Data Appendix

Our dataset comes from the FAA (Federal Aviation Administration) website and contains information regarding bird strikes on airplanes from 1990-2023. From the original, raw dataset, we cleaned it by removing NA values, then created a splinter dataset with a count variable grouped by the year and time of day, giving us a dataframe with year, time of day, and the number of bird strikes for each year and time fo day combination. From there, we split this dataset by the time of day, giving us 4 separate dataframes with the same information for each time of day.

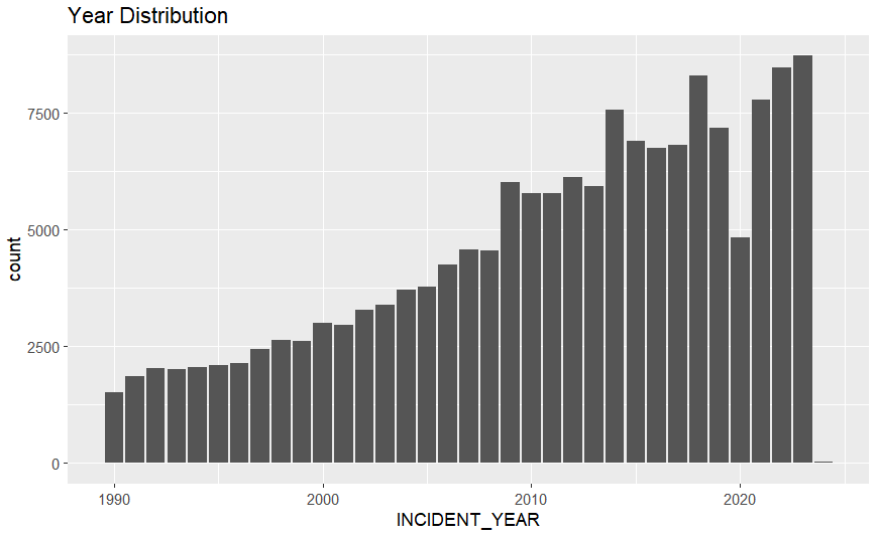
Final Dataset

Our final dataset (before the time of day splits) contained 3 columns. These columns are INCIDENT_YEAR (the year, ex. 1995), TIME_OF_DAY (categorical time of day from the options Dawn, Day, Dusk, Night), and n (the number of bird strikes from that year and time of day combination, ex. 1738).

Variable	Description	Potential Response(s)
INCIDENT_YEAR	The year that the incident occurred	2000, 2001
TIME_OF_DAY	The time of day that the incident occurred (categorical)	Day, Dawn, Dusk, Night
n	The number of strikes that occurred in this year at this time of day	40, 101

INCIDENT_YEAR

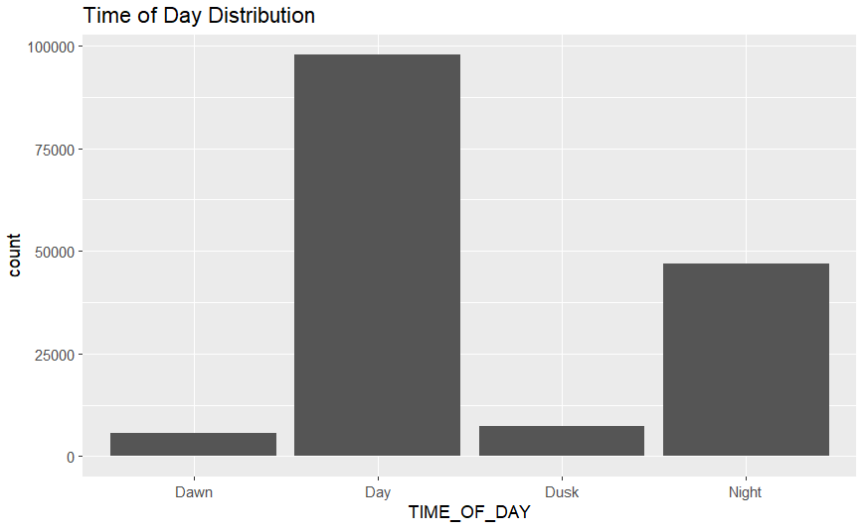
The incident_year column is from 1990-2023 and is a numerical value.



The number of wildlife strikes is increasing over the years.

TIME_OF_DAY

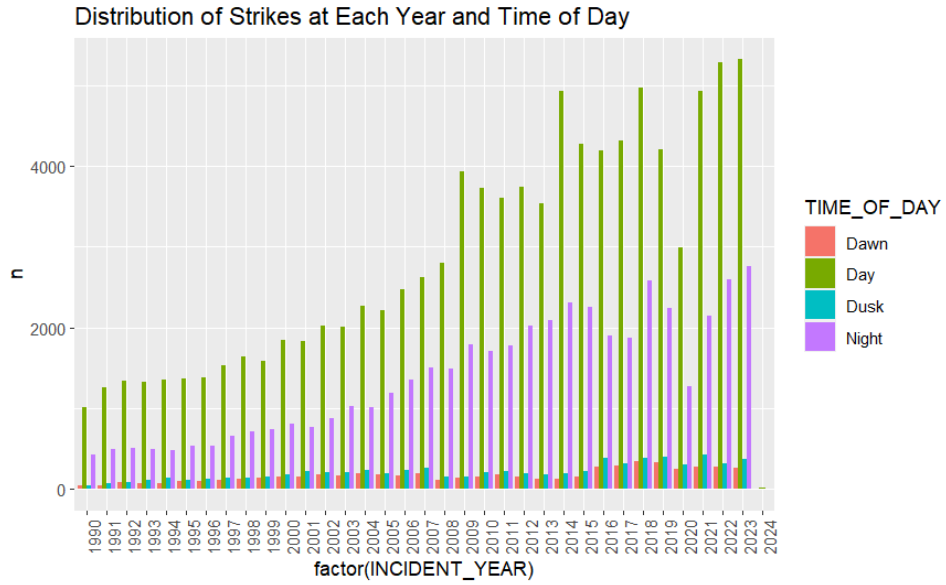
The time_of_day variable is a categorical variable with 4 different options.



There are more wildlife strikes during the day as well as at night.

n

The n variable is the count of wildlife strikes occurring during the specific year and time of day.



The count variable n is represented in the y axis.