# Kia Thefts

Akshay Sharma

2025-07-09

## Load Datasets

```
kia_data <- read_csv("C:/Users/aksha/Downloads/kiaHyundaiThefts.csv")
```

```
## Rows: 552 Columns: 7
## -- Column specification --------------------------------------------------------
## Delimiter: ","
## chr (3): month, city, state
## dbl (4): year, countKiaHyundaiThefts, countOtherThefts, percentKiaHyundai
##
## i Use `spec()` to retrieve the full column specification for this data.
## i Specify the column types or set `show_col_types = FALSE` to quiet this message.
```

```
milwaukee <- read_csv("C:/Users/aksha/Downloads/KiaHyundaiMilwaukeeData.csv")
```

```
## Rows: 48 Columns: 7
## -- Column specification --------------------------------------------------------
## Delimiter: ","
## chr (3): month, city, state
## dbl (4): year, countKiaHyundaiThefts, countOtherThefts, percentKiaHyundai
##
## i Use `spec()` to retrieve the full column specification for this data.
## i Specify the column types or set `show_col_types = FALSE` to quiet this message.
```

```
map_data <- read_csv("C:/Users/aksha/Downloads/carTheftsMap.csv")
```

```
## Rows: 556 Columns: 9
## -- Column specification --------------------------------------------------------
## Delimiter: ","
## chr (4): agency_ori, geo_name, countCarThefts2020, countCarThefts2021
## dbl (3): latitude, longitude, percentChange2019to2022
## num (2): countCarThefts2019, countCarThefts2022
##
## i Use `spec()` to retrieve the full column specification for this data.
## i Specify the column types or set `show_col_types = FALSE` to quiet this message.
```

```
motherboard <- read_excel("C:/Users/aksha/Downloads/Motherboard_VICE_News_Kia_Hyundai_Theft_Data.xlsx")
```

```
## New names:
## * `` -> `...1`
## * `` -> `...3`
## * `` -> `...4`
## * `` -> `...6`
## * `` -> `...7`
## * `` -> `...9`
```

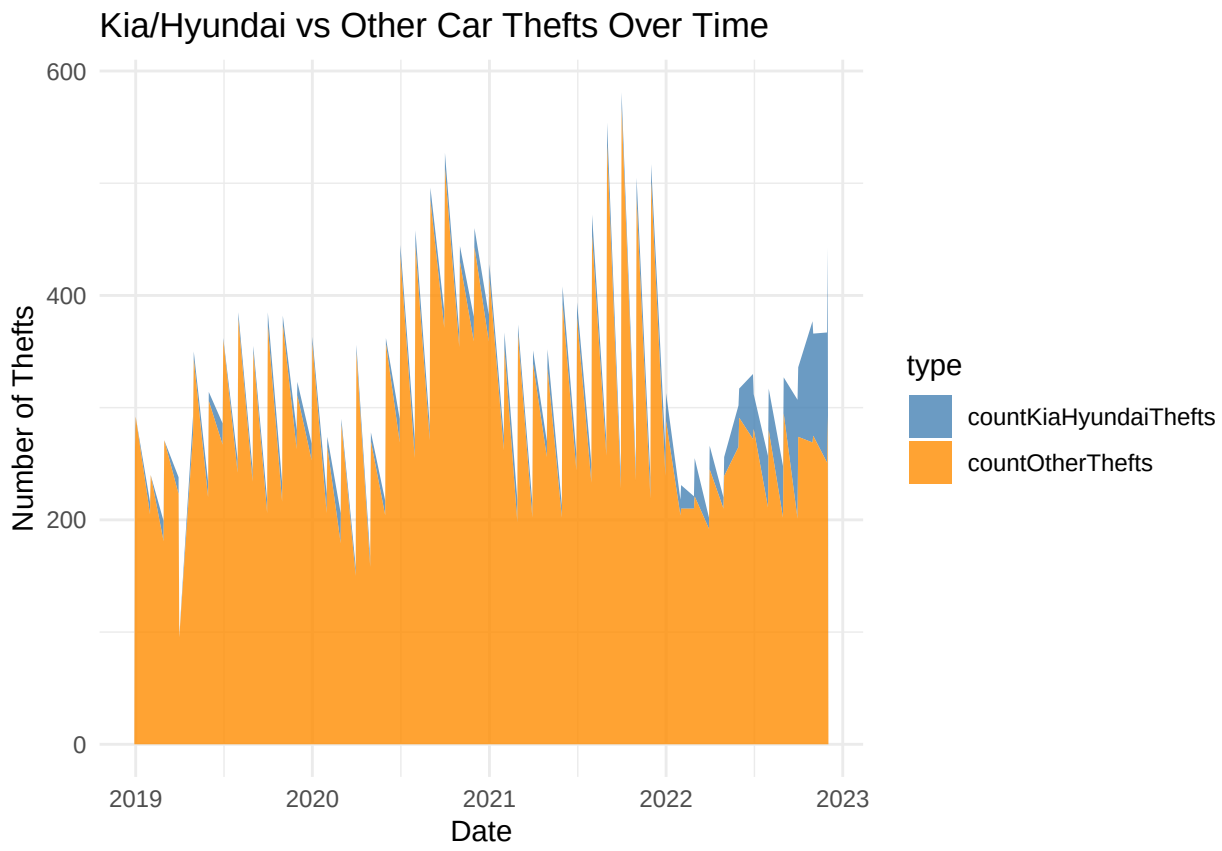
```
## * `` -> `...10`
## * `` -> `...12`
## * `` -> `...13`
## * `` -> `...15`
## * `` -> `...16`
## * `` -> `...18`
## * `` -> `...19`
## * `` -> `...21`
## * `` -> `...22`
## * `` -> `...24`
## * `` -> `...25`
## * `` -> `...27`
## * `` -> `...28`
## * `` -> `...30`
## * `` -> `...31`
## * `` -> `...33`
## * `` -> `...34`
## * `` -> `...36`
## * `` -> `...37`
## * `` -> `...39`
## * `` -> `...40`
## * `` -> `...42`
## * `` -> `...43`
## * `` -> `...45`
## * `` -> `...46`
## * `` -> `...48`
## * `` -> `...49`
## * `` -> `...51`
## * `` -> `...52`
## * `` -> `...54`
## * `` -> `...55`
## * `` -> `...57`
## * `` -> `...58`
## * `` -> `...60`
## * `` -> `...61`
## * `` -> `...63`
## * `` -> `...64`
## * `` -> `...66`
## * `` -> `...67`
## * `` -> `...69`
## * `` -> `...70`
## * `` -> `...72`
## * `` -> `...73`
## * `` -> `...75`
## * `` -> `...76`
## * `` -> `...78`
## * `` -> `...79`
## * `` -> `...81`
## * `` -> `...82`
## * `` -> `...84`
## * `` -> `...85`
## * `` -> `...87`
## * `` -> `...88`
## * `` -> `...90`
```

```
## * `` -> `...91`
## * `` -> `...93`
## * `` -> `...94`
## * `` -> `...96`
## * `` -> `...97`
## * `` -> `...99`
## * `` -> `...100`
## * `` -> `...102`
## * `` -> `...103`
## * `` -> `...105`
## * `` -> `...106`
## * `` -> `...108`
## * `` -> `...109`
## * `` -> `...111`
## * `` -> `...112`
## * `` -> `...114`
## * `` -> `...115`
## * `` -> `...117`
## * `` -> `...118`
## * `` -> `...120`
## * `` -> `...121`
## * `` -> `...123`
## * `` -> `...124`
## * `` -> `...126`
## * `` -> `...127`
## * `` -> `...129`
## * `` -> `...130`
## * `` -> `...132`
## * `` -> `...133`
## * `` -> `...135`
## * `` -> `...136`
## * `` -> `...138`
## * `` -> `...139`
## * `` -> `...141`
## * `` -> `...142`
## * `` -> `...144`
## * `` -> `...145`
## * `` -> `...147`
## * `` -> `...148`
## * `` -> `...150`
## * `` -> `...151`
## * `` -> `...153`
## * `` -> `...154`
## * `` -> `...156`
## * `` -> `...157`
## * `` -> `...159`
## * `` -> `...160`
## * `` -> `...162`
## * `` -> `...163`
## * `` -> `...165`
## * `` -> `...166`
## * `` -> `...168`
## * `` -> `...169`
## * `` -> `...171`
```

```
## * `` -> `...172`
## * `` -> `...174`
## * `` -> `...175`
## * `` -> `...177`
## * `` -> `...178`
## * `` -> `...180`
## * `` -> `...181`
## * `` -> `...183`
## * `` -> `...184`
## * `` -> `...186`
## * `` -> `...187`
## * `` -> `...189`
## * `` -> `...190`
## * `` -> `...192`
## * `` -> `...193`
## * `` -> `...195`
## * `` -> `...196`
## * `` -> `...198`
## * `` -> `...199`
## * `` -> `...201`
## * `` -> `...202`
## * `` -> `...204`
## * `` -> `...205`
## * `` -> `...207`
## * `` -> `...208`
## * `` -> `...210`
## * `` -> `...211`
```

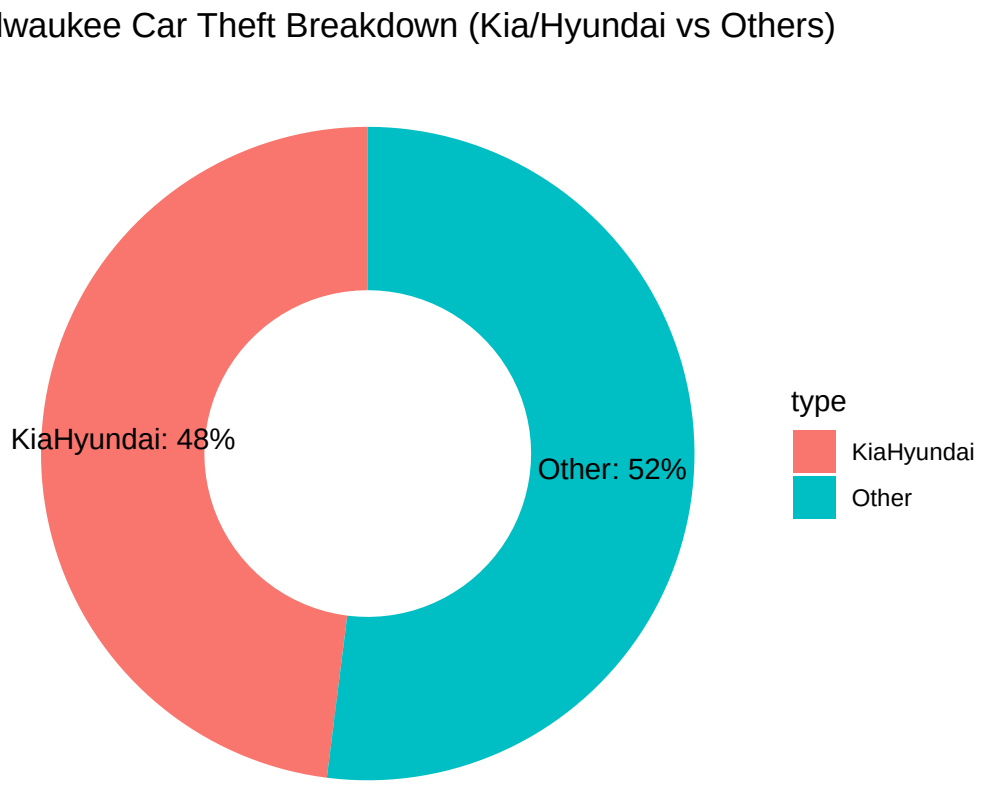
### 1. Stacked Area Chart

```r
kia_data <- kia_data %>%
  mutate(date = as.Date(paste(year, match(month, month.abb), "01", sep = "-"))) %>%
  select(date, countKiaHyundaiThefts, countOtherThefts) %>%
  pivot_longer(cols = starts_with("count"), names_to = "type", values_to = "count")

ggplot(kia_data, aes(x = date, y = count, fill = type)) +
  geom_area(alpha = 0.8) +
  scale_fill_manual(values = c("steelblue", "darkorange")) +
  theme_minimal() +
  labs(title = "Kia/Hyundai vs Other Car Thefts Over Time", x = "Date", y = "Number of Thefts")
```




## 2. Donut Chart

```
milwaukee_summary <- milwaukee %>%
  summarise(KiaHyundai = sum(countKiaHyundaiThefts), Other = sum(countOtherThefts)) %>%
  pivot_longer(cols = everything(), names_to = "type", values_to = "value") %>%
  mutate(percent = value / sum(value) * 100,
         label = paste0(type, ": ", round(percent, 1), "%"))

ggplot(milwaukee_summary, aes(x = 2, y = value, fill = type)) +
  geom_bar(stat = "identity", width = 1) +
  coord_polar("y") +
  theme_void() +
  geom_text(aes(label = label), position = position_stack(vjust = 0.5)) +
  xlim(0.5, 2.5) +
  labs(title = "Milwaukee Car Theft Breakdown (Kia/Hyundai vs Others)")
```



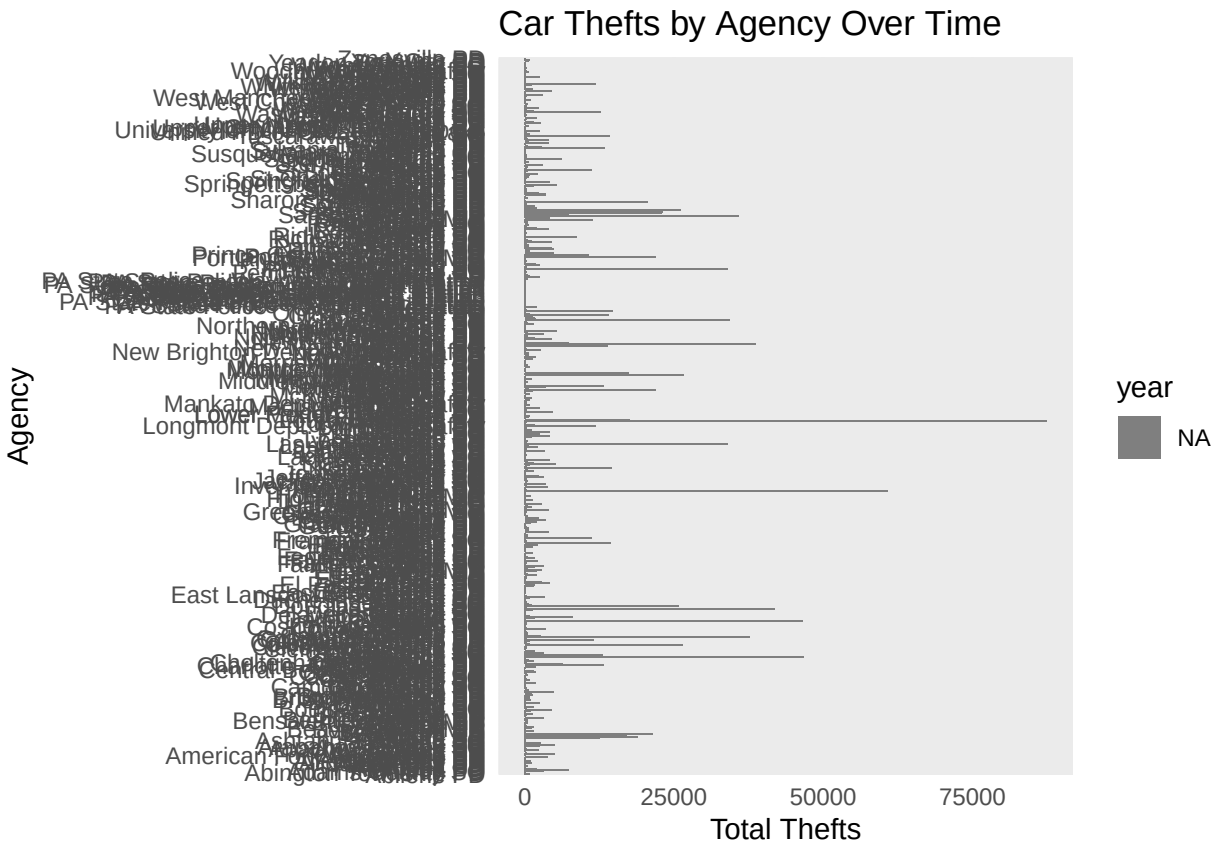

### 3. Stacked Bar Chart

```
map_data_clean <- map_data %>%
  mutate(across(starts_with("countCarThefts"), ~as.numeric(.)))
```

```
## Warning: There were 2 warnings in `mutate()`.
## The first warning was:
## i In argument: `across(starts_with("countCarThefts"), ~as.numeric(.))`.
## Caused by warning:
## ! NAs introduced by coercion
## i Run `dplyr::last_dplyr_warnings()` to see the 1 remaining warning.
```

```
map_data_long <- map_data_clean %>%
  pivot_longer(cols = starts_with("countCarThefts"), names_to = "year", values_to = "count") %>%
  mutate(year = str_extract(year, "\\\\d{4}"))

ggplot(map_data_long, aes(x = geo_name, y = count, fill = year)) +
  geom_bar(stat = "identity") +
  coord_flip() +
  labs(title = "Car Thefts by Agency Over Time", x = "Agency", y = "Total Thefts") +
  theme_minimal()
```

```
## Warning: Removed 9 rows containing missing values or values outside the scale range
## (`geom_bar()`).
```



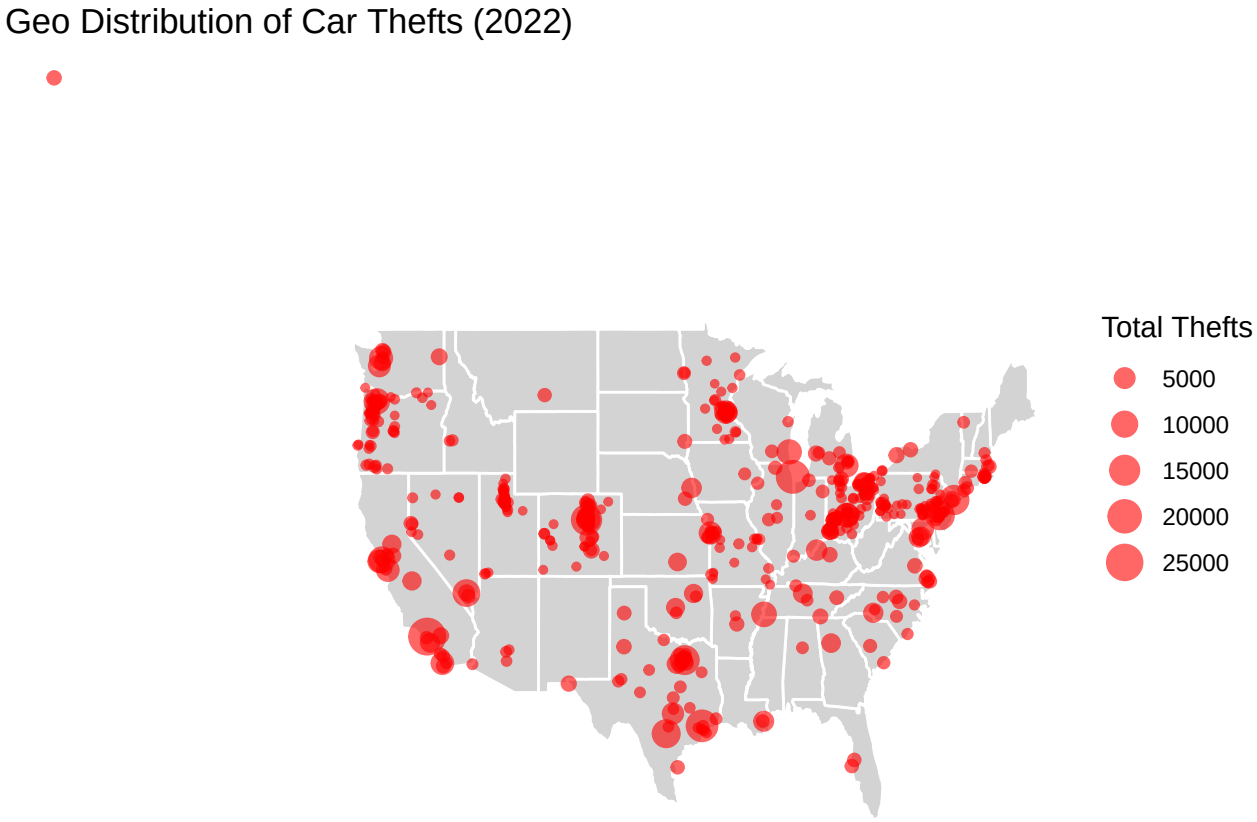

## 4. Tree Map (Skip If Missing)

```
city_cols <- names(motherboard)[str_detect(names(motherboard), "Kia/Hyundais")]

if (length(city_cols) > 0) {
  mother_trim <- motherboard %>%
    slice(2) %>%
    select(all_of(city_cols)) %>%
    pivot_longer(cols = everything(), values_to = "value", names_to = "city") %>%
    mutate(city = str_remove(city, " Kia/Hyundais"),
           value = as.numeric(value)) %>%
    filter(!is.na(value))

  ggplot(mother_trim, aes(area = value, fill = value, label = city)) +
    geom_treemap() +
    geom_treemap_text(colour = "white", place = "centre", grow = TRUE) +
    labs(title = "City-wise Kia/Hyundai Theft Distribution")
} else {
  message("Skipping tree map: No columns with 'Kia/Hyundais' found.")
}
```

```
## Skipping tree map: No columns with 'Kia/Hyundais' found.
```

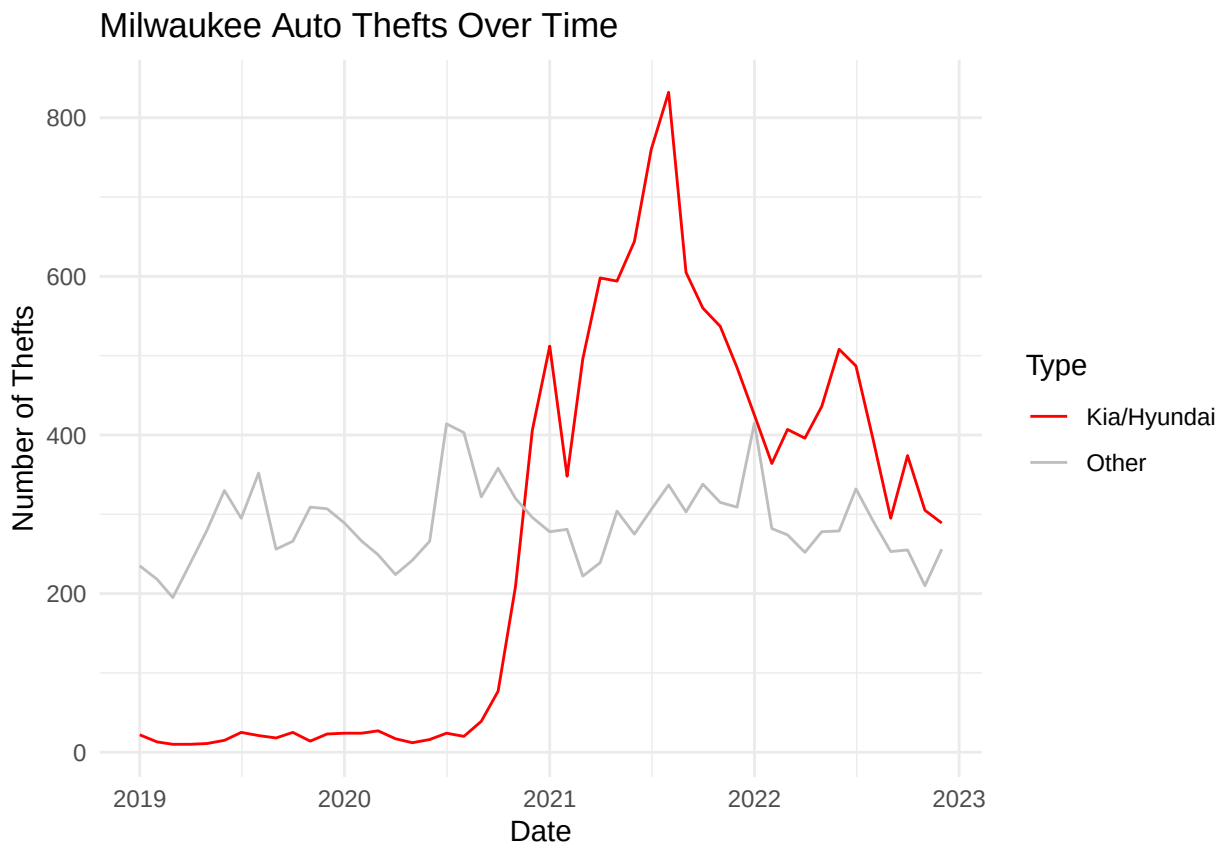
## 5. Geo Map

```
us_map <- map_data("state")

ggplot() +
  geom_polygon(data = us_map, aes(x = long, y = lat, group = group), fill = "lightgray", color = "white
  geom_point(data = map_data, aes(x = longitude, y = latitude, size = as.numeric(countCarThefts2022)),
  theme_void() +
  labs(title = "Geo Distribution of Car Thefts (2022)", size = "Total Thefts")
```




## 6. Line Chart

```
milwaukee <- milwaukee %>%
  mutate(date = as.Date(paste(year, match(month, month.abb), "01", sep = "-")))

ggplot(milwaukee, aes(x = date)) +
  geom_line(aes(y = countKiaHyundaiThefts, color = "Kia/Hyundai")) +
  geom_line(aes(y = countOtherThefts, color = "Other")) +
  scale_color_manual(values = c("Kia/Hyundai" = "red", "Other" = "gray")) +
  labs(title = "Milwaukee Auto Thefts Over Time", x = "Date", y = "Number of Thefts", color = "Type") +
  theme_minimal()
```




## Summary and Ethical Considerations

The following analysis uses publicly available data to examine the sharp increase in Kia and Hyundai vehicle thefts across the United States. This story is aimed at **local policymakers, public safety officials, and journalists**. Our goal is to increase awareness, identify geographic and temporal trends, and drive action—such as software updates, anti-theft measures, and targeted public awareness campaigns.

We presented our findings using an infographic-style dashboard containing six visuals. These included a **stacked area chart**, **donut chart**, **stacked bar**, **tree map**, **geo map**, and a **line chart**. Each visual was selected for its ability to emphasize either proportions, time-based trends, or spatial/geographic disparities. For example, the Milwaukee-focused line chart shows a dramatic rise in Kia/Hyundai thefts starting in late 2020, highlighting a local public safety concern.

The visual design adheres to several Gestalt principles—**proximity and similarity** are used to group related visuals together, while **color and contrast** help distinguish Kia/Hyundai thefts from general trends. Tree maps and geo maps were chosen to capture volume and distribution effectively across space.

From an ethical perspective: - **Data recency**: Most of the data extends through 2022, making it relatively current. - **Data coverage**: Some agencies and cities are missing; thus, findings reflect reported data only. - **Responsible interpretation**: We refrain from inferring causation. We present patterns, not blame. - **Transparency**: Assumptions—such as aggregated monthly trends and city-level comparisons—are clearly defined in both the visuals and summary.

Our **call to action** is straightforward: cities should implement targeted software patches for vulnerable Kia/Hyundai models and support public campaigns to deter theft. Citizens should be aware of whether their vehicle is at risk and consider anti-theft solutions.