Green University of Bangladesh
Department of Computer Science and Engineering (CSE)
Faculty of Sciences and Engineering
Semester: (Fall, Year:2025), B.Sc. in CSE (Day)

Lab Report 1
Course Title: Data Mining Lab
Course Code: CSE 436 **Section:** 221 D2

Lab Experiment Name: Statistical Visualization of Income Data Using Seaborn: A Multi-Plot Analytical Report

Student Details

	Name	ID
1.	Md. Maruf Sarker	221002063

Submission Date : 15 November 2025
Course Teacher's Name : Md. Ataullha

Lab Report Status
Marks: ..... Signature:.....
Comments:..... Date:.....

1. TITLE OF THE LAB REPORT EXPERIMENT

Statistical Visualization of Income Data Using Seaborn: A Multi-Plot Analytical Report

2. OBJECTIVES/AIM

The main objective of this lab was to learn how to visualize data using Python's Seaborn library. I also wanted to understand how age and income might be connected and see if visual graphs can show any patterns clearly.

3. PROCEDURE / ANALYSIS / DESIGN

- First I downloaded the CSV file called **income.csv** that contains Name, Age, and Income.
- I loaded the data in Python using pandas.
- Then I picked 5 different types of plots we were asked to make: heatmap, boxplot, violin plot, barplot, and histogram.
- After making every plot, I saved it in PDF format.
- At the end, I reviewed all the graphs together to understand the final story from the visuals.

4. IMPLEMENTATION

I used **Python**, **Pandas**, **Seaborn**, and **Matplotlib**. All plots were written in VS Code. The code I used is included at the end of the report.

5. TEST RESULT / OUTPUT

Heatmap:

```
9   # Heatmap (Correlation Matrix)
10  plt.figure(figsize=(6,4))    "figsize": Unknown word.
11  corr = data[["Age", "Income($)"]].corr()
12
13  sns.heatmap(corr, annot=True, cmap="Blues", linewidths=0.5)
14  plt.title("Correlation Heatmap: Age vs Income")
15
16  plt.savefig("heatmap.pdf")
17  plt.show()
```

Fig-1: Heatmap (Correlation Matrix)

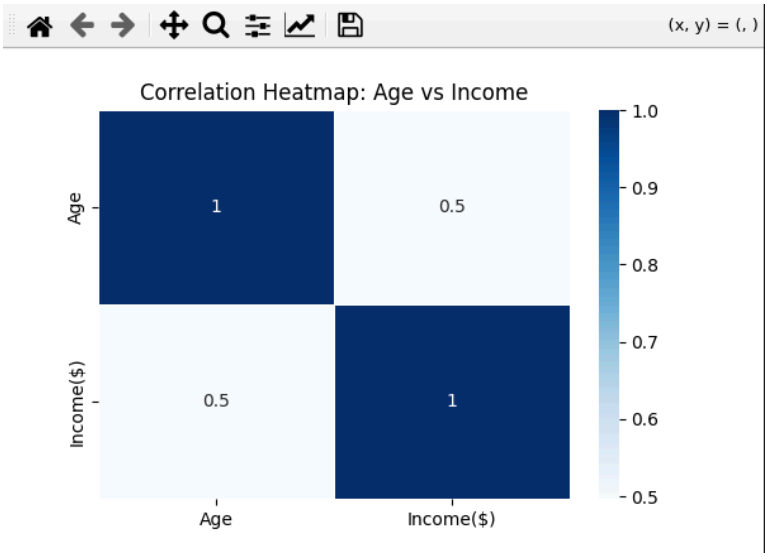

Fig-2: Heatmap (Correlation Matrix)

I think this heatmap kinda shows that age and income are pretty related because the color is strong positive. Looks like as age goes up income also goes up.

Boxplot:

```
19  # Boxplot (Income Distribution)
20  plt.figure(figsize=(6,4))    "figsize": Unknown wor
21  sns.boxplot(y=data["Income($)"], color="skyblue")
22
23  plt.title("Boxplot of Income")
24  plt.ylabel("Income ($)")    "ylabel": Unknown word.
25
26  plt.savefig("boxplot.pdf")
27  plt.show()
```

Fig-3: Boxplot (Income Distribution)

Fig-4: Boxplot (Income Distribution)

The boxplot basically shows income is spread out a lot. Some people earn way more than others and the box is long so income varies a lot.

Violin Plot:

```
30  # Violin Plot (Age and Income Distributions)
31  plt.figure(figsize=(7,5))    "figsize": Unknown word.
32  sns.violinplot(data=data[['Age', 'Income($)']])    "viol
33
34  plt.title("Violin Plot: Age and Income Distributions")
35  plt.xlabel("Variables")    "xlabel": Unknown word.
36
37  plt.savefig("violin.pdf")
38  plt.show()
```

Fig-5: Violin Plot (Age and Income Distribution)

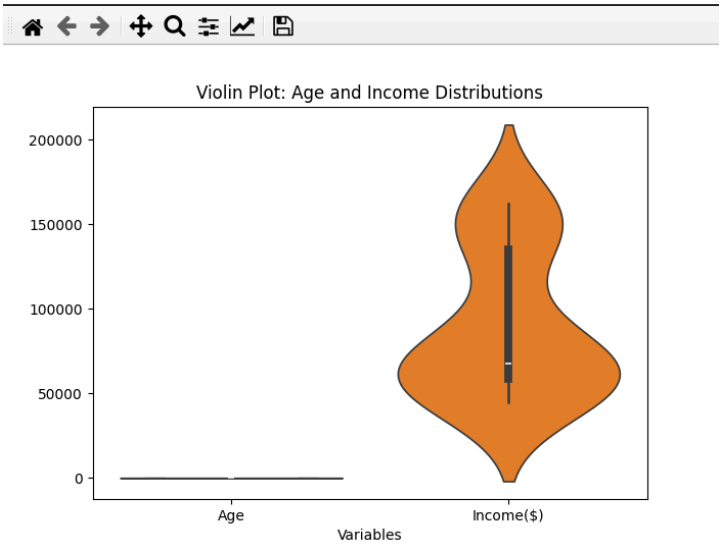

Fig-6: Violin Plot (Age and Income Distribution)

In the violin plot the income part looks wider so I think that means people have very different salaries. The age one is more narrow so ages are closer together.

Barplot:

```
41  # Barplot (Income by Person)    "Barplot": Unknown word.
42  plt.figure(figsize=(12,5))    "figsize": Unknown word.
43  sns.barplot(x="Name", y="Income($)", data=data, palette="viridis")
44
45  plt.xticks(rotation=90)    "xticks": Unknown word.
46  plt.title("Barplot of Income by Person")    "Barplot": Unknown word.
47  plt.ylabel("Income ($)")    "ylabel": Unknown word.
48  plt.xlabel("Name")    "xlabel": Unknown word.
49
50  plt.savefig("barplot.pdf")    "barplot": Unknown word.
51  plt.show()
```

Fig-7: Barplot (Income by Person)

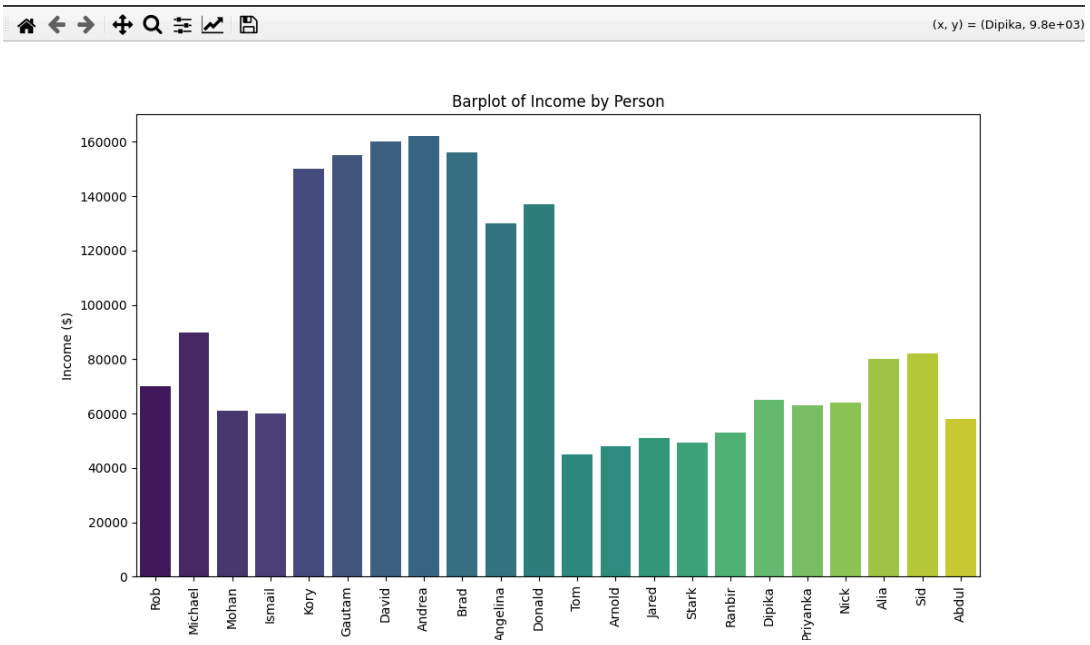

Fig-8: Barplot (Income by Person)

This bar graph shows each person's income. You can clearly see older people like Gautam, Andrea etc have more income than the younger ones.

Histogram:

```
# Histogram (Distribution of Income)
plt.figure(figsize=(7,4))    "figsize": Unknown word.
sns.histplot(data["Income($)"], bins=10, kde=True, color="steelblue")

plt.title("Histogram of Income")
plt.xlabel("Income ($)")    "xlabel": Unknown word.
plt.ylabel("Frequency")    "ylabel": Unknown word.

plt.savefig("histogram.pdf")
plt.show()
```

Fig-9: Histogram (Distribution of Income)

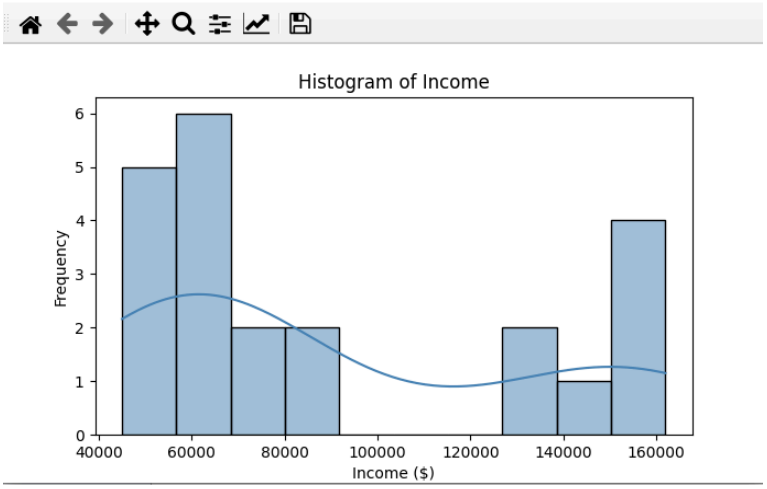

Fig-9: Histogram (Distribution of Income)

The histogram shows two bumps, like two groups of income. Some people earn around 50k and others earn like 130k+. So it's not even distributed.

6. Linkedin Post