

# Rakamin Batch 19A Final Project -- EDA

## Group 1 Members :

- Yosafat R
- Ridho Fajar
- Li'izza Diana M
- Teguh Tri A
- Vito Rihaldijiran
- M Supian Noor
- Raxy Anggala P

## Dataset : Ecommerce dataset

## Summary for Pre Processing :

- **Determine the Threshold for outliers, whether you want to be discarded or left with a certain threshold**
- **Fill in Missing Values, either by imputation or other methods**
- **Perform one hot encoding or use `pd.get_dummies()` to handle features of type string**
- **if necessary, normalization of features is required before entering modeling**

## Summary for Business :

- Single men who use applications using mobile phones tend to be more likely to churn
- The amount of cashback is very influential on the churn rate. the more cashback given, the more likely the customer to churn. with the average of 150 - 200 cashback amount in 300 - 400 users, most likely the possibility in churn will decrease.
- Payments using debit card cause the most churn while payments via cash on delivery cause the least churn, but COD has a higher proportion between churn and non churn users than debit card. the ratio in Debit card between non churned and churned users is +- 1900 : 260 and the COD is +- 250:150
- based on the correlation heatmap, the value of the complaint has the largest positive correlation to churn, but the largest correlation to churn is tenure but the correlation is negative
- There are 2 factors that influence customers to Churn in terms of customer satisfaction, namely, the first in terms of payment methods, where if accumulated in total, customers with payment methods using Debit Cards have the largest churn quantity. However, if only grouped by type of payment method, customers with COD payment methods have the greatest churn potential, which is 28%. So that a review and improvement on the COD payment system is needed. In addition to payment methods, the churn rate for customers who submit complaints is quite high, above 31%.

- The churn trend often occurs in new customers, where customers churn with a ratio above 50% for the data value of Tenure  $\leq 1$ .
- The more devices registered on one Customer ID, the higher the potential for churn.
- As input, improvement in customer service is needed, because many customers who have just made orders, churn.
- There is a possibility that customers are dissatisfied with the cashback program provided, so they churn.
- There is no Churned Users when the tenure is equal or above 22