

AVIATION ACCIDENT ANALYSIS

This project is meant to explore, analyse and visualize aviation accidents and related factors such as reasons, survival rates, fatalities, locations etc. Detailed analyses of following factors:

- The planes crashed per year.
- People aboard per year during crashes.
- People dead per year during crashes.
- People survived per year during crashes.

Dataset - Airplane Crashes and Fatalities since 1908

:This dataset includes:

- All civil and commercial aviation accidents of scheduled and non-scheduled passenger airliners worldwide, which resulted in a fatality (including all U.S. Part 121 and Part 135 fatal accidents).
- All cargo, positioning, ferry and test flight fatal accidents.
- All military transport accidents with 10 or more fatalities.
- All commercial and military helicopter accidents with greater than 10 fatalities.
- All civil and military airship accidents involving fatalities.
- Aviation accidents involving the death of famous people.
- Aviation accidents or incidents of noteworthy interest.

:Attribute Information:

- Date: Date of accident, in the format - January 01, 2001
- Time: Local time, in 24 hr. format unless otherwise specified
- Location: Location of the accident
- Operator: Airline or operator of the aircraft
- Flight: #Flight number assigned by the aircraft operator
- Route: Complete or partial route flown prior to the accident
- Type: Aircraft type
- Registration: ICAO registration of the aircraft
- cn/ln: Construction or serial number / Line or fuselage number

- Aboard: Total aboard (passengers / crew)
- Fatalities: Total fatalities aboard (passengers / crew)
- Ground: Total killed on the ground
- Summary: Brief description of the accident and cause if known

:Missing Attribute Values:

| | |
|---------------|------|
| -Date | 0 |
| -Time | 2219 |
| -Location | 20 |
| -Operator | 18 |
| -Flight # | 4199 |
| -Route | 1706 |
| -Type | 27 |
| -Registration | 335 |
| -cn/ln | 1228 |
| -Aboard | 22 |
| -Fatalities | 12 |
| -Ground | 22 |
| -Summary | 390 |

:Creator: Sauro Grandi

:Date: September, 2016

STEPS INVOLVED

1. Importing the Libraries
2. Importing the Dataset
3. Feature engineering
4. EDA (Count of accidents by month, weekday, hour)
5. EDA (Total Fatalities)
6. Importing the second Dataset

7. EDA (Fatalities vs Year)

8. EDA (Operators)

Libraries used

Pandas

Matplotlib

Numpy

Seaborn

Datetime