

Instructions:

- 1) Background: in order to know our customers in a more holistic way, we would like to put our customers into different segments based on their transaction behaviours.
- 2) Data format:
 - a. Customer and vehicle code: unique identifier of customers and vehicles;
 - b. date_service: the date when the transaction took place;
 - c. value_scaled: scaled value of that transaction.
- 3) Test requirements:
 - a. Please calculate three attributes for each **customer**:
 - i. R (recency): The number of days between the service date and 2020.08.04;
 - ii. F (frequency): the total number of visits of each **customer** in the three-year period;
 - iii. M (monetary): the average amount of value. Note that if there are multiple transactions on the same day for one customer, please add them up as one total transaction for that day.
 - b. Use **k-means** algorithm to perform a customer segmentation using the R, F and M calculated as attributes. You may decide on the number of clusters yourself, with sound rationals.
 - c. Deliverables:
 - i. a sheet with customer code, R, F, M and the segmentation label (which group each customer belongs to) for each customer;
 - ii. a summary of R, F, M for each segment;
 - iii. all the codes you used.