Aviation Wildlife Analysis

Fabian Rüger

This report presents some **basic analyses** of the aviation wildlife dataset available at Kaggle. The dataset contains information about wildlife strikes with military, commercial or civil aircrafts from 1990 to 2023.

```
library(tidyverse)
```

```
# load data
faa_wildlife <- read_csv(
  "../data/raw/faa_wildlife_strikes_1990_2023.csv",
  guess_max = Inf
)

# clean column names
faa_wildlife <- janitor::clean_names(faa_wildlife)
```

The dataset has 100 columns and 288810 observations.

Wildlife Strike Incidents Over Time

Figure 1 shows a positive trend in the number of wildlife strike incidents over time. The number of reported incidents has generally increased from around 2,500 incidents in 1990 to over 15,000 incidents in recent years. This increase could be attributed to various factors such as improved reporting mechanisms, increased air traffic, or changes in wildlife populations near airports.

```
faa_wildlife |>
  group_by(incident_year) |>
  summarise(total_strikes = n(), .groups = "drop") |>
  ggplot(aes(x = incident_year, y = total_strikes)) +
  geom_line(color = "#00688B") +
```

```
  geom_point(color = "#00688B") +
  scale_x_continuous(breaks = seq(1990, 2023, by = 2)) +
  scale_y_continuous(
    breaks = seq(0, 17500, by = 2500),
    labels = scales::label_number(big.mark = ",")
  ) +
  labs(
    title = "Total Wildlife Strike Incidents per Year",
    x = "Year",
    y = "Number of Strikes"
  ) +
  theme_minimal() +
  theme(panel.grid.minor = element_blank())
```

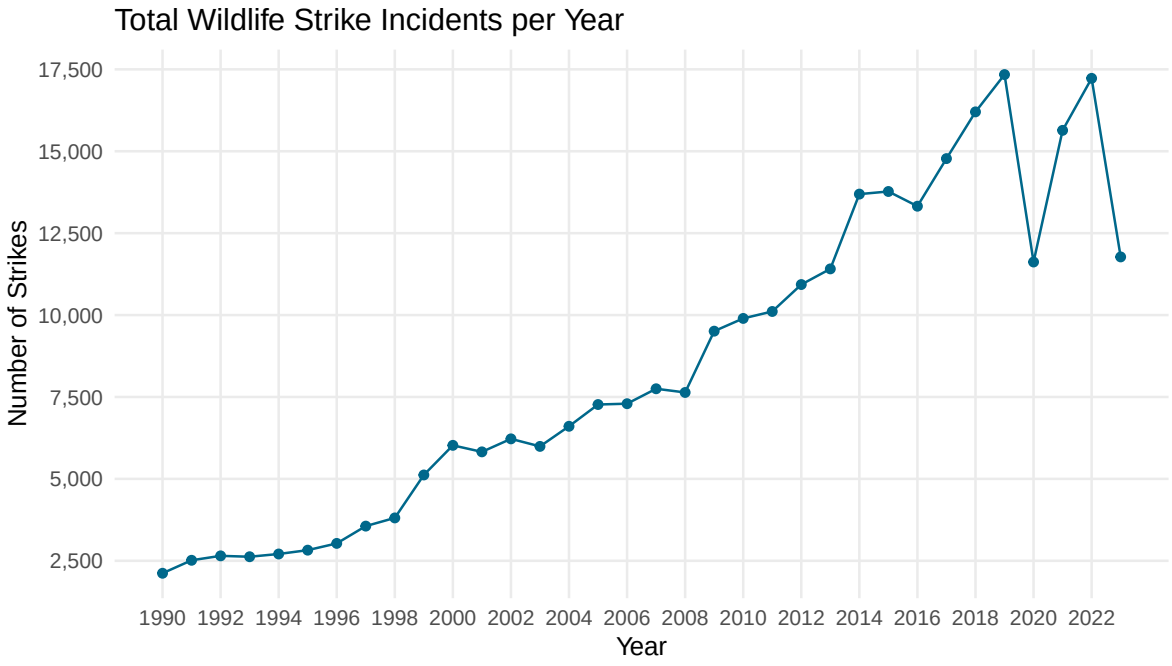

Figure 1: Total Wildlife Strike Incidents per Year

Figure 2 shows the average number of wildlife strike incidents by month. There is a clear seasonal pattern, with the highest number of incidents occurring during the spring and summer months (April to August). This trend may be related to increased wildlife activity during these months, as well as higher air traffic volumes.

```r
faa_wildlife |>
  group_by(incident_month, incident_year) |>
  summarise(n = n()) |>
  summarise(mean_strikes = mean(n)) |>
  ggplot(aes(x = incident_month, y = mean_strikes)) +
  geom_line(color = "#00688B") +
  geom_point(color = "#00688B") +
  scale_x_continuous(breaks = 1:12, labels = month.abb) +
  scale_y_continuous(labels = scales::label_number(big.mark = ",")) +
  labs(
    title = "Average Number of Wildlife Strike Incidents by Month",
    x = "Month",
    y = "Number of Strikes"
  ) +
  theme_minimal() +
  theme(panel.grid.minor = element_blank())
```

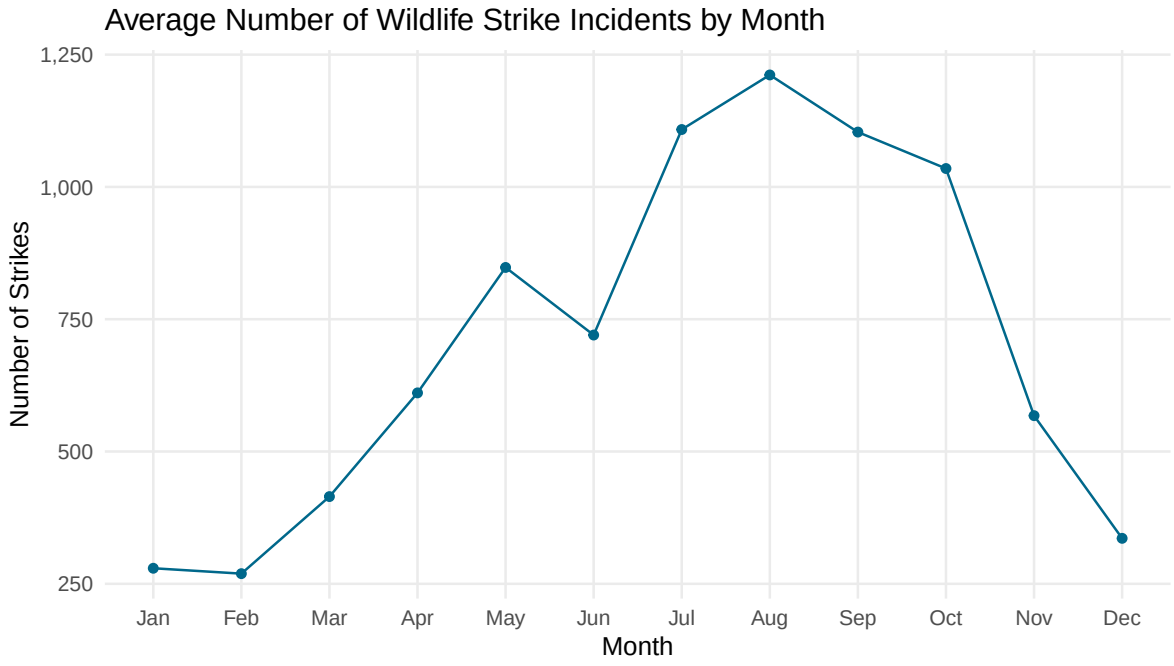

Figure 2: Average Number of Wildlife Strike Incidents by Month