

The background is a stylized illustration of an Alaskan landscape. It features rolling green hills, a large snow-capped mountain on the right, and several dark green evergreen trees. A large, pale sun or moon is on the left, and a few birds are flying in the light blue sky. The overall style is flat and modern.

Alaskan Bush Pilot Safety

Inferences on contributing factors to accidents and fatal injuries

Bush Pilots of DSI

Mike Minikowski, Nolan Smurro, Lisa Paul, Muhammad Hasan, Sophia Joseph



Table of Contents

01

**Problem
Statement**

02

**Cleaning & Feature
Engineering**

03

EDA



04

**Modeling
Process**

05

Results



06

Conclusion



Background

- Our client: Alaska Airmen's Association:
 - A nonprofit established in 1951 to promote general aviation, enhance safety, and support initiatives that benefit pilots
 - Analyze NTSB data concerning aviation incidents/accidents involving single engine aircraft in Alaska
- The Airmen's Association aims to understand the predominant factors contributing to aviation incidents involving single-engine aircraft in Alaska, particularly those leading to fatalities
- With this information - the association can further develop targeted safety initiatives to aid in their mission to advocate for access and safety infrastructure across the state

<https://alaskaairmen.org/about/>



Alaskan Bush Pilots

- Operate throughout the vast and rugged Alaskan landscapes
- Most often fly single engine light aircraft
- Navigate varying weather conditions and unconventional rough landing terrain
- Usually the only feasible line of delivery and transportation to isolated communities
- Crucial to search and rescue operations

Problem Statement

- Can significant contributing factors to serious aviation accidents involving Alaskan Bush Pilots be inferred?



Data Source

- NTSB (National Transportation Safety Board) Aviation Investigation database

Features of Interest

- Information about an event that would be known about a flight before or while it was occurring such as:
- Location (on flight route)
- Details on aircraft itself
- Weather

Omitted from Initial Inference

- Details of an observation that could only be known post-event, such as:
- Probable cause
- Level of damage to aircraft
- Number of injuries

Key Definitions

Term	Definition
Highest Injury Level	Scale for worst injury that occurred as a result of an event with values “None Reported”, “Minor”, “Serious” and “Fatality”
Fatality	Injury resulting in death during the event or within 30 days after
Serious Accident/Event	An event resulting in at least one fatality
IMC / VMC	Relates to weather condition (Visual / Instrument Meteorological Conditions) where VMC is ideal, IMC indicates poor weather
Aircraft Family	A grouping of similar aircraft make/models whose variations are often just a result of a different years’ variation

Cleaning & Feature Engineering



Cleaning

- Non-linear process
 - During EDA and even after evaluating results, additional opportunities to clean/engineer were discovered
- Notable cleaning tasks:
 - Aircraft family grouping using extensive RegEx
 - Replacing nulls with 'unknown' where appropriate
 - Eliminating upper/lower case discrepancies



Feature Engineering

Term	Definition
Created refined target variable	The target variable was engineered such that events with a fatality were the “positive” case 1, otherwise 0
Created “Occurred Near Airport”	Binary value representing whether the event occurred at or within 3 miles of an airport
Bundled skewed categorical values	Categorical values such as make/model/purpose of flight with disproportionate values (some very common values and many uncommon) were bundled to reduce class imbalance. Less common entries were grouped as “uncommon”

EDA



EDA Process



Step 1

Exploring



Step 2

Organizing



Step 3

Picking Features
to Model



Step 4

Graphing and
Visualization



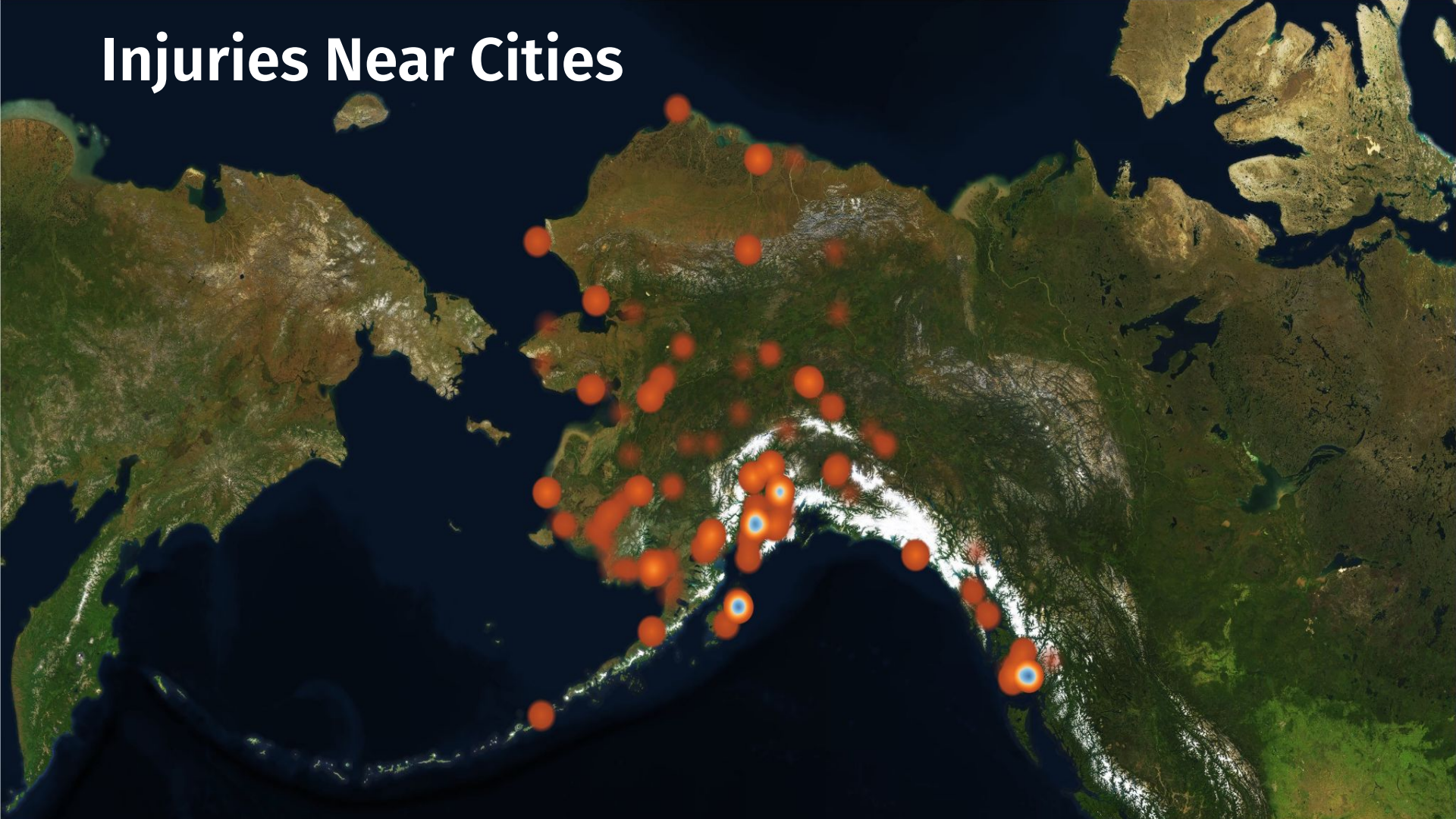
Step 5

Hypothesis

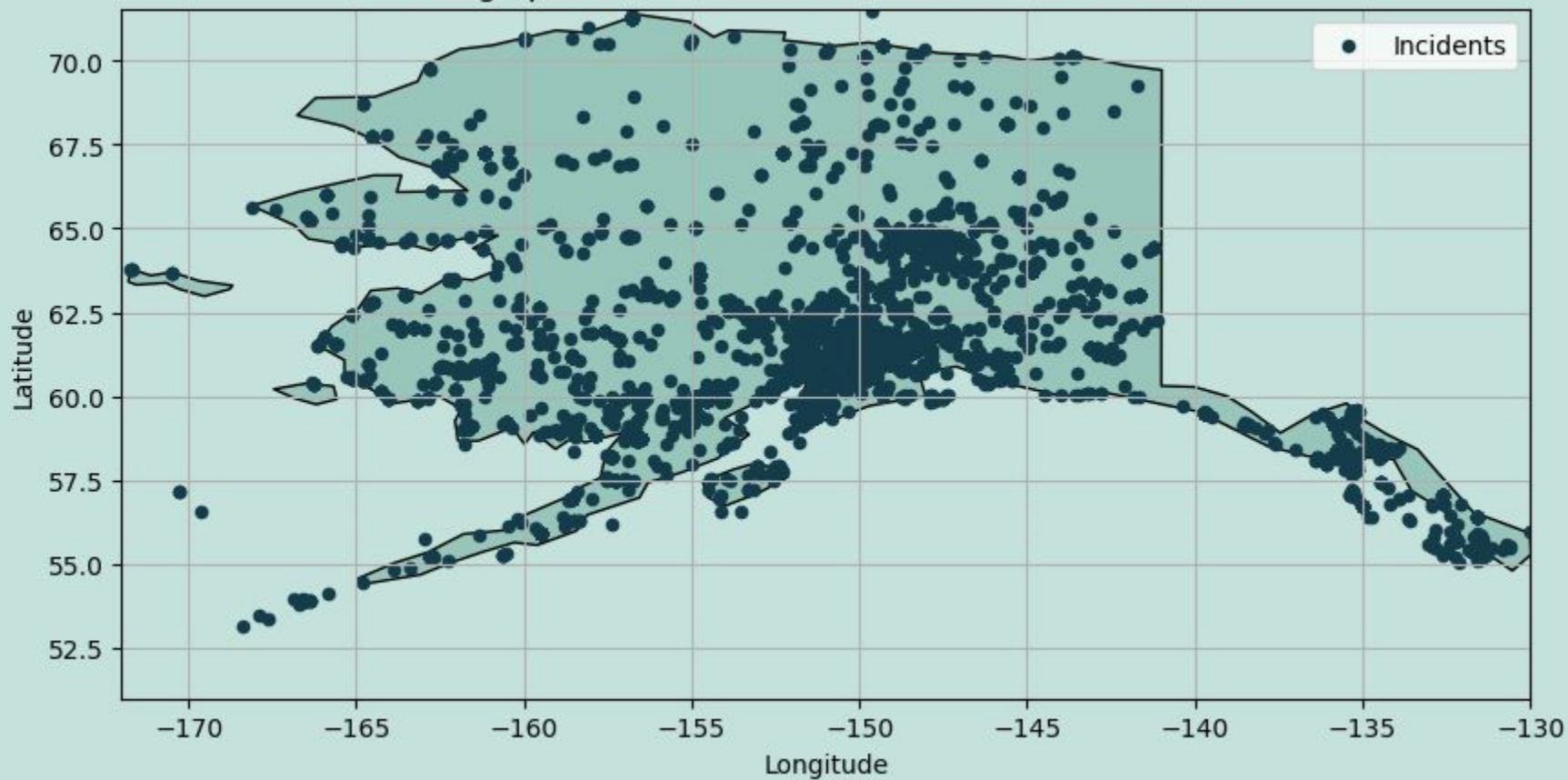
One in every 78 Alaskans is a pilot, which is six times more pilots per capita than anywhere else in the United States. (Source: The Guardian)



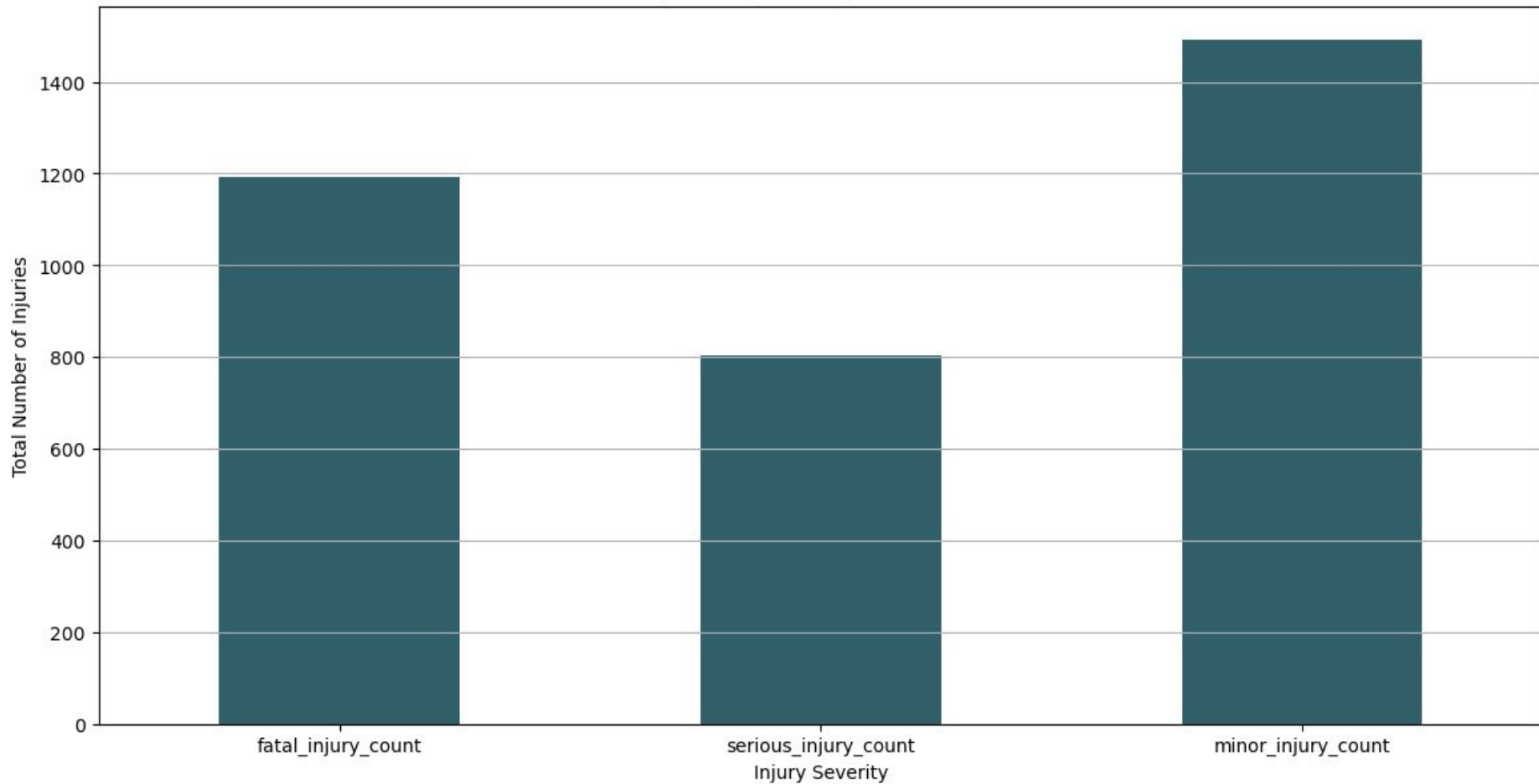
Injuries Near Cities



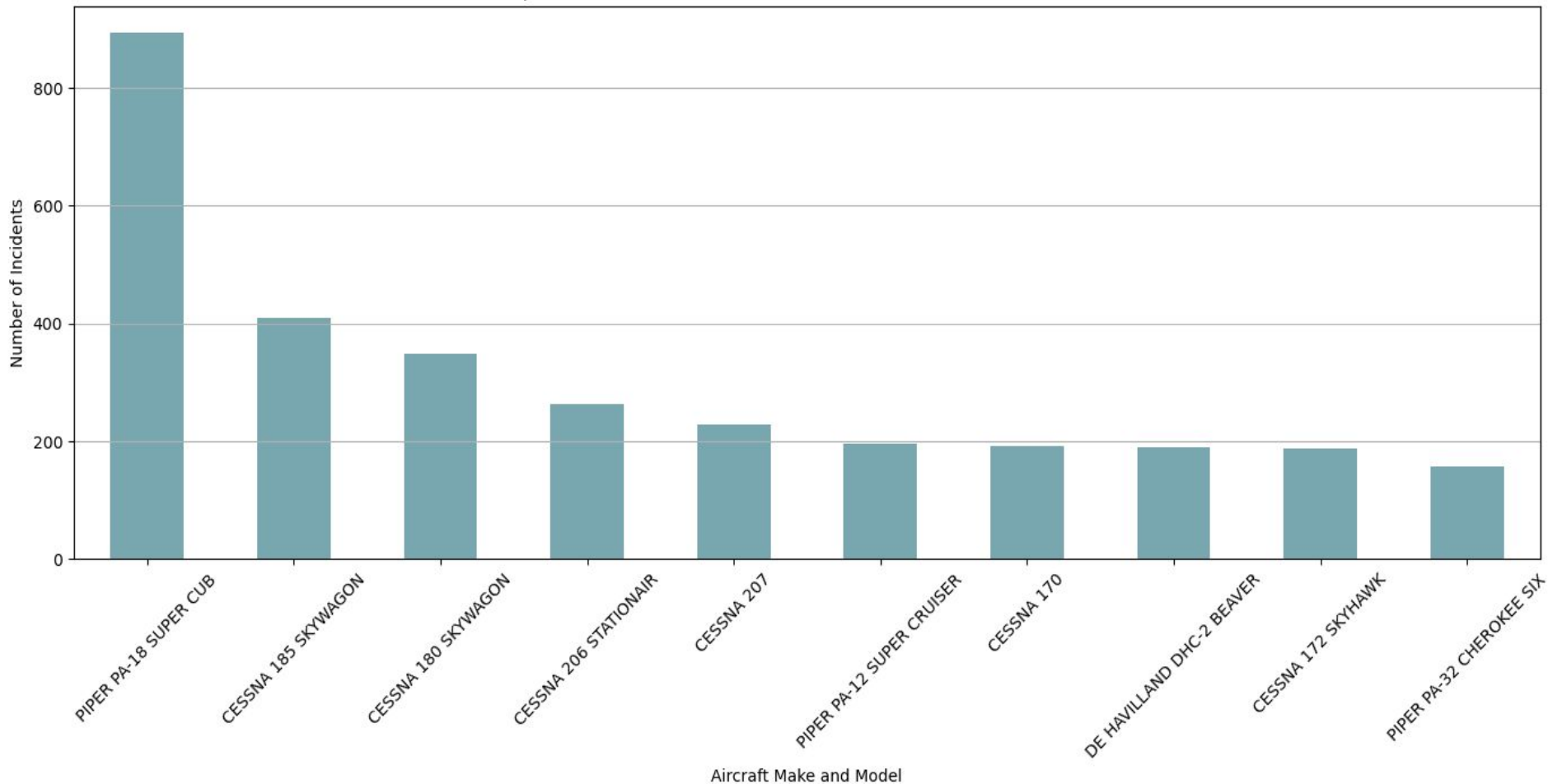
Geographical Distribution of Aviation Events in Alaska



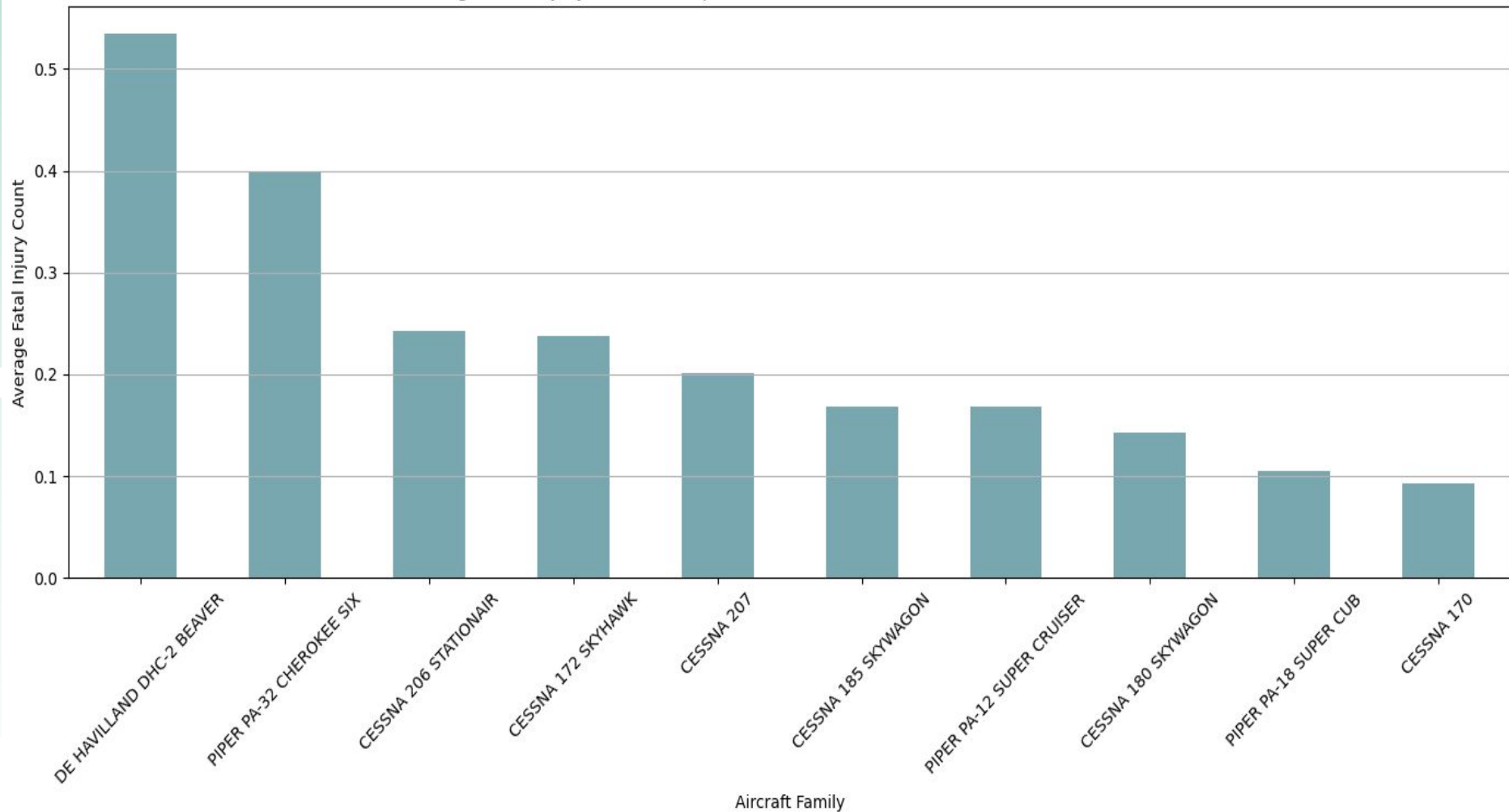
Total Number of Injuries by Severity for Aviation Events in Alaska



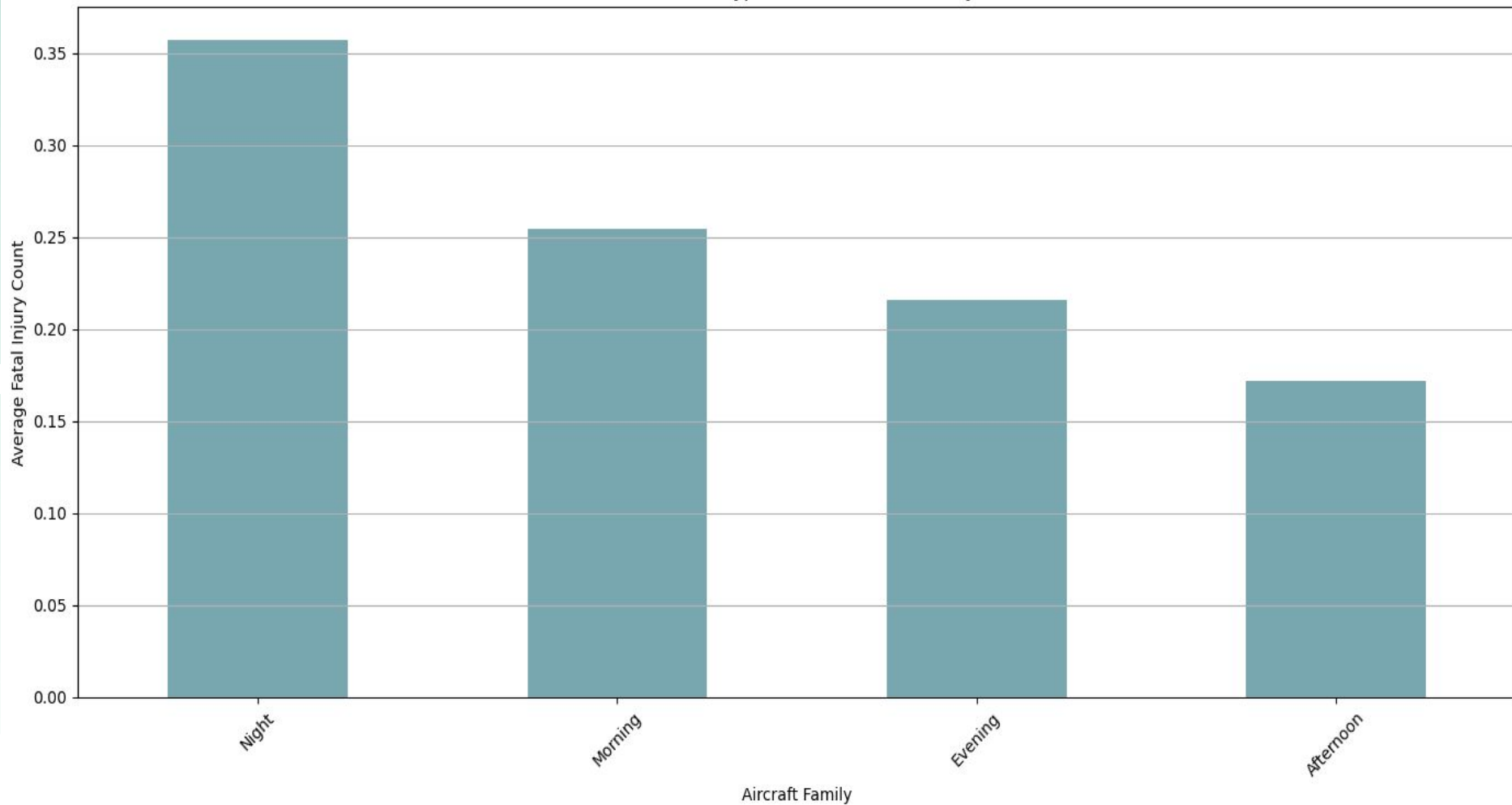
Top 10 Aircraft Makes and Models Involved in Incidents in Alaska



Average Fatal Injury Count for Top 10 Aircraft Families Involved in an event in Alaska



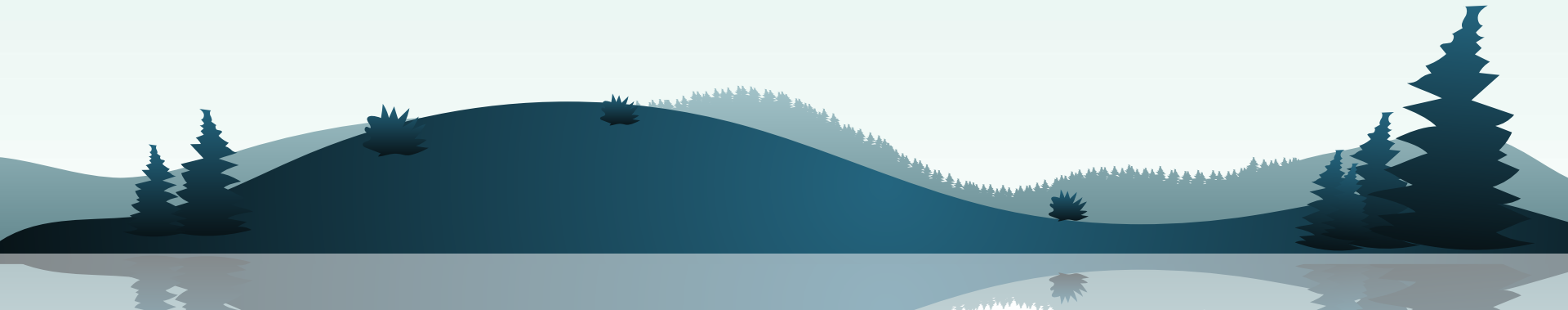
Distribution of Event Types based on time of day in Alaska



Hypothesis

Alternative Hypothesis (H_1): Weather, proximity to an airport/city, time of day and aircraft type do have a correlation with average fatal injuries.

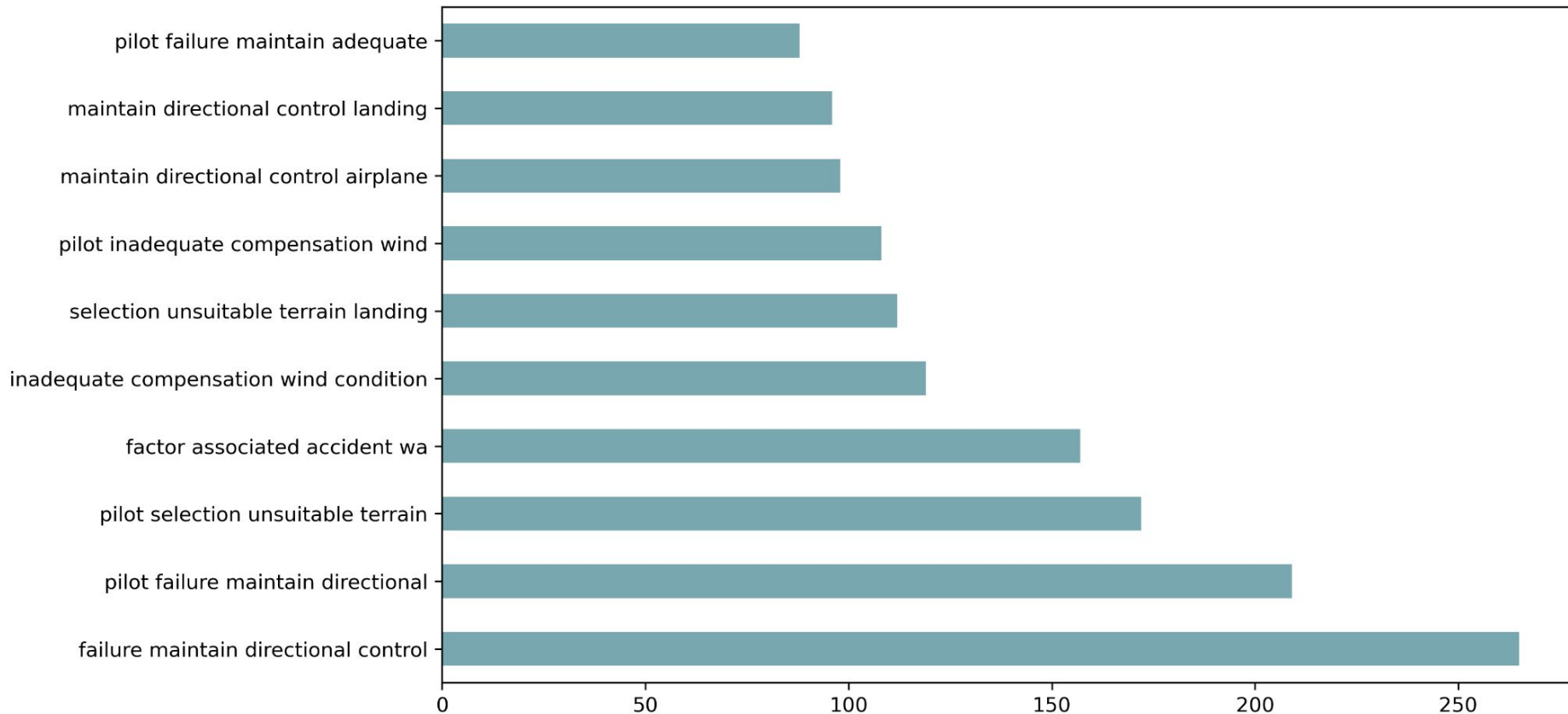
Null Hypothesis (H_0): Weather, proximity to an airport/city, time of day and aircraft type do not have a correlation with average fatal injuries.



Probable Cause vs Highest Injury

Features	Importance
landing	0.016870
flight	0.015610
pilot	0.014021
stall	0.013133
terrain	0.012580
instrument	0.011043
airspeed	0.010541
failure	0.010201
resulted	0.010013
condition	0.009662

Probable Cause vs Highest Injury



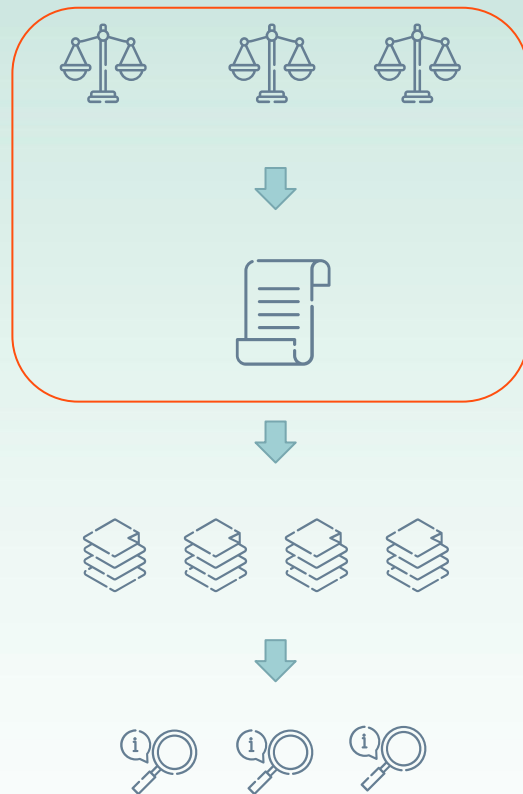
Modeling Process



Modeling Process

Step 1 - statsmodels GLMs

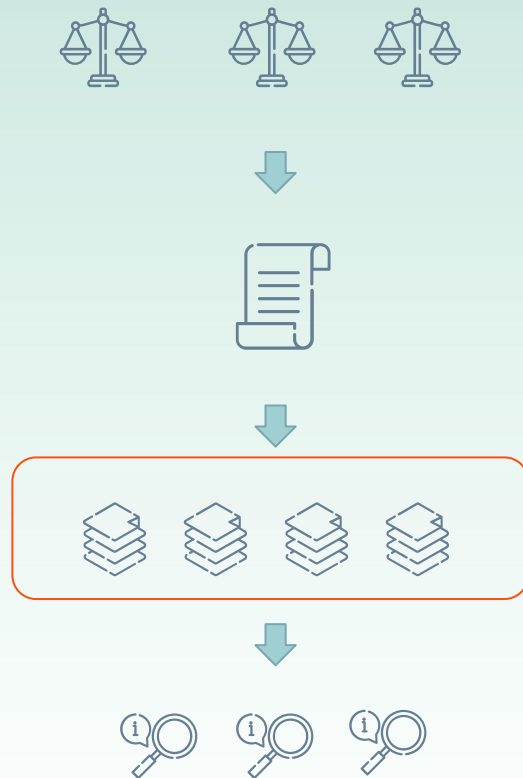
- First needed to determine where to start with predictor variables
- Extracted sources of statistical significance through use of Generalized Linear Models, specifically Binomial
- Using accuracy as a chief metric, GLM models eventually slightly outperformed baseline, which in turn provided p-values
- Expedited this process with a new class that included methods that performed preprocessing, train-test splitting, fitting, and storage of summary results in one method



Modeling Process

Step 2 - Sci-Kit Learn Logistic Regressions

- With p-values, we set up logistic regression models each focused on very small subsets of significant features
- Variations of the same logistic regression model with different reference categorical values were run in order to generate easier to understand inferences
- These product coefficients, which now gave us magnitude and direction (the overall effect in terms of odds) of features like time of day



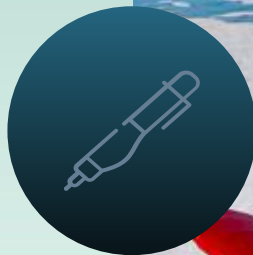
Modeling Process

Step 3 - Inferences

- With statistical significance, coefficients with magnitude and direction in hand, we could make inferences in terms of which factors were likely to contribute to a more serious accident



Results



Parameters

Feature	p_value	Coefficient	Coefficient Reference
weather_condition_VMC	0	0.562	IMC
occurred_near_airport	0.000001	1.111	0
event_time_of_day_Night	0.0001	1.090	Morning

Conclusion

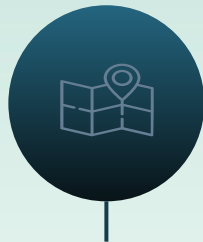


Key Insights/Recommendations



Weather Conditions/Visibility

Flying in VMC conditions leads to a 44% decrease in odds of fatal injury compared to flying in IMC conditions



Proximity to Airport






Accidents at or within 3 miles of an airport increased likelihood of fatal injury by 11%



Time of Day

Flying at night leads to a 9% increase in fatal injury compared to flying in the morning

Possible Considerations

	 Weather Conditions/Visibility	<ul style="list-style-type: none">• Encourage members to seek and maintain their instrument ratings.• Promote workshops and further resources related to operating in IMC conditions.• Promote the use of weather forecasting and flight planning tools.• Advocate for improvements in air traffic control procedures.
	Proximity to Airport	<ul style="list-style-type: none">• Promote the use of ground proximity warning systems (GPWS).• Advocate for improvements to airport infrastructure and runway conditions.• Work with airport authorities to develop and implement safety procedures for low-altitude operations.
	Time of Day	<ul style="list-style-type: none">• Promote workshops and resources on night flying safety.• Promote the use of night vision goggles (NVGs) and other night flying equipment and training where applicable.
	General	<ul style="list-style-type: none">• Host workshops or seminars on fatal crash prevention.• Develop and distribute educational materials on fatal crash risk factors.

*Disclaimer: We are not aviation experts and defer to the association. Text furnished for discussion purposes only

