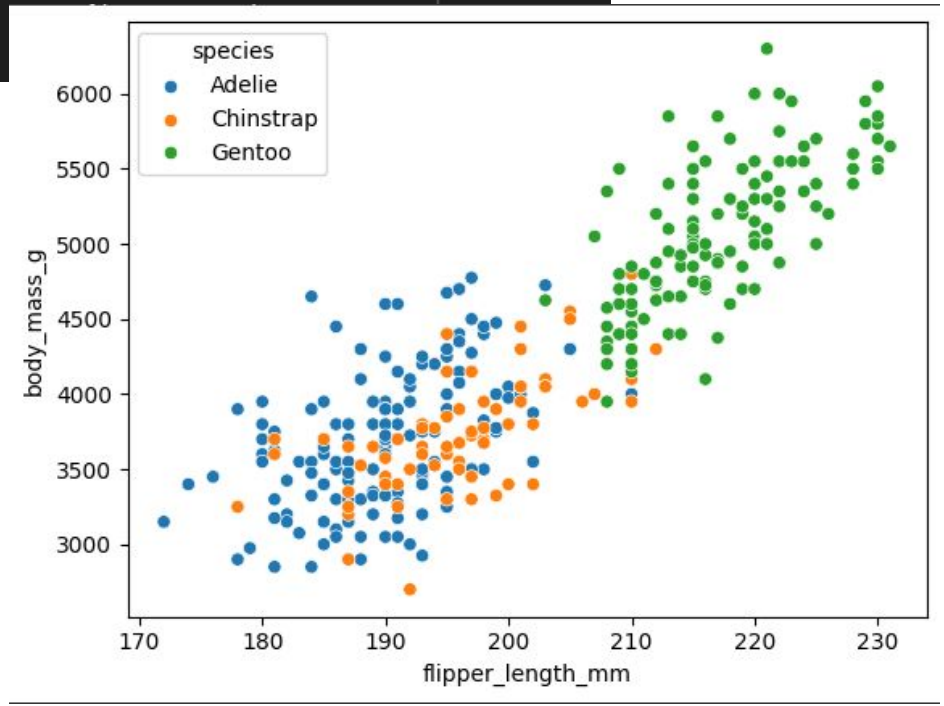
Penguins dataset Exploration

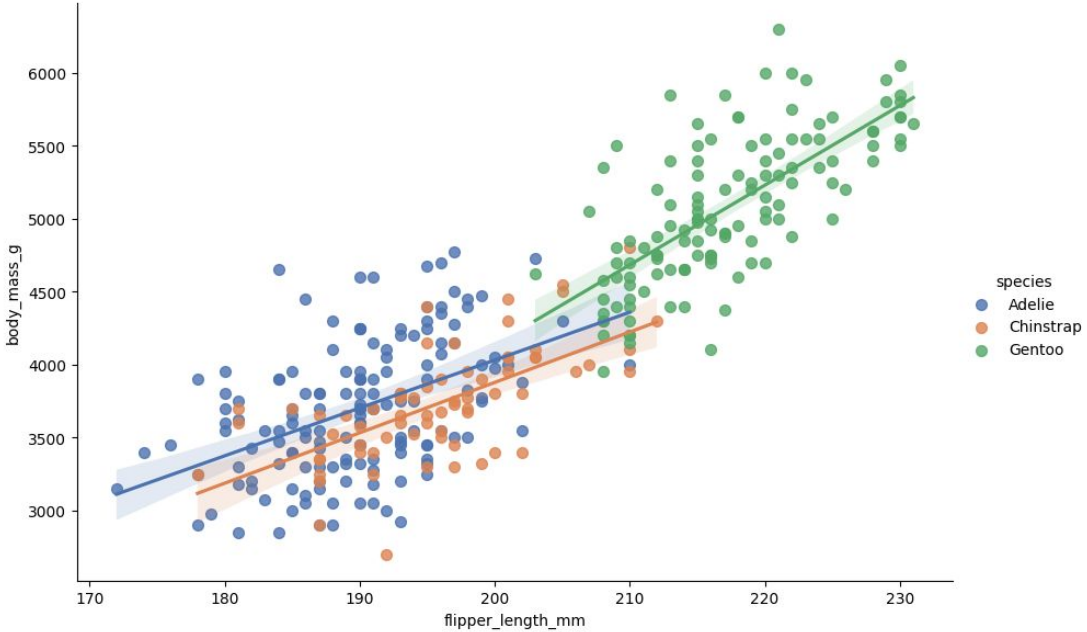
Scatter color: plots points with specified color

```
#Scatter color
import seaborn as sns
import pandas as pd
import matplotlib.pyplot as plt
sns.scatterplot(data=data, x='flipper_length_mm', y='body_mass_g', hue='species')
data = sns.load_dataset('penguins')
plt.show()
```



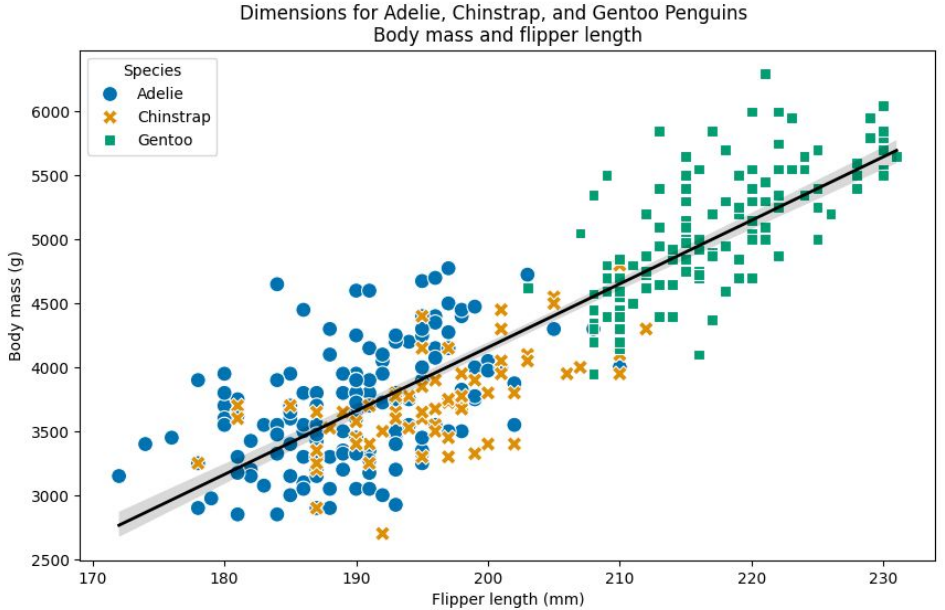
Scatter line: To plot a line through scatter points

```
import seaborn as sns
import matplotlib.pyplot as plt
sns.lmplot(data=data, x='flipper_length_mm', y='body_mass_g', hue='species',
           aspect=1.5, height=6, palette='deep', scatter_kws={'s': 50})
data = sns.load_dataset('penguins')
plt.show()
```



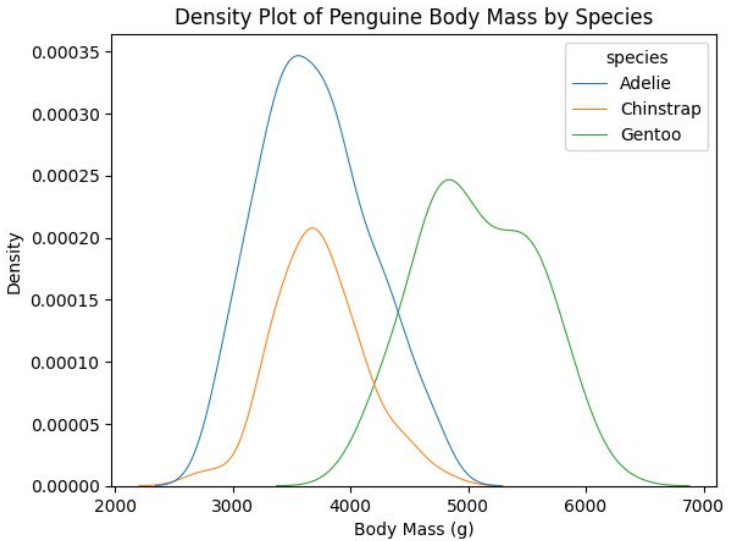
Scatter color and shape: Assign different color & shape for each specifies

```python
#Scatter color and shape :
import seaborn as sns
import matplotlib.pyplot as plt

# Load the penguins dataset
data = sns.load_dataset('penguins')

# Create a scatter plot with color and shape differentiation
plt.figure(figsize=(10, 6))
scatter_plot = sns.scatterplot(data=data, x='flipper_length_mm', y='body_mass_g',
                               hue='species', style='species', palette='colorblind',
                               s=100)  # Adjust marker size with 's'

# Add a linear regression line
sns.regplot(data=data, x='flipper_length_mm', y='body_mass_g',
            scatter=False, color='black', line_kws={'linewidth': 2})

# Customize the plot with labels and title
plt.title('Body mass and flipper length')
plt.suptitle('Dimensions for Adelie, Chinstrap, and Gentoo Penguins', y=0.95)
plt.xlabel('Flipper length (mm)')
plt.ylabel('Body mass (g)')
plt.legend(title='Species')

# Display the plot
plt.show()
```

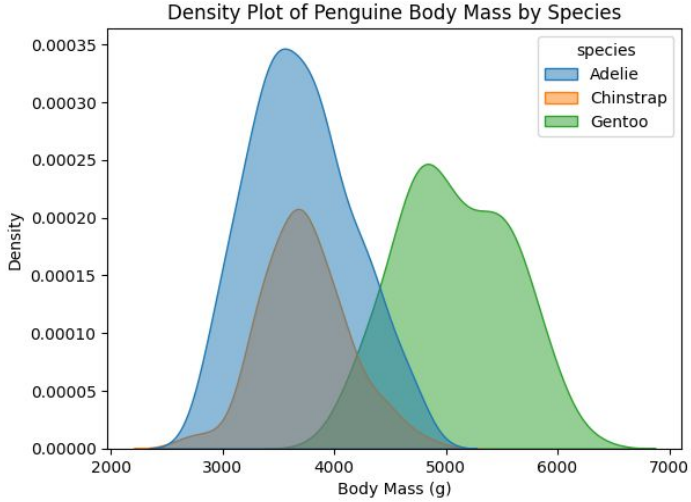


Curved plot: To plot smooth curve(line) between points

```
import seaborn as sns
import matplotlib.pyplot as plt

# Load the penguins dataset
penguins = sns.load_dataset('penguins')

sns.kdeplot(data=penguins, x='body_mass_g', hue='species', linewidth=0.75)

plt.xlabel('Body Mass (g)')
plt.ylabel('Density')
plt.title('Density Plot of Penguine Body Mass by Species')
# Show the plot
plt.show()
```

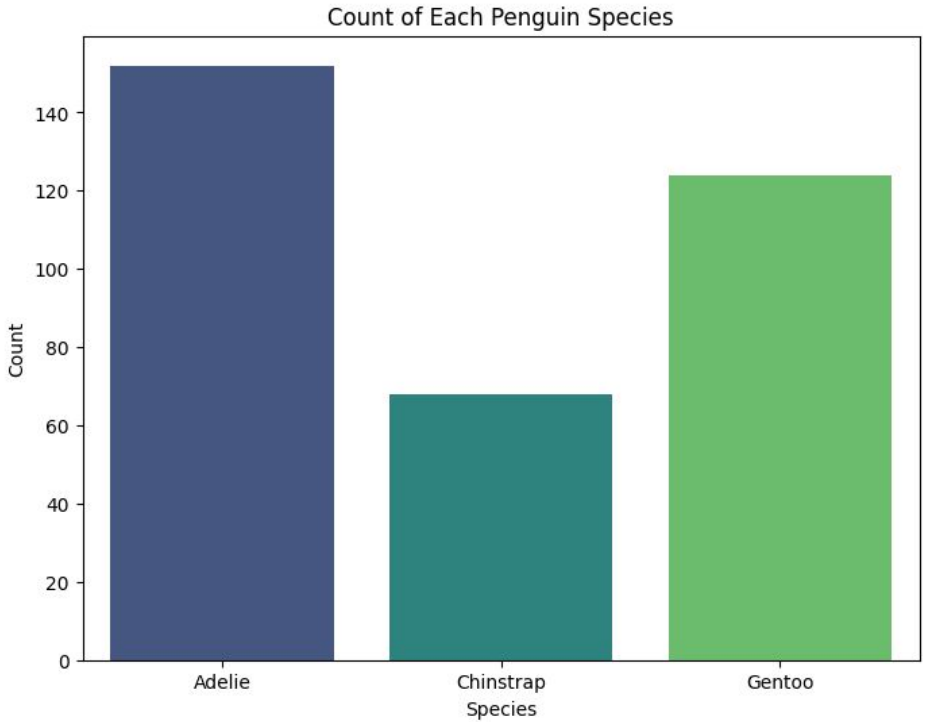



```
import seaborn as sns
import matplotlib.pyplot as plt

# Load the penguins dataset
penguins = sns.load_dataset('penguins')

sns.kdeplot(data=penguins, x='body_mass_g', hue='species', fill=True, alpha=0.5)


plt.xlabel('Body Mass (g)')
plt.ylabel('Density')
plt.title('Density Plot of Penguine Body Mass by Species')
# Show the plot
plt.show()
```



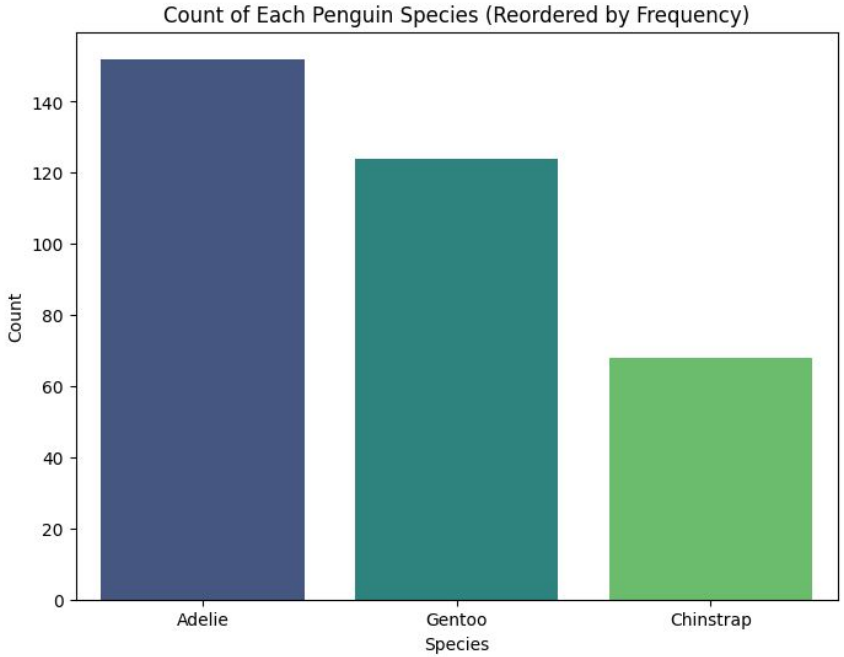
Bar graph: to plot bar graph

```python
#bar graph:
import seaborn as sns
import matplotlib.pyplot as plt

# Load the penguins dataset
data = sns.load_dataset('penguins')

# Create a bar plot of species counts
plt.figure(figsize=(8, 6))
sns.countplot(data=data, x='species', palette='viridis')

# Customize the plot with labels and title
plt.title('Count of Each Penguin Species')
plt.xlabel('Species')
plt.ylabel('Count')

# Display the plot
plt.show()
```



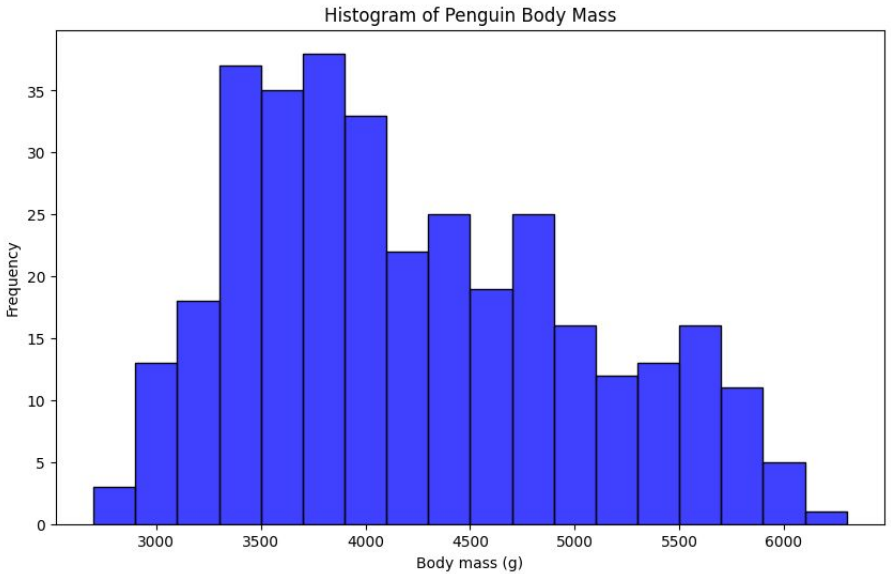
Bar graph descending order: To plot bar graph in descending order

```python
#bar graph descending order:
import seaborn as sns
import matplotlib.pyplot as plt
import pandas as pd

# Load the penguins dataset
data = sns.load_dataset('penguins')

# Reorder 'species' based on frequency
data['species'] = pd.Categorical(data['species'],
                                 categories=data['species'].value_counts().index,
                                 ordered=True)

# Create a bar plot of species counts with reordered categories
plt.figure(figsize=(8, 6))
sns.countplot(data=data, x='species', palette='viridis')

# Customize the plot with labels and title
plt.title('Count of Each Penguin Species (Reordered by Frequency)')
plt.xlabel('Species')
plt.ylabel('Count')

# Display the plot
plt.show()
```



Bar graph with specified width of 200

```python
#bar graph with width=200:
import seaborn as sns
import matplotlib.pyplot as plt

# Load the penguins dataset
data = sns.load_dataset('penguins')

# Create a histogram with a specific bin width
plt.figure(figsize=(10, 6))

# Calculate the number of bins
bin_width = 200
num_bins = int((data['body_mass_g'].max() - data['body_mass_g'].min()) / bin_width)

# Plot the histogram
sns.histplot(data=data, x='body_mass_g', bins=num_bins, kde=False, color='blue')

# Customize the plot with labels and title
plt.title('Histogram of Penguin Body Mass')
plt.xlabel('Body mass (g)')
plt.ylabel('Frequency')

# Display the plot
plt.show()
```



Bar graph with specified width of 20

```python
#bar graph with width=20:
import seaborn as sns
import matplotlib.pyplot as plt
import numpy as np

# Load the penguins dataset
data = sns.load_dataset('penguins')

# Desired bin width
bin_width = 20

# Calculate the number of bins
min_value = data['body_mass_g'].min()
max_value = data['body_mass_g'].max()
num_bins = int((max_value - min_value) / bin_width)

# Create a histogram with the specified bin width
plt.figure(figsize=(10, 6))
sns.histplot(data=data, x='body_mass_g', bins=num_bins, color='blue')

# Customize the plot with labels and title
plt.title('Histogram of Penguin Body Mass')
plt.xlabel('Body mass (g)')
plt.ylabel('Frequency')

# Display the plot
plt.show()
```