Bharatiya Vidya Bhavan's
Sardar Patel Institute of Technology
(Autonomous Institute Affiliated to University of Mumbai)
Bhavan's Campus, Munshi Nagar, Andheri (West), Mumbai-400058-India

Name: Varada Khadake **Roll No. :2021300059**

AIM:

Design Interactive Dashboards and Storytelling using Tableau / Power BI / R (Shiny) / Python (Streamlit/Flask) / D3.js to be performed on the dataset - Disease spread / Healthcare

- Create interactive dashboard - Write observations from each chart given below ● (Advanced - Word chart, Box and whisker plot, Violin plot, Regression plot (linear and nonlinear), 3D chart, Jitter, Line, Area, Waterfall, Donut, Treemap, Funnel ● Basic - Bar chart, Pie chart, Histogram, Timeline chart, Scatter plot, Bubble plot)

DATASET:

This dataset provides comprehensive insights into the distribution and potential spread of avian influenza, commonly known as "Bird Flu," in Ireland. Avian influenza is a highly contagious and often fatal viral disease primarily affecting birds, with wild migratory water birds being the main reservoir of the virus.

The dataset includes information about bird species captured in Ireland from 1980 to 2020, focusing on species targeted for the H5N1 strain of avian flu. Understanding the geographic distribution of these bird species is crucial for assessing the risk of avian influenza introduction into Ireland, especially during migratory seasons when wild birds arrive and congregate on wetlands, potentially mixing with resident species.

This dataset serves as a valuable resource for conducting comprehensive research and risk assessment related to avian influenza in Ireland.

Columns:

_id Scientific_Name Common_Name Date Year Month Day Time Country Country_State_County State County Locality Latitude Longitude Parent_Species target_H5_HPAI

Tables and respective columns:

To design a **star schema** for the avian influenza dataset, we will divide the information into a **fact table** and **dimension tables**. The fact table stores the measurements or events, and the dimension tables store descriptive information.

Fact Table: **Bird_Flu_Fact**

Column Name	Description
_id	Unique identifier for each observation
Scientific_Name	Foreign key to the Bird Species Dimension
Common_Name	Foreign key to the Bird Species Dimension
Date	Date of the bird sighting (split into Year, Month, Day)
Time	Time of the observation
Country	Foreign key to the Location Dimension
Country_State_County	Foreign key to the Location Dimension
Latitude	Latitude of the bird sighting
Longitude	Longitude of the bird sighting
Parent_Species	Foreign key to the Bird Species Dimension
target_H5_HPAI	Indicator of targeted H5N1 strain presence

Dimension Tables

1. **Bird_Species_Dimension**

Column Name	Description
Scientific_Name	Unique scientific name of the bird species
Common_Name	Common name of the bird species
Parent_Species	Parent species group of the bird

2. **Date_Dimension**

Column Name	Description
Date	Full date (YYYY-MM-DD)
Year	Year of observation
Month	Month of observation
Day	Day of observation

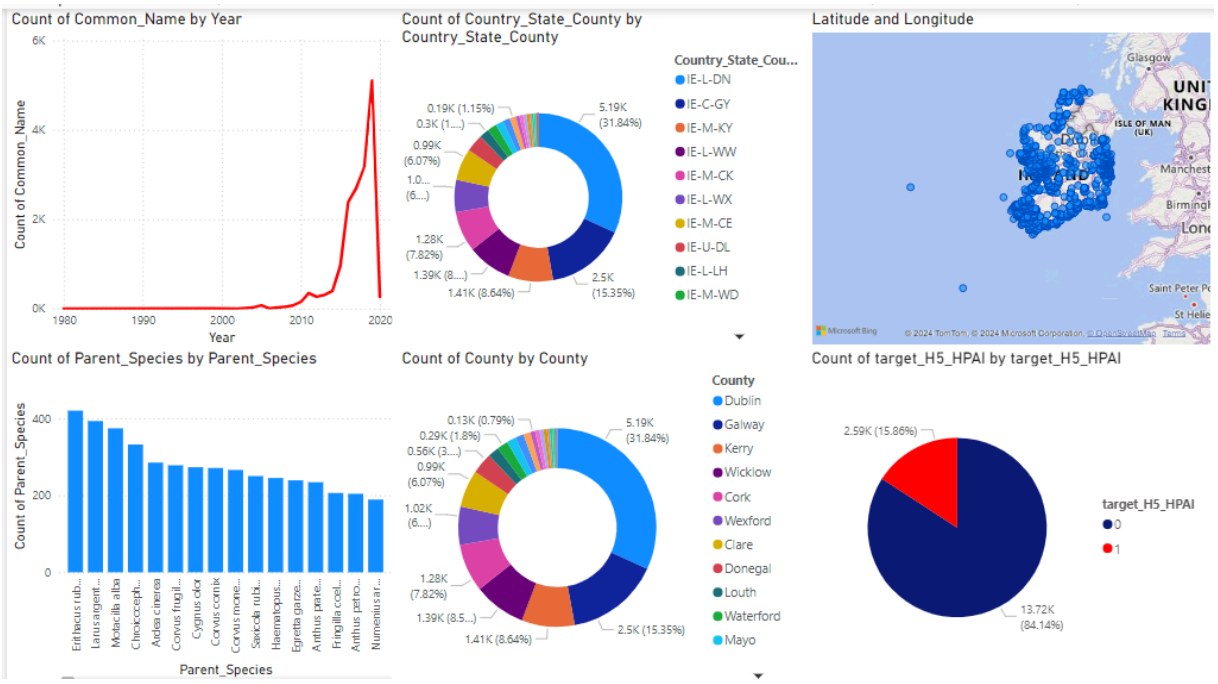
3. **Location_Dimension**

Column Name	Description
Country	Country of observation
State	State where observation was made
County	County where observation was made
Country_State_County	Combined Country, State, and County identifier
Locality	Specific locality of the bird sighting
Latitude	Latitude of the locality
Longitude	Longitude of the locality

Star Schema Structure

- The **Bird_Flu_Fact** table contains the main data points (observations), which refer to the relevant information in the dimension tables.
- The **Bird_Species_Dimension**, **Date_Dimension**, and **Location_Dimension** provide additional descriptive details that relate to the observations in the fact table.

DASHBOARD:



This dashboard provides a comprehensive visual representation of data related to bird species and their possible association with avian influenza (H5N1) cases, focusing on geographic and temporal trends. Here's a breakdown of each visualization:

1. Count of Common Name by Year (Top Left)

- **Description**: This line graph shows the count of bird species (Common_Name) observed over the years from 1980 to 2020.
- **Insight**: There is a sharp increase in bird observations starting around 2018, peaking in 2020. This could indicate heightened monitoring, an increase in the number of birds, or changes in migratory patterns.

2. Count of Country_State_County by Country_State_County (Top Center)

- **Description**: This pie chart breaks down bird sightings based on their geographic locations, with "Country_State_County" as the grouping factor.
- **Insight**: The largest portion of sightings comes from "IE-L-DN" (Dublin), followed by other counties such as Galway (IE-C-GY), and Kerry (IE-M-KY). This gives insight into where the majority of bird species are being observed in Ireland.

3. Latitude and Longitude (Top Right)

- **Description**: This map plots the bird sightings by their geographic coordinates (latitude and longitude) across the UK and Ireland.
- **Insight**: The majority of bird sightings are clustered in Ireland, with dense points indicating areas of frequent observation, particularly along the coast and major wetlands.

4. Count of Parent_Species by Parent_Species (Bottom Left)

- **Description**: A bar chart showing the count of bird observations categorized by their parent species.
- **Insight**: "Fulicarius", "Larus argentatus", and "Haematopus ostralegus" are the most commonly observed parent species, suggesting these species might be of significant interest, possibly due to their association with avian flu.

5. Count of County by County (Bottom Center)

- **Description**: This pie chart breaks down the number of bird sightings by different counties in Ireland.
- **Insight**: Dublin (Dublin) has the highest count, followed by Galway, Kerry, and other counties. The data helps in understanding which regions contribute most to bird observations and potentially where the bird flu monitoring is focused.

6. Count of target_H5_HPAI by target_H5_HPAI (Bottom Right)

- **Description**: A pie chart showing the distribution of the "target_H5_HPAI" variable, which likely indicates whether a bird species is targeted for H5N1 influenza screening.
- **Insight**: The majority (84.14%) of bird species are not targeted for H5N1, while 15.86% of the species have been identified as potential carriers or at risk for the H5N1 strain. This shows that only a subset of the birds is being monitored for avian flu.

Overall Insights:

- **Geographical Distribution**: The data highlights key counties in Ireland (Dublin, Galway, Kerry) where bird species are frequently observed, suggesting these regions are of interest for avian flu monitoring.
- **Temporal Trends**: A sharp rise in bird observations in 2020 might correlate with increased monitoring or changes in bird behavior.
- **Target Species**: Only a small percentage of species are targeted for H5N1 screening, indicating that monitoring efforts are selective and focused on specific species.