**Employee Salary Analysis Report**

By – Ishika Sharma

Roll no. - 202401100300129

Course - Introduction to AI

Date - 11.03.2025

**Introduction**

This report presents an analysis of employee salaries based on key factors such as department, experience, and age. The data includes 20 employees across multiple departments.

**Methodology**

The following methodology was adopted to analyse the given data:

1. Data Collection: The data was imported from a CSV file containing columns for Employee ID, Age, Department, Experience, and Salary.
2. Data Cleaning: The dataset was verified to ensure all required columns were present.
3. Statistical Analysis: Key metrics such as average salary, highest salary, and lowest salary were calculated.
4. Department-Wise Analysis: Average salaries were calculated for each department.
5. Visualization: Various visualizations were created to better understand salary distributions and trends.

Code: import pandas as pd

import matplotlib.pyplot as plt

import seaborn as sns

def analyze\_salaries():

    """

    Function to analyze and visualize employee salaries from a CSV file.

    The user will input the file path once, and it will be reused.

    """

    # Load data from CSV file

    df = pd.read\_csv("employee\_data.csv")

    # Check if required columns exist

    required\_columns = {'EmployeeID', 'Age', 'Department', 'Experience', 'Salary'}

    if not required\_columns.issubset(df.columns):

        print("Error: CSV file must contain columns: EmployeeID, Age, Department, Experience, Salary")

        return

    # Calculate overall statistics

    avg\_salary = df['Salary'].mean()

    max\_salary = df['Salary'].max()

    min\_salary = df['Salary'].min()

    # Calculate department-wise average salary

    dept\_avg\_salary = df.groupby('Department')['Salary'].mean()

    # Display results

    print("\nEmployee Salary Analysis:")

    print(f"Total Employees: {len(df)}")

    print(f"Average Salary: {avg\_salary:.2f}")

    print(f"Highest Salary: {max\_salary}")

    print(f"Lowest Salary: {min\_salary}")

    print("\nDepartment-wise Average Salary:")

    print(dept\_avg\_salary)

    # Visualization

    plt.figure(figsize=(10, 5))

    sns.barplot(x=dept\_avg\_salary.index, y=dept\_avg\_salary.values, palette='viridis')

    plt.xlabel("Department")

    plt.ylabel("Average Salary")

    plt.title("Average Salary by Department")

    plt.xticks(rotation=45)

    plt.show()

    plt.figure(figsize=(10, 5))

    sns.histplot(df['Salary'], bins=10, kde=True, color='blue')

    plt.xlabel("Salary")

    plt.ylabel("Frequency")

    plt.title("Salary Distribution")

    plt.show()

    plt.figure(figsize=(10, 5))

    sns.scatterplot(x=df['Experience'], y=df['Salary'], hue=df['Department'], palette='deep')

    plt.xlabel("Experience (Years)")

    plt.ylabel("Salary")

    plt.title("Salary vs Experience")

    plt.legend(title='Department')

    plt.show()

# Run the analysis function

analyze\_salaries()

Screenshots of output:

  