

```

# Load libraries
library(tidyverse)
library(conflicted)
library(ggplot2)
library(scales)
library(readr)
library(lubridate)
library(readxl)

# Set preferred functions
conflicts_prefer(dplyr::filter)

# Load datasets
multi_city <- read_csv("KiaHyundaiThefts.csv")
car_thefts_map <- read_csv("carTheftsMap.csv")
milwaukee <- read_csv("KiaHyundaiMilwaukeeData.csv")
vice <- read_excel("C:/Users/Amy/Documents/Motherboard VICE News Kia Hyundai Theft Data.xlsx")

# --- National Pie Chart ---
national_totals <- multi_city %>%
  summarise(
    KiaHyundai = sum(countKiaHyundaiThefts, na.rm = TRUE),
    Other = sum(countOtherThefts, na.rm = TRUE)
  )

national_vector <- as.numeric(national_totals[1, ])
labels <- c("Kia/Hyundai", "Other")

png("national_pie_transparent.png", width = 600, height = 600,
    bg = "transparent")
pie(
  national_vector,
  labels = labels,
  col = c("#D55E00", "#009E73"),
  cex.main = 1.5
)
dev.off()

# --- Donut Chart: Milwaukee ---
milwaukee_totals <- milwaukee %>%
  summarise(KiaHyundai = sum(countKiaHyundaiThefts, na.rm = TRUE),
            Other = sum(countOtherThefts, na.rm = TRUE)) %>%
  pivot_longer(everything(), names_to = "Type", values_to = "Count") %>%
  mutate(Percentage = Count / sum(Count),
        ymax = cumsum(Percentage),
        ymin = c(0, head(ymax, -1)))

donut_plot <- ggplot(milwaukee_totals, aes(ymax = ymax, ymin = ymin,
                                             xmax = 4, xmin = 3, fill = Type)) +
  geom_rect() +
  coord_polar(theta = "y") +
  xlim(c(2, 4)) +
  scale_fill_manual(values = c("KiaHyundai" = "#D55E00", "Other" = "#009E73")) +
  labs(title = "Milwaukee Car Theft Share") +
  theme_void() +
  theme(
    plot.title = element_text(face = "bold", size = 16),
    legend.title = element_text(face = "bold"),
    legend.text = element_text(face = "bold")
  )
  ggsave("milwaukee_donut_transparent.png", plot = donut_plot, width = 6,
         height = 6, bg = "transparent")

```

```

# --- Stacked Bar Chart: Top Cities ---
top_cities <- multi_city %>%
  group_by(city) %>%
  summarise(KiaHyundai = sum(countKiaHyundaiThefts, na.rm = TRUE),
            Other = sum(countOtherThefts, na.rm = TRUE)) %>%
  slice_max(KiaHyundai, n = 6) %>%
  pivot_longer(cols = c(KiaHyundai, Other), names_to = "Type",
               values_to = "Count")

bar_plot <- ggplot(top_cities, aes(x = reorder(city, -Count), y = Count,
                                    fill = Type)) +
  geom_bar(stat = "identity") +
  labs(x = "City", y = "Thefts") +
  scale_fill_manual(values = c("KiaHyundai" = "#D55E00", "Other" = "#009E73")) +
  theme(
    panel.grid = element_blank(),
    panel.background = element_rect(fill = NA, color = NA),
    plot.background = element_rect(fill = NA, color = NA),
    axis.text = element_text(face = "bold"),
    axis.title = element_text(face = "bold"),
    plot.title = element_text(face = "bold")
  )

ggsave("bar_topcities_transparent.png", bar_plot, width = 7, height = 5,
       bg = "transparent")

# --- Area Chart: National Monthly Trends ---
multi_city_date <- multi_city %>%
  mutate(date = my(paste(month, year))) %>%
  group_by(date) %>%
  summarise(KiaHyundai = sum(countKiaHyundaiThefts, na.rm = TRUE),
            Other = sum(countOtherThefts, na.rm = TRUE)) %>%
  pivot_longer(cols = c(KiaHyundai, Other), names_to = "Type",
               values_to = "Thefts")

area_plot <- ggplot(multi_city_date, aes(x = date, y = Thefts, fill = Type)) +
  geom_area(alpha = 0.8) +
  labs(title = "Monthly Thefts: All Cities", x = "Date", y = "Thefts") +
  scale_fill_manual(values = c("KiaHyundai" = "#D55E00", "Other" = "#009E73")) +
  theme(
    panel.grid = element_blank(),
    panel.background = element_rect(fill = NA),
    plot.background = element_rect(fill = NA),
    axis.text = element_text(face = "bold"),
    axis.title = element_text(face = "bold"),
    plot.title = element_text(face = "bold")
  )

ggsave("area_national_transparent.png", area_plot, width = 7, height = 5,
       bg = "transparent")

# --- Area Chart: Milwaukee Monthly Trends ---
milwaukee_long <- milwaukee %>%
  mutate(date = my(paste(month, year))) %>%
  select(date, KiaHyundai = countKiaHyundaiThefts, Other = countOtherThefts) %>%
  pivot_longer(cols = c(KiaHyundai, Other), names_to = "Type",
               values_to = "Thefts")

area_mke <- ggplot(milwaukee_long, aes(x = date, y = Thefts, fill = Type)) +
  geom_area(alpha = 0.8) +
  labs(x = "Date", y = "Thefts") +
  scale_fill_manual(values = c("KiaHyundai" = "#D55E00", "Other" = "#009E73")) +
  theme(
    panel.grid = element_blank(),

```

```

panel.background = element_rect(fill = NA),
plot.background = element_rect(fill = NA),
axis.text = element_text(face = "bold"),
axis.title = element_text(face = "bold"),
plot.title = element_text(face = "bold")
)

ggsave("area_milwaukee_transparent.png", plot = area_mke, width = 7, height = 5,
bg = "transparent")

# --- Stacked Bar Chart: Top 10 Cities by Year ---
char_cols <- car_thefts_map %>%
  select(starts_with("countCarThefts")) %>%
  select(where(is.character)) %>%
  names()

car_thefts_map <- car_thefts_map %>%
  mutate(across(all_of(char_cols), parse_number))

thefts_long <- car_thefts_map %>%
  select(geo_name, starts_with("countCarThefts")) %>%
  pivot_longer(
    cols = starts_with("countCarThefts"),
    names_to = "Year",
    names_prefix = "countCarThefts",
    values_to = "Thefts"
  ) %>%
  mutate(Year = as.integer(Year))

top_cities <- car_thefts_map %>%
  arrange(desc(countCarThefts2022)) %>%
  slice(1:10) %>%
  pull(geo_name)

thefts_filtered <- thefts_long %>%
  filter(geo_name %in% top_cities)

cb_palette <- c("#D55E00", "#009E73", "#0072B2", "#F0E442")

stacked_bar_plot <- ggplot(thefts_filtered, aes(x = reorder(geo_name, -Thefts),
                                                y = Thefts,
                                                fill = factor(Year))) +
  geom_bar(stat = "identity") +
  scale_fill_manual(values = cb_palette) +
  labs(x = "City", y = "Number of Thefts", fill = "Year") +
  theme_minimal(base_size = 13) +
  theme(
    panel.grid = element_blank(),
    plot.background = element_rect(fill = NA),
    panel.background = element_rect(fill = NA),
    axis.text.x = element_text(angle = 45, hjust = 1, face = "bold"),
    axis.text.y = element_text(face = "bold"),
    axis.title = element_text(face = "bold"),
    plot.title = element_text(face = "bold", size = 16),
    legend.title = element_text(face = "bold"),
    legend.text = element_text(face = "bold")
  )
)

ggsave("car_thefts_stacked_bar_top10.png", plot = stacked_bar_plot, width = 10,
height = 6, bg = "transparent")

```