Ideas for data analysis

April 29, 2020

In the airport platform, the situations that generate more interest are the high traffic hours and the non-standard situations such as bad weather conditions.

The idea is to select in our dataset the flights whose taxi time is higher (when the dataset will be available, we'll be able to select flight that have a delay with respect to the scheduled time), and see if the delay is caused by a non-standard situation or only by traffic or, on the contrary, detect the non-standard situations and check if on avarage the taxi time is higher.

0.1 High traffic hours

What is the main cause of delay in taxi time?

- Aircrafts encounter on the taxiway and one of them has to stop, causing a waiting time at the taxiway.

  To know how many time this situation occurs it is necessary to detect it. One possible solution is to compare the variable "Cheminement" of aircrafts that are crossing the taxiway at the same time to see if there are some common roads or some crossing roads.

- Aircrafts have to wait at the gate a long time. To detect this situation the data we have are not sufficient.

- Aircrafts have to wait at the runway. One of the columns of our dataset is "TempsAttentePiste".

- The Aircraft leaves from a runway which is too far from the gate where it is parked, and this situation extends the taxi time.

The idea is to "decompose" the time between when the aircraft leaves the gate (and eventually include also the waiting time at the gate) and when it takes off and vice versa in waiting time at the runway, waiting time in the taxiway and real taxi time to see which one is the major cause of delay during high traffic hours.

0.2 Non-standard situations

The non-standard situation that can cause delay are the following.

- Closed taxiway.
- Closed runway. It can be detected thanks to the creation of a new variable that represents the time interval between two flights that use the same runway. When this time is too long, the runway can be considered closed.
- Bad weather conditions. They can be detected integrating the weather data to our dataset.
- De-icing procedures. There is a variable called "TempsDegivrage" in our dataset. It is necessary to understand how this variable impacts the eventual delay of flights.

Once we have detected the non standard situation, we can estimate the impact that each situation has on the taxi time and on the delay of flights.