Rencana Tindakan yang Diusulkan

1. Penilaian & Pemahaman Data Awal

Tugas pertama saya adalah memuat data dan melakukan eksplorasi awal untuk memahami strukturnya serta mengidentifikasi masalah kualitas data yang langsung terlihat. Dari tinjauan cepat pada sampel, saya sudah melihat beberapa masalah umum yang perlu kita perbaiki:

• Nilai Hilang (Missing Values): Sel kosong (misalnya, Item, Payment Method, Location).

• Error Placeholder: Nilai seperti "ERROR" dan "UNKNOWN" pada berbagai kolom.

• Ketidakkonsistenan Tipe Data: Field numerik (Quantity, Price Per Unit, Total Spent) yang berisi kesalahan teks, serta tanggal yang tidak valid.

• Validasi: Memastikan Total Spent konsisten secara matematis dengan Quantity × Price Per Unit.

2. Pembersihan Data (The "Dirty" in the filename)

Saya akan membuat versi dataset yang sudah dibersihkan. Ini akan melibatkan:

• Mengganti nilai "ERROR", "UNKNOWN", dan sel kosong dengan nilai standar NaN (null) agar dapat ditangani dengan benar.

• Mengoreksi atau mengisi nilai hilang jika memungkinkan dan logis (misalnya, mungkin kita tidak bisa menebak Item yang hilang, tetapi kita terkadang bisa menghitung Total Spent yang hilang).

• Menstandarkan kategori pada kolom Payment Method dan Location.

• Memperbaiki nilai Transaction Date yang tidak valid.

• Memvalidasi dan menghitung ulang kolom Total Spent.

3. Analisis Data Eksploratif (EDA)

Setelah data bersih, saya dapat melakukan analisis untuk menjawab pertanyaan-pertanyaan bisnis utama. Saya akan fokus menyediakan visualisasi dan ringkasan yang jelas. Bisakah Anda jelaskan aspek mana yang paling penting bagi Anda? Ini akan membantu saya memprioritaskan analisis.

• Kinerja Penjualan:

- Item apa yang paling laris?

- Berapa total pendapatan dan nilai transaksi rata-rata?

- Bagaimana tren penjualan dari waktu ke waktu (harian, bulanan)?

• Perilaku Pelanggan:

- Metode pembayaran apa yang paling disukai?

- Apakah pelanggan lebih suka makan di tempat (In-store) atau dibawa pulang (Takeaway)?

- Bagaimana distribusi jumlah pesanan?

• Wawasan Operasional:

- Apakah ada item tertentu yang sering memiliki masalah data (mengindikasikan masalah di point-of-sale)?

- Bagaimana distribusi penjualan di berbagai lokasi? (Meskipun Location tampaknya hanya memiliki dua nilai).

4. Hasil Akhir (Delivery)

1. **Script Python** (dalam format Jupyter Notebook atau .py file) yang berisi:
   * Kode untuk data cleaning dan preprocessing.
   * Kode untuk melakukan filtering dan analisis.
   * Comment yang jelas pada setiap langkah.
2. **Summary Report Singkat** (dalam format PDF atau Markdown) yang berisi:
   * Penjelasan singkat tentang proses cleaning yang dilakukan.
   * Jawaban dari semua pertanyaan analisis di atas dalam bentuk tulisan dan visualisasi sederhana (e.g., plot sebaran rating, tabel frekuensi negara).
   * Kesimpulan dan rekomendasi .

Tugas pertama saya adalah membersihkan dataset ini secara menyeluruh. Saya akan segera memulainya sambil menunggu masukan Anda tentang area fokus utama dalam analisis.

Saya sangat antusias untuk membantu Anda mendapatkan wawasan berharga dari data penjualan kafe Anda. Mohon beri tahu saya pertanyaan prioritas Anda.

Hormat saya,

Fitrah Rahmi Putri