**Name: Omkar Shirwadkar**
**UID: 2021600062**
**Batch: D**
**Class: CSE(AI & ML)**

# ADV LAB ESE
# Create Basic and Advanced Charts using R programming language on Animal / Wildlife / Marine dataset

**Aim:**

**To create Basic and Advanced Charts using R programming language on Animal / Wildlife / Marine dataset**

**Dataset:**
https://www.kaggle.com/code/parulpandey/penguin-dataset-the-new-iris

1. **Program and Code:**

### Actual Code

```
install.packages("palmerpenguins")
install.packages("ggplot2")
install.packages("tidyverse")
install.packages("wordcloud")

library(palmerpenguins)
library(ggplot2)
library(tidyverse)
library(wordcloud)

data("penguins")

ggplot(data = penguins, aes(x = species, y = body_mass_g, fill = species)) +
  geom_violin(trim = FALSE) +
  labs(title = "Violin Plot of Body Mass by Species",
       x = "Species",
       y = "Body Mass (g)") +
```

```r
  theme_minimal()

species_count <- table(penguins$species)
wordcloud(names(species_count), species_count, colors = brewer.pal(3, "Dark2"))

island_count <- penguins %>% count(island)

ggplot(island_count, aes(x = "", y = n, fill = island)) +
  geom_bar(stat = "identity", width = 1) +
  coord_polar("y") +
  labs(title = "Penguin Distribution by Island",
       fill = "Island") +
  theme_void()

ggplot(data = penguins, aes(x = species, y = bill_length_mm, fill = species)) +
  geom_boxplot() +
  labs(title = "Box Plot of Bill Length by Species",
       x = "Species",
       y = "Bill Length (mm)") +
  theme_minimal()

ggplot(data = penguins, aes(x = flipper_length_mm, fill = species)) +
  geom_histogram(position = "dodge", bins = 20, alpha = 0.7) +
  labs(title = "Histogram of Flipper Length by Species",
       x = "Flipper Length (mm)",
       y = "Count") +
  theme_minimal()

ggplot(data = penguins, aes(x = bill_length_mm, y = bill_depth_mm, color = species)) +
  geom_point() +
  labs(title = "Scatter Plot of Bill Length vs. Bill Depth by Species",
       x = "Bill Length (mm)",
       y = "Bill Depth (mm)") +
  theme_minimal()
```
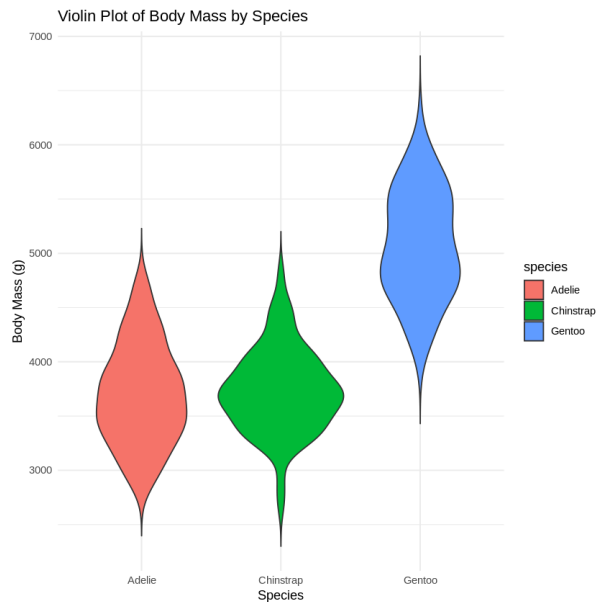
**Visualization and Observations:**

**1. Violin Plot:**



Violin Plot of Body Mass by Species

The violin plot shows that the distribution of body mass varies significantly among penguin species, with the Gentoo species typically having a higher body mass than Adelie and Chinstrap.
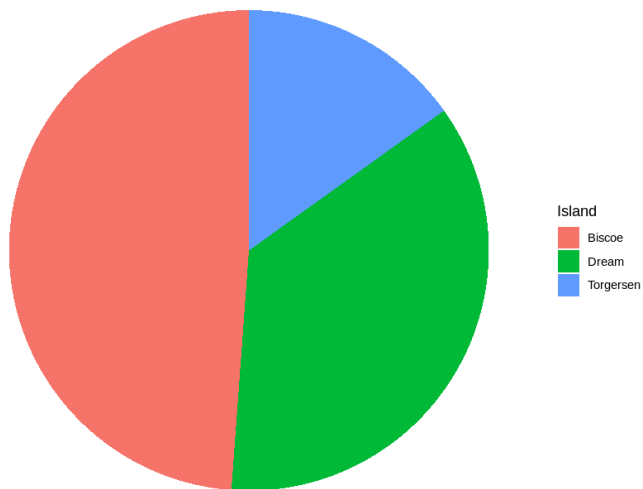
**2. Word Cloud:**



This word cloud displays the species names with font sizes reflecting their frequency in the dataset, showing which species have more data points.
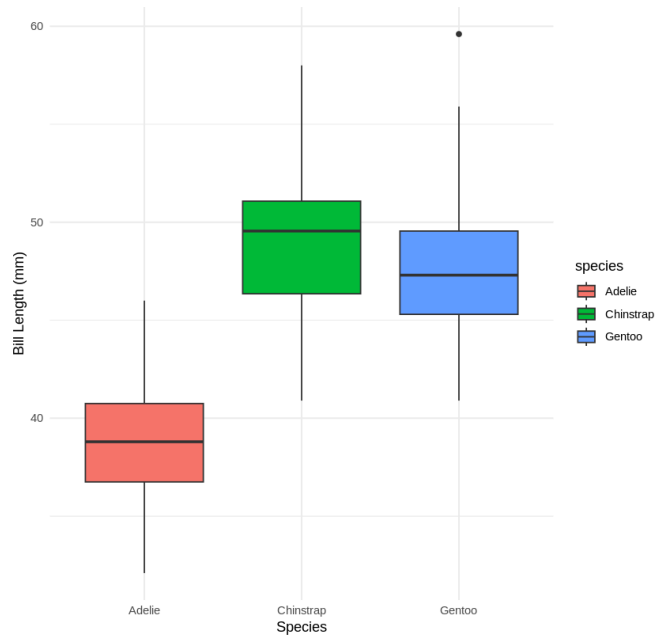
### 3.  Pie Chart:

Penguin Distribution by Island



This pie chart indicates that the majority of penguins are from Biscoe Island, followed by Dream and Torgersen Islands.
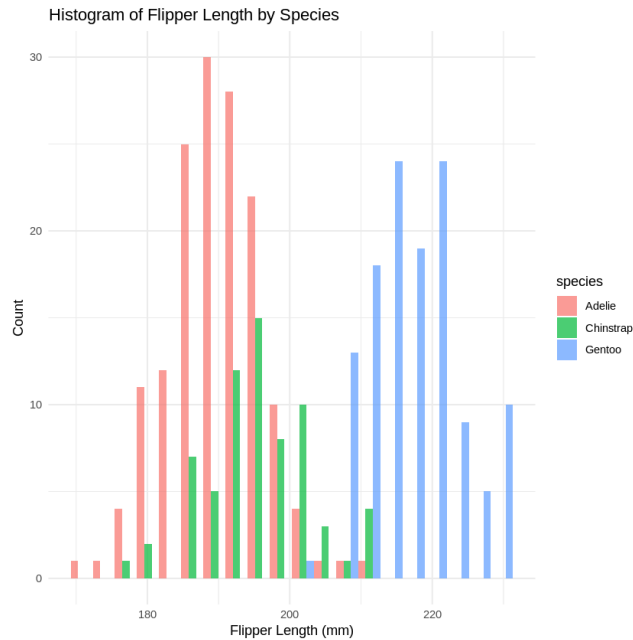
### 4.  Box or Whisker Plot:

Box Plot of Bill Length by Species



The box plot reveals that Gentoo penguins tend to have longer bills, while Adelie penguins have shorter bills. Chinstrap penguins have a moderately long bill, but with less variability.

## 5. Histogram:



Histogram of Flipper Length by Species

The histogram shows that Gentoo penguins have longer flipper lengths, while Adelie and Chinstrap penguins have shorter flippers, with overlapping distributions.

## 6. Scatter Plot:



Scatter Plot of Bill Length vs. Bill Depth by Species

The scatter plot indicates that there is a noticeable separation in bill length and depth among species. Gentoo penguins tend to have longer and shallower bills, while Adelie penguins have shorter, deeper bills.

**Conclusion:**
1. The penguin dataset reveals distinct physical differences across species, with Gentoo penguins generally having greater body mass, longer flippers, and unique bill characteristics compared to Adelie and Chinstrap penguins.
2. Most penguins are found on Biscoe Island, suggesting it may offer a favorable habitat.
3. Variations in bill length and depth among species may indicate adaptations to different diets.
4. These findings underscore the role of morphology and distribution in species adaptation.