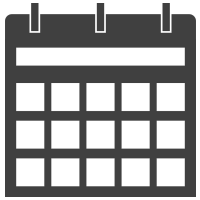
Thematic Campaign Recommendation Based on User Preferences

With Python

By Mega Oceanna

Executive Summary

1



Business Problem
From monthly data (Aug to Sep 2021), we have 7073 users. To make a targeted campaign, customer segmentation is needed to understand the best thematic campaign based on user preferences so we will gain profit due to right targeted campaign.

2

User Analysis
We have total transaction amount up to Rp 21 Billion with Bond Funds as the highest transaction amount (Rp 7.8 Million). However, most of our users are Sink Users (5953 users from 7073) that have lowest invested amount.

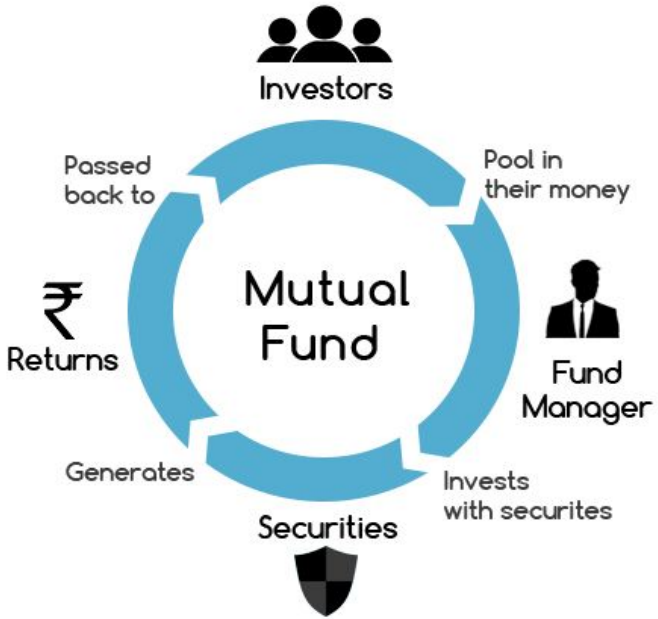
3

Expected Return Analysis
We can gain up to Rp 11 Million expected return if the users transacted with campaign. However, there are some customers who are predicted to be churn and it is apprehensive that this could result in a negative expected return.

1) Business Understanding

We work as a Data Analyst in Indonesian investment start up company which makes an application that users can buy and sell mutual funds.

Cycle of Mutual Funds



Source

PRODUCTS

1) Money Market Funds
Low risk, low return, fit to short term savings (1 year), fit to people with conservative profile.

2) Bond funds
Higher risk and return than money market, fit to medium term savings (3-5 years), fit to people with conservative profile.

3) Target Date Funds
Mix of stocks, bonds, and other investments, fit to medium term savings (3-5 years) and moderat profile.

4) Stock Funds
Highest risk and return, fit to long term savings (more than 5 years), fit to people with aggressive profile.

Source: ojk.go.id & investor.gov of USA

2) *Business Problem and Analytical Objectives*

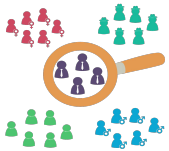
We work as a Data Analyst in Indonesian investment start up company which makes an application that user can buy and sell mutual funds.

Business Problem and Objective

To make a targeted campaign next month for the Marketing Team, we will help them to define the best thematic campaign based on customer segmentation analysis, so we can help the Marketing Team to make the next month campaign runs smoothly, optimize our budget due to right targeted campaign, reduce the loss, and increase our profit. To make it more targeted, we must list down top 30% users of each segment for budget efficiency.

What we need to do

Analyze descriptive statistics to understand the demographic and the trend of each mutual funds



Customer segmentation with k-means clustering to understand which cluster that each user belongs to and give an insight and thematic campaign recommendation based on segmentation

List down top 30% users of each cluster and understand how profitable of each campaign with regression and benefit cost analysis

3) *Analysis Flow*

Summary: We have 127404 rows and 26 columns of cleaned data (transaction data merge with users data. **Click here to access the python syntax in google colab**

01 Business understanding

Understand the business (mutual funds) concept and the objective of the project.

02 Data Preparation (no duplicated and typo values found in our data)

- **Import the data** of transaction (transaction history of each users in August and September 2021), user (user demographic), and churn. After merge with INNER, our users have transaction on 4 August to 30 September registered on 1 August to 28 September 2021.
- **Remove irrelevant data**: total buy and sell transaction will be deleted and use total buy and sell amount from users data.
- **Convert data type:** date and register date to datetime, add age category column (Youth 17-24yo, Adult 25-65yo, Senior >65yo).
- **Missing value**: fill null in amount with 0 (means no transaction). Null in referral code means registered without referral, replace with no referral code.
- **Merge dataset** of transaction and user with INNER to see daily transaction of each users.
- **Remove outlier**: detect in invested amount, total buy, and total sell with exclude 0. Decide to remove outlier since it will affect our K-Means analysis.

03 Exploratory Data Analysis (EDA)

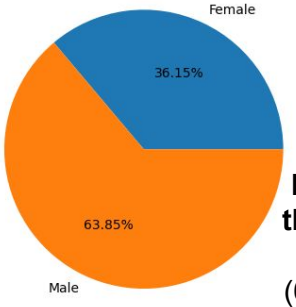
- **Descriptive statistics** to understand the trend
- **Customer segmentation** with K-Means Clustering (unsupervised learning, find the nearest neighbors)
- **Regression analysis** to find top 30% users of each segment

04 Insight and Recommendation

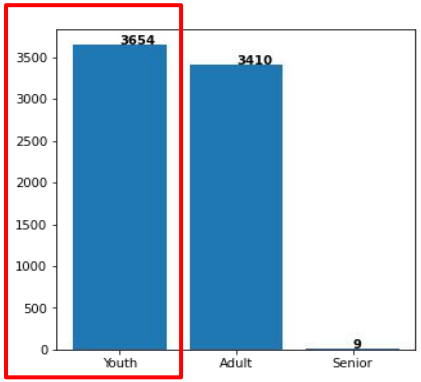
4) *Exploratory Data Analysis (EDA)*

Understand the demographic of the data (python link)

There are 7073 users in August and September 2021

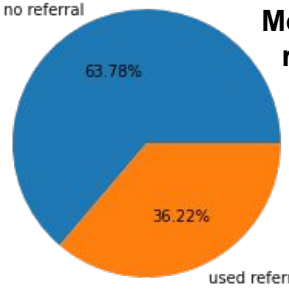


Most of them are Male (63.85%)

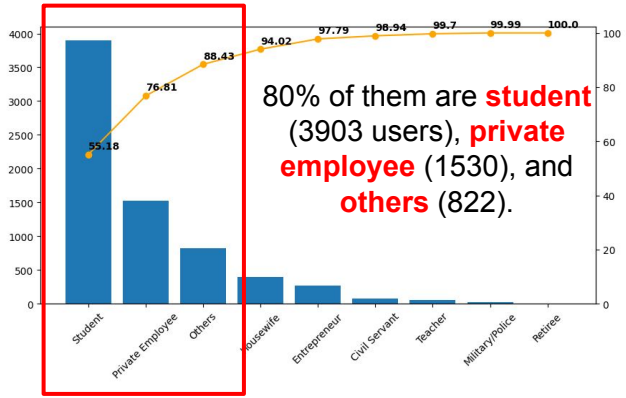


user_age	
count	127404.00
mean	26.48
std	8.19
min	17.00
25%	21.00
50%	24.00
75%	30.00
max	74.00

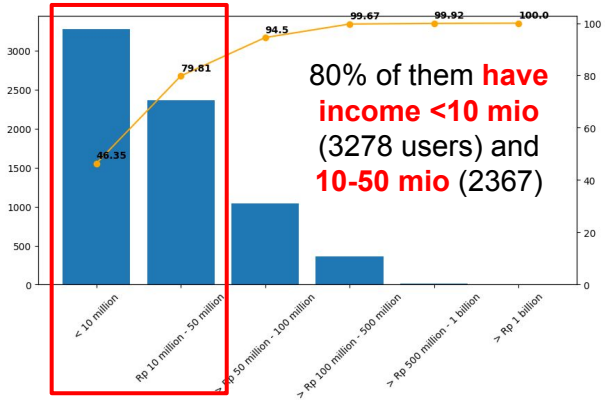
Most of them are Youth 17-24 yo (3654 users). The youngest user is 17 yo and oldest is 74 yo



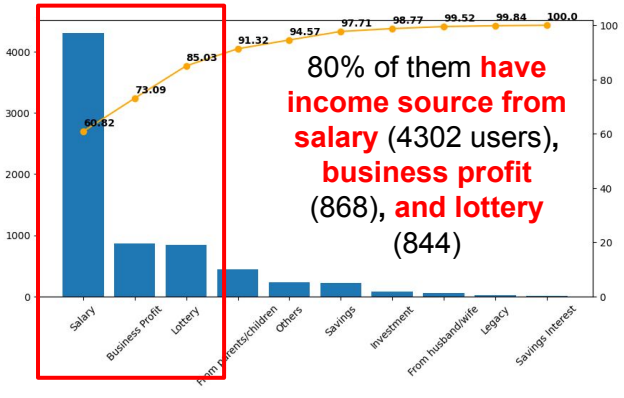
Most of them registered with no referral (63.78%)



80% of them are **student** (3903 users), **private employee** (1530), and **others** (822).



80% of them **have income <10 mio** (3278 users) and **10-50 mio** (2367)

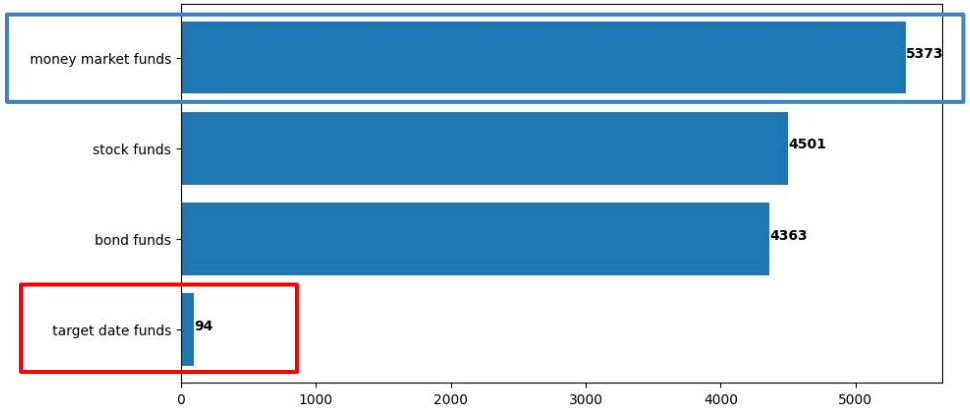


80% of them **have income source from salary** (4302 users), **business profit** (868), **and lottery** (844)

4) Exploratory Data Analysis (EDA)

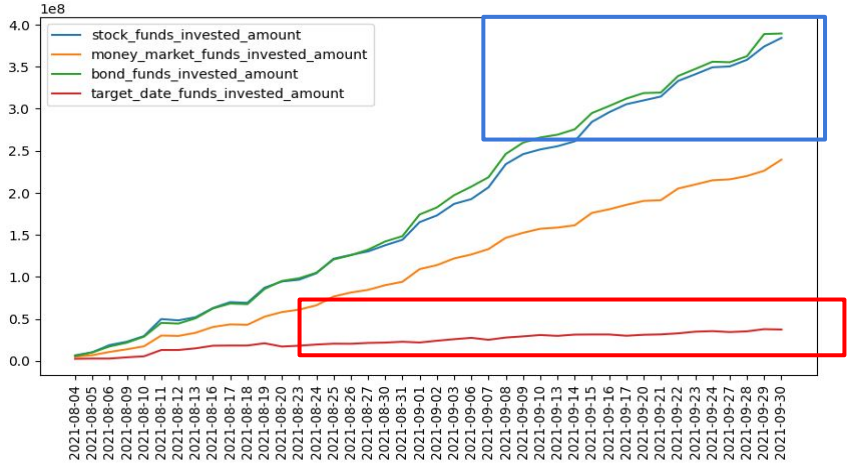
Understand the total users of each mutual funds and trend of transaction (python link)

Number of Users per Mutual Fund



Based on total users, **money market funds (5373 users) is the most favorite mutual fund** to be bought by the users. Unfortunately, **target date funds (94 users) is the most dislike mutual funds**.

Transaction Trend per Mutual Fund



Based on invested amount per product, the **highest invested amount trend is bond funds as well as stock funds** and **the lowest trend is target date funds.**

Additional Information

Most of our users use money market funds, but the total invested amount trend is not the highest. This can be happened because money market funds didn't give a high return (%), but the return is always positive (low risks), so they don't put more money in money market funds. Meanwhile, stock funds is the 2nd highest used by users (4501 users) with significantly increased total invested amount trend. This shows us most of our users have aggressive profile, they tend to invest more money in stock funds to gain the highest return.

4) *Exploratory Data Analysis (EDA)*

Understand the total transaction amount of each mutual fund (python link)

Total transaction amount of all mutual funds is Rp21,260,445,926

Invested Amount of
Bond Funds
36.96% of
total mutual funds
Rp7.858.092.044
from 4363 users

Invested Amount of
Stock Funds
36.01% of
total mutual funds
Rp7.654.943.517
from 4501 users

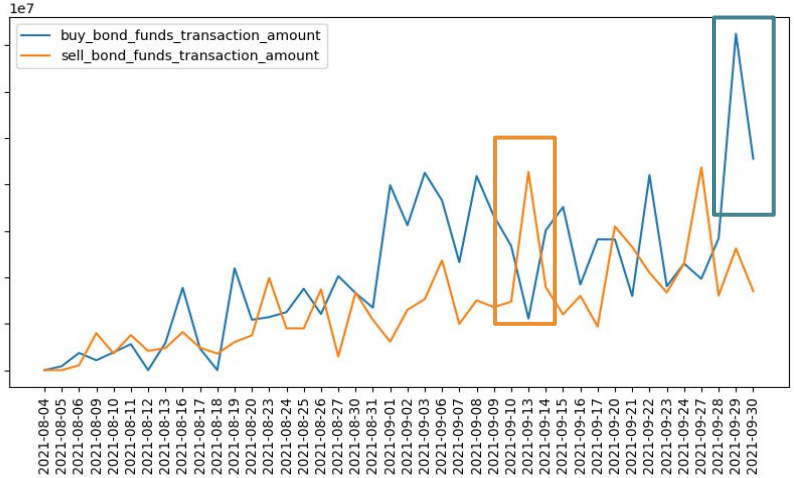
Invested Amount of
Money Market Funds
22.46% of
total mutual funds
Rp4.774.711.365
from 5373 users

Invested Amount of
Target Date Funds
4.58% of
total mutual funds
Rp972.699.000
from 94 users

4) *Exploratory Data Analysis (EDA)*

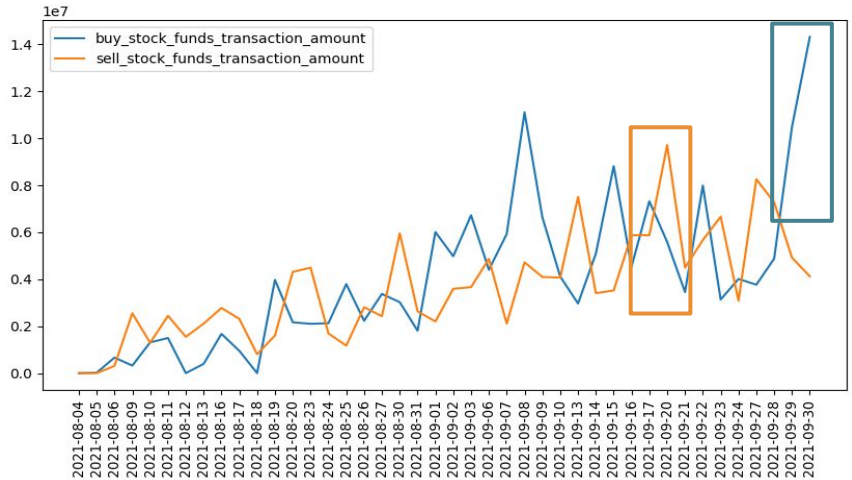
Understand the transaction trend of each mutual fund (python link)

Bond Funds Transaction Trend



Based on bond funds transaction trend, buy and sell have the similar pattern with the **highest value of buy bond funds happened on 2021-09-29** for around Rp14.000.000 and the **highest value of sell happened on 2021-09-13** for around Rp8.000.000

Stock Funds Transaction Trend

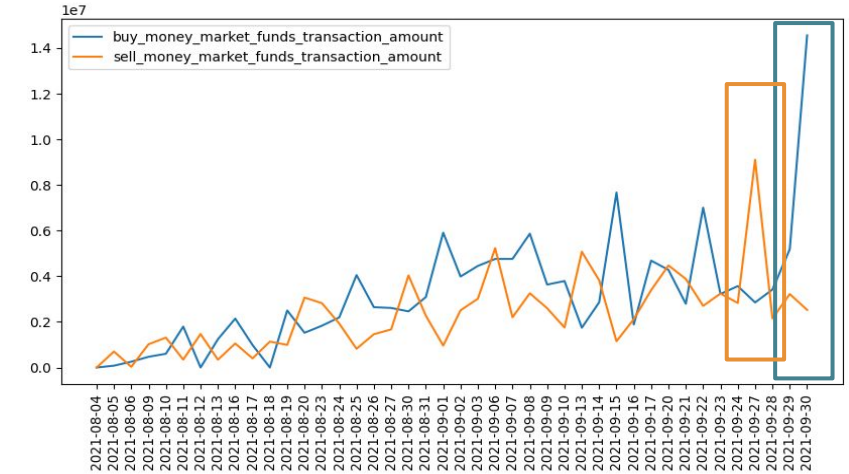


Based on stock transaction trend, buy and sell have the similar pattern with the **highest value of buy stock funds happened on 2021-09-30** for around Rp15.000.000 and **the highest value of sell happened on 2021-09-20** for around Rp9.000.000.

4) *Exploratory Data Analysis (EDA)*

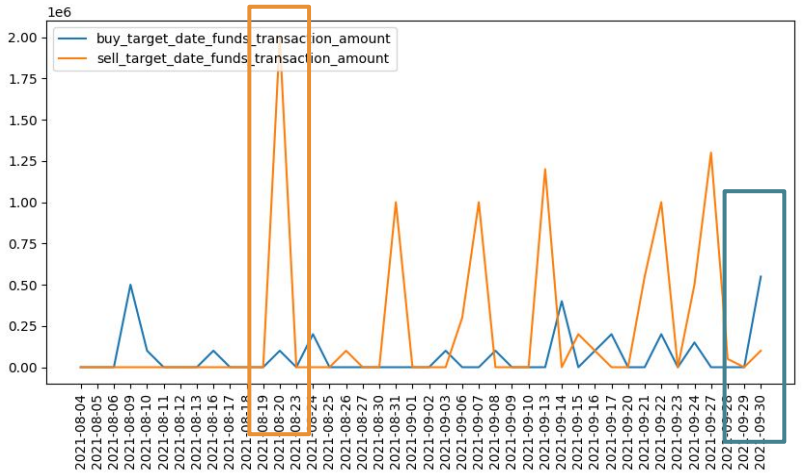
Understand the transaction trend of each mutual fund (python link)

Money Market Funds Transaction Trend



Based on money market funds transaction trend, buy and sell have the similar pattern with the **highest value of buy money market fund happened on 2021-09-30** for around Rp15.000.000 and the **highest value of sell happened on 2021-09-27** for around Rp9.000.000

Target Date Funds Transaction Trend



Based on target date funds transaction trend, sell target date funds have the highest trend than buy, it means our users tend to withdraw their investment in target date funds rather than buy it. The **highest buy happened on 2021-09-30** with the value Rp500.000. The **highest sell happened on 2021-08-20** with the value Rp2.000.000.

Additional Information

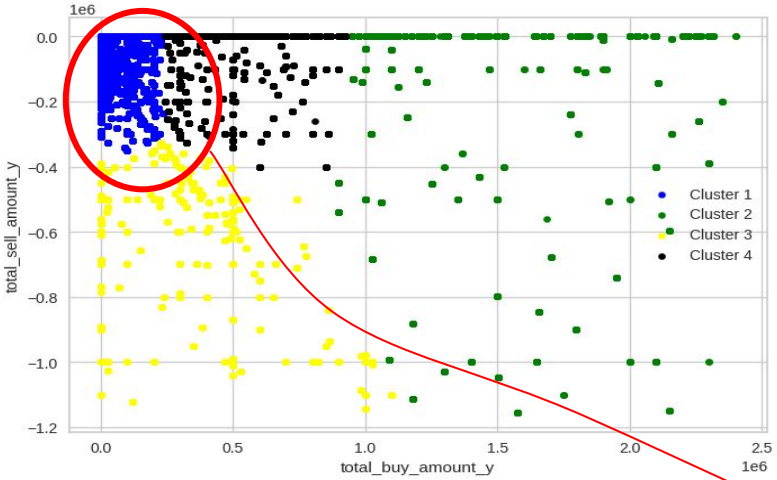
Overall, stock, money market, and bond funds have a similar trend (highest buy in the end of September). Target date funds has the very different trend, sell tends to has highest value rather than buy, meaning that many users tend to withdraw their investment in target date funds rather than buy it.

5) *Thematic Campaign Recommendation*

Segment the users with cluster analysis using k-means and list down top 30% of each segment with profit projection **(python link)**

By using K-Means Clustering for segmentation, we have 4 segments.

Scatter plot of each segment



- Cluster 1 (blue color); low total buy and low total sell amount
- Cluster 2 (green color): high total buy and low to high total sell
- Cluster 3 (yellow color): low to medium total buy and medium to high total sell
- Cluster 4 (black): medium total buy and low total sell

Our priority users (based on total investment):

Cluster 2 - Priority User

3.71% users is Cluster 2 (263 users) with average total invested Rp529.6K

Cluster 4 - Potential User

8.12% users is Cluster 1 (575 users) with average total invested Rp429.06K

Cluster 3 - Risky User

3.98% users is Cluster 3 (282 users) with average total invested Rp226.7K

Cluster 1 - Sink User

84.16% users is Cluster 1 (5953 users) with average total invested Rp118.5K

Most of our users are sink users

5) *Thematic Campaign Recommendation*

Segment the users with cluster analysis using k-means (python link)

1 Cluster 2 - Priority (highest invested amount)

263 crazy rich users that are really aware with investment to gain more return, have highest investment amount, and invest more in bond funds and followed by stock funds.

User Demography

- Male (142 users)
- Adult (25-65 yo with average 31 yo)
- Student and private employee
- Income 10-50 mio from salary, business profit
- Registered without referral (199 users) > used referral (64 users)

Transaction Average Amount

- Bond invested: Rp253,9K
- Stock funds invested: Rp209,2K
- Money market invested: Rp99,7K
- Target date invested: Rp29,7K
- Total buy: Rp1.380,8K
- Total sell: Rp160.4K
- Total invested: Rp592.6K

Thematic Campaign Recommendation

Go exploring your investment with stock! Highest return for your bright future

Recommendation: users who are really aware with investment still need a little assistance. **Private assistance** is really helpful to make them stay with us. Attract them to invest more in stock for highest return by **giving them reward** like **cashback voucher** for stock investment or **vacation deals** to attract them so they will invest more money in stock.

List of Users: Top 30% Users of Priority Users

From top 30% users (79 users), there are **67% (*) users will be churn** (53 users).

**from benefit cost analysis, the campaign will impact the negative expected return (-Rp2.3M). It can be happened due to high churn rate.*

5) *Thematic Campaign Recommendation*

Segment the users with cluster analysis using k-means (python link)

2 Cluster 4 - Potential to Become Priority

575 adult users with standard income who are still in a beginner step of investment and put more invest in bond funds, followed by stock funds.

User Demography

- Male (355 users)
- Adult (25-65 yo with average 28 yo)
- Student and private employee
- Income <10 mio from salary, business profit
- Registered without referral (422 users) > used referral (153 users)

Transaction Average Amount

- Bond invested: Rp163.1K
- Stock funds invested: Rp155.8K
- Money market invested: Rp98,7K
- Target date invested: Rp11.4K
- Total buy: Rp455.3K
- Total sell: Rp37.9K
- Total invested: Rp429.06K

Thematic Campaign Recommendation

Break the glass! Add your investment with stock to achieve the goal fast

Recommendation: feedback report of their return will help them to analyze their investment. If they found it helpful, they will invest more. **Make them set the goal and offer the loyalty programs** to gain voucher when they achieved the target amount with stock invest.

List of Users: Top 30% Users of Potential Users

From top 30% users (173 users), there are **67% (*) users will be churn** (116 users).

**from benefit cost analysis, the campaign will impact the negative expected return (-Rp3.7M). It can be happened due to high churn rate.*

5) *Thematic Campaign Recommendation*

Segment the users with cluster analysis using k-means (python link)

3 Cluster 3 - Risky Users

282 adult users with standard income who start to aware with the importance of investments, has a low transaction amount, interest to gain highest return by investing more in stock funds.

User Demography

- Male (203 users)
- Adult (25-65 yo with average 28 yo)
- Student and private employee
- Income <10 mio from salary, business profit, lottery
- Registered without referral (222 users) > used referral (60 users)

Transaction Average Amount

- Stock funds invested: Rp83.1K
- Bond invested: Rp81.8K
- Money market invested: Rp41.5K
- Target date invested: Rp20.1K
- Total buy: Rp237.0K
- Total sell: Rp650.1K
- Total invested: Rp226.7K

Thematic Campaign Recommendation

The more often you invest, the more you achieve the goal. Invest now!

Recommendation: as a new user who are aware with investment, **investment calendar** will help them to track their investment activity. **Give a voucher after achieving their calendar goal** will also help to attract them to achieve the goal.

List of Users: Top 30% Users of Risky Users

From top 30% users (85 users), there are **48% (*) users will be churn** (41 users)

Campaign Result Calculation:

Rp670.632
expected return

5) *Thematic Campaign Recommendation*

Segment the users with cluster analysis using k-means (python link)

4 Cluster 1 - Sink Users (Lowest Invested Amount)

5953 youth users with standard income who just recognize investment and still in the curious moment, most of them only sign up and don't do the transaction.

User Demography

- Male (3836 users)
- Youth (17-24 yo with average 21 yo)
- Student and private employee
- Income <10 mio from salary, business profit, lottery
- Registered without referral (3668 users) > used referral (2285 users)

Transaction Average Amount

- Stock funds invested: Rp42.5K
- Bond invested: Rp42.2K
- Money market invested: Rp28.0K
- Target date invested: Rp5.6K
- Total buy: Rp31.2K
- Total sell: Rp34.5K
- Total invested: Rp118.5K

Thematic Campaign Recommendation

Reach up to 10% return by investing your money!

Recommendation: users in curious moment sometimes still need to learn what is the function of investment. Assistance will help these users to dig deeper about investment! Send the **weekly material** to them including **investment calculator**. **Notification with the number of return percentage** will also attract these users. **Give a voucher for first investment** as appreciation to them.

List of Users: Top 30% Users of Sink Users

From top 30% users (1786 users), there are **56% (*) users will be churn** (1006 users)

Campaign Result Calculation:

Rp11.930.213
expected return

6) *Summary and Recommendation*

SUMMARY

1

Our platform dominated by youth male with income < Rp10M from salary or business profit and registered without referral code.

2

Based on segmentation, most of our users are Sink Users (lowest total invested amount, 5953 users from 7073 total users)

3

Highest invested amount comes from bond funds (up to Rp7.858M from 4363 users), but the highest total user comes from money market funds (Rp4.774M from 5373 users).

4

We can gain expected return up to Rp11.9M from the campaign with the positive expected return occurred for campaign from Risky Users and Sink Users.

OVERALL RECOMMENDATION

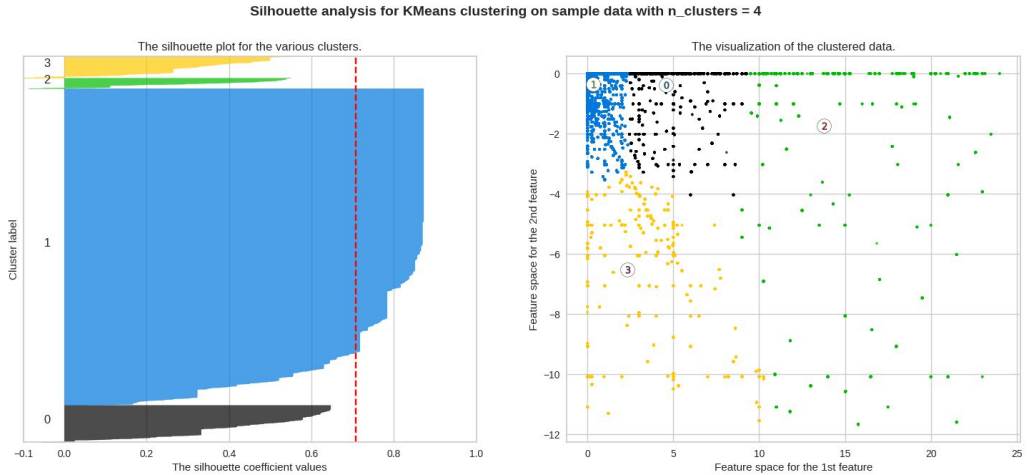
From the difference expected return, we can focus to run a campaign for Risky Users and Sink Users (Risky 48% churn, sink 56% churn) first since they don't have a high churn rate if compared to Priority and Potential Users (Priority 67% churn, Potential 67% churn).

THANK YOU

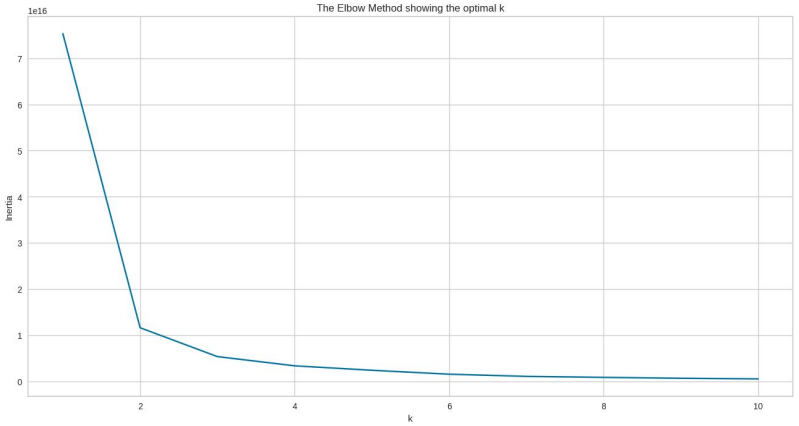
APPENDIX

Segment the users with cluster analysis using k-means [(python link)]

Silhouette Analysis with Cluster Scatterplot



Elbow Method



- We use k-means algorithm for clustering to find groups which have not been explicitly labeled in the data
- Based on elbow method and silhouette, we choose 4 clusters. In Silhouette Analysis, 4 shows a similar form with **avg value of silhouette score: 0.70.**
- Use robust scaler standardization because we have many outliers (before clean up the data) and we want to standardize use interquartile range (IQR) so that it is robust to outliers.
- Based on scatter plot, cluster 1 (sink users in blue color) has low total buy and low total sell amount, cluster 2 (priority users in green color) has high total buy and low to high total sell, cluster 3 (risky users in yellow color) has low to medium total buy and medium to high total sell, and cluster 4 (potential users in black color) has medium total buy and low total sell