

## Assignment -4

Assignment Date	03 November 2022
Student Name	Roohi Fathima K.M
Student Roll Number	812419104707
Maximum Marks	2 Marks

```

+ Code + Text
Double-click (or enter) to edit

IMPORTING LIBRARIES

[ ]
Double-click (or enter) to edit

[ ] import numpy as np
import pandas as pd
import matplotlib.pyplot as plt
Seaborn inline
import seaborn as sns

[ ] from google.colab import drive
drive.mount('/content/drive')

Mounted at /content/drive

[ ] #load the data
df = pd.read_csv('/content/drive/My Drive/Seaborn/seaborn.csv')

[ ] df

+ Code + Text
Connect Editing
[ ] #load the data
df = pd.read_csv('/content/drive/My Drive/Seaborn/seaborn.csv')

[ ] df

sex Length Standard Height Whole weight Shucked weight Viscera weight Shell weight Rings
8 M 0.455 0.365 0.365 0.5940 0.2245 0.1043 0.1500 15
1 M 0.500 0.385 0.385 0.2255 0.8995 0.0485 0.0730 7
2 F 0.330 0.430 0.130 0.4770 0.2080 0.1815 0.2130 9
9 M 0.440 0.365 0.125 0.5160 0.2155 0.1140 0.1550 10
4 I 0.550 0.355 0.380 0.2980 0.8895 0.0095 0.0550 7
-- -- -- -- -- -- -- -- --
4072 F 0.565 0.450 0.155 0.9870 0.3700 0.2290 0.2490 11
4073 M 0.890 0.440 0.130 0.9600 0.4290 0.2140 0.3600 10
4074 M 0.600 0.415 0.220 1.1760 0.5235 0.2675 0.3080 9
4076 F 0.625 0.465 0.150 1.0940 0.5310 0.2643 0.2960 10
4078 M 0.710 0.385 0.180 1.5680 0.5435 0.2760 0.4980 13
4177 rows x 9 columns

[ ] df.head()

max 0.810000 0.650000 1.150000 2.821000 1.488000 0.750000 1.020000 29.000000

[ ] df['age'] = df['Rings']*1.5
df = df.drop('Rings', axis = 1)

[ ] from sklearn.preprocessing import StandardScaler
from sklearn.model_selection import train_test_split, cross_val_score
from sklearn.feature_selection import SelectKBest
from sklearn.metrics import r2_score, mean_squared_error
import warnings
warnings.filterwarnings("ignore", category=DeprecationWarning)

UNIVARIATE ANALYSIS

[ ] sns.histmap(df,smell100)

sns.histmap(df,smell100)

```

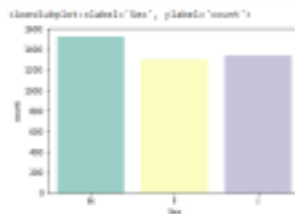


```
df.info()
```

```
<class 'pandas.core.frame.DataFrame'>
Int64Index: 4170 entries, 0 to 4170
Data columns (total 9 columns):
 #   Column      Non-Null Count  Dtype
---  -
 0   Sex         4170 non-null    object
 1   Length      4170 non-null    float64
 2   Flippers    4170 non-null    float64
 3   Depth       4170 non-null    float64
 4   Mass weight 4170 non-null    float64
 5   Shucked weight 4170 non-null    float64
 6   Viscera weight 4170 non-null    float64
 7   Shell weight 4170 non-null    float64
 8   Age         4170 non-null    float64
dtypes: float64(8), object(1)
memory usage: 281.4+ KB
```

```
df["Viscera weight"]
df["Shell weight"]
df["Age"]
dtypes: float64(8), object(1)
memory usage: 281.4+ KB
```

```
sns.countplot(x='Sex', data=df, palette='hls')
```



```
plt.figure(figsize=(10,7))
sns.countplot(x='Sex', y='Age', data=df, hue='Sex')
sns.violinplot(x='Sex', y='Age', data=df)
```

C:\Users\Arind\Anaconda3\lib\site-packages\matplotlib\cbook.py:259: UserWarning: 56.2% of the points cannot be placed; you may want to decrease the size of the markers or use stripplot.  
warnings.warn(msg, UserWarning)

C:\Users\Arind\Anaconda3\lib\site-packages\matplotlib\cbook.py:259: UserWarning: 53.2% of the points cannot be placed; you may want to decrease the size of the markers or use stripplot.  
warnings.warn(msg, UserWarning)

+ Code + Test

Connect + Editing

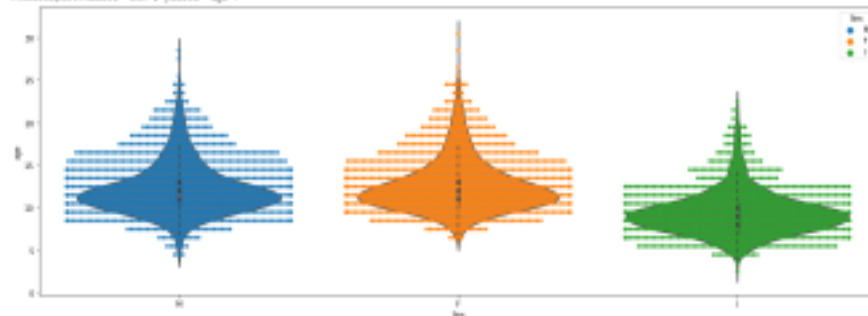
```
sns.violinplot(x='Sex', y='Age', data=df, hue='Sex')
sns.violinplot(x='Sex', y='Age', data=df)
```

C:\Users\Arind\Anaconda3\lib\site-packages\matplotlib\cbook.py:259: UserWarning: 56.2% of the points cannot be placed; you may want to decrease the size of the markers or use stripplot.  
warnings.warn(msg, UserWarning)

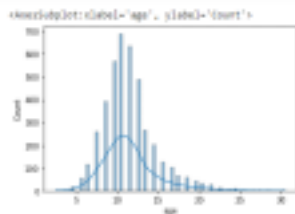
C:\Users\Arind\Anaconda3\lib\site-packages\matplotlib\cbook.py:259: UserWarning: 53.2% of the points cannot be placed; you may want to decrease the size of the markers or use stripplot.  
warnings.warn(msg, UserWarning)

C:\Users\Arind\Anaconda3\lib\site-packages\matplotlib\cbook.py:259: UserWarning: 56.2% of the points cannot be placed; you may want to decrease the size of the markers or use stripplot.  
warnings.warn(msg, UserWarning)

```
<matplotlib.axes._subplots.AxesSubplot>
```



```
sns.histplot(df, age, kde=True)
```



## BIVARIATE ANALYSIS


```
df.info()
```

## MULTI VARIATE ANALYSIS


```
df.info()
```




## • BIVARIATE ANALYSIS

 1.2 cells hidden

## • MULTI VARIATE ANALYSIS

 1.7 cells hidden

## • MISSING VALUES


 1.2 cells hidden

## • OUTLIERS

 1.18 cells hidden

## • CATEGORICAL COLUMNS

## • SCALING THE INDEPENDENT VARIABLE

 1.2 cells hidden


## • SPLITTING THE DATA

 1.2 cells hidden

## • BUILDING MODEL

 1.4 cells hidden

## • TRAINING THE MODEL

 1.3 cells hidden

## • TESTING THE MODEL

 1.8 cells hidden