assignment=24

July 28, 2023

1 Que 1: Name any five plots that we can plot using the Seaborn library. Also, state the uses of each plot.

Scatter Plot: A scatter plot is used to visualize the relationship between two continuous variables. It is a useful tool for identifying patterns or clusters in the data, and it can also show the strength and direction of the correlation between the two variables.

Line Plot: A line plot is a graph that displays the data as a series of points connected by a line. It is useful for visualizing trends over time, such as stock prices or temperature changes, and can also be used to compare multiple groups or categories.

Bar Plot: A bar plot is used to compare the values of different categories or groups. It is useful for showing the distribution of categorical data, such as the frequency of different types of cars on a road or the number of people in different age groups.

Box Plot: A box plot is a graphical representation of the distribution of numerical data. It is used to show the distribution of a dataset, and can also be used to compare multiple datasets or groups. The box represents the interquartile range, while the whiskers extend to show the range of the data.

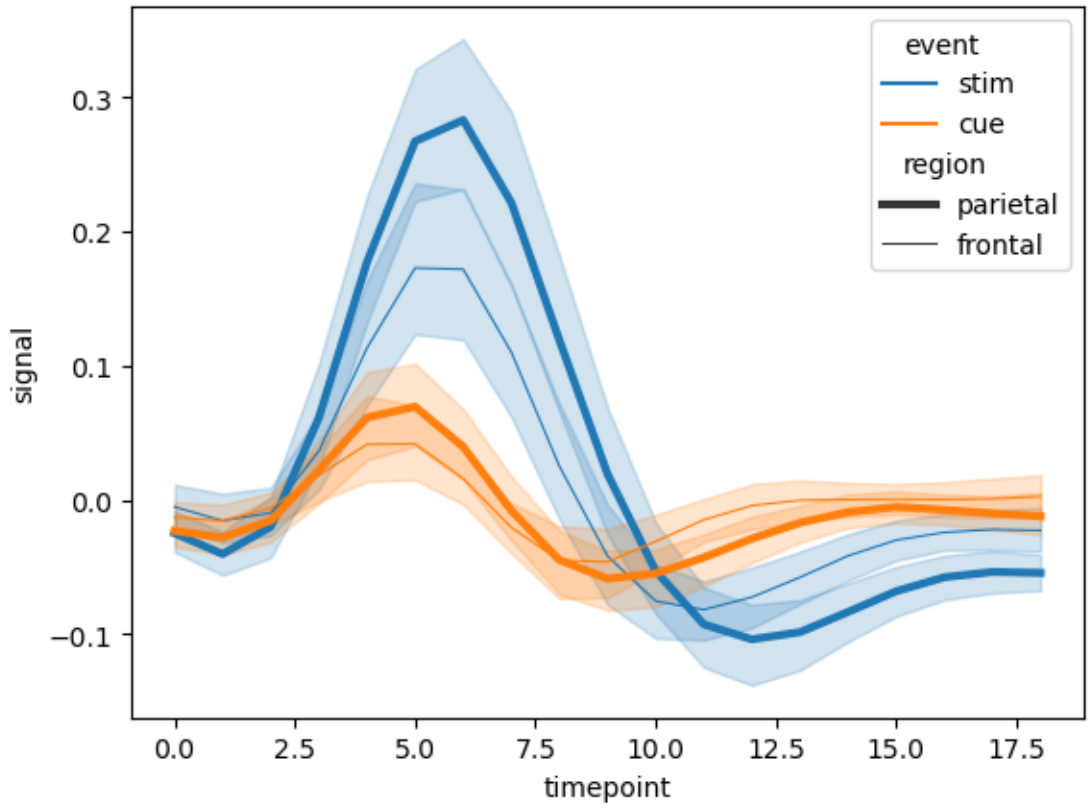
Heatmap: A heatmap is a graphical representation of data where the individual values are represented as colors. It is useful for visualizing large datasets and identifying patterns or trends in the data. It is commonly used in biology, social sciences, and finance to visualize large datasets with many variables.

2 Que 2: Load the "fmri" dataset using the load_dataset function of seaborn. Plot a line plot using x = "timepoint" and y = "signal" for different events and regions.

Note: timepoint, signal, event, and region are columns in the fmri dataset.

[1]:
```
import seaborn as s
fmri = s.load_dataset("fmri")
s.lineplot(x="timepoint", y="signal", hue="event", size="region", data=fmri)
```

[1]:
```
<AxesSubplot: xlabel='timepoint', ylabel='signal'>
```

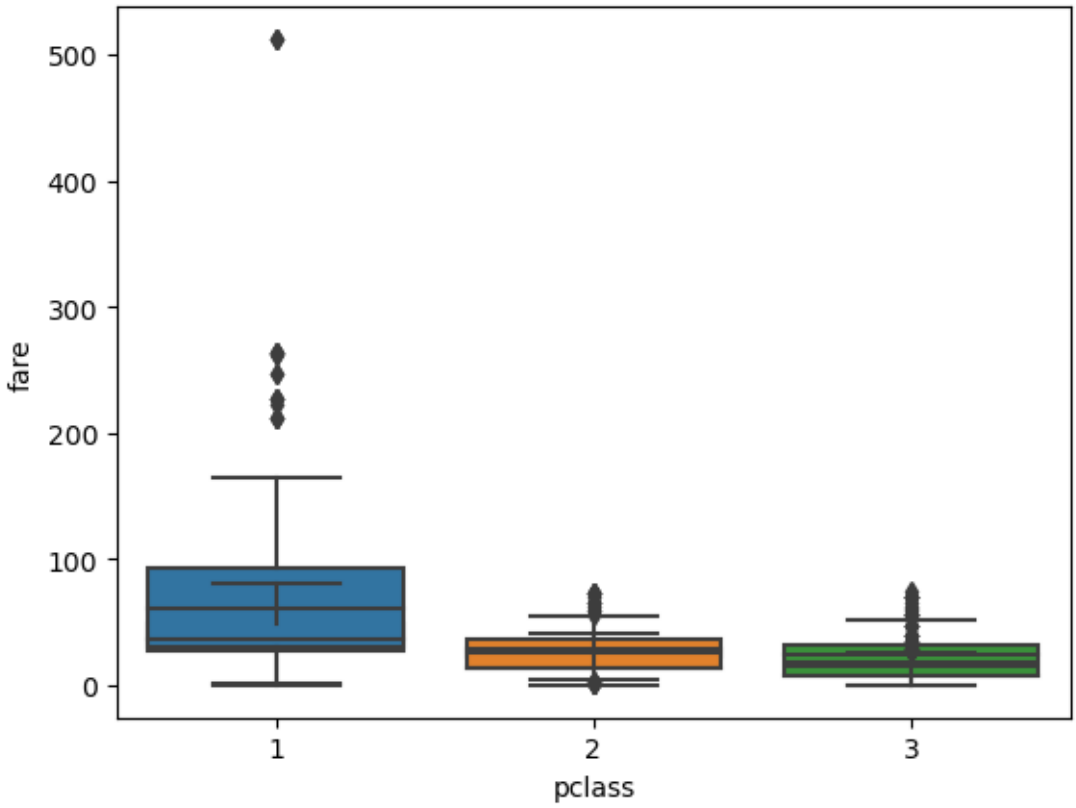


3 Que 3: Load the "titanic" dataset using the load_dataset function of seaborn. Plot two box plots using x = 'pclass', y = 'age' and y = 'fare'.

Note: pclass, age, and fare are columns in the titanic dataset.

[4]:
```
import seaborn as s
titanic = s.load_dataset("titanic")
s.boxplot(x = 'pclass', y = 'age', data= titanic)
s.boxplot(x = 'pclass', y = 'fare', data= titanic)
```

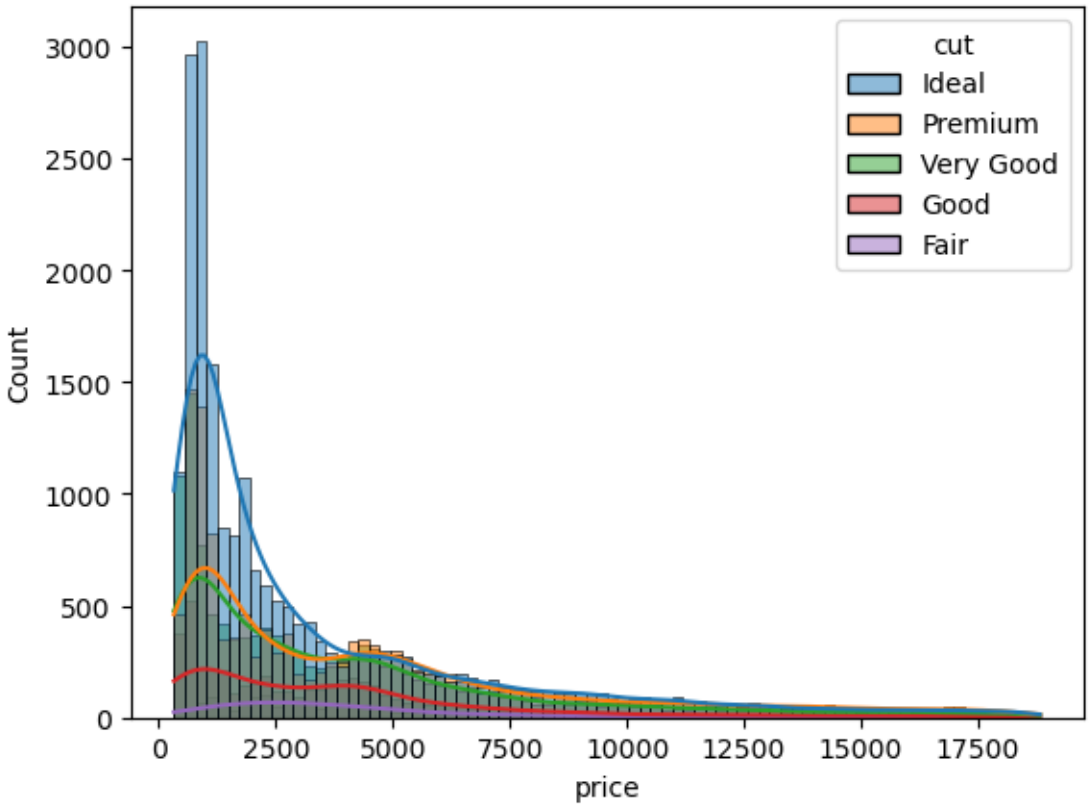
[4]:
```
<AxesSubplot: xlabel='pclass', ylabel='fare'>
```



4 Que 4: Use the "diamonds" dataset from seaborn to plot a histogram for the 'price' column. Use the hue parameter for the 'cut' column of the diamonds dataset.

[5]:
```
import seaborn as sns

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

# Plot histogram with hue
sns.histplot(data=diamonds, x='price', hue='cut', kde=True)
```

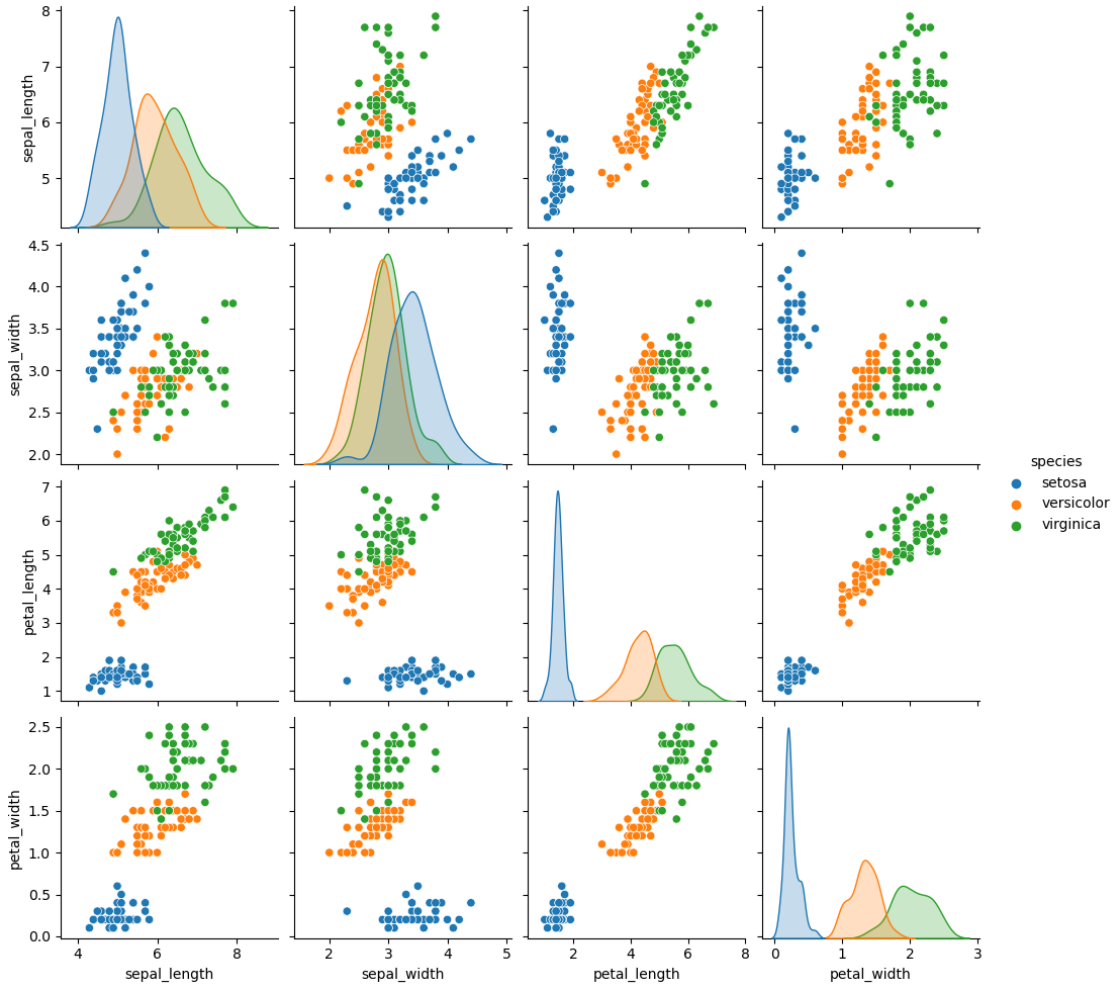
[5]: <AxesSubplot: xlabel='price', ylabel='Count'>



5 Que 5: Use the "iris" dataset from seaborn to plot a pair plot. Use the hue parameter for the "species" column of the iris dataset.

[7]:
```
import seaborn as sns
iris = sns.load_dataset("iris")
sns.pairplot(data=iris, hue="species")
```

[7]: <seaborn.axisgrid.PairGrid at 0x7f067f43f520>



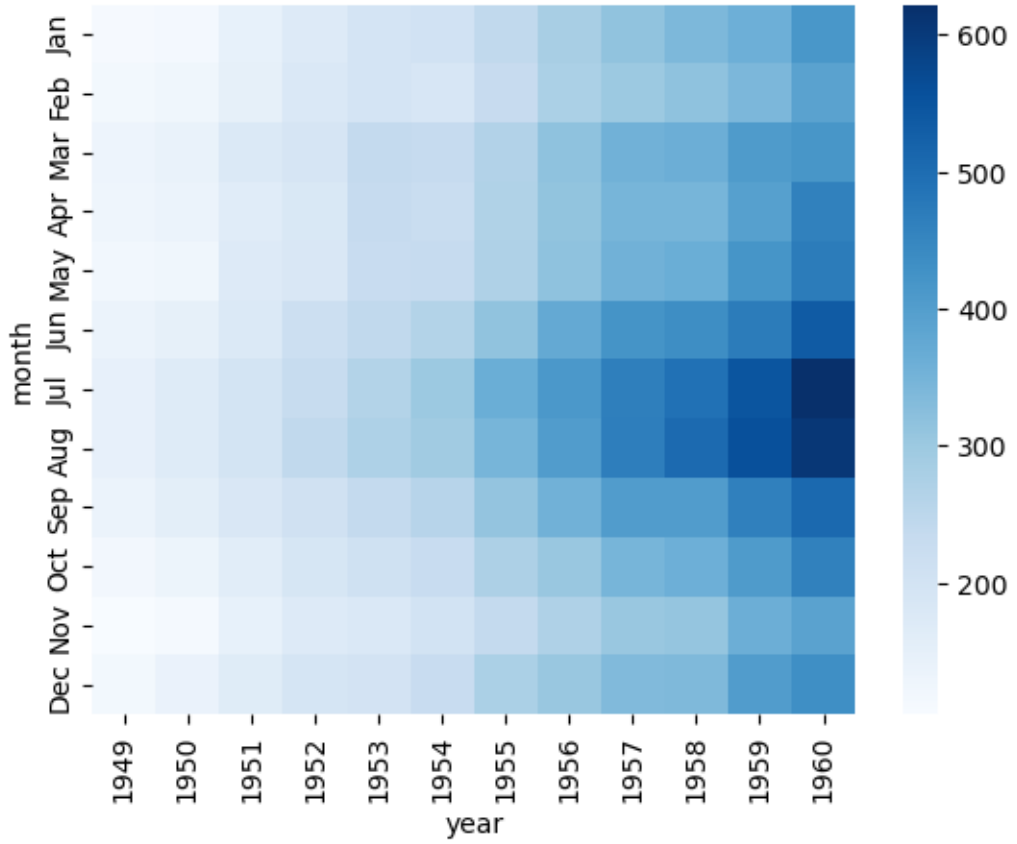
6 Que 6: Use the "flights" dataset from seaborn to plot a heatmap.

[9]:
```
import seaborn as sns

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

# Pivot the dataset to create a matrix

flights_matrix = flights.pivot('month', 'year', 'passengers')

# Create the heatmap
sns.heatmap(flights_matrix, cmap='Blues')
```

```
/tmp/ipykernel_935/3094614333.py:8: FutureWarning: In a future version of pandas
all arguments of DataFrame.pivot will be keyword-only.
```

```
flights_matrix = flights.pivot('month', 'year', 'passengers')
```

[9]: <AxesSubplot: xlabel='year', ylabel='month'>



[ ]: