden patterns (e.g., clustering customers into purchasing groups using k-means algorithms).
- **SQL Clustering & Text Analysis:** Grouping data points using SQL queries and analyzing customer text feedback to identify sentiments.

Analytical Tools & Techniques

Firms utilize a suite of analytical tools: Microsoft Excel for basic calculations and descriptive analysis; SQL for querying database tables; Python and R for complex statistical modeling and machine learning; and PowerBI/Tableau for enterprise dashboarding.

Acing your exams is just a click away!

Visit www.degreeelive.in to download the next module for free.

DegreeLive