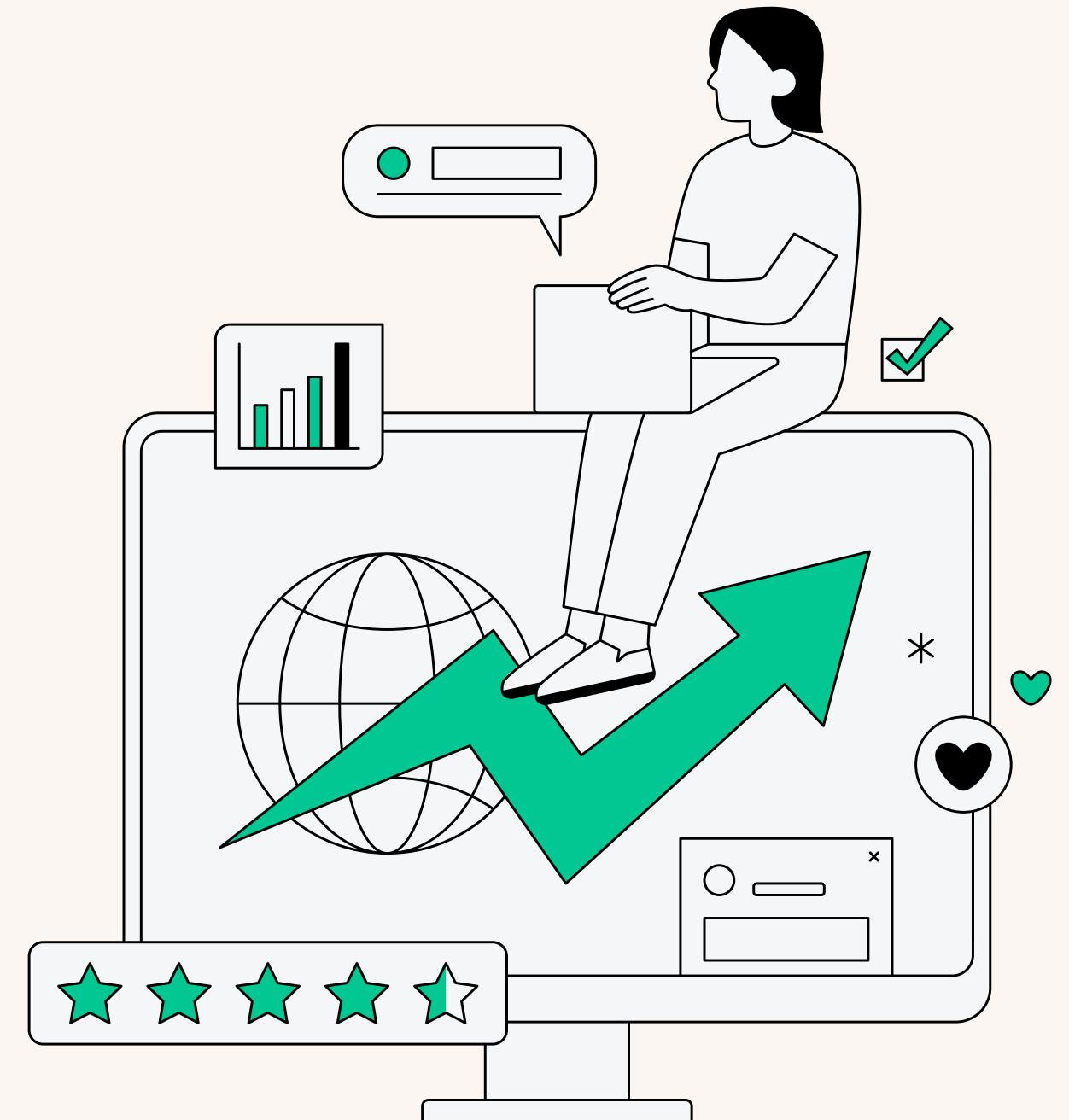


Presented by Tilova Shahrin

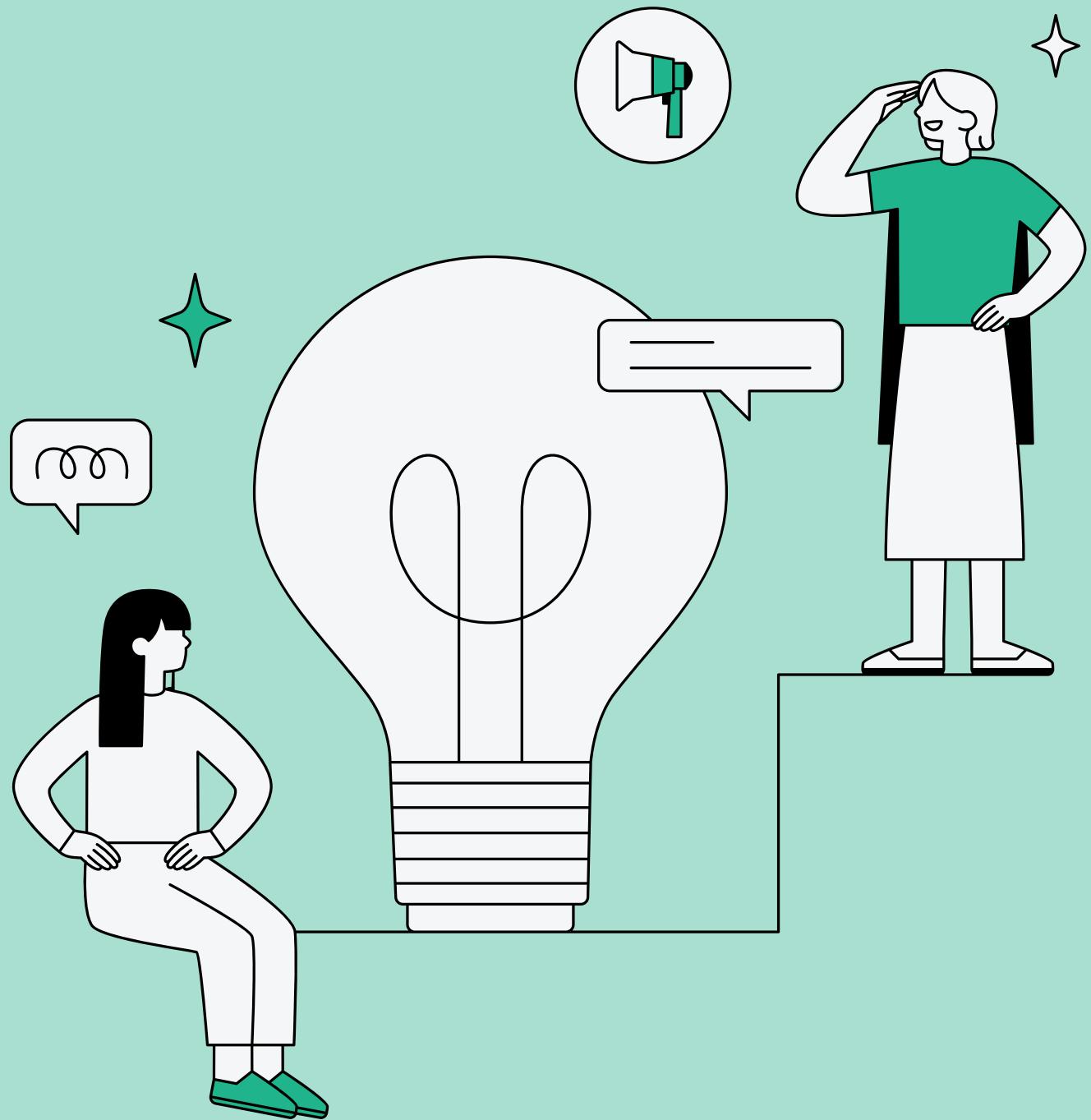
Leveraging Machine Learning for Ticket Data Analysis

github.com/tilovashahrin



Smart Parking Solutions for Citizen Convenience

- Parking tickets are a common inconvenience for citizens, leading to frustration and unnecessary expenses.
- Reduced Parking Violations: By providing accurate information and reminders, citizens are less likely to inadvertently violate parking regulations, leading to a decrease in parking tickets.
- Time and Cost Savings: Citizens save time and money by avoiding fines and towing fees associated with parking violations.
- Improved community satisfaction for local government services and infrastructure.



Dataset Overview and Preprocessing

Parking ticket data collected from
Toronto's open data portal.

- Geocoding:
 - Converting addresses into coordinates (latitude, longitude).
 - Adding city, province, and country for accurate geocoding.
 - Reduced API calls by identifying and processing unique addresses.
 - Incorporated latitude and longitude data into the dataset.
- Extracted year, month, and day from the Date-time of infraction column for temporal analysis.
- Categorized infractions into groups using keywords for classification



Data Analysis Findings

Infraction Time

- Most tickets issued between 11 AM and 12 PM

Infractions by Day of the Week

- Tuesday has the most tickets of the week, followed by Wednesday

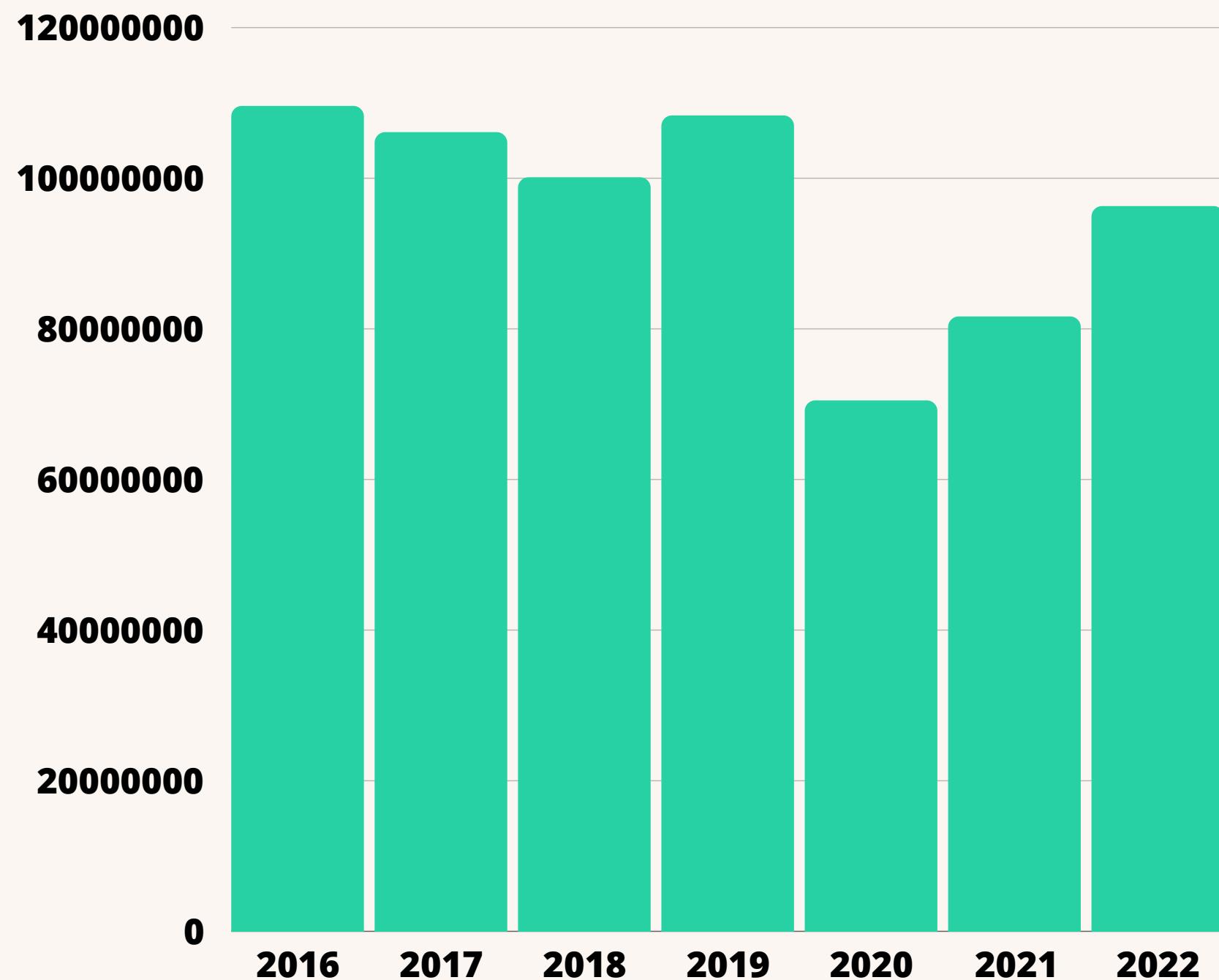
Infraction by Year

- Peak in spring/summer months every year
- Bounce in October, potentially due to holidays

Total Fine

- Increased slightly over the years, with a dip in 2020 likely due to COVID-19

Total Fines by Year

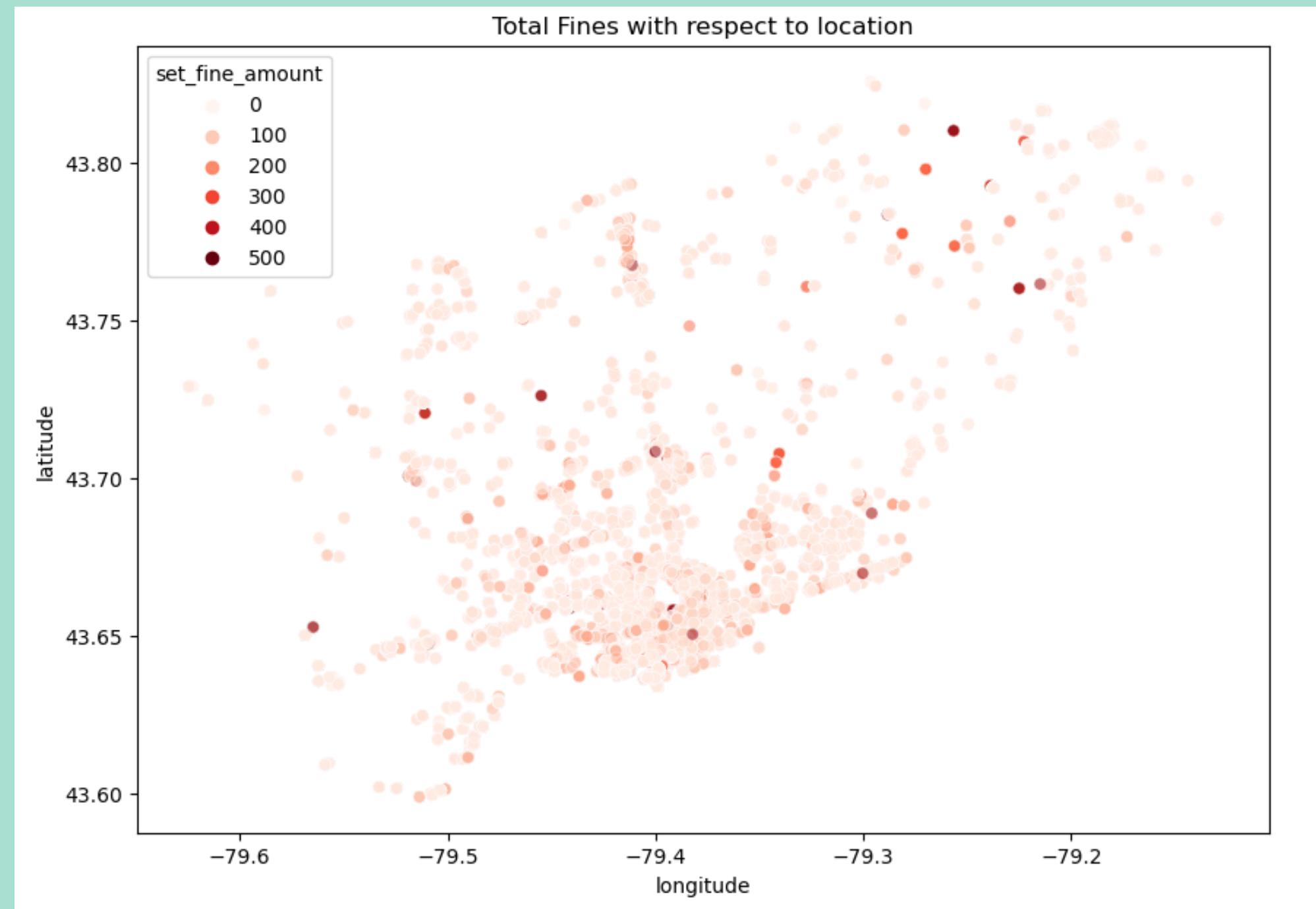


Overview of ticket fines

Description: This plot visualizes parking fine locations in Toronto using latitude and longitude points, with hue indicating fine amounts.

Key Points:

- No distinct patterns in fine amounts.
- More fines near the city center.
- Helps inform urban planning and parking enforcement strategies.



Machine Learning Modeling

K Nearest Neighbors (KNN)

- Evaluated using R-squared score for regression (predicting fine amounts).
- KNN achieved an R-squared score of around 0.37 for predicting fine amounts based on coordinates.

Decision Tree Classifier

- Use location data to forecast types of infractions.
- Types
 - Accessible spots
 - Fire Route
 - Private Property
 - Fee not paid/enough hours for meter

Prophet for Time Series

Prophet successfully forecasted the number of tickets for the next year, capturing seasonal patterns and external events like COVID-19.

Next Steps

Feature Engineering

- Explore spatial and temporal features (neighbourhood demographics).

Advanced Modeling Techniques

- Experiment with other methods such as Random Forest and Gradient Boosting for improved predictive accuracy.

Integration of External Data Sources:

- Incorporate external datasets such as weather data, public events calendars, and road construction schedules to enrich the analysis.



Presented by Sandra Haro

Thank you very much!

www.reallygreatsite.com

