

# AIRPLANE CRASH ANALYSIS

## ***INTRODUCTION***

*Airplanes the most efficient form of modern transportation ,make the world a small place by quickly delivering people and cargo from one place to another . As more and more people use airplanes as a means of transportation the no. of passengers and flight routes have expanded recent years . Taking these impacts into consideration this projects aims to discover and analyze reasons for airplane crashes , raise awareness of flight safety, and better understanding its problems and progress .*

### ***DATA SOURCE:***

*The data is taken from kaggle . The data contains details of crashes between 1908 till 2009.*

*The data set contains 13 unique variables and 5268 data points.*

### ***FEATURES:***

*-Date*

*-time*

*Location*

*-Operator*

*-Type*

## ***METHODOLOGY***

*In this section we are going to walk through the life cycle of data science project.*

*Data collection    ⇒    Data cleaning    ⇒    Data analysis*



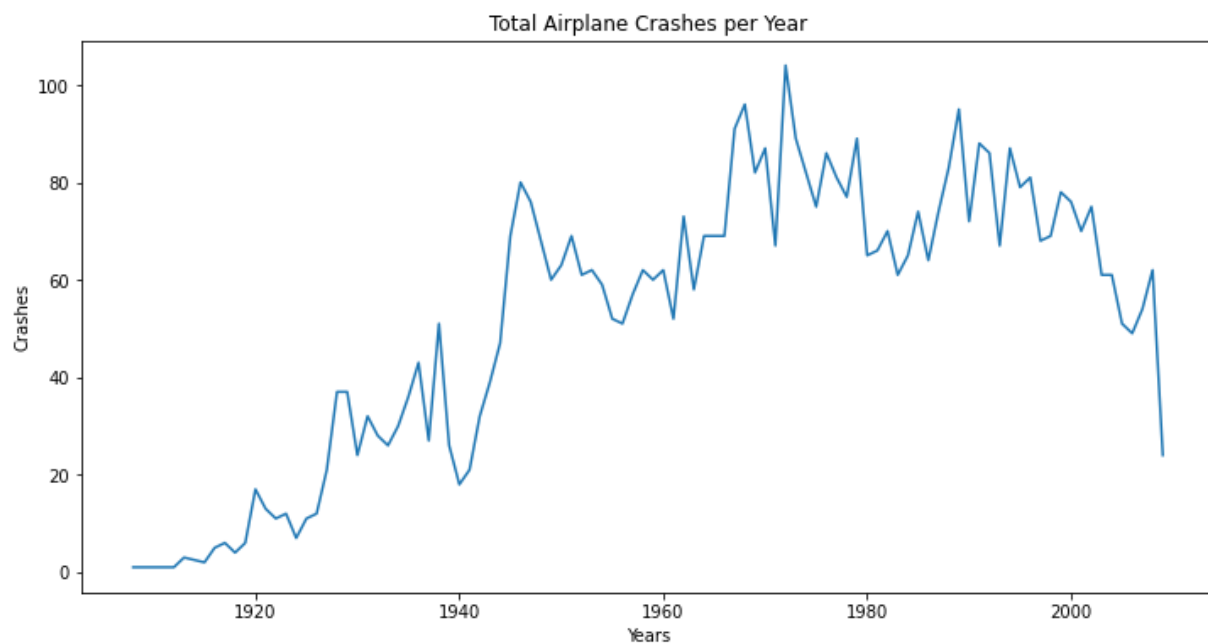
*Data visualizations*

*We have already talked about our data collection process ,in this section we are going to explain our data cleaning methodology, findings of our analysis and data visualization .*

## ***DATA CLEANING AND DATA VISUALIZATION***

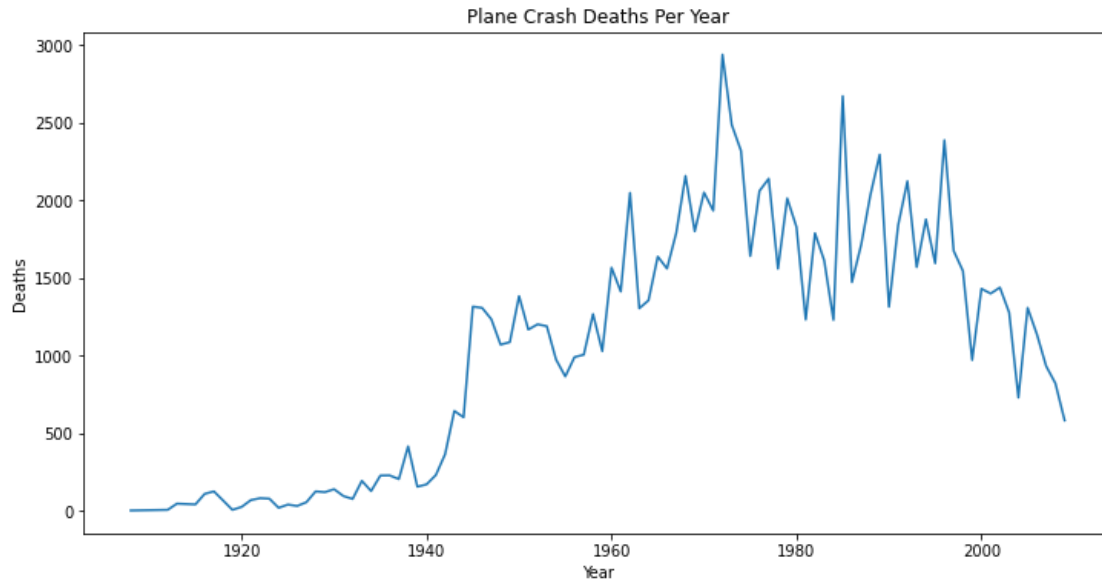
*I started my analysis by looking at the null values in our dataset . found many null values and fill it with appropriate values . And drop the unnecessary values and columns. Now our dataset is ready for visualization.*

*First of all I checked how many crashes per year?*



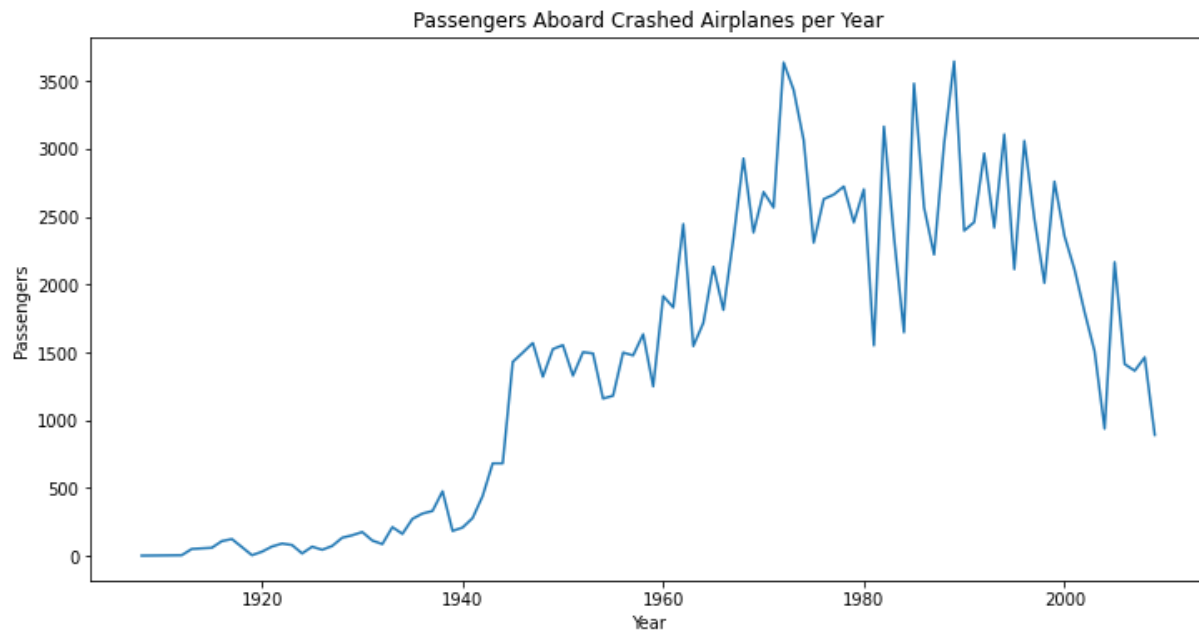
*From the 40's, there's a significant increase in airplane crashes, which must likely be because of World War II (1939 - 1945). The highest peaks are between 1960 and 2000. The year with most accidents is 1972, with 104 occurrence .*

*Now lets analyze How many deaths per year ?*



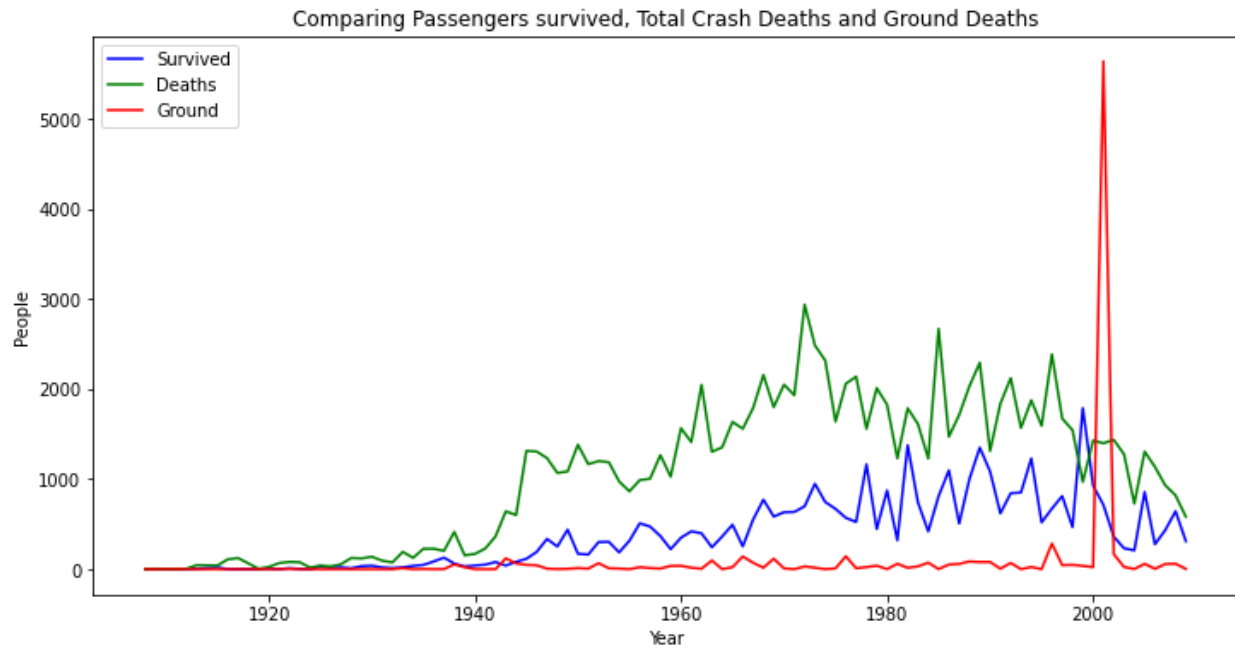
***Here we can see the same pattern, the years that had the most accidents are also the ones with the most fatalities .***

***How many passengers in crashed planes each year ?***



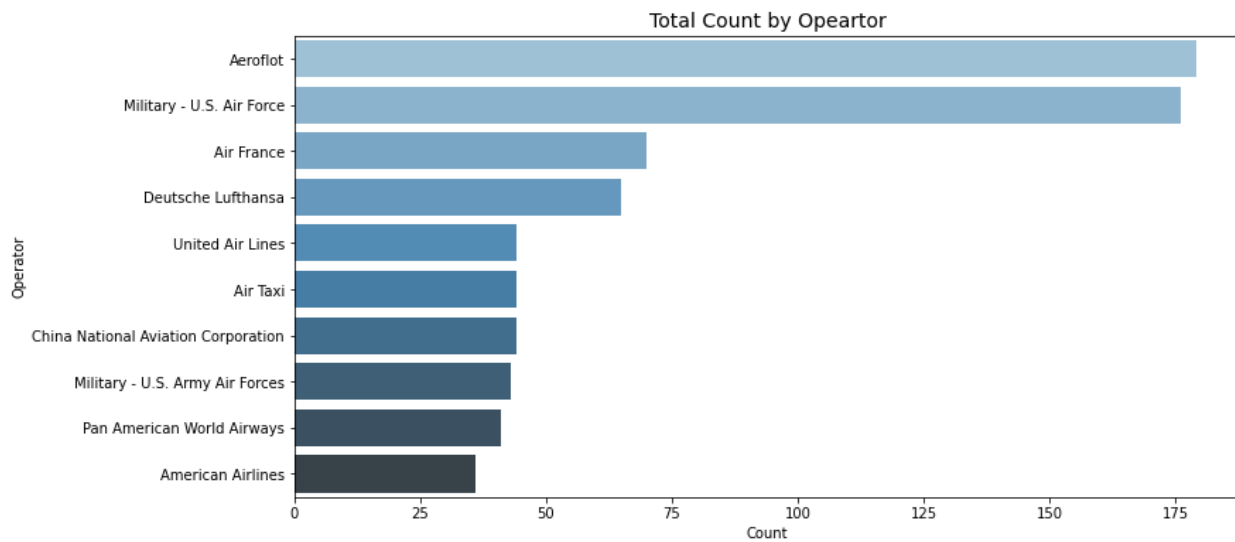
***From the 40's, the number of people aboard airplanes starts to increase. From 1960 to 2000 is where we have most people aboard, the same years with most plane crashes and fatalities .***

***Analysing the values of survived ,fatalities and ground ?***



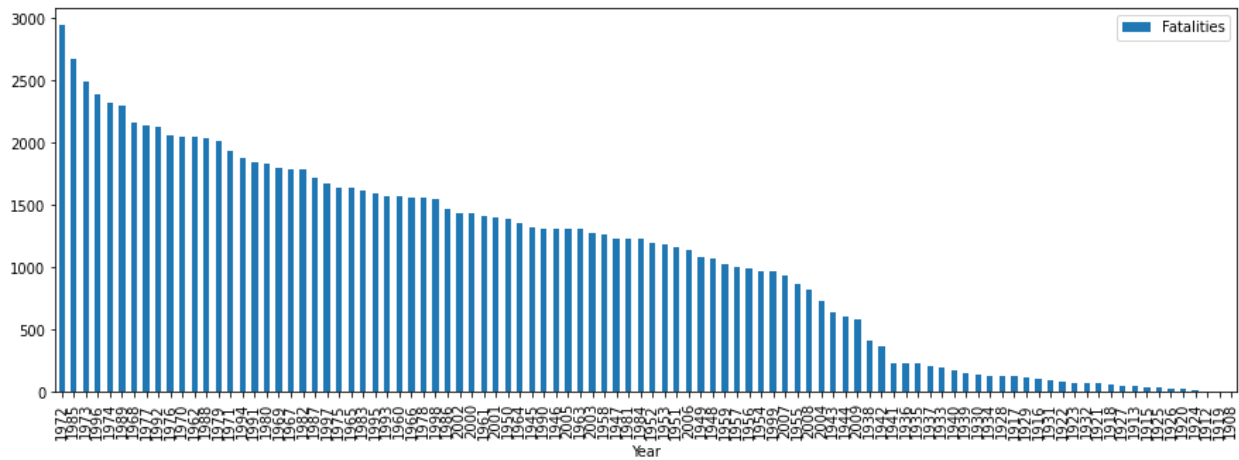
***There's a peak close to 2000, with more than 5000 killed on the ground, way more than the number of fatalities. This dreadful number of people killed on the ground is due to the tragic event of 9/11, where the Twin Towers were brought down by two planes hijacked by terrorists.***

***which was the worst Operator ?***



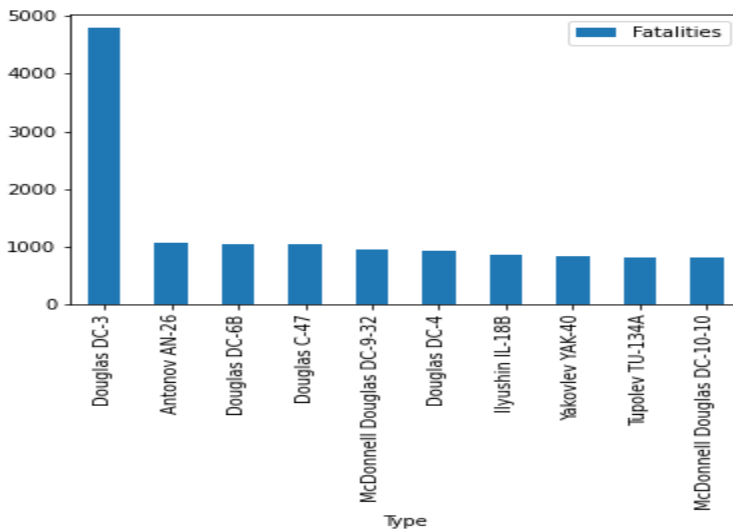
***Aeroflot is the most dangerous operator. People have high chances of dying if they flew with Aeroflot than any other operator .***

***which was the worst year to fly ?***



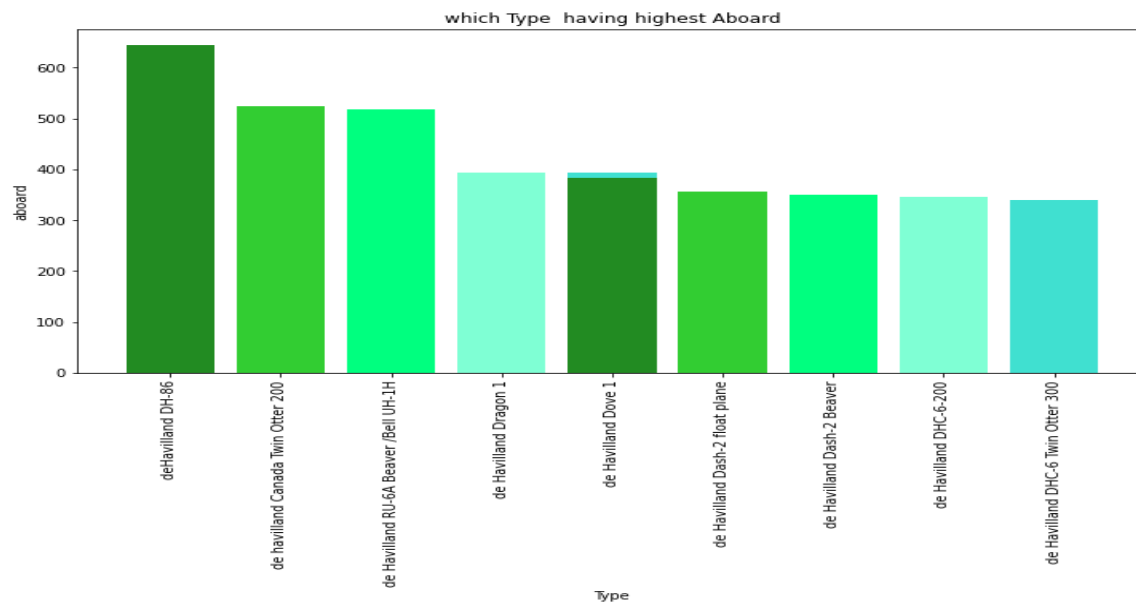
***70s and 80s are among the worst years for airline industry. 1972 is the worst so far, causing maximum no. of fatalities***

***which Type is used in the most crashed planes ?***



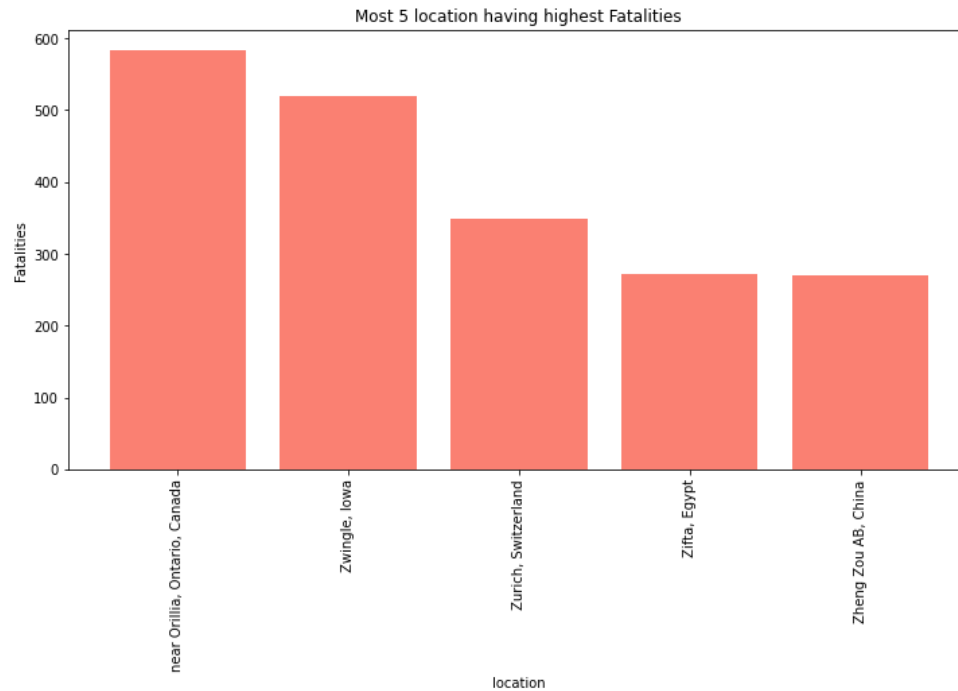
***Douglas DC-3 is atleast 5 times more probable to crash than any other type of aircraft***

***which type of plane has highest aboard ?***



***dehavilland DH has most no.of aboard , from the graph its clear that dehavilland type is most used and its variety types. This type and its variety is most demanding and this type can occupy more aboards than any other.***

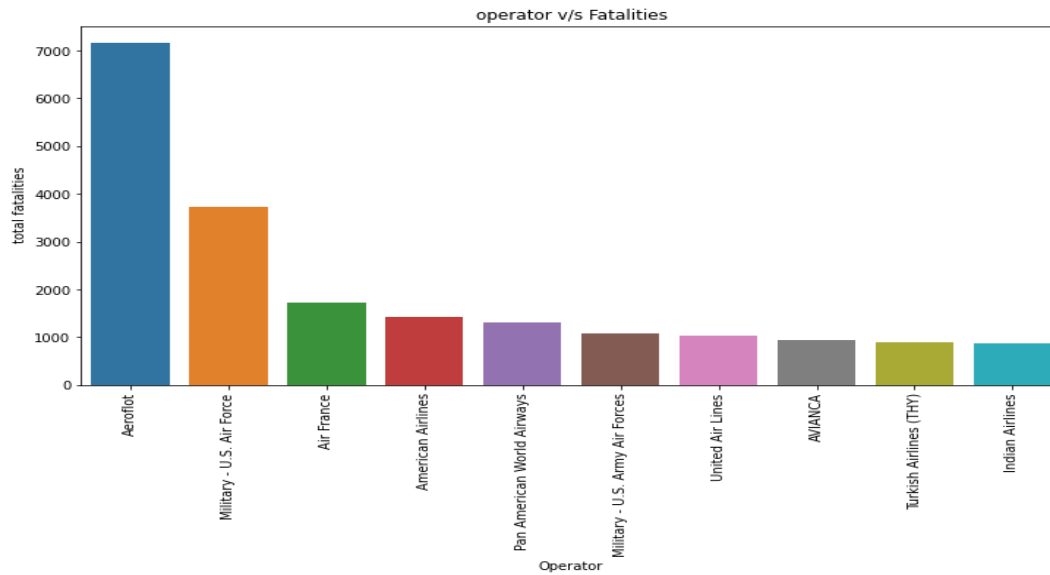
***which location has highest fatalities?***



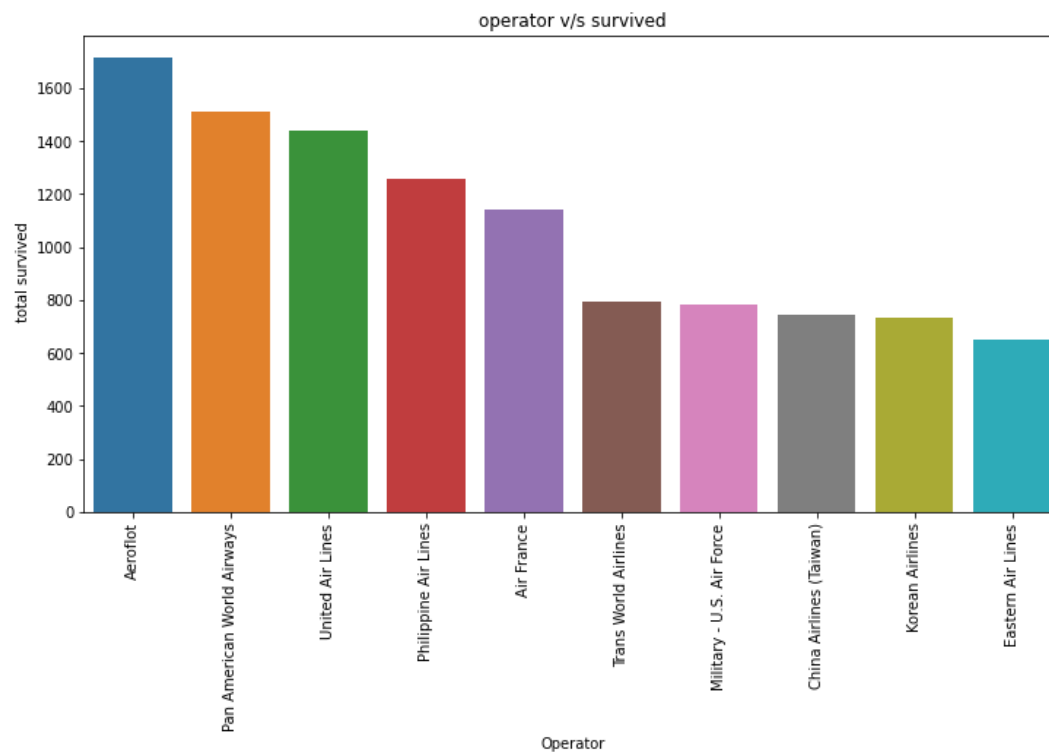
***Canada have the highest no.of crashes occurred and has highest fatalities. And this is most the riskiest location to fly***

***what is the relation between operator and fatalities?***

***The operator aerflot is the most dangerous operator and have high fatalities than others. second most is Military U.S air force***



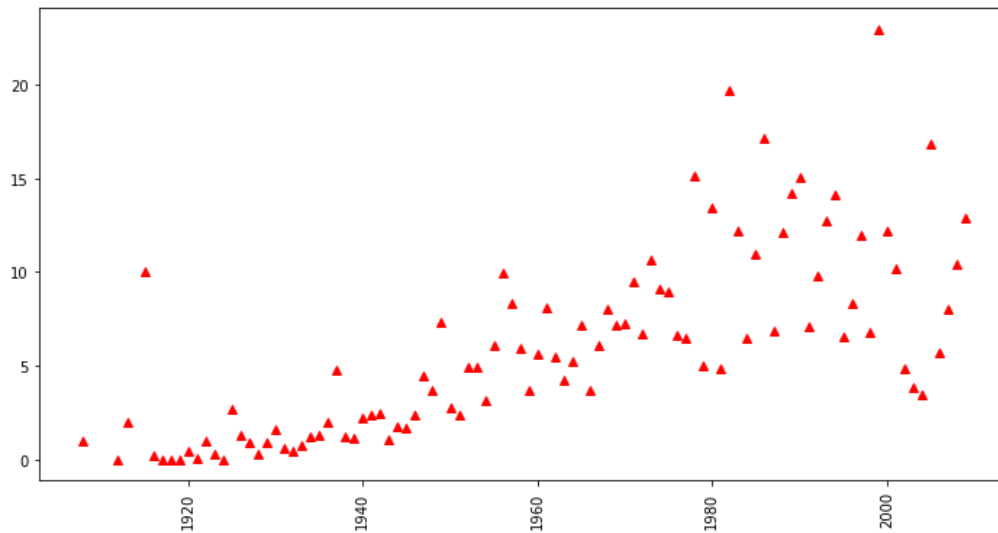
which operator has highest survivors ?





*as we know that aerflot is the worst operator from our previous observations, even though there is a chance to survival. in this graph the aerflot has the highest no. of survivors almost 1700 were survived. second most is pan american airways its around 1500 and so on.*

*which year has highest survived?*



*Earlier the survivor rate is really low and after 1980 there is a change in the pattern. which means after 1980's survivor rate is highly increasing, its a good sign for aviation industry. And people using airways also feel safer. technological development is also a factor.*

## ***CONCLUSIONS***

*This analysis will be helpful for both aviation industry as well as people who choose air ways as their means of transportation. they get better understanding of how crashes happen, is there any hope for survival in case anything happens, and how much safe it is?. We analyze data of around 101 yrs of plane crashes. From this it's clear that early 40's, and 70's to 80's are the worst period of aviation industry. lots of crashes and the fatalities rate also high, and also that time the world war 2 happened it's also a reason for highest crash rates. Aeroflot was the worst operator and the planes have Douglas C3 as their type had the high chances to crash. the year which higher crashes were occurred in 1972 around 104 crashes. Most accidents took place in Canada, may be which is the riskiest*

*location. The survival rates were increased from 1980's till 2009, which means technologies are improved, the fatalities rates were decreased and also crash rates also decreased, it's a good sign. There are so many reasons for air accidents, may be its technical issue, air turbulence, terrorist attacks, or may be pilots error, even though which is an efficient form of modern transportation.*

*The 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, and help raise customer confidence with the use of statistical evidence.*