

# **FINAL PROJECT DATA ANALYST**

**ARDY ANSYAH**

# DAFTAR ISI

- **TENTANG SAYA**
- **PENDIDIKAN**
- **PENGALAMAN KERJA**
- **KETERAMPILAN DAN ALAT**
- **RINGKASAN PROYEK**
- **PROYEK UTAMA**
- **PEMAHAMAN DATA**
- **DASHBOARD DAN VISUALISASI**
- **ANALISIS DATA DAN PEMAHAMAN**
- **REKOMENDASI**
- **KONTAK SAYA**

# TENTANG SAYA

Saya Ardy Ansyah memiliki passion di bidang data dan percaya bahwa data memiliki kekuatan besar dalam membantu pengambilan keputusan yang tepat, Ketertarikan saya pada dunia data analyst dan data science mendorong saya untuk terus belajar, mengembangkan diri, dan memperdalam keterampilan dalam pengolahan data secara kompeten dan efisien

Saya percaya bahwa setiap tantangan adalah peluang untuk tumbuh dan menantikan arah baru yang akan membawa saya pada pengalaman berikutnya.



# PENDIDIKAN



## UNIVERSITAS INDRAPRASTA PGRI

Sarjana Teknik Informatika (2017-2021)

- Merancang dan membangun sistem inventarisasi barang menggunakan Java (NetBeans) untuk mengelola stok, pencatatan masuk/keluar barang, serta pelaporan data secara lebih efisien dan terstruktur.



## Dibimbing

Bootcamp Data Science dan Data Analyst (2025-Sekarang)

- Mempelajari tentang data preprocessing seperti data cleaning, data manipulation, dan Exploratory Data analyst (EDA)
- Mempelajari tentang management database (SQL, ETL, Web Scrapping) dan pembuatan dashboard seperti data storytelling di Power BI & Tableau

# PENGALAMAN KERJA



## BADAN INTELIJEN NEGARA

Pegawai Non Organik (2022-2024)

- Melaksanakan operasi intelijen dan menyusun laporan analisis strategis yang berhasil memenuhi 90% kebutuhan informasi pimpinan dalam pengambilan keputusan operasional

# KETERAMPILAN DAN ALAT



## ANALISIS DATA

- Menganalisis data untuk menghasilkan insight bisnis serta terbiasa melakukan statistik deskriptif & pengujian hipotesis.



## PEMROSESAN DATA

- Mahir melakukan cleaning & manipulation data menggunakan Python dan Microsoft Excel



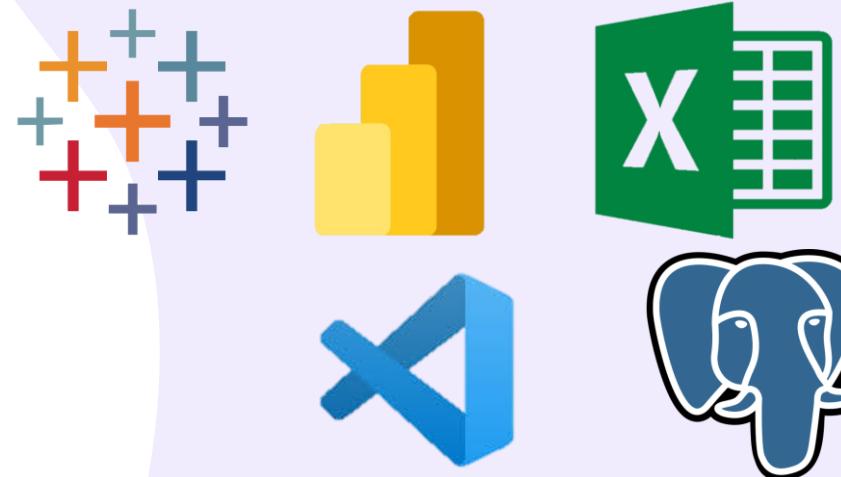
## PEMBUATAN DASHBOARD

- Mahir dalam pembuatan dashboard dengan menggunakan Power BI, Tableau, dan Microsoft Excel

### PROGRAMMING LANGUAGE



### TOOLS



### LIBRARY



# RINGKASAN PROYEK

## Segmentasi Pelanggan pada Automobile Sales Data Menggunakan Analisis RFM

### Deskripsi

Automobile Sales data merupakan dataset yang berisikan informasi transaksi penjualan mobil yang mempunyai 2.747 entri dan 20 kolom, Dataset ini juga mencakup berbagai variabel yang berkaitan dengan detail transaksi, pelanggan, dan produk

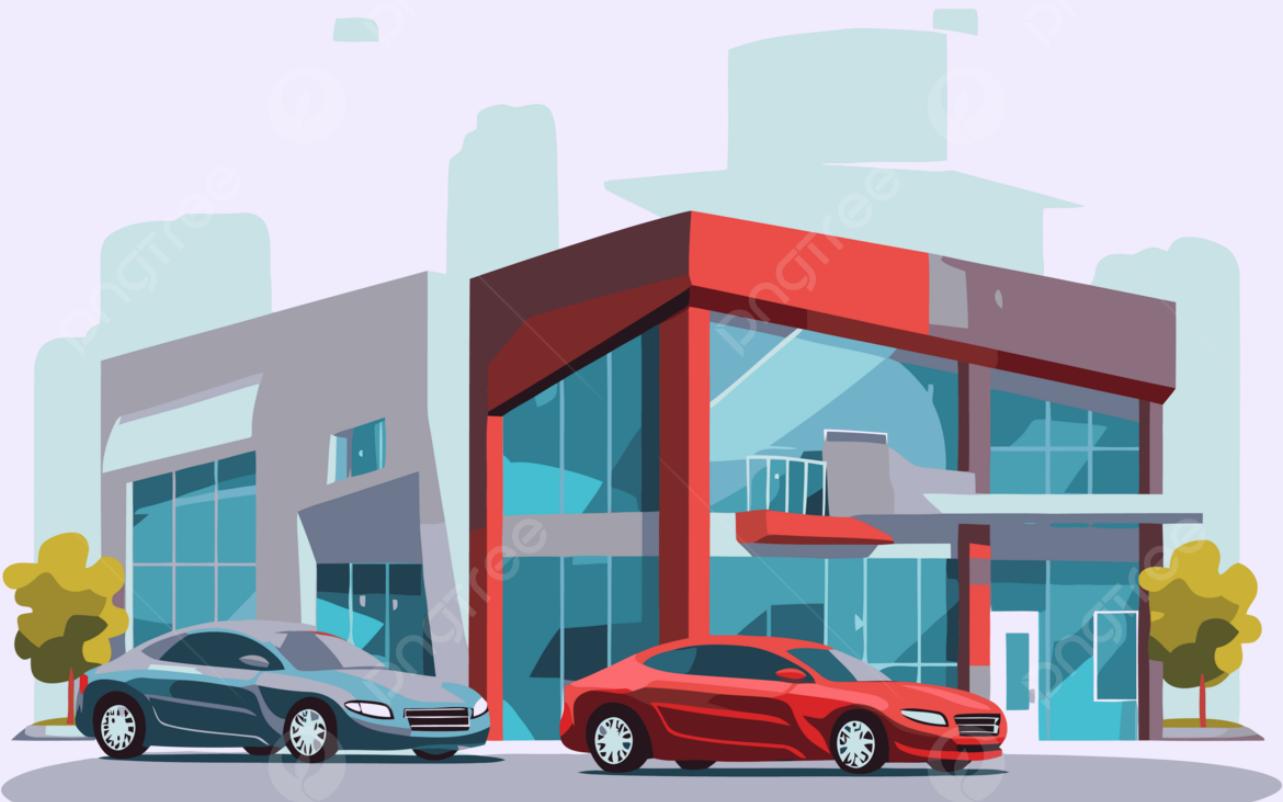
### Use Case Bisnis

Customer Segmentation RFM Analysis

### Hasil Akhir

Mengelompokkan pelanggan dalam beberapa segment berdasarkan atribut seperti perilaku pembelian, demografi pelanggan, dan lokasi geografis setiap pelanggan untuk mendukung strategi pemasaran

kaggle



# PROYEK UTAMA

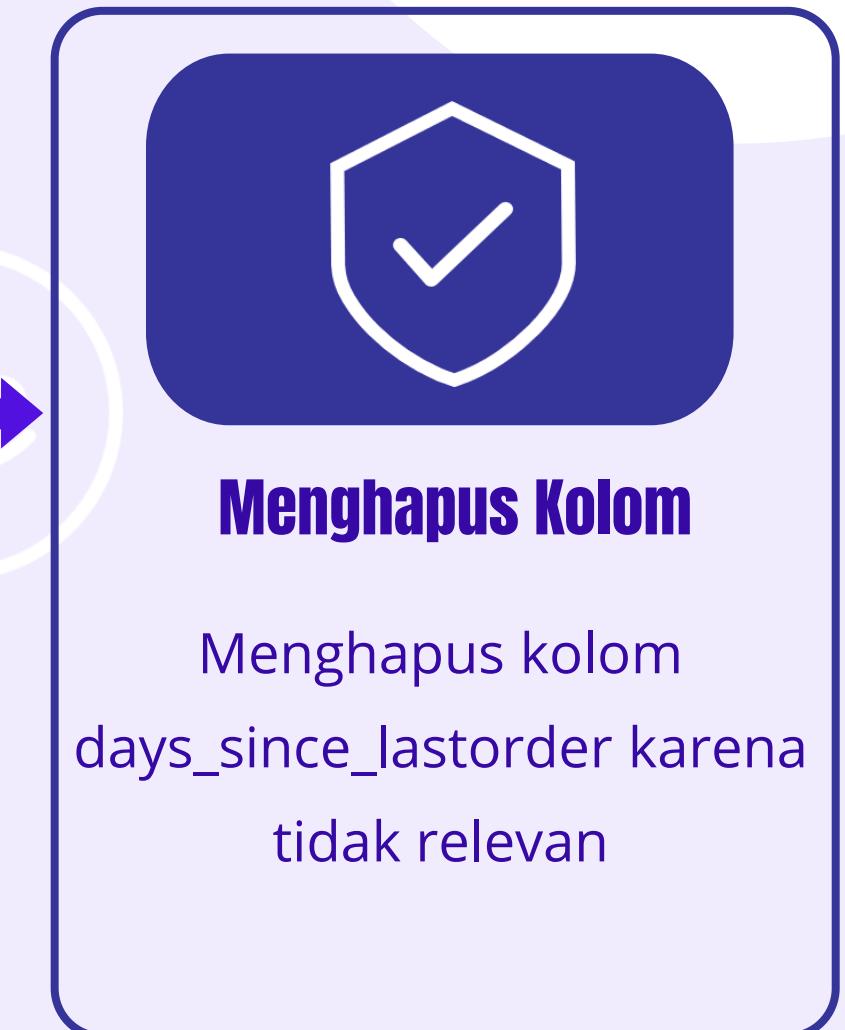
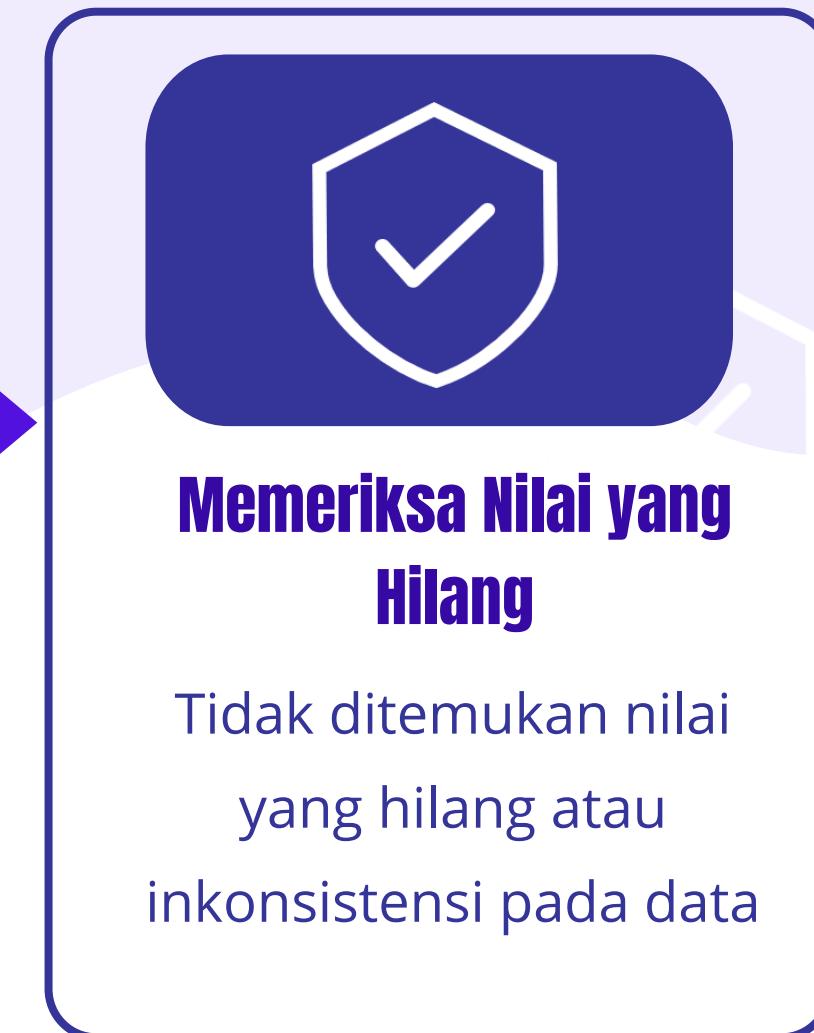
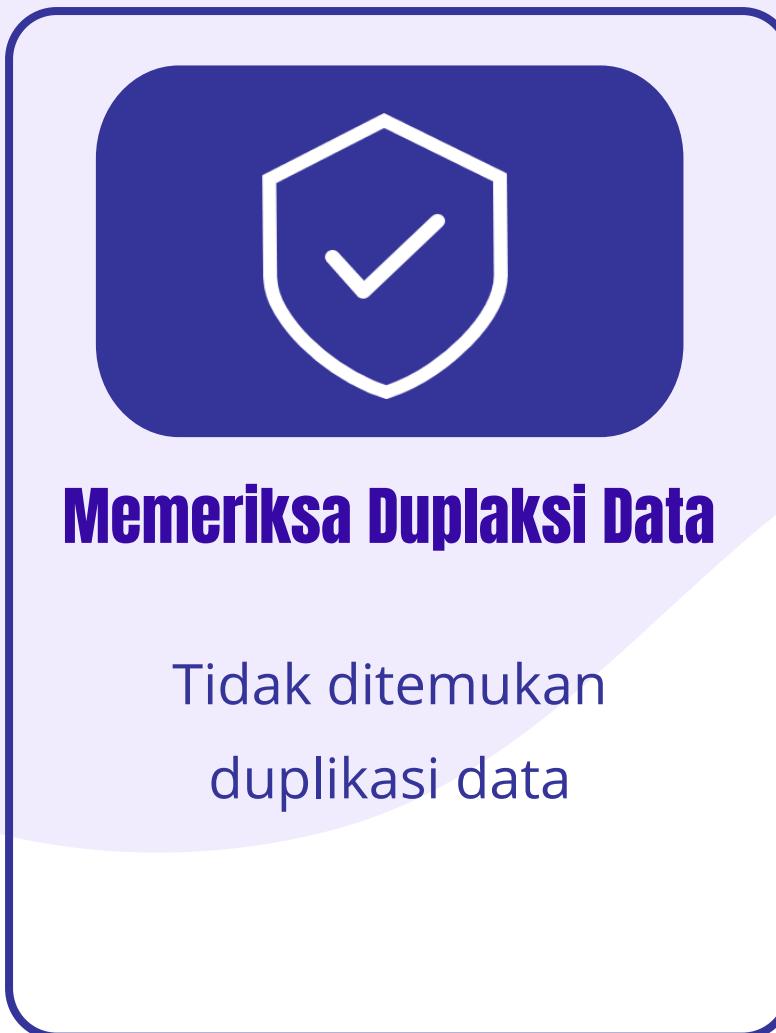
## Latar Belakang

- Dalam setiap industri, memahami perilaku pelanggan sangat penting untuk meningkatkan penjualan, menyusun strategi pemasaran yang tepat dan menjaga loyalitas pelanggan. Banyak perusahaan menghadapi tantangan dalam menentukan segment pelanggan mana yang paling menguntungkan, berpotensi churn, serta mana yang layak untuk dijadikan target utama promosi khusus, dan karena itu lah hasil akhir yang diharapkan dari project ini dapat memberikan insight yang dapat membantu dalam pengambilan keputusan bisnis serta dapat meningkatkan penjualan

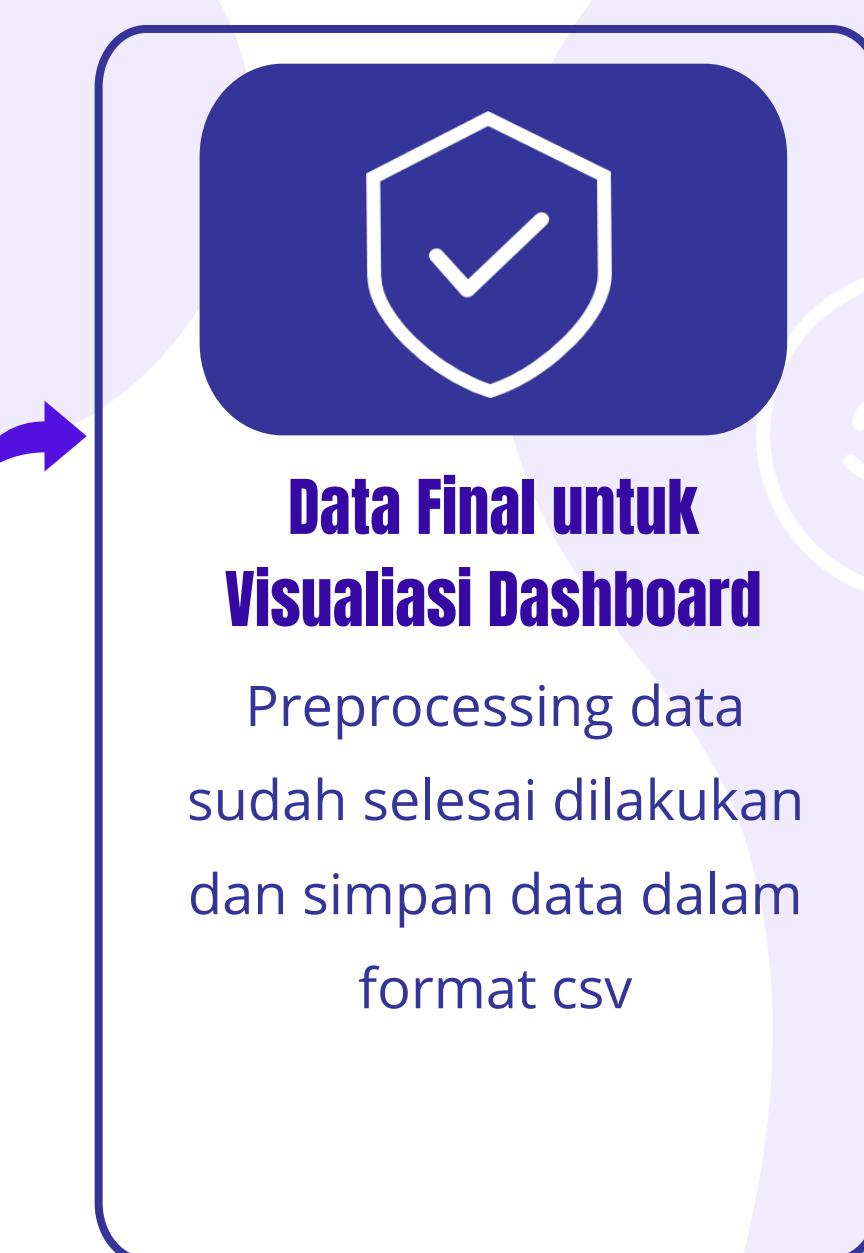
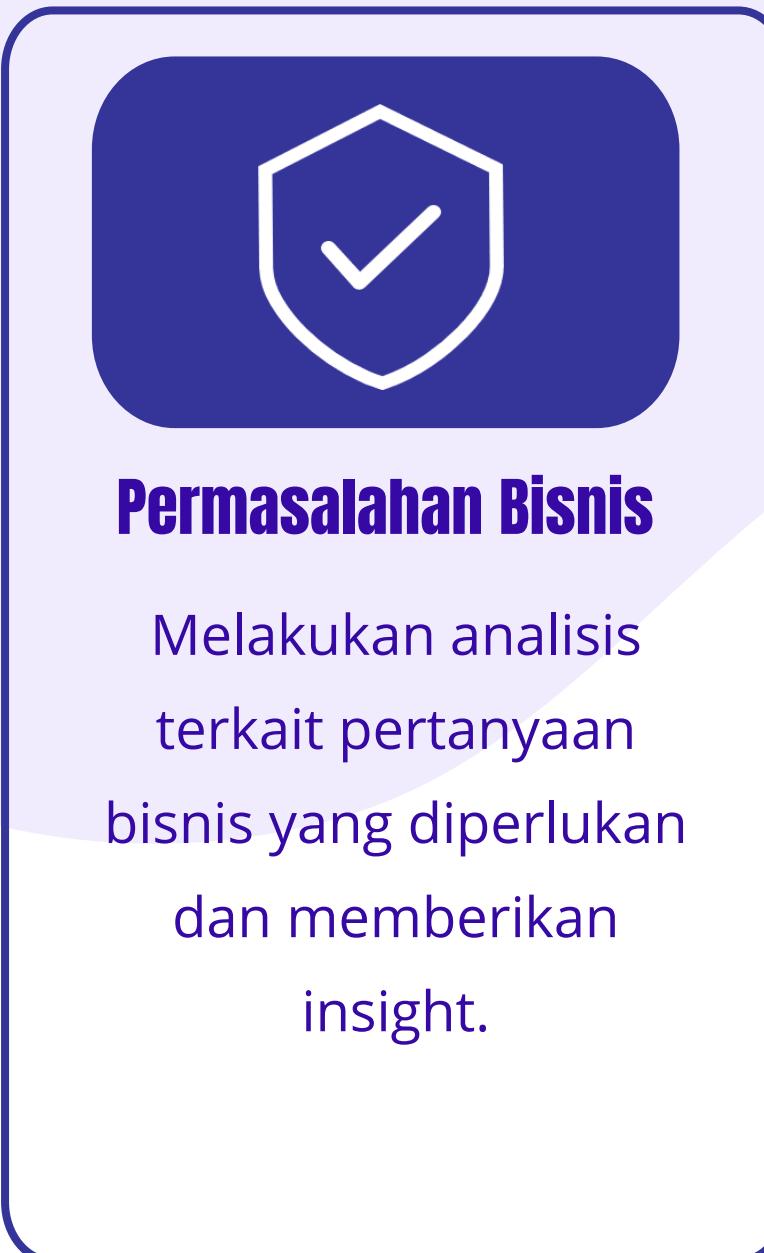
## Permasalahan Bisnis

- Pengaruh Low RFM Score terhadap pendapatan yang diperoleh?
- Apakah ada hubungan antara banyaknya order dengan total pendapatan yang diperoleh?
- Distribusi customer segment jika dilihat dari persentase?
- Negara mana dengan jumlah pelanggan terbanyak?
- Bagaimana trend penjualan dari waktu ke waktu (tahun)?
- Perilaku pembelian pelanggan dalam kaitannya dengan skala pesanan?
- Segmentasi pelanggan dengan penyumbang pendapatan tertinggi?
- Bagaimana karakteristik pelanggan jika dilihat dari skala order?

# PEMAHAMAN DATA



# PEMAHAMAN DATA



# PEMAHAMAN DATA

## Dataset Summary

- Dataset Automobile Sales data.csv diperoleh dari kaggle
- Dataset ini berisi 2747 entri dengan 20 kolom
- Tidak ada duplicate pada dataset
- Tidak dilakukan handling outlier dikarenakan nilai masih dalam batas wajar
- Semua kolom memiliki 2747 nilai yang tidak null, sehingga tidak ditemukan data yang hilang.
- Kolom-kolom kunci mencakup detail transaksi penjualan, informasi pelanggan, detail produk, status pesanan, dan informasi recency.
- Tipe data bervariasi di setiap kolom, termasuk float, int, datetime, dan object.

```
len(df.drop_duplicates()) / len(df)
```

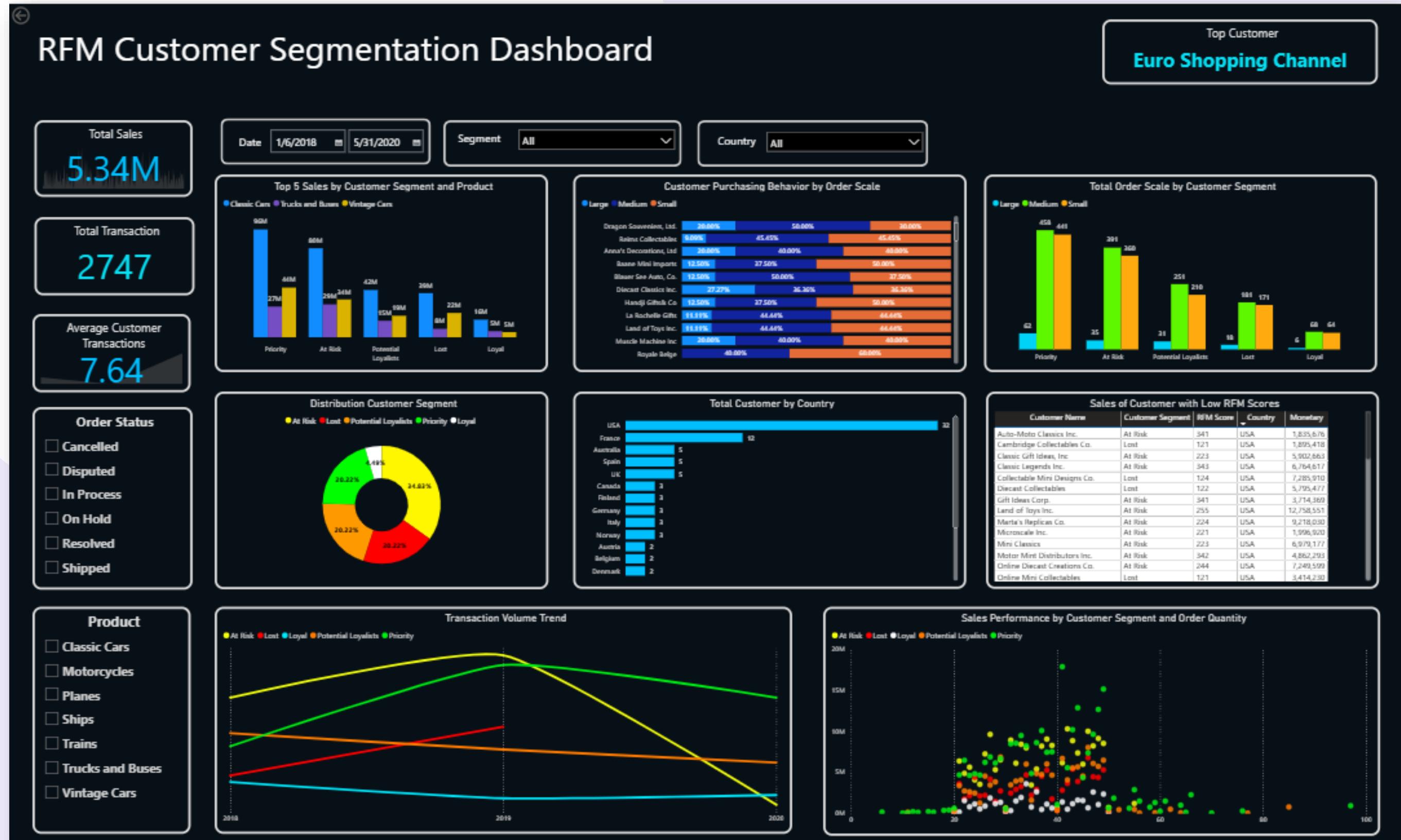
#jika output dari code di cell ini tidak bernilai 1 maka terdapat duplikat

1.0

	count	mean	std	min	25%	50%	75%	max
ORDERNUMBER	2747.0	10259.76	91.88	10100.00	10181.00	10264.00	10334.50	10425.00
QUANTITYORDERED	2747.0	35.10	9.76	6.00	27.00	35.00	43.00	97.00
PRICEEACH	2747.0	101.10	42.04	26.88	68.74	95.55	127.10	252.87
ORDERLINENUMBER	2747.0	6.49	4.23	1.00	3.00	6.00	9.00	18.00
SALES	2747.0	3553.05	1838.95	482.13	2204.35	3184.80	4503.09	14082.80
DAY_SINCE_LASTORDER	2747.0	1757.09	819.28	42.00	1077.00	1761.00	2436.50	3562.00
MSRP	2747.0	100.69	40.11	33.00	68.00	99.00	124.00	214.00

ORDERNUMBER	0
QUANTITYORDERED	0
PRICEEACH	0
ORDERLINENUMBER	0
SALES	0
ORDERDATE	0
DAY_SINCE_LASTORDER	0
STATUS	0
PRODUCTLINE	0
MSRP	0
PRODUCTCODE	0
CUSTOMERNAME	0
PHONE	0
ADDRESSLINE1	0
CITY	0
POSTALCODE	0
COUNTRY	0
CONTACTLASTNAME	0
CONTACTFIRSTNAME	0
DEALSIZE	0

# DASHBOARD DAN VISUALISASI



# ANALISIS DATA DAN PEMAHAMAN

## Objective Analysis

Analisis ini bertujuan untuk memahami perilaku pelanggan berdasarkan segmentasi RFM yang telah dibuat berdasarkan data yang diperoleh dari histori pembelian khususnya dalam aktifitas pelanggan, frekuensi pembelian dan total biaya yang telah dikeluarkan yang mana akan dijadikan sebagai bahan pertimbangan untuk mendukung keputusan strategis dalam pemasaran, peningkatan retensi, pencegahan churn dan peningkatan profitabilitas perusahaan.



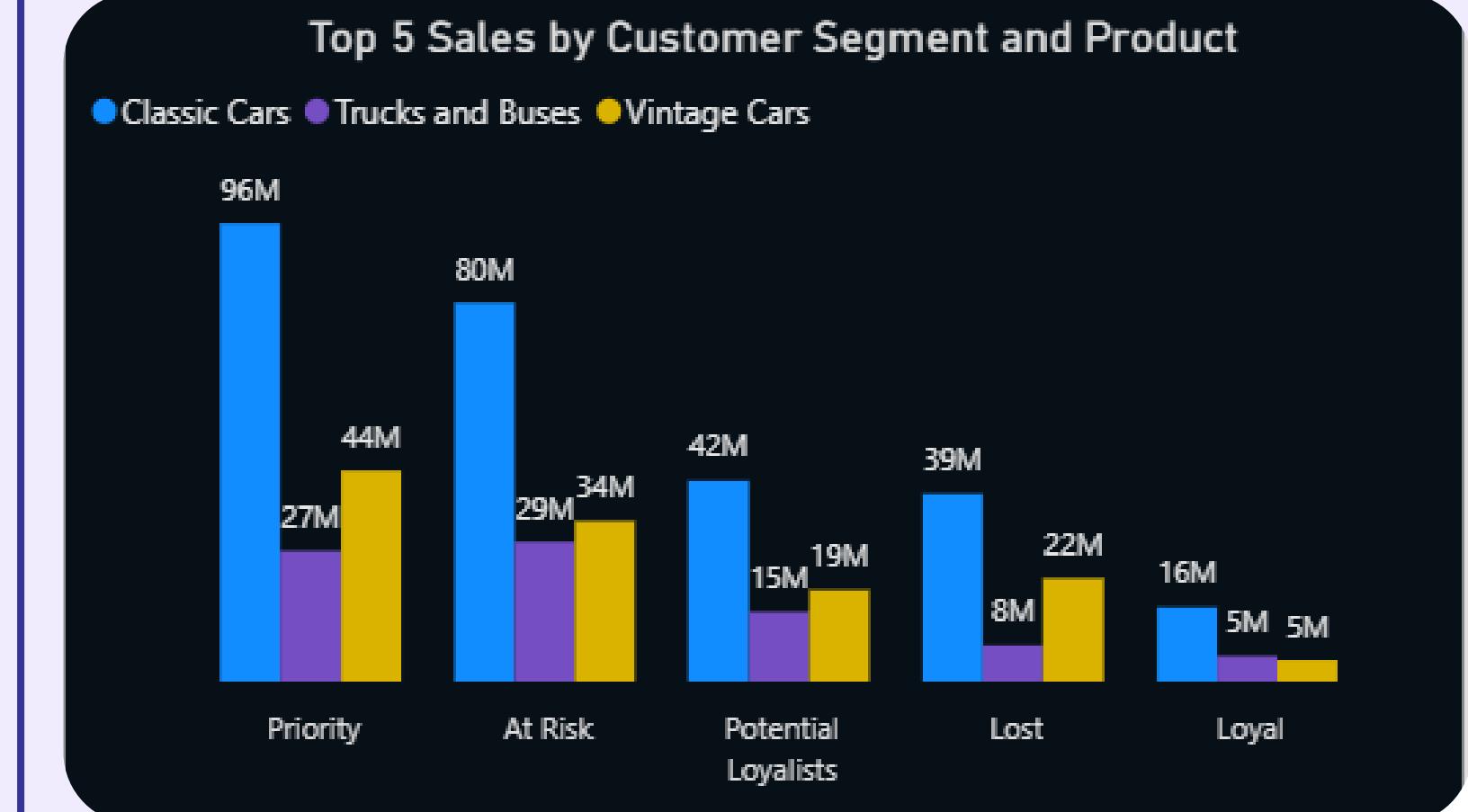
# ANALISIS DATA DAN PEMAHAMAN

## INSIGHT

Segmentasi pelanggan dengan penyumbang pendapatan tertinggi?

- Produk **Classic Cars** mendominasi penjualan di semua segmen pelanggan, terutama **Priority (96M)**, **At Risk (80M)**, **Potential Loyalists (42M)**, Bahkan di segmen **Lost** dan **Loyal**, Classic Cars tetap menjadi kontributor terbesar, sehingga perusahaan harus menjaga kualitas dan ketersediaan produk dengan pendapatan tertinggi di masing-masing segment

## VISUALISASI DASHBOARD



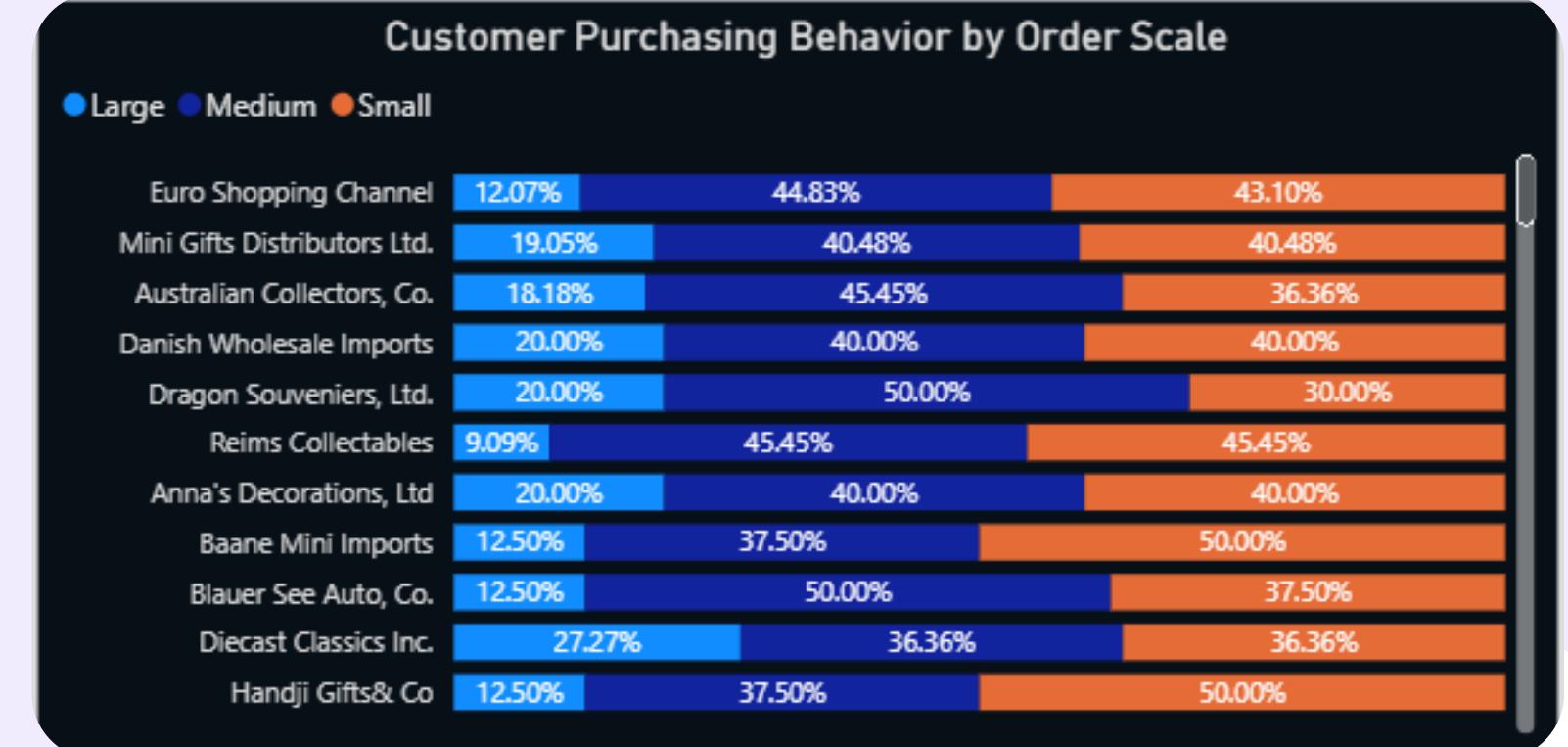
# ANALISIS DATA DAN PEMAHAMAN

## INSIGHT

Perilaku pembelian pelanggan dalam kaitannya dengan skala pesanan?

- Sebagian besar pelanggan melakukan pembelian dalam skala kecil hingga menengah, sementara itu, pembelian skala besar di setiap pelanggan cenderung rendah. maka, perusahaan perlu mengembangkan pendekatan berbeda, seperti memberikan **incentif voucher diskon atau paket bundling** untuk mendorong pelanggan dengan dominasi order kecil hingga menengah agar beralih ke pembelian dalam skala lebih besar, serta menjaga hubungan dengan pemberian layanan premium bagi pelanggan dengan pembelian besar secara konsisten.

## VISUALISASI DASHBOARD



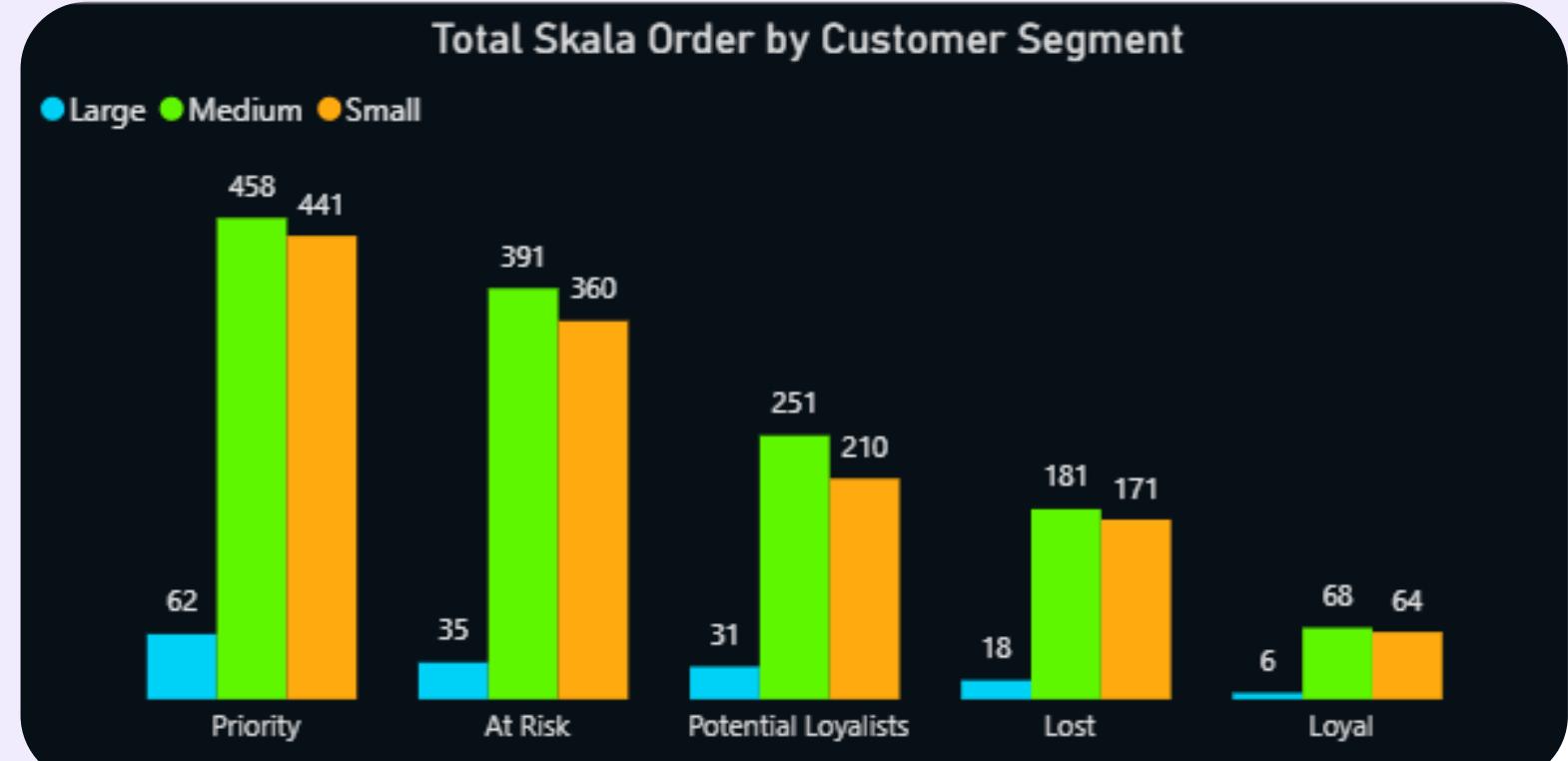
# ANALISIS DATA DAN PEMAHAMAN

## INSIGHT

Bagaimana karakteristik pelanggan jika dilihat dari skala order?

- Order berskala **Medium** dan **Small** mendominasi di semua segmen pelanggan, terutama pada segmen **Priority** (**Medium: 458, Small: 441**) dan **At Risk** (**Medium: 391, Small: 360**). Sementara itu, order Large masih sangat sedikit di seluruh segmen, dengan jumlah tertinggi hanya **62 order** pada **segment Priority**, Perusahaan perlu mendorong peningkatan **order berskala besar (Large)** melalui program seperti **diskon volume, loyalty points untuk pembelian besar, atau bundling produk premium**, terutama untuk segmen **Priority** dan **At Risk** yang memiliki potensi tinggi.

## VISUALISASI DASHBOARD



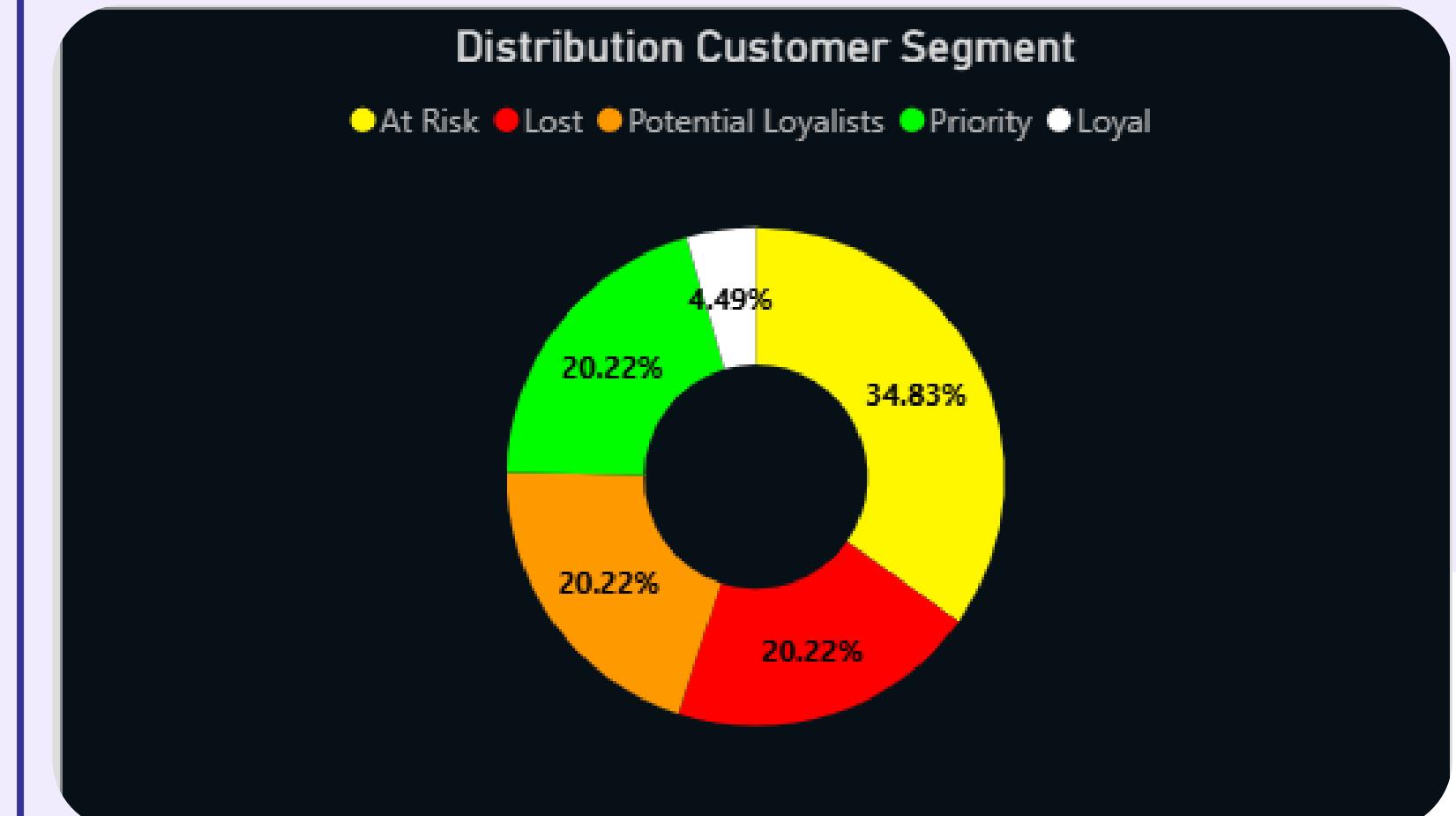
# ANALISIS DATA DAN PEMAHAMAN

## INSIGHT

Distribusi customer segment jika dilihat dari persentasenya?

- Proporsi pelanggan terbesar berada pada **segment At Risk (34.83%)**, yang menunjukkan banyak pelanggan aktif sebelumnya mulai menunjukkan tanda penurunan transaksi. Fokus utama perusahaan sebaiknya diarahkan pada **strategi retensi** seperti **program loyalitas, dan penawaran khusus**. Mengingat tingginya jumlah pelanggan At Risk dan Lost (total lebih dari 50%), perusahaan perlu segera menerapkan tindakan pencegahan churn untuk menjaga basis pelanggan dan mendorong mereka naik ke segment Potential Loyalists dan Loyal.

## VISUALISASI DASHBOARD



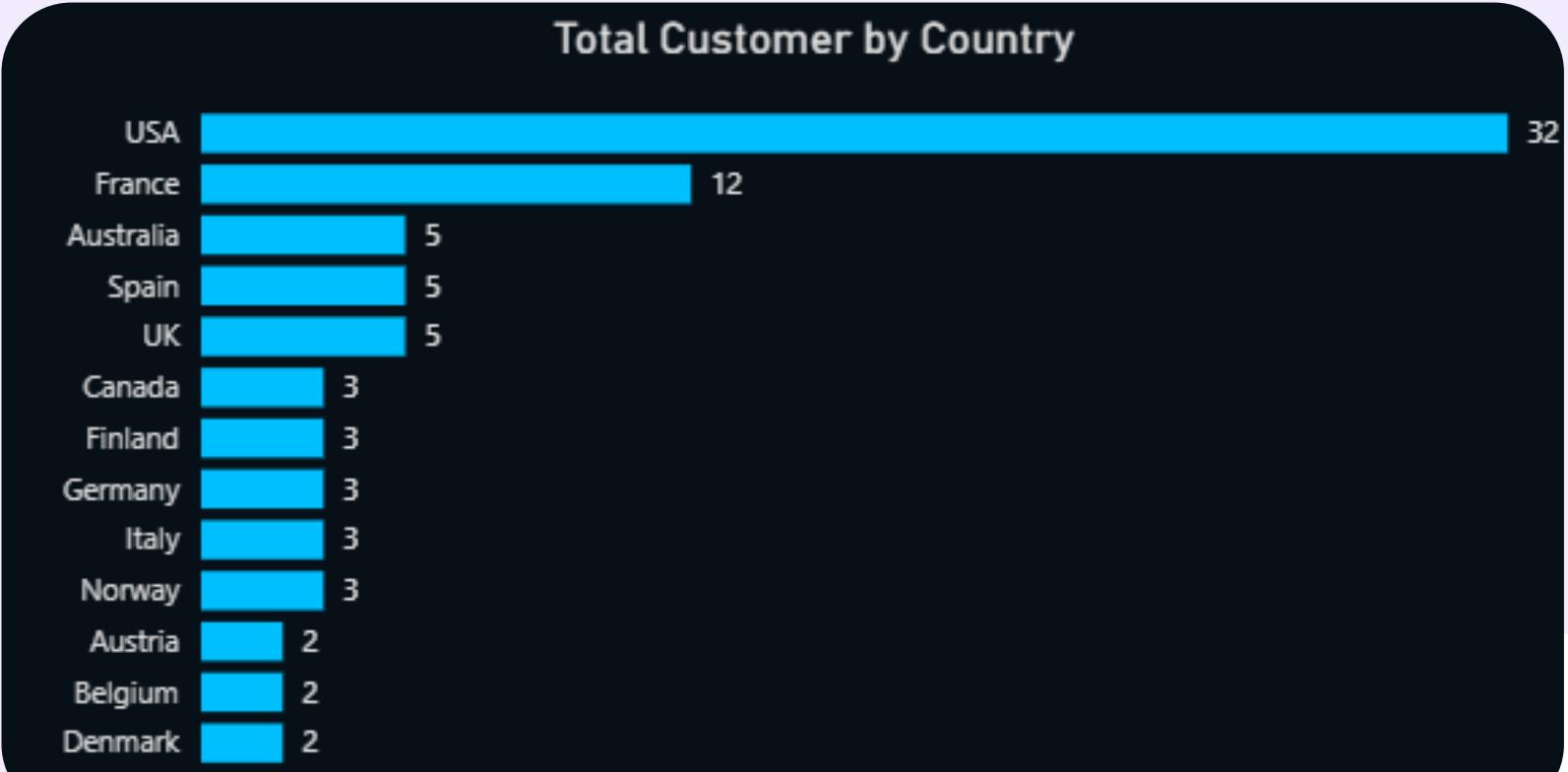
# ANALISIS DATA DAN PEMAHAMAN

## INSIGHT

### Negara mana dengan jumlah pelanggan terbanyak?

- **USA** merupakan pasar utama yang menyumbang jumlah pelanggan terbanyak (**32**). Oleh karena itu, strategi pemasaran dan distribusi harus difokuskan ke pasar ini, **seperti penawaran eksklusif, layanan prioritas, serta promosi lokal** yang sesuai dengan preferensi pelanggan di Amerika Serikat. Di sisi lain, negara-negara seperti France, Australia, dan UK juga menunjukkan potensi pasar menengah yang dapat **ditingkatkan dengan pendekatan kampanye yang lebih agresif**

## VISUALISASI DASHBOARD



# ANALISIS DATA DAN PEMAHAMAN

## INSIGHT

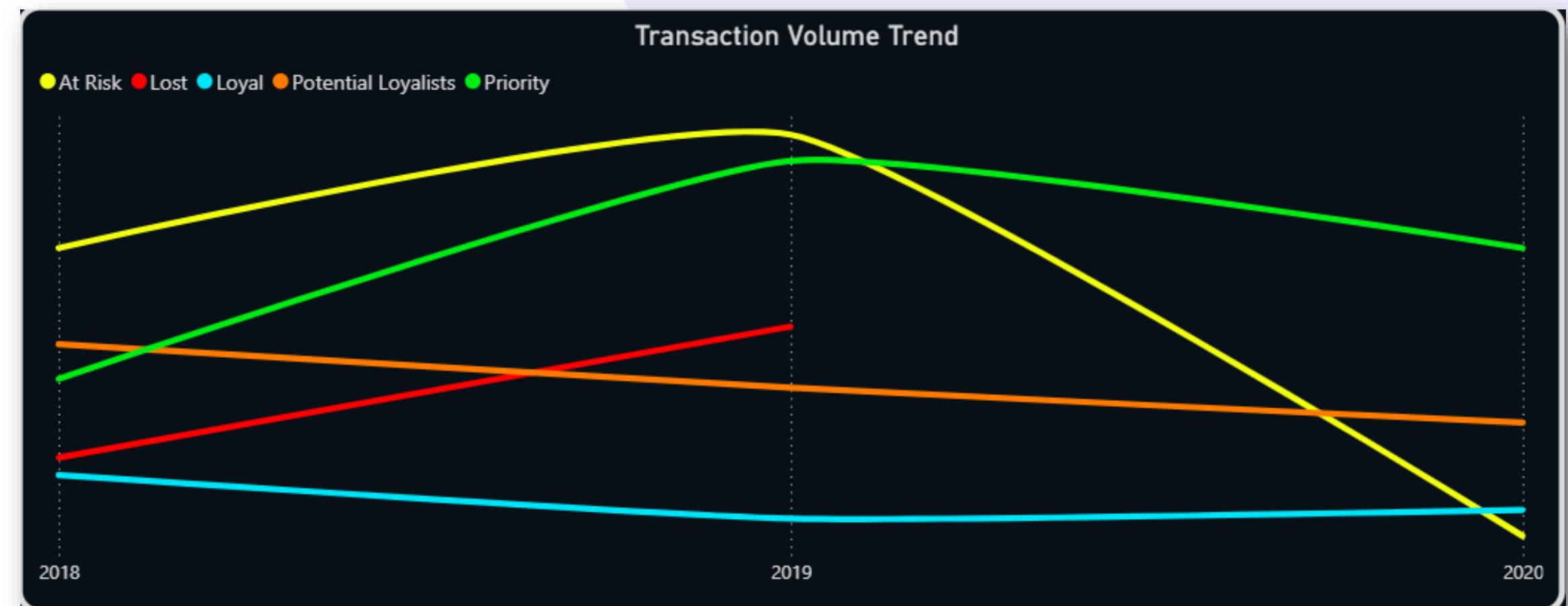
Pengaruh Low RFM Score terhadap pendapatan yang diperoleh?

- Terdapat sejumlah pelanggan dari segmen **At Risk** dan **Lost** yang memiliki **kontribusi penjualan (monetary)** **sangat tinggi**, meskipun RFM Score mereka tergolong rendah. Seluruh pelanggan dalam tabel ini berasal dari **USA**, yang merupakan pasar terbesar perusahaan. Maka Perusahaan perlu menerapkan strategi pemasaran yang menarik minat beli seperti flash sale terhadap produk dengan pendapatan tertinggi, dan promo eksklusif bagi pelanggan yang sudah lama tidak bertransaksi.

## VISUALISASI DASHBOARD

Sales of Customer with Low RFM Scores				
Customer Name	Customer Segment	RFM Score	Country	Monetary
Auto-Moto Classics Inc.	At Risk	341	USA	1,835,676
Cambridge Collectables Co.	Lost	121	USA	1,895,418
Classic Gift Ideas, Inc	At Risk	223	USA	5,902,663
Classic Legends Inc.	At Risk	343	USA	6,764,617
Collectable Mini Designs Co.	Lost	124	USA	7,285,910
Diecast Collectables	Lost	122	USA	5,795,477
Gift Ideas Corp.	At Risk	341	USA	3,714,369
Land of Toys Inc.	At Risk	255	USA	12,758,551
Marta's Replicas Co.	At Risk	224	USA	9,218,030
Microscale Inc.	At Risk	221	USA	1,996,920
Mini Classics	At Risk	223	USA	6,979,177
Motor Mint Distributors Inc.	At Risk	342	USA	4,862,293
Online Diecast Creations Co.	At Risk	244	USA	7,249,599
Online Mini Collectables	Lost	121	USA	3,414,230

# ANALISIS DATA DAN PEMAHAMIAN

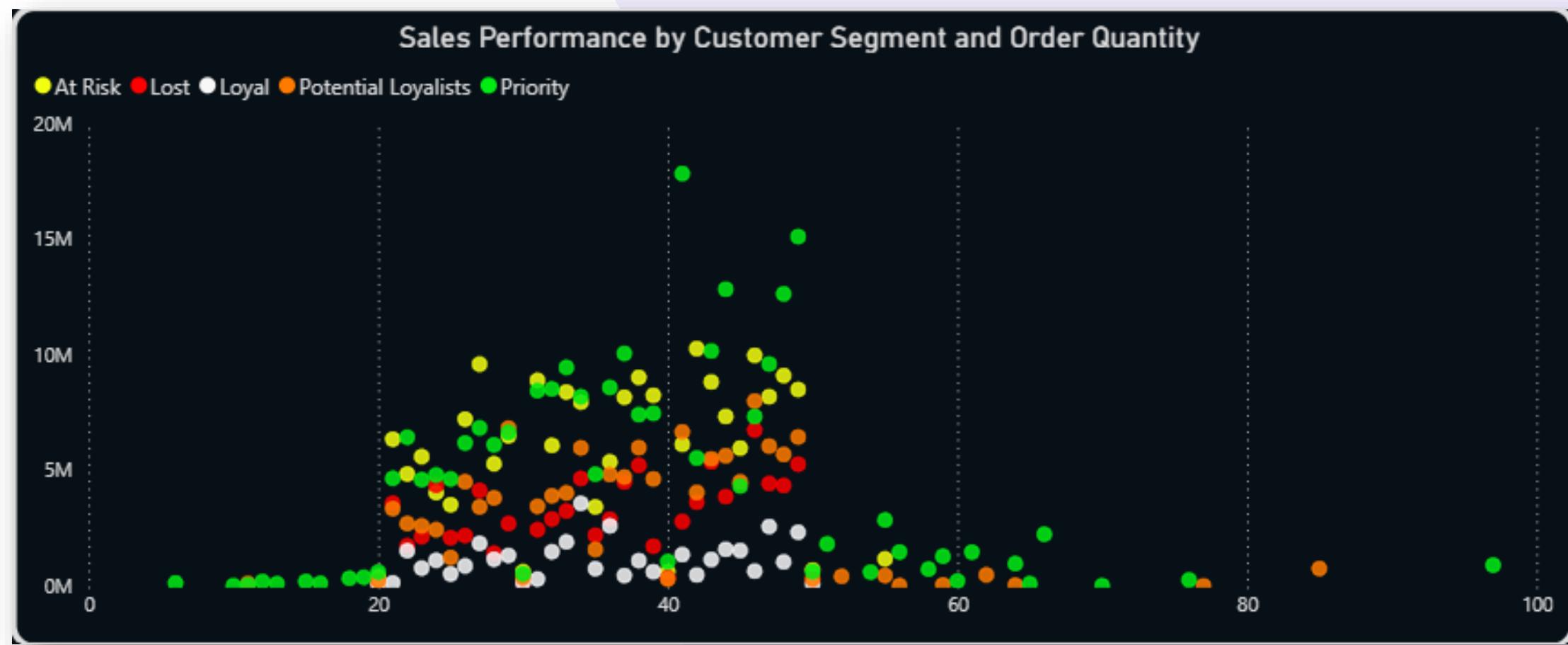


## INSIGHT

Bagaimana trend penjualan dari waktu ke waktu (tahun)?

- **At Risk** mengalami penurunan transaksi yang cukup tajam dari **tahun 2019 hingga 2020** sehingga perusahaan harus segera melakukan **strategi pencegahan churn** seperti komunikasi personal untuk menjaga engagement mereka.
- **Priority** mengalami **peningkatan hingga 2019**, lalu **sedikit menurun di 2020**, tetap konsisten jaga kepercayaan mereka dengan loyalty program eksklusif
- Mengingat tren naik pada **segment Lost**, penting untuk **mengidentifikasi alasan kehilangan pelanggan**, baik melalui survei, feedback, atau analisis perilaku sebelum mereka menjadi tidak aktif.
- **Segment Loyal dan Potential Loyalists** terbilang stabil akan tetapi cenderung menurun perlu upaya untuk meningkatkan retensi di segment ini, seperti mengembangkan **reward system** seperti pemberian **voucher diskon dan cashback**

# ANALISIS DATA DAN PEMAHAMAN



## INSIGHT

Apakah ada hubungan antara banyaknya order dengan total pendapatan yang diperoleh?

- Pelanggan dari **segment Priority** secara konsisten mendominasi penjualan dengan nilai transaksi tertinggi hingga **17M**, maka perusahaan perlu mempertahankan relasi melalui loyalty program eksklusif
- **Segment Loyal dan Lost** justru menunjukkan performa penjualan yang cenderung rendah dan jumlah pesanan yang lebih kecil (**49 unit**), maka perlu dilakukan upaya upselling.
- Segment **At Risk** dan **Potential Loyalists** juga menunjukkan kontribusi signifikan dalam kisaran penjualan **8M-10M**, dengan jumlah pesanan rata-rata di bawah **50 unit**, perusahaan perlu mempercepat konversi menjadi segmen Loyal/Priority dengan strategi retensi seperti program cashback, dan reminder berkala.

# REKOMENDASI

## Fokus pada peningkatan skala order high

Mengacu pada Customer Corrida Auto Replicas, Ltd yang memiliki persentase skala order high paling tinggi serta produk yang sering mereka beli adalah Classic Cars maka perusahaan harus lebih menekankan promosi terhadap produk tersebut karena sangat berpotensi untuk menunjang pendapatan yang lebih tinggi

1

## Kembangkan Strategi Khusus untuk Low Order - High Value Customer

Lakukan penawaran produk premium dengan bundling eksklusif sesuai dengan kebutuhan dan preferensi pelanggan agar meningkatkan kuantitas pesanan

2



# REKOMENDASI

## Tingkatkan Loyalitas Pelanggan dengan Fokus pada Segmen Loyal yang Underperforming

Buat sistem poin/reward berdasarkan nilai transaksi, frekuensi pembelian, atau aktivitas lain (review produk, referral, dsb), serta tawarkan insentif eksklusif khusus segment loyal seperti voucher diskon, pengiriman gratis agar mendorong terjadinya transaksi

3

## Fokus Pendistribusian Stok Produk dengan Pendapatan Tertinggi

Prioritaskan distribusi ke wilayah pasar utama seperti USA yang sebelumnya telah dilakukan analisa produk dengan penjualan terbanyak agar memastikan produk tersebut selalu ready stock dan mengurangi resiko kekecewaan dari pelanggan, terutama pada segment Priority

4



# LET'S CONNECT



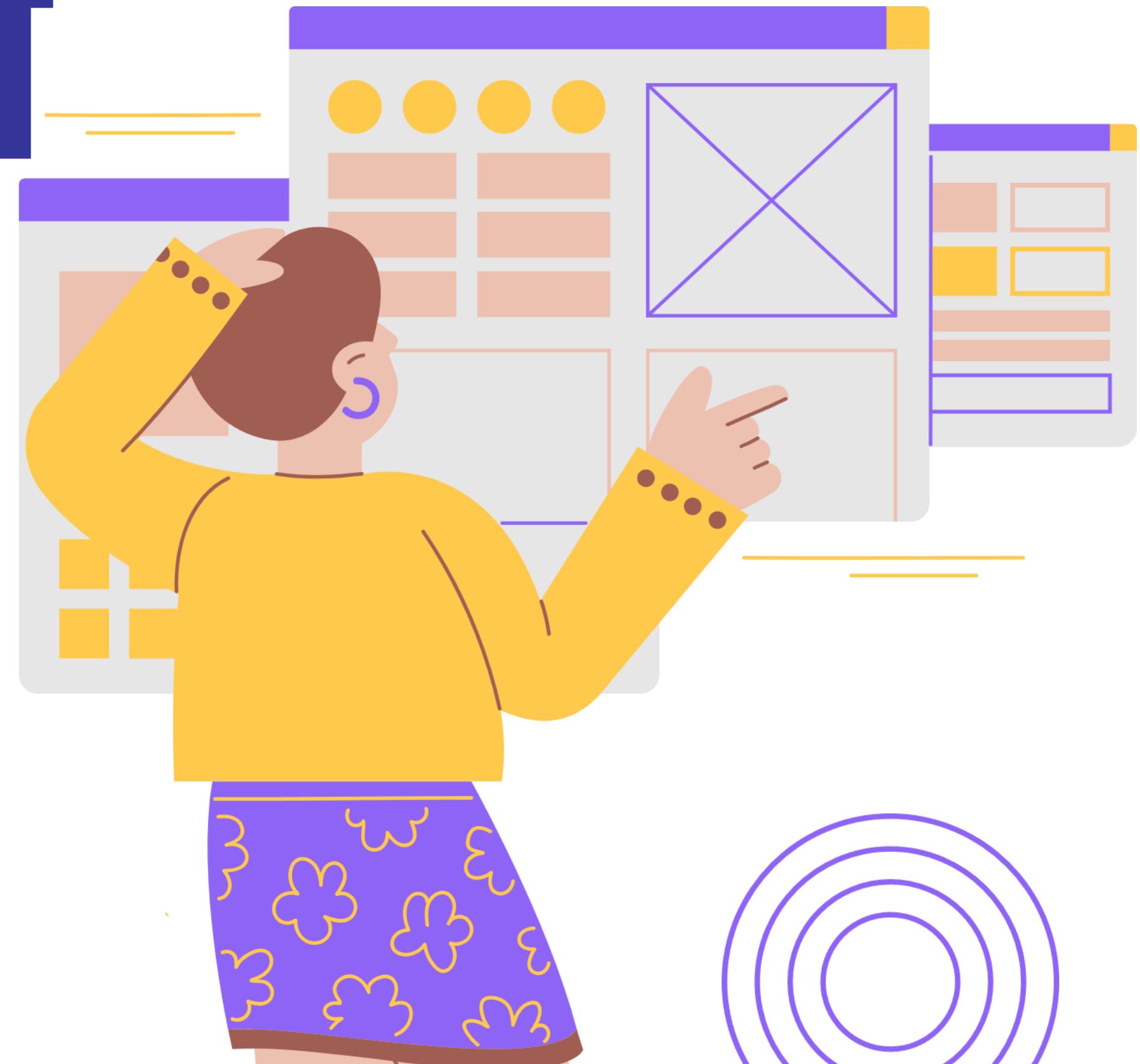
Bogor, Indonesia



ardyansyah3199@gmail.com



Ardy Ansyah



# **THANK YOU!**