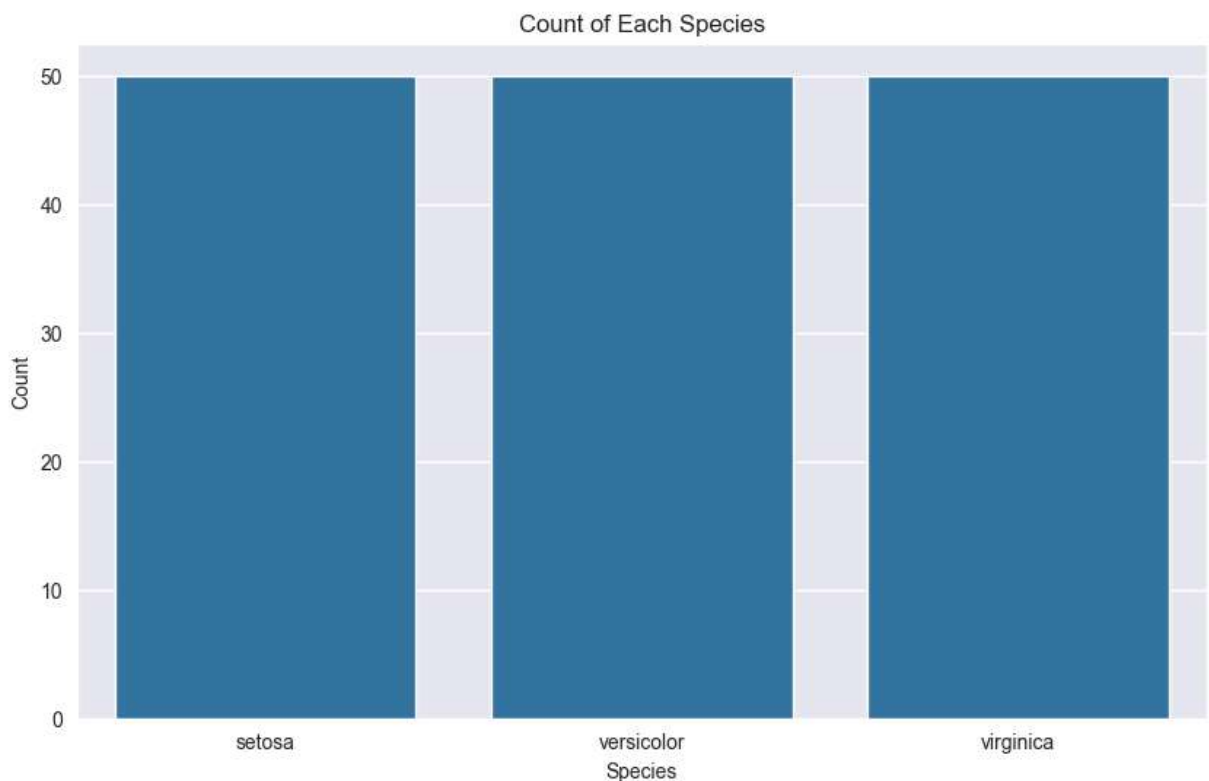


Exercise - 4

```
In [1]: import pandas as pd
import seaborn as sns
import matplotlib.pyplot as plt
from sklearn.datasets import load_iris
```

```
In [2]: # Load the Iris dataset
iris = load_iris()
iris_df = pd.DataFrame(data=iris.data, columns=iris.feature_names)
iris_df['species'] = pd.Categorical.from_codes(iris.target, iris.target_names)
```

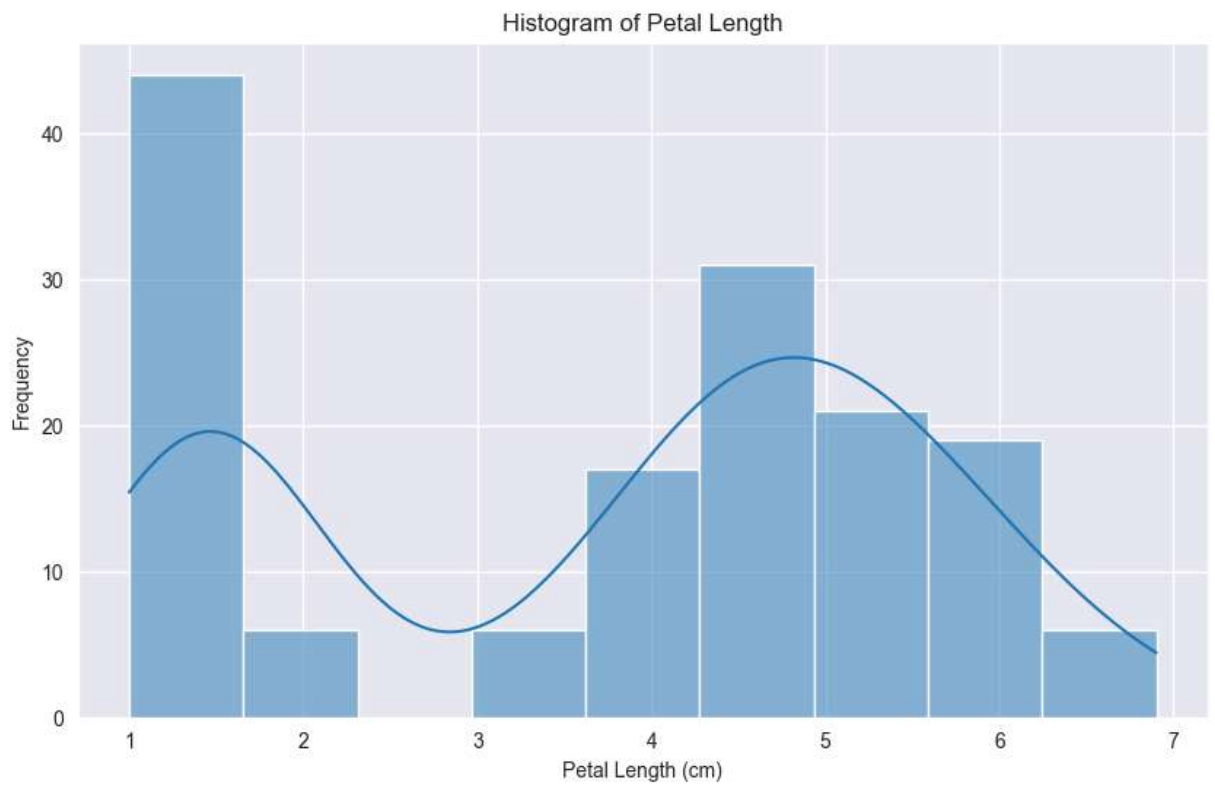
```
In [3]: # Bar chart for the count of each species
plt.figure(figsize=(10, 6))
sns.countplot(x='species', data=iris_df)
plt.title('Count of Each Species')
plt.xlabel('Species')
plt.ylabel('Count')
plt.show()
```



```
In [4]: # Scatter plot for sepal length vs. sepal width
plt.figure(figsize=(10, 6))
sns.scatterplot(x=iris_df['sepal length (cm)'], y=iris_df['sepal width (cm)'], hue=
plt.title('Sepal Length vs. Sepal Width')
plt.xlabel('Sepal Length (cm)')
plt.ylabel('Sepal Width (cm)')
plt.show()
```



```
In [5]: # Histogram for petal length
plt.figure(figsize=(10, 6))
sns.histplot(iris_df['petal length (cm)'], kde=True)
plt.title('Histogram of Petal Length')
plt.xlabel('Petal Length (cm)')
plt.ylabel('Frequency')
plt.show()
```

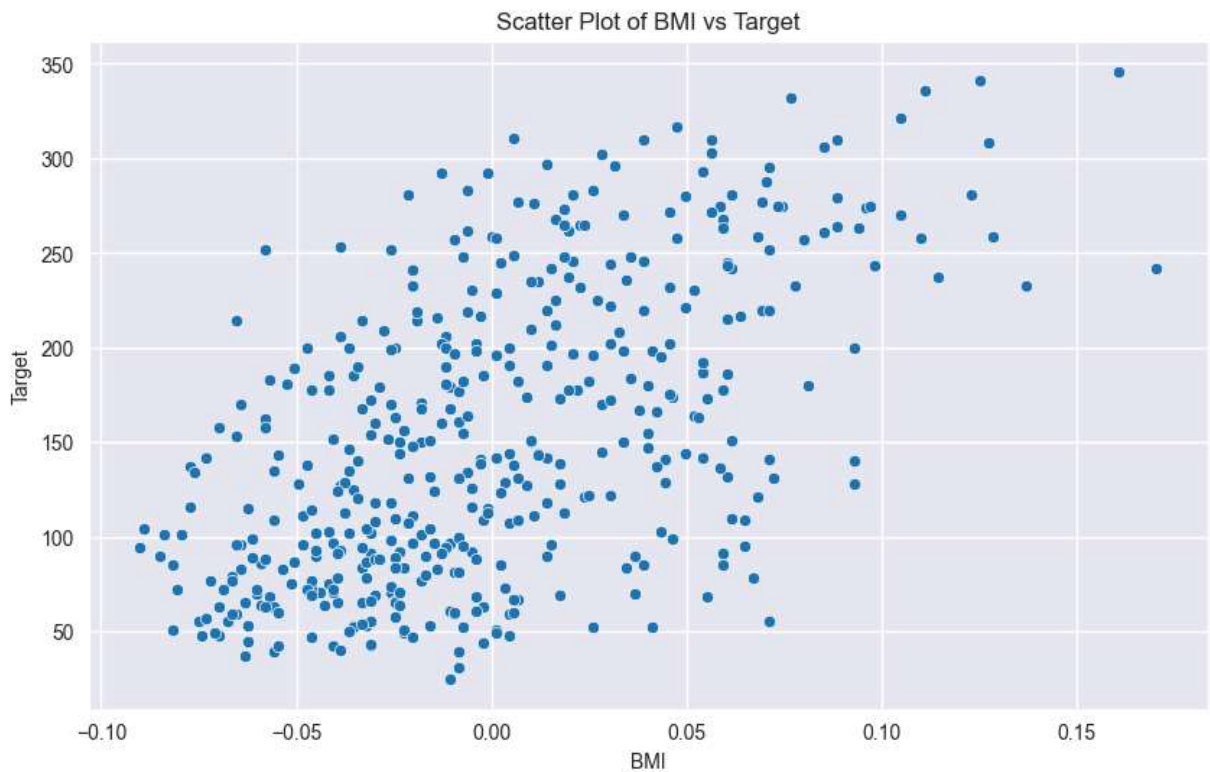


Own Data Set - Diabetes

```
In [6]: import matplotlib.pyplot as plt
import seaborn as sns
import pandas as pd
from sklearn.datasets import load_diabetes

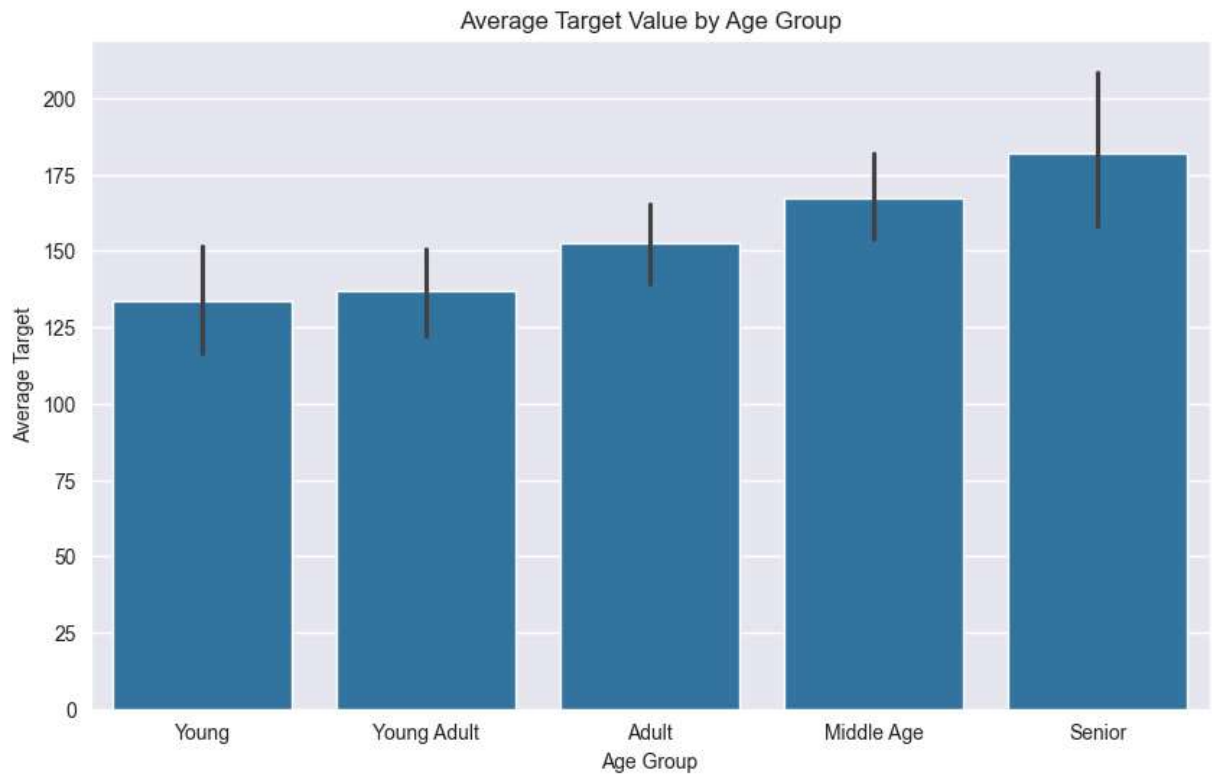
# Load the diabetes dataset
diabetes = load_diabetes()
data = pd.DataFrame(diabetes.data, columns=diabetes.feature_names)
data['target'] = diabetes.target
```

```
In [7]: plt.figure(figsize=(10, 6))
sns.scatterplot(x='bmi', y='target', data=data)
plt.title('Scatter Plot of BMI vs Target')
plt.xlabel('BMI')
plt.ylabel('Target')
plt.show()
```

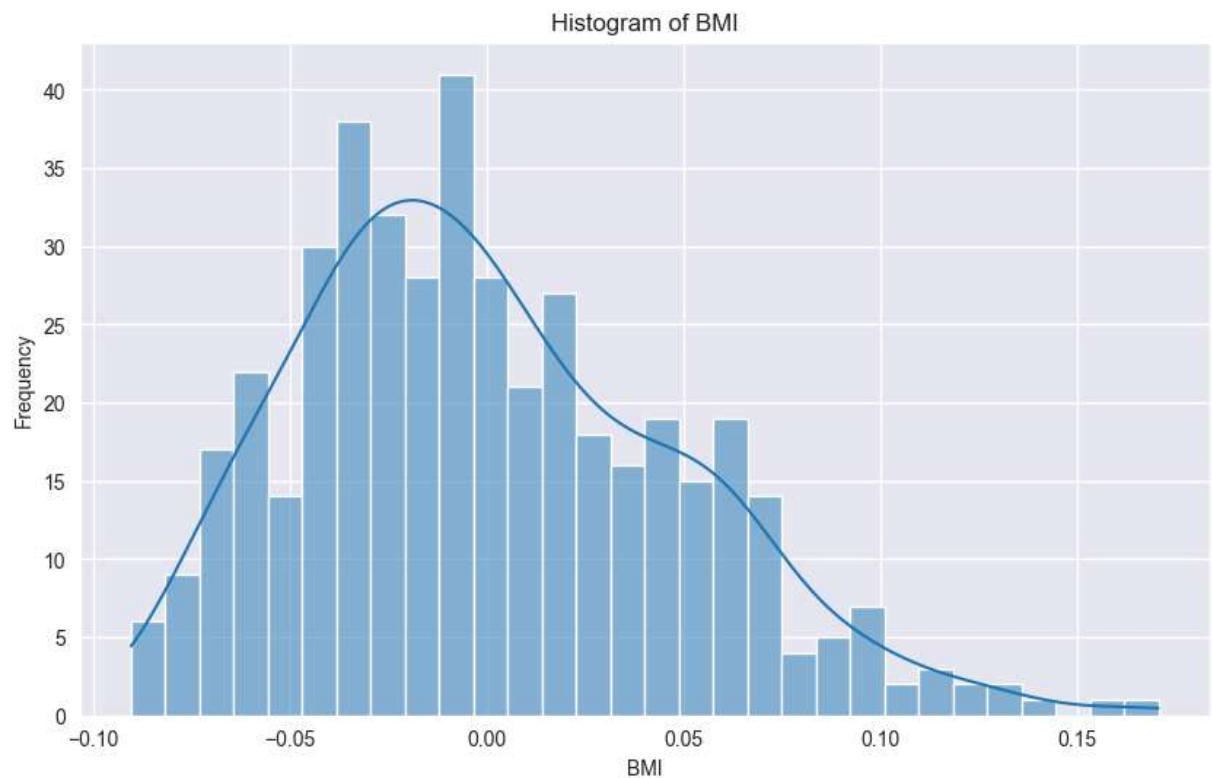


```
In [8]: data['age_group'] = pd.cut(data['age'], bins=5, labels=['Young', 'Young Adult', 'Ad

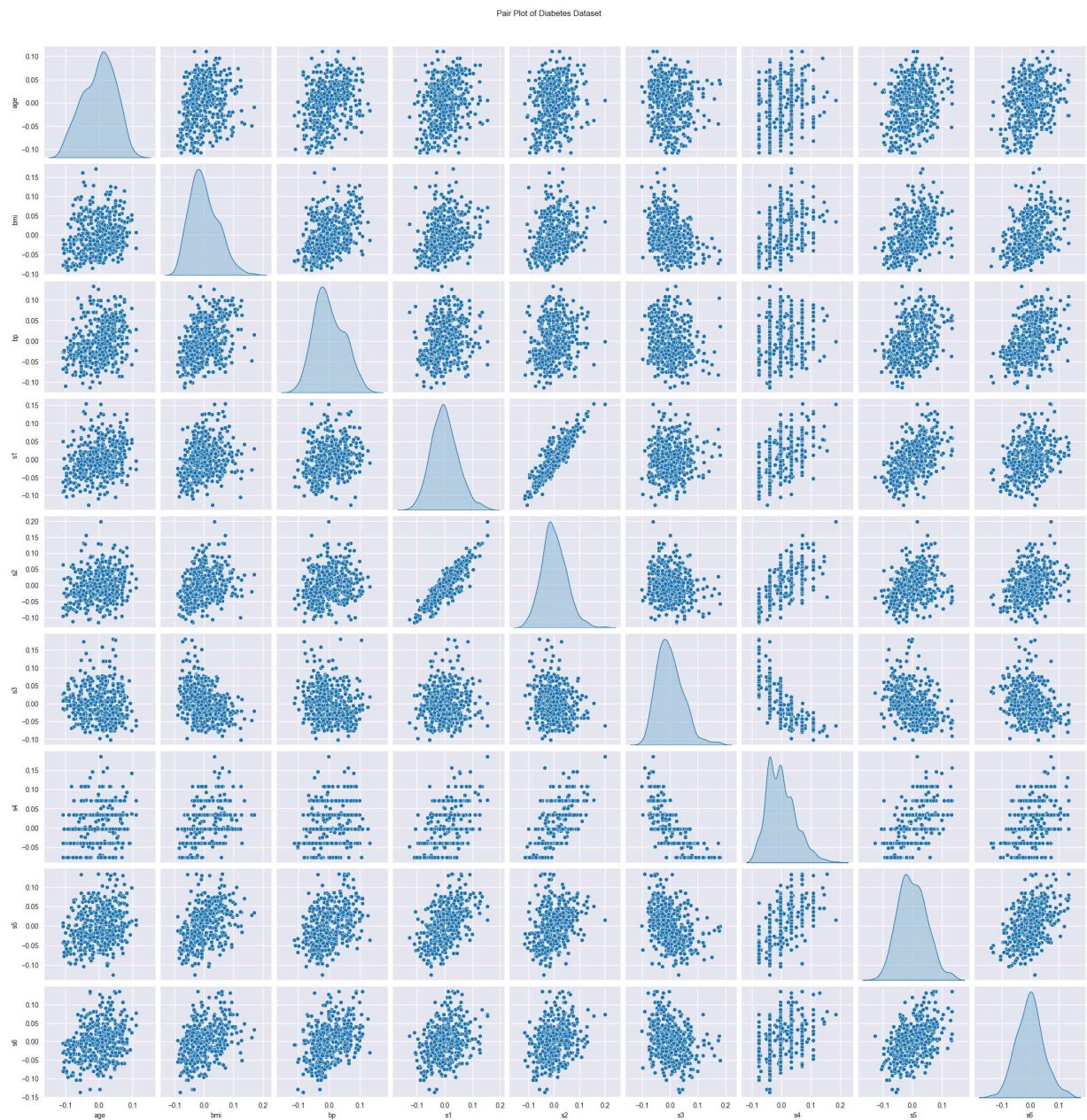
plt.figure(figsize=(10, 6))
sns.barplot(x='age_group', y='target', data=data, estimator=lambda x: sum(x) / len(
plt.title('Average Target Value by Age Group')
plt.xlabel('Age Group')
plt.ylabel('Average Target')
plt.show()
```



```
In [9]: plt.figure(figsize=(10, 6))
sns.histplot(data['bmi'], bins=30, kde=True)
plt.title('Histogram of BMI')
plt.xlabel('BMI')
plt.ylabel('Frequency')
plt.show()
```



```
In [10]: sns.pairplot(data, vars=['age', 'bmi', 'bp', 's1', 's2', 's3', 's4', 's5', 's6'], d
plt.suptitle('Pair Plot of Diabetes Dataset', y=1.02)
plt.show()
```



```
In [12]: data = {
    'Age Group': ['Young', 'Middle Age', 'Old'],
    'Value': [10, 20, 30]
}
df = pd.DataFrame(data)

# Convert non-numeric data to NaN
df['Value'] = pd.to_numeric(df['Value'], errors='coerce')

# Drop rows with NaN values
df = df.dropna()

# Plot heatmap
```

```
sns.heatmap(df[['Value']])  
plt.show()
```

