**Machine Learning**

**Ex: 03 Data Preprocessing**

**Name:** R.Athithraja

**Reg.no:** 2022503702

1. Handling Missing Values

**Code:**

import pandas as pd

df = pd.read\_excel("Salary\_Data\_miss.xlsx")

print(df.isnull().sum())

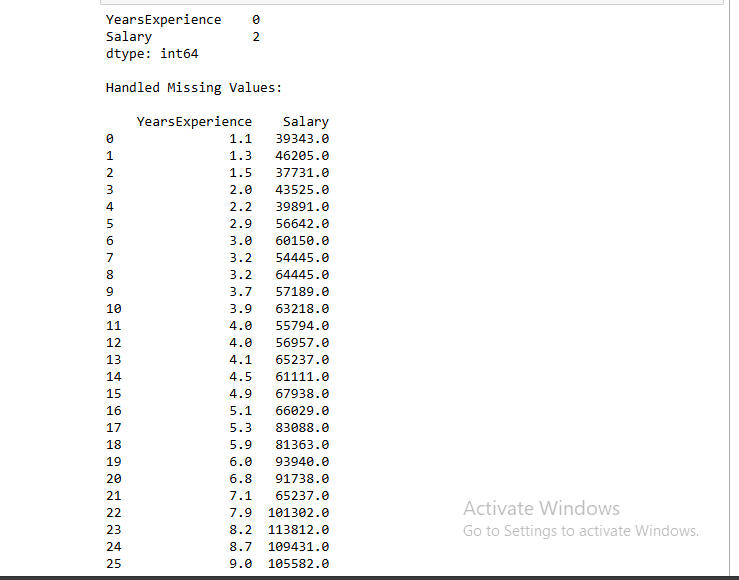
df = df.fillna(df.median())

#df = df.dropna()

print("\nHandled Missing Values: \n")

print(df)

**Output:**



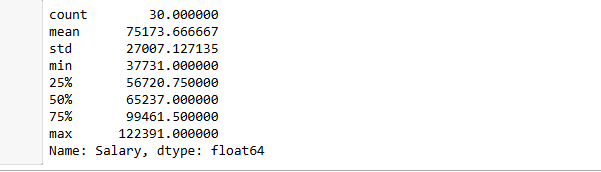
2. Five-Number Summary

**Code:**

summary = df['Salary'].describe()

print(summary)

**Output:**



3. Box Plot

**Code:**

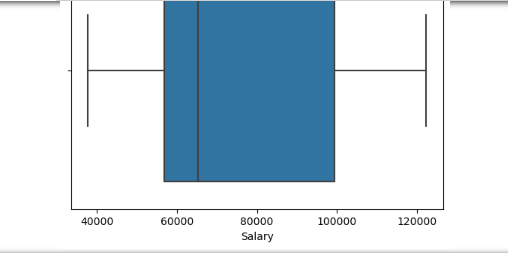
import seaborn as sns

import matplotlib.pyplot as plt

sns.boxplot(x=df['Salary'])

plt.show()

**Output:**



4. Correlation Matrix

**Code:**

corr\_matrix = df.corr()

print(corr\_matrix)

**output:**



5. Covariance Matrix

**Code:**

cov\_matrix = df.cov()

print(cov\_matrix)

**Output:**

