

Toronto Beach Data*

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

Yanzun Jiang

September 27, 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 air temperatures have a linear relationship with water temperatures, while increased water temperatures correlate with higher waterfowl activity in some content. 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

1	Introduction	3
2	Data	3
2.1	Overview	3
2.2	Preview of the Dataset	4
2.3	Data Visualization	4
2.3.1	Distribution of the Variables	4
2.3.2	Changes of Beaches with Time	6
2.3.3	Correlations Between Observed Variables	10
3	Results	10
3.1	Gain from Distribution of the Variables	10
3.2	Gain from Changes of Beaches with Time	12

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

3.3	Gain from Seasonal Changes of Beaches	12
3.4	Gain from Correlations Between Observed Variables	13
4	Discussion	14
4.1	First discussion point	14
4.2	Second discussion point	14
4.3	Third discussion point	14
4.4	Weaknesses and next steps	14
A	Appendix	15
A.1	Graph and Sketches	15
A.2	Data Cleaning	15
A.3	Summary Statistics of the Data	15
	References	16

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 higher air temperatures are correlated with higher water temperatures, while higher water temperatures coincided with an increase in waterfowl activity in some extent (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)).

2.2 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

2.3 Data Visualization

2.3.1 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.

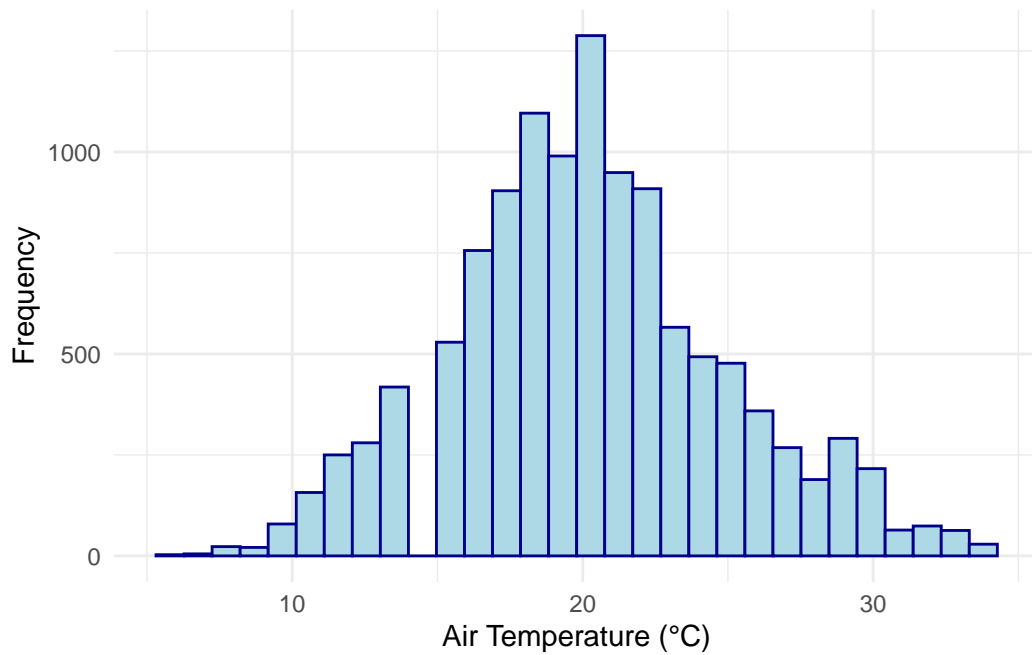


Figure 1: Distribution of Air Temperature at Toronto Beaches

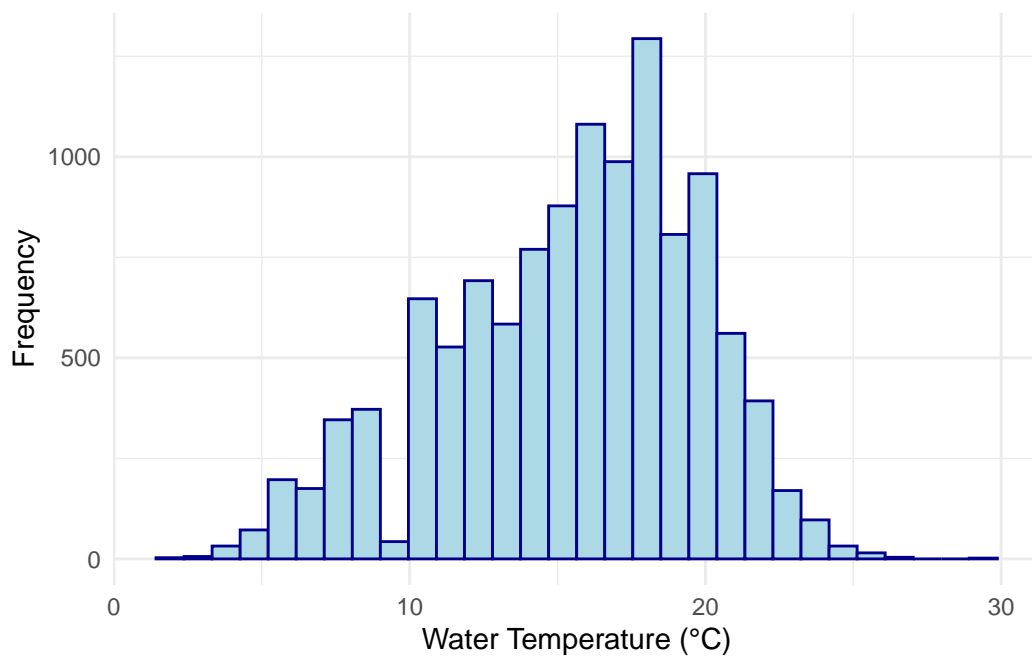


Figure 2: Distribution of Water Temperature at Toronto Beaches

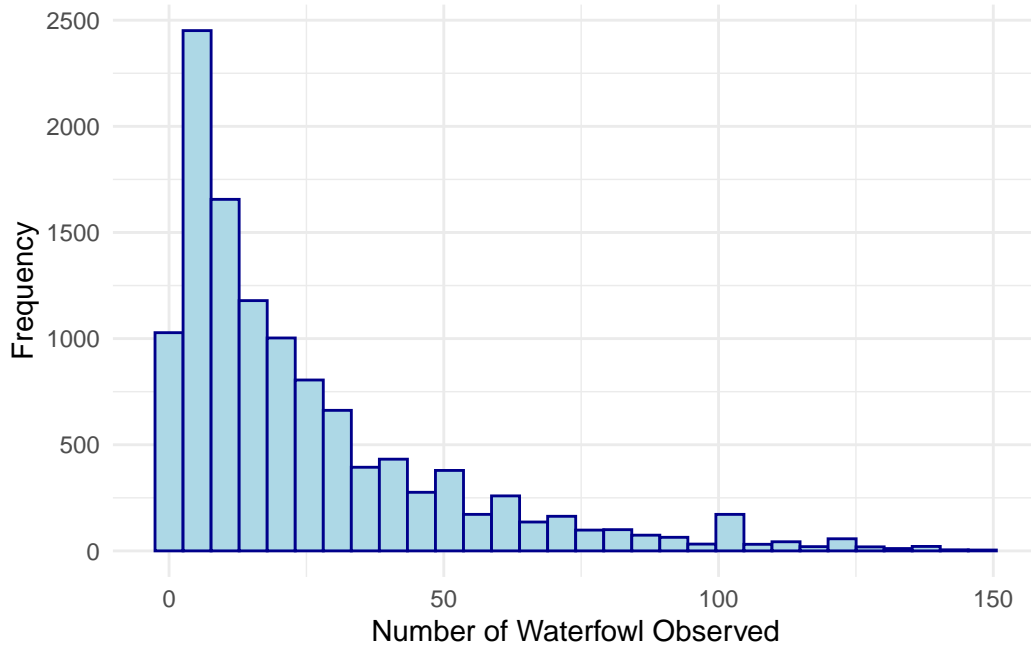


Figure 3: Distribution of Waterfowl Observations at Toronto Beaches

2.3.2 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. 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.

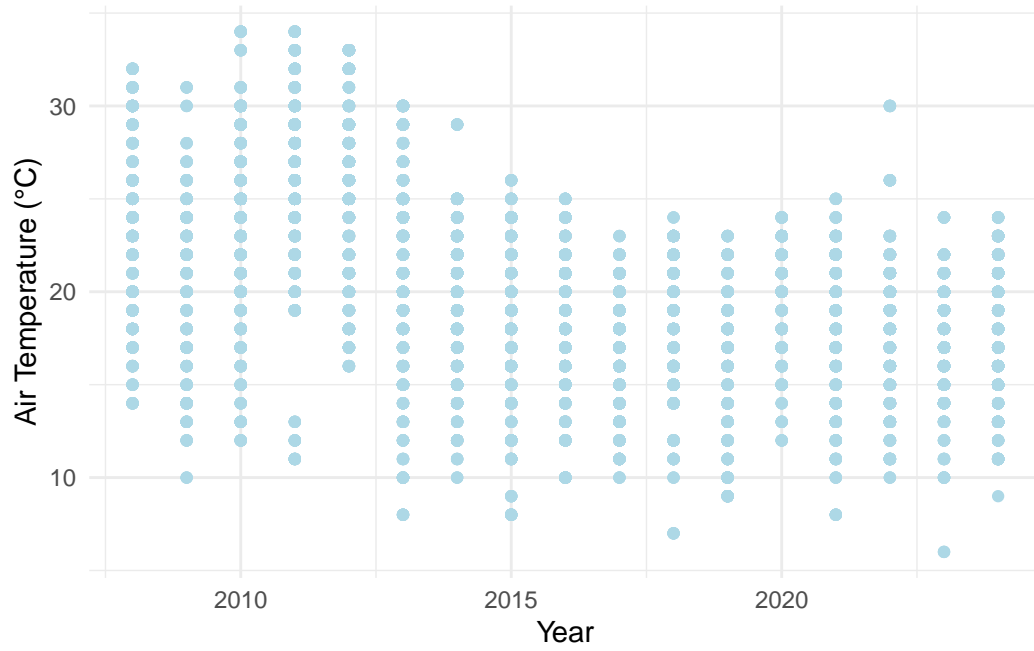


Figure 4: Air Temperature Over Years at Toronto Beaches

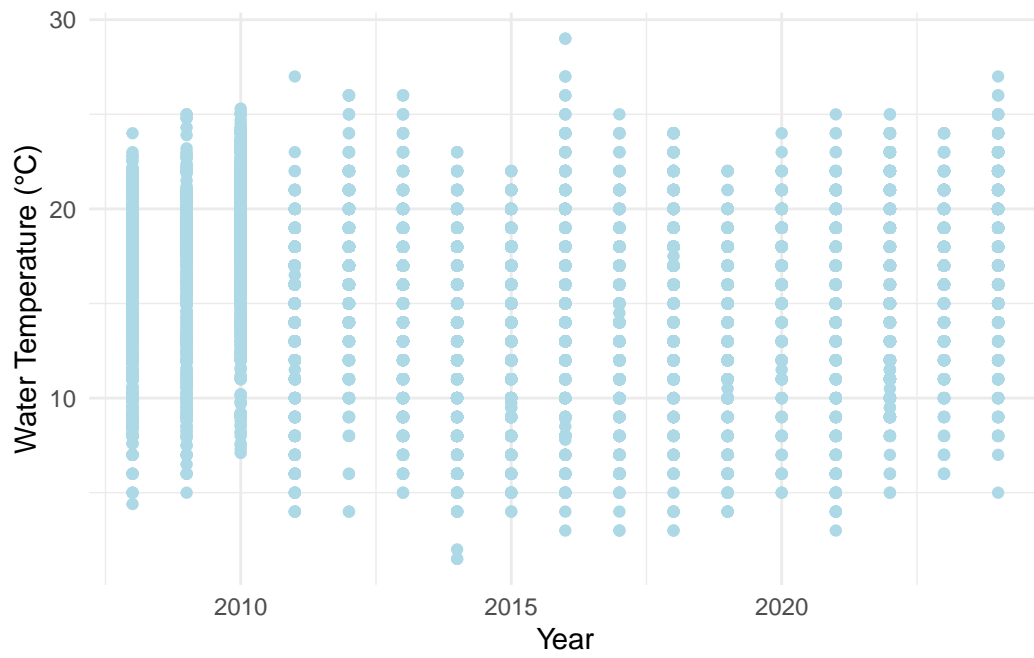


Figure 5: Water Temperature Over Years at Toronto Beaches

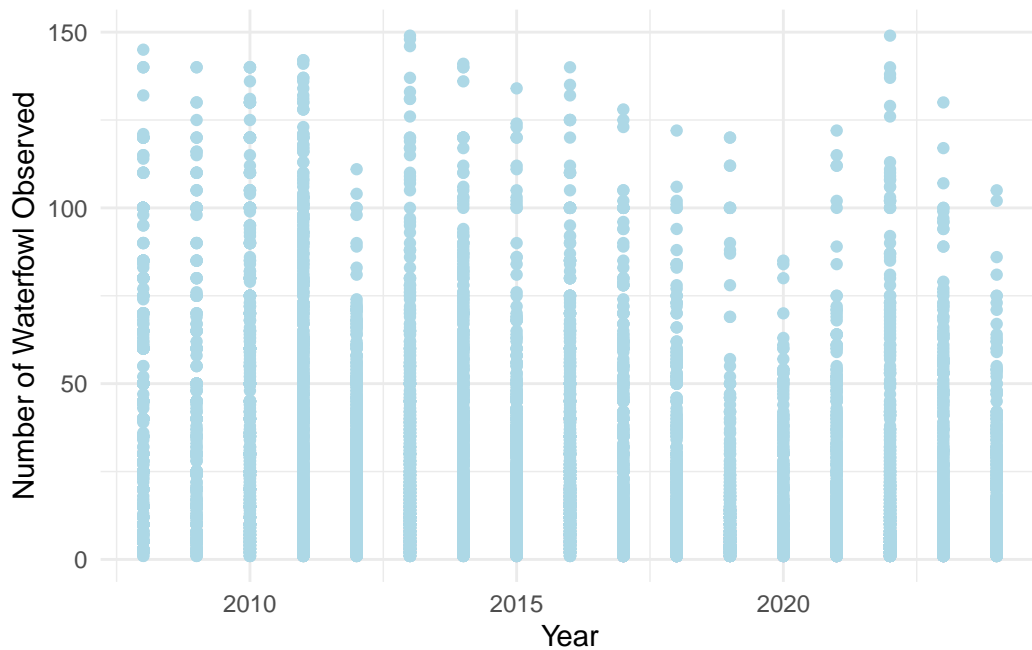


Figure 6: Waterfowl Observations Over Years at Toronto Beaches

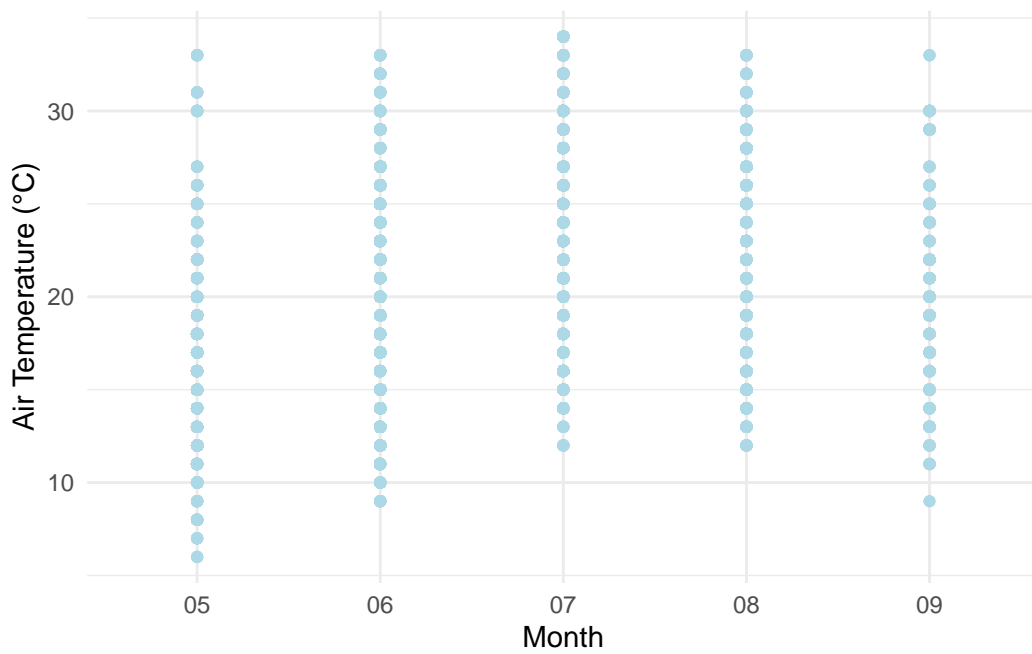


Figure 7: Seasonal Change of Air Temperature at Toronto Beaches

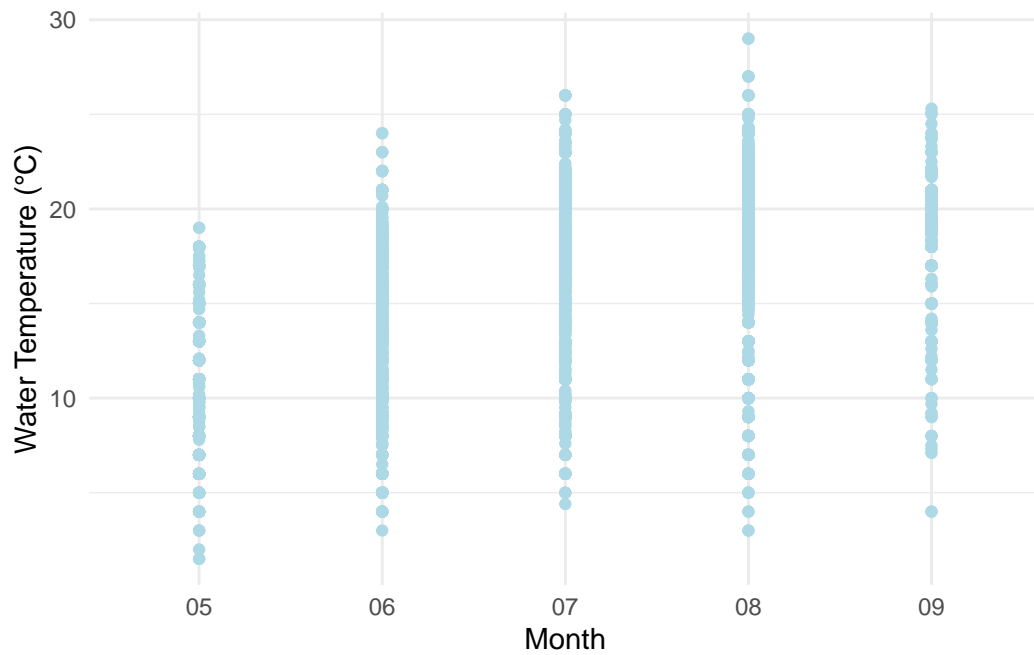


Figure 8: Seasonal Change of Water Temperature at Toronto Beaches

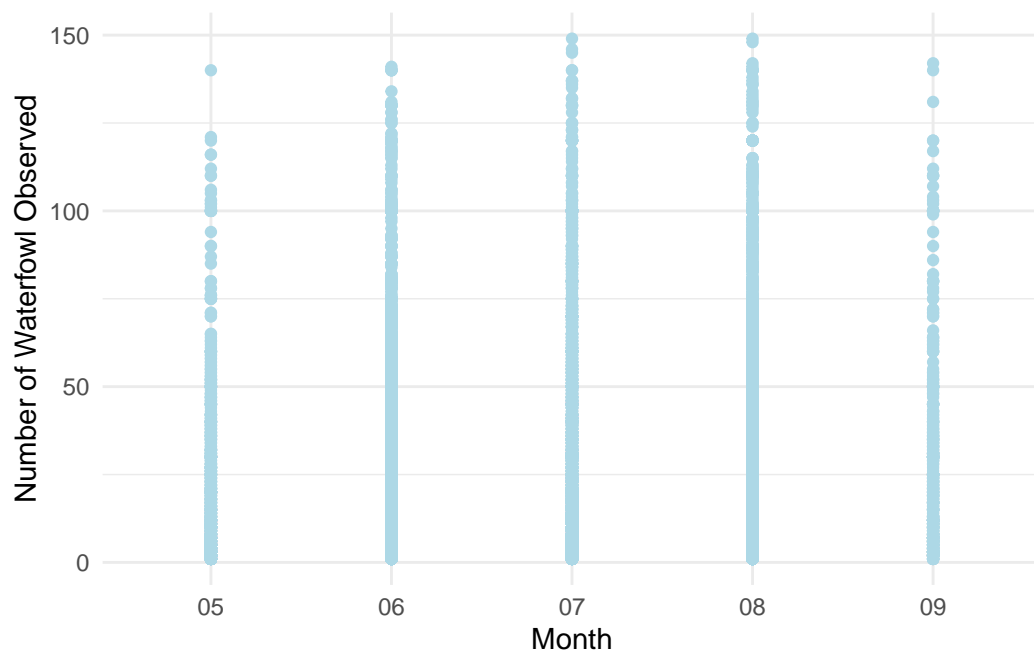


Figure 9: Seasonal Change of Waterfowl Observations at Toronto Beaches

2.3.3 Correlations Between Observed Variables

These scatter plots (Figure 10, Figure 11) explore the relationships between various observed variables, such as air temperature, water temperature, and waterfowl observations, to identify potential correlations. By plotting some variables against some other, we aim to uncover how these environmental factor might influence another. For instance, the relationship between air temperature and water temperature may reveal patterns in heat transfer between the atmosphere and water, while the correlation between water temperature and waterfowl observations could indicate how weather conditions affect wildlife activity. These scatter plots provide a deeper understanding of the interconnectedness of these variables and offer insights into how environmental conditions at Toronto beaches influence each other.

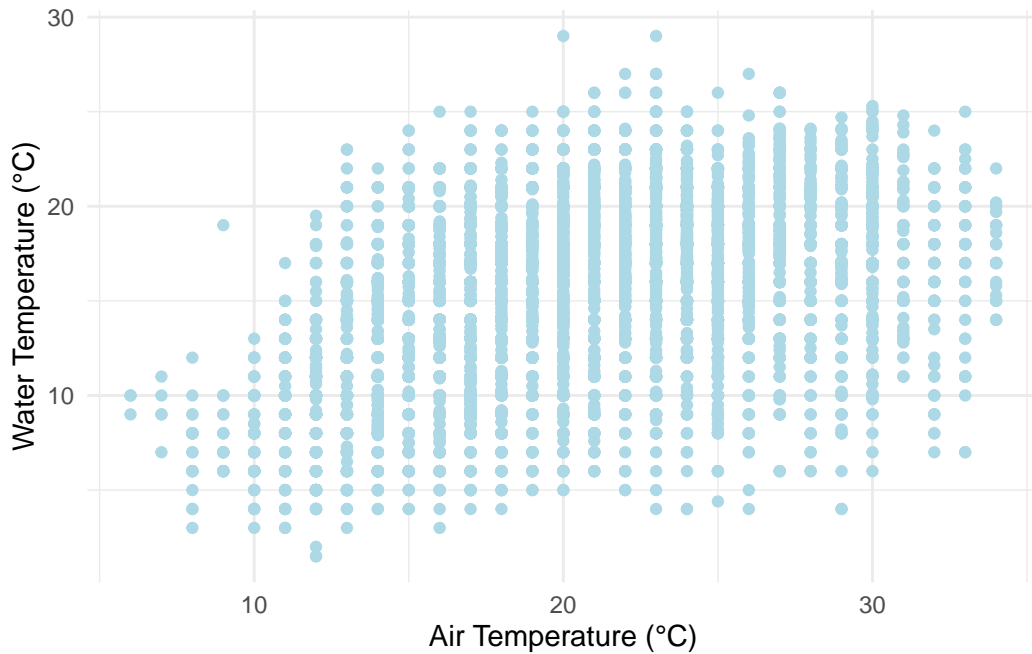


Figure 10: Relationship Between Air Temperature and Water Temperature at Toronto Beaches

3 Results

3.1 Gain from Distribution of the Variables

The histogram of air temperatures, as shown in Figure 1, reveals that most recorded air temperatures at Toronto beaches fall between 12°C and 30°C, with a noticeable peak around 20°C. The distribution is approximately symmetric, indicating no significant skew, meaning that air temperatures are evenly spread around the central value. This suggests that beach

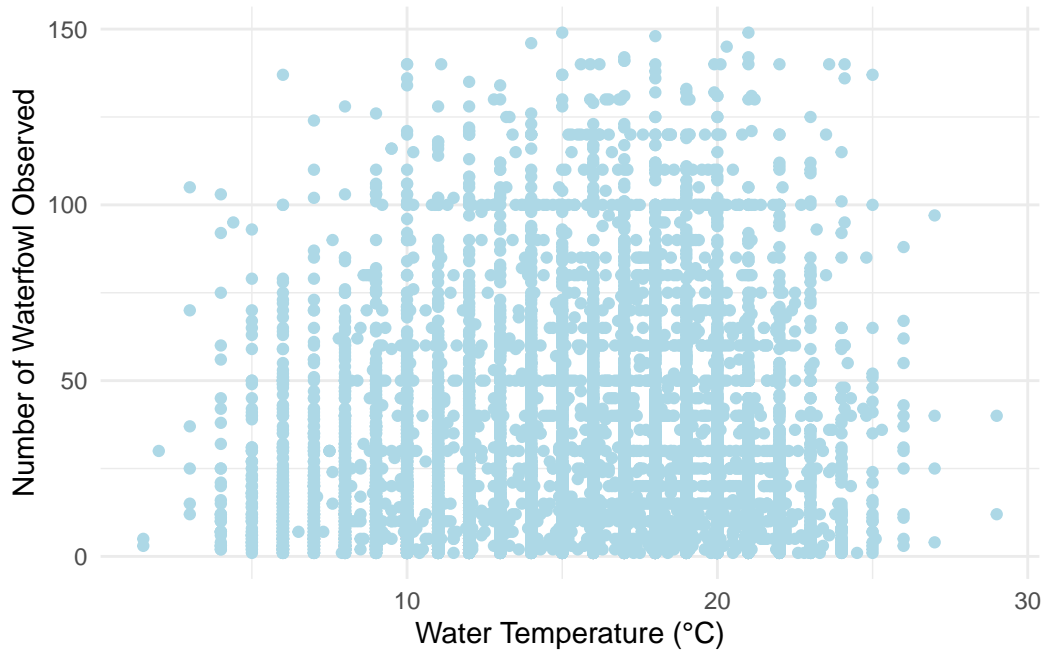


Figure 11: Relationship Between Water Temperature and Waterfowl Observations at Toronto Beaches

monitoring predominantly occurs during periods of moderate and comfortable weather, which is conducive to outdoor activities. Temperatures below 8°C and above 33°C are much less frequent, indicating that extreme weather conditions are rare, emphasizing the temperate climate in Toronto during beach usage periods.

In contrast, the water temperature histogram, seen in Figure 2, is left-skewed, meaning that while most temperatures cluster between 8°C and 25°C, there is a tail extending toward lower temperatures. A peak is observed around 18°C, representing the most common water temperatures during beach season. The left skew indicates that colder water temperatures are infrequent but present, typically occurring outside peak summer months. This distribution reflects how water retains heat more slowly than air, remaining relatively cool even as air temperatures rise, which is characteristic of large bodies of water.

Lastly, the waterfowl observations histogram, as seen in Figure 3, displays a highly right-skewed distribution. Most days had a low count of waterfowl, with the majority of observations falling between 0 and 50 observations. However, a long tail extends to the right, indicating that while large counts of waterfowl are rare, they do occur on certain days. This right skew suggests occasional spikes in bird activity, possibly linked to seasonal migrations or environmental conditions that attract more birds to the beaches. These sporadic peaks could have implications for water quality, as larger groups of waterfowl might increase the risk of contamination.

These histograms (Figure 1, Figure 2, and Figure 3) offer valuable insights into the key environmental variables affecting Toronto beaches. The symmetric air temperature distribution, left-skewed water temperature, and highly right-skewed waterfowl observations each reveal distinct patterns that inform public health monitoring and beach management decisions, ensuring that the environment remains safe and enjoyable for visitors.

3.2 Gain from Changes of Beaches with Time

Figure 4 shows the relationship between air temperature and time. Over the years, we observe a decrease followed by an increase in air temperatures, with potential gradual trends reflecting long-term changes rather than distinct seasonal cycles. This could indicate climate-related variations or other long-term environmental shifts affecting the air temperature at the beaches.

Figure 5 displays water temperature versus time. Similar to air temperature, water temperature shows fluctuations across years, but with a less explicit pattern. The data suggests that the overall trends could be linked to climate change or weather anomalies.

Figure 6 illustrates waterfowl observations over time. The scatter plot reveals variations in bird activity across different years. While no sharp seasonal patterns are apparent due to the yearly scale, the data might point toward longer-term changes in waterfowl populations, potentially related to environmental factors, habitat changes, or migratory patterns.

These scatter plots (Figure 4, Figure 5, and Figure 6) provide an overview of how the environmental conditions at Toronto beaches have evolved over the years. Rather than focusing on short-term seasonal changes, these visualizations highlight potential long-term trends and environmental shifts that could have significant implications for beach management and ecosystem health.

3.3 Gain from Seasonal Changes of Beaches

Figure 7 illustrates the relationship between air temperature and time, focusing on the months from May to October. A clear seasonal pattern emerges, with air temperatures gradually increasing from May, peaking around July, and then declining into September. The seasonal cycle is evident, as air temperatures warm up in late spring, reach their highest points in midsummer, and cool down as fall approaches. This reinforces the influence of seasonal climate variations on beach conditions.

Figure 8 presents air temperature versus time over the same period. It reveals a similar pattern to air temperature, with a peak in August, which is later than air temperature. This might be due to the absorption of heat in water, causing the temperature to fall gradually. This trend aligns with typical summer weather patterns, where water temperatures rise due to prolonged exposure to sunlight and warmer air, creating optimal conditions for beachgoers.

Figure 9 depicts waterfowl observations versus months from May to September. It shows varying levels of bird activity throughout the summer months, with peaks in July or August. These fluctuations could be influenced by breeding seasons, availability of food, or migratory behavior, all of which are affected by seasonal changes. The variability in bird counts reflects how environmental factors and beach conditions interact with wildlife throughout the warmer months.

In summary, these scatter plots (Figure 7, Figure 8, and Figure 9) highlight the clear seasonal trends in both temperature and waterfowl observations at Toronto beaches between May and September. These patterns are important for understanding how environmental conditions shift during the beach season and inform management strategies for maintaining water quality and ecosystem balance during peak usage times.

3.4 Gain from Correlations Between Observed Variables

Figure 10 depicts the relationship between air temperature and water temperature. It shows a clear positive correlation between the two variables: as air temperatures rise, water temperatures also tend to increase. This relationship is expected, as higher air temperatures lead to more intense solar radiation and warming of the water. However, the variance of water temperature given same air temperature is large. There might be other confounders affecting this relationship. Also, this might be due to the slower rate at which water heats and cools compared to air.

Figure 11 illustrates the relationship between water temperature and waterfowl observations. It does not show a linear relationship between the variables. Rather, it reveals that waterfowl counts are generally higher when water temperatures are moderate, with fewer birds observed at the extreme ends of the temperature range. This suggests that waterfowl may be more active or present in larger numbers when water conditions are neither too cold nor too warm, possibly reflecting optimal conditions for feeding or habitat usage during these periods.

These scatter plots (Figure 10 and Figure 11) help clarify how environmental variables at Toronto beaches are interrelated. The strong correlation between air and water temperatures demonstrates how weather patterns influence beach conditions, while the relationship between water temperature and waterfowl observations provides insights into how wildlife activity is shaped by environmental factors. Understanding these connections is crucial for managing beach environments and ensuring both human and ecological health.

4 Discussion

4.1 First discussion point

If my paper were 10 pages, then should 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.

A.3 Summary Statistics of the Data

The summary statistics provides key statistical insights into the environmental variables analyzed at Toronto beaches. For each variable, the summarize tables (Table 2, Table 3, and Table 4) reports the mean, median, variance, minimum value, maximum value, and interquartile range of those variables, offering a concise overview of range, central tendencies, and variability in the data.

Table 2: Summarise Statistics of Air Temperatures

Mean	Median	Variance	Min	Max	IQR
20.09467	20	22.04706	6	34	6

Table 3: Summarise Statistics of Water Temperatures

Mean	Median	Variance	Min	Max	IQR
15.45061	16	18.67936	1.5	29	7

Table 4: Summarise Statistics of Waterfowl Observations

Mean	Median	Variance	Min	Max	IQR
24.82683	15	672.9996	1	149	28

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 Grolemond, et al. 2019. “Welcome to the tidyverse.” *Journal of Open Source Software* 4 (43): 1686. <https://doi.org/10.21105/joss.01686>.