

Airlines Delay Causes Exploration

By Hosam Elbakry

A- Dataset Overview

The data we have here -from Udacity datasets options- is airlines delay causes from 2017 till 2020

The dataset contains **22 column** and **69522 record** or row. Through the project we will start to understand the dataset variables and carryout the necessary wrangling/cleaning and then figure out which variables worthy to be concluded in our analysis.

The insights we are trying to inspect here for example

- what is the main cause of delay?
- what is the best time for travel?
- Is the destination effect on delay?

B- Findings and insights

1. Correlation

We can notice that the highest correlation exists between the following variables in descending order ['late_aircraft_delay', 'carrier_delay', 'nas_delay', 'weather_delay', 'security_delay'].

2. Main delay cause

The most variable cause delay is late_aircraft_delay.

3. Airlines delay

Most of the delay comes from SkyWest Airlines With a percentage 15%.

4. Airports/Destination

I cannot say that there is a certain airport cause delay they are almost the same.

5. Delay over time

The delay causes are declining it is tending to minimum in 2020.

6. Best/worst time to travel

- The Best time to travel is from September to November.
- The worst time to travel is from June to August.