



# Unravelling Bird Diversity in Chicago



## The Data Sleuths

Takashi Osanai - [tosana2@uic.edu](mailto:tosana2@uic.edu) - TakashiOsa

Arka Pal - [apal7@uic.edu](mailto:apal7@uic.edu) - ArkaPal-uic

Raunak Singh - [rsing76@uic.edu](mailto:rsing76@uic.edu) - raunaksingh1497

Lokesh Manideep - [lbogg@uic.edu](mailto:lbogg@uic.edu) - lokeshmanideepb

Repository: <https://github.com/uic-cs418/group-project-the-data-sleuths.git>



# Chicago Bird Diversity

*“The project explores the bird diversity across neighbourhoods in Chicago.”*

## Why care about urban Biodiversity

- Higher biodiversity leads to a healthier ecosystem
- A healthy ecosystem provides many benefits to the community, including Nutrient cycling, Pest and disease control and Mental wellness.

# Why we chose this problem?

- Birds are a great indicator of the health of our environment.
- There is a lot of easily accessible data on birds compared to other wildlife
- Understanding how and why wildlife is distributed can help urban planners design neighborhoods that best supports the ecosystem.



# How will it help stakeholders

Can assist the Chicago City Government in identifying areas with significant declines in diversity.

Can help urban planners to prioritize conservation efforts and implement policies.

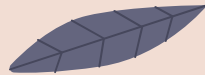
Nonprofit Organizations can utilize our findings to engage communities in bird conservation and raise awareness.



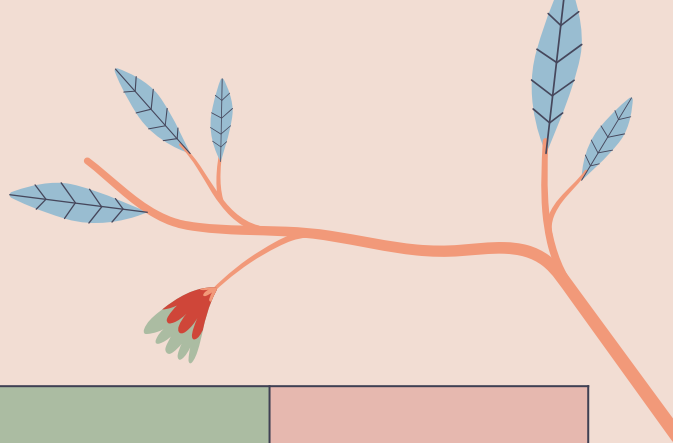


# Hypothesis

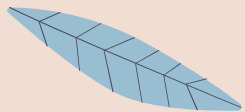
Bird diversity is increasing  
in Chicago.



# Dataset



Datasets	Description	Relevant Features	Size	Resource
<b>Cornell Lab of ornithology -eBird Basic Dataset</b>	Each row represents a single observation	Species Bird Count Coordinates Date	~ 4 Million rows 20 Columns	<a href="https://ebird.org">ebird.org</a>



# Key Insights from the Data



01

## Species

**Ring-billed Gull** and **European Starling** are the most commonly seen birds in Chicago .

02

## Communities

Approximately **60%** of bird observations are concentrated in only five communities, with uptown accounting for nearly **27%** of these observations.

03

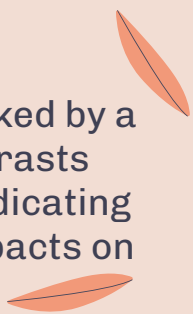
## Observation

**April**, **May**, and **June** have higher bird observations due to breeding and migration patterns in the summer.

04

## Temporal Trend

Peak diversity in **2016**, marked by a higher Shannon Index, contrasts with the decline by **2021**, indicating potential environmental impacts on bird populations.



# Exploratory Data Analysis



## Data Preparation

Prepared Data to focus on key attributes like species, location etc.



## Data Filtering

Filtering out Incomplete and incidental Observations



## Feature Engineering

Created new features 'NATIVE' and 'COUNT'.



## Geospatial Analysis

Used community area shapefiles to enhance understanding of species distribution in Chicago.




## EDA and Visualization

Analyzed the seasonal distribution of species in Chicago.



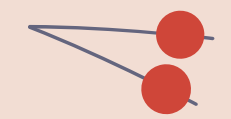


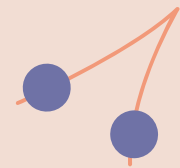
# Visualizations



In our analysis, we've conducted three visualizations to refine our findings.

- 1) Bird Diversity by neighbourhoods.
- 2) Diversity through different seasons.
- 3) Trends in Diversity in the last decade.

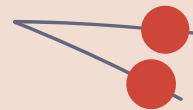




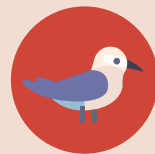
# Diversity through Seasons



There's a decrease in observations from Autumn to Winter.



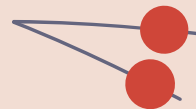
# Trends in Bird Diversity across time by region



Trend in Bird Diversity Across Locations Over the Years



Over the years 2014 to 2023, there is a noticeable trend in bird diversity and abundance, especially in the **South Western** region



# Machine Learning

**Input Features :** Neighborhood, Year, Month



**Target Variable :** Shannon Index (Bird Diversity)

## Algorithms

- Baseline Algorithm(Mean)
- Decision Tree Regressor
- Random Forest Regressor
- Support Vector Regressor
- Gradient Boosting Regressor

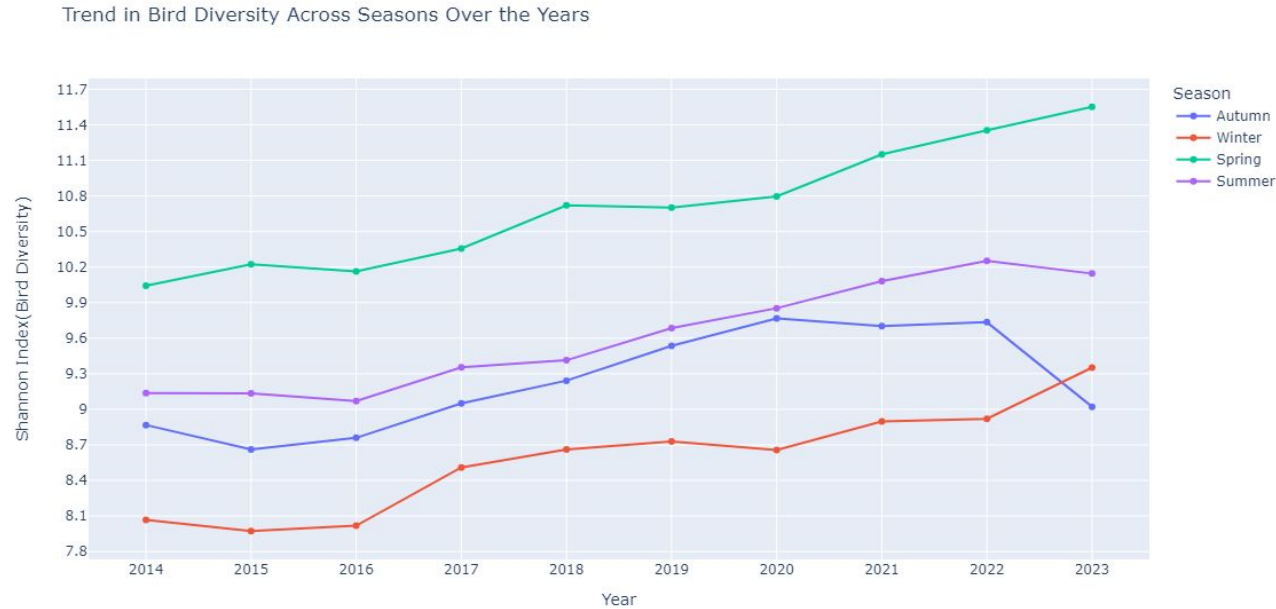
# Machine Learning



Model	MSE	RMSE	MAE	R <sup>2</sup>
Baseline	3.32	1.822	1.491	-0.002
Decision Tree Regressor	1.229	1.109	0.815	0.629
Random Forest Regressor	0.798	0.894	0.665	0.759
SVR	3.295	1.815	1.485	0.006
Gradient Boosting Regressor	1.467 	1.211	0.961	0.557 



# Conclusion





# Questions?

