TASK -1:

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](https://www.google.com/url?q=https%3A%2F%2Fdata.worldbank.org%2Findicator%2FSP.**POP**.TOTL)

Code:

import pandas as pd

import matplotlib.pyplot as plt

import seaborn as sns

df=pd.read\_csv('iris.csv')

print(df.head())

plt.figure(figsize=(8, 6))

sns.histplot(data=iris\_df,x='Petal\_Width', kde=True, bins=20)

plt.title('Distribution of Petal Width in Iris Dataset', fontsize=16)

plt.xlabel('Petal Width (cm)')

plt.ylabel('Frequency')

plt.grid(True)

plt.show()

plt.figure(figsize=(8, 6))

sns.countplot(data=iris\_df, x='Class',hue='Class',legend=False,palette='cool')

plt.title('Distribution of Iris Species', fontsize=16)

plt.xlabel('Species')

plt.ylabel('Count')

plt.grid(axis='y')

plt.show()

plt.figure(figsize=(8, 6))

plt.hist(iris\_df['Sepal\_Length'], bins=10, color='lightgreen', edgecolor='black') # You can adjust the number of bins

plt.title('Distribution of Sepal Length in Iris Dataset')

plt.xlabel('Sepal Length (cm)')

plt.ylabel('Frequency')

plt.grid(axis='y', alpha=0.75) # Add grid for better readability

plt.show()

# Histogram for Petal Width

plt.figure(figsize=(8, 6))

plt.hist(iris\_df['Petal\_Width'], bins=10, color='pink', edgecolor='black')

plt.title('Distribution of Petal Width in Iris Dataset')

plt.xlabel('Petal Width (cm)')

plt.ylabel('Frequency')

plt.grid(axis='y', alpha=0.75)

plt.show()

# Histogram for Petal Length

plt.figure(figsize=(8, 6))

plt.hist(iris\_df['Petal\_Length'], bins=10, color='gold', edgecolor='black')

plt.title('Distribution of Petal Length in Iris Dataset')

plt.xlabel('Petal Length (cm)')

plt.ylabel('Frequency')

plt.grid(axis='y', alpha=0.75)

plt.show()

# Histogram for Sepal Width

plt.figure(figsize=(8, 6))

plt.hist(iris\_df['Sepal\_Width'], bins=10, color='skyblue', edgecolor='black')

plt.title('Distribution of Sepal Width in Iris Dataset')

plt.xlabel('Sepal Width (cm)')

plt.ylabel('Frequency')

plt.grid(axis='y', alpha=0.75)

plt.show()

plt.figure(figsize=(8, 6))

sns.countplot(x='Class', data=iris\_df,hue='Class',legend=False,palette='cool')

plt.title('Distribution of Iris Species', fontsize=16)

plt.xlabel('Species')

plt.ylabel('Count')

plt.grid(axis='y')

plt.show()

average\_sepal\_length = iris\_df.groupby('Class')['Sepal\_Length'].mean().reset\_index()

plt.figure(figsize=(8, 6))

sns.barplot(x='Class', y='Sepal\_Length', data=average\_sepal\_length,hue='Class',legend=False,palette='cool')

plt.title('Average Sepal Length per Iris Species', fontsize=16)

plt.xlabel('Species')

plt.ylabel('Average Sepal Length (cm)')

plt.grid(axis='y')

plt.show()

plt.figure(figsize=(8, 6))

sns.kdeplot(data=iris\_df, x='Petal\_Width', fill=True, color='green')

plt.title('KDE Plot of Petal Width in Iris Dataset', fontsize=14)

plt.xlabel('Petal Width (cm)')

plt.ylabel('Density')

plt.grid(True)

plt.show()

# KDE plot of Sepal Length

plt.figure(figsize=(8, 6))

sns.kdeplot(data=iris\_df, x='Sepal\_Length', fill=True, color='blue')

plt.title('KDE Plot of Sepal Length in Iris Dataset', fontsize=14)

plt.xlabel('Sepal Length (cm)')

plt.ylabel('Density')

plt.grid(True)

plt.show()

# KDE plot of Sepal Width

plt.figure(figsize=(8, 6))

sns.kdeplot(data=iris\_df, x='Sepal\_Width', fill=True, color='purple')

plt.title('KDE Plot of Sepal Width in Iris Dataset', fontsize=14)

plt.xlabel('Sepal Width (cm)')

plt.ylabel('Density')

plt.grid(True)

plt.show()

# KDE plot of Petal Length

plt.figure(figsize=(8, 6))

sns.kdeplot(data=iris\_df, x='Petal\_Length', fill=True, color='red')

plt.title('KDE Plot of Petal Length in Iris Dataset', fontsize=14)

plt.xlabel('Petal Length (cm)')

plt.ylabel('Density')

plt.grid(True)

plt.show()

plt.figure(figsize=(8, 6))

# Use sns.histplot and set kde=True to overlay the KDE curve

sns.histplot(data=iris\_df, x='Petal\_Width', kde=True, bins=20)

plt.title('Distribution of Petal Width with KDE in Iris Dataset', fontsize=16)

plt.xlabel('Petal Width (cm)')

plt.ylabel('Frequency')

plt.grid(True)

plt.show()

# Histogram with KDE for Sepal Length

plt.figure(figsize=(8, 6))

sns.histplot(data=iris\_df, x='Sepal\_Length', kde=True, bins=20, color='lightgreen')

plt.title('Distribution of Sepal Length with KDE in Iris Dataset', fontsize=16)

plt.xlabel('Sepal Length (cm)')

plt.ylabel('Frequency')

plt.grid(True)

plt.show()

# Histogram with KDE for Sepal Width

plt.figure(figsize=(8, 6))

sns.histplot(data=iris\_df, x='Sepal\_Width', kde=True, bins=20, color='salmon')

plt.title('Distribution of Sepal Width with KDE in Iris Dataset', fontsize=16)

plt.xlabel('Sepal Width (cm)')

plt.ylabel('Frequency')

plt.grid(True)

plt.show()

# Histogram with KDE for Petal Length

plt.figure(figsize=(8, 6))

sns.histplot(data=iris\_df, x='Petal\_Length', kde=True, bins=20, color='gold')

plt.title('Distribution of Petal Length with KDE in Iris Dataset', fontsize=16)

plt.xlabel('Petal Length (cm)')

plt.ylabel('Frequency')

plt.grid(True)

plt.show()