Rakamin Academy x KALBE Nutritionals

MACHINE LEARNING PROJECT REGRESSION AND CLUSTERING

BY: DEA DAHLILA

Click here for more details.

About Me

“I'm a Fresh Graduate of Geophysics Engineering with a keen interest in data science and analysis. I'm passionate about extracting meaningful insights from data and using them to drive informed decision-making.”

Let's Connect

ddahlila6@gmail.com

linkedin.com/in/dea-dahlila/

github.com/Dea-dahlila

Background Story

- The inventory team asked to predict the sales volume (quantity) of the total number of Kalbe products. The goal of this project is to keep track of the estimated quantity of products sold so that the inventory team can make sufficient daily inventory stock.

- The marketing team asked to create clusters/segments based on several criteria. the goal is to create customer segments which will later be used to provide personalized promotion and sales treatment.

Exploratory Data Analysis

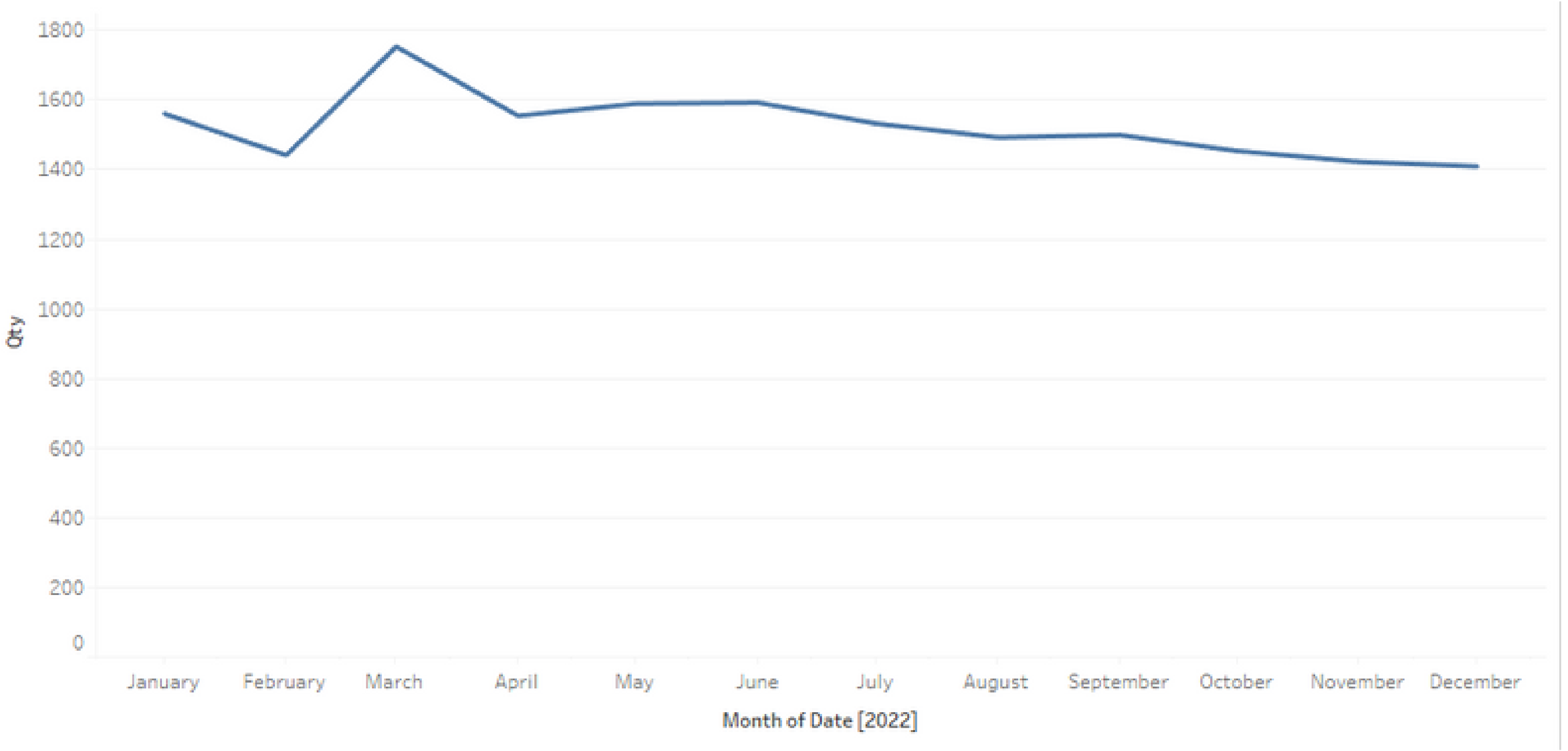
	Marital Status character varying	avg_age_maritalstatus numeric
1		31
2	Married	43
3	Single	29

	gender integer	avg_age_gender numeric
1	0	40.32644628099173 55
2	1	39.1414634146341463

	storename character varying	high_quantity bigint
1	Lingga	2777

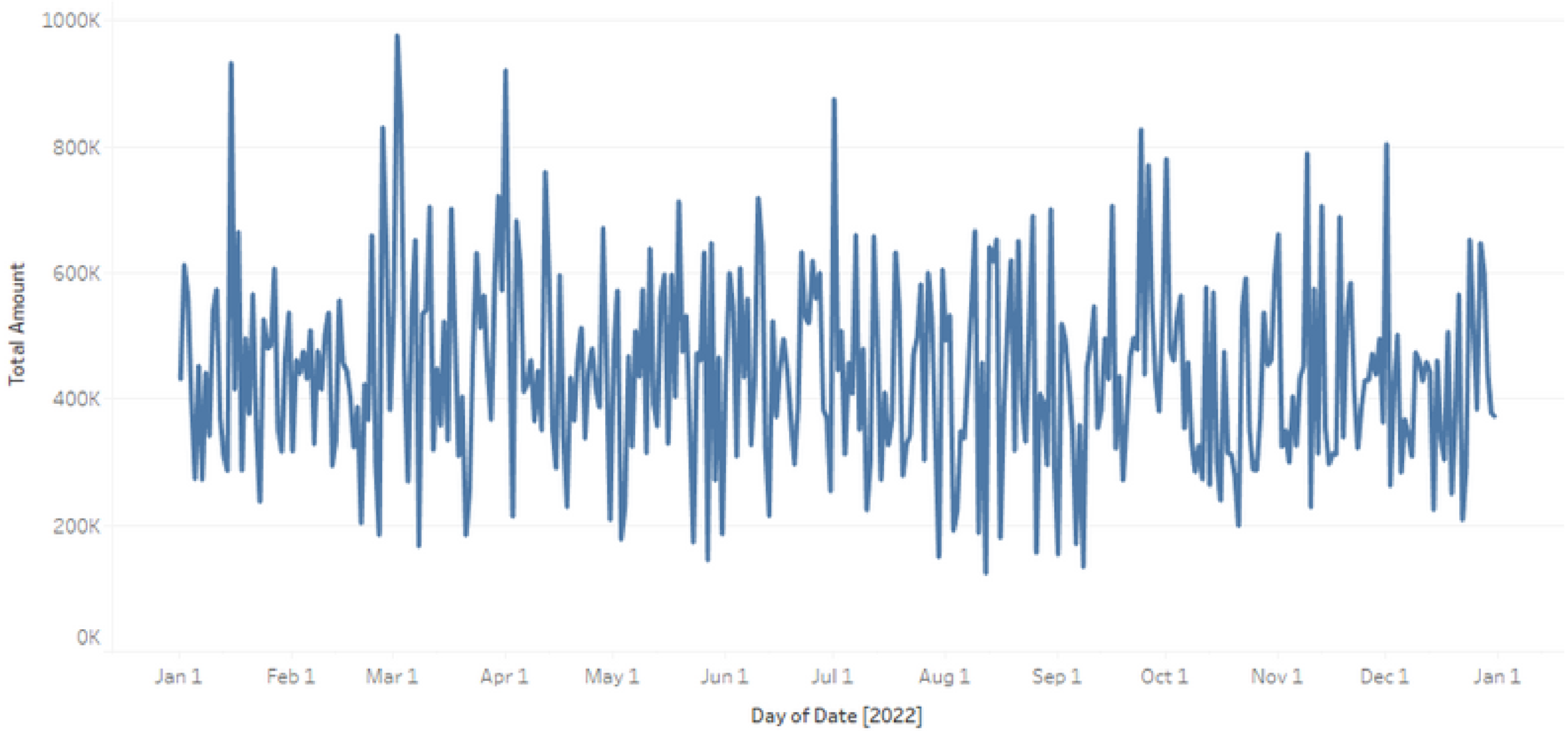
	high_amount bigint	Product Name character varying
1	27615000	Cheese Stick

Total Qty Month by Month



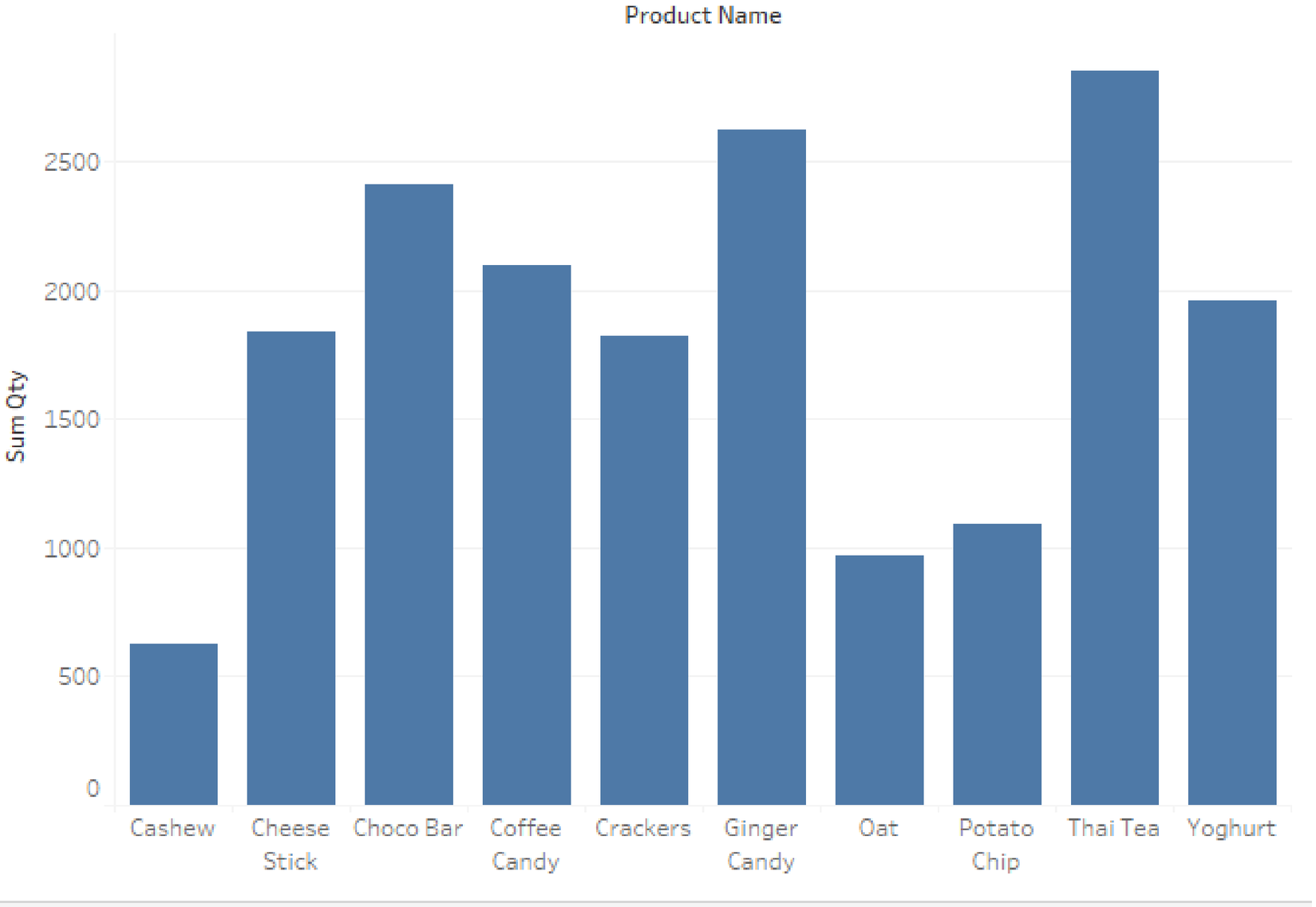
Click here for more details Tablue Dashboard.

Total Amount Day by Day

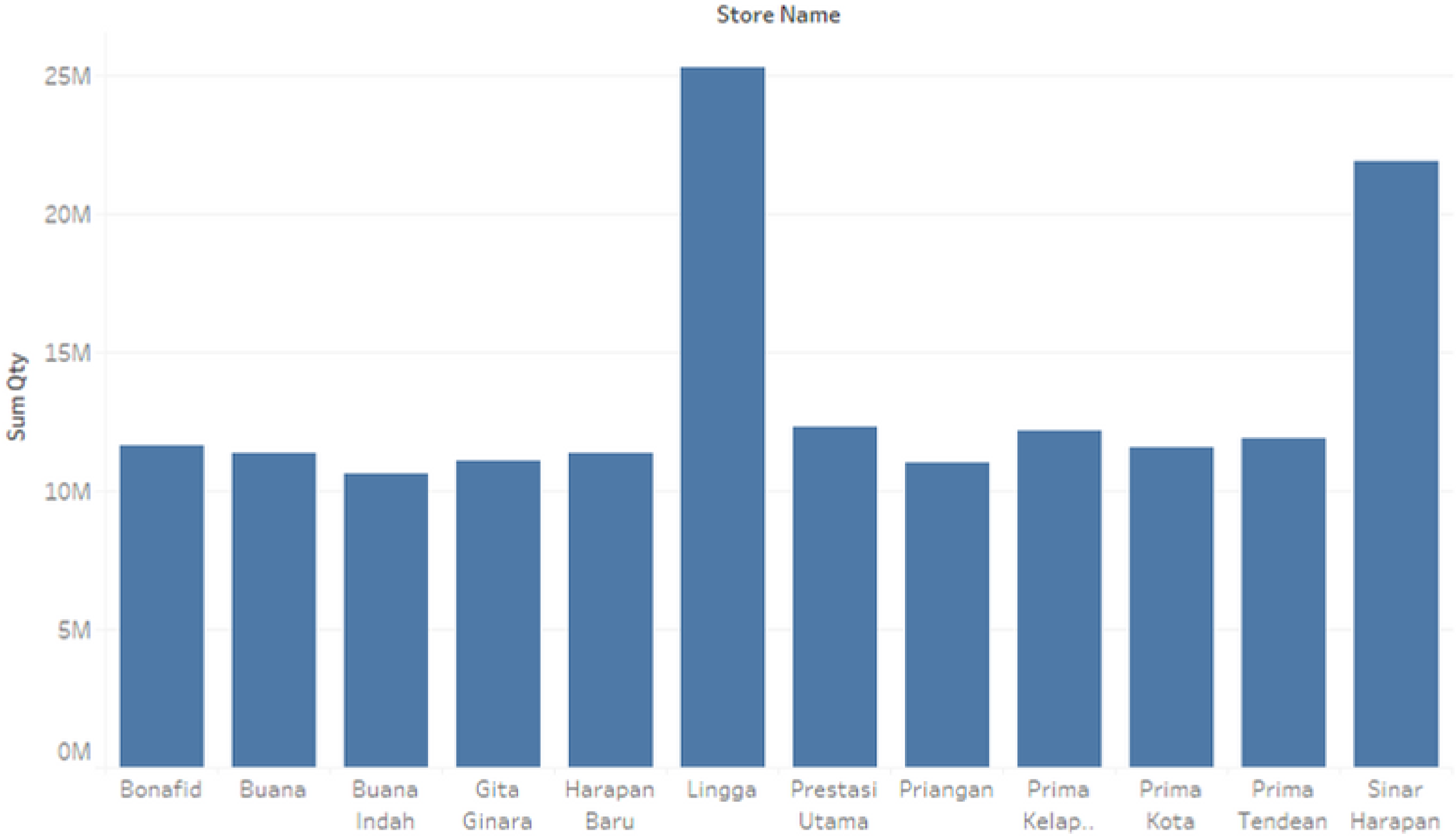


Click here for more details Tablue Dashboard.

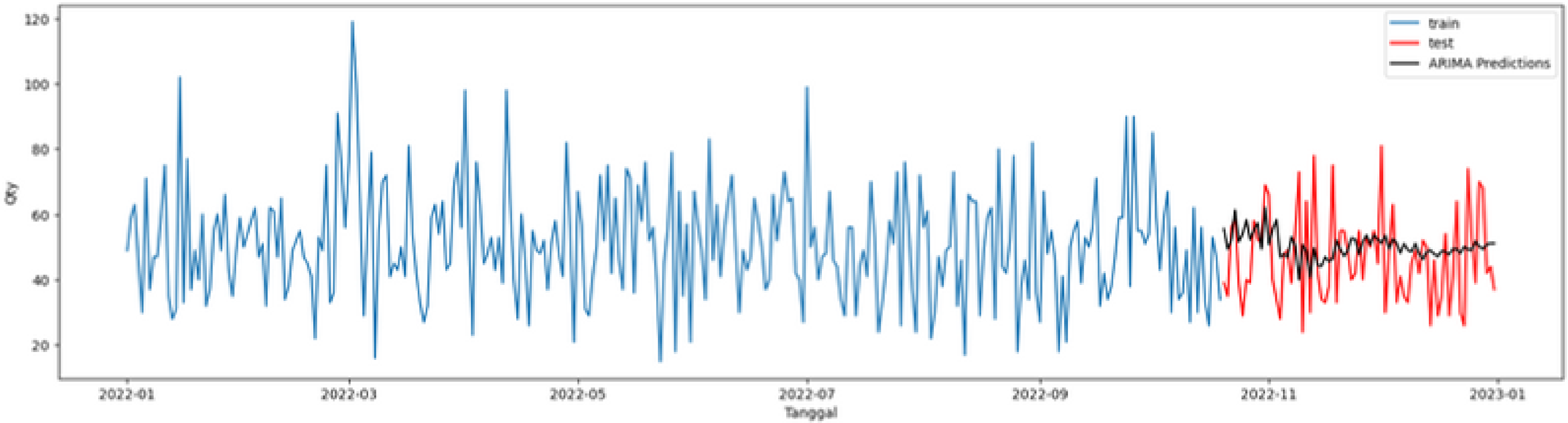
Total Qty by Product Name



Total Qty by Store Name

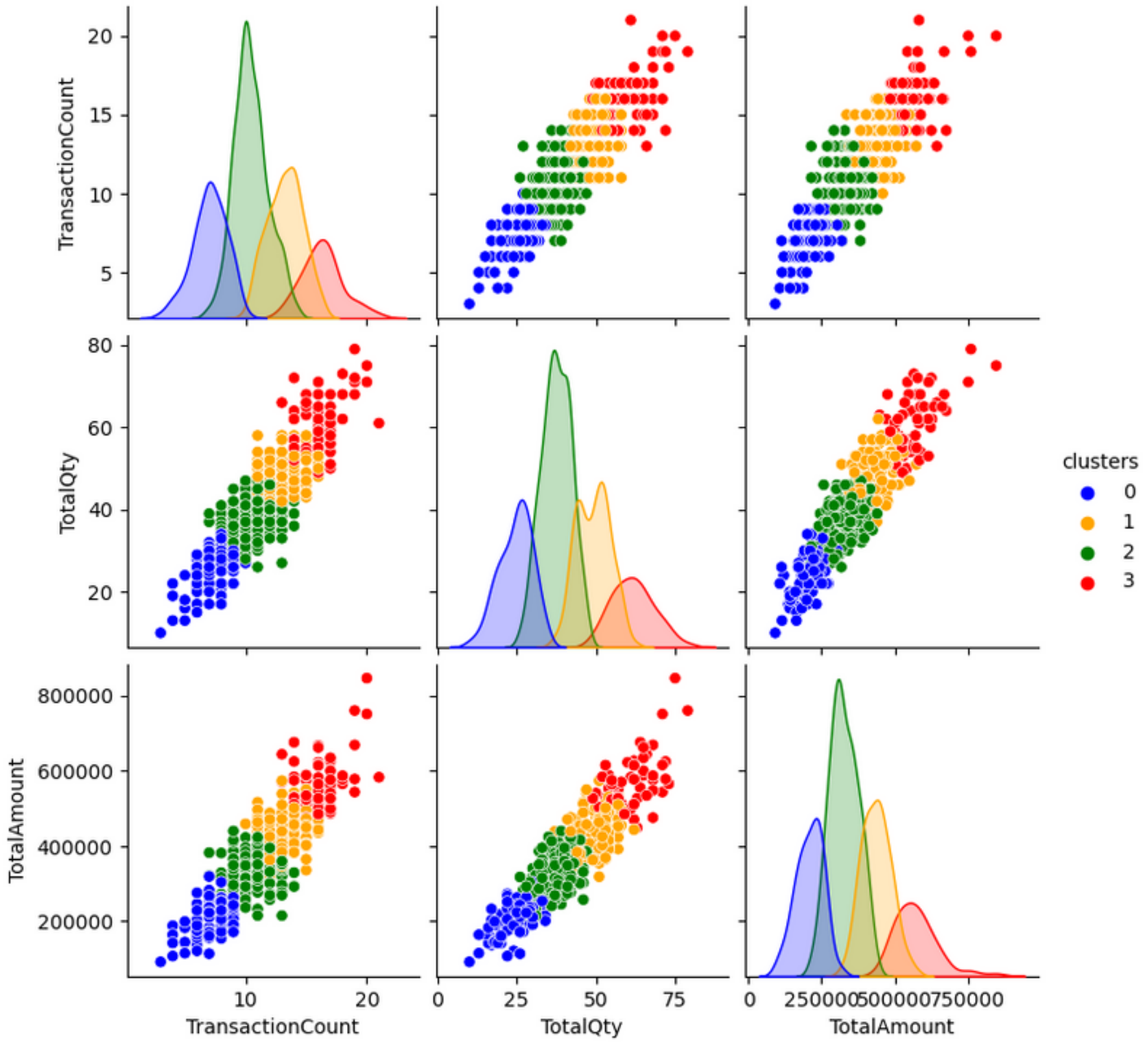


Regression by Arima Model



From the forecasting above, it can be concluded that next month's sales volume is around an average of 50 pcs/day.

Clustering



Cluster 1

Has the highest average number of transactions per customer and also has a relatively high total product quantity and average spending per customer. This shows that customers in this group are customers who tend to make more transactions with large purchases.

Cluster 2

has the largest customers. They tend to make more transactions, even though the total quantity of products purchased and the average spending per customer are relatively low compared to the other groups.

Cluster 3

Cluster 3 has the highest number of transactions and the highest total product quantity per customer. The average spending per customer in this group is also very high, indicating that customers in this group are customers who make many transactions with a large number of products and high spending.

Cluster 0

Has a lower number of transactions and total product quantity compared to other groups. The average spend per customer in this group is also lower. This shows that customers in this group are customers who make fewer transactions with fewer products and lower expenses.

Business Recomendation

- For cluster 0 : try to increase the frequency of transactions in this cluster with strategies such as increasing relevant product offers, sending reminders or special offers (bundles) to customers on a regular basis.

- For cluster 1: Do strong customer retention in this cluster. They already tend to do more transactions with high value. Focus on maintaining customer loyalty by providing superior customer service, exclusive offers and a satisfying shopping experience.

- For cluster 2: Focus on increasing the quantity of products purchased by customers in this cluster. Can implement strategies such as special offers, additional discounts for repeat purchases, or loyalty programs that provide incentives for more transactions. Work to increase the value of spending per customer in this cluster.

- For cluster 3: Continue to strengthen relationships with customers in this cluster. They are the most valuable customers with high transaction value. Consider providing exclusive services, hosting special events, or offering them premium products. Make sure they feel valued and given special attention.

THANK YOU