

Tableau Project Report

Tableau workbooks links:

1- First version before review:

https://public.tableau.com/profile/kariman.fawzi#!/vizhome/AirlineDelayStory_2Modified/Story2

2- Second version after review

<https://public.tableau.com/profile/kariman.fawzi#!/vizhome/AirlineDelayCausesExploration-2ndversionupdated/Story1>

3- Final version:

https://public.tableau.com/profile/kariman.fawzi#!/vizhome/Final_v4_15762582776180/Story3?publish=yes

Summary:

I obtained my dataset of Flight delays from RITA website which contained information on flight delays and the performances. The link to download the dataset is,

https://www.transtats.bts.gov/OT_Delay/OT_DelayCause1.asp

I downloaded the data from January 2004 to December 2018.

Before starting my visualizations, I spent a considerable time in understanding the variables of the dataset and its data types, and to be honest, there are some variables I couldn't figure out what they represent or in what units they are, so I avoided using them.

That would be a great help if you checked on variables definitions on RITA site as the downloaded data has some undefined variables.

I started my exploration with one main simple question in mind, is it extreme weather conditions or security checks that causes the most delays?

From that simple question, I started exploring the data variation over time and experiment with other variables like airport and carrier variations.

Design:

My first design was aiming at exploring overall delays over years and months looking for patterns, then exploring the delay causes over time.

Actually, I found a hard time getting started my story, I don't know my audience or their backgrounds, what should I tell or what I shouldn't. So, finally I decided to start and get some feedback that would help build a clear path.

I created my first trial and sent it to some friends. From their feedback, which was definitely a great help, I started to improve and resend and reimprove again till I reached an acceptable version. Of course, the nicest thing about this process is that there is always a room for improvements.

So, let's jump to my final design, as the first one I go over relationships with time, then carriers & airport variations and ended up with exploring causes.

I have changed charts types till find the ones that convey my message, and I was wrong, it is not weather conditions nor security checks that causes the most delays, actually they are least causes.

Feedback:

I received feedback from my friends through mails and chats. They were mostly instance messages and mostly in Arabic. I don't know how to collect them, so I will mention the main points here:

A- From Shymaa Saad:

- 1- At first version, I start my findings without emphasizing what I'm exploring exactly.
- 2- She mentioned that she as a passenger wants to know which companies have high delay rates to avoid them.

B- Sara Kilany: Sara has a lot of notes and was a great help actually

- 1- She found it hard to see that tiny writing font.
- 2- She noted that I need to add the company and airport names to slides instead of appearing on pointing only, we argued about that for a while.

C- Mahitab Abd El Monem:

- 1- She noted some typos.
- 2- She noted that I haven't clarify that the data belong to US only.

Resources:

1- Tableau documentation:

<https://www.tableau.com/support/help>

2- ASA Sections on Statistical Computing & Statistical Graphics:

<http://stat-computing.org/dataexpo/2009/the-data.html>

3- Bureau of Transportation Statistics:

https://www.transtats.bts.gov/Fields.asp?Table_ID=236