STAT679 - Group 2 **Project Milestone 2** **11/20/2023**

Public Code Link: https://github.com/cjones46/992GroupProject

Data Review

We obtained bird sighting data from the North American Bird Banding Project at https://www.sciencebase.gov/catalog/item/632b2d7bd34e71c6d67bc161, choosing the datasets corresponding to the American Goldfinch and Song Sparrow. These are common migratory birds in the US.

We obtained temperature time series data from the National Centers for Environmental Information at https://www.ncei.noaa.gov/access/monitoring/climate-at-a-glance/statewide/time-series/22/tavg/all/1/2000-2023.

Data cleaning was done in R; code can be found at the public code link at the top of this document.

This report details our development of four distinct data visualizations:

1. Interactive Map with Time-Controlled Visualization
2. Bird Migration Trajectory Map
3. Bivariate Choropleth
4. Interactive Heatmap of Species Richness

Visualization 1

The first visualization we chose to implement was an interactive map where the user can control the time via month and year sliders, and bird sightings are overlaid as points on a map of the USA where color represents the average temperature of each state in the given month. This was implemented using ggplot and Shiny.

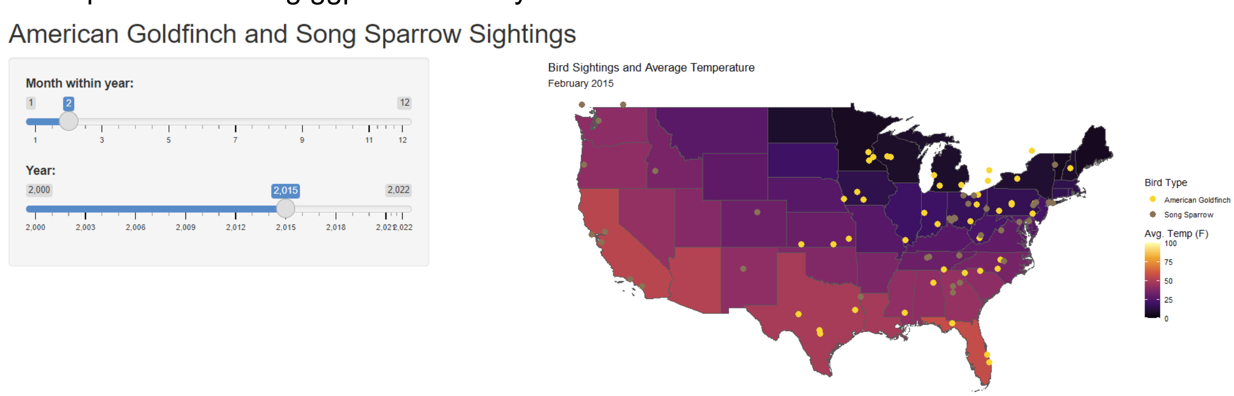


Figure 1.1. Snapshot of interactive map visualization. Sliders on the left allow the user to control the time point shown in the map on the right.

Implementation:

We chose to implement this visualization because encoding a continuous variable over geographic space with color is a classic concept in geospatial visualization, with discrete bird sightings mapping intuitively to points. Color facilitates relative comparison of temperatures better than absolute evaluation; but the former is much more important to this visualization than the latter. We discussed in Milestone 1 how Chen, Wang & Chen (2015)[1] and many others use this principle in their papers - however, a drawback of the Chen, Wang & Chen paper was that they only showed a selection of static snapshots.

Critical Evaluation:

Strengths:

This visualization keeps the visual information in one place, allowing the reader's eyes to stay constant as they change the timepoints - furthermore, they are not limited to a selection of timepoints. Other strengths of this visualization include the month slider enabling the user to see short-term trends in migration and temperature.

Limitations:

However, this visualization also has drawbacks - it does not allow the user to answer all of the queries we outlined in Milestone 1. In particular, the specific migration of a single bird is not visible at all in this visualization. Other queries are merely difficult rather than impossible to answer; teasing out long-term trends in migration and temperature require a lot of interaction with the sliders to parse out.

Possible paths forward:

If we decide to pursue this visualization in the final stage, we would also make further refinements such as enabling zoom/pan on the map and improving the labeling of the sliders.

Visualization 2

Our second visualization in this project is an interactive map that illustrates the migration trajectories of two common migratory birds in the US: the American Goldfinch and the Song Sparrow. Implemented in R Shiny, this visualization utilizes the leaflet package to dynamically display bird migration paths. Users can filter the data based on bird species, bird ID, and date range, offering a comprehensive and interactive exploration of migration patterns.

Implementation:

- Key Features:
    - Interactive Filters: Users can filter data by bird species, ID, and date range.

- Trajectory Visualization: Each bird's trajectory is represented by colored, dashed lines on the map, with different colors signifying different birds.
- Start/End Markers: The start and end of each trajectory are marked with green and red markers, respectively.

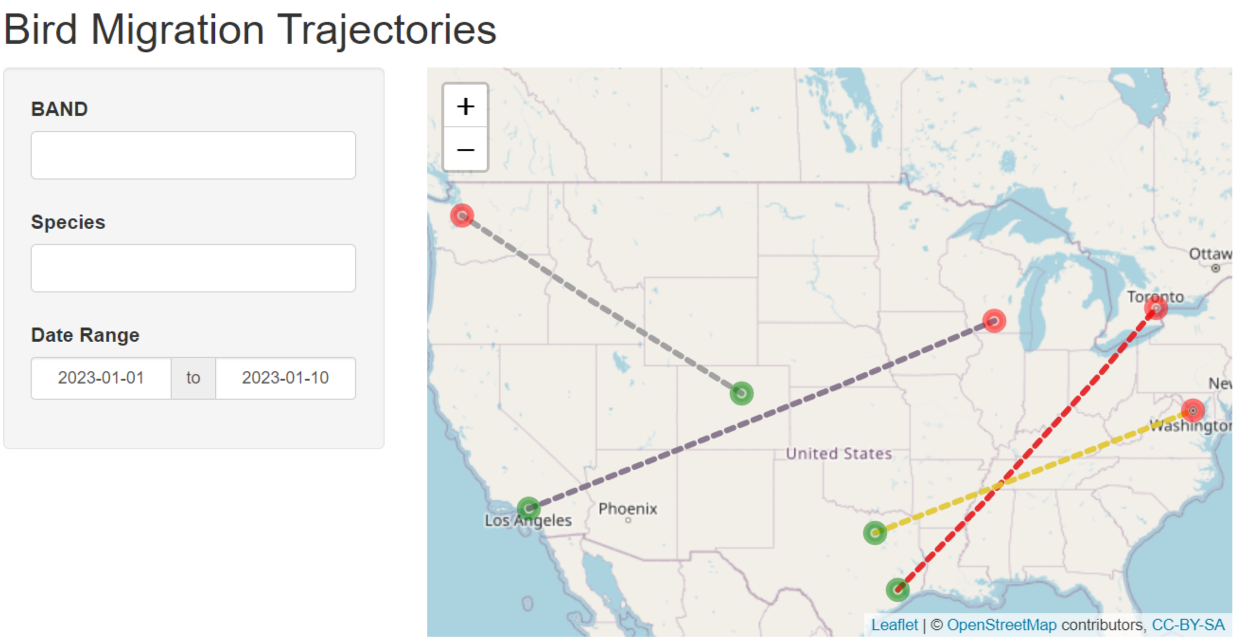


Figure 1.2. Snapshot of interactive map visualization. Each Line represents a bird's movement throughout the filtered date range.

<u>Critical Evaluation:</u>

Strengths:

- Interactivity: Users can interact with the data in real-time, choosing specific birds or time frames to explore.
- Clarity in Data Representation: Different colors and markers make it easy to distinguish between different birds' migration paths.
- Enhanced User Engagement: The ability to manipulate data views encourages user exploration and engagement with the data.

Limitations:

- Data Overload: When multiple birds are selected, the map can become cluttered, making it harder to distinguish individual trajectories.
- Static Temperature Data: The visualization currently focuses only on bird migration without integrating temperature data.

Possible paths forward:

- Enhanced Zoom and Pan Features: Improving map navigation features could enhance user experience.
- Clutter Management: Implementing features like trajectory thinning or dynamic opacity adjustments could help manage visual clutter.

Link for plot: https://github.com/cjones46/992GroupProject/blob/main/Visualisation%202.R

Visualization 3

For our third visualization, we chose an interactive bivariate choropleth. When provided a point in time (via slider) and a species (via drop-down filter), it produces a plot which colors each state based on a two-variable color scheme of increasing population share and temperature. This plot was created in R via the use of the ggplot2, shiny, and biscale packages. The reactive graph follows a relatively simple flow, with much of the work being in the pre-processing. There are three inputs which flow directly into the singular plot output. A static view of the app can be seen in **Figure 1.3**. The associated app behavior and source code can be seen via the “Bivariate Choropleth App” recording and “Project Milestone 2” markdown file respectively at the included Github link.

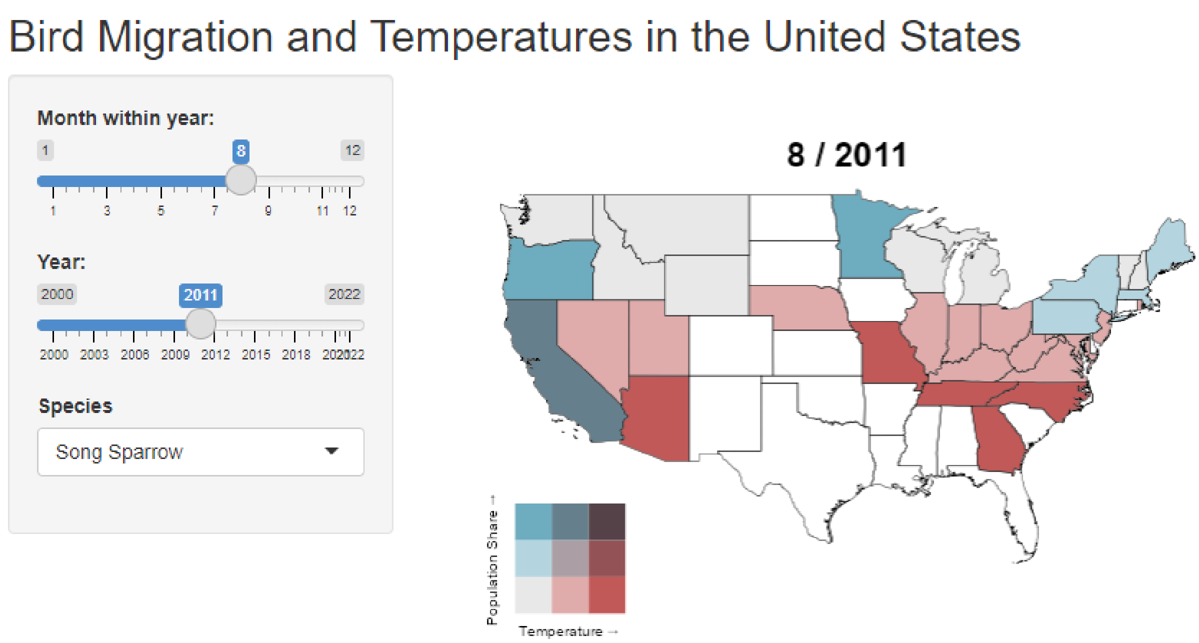


Figure 1.3. Snapshot of interactive map visualization. Bivariate choropleth of temperatures and share of bird population.

<u>Implementation:</u>

This implementation was chosen due to its effectiveness and simplicity. Color-encoding, and choropleths specifically, are an extremely common form of geospatial visualization. While not using the exact same radial tree design as Lu, Ai, Zhang, & He (2017)[2] , this visualization aims to offer similar features in its temporal interactivity and intuitive color schemes, but without the creation of an additional “object” taking up space on the map.

<u>Critical Evaluation:</u>

Strengths:

+ Immediate recognition: Within seconds of looking at the visualization, you know what geographic area you are looking at, what each polygon represents, and a clear legend explains what the color-coding refers to.
+ Interactivity: The data being interactive provides user engagement, the option to encode more information than a static visualization, and the ability to make customized queries.
+ Simplicity/focused view: The plot contains only one species at a time. Compared to other possible migratory visualizations, it is easy to keep track of the relevant elements.

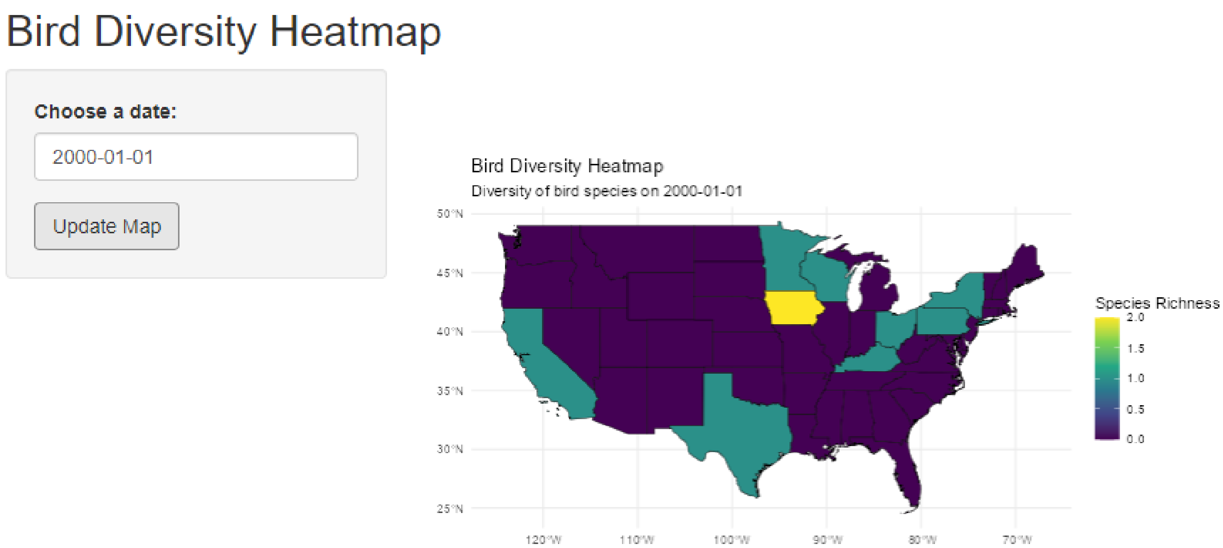
Limitations:

- No precise numeric scale: There is currently no absolute-value representation of birds found within a state; you can’t tell if there were 100 birds or 5 as the max value.
- Query limitation: In the current implementation, only one species and one time period can be selected at a given time.
- Does not show individual patterns: While the bivariate choropleth implementation gives a measure of population and temperature, it does not provide the actual route of any specific bird.

Possible paths forward:

- Multiple visualizations for the same interaction: It could be beneficial as far as getting a full scope of the data to see other visuals when a given selection is made.
- More fluid time view: For example, currently if a user would like to cycle through years chronologically, they must cycle through months 1-12, select the next year, and select month 1 again.
- Additional interaction: Perhaps a user could mouse-over a state at a given point and see the specific observations of birds which compose the selected state, or could see the number of individuals/share of total population at a given time.

Visualization 4



The interactive heatmap was selected to represent bird species richness.
The choice was likely due to its intuitive design, allowing users to quickly grasp variations in species richness across different states.

Implementation:

- The map uses color intensity to indicate the number of species observed, with darker colors representing higher richness.
- The user can select a date to filter sightings, which dynamically updates the map, providing immediate visual feedback.

Critical Evaluation:

Strengths:

Heatmaps effectively communicate density and distribution, which is central to understanding species richness. Interactivity adds the dimension of time, enabling users to see changes over selected dates, making the data exploration dynamic and user-driven.

Limitations:

1.Resolution of Data: Choropleth maps aggregate data into predefined areas (like states), which can mask variations within those areas. Rare or localized species may not significantly impact the overall richness score of a state, leading to underrepresentation.

2.Color Perception: The use of color to represent species richness relies on the user's ability to perceive color differences. This can be problematic for individuals with color vision deficiencies.

3.Static Boundaries: The state boundaries are static and don't account for ecological regions or habitats that cross state lines, which might be more relevant for understanding biodiversity.

Possible paths forward:

Hexbin or Grid Map: Instead of using state boundaries, use a hexagonal or square grid overlay for the map, which can provide a more uniform resolution for data visualization.

Ecological Regions: Overlay ecological regions on the map and calculate species richness within these regions to reflect more natural divisions relevant to bird species.

References

1. Chen, Y., Wang, Q., & Chen, S. (2015). Visualizing spatiotemporal patterns of land use changes using a time-varying cellular automata model. Computers, Environment and Urban Systems, 54, 362-376. https://www.mdpi.com/2073-445X/12/8/1525
2. Lu W, Ai T, Zhang X, He Y. (2017). An Interactive Web Mapping Visualization of Urban Air Quality Monitoring Data of China. Atmosphere, 8(8), 148. https://doi.org/10.3390/atmos8080148