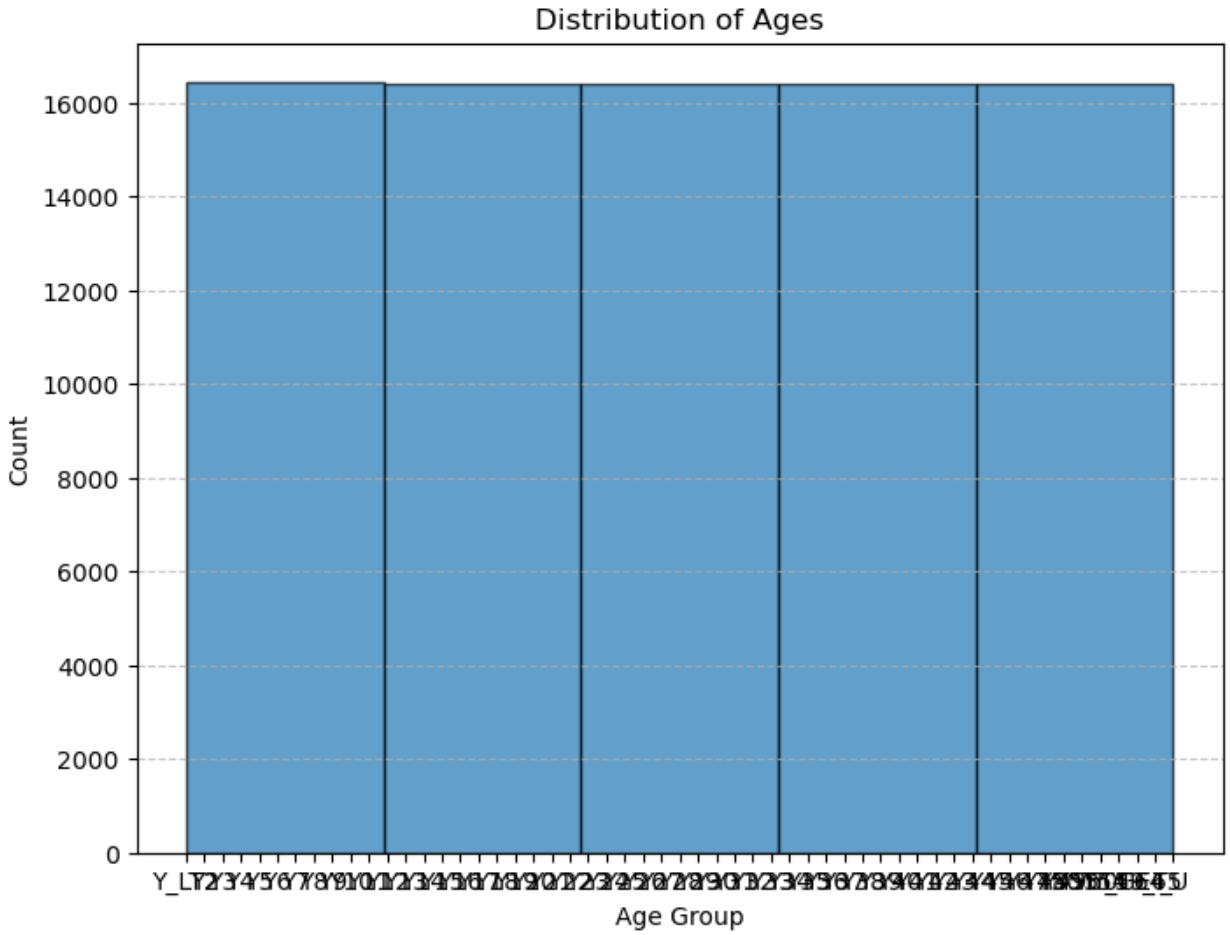
PRODIGY_DS_01

Task-01

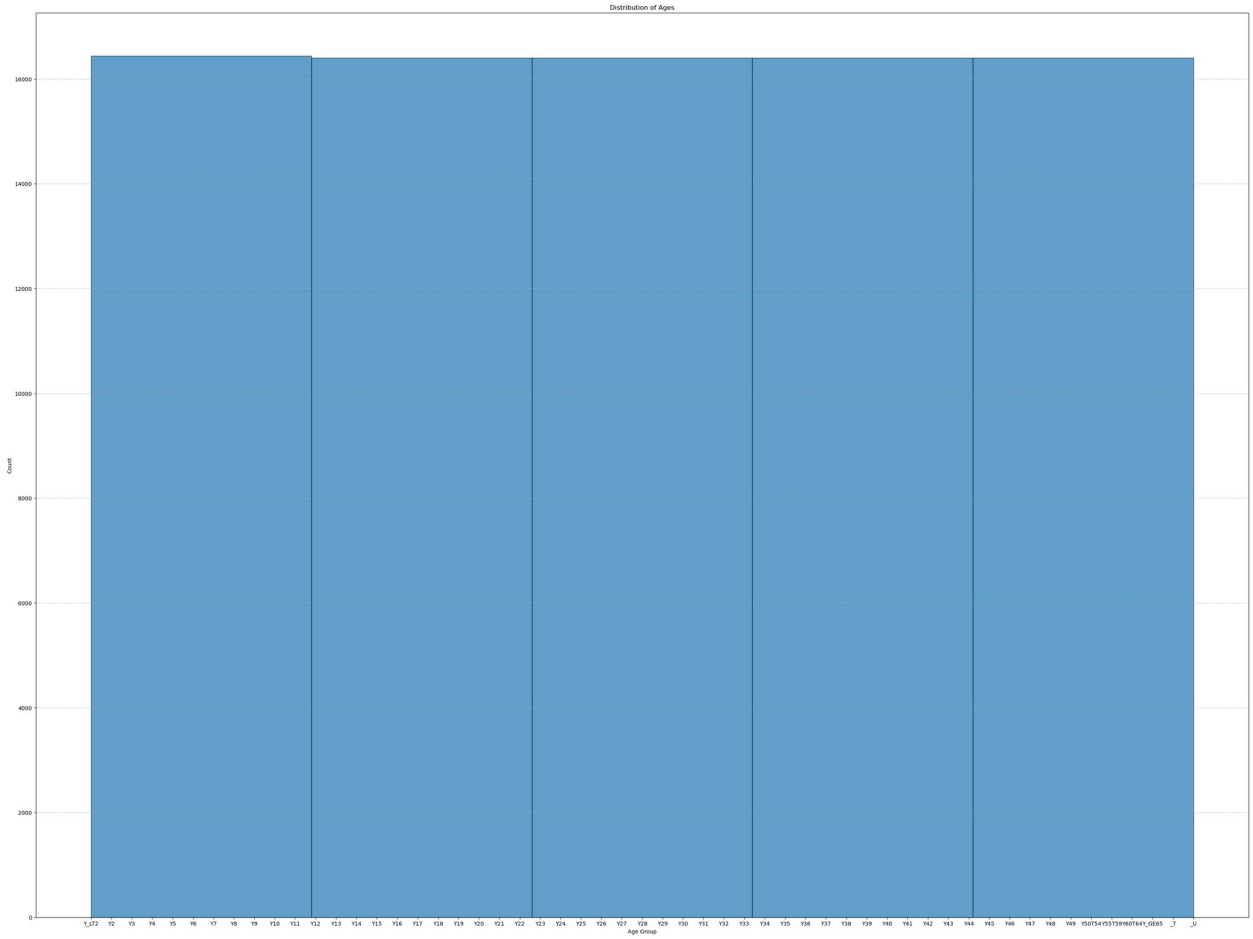
Create a bar chart or histogram to visualize the distribution of a categorical or continuous variable, such as the distribution of ages or genders in a population.

Sample Dataset :- https://data.worldbank.org/indicator/SP.POP.TOTL

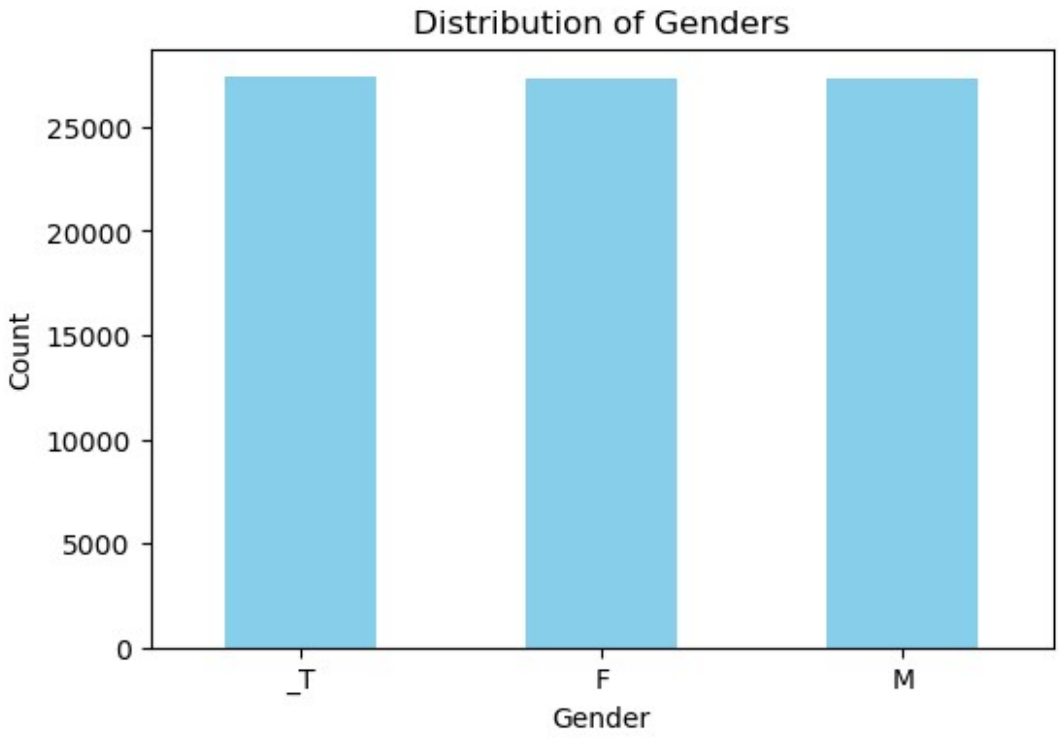
```
import pandas as pd
import matplotlib.pyplot as plt

import seaborn as sns
import numpy as np

# Load your dataset from a CSV file
file_path = 'C:\\Users\\Narthana\\Downloads\\
EDU_DEM_02092023062920970.csv'
df = pd.read_csv(file_path)

# Create a histogram to visualize the distribution of ages
plt.figure(figsize=(8, 6))
plt.hist(df['AGE'], bins=5, edgecolor='k', alpha=0.7)
plt.xlabel('Age Group')
plt.ylabel('Count')
plt.title('Distribution of Ages')
plt.grid(axis='y', linestyle='--', alpha=0.7)
plt.show()
```



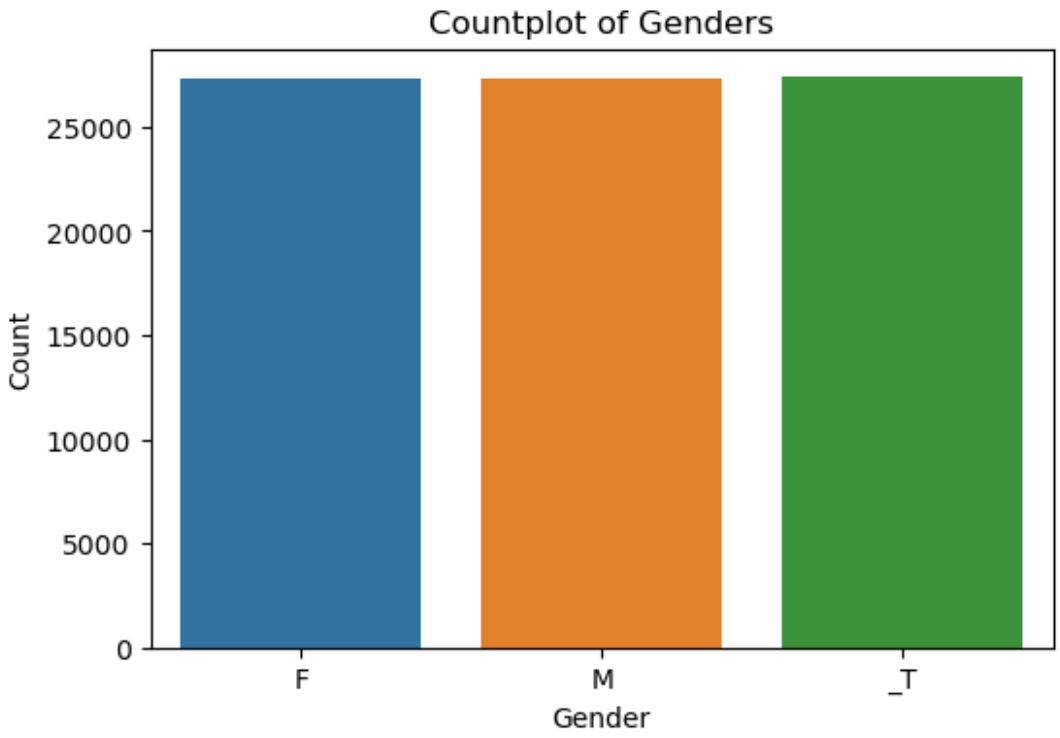

```
# Create a histogram to visualize the distribution of ages with a
larger figure size
plt.figure(figsize=(40, 30))  # Adjust the figure size as needed
plt.hist(df['AGE'], bins=5, edgecolor='k', alpha=0.7)
plt.xlabel('Age Group')
plt.ylabel('Count')
plt.title('Distribution of Ages')
plt.grid(axis='y', linestyle='--', alpha=0.7)
plt.show()
```



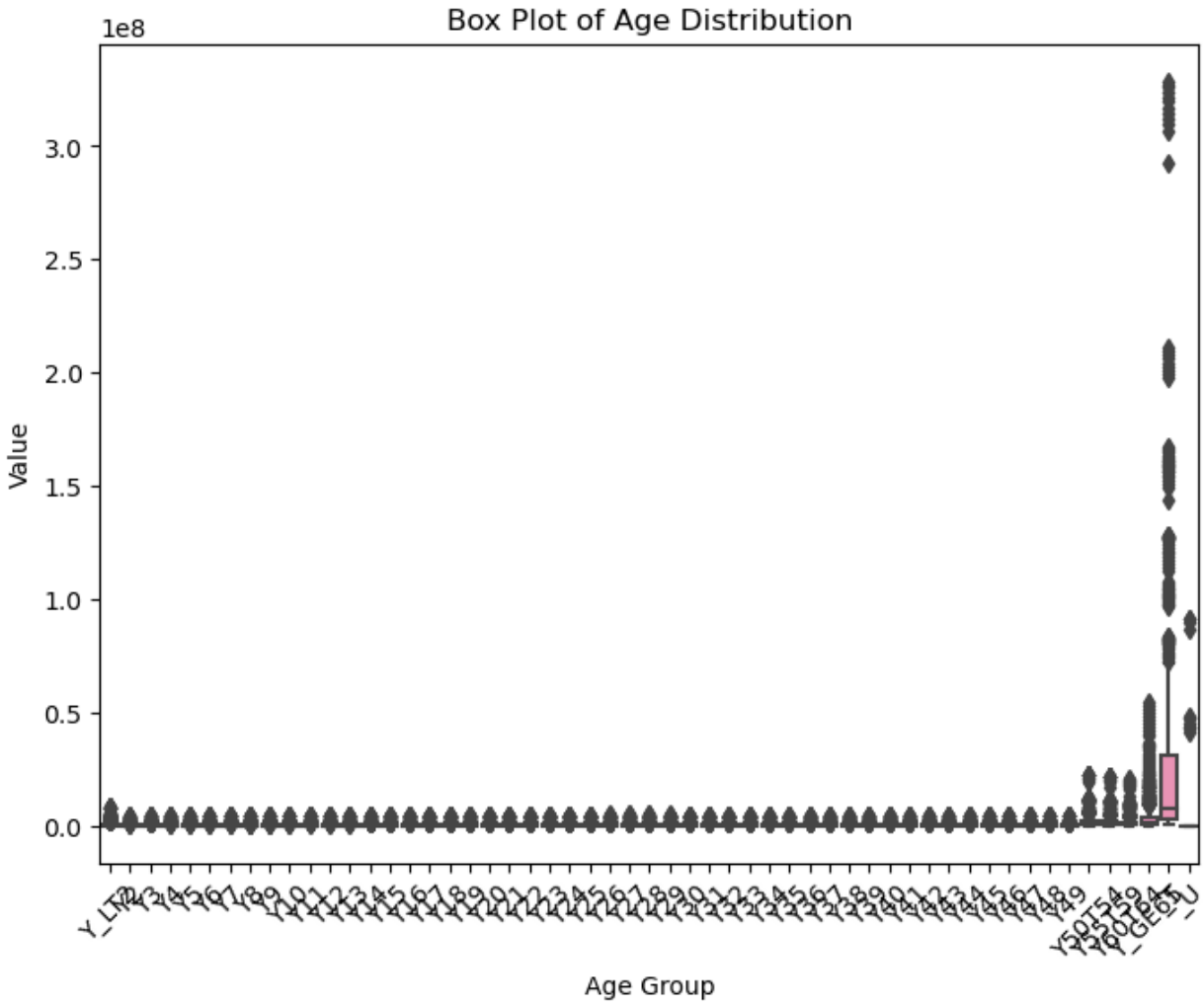

```
# Create a bar chart to visualize the distribution of genders
gender_counts = df['SEX'].value_counts()
plt.figure(figsize=(6, 4))
gender_counts.plot(kind='bar', color='skyblue')
plt.xlabel('Gender')
plt.ylabel('Count')
plt.title('Distribution of Genders')
plt.xticks(rotation=0)
plt.show()
```




```
#Countplot: This plot is useful for visualizing the count of
categorical data, such as the distribution of genders.
plt.figure(figsize=(6, 4))
sns.countplot(x='SEX', data=df)
plt.xlabel('Gender')
plt.ylabel('Count')
plt.title('Countplot of Genders')
plt.show()
```




```
#Box Plot (Box-and-Whisker Plot): This plot shows the distribution of
a dataset, including the median, quartiles, and potential outliers.
plt.figure(figsize=(8, 6))
sns.boxplot(x='AGE', y='Value', data=df)
plt.xlabel('Age Group')
plt.ylabel('Value')
plt.title('Box Plot of Age Distribution')
plt.xticks(rotation=45)
plt.show()
```


Box Plot of Age Distribution
1e8
3.0
2.5
2.0
1.5
1.0
0.5
0.0
Value
Age Group