Title: Bird Strikes and Damage to Aircrafts
Name: Ryan Bali

There are millions of domestic flights annually in the US alone, requiring the use of thousands of different planes. These planes need daily refueling between trips and routine maintenance, but maintenance that results from a bird hitting the plane varies individually between aircrafts. My goal in this project is to analyze US airports where planes are hit most frequently by birds, and to forecast incidents of damage associated with these bird strikes. These predictions have implications for dollars spent on preventative measures (bird deterrence systems, change in flight timings, cancellation of flights, etc.) at high profile airports which will reduce the amount of damage sustained by an aircraft due to a bird strike.

The Federal Aviation Administration (FAA) has published a 30-year history of bird strike occurrences on commercial planes in the US online at https://wildlife.faa.gov. This dataset has 68 columns describing each occurrence of a bird strike, including date, time, airport, runway, state, operator, aircraft details, phase of flight, weather, bird species, etc. As with most public datasets, there are missing fields for many records, so non-essential ones were removed from my analysis.

I selected 6 features for my analysis: 1) state, 2) airports, 3) count of incidents, 4) hour of day, 5) year, and 6) damage caused. Since the FAA started detailed data collection on bird strikes starting in 2010, many years prior to 2010 have a lot of missing fields that may misconstrue conclusions from the data. Thus, I filtered the dataset to only include the years from 2010 to 2019. I also removed data entries that did not have the hour of the bird strike, and sorted the hours to reflect 4 times of the day (6-hour periods) that the bird strikes occurred: dawn, day, dusk, and night. Bird strikes can also occur at any point of an aircraft's flight path, but I wanted to use the data to implement preventative measures at certain airports. To achieve this, I filtered for bird strikes that occurred during 5 distinct points in the aircraft's flight path: approach, climb, descent, landing roll, and takeoff. For this analysis, the original dataset reduced from about 250,000 to about 130,000 rows.

I started by exploring the total occurrences of bird strikes on planes as they vary by state. There is a lot of variability between states and bird strikes within their airports that distort the graph, which makes sense as certain states like New York and Texas are much more frequently traveled to than states like Arkansas and North Dakota. To account for this, I analyzed the top 10 states that have airports with the most bird strikes. The red bar from *Figure 1* shows the total bird strikes from the years 2010 to 2019 for each of the top ten states. The blue bar from *Figure 1* shows the number of bird strikes at the airport with the most strikes in each respective state as a fraction of the total. For example, it can be seen that airports in Colorado experienced just under 6,000 bird strikes from 2010 to 2019, but more than 70% of these bird strikes came from Denver International Airport.

Next, I tried to identify if certain hours of the day were correlated with more bird strikes in these top 10 states. Using sklearn's StandardScaler and KMeans, I created a cluster visualization of the proportion of the bird strikes in a single state relative to the total in that state for all hours of the day. Since I separated the data into four time periods in the day mentioned previously, this created four clusters. It can be seen from the visualization in *Figure 2* that bird strikes are more likely to occur between the 8th and 10th hours of the day as well as the 20th and 22nd hours of the day, corresponding the day and night periods. This makes sense because these are the hours when birds are typically flying to find food or going back to their nest.

Finally, I used sklearn LinearRegression to determine how many bird strikes correspond to actual incidents of damage to aircrafts in an effort to see the full scope of damage birds cause to airplanes. I observed that, over the last decade, the number of bird strikes that inflict damage to planes has been trending upwards. *Figure 3* predicts 717 and 734 bird strikes for the years 2020 and 2021 respectively. These are 2.3% and 4.7% increases respectively compared to the number of bird strikes in the most recent year with full data – 2019. Even though these types of increases may seem low, the damage sustained by aircrafts over these years and the cost of maintenance to repair these damages will add up. Airports, especially the top 10 mentioned earlier, should have silent high frequency emitting deterrents when planes approach landing strips to land or depart in an effort to mitigate these repair costs and extend the lifespan of airplanes.

Figure 1:

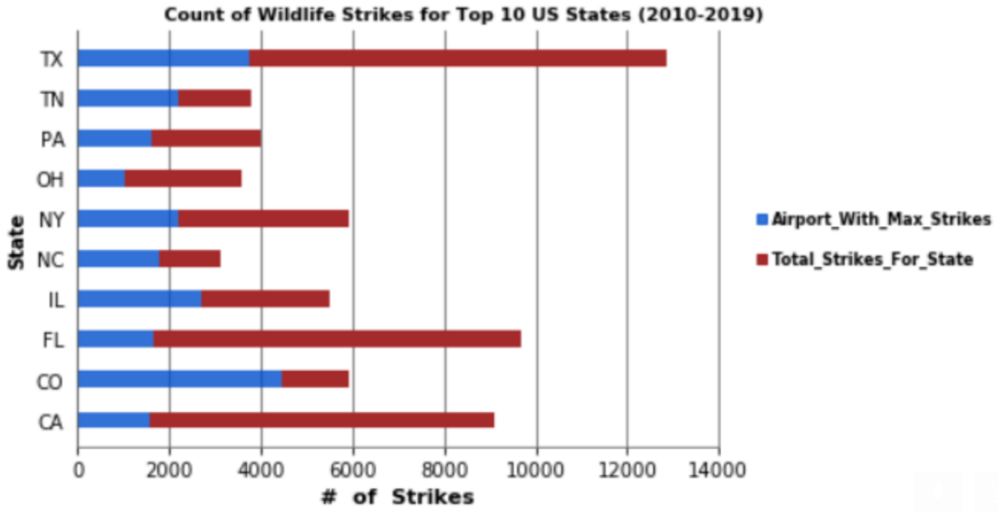


Figure 2:

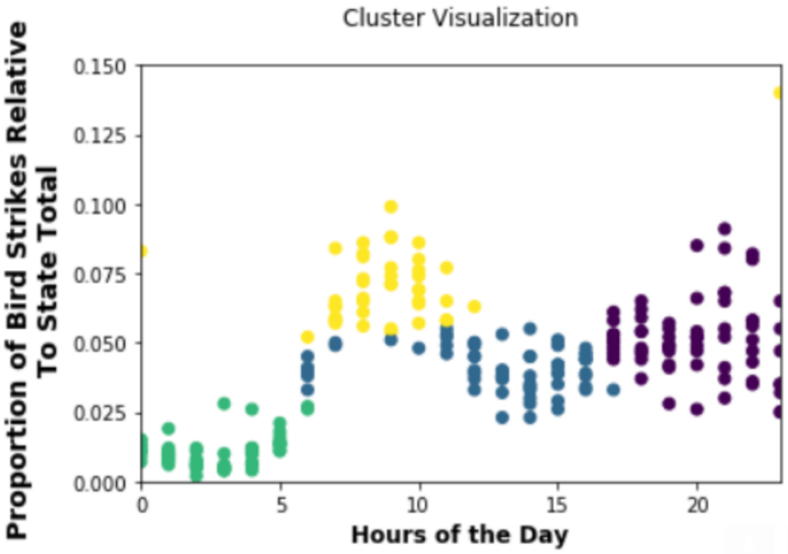
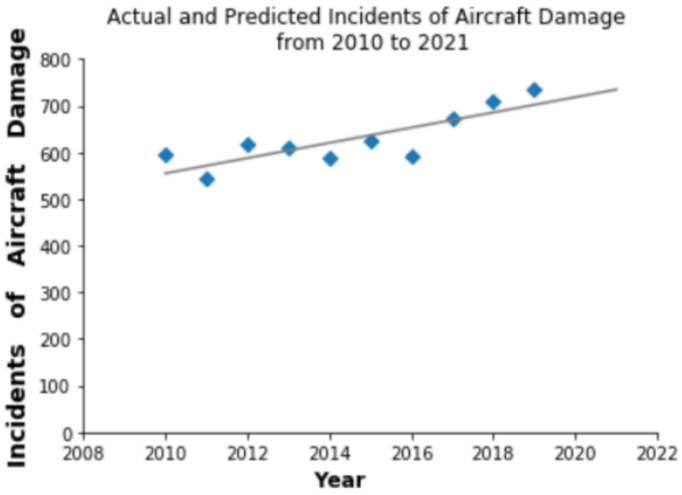


Figure 3:




```
Model Parameters :
    coefficient of determination ->  0.6817360698475199
    intercept ->  [-32250.78181818]
    slope ->  [[16.32121212]]
```