Bird Strikes Data Analysis in Python

Aadit Nandan

Introduction

- We have been given the data of bird strikes ie. The crashing of birds into airplanes. The term is often expanded to represent crashing of any wildlife into an airplane.
- We have to analyse the data to find key metrics and factors along with any meaningful relationship between attributes.
- As usual the first steps would be to read and clean the data.

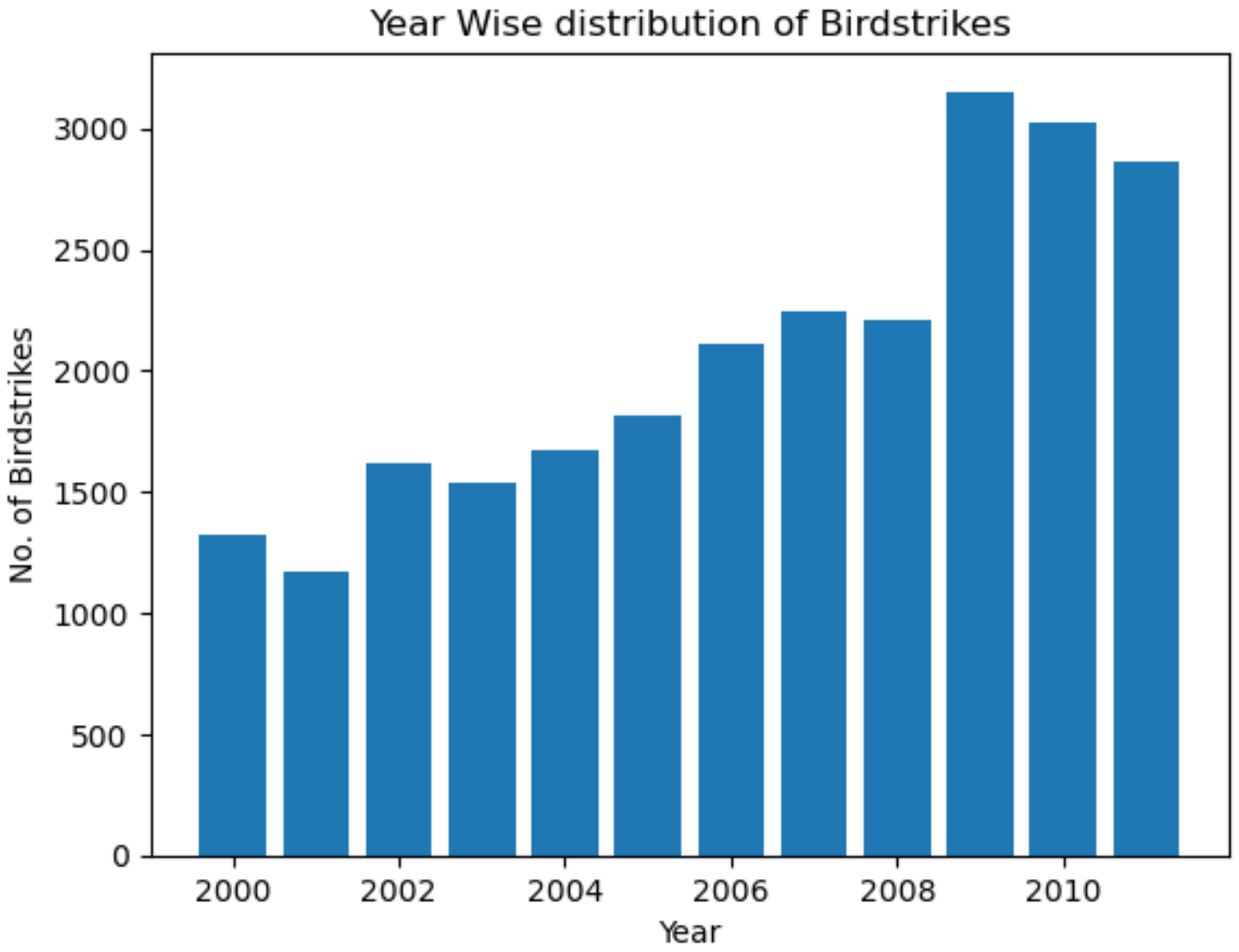
	Record ID	Wildlife: Number Struck Actual	FlightDate	Effect: Indicated Damage	Pilot warned of birds or wildlife?	Cost: Total $	Feet above ground	Number of people injured
count	24747.000000	24747.000000	24747	24747.000000	24747.000000	2.474700e+04	24747.000000	24747.000000
mean	254485.775165	2.689255	2007-01-23 03:03:49.118680832	0.096335	0.430557	5.485157e+03	801.538449	0.000849
min	200011.000000	1.000000	2000-01-02 00:00:00	0.000000	0.000000	0.000000e+00	0.000000	0.000000
25%	225827.000000	1.000000	2004-06-21 00:00:00	0.000000	0.000000	0.000000e+00	0.000000	0.000000
50%	248552.000000	1.000000	2007-07-27 00:00:00	0.000000	0.000000	0.000000e+00	50.000000	0.000000
75%	268974.500000	1.000000	2009-10-31 00:00:00	0.000000	1.000000	0.000000e+00	700.000000	0.000000
max	321909.000000	942.000000	2011-12-31 00:00:00	1.000000	1.000000	1.239775e+07	18000.000000	6.000000
std	35581.910360	12.506021	NaN	0.295056	0.495164	1.231439e+05	1736.743268	0.047986

Basic Figures

- We can get many of the important figures of each column with a single command using the "describe" method.
- This method provides us with the count of entries in each column along with the mean, standard deviation, minimum, maximum and the three major quartiles.
- These figures are extremely important as they allow us to know the range of figures we are operating in for each attribute.

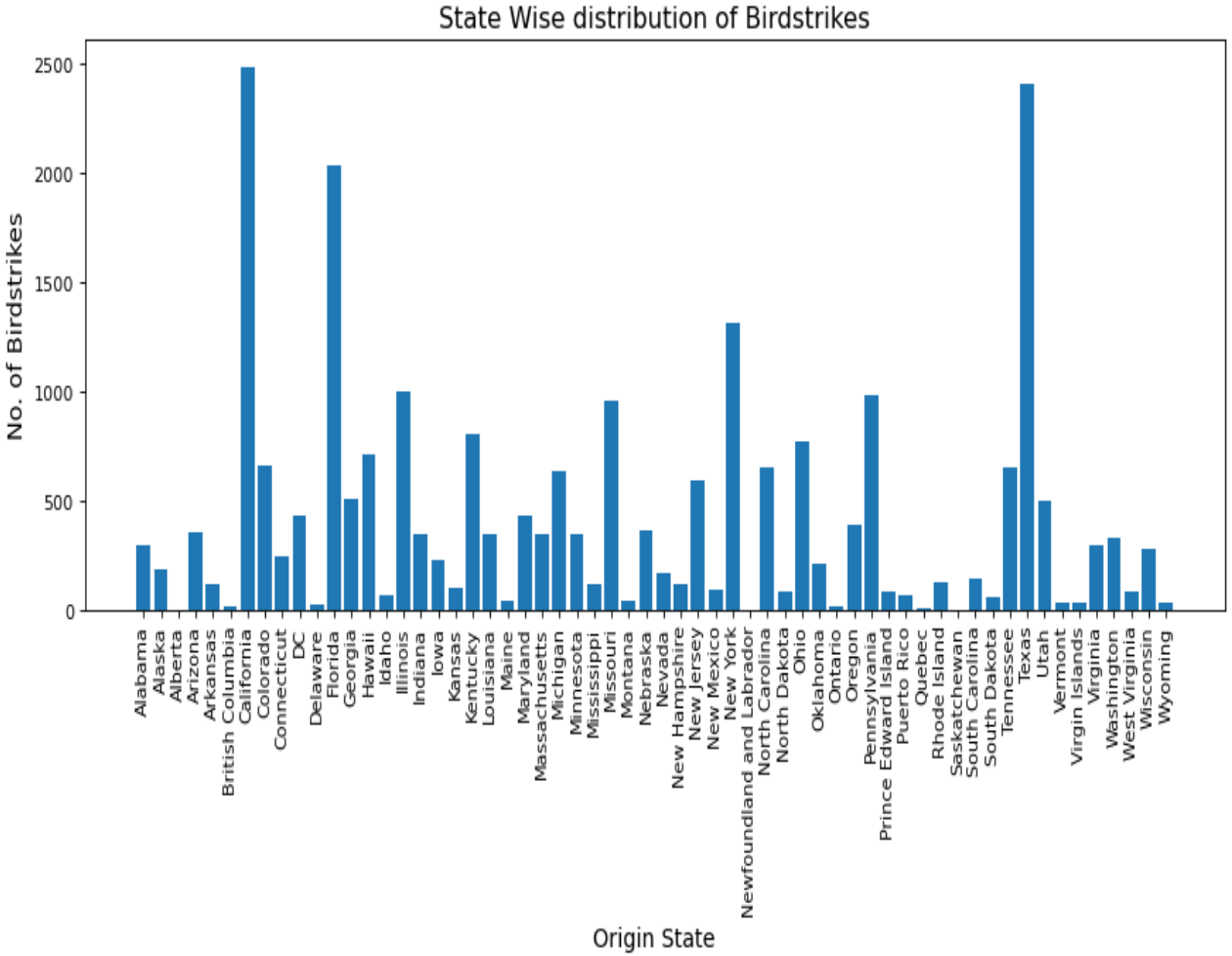
Yearly Distribution

- Grouping the data yearly, we get the Total number of Bird strikes per year.
- This data can be used to plot graphs and observe trends to make conclusions and decisions.
- This is the bar graph of the Total number of Bird strikes year wise.
- We can conclude that 2009 was the year with the most number of bird strikes.



State Wise Distribution

- Grouping the data State wise, we get the Total number of Bird Strikes per State.
- This data can be used to plot graphs and observe trends to make conclusions and decisions.
- This is the bar graph of the Total number of Bird Strikes per State.
- We can conclude that California, Florida and Texas observe the highest numbers of bird strikes.



Top 10 Airlines to encounter bird strikes

- Sorting the data in descending order using the number of bird strikes after grouping airline wise, we can find out the top 10 Airlines which encounter bird strikes.
- Following are those airlines along with the number of bird strikes they observed.

Aircraft: Airline/Operator	No. of Birdstrikes
SOUTHWEST AIRLINES	4628
BUSINESS	3035
AMERICAN AIRLINES	1852
DELTA AIR LINES	1312
AMERICAN EAGLE AIRLINES	920
SKYWEST AIRLINES	891
US AIRWAYS*	780
JETBLUE AIRWAYS	677
UPS AIRLINES	546
US AIRWAYS	515

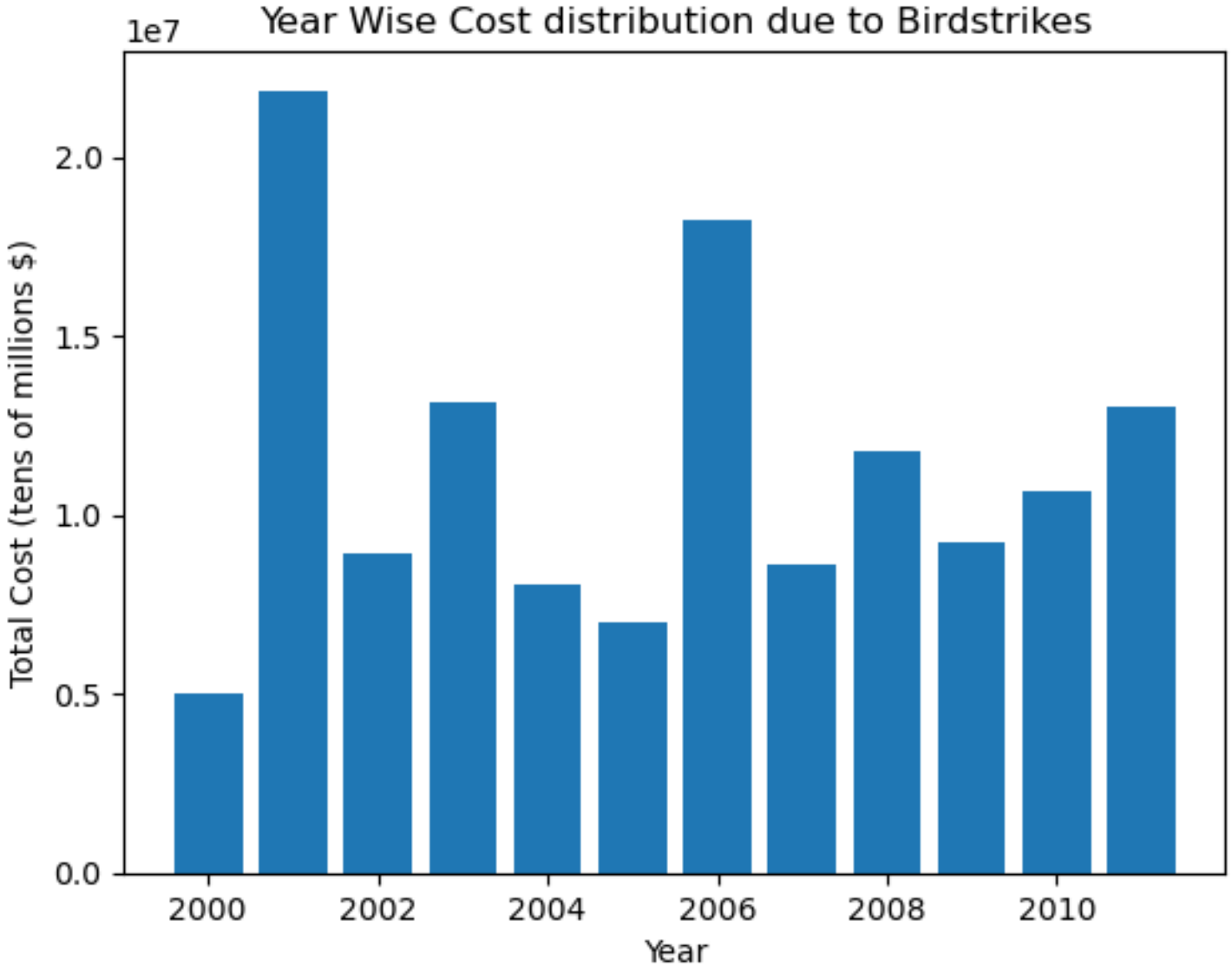
Top Airports with incidents of bird strikes

- Grouping the data airport wise, we can count the number of incidents per airport.
- Then we can sort the data in descending order and get the top 50 airports with incidents of bird strikes.
- Here are the few top airports.

Airport: Name	No. of Birdstrikes
DALLAS/FORT WORTH INTL ARPT	802
SACRAMENTO INTL	676
SALT LAKE CITY INTL	479
DENVER INTL AIRPORT	476
KANSAS CITY INTL	452
PHILADELPHIA INTL	442
ORLANDO INTL	408
BALTIMORE WASH INTL	401
LOUISVILLE INTL ARPT	394
JOHN F KENNEDY INTL	389
CHARLOTTE/DOUGLAS INTL ARPT	367
NASHVILLE INTL	364
LAMBERT-ST LOUIS INTL	363
CHICAGO O'HARE INTL ARPT	331
PORTLAND INTL (OR)	313
NEWARK LIBERTY INTL ARPT	305

Top Airports with incidents of bird strikes

- Grouping the data year wise, we can sum the cost of all individual incidents to get the yearly cost.
- Then we can use the data to plot a bar graph to visualize the data.
- Here is that bar graph.



When do most bird strikes occur

- Grouping the data by the phase of the flight, we can count the number of incidents during each phase.
- Use the data to plot a graph which will be as shown.

When: Phase of flight	No. of Birdstrikes
Approach	10151
Landing Roll	4946
Take-off run	4560
Climb	4247
Descent	763
Taxi	71
Parked	9

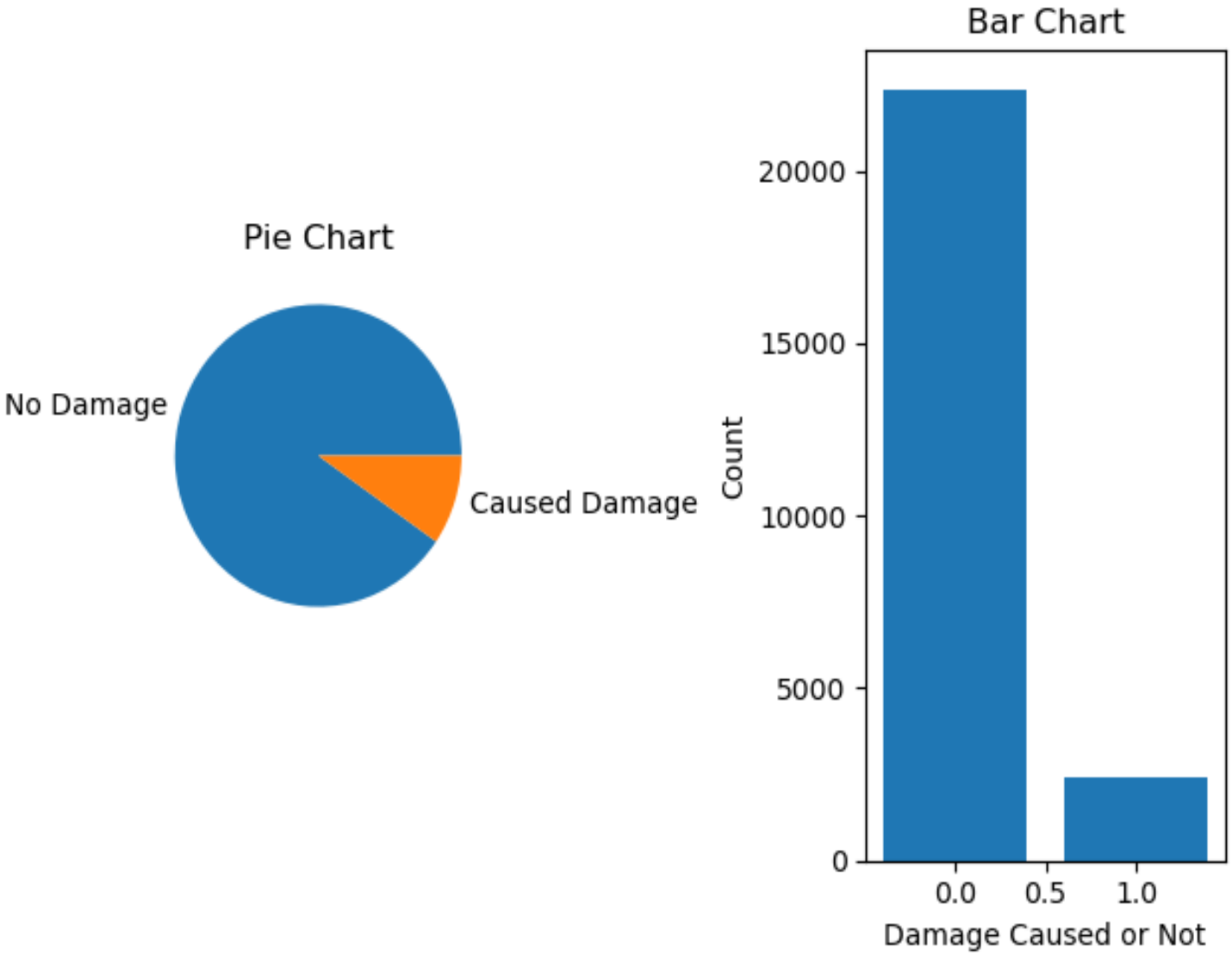
Altitude during bird strikes

- We can group the data phase wise and then calculate the average altitude during the incidents to get an idea of the altitudes.
- We can see the average altitude during each phase here.

When: Phase of flight	Feet above ground
Approach	1004.757856
Climb	1204.549564
Descent	5924.840105
Landing Roll	0.000000
Parked	0.000000

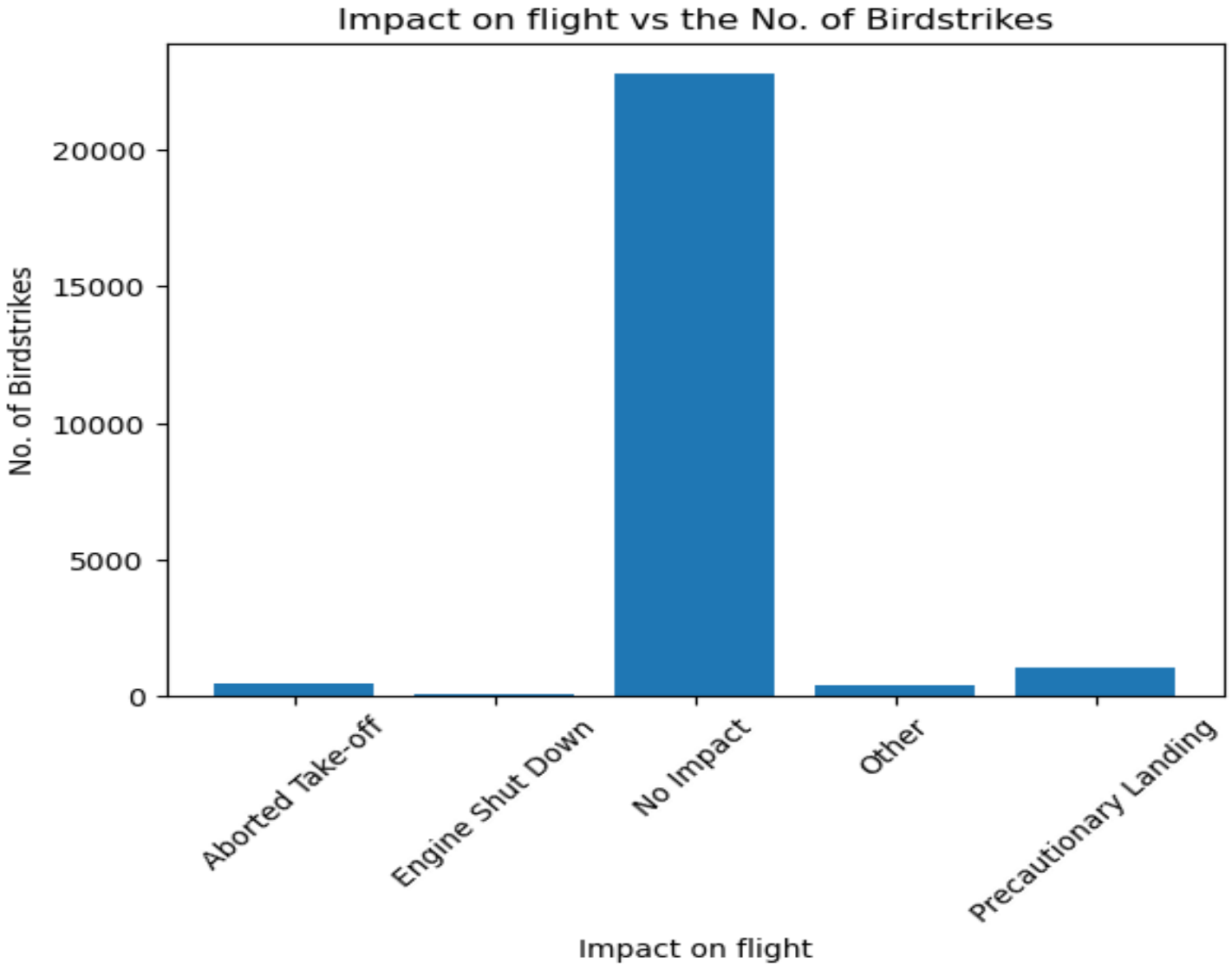
Bird Strikes causing Damage

- We can group the data on the basis of whether damage was caused as a result of the bird strikes and plot the graph using that data.
- Here is the graph.



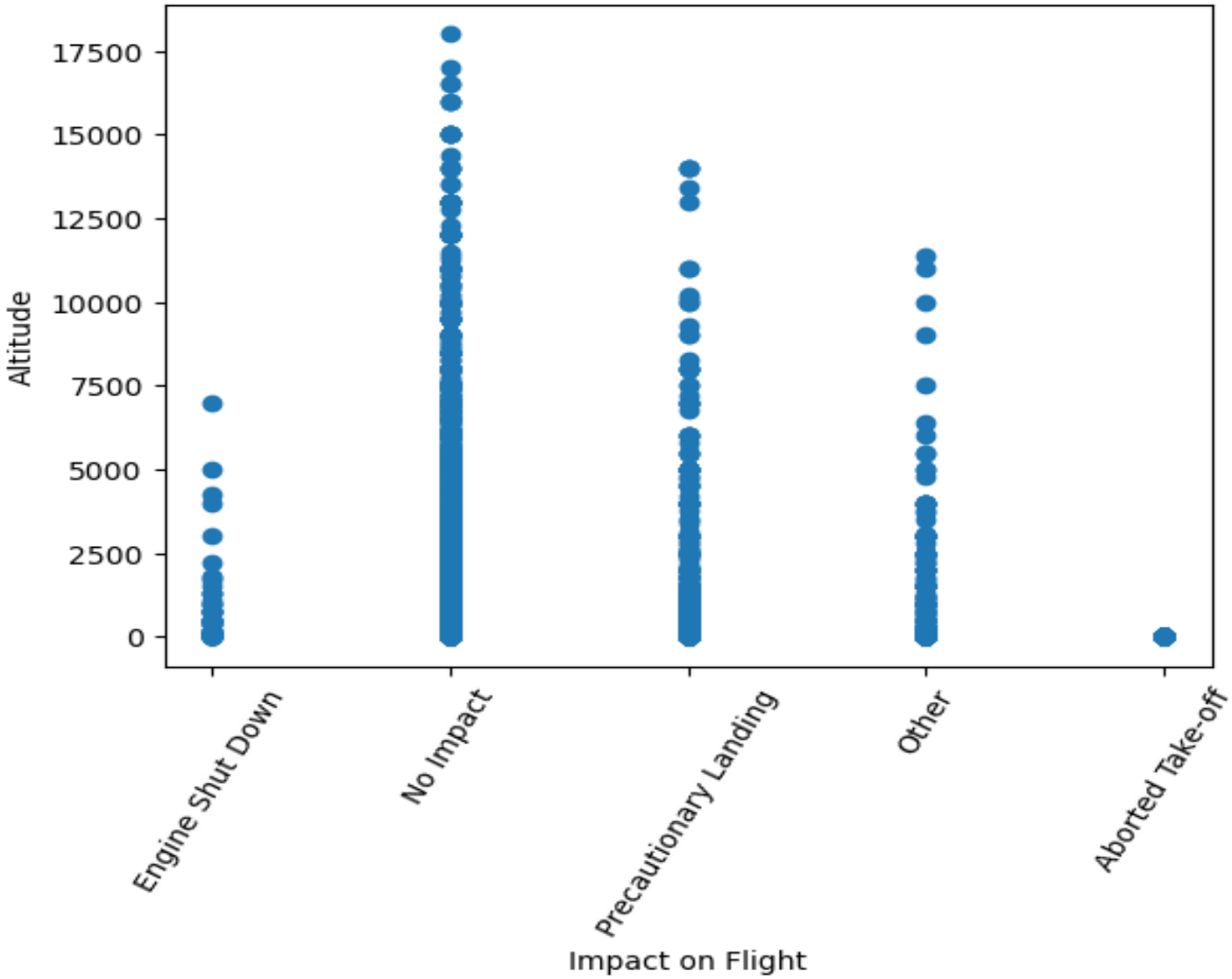
Bird Strike Impact on flight

- We can plot a bar graph that tells us what the impact was on a flight after a bird strike.
- Here is the bar graph.
- We conclude the majority of the flights suffered no consequences of the bird strike.



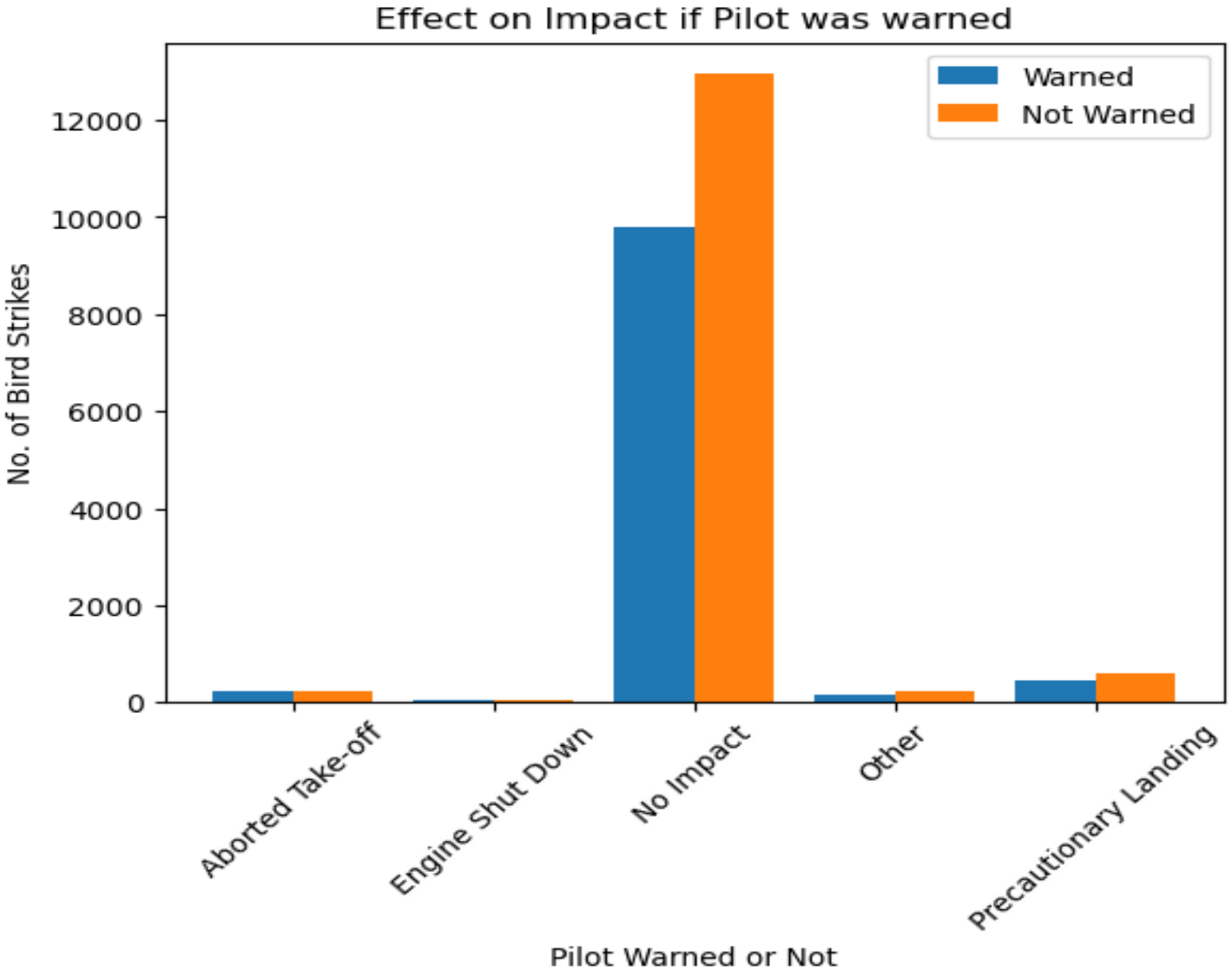
Effect of altitude on the impact on flight

- We can plot a scatterplot of the impact on flight vs the altitude at the time of the incident to see if the two are correlated.
- Here is the graph.



Effect of warning the pilot beforehand

- We can plot a double bar graph to check if warning the pilot beforehand softened the impact on the flight during the incident.
- Here is the graph.
- We conclude that warning the pilot does reduce the impact on airplane during impact.



THANK YOU