T-SNE-iris.ipynb

File Edit View Insert Runtime Tools Help Saved: 2024-05-20 20:00:00

Code Text Copy to Drive

Question 1

Consider a subset of attributes for your project which will apply the whole given visualization which is any pattern of your choice (discussed focus on the following visualizations) Eg.

- Pie Chart/Donut Chart
- Stacked Bar Chart
- Grouped Bar Chart
- Line Chart
- Scatterplot
- Bubble Chart
- Stacked Area Chart
- Grouped Area Chart
- Histogram
- Boxplot

Kaggle URL: https://www.kaggle.com/datasets/uciml/iris Kindly Share the uploaded URL for sample data (at least as well as you)

Kindly share the code for the above visualization (Kindly share the code for the above visualization in the form of a function that takes the input data and the parameters to the function for the visualization) Input/Output/Parameters/Assumptions/Limitations

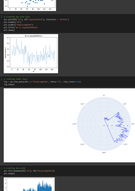
```
!wget -q -O iris.zip "https://www.kaggle.com/datasets/uciml/iris?dataset_version_number=1&select=Iris.csv" -c -x
!unzip -o iris.zip
!ls -lrt
!head -5 Iris.csv
```

```
Archive:  iris.zip
  inflating: Iris.csv
total 4
-rw-r--r-- 1 root root 3716 Jan  1  1980 Iris.csv
Id,SepalLengthCm,SepalWidth,PetalLengthCm,PetalWidth,Species
1,5.1,3.5,1.4,0.2,Iris-setosa
2,4.9,3.0,1.4,0.2,Iris-setosa
3,4.7,3.2,1.3,0.2,Iris-setosa
4,4.6,3.1,1.5,0.2,Iris-setosa
```

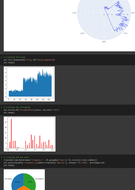
```
# Load the data
df = pd.read_csv('Iris.csv')
df.drop(['Id'], inplace=True)

# Define the function
def plot_iris(df, x='SepalLengthCm', y='PetalLengths', hue='Species', size=None, s=50, palette='Set1', alpha=0.7, figsize=(10, 6), title='Iris Scatter Plot'):
    plt.figure(figsize=figsize)
    sns.scatterplot(data=df, x=x, y=y, hue=hue, size=size, s=s, palette=palette, alpha=alpha)
    plt.title(title)
    plt.show()

# Example usage
plot_iris(df, x='SepalLengthCm', y='PetalLengths', hue='Species', size='PetalWidth', s=100, palette='Dark2', alpha=0.8, figsize=(12, 8), title='Iris Scatter Plot')
```

```
# Load the data
df = pd.read_csv('Iris.csv')
df.drop(['Id'], inplace=True)

# Define the function
def plot_iris(df, x='SepalLengthCm', y='PetalLengths', hue='Species', size=None, s=50, palette='Set1', alpha=0.7, figsize=(10, 6), title='Iris Scatter Plot'):
    plt.figure(figsize=figsize)
    sns.scatterplot(data=df, x=x, y=y, hue=hue, size=size, s=s, palette=palette, alpha=alpha, legend=False)
    plt.title(title)
    plt.show()

# Example usage
plot_iris(df, x='SepalLengthCm', y='PetalLengths', hue='Species', size='PetalWidth', s=100, palette='Dark2', alpha=0.8, figsize=(12, 8), title='Iris Scatter Plot')
```

```
# Load the data
df = pd.read_csv('Iris.csv')
df.drop(['Id'], inplace=True)

# Define the function
def plot_iris(df, x='SepalLengthCm', y='PetalLengths', hue='Species', size=None, s=50, palette='Set1', alpha=0.7, figsize=(10, 6), title='Iris Scatter Plot'):
    plt.figure(figsize=figsize)
    sns.scatterplot(data=df, x=x, y=y, hue=hue, size=size, s=s, palette=palette, alpha=alpha, legend=False)
    plt.title(title)
    plt.show()

# Example usage
plot_iris(df, x='SepalLengthCm', y='PetalLengths', hue='Species', size='PetalWidth', s=100, palette='Dark2', alpha=0.8, figsize=(12, 8), title='Iris Scatter Plot')
```




```
# Load the data
df = pd.read_csv('Iris.csv')
df.drop(['Id'], inplace=True)

# Define the function
def plot_iris(df, x='SepalLengthCm', y='PetalLengths', hue='Species', size=None, s=50, palette='Set1', alpha=0.7, figsize=(10, 6), title='Iris Scatter Plot'):
    plt.figure(figsize=figsize)
    sns.scatterplot(data=df, x=x, y=y, hue=hue, size=size, s=s, palette=palette, alpha=alpha, legend=False)
    plt.title(title)
    plt.show()

# Example usage
plot_iris(df, x='SepalLengthCm', y='PetalLengths', hue='Species', size='PetalWidth', s=100, palette='Dark2', alpha=0.8, figsize=(12, 8), title='Iris Scatter Plot')
```




```
# Load the data
df = pd.read_csv('Iris.csv')
df.drop(['Id'], inplace=True)

# Define the function
def plot_iris(df, x='SepalLengthCm', y='PetalLengths', hue='Species', size=None, s=50, palette='Set1', alpha=0.7, figsize=(10, 6), title='Iris Scatter Plot'):
    plt.figure(figsize=figsize)
    sns.scatterplot(data=df, x=x, y=y, hue=hue, size=size, s=s, palette=palette, alpha=alpha, legend=False)
    plt.title(title)
    plt.show()

# Example usage
plot_iris(df, x='SepalLengthCm', y='PetalLengths', hue='Species', size='PetalWidth', s=100, palette='Dark2', alpha=0.8, figsize=(12, 8), title='Iris Scatter Plot')
```

```
# Load the data
df = pd.read_csv('Iris.csv')
df.drop(['Id'], inplace=True)

# Define the function
def plot_iris(df, x='SepalLengthCm', y='PetalLengths', hue='Species', size=None, s=50, palette='Set1', alpha=0.7, figsize=(10, 6), title='Iris Scatter Plot'):
    plt.figure(figsize=figsize)
    sns.scatterplot(data=df, x=x, y=y, hue=hue, size=size, s=s, palette=palette, alpha=alpha, legend=False)
    plt.title(title)
    plt.show()

# Example usage
plot_iris(df, x='SepalLengthCm', y='PetalLengths', hue='Species', size='PetalWidth', s=100, palette='Dark2', alpha=0.8, figsize=(12, 8), title='Iris Scatter Plot')
```

Question 2

Identify the type of correlation using the Kendall correlation for Iris dataset of your choice and find the Dominance factor from the table.

```
# Load the data
df = pd.read_csv('Iris.csv')
df.drop(['Id'], inplace=True)

# Calculate Kendall correlation matrix
corr = df.corr(method='kendall')

# Print the correlation matrix
print(corr)
```

```
               SepalLengthCm  SepalWidth  PetalLengths  PetalWidth
SepalLengthCm       1.000000   -0.072531      0.718264    0.656037
SepalWidth         -0.072531    1.000000     -0.176183   -0.153776
PetalLengths        0.718264   -0.176183      1.000000    0.801454
PetalWidth          0.656037   -0.153776      0.801454    1.000000
```

```
# Load the data
df = pd.read_csv('Iris.csv')
df.drop(['Id'], inplace=True)

# Calculate Kendall correlation matrix
corr = df.corr(method='kendall')

# Print the correlation matrix
print(corr)

# Find the dominance factor
dominance = corr.abs.sum(axis=1)
print(dominance)
```

```
               SepalLengthCm  SepalWidth  PetalLengths  PetalWidth
SepalLengthCm       1.000000   -0.072531      0.718264    0.656037
SepalWidth         -0.072531    1.000000     -0.176183   -0.153776
PetalLengths        0.718264   -0.176183      1.000000    0.801454
PetalWidth          0.656037   -0.153776      0.801454    1.000000
SepalLengthCm    2.446832
SepalWidth       1.402492
PetalLengths     2.696855
PetalWidth       2.411317
dtype: float64
```

Question 3

Identify the range of correlation using the Pearson correlation for Iris dataset of your choice and find the Dominance factor from the table.

```
# Load the data
df = pd.read_csv('Iris.csv')
df.drop(['Id'], inplace=True)

# Calculate Pearson correlation matrix with the correlation
corr = df.corr(method='pearson')

# Print the correlation matrix
print(corr)
```

```
               SepalLengthCm  SepalWidth  PetalLengths  PetalWidth
SepalLengthCm       1.000000   -0.109369      0.871754    0.817954
SepalWidth         -0.109369    1.000000     -0.420516   -0.356544
PetalLengths        0.871754   -0.420516      1.000000    0.962757
PetalWidth          0.817954   -0.356544      0.962757    1.000000
```

```
# Load the data
df = pd.read_csv('Iris.csv')
df.drop(['Id'], inplace=True)

# Calculate Pearson correlation matrix
corr = df.corr(method='pearson')

# Print the correlation matrix
print(corr)

# Find the dominance factor
dominance = corr.abs.sum(axis=1)
print(dominance)
```

```
               SepalLengthCm  SepalWidth  PetalLengths  PetalWidth
SepalLengthCm       1.000000   -0.109369      0.871754    0.817954
SepalWidth         -0.109369    1.000000     -0.420516   -0.356544
PetalLengths        0.871754   -0.420516      1.000000    0.962757
PetalWidth          0.817954   -0.356544      0.962757    1.000000
SepalLengthCm    2.808677
SepalWidth       1.286429
PetalLengths     3.255027
PetalWidth       3.137255
dtype: float64
```

Question 4

Find the correlation between the pairs (excluding the Species) for the Iris dataset given below.

```
# Load the data
df = pd.read_csv('Iris.csv')
df.drop(['Id'], inplace=True)

# Calculate correlation matrix
corr = df.corr(method='pearson')

# Print the correlation matrix
print(corr)
```

```
               SepalLengthCm  SepalWidth  PetalLengths  PetalWidth
SepalLengthCm       1.000000   -0.109369      0.871754    0.817954
SepalWidth         -0.109369    1.000000     -0.420516   -0.356544
PetalLengths        0.871754   -0.420516      1.000000    0.962757
PetalWidth          0.817954   -0.356544      0.962757    1.000000
```

```
# Load the data
df = pd.read_csv('Iris.csv')
df.drop(['Id'], inplace=True)

# Calculate correlation matrix
corr = df.corr(method='pearson')

# Print the correlation matrix
print(corr)

# Find the dominance factor
dominance = corr.abs.sum(axis=1)
print(dominance)
```

```
               SepalLengthCm  SepalWidth  PetalLengths  PetalWidth
SepalLengthCm       1.000000   -0.109369      0.871754    0.817954
SepalWidth         -0.109369    1.000000     -0.420516   -0.356544
PetalLengths        0.871754   -0.420516      1.000000    0.962757
PetalWidth          0.817954   -0.356544      0.962757    1.000000
SepalLengthCm    2.808677
SepalWidth       1.286429
PetalLengths     3.255027
PetalWidth       3.137255
dtype: float64
```

Question 5

Find the correlation between pairs (excluding the Species) for the Iris dataset given below.

```
# Load the data
df = pd.read_csv('Iris.csv')
df.drop(['Id'], inplace=True)

# Calculate correlation matrix
corr = df.corr(method='pearson')

# Print the correlation matrix
print(corr)

# Find the dominance factor
dominance = corr.abs.sum(axis=1)
print(dominance)
```

```
               SepalLengthCm  SepalWidth  PetalLengths  PetalWidth
SepalLengthCm       1.000000   -0.109369      0.871754    0.817954
SepalWidth         -0.109369    1.000000     -0.420516   -0.356544
PetalLengths        0.871754   -0.420516      1.000000    0.962757
PetalWidth          0.817954   -0.356544      0.962757    1.000000
SepalLengthCm    2.808677
SepalWidth       1.286429
PetalLengths     3.255027
PetalWidth       3.137255
dtype: float64
```

```
# Load the data
df = pd.read_csv('Iris.csv')
df.drop(['Id'], inplace=True)

# Calculate correlation matrix
corr = df.corr(method='pearson')

# Print the correlation matrix
print(corr)

# Find the dominance factor
dominance = corr.abs.sum(axis=1)
print(dominance)
```

```
               SepalLengthCm  SepalWidth  PetalLengths  PetalWidth
SepalLengthCm       1.000000   -0.109369      0.871754    0.817954
SepalWidth         -0.109369    1.000000     -0.420516   -0.356544
PetalLengths        0.871754   -0.420516      1.000000    0.962757
PetalWidth          0.817954   -0.356544      0.962757    1.000000
SepalLengthCm    2.808677
SepalWidth       1.286429
PetalLengths     3.255027
PetalWidth       3.137255
dtype: float64
```

Question 6

Find the correlation between pairs (excluding the Species) for the Iris dataset given below.

```
# Load the data
df = pd.read_csv('Iris.csv')
df.drop(['Id'], inplace=True)

# Calculate correlation matrix
corr = df.corr(method='pearson')

# Print the correlation matrix
print(corr)

# Find the dominance factor
dominance = corr.abs.sum(axis=1)
print(dominance)
```

```
               SepalLengthCm  SepalWidth  PetalLengths  PetalWidth
SepalLengthCm       1.000000   -0.109369      0.871754    0.817954
SepalWidth         -0.109369    1.000000     -0.420516   -0.356544
PetalLengths        0.871754   -0.420516      1.000000    0.962757
PetalWidth          0.817954   -0.356544      0.962757    1.000000
SepalLengthCm    2.808677
SepalWidth       1.286429
PetalLengths     3.255027
PetalWidth       3.137255
dtype: float64
```

```
# Load the data
df = pd.read_csv('Iris.csv')
df.drop(['Id'], inplace=True)

# Calculate correlation matrix
corr = df.corr(method='pearson')

# Print the correlation matrix
print(corr)

# Find the dominance factor
dominance = corr.abs.sum(axis=1)
print(dominance)
```

```
               SepalLengthCm  SepalWidth  PetalLengths  PetalWidth
SepalLengthCm       1.000000   -0.109369      0.871754    0.817954
SepalWidth         -0.109369    1.000000     -0.420516   -0.356544
PetalLengths        0.871754   -0.420516      1.000000    0.962757
PetalWidth          0.817954   -0.356544      0.962757    1.000000
SepalLengthCm    2.808677
SepalWidth       1.286429
PetalLengths     3.255027
PetalWidth       3.137255
dtype: float64
```

Question 7

Find the correlation between pairs (excluding the Species) for the Iris dataset given below.

```
# Load the data
df = pd.read_csv('Iris.csv')
df.drop(['Id'], inplace=True)

# Calculate correlation matrix
corr = df.corr(method='pearson')

# Print the correlation matrix
print(corr)

# Find the dominance factor
dominance = corr.abs.sum(axis=1)
print(dominance)
```

```
               SepalLengthCm  SepalWidth  PetalLengths  PetalWidth
SepalLengthCm       1.000000   -0.109369      0.871754    0.817954
SepalWidth         -0.109369    1.000000     -0.420516   -0.356544
PetalLengths        0.871754   -0.420516      1.000000    0.962757
PetalWidth          0.817954   -0.356544      0.962757    1.000000
SepalLengthCm    2.808677
SepalWidth       1.286429
PetalLengths     3.255027
PetalWidth       3.137255
dtype: float64
```

```
# Load the data
df = pd.read_csv('Iris.csv')
df.drop(['Id'], inplace=True)

# Calculate correlation matrix
corr = df.corr(method='pearson')

# Print the correlation matrix
print(corr)

# Find the dominance factor
dominance = corr.abs.sum(axis=1)
print(dominance)
```

```
               SepalLengthCm  SepalWidth  PetalLengths  PetalWidth
SepalLengthCm       1.000000   -0.109369      0.871754    0.817954
SepalWidth         -0.109369    1.000000     -0.420516   -0.356544
PetalLengths        0.871754   -0.420516      1.000000    0.962757
PetalWidth          0.817954   -0.356544      0.962757    1.000000
SepalLengthCm    2.808677
SepalWidth       1.286429
PetalLengths     3.255027
PetalWidth       3.137255
dtype: float64
```

Question 8

Find the correlation between pairs (excluding the Species) for the Iris dataset given below.

```
# Load the data
df = pd.read_csv('Iris.csv')
df.drop(['Id'], inplace=True)

# Calculate correlation matrix
corr = df.corr(method='pearson')

# Print the correlation matrix
print(corr)

# Find the dominance factor
dominance = corr.abs.sum(axis=1)
print(dominance)
```

```
               SepalLengthCm  SepalWidth  PetalLengths  PetalWidth
SepalLengthCm       1.000000   -0.109369      0.871754    0.817954
SepalWidth         -0.109369    1.000000     -0.420516   -0.356544
PetalLengths        0.871754   -0.420516      1.000000    0.962757
PetalWidth          0.817954   -0.356544      0.962757    1.000000
SepalLengthCm    2.808677
SepalWidth       1.286429
PetalLengths     3.255027
PetalWidth       3.137255
dtype: float64
```

```
# Load the data
df = pd.read_csv('Iris.csv')
df.drop(['Id'], inplace=True)

# Calculate correlation matrix
corr = df.corr(method='pearson')

# Print the correlation matrix
print(corr)

# Find the dominance factor
dominance = corr.abs.sum(axis=1)
print(dominance)
```

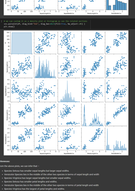
```
               SepalLengthCm  SepalWidth  PetalLengths  PetalWidth
SepalLengthCm       1.000000   -0.109369      0.871754    0.817954
SepalWidth         -0.109369    1.000000     -0.420516   -0.356544
PetalLengths        0.871754   -0.420516      1.000000    0.962757
PetalWidth          0.817954   -0.356544      0.962757    1.000000
SepalLengthCm    2.808677
SepalWidth       1.286429
PetalLengths     3.255027
PetalWidth       3.137255
dtype: float64
```

Question 9

Find the correlation between pairs (excluding the Species) for the Iris dataset given below.

```
# Load the data
df = pd.read_csv('Iris.csv')
df.drop(['Id'], inplace=True)

# Calculate correlation matrix
corr = df.corr(method='pearson')

# Print the correlation matrix
print(corr)

# Find the dominance factor
dominance = corr.abs.sum(axis=1)
print(dominance)
```

```
               SepalLengthCm  SepalWidth  PetalLengths  PetalWidth
SepalLengthCm       1.000000   -0.109369      0.871754    0.817954
SepalWidth         -0.109369    1.000000     -0.420516   -0.356544
PetalLengths        0.871754   -0.420516      1.000000    0.962757
PetalWidth          0.817954   -0.356544      0.962757    1.000000
SepalLengthCm    2.808677
SepalWidth       1.286429
PetalLengths     3.255027
PetalWidth       3.137255
dtype: float64
```




```
# Load the data
df = pd.read_csv('Iris.csv')
df.drop(['Id'], inplace=True)

# Calculate correlation matrix
corr = df.corr(method='pearson')

# Print the correlation matrix
print(corr)

# Find the dominance factor
dominance = corr.abs.sum(axis=1)
print(dominance)
```

```
               SepalLengthCm  SepalWidth  PetalLengths  PetalWidth
SepalLengthCm       1.000000   -0.109369      0.871754    0.817954
SepalWidth         -0.109369    1.000000     -0.420516   -0.356544
PetalLengths        0.871754   -0.420516      1.000000    0.962757
PetalWidth          0.817954   -0.356544      0.962757    1.000000
SepalLengthCm    2.808677
SepalWidth       1.286429
PetalLengths     3.255027
PetalWidth       3.137255
dtype: float64
```

Question 10

Find the correlation between pairs (excluding the Species) for the Iris dataset given below.

```
# Load the data
df = pd.read_csv('Iris.csv')
df.drop(['Id'], inplace=True)

# Calculate correlation matrix
corr = df.corr(method='pearson')

# Print the correlation matrix
print(corr)

# Find the dominance factor
dominance = corr.abs.sum(axis=1)
print(dominance)
```

```
               SepalLengthCm  SepalWidth  PetalLengths  PetalWidth
SepalLengthCm       1.000000   -0.109369      0.871754    0.817954
SepalWidth         -0.109369    1.000000     -0.420516   -0.356544
PetalLengths        0.871754   -0.420516      1.000000    0.962757
PetalWidth          0.817954   -0.356544      0.962757    1.000000
SepalLengthCm    2.808677
SepalWidth       1.286429
PetalLengths     3.255027
PetalWidth       3.137255
dtype: float64
```

```
# Load the data
df = pd.read_csv('Iris.csv')
df.drop(['Id'], inplace=True)

# Calculate correlation matrix
corr = df.corr(method='pearson')

# Print the correlation matrix
print(corr)

# Find the dominance factor
dominance = corr.abs.sum(axis=1)
print(dominance)
```

```
               SepalLengthCm  SepalWidth  PetalLengths  PetalWidth
SepalLengthCm       1.000000   -0.109369      0.871754    0.817954
SepalWidth         -0.109369    1.000000     -0.420516   -0.356544
PetalLengths        0.871754   -0.420516      1.000000    0.962757
PetalWidth          0.817954   -0.356544      0.962757    1.000000
SepalLengthCm    2.808677
SepalWidth       1.286429
PetalLengths     3.255027
PetalWidth       3.137255
dtype: float64
```

Question 11

Find the correlation between pairs (excluding the Species) for the Iris dataset given below.

```
# Load the data
df = pd.read_csv('Iris.csv')
df.drop(['Id'], inplace=True)

# Calculate correlation matrix
corr = df.corr(method='pearson')

# Print the correlation matrix
print(corr)

# Find the dominance factor
dominance = corr.abs.sum(axis=1)
print(dominance)
```

```
               SepalLengthCm  SepalWidth  PetalLengths  PetalWidth
SepalLengthCm       1.000000   -0.109369      0.871754    0.817954
SepalWidth         -0.109369    1.000000     -0.420516   -0.356544
PetalLengths        0.871754   -0.420516      1.000000    0.962757
PetalWidth          0.817954   -0.356544      0.962757    1.000000
SepalLengthCm    2.808677
SepalWidth       1.286429
PetalLengths     3.255027
PetalWidth       3.137255
dtype: float64
```

```
# Load the data
df = pd.read_csv('Iris.csv')
df.drop(['Id'], inplace=True)

# Calculate correlation matrix
corr = df.corr(method='pearson')

# Print the correlation matrix
print(corr)

# Find the dominance factor
dominance = corr.abs.sum(axis=1)
print(dominance)
```

```
               SepalLengthCm  SepalWidth  PetalLengths  PetalWidth
SepalLengthCm       1.000000   -0.109369      0.871754    0.817954
SepalWidth         -0.109369    1.000000     -0.420516   -0.356544
PetalLengths        0.871754   -0.420516      1.000000    0.962757
PetalWidth          0.817954   -0.356544      0.962757    1.000000
SepalLengthCm    2.808677
SepalWidth       1.286429
PetalLengths     3.255027
PetalWidth       3.137255
dtype: float64
```

Question 12

Find the correlation between pairs (excluding the Species) for the Iris dataset given below.

```
# Load the data
df = pd.read_csv('Iris.csv')
df.drop(['Id'], inplace=True)

# Calculate correlation matrix
corr = df.corr(method='pearson')

# Print the correlation matrix
print(corr)

# Find the dominance factor
dominance = corr.abs.sum(axis=1)
print(dominance)
```

```
               SepalLengthCm  SepalWidth  PetalLengths  PetalWidth
SepalLengthCm       1.000000   -0.109369      0.871754    0.817954
SepalWidth         -0.109369    1.000000     -0.420516   -0.356544
PetalLengths        0.871754   -0.420516      1.000000    0.962757
PetalWidth          0.817954   -0.356544      0.962757    1.000000
SepalLengthCm    2.808677
SepalWidth       1.286429
PetalLengths     3.255027
PetalWidth       3.137255
dtype: float64
```



```
# Load the data
df = pd.read_csv('Iris.csv')
df.drop(['Id'], inplace=True)

# Calculate correlation matrix
corr = df.corr(method='pearson')

# Print the correlation matrix
print(corr)

# Find the dominance factor
dominance = corr.abs.sum(axis=1)
print(dominance)
```

```
               SepalLengthCm  SepalWidth  PetalLengths  PetalWidth
SepalLengthCm       1.000000   -0.109369      0.871754    0.817954
SepalWidth         -0.109369    1.000000     -0.420516   -0.356544
PetalLengths        0.871754   -0.420516      1.000000    0.962757
PetalWidth          0.817954   -0.356544      0.962757    1.000000
SepalLengthCm    2.808677
SepalWidth       1.286429
PetalLengths     3.255027
PetalWidth       3.137255
dtype: float64
```

Kindly share the above URL and code that :

- Generates Sample Data for the Iris Dataset (Kindly share the code for the above visualization in the form of a function that takes the input data and the parameters to the function for the visualization)
- Generate Sample Data for the Iris Dataset (Kindly share the code for the above visualization in the form of a function that takes the input data and the parameters to the function for the visualization)
- Generate Sample Data for the Iris Dataset (Kindly share the code for the above visualization in the form of a function that takes the input data and the parameters to the function for the visualization)
- Generate Sample Data for the Iris Dataset (Kindly share the code for the above visualization in the form of a function that takes the input data and the parameters to the function for the visualization)
- Generate Sample Data for the Iris Dataset (Kindly share the code for the above visualization in the form of a function that takes the input data and the parameters to the function for the visualization)

Question 4

Find the correlation between pairs (excluding the Species) for the Iris dataset given below.


```
# Load the data
df = pd.read_csv('Iris.csv')
df.drop(['Id'], inplace=True)

# Calculate correlation matrix
corr = df.corr(method='pearson')

# Print the correlation matrix
print(corr)

# Find the dominance factor
dominance = corr.abs.sum(axis=1)
print(dominance)
```

```
               SepalLengthCm  SepalWidth  PetalLengths  PetalWidth
SepalLengthCm       1.000000   -0.109369      0.871754    0.817954
SepalWidth         -0.109369    1.000000     -0.420516   -0.356544
PetalLengths        0.871754   -0.420516      1.000000    0.962757
PetalWidth          0.817954   -0.356544      0.962757    1.000000
SepalLengthCm    2.808677
SepalWidth       1.286429
PetalLengths     3.255027
PetalWidth       3.137255
dtype: float64
```



```
# Load the data
df = pd.read_csv('Iris.csv')
df.drop(['Id'], inplace=True)

# Calculate correlation matrix
corr = df.corr(method='pearson')

# Print the correlation matrix
print(corr)

# Find the dominance factor
dominance = corr.abs.sum(axis=1)
print(dominance)
```

```
               SepalLengthCm  SepalWidth  PetalLengths  PetalWidth
SepalLengthCm       1.000000   -0.109369      0.871754    0.817954
SepalWidth         -0.109369    1.000000     -0.420516   -0.356544
PetalLengths        0.871754   -0.420516      1.000000    0.962757
PetalWidth          0.817954   -0.356544      0.962757    1.000000
SepalLengthCm    2.808677
SepalWidth       1.286429
PetalLengths     3.255027
PetalWidth       3.137255
dtype: float64
```

Trouble submitting your feedback? View the Help Center for more ways to get assistance.