Name: Vinayak V Thayil

Roll No:AM.EN.U4CSE21161

```python
import pandas as pd
import matplotlib.pyplot as plt
import seaborn as sns



penguins = sns.load_dataset("penguins")
flights = sns.load_dataset("flights")
titanic = sns.load_dataset("titanic")
car_crashes = sns.load_dataset("car_crashes")
fmri = sns.load_dataset("fmri")
diamonds = sns.load_dataset("diamonds")



penguins_grouped = penguins[['species', 'bill_length_mm']].groupby('species').mean().reset_index()



plt.bar(penguins_grouped['species'], penguins_grouped['bill_length_mm'])
plt.title('Average penguin bill length by species')
plt.show()
```
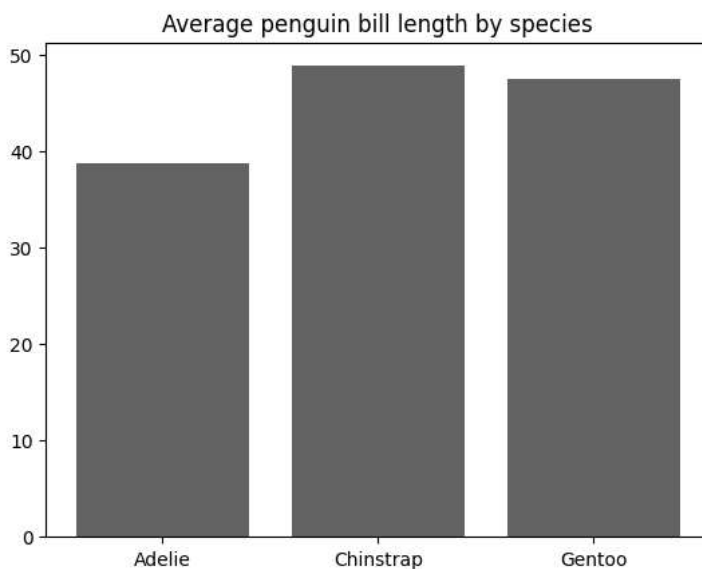


```python
# Data preparation
flights_grouped = flights[['year', 'passengers']].astype({'year': 'string'}).groupby('year').sum().reset_index()
# Creating a line plot
plt.plot(flights_grouped['year'], flights_grouped['passengers'])
plt.title('Total number of passengers by year')
plt.show()
```
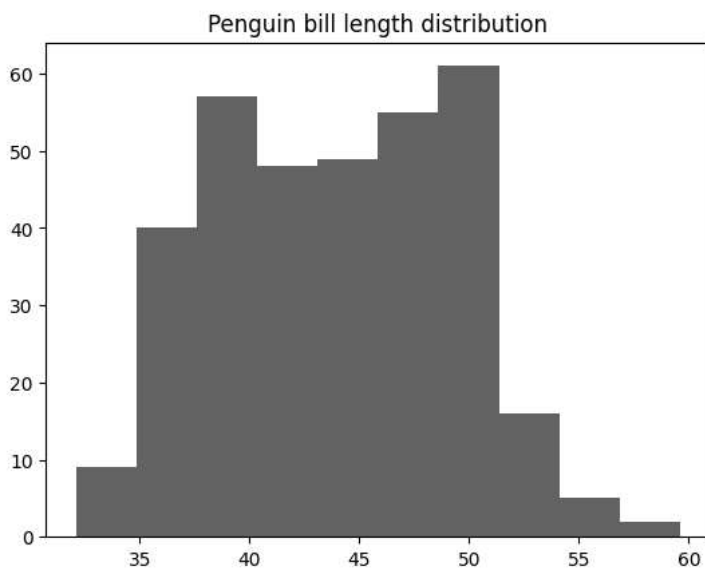
Total number of passengers by year

```
# Creating a scatter plot
plt.scatter(penguins['bill_length_mm'], penguins['bill_depth_mm'])
plt.title('Penguin bill length vs. bill depth')
plt.show()
```

Penguin bill length vs. bill depth
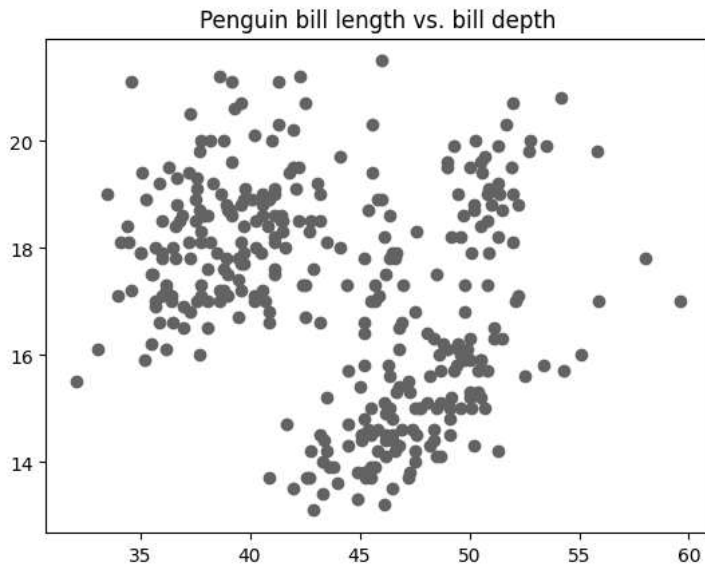

```
# Creating a histogram
plt.hist(penguins['bill_length_mm'])
plt.title('Penguin bill length distribution')
plt.show()
```

Penguin bill length distribution


```
# Data preparation
penguins_cleaned = penguins.dropna()
# Creating a box plot
plt.boxplot(penguins_cleaned['bill_length_mm'])
plt.title('Penguin bill length distribution')
plt.show()
```
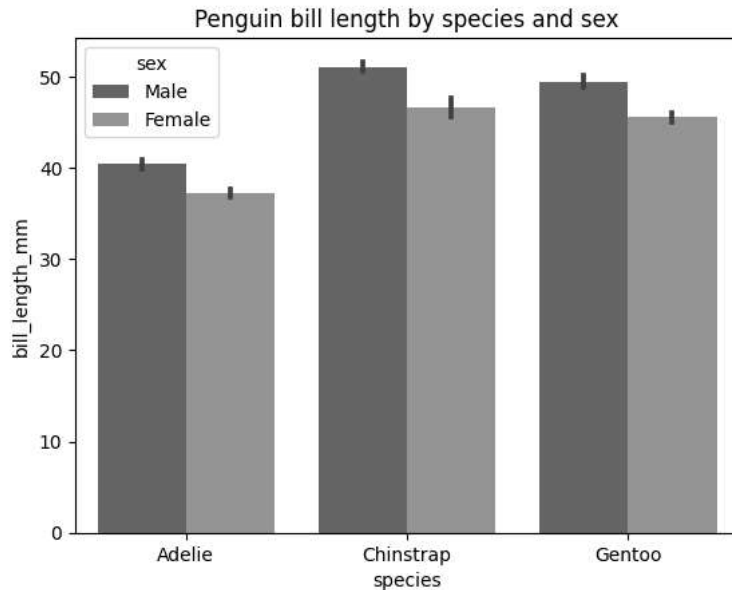
Penguin bill length distribution
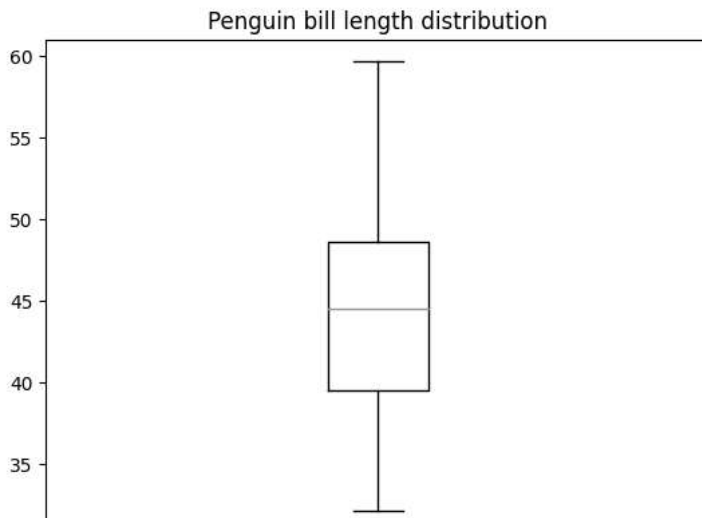

```python
# Creating a box plot
sns.boxplot(data=penguins, y='bill_length_mm')
plt.title('Penguin bill length distribution')
plt.show()
```
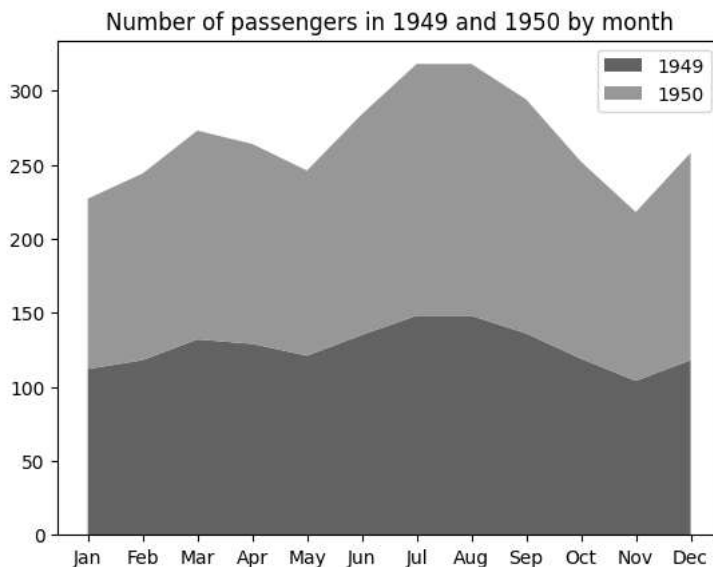
Penguin bill length distribution
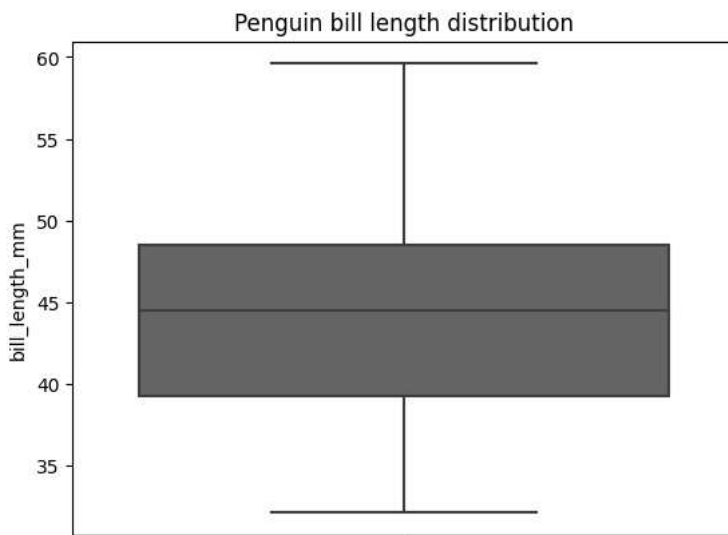

```python
# Data preparation
titanic_grouped = titanic.groupby('class')['pclass'].count().reset_index()
# Creating a pie chart
plt.pie(titanic_grouped['pclass'], labels=titanic_grouped['class'])
plt.title('Number of passengers by class')
plt.show()
```

Number of passengers by class

```
# Creating a grouped bar chart
sns.barplot(data=penguins, x='species', y='bill_length_mm', hue='sex')
plt.title('Penguin bill length by species and sex')
plt.show()
```

Penguin bill length by species and sex



```
# Data preparation
flights_grouped = flights.groupby(['year', 'month']).mean().reset_index()
flights_49_50 = pd.DataFrame(list(zip(flights_grouped.loc[:11, 'month'].tolist(), flights_grouped.loc[:11, 'passengers'].tolist(), flights_gr
# Creating a stacked area chart
plt.stackplot(flights_49_50['month'], flights_49_50['1949'], flights_49_50['1950'], labels=['1949', '1950'])
plt.title('Number of passengers in 1949 and 1950 by month')
plt.legend()
plt.show()
```

Number of passengers in 1949 and 1950 by month


```
# Creating multiple box plots
sns.boxplot(data=penguins, x='bill_length_mm', y='species')
plt.title('Penguin bill length distribution by species')
plt.show()
```

Penguin bill length distribution by species
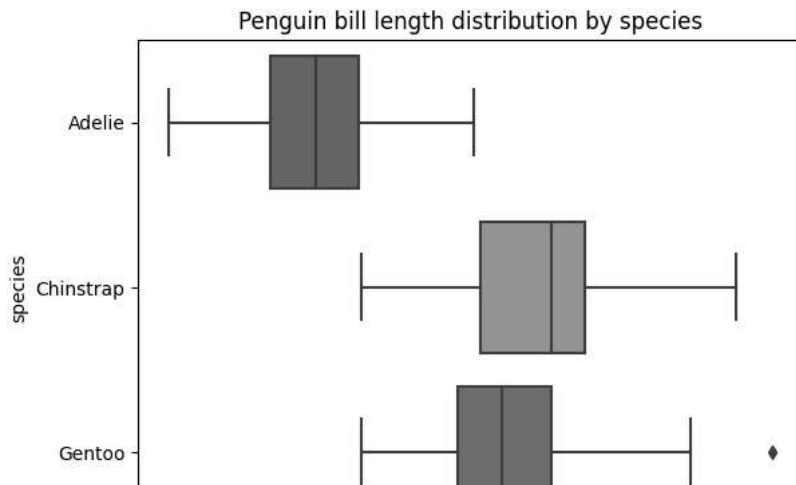
```
sns.violinplot(data=penguins, x='bill_length_mm', y='species')
plt.title('Penguin bill length distribution by species')
plt.show()
```


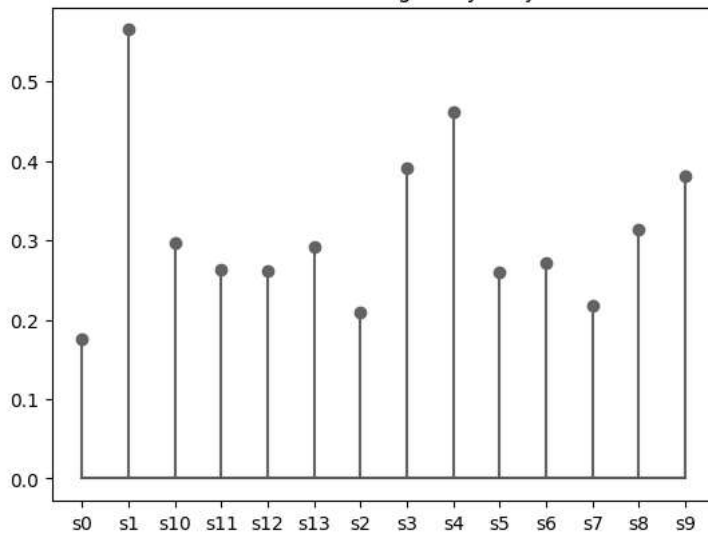Penguin bill length distribution by species
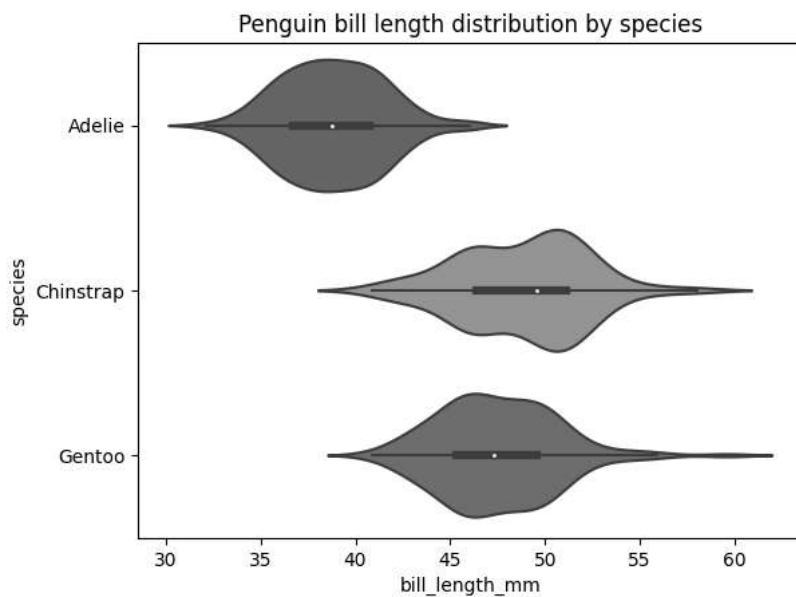
```
# Data preparation
from sklearn import preprocessing
car_crashes_cleaned = car_crashes.drop(labels='abbrev', axis=1).iloc[0:10]
min_max_scaler = preprocessing.MinMaxScaler()
car_crashes_normalized = pd.DataFrame(min_max_scaler.fit_transform(car_crashes_cleaned.values), columns=car_crashes_cleaned.columns)
# Creating a heatmap
sns.heatmap(car_crashes_normalized, annot=True)
plt.title('Car crash heatmap for the first 10 car crashes')
plt.show()
```

Car crash heatmap for the first 10 car crashes

|						
0	0.78	0.98	0.96	0.8	0.59	0	0.61
1	0.74	1	0.69	0.69	0.72	0.55	0.42
2	0.77	0.83	0.85	0.64	0.78	0.23	0
3	1	0.37	1	1	1	0.087	0.57
4	0.37	0.41	0.42	0.33	0.31	0.19	0.98

```
# Data preparation
fmri_grouped = fmri.groupby('subject')[['subject', 'signal']].max()
# Creating a stem plot
plt.stem(fmri_grouped['subject'], fmri_grouped['signal'])
plt.title('FMRI maximum signal by subject')
plt.show()
```

FMRI maximum signal by subject

```
# Creating a strip plot
sns.stripplot(data=penguins, x='bill_length_mm', y='species')
plt.title('Penguin bill length distribution by species')
plt.show()
```

Penguin bill length distribution by species
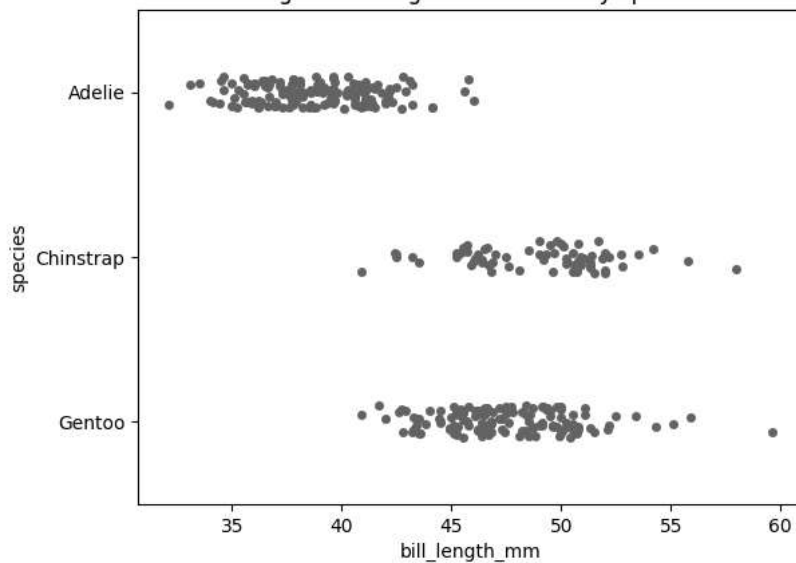
```
# Creating a swarm plot
sns.swarmplot(data=penguins, x='bill_length_mm', y='species')
```

```
plt.title('Penguin bill length distribution by species')
plt.show()
```


Penguin bill length distribution by species