# DSI 03

Capstone Project
Partner Insight: Broadband new acquisition

# Background

Partner : Personal selling is a face-to-face selling technique by which a salesperson uses his or her interpersonal skills to persuade a customer in buying a particular product. The salesperson tries to highlight various features of the product to convince the customer that it will only add value. However, getting a customer to buy a product is not the motive behind personal selling every time. Often companies try to follow this approach with customers to make them aware of a new product.

Partner sales target : End Customer, SME (Small and medium-sized enterprises)

Partner Benefit : Salary, Commission

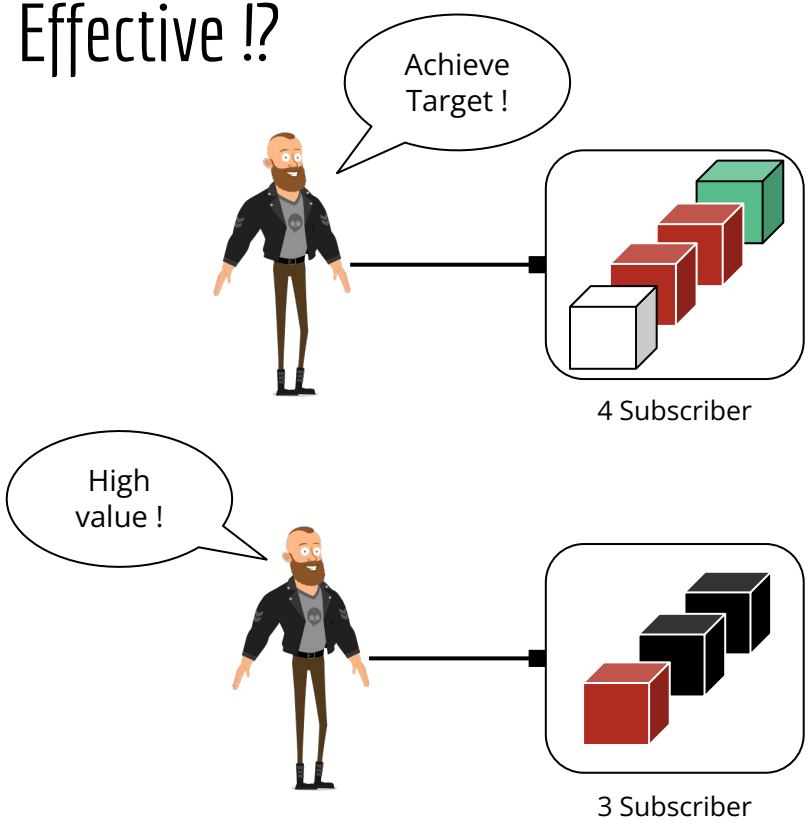
# Problem Statement

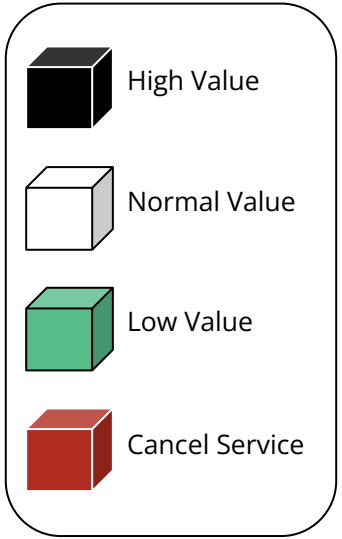
The XYZ Company is a communications satellite service company with thousands of regional sales teams. Each month, partner leave the area to sell products to customers at various locations at all times

Because there are thousands of active partner “Effective” cluster of partner and not effective based on measurement when achieving target. Therefore may not be enough

Company wants to find a way to divide the cluster of partners. For purpose maintain good partners and focus on  ineffective partner.

# Partner Effective !?



# Data availability and Limitation

## availability

1. Partner Information
2. Information successful sales rate per partner
3. Sales revenue (Based on the price of the package) that has not yet been issued a bill
4. Information for cancellation of services that comes from the partner's offering

## Not availability

1. Customer LifeTime / Customer Value
2. Partner complaint

## limitation

1. Data compensation and salary
2. Cost per partner condition is not clear (hidden cost)

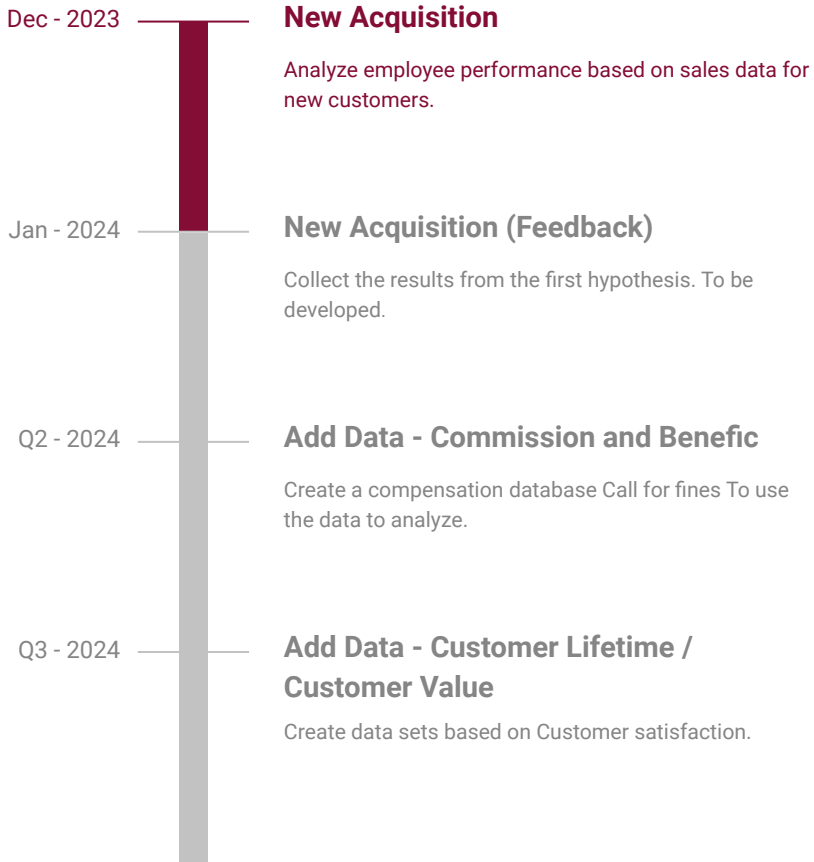


Update: 20/12/2023

# Project Vision

Due to limitations and the availability of information, team will divide project into phase.

For in this phase we focus on New acquisition data for proof of concept




Update: 21/12/2023

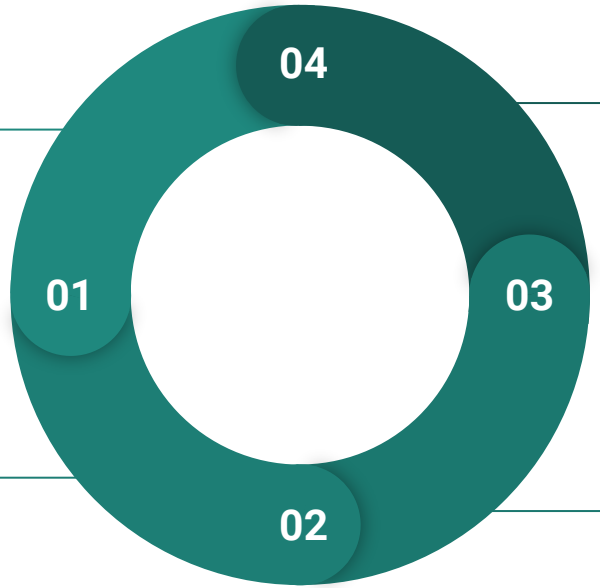
# Develop Process

### EDA

Data analysis.

### Feature Engineer

Create variables and functions.




### Interpret Result

Translate the grouping results from the acquired model. To use Improve data sets and models.

### Modeling

Use clustering model and check result

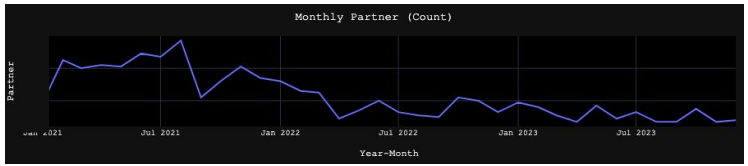
# EDA (Exploratory data analysis)

Study the nature and format of data to create models.

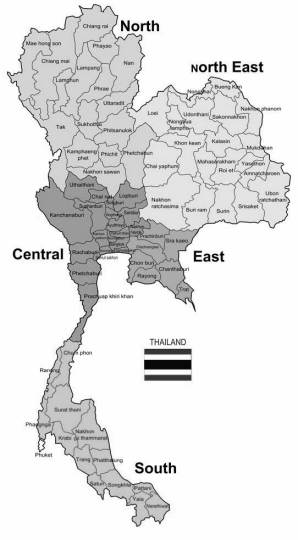
The data obtained is from January 2021 to the current month (December 2023).

Data includes number of sales information. Total revenue, total service cancellation and Partner details


Performance from New acquisition in each month




Number of partner each month




Partner's workspace

# Feature Engineer

The information obtained can be used at all, the Feature Engineer process is only to delete some partner data and create a feature that is about time.

Feature Engineer

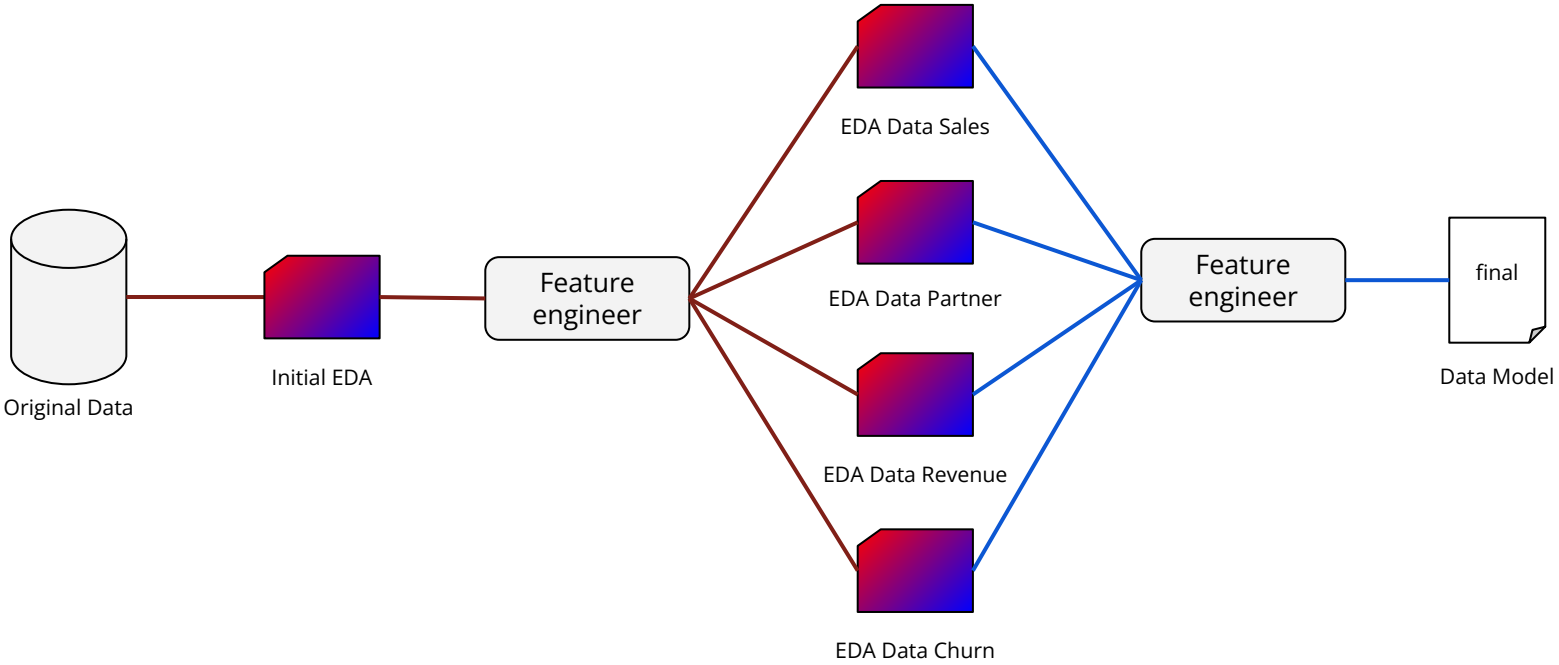
- Eliminate sales data sold by AE (Account executive).
- Add a create set of function about date.

➔ *AE are a group of salespeople who sell products to departments, organizations, and large companies.*

➔ *Number of sale from AE is less but revenue is more.*

➔ *SMB: Small and Midsize Business.*

# Data Processing



Dataset: Partner

| Column | Description |
|---|---|
| partner_code | Partner code registered with the company |
| partner_status | Status of partner : Active, Inactive |
| partner_aging | Working age of employees (unit : Day) |
| partner_ageing_mth | Working age of employees (unit : Month) |
| partner_ageing_yr | Working age of employees (unit : Year) |
| total_revenue | Number of revenue from product sold per partner |
| total_churn | The number of customers who cancel the service from the goods that the partner sells |

Dataset: Sales

| Column | Description |
|---|---|
| sale_rate_m4 | The number of products sold in month per partner back 4 months (e.g. current December back 4 month = October) |
| sale_rate_m3 | The number of products sold in month per partner back 3 months (e.g. current December back 4 month = August) |
| sale_rate_m2 | The number of products sold in month per partner back 2 months (e.g. current December back 4 month = September) |
| sale_rate_m1 | The number of products sold in month per partner back 1 months (e.g. current December back 4 month = November) |
| sale_flag_m4 | Flag True or False (True = have sold in that month, False = not sold in that month) |
| sale_flag_m3 | |
| sale_flag_m2 | |
| sale_flag_m1 | |
| sales_month | Total number of month sold per partner |
| sales_consecutive | Number of consecutive sales months per partner |

# Dataset for Modeling : Partner and Sales

Dataset: Revenue

| Column | Description |
|---|---|
| rev_m4 | The number of revenue from product sold per partner back 4 months (e.g. current December back 4 month = October) |
| rev_m3 | The number of revenue from product sold per partner back 3 months (e.g. current December back 4 month = August) |
| rev_m2 | The number of revenue from product sold per partner back 2 months (e.g. current December back 4 month = September) |
| rev_m1 | The number of revenue from product sold per partner back 1 months (e.g. current December back 4 month = November) |
| rev_flag_m4 | Flag True or False (True = have sold in that month, False = not sold in that month) |
| rev_flag_m3 | |
| rev_flag_m2 | |
| rev_flag_m1 | |
| rev_month | Total revenue of month sold per partner |
| rev_consecutive | Number of consecutive revenue from sold per partner |

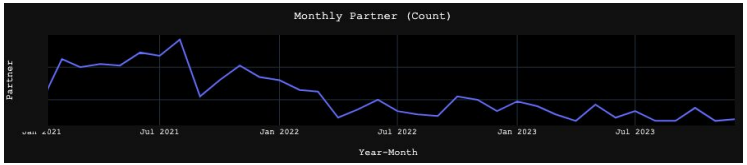
Dataset: Churn

| Column | Description |
|---|---|
| churn_m4 | The number of customers who cancel the service per partner back 4 months (e.g. current December back 4 month = October) |
| churn_m3 | The number of customers who cancel the service per partner back 3 months (e.g. current December back 4 month = August) |
| churn_m2 | The number of customers who cancel the service per partner back 2 months (e.g. current December back 4 month = September) |
| churn_m1 | The number of customers who cancel the service per partner back 1 months (e.g. current December back 4 month = November) |
| churn_flag_m4 | Flag True or False (True = have sold in that month, False = not sold in that month) |
| churn_flag_m3 | |
| churn_flag_m2 | |
| churn_flag_m1 | |
| churn_month | Total customers cancel per partner |
| churn_consecutive | Number of consecutive customer cancel per partner |

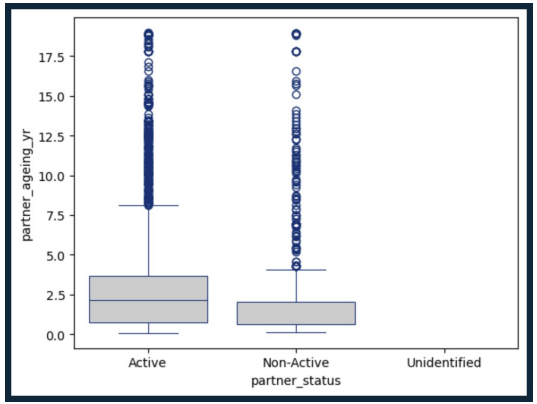
# Dataset for Modeling : Revenue and Churn

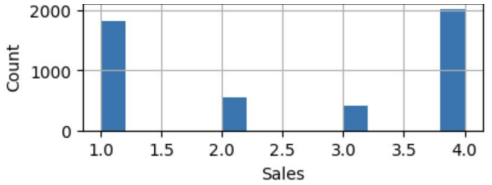
# Second EDA


Performance from New acquisition in each month




Number of partner each month




| | partner_code | churn_m4 | churn_m3 | churn_m2 | churn_m1 |
|---|---|---|---|---|---|
| 0 | 01009075 | NaN | 0.0 | 0.0 | 0.0 |
| 1 | 01010104 | NaN | 0.0 | NaN | NaN |
| 2 | 01011451 | 0.0 | 0.0 | NaN | 0.0 |
| 3 | 01011709 | NaN | NaN | NaN | 0.0 |
| 4 | 01015170 | 0.0 | 0.0 | 0.0 | NaN |

# Final Data Model

| Data Set 1 | Description |
|---|---|
| partner_code | Partner code registered with the company |
| partner_ageing_yr | Working age of employees (unit : Year) |
| sales_consecutive | Number of consecutive sales months per partner |
| rev_consecutive | Number of consecutive revenue from sold per partner |
| churn_consecutive | Number of consecutive customer cancel per partner |

| partner_code | partner_ageing_yr | sales_month | sales_consecutive | rev_month | rev_consecutive | churn_month | churn_consecutive |
|---|---|---|---|---|---|---|---|
| 39089437 | 2 | 4 | 4 | 4 | 4 | 0 | 0 |
| 80101224 | 2 | 2 | 2 | 2 | 2 | 0 | 0 |
| 39096384 | 2 | 4 | 4 | 4 | 4 | 0 | 0 |
| 39097318 | 2 | 4 | 4 | 4 | 4 | 1 | 1 |
| 39094136 | 2 | 4 | 4 | 4 | 4 | 0 | 0 |

Contain of 1,690 partner and 7 feature for clustering

# Final Data Model

| Data Set 1 | Description |
|---|---|
| partner_ageing_yr | Working age of employees (unit : Year) |
| sales_consecutive | Number of consecutive sales months per partner |
| rev_consecutive | Number of consecutive revenue from sold per partner |
| churn_consecutive | Number of consecutive customer cancel per partner |

| Data Set 2 | Description |
|---|---|
| partner_ageing_mth | Working age of employees (unit : Month) |
| sale_rate_m4 | The number of products sold in month per partner back 4 months (e.g. current December back 4 month = October) |
| sale_rate_m3 | The number of products sold in month per partner back 3 months (e.g. current December back 4 month = August) |
| sale_rate_m2 | The number of products sold in month per partner back 2 months (e.g. current December back 4 month = September) |
| sale_rate_m1 | The number of products sold in month per partner back 1 months (e.g. current December back 4 month = November) |
| rev_m4 | The number of revenue from product sold per partner back 4 months (e.g. current December back 4 month = October) |
| rev_m3 | The number of revenue from product sold per partner back 3 months (e.g. current December back 4 month = August) |
| rev_m2 | The number of revenue from product sold per partner back 2 months (e.g. current December back 4 month = September) |
| rev_m1 | The number of revenue from product sold per partner back 1 months (e.g. current December back 4 month = November) |
| churn_m4 | The number of customers who cancel the service per partner back 4 months (e.g. current December back 4 month = October) |
| churn_m3 | The number of customers who cancel the service per partner back 3 months (e.g. current December back 4 month = August) |
| churn_m2 | The number of customers who cancel the service per partner back 2 months (e.g. current December back 4 month = September) |
| churn_m1 | The number of customers who cancel the service per partner back 1 months (e.g. current December back 4 month = November) |

# Modeling

# K Mean

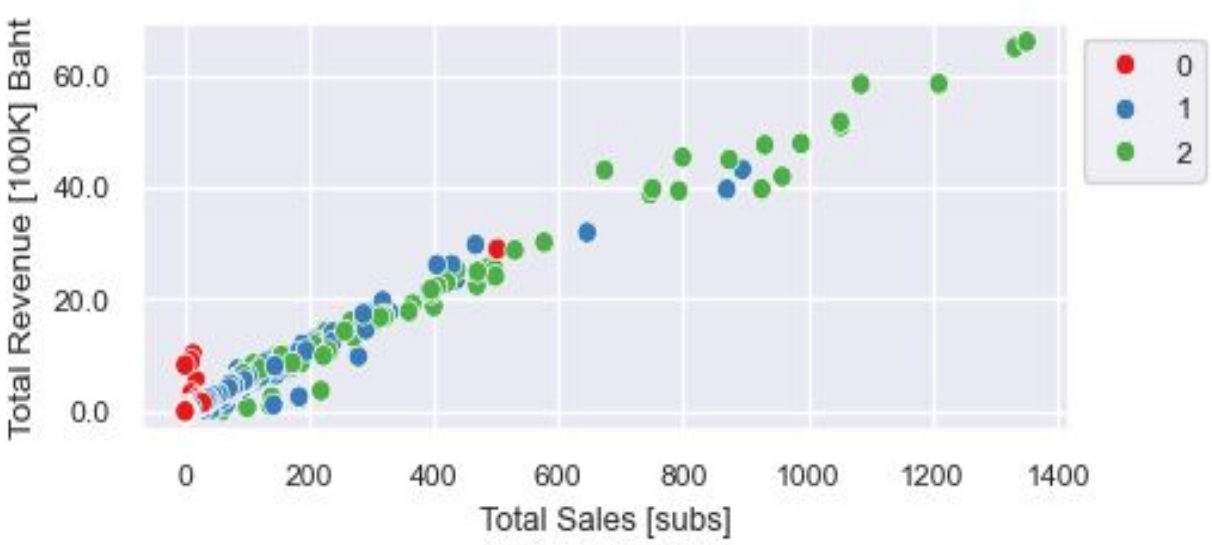
The concentration density is at the sales range from 1 to 550 subscriber.

The model can be divided into 3 cluster of partner, with the following amounts

Cluster 0 : 60%

Cluster 1 : 32%

Cluster 2 : 0.8%

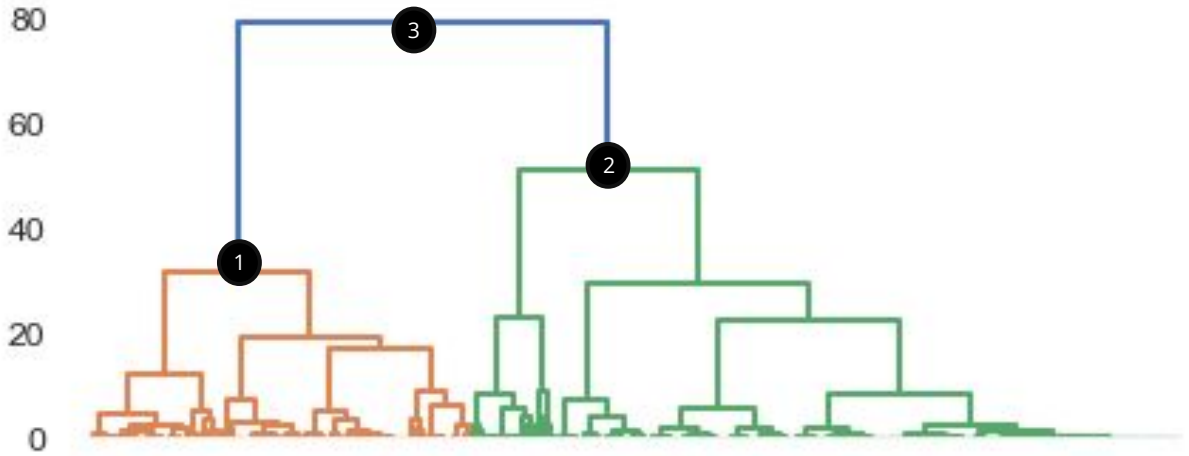



*Silhouette Score: 0.620*

*! label not defined yet*

# Agglomerative

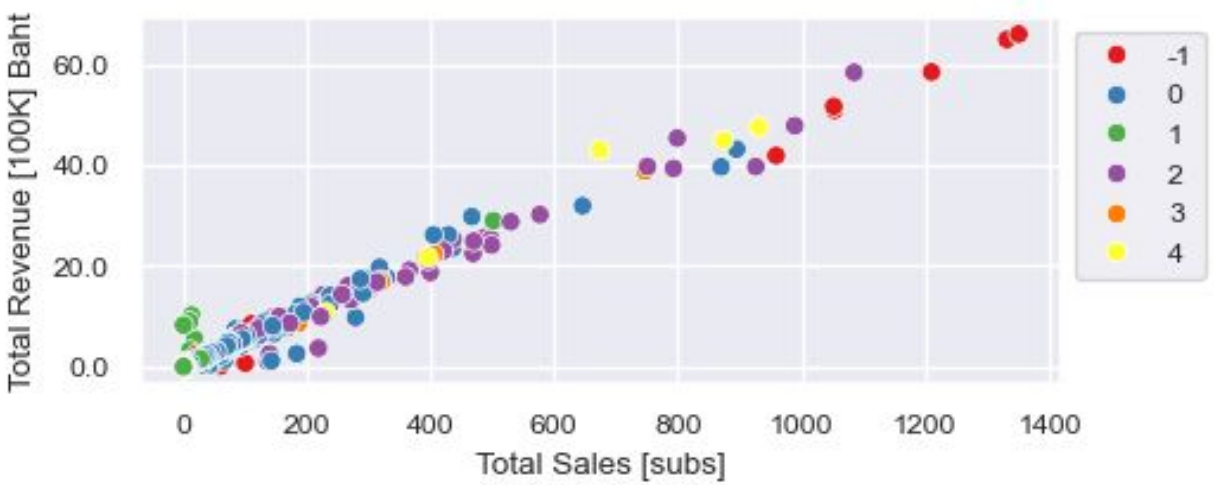
From the model, can separate partner into 3 cluster, with the same result as the K Mean model.



# DBSCAN

The DBSCAN model clearly differs from the two models above. It can divide up to 6 cluster of partners, which 1 outlier.

| Cluster | Percent |
|---|---|
| -1 | 0.8% |
| 0 | 60% |
| 1 | 31% |
| 2 | 0.5% |
| 3 | 0.5% |
| 4 | 0.3% |




*Silhouette Score: 0.610*

*! label not defined yet*

# DBSCAN

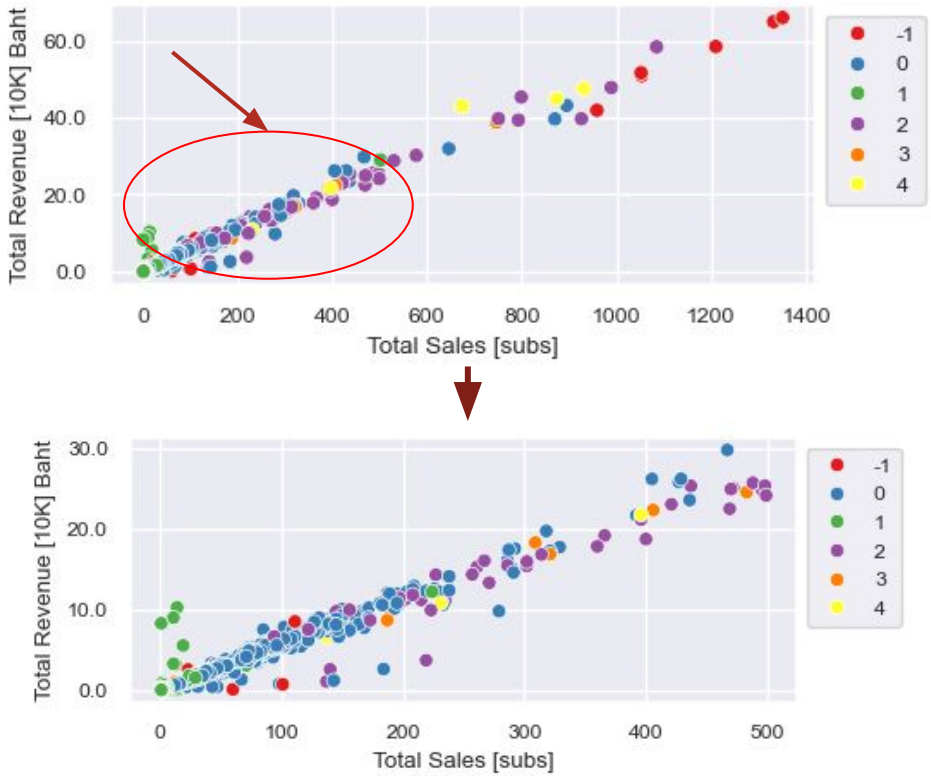
The model can divide the groups of partners in more detail than other

# Result

As a result of the model, we find that most partner are concentrated in sales ranges from 1 to 500 subs,

And considering deep down, we will see the density of partner and cluster clearly.

Which lead to the action plan.



# Recommendation and Action Plan

**Recommend.**

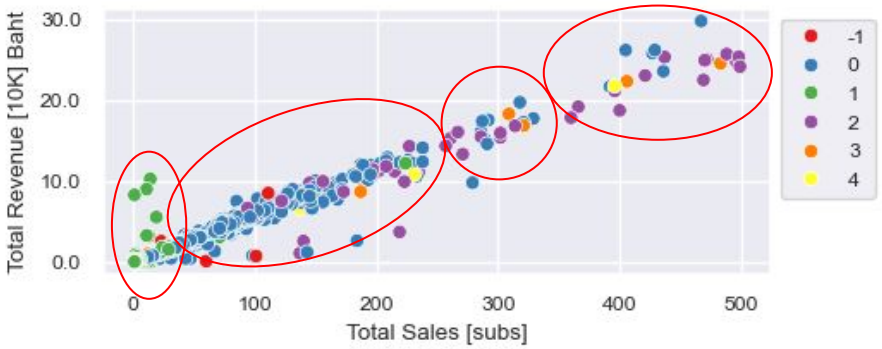
From the density in previous slide.

The team has divided the sales range into 4 parts, each of which consists of a cluster of partner in different levels, That make company easy to focus.

**Action Plan.**

Partner cluster 1: Has a small sales volume but products are high price. Companies should push this group to increase sales rate.

Partner Cluster 0: Has a large sales volume but low price and is a large group. Companies could help to find a package or improve the this group to sold high value



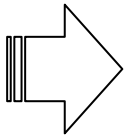

| Cluster | Label |
|---|---|
| **-1** | Sales Less, Low Revenue+, With Churn |
| **0** | Sales More, Low Revenue, No Churn |
| **1** | Sales Less, High Revenue, No Churn |
| **2** | Sales More, Low Revenue, With Churn |
| **3** | Sales Less, Low Revenue, With Churn |
| **4** | Sales Less, Low Revenue, With Churn |

*Sample label (subject may be change)*

# Model Improve Plan

Quick Win: Develop additional processes to determine the characteristics of each cluster. Another improve follow Project vision

| Cluster | Partner Aging | total_sales | total_revenue | total_churn |
|---|---|---|---|---|
| -1 | 3 | 7,000 | 3,000,000 | 20 |
| 0 | 2 | 60,000 | 41,111,000 | - |
| 1 | 2 | 1,000 | 1,000,000 | - |
| 2 | 2 | 19,000 | 10,100,200 | 100 |
| 3 | 2 | 2,000 | 1,011,010 | 10 |
| 4 | 2 | 4,000 | 1,222,000 | 10 |



| Cluster | Label |
|---|---|
| **-1** | Sales Less, Low Revenue+, With Churn |
| **0** | Sales More, Low Revenue, No Churn |
| **1** | Sales Less, High Revenue, No Churn |
| **2** | Sales More, Low Revenue, With Churn |
| **3** | Sales Less, Low Revenue, With Churn |
| **4** | Sales Less, Low Revenue, With Churn |

*Sample label (subject may be change)*

# Thank You

# Enhance