BONDA NAVEEN KUMAR

Bird Species Observation Analysis

The project analyses bird species distribution and diversity in two distinct habitats: forests and grasslands. Using observational data, the goal is to understand how environmental and spatial factors affect bird populations and behaviours. Insights derived support wildlife conservation, land management, eco-tourism, sustainable agriculture, and policymaking.

Multiple datasets representing forest and grassland observations were imported, each containing detailed records of bird sightings, environmental factors, and observation conditions.

Cleaning steps included:

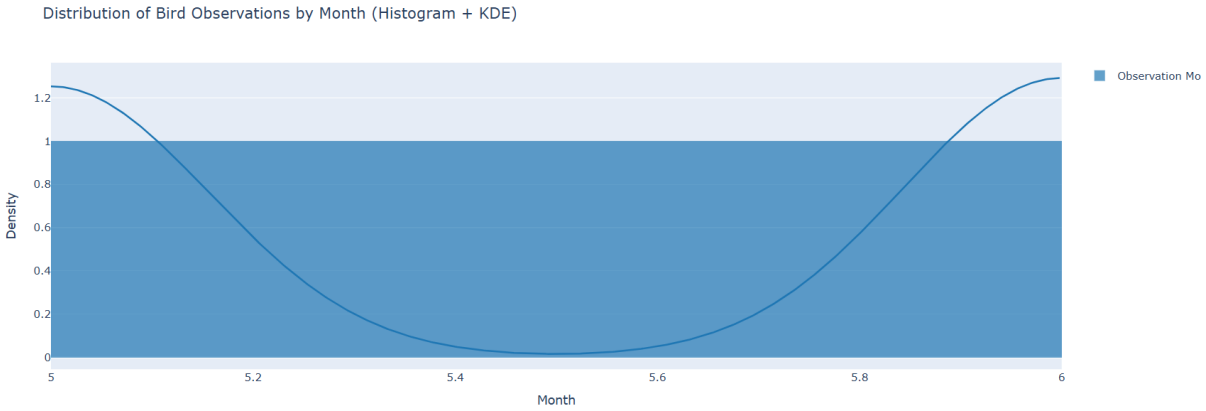
- Standardizing column names to lowercase with underscores for consistency.
- Removing duplicate rows to avoid data redundancy.
- Handling missing values by imputing median values for numerical columns and mode values for categorical columns.
- Converting date and time columns into proper datetime formats.
- Normalizing categorical variables by stripping whitespace and capitalizing first letters.
    - The cleaned datasets for forests and grasslands were saved separately for analysis.
    - Both datasets were then **combined into a single dataset** to facilitate unified analysis and comparative visualization.

BONDA NAVEEN KUMAR

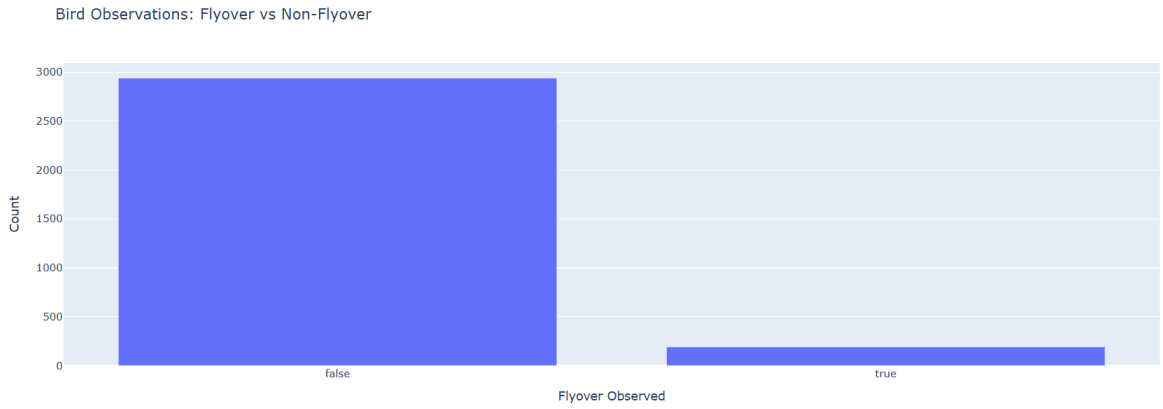
EDA :

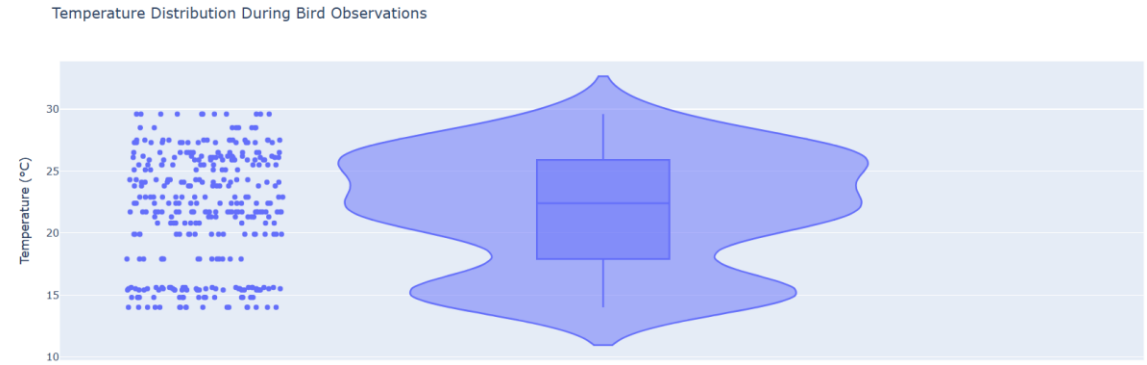
- **Species Distribution and Identification:**
  - Visualized counts of observations by identification methods to understand how birds were detected (e.g., Singing, Calling, Visual).
  - Analyzed temperature distributions overall and by identification method to explore environmental conditions during observations.
- **Habitat and Environmental Conditions:**
  - Examined temperature variations across habitats (forest vs grassland) and how they correlated with bird activity.
  - Created violin and box plots to observe the spread and median of environmental variables such as temperature, humidity, and distance of birds.
- **Behavioral and Spatial Patterns:**
  - Investigated flyover frequencies by habitat and species.
  - Explored interval lengths of observation times across habitat types to understand observation effort.
  - Conducted species count analysis and sex ratio studies to explore bird population dynamics.
- **Advanced Visualizations:**
  - Created parallel categories plots to reveal relationships among identification method, sex, habitat, and flyover observations.
  - Developed 3D scatter plots and correlation heatmaps combining environmental and behavioral data to uncover multi-dimensional patterns.

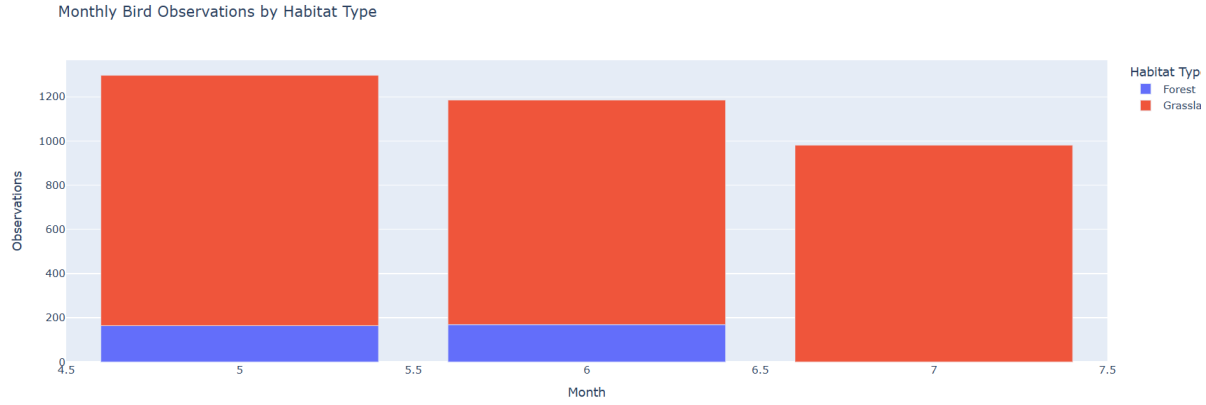
These analyses provided actionable insights into which species prefer certain habitats, how environmental factors influence bird activity, and the timing and nature of observations.

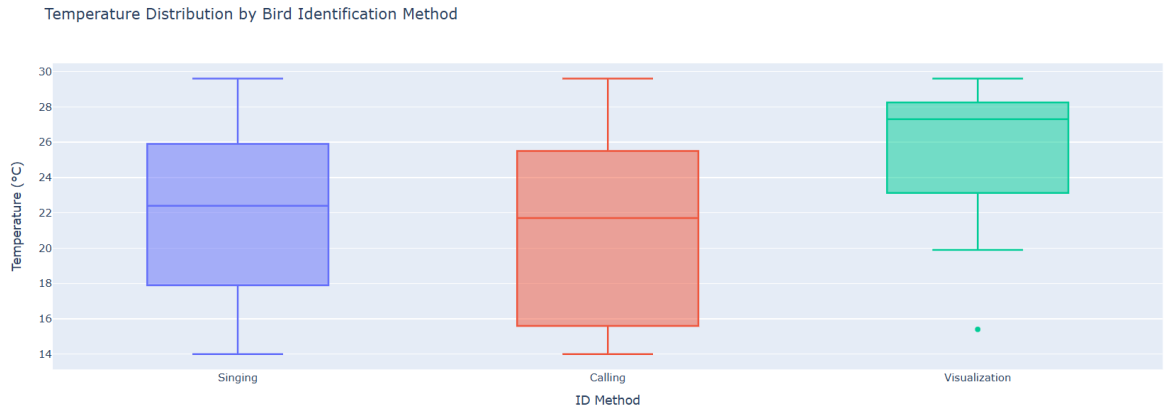

Distribution of Bird Observations by Month (Histogram + KDE)
Observation Mo
Density
1.2
1
0.8
0.6
0.4
0.2
0
5
5.2
5.4
5.6
5.8
6
Month

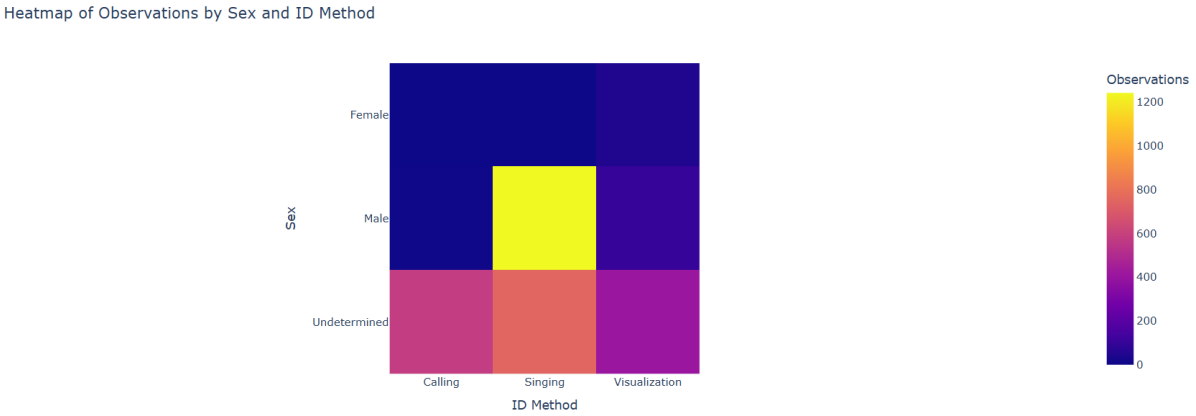

Count of Observations by Identification Method
count
200
150
100
50
0
Singing
Calling
Visualization
ID Method

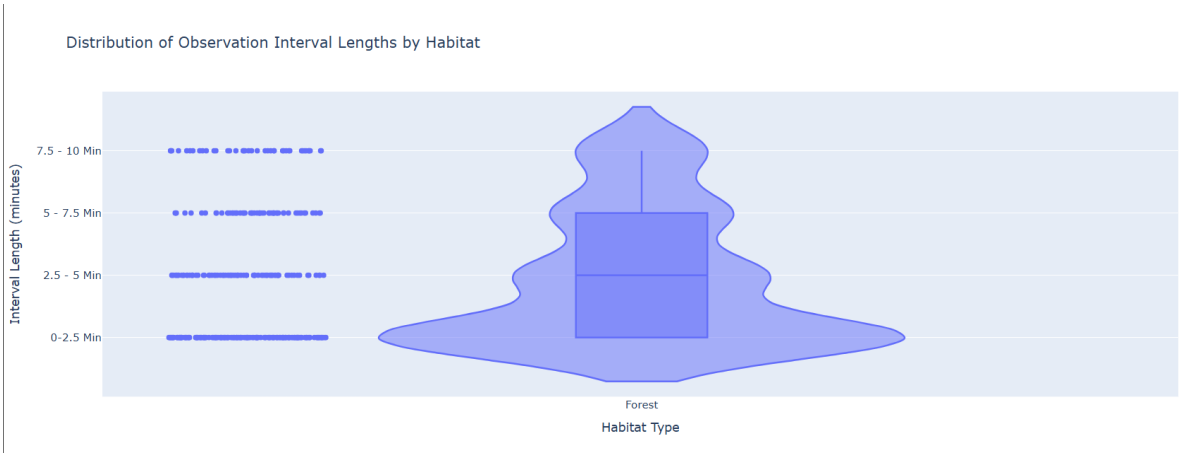

Bird Observations: Flyover vs Non-Flyover
Count
3000
2500
2000
1500
1000
500
0
false
true
Flyover Observed


Temperature Distribution During Bird Observations
Temperature (°C)
30
25
20
15
10


Monthly Bird Observations by Habitat Type
Habitat Typ
Forest
Grassla
Observations
0
200
400
600
800
1000
1200
4.5
5
5.5
6
6.5
7
7.5
Month

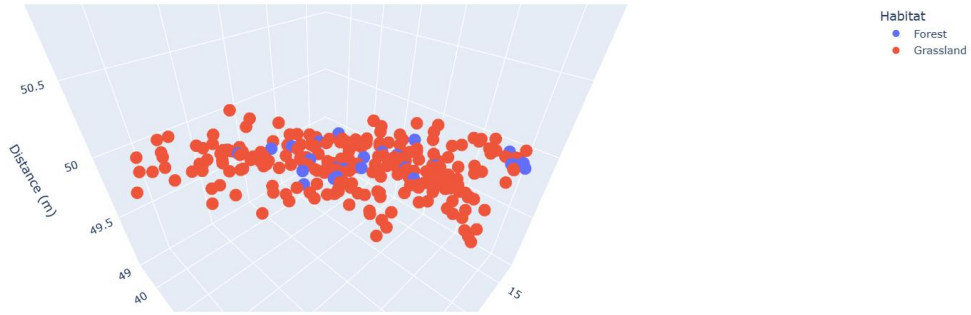

Temperature Distribution by Bird Identification Method
Temperature (°C)
14
16
18
20
22
24
26
28
30
Singing
Calling
Visualization
ID Method


Heatmap of Observations by Sex and ID Method
Observations
1200
1000
800
600
400
200
0
Female
Male
Undetermined
Sex
Calling
Singing
Visualization
ID Method

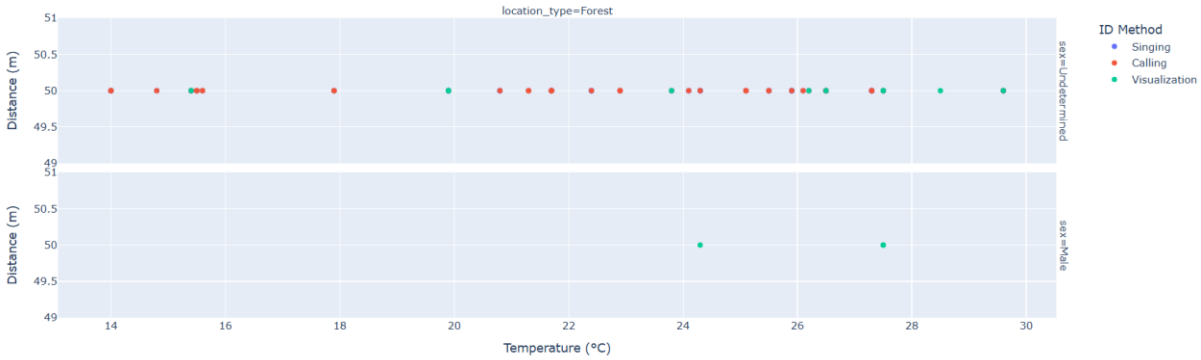

Distribution of Observation Interval Lengths by Habitat
Interval Length (minutes)
7.5 - 10 Min
5 - 7.5 Min
2.5 - 5 Min
0-2.5 Min
Forest
Habitat Type

BONDA NAVEEN KUMAR

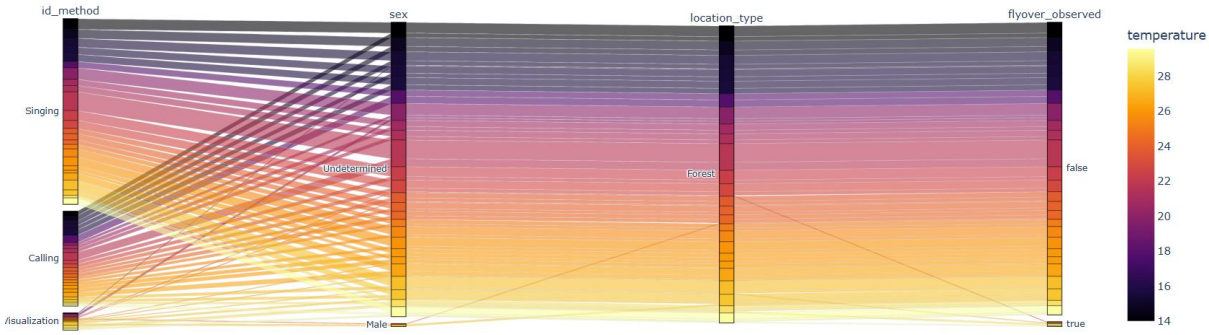
3D Scatter: Temperature, Humidity & Observation Distance by Habitat



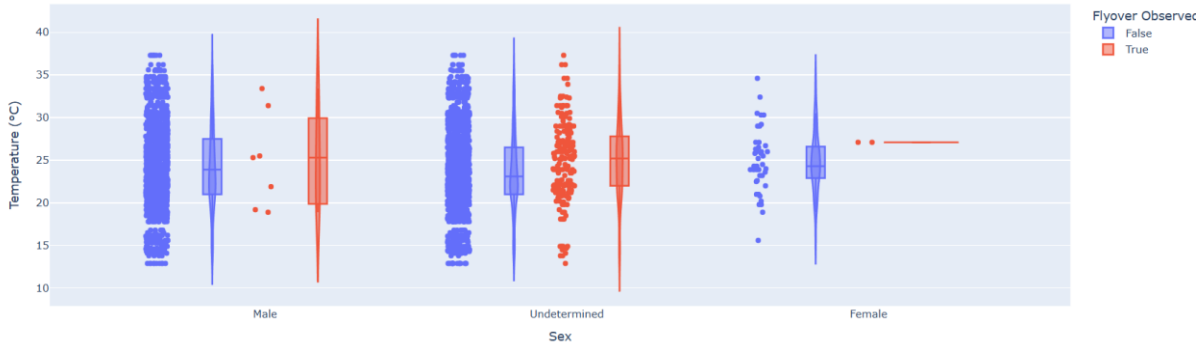
Temperature vs Distance Faceted by Habitat and Sex

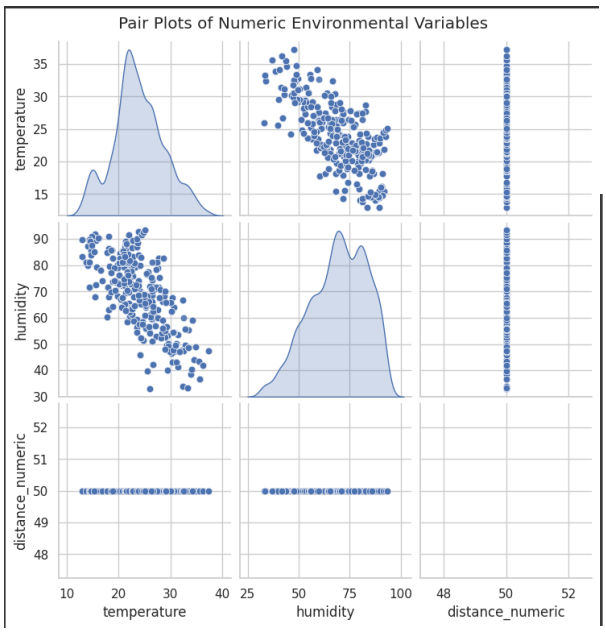


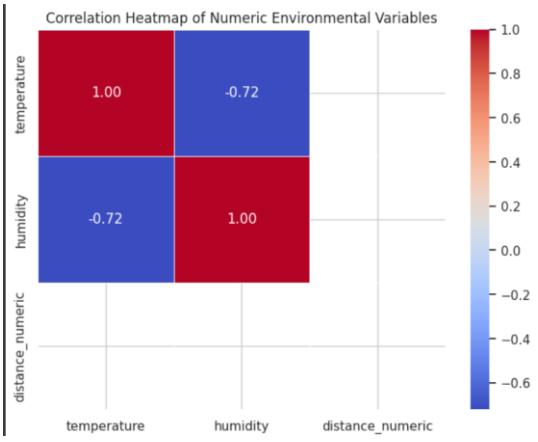
Parallel Categories: ID Method, Sex, Habitat & Flyover



Temperature Distribution by Sex and Flyover Observed







POWER BI :

To translate EDA insights into an interactive and stakeholder-friendly format, a Power BI dashboard was designed with the following steps:

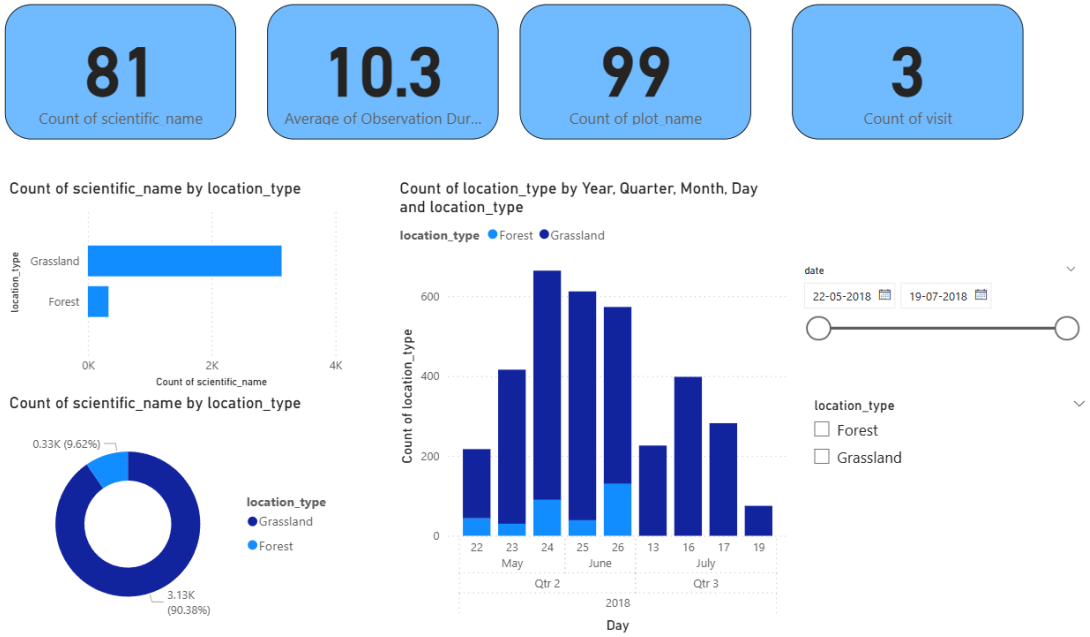
- Data Modeling:
  - Loaded the cleaned combined dataset.
  - Created calculated columns and DAX measures for key metrics like Observation Duration (time difference between Start_Time and End_Time), Flyover Percentage, Species Count, and bucketed Distance categories.
- Visualizations Designed:

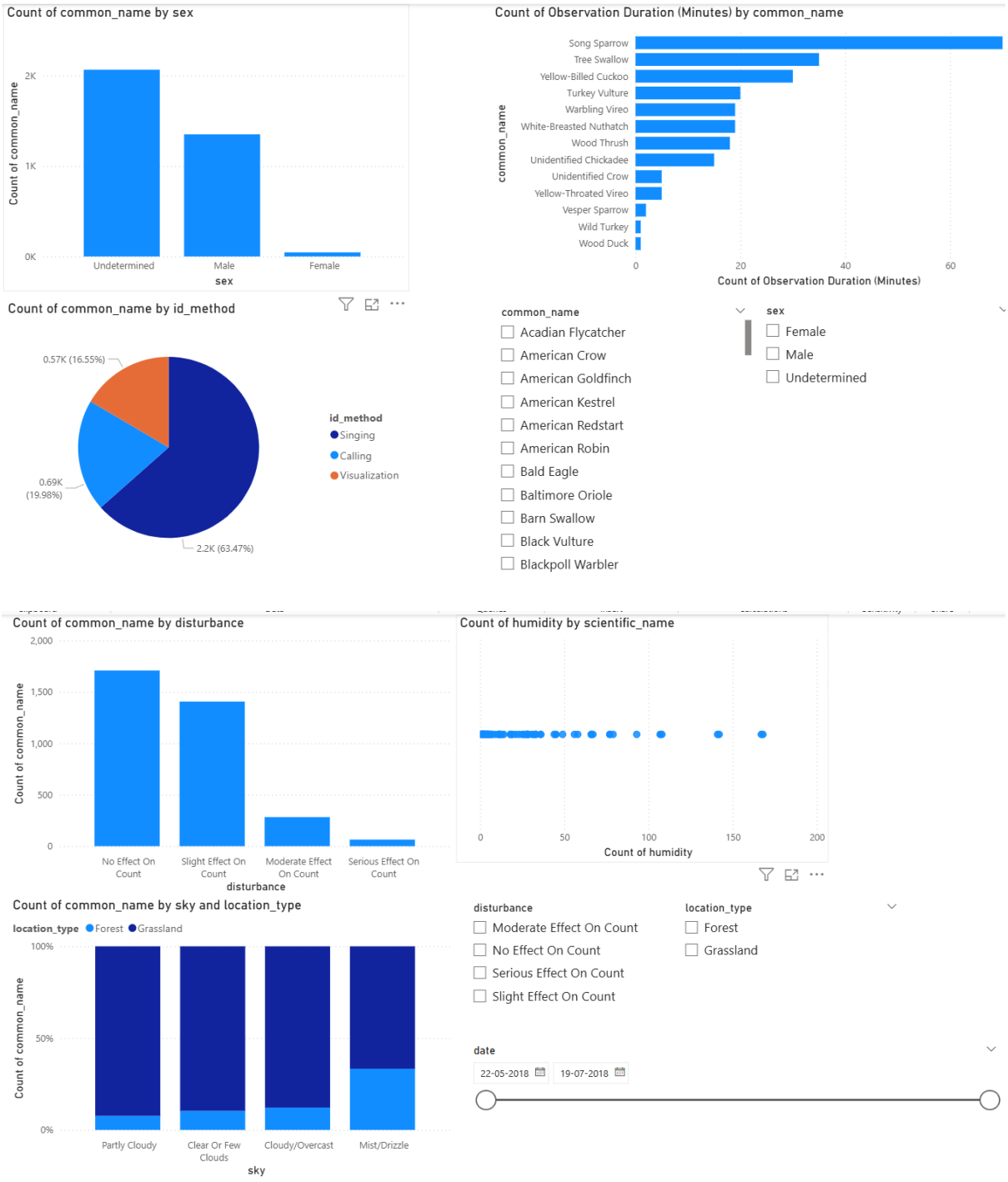
1. Observation Duration Analysis
   - Visualized observation duration across habitats and identification methods using box and violin plots.
2. Species Distribution by Habitat and Identification Method
   - Multi-level stacked bar charts to compare species counts by habitat and ID method.
3. Environmental Conditions Impact
   - Scatter plots and correlation matrices illustrating relationships between temperature, humidity, and bird distance.
4. Scatter Plot of Distance vs Flyover Frequency
   - Scatter plot with bucketed distance on the X-axis.

- Y-axis as percentage of Flyover Observed TRUE (calculated via DAX).
- Size representing Species Count.
- Color indicating Habitat (Location_Type).

5. Temporal and Spatial Trends
   - Heatmaps and line charts showing bird observation frequencies by year, month, and plot location.

- Interactivity and Filters:
  - Added slicers for Year, Location_Type, Species, and Identification Method.
  - Enabled cross-filtering between charts for deeper exploration.
- Insights Delivered:
  - Identified key hotspots of bird activity.
  - Highlighted environmental factors strongly influencing bird presence.
  - Provided clear visuals for temporal trends and behavioral patterns.





Key Findings and Business Insights

- Forest and grassland habitats host distinct bird communities influenced by temperature, distance, and observation methods.
- Flyover behavior varies with distance and habitat type, aiding understanding of bird movement patterns.
- Observation durations and identification methods impact data completeness, emphasizing the need for standardized protocols.

- Temporal trends reveal peak seasons for bird activity, important for conservation planning and eco-tourism scheduling.
- Data-driven identification of biodiversity hotspots supports targeted habitat protection and restoration.
- Insights can guide sustainable agricultural practices by minimizing disruption to sensitive bird populations.

This project successfully integrated multi-habitat observational data, cleaned and combined diverse datasets, performed comprehensive exploratory analyses, and developed a robust Power BI dashboard for interactive visualization. The results provide actionable insights for wildlife conservation, land management, eco-tourism, and policy support. The dashboards empower stakeholders to make informed decisions and track bird population health over time, aiding biodiversity monitoring and ecosystem sustainability.