

Aviation Accidents and the Patterns

Pradip Karki

Washington University at St. Louis
School of Engineering & Applied Science
St. Louis, U.S.A.

Pradip.Karki@wustl.edu

Abstract— In today’s world, big data and big data analytics plays a greater role to improve knowledge-based tasks, and to uncover hidden patterns and other insights. With today’s technology, it is possible to analyze the data and get answers from it almost immediately. This paper is intended to discover the aviation accident patterns on each decade. The paper also focuses on independent variables: weather, scheduled or non-scheduled flights, phase of flight, types of engine, number of engines used and amateur built, and looks into the statistical significance of patterns in aviation accidents. This paper utilizes factor analysis to describe variability among observed, correlated variables and correlation analysis to observe the association between continuous variables. The paper uses descriptive crosstabs and explores row percentages, column percentages and cell percentages to compute statistical results.

The statistical result shows the aviation accident patterns and affects of independent variables on accidents. Finally it concluded that cruise flight is smoother and safer than 1980s. Non-scheduled flights are responsible for 76 percent of aviation accidents where total number of fatal injuries is more than or equal to 1. Approximately 7.4% of general aviation accidents were weather related yet weather related flights account for 22% of fatal injuries where total number of fatal injuries is more than or equal to one. Almost 83% of aviation accidents occurred in single-engine piston aircraft.

Index Terms— aviation, accident, pattern, NTSB, SPSS, descriptive statistics, predictive modeling, python, Chi-Square, cross tabulations

I. Introduction

Aviation accidents involving airplane, helicopter, air balloon, jet plane etc. occurred from time to time. Aviation accidents are unpredictable which result the greater loss of life. Aviation accidents can be caused by several factors including human error, unpredictable bad weather, engine failure, aviation component failure, fires in the cargo, design flaws, bird strikes, etc. The paper focuses on identifying historical patterns on aviation accidents on specific decade as well as statistical analysis of several factors of aviation accidents and the number of injuries caused by it.

II. Data Collection Method

For this paper, data was downloaded from National Transportation Safety Board (NTSB) website. NTSB aviation accident database contains information from 1962 and later about civil aviation accidents and selected incidents within the United States, its territories and possessions, and in international waters. [6]

The total number of unique columns is 31 and database has 77975 rows.

An overview of the initial data is shown below:

	Name	Type	Width	Decimals	Label	Values	I
1	Event_Id	String	15	0	Unique Identification for Each Event	None	N
2	Investigation_Type	String	10	0	Type of Event	None	N
3	Accident_Number	String	13	0	NTSB Number	None	N
4	Event_Date	Date	10	0	Event Date	None	N
5	Location	String	39	0	Event Location Nearest City	None	N
6	Country	String	15	0	Event Country	None	N
7	Latitude	Numeric	12	6	Event Location Latitude	None	N
8	Longitude	Numeric	13	6	Event Location Longitude	None	N
9	Airport_Code	String	6	0	Event Location Nearest Airport ID	None	N
10	Airport_Name	String	32	0	Event Location Airport	None	N
11	Injury_Severity	String	13	0	Event Highest Injury	None	N
12	Aircraft_Damage	String	13	0	Damage	None	N
13	Aircraft_Category	String	14	0	Aircraft Category	None	N
14	Registration_Number	String	8	0	Aircraft Registration Number	None	N
15	Make	String	32	0	Aircraft Manufacturer's Full Name	None	N
16	Model	String	19	0	Aircraft Model	None	N
17	Amateur_Built	String	5	0	Aircraft is a homebuilt (Y/N).	None	N
18	Number_of_Engines	Numeric	3	0	Number of Engines	None	N
19	Engine_Type	String	15	0	Engine Type	None	N
20	FAR_Description	String	31	0	Federal Aviation Reg. Part	None	N
21	Schedule	String	6	0	an air carrier operation was scheduled or not	None	N
22	Purpose_of_Flight	String	25	0	Type of Flying (Per_Bus / Primary)	None	N
23	Air_Carrier	String	37	0	Operator Name& Operator Is Doing Business As	None	N
24	Total_Fatal_Injuries	Numeric	4	0	Injury Total Fatal	None	N
25	Total_Serious_Injuries	Numeric	3	0	Injury Total Serious	None	N
26	Total_Minor_Injuries	Numeric	3	0	Injury Total Minor	None	N
27	Total_Uninjured	Numeric	5	0	Non-Injury Total	None	N
28	Weather_Condition	String	5	0	Basic weather conditions	None	N
29	Broad_Phase_of_Flight	String	13	0	Phase of Flight	None	N
30	Report_Status	String	16	0	Latest Report Level	None	N
31	Publication_Date	Date	10	0	Publication data of the Latest Report Level	None	N

Figure 1. Variable view of data in SPSS software

III. Data Analysis Method

IBM SPSS (Statistical Package for the Social Science) has been used to conduct the data analysis on aviation accidents.

The IBM SPSS is capable to perform linear model, non-linear model, simulation capabilities, decision tree, descriptive analysis and so many others. Descriptive analysis, factor analysis, correlation analysis, cross tabulation analysis and cluster analysis are used during the process to conduct the data

analytics on National Transportation Safety Board (NTSB) accident database.

A. Data Overview

To analyze the data overall on aviation accidents, descriptive statistics has been used on the phase of flight, it can be concluded takeoff and landing are the phases of light most aviation accidents occur.

Phase of Flight				
	Frequency	Percent	Valid Percent	Cumulative Percent
Valid	6193	7.9	7.9	7.9
APPROACH	7591	9.7	9.7	17.7
CLIMB	2246	2.9	2.9	20.6
CRUISE	10641	13.6	13.6	34.2
DESCENT	2170	2.8	2.8	37.0
GO-AROUND	1580	2.0	2.0	39.0
LANDING	18752	24.0	24.0	63.1
MANEUVERING	9620	12.3	12.3	75.4
OTHER	151	.2	.2	75.6
STANDING	1184	1.5	1.5	77.1
TAKEOFF	14928	19.1	19.1	96.3
TAXI	2285	2.9	2.9	99.2
UNKNOWN	634	.8	.8	100.0
Total	77975	100.0	100.0	

Figure 2. Aviation accidents during certain Phase of Flight

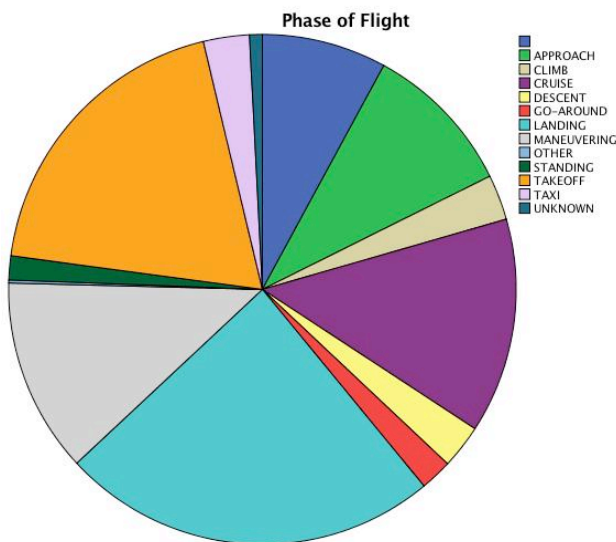


Figure 3. Graphical representation of aviation accidents during certain Phase of Flight

B. Accident pattern analysis method by Decade

To analyze the aviation accident on each decade and evaluate the pattern of accidents by decade, the new column named *Decade_Identifier* has been added for each accident entry. For instance, if the aviation accident was occurred on 02/16/2016, it will be marked as 2010. Similarly if the accident was occurred on 07/25/1987, it will be marked as 1980. Python script was written to identify the decade and season of the year based on the event date.

Sample Python code:

```
import os
f1 = open(os.path.expanduser("../AviationData.txt"))
f2 = open("../newAviationData.txt", 'w')
with f1:
    for line in f1:
        words = line.split(" | ")
        find_this1 = "/193"
```

32	Decade	Numeric	6	0	Time frame in Decade
33	Season_of_Year	String	9	0	Recoded
34	Season_of_Year_ID	Numeric	1	0	Recoded
35	Aircraft_Category_ID	Numeric	2	0	Recoded
36	Make_ID	Numeric	4	0	Recoded
37	Model_ID	Numeric	5	0	Recoded
38	Amateur_Built_ID	Numeric	1	0	Recoded
39	Engine_Type_ID	Numeric	2	0	Recoded
40	Purpose_of_Flight_ID	Numeric	2	0	Recoded
41	Weather_Condition_ID	Numeric	1	0	Recoded
42	Broad_Phase_of_Flight_ID	Numeric	2	0	Recoded
43	Airport_Name_ID	Numeric	5	0	Recoded
44	Aircraft_Damage_ID	Numeric	1	0	Recoded

Figure 4. Additional variables plus recoded variables added in Data Set

Sometimes it is important to change a continuous variable into a categorical variable and in SPSS, this type of transformation can be done through recoding. To perform the statistical analyse, some of the variables were recoded and new variables were created to put the recoded data into. For instance, Season of Year has been recoded into Season_of_Year_ID which holds 1 = "Winter" 2 = "Spring" 3 = "Summer", 4 = "Fall" and 5 = "Season_" (unknown.) Similarly, other variables like Aircraft_Category, Make, Model, Amateur_Built, Engine_Type, Purpose_of_Flight, Weather_Condition, Broad_Phase_of_Flight, Airport_Name, and Aircraft_Damage were recoded too.

After recoding, the factor analysis was performed to see the relationship among selected variables. Some of the independent variables were removed which didn't have impacts on aviation accidents after performing factor analysis to produce the factors of correlated independent variables. The variables can be seen in the structure matrix below.

Structure Matrix

	Component		
	1	2	3
Amateur_Built_ID	-.516	.271	.171
Weather_Condition_ID	-.240	-.051	.718
Broad_Phase_of_Flight_ID	.051	.035	.728
Total_Fatal_Injuries	.095	.670	-.226
Total_Serious_Injuries	.056	.741	-.009
Engine_Type_ID	.688	.132	.094
Decade	-.062	.376	.230
Number_of_Engines	.703	.099	-.166

Extraction Method: Principal Component Analysis.
Rotation Method: Oblimin with Kaiser Normalization.

Figure 5. Factor analysis

Finally, the Structure Matrix shows the factor loadings for each variable. Based on these factor loadings, the factors represent:

Variable	Factor
Amateur_Built_ID	Factor 1 - Make
Engine_Type_ID	Factor 1 - Make
Number_of_Engines	Factor 1 - Make
Total_Fatal_Injuries	Factor 2 - Injury
Total_Serious_Injuries	Factor 2 - Injury
Decade	Factor 2 - Injury
Weather_Condition_ID	Factor 3 -Weather
Broad_Phase_of_Flight_ID	Factor 3 - Weather

Figure 6. Factor analysis

C. Correlation analysis on factor variables

The paper uses the correlation analysis to closely observe the relationship among independent variables.

By looking at the chart below, it can be observed the variable Total Number of Fatal Injuries has a 0.265 significant to the variable Total Number of Serious Injuries.

GGraph

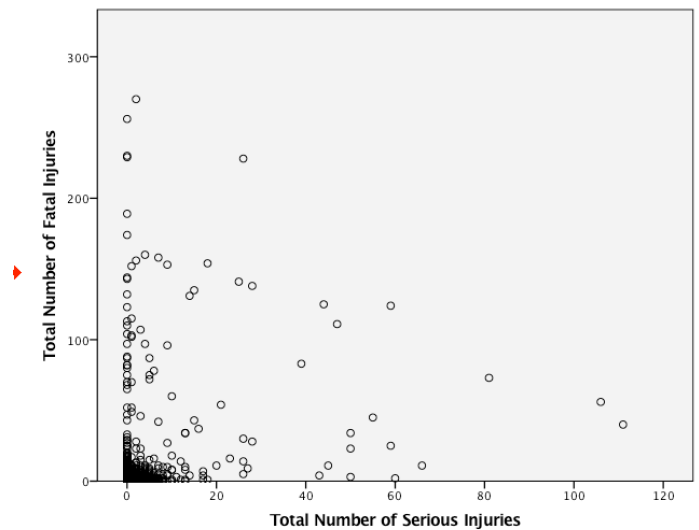


Figure 8. Correlation among total number of fatal injuries and total number of serious injuries in Scatter graph

Using Scatter point graph, the relationship between total number of fatal injuries and total number of serious injuries can be observed. Total number of fatal injuries and total number of serious injuries tend to be around same number when there are less than 30 injuries but tend to go far away when there are more injuries. Sometimes, there are high fatal injuries but low serious injuries. Similarly sometimes, there are high serious injuries and low fatal injuries.

Correlations

		Amateur_Built_ID	Engine_Type_ID	Number_of_Engines	Total Number of Fatal Injuries	Total Number of Serious Injuries	Time frame in Decade	Weather_Condition_ID	Broad_Phase_of_Flight_ID
Amateur_Built_ID	Pearson Correlation	1	-.095**	-.106**	-.010*	.009*	.079**	.071**	.019**
	Sig. (2-tailed)		.000	.000	.014	.033	.000	.000	.000
	N	77406	74490	73810	55501	53533	77405	75762	71640
Engine_Type_ID	Pearson Correlation	-.095**	1	.208**	.078**	.081**	.034**	-.063**	-.022**
	Sig. (2-tailed)	.000		.000	.000	.000	.000	.000	.000
	N	74490	74671	72969	54470	52914	74671	74040	70746
Number_of_Engines	Pearson Correlation	-.106**	.208**	1	.109**	.044**	-.015**	-.153**	-.023**
	Sig. (2-tailed)	.000	.000		.000	.000	.000	.000	.000
	N	73810	72969	74031	53381	51936	74031	73217	70353
Total Number of Fatal Injuries	Pearson Correlation	-.010*	.078**	.109**	1	.265**	.100**	-.116**	-.036**
	Sig. (2-tailed)	.014	.000	.000		.000	.000	.000	.000
	N	55501	54470	53381	55691	50532	55691	54882	52926
Total Number of Serious Injuries	Pearson Correlation	.009*	.081**	.044**	.265**	1	.165**	-.039**	-.015**
	Sig. (2-tailed)	.033	.000	.000	.000		.000	.000	.000
	N	53533	52914	51936	50532	53606	53606	53333	51933
Time frame in Decade	Pearson Correlation	.079**	.034**	-.015**	.100**	.165**	1	.043**	.049**
	Sig. (2-tailed)	.000	.000	.000	.000	.000		.000	.000
	N	77405	74671	74031	55691	53606	77974	75941	71782
Weather_Condition_ID	Pearson Correlation	.071**	-.063**	-.153**	-.116**	-.039**	.043**	1	.161**
	Sig. (2-tailed)	.000	.000	.000	.000	.000	.000		.000
	N	75762	74040	73217	54882	53333	75941	75941	71365
Broad_Phase_of_Flight_ID	Pearson Correlation	.019**	-.022**	-.023**	-.036**	-.015**	.049**	.161**	1
	Sig. (2-tailed)	.000	.000	.000	.000	.000	.000	.000	
	N	71640	70746	70353	52926	51933	71782	71365	71782

** . Correlation is significant at the 0.01 level (2-tailed).

* . Correlation is significant at the 0.05 level (2-tailed).

Figure 7. Correlation analysis on variables

IV. Data Analysis Method

Cruise flight is smoother and safer than 1980s, aircrafts industry saw most accidents while preparing for landing

Almost 64% of aviation accidents occurred during the landing, the approach, the maneuvering and the take off stage on 2010s. Most accidents took place during the departure (take off / climb) and arrival (approach / descent / maneuvering / landing) stages. Aircrafts industry saw most accidents while preparing for landing, and occurred during decent, approach and landing phase of flight. This accounts for 41% of aviation accidents.

What's interesting is there is not a noticeable change in the pattern of aviation accidents from 1980s to 2010s except for the aviation accidents in cruise mode.

In 1980s, 16.4% of aviation accidents were cruise flight related. In 2010s, cruise accidents only account for 2.7%. Cruise flight is smoother and safer than 1980s. Thanks to the technological advancement!

On the chart below for the cross tabulation between Phase of Flight and Decade, aviation accidents during each phase of flight on each decade can be observed clearly.

Figure 9. Crosstabulation between Phase of Flight and Decade

Broad_Phase_of_Flight * Time frame in Decade Crosstabulation

		Time frame in Decade				Total
		1980	1990	2000	2010	
Broad_Phase_of_Flight	Count	3	106	1103	1965	3177
	% within Time frame in Decade	0.1%	63.5%	17.9%	23.0%	17.6%
APPROACH	Count	356	6	492	680	1534
	% within Time frame in Decade	11.1%	3.6%	8.0%	8.0%	8.5%
CLIMB	Count	91	2	101	150	344
	% within Time frame in Decade	2.8%	1.2%	1.6%	1.8%	1.9%
CRUISE	Count	526	7	366	230	1129
	% within Time frame in Decade	16.4%	4.2%	5.9%	2.7%	6.2%
DESCENT	Count	91	2	146	179	418
	% within Time frame in Decade	2.8%	1.2%	2.4%	2.1%	2.3%
GO-AROUND	Count	56	2	152	160	370
	% within Time frame in Decade	1.7%	1.2%	2.5%	1.9%	2.0%
LANDING	Count	865	10	2029	2360	5264
	% within Time frame in Decade	27.0%	6.0%	32.9%	27.6%	29.1%
MANEUVERING	Count	328	10	401	735	1474
	% within Time frame in Decade	10.2%	6.0%	6.5%	8.6%	8.1%
OTHER	Count	11	0	9	16	36
	% within Time frame in Decade	0.3%	0.0%	0.1%	0.2%	0.2%
STANDING	Count	28	0	77	148	253
	% within Time frame in Decade	0.9%	0.0%	1.2%	1.7%	1.4%
TAKEOFF	Count	681	18	1085	1636	3420
	% within Time frame in Decade	21.2%	10.8%	17.6%	19.2%	18.9%
TAXI	Count	126	3	187	225	541
	% within Time frame in Decade	3.9%	1.8%	3.0%	2.6%	3.0%
UNKNOWN	Count	47	1	23	59	130
	% within Time frame in Decade	1.5%	0.6%	0.4%	0.7%	0.7%

Non-scheduled flights account for 76 percent of aviation accidents where total number of fatal injuries is more than or equal to one

Almost 76% of aviation accidents were reported on non-scheduled flights where total number of fatal injuries was more than or equal to one. Around 24% of aviation accidents were reported on scheduled flights where total number of fatal injuries was more than or equal to one.

From the analyzed dataset, only around 44% of aviation accidents were reported on non-scheduled flights where total number of fatal injuries was more than or equal to one. More than half of the accidents i.e. 56% were reported on scheduled flights where total number of fatal injuries was more than or equal to one.

The percentage of aviation accidents were not found similar on scheduled and non-schedule flights where total number of fatal injuries or total number of serious injuries was more than or equal to one, which was not expected. The paper needs to analyze more on this.

Schedule	Not Scheduled	Count	977
		% within Total Number of Fatal Injuries	76.1%
	Scheduled	Count	307
		% within Total Number of Fatal Injuries	23.9%
Total		Count	1284
		% within Total Number of Fatal Injuries	100.0%

Schedule	Not Scheduled	Count	546
		% within Total Number of Serious Injuries	43.6%
	Scheduled	Count	706
		% within Total Number of Serious Injuries	56.4%
Total		Count	1252
		% within Total Number of Serious Injuries	100.0%

Figure 10. Crosstabulation of scheduled flight and non-scheduled flight on total number of serious injuries and total number of fatal injuries

The statistical result shows the pattern of accidents for scheduled and non-scheduled flights on each decade is consistent. Each decade, non-scheduled flights account for more than 70 percent of accidents.

Schedule * Time frame in Decade Crosstabulation							
Schedule			Time frame in Decade				Total
			1980	1990	2000	2010	
NSCH	Count		252	264	333	128	977
	% within Time frame in Decade		78.3%	71.5%	76.9%	80.0%	76.1%
	Count		70	105	100	32	307
	% within Time frame in Decade		21.7%	28.5%	23.1%	20.0%	23.9%
Total	Count		322	369	433	160	1284
	% within Time frame in Decade		100.0%	100.0%	100.0%	100.0%	100.0%

Figure 11. Crosstabulation of scheduled flight and non-scheduled flight on each decade

How weather and amateur built affect aviation accident?

Definitions taken from the website <http://aviation.about.com>, “VMC (Visual Flight Rules): The definition of VMC, according to the FAA, is: meteorological conditions expressed in terms of visibility, distance from cloud, and ceiling equal to or better than specified minima.

IMC (Instrument Flight Rules): The definition of IMC, according to the FAA, is: meteorological conditions expressed in terms of visibility, distance from cloud, and ceiling less than the minima specified for visual meteorological conditions.” [8]

Weather_Condition_ID * Time frame in Decade * Amateur_Built_ID Crosstabulation											
Amateur_Built_ID				Time frame in Decade							Total
				1940	1960	1970	1980	1990	2000	2010	
No	Weather_Condition_ID	IMC	Count	0	0	2	2034	1726	1268	440	5470
			% within Time frame in Decade	0.0%	0.0%	66.7%	8.8%	8.3%	7.5%	5.7%	8.0%
		UNK	Count	1	1	0	347	387	55	74	865
		% within Time frame in Decade	100.0%	100.0%	0.0%	1.5%	1.9%	0.3%	1.0%	1.3%	
	VMC	Count	0	0	1	20668	18687	15475	7196	62027	
		% within Time frame in Decade	0.0%	0.0%	33.3%	89.7%	89.8%	92.1%	93.3%	90.7%	
	Total	Count	1	1	3	23049	20800	16798	7710	68362	
		% within Time frame in Decade	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	
	Yes	Weather_Condition_ID	IMC	Count				26	35	52	18
			% within Time frame in Decade				1.6%	1.6%	2.2%	1.5%	1.8%
UNK			Count				18	45	3	3	69
		% within Time frame in Decade				1.1%	2.1%	0.1%	0.2%	0.9%	
VMC		Count				1589	2080	2311	1220	7200	
		% within Time frame in Decade				97.3%	96.3%	97.7%	98.3%	97.3%	
Total		Count				1633	2160	2366	1241	7400	
		% within Time frame in Decade				100.0%	100.0%	100.0%	100.0%	100.0%	
Total		Weather_Condition_ID	IMC	Count	0	0	2	2060	1761	1320	458
			% within Time frame in Decade	0.0%	0.0%	66.7%	8.3%	7.7%	6.9%	5.1%	7.4%
	UNK		Count	1	1	0	365	432	58	77	934
		% within Time frame in Decade	100.0%	100.0%	0.0%	1.5%	1.9%	0.3%	0.9%	1.2%	
	VMC	Count	0	0	1	22257	20767	17786	8416	69227	
		% within Time frame in Decade	0.0%	0.0%	33.3%	90.2%	90.4%	92.8%	94.0%	91.4%	
	Total	Count	1	1	3	24682	22960	19164	8951	75762	
		% within Time frame in Decade	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	

Figure 12. Crosstabulation of scheduled flight and non-scheduled flight on each decade depending on airplane built amateur or not

By looking at the chart above, approximately 7.4% of general aviation accidents were weather related. When airplanes were amateur built, aviation accidents related to weather dropped down to 1.8%. When airplanes were not amateur built, aviation accidents related to weather jumped to 8%. From 1980s to 2010s, weather related aviation accidents had been dropped down from 8.3% to 5.1%. This percentage dropped is 8.8 to 5.7 when airplanes were not amateur built.

In 2010s, 5.1% of general aviation accidents were weather related. In 2010s, when airplanes were amateur built, these general aviation accidents related to weather dropped down to 1.5%. When airplanes were not amateur built, general aviation accidents related to weather jumped to 5.7%.

Weather related flights account for 22% of fatal injuries where total number of fatal injuries is more than or equal to one

Total Number of Fatal Injuries * Weather_Condition_ID Crosstabulation

		Weather_Condition_ID			Total
		IMC	UNK	VMC	
Total	Count	3262	480	11050	14792
	% within Weather_Condition_ID	100.0%	100.0%	100.0%	100.0%

Chi-Square Tests

	Value	df	Asymptotic Significance (2-sided)	Monte Carlo Sig. (2-sided)		Monte Carlo Sig. (1-sided)	
				Significance	99% Confidence Interval	Significance	99% Confidence Interval
				Lower Bound	Upper Bound	Lower Bound	Upper Bound
Pearson Chi-Square	877.289 ^a	188	.000	.000 ^b	.000	.000	
Likelihood Ratio	599.197	188	.000	.000 ^b	.000	.000	
Fisher's Exact Test	710.990		.000	.000 ^b	.000	.000	
Linear-by-Linear Association	67.468 ^c	1	.000	.000 ^b	.000	.000	.000
N of Valid Cases	14792						

a. 255 cells (89.5%) have expected count less than 5. The minimum expected count is .03.
b. Based on 10000 sampled tables with starting seed 166864539.
c. The standardized statistic is -8.214.

Figure 13. Crosstabulation of weather condition and total number of fatal injuries

The fatal injuries rate is high for the weather related aviation accidents. These results indicate that there is statistically significant relationship between the total number of

fatal injuries and weather condition (chi-square = 877.829). Out of 14792 fatal injuries, 3262 fatal injuries were due to weather related aviation accidents. That means 22% of total fatal injuries were related with weather aviation accidents. Weather related aviation accidents much less likely to happen i.e. 5.1% in 2010s but it accounted for around 22% of total fatal injuries in 2010s.

Almost 83% of aviation accidents occurred in single-engine piston aircraft

Aircraft engines are almost always either lightweight piston engines or gas turbines. Turbo jet engine, turbo prop engine, turbo fan engine and turbo shaft engine are variants of turbine engine.

A reciprocating engine, also known as a piston engine, is a heat that uses one or more reciprocating pistons to convert pressure into a rotating motion. [7]

Recorded * Number_of_Engines Crosstabulation							
			Number_of_Engines				Total
			1	2	3	4	
Recorded	Electric	Count	2	0	0	0	2
		% of Total	0.0%	0.0%	0.0%	0.0%	0.0%
REC, TJ, REC,	REC, TJ, REC,	Count	0	0	0	1	1
		% of Total	0.0%	0.0%	0.0%	0.0%	0.0%
REC, TJ, TJ	REC, TJ, TJ	Count	0	0	0	2	2
		% of Total	0.0%	0.0%	0.0%	0.0%	0.0%
Reciprocating	Reciprocating	Count	13696	1195	0	7	14898
		% of Total	82.8%	7.2%	0.0%	0.0%	90.1%
TF, TJ	TF, TJ	Count	0	4	0	0	4
		% of Total	0.0%	0.0%	0.0%	0.0%	0.0%
TJ, REC, REC,	TJ, REC, REC,	Count	0	0	0	1	1
		% of Total	0.0%	0.0%	0.0%	0.0%	0.0%
Turbo Fan	Turbo Fan	Count	10	520	38	39	607
		% of Total	0.1%	3.1%	0.2%	0.2%	3.7%
Turbo Jet	Turbo Jet	Count	21	81	2	10	114
		% of Total	0.1%	0.5%	0.0%	0.1%	0.7%
Turbo Prop	Turbo Prop	Count	526	367	0	3	896
		% of Total	3.2%	2.2%	0.0%	0.0%	5.4%
Turbo Shaft	Turbo Shaft	Count	7	2	0	0	9
		% of Total	0.0%	0.0%	0.0%	0.0%	0.1%
Total	Total	Count	14262	2169	40	63	16534
		% of Total	86.3%	13.1%	0.2%	0.4%	100.0%

Figure 13. Crosstabulation of aircraft engine with number of engines

Turbo prop and turbo fan engine aircrafts had accident rate of 5.4% and 3.7% respectively. Reciprocating engine (piston engine) accounted for 90.1% of the accidents. Popular commercial planes Boeing 737 and Airbus A320 both use turbo fan engines. The paper cannot conclude necessarily what kind of turbo engines is safer from the analyzed data. One of the reason is the current dataset doesn't have information about aircraft engines and its used in the aviation market. Current dataset also doesn't know if failures of aircraft engines were the cause for some of these accidents.

From the analyzed dataset, it can be concluded that most of the accidents occurred on single-engine piston aircrafts i.e. 82.8% and double-engine piston aircrafts i.e. 7.2%. It could be the reason that there are a lot of piston engine based aircrafts in the aviation world. Lots of small planes use piston engine for fuel efficiency.

IV. Future Work

To analyze aircraft engines on engine failure aviation accidents

Various aircraft engines are used in various aircrafts depending on the size of the plane, fuel efficiency and

maintenance cost. This research can be elaborated to further analyze the role of aircraft engines on engine failure aviation accidents. Collecting separate data of engine-failure aircraft accidents as well as the list of planes with aircraft engines information, and integrating with the current dataset can help achieving this analysis.

V. Conclusion

The trend and pattern of aviation accidents on phase of flight varies significantly by decade. Other factors didn't vary significantly by decade. The statistical result shows that cruise flight is smoother and safer than 1980s; the credit goes to advancement in the aviation technology and development of high-tech cockpit equipment in modern aircrafts. Non-scheduled flights are responsible for 76 percent of aviation accidents where total number of fatal injuries is more than or equal to 1. Approximately 7.4% of general aviation accidents were weather related yet weather related flights account for 22% of fatal injuries where total number of fatal injuries is more than or equal to one. Almost 83% of aviation accidents occurred in single-engine piston aircraft.

The statistical result shows various aviation accident patterns and affects of independent variables on accidents.

References

- [1] "1001 Crash - Worldwide aviation safety", *1001crash.com*, 2016. [Online]. Available: <http://www.1001crash.com/index-page-accueil-1g-2.html>. [Accessed: 24- Apr- 2016].
- [2] "Accident statistics", *Planecrashinfo.com*, 2016. [Online]. Available: <http://www.planecrashinfo.com/cause.htm>. [Accessed: 18- Apr- 2016].
- [3] "Aircraft engine", *Wikipedia*, 2016. [Online]. Available: https://en.wikipedia.org/wiki/Aircraft_engine. [Accessed: 24- Apr- 2016].

- [4] "Home - SPSS Tutorials - LibGuides at Kent State University", *Libguides.library.kent.edu*, 2016. [Online]. Available: <http://libguides.library.kent.edu/SPSS>. [Accessed: 24- Apr- 2016].
- [5] J. Hui bin and W. Lei. Aviation Accident Causing-Analysis and Prevention Strategy on Organizational Level. *Emergency Management and Management Sciences (ICEMMS)*, 2010.
- [6] "NTSB Aviation Accident Database & Synopses", *Ntsb.gov*, 2016. [Online]. Available: http://www.nts.gov/_layouts/ntsb.aviation/index.aspx. [Accessed: 18- Apr- 2016].
- [7] "Reciprocating engine", *Wikipedia*, 2016. [Online]. Available: https://en.wikipedia.org/wiki/Reciprocating_engine. [Accessed: 26- Apr- 2016].
- [8] S. Houston and S. Houston, "What Do VMC and IMC Mean?", *About.com Money*, 2016. [Online]. Available: <http://aviation.about.com/od/Regulations/fl/What-Do-VMC-and-IMC-Mean.htm>. [Accessed: 24- Apr- 2016].

Pradip Karki was born in Nepal. He moved to U.S. in 2008. He received his bachelor's degree in computer science from College of Staten Island - CUNY, New York City in 2012. He is currently enrolled in graduate program pursuing Master of Information Science in Washington University at St. Louis. Pradip chose his first professional job as a Software Engineer at the startup company, Better Mobile Security in 2012. In 2013, he joined the Engineering, Operations & Technology (EO&T) department at The Boeing Company, as a Programmer Analyst. Since then, he is with The Boeing Company. His current research interests include big data, machine learning, cryptography, and cyber security.