Project Title :- EDA and data visualization of Bird Strikes between 2000 – 2011

Dataset :- Bird Strikes Dataset (2000 – 2011)

By :- Sarthak Priyank Verma (UMIP8357)

Project Link :- Bird Strike

Problem Statement :-

Transport and communication, vital sectors in analytics, face pressing concerns regarding environmental impact and safety, particularly in urban areas with growing populations and vehicle numbers. As solutions are sought, emerging technologies like artificial intelligence, notably Multi-Agent Systems (MAS), are being recognized for their potential in addressing these challenges effectively. Bird strikes, defined as collisions between birds and aircraft, pose significant safety risks, often causing damage to aircraft structures and engine failures, with potentially fatal consequences. These strikes are most likely during take-off, initial climb, approach, and landing phases due to the higher presence of birds at lower altitudes. The following document visually presents FAA-collected data on bird strikes from 2000 to 2011.

Expected Outcomes :-

- **Trends Over Time:**
  Identification of trends in bird strikes over the years, including any increase or decrease in frequency.
  Analysis of seasonal variations in bird strikes.
- **Cost Implications:**
  Assessment of the total cost incurred due to bird strikes over the years.
  Understanding the financial impact of bird strikes on airlines and aviation industry stakeholders.
- **Impact on Flights:**
  Evaluation of the effect of bird strikes on flight operations, including delays, diversions, and cancellations.
  Analysis of the severity of incidents, including injuries to passengers and damage to aircraft.
- **Factors Influencing Strikes:**
  Identification of factors such as altitude, time of day, and geographical location that influence the likelihood of bird strikes.
  Assessment of the effectiveness of prior warnings in mitigating the impact of bird strikes.
- **Preventive Measures:**
  Insights into the effectiveness of measures such as wildlife management programs and pilot warnings in reducing bird strike incidents.

Targeted Variables :-

- **Date and Time:**
  Recording when the bird strike occurred, including the specific date and time of day.
- **Location:**
  Identifying where the bird strike incident took place, including the airport code or geographical coordinates.
- **Flight Information:**
  Capturing details about the flight involved, such as the airline, flight number, and aircraft type.
- **Bird Species:**
  Documenting the species of bird involved in the strike, if identifiable.
- **Flight Phase:**
  Categorizing the phase of flight during which the strike occurred, such as takeoff, landing, or cruise.
- **Altitude:**
  Noting the altitude of the aircraft at the time of the bird strike.
- **Weather Conditions:**
  Describing the weather conditions prevailing at the time of the incident, including factors like visibility, wind speed, and precipitation.
- **Damage Assessment:**
  Assessing the extent of damage to the aircraft resulting from the bird strike.

- **Cost:**
  Estimating the financial cost associated with the incident, including repair costs and operational impacts.
- **Injuries or Fatalities:**
  Recording any injuries to passengers or crew and any fatalities resulting from the bird strike.

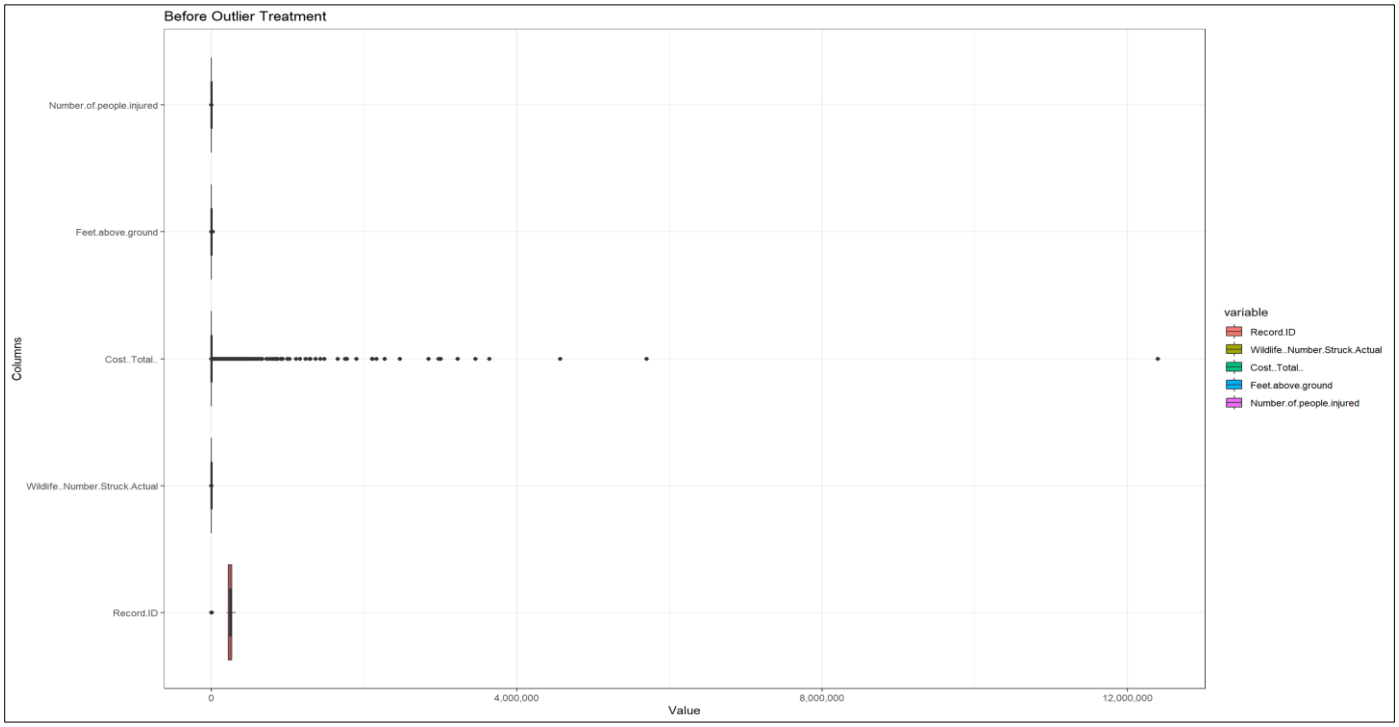
Libraries Used :-

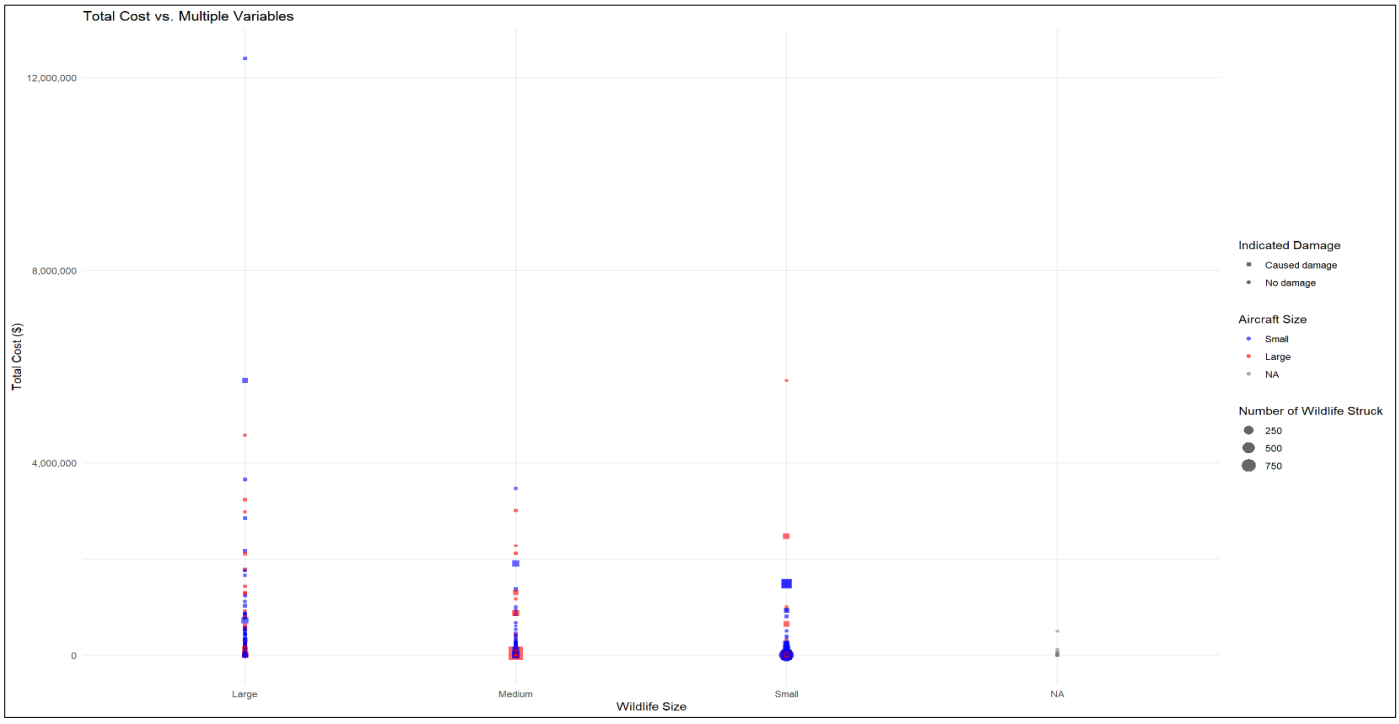
- **dplyr:** for data manipulation tasks like filtering, selecting, mutating, and summarizing data.
- **plotly:** for creating interactive plots and visualizations.
- **tidyr:** for data tidying tasks like reshaping data frames, gathering, and spreading columns.
- **ggplot2:** for creating static plots and visualizations based on the grammar of graphics.
- **reshape2:** simplifies data reshaping and manipulation tasks in R. (melt function)
- **stringr:** used for string manipulation tasks.
- **lubridate:** used for handling date-time objects and operations.
- **gridExtra:** used to make extra grids to combine the plots side by side.

Column Description (Metadata) :-

- **Record ID:** Unique identifier for each record in the dataset.
- **Aircraft: Type:** Type of aircraft involved in the bird strike incident.
- **Airport: Name:** Name of the airport where the bird strike incident occurred.
- **Altitude bin:** Categorized altitude range of the aircraft at the time of the bird strike.
- **Aircraft: Make/Model:** Make and model of the aircraft involved.
- **Wildlife: Number struck:** Number of wildlife struck as reported.
- **Wildlife: Number Struck Actual:** Actual number of wildlife struck as determined.
- **Effect: Impact to flight:** Impact of the bird strike on the flight operations.
- **FlightDate:** Date of the bird strike incident.
- **Effect: Indicated Damage:** Indicated damage caused by the bird strike.
- **Aircraft: Number of engines?:** Number of engines on the aircraft involved.
- **Aircraft: Airline/Operator:** Airline or operator of the aircraft.
- **Origin State:** State of origin of the flight.
- **When: Phase of flight:** Phase of flight during which the bird strike occurred.
- **Conditions: Precipitation:** Precipitation conditions at the time of the bird strike.
- **Remains of wildlife collected?:** Whether remains of wildlife were collected at the scene.
- **Remains of wildlife sent to Smithsonian:** Whether remains of wildlife were sent to the Smithsonian Institution.
- **Remarks:** Additional remarks or notes about the bird strike incident.
- **Wildlife: Size:** Size category of the wildlife involved.
- **Conditions: Sky:** Sky conditions at the time of the bird strike.
- **Wildlife: Species:** Species of wildlife involved in the bird strike.
- **Pilot warned of birds or wildlife?:** Whether the pilot was warned of birds or wildlife before the incident.
- **Cost: Total $:** Total cost associated with the bird strike incident.
- **Feet above ground:** Altitude of the aircraft above ground at the time of the bird strike.
- **Number of people injured:** Number of people injured as a result of the bird strike.
- **Is Aircraft Large?:** Indicator of whether the aircraft is large.

CHECK THE OUTLIERS

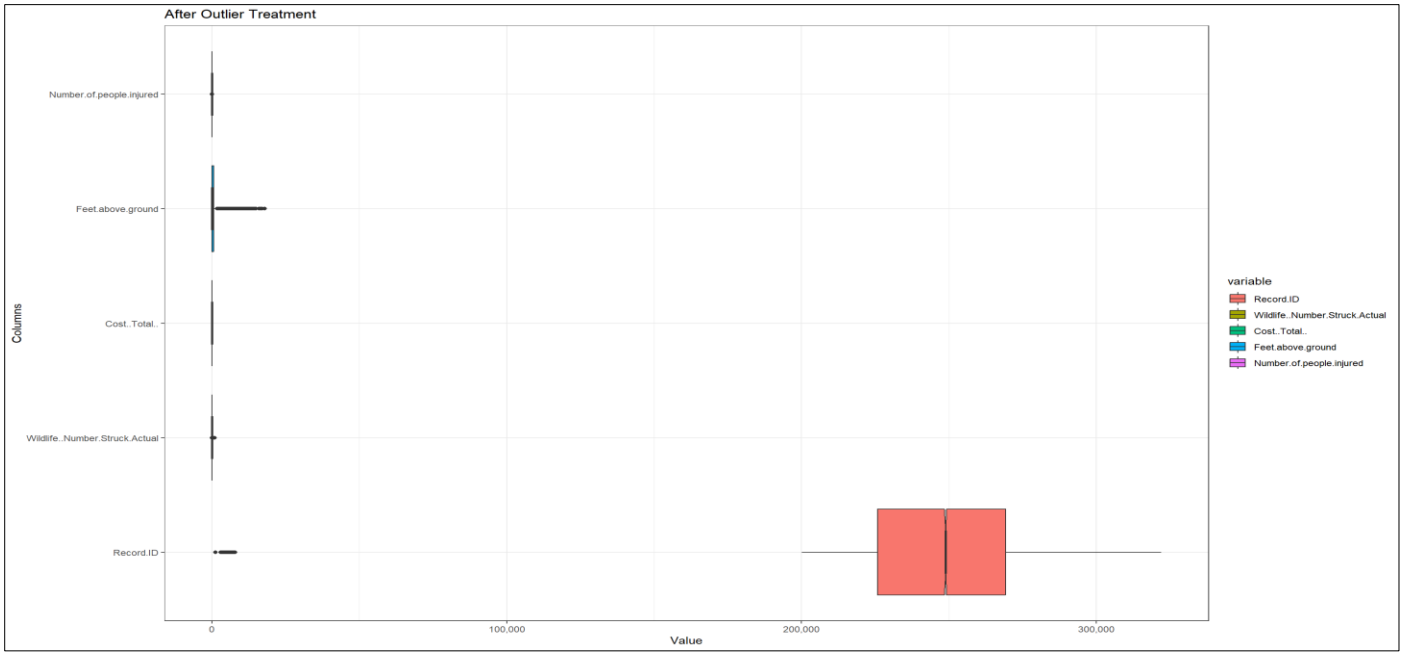




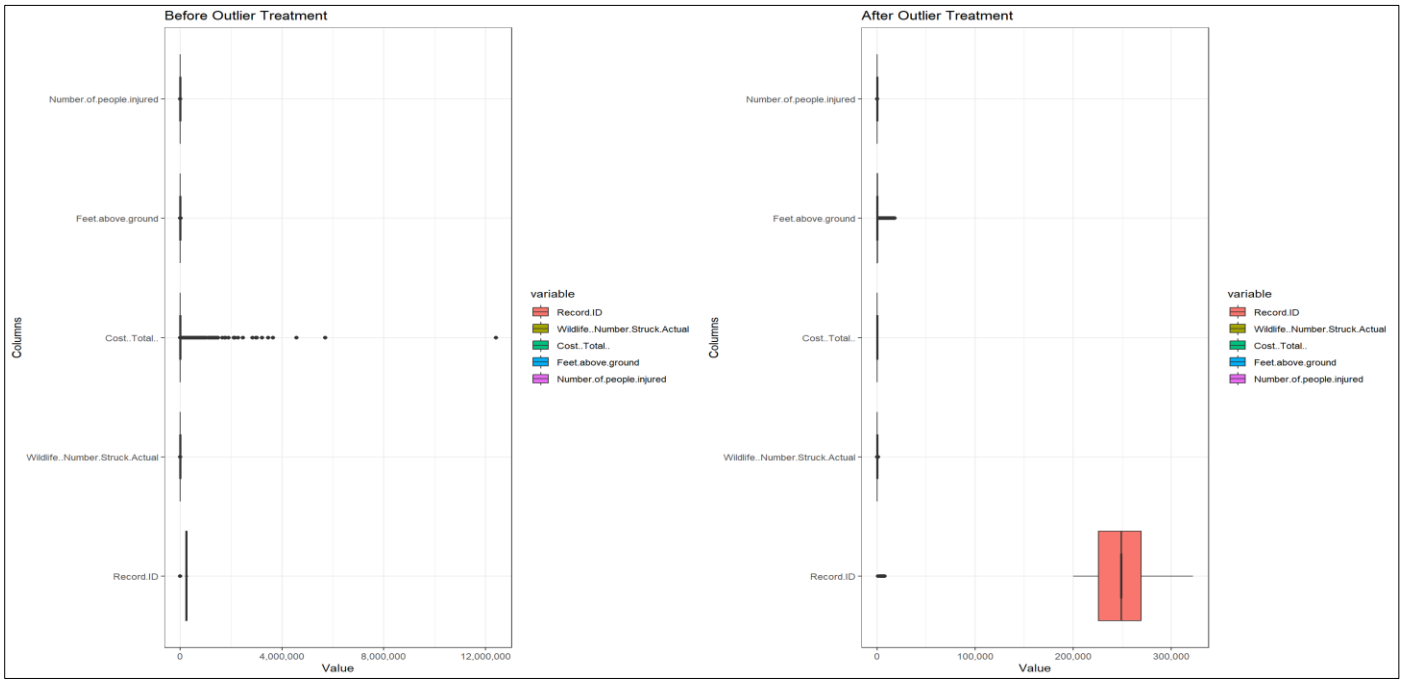
- The total cost of bird strikes appears to be influenced by multiple variables, including the indicated damage, aircraft size, and the number of wildlife struck.
- There are outliers with extremely high costs, suggesting some strikes have had catastrophic financial impacts that may warrant further investigation.

Treating the outliers

After treatment boxplot

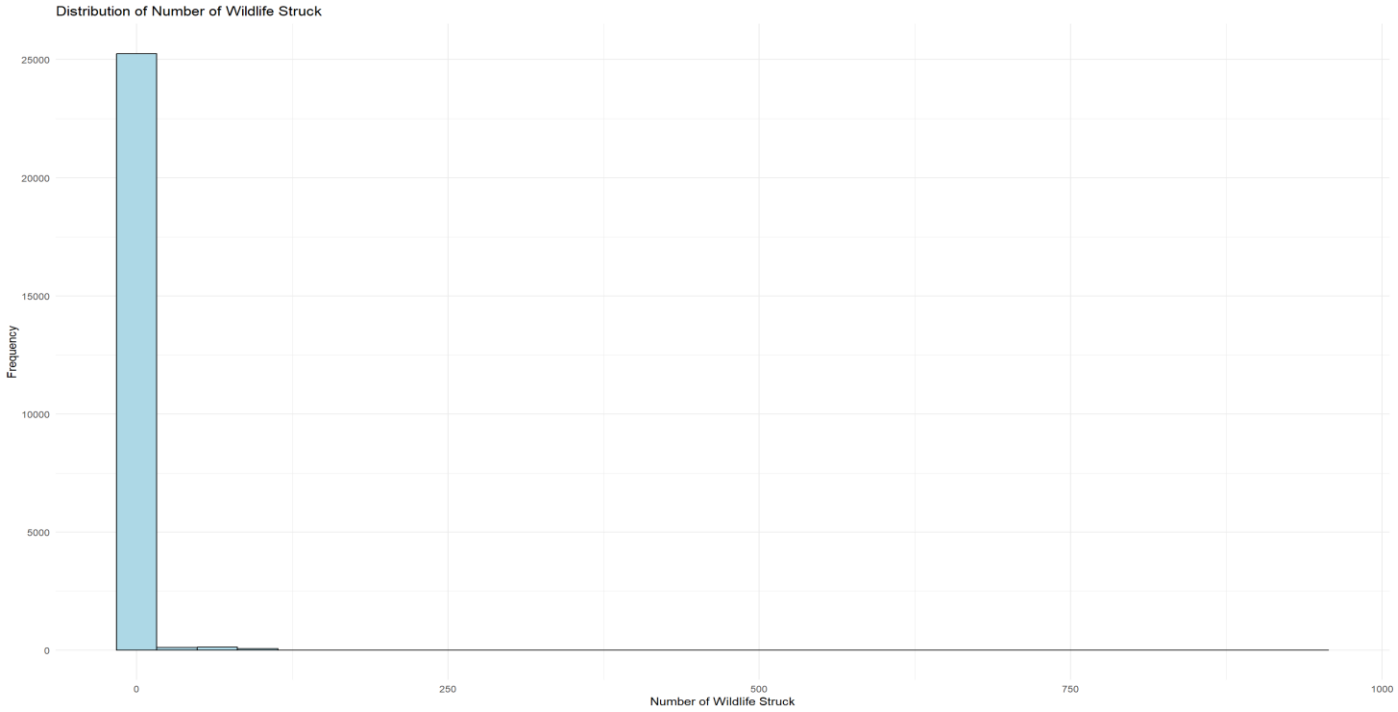


Combine plots side by side (Treatment Comparison)

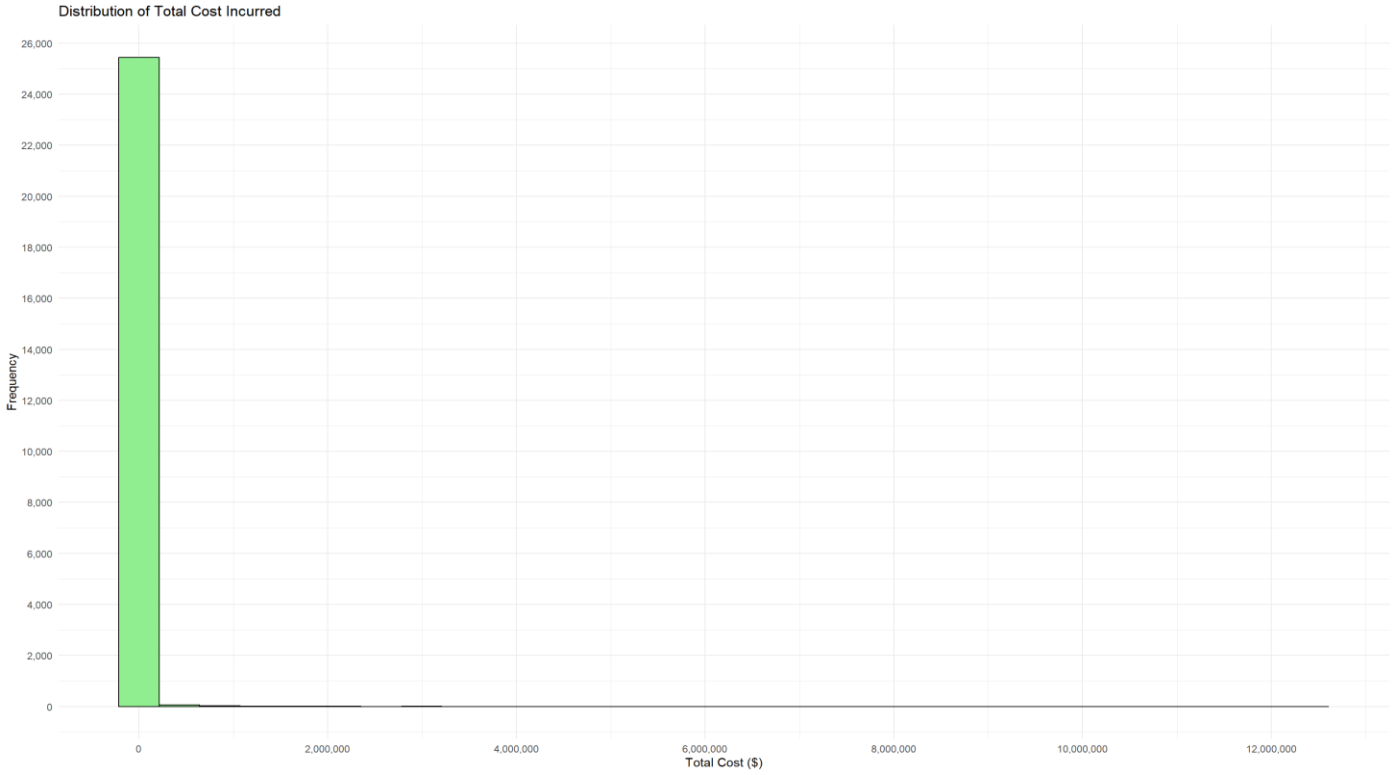


- Before outlier treatment, the data had extreme values for variables like cost and the number of wildlife struck, which could skew analyses.
- After outlier treatment, the data appears more condensed and manageable, allowing for more accurate analyses and insights.

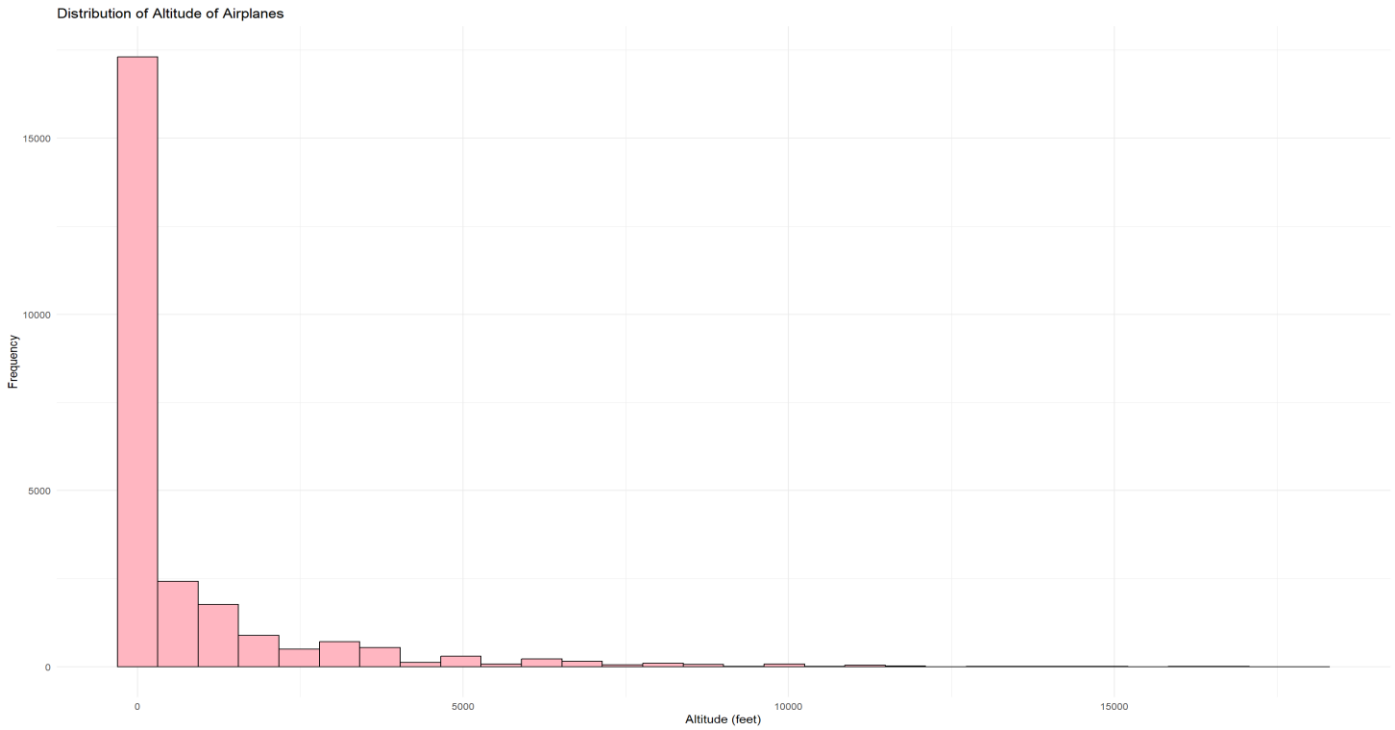
Univariate Analysis



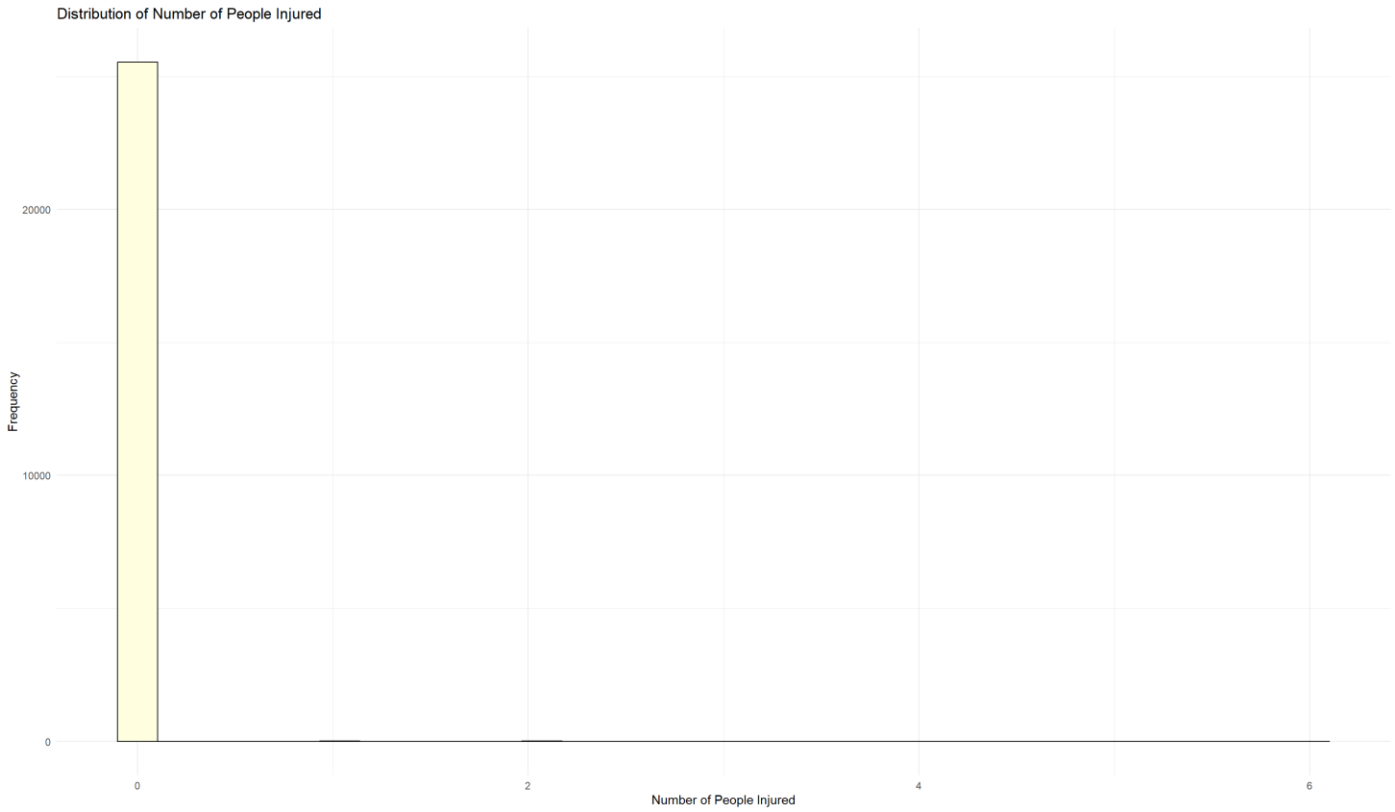
- The chart shows a very high frequency of incidents where a single wildlife (likely a bird) was struck by an aircraft.
- This highlights the prevalence of the bird strike issue and the need to address it effectively.



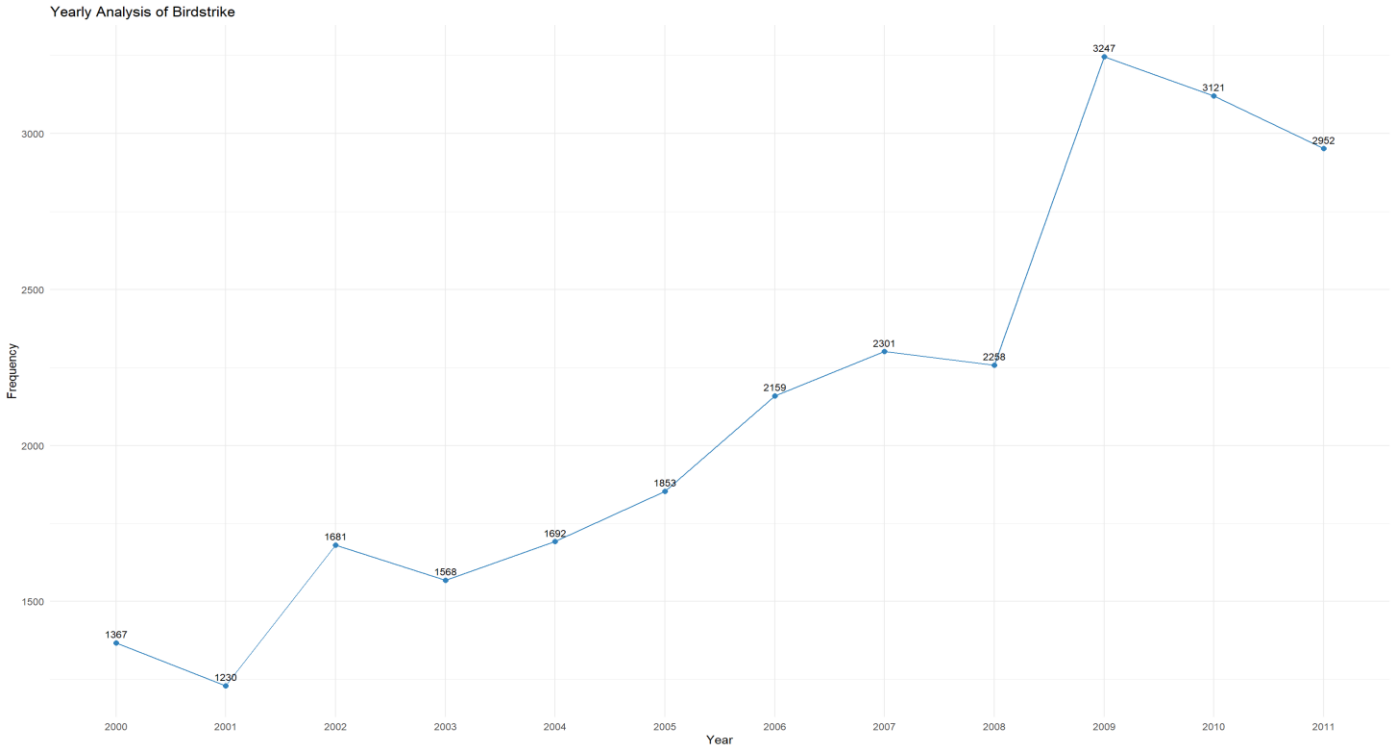
- There is a significant peak in the distribution, indicating that a large number of bird strike incidents resulted in a total cost between $8,000,000 and $10,000,000.
- This demonstrates the substantial economic impact of bird strikes on the aviation industry.



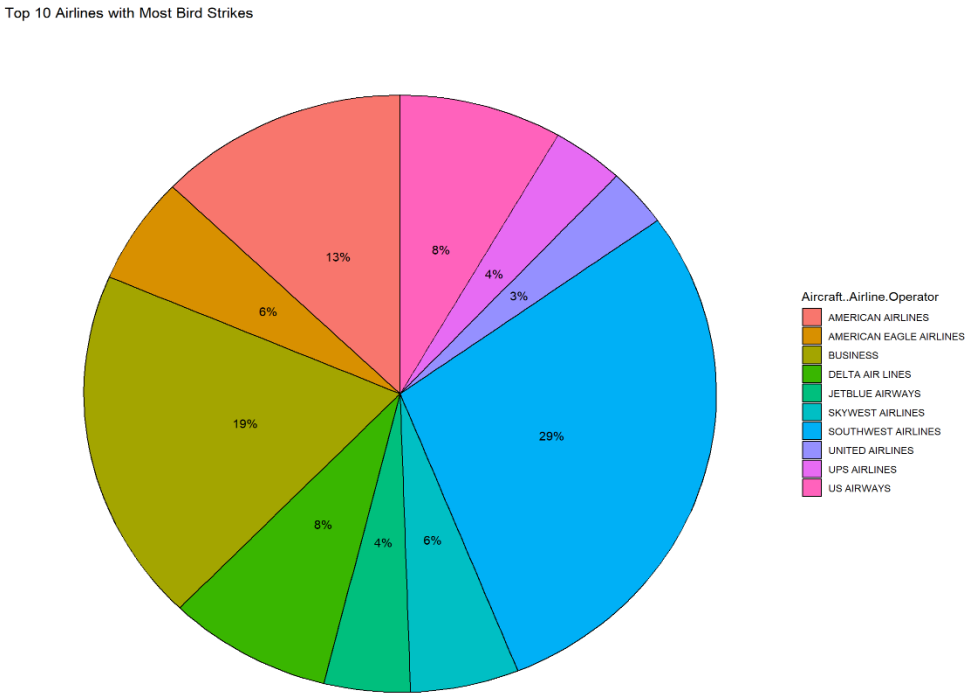
- The distribution shows a very high frequency of bird strikes occurring at lower altitudes, with a sharp drop-off as the altitude increases.
- This aligns with the problem statement mentioning that bird strikes are most likely during take-off, initial climb, approach, and landing phases when aircraft operate at lower heights.



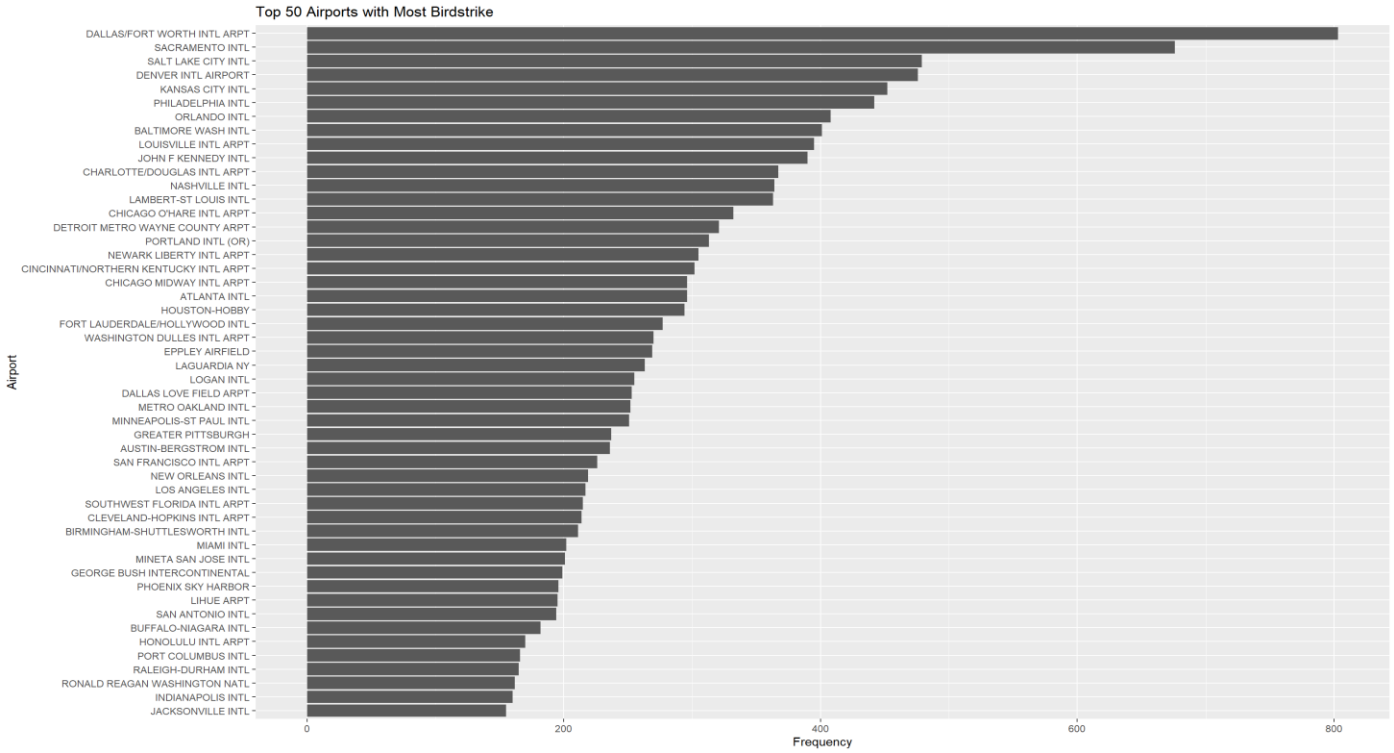
- The chart reveals a single peak, indicating that in most bird strike incidents, two people were injured.
- While not as severe as potential fatalities, injuries resulting from bird strikes highlight the safety concerns for passengers and crew.



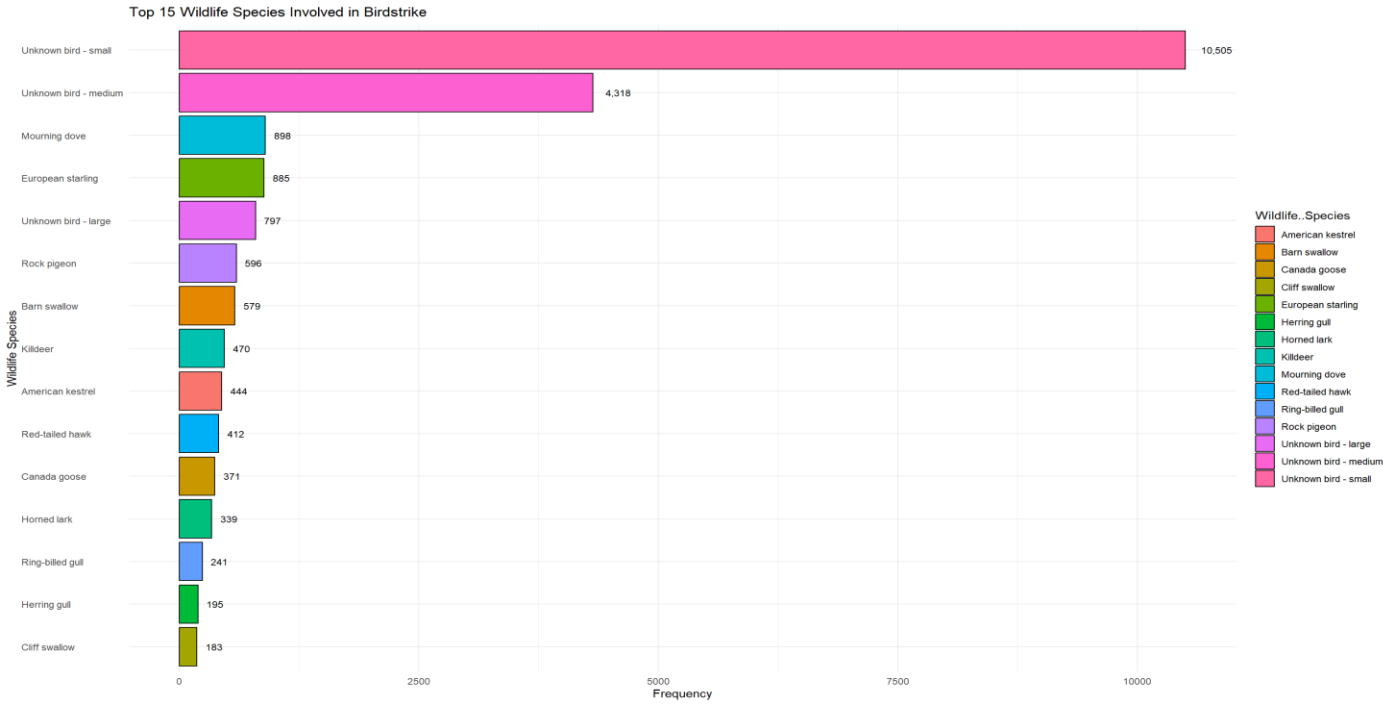
- The line graph shows a steady increase in the number of bird strikes over the years from 2000 to 2011.
- This trend suggests that the problem is becoming more prevalent and underscores the urgency of finding effective solutions to mitigate bird strike risks.



- Southwest Airlines experienced the highest number of bird strikes (29%) among the top 10 airlines shown.
- Other major US airlines like American Airlines, Delta Air Lines, and United Airlines also had significant shares of reported bird strikes.
- This data highlights the importance of addressing bird strike risks for major commercial airlines operating in the US.

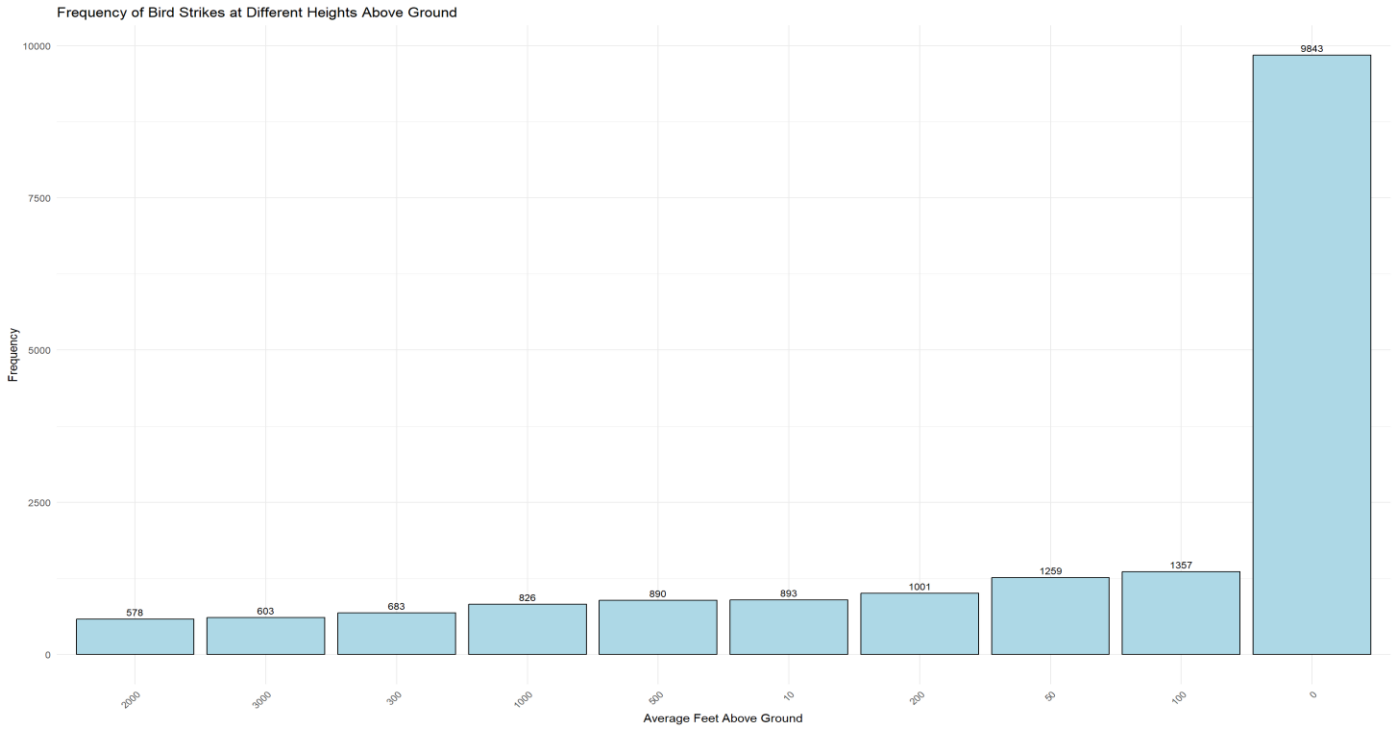


- Dallas/Fort Worth International Airport had the highest frequency of bird strikes among the top 50 airports shown.
- Several other major US hub airports like Denver, Chicago O'Hare, and Atlanta also ranked high in terms of bird strike incidents.
- This suggests that bird strike mitigation efforts should be focused on these busy airports with heavy air traffic.

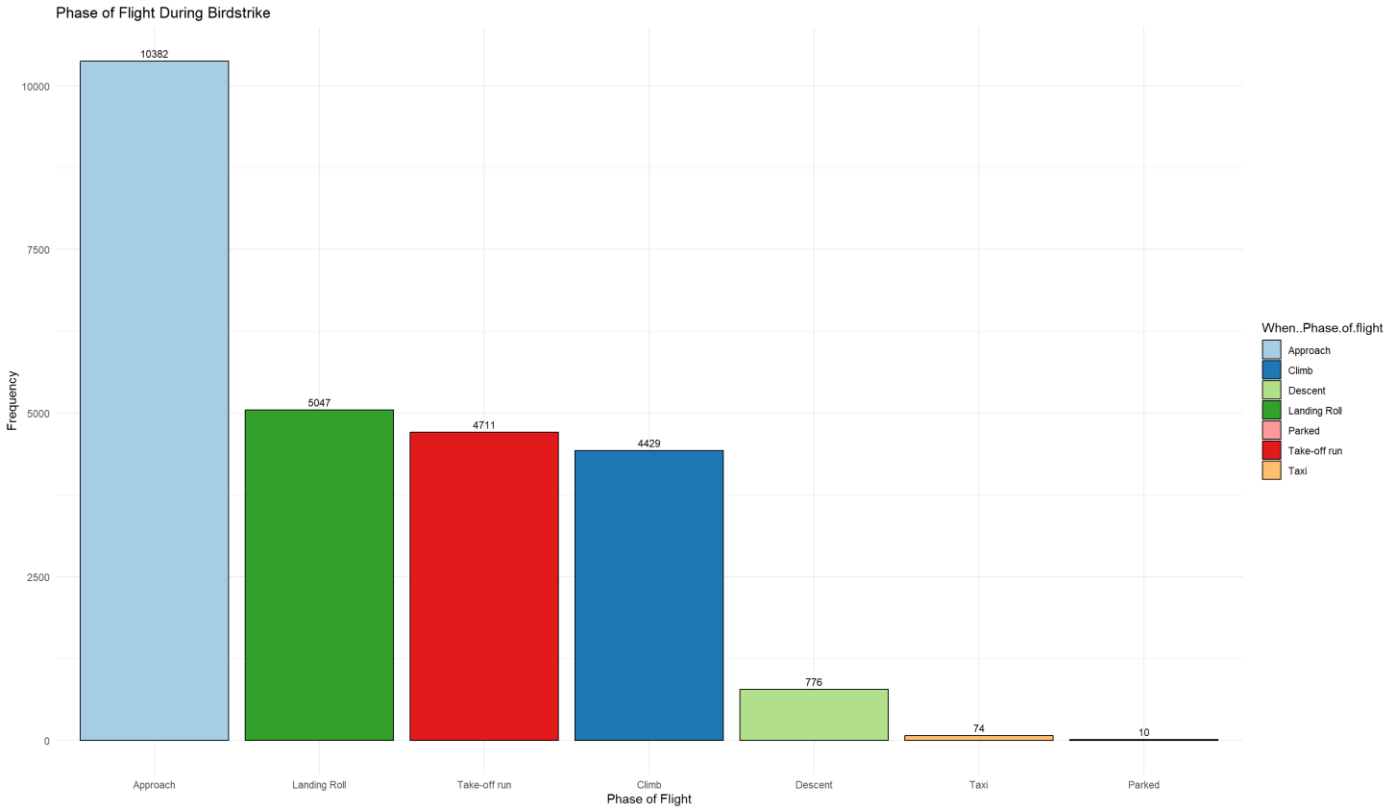


- Unknown bird species (small, medium, and large) collectively accounted for the highest number of bird strikes, indicating the challenge in identifying the specific species involved.
- Among identified species, mourning doves, European starlings, and rock pigeons were among the most frequent birds involved in strikes.
- This information can help target specific bird species and their habitats around airports for mitigation measures.

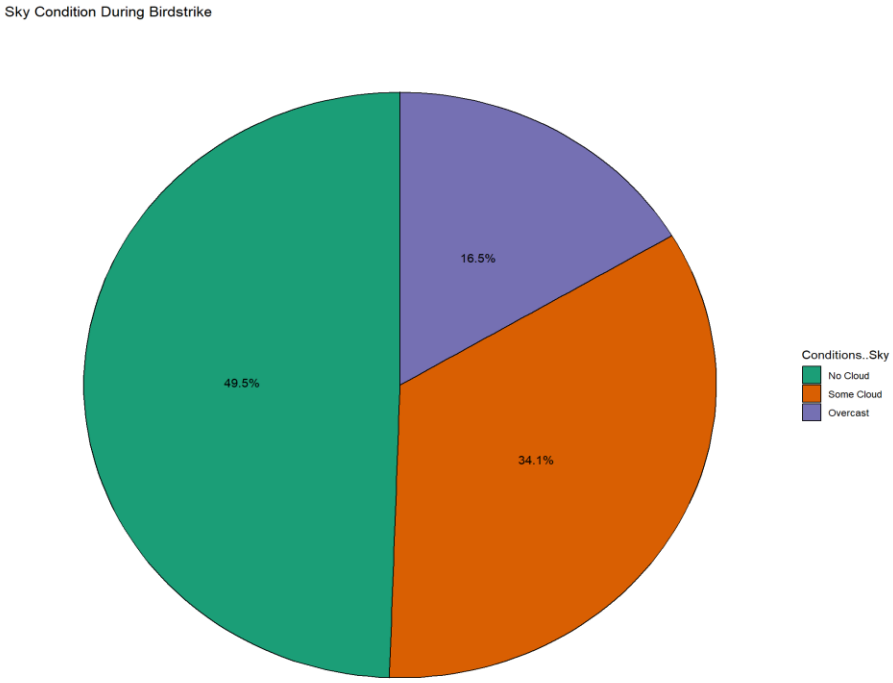
- The vast majority of bird strikes had no significant impact on flights, as indicated by the large green "None" bar.
- However, a small number of incidents did result in aborted take-offs, precautionary landings, or engine shutdowns, highlighting the potential safety risks.



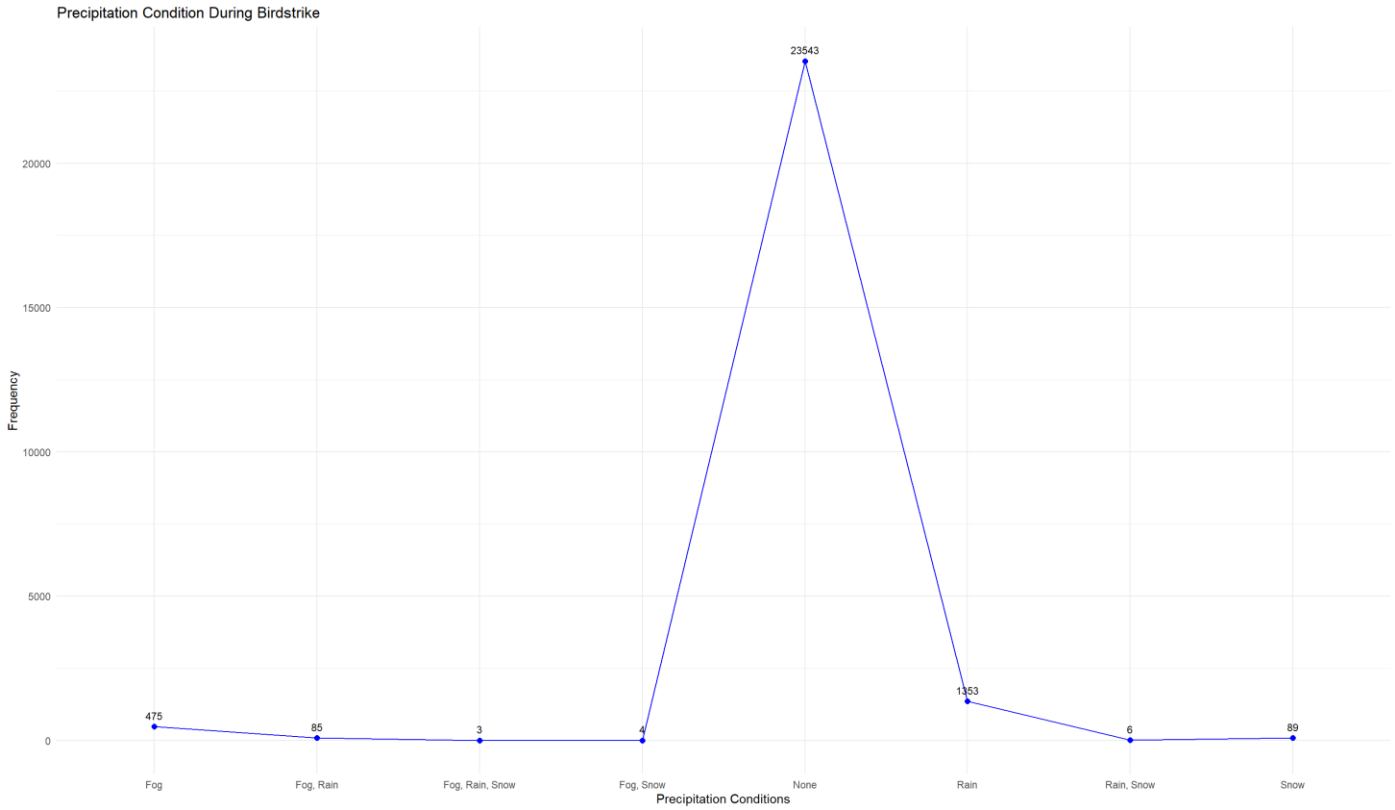
- The highest frequency of bird strikes occurred at heights below 500 feet above ground level, with a peak around take-off and landing phases.
- This aligns with the problem statement's mention of bird strikes being most likely during take-off, initial climb, approach, and landing phases.
- Implementing mitigation measures at lower altitudes around airports could help reduce the risk of bird strikes during these critical flight phases.



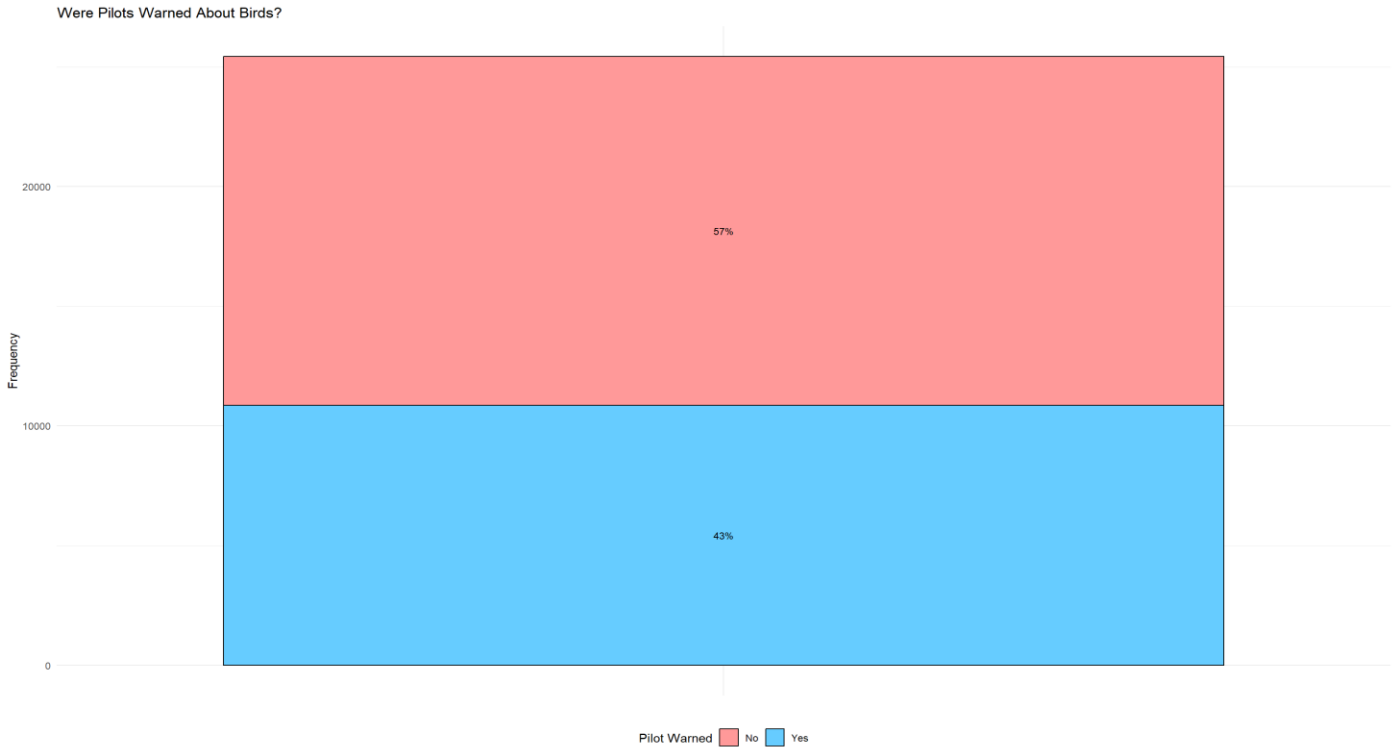
- The approach phase of flight has the highest frequency of bird strikes, with over 10,000 incidents recorded.
- Landing roll and take-off run phases also have relatively high occurrences of bird strikes, indicating the vulnerability during ground operations.



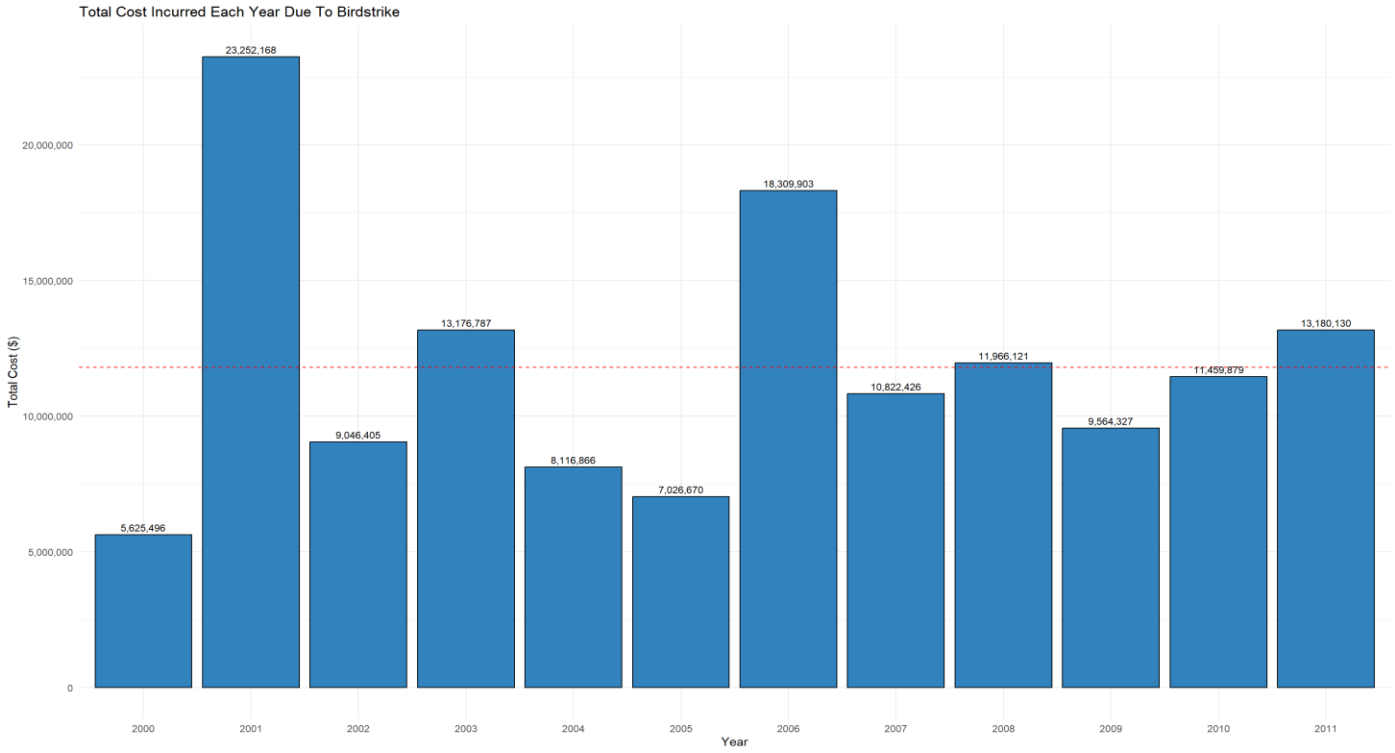
- Nearly half (49.5%) of the bird strikes occurred when the sky was clear with no cloud cover.
- However, a significant portion (34.1%) also happened under some cloud conditions, suggesting birds can still pose a threat in non-ideal weather.



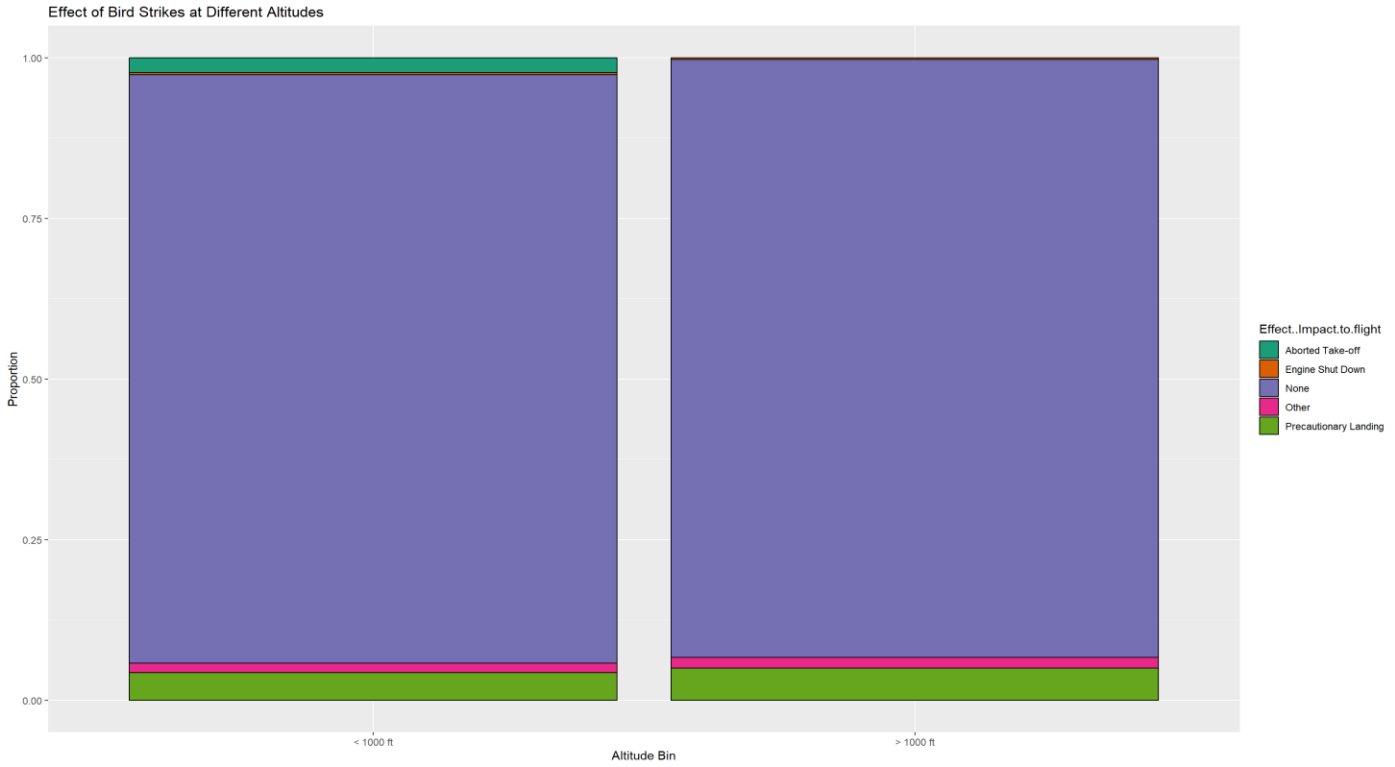
- The vast majority of bird strikes (over 23,000 incidents) occurred in the absence of precipitation like fog, rain, or snow.
- However, a considerable number (around 1,100) did occur during foggy, snowy conditions, highlighting the need for vigilance even in adverse weather.



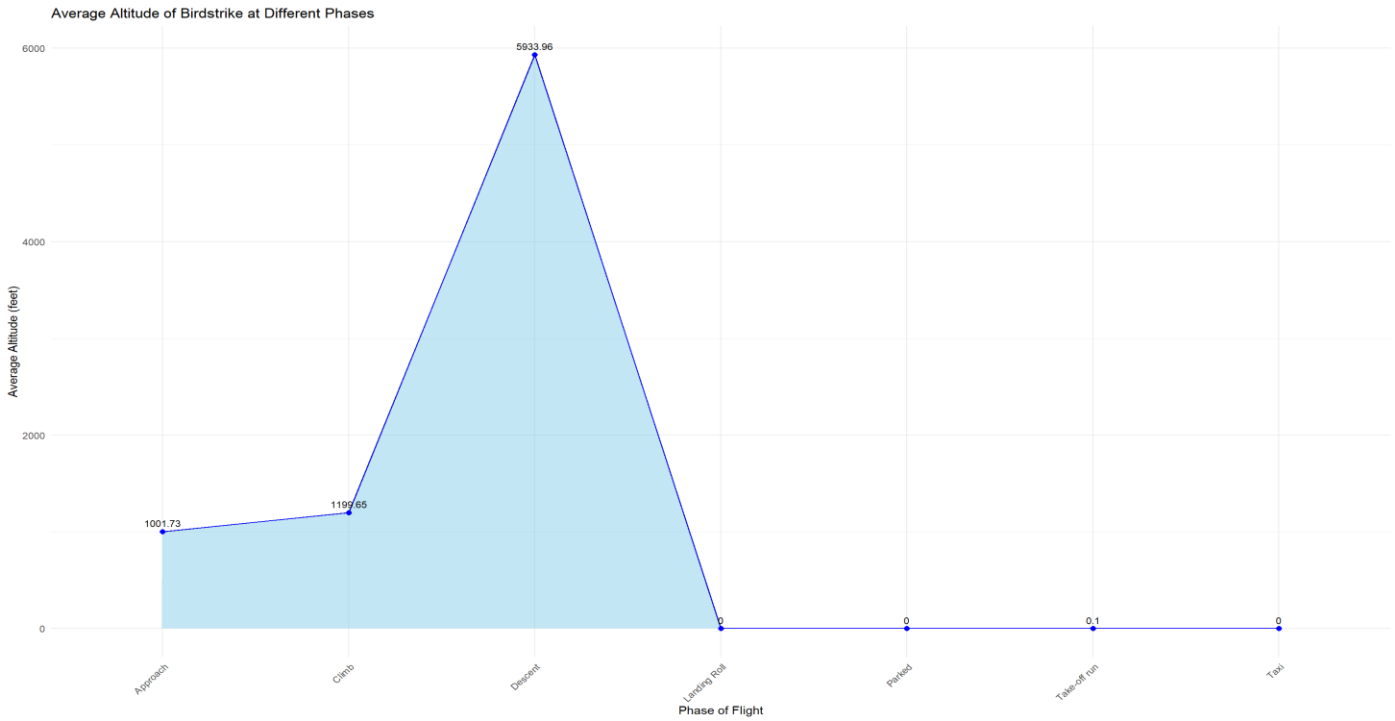
- In 57% of the cases, pilots were not warned about the presence of birds in the area, indicating a potential gap in bird hazard communication and awareness.
- Only 43% of the pilots received warnings, suggesting room for improvement in bird strike prevention measures.



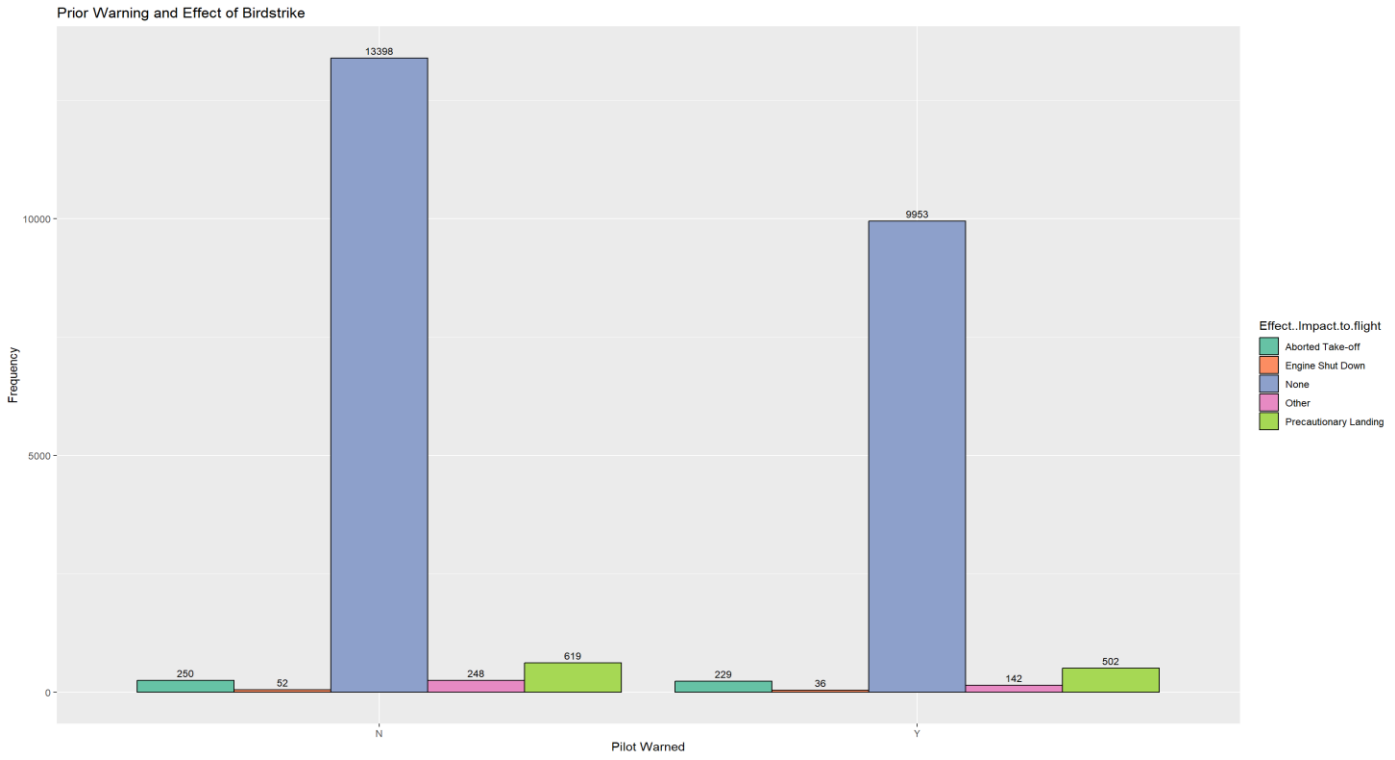
- The cost of bird strikes has been substantial, with several years exceeding $10 million in damages.
- The highest cost was incurred in 2000, at over $23 million, underlining the significant economic impact of bird strikes on the aviation industry.
- While costs fluctuate yearly, the overall trend suggests that bird strikes remain an ongoing concern with substantial financial implications.



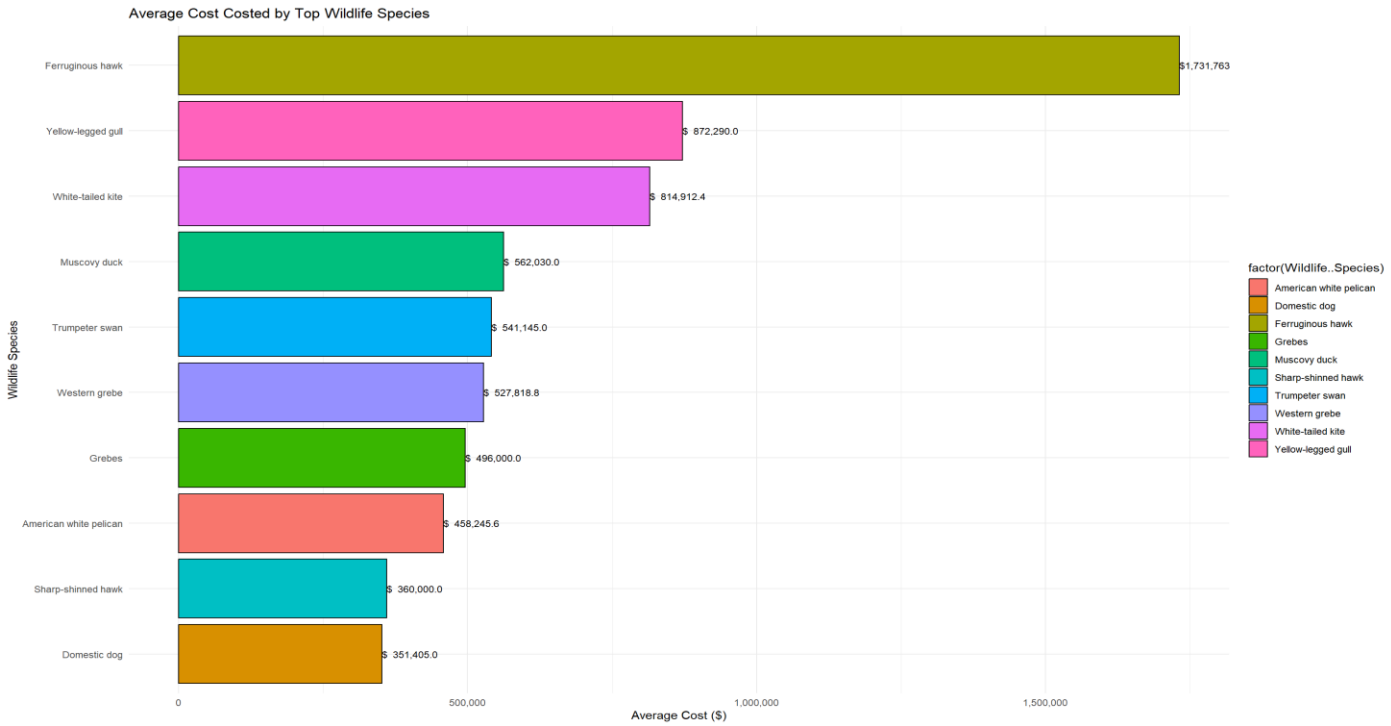
- For altitudes below 1000 feet, a significant proportion of bird strikes result in aborted take-offs, suggesting that low altitude is a critical phase where pilots need to be extra vigilant.
- Above 1000 feet, while aborted take-offs are less common, there are still instances of engine shutdowns and other impact to flights due to bird strikes.



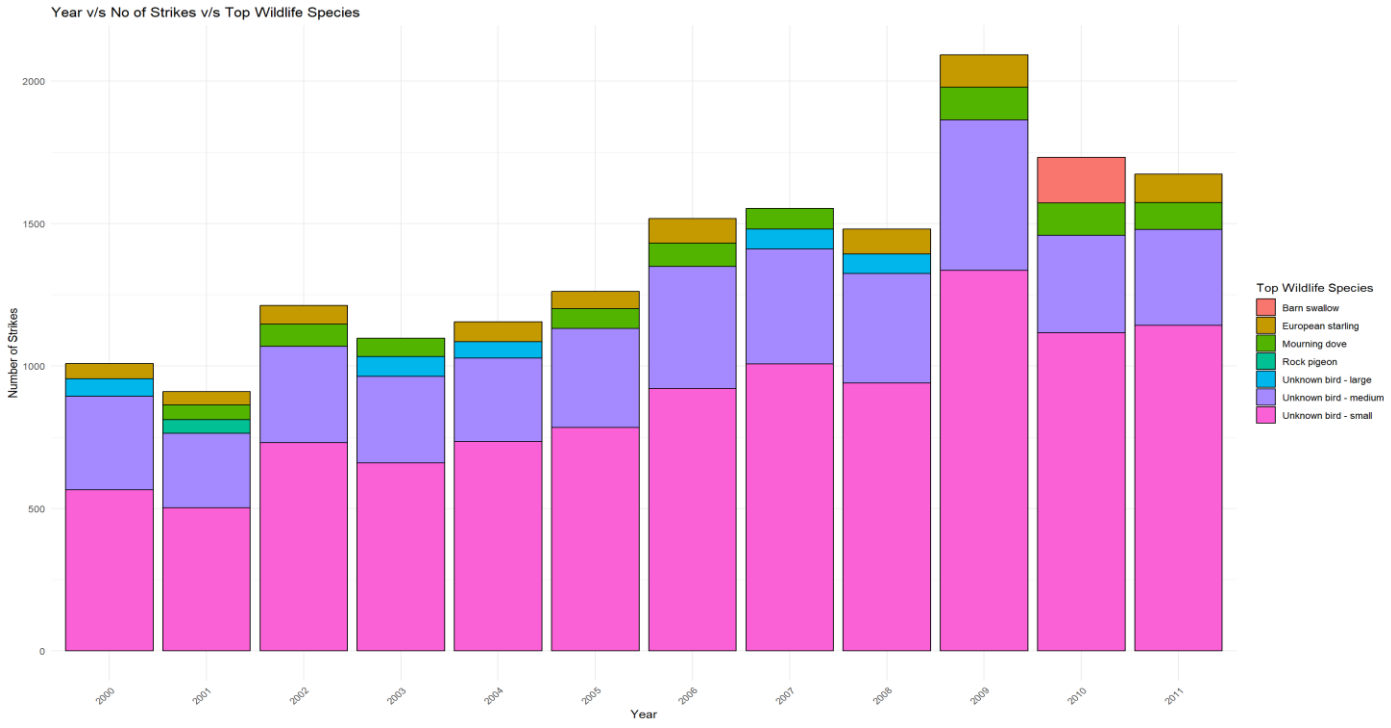
- The highest average altitude for bird strikes is during the climb phase, reaching up to around 5,933 feet, indicating that this is a vulnerable phase where collisions with birds are more likely.
- The approach and landing phases also have relatively high average altitudes for bird strikes, around 1,000-1,150 feet.



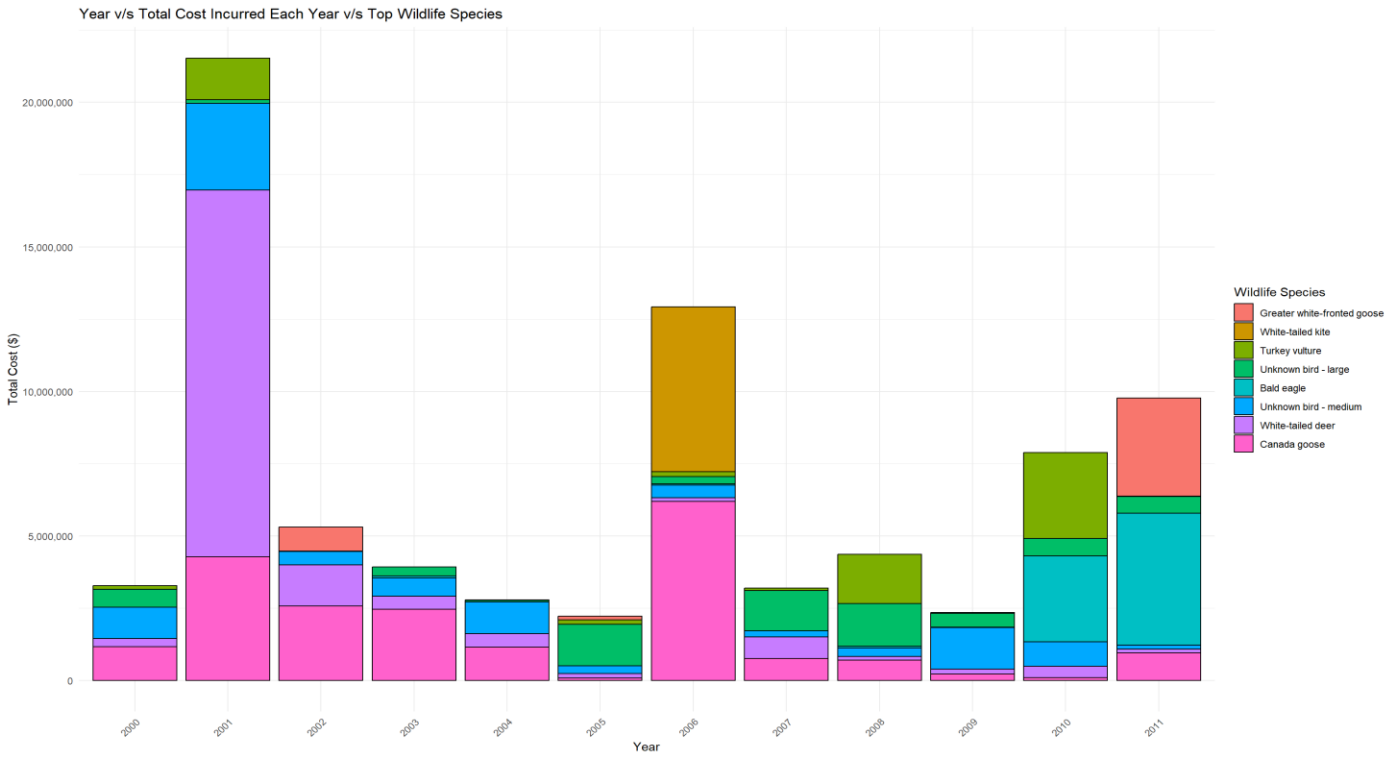
- When pilots are not warned about potential bird activity, there is a significantly higher likelihood of aborted take-offs or engine shutdowns resulting from bird strikes.
- Providing prior warning to pilots seems to mitigate the impact of bird strikes, with fewer instances of aborted take-offs or engine shutdowns.



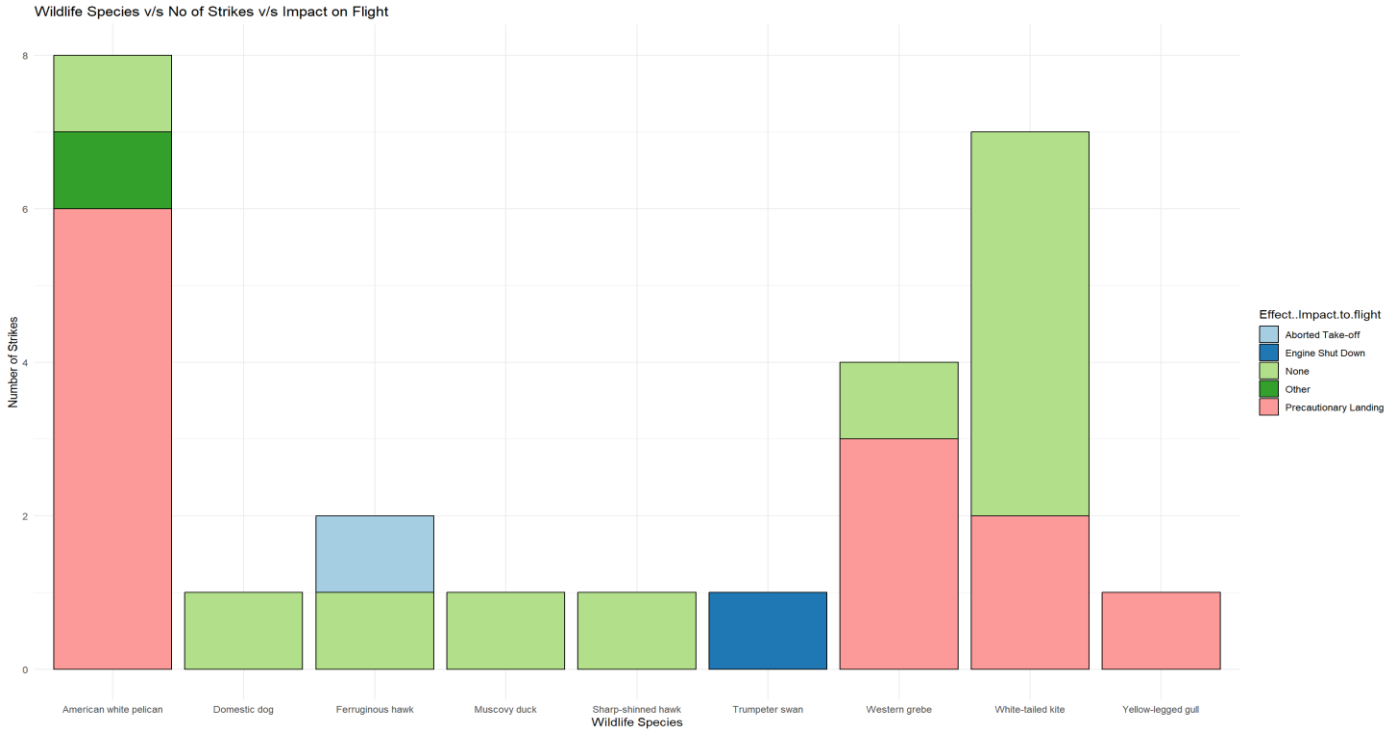
- Certain bird species, such as the ferruginous hawk, yellow-legged gull, and white-tailed kite, tend to cause higher average costs when involved in bird strikes, likely due to their size and potential for more significant damage.
- Other species like the muscovy duck, trumpeter swan, and western grebe also contribute to substantial average costs resulting from bird strikes.



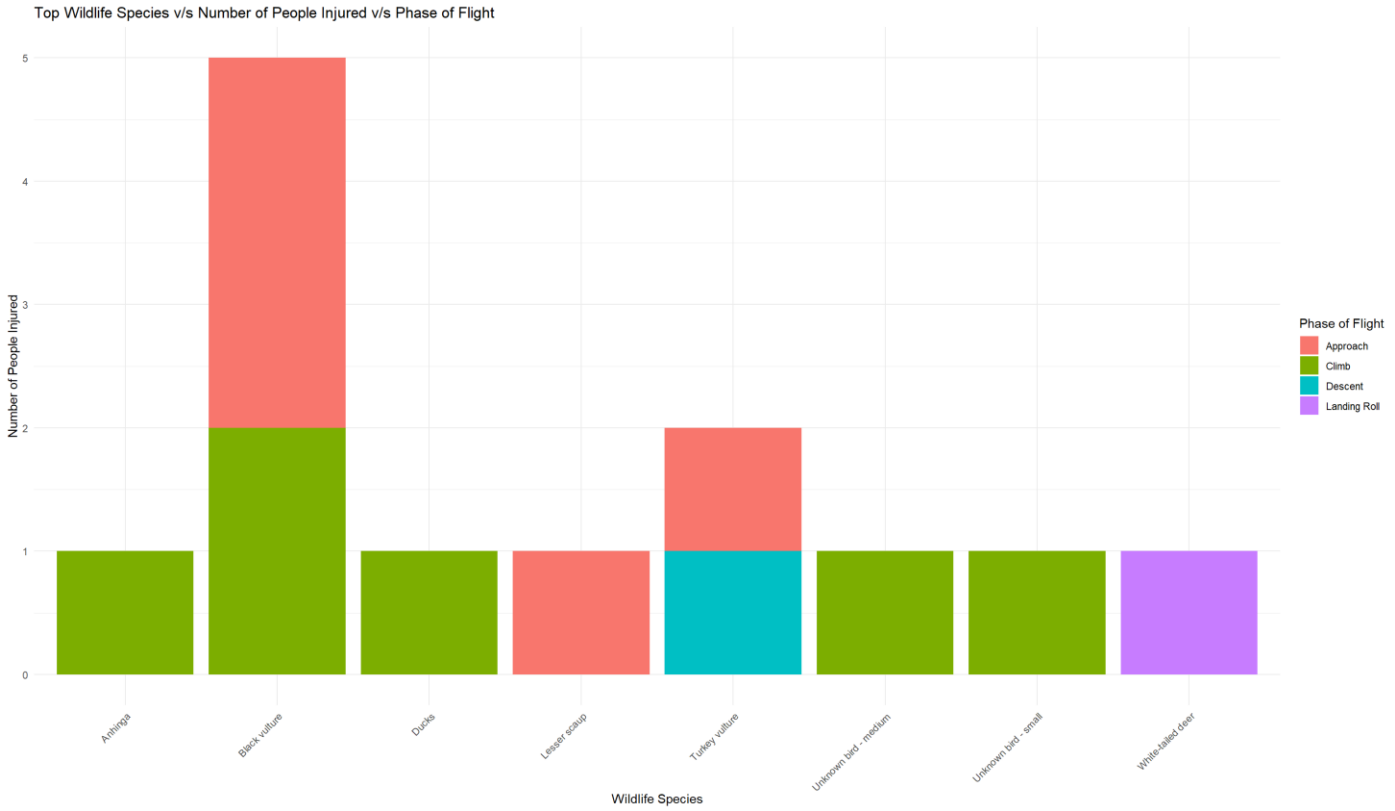
- The number of bird strikes involving various bird species fluctuates over the years, with certain species like barn swallows, European starlings, and mourning doves being more consistently involved.
- The data suggests that monitoring and mitigating the presence of these specific bird species around airports could be crucial in reducing the risk of bird strikes.



- The greater white-fronted goose has caused the highest total cost incurred due to bird strikes among the wildlife species shown, indicating it poses a significant threat to aircraft safety and financial impact.
- The total cost incurred has fluctuated over the years, with some years having higher costs than others, suggesting the need for consistent mitigation measures.



- The trumpeter swan and western grebe have resulted in the most severe impacts on flights, such as aborted take-offs and engine shutdowns, highlighting the potential safety risks posed by these species.
- Smaller birds like the American white pelican and domestic dog seem to have less severe impacts on flights, but their strikes should still be addressed.



- The majority of people injured due to bird strikes occur during the approach phase of flight, which aligns with the problem statement's mention of increased risk during landing.
- The white-tailed deer and unknown wildlife species also contribute to injuries, emphasizing the need to consider strikes from non-avian wildlife as well.

Overall Summary : -

1. Bird strikes are a significant issue for the aviation industry, with increasing frequency over time and substantial economic impact.
2. Most strikes occur at lower altitudes during takeoff, landing, and approach phases.
3. Unknown bird species are common, but certain identified species like mourning doves are frequently involved.
4. Strikes typically have no major impact, but some cause aborted takeoffs, engine shutdowns, injuries, and even fatalities.
5. Pilot warnings about bird activity can help mitigate the severity of strikes.
6. Certain larger bird species cause more damage and higher costs than smaller birds.

Important Insights:

1. **Trends Over Time:**
   A steady increase in bird strikes from 2000 to 2011, indicating a growing problem.
   Most incidents occur during take-off, initial climb, approach, and landing phases.
2. **Cost Implications:**
   Significant financial costs incurred due to bird strikes, with some incidents exceeding $10 million.
   Outliers with extremely high costs suggest catastrophic impacts on the aviation industry.
3. **Impact on Flights:**
   Majority of bird strikes have no significant impact on flights, but some result in aborted take-offs, precautionary landings, or engine shutdowns.
   Approach phase has the highest frequency of bird strikes, with landing roll and take-off run phases also affected.
4. **Factors Affecting Strikes:**
   Altitude, time of day, and weather conditions influence the likelihood of bird strikes.
   Pilots warned of bird activity have fewer instances of aborted take-offs or engine shutdowns.
5. **Strategies Based on Goals:**
   Implement wildlife management programs and enhance pilot warnings to mitigate bird strike risks.
   Focus on identifying and monitoring specific bird species known to cause high costs and severe flight impacts.

Insight Summary:

1. **Transport Safety Enhancement:**
   The analysis identified trends and factors influencing bird strikes, enabling the development of targeted strategies to enhance aviation safety.
2. **Financial Impact Mitigation:**
   Understanding the financial costs associated with bird strikes helped in proposing measures to mitigate economic losses and operational disruptions.
3. **Operational Efficiency Improvement:**
   Insights into the impact of bird strikes on flights informed strategies to improve operational efficiency by reducing delays, diversions, and cancellations.
4. **Risk Reduction Strategies:**
   Factors affecting bird strikes, such as altitude, weather conditions, and geographic location, guided the development of risk reduction strategies tailored to specific environments.

5. **Collaborative Approach:**
   Collaboration with industry stakeholders, wildlife experts, and government agencies was emphasized to develop comprehensive and effective strategies for bird strike prevention and mitigation.

Additional Insights : -

1. Bird strikes are becoming more prevalent and pose a safety and financial threat to the aviation industry.
2. Lower flight phases are most vulnerable to bird strikes.
3. Specific bird species can be targeted for mitigation efforts around airports.
4. Improved communication and warnings about bird hazards can benefit pilots.
5. The impact of bird strikes varies depending on the species and altitude.

Key Metrics:

1. Total number of bird strikes over the years.
2. Financial costs incurred due to bird strikes.
3. Severity of incidents, including injuries and damage to aircraft.
4. Distribution of bird strikes by flight phase and altitude.
5. Effectiveness of prior warnings in mitigating bird strike impacts.

Factors Affecting:

1. **Altitude:** Lower altitudes during take-off and landing phases are more prone to bird strikes.
2. **Time of Day:** Bird strikes are influenced by visibility and lighting conditions.
3. **Weather Conditions:** Clear skies or cloudy conditions can affect bird activity and strike frequency.
4. **Pilot Warnings:** Providing prior warnings can mitigate the impact of bird strikes on flights.
5. **Geographic Location:** Certain airports may experience higher bird strike frequencies due to habitat proximity.

Strategies To Implement:

1. **Wildlife Management:**
   Implement measures to control bird populations near airports and reduce attractiveness of airport surroundings to birds.
2. **Pilot Awareness:**
   Enhance pilot training and communication to increase awareness of bird strike risks and proper procedures for mitigation.
3. **Technological Solutions:**
   Explore advanced radar systems and bird detection technologies to improve early warning and avoidance strategies.
4. **Data Sharing:**
   Establish a centralized database for sharing bird strike incident data to facilitate research and collaborative efforts in bird strike prevention.
5. **Collaborative Efforts:**
   Collaborate with aviation industry stakeholders, wildlife experts, and government agencies to develop comprehensive strategies for bird strike prevention and mitigation.