

# AIRPLANE CRASHES ANALYSIS (YEAR 1908 - 2009)

Author: Zoe Ni  
Date: 2024/4/09

Note: Codes and results are original

# OBJECTIVE

Act as a safety agent, aiming to draw observations about trends, main contributors (airlines/aircraft types), and reasons of airplane crashes. In this presentation, I try to investigate the following topics:

1. Time trend of total number of crashes, total fatality, and ratio (fatality/aboard)
2. Top 10 crashes locations/operators/aircraft types sorted by total fatality count
  - Distinguished by Military and Commercial operator
3. Main reasons for crashes: extracting text information from Summary of crashes
  - Distinguished by Military and Commercial operator
4. A video demo of interactive dashboard, focusing reasons on Commercial crashes.

# DATA SOURCE

- There are 5268 observations (worldwide crashes) and 13 features, spanning between year 1908 -2009.
- Features consist of: (usable features are highlighted)
  1. **Date** (date the crash had taken place)
  2. Time (time the crash had taken place)
  3. **Location** (of crashes)
  4. **Operator**
  5. Flight #
  6. Route
  7. **Type** (of aircraft)
  8. Registration
  9. cn/ln - ?
  10. **Aboard** (number of people aboard)
  11. **Fatalities** (lethal outcome)
  12. Ground (saved people)
  13. **Summary** (brief summary of the crash)
- There are in total 4,469 commercial flight crashes, and 781 military crashes
  - Why are we separating these categories? Because the locations and reasons for crashes could be drastically different, which will be investigated in more details later.

	Commercial	Military
Total crashes	4,469	781
Total fatalities	85,660	19,663
Total aboard	122,932	21,450
Fatalities per aboard	0.696	0.916

# SAMPLE DATA

- The first 10 records:

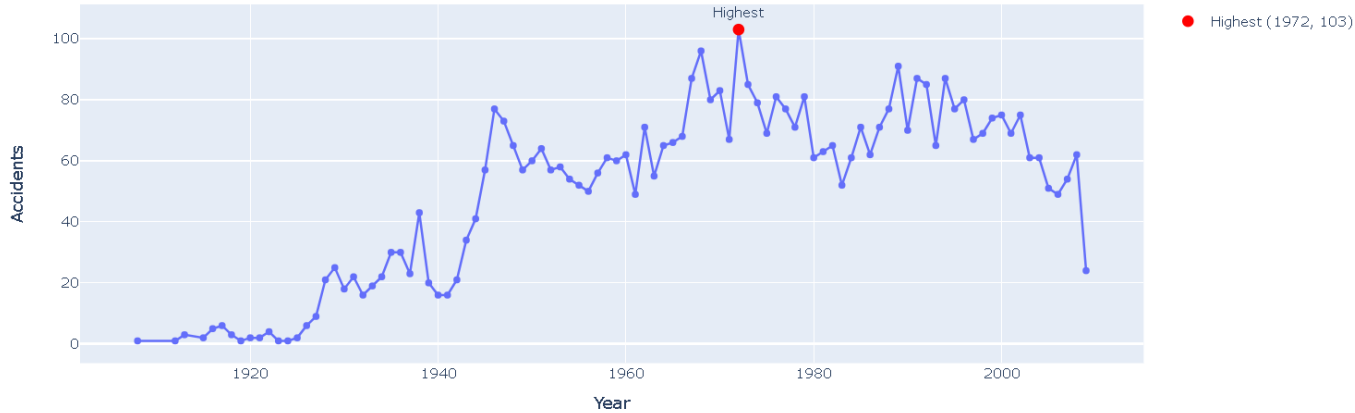
	A	B	C	D	E	F	G	H	I	J	K	L	
1	Date	Time	Location	Operator	Flight #	Route	Type	Registratcn/ln	Aboard	Fatalities	Ground	Summary	
2	09/17/1908	17:18	Fort Myer,	Military - U.S. Army		Demonstr	Wright Flyer III		1	2	1	0	During a demonstration flight, a U.S. Army flyer flown by Orville Wright nose-dived into the ground from a height of approxi
3	07/12/1912	06:30	AtlantiCity	Military - U.S. Navy		Test flight	Dirigible			5	5	0	First U.S. dirigible Akron exploded just offshore at an altitude of 1,000 ft. during a test flight.
4	08/06/1913		Victoria, B	Private	-		Curtiss seaplane			1	1	0	The first fatal airplane accident in Canada occurred when American barnstormer, John M. Bryant, California aviator was kille
5	09/09/1913	18:30	Over the M	Military - German Navy			Zeppelin L-1 (airship)			20	14	0	The airship flew into a thunderstorm and encountered a severe downdraft crashing 20 miles north of Helgoland Island into th
6	10/17/1913	10:30	Near Joha	Military - German Navy			Zeppelin L-2 (airship)			30	30	0	Hydrogen gas which was being vented was sucked into the forward engine and ignited causing the airship to explode and bu
7	03/05/1915	01:00	Tienen, B	Military - German Navy			Zeppelin L-8 (airship)			41	21	0	Crashed into trees while attempting to land after being shot down by British and French aircraft.
8	09/03/1915	15:20	Off Cuxha	Military - German Navy			Zeppelin L-10 (airship)			19	19	0	Exploded and burned near Neuwerk Island, when hydrogen gas, being vented, was ignited by lightning.
9	07/28/1916		Near Jamt	Military - German Army			Schutte-Lanz S-L-10 (airship)			20	20	0	Crashed near the Black Sea, cause unknown.
10	09/24/1916	01:00	Billericay,	Military - German Navy			Zeppelin L-32 (airship)			22	22	0	Shot down by British aircraft crashing in flames.

- After dropping irrelevant features, we can sort by the Fatalities (top 10):

	Date	Time	Location	Operator	Type	Aboard	Fatalities	Ground	Summary
2963	1977-03-27	17:07	Tenerife, Canary Islands	Pan American World Airways / KLM	Boeing B-747-121 / Boeing B-747-206B	644.0	583.0	0.0	Both aircraft were diverted to Tenerife becaus...
3568	1985-08-12	18:56	Mt. Osutaka, near Ueno Village, Japan	Japan Air Lines	Boeing B-747-SR46	524.0	520.0	0.0	The aircraft suffered an aft pressure bulkhead...
4455	1996-11-12	18:40	Near Charkhidadri, India	Saudi Arabian Airlines / Kazastan Airlines	Boeing B-747-168B / Ilyushin IL-76TD	349.0	349.0	0.0	Midair collision 17 km W of New Delhi. The Sa...
2726	1974-03-03	11:41	Near Ermenonville, France	Turkish Airlines (THY)	McDonnell Douglas DC-10-10	346.0	346.0	0.0	The aircraft crashed shortly after takeoff fro...
3562	1985-06-23	07:15	AtlantiOcean, 110 miles West of Ireland	Air India	Boeing B-747-237B	329.0	329.0	0.0	The aircraft broke up in flight and crashed in...
3240	1980-08-19	19:08	Near Riyadh, Saudi Arabia	Saudi Arabian Airlines	Lockheed 1011-200 TriStar	301.0	301.0	0.0	The flight experienced a fire in the aft cargo...
3775	1988-07-03	10:55	Over the Persian Gulf, near Bandar Abbas, Iran	Iran Air	Airbus A300B2-203	290.0	290.0	0.0	The civilian Iranian airliner was shot down by...
4916	2003-02-19	17:30	Near Shahdad, Iran	Military - IslamiRevolution's Guards Corps	Ilyushin IL-76MD	275.0	275.0	0.0	The plane crashed into an 11,500 ft. mountain...
3137	1979-05-25	15:04	Chicago O'Hare, Illinois	American Airlines	McDonnell Douglas DC-10-10	271.0	271.0	2.0	During takeoff just, as the plane lifted from ...

# OBJECTIVE 1 - Time trend of total number of crashes, total fatality, and ratio

Number of accidents by year

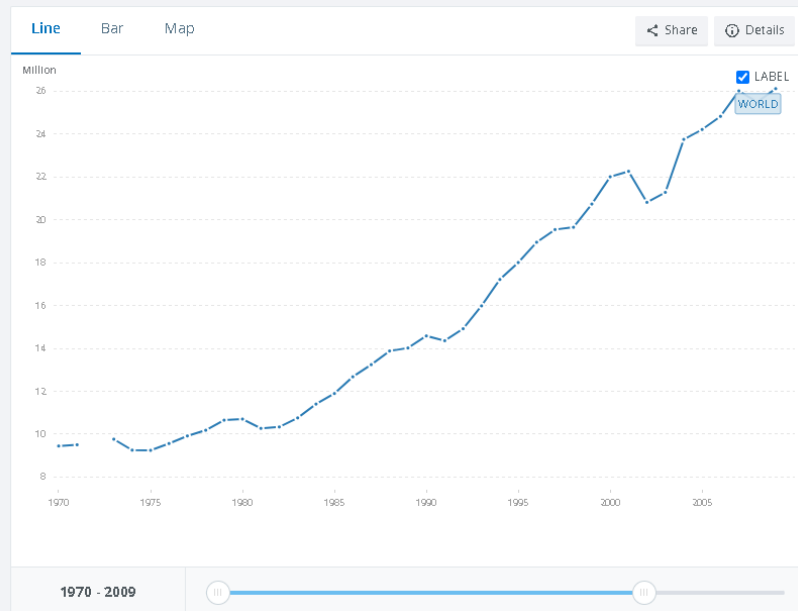


- The year 1972 experienced the highest number of crashes at 103. The total fatality also seems to peak between years 1960 and 1990.
- The ratio figure shows that death/aboard ratio seems to trend downwards after 1960.
- The bottom left figure shows that flight accessibility increases steadily. Therefore, we cannot conclude that crashes occur more often between years 1960 and 1990.

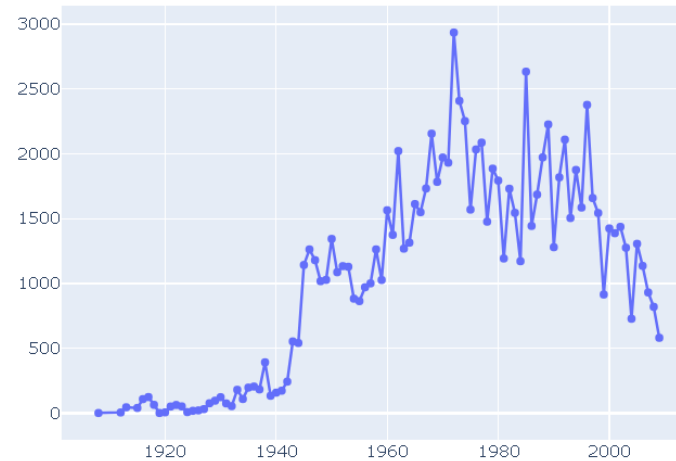
Air transport, registered carrier departures worldwide

International Civil Aviation Organization, Civil Aviation Statistics of the World and ICAO staff estimates.

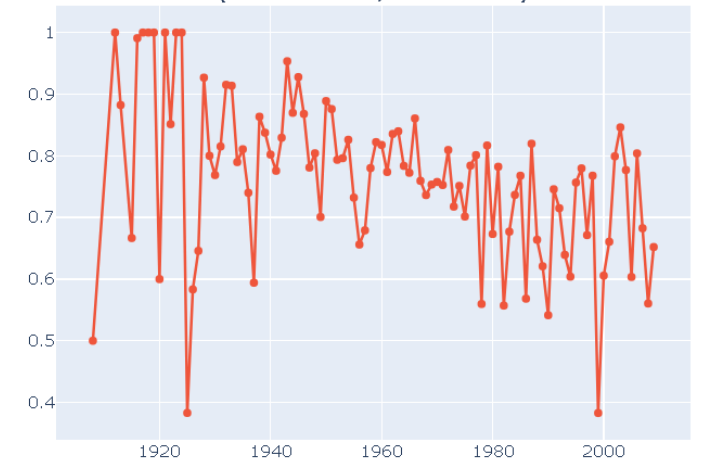
License: CC BY-4.0



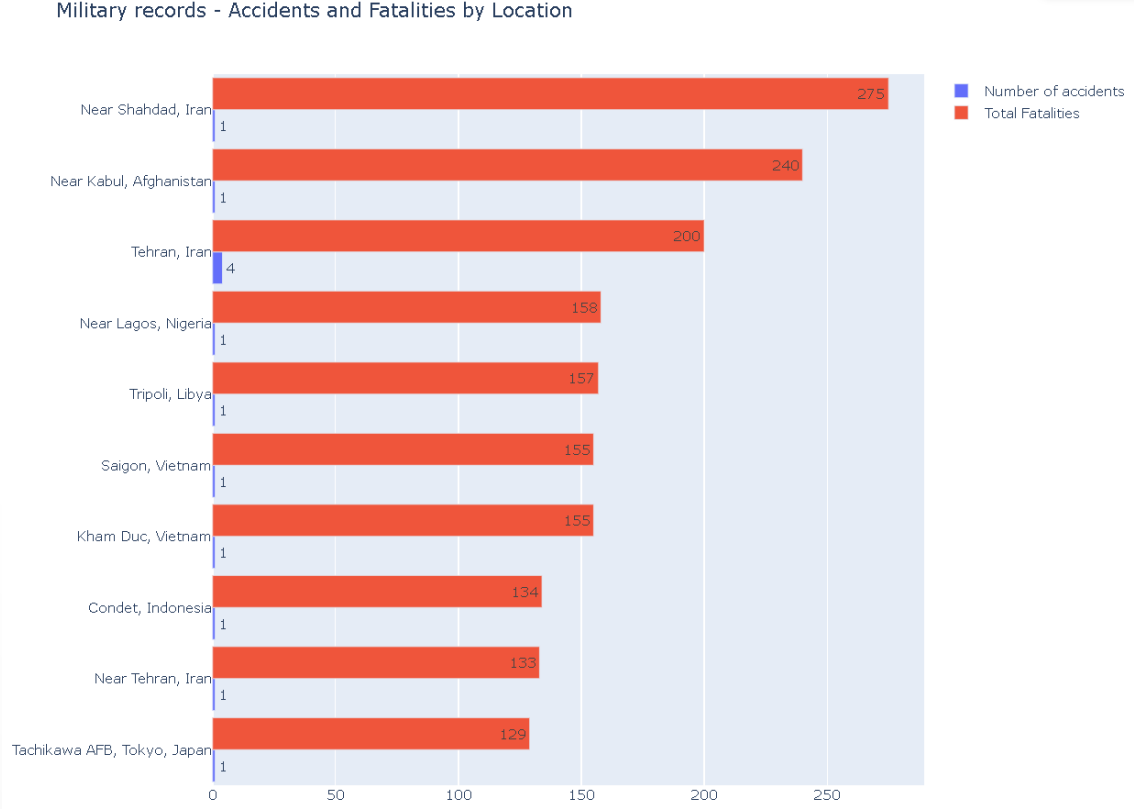
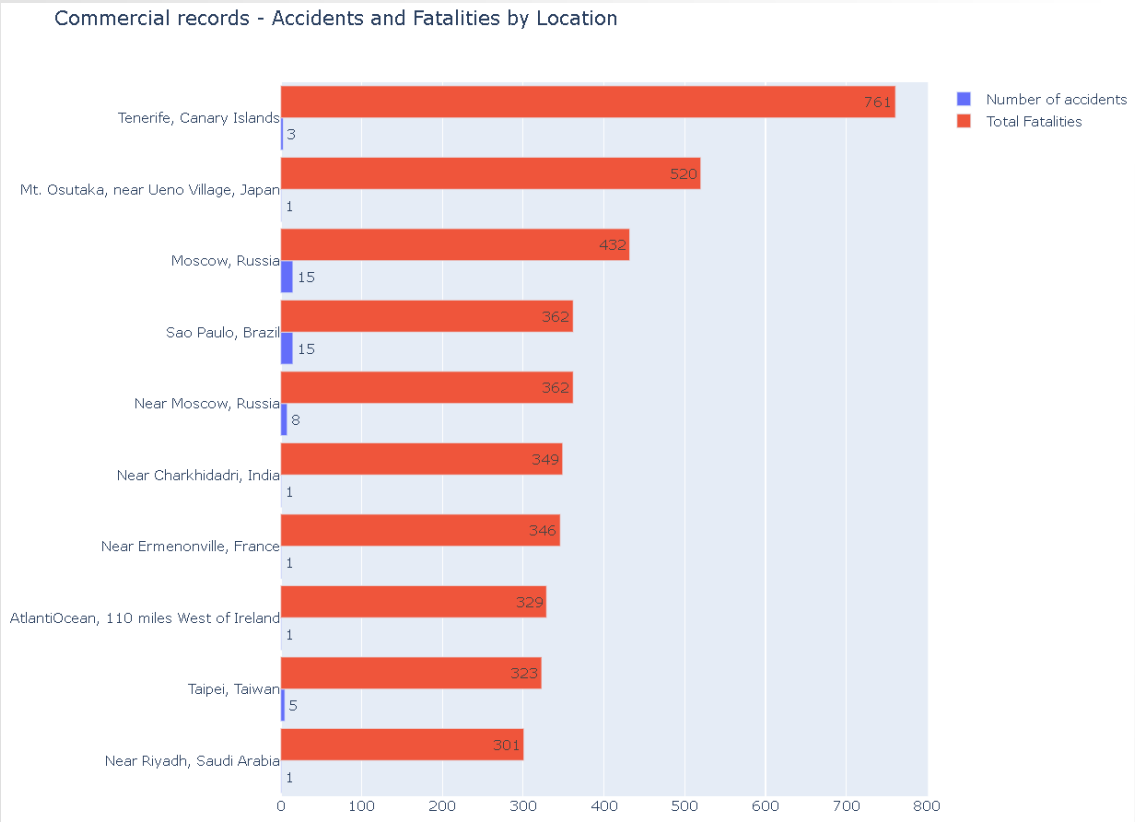
Total fatalities trend



Ratio (total fatalities/total aboard) trend

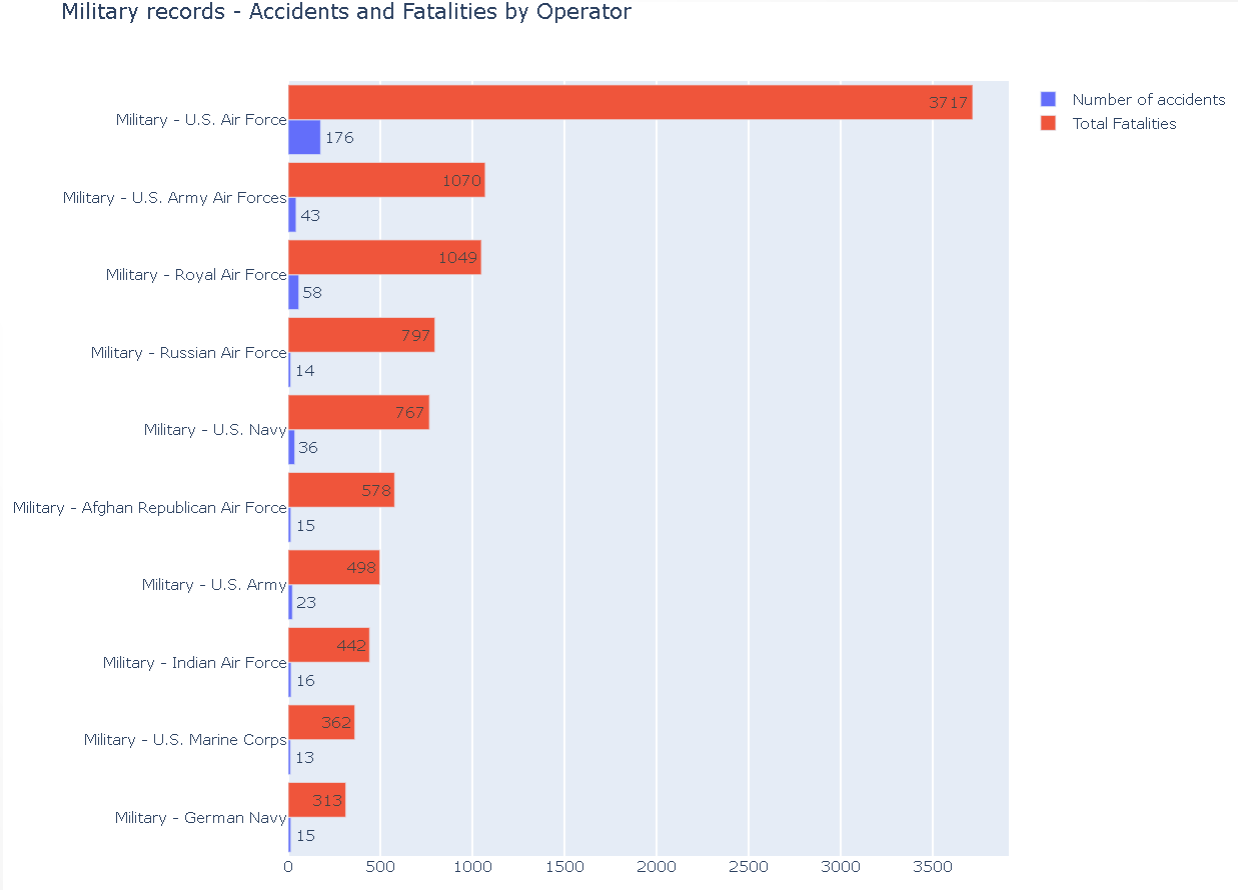
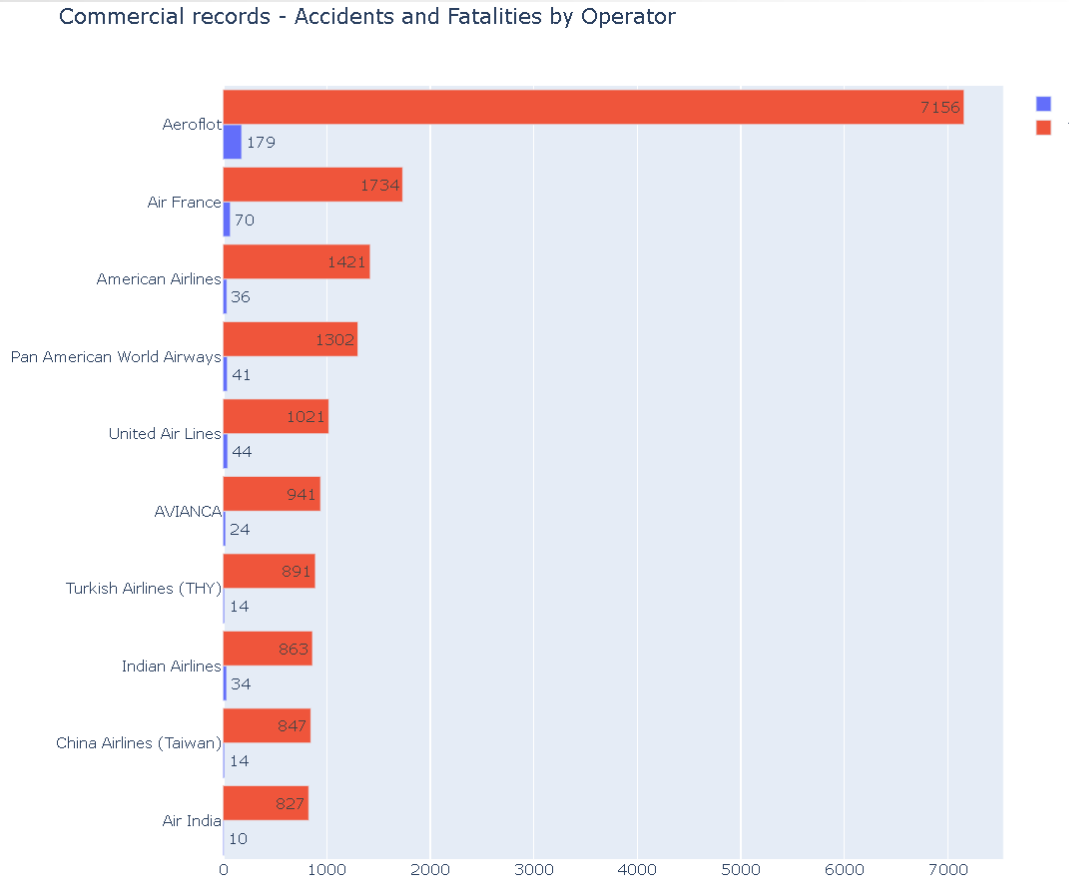


# OBJECTIVE 2- Top 10 accident locations sorted by total fatality



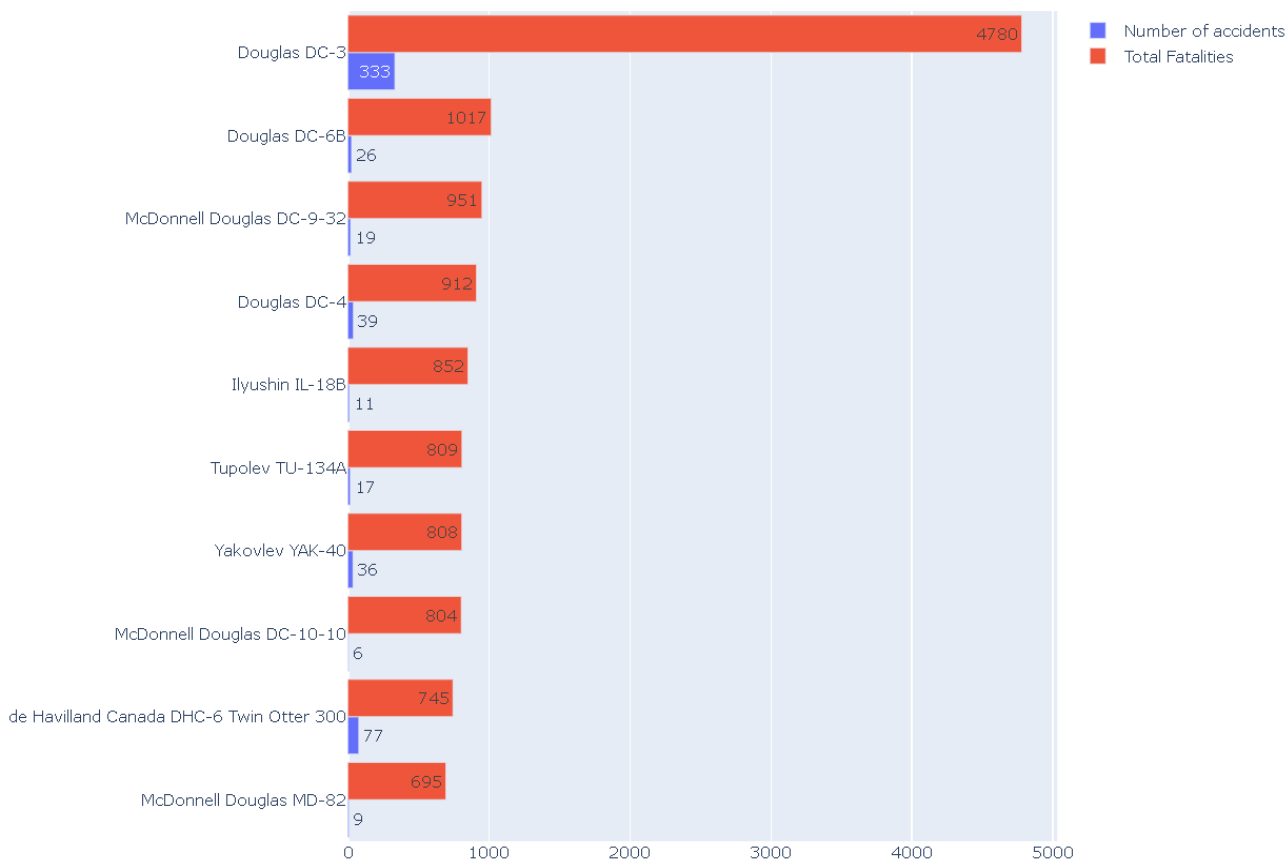
- Commercial records:
  - Tenerife tops the list with 3 crashes. One notable crashes occurred due to collision of two airplanes, causing 583 fatalities out of 644 passengers. The KLM plane began its takeoff roll without permission with the Pan Am plane still on the runway. Two other deadly crashes are cause by airplanes crashing into high terrain.

# OBJECTIVE 2- Top 10 operators sorted by total fatality

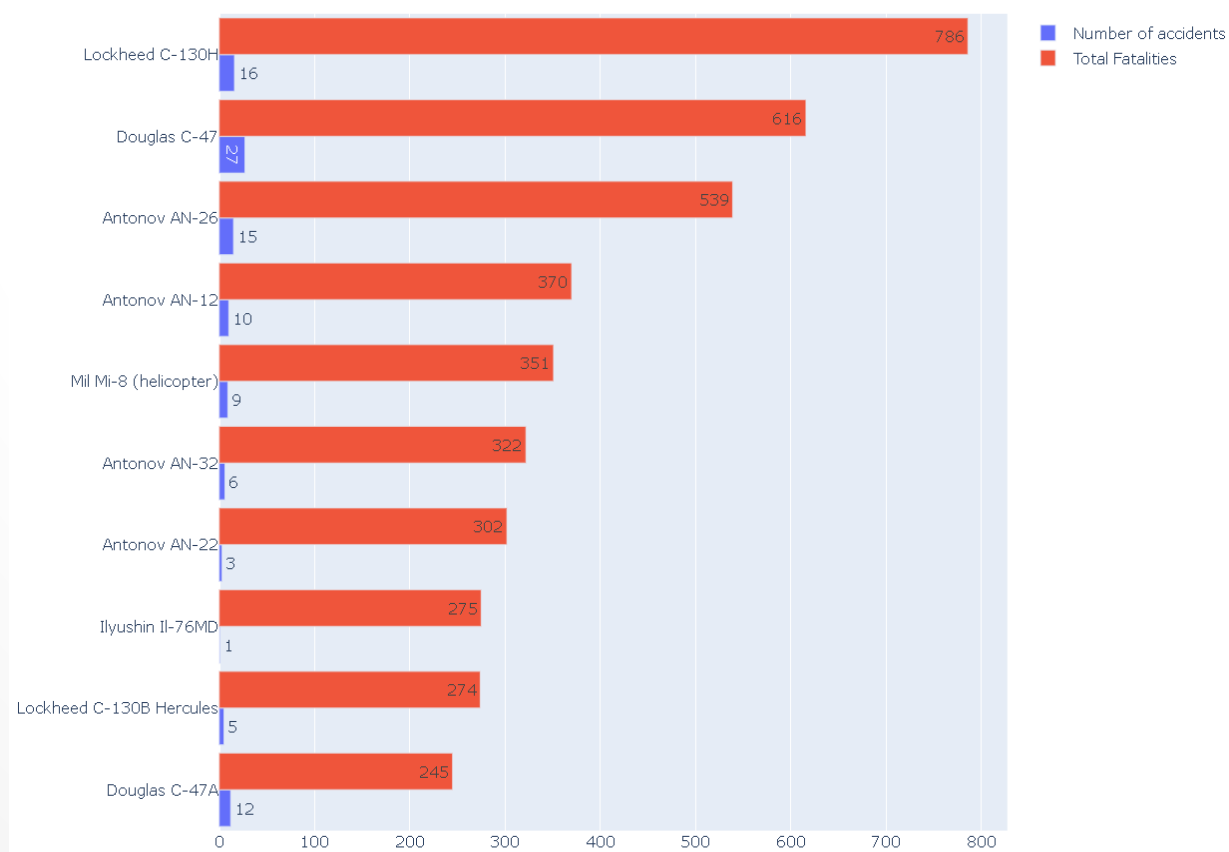


# OBJECTIVE 2- Top 10 aircraft types sorted by total fatality

Commercial records - Accidents and Fatalities by Type



Military records - Accidents and Fatalities by Type





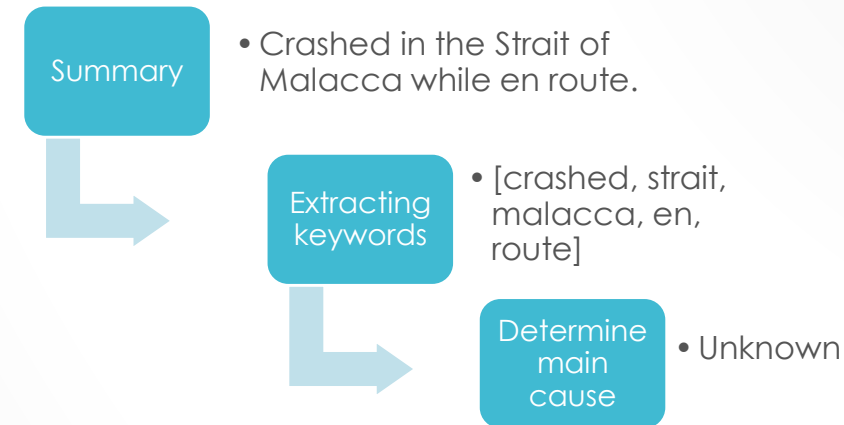
# OBJECTIVE 3 - REASONS FOR CRASHES

## Methodology:

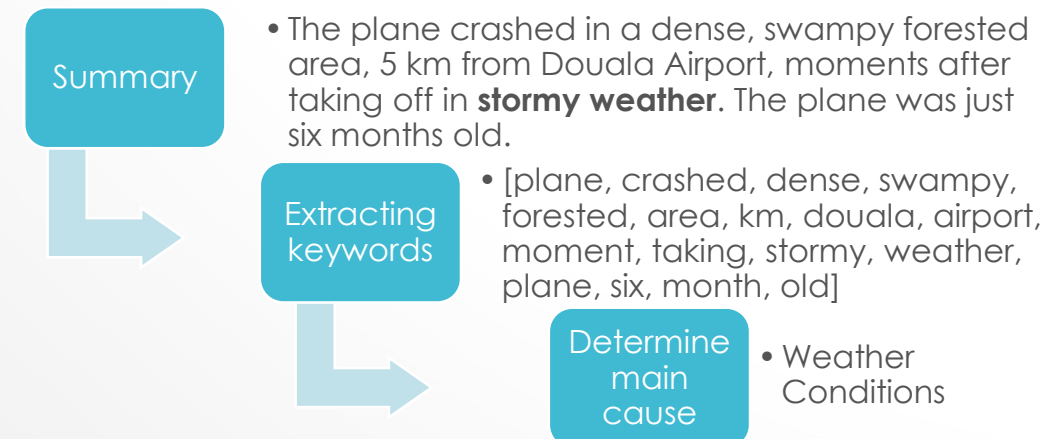
Utilizing a human language data library called *nltk*, we can transform the summary of the crash into a list of keywords, and determine the main cause of crashes by selecting the category with highest occurrences of keywords defined by the author. If the summary doesn't include any keywords, the reason is classified as 'Unknown'

Categories	Contain Keywords
Shot down	'shot', 'missiles', 'missile'
Weather Conditions	'stormy', 'sandstorm', 'snow', 'rainstorm', 'typhoon', 'rain', 'thunderstorm', 'thunderstorms', 'fog', 'visibility', 'weather', 'conditions', 'condition', 'icing', 'storm'
Pilot Error	'suicide', 'misjudgment', 'judgment', 'procedure', 'negligence', 'fatigue', 'neglected', 'overconfidence', 'communication', 'pilot', 'captain'
Mechanical Failure	'disintegrated', 'structural', 'cracks', 'system', 'software', 'tire', 'wheel', 'malfunctioned', 'mechanical', 'malfunctioning', 'engine', 'hydraulic', 'malfunction', 'propeller', 'fuel'
Collision/Geographical	'collision', 'collided', 'trees', 'terrain', 'struck', 'mountain', 'hill', 'terrain', 'trees'
Terror Attacks	'terror', 'explosives', 'explosive', 'terrorists', 'bombs', 'bomb', 'hijacker'

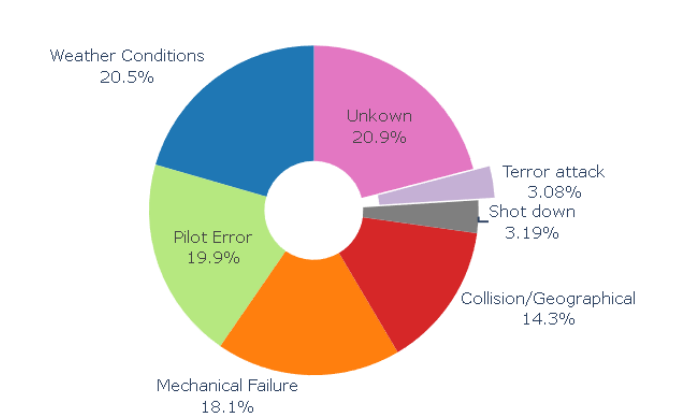
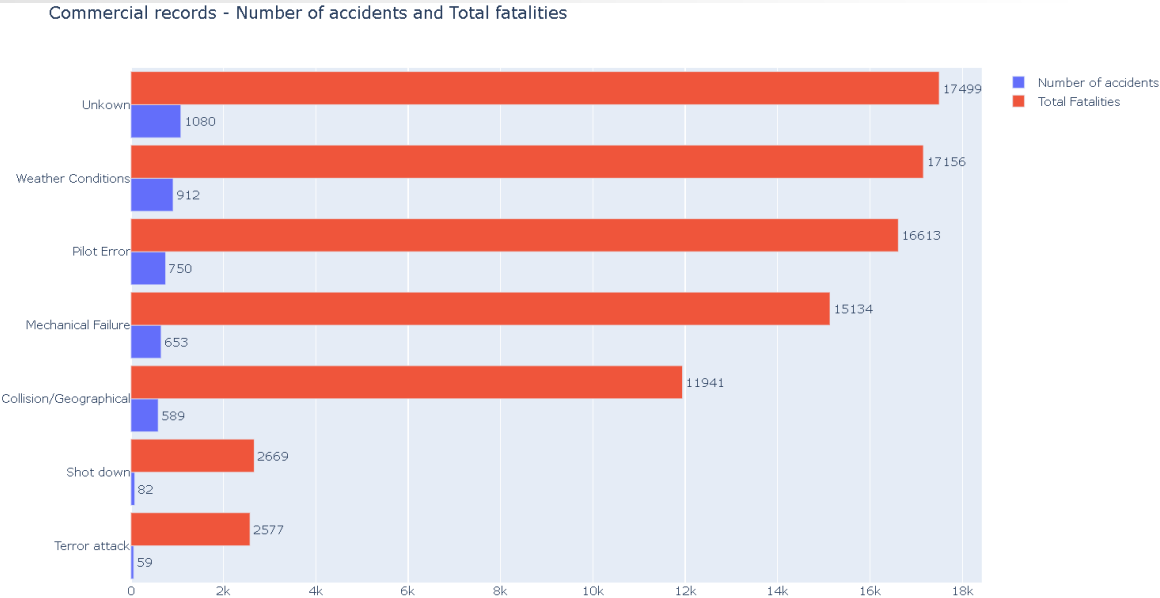
- An example of 'Unknown' cause:



- An example of 'Weather Conditions' cause:

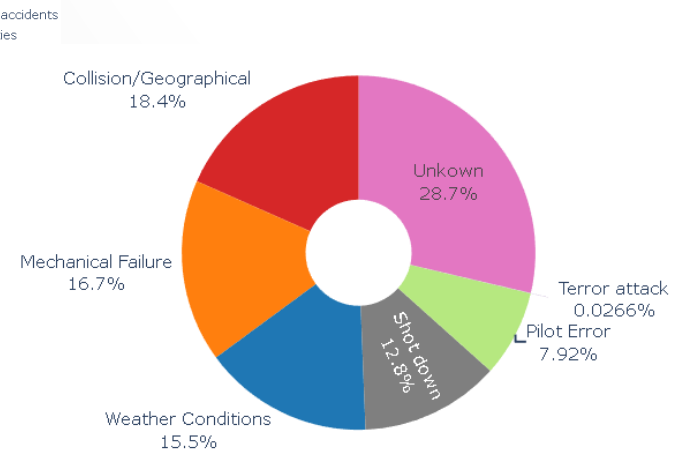
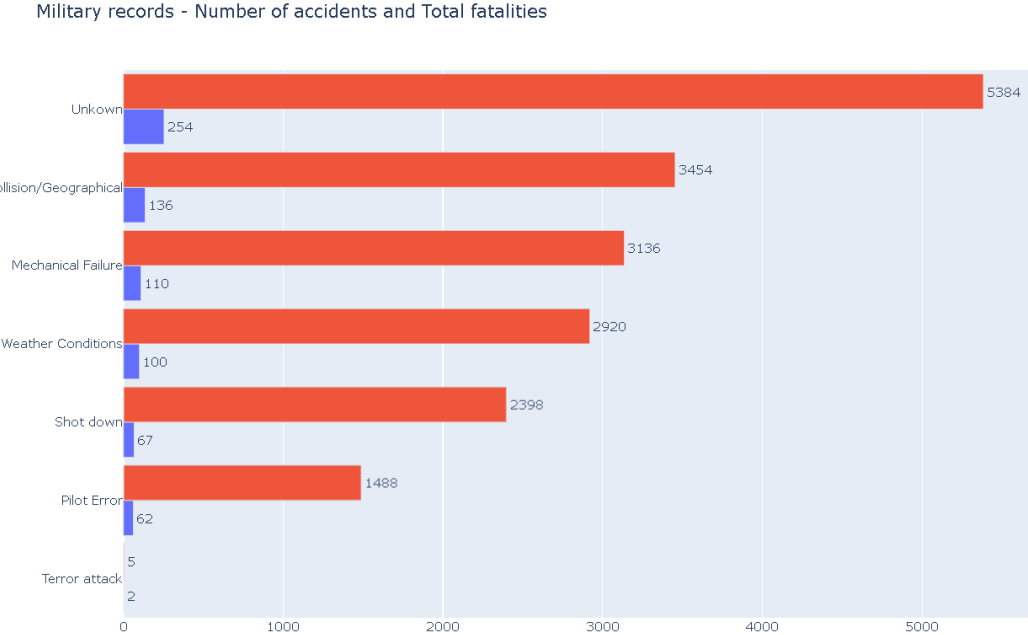


# OBJECTIVE 3 - CONT



## Commercial:

- ‘Weather Conditions’ leads to highest fatalities after reason ‘Unknown’, followed by ‘Pilot error’
- ‘Shot down’ is less likely in commercial flights.



## Military:

- ‘Collision/Geographical’ leads to highest fatalities after reason ‘Unknown’, followed by ‘Mechanical Failure’
- Aircraft shot down is more likely at 12.8%, while ‘Pilot error’ is less likely in military flights crashes.

## OBJECTIVE 4 - REASONS ON COMMERCIAL CRASHES.

During the demo in the next slide, I will perform the following steps/tasks:

1. Run the Python code for the dashboard-> open the dashboard
2. Browse the Commercial crash table
3. Filter and sort on the table -> show pivot table and graphs based on filtered data
4. Choose a crash reason from the Dropdown and the table is updated as requested.

## FURTHER RESEARCH

1. It is sometimes difficult to determine a single reason for a crash. It could be a combination of plausible reasons.
2. Lack of measure of accuracy of the classification of main reasons for crashes
3. Main reasons 'Unknown' tops the list for both commercial and military flights, one could investigate and try to reduce the count, however it could lead to overfitting the data.