

Machine Learning Project

Data Scientist

Presented by
Yoga Aria Sena



Yoga Aria Sena

About Me

I am an enthusiastic individual deeply intrigued by the world of data science. My journey is fueled by a profound interest in unraveling insights from data and transforming them into meaningful solutions. Aspiring to become a skilled data scientist, I am dedicated to expanding my expertise in programming, statistical analysis, and machine learning algorithms. My aspiration is to harness the power of data to create innovative solutions that make a tangible impact on the world.

My Experience

- Project Based Virtual Internship:
Data Scientist Kalbe Nutritionals x
Rakamin Academy
(August 2023 – Present)

About Rakamin Academy



Rakamin Academy adalah Platform akselerasi karir, tempat bagi siapapun untuk belajar dan membangun karir di bidang digital dan teknologi.

Rakamin Academy menyediakan layanan Pendidikan yang paling terjangkau (efisien secara model bisnis), berdampak (dihubungkan langsung dengan ahli, praktik standar industri), dan bermakna (mengajarkan soft skills, membantu pengembangan karir, membangun komunitas).

About Kalbe Nutritionals



Kalbe Nutritionals adalah anak perusahaan dari PT Kalbe Farma Tbk. yang didirikan pada tanggal 2 Desember 1982 dan bergerak di bidang kesehatan.

Produk-produk yang dihasilkan oleh Kalbe Nutritionals berupa makanan dan minuman kesehatan yang menjangkau di setiap titik kritis tahap pertumbuhan dan perkembangan manusia, seperti susu untuk bayi, anak-anak, remaja, ibu hamil dan menyusui, beberapa kebutuhan khusus kaum manula, serta biskuit dan sereal bayi .

Kalbe Nutritionals berkomitmen untuk memberikan solusi nutrisi terbaik di setiap tahap kehidupan manusia, sehingga secara bersama-sama bisa membantu menyehatkan bangsa Indonesia.

Case Study

Kalbe Nutritionals - Sales Prediction & Customer Segmentation

- Backgroud Story:

Saya adalah seorang Data Scientist di Kalbe Nutritionals dan sedang mendapatkan project baru dari tim inventory dan tim marketing. Dari tim inventory, saya diminta untuk dapat membantu memprediksi jumlah penjualan (quantity) dari total keseluruhan product Kalbe.

- Tujuan :
Prediksi harian penjualan & segmentasi pelanggan.
- Analisis Penjualan :
Model prediksi menggunakan Python.
- Segmentasi Pelanggan :
Klasterisasi data pelanggan.
- Alat
Python, Jupyter Notebook, Tableau, DBeaver

Exploratory Data Analysis Using DBeaver



Rata-Rata Umur Customer Berdasarkan Marital Status

```
-- Rata-Rata Umur Customer Berdasarkan Marital Status  
select "Marital Status", avg(age) as "Average Age"  
from customer c  
group by "Marital Status";
```

	ABC Marital Status ▼	123 Average Age ▼
1		31
2	Married	43
3	Single	29

- Rata-rata umur customer yang Married yaitu 43 tahun
- Rata-rata umur customer yang Single yaitu 29 tahun
- Tanpa Keterangan memiliki rata-rata 31 tahun

Rata-Rata Umur Customer Berdasarkan Gender

```
-- Rata-Rata Umur Customer Berdasarkan Gender
select gender, round(avg(age)) as "Average Age"
from customer c
group by gender;
```

	123 gender ▼	123 Average Age ▼
1	0	40
2	1	39

- Rata-rata umur customer Wanita yaitu 40 tahun
- Rata-rata umur customer Pria yaitu 39 tahun

Store dengan Total Quantity Terbanyak

```
-- 3 Store dengan Quantity Terbanyak
select
  s.storename,
  sum(t.qty) as "Total Quantity"
from store s
full join "Transaction" t
on s.storeid = t.storeid
group by s.storename
order by "Total Quantity" desc
limit 1;
```

- Lingga merupakan store dengan total quantity terbanyak yaitu sejumlah **2777**.

	ABC storename ▼	123 Total Quantity ▼
1	Lingga	2,777

Produk Terlaris dengan Total Amount Terbanyak

```
-- 4 Produk Terlaris dengan Total Amount Terbanyak
select
  p."Product Name",
  sum(t.totalamount) as "Total Amount"
from product p
full join "Transaction" t
on p.productid = t.productid
group by p."Product Name"
order by "Total Amount" desc
limit 1;
```

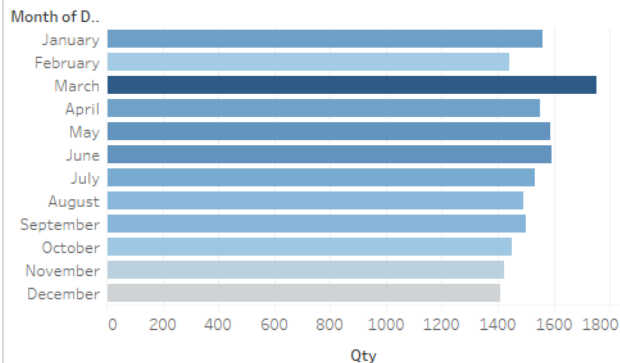
- Cheese Stick merupakan store dengan total amount terbanyak yaitu sejumlah **27,615,000**.

	ABC Product Name ▼	123 Total Amount ▼
1	Cheese Stick	27,615,000

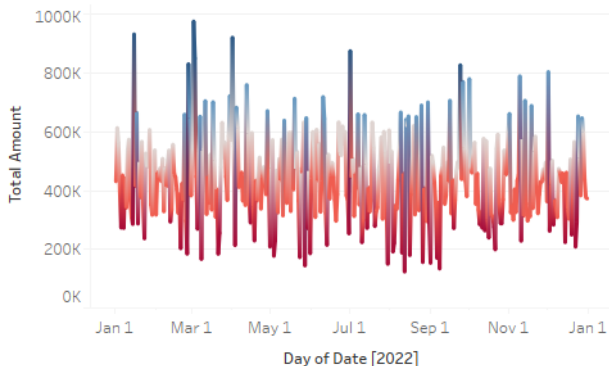
Sales Performance Overview Using Tableau



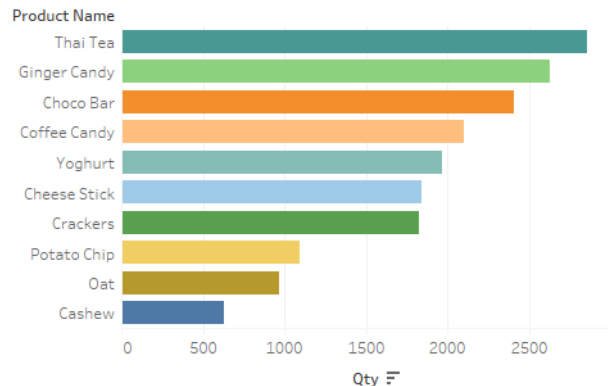
Qty Monthly Breakdown



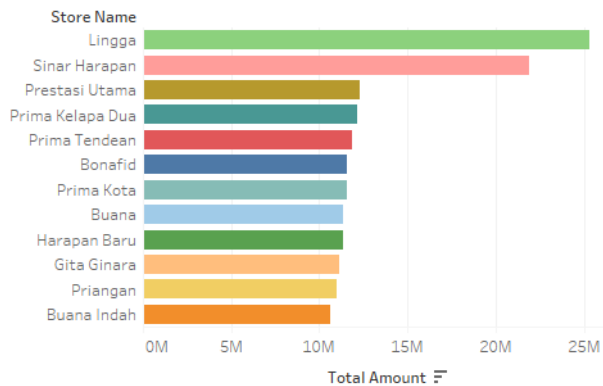
Daily Total Amount



Sales by Product (Qty)

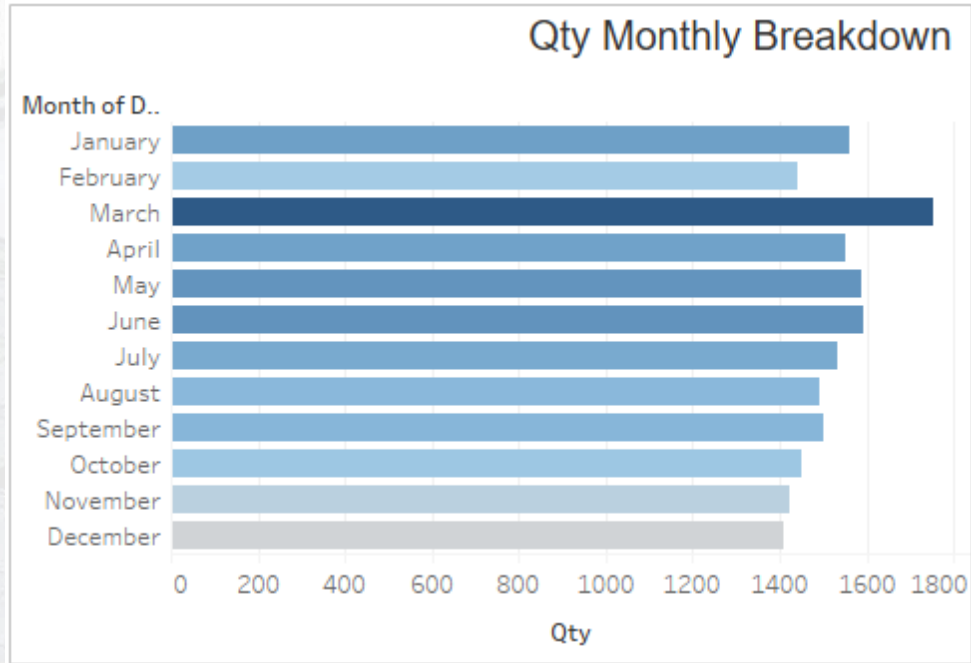


Sales by Store (Total Amount)



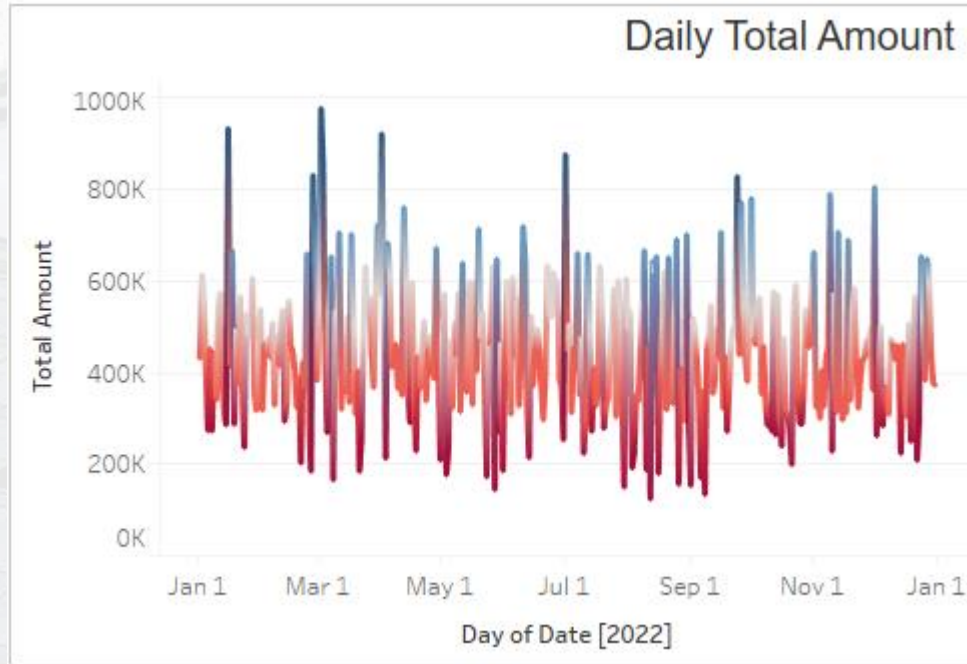
Link :
[FinalTask_Kalbe_DS_](#)
[Yoga Aria Sena |](#)
[Tableau Public](#)

Jumlah Qty Dari Bulan ke Bulan



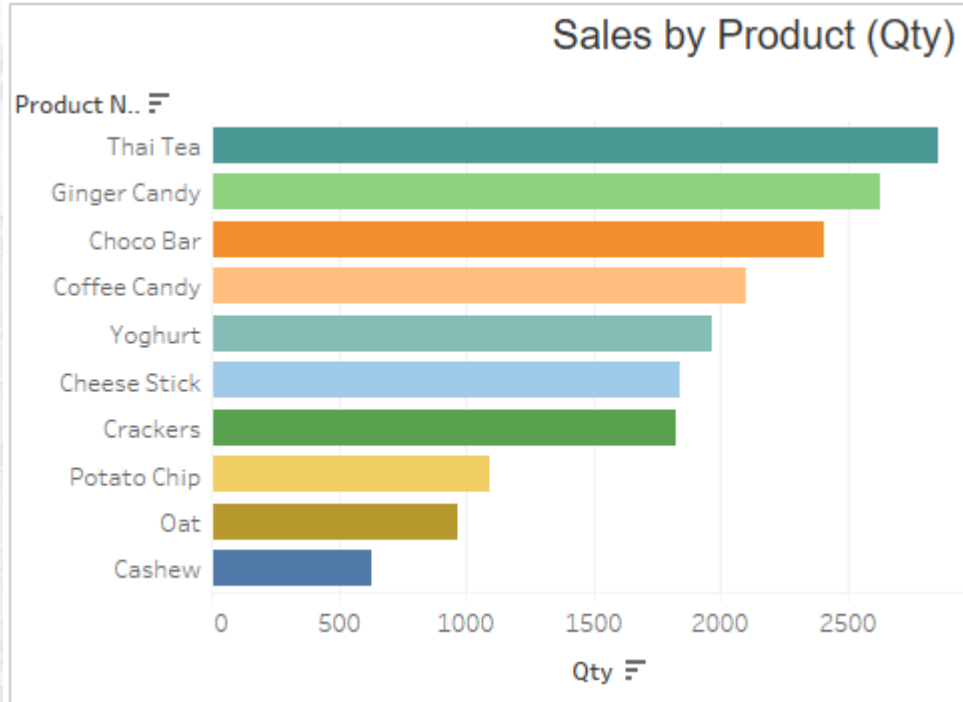
Jumlah Quantity terbanyak yaitu pada bulan Maret dan terendah pada bulan Desember

Jumlah Total Amount Hari ke Hari



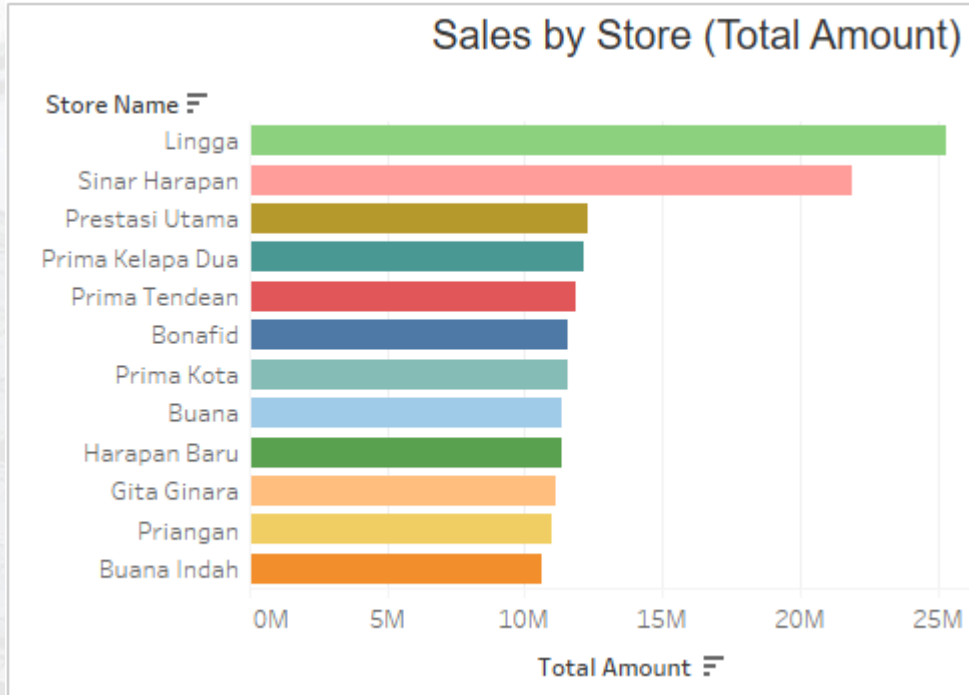
Jumlah total amount tertinggi yaitu sebesar **976,500** pada 2 Maret 2022, sedangkan jumlah total amount terendah yaitu sebesar **123,600** pada 12 Agustus 2022.

Jumlah Penjualan (Qty) by Product



Jumlah penjualan (Qty) by product tertinggi yaitu Thai Tea sejumlah **2,853**. Sedangkan jumlah penjualan (Qty) by product terendah yaitu Cashew sejumlah **627**.

Jumlah Penjualan (Total Amount) by Store Name



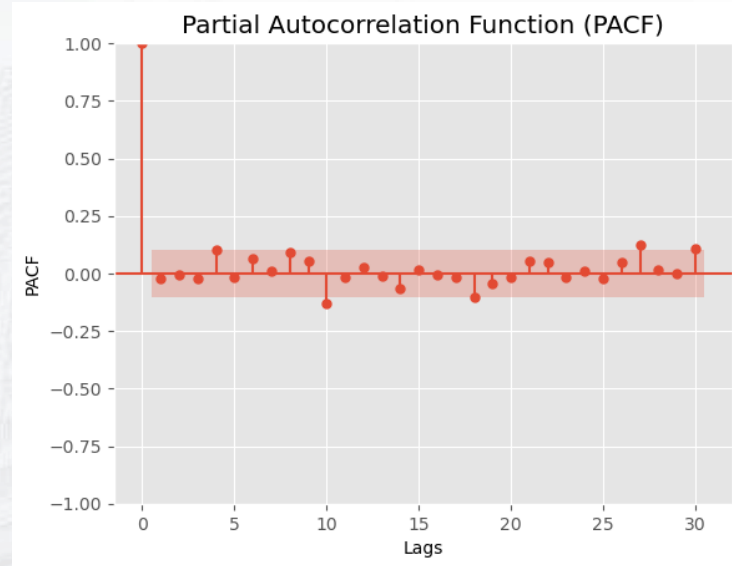
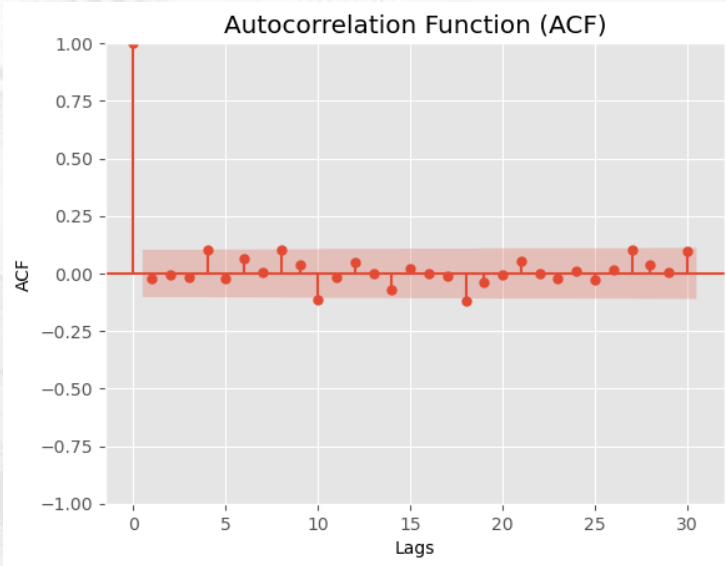
Jumlah penjualan (Total Amount) by store name tertinggi yaitu Lingga sejumlah **25,294,100**.

Sedangkan jumlah penjualan (Total Amount) by store name terendah yaitu Buana Indah sejumlah **10,629,900**.

Machine Learning Regression (Time Series) Using ARIMA Model



ARIMA MODEL PARAMETER OPTIMIZATION

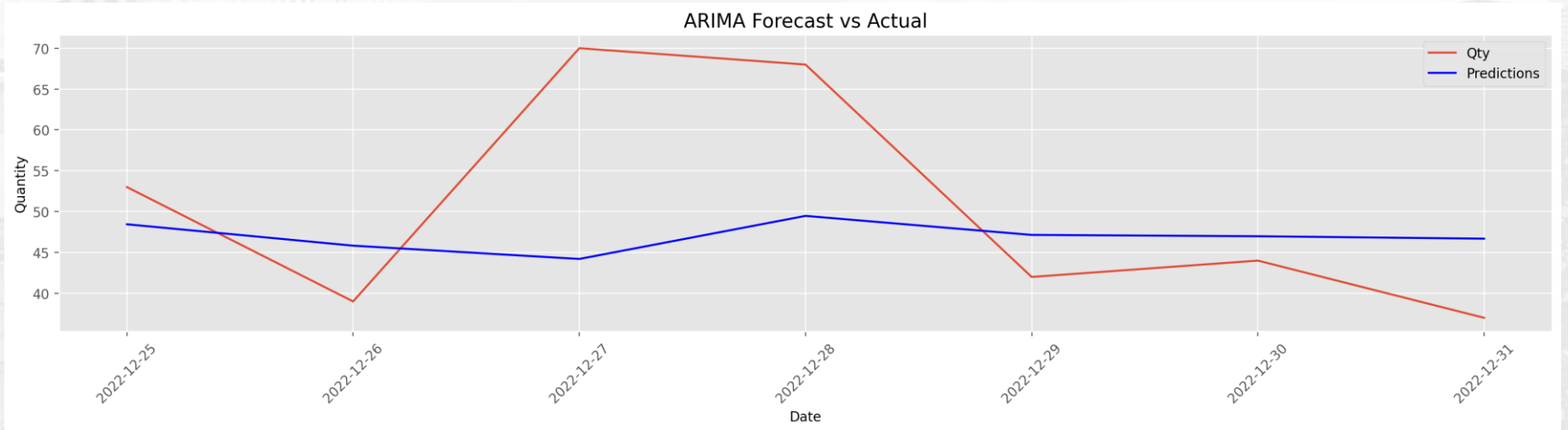


$P = 4$

$D = 1$

$Q = 1$

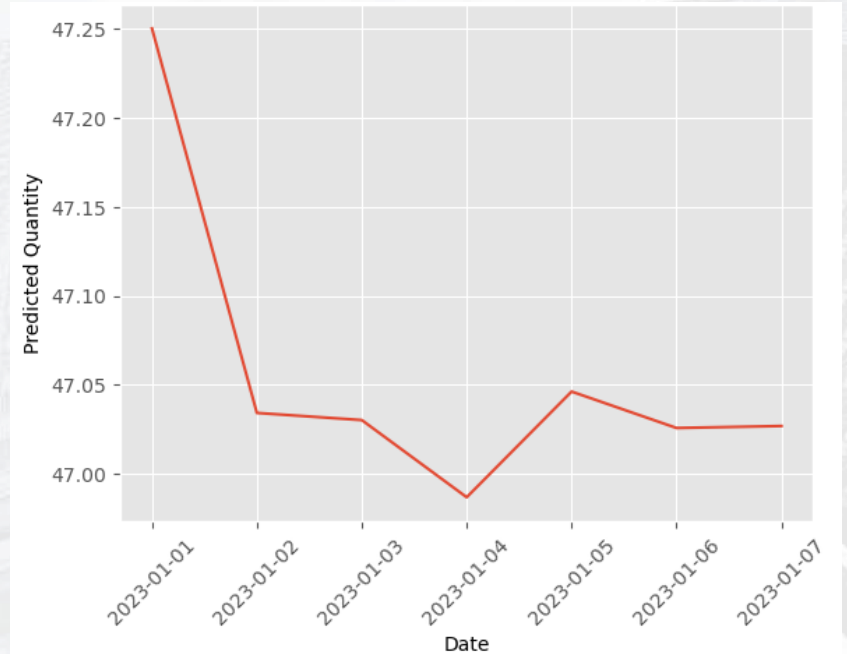
MODEL EVALUATION



Mean Squared Error : 172.19
RMSE : 13.12

Quantity Predictions for Jan 01-07, 2023

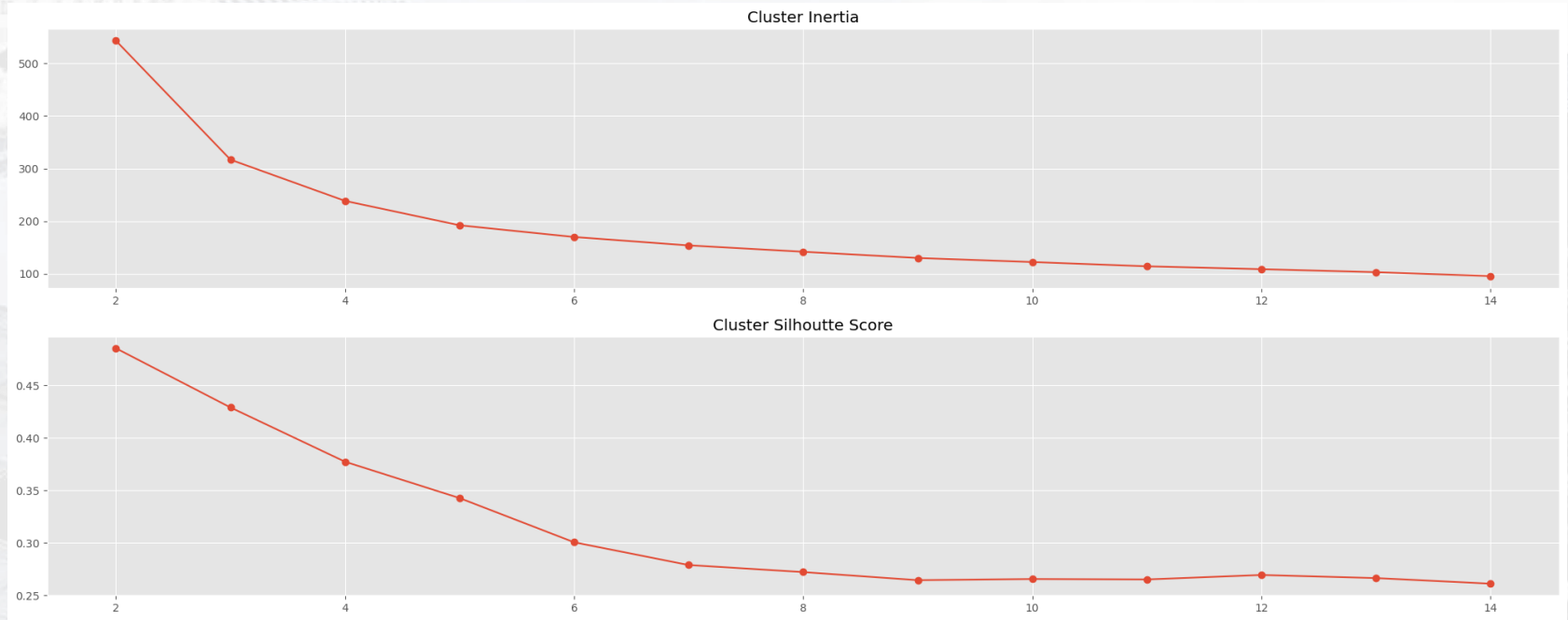
	predicted_mean
2023-01-01	47.250296
2023-01-02	47.034333
2023-01-03	47.030323
2023-01-04	46.986969
2023-01-05	47.046332
2023-01-06	47.025875
2023-01-07	47.026980



Machine Learning Clustering (K-Means)

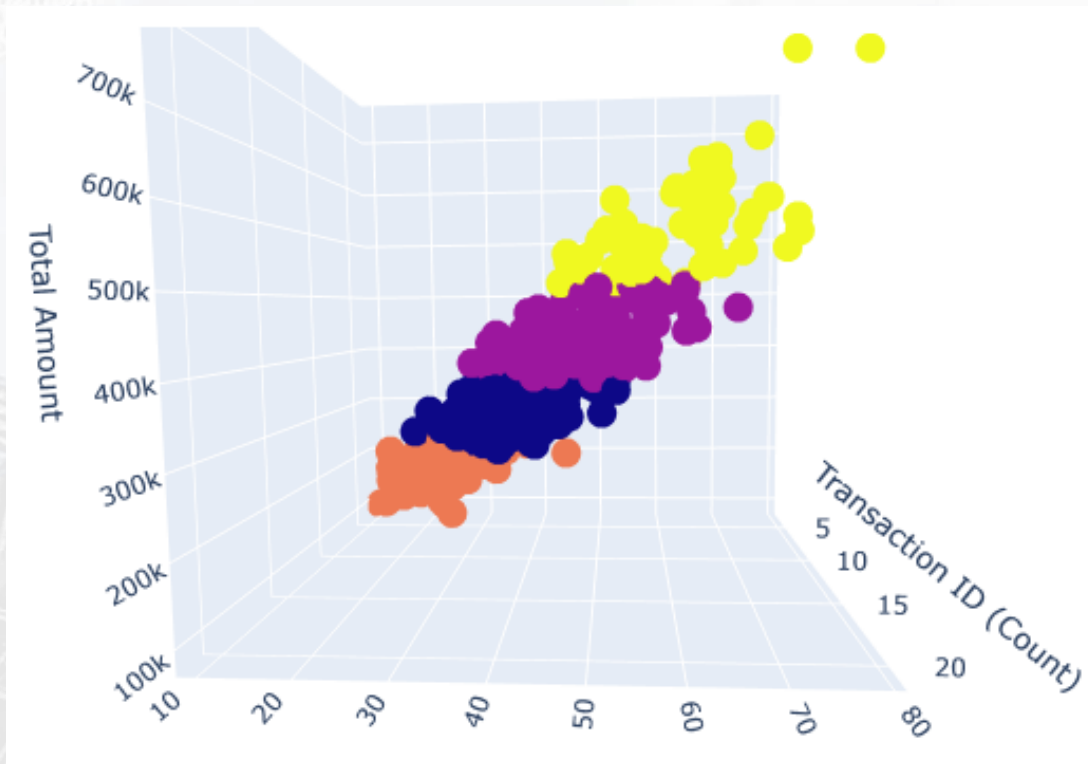


Cluster Inertia & Cluster Silhouette Score



n_cluster = 4

Customer Segmentation by Transaction ID (Count), Total Quantity, Total Amount



Result & Video Presentation



[yogaariasena/FinalTask_Kalbe_DS_Yoga-Aria-Sena \(github.com\)](https://github.com/yogaariasena/FinalTask_Kalbe_DS_Yoga-Aria-Sena)



[FinalTask_Kalbe_DS_Yoga Aria Sena - Google Drive](#)



Thank You



Rakamin
Academy



KALBE
Nutritional