5-Day Workshop Plan

Data Science

Day 1: Introduction to Data Science and Python

Session Overview: Day 1 will provide an introduction to the field of data science, basic data analysis concepts, and Python for data science.

Schedule:

- **10:00 AM - 10:15 AM: Welcome & Introduction**
- **Overview of the workshop, introduction to data science, and key skills.**
- **10:15 AM - 11:15 AM: Introduction to Python for Data Science**
- **Learn Python basics for data science: variables, data types, lists, and libraries like NumPy and Pandas.**
- **11:15 AM - 11:30 AM: Break**
- **11:30 AM - 1:00 PM: Data Manipulation with Pandas**
- **Introduction to Pandas for data manipulation: loading, cleaning, and transforming datasets.**
- **2:00 PM - 2:15 PM: Recap and Q&A**
- **Review key concepts from Python and Pandas.**
- **2:15 PM - 3:15 PM: Data Cleaning and Preprocessing**
- **Learn techniques for cleaning and preprocessing data: handling missing values, duplicates, and outliers.**
- **3:15 PM - 3:30 PM: Break**
- **3:30 PM - 5:00 PM: Hands-on: Data Cleaning Practice**
- **Practical session: Clean and preprocess a sample dataset using Pandas.**

sltechhsolutions@gmail.com

91-8341475919

5-Day Workshop Plan

Data Science

Day 2: Data Visualization and Exploratory Data Analysis (EDA)

Session Overview: Day 2 will focus on data visualization and EDA techniques using libraries like Matplotlib and Seaborn.

Schedule:

- **10:00 AM - 10:15 AM: Recap of Day 1**
- **Quick review of Python, Pandas, and data cleaning concepts.**
- **10:15 AM - 11:15 AM: Introduction to Data Visualization**
- **Overview of data visualization principles and tools: Matplotlib and Seaborn.**
- **11:15 AM - 11:30 AM: Break**
- **11:30 AM - 1:00 PM: Creating Visualizations with Matplotlib and Seaborn**
- **Learn how to create different types of plots: histograms, scatter plots, line plots, and box plots.**
- **2:00 PM - 2:15 PM: Recap and Q&A**
- **Clarify any doubts related to data visualization techniques.**
- **2:15 PM - 3:15 PM: Exploratory Data Analysis (EDA)**
- **Introduction to EDA: summarizing datasets, understanding distributions, correlations, and trends.**
- **3:15 PM - 3:30 PM: Break**
- **3:30 PM - 5:00 PM: Hands-on: EDA Practice**
- **Participants perform EDA on a given dataset, exploring and visualizing trends.**

sltechhsolutions@gmail.com

91-8341475919

5-Day Workshop Plan

Data Science

Day 3: Introduction to Statistics for Data Science

Session Overview: Day 3 will introduce statistical concepts and techniques commonly used in data science for analyzing data.

Schedule:

- **10:00 AM - 10:15 AM: Recap of Day 2**
- **Review of data visualization and EDA techniques.**
- **10:15 AM - 11:15 AM: Descriptive Statistics**
- **Learn about measures of central tendency (mean, median, mode), dispersion (variance, standard deviation), and data distribution.**
- **11:15 AM - 11:30 AM: Break**
- **11:30 AM - 1:00 PM: Probability and Distributions**
- **Introduction to probability concepts and common probability distributions: Normal, Binomial, and Poisson.**
- **2:00 PM - 2:15 PM: Recap and Q&A**
- **Clarify doubts regarding descriptive statistics and probability.**
- **2:15 PM - 3:15 PM: Hypothesis Testing**
- **Learn about hypothesis testing, p-values, and types of errors.**
- **3:15 PM - 3:30 PM: Break**
- **3:30 PM - 5:00 PM: Hands-on: Statistical Analysis**
- **Perform statistical tests on sample data, including t-tests and chi-square tests.**

sltechhsolutions@gmail.com

91-8341475919

5-Day Workshop Plan

Data Science

Day 4: Introduction to Machine Learning

Session Overview: Day 4 will introduce machine learning concepts and algorithms, focusing on supervised learning.

Schedule:

- **10:00 AM - 10:15 AM: Recap of Day 3**
- **Review of statistical concepts and hypothesis testing.**
- **10:15 AM - 11:15 AM: Introduction to Machine Learning**
- **Learn the basics of machine learning, types of learning (supervised, unsupervised), and key algorithms.**
- **11:15 AM - 11:30 AM: Break**
- **11:30 AM - 1:00 PM: Supervised Learning: Regression**
- **Introduction to regression models: Linear Regression, understanding relationships in data.**
- **2:00 PM - 2:15 PM: Recap and Q&A**
- **Address any questions on machine learning and regression.**
- **2:15 PM - 3:15 PM: Supervised Learning: Classification**
- **Learn classification models: Logistic Regression, Decision Trees, and K-Nearest Neighbors (KNN).**
- **3:15 PM - 3:30 PM: Break**
- **3:30 PM - 5:00 PM: Hands-on: Implementing Regression and Classification**
- **Implement linear regression and classification models using scikit-learn.**

sltechhsolutions@gmail.com

91-8341475919

5-Day Workshop Plan

Data Science

Day 5: Advanced Machine Learning and Model Evaluation

Session Overview: Day 5 will cover model evaluation techniques, advanced machine learning algorithms, and model deployment.

Schedule:

- **10:00 AM - 10:15 AM: Recap of Day 4**
- **Review of regression, classification, and machine learning concepts.**
- **10:15 AM - 11:15 AM: Model Evaluation Techniques**
- **Learn about model evaluation metrics: accuracy, precision, recall, F1-score, confusion matrix, and cross-validation.**
- **11:15 AM - 11:30 AM: Break**
- **11:30 AM - 1:00 PM: Advanced Machine Learning: Ensemble Methods**
- **Introduction to ensemble methods like Random Forests, Gradient Boosting, and XGBoost.**
- **2:00 PM - 2:15 PM: Recap and Q&A**
- **Clarify doubts on model evaluation and ensemble methods.**
- **2:15 PM - 3:15 PM: Model Deployment and Final Project Overview**
- **Overview of deploying machine learning models to production using tools like Flask or FastAPI.**
- **3:15 PM - 3:30 PM: Break**
- **3:30 PM - 5:00 PM: Hands-on: Final Project**
- **Participants work on their final project, building and evaluating a machine learning model and preparing it for deployment.**

sltechhsolutions@gmail.com

91-8341475919