

Retail Store Data Analysis

Final Project Intermediate - Data Analyst #25

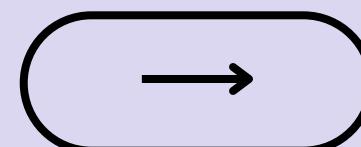
Nicholas Glenn Manurung





Latar Belakang

Dalam era digital saat ini, bisnis ritel menghadapi tantangan dan peluang besar dalam memahami perilaku konsumen yang dinamis. Dengan maraknya e-commerce dan meningkatnya ekspektasi pelanggan, perusahaan harus mampu mengidentifikasi produk unggulan, mempertahankan loyalitas pelanggan, serta mengoptimalkan strategi harga dan promosi. Oleh karena itu, analisis data penjualan dilakukan untuk menggali pola-pola penting dalam perilaku konsumen



Tujuan Analisis



Data Explanation

Analisis ini menggunakan dataset SalesTransaction.csv yang memiliki beberapa kolom sebagai berikut:

**TransactionNo**

Nomor unik untuk setiap transaksi penjualan

Date

Tanggal terjadinya transaksi

ProductNo

Kode unik untuk setiap produk yang dijual

ProductName

Nama atau deskripsi produk yang dijual

Price

Harga satuan produk saat transaksi

Quantity

Jumlah unit produk yang dibeli dalam transaksi

CustomerNo

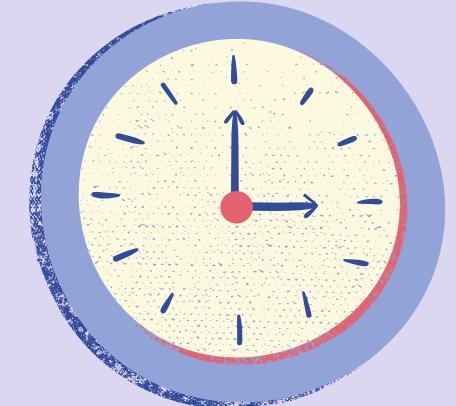
ID unik untuk setiap pelanggan

Country

Negara asal dari pelanggan

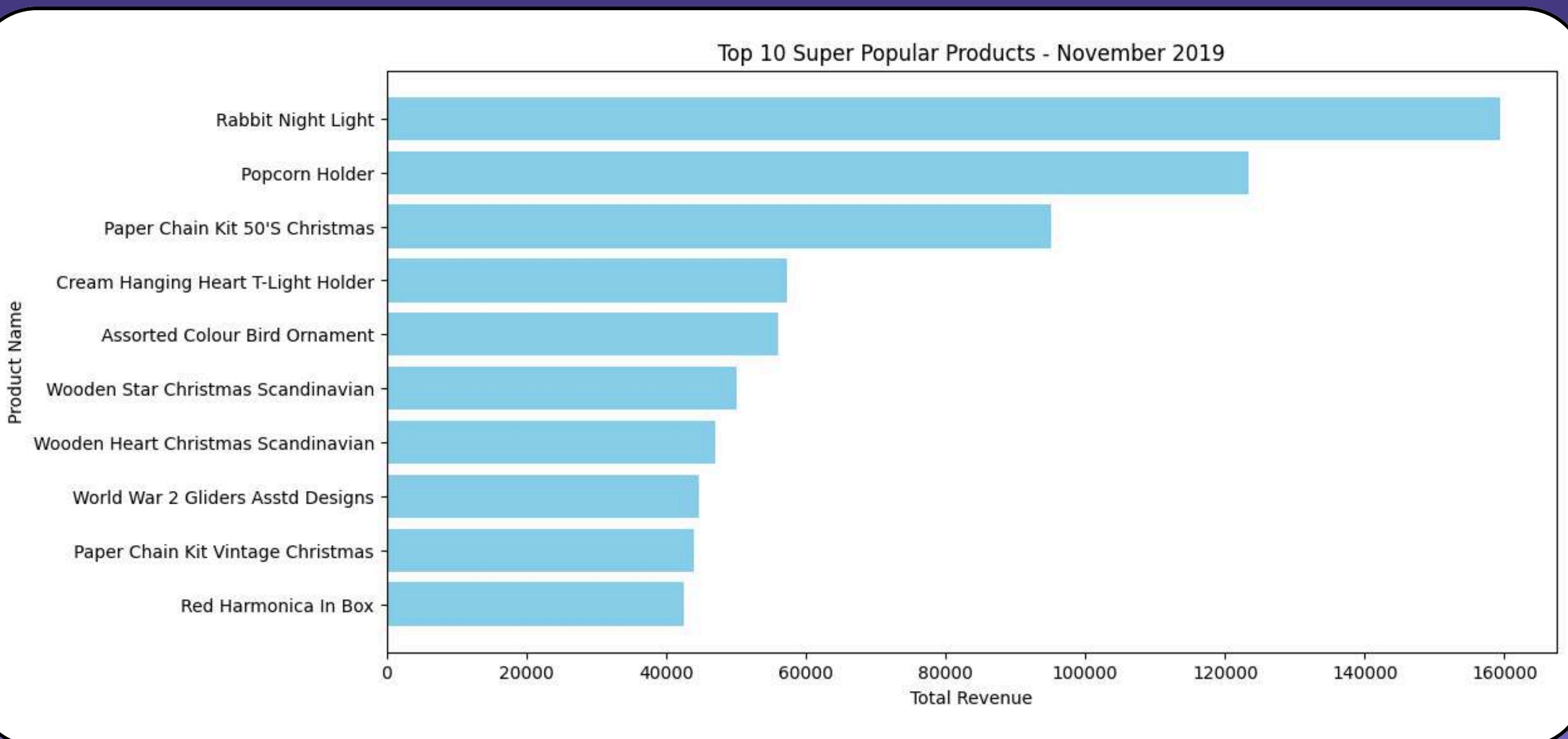


Analisis



Top Contributing products

Top 10 Tipe barang masuk ke kategori super populer pada bulan November 2019



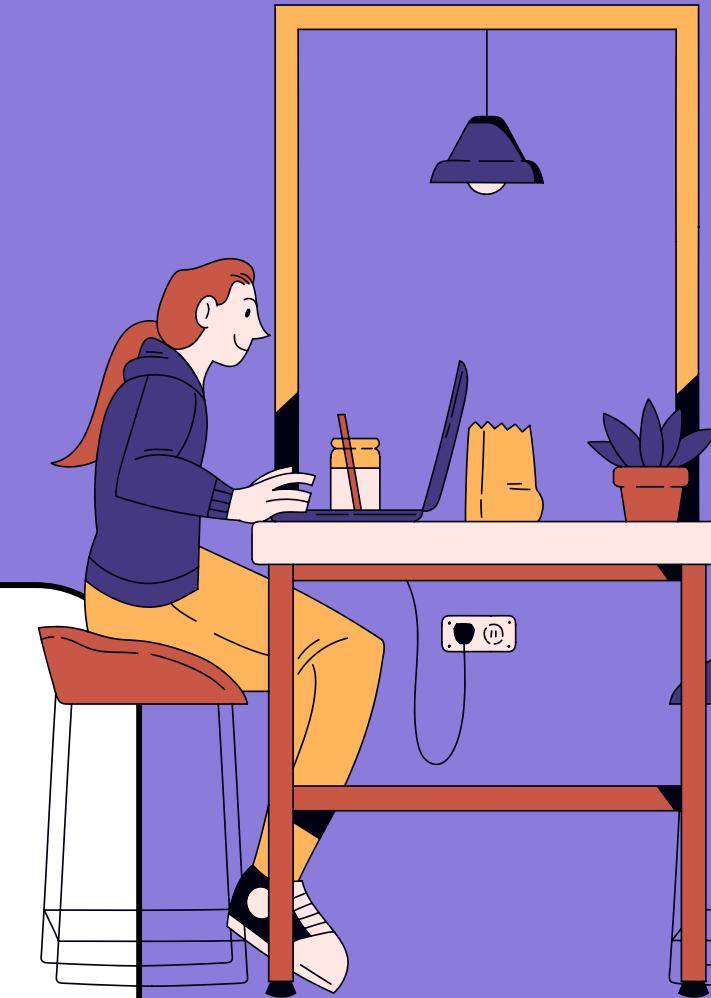
- Bisa kita lihat bahwa "Rabbit Night Light" merupakan Produk dengan Revenue terbanyak di November 2019
- Berdasarkan analisis nama produk, banyak di antaranya memiliki nuansa musiman, terutama bertema Natal dan liburan akhir tahun, seperti Christmas Decoration.



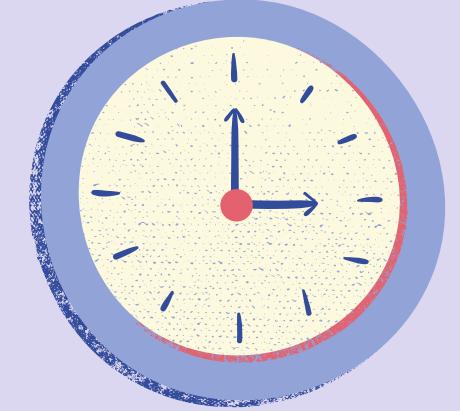
Analisis

Segmentasi produk

1. Total Revenue dihitung dari hasil perkalian antara harga satuan dan jumlah pembelian suatu produk.
2. Produk diklasifikasikan berdasarkan volume transaksi:
 - 20% teratas = Popular,
 - 20–80% = Normal,
 - sisanya = Low.
3. Produk juga diklasifikasikan berdasarkan total revenue:
 - 20% teratas = Popular,
 - 20–80% = Normal,
 - sisanya = Low.
4. Segmentasi akhir produk ditentukan dari kombinasi dua klasifikasi di atas:
 - Super Popular = Popular di volume & revenue,
 - Popular = Salah satu Popular, yang lain Normal,
 - Low = Keduanya Low,
 - Sisanya = Normal.



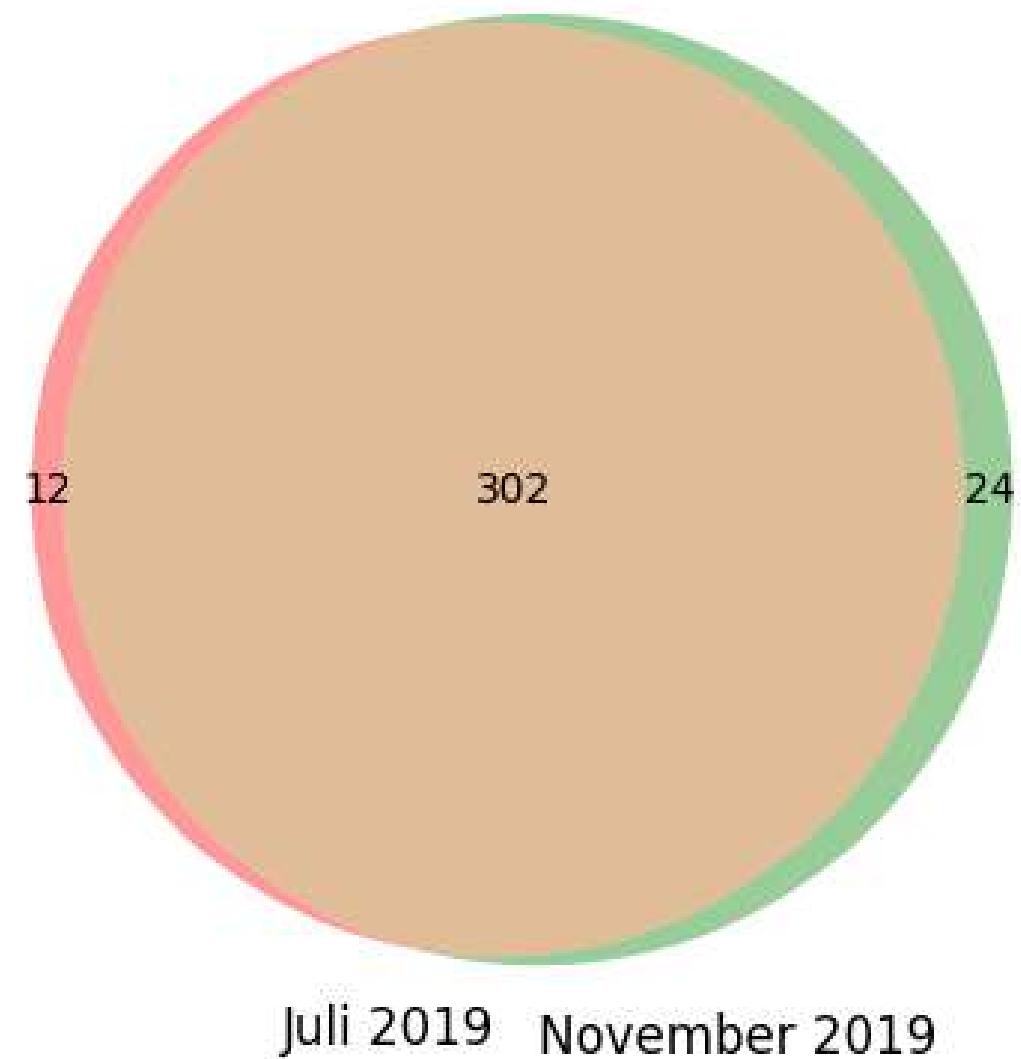
Analisis



Top Contributing products

Barang yang populer di November 2019 yang sama dengan barang yang populer di Juli 2019

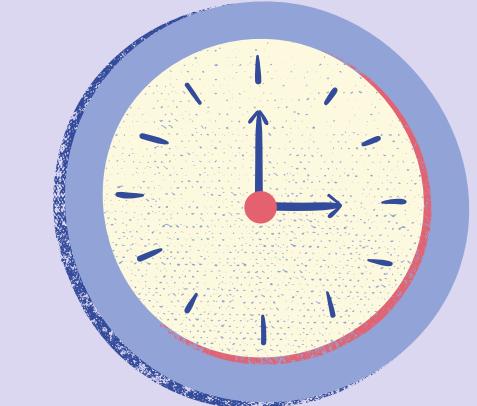
Diagram Venn Produk Populer Juli vs November 2019



- Dari total produk populer di November 2019 dan Juli 2019, terdapat 302 produk yang sama. Artinya, sebagian besar produk populer di bulan November 2019 juga sudah populer sejak bulan Juli 2019.
- Hanya terdapat 24 produk baru di November 2019 yang tidak muncul sebagai populer di Juli, dan 12 produk di Juli 2019 yang tidak lagi masuk daftar populer di November. Ini menunjukkan bahwa pergeseran minat konsumen tidak terlalu drastis.

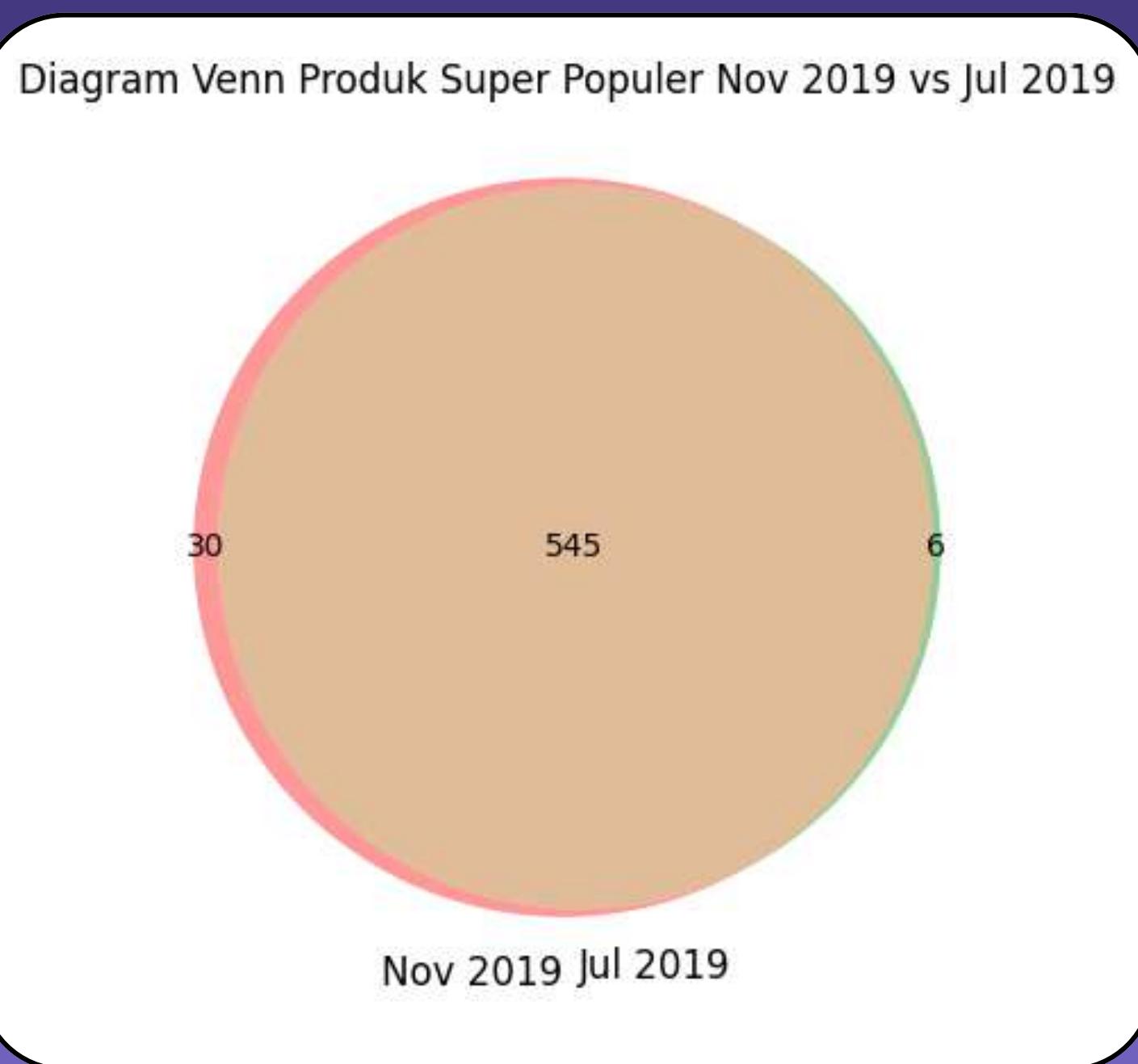


Analisis



Top Contributing products

Barang yang masuk kategori super populer di November 2019 yang sama dengan barang super populer di July 2019

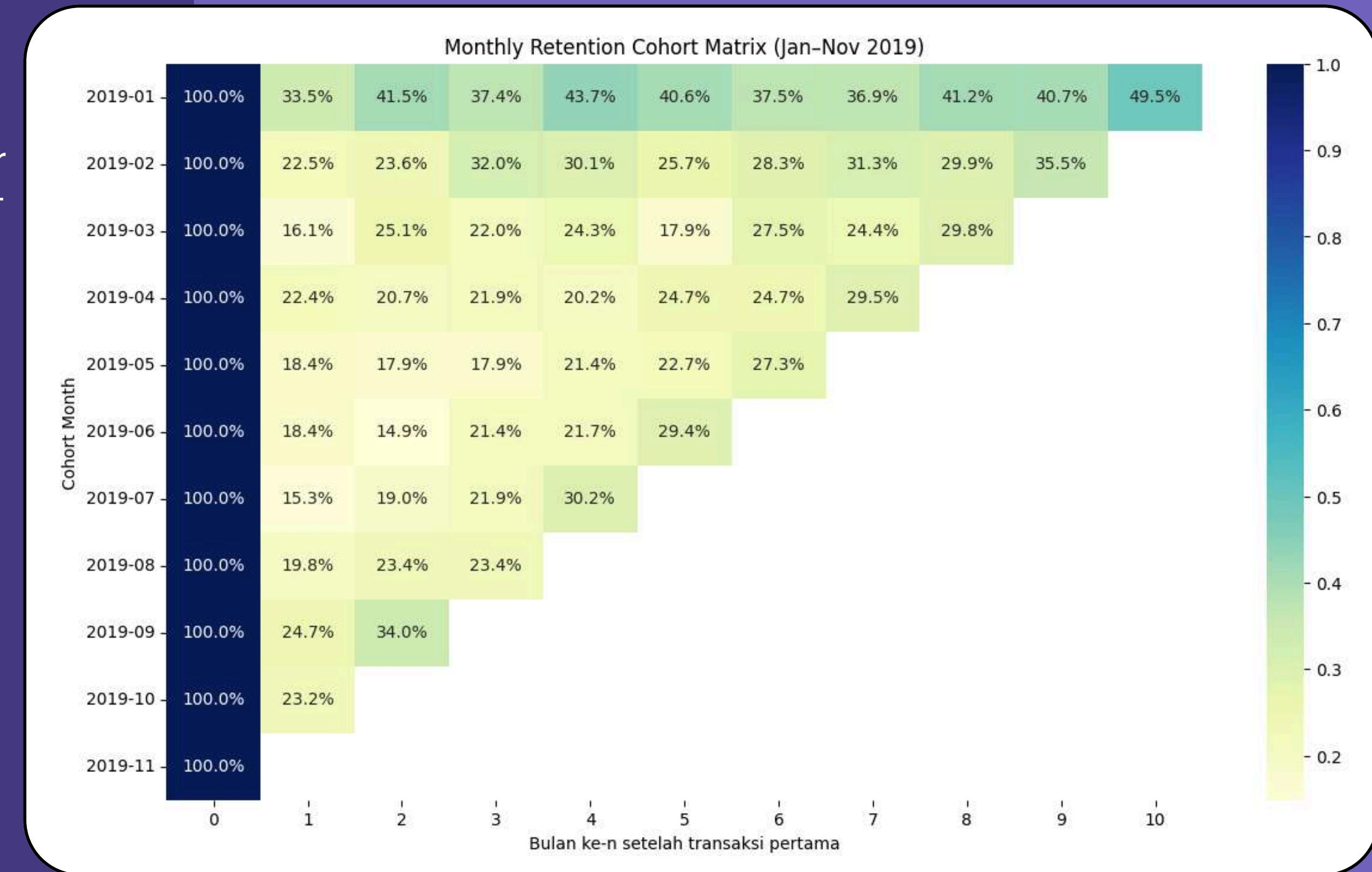


- Terdapat 545 produk yang tetap berada dalam kategori Super Populer di kedua bulan, menunjukkan adanya konsistensi preferensi pelanggan terhadap produk-produk tertentu selama rentang waktu tersebut.
- Sedikit perbedaan terjadi antar bulan. Hanya terdapat 30 produk Super Populer di November yang tidak termasuk dalam Super Populer di Juli & 6 produk Super Populer di Juli yang tidak termasuk dalam Super Populer di November.

Analisis

Retention Cohort Analysis

Matriks Kohort Retensi per
Bulan (Januari – November 2019):



Analisis

Retention Cohort Analysis

Karakteristik Pertumbuhan Customer Baru

Dari Matriks Retensi Cohort sebelumnya, kita bisa melihat banyaknya customer baru yang pertama kali bertransaksi tiap bulan:

Jan: 827

Feb: 569

Mar: 610

Apr: 397

Mei: 374

Jun: 309

Jul: 242

Agu: 192

Sep: 312

Okt: 353

Nov: 334



- Terlihat bahwa jumlah customer baru menurun secara bertahap dari Januari (827) hingga Juli (242), menunjukkan tren perlambatan akuisisi customer baru di semester pertama.
- Namun, mulai Agustus hingga November, terdapat sedikit peningkatan kembali ($192 \rightarrow 312 \rightarrow 353 \rightarrow 334$), yang bisa mengindikasikan adanya produk yang disukai yang dijual di akhir tahun.

Analisis

Retention Cohort Analysis

Bulan dengan retention paling tinggi

Retention tertinggi dilihat dari persentase customer yang tetap kembali di bulan ke-1, ke-2, dst. Kita fokus ke bulan ke-1 dan bulan ke-2 sebagai indikator retensi jangka pendek. Dari Matriks Retensi Cohort sebelumnya, kita bisa melihat bulan di tahun 2019 dengan retensi tertinggi di samping:

- Januari 2019 menghasilkan retensi paling tinggi baik di bulan ke-1 (33.5%) maupun bulan ke-2 (41.5%).
- September 2019 juga menunjukkan performa yang sangat baik di bulan ke-2 (34%).

1. Januari

- Retensi Bulan ke-1: 33.5%
- Retensi Bulan ke-2: 41.5%

2. Februari

- Retensi Bulan ke-1: 22.5%
- Retensi Bulan ke-2: 23.6%

3. Maret

- Retensi Bulan ke-1: 16.1%
- Retensi Bulan ke-2: 25.1%

4. April

- Retensi Bulan ke-1: 22.4%
- Retensi Bulan ke-2: 20.7%

5. Mei

- Retensi Bulan ke-1: 18.5%
- Retensi Bulan ke-2: 17.9%

6. Juni

- Retensi Bulan ke-1: 18.5%
- Retensi Bulan ke-2: 14.9%

7. Juli

- Retensi Bulan ke-1: 15.3%
- Retensi Bulan ke-2: 19.0%

8. Agustus

- Retensi Bulan ke-1: 19.8%
- Retensi Bulan ke-2: 23.4%

9. September

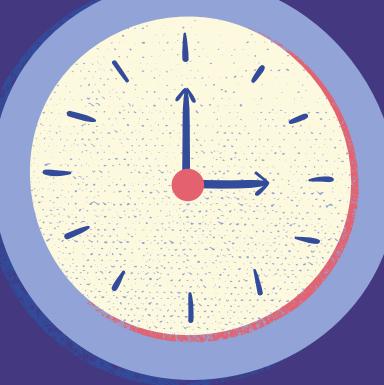
- Retensi Bulan ke-1: 24.7%
- Retensi Bulan ke-2: 34.0%

10. Oktober

- Retensi Bulan ke-1: 23.2%
- Retensi Bulan ke-2: -

11. November

- Retensi Bulan ke-1: -
- Retensi Bulan ke-2: -





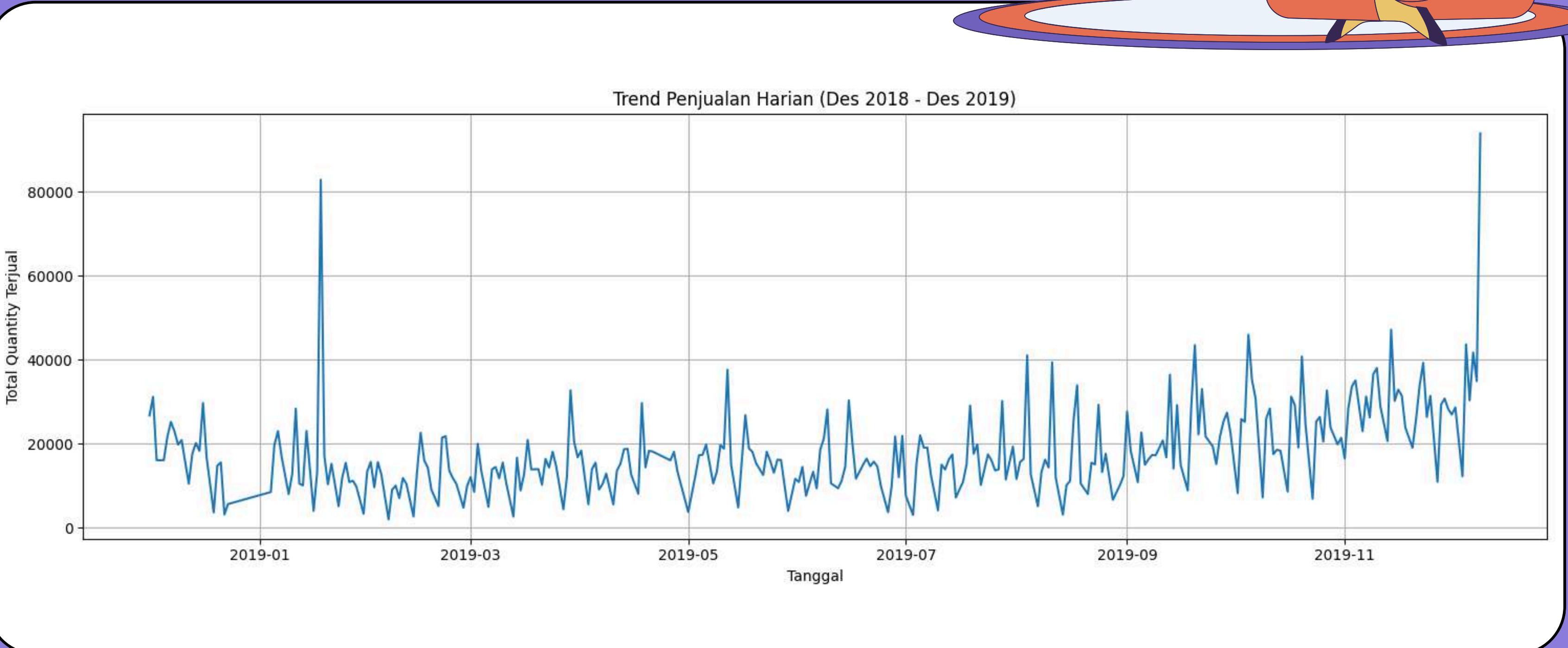
Analisis

Time Series Analysis

Grafik Trend Penjualan Harian (Des 2018 – Des 2019)



Trend Penjualan Harian (Des 2018 - Des 2019)



Analisis

Time Series Analysis

Trend penjualan harian



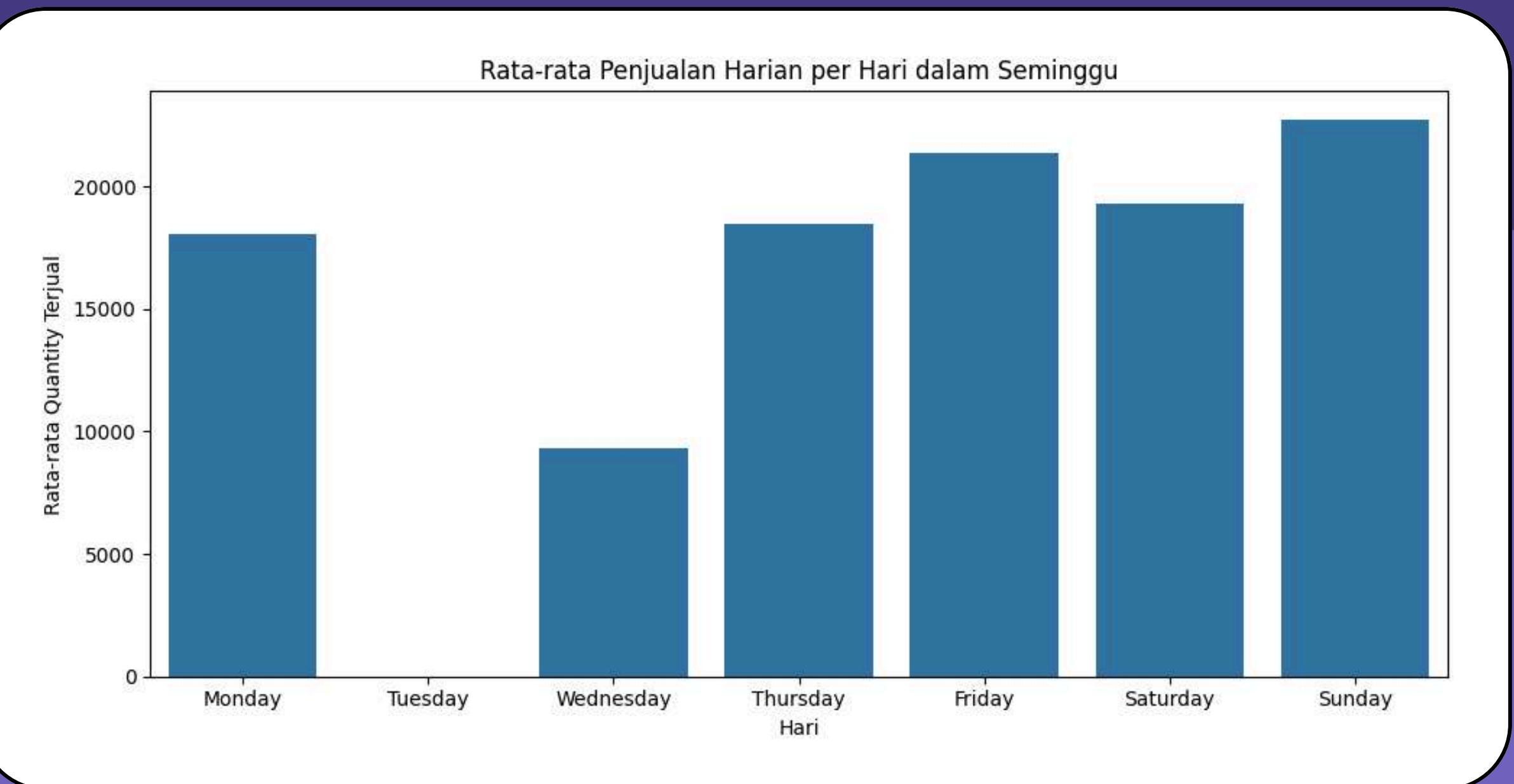
- Grafik menunjukkan adanya tren kenaikan penjualan harian, terutama dalam periode sekitar September 2019 hingga Desember 2019
- Rata-rata penjualan harian adalah sekitar 18.278 unit, dengan nilai maksimum mencapai 93.925 unit di bulan Desember 2019
- Tidak terlihat pola penurunan atau stagnasi yang signifikan — menunjukkan pertumbuhan penjualan yang positif dan sehat sepanjang tahun.
- Terjadi lonjakan tren quantity pada tanggal 18 Januari 2019 dan 9 Desember 2019 dengan quantity di atas 80.000 unit



Analisis

Time Series Analysis

Trend penjualan per harinya dari Senin hingga Minggu



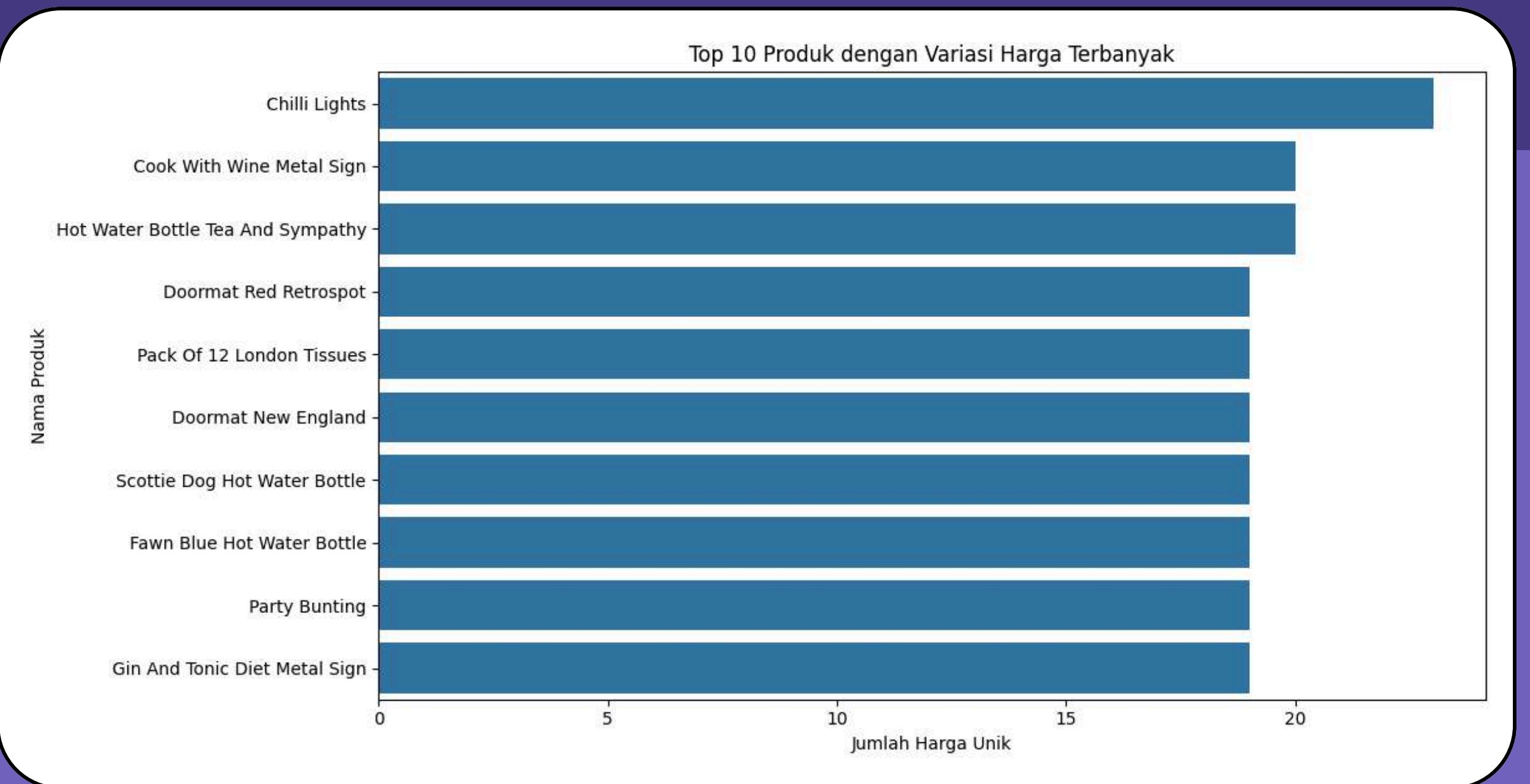
- Hari dengan rata-rata penjualan tertinggi adalah Minggu dan Jumat (masing-masing sekitar >20.000 unit).
- Hari dengan penjualan paling rendah adalah Rabu (sekitar <10.000 unit)
- Data Selasa (Tuesday) hilang atau tidak terekam, sehingga perlu diperiksa ulang data mentahnya.



Analisis

Price Elasticity

Top 10 produk yang terjual dengan setidaknya 10 variasi harga yang berbeda.



- Produk dengan variasi harga terbanyak adalah "Chilli Lights" dengan 23 variasi harga berbeda, diikuti oleh "Cook With Wine Metal Sign" dan "Hot Water Bottle Tea And Sympathy", masing-masing dengan 20 variasi harga.
- Mayoritas dari produk dengan variasi harga tinggi adalah produk dekoratif atau hadiah musiman, seperti lampu, signage logam, botol air panas, dan aksesoris rumah.



Analisis

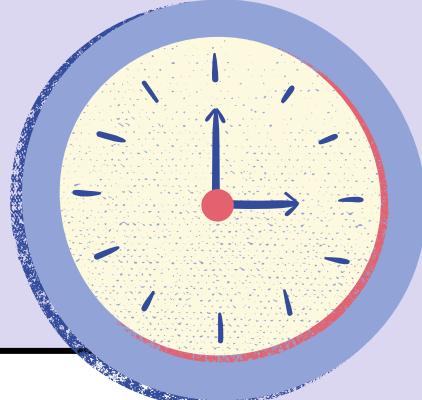
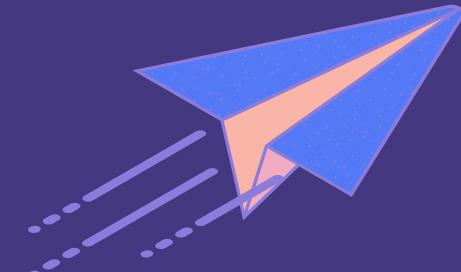
Price Elasticity

Korelasi antara harga dengan quantity untuk setiap barang

1. Dari 700 produk yang memiliki setidaknya 10 variasi harga, terdapat beberapa produk yang menunjukkan korelasi negatif kuat antara harga dan jumlah pembelian. Produk dengan korelasi paling negatif adalah:

- Set Of 2 Tea Towels Ping Microwave (-0.476)
- Squarecushion Cover Pink Union Jack (-0.462)
- Pizza Plate In Box (-0.450)

2. Nilai korelasi negatif menunjukkan bahwa ketika harga produk naik, kuantitas penjualan turun secara signifikan, menandakan bahwa produk tersebut sangat sensitif terhadap perubahan harga atau memiliki price elasticity tinggi.



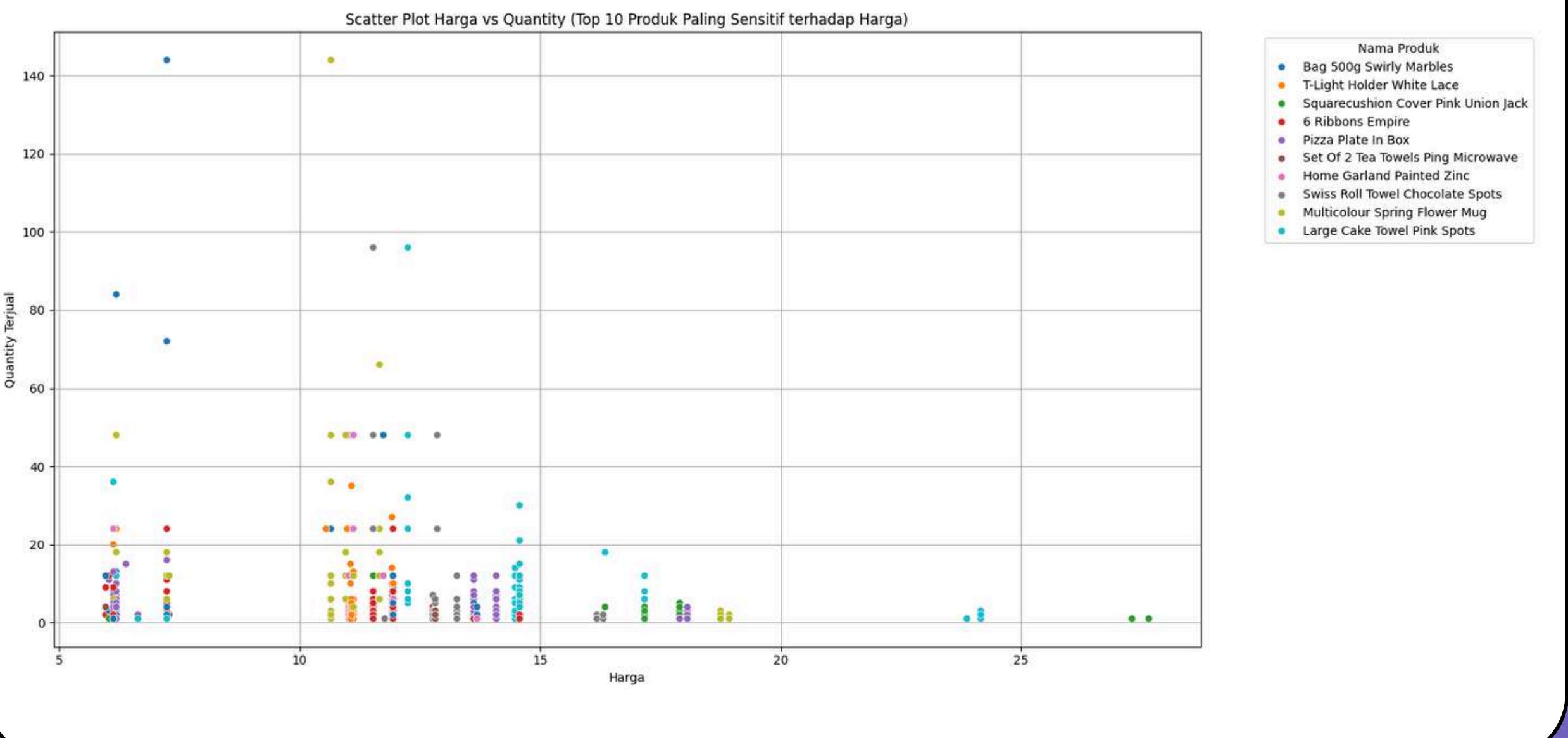
ProductName	Correlation
Set Of 2 Tea Towels Ping Microwave	-0.476031
Squarecushion Cover Pink Union Jack	-0.461947
Pizza Plate In Box	-0.449962
Large Cake Towel Pink Spots	-0.439484
Bag 500g Swirly Marbles	-0.436841
Multicolour Spring Flower Mug	-0.411175
Swiss Roll Towel Chocolate Spots	-0.400704
T-Light Holder White Lace	-0.392285
Home Garland Painted Zinc	-0.351573
6 Ribbons Empire	-0.348786



Analisis

Price Elasticity

Scatter plot antara harga dan quantity



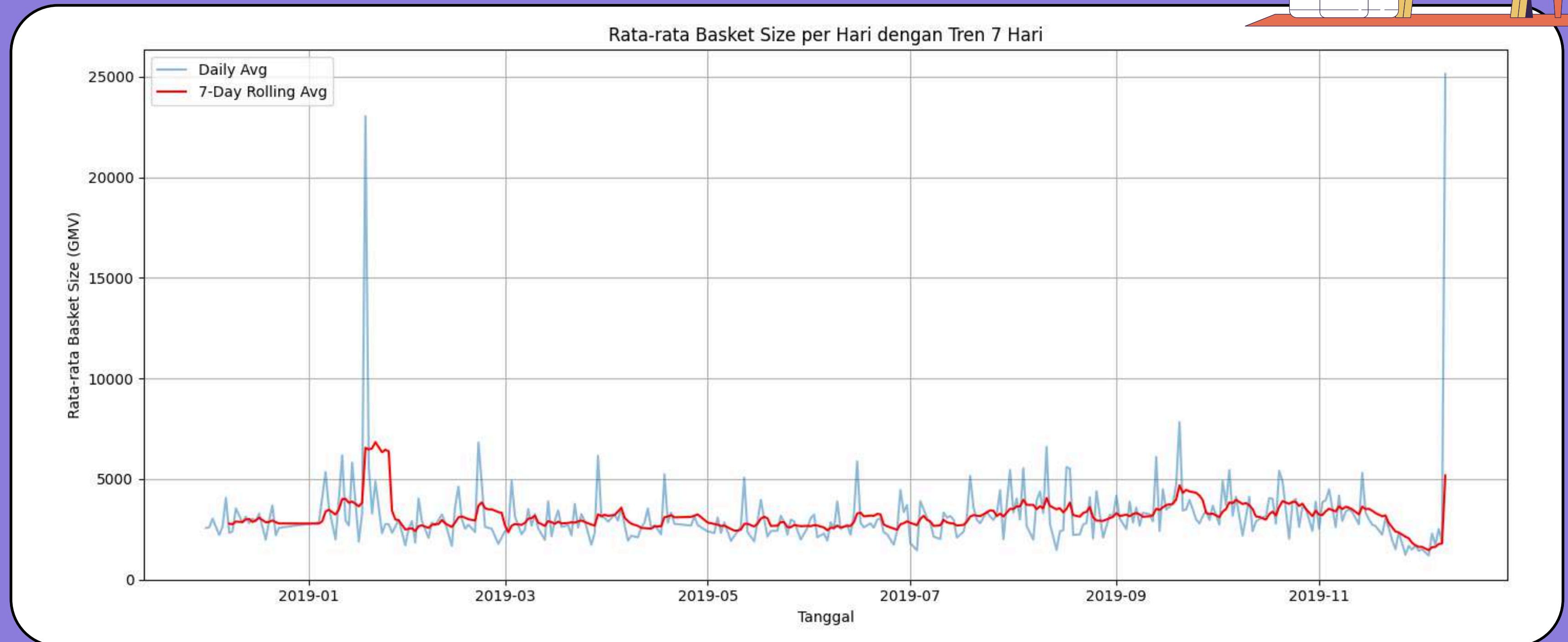
- Produk-produk yang masuk dalam Top 10 sensitif terhadap harga cenderung menunjukkan penurunan quantity yang signifikan ketika harga naik.
- Scatter plot menunjukkan pola menurun pada banyak titik antara harga dan quantity.
- Apabila terjadi kenaikan harga, maka jumlah unit terjual (quantity) cenderung menurun tajam.



Analisis

Basket Size Analysis

Grafik Rata-rata Basket Size per Hari dengan Tren 7 Hari (Des 2018 – Des 2019)



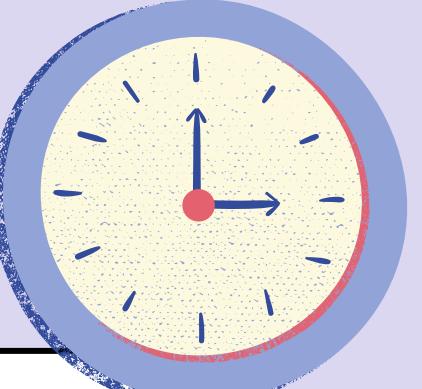
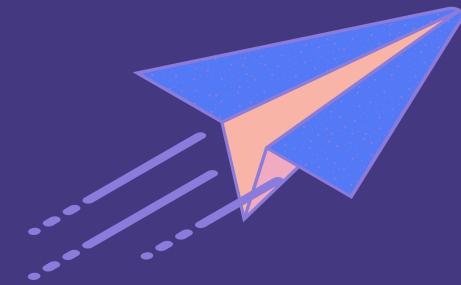


Analisis

Basket Size Analysis

Rata-rata basket size dari seluruh transaksi per harinya

- Rata-rata basket size per hari adalah sekitar 3.212
- Berdasarkan grafik sebelumnya, basket size harian (daily avg) sangat fluktuatif dengan banyak lonjakan tajam dan penurunan mendadak. Ini menunjukkan bahwa beberapa hari memiliki transaksi dengan nilai sangat besar (outlier), kemungkinan karena promo besar, campaign, atau event tertentu.
- Rentang nilai yang lebar dengan yang terendah sekitar 1.190 dan yang tertinggi mencapai 25.137. Perbedaan signifikan antara kuartil bawah ($Q_1 = 2.405$) dan kuartil atas ($Q_3 = 3.496$) juga menunjukkan fluktuasi harian yang cukup besar.



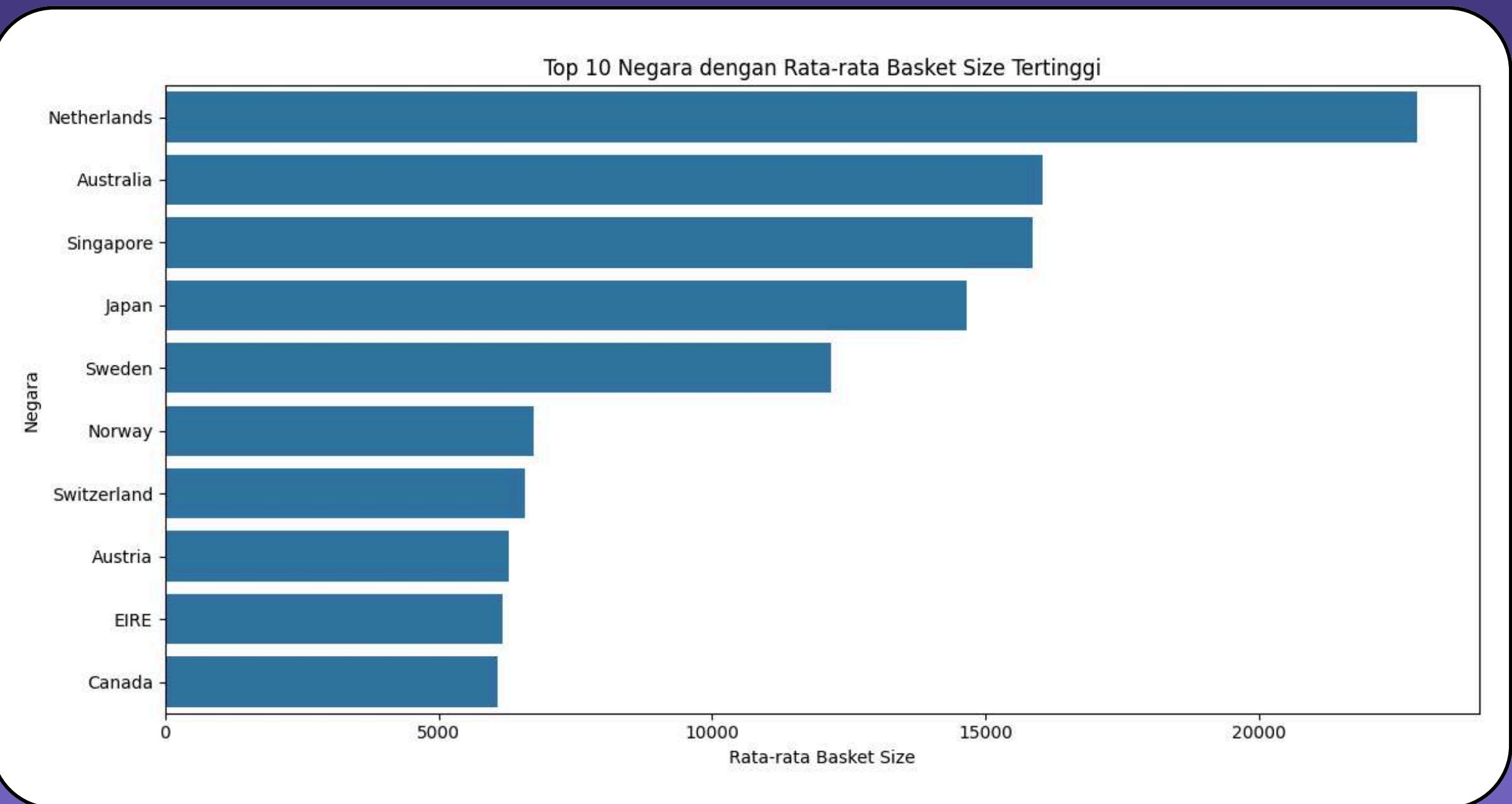
AverageBasketSize	
count	305.000000
mean	3212.121463
std	1987.263136
min	1189.638939
25%	2404.940141
50%	2886.865041
75%	3495.957397
max	25137.226591



Analisis

Basket Size Analysis

Perbedaan ukuran basket dari negara yang berbeda



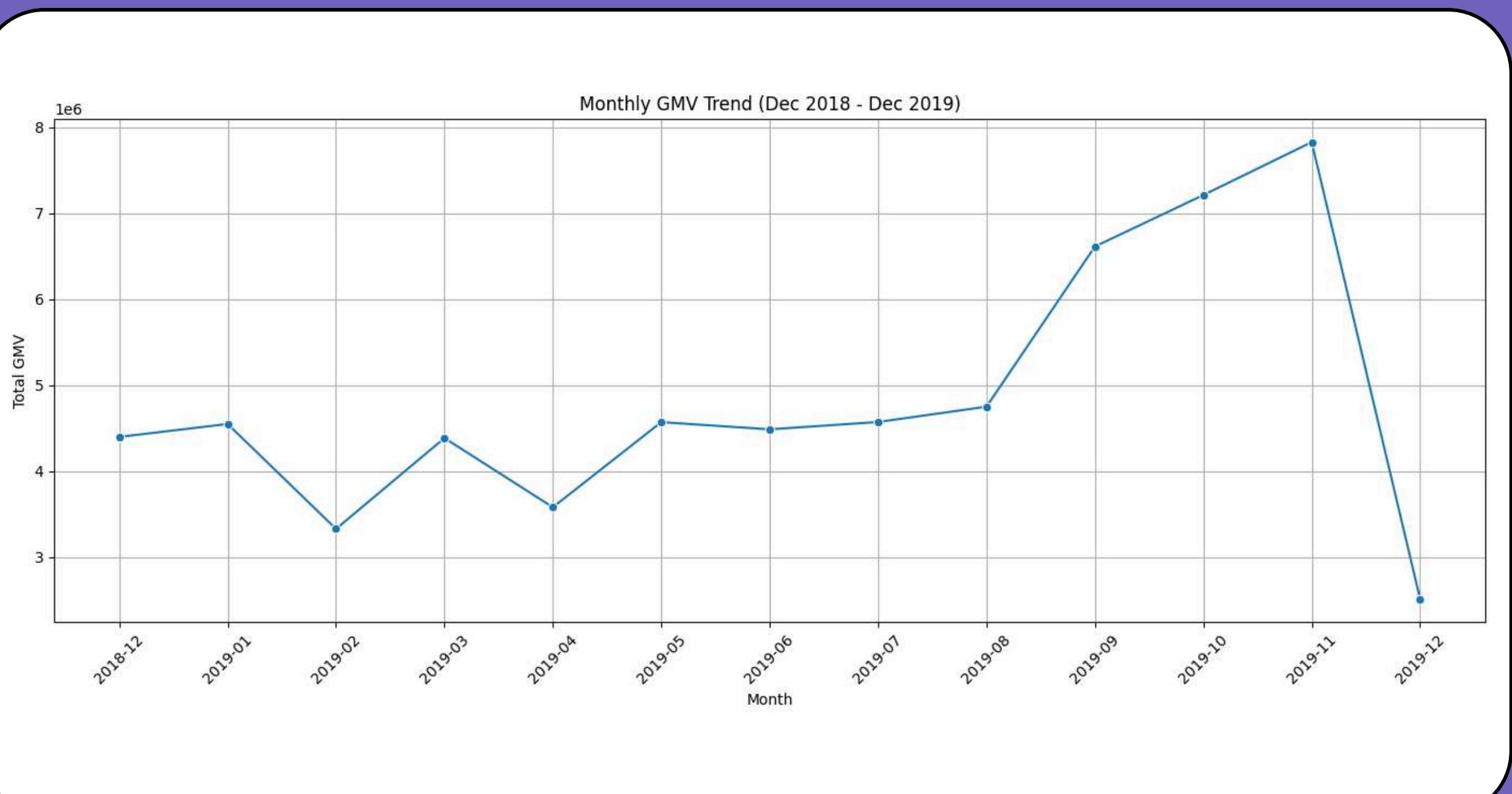
- Dari Grafik, kita bisa melihat bahwa terdapat perbedaan yang cukup signifikan dalam basket size dari negara yang berbeda.
- Negara seperti Belanda, Australia, dan Singapura memiliki basket size yang jauh lebih besar dibandingkan negara seperti EIRE atau Kanada.



Analisis

Crucial Questions

Tren Sales Tiap Bulan



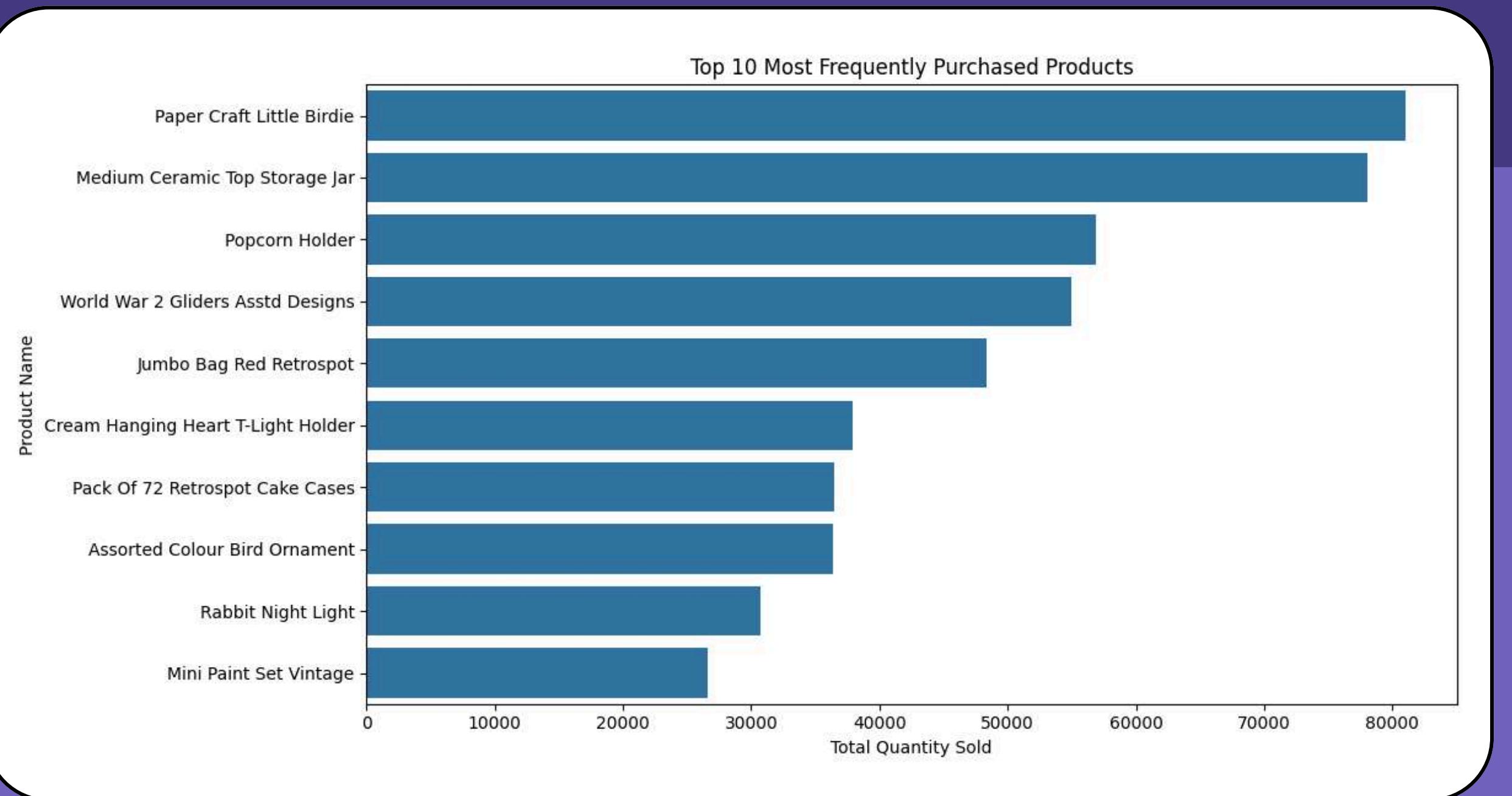
- Bulan penjualan tertinggi terjadi di November 2019 (Quantity = 760.175, GMV = Rp7.828.489,53) & Oktober 2019 dan September 2019 juga menunjukkan lonjakan signifikan.
- Bulan dengan GMV terendah adalah Desember 2019, menurun drastis setelah puncak November.
- Tren meningkat stabil dari Januari hingga puncak di November 2019, lalu drop tajam di Desember.



Analisis

Crucial Questions

Produk yang paling sering dibeli



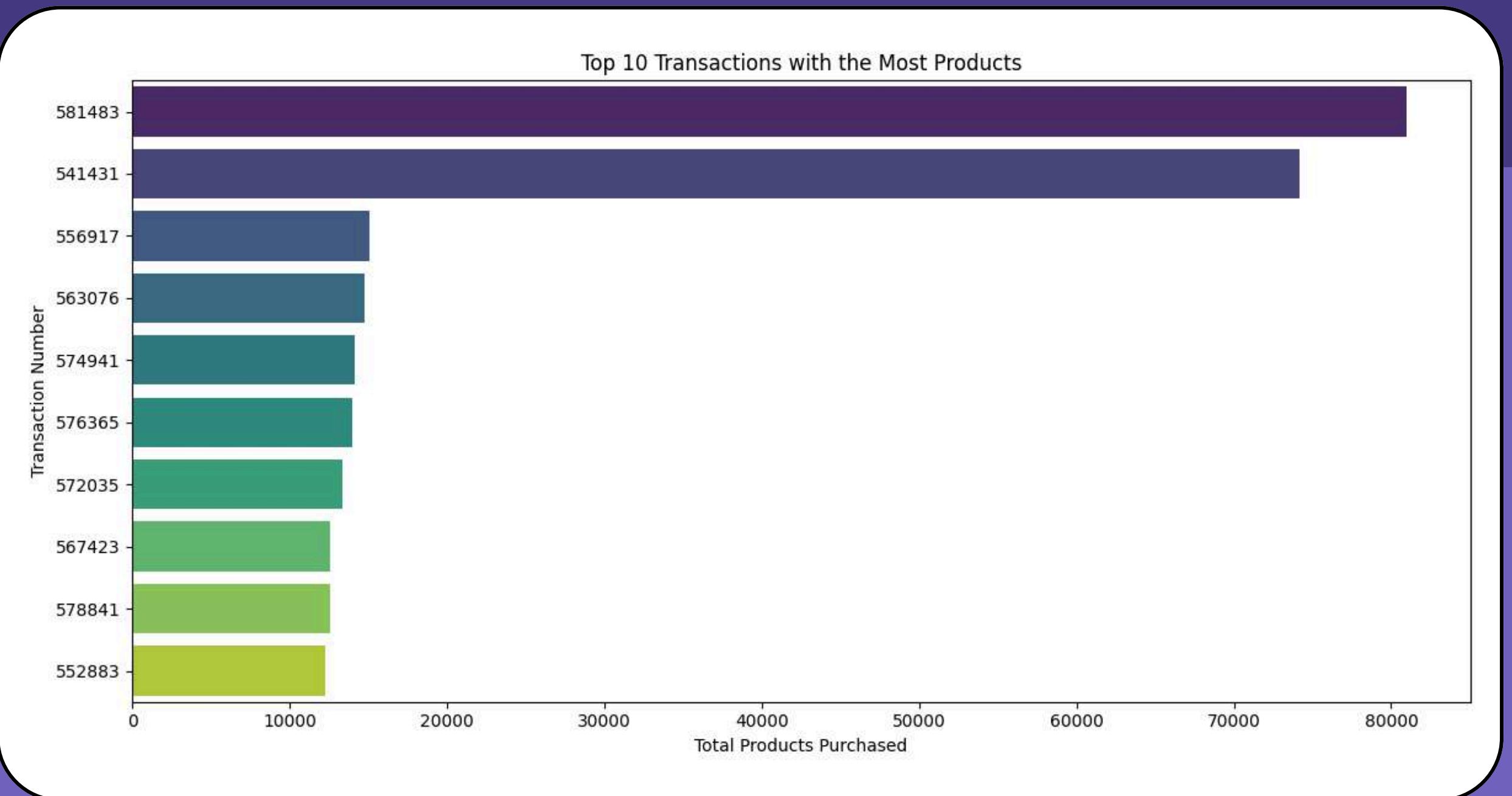
1. Produk dengan quantity penjualan tertinggi:
 - Paper Craft Little Birdie: 80.995 unit
 - Medium Ceramic Top Storage Jar: 78.033 unit
 - Popcorn Holder: 56.902 unit
2. Sebagian besar produk adalah barang kecil, lucu, dan dekoratif, sehingga cocok untuk gift atau impulse buying.



Analisis

Crucial Questions

Banyak produk yang dibeli pelanggan dalam setiap transaksi

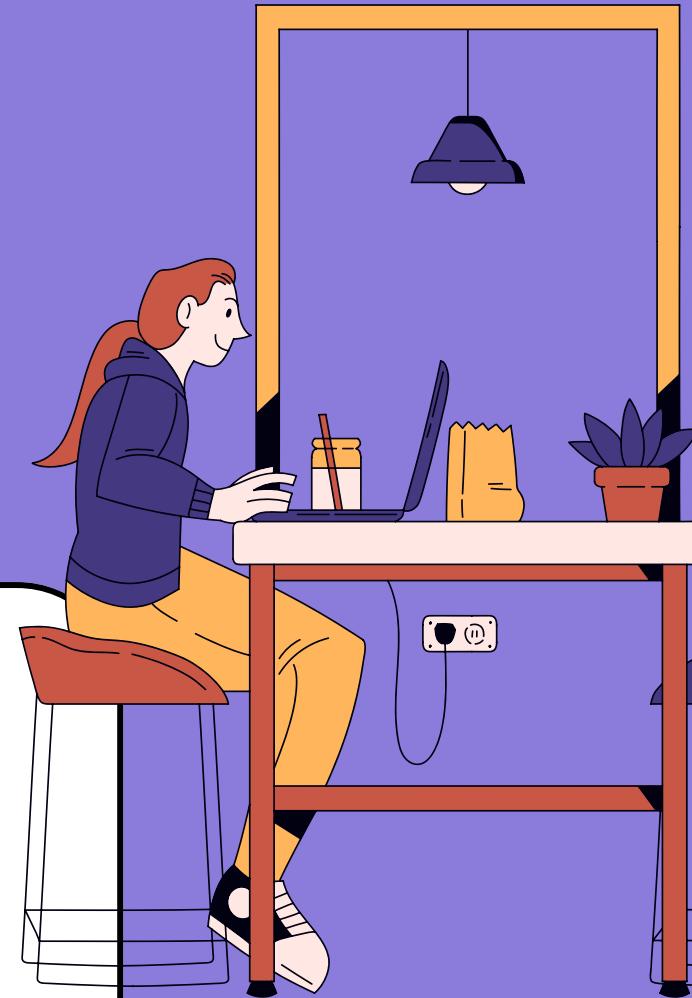


- Dalam satu transaksi, jumlah produk yang dibeli oleh pelanggan bisa sangat besar.
- Contohnya, transaksi dengan ID 581483 mencatatkan 80.995 produk dalam satu kali pembelian, yang merupakan jumlah tertinggi.
- Transaksi-transaksi lain dalam 10 besar juga menunjukkan pembelian ribuan produk dalam satu transaksi.

Analisis

Segmentasi Customers Berdasarkan RFM Score

1. Masing-masing pelanggan diberi skor 1–4 untuk Recency, Frequency, dan Monetary berdasarkan kuartil:
 - Recency: Semakin baru, semakin tinggi skornya.
 - Frequency & Monetary: Semakin banyak, semakin tinggi skornya.
2. Segmentasi Berdasarkan RFM Score dilakukan berdasarkan total dari skor R+F+M:
 - 9–12: 'Best Customers'
 - 7–8: 'Loyal Customers'
 - 5–6: 'Potential Loyalist / At Risk'
 - <5: 'Needs Attention / Low Value'



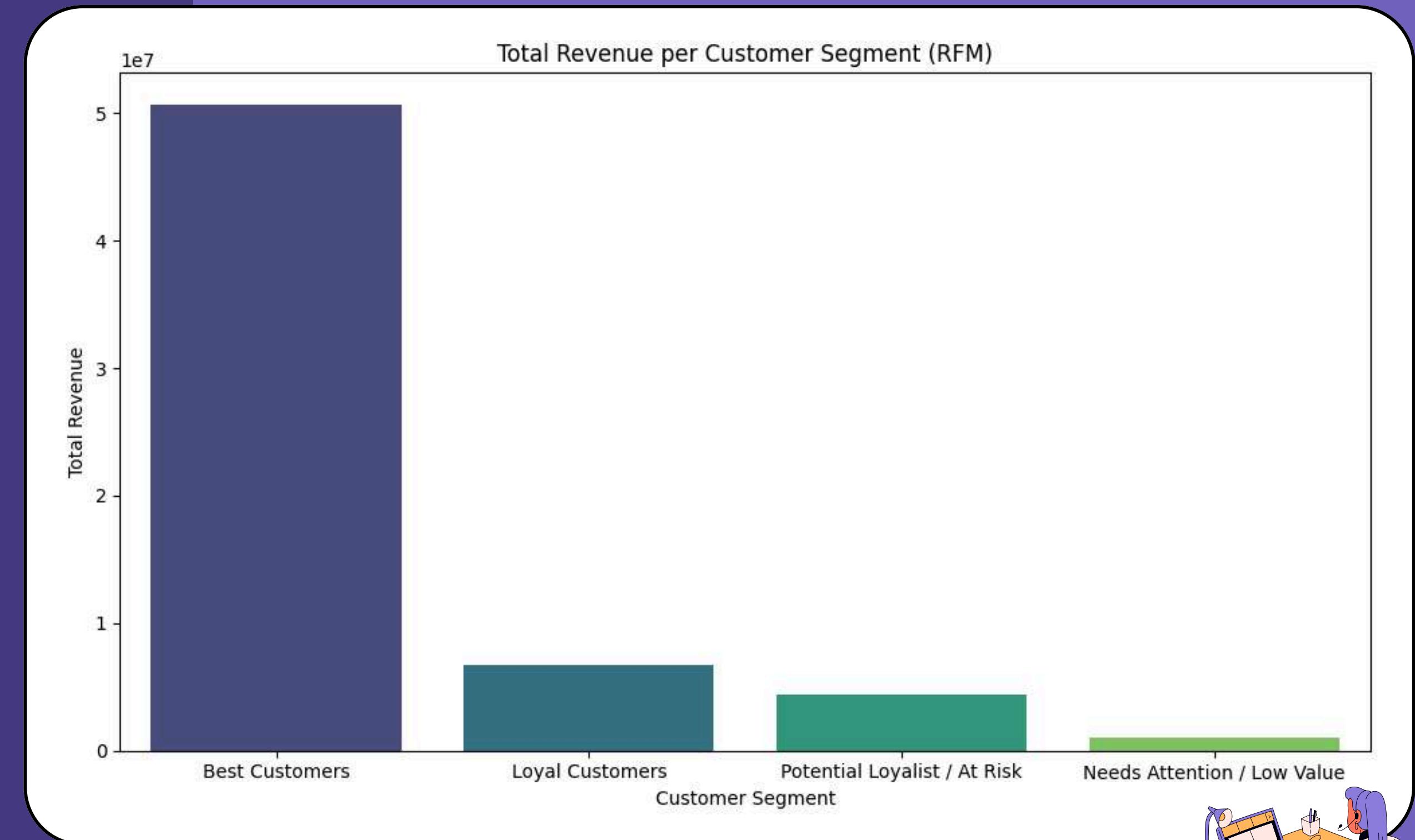


Analisis

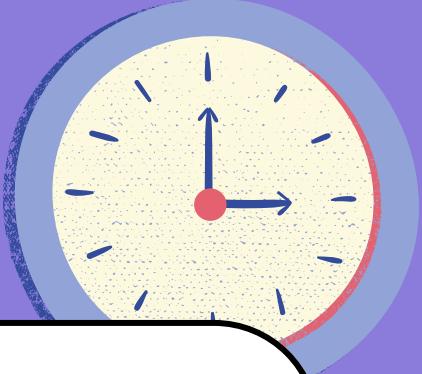
Crucial Questions

Segmen Customers yang Paling Menguntungkan

- Segmen "Best Customers" merupakan kontributor terbesar terhadap total revenue dengan nilai sebesar 50.62 juta.
- Disusul oleh "Loyal Customers" dengan kontribusi revenue sebesar 6.71 juta.
- Segmen lain seperti "Potential Loyalist / At Risk" dan "Needs Attention / Low Value" memberikan kontribusi jauh lebih kecil terhadap total pendapatan.



Kesimpulan



1. Produk dengan Kontribusi Tertinggi terhadap Penjualan

- Sebanyak **575 produk** masuk kategori Super Populer pada November 2019.
- Terdapat **302 produk** yang tetap populer sejak Juli 2019, menunjukkan konsistensi minat pelanggan.
- Produk musiman **bertema Natal** (seperti Christmas Decoration, Festive Bunting, Toilet Roll Natal) memberikan kontribusi besar pada penjualan akhir tahun.

2. Evaluasi Pola Retensi Pelanggan

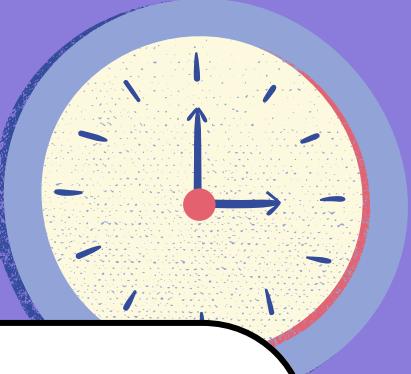
- Jumlah pelanggan baru menurun dari **827 (Januari)** ke **242 (Juli)**.
- Meningkat kembali pada Agustus–November hingga mencapai **334 pelanggan baru** di November.
- Januari 2019 menunjukkan **retensi terbaik**:
 - a. 33,5% di bulan ke-1
 - b. 41,5% di bulan ke-2.

3. Tren Penjualan Harian dan Bulanan

- Penjualan harian rata-rata: **18.278 unit**, dengan maksimum **93.925 unit** (9 Desember 2019).
- Lonjakan besar terjadi pada **18 Januari** dan **9 Desember** 2019.
- Penjualan mingguan tertinggi terjadi pada **Minggu dan Jumat** (>20.000 unit/hari), dan terendah pada Rabu (<10.000 unit/hari).



Kesimpulan



4. Elastisitas Harga terhadap Pembelian

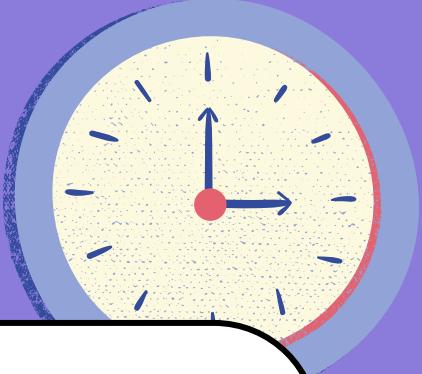
- Terdapat **700 produk** dengan minimal 10 variasi harga.
- Produk dengan **korelasi negatif tertinggi** (elastisitas tinggi):
 - a. Set Of 2 Tea Towels Ping Microwave: -0.476
 - b. Squarecushion Cover Pink Union Jack: -0.462
 - c. Pizza Plate In Box: -0.450
- Artinya, kenaikan harga secara signifikan **menurunkan penjualan** produk-produk tersebut.

5. Analisis Basket Size

- Rata-rata basket size per hari: **3.212 unit**.
- Rentang harian sangat fluktuatif: **1.190 – 25.137** unit.
- Negara dengan basket size terbesar: **Belanda, Australia, Singapura**.



Rekomendasi Strategi



1. Maksimalkan Potensi “Best Customers” dengan Program Loyalitas Eksklusif

- **Alasan**

- a. Mereka menyumbang Rp50,62 juta dari total revenue (kontributor terbesar).
- b. Memiliki potensi repeat order tinggi dan lebih mudah dipertahankan.

- **Aksi**

- a. Buat tiered loyalty program (misal: Gold, Silver, Bronze).
- b. Berikan benefit eksklusif seperti diskon spesial, akses early bird, atau hadiah ulang tahun.
- c. Pertahankan engagement melalui email dan promo eksklusif.

2. Optimalkan Hari Penjualan Tinggi (Minggu & Jumat) untuk Flash Sale dan Promosi

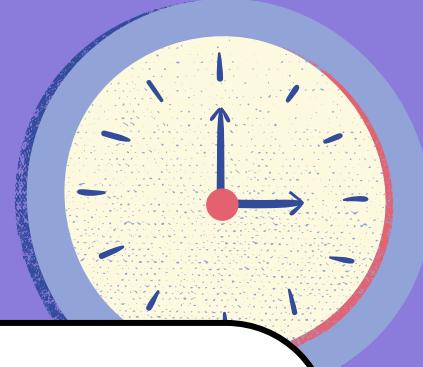
- **Alasan**

- a. Hari Minggu dan Jumat menunjukkan rata-rata penjualan tertinggi (>20.000 unit).
- b. Peluang besar untuk meningkatkan volume penjualan harian dengan promosi terarah.

- **Aksi**

- a. Jalankan Flash Sale hanya di Jumat dan Minggu dengan durasi terbatas.
- b. Gunakan email & media sosial untuk mendorong traffic di dua hari tersebut.

Rekomendasi Strategi



3. Manfaatkan Produk Musiman (Terutama Natal) untuk Kampanye Penjualan yang Terjadwal

- **Alasan**

- a. Banyak produk Super Populer di bulan November adalah produk bertema Natal/liburan.
- b. Terjadi lonjakan penjualan signifikan di Q4 (Sep-Nov).

- **Aksi**

- a. Rancang kampanye tematik jauh sebelum musim puncak (misal: "Christmas in July" promo).
- b. Sediakan bundling dan diskon spesial untuk produk musiman yang terbukti laku.





Growia
Indonesia

Thank You!



nicholas.manurung2002@gmail.com



082168235025

