

Flight Risk & Safety Analysis



Analyzing Aircraft Accident Trends with Data

This project explores over 100 years of global aircraft accidents using Python, Excel, and Power BI to uncover fatality patterns, crash trends, and aviation safety insights.

Start Presentation



BY: SUGANDHA SAINI

Date

A horizontal timeline visualization. The word "Year" is at the top left. Two rectangular boxes are positioned side-by-side. The left box contains the year "1908". The right box contains the year "2009". A thick white horizontal line connects the two boxes, ending in circular markers at each end.

Flight Risk: Aircraft Accidents & Safety Analysis

105K

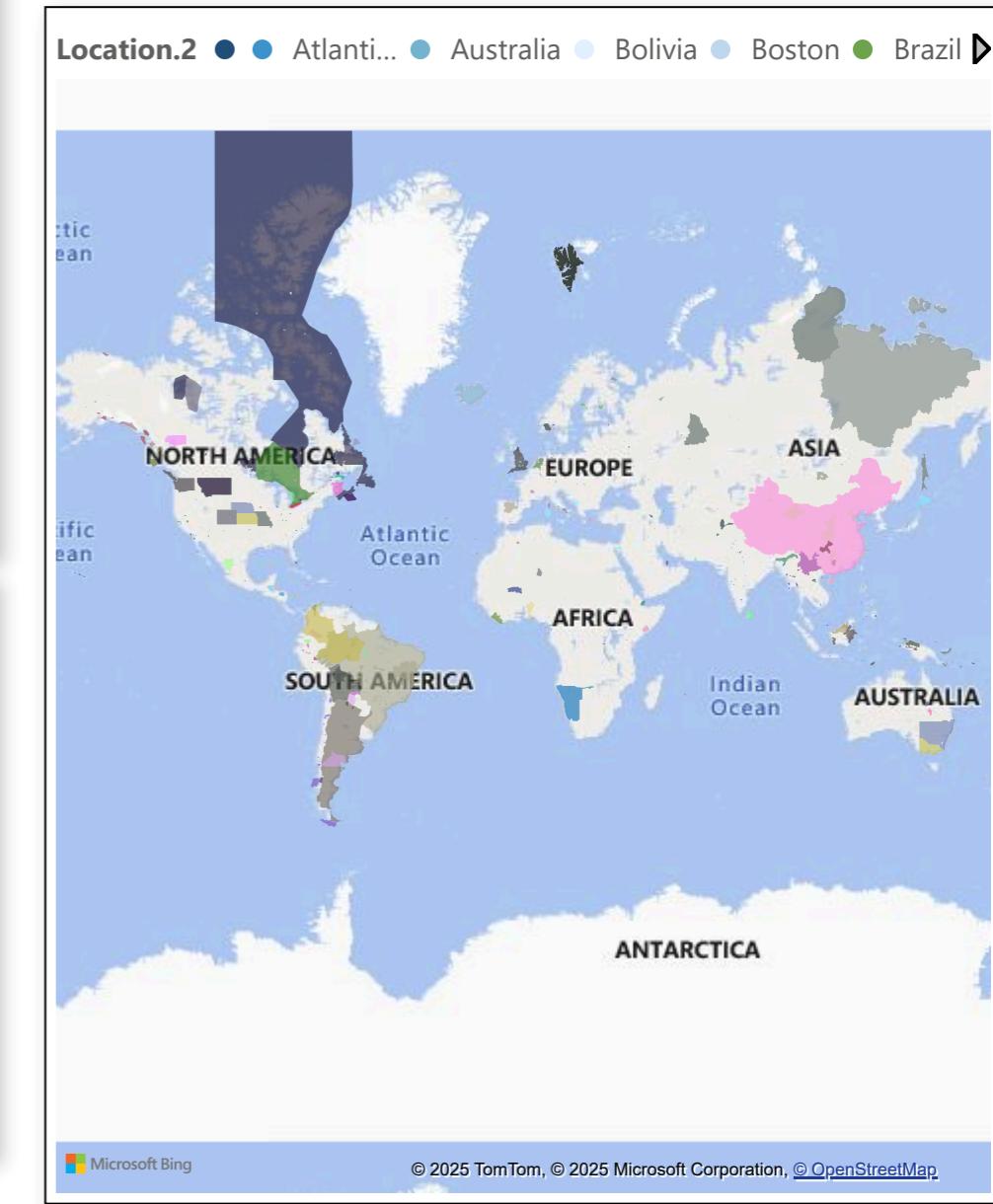
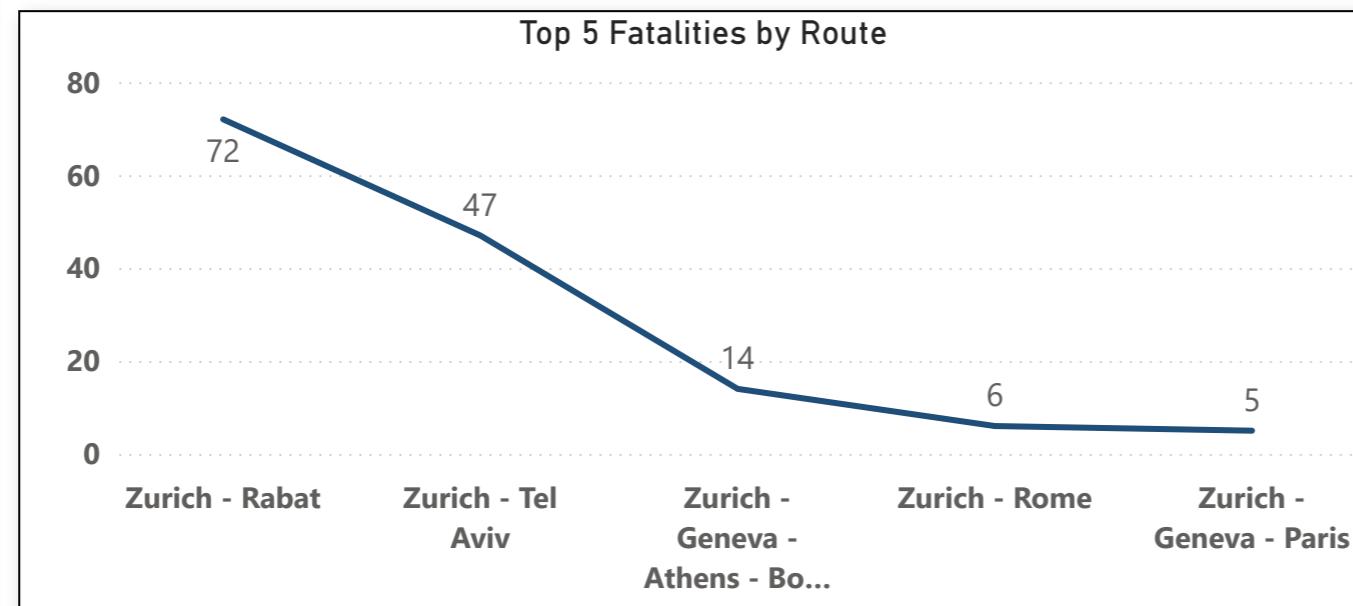
Sum of Fatalities 

144K

8440
Sum of Ground 

73.0%
Fatality Rate

39K
Survivors 



Date

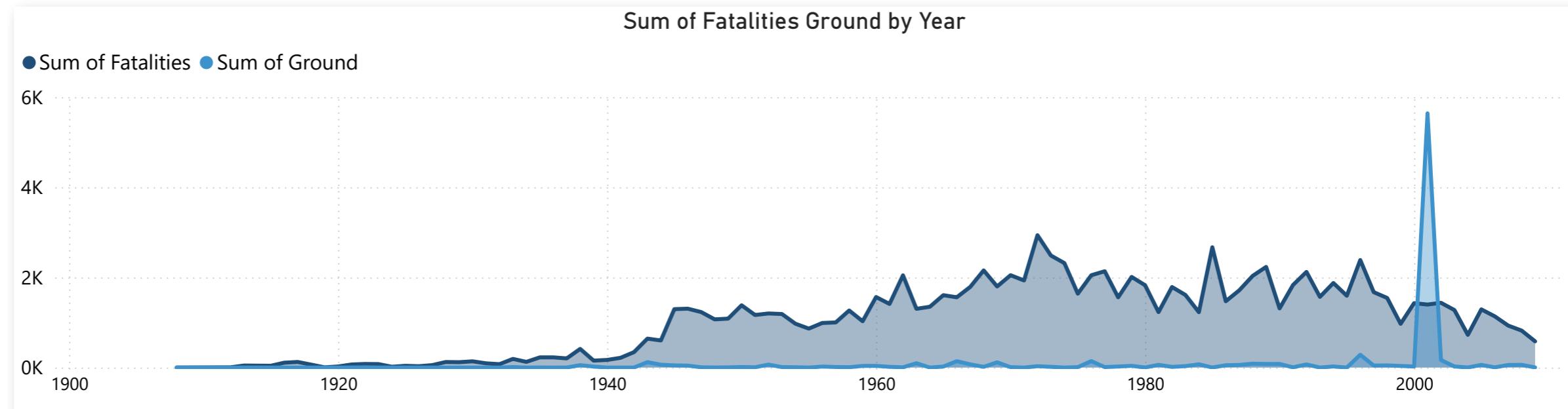
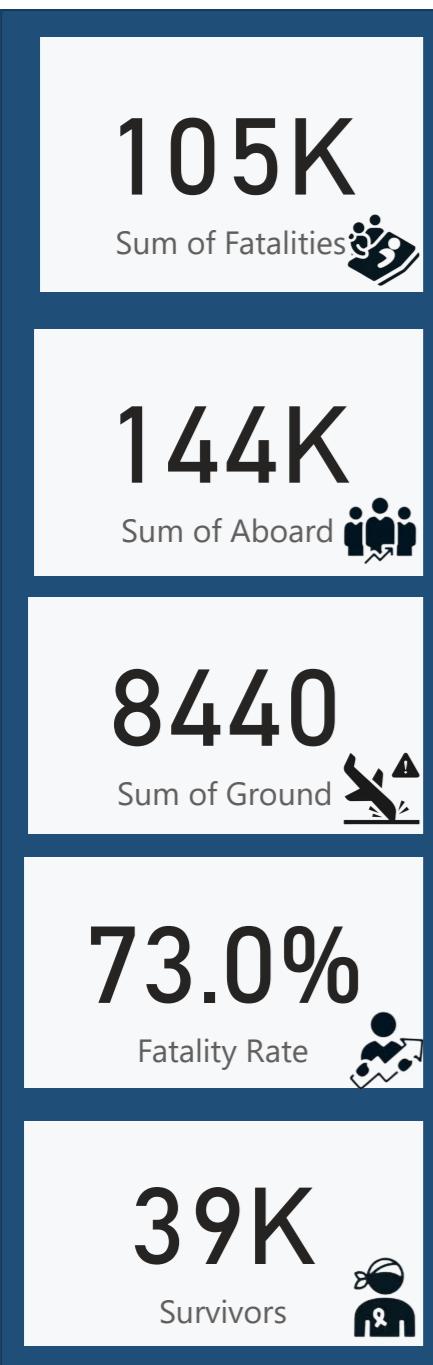
Year

Flight Risk: Aircraft Accidents & Safety Analysis



Operator

Crash Hour



Date	Sum of Aboard	Sum of Ground	Fatality Rate	Route	Location	Time	Survivors
17-09-1908	2	0	50.0%	Demonstration	Fort Myer, Virginia	17:18:00	1
12-07-1912	5	0	100.0%	Test flight	Atlantic City, New Jersey	06:30:00	0
06-08-1913	1	0	100.0%	unknown	Victoria, British Columbia, Canada	00:00:00	0
09-09-1913	20	0	70.0%	unknown	Over the North Sea	18:30:00	6
17-10-1913	30	0	100.0%	unknown	Near Johannisthal, Germany	10:30:00	0
05-03-1915	41	0	51.2%	unknown	Tienen, Belgium	01:00:00	20
03-09-1915	19	0	100.0%	unknown	Off Cuxhaven, Germany	15:20:00	0
Total	144304	8440	73.0%				39016

```
[2]: import pandas as pd  
import numpy as np
```

```
[15]: df=pd.read_csv("Airplane_Crashes_and_Fatalities_Since_1908.csv")  
print(df.head())  
print(df.columns)
```

```
index      Date    Time          Location \
0     0  09/17/1908  17:18  Fort Myer, Virginia
1     1  07/12/1912  06:30  Atlantic City, New Jersey
2     2  08/06/1913    NaN  Victoria, British Columbia, Canada
3     3  09/09/1913  18:30        Over the North Sea
4     4  10/17/1913  10:30  Near Johannisthal, Germany
```

```
Operator Flight #          Route          Type \
0  Military - U.S. Army    NaN  Demonstration  Wright Flyer III
1  Military - U.S. Navy    NaN  Test flight    Dirigible
2       Private           -    NaN  Curtiss seaplane
3  Military - German Navy  NaN    NaN  Zeppelin L-1 (airship)
4  Military - German Navy  NaN    NaN  Zeppelin L-2 (airship)
```

```
Registration cn/In  Aboard  Fatalities  Ground \
0         NaN     1     2.0      1.0     0.0
1         NaN   NaN     5.0      5.0     0.0
2         NaN   NaN     1.0      1.0     0.0
3         NaN   NaN    20.0     14.0     0.0
4         NaN   NaN    30.0     30.0     0.0
```

Summary

- 0 During a demonstration flight, a U.S. Army fly...
- 1 First U.S. dirigible Akron exploded just offsh...
- 2 The first fatal airplane accident in Canada oc...
- 3 The airship flew into a thunderstorm and encou...

jupyter FlightRiskAnalysis Last Checkpoint: 4 days ago



File Edit View Run Kernel Settings Help

Trusted

File + X □ ▶ ■ C ▶ Code ▾ JupyterLab ▾ Python 3 (ipykernel) ○ ≡ ≡

```
3 Military - German Navy      NaN      NaN Zeppelin L-1 (airship)
4 Military - German Navy      NaN      NaN Zeppelin L-2 (airship)
```

```
Registration cn/In Aboard Fatalities Ground \
0      NaN    1    2.0      1.0    0.0
1      NaN    NaN   5.0      5.0    0.0
2      NaN    NaN   1.0      1.0    0.0
3      NaN    NaN  20.0     14.0    0.0
4      NaN    NaN  30.0     30.0    0.0
```

Summary

```
0 During a demonstration flight, a U.S. Army fly...
1 First U.S. dirigible Akron exploded just offsh...
2 The first fatal airplane accident in Canada oc...
3 The airship flew into a thunderstorm and encou...
4 Hydrogen gas which was being vented was sucked...
```

```
Index(['index', 'Date', 'Time', 'Location', 'Operator', 'Flight #', 'Route',
       'Type', 'Registration', 'cn/In', 'Aboard', 'Fatalities', 'Ground',
       'Summary'],
      dtype='object')
```

```
[8]: df['Date']=pd.to_datetime(df['Date'],errors='coerce')
df=df.dropna(subset=['Date','Location','Fatalities'])
```

3

```
[10]: df.loc[:, 'Operator'] = df['Operator'].fillna('Unknown')
df.loc[:, 'Type'] = df['Type'].fillna('Unknown')
```

```
[12]: df.loc[:, 'Year'] = df['Date'].dt.year
```

```
[14]: fatalities_np = df['Fatalities'].to_numpy()
total_fatalities = np.sum(fatalities_np)
avg_fatalities = np.mean(fatalities_np)
```

```
fatality_std = np.std(fatalities_np)
high_fatal = df[df['Fatalities'] > 100]
```