



FLIXBUS



Case Study Marketing Intelligence



Scenario

- You were given the task to estimate how many tickets will we sell during the next weeks in a given country and through a specific channel.
- You have as **input** the data of company sales during the last year, split in three tables:

Orders_date	
column	content
id	Order identifier
date	Date of the order
channel_id	Channel through which the order was done

Orders_tickets	
column	content
id	Order identifier
n_tickets	Quantity of tickets in the order
type	Type of product

Orders_country	
column	content
id	Order identifier
country_1	First country
country_2	Second country

- You can download the files from:
 - <https://drive.google.com/open?id=1TGSaZxe0uNdGrE1OD2OEyAHolrDaTOnR>

Notes:

- One order can contain several tickets.
- Some orders are assigned to two countries, with country_1 being more important. If there is no value in “country_1”, use “country_2”.

Task

- The **goal** is to build a **script** that takes the information and estimates roughly how many tickets we will sell in the 10 days after the end of the input data.
- The expected **output of the script** is a csv file containing the following columns:
 - date, country, channel, n_tickets
- Please send us a **short presentation** of your process including the following points:
 - How did you do it?
 - Which statistics / criteria / assumptions did you use / make?
 - What difficulties did you find during the process?
 - A screenshot of some part of your code 😊
- Feel free to use the methods / tools / languages you want!
- Also, feel free to make assumptions whenever you think something is not clearly specified.

Note: We want to see your process for approaching such a case, even if the accuracy is not the highest you know you could get ;) Please implement a quick solution and have fun!
