

|  |  |
| --- | --- |
| Report of Café Sales | |
| Client name | DeepSeek |
| ANALYst name | Ftrah Rahmi Putri |
| Date | 27-29 Aug 2025 |

## Executive summary

Berdasarkan dataset Penjualan Café, kami mengidentifikasi bahwa **Juice** merupakan item terlaris dengan total penjualan 6.435 unit sepanjang tahun 2023, dengan rata-rata penjualan harian dan bulanan masing-masing sebesar 69 dan 178 unit. Pelanggan paling sering menggunakan **digital wallet** sebagai metode pembayaran dan lebih memilih **takeaway** dibandingkan menikmati minumannya di tempat (*in-store*).

Distribusi penjualan sepanjang tahun 2023 cenderung stabil dengan beberapa kali pertumbuhan, menghasilkan total pendapatan (*revenue*) sebesar **$87,768**. Meskipun demikian, terdapat **2.845 issue** dalam dataset, di antaranya adalah transaksi dengan nilai $0, serta data yang tercatat sebagai *error* atau *unknown* pada transaksi maupun nama item.

## Background and Objectives

Analisis ini dilakukan untuk membantu client dalam:

* **Sales Performance:**
  + Mengidentifikasai item yang paling laris selama setehun.
  + Mengidentifikasi total dan rata-rata pendapatan (*revenue*) Café di tahun2023.
  + Mengidentifikasi tren penjualan harian dan bulanan.
* **Customer Behavior:**
  + Menentukan metode pembayaran yang paling disukai pelanggan.
  + Mengungkapkan kecenderungan pola konsumsi (takeaway vs. dine-in).
  + Memetakan permintaaan pesanan dari pelanggan.
* **Operational Insights:**
  + Menganalisis item yang sering bermasalah dalam penjualan
  + Memetakan preferensi lokasi konsumsi pelanggan.

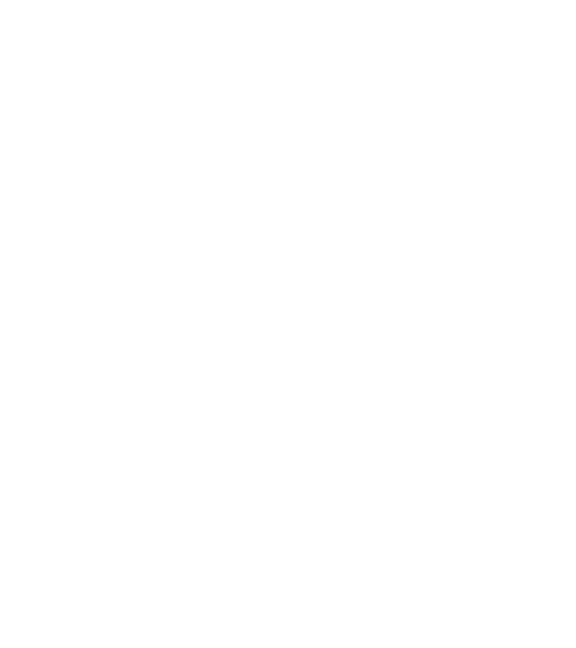
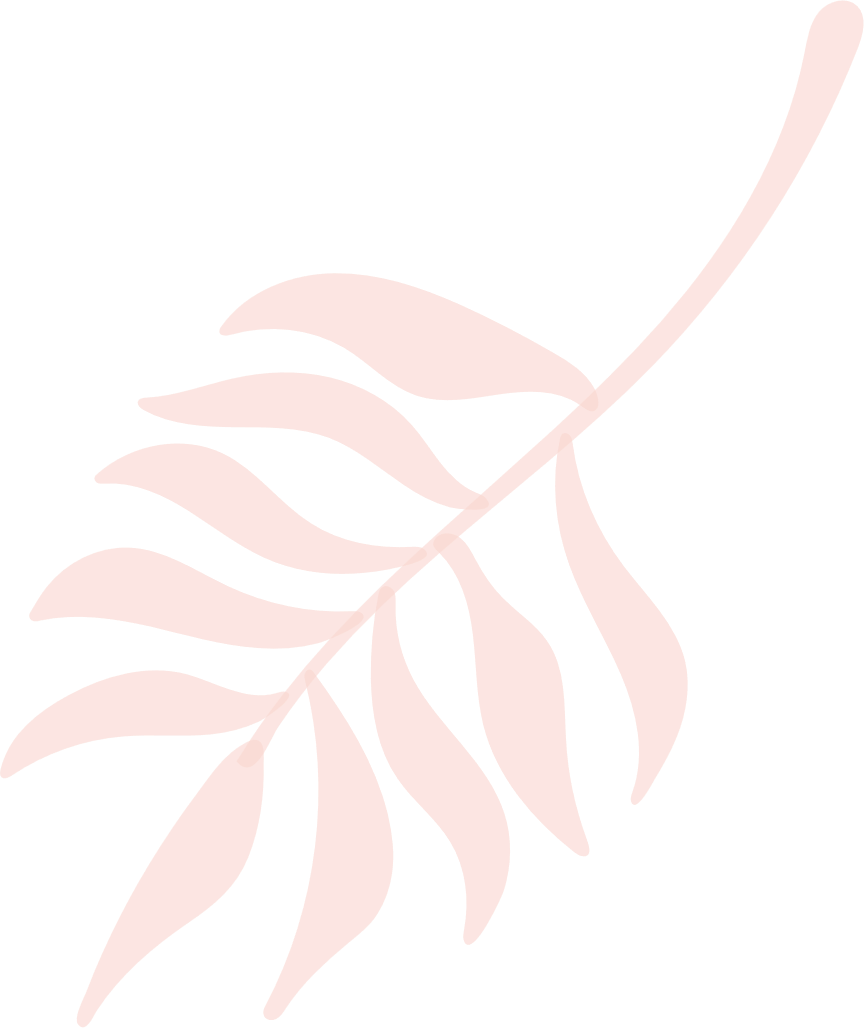
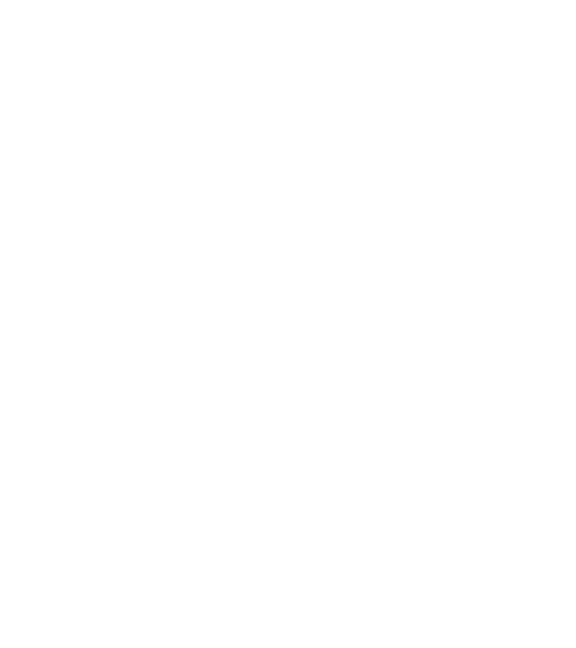
## Metodology

1. Pendahuluan

Tujuan proyek ini adalah menganalisis performa penjualan café selama tahun 2023 untuk mengidentifikasi:

* + Produk terlaris dan dan kinerja pendapatan.
  + Perilaku konsumen
  + Masalah operasional dalam transaksi.

1. Sumber Data
   * Sumber Data : Kaggle Dataset ([Cafe Sales - Dirty Data for Cleaning Training](https://www.kaggle.com/datasets/ahmedmohamed2003/cafe-sales-dirty-data-for-cleaning-training))
   * Periode Data : 1 Januari – 31 Desember 2023
   * Tata Kelola: Data telah dianomisasi dan diproses sesuai regulasi GDPR
2. Persiapan dan Pemrosesan Data
   * **Integrasi** : Mengimpor dan memvalidasi data mentah menggunakan SQL untuk memastikan struktur, tipe data, dan penanganan error sesuai kebutuhan analisis.
   * **Pembersihan**:
     + Penanganan Placeholder Error: diidentifikasi "**ERROR**" dan "**UNKNOWN**" dari berbagai kolom dan diubah menjadi nilai **Null** (2845 records).
     + Penanganan Missing values:
       - Kolom **quantitiy** dan **price\_unit**: diisi dengan ***median***
       - Kolom **total\_spend**: Diisi dengan hasil perkalian antara **quantity** dan **price\_unit**.
       - Kolom **transaction\_date**: Diisi dengan ***median***
       - Kolom Kategorical lainnya dengan ***mode*** (nilai paling sering muncul)
     + Outlier : Tidak dilakukan penanganan khusus karena nilai **total\_spend** merupakan hasil perkalian alami antara **quantity** dan **price\_unit**.
   * **Transformasi**
     + Mengubah tipe data **transaction\_date** dari teks (*object*) menjadi format datetime.
     + Membuat metrik baru: **trans\_yea**r, **trans\_month**, dan **trans\_date**.
   * **Teknik Statistik**
     + Descriptive Statistic (median, mode).
     + Analisis deret waktu (*time series*) untuk identifikasi tren.
     + Analisis frekuensi untuk preferensi konsumen.
   * **Segmentasi:** Segmentasi pelanggan berdasarkan perilaku pembelian.
   * **Alat**
     + SQL untuk data query
     + Python (Pandas, Numpy, Matplotlib, Seaborn, Scripy, mysql.connector)
     + Excel untuk validasi data

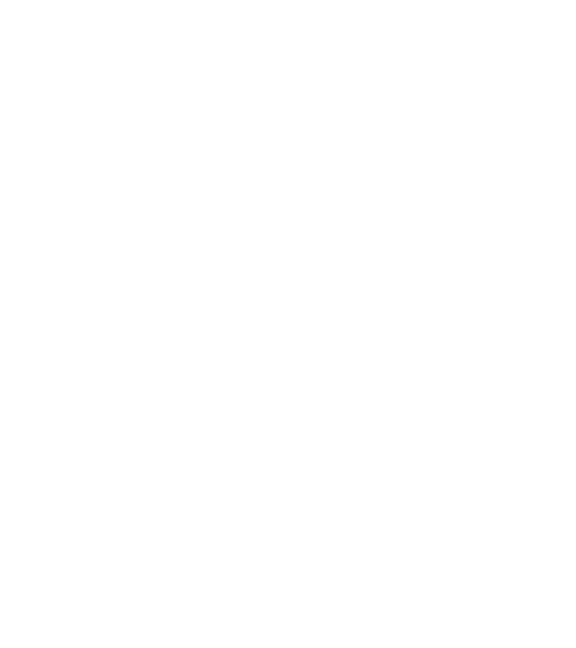
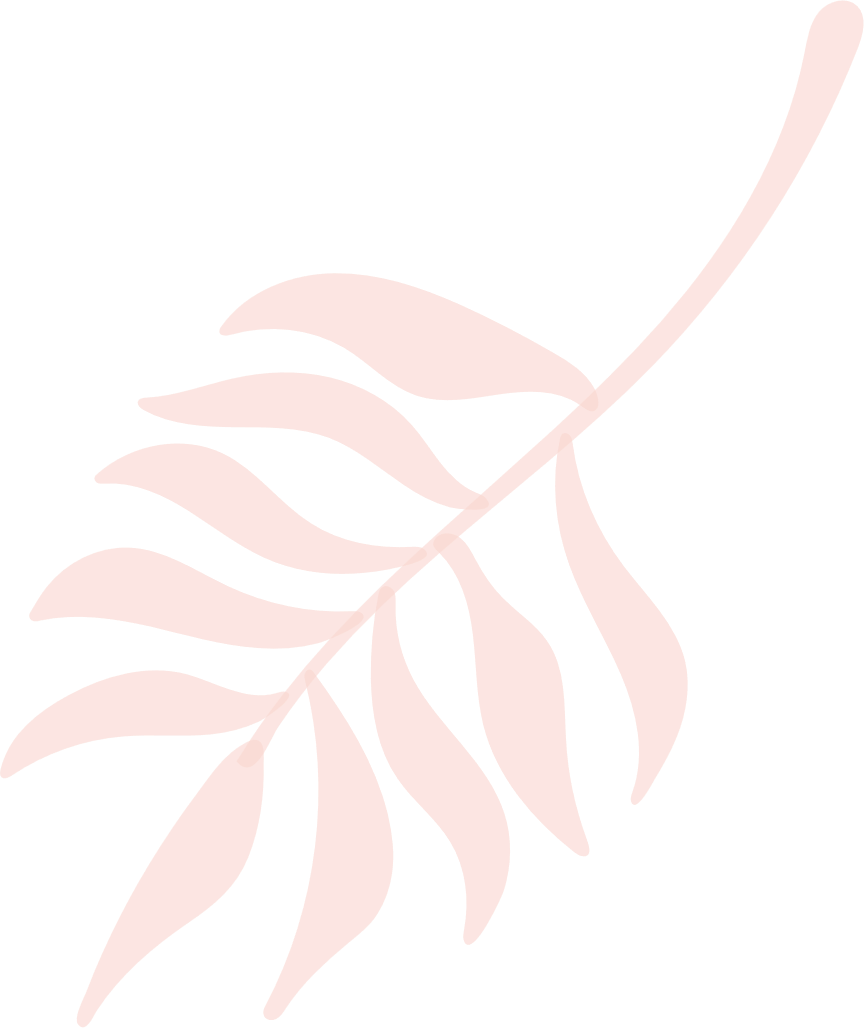
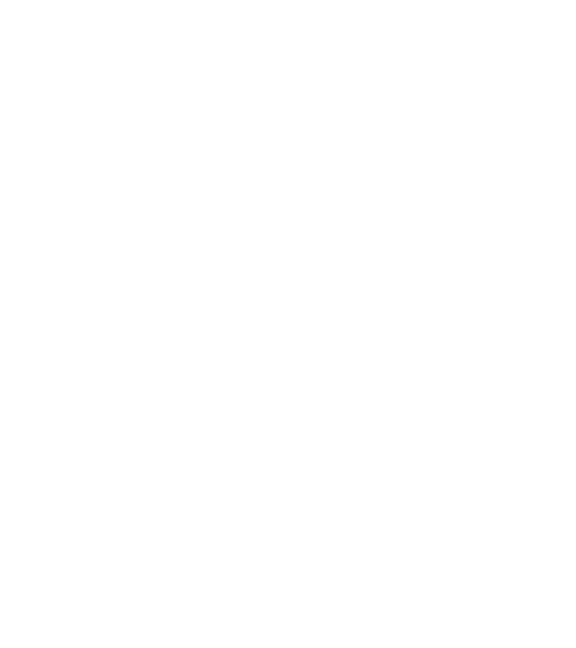


## Findings

1. **Produk Terlaris**: Juice adalah item paling populer (6.435 unit terjual).
2. **Preferensi Pembayaran**: Digital wallet dominan (54.69%), diikuti kartu Kredit (22.73%) dan tunai (22.58%).
3. **Pola Konsumsi**: 70% pelanggan memilih *take away*, 30% *dine-in*.
4. **Tren Penjualan**: Stabil sepanjang tahun dengan puncak pada pertengahan tahun.
5. **Issue Operasional**: Terdapat 2.845 transaksi bermasalah (data **Null**, **ERROR** dan **UNKNOWN**).

## Recommendation

1. **Optimasi Produk Unggulan**
   1. Tingkatkan stok dan promosi untuk produk **Juice.**
   2. Kembangkan varian baru berdasarkan preferensi konsumen.
2. **Enhanced Digital Experience**
   1. Tingkatkan fitur digital wallet untuk kemudahan transaksi.
   2. Implementasi loyalty program digital.
3. **Improve Operational Efficiency**
   1. Implementasi data validation system untuk mengurangi error.
   2. Training staff untuk input data yang akurat.
4. **Customer Experience Development**
   1. Optimasi layanan take away dengan packaging yang better.
   2. Penataan ulang space untuk dine-in experience.
5. **Continuous Monitoring**
   1. Develop dashboard real-time untuk monitoring performa.
   2. Setup regular monthly reporting system.



**Catatan**: Rekomendasi didasarkan pada data tahun 2023. Perlu evaluasi berkala untuk menyesuaikan dengan dinamika pasar terkini.