

Customer Segmentation of B2B Vehicles Sales Dataset

Batch 34 | Bootcamp Data Science

Valerius Evan Ligasetiawan



Valerius Evan L

Student DS34+

-

LinkedIn

[Valerius Evan](#)

My other profile

- [Github](#)

Outline



Business Use Case

- Sales Performance
- Customer Segmentation and Profiling
- Product Performance Over Time
- Geographic Insights



Data Preprocessing

- Need to use Colab [here](#)
- Check null, outlier, duplicate dan format data tertentu
- Bisa di Collab/Python atau di Power BI, tergantung kebutuhan



Data Visualization

- Visualize business use case and RFM segmentation
- Building a dashboard

Where You Get the Materials

The whole materials can be found [here](#)

Prepare your own drive folder that at least contain the following items to get better learning experience:

- Copy of *SalesDataSampleEDA.ipynb*
- Copy of *sales_data_sample.csv*

A large black parallelogram is positioned on the left side of the slide. Below it, two overlapping parallelograms in shades of orange and yellow are also positioned on the left, creating a layered, abstract design.

Business Use Case

Executive Summary

- Penjualan B2B kendaraan menunjukkan nilai total yang tinggi, terutama didorong oleh segmen pelanggan tertentu dengan volume order besar.
- Namun, retensi pelanggan tidak merata: segmen **Champions** dan **Loyal** konsisten berkontribusi pada pendapatan, sementara segmen lain seperti **At Risk** dan **Hibernating** memperlihatkan penurunan aktivitas pembelian.
- Pola pipeline juga beragam, ada pelanggan dengan transaksi kecil namun berulang secara stabil, serta pelanggan dengan pembelian besar tetapi jarang melakukan repeat order.
- Kondisi ini menunjukkan adanya potensi risiko **churn**, khususnya dari pelanggan dengan nilai transaksi tinggi tetapi jarang melakukan pembelian kembali dalam periode terakhir.



Business Use Case

- Dalam kasus ini, kita punya dataset yang merupakan car dealer b2b
- Bagaimana penjualan total tiap periode tertentu?
- Customer mana yang paling berkontribusi untuk revenue dan frekuensi pembelian (juga deal size)?
- Kota/negara/teritori mana yang paling berkontribusi pada pendapatan dealer?



A large black parallelogram is positioned on the left side of the slide. Overlapping its bottom edge are two smaller parallelograms: a light orange one on top and a darker orange one below it, both pointing towards the right.

Data Preprocessing

Data Understanding and Method

- **Dataset:** Order-level B2B vehicle sales (2003–2005).
- **Method:** RFM (Recency, Frequency, Monetary) → scoring → labeling segment.
- **Preprocessing:** No null, no duplicate, but there is outlier, so it is preprocessed using median. Date/time will be formatted in Power BI and customer ID is customized based on trial and error.



Abstract geometric shapes consisting of a large black parallelogram, a medium yellow parallelogram, and a small orange parallelogram, all slanted to the right.

Data Visualization

Periodic Sales

- Monthly



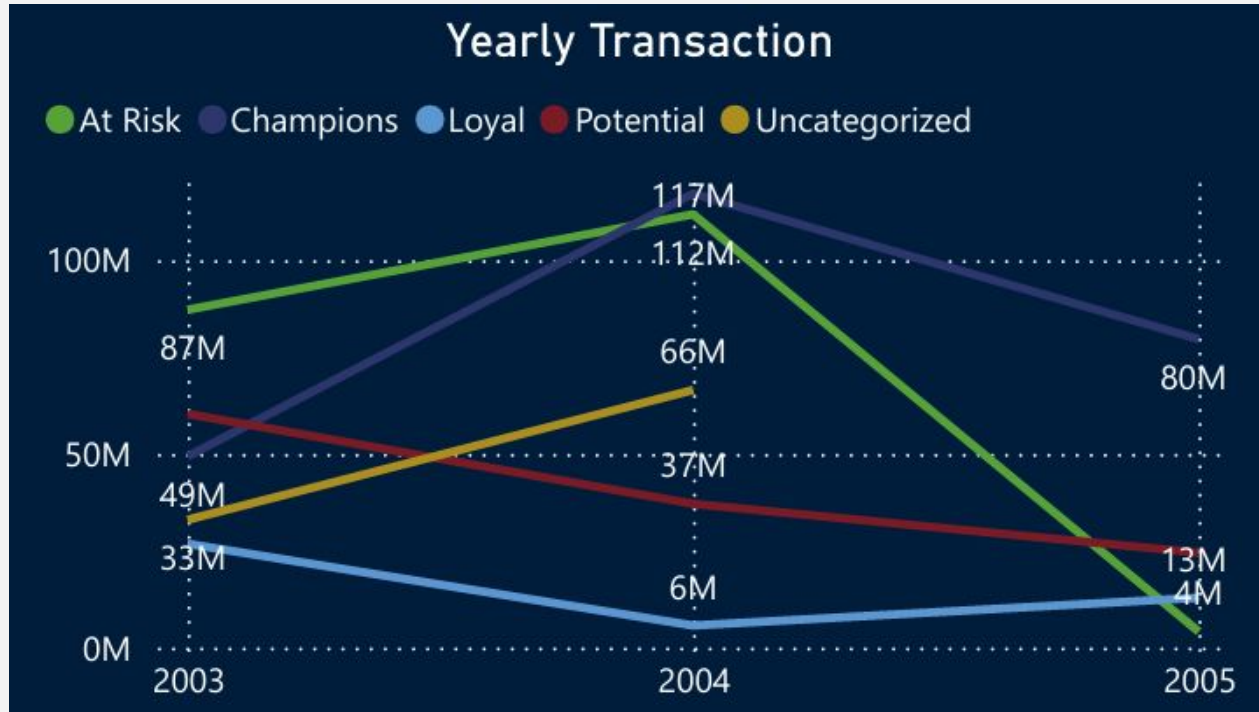
Periodic Sales

- Quarterly



Periodic Sales

- Yearly



Periodic Sales

- Month and Year



Product Line Revenue (sum of sales)



Product Line Revenue (sum of sales)

ProductLine Yearly Contribution

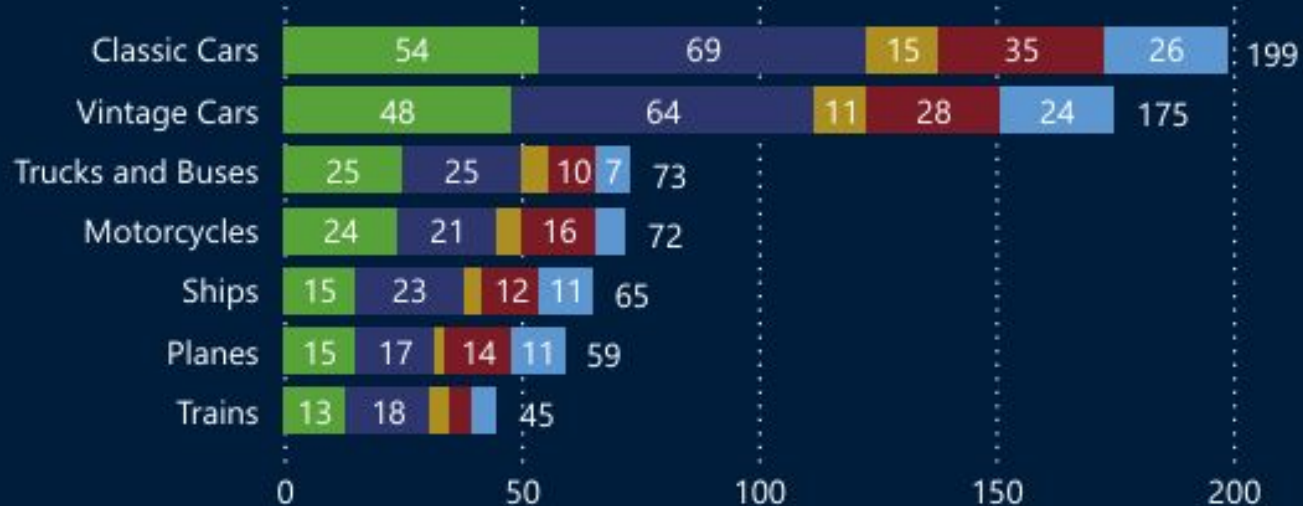
PRODUCTLINE	2003	2004	2005	Total
Classic Cars	107873913	126578910	45047479	279500302
Motorcycles	24647643	42082950	15342490	82073083
Planes	18750163	35247656	13565676	67563495
Ships	16487022	24812813	10040506	51340341
Trains	5794113	8722669	3066332	17583114
Trucks and Buses	35214621	39126359	11891397	86232377
Vintage Cars	47988444	61799101	22235502	132023047
Total	256755919	338370458	121189382	716315759



Product Line Orders

of Order per ProductLine

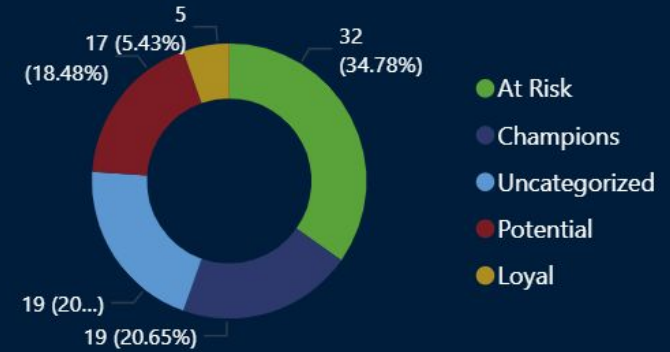
● At Risk ● Champions ● Loyal ● Potential ● Uncategorized



Customer Segmentation

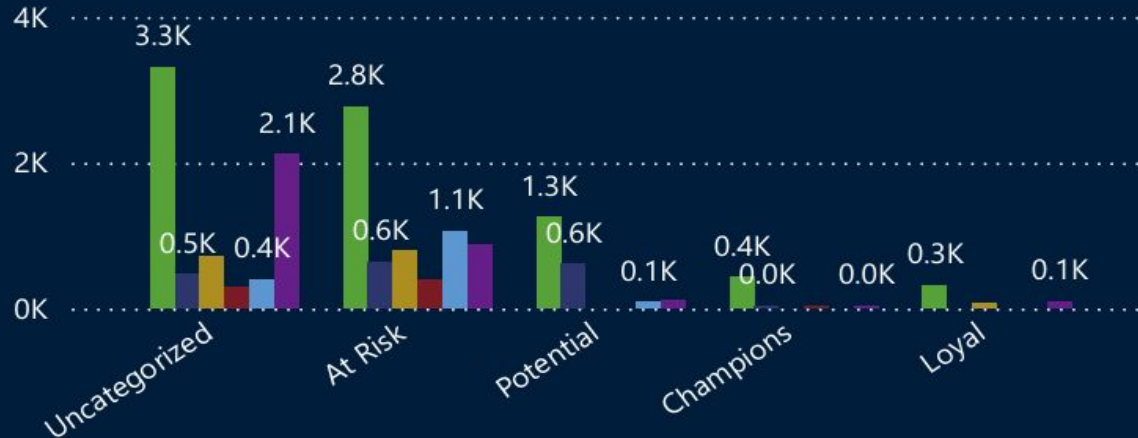
- Customer mana yang paling berkontribusi untuk revenue dan frekuensi pembelian (juga deal size)?

Customer per Segment



Recency by Segment

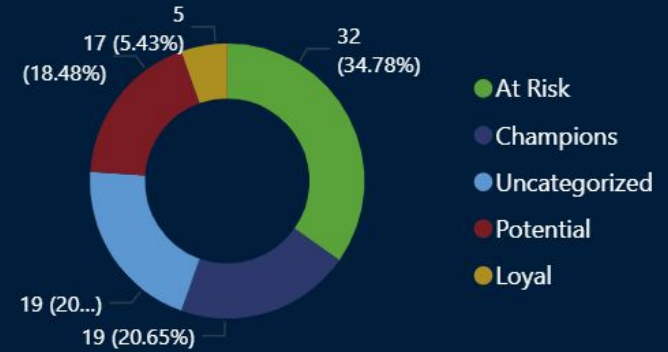
Classic Cars Motorcycles Planes Ships Trucks and Buses Vintage Cars



Customer Segmentation

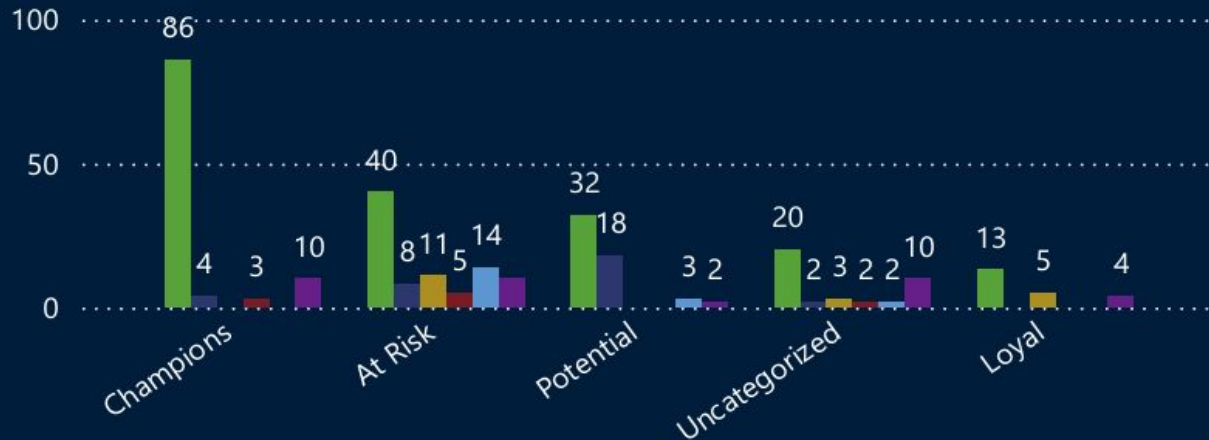
- Customer mana yang paling berkontribusi untuk revenue dan frekuensi pembelian (juga deal size)?

Customer per Segment



Frequency by Segment

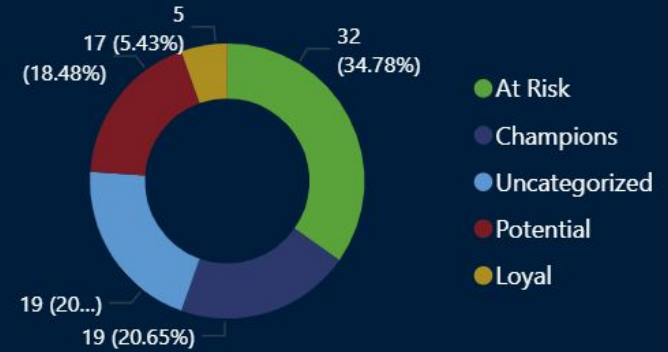
Classic Cars Motorcycles Planes Ships Trucks and Buses Vintage Cars



Customer Segmentation

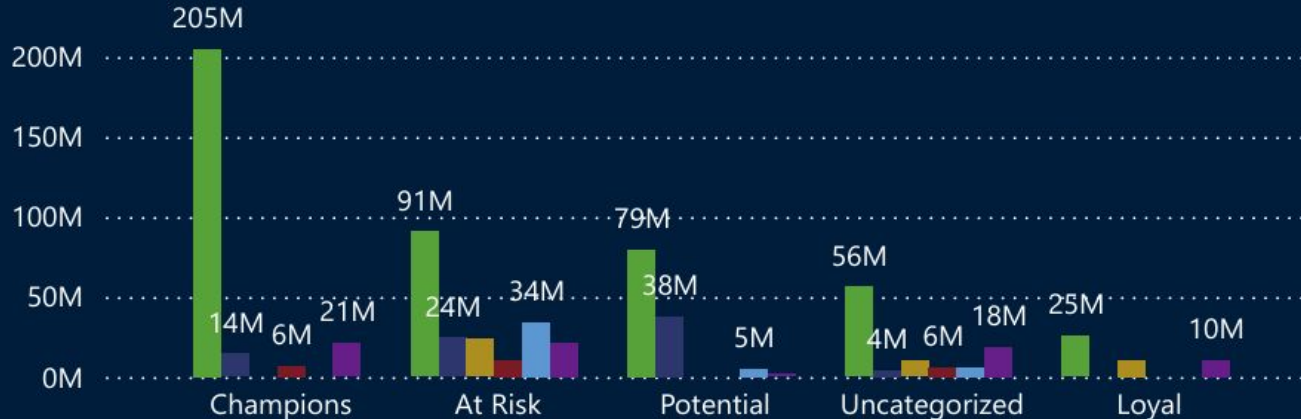
- Customer mana yang paling berkontribusi untuk revenue dan frekuensi pembelian (juga deal size)?

Customer per Segment



Monetary by Customer Segment

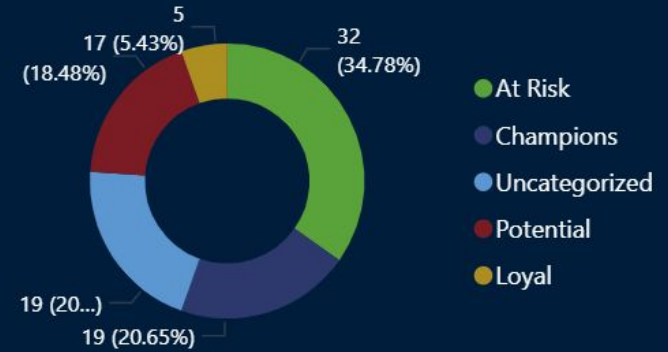
● Classic Cars ● Motorcycles ● Planes ● Ships ● Trucks and Buses ● Vintage Cars



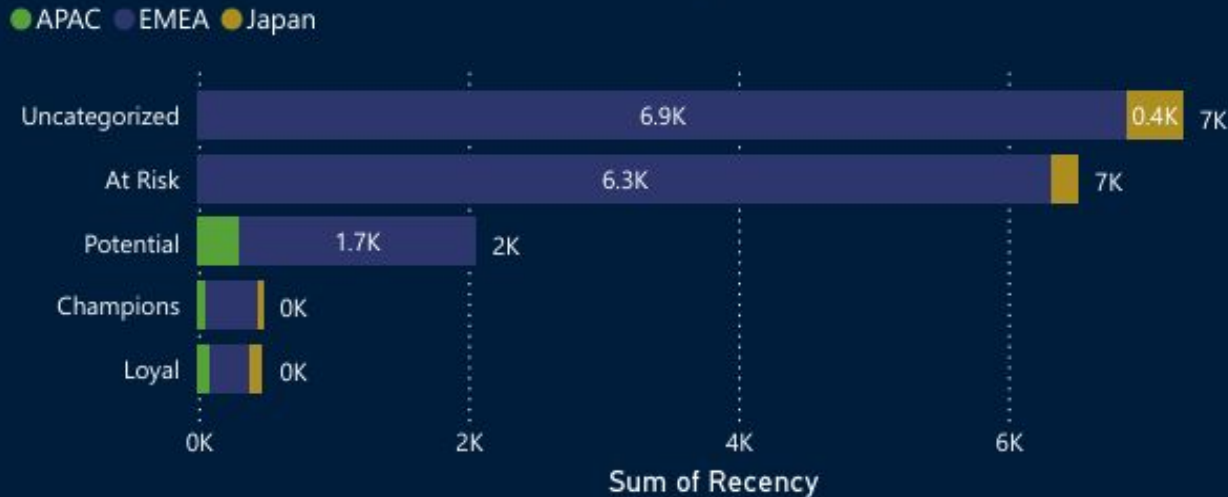
Territorial Segmentation

- Customer territory mana yang paling sering berkunjung?

Customer per Segment



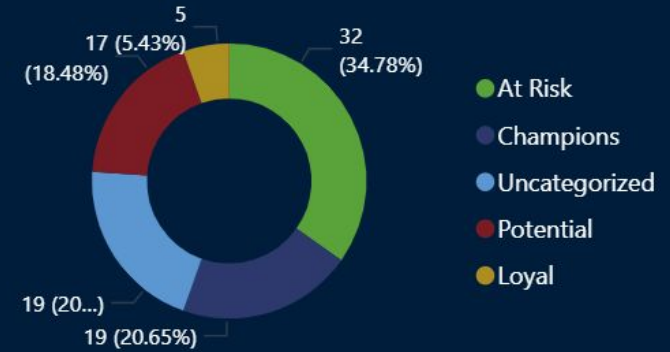
Recency by Segment



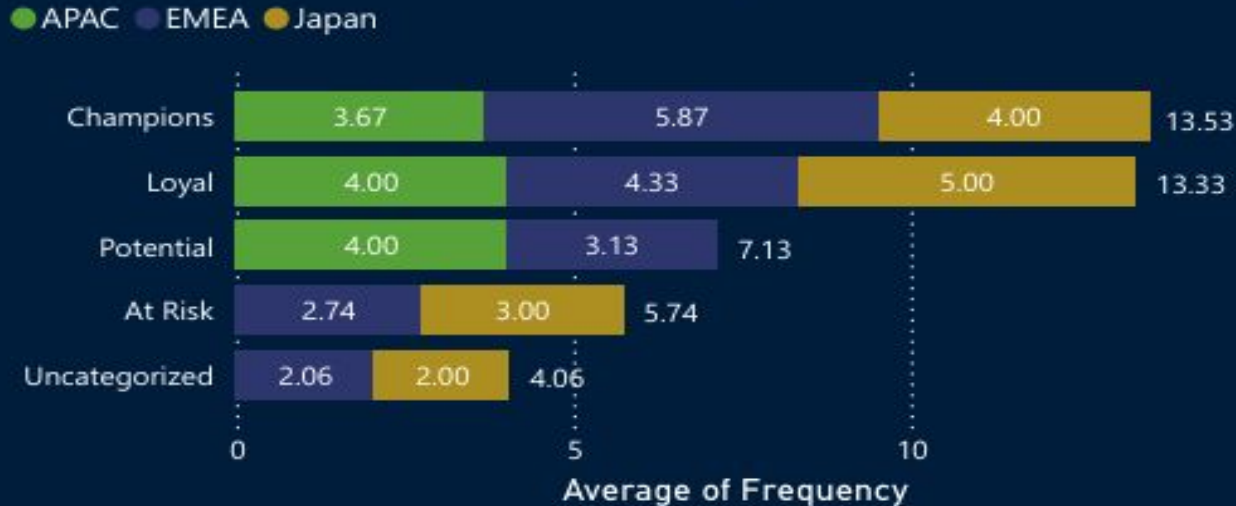
Territorial Segmentation

- Customer territory mana yang paling sering melakukan order?

Customer per Segment



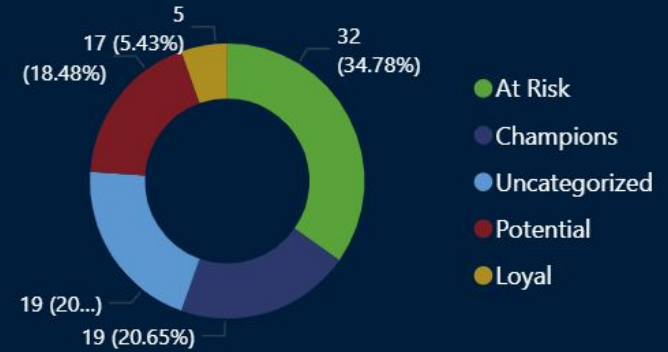
Frequency by Segment



Territorial Segmentation

- Customer territory mana yang paling berkontribusi untuk revenue?

Customer per Segment



Monetary by Segment



Geographic Insight

- Kota/negara/teritori mana yang paling berkontribusi pada pendapatan dealer?



Total Sales by City

● At Risk ● Champions ● Loyal ● Potential ● Uncategorized



Total Sales by Country

● At Risk ● Champions ● Loyal ● Potential ● Uncategorized



Total Sales by Territory

● At Risk ● Champions ● Loyal ● Potential ● Uncategorized



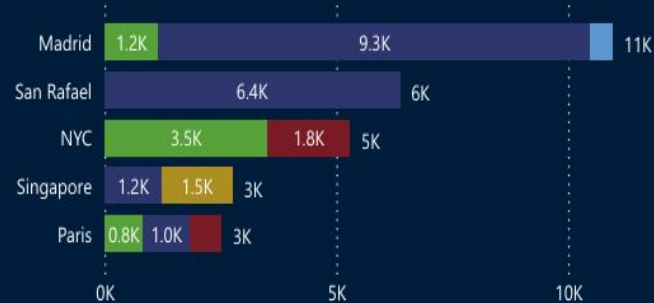
Geographic Insight

- Kota/negara/teritori mana yang paling sering melakukan order ke dealer?



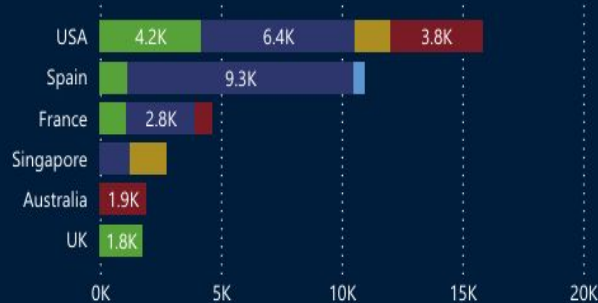
Total Order by City

At Risk Champions Loyal Potential Uncategorized



Total Order by Country

At Risk Champions Loyal Potential Uncategorized



Total Sales by Territory

At Risk Champions Loyal Potential Uncategorized



Executive Summary

1. **Total Sales 2003–2005: 716M**, peak di 2004.
2. **Customer Segmentation:**
 - a. Champions & Loyal = kontributor utama revenue.
 - b. At Risk = high monetary, low retention → risiko churn.
 - c. Potential = peluang naik ke Loyal.
3. **Top Product Lines:**
 - a. Classic Cars & Vintage Cars = >400M.
 - b. Trucks = kontribusi terendah.



Executive Summary

4. Geographic Insights:

- a. Negara: USA (~130M), Spain (~75M), France (~34M).
- b. Kota: Madrid (75M), San Rafael (46M), NYC (41M).
- c. Regional: APAC tertinggi (~129M).

5. Key Focus:

- a. Retain At Risk, konversi Potential → Loyal.
- b. Optimalkan produk unggulan, tingkatkan promosi produk
- c. Perkuat strategi di USA, Spain, APAC.



A large, stylized graphic on the left side of the slide. It consists of a blue outline of a person's head and shoulders. Inside the head is a yellow circle with an orange ring and a yellow center. Inside the torso is a large yellow circle with an orange ring and a yellow center.

Terima Kasih