Statistical Analysis for Possible Fuel Stops for Private Aviation

Trent Bazemore, Samantha Craig, Emily Diegel, Isabelle Hoare, John Scheer
Mihhail Berezovski Ph.D
Embry-Riddle Aeronautical University (ERAU)
OneSky Flight LLC, Michael Campobasso

Introduction

OneSky Flight is an aviation company under an umbrella of companies offering technology services for other private jet companies. OneSky and its established jet brands fly high paying customers all over the world, and it is important to know the expected time spent traveling. A challenge this company faces is having the ability to determine if a fuel stop is necessary upon the booking of a flight. When a customer books a flight, OneSky wants booking services to have the capability to inform customers of the likelihood of a fuel stop, and offer the option to upgrade their aircraft to reduce travel time. The goal is to minimize the need for last minute upgrades to bookings in order to increase customer satisfaction.

Data Analysis

Approximately 210,000 rows of data are used for analysis. 1.32% of recorded flights required a fuel stop. The variables provided in the original dataset include flight ID number, starting and ending airport codes, start and end of flight date and time, number of passengers, estimated flight hours, aircraft category name, and fuel stop status. The four categories of aircraft that required fuel stops are Light Jet, Midsize Jet, Super Midsize Jet, and Large Cabin. Of these variables, aircraft category is important to focus on for booking purposes. Figures 1 and 2 demonstrate that Light Jet category aircraft are the most frequently booked flights, but Midsize Jet's had the highest percentage of fuel stop flights at 2.08%.

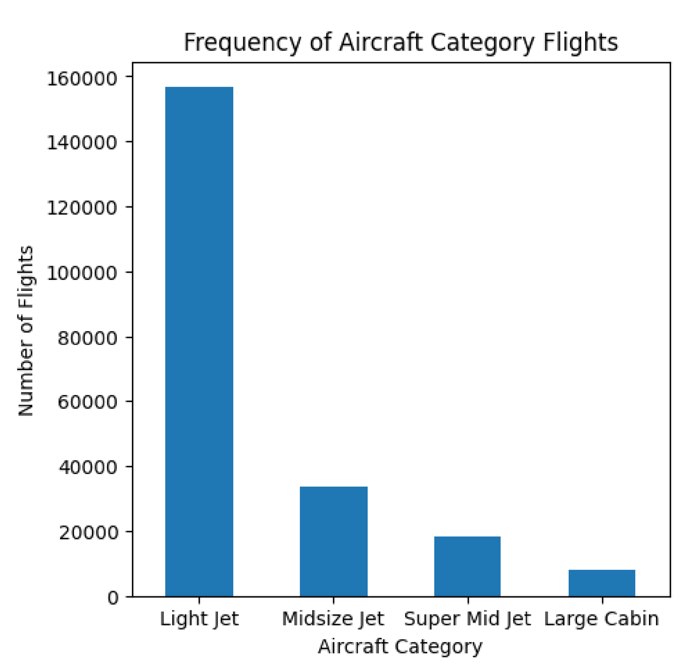

Figure 1: Frequency of flights based on aircraft category.

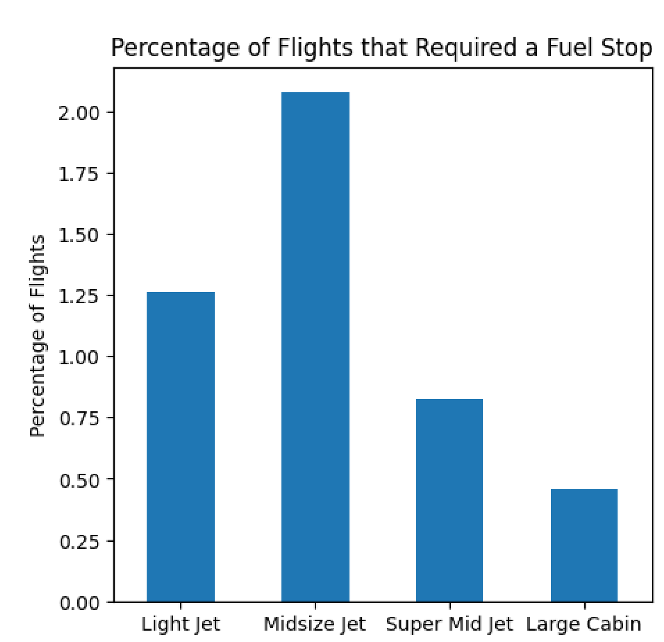

Figure 2: Percentage of flights based on aircraft category that required a fuel stop.

Figures 3 and 4 demonstrate the amount of time spent flying across all categories of aircraft. From frequency, it can be seen that most flights take place between 1-3 hours, and this remains true for fuel stop flights as well. To mitigate the skew cause by frequency, when observing the percentage of flights that required fuel stops, it can be seen that the data peaks between 6 and 9 hours of flight time.

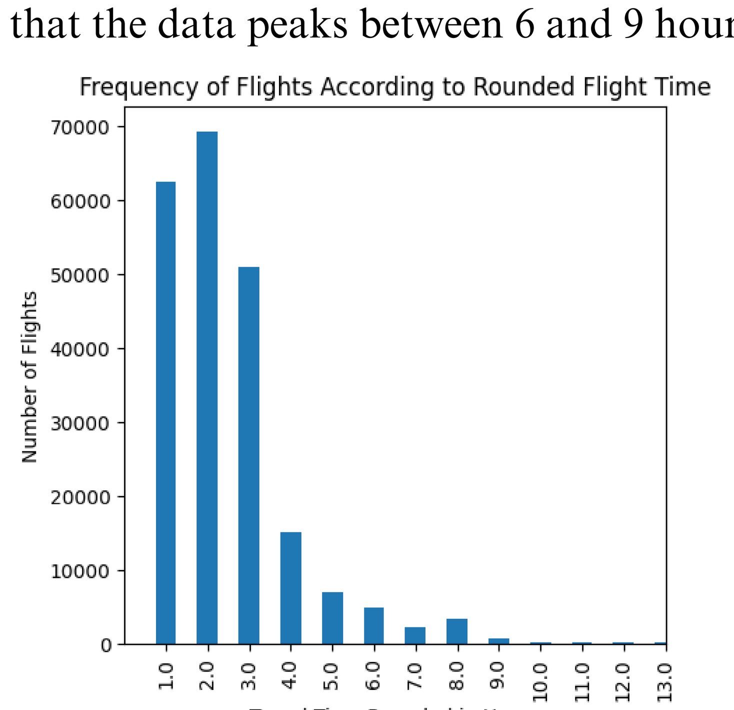

Figure 3: Frequency of flights based on rounded flight time.

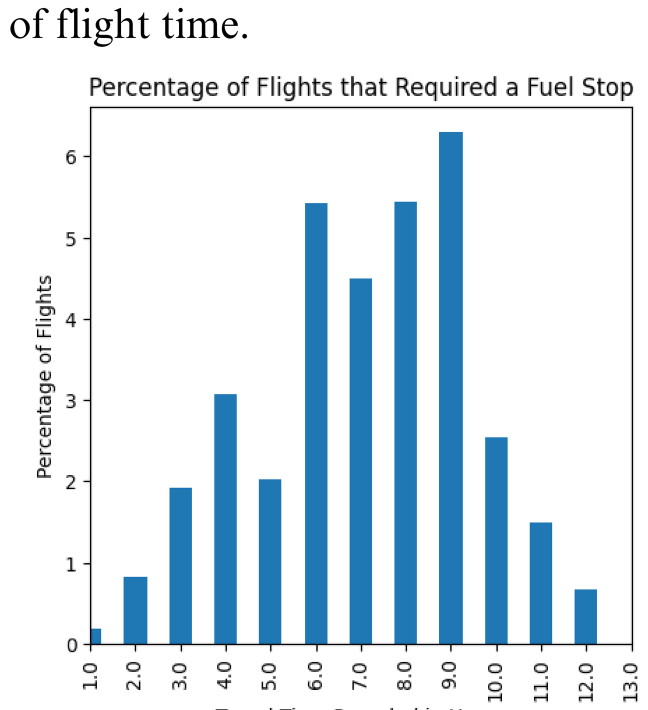

Figure 4: Percentage of flights based on rounded flight time that required a fuel stop.

Methodology

Feature importance of type weight, gain and cover for the XGBoost model. Looking at the XGBoost Model, distance is the most important feature in determining if a flight has a fuel stop.

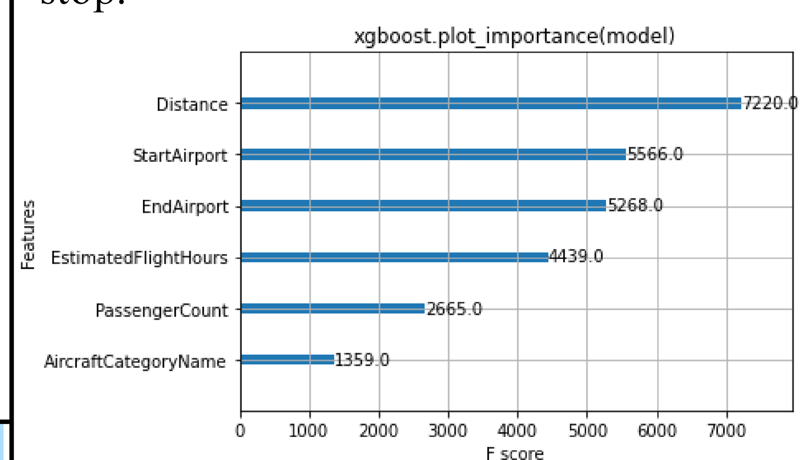

Figure 5: Feature importance of type weight. The number of times a feature is used to split the data across all trees.

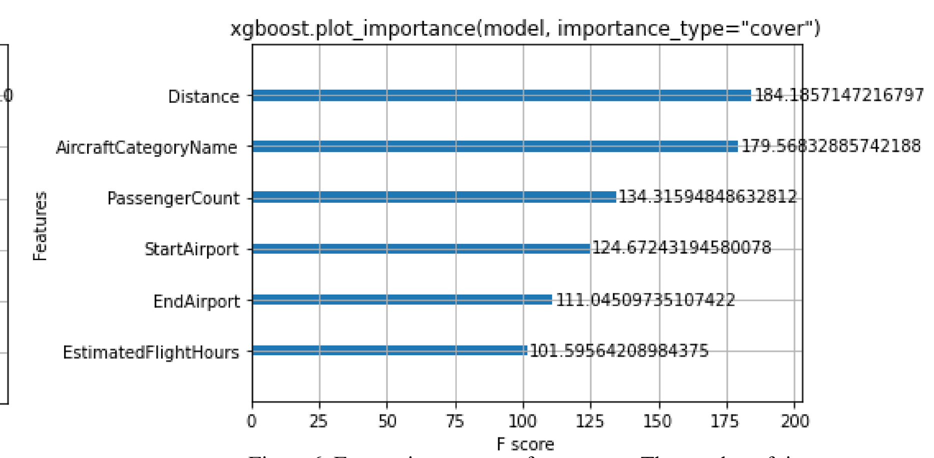

Figure 6: Feature importance of type cover. The number of times a feature is used to split the data across all trees weighted by the number of training data points that go through those splits

Location may play a determining factor regarding the likelihood of a fuel stop. In figure 7, a choropleth map is used to illustrate the percentage of flights arriving and departing from different U.S. states. States colored in red had no flights entering or leaving that required a fuel stop. Hawaii has the largest percentage of departing flights that required a fuel stop, while Alaska, West Virginia, and North Dakota had the lowest at 0%. For arriving flights, Alaska had the largest percentage of fuel stop flights, while North Dakota is lowest at 0%.

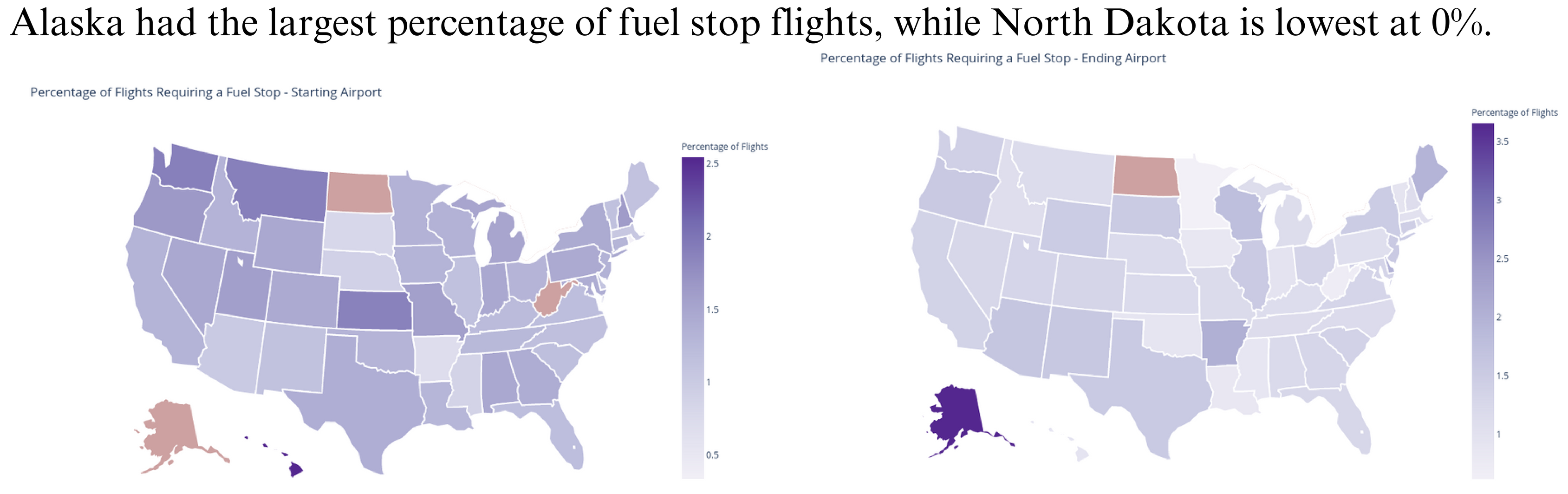

Figure 7: Percentage of flights requiring a fuel stop associated with state based on departing airport (left) and arriving airport (right).

Distance is a influential factor in determining if a flight will have a fuel stop. The graph shown in figure 8 provides insight on what distances seem to generate the most amount of fuel stops. It is shown that the distances ranging from (1000-2000) KM produce the most fuel stops for the Light Jet category.

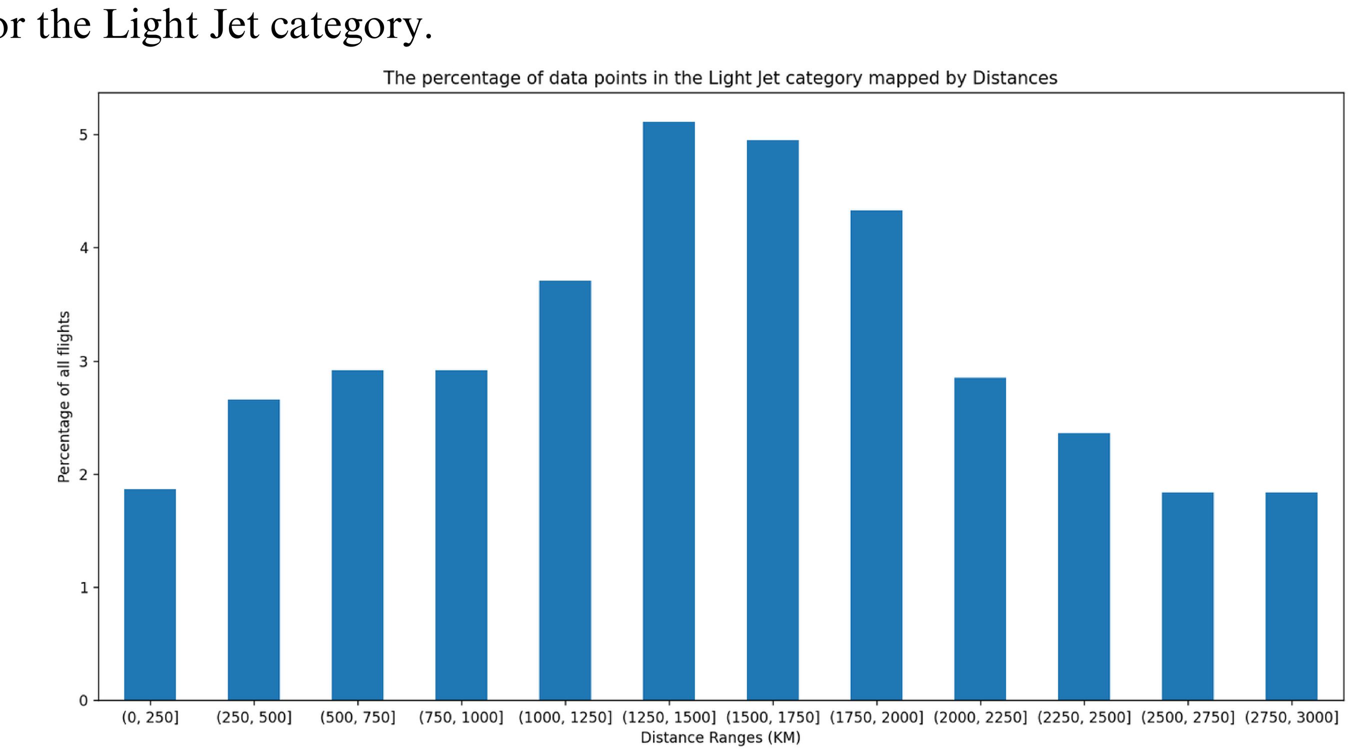

Figure 8: is a representation of the Light Jet Aircraft percentage of fuel stops in the y-axis and the Distances in KM on the x-axis

Understanding the thrust, speed, and acceleration of the aircraft is a variable that drastically impacts fuel consumption. This data was not provided and needed to be calculated, but with the data given, the calculations become skewed. Figure 9 demonstrates that on average, Light Jet, and Large Cabin aircraft are traveling longer distances for flights associated with fuel stops, but the opposite is true for Midsize Jets and Super Midsize Jets. When calculating distance over time, speed is impacted by the fuel stop flags, making speed calculations unreliable, as illustrated by figure 10, with lower speeds associated with all fuel stop flights.

Methodology Cont.

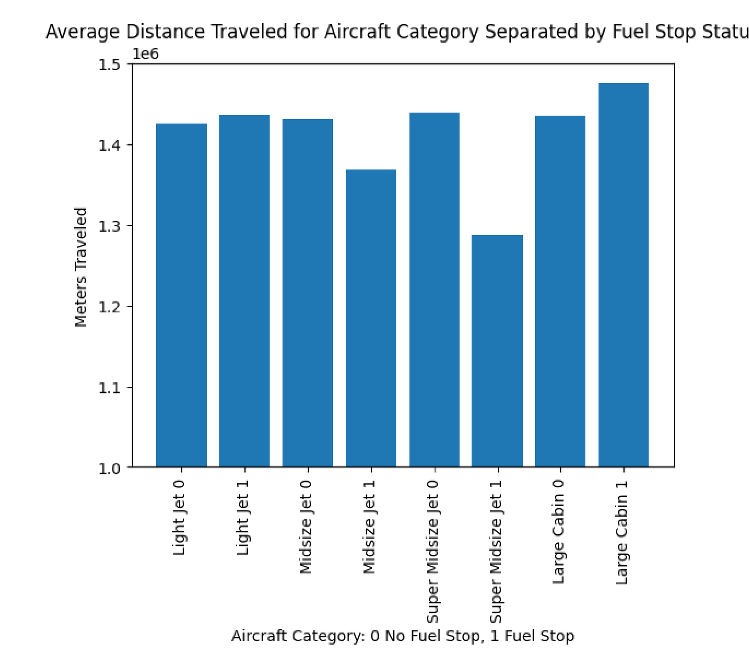

Figure 9: Average distances traveled in meters classified by aircraft category and fuel stop status, 0 being no fuel stop, 1 being positive fuel stop flag.

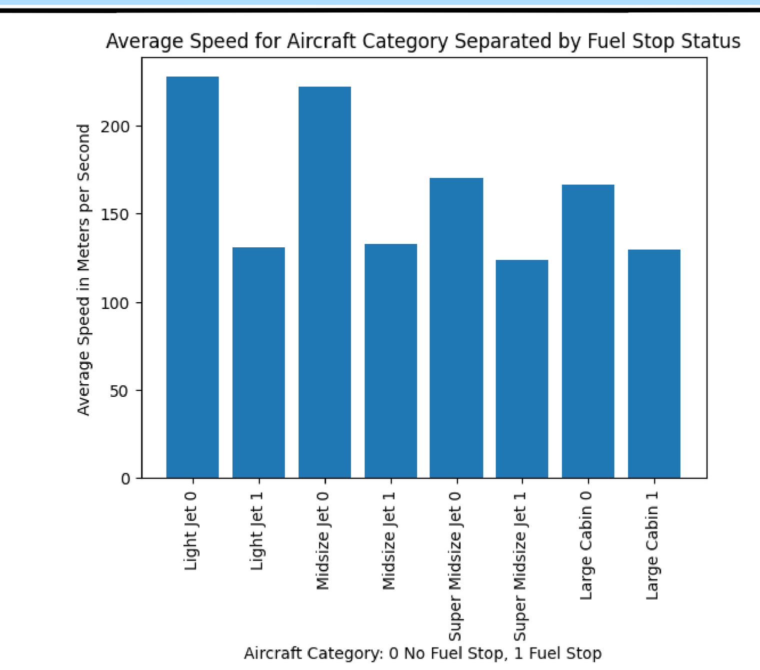

Figure 10: Average speed in meters per second classified by aircraft category and fuel stop status, 0 being no fuel stop, 1 being positive fuel stop flag.

Results

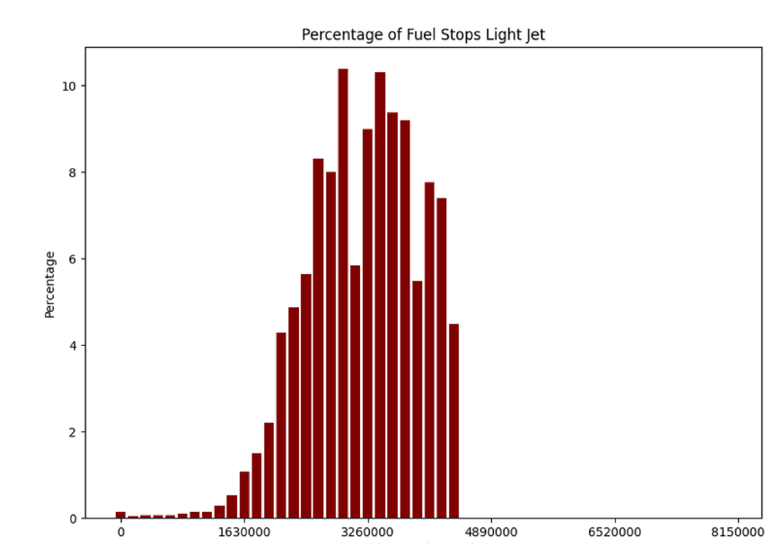

Figure 11: Likelihood of a fuel stop for specific distances in the light jet.

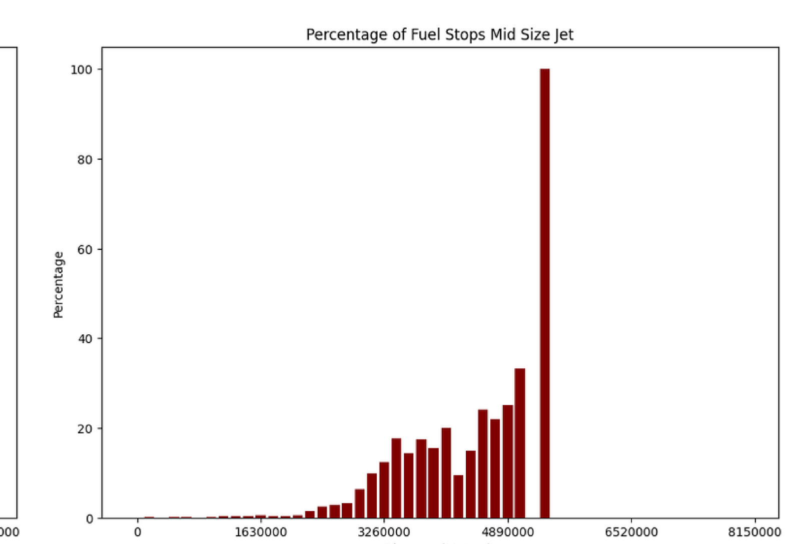

Figure 12: Likelihood of a fuel stop for specific distances in the mid size jet.

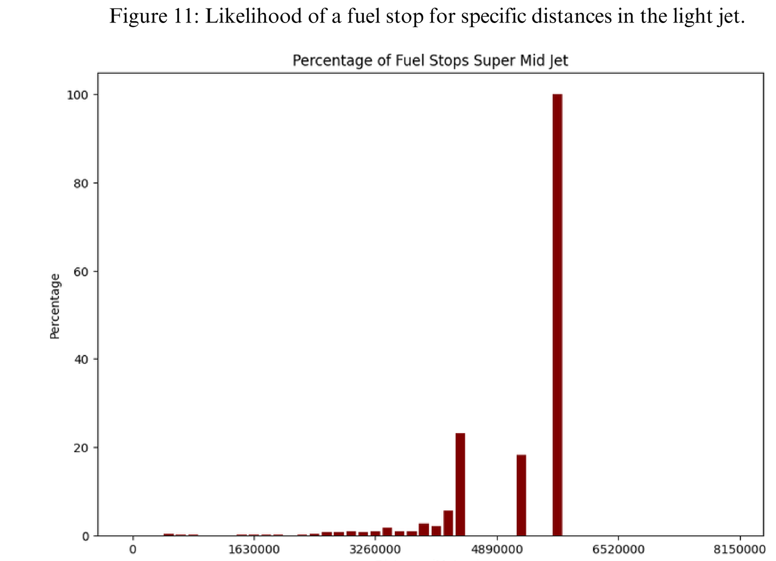

Figure 13: Likelihood of a fuel stop for specific distances in the super mid jet.

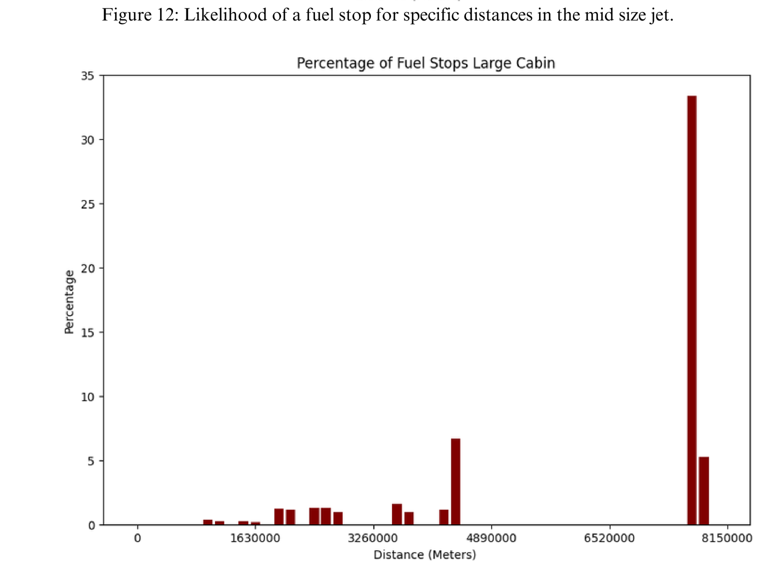

Figure 14: Likelihood of a fuel stop for specific distances in the large cabin.

The figures above demonstrate the likelihood of a fuel stop in specific types of aircrafts for a range of distances. For the light jet, the likelihood of having a fuel stop increases to around 8 percent for distances between 1630000 and 4800000 meters. For the mid size jet, distances above 3260000 show a higher likelihood of a fuel stop ranging between 20 and 40 percent. In the super mid jet, distances above 4000000 meters tend to jump up into the 20 percent likelihood of having a fuel stop. In figure 14, there is no reasonable trend for the large cabin in fuel stops, and as it is understood, the large cabin tends to have least amount of fuel stops as it is a bigger plane.

Further Research

More variables and data are needed to fully develop a program to predict fuel stops accurately for private aviation. Access to thrust and acceleration data would be pertinent to understanding fuel consumption of the different aircraft categories. It is also difficult to apply machine learning algorithms to the data, because so few fuel stops occur, classifier models and feature importance models become overfit. More time and data utilized to train a model could assist in creating the desired program OneSky is seeking to improve their customer satisfaction.