

Rowdy Datathon 2024

Theme: Decline in Population of Monarch Butterflies

1. Introduction

As a duo, we undertook an ambitious project to analyze monarch butterfly populations and explore the factors contributing to their decline. This initiative was part of the 2024 Rowdy Datathon at the University of Texas at San Antonio, and it required us to dive deep into data collection, analysis, and interpretation to propose strategies that could help reverse the decline in pollinators—specifically monarch butterflies.

2. Problem Description

We were tasked with investigating the alarming decline in monarch butterfly populations, particularly their migration through Texas. The challenge was to identify the environmental and human-driven factors behind this decline and propose actionable strategies. Our approach involved collecting and analyzing data on monarch migration patterns, pesticide usage, air quality, and climate factors. It was a challenging but rewarding task that stretched our data analysis skills and understanding of ecological issues.



Fig 1. An Adult Monarch Butterfly Feeding on the Nectar of a Flower

3. Motivation

Our shared motivation stemmed from a deep understanding of the ecological importance of monarch butterflies. We quickly realized that their decline had implications far beyond just one species—it threatened entire ecosystems, agricultural sectors, and food security. As we delved deeper, the connection between pollinators and broader environmental issues became evident, fueling our desire to uncover actionable insights.

One of the primary motivators for us was the potential impact on agriculture in Texas. Monarchs pass through key agricultural regions during their migration, and the consequences of their decline would be devastating for local food production. This challenge brought us closer as a team, and we were motivated to use our combined skills to find meaningful insights.

4. Data Collection and Analysis

As a team, we gathered data from several sources to form a comprehensive picture of the situation. We used Journey North to track monarch sightings, EPA's Air Quality System for environmental data, and the USDA Pesticide Data Program (PDP) for pesticide-related information.

Our roles complemented each other perfectly: while one of us focused on web scraping and cleaning large datasets, the other worked on building pipelines and geocoding city-level data to county-level insights using the Nominatim OpenStreetMap API. This teamwork helped us overcome some of the more challenging aspects of data collection, like normalizing inconsistent formats and filling in missing location data.

Together, we cleaned and prepared datasets, enabling us to explore correlations between monarch butterfly declines and environmental stressors like pesticide use and air quality. This was the first time we both worked with such large datasets, and it required a lot of patience and collaboration, but it was satisfying to see it all come together.

5. Challenges Faced

One of the biggest challenges we faced was dealing with data at different granularity levels. We had city-level data for monarch sightings but needed county-level information for deeper analysis. We used the Nominatim OpenStreetMap API to infer this, which was time-consuming but essential to our analysis. Another challenge was handling the large datasets we scraped from platforms like Journey North, requiring careful filtering and automation to ensure only relevant data was processed.

We also ran into performance bottlenecks while processing some of the larger datasets. However, by pooling our skills and troubleshooting together, we were able to optimize our code and make sure it ran smoothly. Working as a team helped us think through problems from different perspectives, ultimately leading to better solutions.

6. Findings

Our analysis uncovered several key findings. First, we identified a clear correlation between regions with high pesticide use and a decline in monarch butterfly sightings. Areas with intensive agricultural practices and genetically modified crops seemed to be the hardest hit. Air quality data showed that poor conditions during key migration months further contributed to the decline as shown in Fig 2.

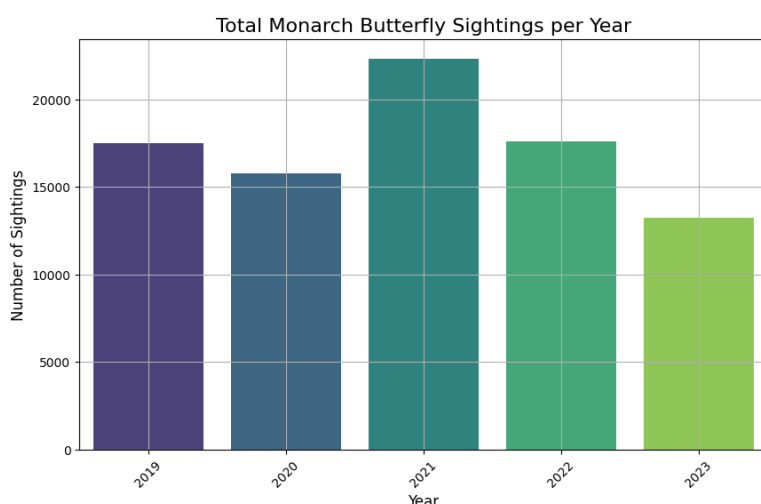


Fig 2. Sightings of a Monarch Butterfly Population 2019 through 2023

We also discovered that climate change is intensifying these effects. In some regions, monarchs arrived earlier than usual due to warmer temperatures, which disrupted their migration and breeding patterns as shown in Fig 3. These disruptions had cascading effects, such as monarchs arriving before milkweed—a crucial plant for their lifecycle—was available.



Fig 3. Sightings of a Monarch Butterfly Population 2019 through 2023

Based on the air quality analysis, we found that the further stages of migration namely through Mexico and California influenced the decrease in the population of Monarch Butterflies due to the high AQI index in the counties of these 2 states.

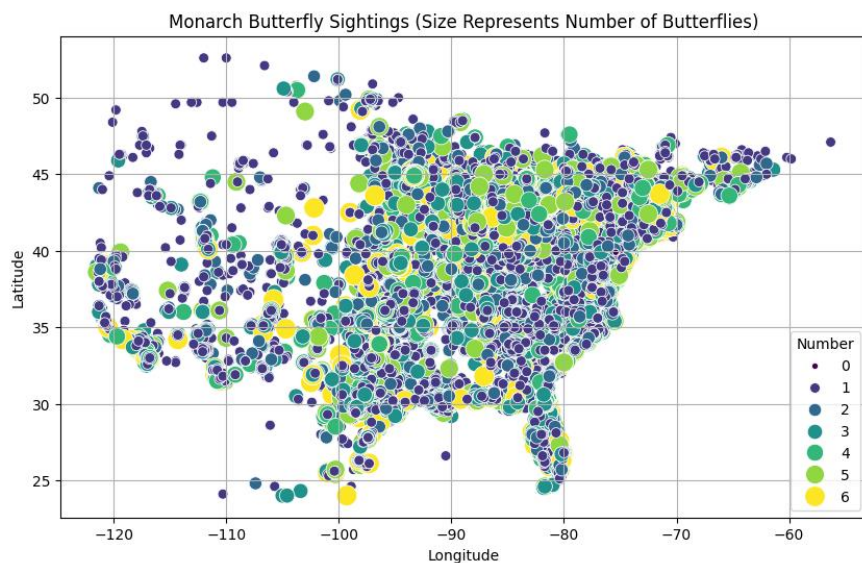


Fig 3. Monarch Butterfly Sightings 2019 through 2023 Across the Country

The above figure demonstrates the sighting of Monarch Butterflies across the country that are reported by volunteers spanning multiple cities.

7. Personal Reflections

Working on this project as a duo allowed us to push our individual strengths while learning from each other. The experience wasn't just about analyzing data; it was about collaboration, problem-solving, and applying our skills to address real-world ecological challenges.

There were moments when the complexity of the data seemed overwhelming, but we managed to keep each other motivated. We both learned the importance of optimizing workflows, leveraging APIs, and presenting complex data in meaningful ways. This project solidified our belief that data science can have a tangible impact on pressing global issues like biodiversity loss.

8. Conclusion

The decline of monarch butterflies is a multi-faceted issue with far-reaching consequences. Through our combined efforts, we identified several critical factors contributing to this decline, including pesticide use, air quality, and climate change. Our findings emphasize the need for immediate action in conservation efforts.

As a team, we're proud of the work we've done on this project and look forward to applying the lessons we've learned to future challenges. The collaboration and technical skills we honed during this project will serve as a strong foundation for our future endeavors in ecological data science.

Christina R Duthie | Ubaidullah Khan
The University of Texas at San Antonio