

# DATA SCIENCE

in collaboration with ST IIT Bombay

## COURSE CURRICULUM



**BSATES EdTech Foundation**

Section-8 EdTech Company

CIN : U88900DL2024NPL440152



## Duration

2 Days

14 Hours of Learning

## Introduction

This workshop has been designed to provide a holistic learning experience. The curriculum balances theoretical knowledge and hands-on practice, ensuring participants gain both foundational understanding and practical expertise. Interactive sessions, real-world use cases, and collaborative hackathons help embed key concepts effectively. The structured flow, from basic concepts to advanced applications, caters to diverse learning paces while promoting teamwork and problem-solving. This methodology ensures a robust, engaging, and outcome-driven learning journey for all participants.

## Venue

Indian Institute of Technology IIT Delhi



**BSATES Ed-Tech Foundation**

National HQ.- BSAITM Bhawan, Plot No. 13, Bodhella, Vikaspuri, New Delhi-110018



## DAY 1: FUNDAMENTALS OF DATA SCIENCE

### Session 1: Welcome & Icebreaker

#### Welcome & Introduction (15 minutes)

Brief overview of workshop agenda and objectives. Set the stage for learning and collaboration.

#### Interactive Discussion (45 minutes)

- Real-world applications of Data Science: Exploring how data is used for various tasks.
- Brainstorming potential use cases from participants' industries.
- Overview of data collection, cleaning, analysis, and visualization. Introduction to data pipelines and organizational data flow
- Roles in Data Science: Data Engineer, Data Analyst, Data Scientist

### Session 2: Intro to Google Sheets, basic features and Dashboarding

#### Core Concepts Overview

- Learning Google Sheets (45 minutes) Basic navigation of GoogleSheets. Core formulas: SUM, AVERAGE, COUNT, IF, VLOOKUP, INDEX-MATCH. Conditional formatting and data validation. Creating charts (bar, pie, line charts). Data filtering and sorting.
- Exercise 1 (15 minutes) Use a dataset to solve different data analytics problems. Discuss the roles of Data Engineers, Analysts, and Scientists in each step.
- Exercise 2 (30 minutes) Analyze a pre-provided dataset (e.g., e-commerce sales or hospitalpatient data). Visualize insights using Google Sheets.
- Exercise 3 (30 minutes) Create a dashboard for the given dataset using charts and calculated fields.

### Session 3: Power BI Data Analysis (1.5 Hours)

- Basics of Power BI
  1. Introduction to Power BI: Interface and functionalities.
  2. When to use Google Sheets vs. Power BI.
- Creating Dashboards in Power BI
  1. Different types of charts and when to use them.
  2. Hands-on exercise: Create a dashboard from scratch using a sample dataset.

### Session 4: Hackathon (1.5 Hours)

#### Team Activity

- Participants work on a data cleaning and visualization project.
- Start analyzing small data chunks in Google Sheets.
- Transition insights to Power BI dashboards.

#### Example Project

- Analyze customer feedback data to identify sentiment trends.
- Create dashboards and visualizations to present insights.

### Session 5: Wrap up & Reflection

#### Recap of Day 1 Concepts

Review the day's activities and key learnings.

#### Q&A Session

Open discussion for clarifications.

#### Preview of Day 2

Overview of advanced Data Science topics and Machine Learning..





## DAY 2: DATA SCIENCE AND REAL-WORLD APPLICATIONS

### Session 1: Foundation of Data Science (3 Hours)

#### Introduction to Data Science

- What is Data Science? Definition and importance.
- Key components: Statistics, Programming, Domain Knowledge.
- Applications across industries (e.g., Finance, Healthcare, Retail)

#### Data Science Workflow

- Problem definition, data collection, cleaning, analysis, modeling, evaluation, deployment.
- Real-world workflow examples.

#### Exploring Data Types of data: Structured vs. Unstructured.

- Basic statistics for data analysis: Mean, Median, Mode, Variance, Standard Deviation.
- Visualizing data: Using tools like Matplotlib or Power BI.

### Session 2: Beginner Friendly Machine Learning Concepts (3 Hours)

#### Supervised Learning

- Regression: Predicting continuous outcomes
- Classification: Predicting categorical outcomes
- Hands-On Exercise: Build a simple linear regression model using scikit-learn.

#### Basic ML Workflow

- Splitting data into training and test sets.
  - Training models and evaluating performance metrics.
1. Regression
  2. Classification

#### Unsupervised Learning Overview

- Clustering: Grouping unlabeled data (e.g., customer segmentation).
- Hands-On Exercise: Use K-Means clustering on a small dataset (e.g., Iris dataset).

#### Mini Project

- Build a classification model with the Industry dataset.
- Preprocess data, train a decision tree or logistic regression model, evaluate performance. WORKSHOP OUTCOMES

### Workshop Outcomes

- Understand the basics of Data Science and Analytics.
- Gain hands-on experience with tools like Google Sheets, Power BI, and scikit-learn.
- Solve real-world data problems and present insights effectively.
- Learn the fundamentals of Machine Learning.

### What Next ?

#### Continuing Education

- Explore advanced topics like deep learning, big data, and cloud computing.
- Use platforms like Kaggle or Coursera to enhance skills.

#### Practical Applications

- Apply concepts in your domain to improve workflows and decision-making.

#### Building a Portfolio

- Document projects and build dashboards/models to showcase expertise.