· **Course Title:**

DATA SCIENCE FOUNDATION (DSF)
Course Code: DS2010
Duration: 5 Days (40 hours)
Level: Foundation / Beginner
Target Audience: Aspiring data analysts, business professionals, IT staff, and students interested in data science.

· **Course Objective:**

The objective of this course is to introduce participants to the core principles and practices of data science. It provides foundational knowledge in data handling, statistical analysis, data visualization, and the use of tools like Python and Excel for data manipulation. By the end of the course, participants will be prepared to pursue further study in data science or apply basic data-driven decision-making in their roles.

✅**Course Outcomes:**

Upon successful completion of this course, participants will be able to:

1. Understand key concepts, processes, and lifecycle of data science.
2. Collect, clean, and prepare data for analysis.
3. Apply basic statistical and exploratory techniques.
4. Use data visualization tools to present insights clearly.
5. Gain hands-on experience using Python and Excel for data analysis.
6. Understand the fundamentals of machine learning and its applications.

· **5-Day Course Outline:**

· **Day 1: Introduction to Data Science**

- What is Data Science?
- Roles: Data Analyst vs. Data Scientist vs. Data Engineer
- The Data Science Lifecycle
- Overview of Tools: Excel, Python, Jupyter Notebook, SQL

- Introduction to Data Ethics and Privacy

· **Day 2: Data Collection & Preparation**

- Types and Sources of Data
- Structured vs. Unstructured Data
- Data Cleaning Techniques
- Data Wrangling with Excel and Python (Pandas)
- Missing Value Handling and Outliers

· **Day 3: Exploratory Data Analysis (EDA)**

- Descriptive Statistics
- Mean, Median, Mode, Variance, Standard Deviation
- Correlation and Covariance
- Hands-on EDA in Excel and Python
- Basic Data Visualization (Bar, Line, Histogram, Box Plot)

· **Day 4: Data Visualization Techniques**

- Principles of Effective Visualization
- Using Python Libraries (Matplotlib, Seaborn)
- Dashboarding in Excel (Pivot Tables, Charts)
- Case Study: Visual Storytelling with Data
- Identifying Patterns and Trends

· **Day 5: Introduction to Machine Learning**

- What is Machine Learning?
- Supervised vs Unsupervised Learning
- Simple Linear Regression Concept
- Introduction to Scikit-learn in Python
- Course Project: Mini Data Science Challenge

Course Objective:

The objective of this course is to introduce participants to the core principles and practices of data science. It provides foundational knowledge in data handling, statistical analysis, data visualization, and the use of tools like Python and Excel for data manipulation. By the end of the course, participants will be prepared to pursue further study in data science or apply basic data-driven decision-making in their roles.

Target audience:

Malaysian citizen who have been retrenched or are seeking new job opportunities.

Course outline:

Part 1: Introduction to Data Science

Part 2: Data Collection & Preparation

Part 3: Exploratory Data Analysis (EDA)

Part 4: Data Visualization Techniques

Part 5: Introduction to Machine Learning

Course outcomes:

Upon successful completion of this course, participants will be able to:

1. Understand key concepts, processes, and lifecycle of data science.
2. Collect, clean, and prepare data for analysis.
3. Apply basic statistical and exploratory techniques.
4. Use data visualization tools to present insights clearly.
5. Gain hands-on experience using Python and Excel for data analysis.
6. Understand the fundamentals of machine learning and its applications.