Air Crash Analysis of Top Airlines in the US

Rahul Shrestha
The University of Findlay

Introduction

- Air Travel is one of the most convenient means of transportation used worldwide
- While the safest means to travel, with only 0.07 deaths per one billion passenger miles, air crashes are probably the most feared type of transportation accidents
- This research will help us determine which factor(s) to look at most carefully to make the next flight safer

Goals

- My goal in this project is to predict the most probable cause behind an aircraft accident
- For this, I will be running a multinomial logistic regression with 7 independent variables (day, region, season, lighting, weather, aircraft make and airframe hours), that would influence the dependent variable
- The cause of the accident, our dependent variable, will be one of four categorized causes: environmental, equipment, airport or flight crew

Customers

- Airline corporations
- Pilots & crew members
- Passengers
- Aircraft manufacturers
- Air traffic controllers

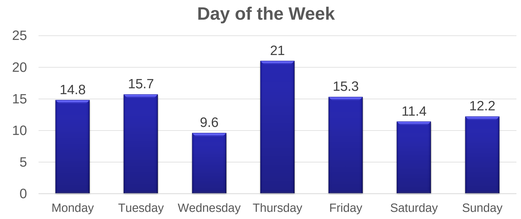
Operationalized Research Questions

- Which Airlines have the least equipment faults?
- Which Airlines have the least flight crew faults?
- Can we predict a future cause of accidents?
- Which region in the US has the least number of accidents due to airport faults?
- Which day of the week has the least number of accidents?

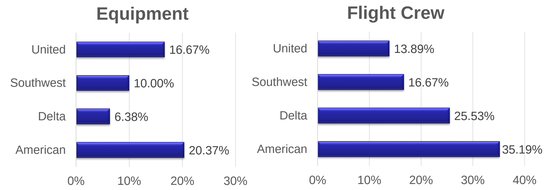
Data, Tools and Methods

- Data source: NTSB website
- MS-Access for queries and preliminary filtering of data
- MS-Excel for advanced filtering and chart creation
- SPSS for regressions
- R studio for regression and prediction model generation
- Data from years 1988 to 2017 (30 years) is used to generate the predictive model
- The model is then tested on the available data to determine its accuracy
- The model is then used to predict causes for accidents in the future

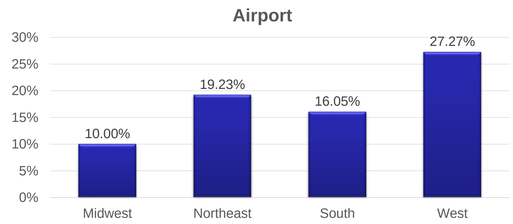
Outcomes



- Thursday is the worst day of the week to fly with 21% of accidents occurring on that day
- Wednesday is the safest with only 9.6% of accidents occurring on that day



- Delta airlines has the most dependable equipment and United airlines the most dependable flight crews
- American airlines does bad in both equipment and flight crew reliability
- Southwest airlines is the best airlines overall



- Midwest has the best airports with only 10% of the accidents occurring due to the fault of airports
- West has the worst airports as 27% of accidents occur due to a fault of the airport

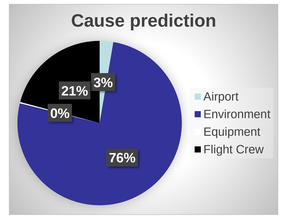
Prediction

Let's imagine a flight on a Saturday evening, departing from Boston, in January, with cloudy weather, on a Boeing aircraft with 12,000 hours of flight time.

daycat	region	season	light	weather	make	airhrs
WE	NE	WI	NITE	IMC	Boeing	12000

My multinomial logistic regression model predicted around a 76% for an accident caused by the Environment.

Airport	2.67%
Environment	76.19%
Equipment	0.25%
Flight Crew	20.87%



Conclusions

- Southwest airlines is the best airlines overall
- This model can be used to predict the cause of an accident if we know all the independent variables involved
- The Midwest region has the best airports in the US
- Wednesday is the best day to fly, Thursday the worst

References

- Aviation Accident Database & Synopses. Retrieved June 1, 2019 from https://www.ntsb.gov/_layouts/ntsb.aviation/index.aspx
- Mike Crowson. (2019, July 11). Multinomial logistic regression using SPSS (July 2019) [Video]. YouTube. https://www.youtube.com/watch?v=1BL5cL8_Cyc
- Dr. Bharatendra Rai. (2015, November 12). Multinomial Logistic Regression with R: Categorical Response Variable at Three Levels [Video]. YouTube. https://www.youtube.com/watch?v=fDjKa7yWk1U&t=12s

Thank you, Dr. Leventhal, Dr. Geise, Prof. Samimi, Jeffrey Zimmer and John Farhat for all your support.