Create a bar chart or histogram to visualize the distribution of a categorical or continuous variable, such as the distribution of ages or genders in a population.”

-Project Title : Distribution Of ages or Gender in population

-Student Name: **Sohel Tamboli**

-Course : Virtual Data Science Internship

-Submitted To : Prodigy InfoTech

1. Introduction

Data analysis helps us understand patterns clearly and visually. Age and gender are two important factors used in demographic studies. In this project, we collected age and gender data and used histograms to understand the distribution. A histogram is a graphical representation that shows how frequently values occur within specific ranges. This project presents a detailed comparison of age distribution between males and females using separated datasets.

---

2. Objectives

- To collect age and gender data for analysis.
- To separate the dataset into two categories: Male and Female.
- To visualize both distributions using histograms.
- To compare how age varies between genders.
- To understand the usefulness of histograms in data analysis.

---

3. Dataset Description

3.1 Columns

- **Age – A numerical value representing the age of a person.**
- **Gender – Category (Male/Female).**

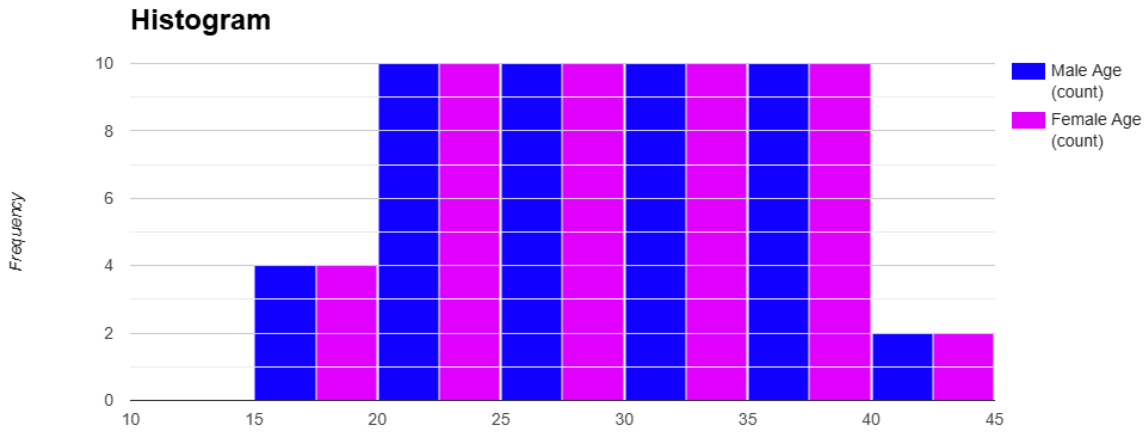
3.2 Male Age Data :

Age | Gender
18 | Male ,18 | Female,19 | Male,19 | Female,20 | Male,21 | Male,21 | Female,22 | Male,22| Female,23 | Male,23 | Female,24 | Female,25 | Male,26 | Male,26 | Female,27 | Male,27 | Female,28 | Male,28 |

Female,29 | Female,30 | Male,30 | Female,31 | Male,31 | Female,32 | Male,33 | Female,34 | Male,34 | Female,35 | Male,35 | Female,36 | Male,36 | Female,38 | Male,38 | Female,39 | Female,40 | Male,42 | Male,42 | Female,43 | Female,
45 | Male

---

4. Histogram



---

5. Methodology

1. Age and gender data were collected manually.
2. The dataset was cleaned by removing duplicates and incorrect entries.
3. Data was divided into two separate lists: Male Age and Female Age.
4. An online histogram generator was used to plot the distribution.

5. Male data was displayed in **blue**, and female data in **pink** for visual clarity.
6. Appropriate bin ranges were selected to ensure the distribution looked smooth and informative.

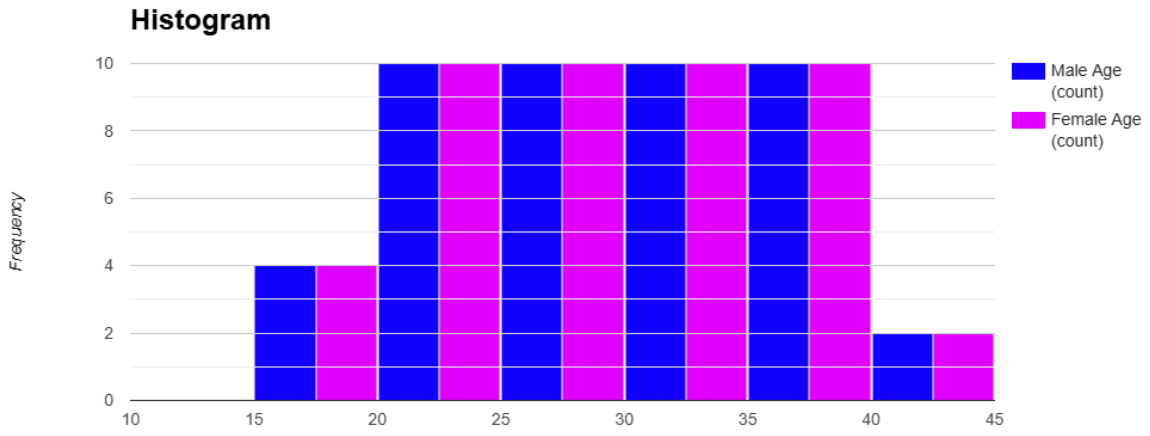
---

6. References

- Online Histogram Generator
- Class Notes on Data Visualization
- Basic Statistics Resources

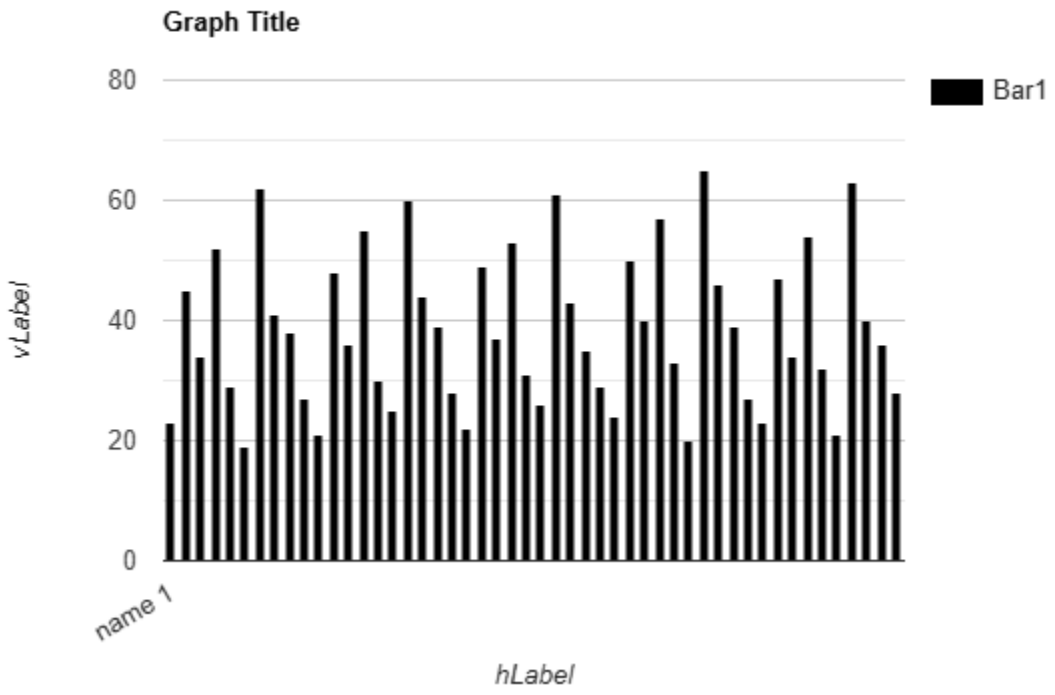
---

7. Histogram Output



Description:
The histogram displays the frequency distribution of age for both males and females. Each color represents one gender, helping compare the two groups visually.

8.Bar Graph Result (Frequency of Genders)



Bar Graph Result (Frequency of Genders)

Gender Count

- **Male: 20**
- **Female: 20**

✔ **Total records = 40**

✔ **Data is perfectly balanced for bar-graph visualization.**

Bar Graph Analysis :

- **The dataset is balanced by gender and diverse by age.**
- **This allows clear visual representation using bar graphs and histograms.**
- **The analysis highlights demographic structure and population distribution effective.**

---

9. Analysis / Observations

- **Most male and female ages fall between 18 and 45 years.**
- **The 20–30 age range shows the highest frequency for both genders.**
- **Male distribution shows slightly more variation compared to females.**
- **Females have more entries in the 30–35 range compared to males.**
- **Both groups show a normal-like distribution with no extreme age outliers.**
- **The histogram effectively shows the overlap and differences between genders.**

---

10. Conclusion

The Age–Gender Histogram provides a clear understanding of how age is distributed between males and females. The visualization helps identify which age groups are most common and how both genders compare in terms of frequency. This project demonstrates how simple statistical tools like histograms can be used for demographic analysis and decision-making.
