Weeks 5 & 6 Exercise – Kia Thefts

Sreenivasulu Somu
Department of Data Science, Bellevue University
DSC640-T301 Data Presentation & Visualization (2255-1)
Benjamin Schneider
Apr 20, 2025

Data Science Analysis and Visualization of Kia Thefts

Audience

The primary audience for this analysis includes city policymakers, law enforcement agencies, automotive industry stakeholders (especially Kia and Hyundai), and data-driven journalists. Secondary audiences are insurance companies, urban planners, and concerned citizens seeking to understand the surge in vehicle thefts, particularly for Kia and Hyundai models.

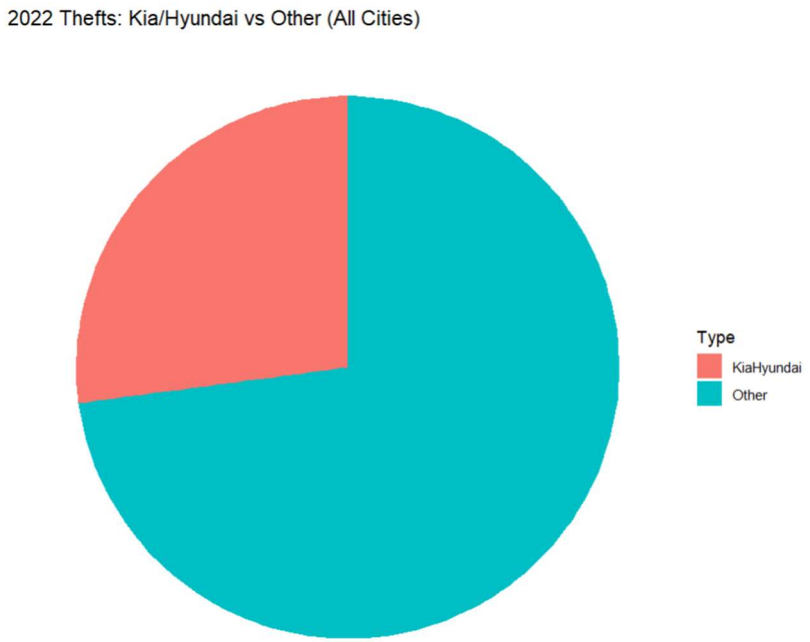
Purpose

The purpose of this project is to provide a comprehensive, data-driven overview of the Kia/Hyundai theft phenomenon across multiple U.S. cities from 2019 to 2022. By leveraging four diverse datasets (city-level CSVs, national agency data, and an Excel compilation), the analysis aims to:

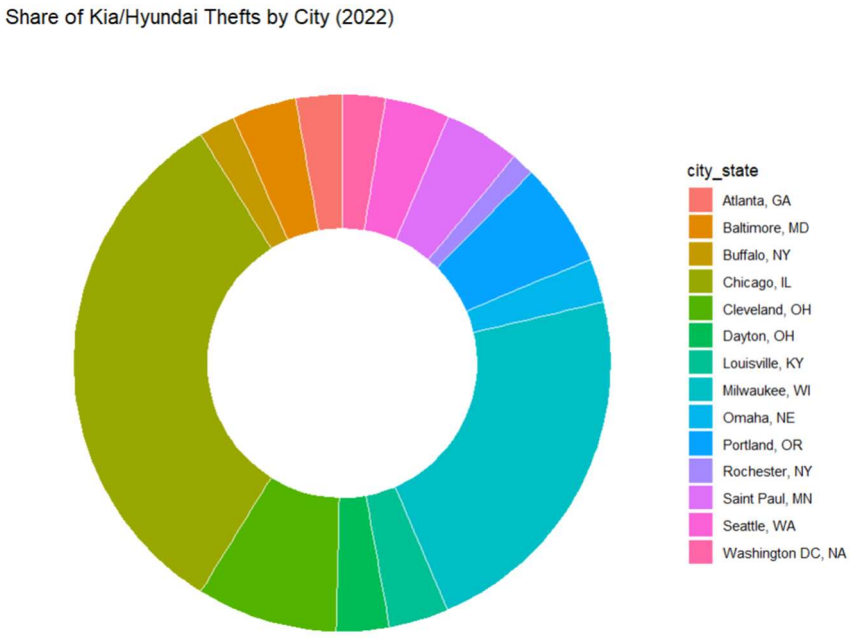
- Quantify the scale and growth of thefts.
- Identify geographic and temporal hotspots.
- Compare Kia/Hyundai thefts to other vehicles.
- Inform prevention strategies and policy responses.

Chart Explanations

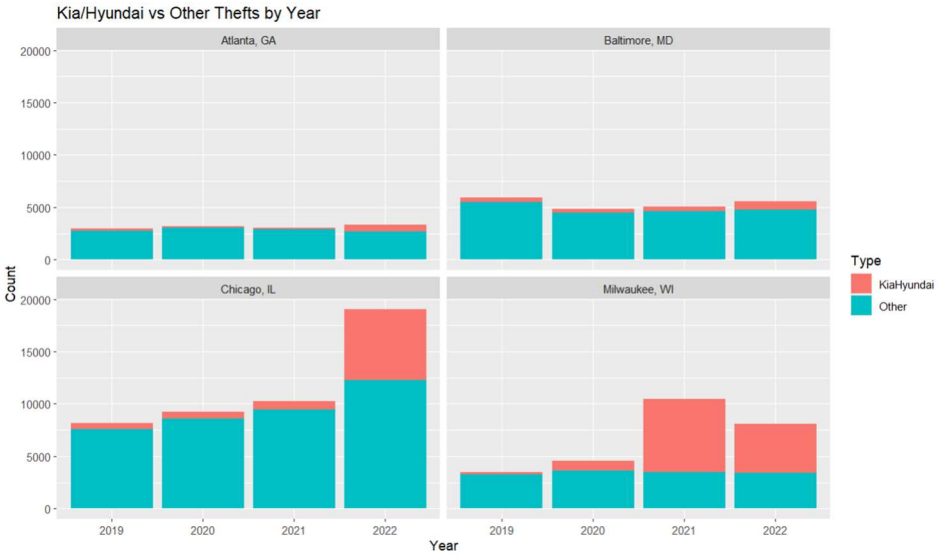
Pie Chart: Shows the 2022 proportion of all reported vehicle thefts that were Kia/Hyundai versus all other makes, highlighting the dramatic increase in thefts of these brands.



Donut Chart: Displays the share of Kia/Hyundai thefts by city for 2022, allowing quick visual comparison of which cities are most affected by the surge in these thefts.

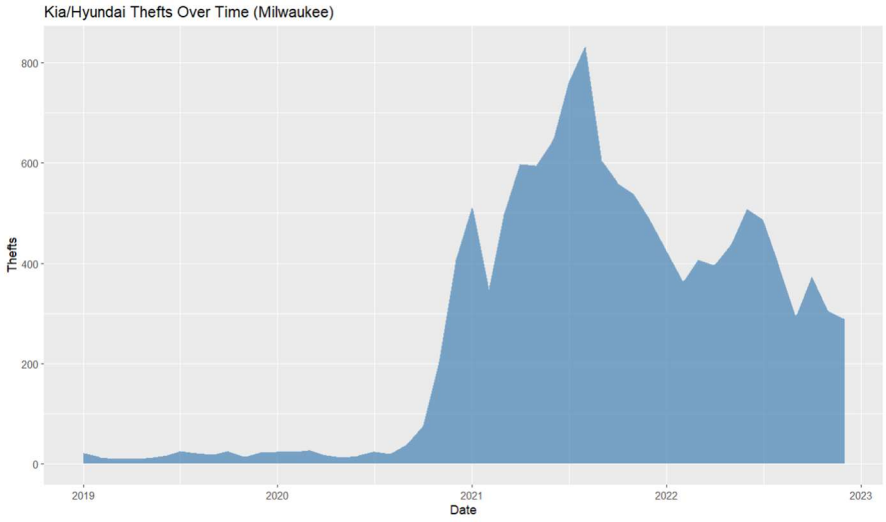


Stacked Bar Chart: Compares annual Kia/Hyundai thefts to other vehicle thefts in major cities, revealing both the rise in total thefts and the shifting composition toward these brands.

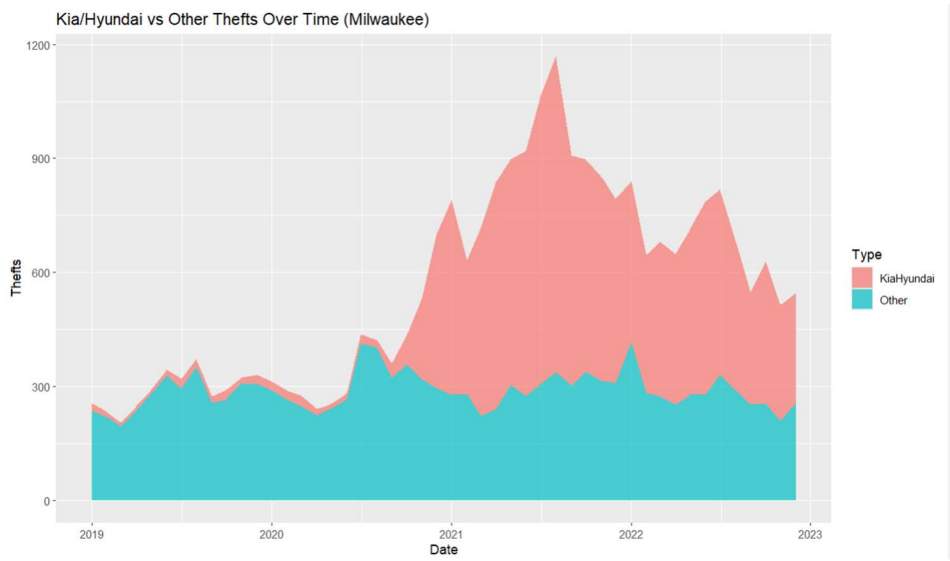


Tree Map: Visualizes the top 30 law enforcement agencies by total car thefts in 2022, with area size representing theft volume, making it easy to spot major hotspots.

Area Chart: Tracks monthly Kia/Hyundai thefts in Milwaukee over four years, clearly illustrating the time trend and the explosive spike in late 2020 and 2021.



Stacked Area Chart: Shows the monthly evolution of Kia/Hyundai versus other vehicle thefts in Milwaukee, highlighting how Kia/Hyundai thefts overtook other brands during the studied period.



Medium Selection

- The PDF output format is chosen for accessibility and clarity, suitable for stakeholders, policymakers, and the public.
- This will ensure that charts are easily shared, printed, and viewed across devices.
- Visuals are designed for non-technical and technical audiences, supporting informed decision-making.
- Ethical transparency is maintained by using aggregate, anonymized data, clear labeling, and honest representation of findings.
- High-quality rendering of R-generated charts and preserves layout integrity.
- Ideal for sharing complex visual data with policymakers, law enforcement, and stakeholders.

Design Strategy

- Uses tidyverse for efficient data wrangling and ggplot2 for consistent, high-quality visualizations.
- Selected chart types (pie, donut, stacked bar, tree map, area, stacked area) to match data structure and analytical goals.
- Applied clear labeling, color coding, and legends for accessibility and interpretability.
- Integrated multiple datasets for comprehensive, comparative insights across cities and years.
- Consistent color palettes distinguish Kia/Hyundai thefts from others, aiding rapid interpretation.

Ethical Considerations

- **Privacy:** All data used is aggregated at city or agency level, with no personally identifiable information, protecting individual privacy.
- **Accuracy:** Data cleaning and validation steps were taken to avoid misrepresentation.

- **Transparency:** Methods and sources are clearly documented, enabling reproducibility.
- **Impact:** Care was taken to avoid stigmatizing specific communities or brands; the focus is on systemic trends and prevention, not blame.