```
#step 1: importing the necessary libraries
import numpy as np
import pandas as pd
import matplotlib.pyplot as plt
import seaborn as sn
data=pd.read_csv('insurance.csv')
data

#step 2: checking the shape and the DataTypes of the DataSet
data.shape

#step 3: Dealing with missing values
data.isna().sum()

#step 4: Exploring the Relationship between the Feature and Target Columns
fig,axs=plt.subplots(1,3,sharey=True)
data.plot(kind='scatter',x='age',y='charges',ax=axs[0],figsize=(16,8))
data.plot(kind='scatter',x='children',y='charges',ax=axs[1])
data.plot(kind='scatter',x='bmi',y='charges',ax=axs[2])

import seaborn as sns
import matplotlib.pyplot as plt
sns.countplot(data=data,x='region',palette='pastel')
plt.title('Regions')
plt.show()

import seaborn as sns
import matplotlib.pyplot as plt
sns.countplot(data=data,x='sex',palette='pastel')
plt.title('sex')

import seaborn as sns
import matplotlib.pyplot as plt
sns.countplot(data=data,x='smoker',palette='pastel')
plt.title('smoker')
plt.show()

#Step:5 Plotting of Feature vs Feature Plots
import seaborn as sns
import matplotlib.pyplot as plt
plt.figure(figsize=(10,6))
sns.boxplot(x='children',y='age',data=data,palette='pastel')
plt.show()

#draw barplot age vs charges
import pandas as pd
import matplotlib.pyplot as plt

# Read the CSV file
df = pd.read_csv('insurance.csv')

# Extract the necessary columns
age = df['age']
charges = df['charges']
```

```
# Create the bar chart
plt.bar(age, charges, color='skyblue')
plt.xlabel('age')
plt.ylabel('charges')
plt.title('Bar Chart from CSV Data')
plt.grid(True)
plt.show()

#draw barplot region vs charges
import pandas as pd
import matplotlib.pyplot as plt

# Read the CSV file
df = pd.read_csv('insurance.csv')

# Extract the necessary columns
region = df['region']
charges = df['charges']

# Create the bar chart
plt.bar(region, charges, color='skyblue')
plt.xlabel('region')
plt.ylabel('charges')
plt.title('Bar Chart from CSV Data')
plt.grid(True)
plt.show()

#draw barplot sex vs charges
import pandas as pd
import matplotlib.pyplot as plt

# Read the CSV file
df = pd.read_csv('insurance.csv')

# Extract the necessary columns
sex = df['sex']
charges = df['charges']

# Create the bar chart
plt.bar(sex, charges, color='skyblue')
plt.xlabel('sex')
plt.ylabel('charges')
plt.title('Bar Chart from CSV Data')
plt.grid(True)
plt.show()

#draw barplot bmi vs charges
import pandas as pd
import matplotlib.pyplot as plt

# Read the CSV file
df = pd.read_csv('insurance.csv')

# Extract the necessary columns
bmi = df['bmi']
```

```
charges = df['charges']

# Create the bar chart
plt.bar(bmi, charges, color='skyblue')
plt.xlabel('bmi')
plt.ylabel('charges')
plt.title('Bar Chart from CSV Data')
plt.grid(True)
plt.show()

#draw barplot smoker vs charges
import pandas as pd
import matplotlib.pyplot as plt

# Read the CSV file
df = pd.read_csv('insurance.csv')

# Extract the necessary columns
smoker = df['smoker']
charges = df['charges']

# Create the bar chart
plt.bar(bmi, charges, color='skyblue')
plt.xlabel('smoker')
plt.ylabel('charges')
plt.title('Bar Chart from CSV Data')
plt.grid(True)
plt.show()

# step6: draw scatterplot age vs charges
import pandas as pd
import matplotlib.pyplot as plt

# Read the CSV file
df = pd.read_csv('insurance.csv')

# Filter data for smokers and non-smokers
smokers = df[df['smoker'] == 'yes']
non_smokers = df[df['smoker'] == 'no']

# Create the scatter plot
plt.scatter(smokers['age'], smokers['charges'], color='red', label='Smokers', alpha=0.5)
plt.scatter(non_smokers['age'], non_smokers['charges'], color='blue', label='Non-Smokers', alpha=0.5)

plt.xlabel('Age')
plt.ylabel('Charges')
plt.title('Scatter Plot of Age vs Charges (Considering Smoker Status)')
plt.legend()
plt.show()

#draw scatterplot region vs charges
import pandas as pd
import matplotlib.pyplot as plt

# Read the CSV file
```

```
df = pd.read_csv('insurance.csv')

# Filter data for smokers and non-smokers
smokers = df[df['smoker'] == 'yes']
non_smokers = df[df['smoker'] == 'no']

# Create the scatter plot
plt.scatter(smokers['region'], smokers['charges'], color='red', label='Smokers', alpha=0.5)
plt.scatter(non_smokers['region'], non_smokers['charges'], color='blue', label='Non-Smokers', alpha=0.5)

plt.xlabel('Region')
plt.ylabel('Charges')
plt.title('Scatter Plot of Age vs Charges (Considering Smoker Status)')
plt.legend()
plt.show()

#draw scatterplot sex vs charges
import pandas as pd
import matplotlib.pyplot as plt

# Read the CSV file
df = pd.read_csv('insurance.csv')

# Filter data for smokers and non-smokers
smokers = df[df['smoker'] == 'yes']
non_smokers = df[df['smoker'] == 'no']

# Create the scatter plot
plt.scatter(smokers['sex'], smokers['charges'], color='red', label='Smokers', alpha=0.5)
plt.scatter(non_smokers['sex'], non_smokers['charges'], color='blue', label='Non-Smokers', alpha=0.5)

plt.xlabel('Sex')
plt.ylabel('Charges')
plt.title('Scatter Plot of Age vs Charges (Considering Smoker Status)')
plt.legend()
plt.show()

#draw scatterplot bmi vs charges
import pandas as pd
import matplotlib.pyplot as plt

# Read the CSV file
df = pd.read_csv('insurance.csv')

# Filter data for smokers and non-smokers
smokers = df[df['smoker'] == 'yes']
non_smokers = df[df['smoker'] == 'no']

# Create the scatter plot
plt.scatter(smokers['bmi'], smokers['charges'], color='red', label='Smokers', alpha=0.5)
plt.scatter(non_smokers['bmi'], non_smokers['charges'], color='blue', label='Non-Smokers', alpha=0.5)

plt.xlabel('BMI')
plt.ylabel('Charges')
plt.title('Scatter Plot of Age vs Charges (Considering Smoker Status)')
```

```
plt.legend()
plt.show()
```