Every visualization included in this project follows this protocol.

# The visualization protocol consists of the following:

1. Timespan:

The data used was collected during the year 2018. We need to keep into account the fact that with the recent COVID-19 pandemic it may be possible that this data does not reflect anymore the current situation.

1. Data source:

The data was collected from the United States Department of Transportation, more precisely on the Bureau of Transportation Statistics section, it can be found at this link: [Bureau of Transportation Statistics](https://www.transtats.bts.gov/Homepage.asp)

1. Link to the dataset used:

The specific dataset was gathered from a Kaggle page containing a collection of Air Traffic Data with a span from 2009 to 2018, it can be found here: [Airline Delay and Cancellation Data, 2009 - 2018](https://www.kaggle.com/datasets/yuanyuwendymu/airline-delay-and-cancellation-data-2009-2018)

1. Metadata of the dataset:

The Features of the main dataset are:

FL\_DATE (date of the flight),

OP\_CARRIER (name of the carrier operator),

OP\_CARRIER\_FL\_NUM (flight number of the carrier operator),

ORIGIN (origin airport IATA code),

DEST (destination airport IATA code),

DEP\_TIME (time of departure),

DEP\_DELAY (departure delay),

TAXI\_OUT (time spent on the runway),

WHEELS\_OFF (time when wheels are of the ground),

WHEELS\_ON (time when wheels touch the ground),

TAXI\_IN (time spent on the arrival runway),

ARR\_TIME (time of arrival at the gate),

ARR\_DELAY (delay on arrival),

CANCELLED (true or false),

CANCELLATION\_CODE (determines the cancellation reason),

DIVERTED (true or false),

ACTUAL\_ELAPSED\_TIME (total time elapsed),

AIR\_TIME (time in the air),

DISTANCE (In kilometres),

DELAY\_REASON (carrier\_delay, weather\_delay, nas\_delay, security\_delay, late\_aircraft\_delay).

1. Short abstract of the data visualization process:

The dataset was processed using python, it didn’t require any pre-processing and alongside the Airports Dataset containing IATA code and coordinates available on the Bureau of Transportation Statistics site, for each of the visualizations we saved a new file containing only the necessary columns for the specific use case.

1. An ordered list of all the actions performed and parameters and scripts used to transform the raw data into the final visualization is available in the MAP PROTOCOL section of each map/plot.