

AlphaROC - Data Challenge

Data

The data contains booking records of flight tickets. You will need to understand some airline terminologies in order to understand this dataset. You may find the following webpage useful in explaining the terminologies.

<https://www.t2rl.net/airline/segments#:~:text=A%20Segment%20is%20the%20operation,have%202%20or%20more%20legs.>

Questions

1. Generate some statistics and plots to describe the dataset. Please feel free to clean/preprocess the dataset in any way you think necessary.
2. List the top 20 airports by the number of inbound flights and outbound flights, respectively, in 2019.
3. Draw a graph to show the temporal change of the proportion of
 - a) US domestic flights vs
 - b) international flights departing from the US vs
 - c) international flights arriving to the US
4. What are the changes before and after Covid? How would you quantify the effect of Covid to the airline industry using this dataset? Is there any airline company stood out among all airline companies after Covid?