

# Data Science Fundamentals

## Introduction to Data Science

### What is Data Science?

Data Science is an interdisciplinary field that uses scientific methods, processes, algorithms, and systems to extract knowledge and insights from structured and unstructured data. It combines statistics, mathematics, programming, and domain expertise.

### The Data Science Process

#### 1. Problem Definition

- Understand the business problem

#### 2. Data Collection

- Identify data sources

#### 3. Data Cleaning and Preparation

- Handle missing values

#### 4. Exploratory Data Analysis (EDA)

- Statistical analysis

#### 5. Feature Engineering

- Create new features

#### 6. Model Building

- Select a model

#### 7. Model Evaluation

- Use appropriate metrics

#### 8. Deployment and Monitoring

- Deploy the model

### Key Skills for Data Scientists

## Programming

- Python

## Statistics and Mathematics

- Probability

## Machine Learning

- Supervised

## Data Visualization

- Matplotlib

## Domain Knowledge

- Understanding

## Common Tools and Technologies

### Programming Languages

- Python

### Big Data Tools

- Apache

### Databases

- SQL

### Visualization Tools

- Matplotlib

## Career Paths in Data Science

- Data Analyst

- Data Scientist