

Machine Learning Project

**Kalbe Nutritional Data Scientist
Project Based Internship Program**

Presented by
Muhammad Nur Ishlaahudin Sya'bani



**Muhammad Nur
Ishlaahudin
Sya'bani**

Tentang saya

Saya mahasiswa semester akhir di Universitas Pembangunan Nasional "Veteran" Yogyakarta dengan program studi Informatika. Saya memiliki minat dibidang data. Pengalaman saya didalam bidang data masih minim tetapi saya optimis ketika mengikuti suatu kegiatan dalam bidang data yang akan menunjang karir saya kedepannya.

Pengalaman

- Mengikuti Studi Independent di Startup Campus sebagai Data Analyst
- Mengikuti Studi Independen di Bangkit 2023 by Google, GoTo, Traveloka sebagai Machine Learning Engineer
- Mengikuti SAP Analytic Cloud ASEAN Foundation sebagai Data Scientist

Case Study

Seorang Data Scientist di Kalbe Nutritionals yang sedang mendapatkan project baru dari tim inventory dan tim marketing. Dari tim inventory, Data Scientist diminta untuk dapat membantu memprediksi jumlah penjualan (quantity) dari total keseluruhan product Kalbe. Tujuannya adalah untuk mengetahui perkiraan quantity product yang terjual sehingga tim inventory dapat membuat stock persediaan harian yang cukup. Prediksi yang dilakukan harus harian. Dari tim marketing diminta untuk membuat cluster/segment customer berdasarkan beberapa kriteria. Tujuannya adalah untuk membuat segment customer. Segment customer ini nantinya akan digunakan oleh tim marketing untuk memberikan personalized promotion dan sales treatment.

Hasil

Exploration Data Analysis

1. Rata-rata umur customer jika dilihat dari marital statusnya

	ABC marital_status	123 rata_rata_umur
1		31.3333333333
2	Married	43.0382352941
3	Single	29.3846153846

Diketahui bahwa rata-rata umur untuk yang sudah menikah adalah 43,1 tahun. Sedangkan untuk yang masih single atau belum menikah memiliki rata-rata 29,3 tahun. Untuk status pernikahan yang tidak diketahui memiliki rata-rata umur 31,3 tahun.

2. Rata-rata umur customer jika dilihat dari gender nya

	123 gender	123 rata_rata_umur
1	0	40.326446281
2	1	39.1414634146

Diketahui bahwa rata-rata umur untuk wanita adalah 40,3 tahun. Sedangkan untuk laki-laki adalah 39,4 tahun. Keterangan untuk 0 adalah sebagai gender wanita, sedangkan 1 adalah sebagai gender laki-laki.

Exploration Data Analysis

3. Nama store dengan total quantity terbanyak

	ASC store_name ▼	123 total_quantity ▼
1	Lingga	2,777

Toko dengan nama Lingga menjadi toko dengan total quantity terbanyak dari semua toko.

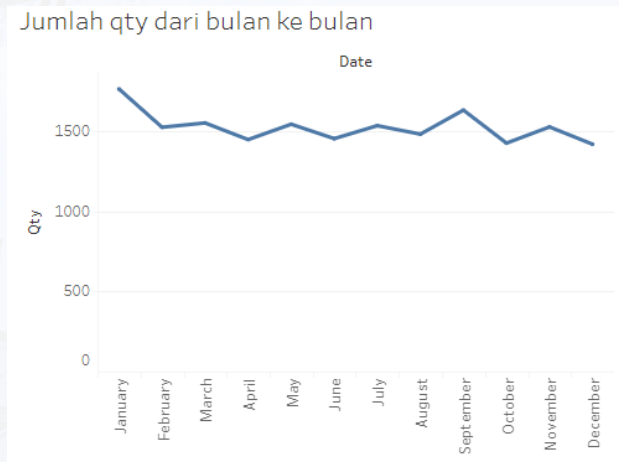
4. Nama produk terlaris dengan total amount terbanyak

	ASC product_name ▼	123 total_amount ▼
1	Cheese Stick	27,615,000

Produk yang paling banyak dibeli adalah cheese stick dengan jumlah total amount Rp 27.615.000 .

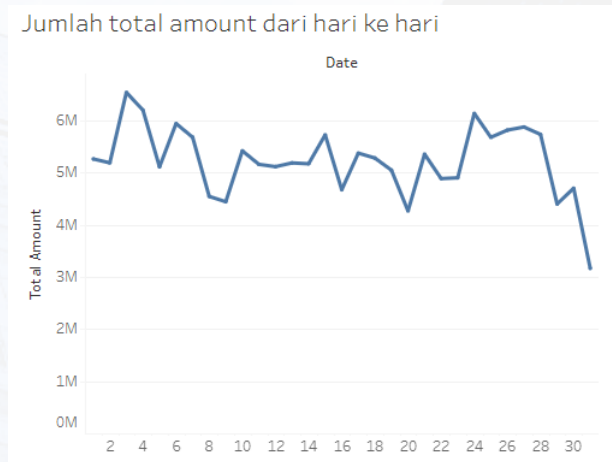
Visualisasi Data

1. Jumlah quantity dari bulan ke bulan



Jumlah quantity dari bulan ke bulan cenderung naik turun atau fluktuatif. Ini ditunjukkan dari bulan Januari ke Februari yang turun tetapi dari Februari ke Maret ada kenaikan.

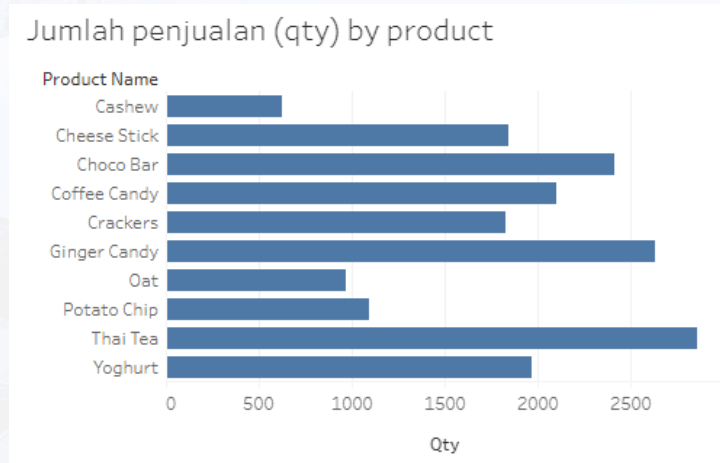
2. Jumlah total amount dari hari ke hari



Jumlah total amount setiap harinya naik turun atau fluktuatif. Bisa dilihat dari tanggal 1 ke 2 ada penurunan tetapi dari tanggal 2 ke 3 ada kenaikan yang cukup signifikan.

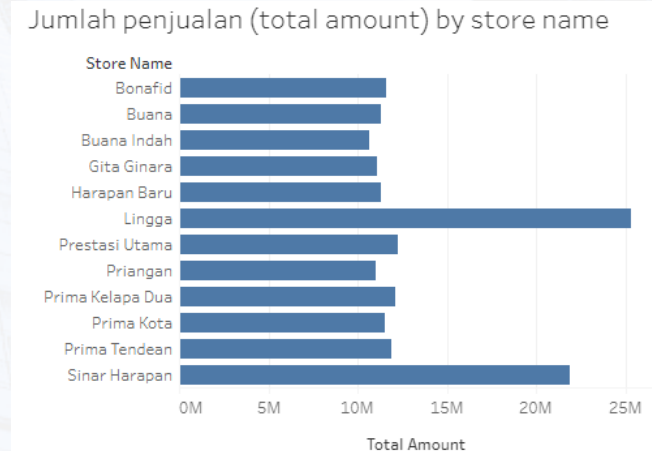
Visualisasi Data

3. Jumlah penjualan (qty) by product



Jumlah penjualan atau quantity paling banyak yaitu Thai Tea dengan 2.853 Thai Tea yang terjual. Sedangkan penjualan paling sedikit adalah Cashew dengan total penjualan 627.

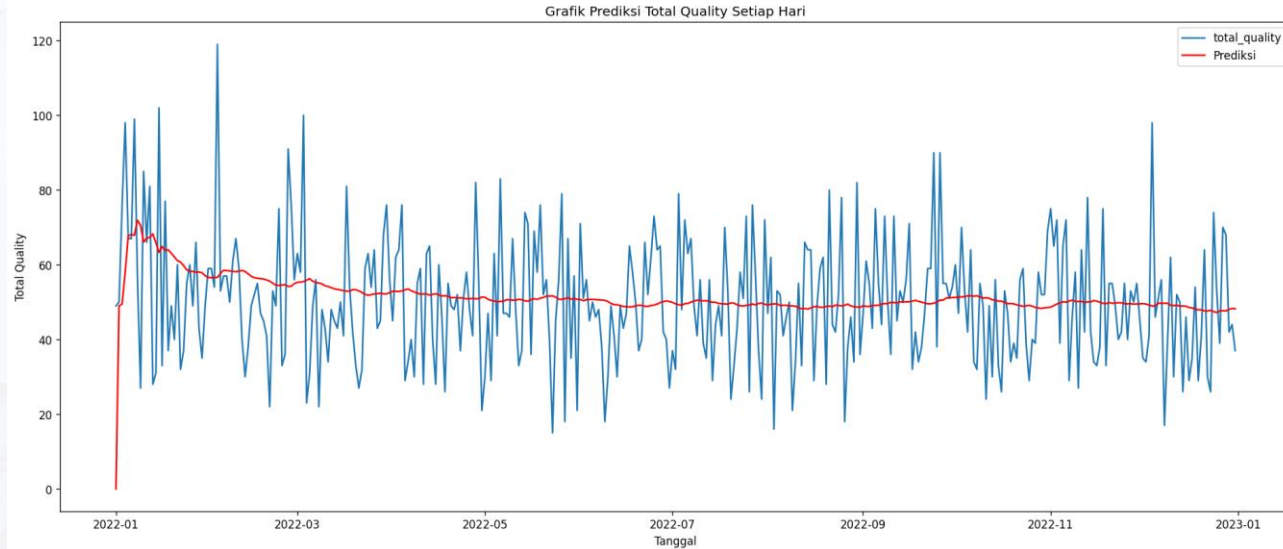
4. Jumlah penjualan (total amount) by store name



Toko dengan penjualan terbanyak adalah Harapan Baru dengan total penjualan Rp 25.294.100 . Sedangkan toko dengan penjualan paling sedikit yaitu Buana Indah dengan total Rp 10.629.900 .

Machine Learning Regression (Time Series)

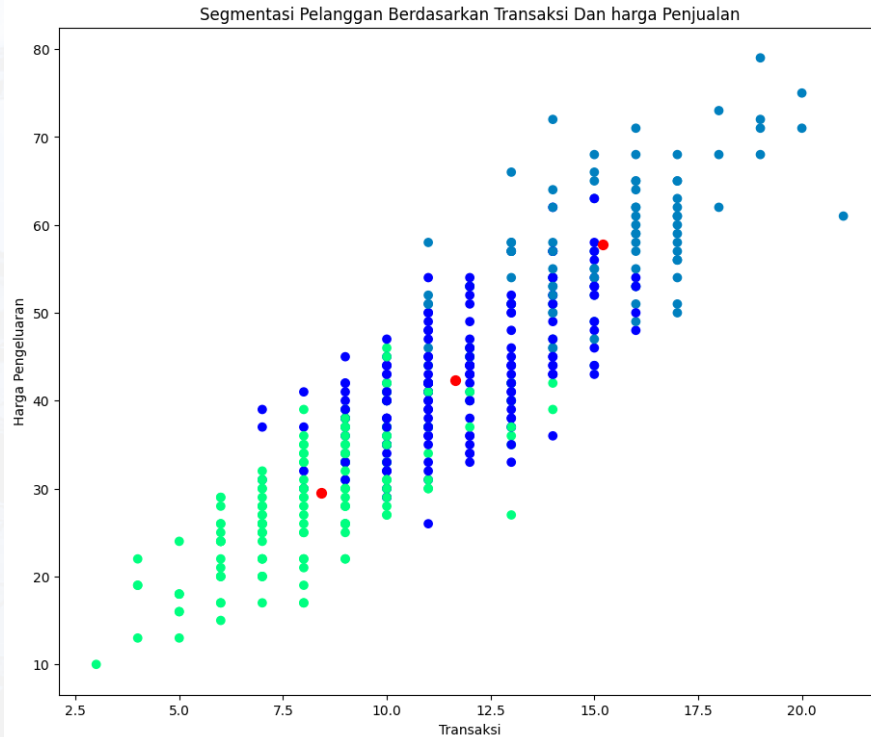
Metode ARIMA



Dari grafik visualisasi menggunakan metode time series ARIMA, dapat disimpulkan bahwa jumlah penjualan (total quantity) dari hari ke hari cenderung fluktuatif atau naik turun. Naik turunnya penjualan juga tidak terlalu signifikan. Penjualan juga melampaui prediksi yang telah dibuat. Artinya penjualan tersebut sudah bagus. Mungkin masih ada perbaikan untuk tim marketing agar penjualan naik sedikit demi sedikit.

Machine Learning Clustering

Metode Kmeans Clustering



Dari grafik visualisasi menggunakan metode clustering yaitu K-Means Clustering, dapat disimpulkan bahwa segmentasi pelanggan dikelompokkan menjadi 3 kelompok. Masing-masing kelompok sudah memiliki centroid atau pusatnya dengan warna merah.

1. Kelompok pertama yaitu warna hijau muda yang mana transaksinya antara 0 sampai 13 transaksi dan untuk harga penjualan berkisar antara 0 sampai 450.000 .
2. Kelompok kedua yaitu warna biru yang mana transaksinya antara 7 sampai 16 dan harga penjualannya antara 250.000 sampai 610.000 .
3. Kelompok ketiga yaitu warna cyan dengan transaksi antara 10,5 sampai 21 dan harga penjualannya antara 450.000 sampai 790.000 .

Link Dokumentasi

Link Github :

<https://github.com/Ishlaah/Project-Kalbe-Nutritionals.git>

Link Tableau :

<https://public.tableau.com/app/profile/ishlaahudin.sya.bani/viz/ProjectKableNutritionals/Dashboard1>

Link Collabs :

https://colab.research.google.com/drive/1RJFZq7aYogZh_Zt0eU0vq_ZSpmBYNLLH?usp=sharing

<https://colab.research.google.com/drive/1b8g0mBBcmXIKDn6lI4ySe-TtIWkYabdY?usp=sharing>

Link Video Presentasi

Link Youtube : <https://youtu.be/9QSRiPZWgQE>

Thank You



Rakamin
Academy



KALBE
Nutritional