Data visualization and fundamental of data visualization

Graphic system Plots convert code into graph that is called data visualization. the most important part of data science is to represent the data into graph

what is the grammer of graphics?

Major Parts

1- Mapping (Data)

2- Aestheric (Color, Shape, Size)

3- Geometric (Ojects: Line, Bar, Points, Box, Map)

"word like an artist or be a Copycat"

Python libraries of Data Science

1- Pandas (data structure and tools 2D dataframes)

2- Numpy (Arrays and matrices)

3- Scipy (Optimization and solving differential equation)

4- Matplotlib (Plots, Graphs and figures)

5- Seaborn (heat mpas, times series and other plots)

Machine learning Libraries

ML algorithmic development

1- Scikit-learn (classification, analysis and son on..)

2- statsmodels (explore data, estimation of stat modes)

variable type matters

visualization depends on the variable type

1- Categorical variable

```
-Counts

-Male vs Female
```

```
-True vs False

-0 vs 1

- yes vs NO
```

2- Countinous variable

```
- scatter plot

- statistical proportions (means and their comparison)
```

think before you start

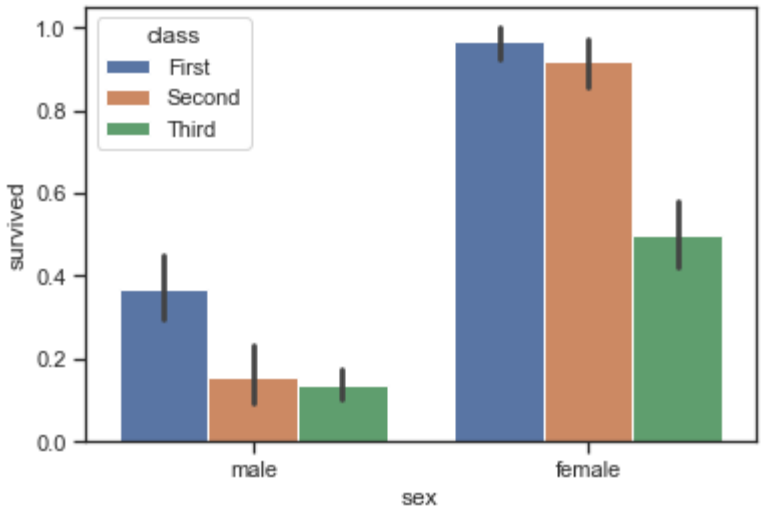
- generate a hypothesis (e.g how much male and female in course)
- find a story in your data
- clean/tidy up data
- remove duplicates
- handle missing values
- find a proper variable to drop duplicates in data

Bar Plots

In [ ]:
```
import seaborn as sns
import matplotlib.pyplot as plt

sns.set_theme(style="ticks", color_codes=True)
```

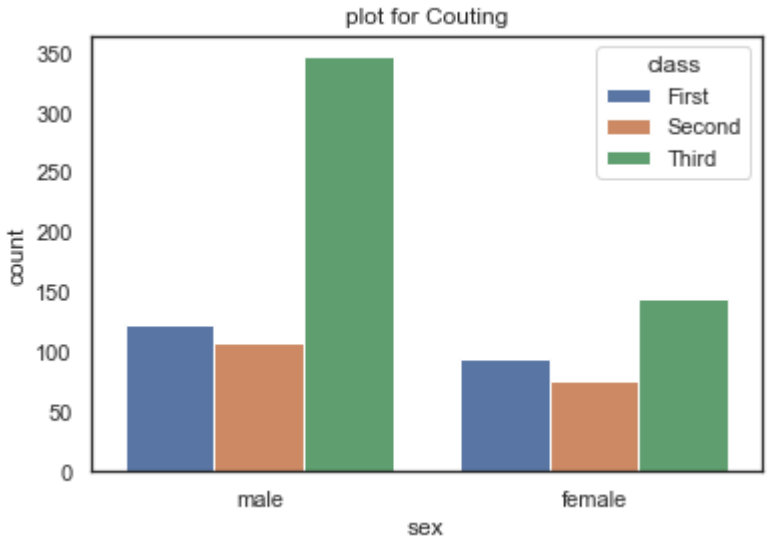
In [ ]:
```
#import data set
titanic = sns.load_dataset("titanic")

#google how many data set in seaborn??
```

In [ ]:
```
# graph
sns.barplot(x="sex", y= "survived", hue="class" ,data=titanic)
plt.show()
```



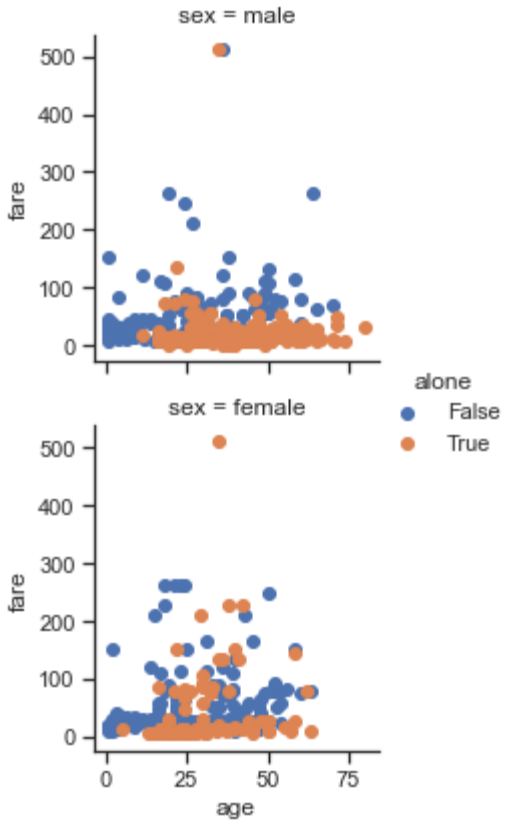
In [ ]:

```
p1 = sns.countplot(x="sex", data=titanic, hue= "class")
p1.set_title("plot for Couting")
plt.show()
```



In [ ]:

```
# scatterplot
sns.set_theme(style="ticks", color_codes =True)
g = sns.FacetGrid(titanic, row="sex", hue="alone")
g= (g.map(plt.scatter,"age","fare").add_legend())
plt.show()
```


sex = male
sex = female
fare
age
alone
False
True
0
100
200
300
400
500
0
25
50
75