DegreeLive

B.Com Honours | Semester I | Calicut University

Business Analytics for Decision Making

Course Code: COM1MN110 • Module 2 Notes

1. The Business Data Analytics Process

Executing an analytics project successfully requires a structured workflow to ensure that the data processed aligns with real-world business objectives. Without a rigorous lifecycle, data scientists risk producing statistically sound models that solve the wrong business problems. This module covers the key stages of the Business Data Analytics (BDA) process, the role of data science in management, and the core characteristics of Big Data (the 3Vs).

The 5 Stages of the BDA Process

1. **Identify Research Questions (Business Problem Formulation):** Defining the business objective clearly. E.g., "Why are subscribers canceling services?" rather than just "Analyze subscription data."
2. **Source Data:** Identifying, locating, and extracting relevant data from internal databases (ERP, CRM) and external sources (public reports, APIs).
3. **Clean and Prepare Data:** Handling missing values, removing duplicate entries, treating outliers, and converting data formats (the most time-consuming stage).
4. **Analyse Data:** Applying descriptive, predictive, or prescriptive statistical models (e.g., running regression or clustering algorithms) and visualizing relationships.
5. **Interpret and Report Results:** Translating complex statistical findings into clear, non-technical business recommendations using charts and dashboards.

The Art of Data Science

Data science is the interdisciplinary field that combines scientific methods, algorithms, systems, and domain knowledge to extract knowledge and insights from structured and unstructured data.

In management decision-making, data science serves as a bridge, transforming raw data assets into corporate intelligence. It integrates three core competencies: **Mathematical/Statistical Expertise** (for model design), **Technology/Programming Skills** (for database queries and scripting), and **Business Acumen** (for defining goals and interpreting outcomes).

The 3Vs of Big Data

Modern analytical processes must handle "Big Data," characterized by three defining dimensions:

Volume

The scale of data generated, ranging from gigabytes to petabytes. Requires distributed storage architectures (like Hadoop, cloud data lakes) rather than standard spreadsheets.

Velocity

The speed at which new data is generated and must be processed. E.g., real-time credit card fraud detection, stock market feeds, or social media streams.

Variety

The structural diversity of data. Includes structured data (relational tables), semi-structured data (XML, JSON), and unstructured data (emails, videos, PDF documents).