**EXPLORATORY DATA ANALYSIS OF PENGUIN DATA SET**

Penguins are **birds**, although they are flightless birds. Lots of people think penguins are mammals rather than birds because they can't fly, and we see them swimming underwater or waddling on land instead.

**DATA SET**

**THE DATA TAKEN IS ABOUT THE PENGUINE WHICH STAY NEAR ANTARTIC CIRCLE OR NEAR COLD REGION**

**THE GIVEN DATA SET TELL US ABOUT DIFFERENT TYPES OF SPECIES OF PENGUIN WHICH LIVES IN DIFFERENT PART OF ISLANDS AND WE HAVE DONE EDA ON THIS SO THAT WE CAN DIFFERENTIATE BETWEEN THESE SPECIES TO HAVE A GOOD INSIGHT OF PARAMETERS AND OTHER INFORMATION REGARDING PENGUINE**

**THIS KIND OF ANALYSIS IS DONE USING THE DATA BY GETTING THE FACTORS AS FOLLOWS**

1. **CULMEN LENGTH**
2. **CULMEN DEPTH**
3. **FLIPPER LENGTH**
4. **BODY MASS**

**TOOLS USED**

**PYTHON | JUYPYTER NOTEBOOK | NUMPY | PANDAS | SEABORN | MATPLOTLIB | PLOTLY |**

**DATA DESCRIPTION**

The dataset consists of 7 columns.

* **species**: penguin species (Chinstrap, Adelie, or Gentoo)
* **culmen\_length\_mm**: The culmen is "the upper ridge of a bird's beak"
* **culmen\_depth\_mm**: Same As Above
* **flipper\_length\_mm**: Penguins wings are called flippers. They are flat, thin, and broad with a long, tapered shape and a blunt, rounded tip
* **body\_mass\_g**: body mass (g)
* **island**: island name (Dream, Torgersen, or Biscoe) in the Palmer Archipelago (Antarctica)
* **sex**: penguin sex

**DATA CLEANING**

**THE CLEANING I MADE IN THIS ARE AS FOLLOWS**

* **REMOVED THE DUPLICATES FROM THE DATA SET**
* **REMOVED UNWANTED COLUMN ACCORDINGLY TO HAVE SPECIFIC DATA**
* **CONVERTED SPECIES AND ISLAND COLUMN FOR NUMERICAL REPRESENTATION**

**EDA ON DATA**

**I HAVE SEEN THE DIFFERENT TYPE OF TRENDS OF DATA AND FOLLOWING ARE THE ANALYISIS I HAVED MADE**

* **BAR CHART TO GET THE NO OF SPECIES**
* **DISTRIBUTION OF SPECIES REGARDING TO THEIR SPECIFICATION**
* **COMPARISION BETWEEN THE PARAMETER OF PENGUIN**
* **FLIPPER LENGTH DISTRIBUTION**
* **KERNAL DISTRBUTION ESTIMATATION OF SPECIES**
* **PIE CHART FOR NO. OF SPECIES**