**Experiment -16**

**Aim:**Visualize the datasets using matplotlib in python/R.(Histogram, Box plot, Bar chart, Pie chart etc.,)

**Program:**

import pandas as pd

import matplotlib.pyplot as plt

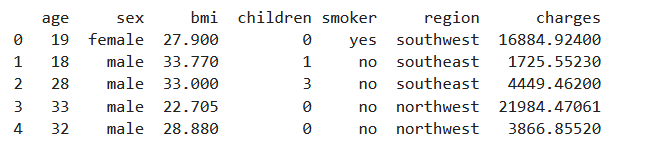
import numpy as np

# Load the dataset

df = pd.read\_csv("insurance.csv")

print(df.head())

Dataset:



#**histogram visualization**

plt.figure(figsize=(6, 4))

plt.hist(df['age'], bins='auto', color='skyblue', edgecolor='black')

plt.title('Age Distribution')

plt.xlabel('Age')

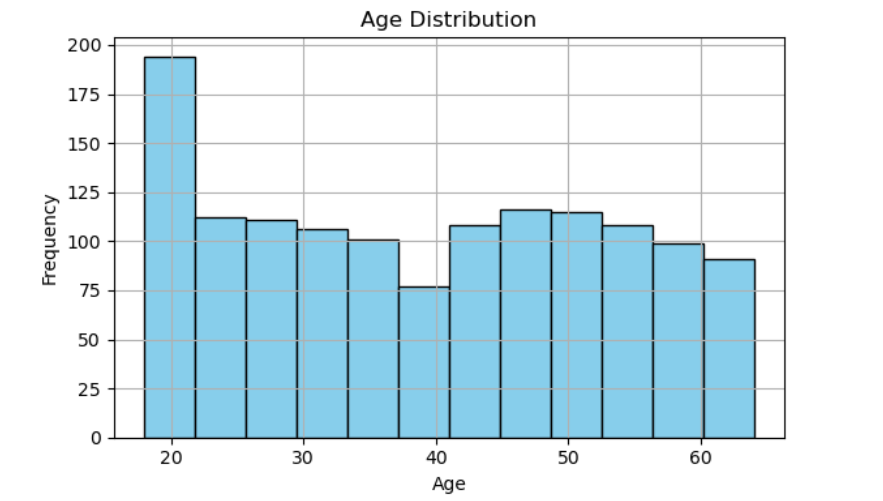
plt.ylabel('Frequency')

plt.grid(True)

plt.tight\_layout()

plt.show()

**Output**:



**# 2. Bar Chart: Count of Smokers**

smoker\_counts = df['smoker'].value\_counts()

plt.figure(figsize=(6, 4))

smoker\_counts.plot(kind='bar', color='lightgreen', edgecolor='black')

plt.title('Smoker vs Non-Smoker Count')

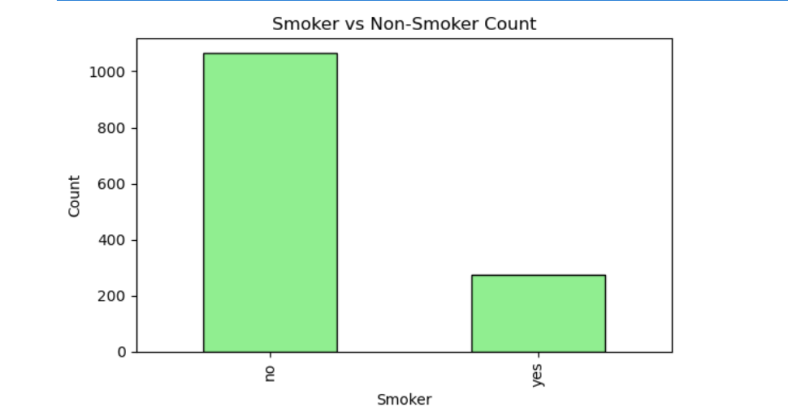
plt.xlabel('Smoker')

plt.ylabel('Count')

plt.tight\_layout()

plt.show()

Output:



**# 3. Pie Chart: Distribution by Region**

region\_counts = df['region'].value\_counts()

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

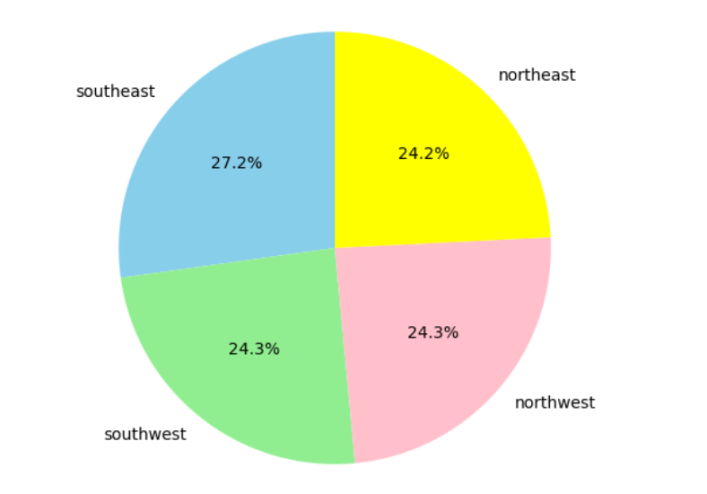
region\_counts.plot.pie(autopct="%1.1f%%",startangle=90,colors=['skyblue','lightgreen','pink','Yellow'])

plt.title('Distribution by Region')

plt.ylabel("")

plt.show()

**Output:**



**# 4. Boxplot: Charges by Smoker Status**

plt.figure(figsize=(6, 4))

df.boxplot(column='charges', by='smoker')

plt.title('Charges Distribution by Smoker')

plt.suptitle("") # remove default title

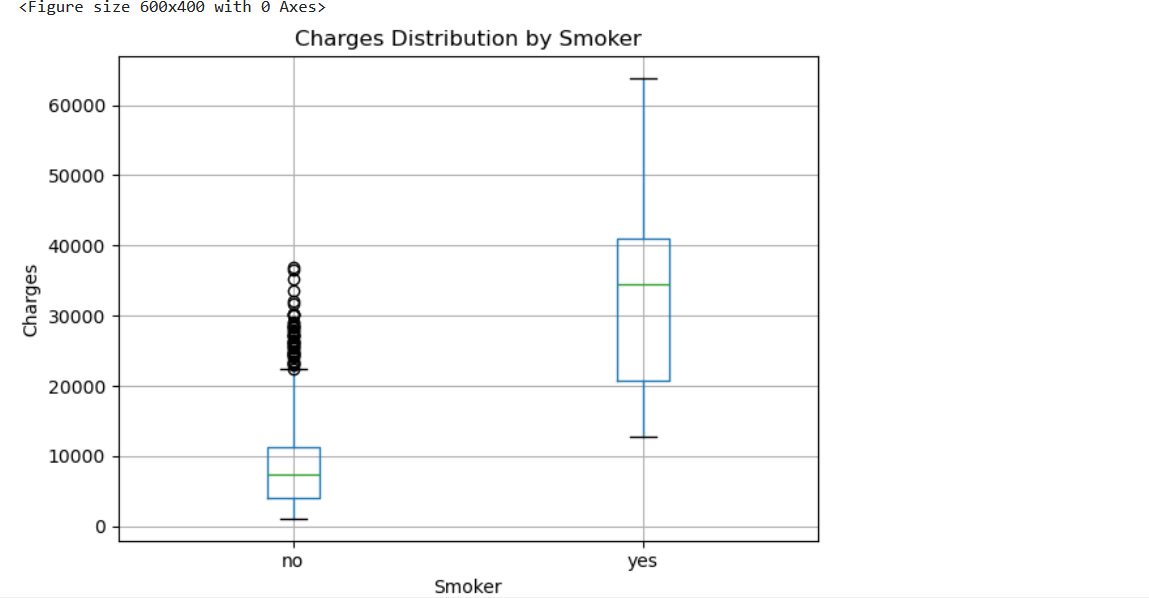
plt.xlabel('Smoker')

plt.ylabel('Charges')

plt.tight\_layout()

plt.show()

**Output:**



# Extract the column (example: charges)

charges = df['charges']

**5.# Calculate quartiles and mean**

q1 = np.percentile(charges, 25)

q2 = np.median(charges)

q3 = np.percentile(charges, 75)

mean = np.mean(charges)

# Print results

print("Boxplot Statistics for Charges:")

print("Quartile 1 (Q1):", q1)

print("Median (Q2):", q2)

print("Quartile 3 (Q3):", q3)

print("Mean:", mean)

**Output:**

