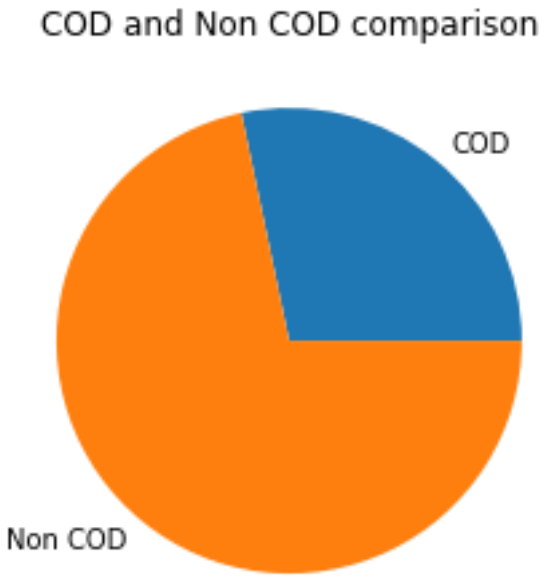
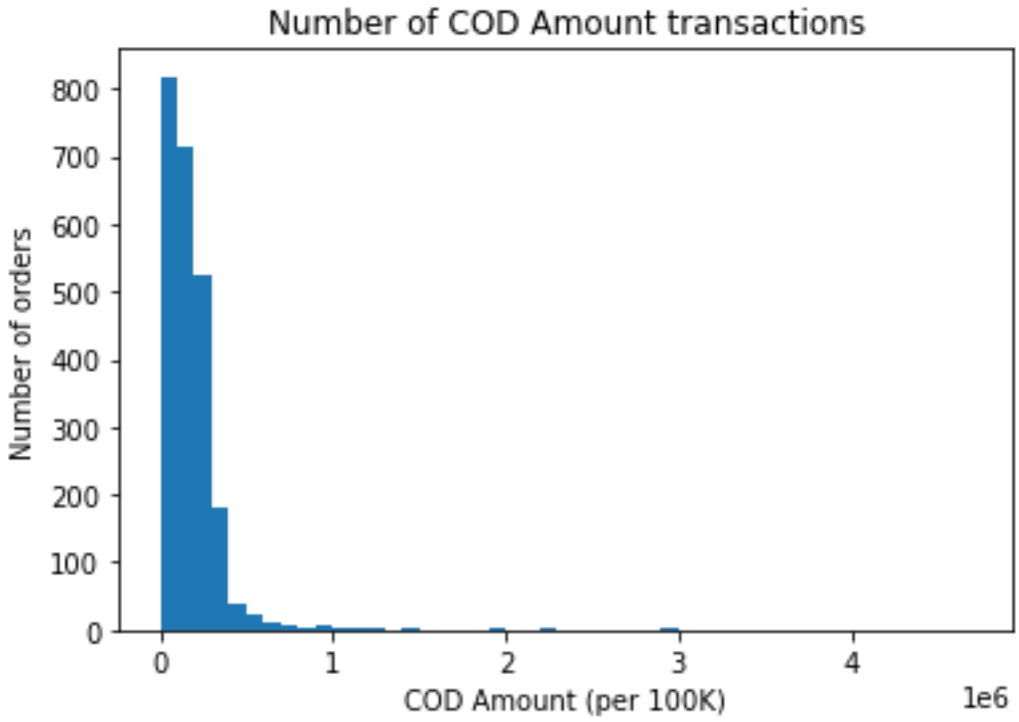
COD Analysis

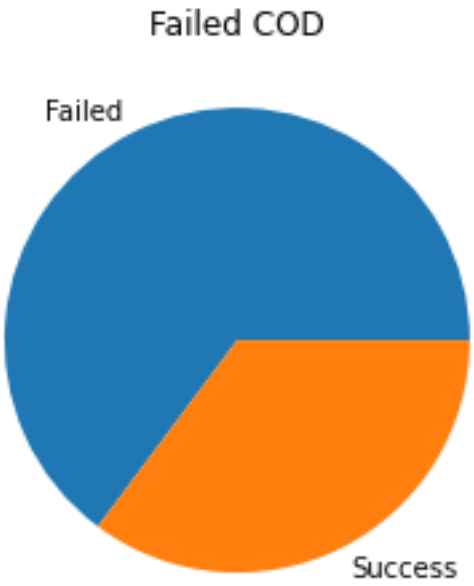
COD is an interesting part to be investigated because it can be an opportunity to increase our user base, or it can be our downfall because its payment is risky for us.



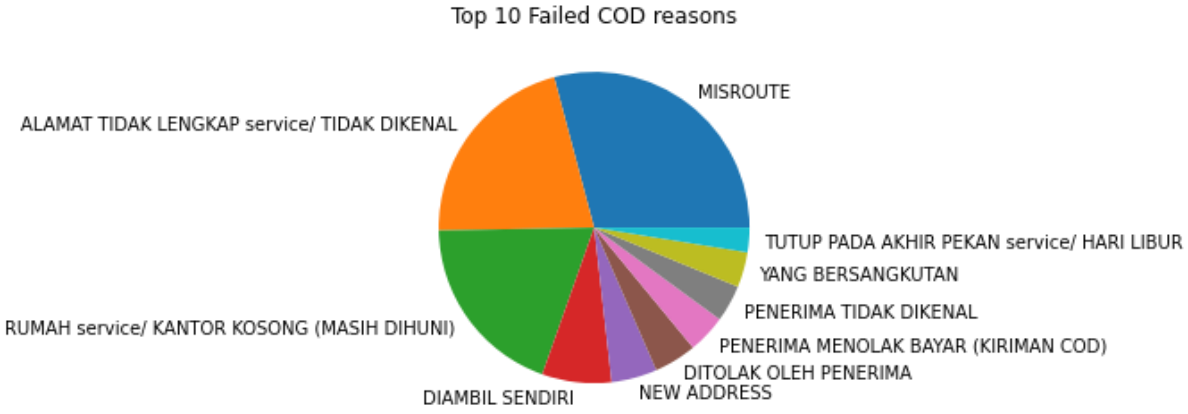
As we can see above, our COD user number is less than non-COD. We might want to increase our COD user base and reach those customers who might prefer COD transactions or cannot use non-COD transactions.



Most of our COD customers paid less than 300K. To gain more profit, we can add promotion for those who have transaction more than 300K. This hopefully will attract them to add more when they are at the border between 290K and 300K.



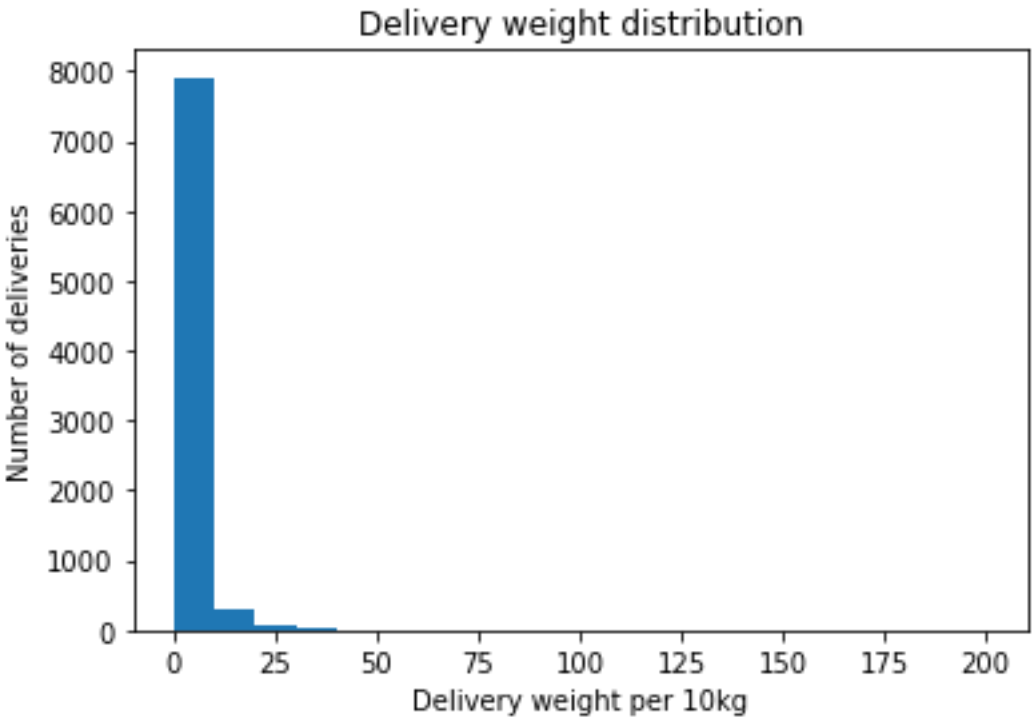
But there's a problem. Most of our COD transactions are failed. Why?



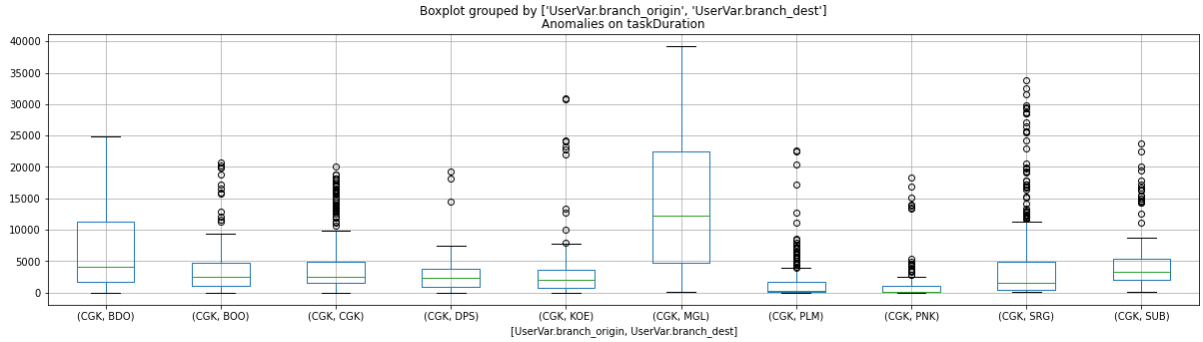
Some of our failed COD caused by the customer itself, there's nothing we can do about it. But the highest number of reasons was caused by misroute. This is something that we should solve. We should improve our routing quality so that our courier will be less likely to be misrouted.

Delivery Analysis

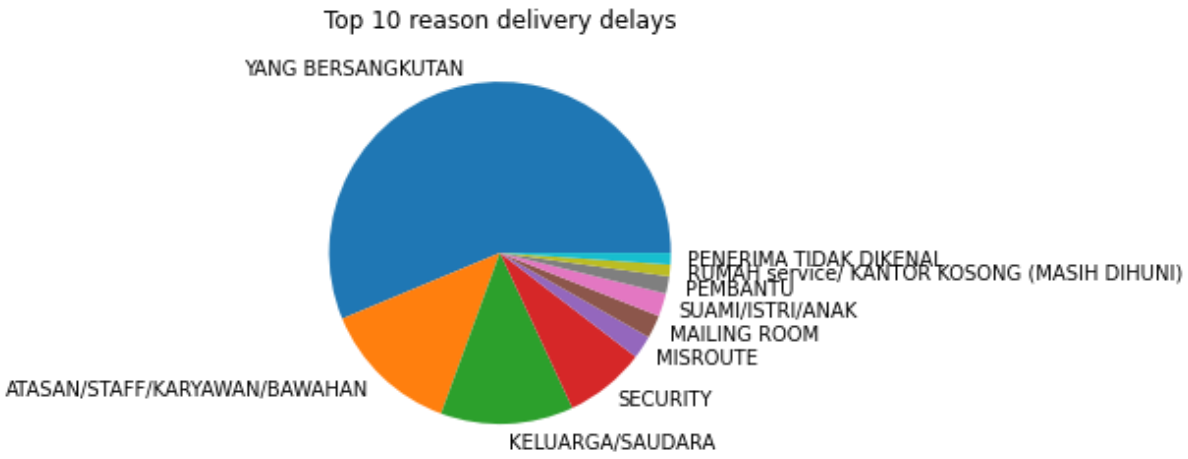
We also need to look at our delivery situation. We might be able to improve our delivery process or prevent any problem in the future.



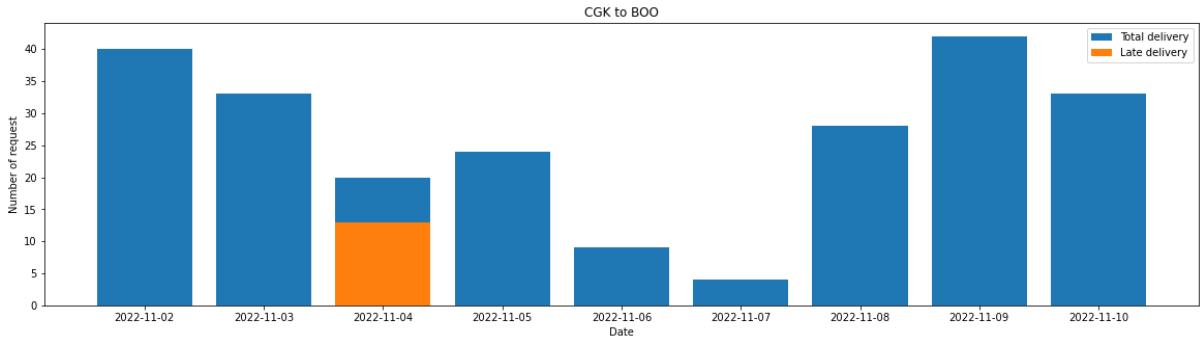
Most of our delivery on these 10 days are lightweighted (less than 10kg) which was delivered by (assumption) bike. We should expect those bikes need some maintenance soon to keep up with the deliveries.

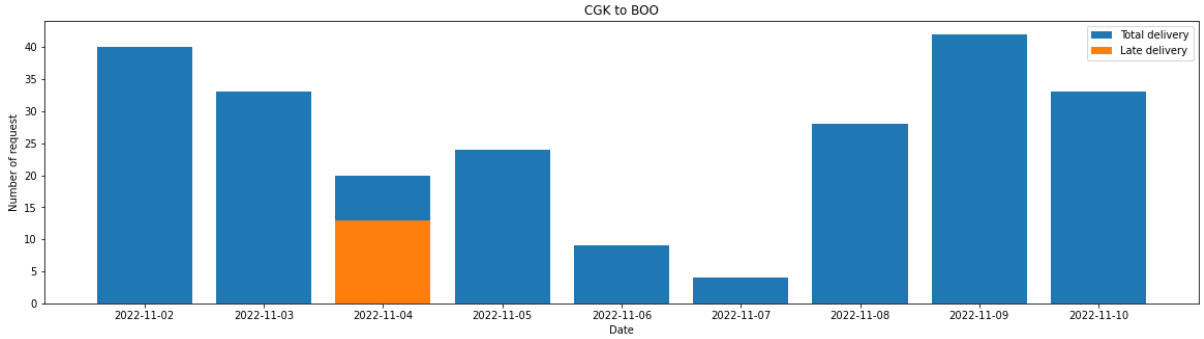


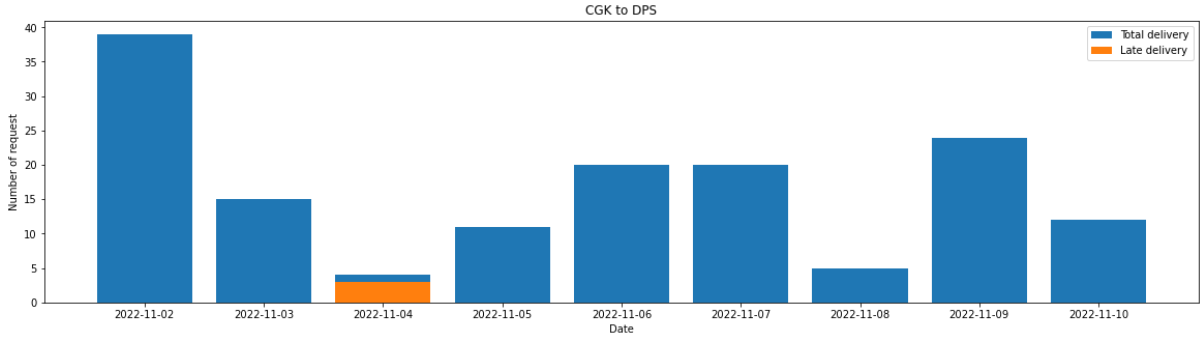
If we look at completed delivery, some of them have “abnormal” delivery duration. Among those abnormal delivery, all of them took longer time than “normally” it should be.

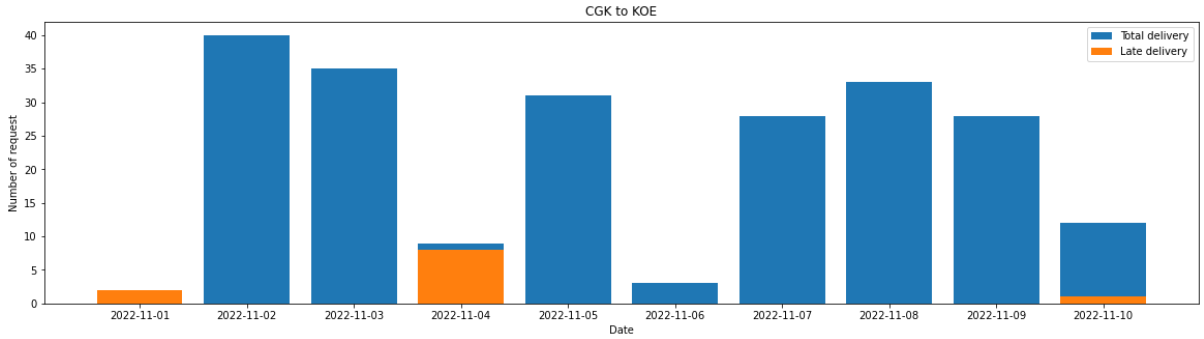


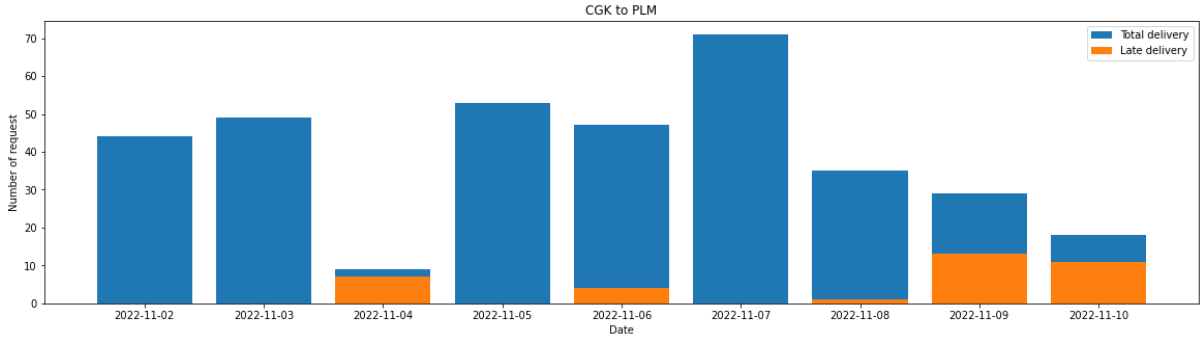
Among those delays, it didn't have any problem except a few misroutes and rumah/kantor kosong. So, what is probably the cause?

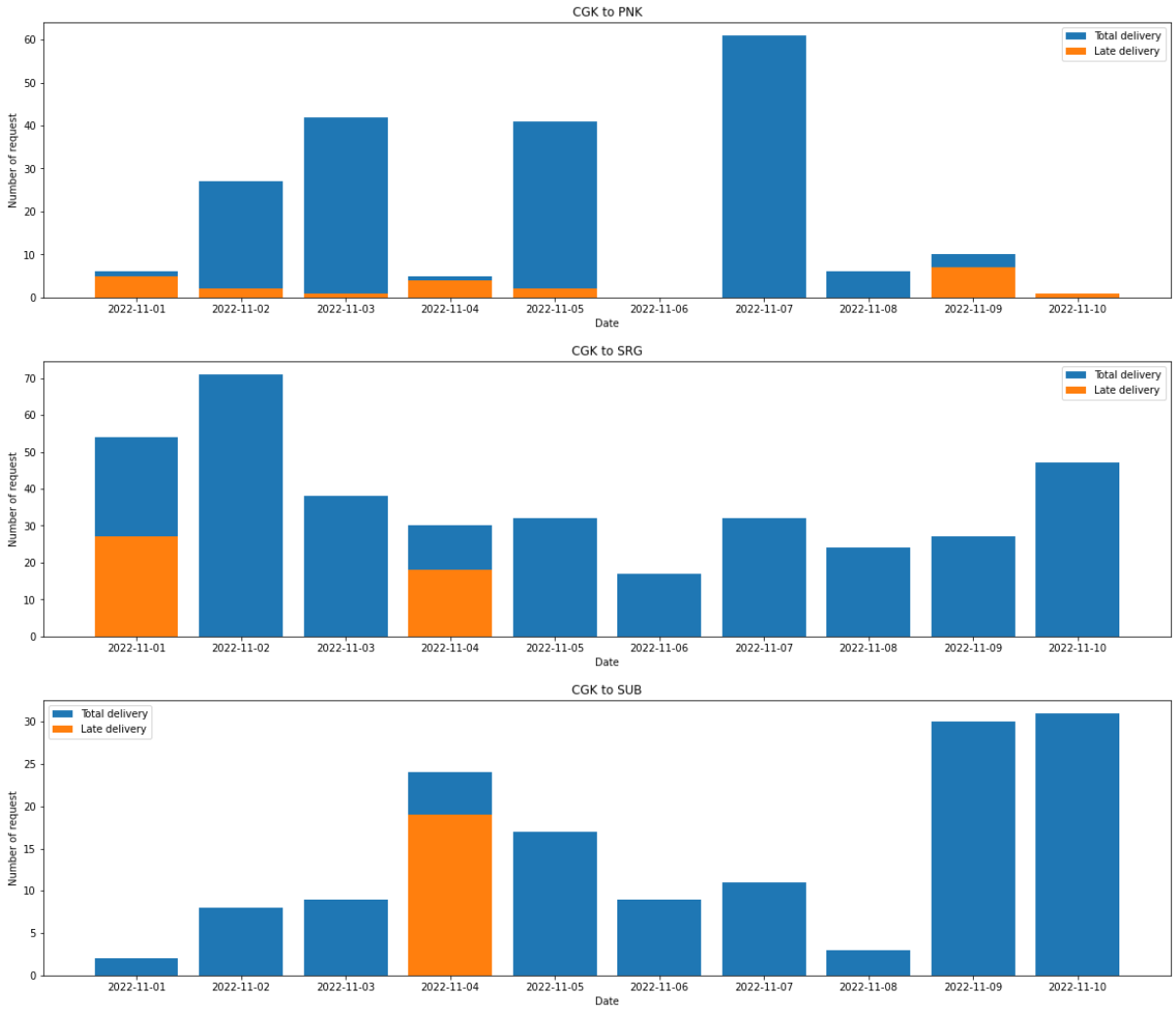








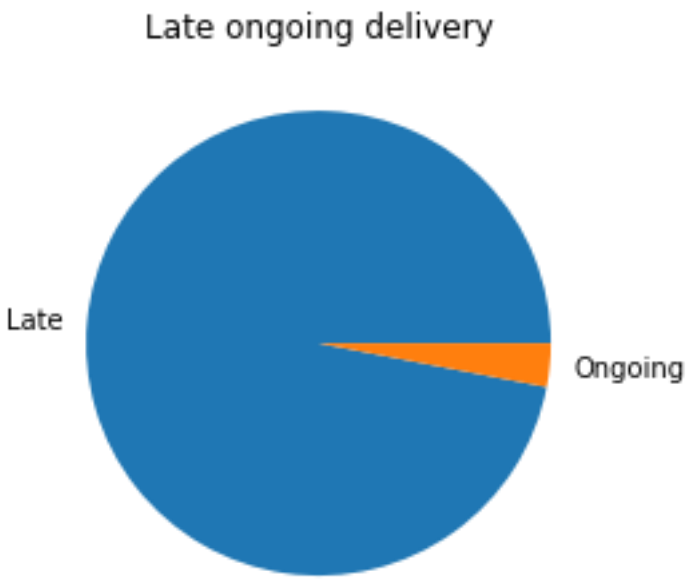




Most of the failed deliveries were requested on 22-11-04. Coincidentally, number order on that day were also the most. Those delays might be caused by overload on that day on CGK branch. We might want to look back at that day, whether there's any event or something happened that makes those number of requests.

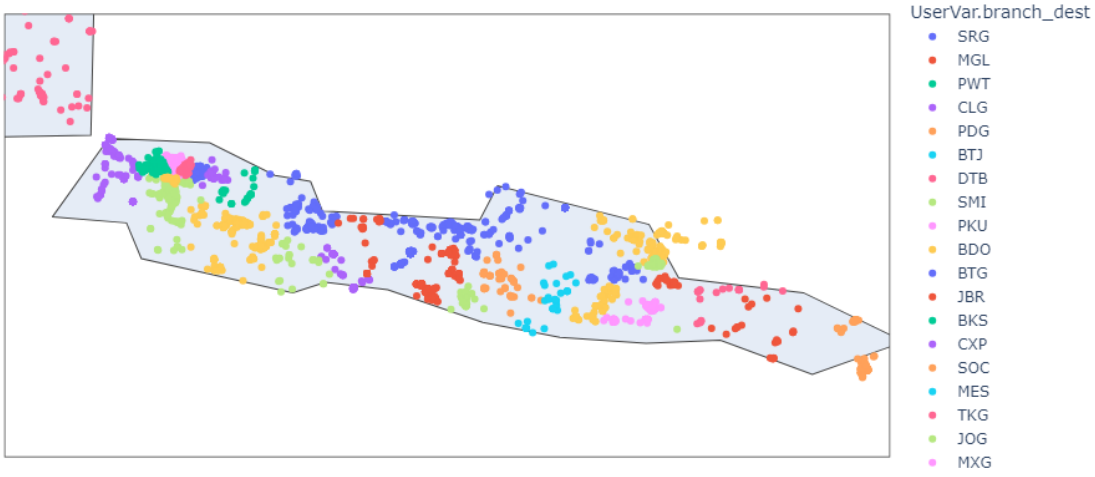
Delivery estimation time

We can train model to predict task duration (from task created time until task completed time) to estimate when delivery supposed to be done. We can train the model using "done" taskStatus records and estimate "ongoing" data to evaluate how many deliveries that already late. I used some "done" records for training and few for validation. I used "taskCreatedDay", "taskAssignedTo", "cod", "branch_dest", "receiver_city", "weight", and "branch_origin" as the feature for training. Those which are text type data like branch, city, and assignee were encoded using word2vec. Cod only used true or false value for whether it is cod transaction or not. I evaluated four models: LightGBM, XGBoost, Random Forest, and MLP using MSE as metric. LightGBM had better score and was used for predict "ongoing" task duration.



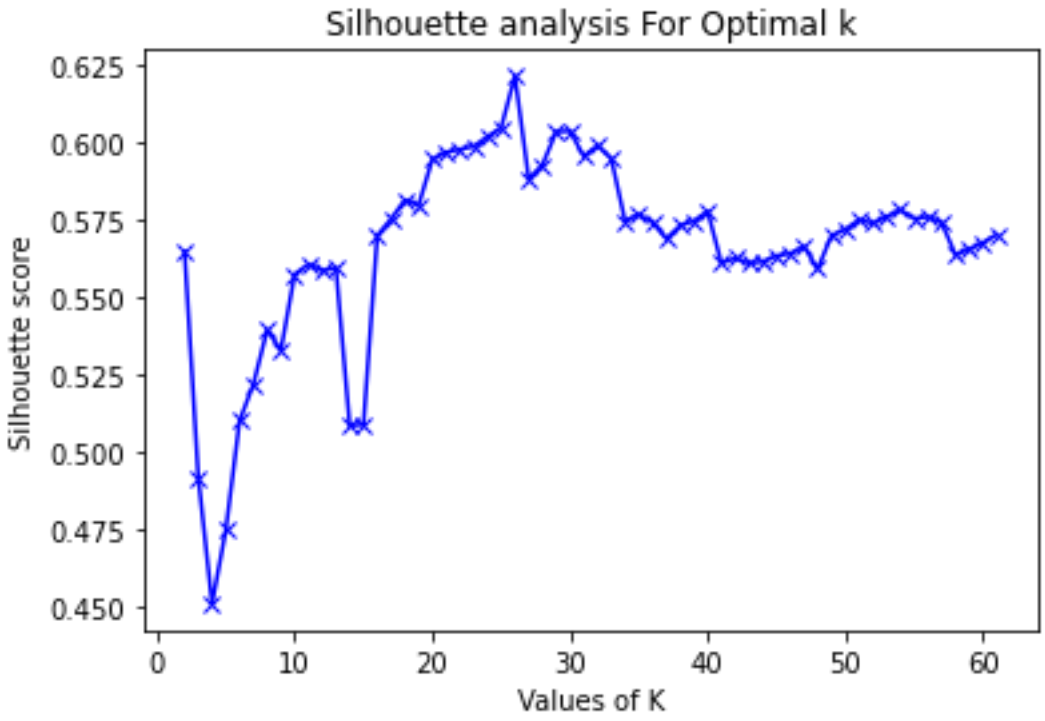
Most of our ongoing delivery are already late (given current time is latest time on taskCreatedTime and taskCompletedTime). We should boost our delivery speed soon.

New Branch Segmentation Proposal

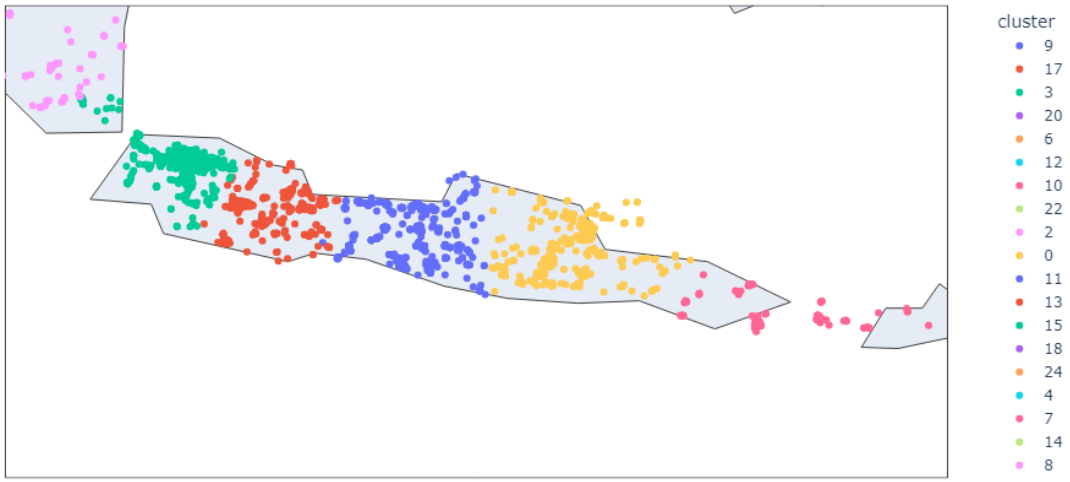


We have 60-ish number of branches currently. The coverage area of each branch is also not clear enough. Those numbers of branches also make more operation cost needed. Based on receiver location (taskLocationDone) we can cluster them geographically and found better segmentation.

I used KMeans to cluster the location latitude and longitude. I used silhouette analysis to determine optimal number of clusters.



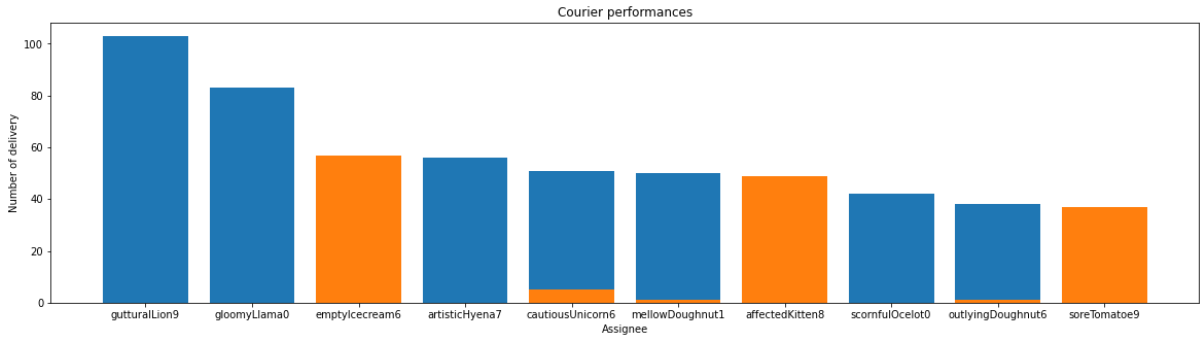
We found that 26 clusters are the optimal number for our KMeans. I clustered the location and proposed new 26 branch segmentation.



This segmentation might need some adjustment based on transportation limitations like cross island delivery.

Courier Performance

Some couriers completed more deliveries than others. But whether it is a successful delivery or not we should evaluate it first.



Three of them have high successful rate with their high number of deliveries. We should reward them accordingly to increase their morale. But some others always failed despite their high number of deliveries. We should keep an eye on them and take appropriate action.