

Analyzing eCommerce Business Performance with SQL



Created by:

Devia Febyanti

devia.febyanti99@gmail.com

<https://www.linkedin.com/in/devia-febyanti/>

Experienced data operator with proven work history in the field of education. Currently motivated to become a Data Scientist to help driving financial inclusion through the use of big data and machine learning.



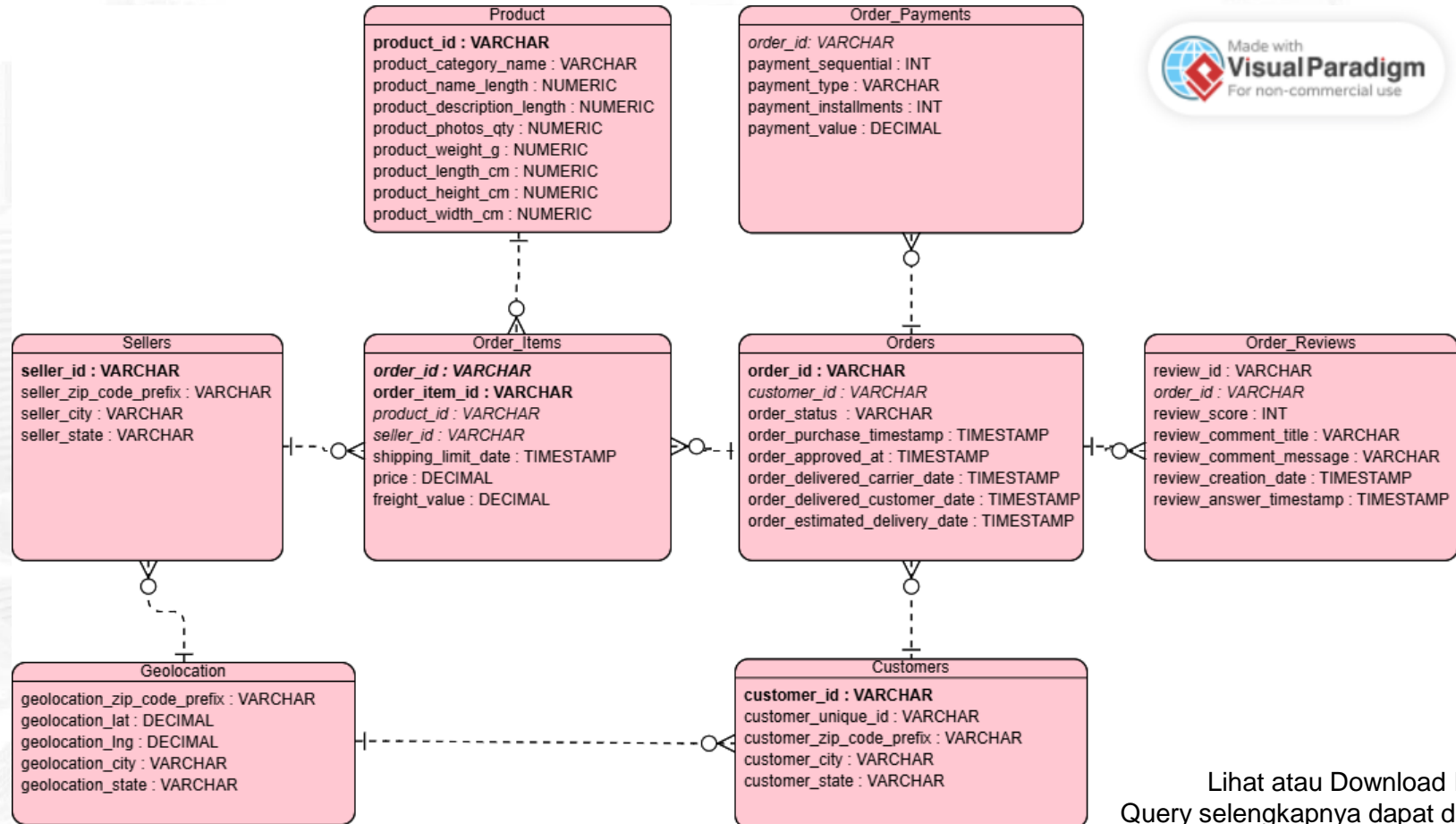
Bachelor of Geophysics
2018-2022

“Dalam suatu perusahaan mengukur performa bisnis sangatlah penting untuk melacak, memantau, dan menilai keberhasilan atau kegagalan dari berbagai proses bisnis. Oleh karena itu, dalam paper ini akan menganalisa performa bisnis untuk sebuah perusahaan eCommerce, dengan memperhitungkan beberapa metrik bisnis yaitu pertumbuhan pelanggan, kualitas produk, dan tipe pembayaran.”

Berikut adalah ringkasan dari apa yang telah dilakukan:

1. Desain Basis Data (Database Schema) → Saya merancang basis data menggunakan SQL untuk merepresentasikan berbagai entitas seperti pelanggan, pesanan, produk, penilaian, dan lokasi geografis. Setiap tabel memiliki struktur yang ditentukan dengan primary key dan foreign key yang sesuai.
2. Impor Data (Data Import) → Data dari berbagai dataset yang diberikan (seperti data pelanggan, pesanan, produk, dan lainnya) diimpor ke dalam database menggunakan perintah COPY dalam SQL. Proses ini memastikan bahwa data mentah dari dataset CSV diintegrasikan dengan benar ke dalam struktur basis data yang telah dirancang.
3. Perbaikan Masalah Data (Data Cleansing) → Saya mengidentifikasi dan memperbaiki masalah seperti duplikasi data atau nilai yang tidak valid dalam beberapa tabel, seperti pada :
 - Entitas product memiliki kolom pertama yang tidak memiliki nama kolom tetapi berisi data indeks/baris. Solusi saya dengan memberikan nama kolom pertama sebagai row_index agar proses COPY dalam SQL berhasil. Kemudian setelah berhasil saya melakukan drop column row_index tersebut karena tidak dibutuhkan.
 - Entitas product yang memiliki format tanggal dan waktu dalam CSV yang tidak sesuai dengan format yang diharapkan oleh PostgreSQL. Solusi saya dengan mengubah pengaturan datestyle di PostgreSQL.
 - Entitas order_reviews yang memiliki review_id tidak unik. Solusi saya dengan mengubah nilai review_id yang tidak unik dengan menambahkan nomor urut di belakangnya untuk memastikan keunikan.
4. Pemodelan dan Analisis Hubungan Antar Tabel → Saya memodelkan hubungan antar tabel dalam bentuk Entity Relationship Diagram (ERD), yang mencakup hubungan: One-to-Many (Product & Order_Items, Sellers & Order_Items, Orders & Order_Reviews, Geolocation & Customers) dan Many-to-One (Order_Payments & Orders, Order_Items & Orders, Sellers & Geolocation, Orders & Customers).

Data Preparation - ERD (Entity Relationship Diagram)



Lihat atau Download ERD [disini](#)
Query selengkapnya dapat dilihat [disini](#)

Annual Customer Activity Growth Analysis

Berikut metrik-metrik yang telah berhasil ditampilkan menjadi satu tampilan tabel:

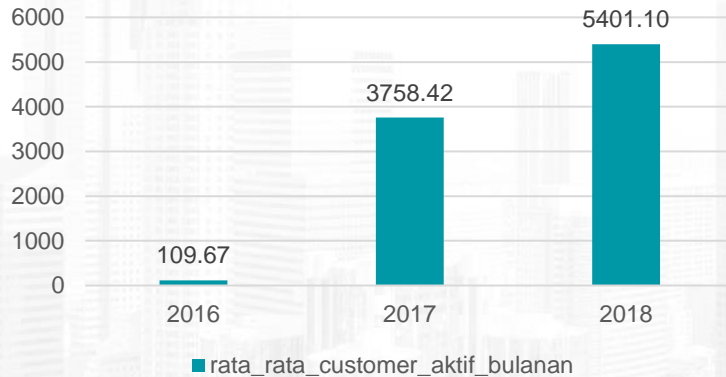
tahun	rata_rata_customer_aktif_bulanan	jumlah_customer_baru	jumlah_customer_repeat_order	rata_rata_order_per_customer
2016	109.6666667	329	NULL	1
2017	3758.416667	45101	NULL	1
2018	5401.1	54011	NULL	1

Berdasarkan tabel, dapat diketahui bahwa:

1. Terjadi peningkatan rata-rata Monthly Active User (MAU) setiap tahun.
2. Terjadi peningkatan jumlah pelanggan baru setiap tahun.
3. Tidak ada customer yang melakukan repeat order.
4. Rata-rata customer hanya melakukan order 1 kali setiap tahun.

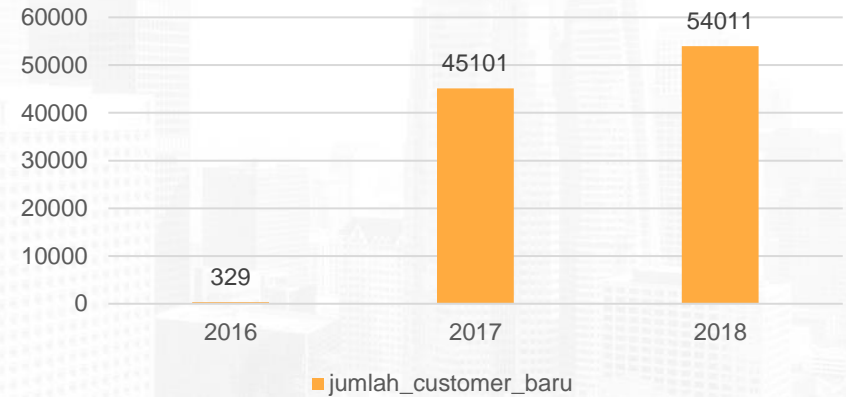
Annual Customer Activity Growth Analysis

Rata-rata Customer Aktif Bulanan



- Terdapat lonjakan besar dalam rata-rata customer aktif bulanan dari tahun 2016 ke 2017 dengan peningkatan lebih dari 34 kali lipat. Ini menunjukkan bahwa ada pergeseran besar dalam jumlah pelanggan yang aktif setiap bulannya. Kemungkinan besar, ini bisa disebabkan oleh peluncuran kampanye pemasaran yang sangat efektif, peningkatan dalam penawaran produk atau layanan, atau mungkin akuisisi pelanggan yang lebih agresif.
- Dari 2017 ke 2018, rata-rata customer aktif bulanan meningkat lagi sekitar 43.8%. Ini menunjukkan bahwa perusahaan berhasil mempertahankan pertumbuhan yang kuat dan terus menarik serta mempertahankan pelanggan aktif.

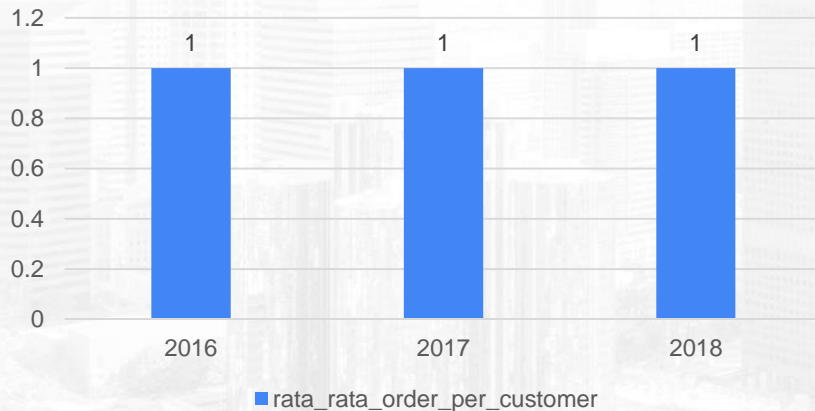
Jumlah Customer Baru



- Jumlah customer baru meningkat tajam dari tahun 2016 ke tahun 2017. Menunjukkan peningkatan lebih dari 137 kali lipat.
- Pada tahun 2018, jumlah customer baru meningkat lagi sekitar 20% dari tahun 2017. Meskipun pertumbuhan tidak secepat lonjakan dari tahun 2016 ke 2017, peningkatan ini masih signifikan dan menunjukkan bahwa perusahaan terus menarik banyak pelanggan baru setiap tahunnya.

Annual Customer Activity Growth Analysis

Rata-rata Order per Customer



Setiap customer rata-rata hanya melakukan satu kali order per tahun selama periode 2016-2018.

- Menunjukkan bahwa pelanggan cenderung tidak kembali untuk melakukan pembelian berulang dalam setahun. Ini bisa menjadi indikasi bahwa pelanggan tidak cukup terlibat atau merasa tidak ada kebutuhan untuk melakukan pembelian tambahan dalam periode waktu tersebut.
- Frekuensi order yang rendah dapat mengindikasikan adanya masalah dengan kepuasan pelanggan. Pelanggan mungkin tidak puas dengan produk atau layanan yang mereka terima, sehingga mereka tidak terdorong untuk kembali dan melakukan pembelian ulang.
- Rata-rata frekuensi order yang rendah juga bisa menunjukkan kurangnya program loyalitas atau insentif yang efektif untuk mendorong pelanggan melakukan pembelian berulang.

Kesimpulan

Berdasarkan visualisasi dan interpretasi, dapat disimpulkan bahwa terdapat pertumbuhan yang positif dalam berbagai aspek terkait pelanggan sebagai berikut:

1. Peningkatan rata-rata Monthly Active User (MAU) setiap tahunnya menunjukkan keterlibatan pelanggan yang meningkat setiap bulannya.
2. Jumlah pelanggan baru yang meningkat setiap tahun menunjukkan efektivitas strategi akuisisi pelanggan.
3. Data menunjukkan bahwa tidak ada pelanggan yang melakukan repeat order setiap tahunnya. Ini merupakan indikasi serius bahwa tidak ada loyalitas pelanggan dan pelanggan cenderung tidak kembali setelah pembelian pertama.
4. Setiap customer rata-rata hanya melakukan satu kali order per tahun selama periode 2016-2018. Hal ini menunjukkan bahwa pelanggan cenderung tidak melakukan pembelian berulang dalam setahun.

Annual Product Category Quality Analysis

Berikut ini informasi-informasi yang telah didapatkan ke dalam satu tampilan tabel:

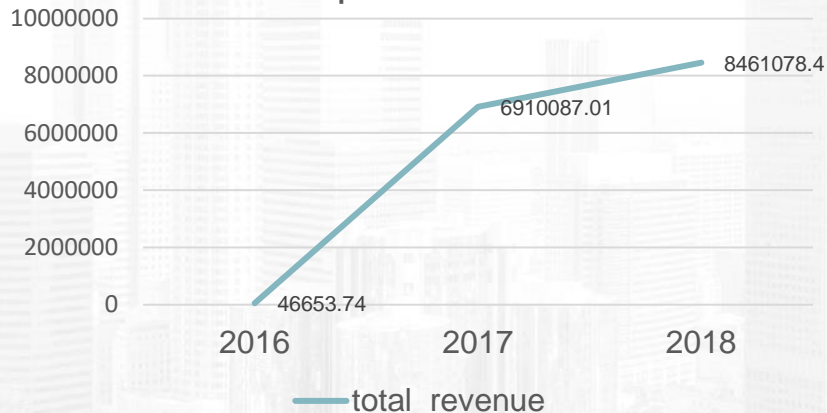
year	total_revenue	total_cancel_orders	top_revenue_category	top_cancel_category
2016	46653.74	26	furniture_decor	toys
2017	6910087.01	265	bed_bath_table	sports_leisure
2018	8461078.4	334	health_beauty	health_beauty

Berdasarkan tabel, dapat diketahui bahwa:

1. Terjadi peningkatan pendapatan/revenue perusahaan total setiap tahun.
2. Terjadi peningkatan jumlah cancel order total setiap tahun.
3. Kategori produk yang memiliki revenue tertinggi setiap tahunnya dari 2016 hingga 2018 adalah furniture_decor, bed_bath_table, dan health_beauty.
4. Kategori produk yang memiliki jumlah cancel order tertinggi setiap tahunnya dari 2016 hingga 2018 adalah toys, sports_leisure, dan health_beauty.

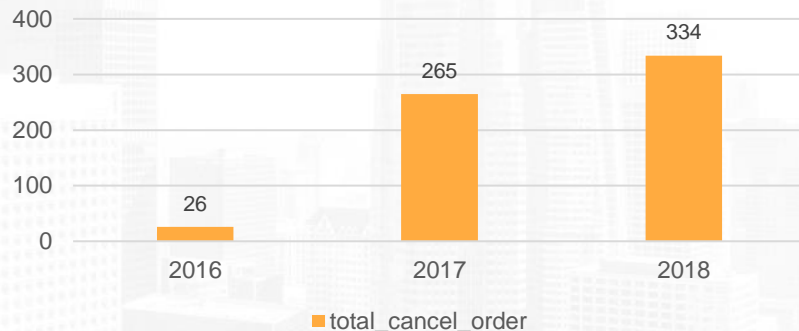
Annual Product Category Quality Analysis

Total Pendapatan Perusahaan



- Terdapat lonjakan besar dalam total revenue perusahaan dari tahun 2016 ke 2017 dengan peningkatan lebih dari 148 kali lipat. Peningkatan ini menunjukkan adanya pertumbuhan yang sangat besar dalam bisnis atau kegiatan yang menghasilkan pendapatan.
- Dari 2017 ke 2018, total revenue perusahaan meningkat lagi sekitar 22.4%. Ini menunjukkan pertumbuhan yang lebih stabil namun masih cukup signifikan.

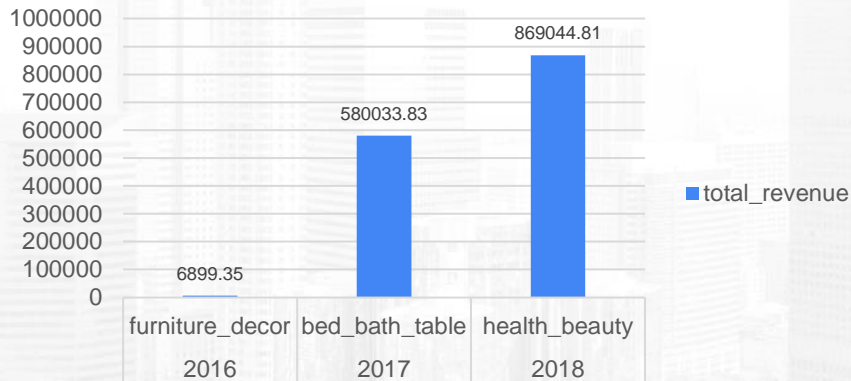
Jumlah Cancel Order



- Terdapat peningkatan yang sangat besar dalam jumlah pesanan yang dibatalkan dari tahun 2016 ke 2017. Hal ini mungkin menunjukkan adanya masalah dalam kualitas produk atau layanan, perubahan kebijakan yang menyebabkan lebih banyak pembatalan, atau peningkatan volume pesanan secara keseluruhan sehingga pembatalan juga meningkat.
- Dari tahun 2017 ke 2018 peningkatan jumlah pesanan yang dibatalkan masih terjadi, tetapi tidak sebesar peningkatan dari tahun sebelumnya. Ini bisa menunjukkan bahwa meskipun ada upaya untuk mengurangi jumlah pembatalan, masih ada beberapa faktor yang menyebabkan pembatalan tetap tinggi.

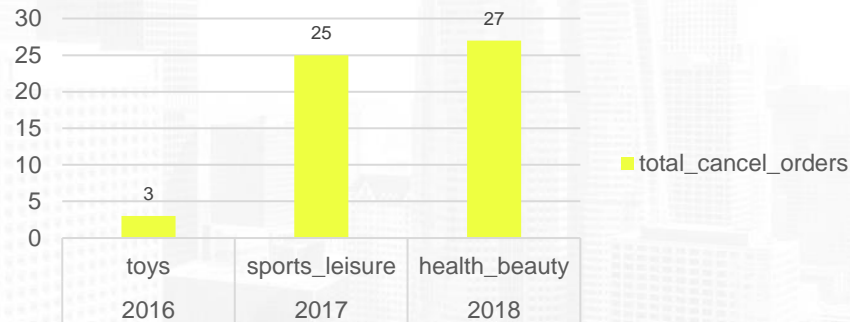
Annual Product Category Quality Analysis

Kategori Teratas Berdasarkan Pendapatan



- Pada tahun 2016, produk-produk yang berkaitan dengan perabotan dan dekorasi rumah adalah yang paling laris.
- Pada tahun 2017 terdapat perubahan signifikan dalam tren pembelian, di mana produk-produk yang berkaitan dengan tempat tidur, kamar mandi, dan meja menjadi sangat populer.
- Pada tahun 2018, produk-produk kesehatan dan kecantikan menjadi sangat diminati.

Kategori Teratas Berdasarkan Cancel Order



- Pada tahun 2016, mungkin kategori Toys memiliki beberapa masalah, seperti stok yang tidak mencukupi atau masalah kualitas produk yang menyebabkan pembatalan pesanan.
- Pada tahun 2017, terdapat peningkatan signifikan dalam pembatalan di kategori Sports & Leisure dibandingkan dengan tahun sebelumnya. Kemungkinan penyebabnya bisa termasuk masalah dalam pemenuhan pesanan, kualitas produk, atau perubahan dalam minat konsumen setelah pemesanan.
- Pada tahun 2018, meskipun kategori Health & Beauty populer, juga menghadapi tantangan dengan jumlah pesanan yang dibatalkan tertinggi. Ini bisa disebabkan oleh ekspektasi yang tinggi dari konsumen terhadap produk kesehatan dan kecantikan yang tidak terpenuhi.

Kesimpulan

Berdasarkan visualisasi dan interpretasi, dapat disimpulkan bahwa :

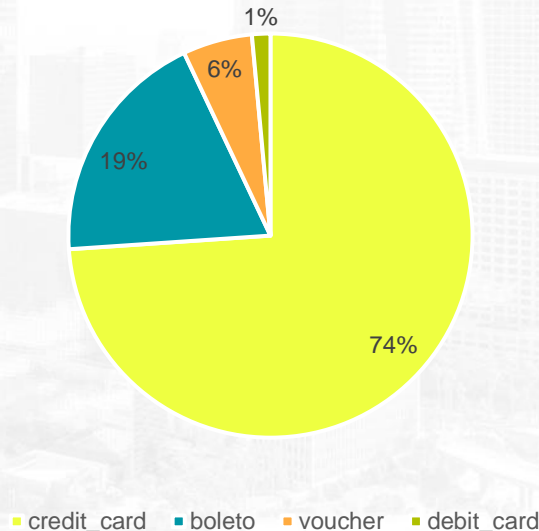
1. Bisnis telah berkembang pesat selama periode tiga tahun.berdasarkan total pendapatan yang diperoleh setiap tahun.
2. Peningkatan pembatalan yang lebih moderat mungkin menunjukkan upaya perbaikan telah dilakukan, namun masih perlu ditingkatkan lebih lanjut dan seningkatan signifikan dalam pembatalan pesanan perlu dianalisis untuk mengidentifikasi dan mengatasi penyebab utamanya.
3. Setiap tahun menunjukkan perubahan signifikan dalam kategori produk yang paling laris, mencerminkan dinamika pasar dan perubahan preferensi konsumen.
4. Setiap tahun ada kategori yang berbeda mengalami masalah dengan pesanan yang dibatalkan terbanyak.

Jumlah pengguna masing-masing tipe pembayaran:

payment_type	usage_count
credit_card	76795
boleto	19784
voucher	5775
debit_card	1529

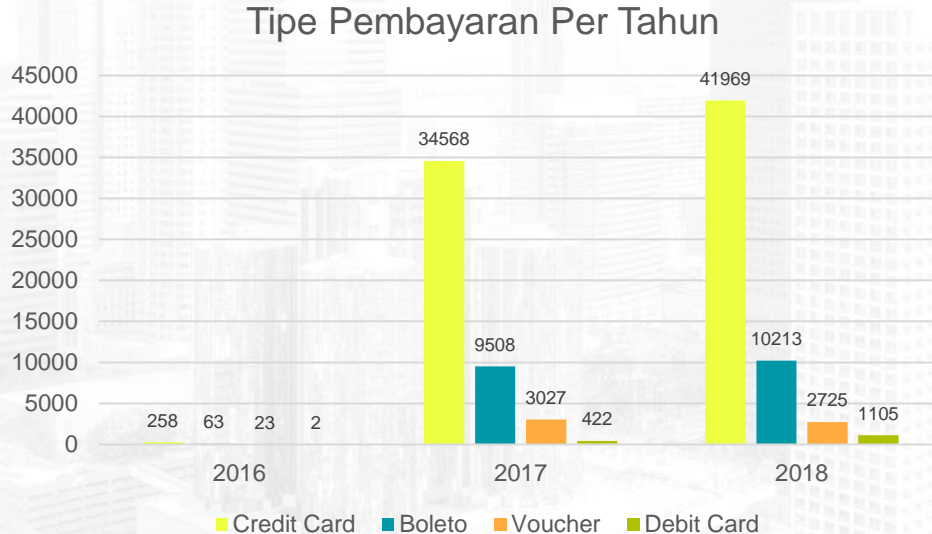
- Credit Card (Kartu Kredit) digunakan dalam 74% dari total transaksi. Ini menunjukkan bahwa sebagian besar pelanggan lebih memilih menggunakan kartu kredit sebagai metode pembayaran mereka.
- Boleto digunakan dalam 19% dari total transaksi. Boleto adalah metode pembayaran yang populer di Brasil, yang memungkinkan pembayaran melalui slip pembayaran yang dapat dibayar di bank atau outlet yang bermitra.
- Voucher digunakan dalam 1% dari total transaksi. Voucher mungkin merupakan metode pembayaran yang kurang umum, tetapi masih digunakan oleh sebagian pelanggan.
- Debit Card (Kartu Debit) digunakan dalam 6% dari total transaksi. Kartu debit adalah metode pembayaran lain yang cukup umum, meskipun kurang digunakan dibandingkan dengan kartu kredit.

Tipe Pembayaran



Credit card metode pembayaran yang paling diminati oleh pengguna

Jumlah pengguna masing-masing tipe pembayaran setiap tahun:



- Credit Card memiliki tren penggunaan yang meningkat dari tahun ke tahun, dengan jumlah transaksi yang signifikan pada tahun 2017 dan 2018.
- Boleto memiliki jumlah penggunaan tidak sebesar credit card, namun penggunaan boleto tetap cukup signifikan, khususnya di Brasil di mana metode ini umum digunakan.
- Voucher memiliki penggunaan yang relatif stabil atau menurun sedikit dari tahun 2016 hingga 2018.
- Debit Card mengalami peningkatan penggunaan dari tahun ke tahun, tetapi jumlahnya masih jauh lebih rendah dibandingkan dengan credit card dan boleto.

Credit card metode pembayaran yang paling diminati oleh pengguna dari tahun ke tahun

Kesimpulan:

1. Persentase tipe pembayaran mencerminkan preferensi pelanggan terhadap berbagai metode pembayaran yang tersedia, dengan credit card menjadi yang paling umum digunakan.
2. Pola persentase tipe pembayaran per tahun memberikan gambaran tentang preferensi pembayaran pelanggan dari waktu ke waktu, dengan credit card menjadi pilihan utama yang dominan, diikuti oleh boleto, kemudian voucher dan debit card yang memiliki penggunaan yang lebih rendah.