Rakamin Academy X KALBE Nutritionals

Machine Learning Project

Kalbe Nutritionals Data Scientist Project Based Internship Program

Presented by
Kristy Natasha Yohanes

Kristy Natasha Yohanes

About Me

Hello! I'm a 2023 graduate from ITB with a Bachelor of Science (B.S.) in Meteorology. I'm deeply passionate about the world of data and analytics. I find joy in tackling machine learning projects, coding challenges, and data science competitions, all while leveraging the power of Python.

Let's connect! https://linkedin.com/in/kristynatasha/

Work Experience

Researcher (Machine Learning Weather Forecaster)

Atmospheric Department ITB

- Led the development of a machine-learning forecasting model and co-authored a scientific paper on ANN-ARIMA weather forecasting.

Data Science Research Intern

BRIN Aviation and Space Research Organization

- Applied advanced data science techniques and statistical modeling to gather and analyze extensive weather data, deriving actionable insights to inform strategic initiatives.

Data Collector & Analyst

Community Service Program (PPM) by Garda Caah

- Led the development of a machine-learning forecasting model and co-authored a scientific paper on ANN-ARIMA weather forecasting.

Case Study

01

Task from
INVENTORY TEAM

- Find out the estimated quantity of products sold so that the inventory team can create sufficient daily inventory.
- Make daily predictions.

02

Task from
MARKETING TEAM

- Create customer segments.
- This customer segment will later be used by the marketing team to provide personalized promotion and sales treatment.

Exploratory Data Analysis

using DBeaver with a PostgreSQL database

Marital Status	Gender
Married (avg. 43 years)	Man (avg. 39 years)
Single (avg. 29 years)	Woman (avg. 40 years)

Store	Quantity
Lingga	2,78K
Sinar Harapan	2,59K
Prima Kota	1,40K

Product	Total Amount
Cheese Stick	27,6M IDR
Choco Bar	21,2M IDR
Coffee Candy	19,7M IDR

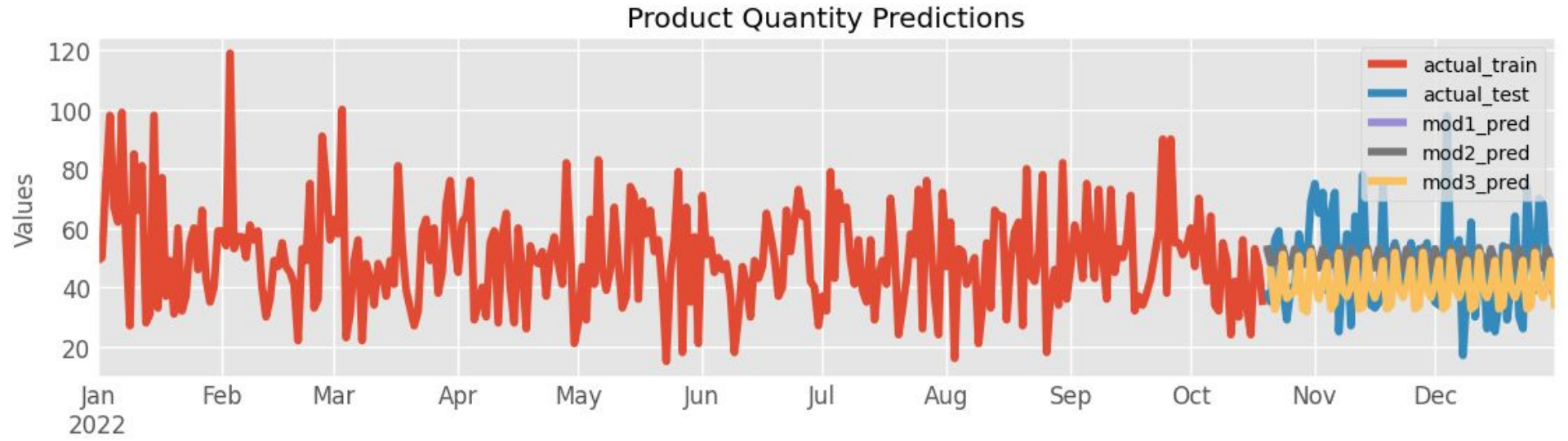
Data Visualization

using Tableau Public

https://public.tableau.com/app/profile/kristy.natasha/viz/KalbeDSIntenship/Dashboard#1

Predictive Analytics

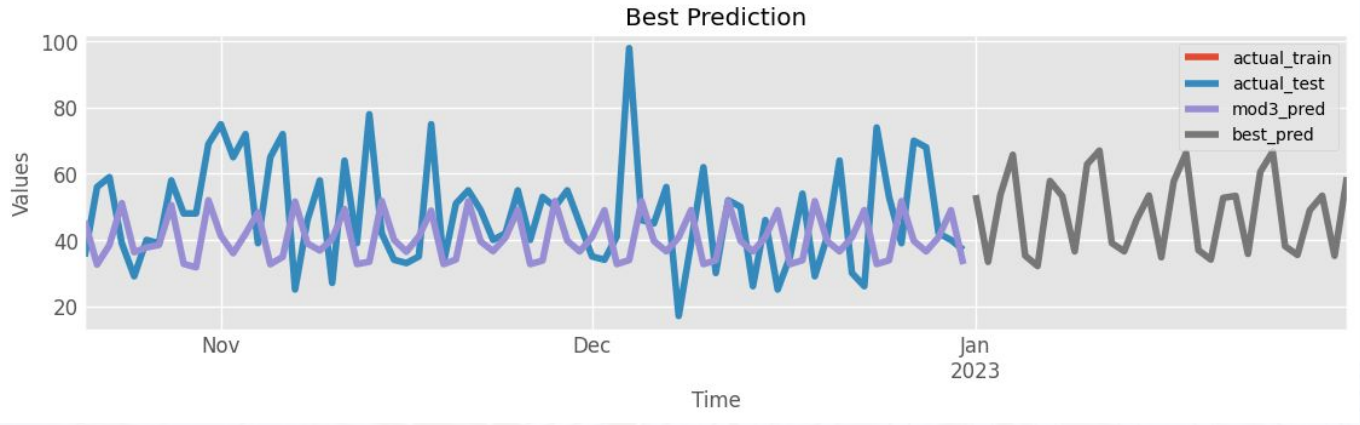
using machine learning regression (time series model ARIMA) with Python



https://colab.research.google.com/drive/14rTMcSWFT9lCfjn_bCG-zvqdR2chq9sB

Predictive Analytics

using machine learning regression (time series model ARIMA) with Python

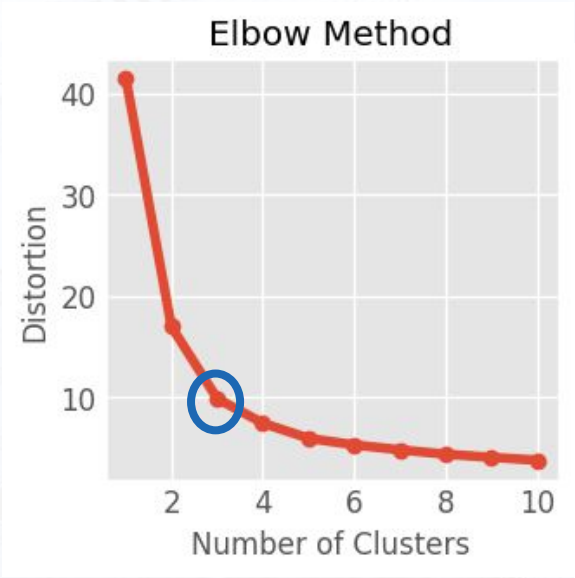


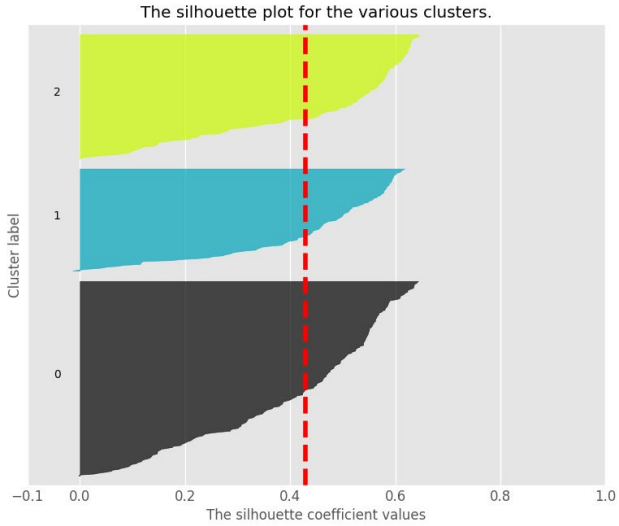
	predicted_mean
count	31.000000
mean	48.233205
std	11.955421
min	32.113562
25%	36.149794
50%	52.715606
75%	57.762546
max	67.016036

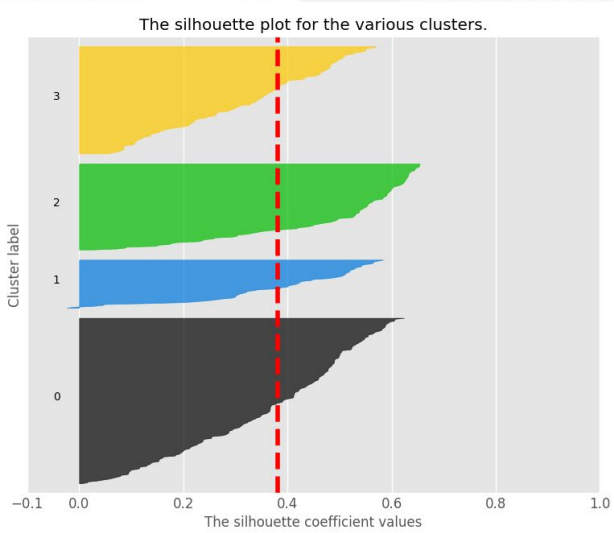
Best forecast result of quantity of product needed in January 2023

Data Clustering

using KMeans library in Python







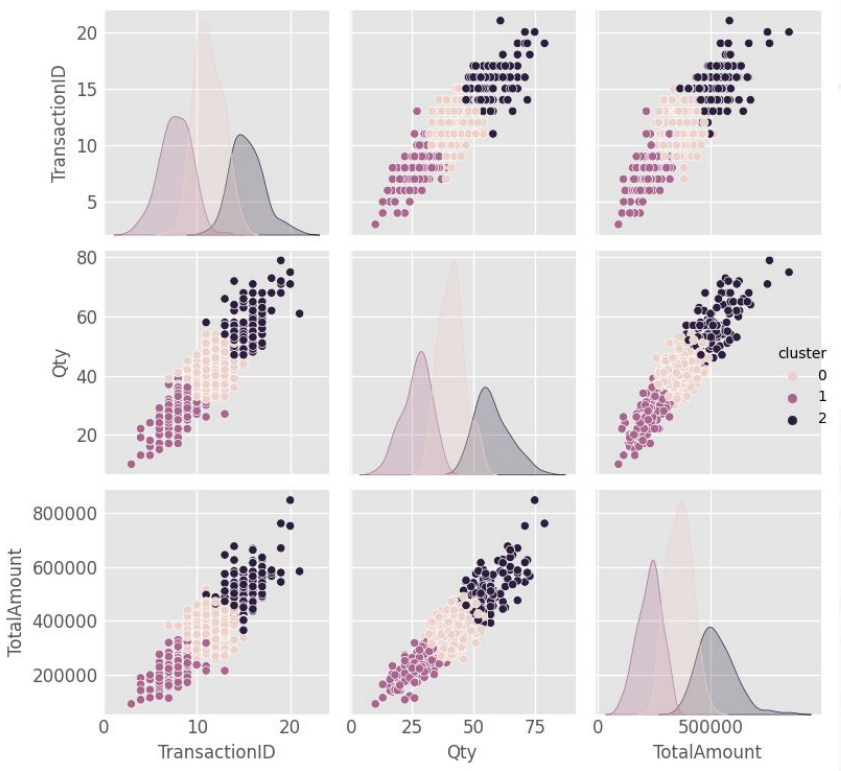
Optimal number of clusters = 3

For n_clusters = 3 The average silhouette_score is : 0.4294669050463297
For n_clusters = 4 The average silhouette_score is : 0.38109175331136835

n = 3 has better proportional distribution than n = 4

Data Clustering

using KMeans library in Python



Customer Segmentation

cluster	0	1	2
TransactionID	11.253659	7.702290	15.370370
Qty	41.004878	26.725191	57.574074
TotalAmount	360908.292683	228550.381679	524466.666667

Customer Profile

High Spenders

Customers in this cluster are the highest spenders, making a large number of transactions and purchasing substantial quantities of products. They are the most consumptive group.

- VIP Treatment (exclusive perks, early access to promotions)
- Premium Products
- Referral Programs
- Personalization

Moderate Shoppers

This cluster consists of customers with a moderate level of consumption. They make a good number of transactions and purchase reasonably-sized quantities of products.

- Retention and Upselling (loyalty programs, offer exclusive discounts, rewards)
- Cross-selling
- Personalization

Budget Shoppers

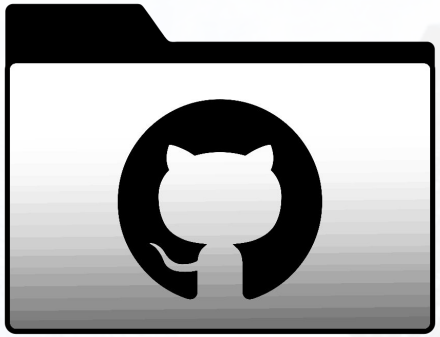
Customers in this cluster are budget-conscious shoppers. They make fewer transactions and opt for smaller quantities. They are the least consumptive group.

- Customer Engagement
- Product Bundles
- Feedback and Surveys

RESULT DOCUMENTATION

Video Presentation

drive.google.com/drive/folders/1eS8P3QKZq6Ipvq6-gVvfz9LvlwK0c3qs



Project Repository

github.com/kristynatasha/FMCG-Data-Modeling

CONTACT INFO

GitHub

github.com/kristynatasha

LinkedIn

linkedin.com/in/kristynatasha/

(+62) 8788-658-3513 | kristynatasha011@gmail.com | Jakarta, Indonesia

Thank You