Lab Examination Question Paper Sets

Set – 1

1. Create NumPy arrays using intrinsic objects and random functions. Perform reshaping and joining operations.
2. Import a CSV file to a Pandas DataFrame. Handle missing values, detect outliers, and compute statistical summaries.
3. Plot histograms and density plots for numeric columns to analyze distribution patterns.
4. **Implement a K-Nearest Neighbors (KNN) classifier on the Iris dataset and evaluate accuracy using confusion matrix.**

Set – 2

5. Create a Pandas DataFrame from a dictionary and perform ranking, sorting, and aggregation operations.
6. Detect and handle skewness in a dataset, and normalize the data using Min-Max scaling.
7. Plot scatter plots and correlation heatmaps to identify relationships between features.
8. **Build a Support Vector Machine (SVM) classifier with a linear kernel and visualize decision boundaries.**

Set – 3

9. Create NumPy arrays from lists and perform indexing, slicing, and universal function operations.
10. Import a dataset, perform label encoding for categorical data, and handle missing values using imputation.
11. Visualize feature relationships using scatter plots with regression lines and bar plots with error bars.
12. **Develop a Decision Tree Classifier using Gini Index.**

Set – 4

13. Use Pandas to import a dataset and detect outliers using the IQR method. Replace extreme values with median.
14. Perform correlation and feature importance analysis using statistical and visual methods.
15. Generate subplots for line, bar, and density plots of multiple variables.
16. **Train and evaluate a Multiple Linear Regression model. Plot residuals and regression line.**

Set – 5

17. Create NumPy arrays and perform arithmetic and statistical operations using universal functions.
18. Import a CSV file, drop duplicate records, and compute group-wise statistics.

19. Plot overlapping histograms and compare distributions among features.
20. **Implement a Logistic Regression model for binary classification and display confusion matrix and ROC curve.**

Set – 6

21. Create a Pandas DataFrame from a CSV file and apply transformation using apply() and map().
22. Handle missing data, detect outliers, and scale numerical features using standardization.
23. Create a line plot and scatter plot with regression line showing trends between two key variables.
24. **Build a KNN Classifier with different k-values and compare accuracy scores using plots.**

Set – 7

25. Perform element-wise computations and matrix operations using NumPy arrays.
26. Create bar plots and histograms for price-related features.
27. Plot bar and density plots to compare feature distributions across categories.
28. **Build a Linear Regression model to predict house prices and plot residuals.**

Set – 8

29. Create a Pandas DataFrame and perform advanced string manipulations on textual columns.
30. Handle missing data using imputation and apply feature scaling for numeric columns.
31. Plot scatter plots and line charts to visualize relationships between numerical features.
32. **Implement a Support Vector Machine (SVM) model for classification and visualize results.**

Extra Questions

33. Perform data preprocessing on the House Pricing dataset from Kaggle (handle missing values, feature encoding).
34. Analyze relationships among variables using correlation matrices and heatmaps.

35. Import a dataset, detect skewness, and apply log or box-cox transformation to normalize data.
36. Plot scatter plots with regression lines and density plots to visualize results.