Toronto Beach Data*

Data Analysis of the Relationship Between Date of Observation, Air Temperatures, Water Temperatures, and Waterfowl Activities

Yanzun Jiang

September 26, 2024

This report analyzes environmental data collected in the summer of 2023 regarding Toronto's beaches, focusing on variables such as air temperature, water temperature, and waterfowl observations. The findings reveal that #TODO increased water temperatures correlate with higher waterfowl activity, while rainy conditions tend to reduce sightings.# These insights emphasize the importance of understanding how weather patterns affect wildlife behavior and beach usability, suggesting that public awareness and management strategies should be adapted accordingly. Future research should explore the impact of these environmental factors on beachgoer experiences and the broader ecological implications.

Table of contents

*Code and data supporting this analysis is available at: https://github.com/Stary54264/Toronto-Beach-Data

1 Introduction

In 2014, Toronto experienced an extremely cold summer that impacted its beaches (Gough and Sokappadu (2016)). During this period, data on key environmental factors, such as air temperature, water temperature, and observations of waterfowl activity, was collected at various beaches across the city (Gelfand (2022)). These factors are critical as they directly influence the ecological balance of the beaches, the safety of recreational activities, and the overall enjoyment of beachgoers (R.-Toubes, Araújo-Vila, and Fraiz-Brea (2020)).

Consequently, analyzing this environmental data is essential to understanding how these variables affect the behavior of waterfowl and the suitability of the beaches for public use. According to Mallory, Venier, and McKenney (2003), waterfowl presence can be influenced by temperature changes, precipitation, and human activity on the beaches. This study examines the relationships between air temperature, water temperature, and waterfowl observations, and how these conditions might correlate with fluctuations in bird activity and potential impacts on beachgoers.

To conduct this analysis, a dataset collected by Gelfand (2022) was utilized, as described in Section 2. Based on the initial findings, it was observed that #TODO higher water temperatures coincided with an increase in waterfowl activity, while rainy days led to a decrease in sightings (Section 3).# As discussed in Section 4, #TODO these trends highlight the potential for waterfowl to congregate under specific environmental conditions, potentially affecting beach safety and cleanliness. Further recommendations include improving public awareness of how wildlife interactions and weather patterns influence beach environments.# This paper's structure includes an overview of the data and methods in Section 2, the results of the analysis in Section 3, a detailed discussion of the results in Section 4, and supplementary insights in Section A.

2 Data

2.1 Overview

The dataset used in this analysis is the "Toronto Beaches Observations" sourced from Toronto's Open Data platform (Gelfand (2022)). It records various aspects of the beaches in Toronto from 2008 to present. In Toronto, water quality at beaches is regularly monitored to ensure public safety, and environmental variables, such as air and water temperature, and the presence of waterfowl, are tracked for ecological and recreational purposes (Gelfand (2022)).

The variables analyzed in this study include "Date," which records the day each measurement was taken; "Air Temperature," which represents the temperature of the atmosphere at the

beach site; "Water Temperature," indicating the temperature of the water in designated swimming areas; and "Waterfowl Observations," which track the number of birds spotted near the water (Gelfand (2022)).

The dataset was accessed using the `opendatatoronto` package (Gelfand (2022)). For the analysis, the R programming language was employed (R Core Team (2023)), utilizing the `tidyverse` (Wickham et al. (2019)) package for data cleaning, transformation, visualization, and the `janitor` package (Firke (2023)) for initial cleaning and formatting of the raw data. Afterward, the cleaned dataset was processed and tested using additional functions from the `tidyverse` package (Wickham et al. (2019)).

2.1.1 Preview of the Dataset

The preview of the dataset (Table 1) spans from 2008 to present. The dataset provides a comprehensive overview of key environmental variables observed at Toronto's beaches, including the date of observation, air and water temperatures, and waterfowl activity. These variables are crucial for understanding long-term trends in environmental conditions and their impact on beach safety, water quality, and wildlife behavior. This preview serves as an introduction to the dataset's structure and quality, setting the stage for in-depth analysis and trend exploration over time.

Table 1: Preview of Data

year	month	day	air_temp	water_temp	water_fowl	date
2010	08	03	31	22.6	12	2010-08-03
2010	08	03	31	21.9	30	2010-08-03
2010	08	03	31	24.3	20	2010-08-03
2010	08	03	31	21.3	12	2010-08-03
2010	08	03	31	21.3	30	2010-08-03
2010	08	03	30	21.3	10	2010-08-03

2.1.2 Distribution of the Variables

These histograms (Figure 1, Figure 2, and Figure 3) visualizes the distribution of air temperatures, water temperature, and waterfowl observations across the dataset from 2008 to the present. These graphs helps identify patterns, outliers, and the overall spread of the data. By examining these distributions, we can gain insights into how they behave over time and how it may influence other variables, such as water quality or wildlife observations.

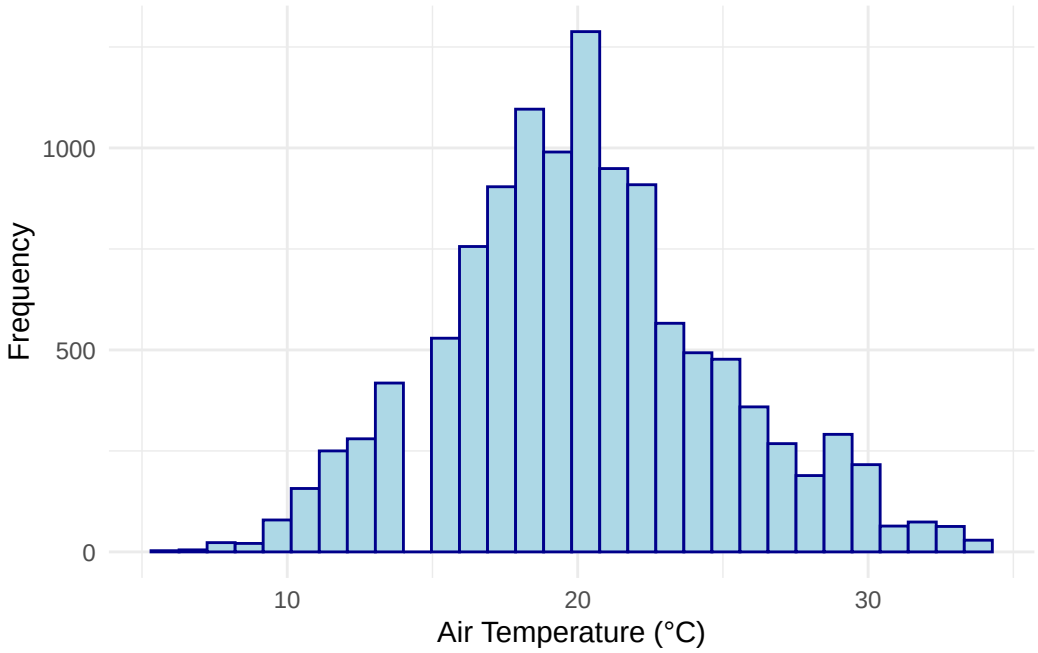

Figure 1: Distribution of Air Temperature at Toronto Beaches

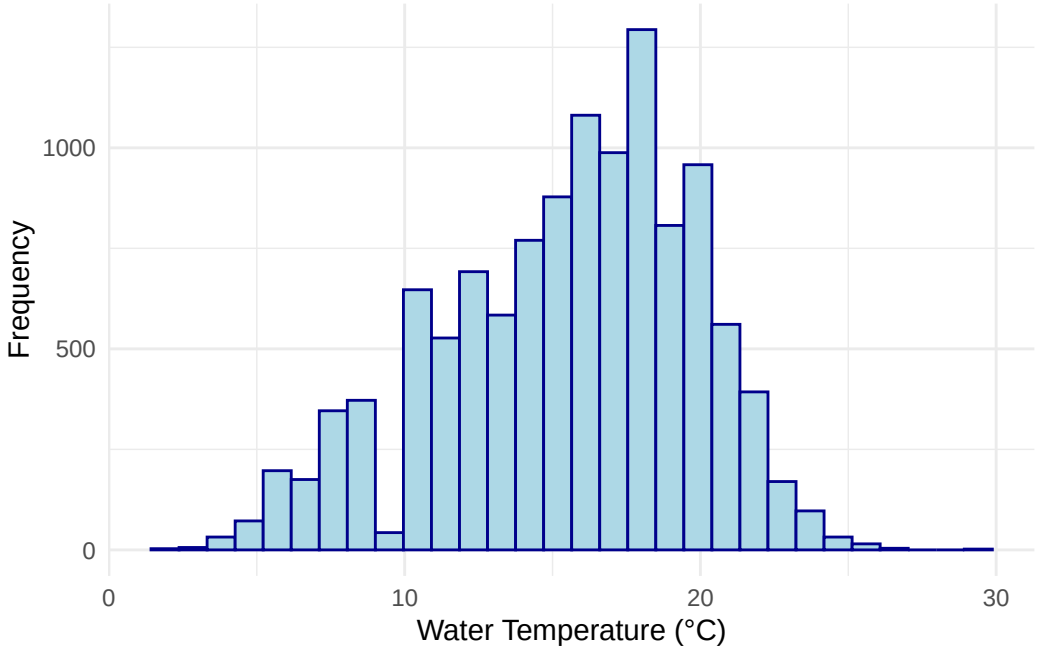

Figure 2: Distribution of Water Temperature at Toronto Beaches

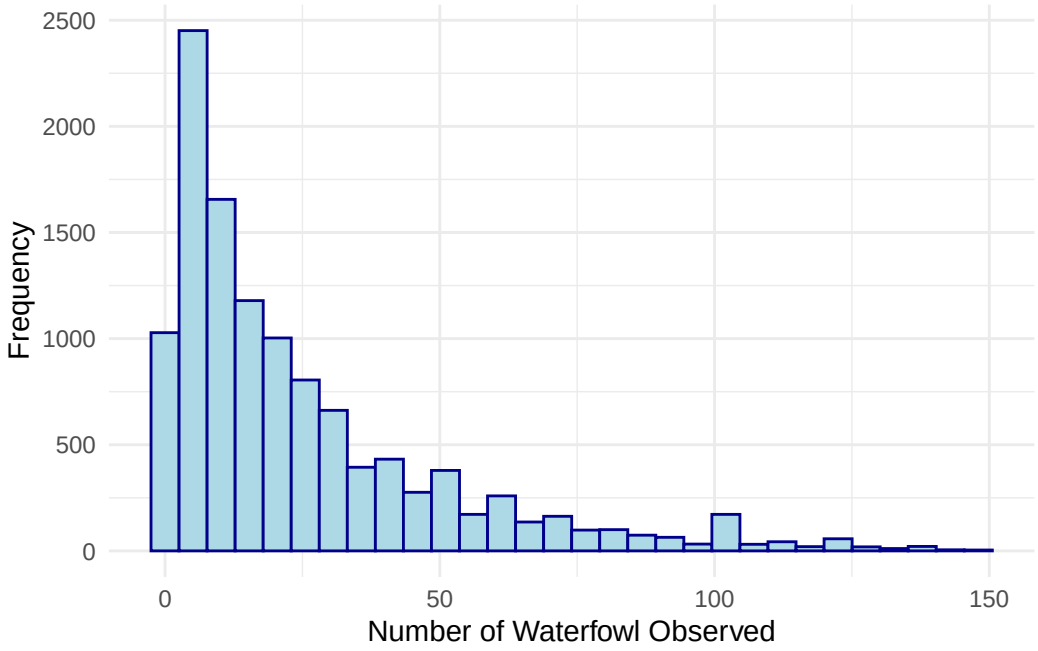

Figure 3: Distribution of Waterfowl Observations at Toronto Beaches

2.2 Relationships Between the Variables

2.2.1 Changes of Beaches with Time

These scatter plots (Figure 4, Figure 5, and Figure 6) visualize the relationship between each environmental variable and time, spanning from 2008 to the present. By plotting air temperature, water temperature, and waterfowl observations against time, we aim to identify trends, and potential long-term changes in these variables at Toronto's beaches. Each scatter plot provides insights into how these environmental factors have fluctuated over the years, revealing trends in beach weather conditions and wildlife presence. These time-based visualizations will aid in understanding any periodic or gradual changes and their potential impact on beach safety and ecological health.

These scatter plots (Figure 7, Figure 8, and Figure 9) illustrate the seasonal variations in air temperature, water temperature, rainfall occurrence, and waterfowl observations at Toronto beaches from 2008 to the present. By plotting these variables over time, distinct seasonal patterns emerge, with warmer temperatures and higher waterfowl activity typically observed during the summer months, and cooler temperatures in winter. The scatter plots help visualize how these environmental factors fluctuate throughout the year, reflecting natural seasonal cycles. Observing these recurring trends provides valuable insights into the seasonal dynamics affecting beach ecosystems and recreational conditions.

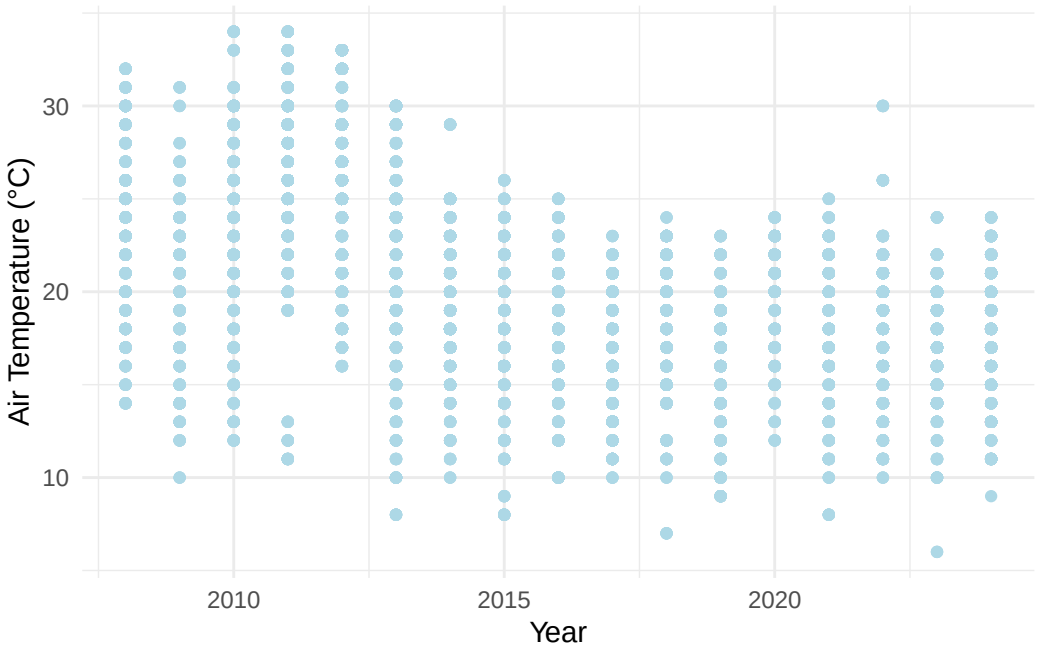

Figure 4: Air Temperature Over Years at Toronto Beaches

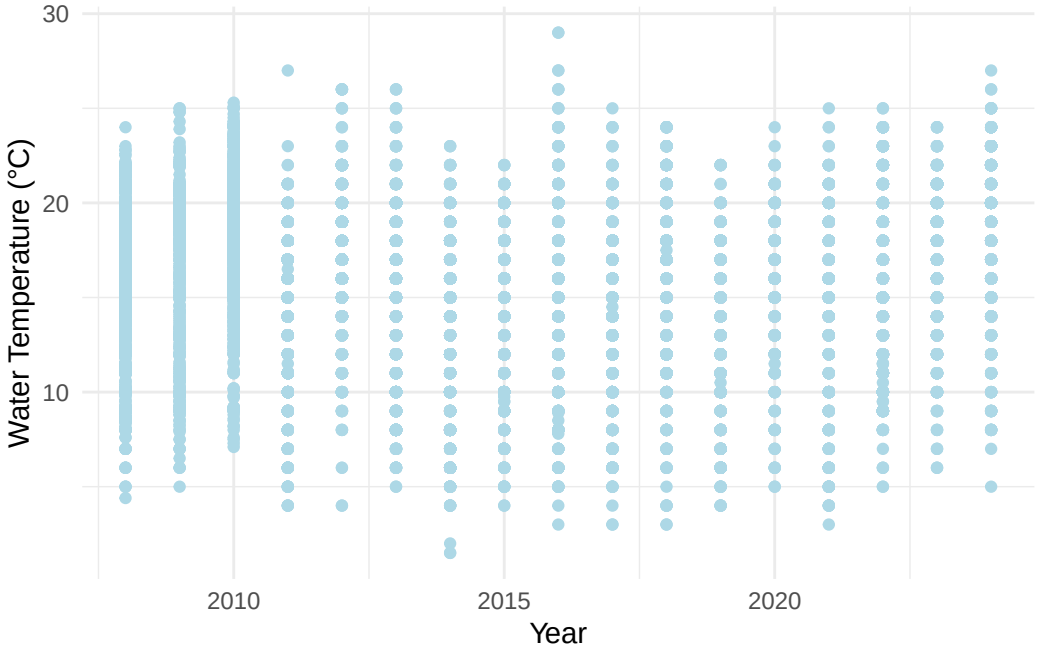

Figure 5: Water Temperature Over Years at Toronto Beaches

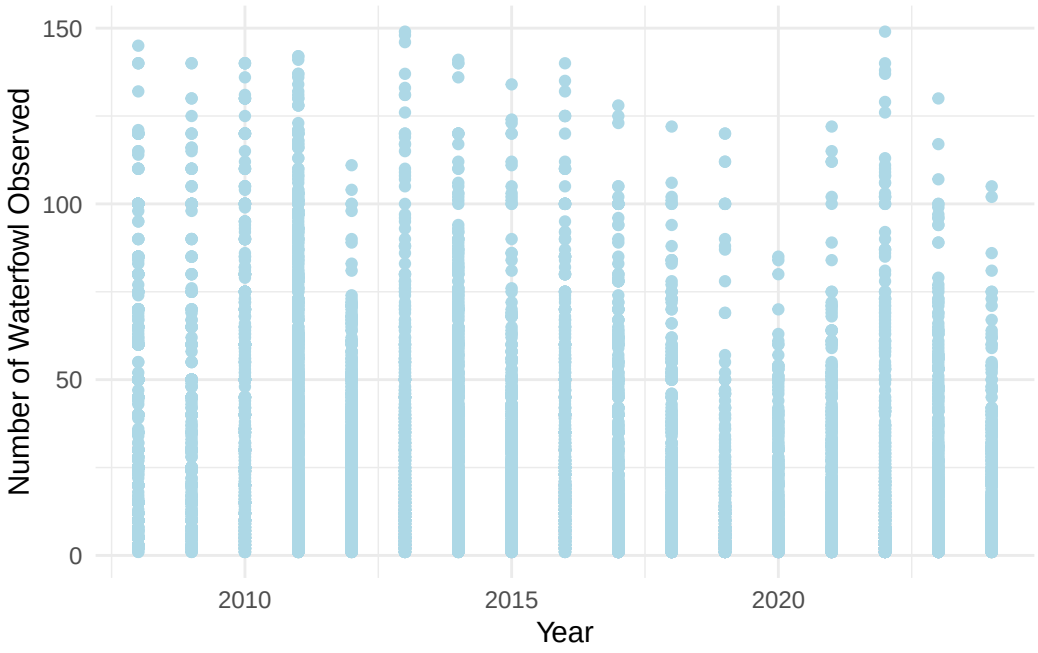

Figure 6: Waterfowl Observations Over Years at Toronto Beaches

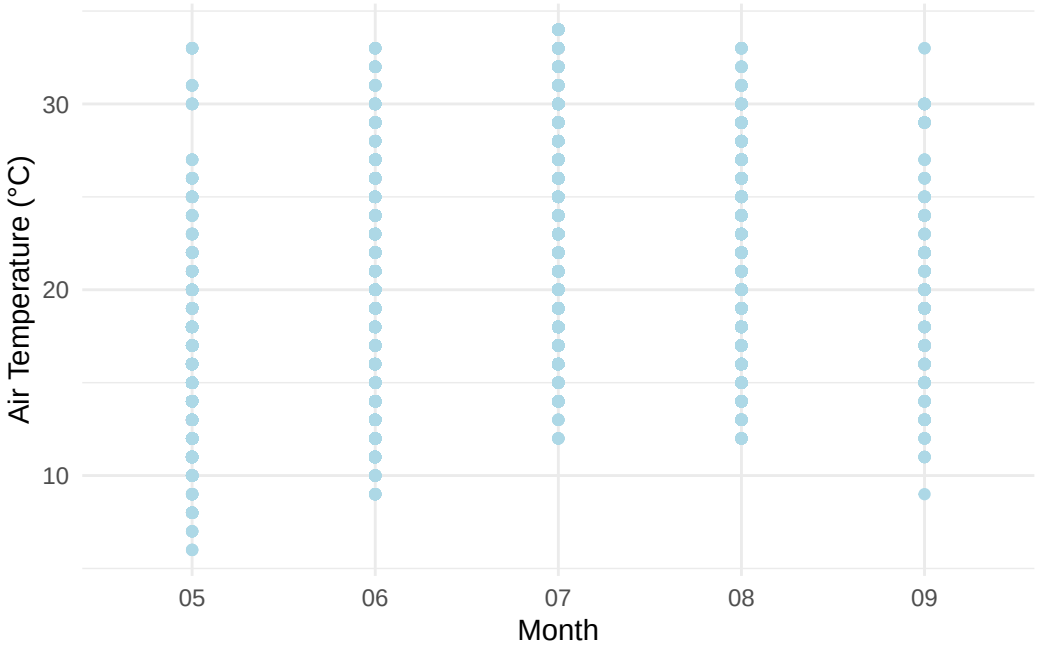

Figure 7: Seasonal Change of Air Temperature at Toronto Beaches

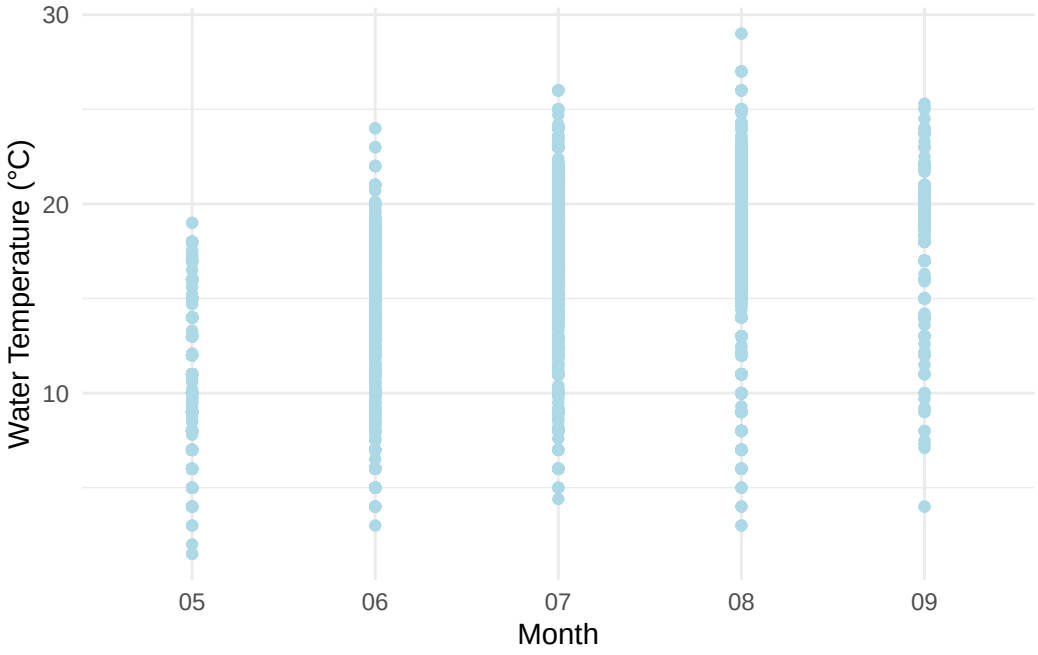

Figure 8: Seasonal Change of Water Temperature at Toronto Beaches

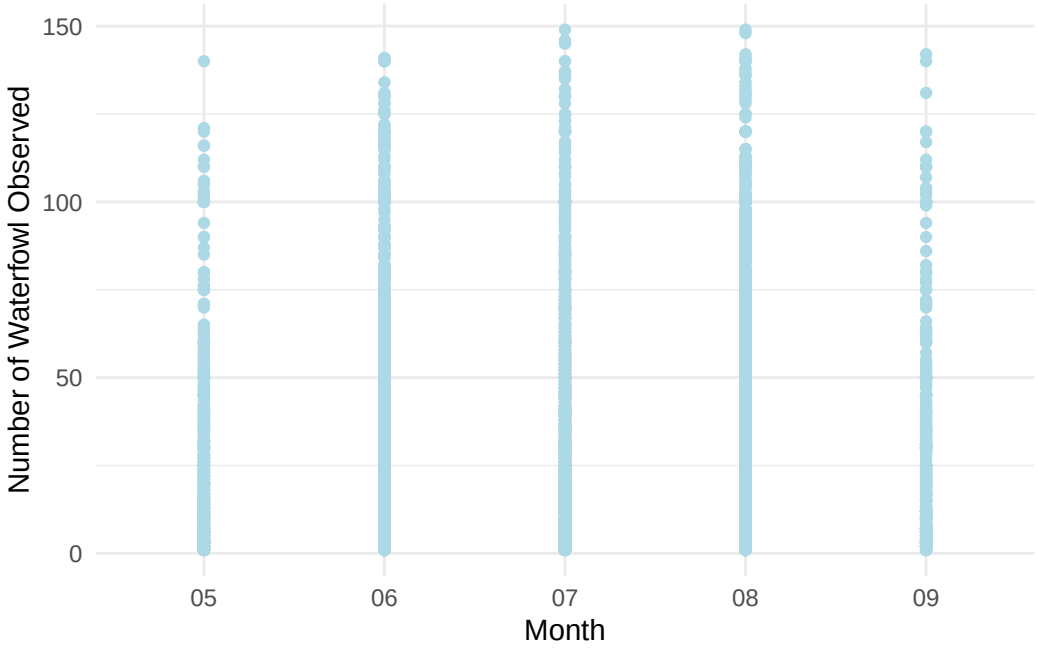

Figure 9: Seasonal Change of Waterfowl Observations at Toronto Beaches

3 Results

From Figure 1, we can clearly see the data is unimodal and is not skewed.

From Figure 2, we can clearly see the data is unimodal, which is the same as air temperatures, but it is left-skewed.

From Figure 3, we can clearly see the data is unimodal but highly right-skewed.

4 Discussion

4.1 First discussion point

If my paper were 10 pages, then should be be at least 2.5 pages. The discussion is a chance to show off what you know and what you learnt from all this.

4.2 Second discussion point

4.3 Third discussion point

4.4 Weaknesses and next steps

Weaknesses and next steps should also be included.

A Appendix

A.1 Graph and Sketches

Sketches depicting both the desired dataset and the graphs generated in this analysis are available in the GitHub Repository.

A.2 Data Cleaning

The data cleaning process involved renaming variables, tidying the dates, filtering out useless columns from the raw dataset, filtering out observations with NAs, and removing the outliers using z-scores.

References

Firke, Sam. 2023. *Janitor: Simple Tools for Examining and Cleaning Dirty Data*. https://CRAN.R-project.org/package=janitor.

Gelfand, Sharla. 2022. *Opendatatoronto: Access the City of Toronto Open Data Portal*. https://CRAN.R-project.org/package=opendatatoronto.

Gough, William A, and Srishtee Sokappadu. 2016. "Climate Context of the Cold Summer of 2014 in Toronto, ON, Canada." *Theoretical and Applied Climatology* 126: 183–89.

Mallory, ML, LA Venier, and D McKenney. 2003. "Winter Weather and Waterfowl Surveys in North-Western Ontario, Canada." *Journal of Biogeography* 30 (3): 441–48.

R Core Team. 2023. *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing. https://www.R-project.org/.

R.-Toubes, Diego, Noelia Araújo-Vila, and José Antonio Fraiz-Brea. 2020. "Influence of Weather on the Behaviour of Tourists in a Beach Destination." *Atmosphere* 11 (1): 121.

Wickham, Hadley, Mara Averick, Jennifer Bryan, Winston Chang, Lucy D'Agostino McGowan, Romain François, Garrett Grolemund, et al. 2019. "Welcome to the tidyverse." *Journal of Open Source Software* 4 (43): 1686. https://doi.org/10.21105/joss.01686.