Crashed Airplanes

Alanoud Almutairi, Zahra Sameer and Alnirah Alqahtani

Abstract:

In the fourth project of SDAIA T5 Bootcamp we decided to analysis several airplane crashes over the last centuries, the world has seen several airplane crashes both in the sky and on the ground. The development of technology aims to decrease technological difficulties and human errors. However, fatalities and aircraft damages caused by crashes do not cease to exist.

Our results found in this project will benefit the ongoing investigations into this important topic. Understanding what factors cause airplane crashes helps aviation industries make continuous improvement in flight safety.

Design:

Exploratory Data Analysis (EDA) to determine the common cause/reason of airplane crash, countries with maximum/minimum airplane crashes, fatalities vs survived ratio and any other interesting trend, finding most dangerous location airplane crashed in it, select best and worst type of airplanes depending on number of survivor and fatalities. In addition, to develop a Model determining the following;

- NMF
- LDA

Data:

- This dataset presents number of crashed airplanes since 1908, where we have 5268 record, Also we have 13 features in this dataset.

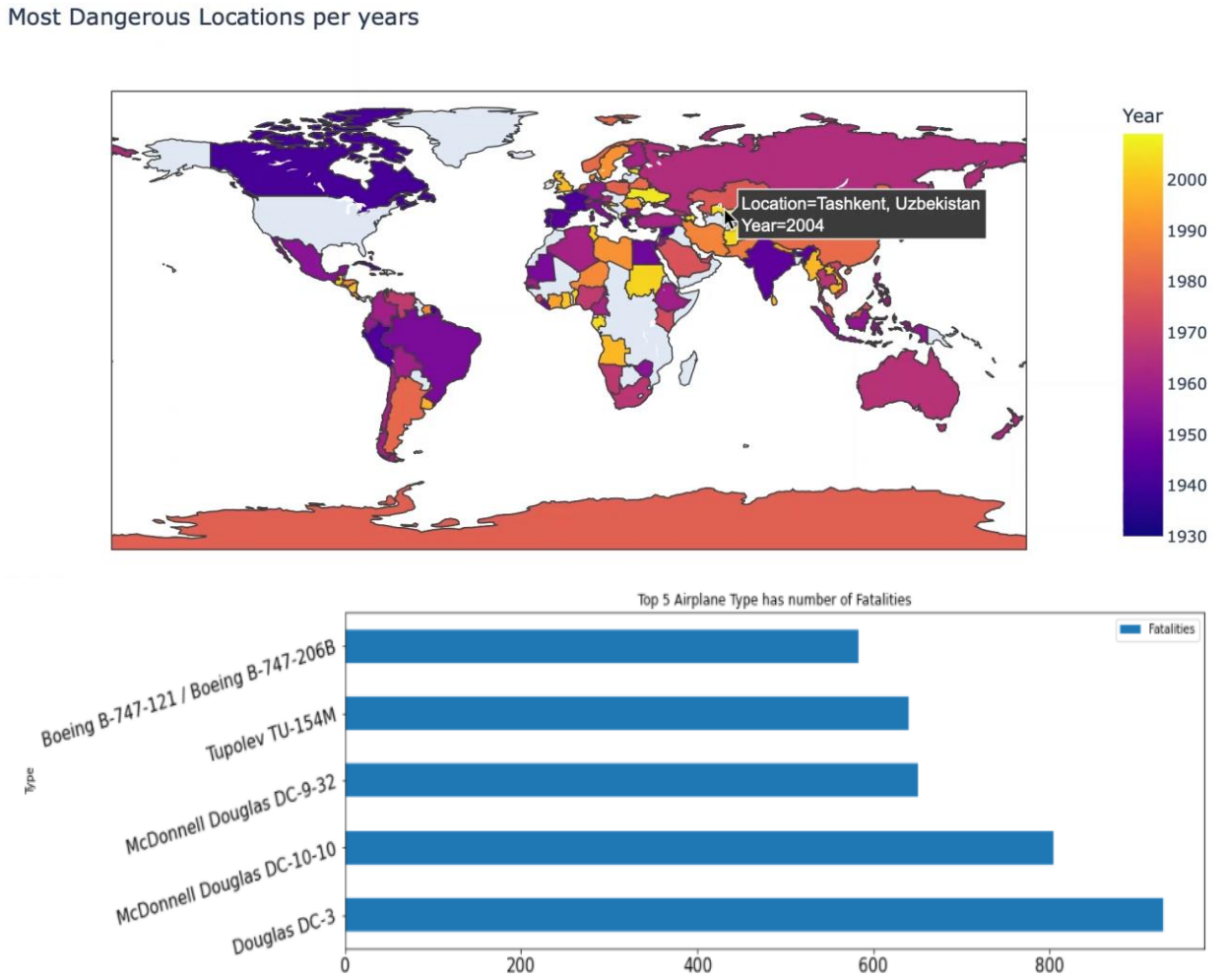
Limitations:

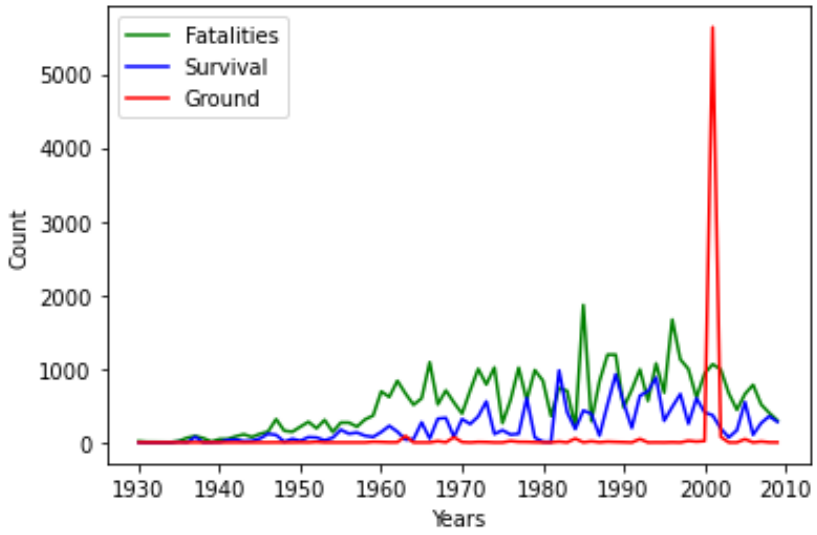
- Although dataset have a large amount of data, they have quite several unrecorded values, therefore some columns had to be either removed or filled with correct values.

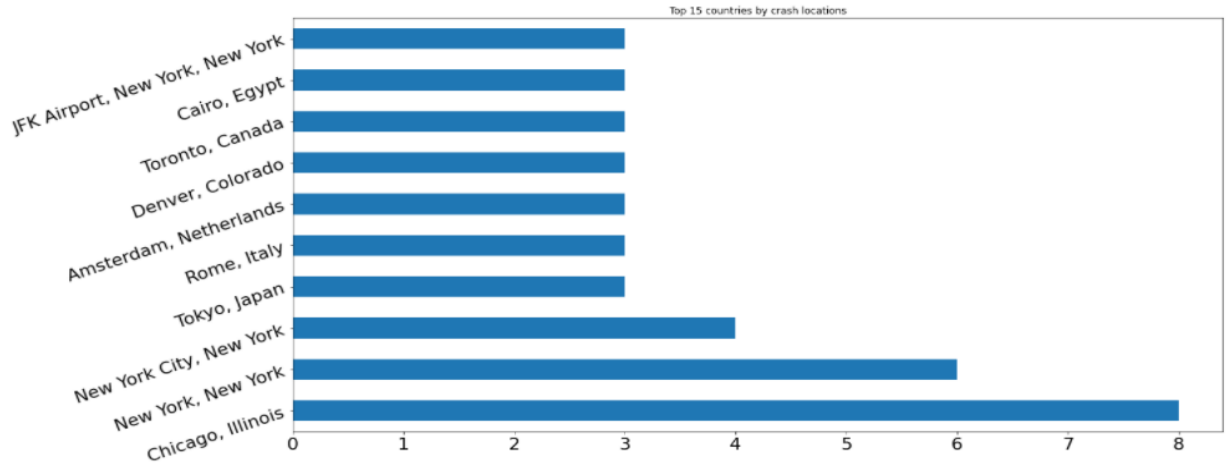
Tools:

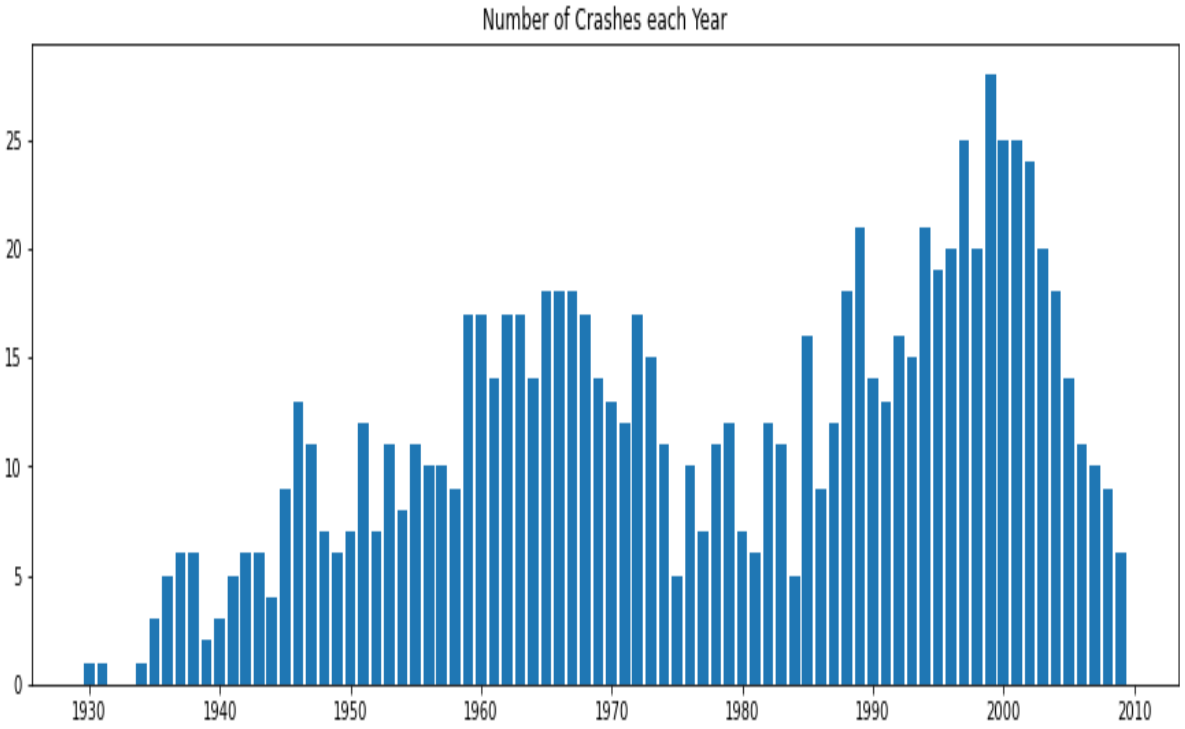
- Technologies : Python, Jupyter Notebook
- Libraires : Pandas, Numpy , Seaborn , Sklearn, Matplot ,WordCloud and number of python libraries.

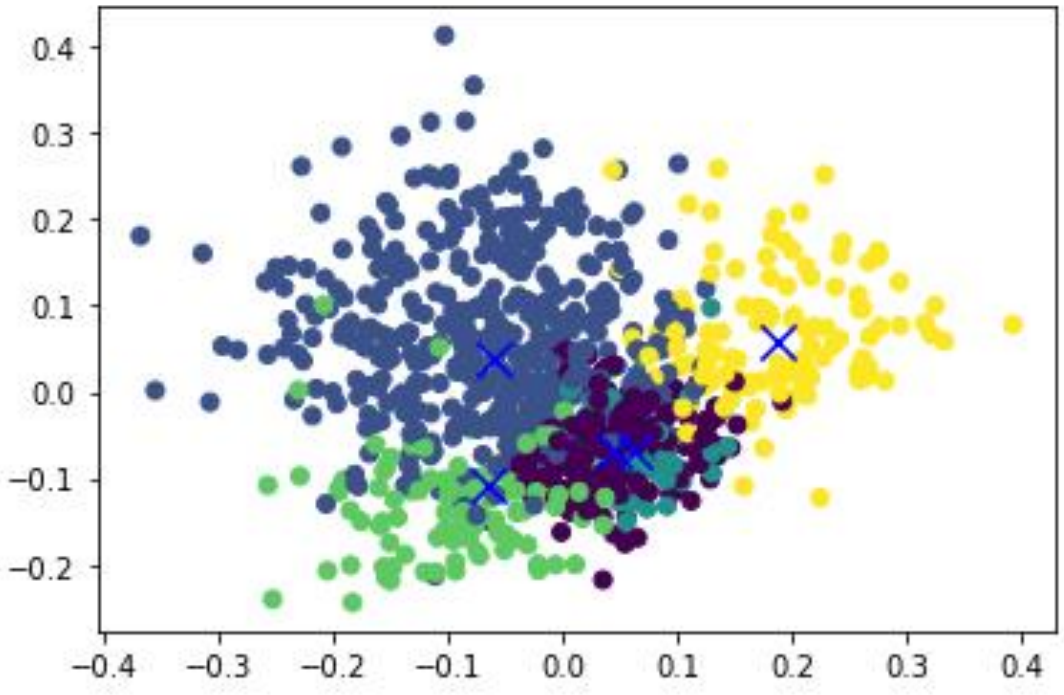
Communication:

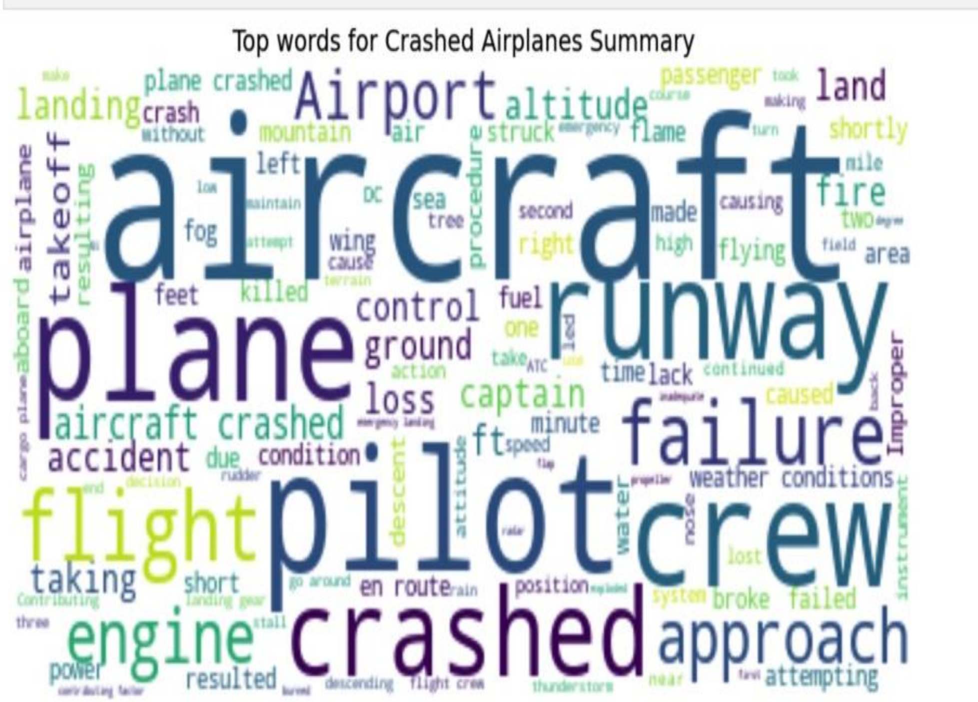


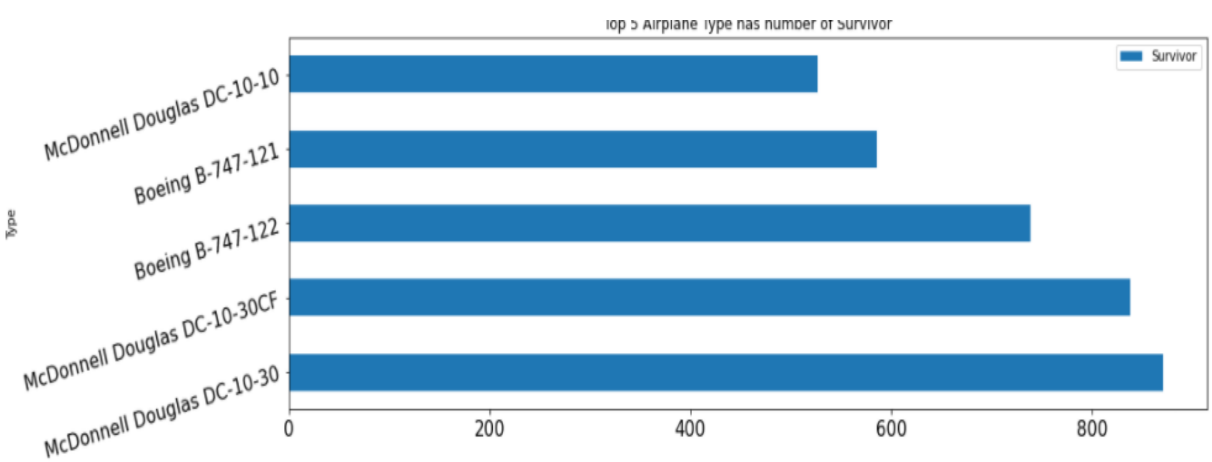

Fatalities
Survival
Ground
Count
Years
0
1000
2000
3000
4000
5000
1930
1940
1950
1960
1970
1980
1990
2000
2010


Top 15 countries by crash locations
JFK Airport, New York, New York
Cairo, Egypt
Toronto, Canada
Denver, Colorado
Amsterdam, Netherlands
Rome, Italy
Tokyo, Japan
New York City, New York
New York, New York
Chicago, Illinois
0
1
2
3
4
5
6
7
8


Number of Crashes each Year
0
5
10
15
20
25
1930
1940
1950
1960
1970
1980
1990
2000
2010


0.4
0.3
0.2
0.1
0.0
−0.1
−0.2
−0.4
−0.3
−0.2
−0.1
0.0
0.1
0.2
0.3
0.4


Top words for Crashed Airplanes Summary
aircraft
plane
runway
pilot
crew
crashed
flight
failure
approach
engine
Airport
aircraft crashed
accident
takeoff
landing
altitude
ground
control
captain
loss
land
fire
plane crashed
crash
mountain
airplane
aboard
resulting
Improper
weather conditions
instrument
fuel
left
fog
feet
killed
taking
short
resulted
power
en route
position
broke
failed
attempting
descent
attitude
water
nose
lost
flying
area
shortly
passenger
flame
struck
air
procedure
sea
tree
wing
cause
second
right
made
high
causing
mile
two
one
take
time
lack
continued
caused
minute
speed
ft
due
condition
field
system
near
thunderstorm
flight crew
descending
stall
three


Top 5 Airplane Type has number of Survivor
Survivor
Type
McDonnell Douglas DC-10-10
Boeing B-747-121
Boeing B-747-122
McDonnell Douglas DC-10-30CF
McDonnell Douglas DC-10-30
0
200
400
600
800

NMF- Model:

	pilot error	On Fire	shot down	poor weather	crash	dominant_topic
Doc0	0.04	0.01	0.00	0.02	0.11	4
Doc1	0.04	0.01	0.08	0.10	0.01	3
Doc2	0.00	0.01	0.01	0.18	0.00	3
Doc3	0.00	0.15	0.00	0.00	0.00	1
Doc4	0.16	0.00	0.04	0.02	0.01	0
...	...	...	...	...	...	...
Doc655	0.00	0.00	0.05	0.09	0.11	4
Doc656	0.02	0.02	0.22	0.00	0.03	2
Doc657	0.00	0.00	0.01	0.03	0.10	4
Doc658	0.06	0.00	0.03	0.08	0.04	3
Doc659	0.00	0.23	0.00	0.00	0.04	1

660 rows × 6 columns

	pilot error	On Fire	crash	poor weather	shot down	dominant_topic
Doc0	0.126	0.000	0.145	0.004	0.045	2
Doc1	0.041	0.025	0.075	0.094	0.000	3
Doc2	0.153	0.000	0.000	0.000	0.023	0
Doc3	0.007	0.000	0.004	0.155	0.000	3
Doc4	0.000	0.042	0.199	0.000	0.000	2
Doc5	0.091	0.248	0.031	0.000	0.003	1
Doc6	0.045	0.113	0.000	0.033	0.064	1
Doc7	0.052	0.000	0.048	0.000	0.054	4
Doc8	0.037	0.012	0.178	0.072	0.000	2
Doc9	0.001	0.006	0.000	0.204	0.000	3

LDA- Model:

	pilot error	On Fire	crash	dominant_topic
Doc0	0.07	0.07	0.87	2
Doc1	0.06	0.06	0.88	2
Doc2	0.06	0.06	0.89	2
Doc3	0.08	0.09	0.83	2
Doc4	0.06	0.06	0.89	2
Doc5	0.08	0.08	0.84	2
Doc6	0.06	0.06	0.88	2
Doc7	0.06	0.06	0.89	2
Doc8	0.05	0.05	0.90	2
Doc9	0.06	0.06	0.87	2

	pilot error	On Fire	crash	dominant_topic
Doc0	0.07	0.07	0.87	2
Doc1	0.06	0.06	0.88	2
Doc2	0.09	0.09	0.82	2
Doc3	0.09	0.09	0.83	2
Doc4	0.06	0.06	0.88	2