# Course Outline: Detailed Notes

## Module 1: Introduction to Data Science

What is Data Science?  
Definition: Data Science is the interdisciplinary field of using scientific methods, algorithms, and systems to extract knowledge and insights from structured or unstructured data.  
Components:  
Data Collection: Gathering information from various sources. Real-time Example: Retail companies collecting customer data from online purchases, store visits, and loyalty programs to better understand buying patterns.  
Data Processing: Organizing and preparing data for analysis. Real-time Example: Financial institutions processing transaction data nightly to update customer balances and detect fraudulent transactions.  
Data Analysis: Exploring and analyzing data to find patterns or anomalies. Real-time Example: Healthcare providers analyzing patient data to identify which treatments lead to better patient outcomes.  
Predictive Modeling: Using statistical models to predict or forecast future events. Real-time Example: E-commerce platforms using predictive models to forecast product demand based on seasonal trends and promotional activities.  
Data Visualization: Presenting data visually to make the interpretation easier. Real-time Example: Energy companies using dashboards to visualize consumption patterns across different regions to manage supply efficiently.

## Module 2: Data Extraction, Wrangling, & Visualization

Data Analysis Pipeline  
Explanation: Steps taken from collecting data to making decisions based on that data.  
Real-time Example: A marketing firm analyzes customer data to optimize ad placements in real-time.  
Python Snippet:  
import pandas as pd  
data = pd.read\_csv('customer\_data.csv')  
print(data.head())