

# **Analysis of Employee Salaries**

Date : 10 March , 2025

Name: Nikhil

Branch : CSE (AI&ML)

Section: B

Roll No: 202401100400125

---

## Introduction

This analysis's goal is to use a dataset to investigate relationships between employee pay and job positions. Understanding how pay differs by department, experience, and work position is made easier by the analysis. Matplotlib will be used for visualization, and Python and Pandas will be used for data manipulation.

---

## Methodology:-

In order to prepare the dataset, we will first construct a CSV file with the following information about each employee: Employee ID, Name, Job Position, Department, Salary, and Experience.

- 1. Data Loading:** A Pandas DataFrame will be loaded with the CSV file.
- 2. Data Cleaning:** We'll look for discrepancies and missing values.
- 3. Exploratory Data Analysis (EDA):** Using histograms and other statistical tools, we will produce statistical summaries and illustrate trends.

## References & Credits -

Used Libaried - matplotlib , seaborn ,pandas

## CODE

```
import pandas as pd
import seaborn as sns
import matplotlib.pyplot as plt
```

```
# Step 1: Load the csv file(The csv file should exist in the same
folder as the script)
file_name = "data.csv"

try:
    data = pd.read_csv(file_name)
    print(f"File loaded Successfully\n")
except FileNotFoundError:
    print(f"'{file_name}' not found Please place 'data.csv' in the same
folder as this script.")
    exit()

# Step 2:basic statistics of Salary and Experience
print("Provided Statistics of Salary and Experience:")
print(data[["Salary", "Experience"]].describe(), "\n")

# Step 3: Plot correlation heatmap And also filter numeric values
plt.figure(figsize = (8, 5))
sns.heatmap(data[["Salary", "Experience"]].corr(), annot=True,
cmap="coolwarm", fmt=".2f")
plt.title(" Correlation Heatmap of Salary and Experience")
plt.show()

# Step 4: Salary Distribution
plt.figure(figsize = (8, 5))
sns.histplot(data["Salary"], bins = 10, kde = True, color = "blue")
plt.title(" Salary Distribution of Employees")
plt.xlabel("Salary")
plt.ylabel("Number of Employees")
plt.show()

# Step 5: Salary vs Experience Scatter Plot
plt.figure(figsize = (8, 5))
sns.scatterplot( x = data["Experience"], y = data["Salary"], hue =
data["Department"], s=100)
plt.title("Salary vs Experience")
plt.xlabel("Years of Experience")
plt.ylabel("Salary")
plt.legend(title = "Department", bbox_to_anchor = (1, 1), loc = "upper
left")
plt.show()
```

## OUTPUT

File loaded Successfully

Provided Statistics of Salary and Experience:

	Salary	Experience
count	1000.000000	1000.000000
mean	101077.152000	10.484000
std	28526.403317	5.808397
min	50062.000000	1.000000
25%	76863.250000	5.000000
50%	103198.000000	10.000000
75%	124136.750000	15.000000
max	149545.000000	20.000000



