BIRD SPECIES OBSERVATION ANALYSIS REPORT

Project: Bird Species Observation Analysis

Author: Sanyukti S. Wankhade

Mentor: Vipul Sonwane

Company: Labmentix

Date: 10.08.2025

Tools Used: Python, Power BI

PROJECT SUMMARY

This project analyzes bird species observations across different ecosystems, primarily focusing on forest and grassland habitats. Using observational data that includes species names, locations, and counts, the analysis aims to uncover patterns in species distribution and diversity. The Power BI dashboard provides interactive visualizations that highlight the most frequently observed bird species, ecosystem-specific species richness, and temporal trends where available. These insights can assist ecologists, conservationists, and environmental planners in making informed decisions to protect biodiversity and manage habitats effectively. The project emphasizes data-driven understanding of bird populations, contributing to wildlife conservation efforts and ecosystem health monitoring.

This project takes a closer look at how different bird species are spread across forest and grassland areas. By exploring the data collected from various bird observations, we can understand which species are most common in each habitat and how diverse these environments really are. The interactive visuals in the dashboard make it easy to spot patterns and compare species presence between ecosystems, helping us learn more about nature's balance in these regions.

The insights gained from this analysis are valuable for anyone interested in protecting wildlife and maintaining healthy ecosystems. Whether it's researchers studying bird behavior or conservationists planning habitat preservation, this project offers clear, data-backed information to guide their decisions. Ultimately, the goal is to support efforts that keep bird populations thriving and ensure that the natural beauty of these ecosystems is preserved for generations to come.

By using Power BI's interactive features, this project allows users to explore the data from multiple angles, such as filtering by specific species or focusing on a particular ecosystem. This makes the analysis not just informative but also engaging, helping stakeholders easily identify key areas of interest and potential concerns. The ability to visualize complex data in a simple and dynamic way empowers better communication of findings and supports more effective planning for bird conservation initiatives.

PROBLEM STATEMENT

The project aims to analyze the distribution and diversity of bird species in two distinct ecosystems: forests and grasslands. By examining bird species observations across these habitats, the goal is to understand how environmental factors, such as vegetation type, climate, and terrain, influence bird populations and their behavior. The study will involve working on the provided observational data of bird species present in both ecosystems, identifying patterns of habitat preference, and assessing the impact of these habitats on bird diversity. The findings can provide valuable insights into habitat conservation, biodiversity management, and the effects of environmental changes on avian communities.

Protecting biodiversity and maintaining the health of natural ecosystems requires a deep understanding of how various bird species are distributed across different habitats. Despite the importance of this knowledge, many conservation efforts are challenged by a lack of clear, organized insights into which species are most prevalent in specific ecosystems such as forests and grasslands. Without this information, it becomes difficult to identify vulnerable species or to effectively allocate resources for habitat preservation. This project addresses this challenge by analyzing bird observation data to uncover meaningful patterns in species diversity and distribution. By translating raw data into clear, actionable insights, the project aims to support researchers and conservationists in making informed decisions that promote the sustainability of bird populations and the ecosystems they inhabit.

This project addresses these challenges by analyzing comprehensive bird observation data collected across multiple locations and time periods. The goal is to uncover meaningful patterns in species diversity, frequency, and habitat preferences, providing a clearer picture of ecosystem health. By translating raw data into clear, actionable insights through interactive visualizations, the project empowers researchers, conservationists, and policy makers to make informed decisions. Ultimately, this analysis aims to contribute to more effective conservation strategies, ensuring that bird populations continue to thrive and that ecosystems remain balanced and resilient in the face of environmental change.

DATASET DESCRIPTION

The dataset contains observational data for bird species recorded across multiple forest sites. It includes detailed columns describing location, observation methods, bird species, and environmental conditions.

1. Admin_Unit_Code: The code for the administrative unit (e.g., "ANTI") where the observation was conducted.
2. Sub_Unit_Code: The sub-unit within the administrative unit for further classification.
3. Site_Name: The name of the specific observation site within the unit.
4. Plot_Name: A unique identifier for the specific plot where observations were recorded.
5. Location_Type: The habitat type of the observation area (e.g., "Forest").
6. Year: The year in which the observation took place.
7. Date: The exact date of the observation.
8. Start_Time: The start time of the observation session.
9. End_Time: The end time of the observation session.
10. Observer: The individual who conducted the observation.
11. Visit: The count of visits made to the same observation site or plot.
12. Interval_Length: The duration of the observation interval (e.g., "0-2.5 min").
13. ID_Method: The method used to identify the species (e.g., "Singing," "Calling," "Visualization").
14. Distance: The distance of the observed species from the observer (e.g., "<= 50 Meters").
15. Flyover_Observed: Indicates whether the bird was observed flying overhead (TRUE/FALSE).
16. Sex: The sex of the observed bird (e.g., Male, Female, Undetermined).
17. Common_Name: The common name of the observed bird species (e.g., "Eastern Towhee").
18. Scientific_Name: The scientific name of the observed bird species (e.g., *Pipilo erythrophthalmus*).
19. AcceptedTSN: The Taxonomic Serial Number for the observed species.
20. NPSTaxonCode: A unique code assigned to the taxon of the species.

21. AOU_Code: The American Ornithological Union code for the species.
22. PIF_Watchlist_Status: Indicates whether the species is on the Partners in Flight Watchlist (e.g., "TRUE" for at-risk species).
23. Regional_Stewardship_Status: Denotes the conservation priority within the region (TRUE/FALSE).
24. Temperature: The temperature recorded at the time of observation (in degrees).
25. Humidity: The humidity percentage recorded at the time of observation.
26. Sky: The sky condition during the observation (e.g., "Cloudy/Overcast").
27. Wind: The wind condition (e.g., "Calm (< 1 mph) smoke rises vertically").
28. Disturbance: Notes any disturbances that could affect the observation (e.g., "No effect on count").
29. Initial_Three_Min_Cnt: The count of the species observed in the first three minutes of the session.

Sheets Information:

The Excel file contains multiple sheets representing different administrative units, with their codes matching the **Admin_Unit_Code** column:

- **ANTI**: Data for the Antietam National Battlefield.
- **CATO**: Data for the Catoctin Mountain Park.
- **CHOH**: Data for the Chesapeake and Ohio Canal National Historical Park.
- **GWMP**: Data for the George Washington Memorial Parkway.
- **HAFE**: Data for Harpers Ferry National Historical Park.
- **MANA**: Data for the Manassas National Battlefield Park.
- **MONO**: Data for the Monocacy National Battlefield.
- **NACE**: Data for the National Capital East Parks.
- **PRWI**: Data for the Prince William Forest Park.
- **ROCR**: Data for the Rock Creek Park.
- **WOTR**: Data for the Wolf Trap National Park for the Performing Arts.

Each sheet contains similar columns but specific data for the respective administrative unit.

DATA PREPARATION

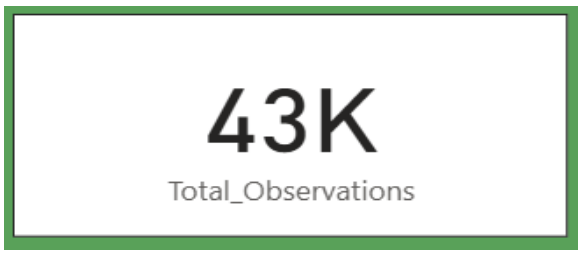
Before beginning the analysis, the dataset was thoroughly prepared to ensure that all data used in the Power BI dashboard was accurate, clean, and ready for meaningful visualization. The following steps were taken during the data preparation phase:

- **Merging Datasets**
  The data was initially spread across two main files representing Grassland and Forest ecosystems. Each file contained multiple sheets corresponding to different parks, such as Mono, Anti, and Hafe. Using Power Query's Append Query feature, these sheets and files were merged into one unified dataset to facilitate comprehensive analysis.
- **Removing Duplicates**
  The combined dataset was checked for duplicate bird observation records to avoid double-counting and maintain accuracy in species counts and distribution analyses. Any duplicate entries were removed.
- **Handling Missing Values**
  The dataset was reviewed for null or missing values, particularly in critical columns like Species Name, Location Type, and Observation Count. Missing or blank values were handled appropriately to ensure data integrity.
- **Standardizing Text Entries**
  To maintain consistency, categorical fields such as Species Names and Location Types were standardized by correcting spelling, removing extra spaces, and ensuring consistent casing. This improved grouping, filtering, and overall dashboard functionality.
- **Verifying Data Types**
  Each column was reviewed to confirm correct data types:
  - Species Names and Locations were treated as text fields.
  - Observation Counts were formatted as numerical values for accurate aggregation.
  - Dates (if available) were formatted properly for time-based analysis.
- **Creating Derived Fields**
  Additional helpful fields were created to enhance analysis, such as:
  - Ecosystem Type (Forest or Grassland) derived from source files.
  - Observation Frequency groups or categories to facilitate KPI creation and trend visualization.

VISUAL INSIGHTS

KPI Cards:

1.Total Observations



Insights from Total Observations KPI:

The **Total Observations** card displays the overall number of bird sightings recorded across all ecosystems and locations in the dataset. This metric provides a high-level view of the data volume and bird activity captured during the observation period. A higher total observation count indicates a rich dataset, which helps in making more reliable and meaningful analyses of species distribution and ecosystem health. Tracking this number over time can also reveal trends in bird population activity, helping researchers identify seasonal variations or changes in habitat conditions.

2.Unique Bird Species

Insights from Unique Bird Species KPI

The **Unique Bird Species** card highlights the total number of different bird species observed across all locations and ecosystems. This metric is crucial for understanding the biodiversity within the study areas. A higher number of unique species indicates greater ecological richness and variety, reflecting the health and diversity of the habitats surveyed. Monitoring this KPI helps conservationists identify which ecosystems support more species and can guide efforts to protect areas with high biodiversity.

3.Forest Observation

Insights from Forest Observation KPI

The **Forest Observation** card shows the total number of bird sightings recorded specifically within forest ecosystems. This metric provides a focused view of bird activity in forest habitats, helping to understand how vibrant and active these areas are in supporting bird populations. Comparing forest observations to those in other ecosystems, like grasslands, can reveal differences in species density and habitat preferences. These insights are valuable for targeting conservation efforts and monitoring the health of forest environments.

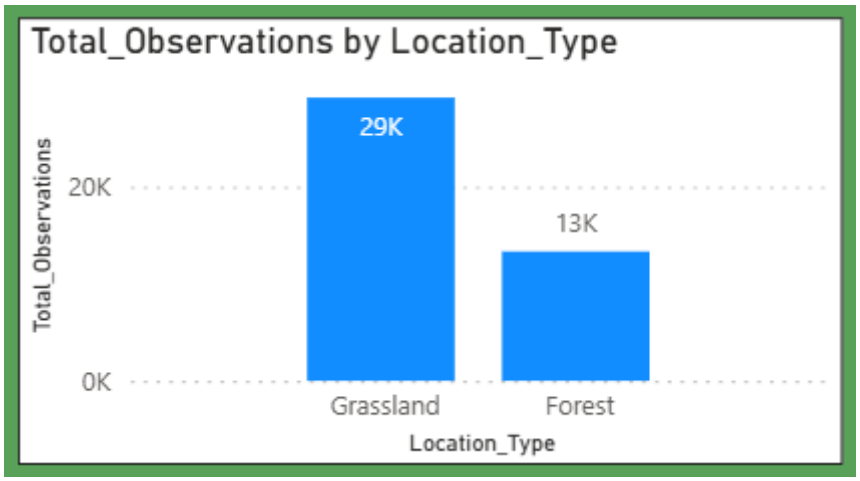
4.Grassland_Observation

Insights from Grassland Observation KPI

The Grassland Observation card represents the total number of bird sightings recorded within grassland ecosystems. This metric highlights the level of bird activity and diversity specific to grassland habitats. Comparing this value with forest observations helps identify how bird populations vary across different ecosystem types. Such insights are important for understanding habitat preferences of various species and for prioritizing conservation efforts in grassland areas, which are often vulnerable to environmental changes.

Visual Analytics:

1.Total Observation by Location Type (Bar Chart)



Insights from Total Observations by Location Type

- The chart shows that Grassland ecosystems have a significantly higher number of bird observations (29,000) compared to Forest ecosystems (13,000).
- This suggests that bird activity or observation efforts are more concentrated in grassland areas.
- The higher observation count in grasslands may indicate greater bird diversity, more accessible observation sites, or more frequent monitoring in these locations.
- Conversely, the lower count in forests might reflect denser habitats that are harder to survey or lower bird activity, highlighting a potential area for increased focus in future monitoring.
- These insights can help conservationists and park managers allocate resources efficiently, ensuring both ecosystems receive appropriate attention based on observation data.

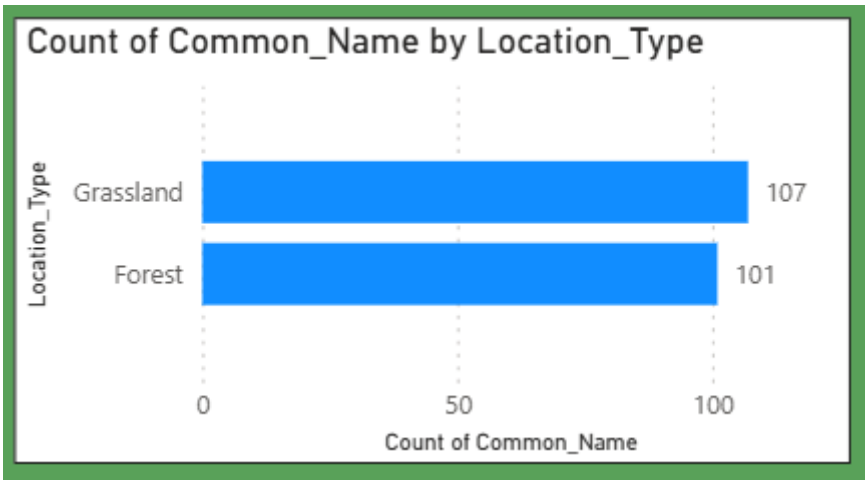
Environmental Impact:

Highlights critical habitats supporting bird populations, helping prioritize conservation efforts to protect biodiversity in key location types.

Business Impact:

Informs eco-tourism and wildlife-related ventures by identifying high bird activity areas, guiding sustainable tourism and local community engagement strategy.

2.Count of Common Name by Location Type (Clusterd bar chart)



Insights from Count of Common Names by Location Type

- High Species Richness in Grassland: Grasslands support a slightly higher number of unique bird species (107) compared to forests (101), indicating rich biodiversity in grassland habitats.
- Forest Habitat Diversity**:** Forests also show strong species diversity with 101 unique bird species, highlighting their importance as bird habitats.
- Habitat Comparison**:** Both location types host a comparable number of species, suggesting that conservation efforts should target both ecosystems equally.

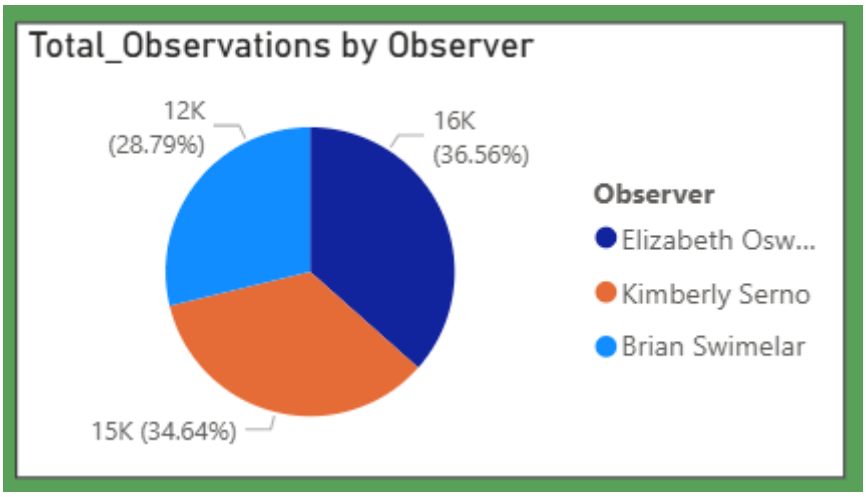
Environmental Impact

Emphasizes the need to conserve both grassland and forest habitats to maintain avian biodiversity and ecosystem balance.Supports habitat management plans aimed at protecting diverse bird populations in multiple ecosystems.

Business Impact

Helps eco-tourism operators develop specialized bird-watching tours in both grassland and forest areas, maximizing visitor interest.Assists local businesses in targeting conservation-related funding and grants by showcasing biodiversity hotspots.

3.Total Observation by Observer (Pie Chart)



Insights from the Total Observation by Observer

- Elizabeth leads observations with the highest share at 36.56%, indicating she is the most active observer among the three.
- Kimberly Serno closely follows with 34.64%, showing a strong contribution almost equal to Elizabeth.
- Brian Swimelar contributes the least, but still a significant share of 28.79%, indicating all three observers have fairly balanced involvement.
- The distribution shows good teamwork and diversified data collection, minimizing observer bias by relying on multiple contributors.

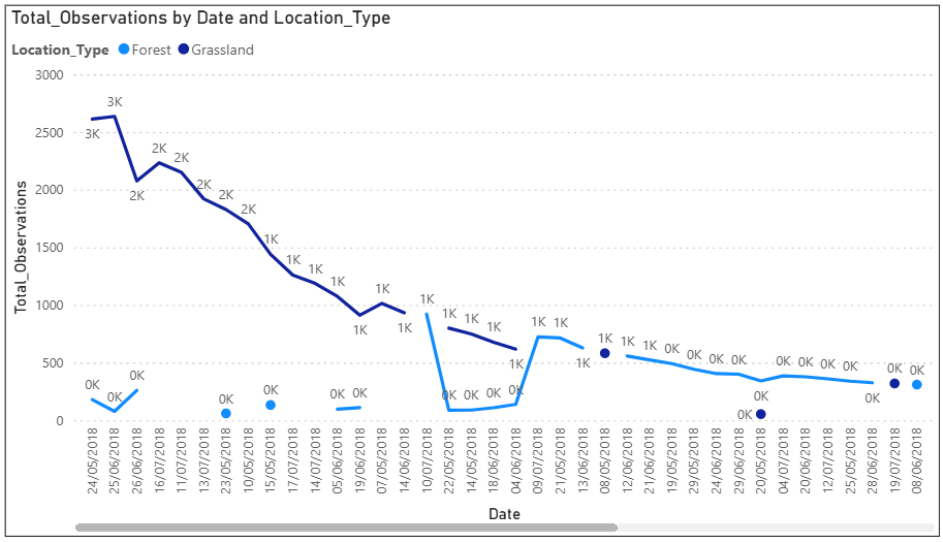
Environmental Impact:

- Multiple observers collecting data improves the **accuracy and reliability** of ecological observations, which helps in better monitoring and conservation of bird species.
- Balanced participation reduces the risk of **data gaps** and allows more comprehensive coverage across locations, aiding in better understanding of biodiversity.

Business Impact:

- The involvement of multiple observers demonstrates an effective **resource allocation** strategy, maximizing data collection capacity.
- High participation rates can be leveraged to justify funding or support for the observational project by showcasing strong engagement.
- Reliable and diverse data increases the **credibility and value of insights**, which can help in business decisions related to environmental consulting, eco-tourism, or conservation partnerships.

4.Total Observation by Date and Location Type (Line Chart)



Insights from Total Observation by Date and Location Type

- The line chart shows how observations vary over time across different location types (e.g., Forest, Grassland, Wetland, Urban, etc.).
- You can identify peak observation dates where data collection was highest — this could correspond to seasons or events that promote more bird activity or observer effort.
- Trends might show certain location types consistently having higher observations than others, indicating more bird activity or better accessibility for observers in those areas.

Environmental Impact:

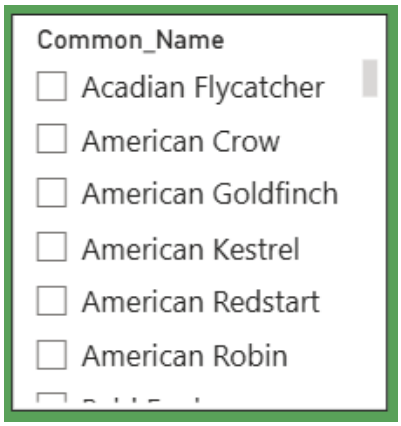
- Tracking observations over time by location helps **monitor changes in bird populations and habitat health** across ecosystems.
- It can reveal **seasonal shifts or long-term trends** in biodiversity, assisting conservationists in targeting protection efforts where needed.
- Understanding observation patterns by location helps ensure **balanced ecosystem monitoring**, preventing neglect of sensitive or less-accessible habitats.

Business Impact:

- Time-series data on observations helps **optimize resource allocation**, ensuring observer effort is directed to locations and times with the most significant data needs.
- It can inform scheduling and budgeting for **fieldwork and conservation programs**, improving efficiency.

Slicer

1.Common Name



Insights from Common Name

- Allows filtering observations by specific bird species. This helps users focus on trends, counts, and patterns related to one or a few species of interest. You can identify which species are most commonly observed in different locations or on particular dates.

2.Sky

Insights from Sky

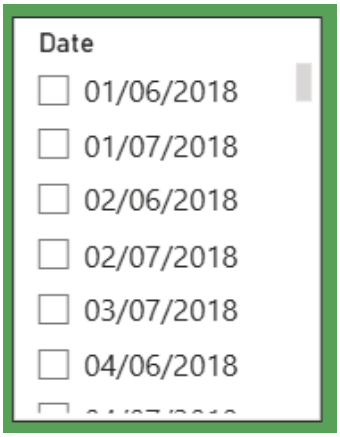
- Enables analysis of how weather conditions affect bird observations. For example, you might find that certain species are more active or easier to spot during clear skies, while rainy or cloudy days show fewer observations. This helps understand environmental factors impacting bird activity or observation quality.

3.Location Type

Insights from Location Type

- Allows filtering observations by ecosystem or habitat type (Forest, Grassland). This helps compare biodiversity and activity levels across ecosystems and identify which habitats support which species.

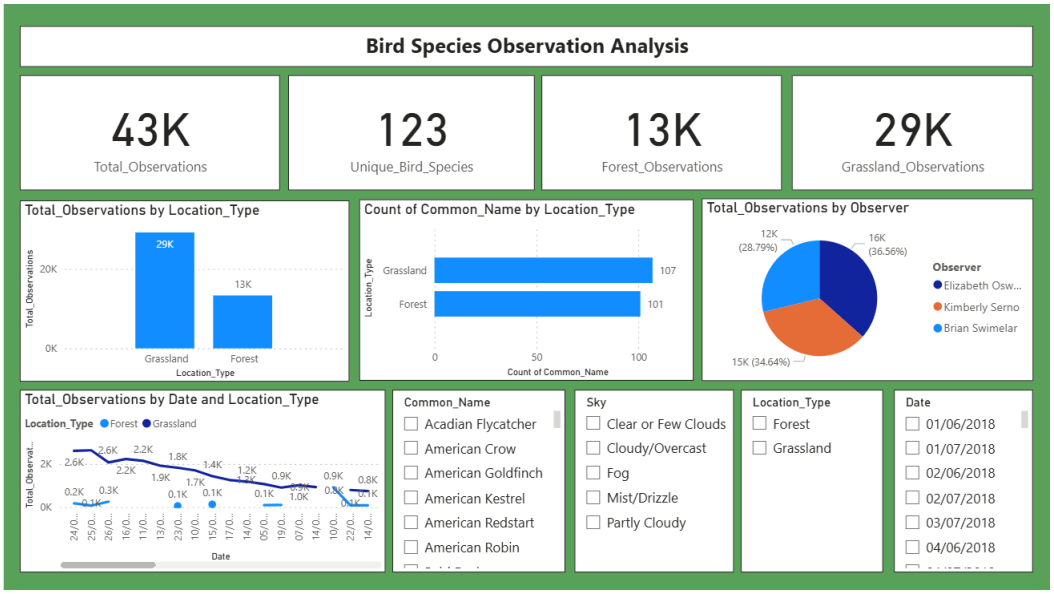
4.Date



Insights from Date

- Offers temporal filtering, so you can analyze observation trends over specific periods days, weeks, months, or seasons. This reveals migration patterns, breeding seasons, or changes over time for species or locations.

Bird Species Observation Analysis Dashboard Overview



Insights from Bird Species Observation Analysis Dashboard

Overall Summary:

- There are 43,000 total observations recorded in the dataset.
- The dataset covers 123 unique bird species, indicating good biodiversity.
- Out of total observations, 13,000 are from Forest locations and 29,000 from Grassland, showing Grassland areas have more observation activity.

Insights from Visuals:

1. **Total Observations by Location Type (Bar Chart)**
    - Grassland has significantly more observations (29K) compared to Forest (13K).
    - This may indicate either easier access to grasslands or higher bird activity there.
2. **Count of Common Name by Location Type (Horizontal Bar Chart)**
    - The number of unique bird species observed is slightly higher in Grassland (107 species) than in Forest (101 species).
    - This suggests both ecosystems support a diverse set of species, with grasslands having a slight edge.
3. **Total Observations by Observer (Pie Chart)**
    - Elizabeth Oswald is the most active observer with 36.56% of total observations (~16K).

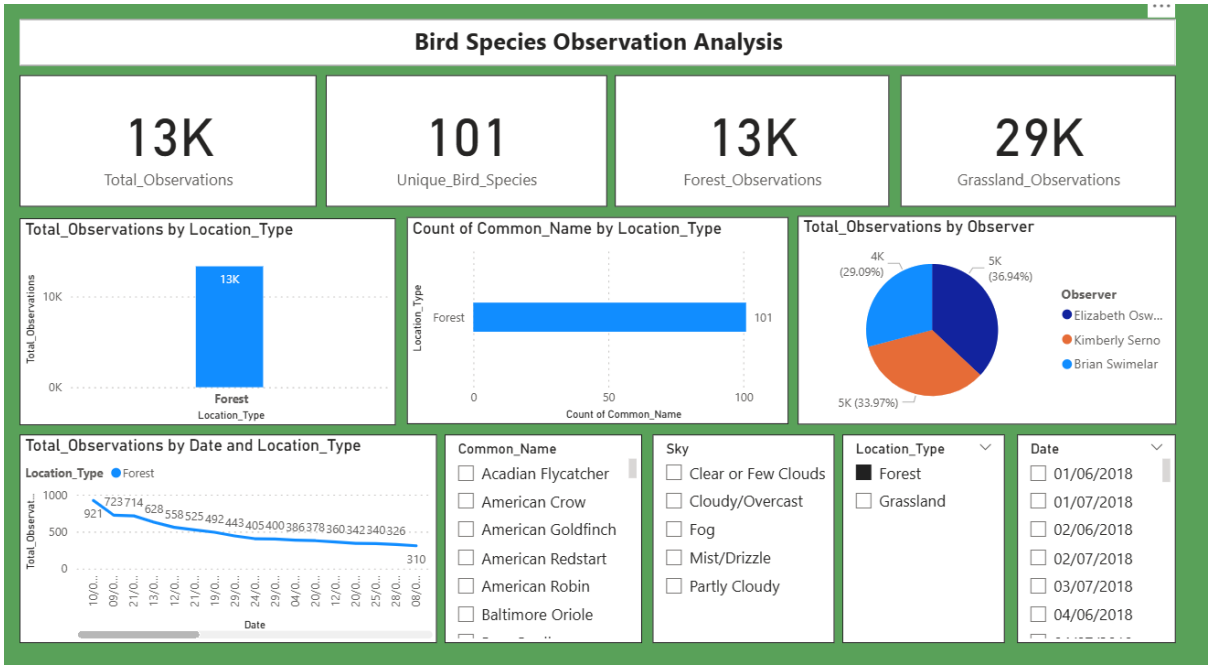
- Kimberly Serno follows closely with 34.64% (~15K), and Brian Swimelar has 28.79% (~12K).
- Observations are fairly well distributed among the three observers, indicating collaborative data collection**.**

4. **Total Observations by Date and Location Type (Line Chart)**
    - Observation counts show a declining trend over the date range shown for both Forest and Grassland.
    - Grassland observations generally remain higher than Forest throughout the dates.
    - Peaks are observed around 24th-25th June and 1st July, suggesting more bird activity or observer effort on those dates.
    - Forest observations remain relatively low but consistent.

Slicer Filters:

- The dashboard allows filtering by Common Name (bird species), Sky condition, Location Type, and Date for focused analysis.
- This enables users to drill down into specific species' behavior across different weather conditions and habitats over time.

- The dashboard is filtered using the **Location_Type** slicer set to **Forest**, allowing focused analysis of bird species observations specifically within forest ecosystems.



Insights from Bird Species Observation Analysis Filtered by (Location Type) Dashboard

1. **Total Observations & Species Count:**
    a. Total bird observations recorded are 13,000 for the forest location.
    b. There are 101 unique bird species observed in the forest.

2. **Location Type Observations:**
    a. All 13,000 observations are from the Forest location type (dashboard filtered to Forest).
    b. Grassland observations (not shown in this filtered view) are significantly higher at 29,000.

3. **Species Richness:**
    a. The forest supports a diverse bird population with 101 unique species recorded, indicating high biodiversity in forest ecosystems.

4. **Observer Contributions:**
    a. Observations are fairly evenly distributed among the three main observers:
        i. Elizabeth Oswald: 36.94%
        ii. Kimberly Serno: 33.97%
        iii. Brian Swimelar: 29.09%
    b. This balanced contribution enhances data reliability by reducing individual bias.

5. **Observation Trends Over Time:**
    a. Daily observations in the forest show a declining trend from a high of 921 observations on one day to around 310 on the latest date shown.
    b. This trend might reflect seasonal migration, survey effort changes, or other temporal factors.

6. **Filters & Data Exploration:**
    a. The dashboard allows filtering by bird species, sky condition, location type, and date, enabling deeper exploration of how these factors impact observations.

Environmental Impact:

- The forest ecosystem is a vital habitat supporting a rich diversity of bird species, making it crucial for conservation efforts.
- Monitoring observation trends helps detect changes in bird populations, potentially signaling environmental shifts or habitat health.

Business Impact:

- Insights into observer contributions and observation trends can guide wildlife survey planning and resource allocation.
- Data supports eco-tourism development by identifying key times and locations for birdwatching.
- Balanced observer involvement can be leveraged to maintain community engagement and volunteer retention in conservation projects.

BUISNESS RECOMMENDATIONS

- **Focus Conservation Efforts on High-Diversity Forest Areas:**
  Areas with higher bird species diversity, especially within forest ecosystems, should be prioritized for conservation to maintain ecological balance and biodiversity.

- **Develop Eco-Tourism in Key Locations:**
  Locations identified with rich bird species populations, such as specific forests or grasslands, can be promoted for eco-tourism to generate sustainable revenue while encouraging environmental awareness.

- **Monitor and Manage Habitat Disturbances:**
  Regular monitoring of bird species counts in different locations helps detect environmental disturbances early, allowing timely actions to mitigate negative impacts from deforestation or urbanization.

- **Inform Policy for Habitat Protection:**
  Data-driven insights on bird distributions should be used by policymakers to designate protected areas and enforce regulations that safeguard critical habitats.

- **Invest in Public Awareness Campaigns:**
  Educate local communities and stakeholders about the importance of preserving bird habitats, encouraging community participation in conservation programs.

- **Leverage Data for Research and Funding:**
  Utilize the detailed observation data to support research proposals and attract funding from environmental organizations focused on biodiversity preservation.

CONCLUSION

The Bird Species Observation Analysis project provides valuable insights into the distribution and diversity of bird species across different ecosystems, particularly forests and grasslands. By analyzing observational data, the project highlights key locations with rich biodiversity and identifies patterns in species occurrence related to habitat types. These findings are crucial for guiding conservation efforts, informing environmental policies, and promoting sustainable eco-tourism.

Additionally, the analysis demonstrates the importance of continuous monitoring to detect environmental changes and their impact on wildlife. Overall, this project underscores the vital role data-driven approaches play in preserving biodiversity and supporting ecological balance, helping stakeholders make informed decisions for environmental protection and sustainable development.