

▼ Problem Statement: Abalone Age Prediction

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

▼ Description :

Predicting the age of abalone from physical measurements. The age of abalone is determined by cutting the shell through the cone, staining it, and counting the number of rings through a microscope -- a boring and time-consuming task. Other measurements, which are easier to obtain, are used to predict age. Further information, such as weather patterns and location (hence food availability) may be required to solve the problem.

▼ Importing Modules

```
import pandas as pd
import seaborn as sns
import matplotlib.pyplot as plt
import numpy as np
```

▼ 1. Dataset has been downloaded

```
#Name of the dataset: abalone.csv
```

▼ 2. Load the dataset into the tool

```
data=pd.read_csv("abalone.csv")
data.head()
```

	Sex	Length	Diameter	Height	Whole weight	Shucked weight	Viscera weight	Shell weight	Rings
0	M	0.455	0.365	0.095	0.5140	0.2245	0.1010	0.150	15
1	M	0.350	0.265	0.090	0.2255	0.0995	0.0485	0.070	7
2	F	0.530	0.420	0.135	0.6770	0.2565	0.1415	0.210	9

▼ Let's know the shape of the data

```
data.shape
```

```
(4177, 9)
```

One additional task is that, we have to add the "Age" column using "Rings" data.

We just have to add '1.5' to the ring data

```
Age=1.5+data.Rings
data["Age"]=Age
data=data.rename(columns = {'Whole weight':'Whole_weight','Shucked weight': 'Shucked_weight',
                             'Shell weight': 'Shell_weight'})
data=data.drop(columns=["Rings"],axis=1)
data.head()
```

	Sex	Length	Diameter	Height	Whole_weight	Shucked_weight	Viscera_weight	Shell_w
0	M	0.455	0.365	0.095	0.5140	0.2245	0.1010	
1	M	0.350	0.265	0.090	0.2255	0.0995	0.0485	
2	F	0.530	0.420	0.135	0.6770	0.2565	0.1415	
3	M	0.440	0.365	0.125	0.5160	0.2155	0.1140	
4	I	0.330	0.255	0.080	0.2050	0.0895	0.0395	

3. Perform Below Visualizations.

▼ (i) Univariate Analysis

The term univariate analysis refers to the analysis of one variable. You can remember this because the prefix "uni" means "one." There are three common ways to perform univariate analysis

on one variable: 1. Summary statistics – Measures the center and spread of values.

###

► Histogram

[] ↳ 3 cells hidden

► Boxplot

[] ↳ 1 cell hidden

► Countplot

[] ↳ 1 cell hidden

► (ii) Bi-Variate Analysis

[] ↳ 11 cells hidden

► (iii) Multi-Variate Analysis

[] ↳ 4 cells hidden

► 4. Perform descriptive statistics on the dataset

[] ↳ 1 cell hidden

► 5. Check for Missing values and deal with them

[] ↳ 1 cell hidden

► 6. Find the outliers and replace them outliers

[] ↳ 3 cells hidden

► 7. Check for Categorical columns and perform encoding

[] ↳ 2 cells hidden

▶ 8. Split the data into dependent and independent variables

[] ↳ 2 cells hidden

▶ 9. Scale the independent variables

[] ↳ 1 cell hidden

▶ 10. Split the data into training and testing

[] ↳ 7 cells hidden

▶ 11. Build the Model

[] ↳ 4 cells hidden

▶ 12. Train the Model

[] ↳ 2 cells hidden

▶ 13. Test the Model

[] ↳ 1 cell hidden

▶ 14. Measure the performance using Metrics

[] ↳ 2 cells hidden

[Colab paid products](#) - [Cancel contracts here](#)

