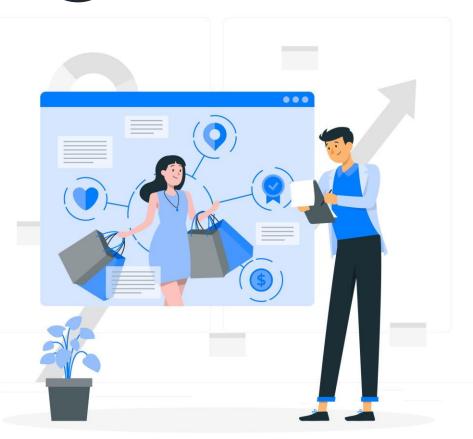
Customer Segmentation by RFM Analysis

Femy Rahma F.

Outline





Project Background

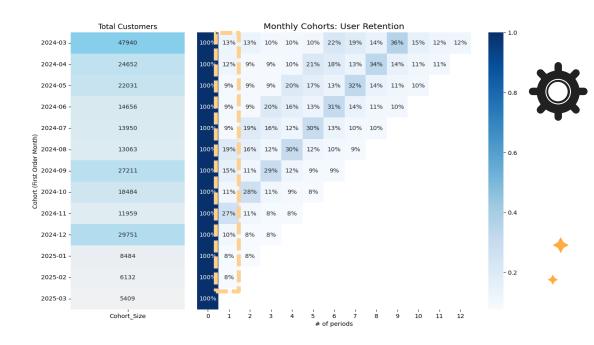
Project ini bertujuan melakukan analisis RFM pada data penjualan store fashion retail untuk mengelompokkan pelanggan berdasarkan unsur recency, frequency, dan monetary, serta mengidentifikasi pelanggan bernilai tinggi. Wawasan dan rekomendasi analisis ini diharapkan dapat meningkatkan strategi pemasaran, meningkatkan kepuasan pelanggan, dan mendongkrak penjualan.

Specific Objective



Business Problem

Analisis cohort mengungkap bahwa 90% pelanggan hanya bertransaksi sekali. Retensi pelanggan baru sangat rendah (~10% di bulan ke-1), mengindikasikan churn tinggi. Fluktuasi cohort size dan lonjakan retensi pada beberapa periode menandakan efektivitas promosi musiman.

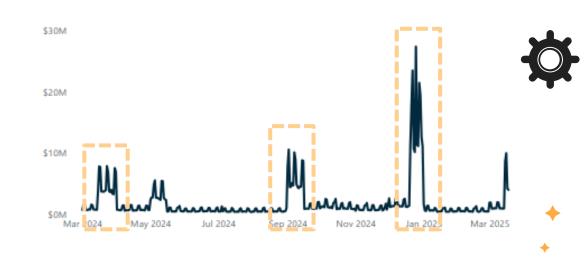




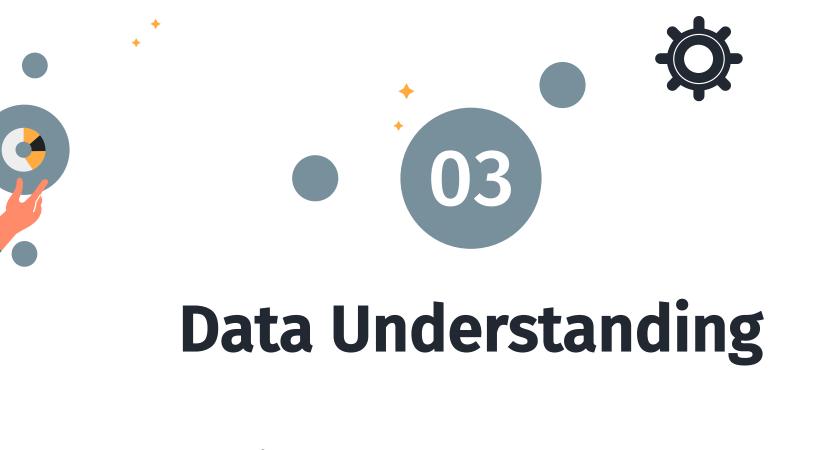
Business Problem

Sementara **total penjualan menunjukkan lonjakan musiman yang diikuti penurunan drastis.**

Untuk memperbaiki churn tinggi dan menstabilkan omzet, kita akan memerlukan analisis lebih lanjut mengingat pentingnya karakteristik pelanggan.









Dataset Information



Kumpulan data dari <u>kaggle</u> yang mensimulasikan data transaksional selama dua tahun untuk fashion retail multinasional, berisi 4 juta+ catatan penjualan dari 35 toko yang terdapat di 7 negara yaitu Amerika Serikat, Tiongkok, Jerman, Inggris Raya, Prancis, Spanyol, dan Portugal. Pada analisis ini data difokuskan untuk transaksi penjualan selama satu tahun (2024 – 2025) di Amerika Serikat.



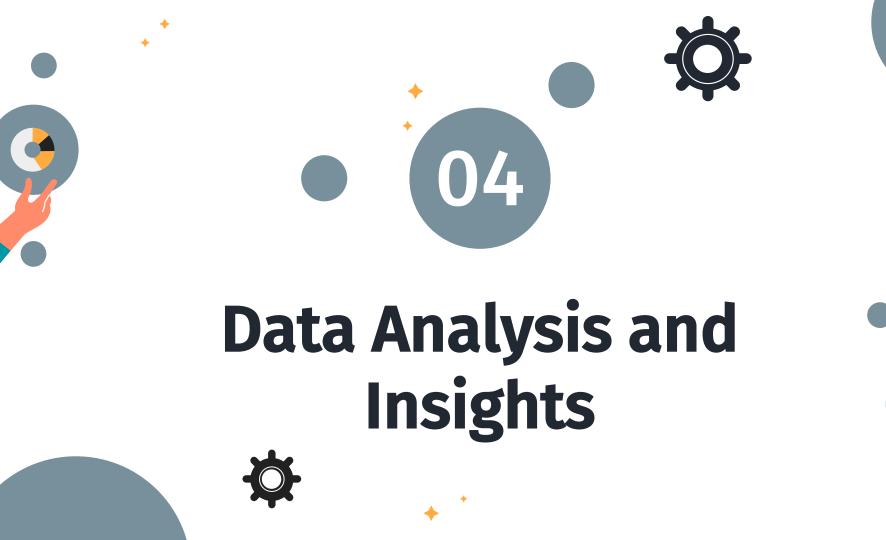
837.884 rows



20 columns

Columns Descriptions

Invoice ID	Kode unik untuk setiap transaksi	Line Total	Total biaya untuk baris item setelah menerapkan diskon
Line	Nomor urut yang mewakili posisi produk dalam faktur	Store ID	Kode unik untuk toko tempat transaksi terjadi
Customer ID	Kode unik untuk pelanggan yang melakukan pembelian	Employee ID	Kode unik untuk karyawan yang memproses transaksi
Product ID	Kode unik untuk produk yang dibeli	Currency	Kode mata uang ISO tiga huruf yang mewakili mata uang yang digunakan untuk transaksi
Size	Varian ukuran produk	Currency Symbol	Simbol yang terkait dengan mata uang transaksi
Color	Variasi warna produk	SKU	Kode inventaris unik yang menggabungkan ID
Unit Price	Harga satu unit produk sebelum		Produk, Ukuran, dan Warna
	diskon diterapkan	Transaction Type	Menentukan apakah transaksi tersebut adalah
Quantity	Jumlah unit produk yang dibeli		Penjualan atau Pengembalian
	dalam baris item faktur	Payment Method	Metode yang digunakan untuk menyelesaikan
Date	Tanggal dan waktu transaksi		transaksi
Discount	Diskon diterapkan pada baris	Invoice Total	Mengacu pada nilai total transaksi (ID Faktur)
	item	Sub Category	Klasifikasi produk lebih spesifik dalam kategori



Data Cleaning & Preprocessing



Check Duplicate

Terdapat 138 dapat duplicate dan telah dihilangkan (drop duplicate)



2



Check Missing Value

Terdapat 53.347 missing value pada kolom 'Size' dan '566.227' pada kolom 'Color'. Kolom kosong diisikan dengan 'UNKNOWN'.



Convert Data Type

Mengubah tipe kolom 'Date' dari object menjadi datetime.



Check Outliers



Ditemukan outliers (nilai ekstrem) pada beberapa kolom, akan tetapi untuk analisis RFM keberadaan outliers ini masih penting.

New Table → "RFM_Table"

recency

Jumlah hari sejak transaksi terakhir pelanggan hingga tanggal reference (31/03/2025)



Jumlah transaksi (invoice) yang dilakukan pelanggan dalam periode analisis

monetary

Total nilai pembelanjaan (sum of spend) pelanggan dalam periode analisis



Customer ID

Kode unik untuk pelanggan yang melakukan pembelia

Recency_score

Skor 1– 3 yang mengukur "segar"nya pelanggan

frequency_score

Skor 1– 3 yang mencerminkan seberapa sering pelanggan bertransaksi

monetary_score

Skor 1– 3 berdasarkan total spend pelanggan



Total Sales

\$748M

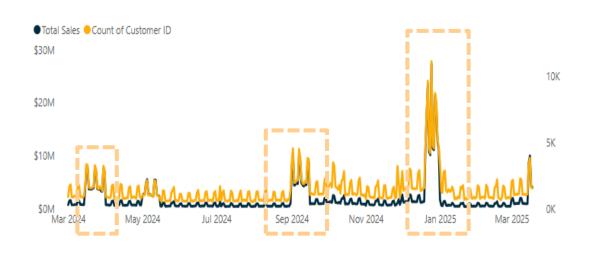
Total Quantity Sold

873K

Total Transactions

794K

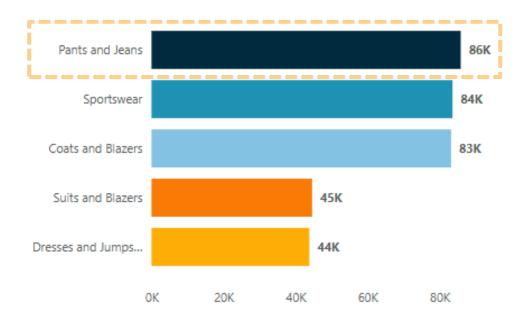
Bagaimana pola perkembangan penjualan dan jumlah pelanggan sepanjang waktu?



Korelasi erat antara jumlah pelanggan & penjualan, setiap kali jumlah pelanggan bertambah, total sales juga meningkat. Menunjukkan volume pelanggan menjadi driver utama revenue.

Spike musiman terlihat pada beberapa bulan Apr 2024, kemungkinan Spring/Summer launch atau promo Lebaran. Sep 2024 yang cocok dengan Back-to-School/Autumn campaign. Jan 2025, konsisten dengan Holiday/Year-End Sale.

Product apa yang menjadi unggulan?

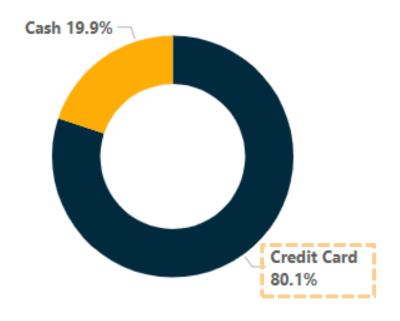


Pants & Jeans memberikan penjualan tertinggi dengan ~86 K unit terjual, menunjukkan ini merupakan core product fashion retail perusahaan.

Sportswear (84 K) dan Coats & Blazers (83 K) hampir setara di urutan kedua dan ketiga, menunjukkan adanya minat tinggi pada activewear dan outerwear, terutama di musim tertentu.

Terdapat **penurunan tajam** pada **product Suits & Blazers** (45 K) **dan Dresses & Jumpsuits** (44 K), menunjukkan segmen ini bersifat lebih spesifik (berfungsi kebutuhan tertentu).

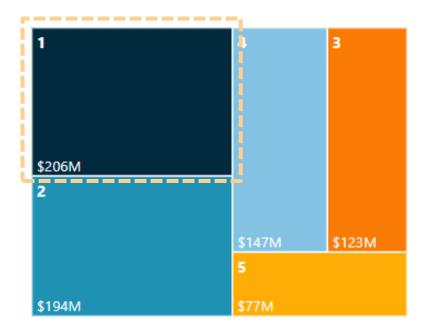
Metode pembayaran seperti apa yang disukai pelanggan?



Credit Card (80 %) mendominasi pembayaran, menunjukkan pelanggan merasa nyaman. Hal ini mungkin karena kemudahan, reward points, atau promosi "cashback".

Pembayaran Cash yang masih diminati sebagian pelanggan (20 %), sebagian segmen pelanggan masih memprioritaskan anonimitas atau tidak memiliki kartu kredit.

Apakah setiap store memberikan performance yang bagus?



Dua Outlet unggulan yaitu Store 1 dan Store 2 bersama-sama menyumbang lebih dari setengah total penjualan (400 M dari ~747 M). Jarak penurunan yang cukup signifikan pada penjualan Store 4 (147 M) dan Store 3 (123 M) dan penjualan terendah berada di Store 5 (77 M) hanya menyumbang ~10 %, terlihat under-performance dibanding outlet lain.

"Bagaimana dengan perilaku pelanggan? Apakah berpengaruh terhadap performa penjualan?"



Apa itu RFM (Recency, Frequency, Monetary)?



Recency

Seberapa baru pelanggan telah melakukan pembelian atau berinteraksi dengan perusahaan.



Frequency

Seberapa sering pelanggan melakukan pembelian atau berinteraksi dalam periode waktu tertentu.



Monetary

Berapa besar total nilai pembelian yang dilakukan oleh pelanggan dalam periode waktu tertentu.



Overall Customer Performance Summary

Total Customer

243.67K

Average Recency (Days)

135

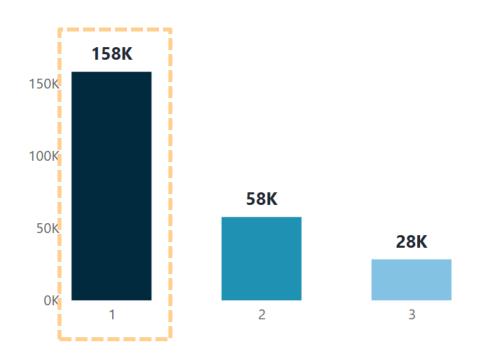
Average Frequency (Times)

3

Average Monetary

\$3.07K

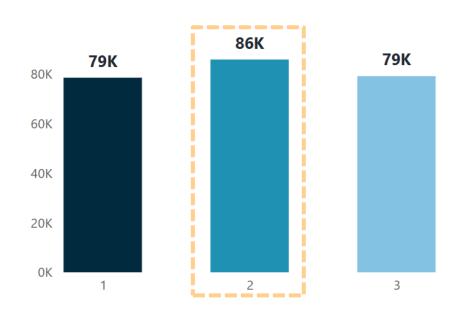
Bagaimana kecenderungan pelanggan kembali berbelanja?



Mayoritas melakukan transaksi belanja terakhir >90 hari lalu. Sekitar 60 % (158 K) pelanggan pelanggan tergolong Recency Score 1, menunjukkan basis pelanggan yang besar tapi tidak aktif (mengindikasikan churn tinggi).

Score	Recency	
1	> 90 hari	
2	31 – 90 hari	
3	≤ 30 hari	

Seberapa sering pelanggan berbelanja?

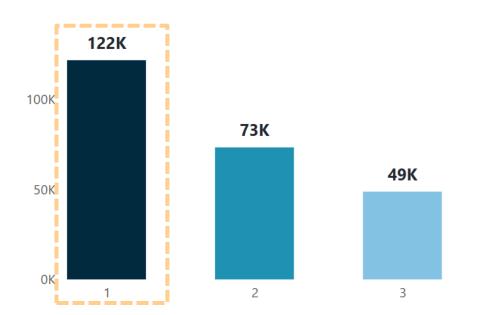


Sebagian besar pelanggan melakukan 2 – 3 kali transaksi dalam setahun.

Pelanggan kategori "occasionally engaged" (Skor 2) memuncaki distribusi dengan jumlah 86 K pelanggan. Kelompok ini penting karena sangat berpeluang menjadi "loyal customer".

Score	Frequency		
1	1		
2	2 – 3		
3	≥ 4		

Seberapa besar uang yang pelanggan belanjakan?



Mayoritas pelanggan membelanjakan uangnya dengan total < \$1.739. Low-Value Customers (Score 1) mengisi 122 K dari pelanggan (sekitar 54 %) yang menyumbang proporsi revenue paling kecil.

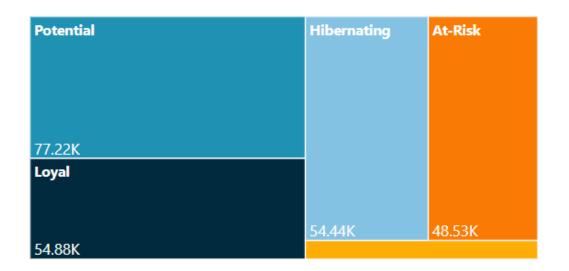
Score	Monetary	
1	< \$1.739	
2	\$1.740 – \$4.879	
3	≥ \$4.880	



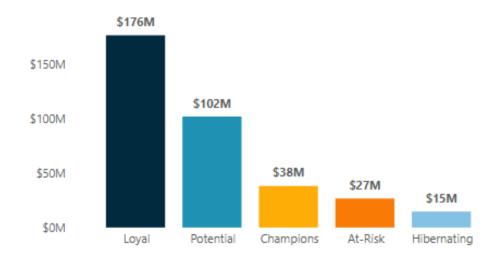
RFM Segmentation Descriptions

Segmen	R	F	M	Deskripsi
Champions	3	3	3	Pelanggan terbaik, baru bertransaksi, paling sering, dan nilai belanja tertinggi
Loyal	2-3	3	2-3	Pelanggan rutin & bernilai tinggi, recency cukup baik
Potential	2-3	2-3	2-3	Pelanggan kelas menengah di semua metrik, berpotensi naik ke Loyal atau Champions
At-Risk	1	2-3	2-3	Dulunya aktif & spend tinggi, kini sudah lama tidak bertransaksi (>90 hari)
Hibernating	1	1	1	Pelanggan paling "dingin", jarang transaksi & spend rendah

Customer Sementation by RFM



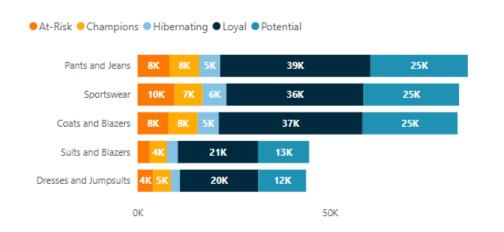
Segmen potential berjumlah paling besar (~77 K pelanggan), menandakan basis pelanggan "medium" di semua metrik.
Segmen Loyal dan Hibernating hampir seimbang (keduanya ~55 K), dan disusul dengan At-Risk (~48 K), sementara Champions hanya minoritas kecil.



Loyal Customer Segmen Paling Bernilai (Sales Tertinggi)

- Loyal segment mampu menghasilkan penjualan tertinggi (\$176 M).
- Potential memiliki jumlah pelanggan terbesar, namun kontribusi penjualan masih lebih rendah, menunjukkan potensi besar yang bisa dioptimalkan.
- Hibernating dan At-Risk memiliki basis pelanggan cukup besar namun kontribusi rendah.
- Champions, meskipun jumlahnya kecil, tetap berkontribusi signifikan sebesar \$38M, menandakan nilai pembelian yang tinggi per pelanggan.

Loyal dan Potential Konsisten Mendominasi Volume Pembelian



Segmen Loyal dan Potential mendominasi jumlah (quantity) penjualan di semua kategori produk unggulan, menjadikannya pendorong utama performa bisnis.

Segmen **Hibernating** memiliki **kontribusi kecil** di semua produk, namun tetap **menunjukkan minat** terhadap kategori popular yang dapat dimanfaatkan dalam kampanye aktivasi ulang.

Recommendation

Fokus pada Retensi Segmen Loyal

- Memperkuat retensi melalui program loyalitas, akses eksklusif produk baru, poin ganda, dan bundling produk premium.
- **Impact**: Peningkatan Customer Lifetime Value, Pengurangan biaya akuisisi, Stabilitas pendapatan.

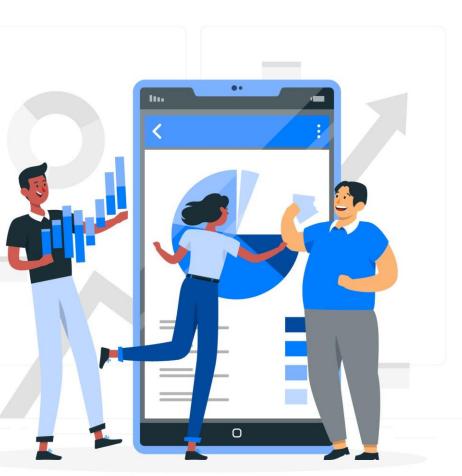
Konversi Segmen Potential Menjadi Loyal

- Melakukan pendekatan personalisasi, edukasi produk, diskon progresif, serta cross-sell produk komplementer untuk mendorong naik kelas menjadi pelanggan Loyal atau Champions.
- **Impact:** Peningkatan konversi, Nilai pelanggan bertumbuh, Membangun loyalitas jangka Panjang.

Reaktivasi Segmen At-Risk dan Hibernating

- Kampanye reaktivasi seperti voucher eksklusif, penawaran spesial kategori favorit, email "We miss you", serta drip campaign ringan untuk menghidupkan kembali minat pelanggan.
- Impact: Mengurangi churn rate, Peningkatan engagement.





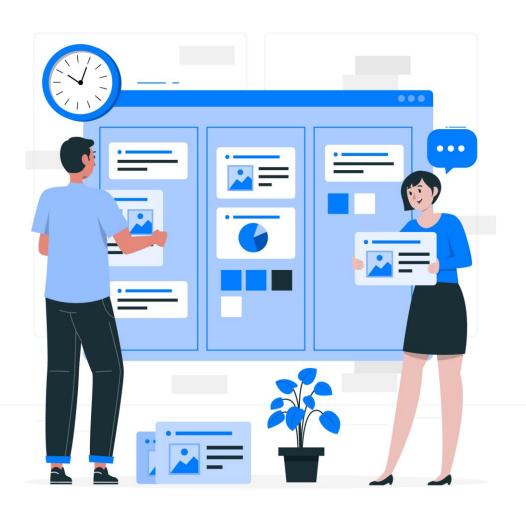
Optimalisasi Peluang dari Segmen Champions

- Menerapkan pendekatan VIP management seperti program ambassador, undangan private sale, dan benefit loyalty tertinggi.
- **Impact:** Meningkatkan loyalitas, Mendorong brand advocacy, Peningkatan pembelian premium.

Optimalisasi Produk Favorit di Segmen Utama

- Menjadikan produk produk favorit sebagai anchor untuk bundling, program reward, atau campaign promosi,
- **Impact:** Mendorong pembelian berulang, Memperluas minat ke kategori lain melalui cross-sell.

Dashboard



Sales Overview Dashboard

Sales Overview Customer Segmentation All Sub Category Store ID

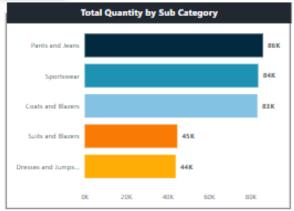
All All All

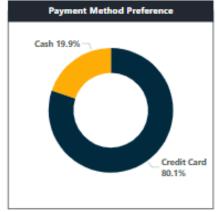
Total Sales

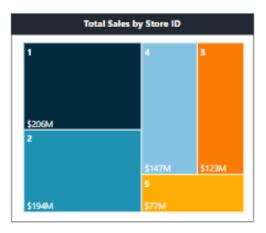
\$748M

Total Quantity Sold

873K







Total Transactions

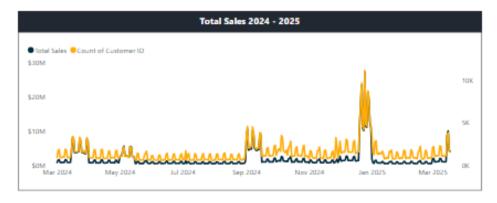
794K

Total Store

5

Top	10	Customer	h۱	/ Sales
IOP	10	Customer	u)	Sales

- Francisco System					
Customer ID	Total Sales	Total Quantity	Total Transaction		
91924	\$51,078	25	16		
97441	\$50,500	11	7		
22666	\$46,367	20	16		
90269	\$44,649	15	13		
16702	\$43,512	23	20		
221017	\$41,567	15	13		
243134	\$40,754	17	16		
116168	\$40,598	32	30		
95887	\$40,466	30	24		
221749	\$40,421	20	14		



Customer Segmentation Dashboard

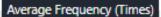


Total Customer

243.67K

Average Recency (Days)

135



Average Monetary

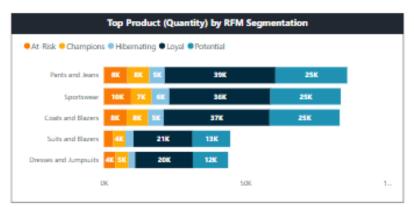
\$3.07K

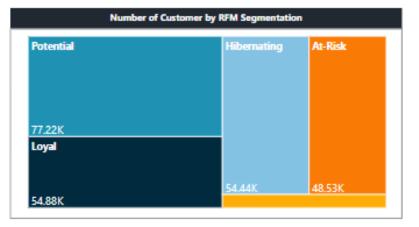




Top 10	Customer	by F	Frequency	
--------	----------	-------------	-----------	--

Customer ID	Recency	Frequency	Monetary
216945	22	35	\$24,893
27272	87	33	\$26,911
36297	44	33	\$31,526
38505	30	33	\$29,432
116168	93	30	\$40,598
218151	25	30	\$17,120
26227	14	29	\$19,020
27376	21	28	\$16,465
29201	23	27	\$28,356
29381	43	27	\$21,579
82844	32	27	\$20,081







Terimakasih.

Drive Reference:

https://drive.google.com/drive/folders/1q3ysa HZnqjmaZSxTRd0m2OVsvC1eTx7L?usp=sharing

₩ femyrahmaf@gmail.com

in linkedin.com/in/femyfitria