



## Data Analysis Competition (DAC) 2021

### Frequently Asked Questions :

1. In the question sheet stated that there are 10.000 customers, but we have many more numbers of rows? What does it mean?
  - The amount of the customers are based on the unique of the variable "buyer\_id" with the range of ID 0-10.000
2. There are 2 missing variables in "test.csv", what should we do?
  - If you think you find inconsistent information regarding the data test.csv and train.csv, you are required to decide on your own and give your justification why you take such an action
3. Does the number of rows should be same as the sample submission?
  - Sample submission is only an example of submission, the number of rows submission is adjusted to the number of rows variable ID in data test
4. What is the meaning of the variable cnt ?
  - As states in the question sheet, variable cnt is the number of similar events from the user session at that time. Means that user sessions have some repetitions on those actions
5. In the question sheet, we are asked to determine the regency cluster chosen by the customer based on several clusters according to creativity. but in another sentence, there is a statement that in the question a residence cluster has been provided based on price, customer star rating, geographical locations, what do you mean? Is it okay to make a cluster according to our creativity in the regency cluster or use the regency cluster that is already available in the data?
  - Regency clusters are adjusted to the clusters available in the data, the purpose of creativity is how you classify the clusters that have been provided
6. What are the data train and data test for? which data should we process?
  - Data train forms a model which is then applied to the data test in predicting the cluster\_regency variable. it can be seen in the sample submission

For all the participants, just ignore the countdown in the website and the submission deadline is the same as stated in the question sheet at 23.59 PM GMT+7 19<sup>th</sup> July 2021.