

# **LAPORAN TUGAS BESAR**

## **DATA WAREHOUSE DAN BUSINESS INTELLIGENCE**

Dosen Pengampu: Arif Rahman Hakim S.T.,M.T,

**Kelompok 7**



**Telkom**  
University

Disusun Oleh:

Prana Ichlasul Kautsar	102042300116
Achmad Vickram Ritonga	102042300085
Gading Ainunzaky Widianto	102042300090
Nicodemus Benaya Gavia L. T.	102042300093
Fredy Arya Hutama	102042300099
M. Affan Febryan T.	102042300108

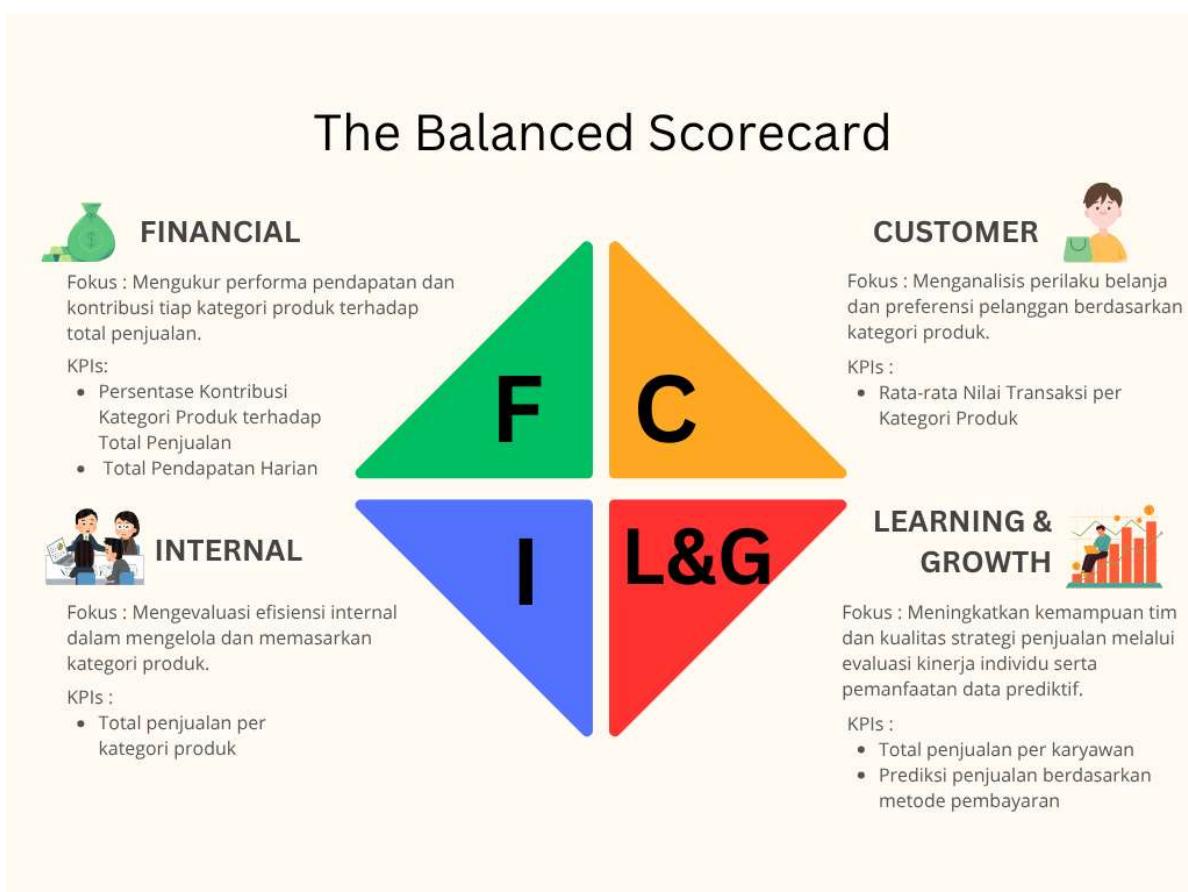
**PROGRAM STUDI SISTEM INFORMASI**  
**FAKULTAS REKAYA INDUSTRI**  
**UNIVERSITAS TELKOM KAMPUS JAKARTA**  
**2025**

## A. Perumusan Objective, Balanced Scorecard, dan KPI

Analisis ini memiliki beberapa objektif kunci yang saling terkait dan mendukung tujuan bisnis. Pertama, **memprediksi keuntungan bisnis** menjadi fokus utama, yang melibatkan analisis data historis untuk mengidentifikasi faktor-faktor pendorong keuntungan. Kedua, terdapat objektif untuk **memprediksi jumlah penjualan dari setiap produk**, yang akan memberikan wawasan berharga untuk manajemen inventaris dan strategi pemasaran. Selanjutnya, perumusan objektif juga mencakup **analisis perilaku pelanggan**, termasuk pola pembelian, preferensi, dan segmentasi, yang krusial untuk strategi retensi dan akuisisi pelanggan. Terakhir, analisis ini bertujuan untuk **menganalisis kinerja karyawan dan kontribusinya terhadap penjualan**, memberikan dasar untuk evaluasi kinerja dan program insentif yang efektif. Secara keseluruhan, objektif ini dirancang untuk memberikan wawasan yang dapat ditindaklanjuti, membantu Faeyza Printing dalam pengambilan keputusan strategis untuk peningkatan profitabilitas dan efisiensi operasional.

### Objective

Mengoptimalkan pendapatan dan efisiensi operasional melalui analisis kontribusi produk, preferensi pelanggan, dan performa penjualan sebagai dasar pengambilan keputusan strategis



Goals	KPI	Data	Perhitungan	Chart
Memastikan setiap kategori produk mencapai minimal target penjualan bulanan untuk menjaga kestabilan finansial usaha.	<p>Total penjualan per kategori produk</p> <p>TARGET:</p> <ul style="list-style-type: none"> <li>• Alat Tulis: <math>\geq</math> Rp10.000.000 per bulan</li> <li>• Perlengkapan: <math>\geq</math> Rp7.500.000 per bulan</li> <li>• Layanan: <math>\geq</math> Rp5.000.000 per bulan</li> <li>• Alat kantor: <math>\geq</math> Rp. 2.000.000</li> <li>• peralatan digital: <math>\geq</math> Rp 500.000</li> <li>• Peralatan cetak: <math>\geq</math> Rp 6.000.000</li> </ul>	Diperoleh dari table fakta penjualan → pada atribut kategori dan total_harga	SUM(total_harga)	Bar Chart
Mengukur produktivitas individu karyawan dan mendorong kontribusi optimal terhadap performa usaha.	<p>Total penjualan per karyawan</p> <p>TARGET: Minimal Rp2.500.000 per karyawan per bulan</p>	Diperoleh dari table fakta penjualan → pada atribut total_harga dan nama_karyawan	SUM(total_harga)	Bar Chart
Mengembangkan strategi penjualan berbasis tren metode pembayaran agar dapat diarahkan ke channel yang paling potensial.	<p>prediksi penjualan berdasarkan metode pembayaran</p> <p>Target:</p> <ul style="list-style-type: none"> <li>• QRIS: naik 10% dibanding bulan sebelumnya</li> <li>• Tunai: stabil atau turun <math>\leq</math> 5%</li> <li>• Transfer/Bank: naik <math>\geq</math> 15%</li> <li>• debit: <math>\leq</math> 35%</li> </ul>	Diperoleh dari table fakta penjualan → pada atribut total_harga dan metode_pembayaran	<p>Actual: SUM(total_harga)</p> <p>Targer: <math>((\text{bulan\_ini} - \text{bulan\_lalu}) / \text{bulan\_lalu}) \times 100\%</math></p>	Donut Chart
Menjaga kestabilan arus kas harian agar	<p>Total Pendapatan Harian</p> <p>TARGET:</p>	Diperoleh dari table fakta penjualan → pada atribut	SUM(total_harga) Time Series	Time Series

operasional berjalan lancar tanpa kendala keuangan jangka pendek	<ul style="list-style-type: none"> <li>Rata-rata harian minimal Rp 500.000</li> </ul>	total_harga dan tanggal		
Meningkatkan nilai transaksi rata-rata per kategori agar menghasilkan pendapatan lebih besar dari jumlah transaksi yang stabil.	<p>Rata-rata Nilai Transaksi per Kategori Produk</p> <p>TARGET:</p> <ul style="list-style-type: none"> <li>Alat Tulis: = Rp 100.000</li> <li>Perlengkapan: = Rp 75.000</li> <li>Layanan: = Rp 200.000</li> <li>Alat kantor: = Rp. 100.000</li> <li>peralatan digital: <math>\geq</math> Rp 60.000</li> <li>peralatan ccetak: <math>\geq</math> Rp 80.000</li> </ul>	Diperoleh dari table fakta penjualan → pada atribut total_harga dan kategori	AVG(total_harga)	Bar Chart
Menyeimbangkan strategi promosi dan stok berdasarkan kontribusi kategori terhadap total penjualan	<p>Percentase Kontribusi Kategori Produk terhadap Total Penjualan</p> <p>TARGET :</p> <ul style="list-style-type: none"> <li>Alat Tulis: = 25%</li> <li>Perlengkapan: = 10% - 15%</li> <li>Layanan: = Rp 205% - 35%</li> <li>Alat kantor: = 20%</li> <li>Peratalan digital: = 10% - 13%</li> <li>peralatan ccetak: = 10%</li> </ul>	Diperoleh dari table fakta penjualan → pada atribut total_harga dan kategori	SUM(total_harga)	Donut chart

## B. Analisis Sumber Data

Analisis ini didasarkan pada dataset komprehensif yang diperoleh dari Zildan Chesta, putra pemilik Faeyza Printing, sebuah bisnis fotokopi dan percetakan yang telah beroperasi selama lima tahun di Tangerang. Data ini, meskipun telah melalui proses modifikasi dan rekayasa minor untuk mempermudah analisis dan memperkaya eksplorasi, secara representatif mencerminkan operasional inti bisnis tersebut. Sumber data utama terbagi menjadi empat dimensi kunci, yaitu

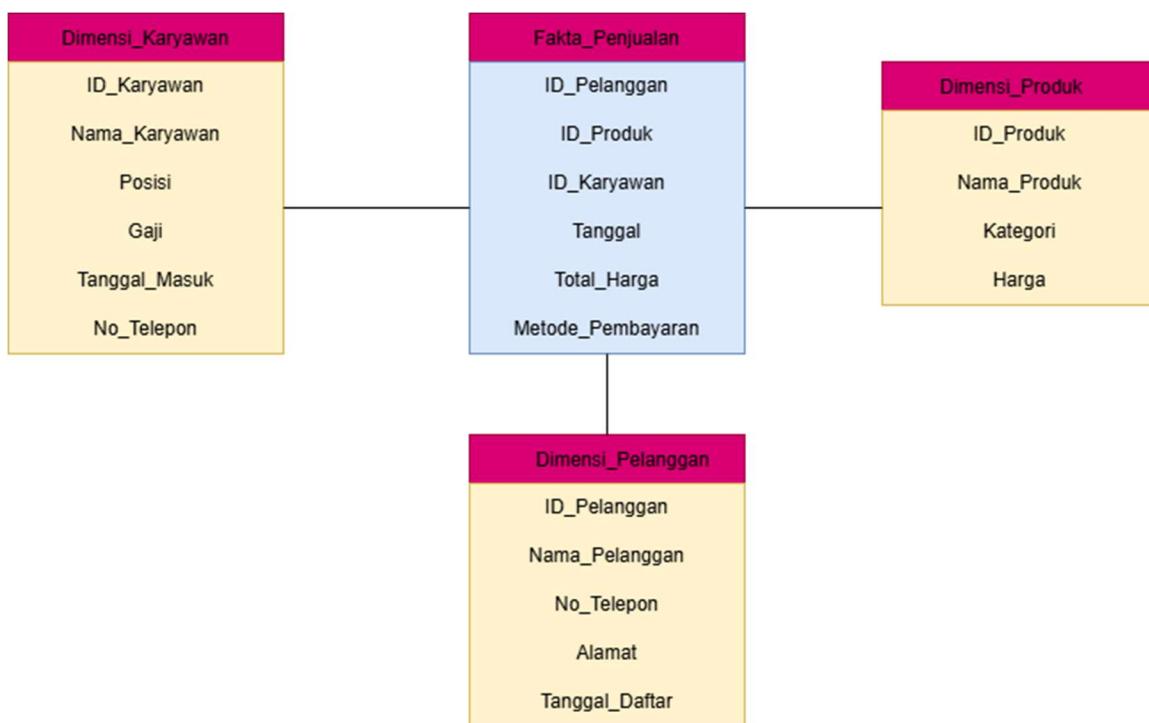
Dimens\_Penjualan, Dimens\_Karyawan, Dimens\_Pelanggan, dan Dimens\_Produk. Data Dimens\_Penjualan menjadi pondasi utama analisis, memuat catatan transaksi yang menjadi cerminan kinerja finansial dan operasional. Sementara itu, Dimens\_Karyawan memberikan detail mengenai sumber daya manusia, Dimens\_Pelanggan menyajikan profil dan riwayat transaksi konsumen, dan Dimens\_Produk menguraikan inventaris layanan dan barang yang ditawarkan Faeyza Printing. Ketersediaan empat dimensi ini memungkinkan analisis holistik terhadap berbagai aspek bisnis, mulai dari efisiensi operasional hingga pola perilaku konsumen.



Sebelum melangkah pada analisis yang lebih mendalam, tahap Exploratory Data Analysis (EDA) telah dilakukan secara sistematis untuk memahami karakteristik intrinsik data, mengidentifikasi potensi anomali, serta menggali pola awal yang mungkin tersembunyi. Proses EDA dimulai dengan pemeriksaan struktur data pada setiap dimensi, memastikan konsistensi tipe data (numerik, kategorikal, tanggal) dan mengidentifikasi potensi kebutuhan transformasi. Selanjutnya, statistik deskriptif dihitung untuk variabel numerik, seperti rata-rata, median, modus, standar deviasi, serta nilai minimum dan maksimum, guna memahami sebaran dan tendensi sentral data. Untuk variabel kategorikal, analisis frekuensi dilakukan untuk mengidentifikasi dominasi kategori tertentu. Salah satu fokus utama EDA adalah identifikasi dan penanganan nilai hilang (missing values) serta duplikasi data. Kehadiran nilai hilang dapat mengindikasikan masalah pada proses pengumpulan data dan berpotensi bias dalam analisis, sementara duplikasi dapat menyebabkan perhitungan yang tidak akurat.

Melalui proses EDA ini, beberapa observasi awal telah teridentifikasi. Potensi masalah kualitas data, seperti ketidakseragaman format pada kolom-kolom kunci yang akan digunakan untuk penggabungan antar tabel (misalnya, ID Pelanggan atau ID Produk), menjadi perhatian utama yang memerlukan standardisasi lebih lanjut. Dalam konteks penjualan, terlihat adanya indikasi pola musiman atau tren tertentu yang akan menjadi subjek analisis deret waktu di tahap berikutnya untuk memahami fluktuasi penjualan Faeyza Printing. Selain itu, data pelanggan menunjukkan potensi untuk segmentasi berdasarkan frekuensi atau nilai pembelian, yang dapat menginformasikan strategi pemasaran yang lebih bertarget. Demikian pula, analisis produk awal dapat menyoroti produk atau layanan unggulan, serta yang kurang diminati, memberikan wawasan untuk optimalisasi portofolio produk. Secara keseluruhan, EDA ini tidak hanya berfungsi sebagai validasi awal terhadap kualitas data, tetapi juga sebagai peta jalan untuk eksplorasi lebih lanjut, menyoroti area-area yang menjanjikan untuk analisis mendalam dan pengembangan wawasan bisnis bagi Faeyza Printing.

### C. Perancangan Star Schema



1. Tabel Fakta: Fakta\_Penjualan Tabel fakta merupakan pusat dari skema bintang ini yang menyimpan data transaksional dan numerik yang akan dianalisis. Tabel ini mencatat setiap transaksi penjualan yang terjadi.

Atribut:

- ID\_Pelanggan (FK) – relasi ke Dimensi\_Pelanggan
  - ID\_Prodak (FK) – relasi ke Dimensi\_Prodak

- ID\_Karyawan (FK) – relasi ke Dimensi\_Karyawan
- Tanggal – waktu terjadinya transaksi
- Total\_Harga – nilai total transaksi (measure)
- Metode\_Pembayaran – jenis pembayaran (tunai, kartu, dll.)

Measure utama dalam tabel fakta ini adalah Total\_Harga, yang dapat digunakan untuk menghitung penjualan total, rata-rata penjualan per produk, dan performa penjualan karyawan.

2. Dimensi\_Pelanggan Menyimpan informasi detail mengenai pelanggan yang melakukan transaksi.

Atribut:

- ID\_Pelanggan (PK)
- Nama\_Pelanggan
- No\_Telepon
- Alamat
- Tanggal\_Daftar

3. Dimensi\_Produk Menyimpan data produk yang dijual dalam transaksi.

Atribut:

- ID\_Prodak (PK)
- Nama\_Prodak
- Kategori
- Harga

4. Dimensi\_Karyawan Berisi informasi tentang karyawan yang melayani atau memproses transaksi penjualan.

Atribut:

- ID\_Karyawan (PK)
- Nama\_Karyawan
- Posisi
- Gaji
- Tanggal\_Masuk
- No\_Telepon

## D. Implementasi Star Schema dalam RDBMS

Setelah perancangan model star schema, tahap selanjutnya adalah mengimplementasikannya ke dalam sistem manajemen basis data relasional (RDBMS). Implementasi ini dilakukan menggunakan perintah Data Definition Language (DDL) untuk membuat struktur tabel dalam database yang akan menyimpan data warehouse.

### 1. Tabel Dimensi

Tabel dimensi menyimpan atribut deskriptif yang mendukung analisis. Masing-masing tabel memiliki primary key sebagai pengenal unik yang akan digunakan sebagai foreign key dalam tabel fakta.

```
CREATE TABLE Dimensi_Pelanggan (
    Id_Pelanggan VARCHAR(10) PRIMARY KEY,
    Nama_Pelanggan VARCHAR(100),
    No_Telepon VARCHAR(20),
    Alamat TEXT,
    Tanggal_Daftar DATE
);
```

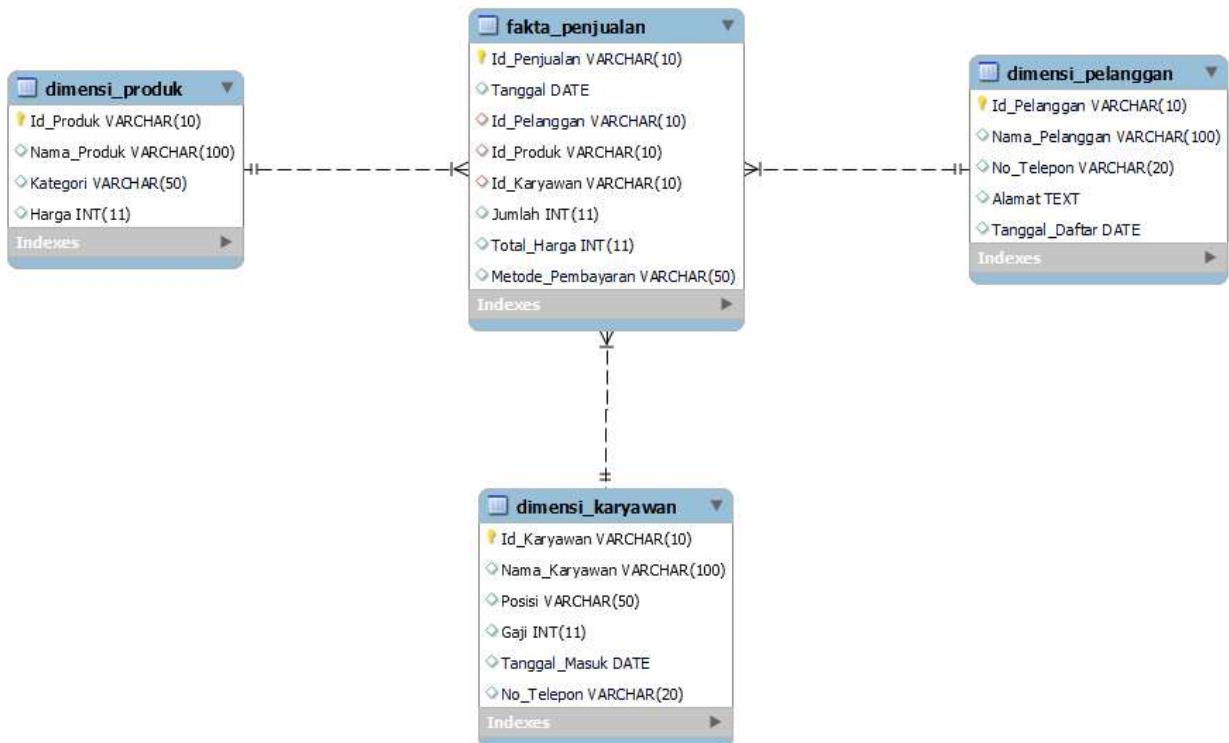
```
CREATE TABLE Dimensi_Produk (
    Id_Produk VARCHAR(10) PRIMARY KEY,
    Nama_Produk VARCHAR(100),
    Kategori VARCHAR(50),
    Harga INT
);
```

```
CREATE TABLE Dimensi_Karyawan (
    Id_Karyawan VARCHAR(10) PRIMARY KEY,
    Nama_Karyawan VARCHAR(100),
    Posisi VARCHAR(50),
    Gaji INT,
    Tanggal_Masuk DATE,
    No_Telepon VARCHAR(20)
);
```

## 2. Tabel Fakta

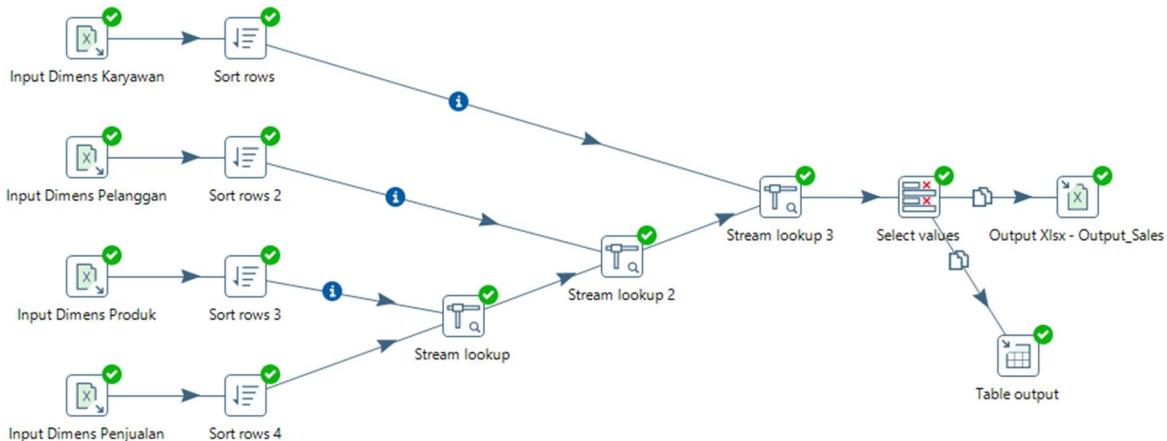
Tabel fakta menyimpan data transaksi penjualan yang menjadi pusat analisis. Di dalamnya terdapat foreign key yang terhubung ke masing-masing tabel dimensi, serta kolom measure seperti Jumlah dan Total\_Harga yang digunakan untuk agregasi analitik.

```
CREATE TABLE Fakta_Penjualan (
    Id_Penjualan VARCHAR(10) PRIMARY KEY,
    Tanggal DATE,
    Id_Pelanggan VARCHAR(10),
    Id_Produk VARCHAR(10),
    Id_Karyawan VARCHAR(10),
    Jumlah INT,
    Total_Harga INT,
    Metode_Pembayaran VARCHAR(50),
    FOREIGN KEY (Id_Pelanggan) REFERENCES Dimensi_Pelanggan(Id_Pelanggan),
    FOREIGN KEY (Id_Produk) REFERENCES Dimensi_Produk(Id_Produk),
    FOREIGN KEY (Id_Karyawan) REFERENCES Dimensi_Karyawan(Id_Karyawan)
);
```



## E. Implementasi Proses ETL

Setelah fase eksplorasi data awal yang telah mengidentifikasi karakteristik dan potensi pola pada dataset Faeyza Printing, langkah krusial berikutnya dalam analisis ini adalah implementasi proses *Extract, Transform, and Load* (ETL). Untuk menjalankan proses ETL ini secara efisien dan terotomatisasi, digunakan **Pentaho Data Integration (PDI)**, yang juga dikenal sebagai Kettle. PDI adalah perangkat lunak ETL *open-source* yang menyediakan lingkungan grafis intuitif untuk merancang alur kerja data tanpa perlu penulisan kode yang ekstensif. Penggunaan PDI memungkinkan penggabungan data dari berbagai sumber, pembersihan, dan transformasi data secara sistematis, guna mengintegrasikan dan mempersiapkan data dari berbagai dimensi untuk analisis yang lebih mendalam. Proses ETL ini dirancang untuk mengkonsolidasikan data penjualan dengan informasi karyawan, pelanggan, dan produk, menghasilkan sebuah tabel fakta terpadu yang kaya akan konteks bisnis.



Tahap **Ekstraksi** dimulai dengan mengimpor masing-masing file dimensi, yaitu Dimens\_Karyawan.xlsx, Dimens\_Pelanggan.xlsx, Dimens\_Produk.xlsx, dan Dimens\_Penjualan.xlsx, ke dalam lingkungan PDI menggunakan langkah *Microsoft Excel Input*. Setiap *dataset* diekstraksi secara independen, mempertahankan integritas data sumbernya. Setelah ekstraksi, tahap **Transformasi** menjadi inti dari proses ini, di mana data mentah diolah dan diperkaya. Langkah pertama dalam transformasi adalah pengurutan baris (*Sort Rows*) pada masing-masing dimensi. Pengurutan ini penting untuk mengoptimalkan kinerja operasi pencarian (*lookup*) dan memastikan konsistensi data saat proses penggabungan, terutama dalam skenario *stream lookup*.

Kemudian, inti dari transformasi data adalah penggunaan teknik *lookup* untuk menggabungkan dan melengkapi data dari tabel dimensi ke dalam tabel fakta Dimens\_Penjualan. Proses ini dimulai dengan memperkaya tabel fakta Dimens\_Penjualan menggunakan Dimens\_Produk. Metode *Stream Lookup* diterapkan, di mana *step* Dimens\_Penjualan berperan sebagai *stream* utama dan *step* Dimens\_Produk sebagai *lookup*.

*stream*. Pencocokan dilakukan berdasarkan ID Produk yang tersedia di kedua tabel. Dari Dimens\_Produk, atribut-atribut relevan seperti Nama Produk, Kategori Produk, dan Harga Satuan diekstraksi dan ditambahkan ke setiap baris transaksi yang sesuai dalam Dimens\_Penjualan. Pendekatan serupa kemudian diterapkan secara iteratif untuk dimensi lainnya. Dimens\_Pelanggan dihubungkan ke tabel fakta menggunakan ID Pelanggan untuk menarik informasi seperti Nama Pelanggan, Tipe Pelanggan, dan Kota. Demikian pula, Dimens\_Karyawan diintegrasikan melalui ID Karyawan untuk menambahkan detail Nama Karyawan dan Jabatan, sehingga setiap transaksi penjualan tidak hanya mencakup detail produk, tetapi juga informasi lengkap mengenai pelanggan yang membeli dan karyawan yang melayani.

Setelah semua dimensi berhasil digabungkan dan melengkapi tabel fakta, tahap selanjutnya adalah penyaringan kolom menggunakan operasi *Select Values*. Pada tahap ini, hanya kolom-kolom yang relevan dan esensial untuk analisis lebih lanjut yang dipilih dari keseluruhan gabungan tabel. Pemilihan kolom ini didasarkan pada kebutuhan analisis yang telah diidentifikasi pada fase EDA dan relevansinya terhadap tujuan bisnis Faeyza Printing, sehingga menghindari redundansi data dan meningkatkan efisiensi penyimpanan serta pemrosesan. Kolom-kolom yang tidak diperlukan dibuang untuk menciptakan tabel keluaran yang ringkas dan fokus. Tahap terakhir dalam proses ETL adalah **Pemuatan (Load)**. Data yang telah diintegrasikan dan ditransformasi kemudian diekspor menjadi file Excel tunggal dengan nama Output\_Sales.xlsx menggunakan langkah *Microsoft Excel Writer* dalam PDI. File ini berfungsi sebagai *dataset* bersih dan siap pakai untuk analisis data lanjutan, *data mining*, dan pelaporan, merepresentasikan representasi data yang komprehensif dan terkonsolidasi dari seluruh operasional penjualan Faeyza Printing. Kita juga menambahkan Alur Output kedua yaitu database SQL pada localhost yang diintegrasikan dengan PDI, dengan secara otomatis membuat table untuk output\_sales pada database yang terkoneksi.

## F. Implementasi Data Mining

### 1. Penerapan Regresi: Prediksi Total Harga Penjualan

Tujuan dari penerapan ini adalah untuk membangun model yang dapat memprediksi Total\_Harga (nilai penjualan). Untuk memastikan akurasi, model ini dibangun dengan menggabungkan data transaksi penjualan dengan data detail produk dari dataset terpisah.

#### A. Fungsi Kode

Kode ini dimulai dengan menggabungkan dua sumber data sebelum membangun model regresi.

#### 1. Penggabungan Data:

- a. pd.read\_excel("dimens\_produk.xlsx"): Memuat dataset eksternal yang berisi informasi detail produk, termasuk **Harga** (harga satuan) dan **Kategori**.

```
import pandas as pd

# Muat dataset utama |
df_penjualan = pd.read_excel("dimens_penjualan.xlsx")

# Muat dataset dimensi produk
try:
    df_produk = pd.read_excel("dimens_produk.xlsx")

    if 'Id_produk' in df_produk.columns:
        df_produk.rename(columns={'Id_produk': 'Id_Produk'}, inplace=True)

    print("Dataset Dimensi Produk berhasil dimuat:")
    display(df_produk.head())

except FileNotFoundError:
    print("File 'dimens_produk.xlsx' tidak ditemukan.")
```

- b. pd.merge(df\_penjualan, df\_produk, on='Id\_Produk', how='left'): Menggabungkan data transaksi penjualan dengan data produk menggunakan Id\_Produk sebagai kunci. Metode left merge memastikan semua data penjualan tetap ada, sementara informasi produk yang relevan ditambahkan ke setiap baris transaksi.

```

# Gabungkan df_penjualan dengan df_produk
df_merged = pd.merge(df_penjualan, df_produk[['Id_Produk', 'Kategori', 'Harga']], on='Id_Produk', how='left')

# Cek hasil penggabungan
print("Dataset setelah digabungkan:")
display(df_merged.head())

# Cek jika ada produk yang tidak memiliki data harga setelah merge
missing_prices = df_merged[df_merged['Harga'].isnull()]
if not missing_prices.empty:
    print("\nPeringatan: Ditemukan transaksi dengan Id_Produk yang tidak ada di dimens_produk:")
    display(missing_prices)

```

## 2. Persiapan Fitur dan Target:

- fitur\_baru = ['Jumlah', 'Metode\_Pembayaran', 'Hari', 'Harga', 'Kategori']: Daftar fitur yang digunakan untuk prediksi mencakup Jumlah barang, Metode\_Pembayaran, Hari, **Harga** (harga satuan), dan **Kategori**.

```

# Pastikan tidak ada nilai kosong di kolom 'Harga' sebelum melatih model
# hapus barisnya jika ada
df_merged.dropna(subset=['Harga'], inplace=True)

# Membuat fitur baru
df_merged['Tanggal'] = pd.to_datetime(df_merged['Tanggal'])
df_merged['Hari'] = df_merged['Tanggal'].dt.day_name()

```

- $X = df\_merged[fitur\_baru]$ : Mendefinisikan matriks fitur ( $X$ ) yang akan digunakan.

```

# 1. Pilih Fitur dan Target baru
# 'Harga' (harga satuan) dan 'Kategori' sebagai fitur baru
fitur_baru = ['Jumlah', 'Metode_Pembayaran', 'Hari', 'Harga', 'Kategori']
X = df_merged[fitur_baru]
y = df_merged['Total_Harga']

```

- $y = df\_merged['Total_Harga']$ : Menetapkan Total\_Harga sebagai variabel target ( $y$ ) yang akan diprediksi.

## 3. Preprocessing dan Pemodelan:

- ColumnTransformer: Sebuah pipeline preprocessing diterapkan untuk menangani tipe data yang berbeda.

```
# 3. Pipeline Preprocessing |
numerical_features = ['Jumlah', 'Harga']
categorical_features = ['Metode_Pembayaran', 'Hari', 'Kategori']

preprocessor = ColumnTransformer(
    transformers=[
        ('num', StandardScaler(), numerical_features),
        ('cat', OneHotEncoder(handle_unknown='ignore'), categorical_features)
    ])
```

- b. StandardScaler diterapkan pada fitur numerik (Jumlah dan Harga) untuk menyamakan skalanya.
  - i. OneHotEncoder diterapkan pada fitur kategorikal (Metode\_Pembayaran, Hari, dan Kategori) untuk mengubahnya menjadi format numerik.
- c. RandomForestRegressor: Model *ensemble learning* ini dipilih untuk melakukan tugas prediksi karena kemampuannya menangani hubungan yang kompleks.

## B. Hasil dan Insight yang Diperoleh

### C. Hasil Evaluasi:

```
\nMelatih model dengan fitur yang diperkaya...
Model selesai dilatih.
\n--- Hasil Evaluasi Model ---
R2 Score: 0.4378
RMSE: Rp 245,479.82
```

**R2 Score: 0.4378:** Hasil evaluasi model menunjukkan R2 Score sebesar 0.4378. Ini berarti model mampu menjelaskan sekitar **43.8%** dari variabilitas Total\_Harga.

- a. **RMSE: Rp 245,479.82:** Root Mean Squared Error (RMSE) atau rata-rata kesalahan prediksi model adalah sekitar Rp 245,479.

### 1. Insight Utama:

- a. **Faktor Penentu Harga:** Analisis menunjukkan bahwa **harga satuan** dan **kategori produk** merupakan faktor prediktor yang signifikan

untuk Total\_Harga. Model dapat belajar dari fitur-fitur ini untuk membuat estimasi yang cukup baik.

- b. **Potensi Peningkatan:** R<sup>2</sup> Score sebesar 0.4378 mengindikasikan bahwa model sudah cukup baik, namun masih ada ruang untuk peningkatan. Sekitar 56% variasi harga belum bisa dijelaskan, yang kemungkinan disebabkan oleh faktor lain seperti **promosi/diskon** atau **karakteristik pelanggan** yang belum dimasukkan ke dalam model.

#### D. Rekomendasi Implementasi

Berdasarkan performa model, berikut adalah rekomendasi yang dapat dipertimbangkan:

1. **Perencanaan Inventaris dan Keuangan:** Gunakan model ini sebagai alat bantu untuk membuat **proyeksi pendapatan** berdasarkan perkiraan volume penjualan untuk setiap kategori produk, yang dapat membantu dalam manajemen arus kas.
2. **Analisis Penetapan Harga (Pricing):** Manfaatkan model untuk melakukan analisis skenario "what-if". Misalnya, simulasi dampak kenaikan harga pada produk tertentu terhadap total pendapatan.
3. **Pengembangan Model Lanjutan:** Untuk meningkatkan akurasi di masa depan, pertimbangkan untuk menambahkan fitur baru seperti **data diskon**, **informasi demografis pelanggan**, atau melakukan *hyperparameter tuning* pada model.

## 2. Penerapan Clustering: Segmentasi Pelanggan

Tujuan dari penerapan ini adalah untuk mengelompokkan pelanggan ke dalam segmen-semen yang berbeda berdasarkan perilaku pembelian mereka, tanpa adanya label sebelumnya.

#### A. Fungsi Kode

##### 2. Agregasi Data:

- a. df.groupby('Id\_Pelanggan').agg(...): Mengubah data dari level transaksi menjadi level pelanggan. Untuk setiap pelanggan, kode ini menghitung tiga metrik utama:

- i. Jumlah: Total barang yang dibeli.
- ii. Total\_Harga: Total uang yang dibelanjakan (Monetary).
- iii. Frekuensi: Jumlah transaksi yang dilakukan (Frequency). Ini adalah bentuk sederhana dari analisis RFM (*Recency, Frequency, Monetary*).

```

customer_df = df.groupby('Id_Pelanggan').agg(
    Jumlah=('Jumlah', 'sum'),
    Total_Harga=('Total_Harga', 'sum'),
    Frekuensi=('Id_Penjualan', 'count')
).reset_index()

print("Data Perilaku Pelanggan:")
print(customer_df.head())

```

✓ 0.0s

### 3. Penentuan Jumlah Cluster Optimal:

- a. StandardScaler: Menyamakan skala dari ketiga fitur di atas agar tidak ada fitur yang mendominasi proses clustering karena perbedaan skala (misalnya, Total\_Harga yang bernilai ratusan ribu tidak mendominasi Frekuensi yang bernilai kecil).
- b. **Elbow Method:** Kode ini menghitung *inertia* (jarak total data poin ke pusat cluster terdekatnya) untuk jumlah cluster (k) dari 1 hingga 10. Plot dari metode ini menunjukkan "siku" (elbow) pada **k=3**, yang mengindikasikan bahwa 3 adalah jumlah cluster yang paling optimal untuk data ini.

```

# Pilih fitur numerik untuk clustering
features = customer_df[['Jumlah', 'Total_Harga', 'Frekuensi']]

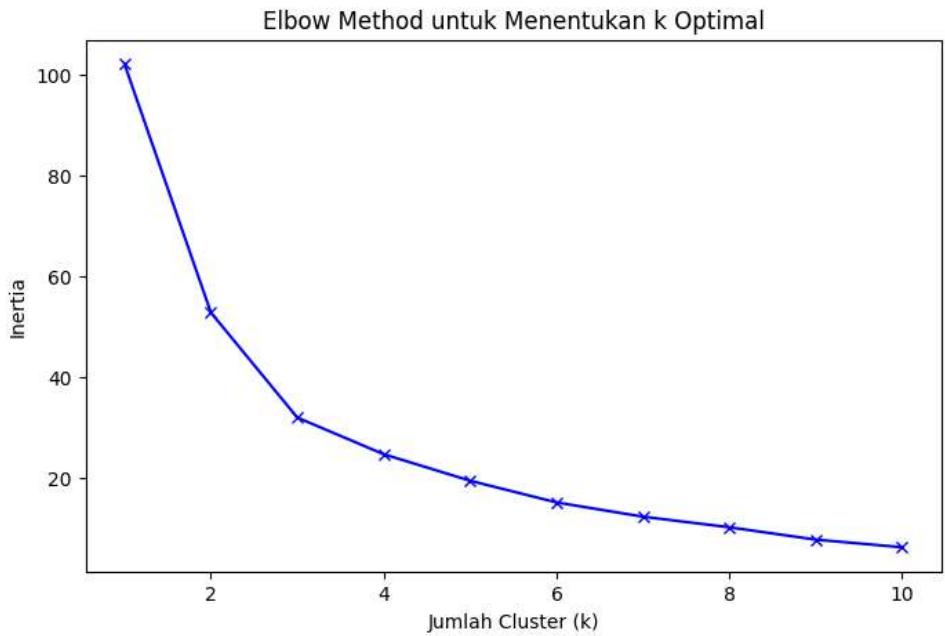
# Scaling fitur sangat penting untuk KMeans agar bobot setiap fitur seimbang
scaler = StandardScaler()
features_scaled = scaler.fit_transform(features)

# Menghitung inertia untuk berbagai nilai k (jumlah cluster)
inertia = []
K = range(1, 11) # Mencoba 1 sampai 10 cluster
for k in K:
    kmeans = KMeans(n_clusters=k, init='k-means++', n_init=10, random_state=42)
    kmeans.fit(features_scaled)
    inertia.append(kmeans.inertia_)

# Plot Elbow Method
plt.figure(figsize=(8, 5))
plt.plot(K, inertia, 'bx-')
plt.xlabel('Jumlah Cluster (k)')
plt.ylabel('Inertia')
plt.title('Elbow Method untuk Menentukan k Optimal')
plt.show()

```

✓ 0.2s



#### 4. Pemodelan dan Pelabelan:

- KMeans(n\_clusters=optimal\_k, ...): Menjalankan algoritma K-Means dengan k=3 yang sudah ditentukan.
- kmeans.fit\_predict(...): Melatih model pada data pelanggan dan memberikan label cluster (0, 1, atau 2) untuk setiap pelanggan.

```

# Menjalankan KMeans dengan k optimal
kmeans = KMeans(n_clusters=optimal_k, init='k-means++', n_init=10, random_state=42)
customer_df['Cluster'] = kmeans.fit_predict(features_scaled)

print("\nData Pelanggan dengan Label Cluster:")
print(customer_df.head())

# Menganalisis karakteristik setiap cluster
cluster_summary = customer_df.groupby('Cluster').agg({
    'Jumlah': 'mean',
    'Total_Harga': 'mean',
    'Frekuensi': 'mean',
    'Id_Pelanggan': 'count'
}).rename(columns={'Id_Pelanggan': 'Jumlah_Pelanggan'})

print("\nRingkasan Karakteristik Setiap Cluster:")
print(cluster_summary)

```

#### B. Hasil dan Insight yang Diperoleh

- Hasil Clustering:** Pelanggan berhasil dikelompokkan menjadi 3 segmen berbeda dengan karakteristik unik.

Data Perilaku Pelanggan:				
	Id_Pelanggan	Jumlah	Total_Harga	Frekuensi
0	C001	65	95000	2
1	C002	55	1050000	2
2	C003	50	260000	2
3	C004	30	910000	2
4	C005	18	300000	2

- **Analisis Karakteristik Cluster:**

Data Pelanggan dengan Label Cluster:					
	Id_Pelanggan	Jumlah	Total_Harga	Frekuensi	Cluster
0	C001	65	95000	2	1
1	C002	55	1050000	2	0
2	C003	50	260000	2	1
3	C004	30	910000	2	0
4	C005	18	300000	2	1

Ringkasan Karakteristik Setiap Cluster:				
Cluster	Jumlah	Total_Harga	Frekuensi	Jumlah_Pelanggan
0	38.500000	1.105000e+06	2.0	6
1	51.000000	2.881500e+05	2.0	10
2	27.222222	3.023472e+05	1.0	18

- **Cluster 0 (Pelanggan Loyal Bernilai Tinggi):**

- **Karakteristik:** Meskipun jumlah barang yang dibeli tidak setinggi cluster 1, mereka memiliki Total\_Harga rata-rata yang **paling tinggi** (sekitar 1.1 juta).
- **Insight:** Ini adalah segmen pelanggan paling berharga Anda. Mereka mungkin membeli produk-produk mahal atau dalam jumlah besar per transaksi.

- **Cluster 1 (Pelanggan Sering Beli, Nilai Sedang):**

- **Karakteristik:** Memiliki Jumlah barang rata-rata **paling tinggi** (51), namun Total\_Harga rata-ratanya lebih rendah dari Cluster 0 (sekitar 288 ribu).
- **Insight:** Ini adalah pelanggan yang sering bertransaksi namun mungkin membeli produk dengan harga lebih murah atau dalam jumlah lebih sedikit per transaksi. Mereka adalah pelanggan reguler Anda.

- **Cluster 2 (Pelanggan Jarang Beli/Baru):**
  - **Karakteristik:** Memiliki Frekuensi paling rendah (rata-rata 1 transaksi), dengan Jumlah dan Total\_Harga yang juga rendah.
  - **Insight:** Ini kemungkinan adalah pelanggan baru, pelanggan yang hanya sekali beli, atau pelanggan yang tidak aktif lagi.

### **C. Rekomendasi Implementasi**

Segmentasi ini sangat berharga untuk strategi pemasaran yang ditargetkan:

1. **Cluster 0 (High Value):**
  - a. **Implementasi:** Buat program loyalitas eksklusif, tawarkan akses awal ke produk baru, atau berikan layanan pelanggan premium. Tujuannya adalah untuk **mempertahankan (retention)** mereka.
2. **Cluster 1 (Regulars):**
  - a. **Implementasi:** Tawarkan promosi *cross-selling* atau *up-selling*. Misalnya, tawarkan produk pelengkap dari yang biasa mereka beli. Tujuannya adalah untuk **meningkatkan nilai transaksi (increase average order value)**.
3. **Cluster 2 (New/Lapsed):**
  - a. **Implementasi:** Kirimkan kampanye re-aktivasi melalui email atau SMS dengan penawaran khusus (misalnya, diskon untuk pembelian kedua) untuk mendorong mereka bertransaksi lagi. Tujuannya adalah untuk **mengonversi atau mengaktifkan kembali**.

## G. Perancangan Dashboard KPI

### 1. Pengantar dashboard - Jumlah Penjualan per-Produk (Tabel)

**Tabel Jumlah Penjualan per Produk**

<b>Nama_Produk</b>	<b>Jumlah</b> ▾
1. Fotokopi Hitam Putih	150
2. Fotokopi Warna	100
3. Print Dokumen Hitam Putih	75
4. Spidol Snowman	65
5. Gantungan ID Card	60
6. Buku Folio Bergaris	55
7. Binder Klip 32mm	50
8. Lem Kertas Fox	50
9. Kertas HVS F4 70gsm	50
10. Cetak Brosur Full Color	45
11. Amplop Coklat A4	45
12. Print Stiker Label	40
13. Isi Staples 24/6	40
14. Jilid Spiral 30mm	30
15. Map Plastik Folio	30

1 - 40 / 40 < >

**Tabel Jumlah Penjualan per Produk**

<b>Nama_Produk</b>	<b>Jumlah</b> ▾
16. Cutter Besar	30
17. Poster A3	30
18. Print Dokumen Warna	30
19. Jilid Spiral 25mm	25
20. Pulpen Pilot G2	25
21. Cetak Foto 4R	22
22. Buku Kwitansi	20
23. Laminating A4	20
24. Fotokopi Buku	20
25. Papan Nama Meja	20
26. Scan Dokumen	15
27. Kertas A4 80gsm	15
28. Laminating KTP	15
29. Pita Printer Epson LQ	15
30. Copy CD/DVD	15

1 - 40 / 40 < >

31.	Stempel Nama	15
32.	Gunting Kertas	10
33.	Tinta Printer Canon	10
34.	Flashdisk 16GB Sandisk	10
35.	Stempel Flash	8
36.	ID Card PVC	6
37.	Print Banner Outdoor	5
38.	Stapler Besar	5
39.	Cartridge HP 678	5
40.	Printer Canon G2010	2

1 - 40 / 40 < >

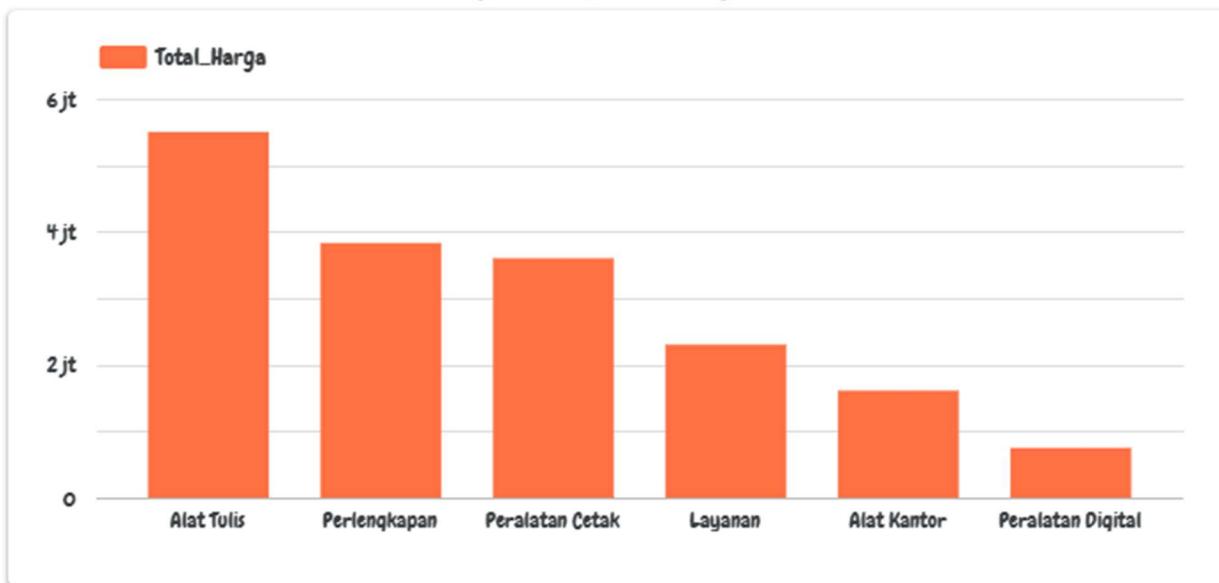
Visualisasi ini disajikan dalam bentuk tabel dua kolom yang menampilkan daftar produk beserta total jumlah penjualannya, diurutkan dari produk yang paling banyak terjual hingga yang paling sedikit. Kolom “Jumlah” diperkaya dengan gradasi warna biru, di mana intensitas warna merepresentasikan tingkat penjualan: semakin gelap warna biru, semakin tinggi jumlah produk terjual. Pendekatan ini memberikan kemudahan visual dalam mengenali tren secara cepat dan intuitif.

Dari visualisasi tersebut, terlihat jelas bahwa produk dengan penjualan tertinggi adalah Fotokopi Hitam Putih (150 unit), disusul oleh Fotokopi Warna (100 unit), dan Print Dokumen Hitam Putih (75 unit). Sebaliknya, beberapa produk seperti Printer Canon G2010, Cartridge HP 678, dan Stapler Besar hanya mencatatkan penjualan di bawah lima unit, yang menunjukkan tingkat permintaan yang rendah untuk produk-produk tersebut.

Melalui penggunaan gradasi warna, pengguna dapat mengidentifikasi outlier atau produk unggulan dan kurang laku secara sekilas, tanpa harus mencermati satu per satu angka yang tercantum.

## 2. KPI Total Penjualan per Kategori Produk

**Total Penjualan per Kategori Produk**



Visualisasi ini menampilkan total penjualan per kategori produk dalam bentuk diagram batang vertikal. Setiap batang mewakili satu kategori produk dengan tinggi batang menunjukkan besarnya nilai total penjualan dalam rupiah. Sumbu Y menggunakan skala jutaan rupiah untuk memudahkan interpretasi, dan data yang ditampilkan merupakan agregasi dari nilai Total\_Harga untuk masing-masing kategori. Visualisasi ini memberikan gambaran yang jelas mengenai kategori produk mana yang memberikan kontribusi terbesar terhadap total penjualan dalam periode tertentu.

Dari visualisasi tersebut, terlihat bahwa kategori “Alat Tulis” merupakan penyumbang penjualan tertinggi dengan total sekitar Rp5.500.000. Diikuti oleh kategori “Perlengkapan” dan “Peralatan Cetak” yang masing-masing mencatatkan penjualan sekitar Rp3.830.000 dan Rp3.600.000. Sementara itu, kategori dengan kontribusi penjualan terendah adalah “Peralatan Digital”, meskipun tetap berhasil mencatatkan penjualan sebesar Rp750.000. Pola ini menunjukkan bahwa kebutuhan rutin seperti alat tulis dan perlengkapan tetap menjadi andalan penjualan, sedangkan kategori seperti layanan dan peralatan digital cenderung lebih terbatas kontribusinya, kemungkinan karena bersifat musiman atau memiliki frekuensi pembelian rendah.

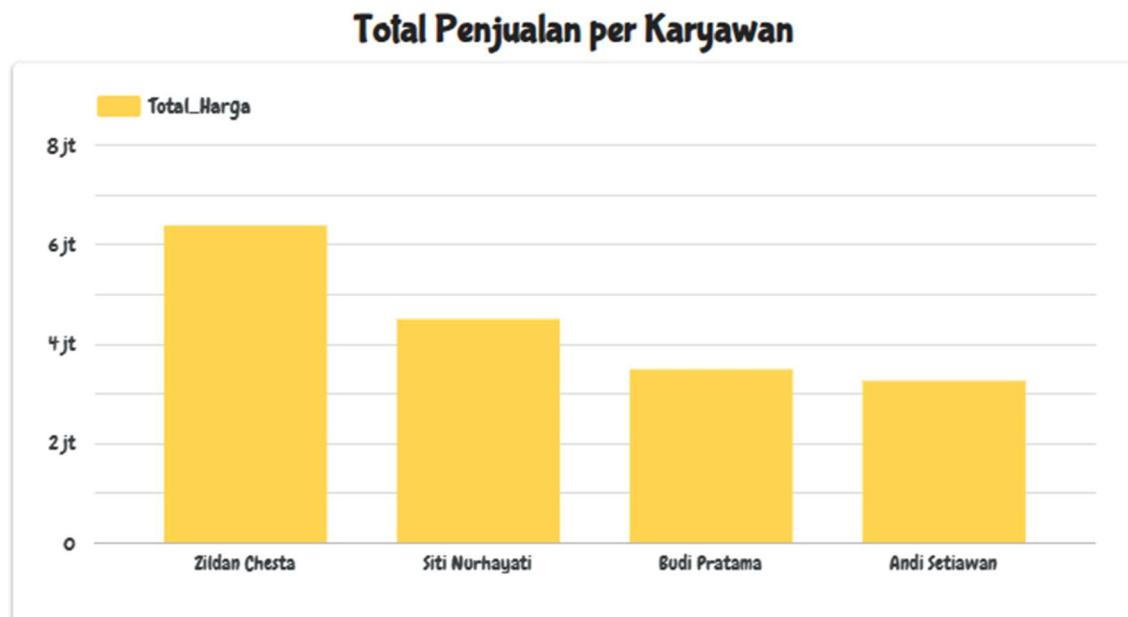
Jika dibandingkan dengan target yang telah ditentukan, maka dapat disimpulkan bahwa sebagian besar kategori belum berhasil mencapai target bulanan. Target untuk kategori Alat Tulis adalah Rp10.000.000, namun realisasinya baru mencapai Rp5.500.000.

Kategori Perlengkapan ditargetkan Rp7.500.000 namun hanya tercapai Rp3.830.000, dan Peralatan Cetak yang ditargetkan Rp6.000.000 hanya mencapai Rp3.600.000. Kategori Layanan dan Alat Kantor juga belum mencapai targetnya masing-masing yaitu Rp5.000.000 dan Rp2.000.000. Satu-satunya kategori yang telah mencapai target adalah Peralatan Digital yang menargetkan Rp500.000 dan berhasil melampaunya dengan penjualan sebesar Rp750.000.

Berdasarkan data tersebut, dapat disimpulkan bahwa dari enam kategori produk, lima di antaranya belum mencapai target penjualan yang telah ditetapkan. Meskipun secara nominal Alat Tulis dan Perlengkapan memberikan kontribusi terbesar, pencapaiannya masih jauh dari yang diharapkan. Hal ini menjadi indikasi bahwa performa penjualan produk secara umum masih di bawah ekspektasi yang telah ditentukan oleh target perusahaan.

Oleh karena itu, dibutuhkan strategi perbaikan dan peningkatan penjualan pada kategori-kategori utama seperti Alat Tulis, Perlengkapan, dan Peralatan Cetak. Rekomendasi yang dapat dilakukan adalah melakukan promosi lebih agresif, bundling produk, diskon, atau kampanye pemasaran berbasis kebutuhan rutin. Selain itu, analisis lebih lanjut terhadap penyebab tidak tercapainya target, seperti harga, stok produk, daya beli pelanggan, atau efektivitas pemasaran, juga sangat penting untuk mendorong peningkatan penjualan ke depannya.

### 3. KPI Total Penjualan per Karyawan



Visualisasi ini menampilkan total nilai penjualan yang berhasil dicapai oleh masing-masing karyawan dalam bentuk diagram batang vertikal. Pada sumbu-X

ditampilkan nama-nama karyawan, sedangkan sumbu-Y menunjukkan jumlah total penjualan dalam satuan jutaan rupiah. Setiap batang berwarna kuning digunakan untuk memperjelas perbandingan antar karyawan dalam hal performa penjualan. Selain itu, terdapat indikator tambahan berupa legenda yang menjelaskan bahwa nilai yang ditampilkan adalah hasil agregasi dari atribut Total\_Harga.

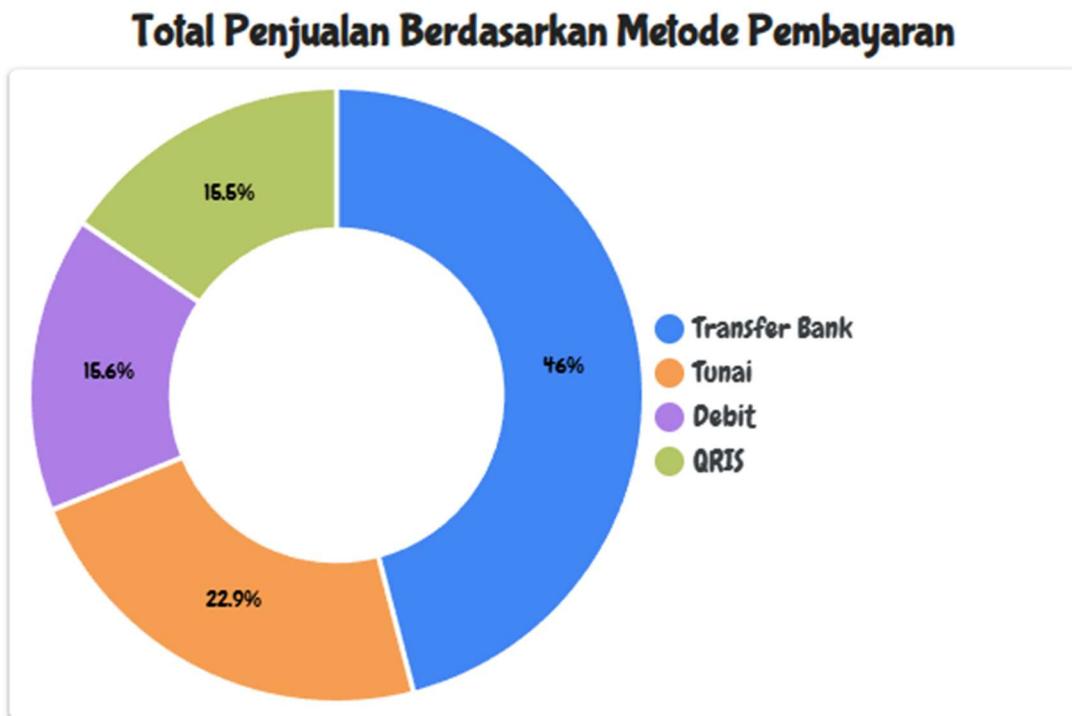
Dari visualisasi ini diperoleh beberapa insight yang menarik. Zildan Chesta menempati posisi teratas dengan total penjualan sebesar **Rp6.372.000**, yang menunjukkan performa yang sangat baik dan kontribusi besar terhadap pencapaian total penjualan perusahaan. Disusul oleh Siti Nurhayati yang juga memiliki performa kuat, dengan total penjualan mencapai **Rp4.490.000**. Sementara itu, Budi Pratama dan Andi Setiawan masing-masing mencatatkan total penjualan sebesar **Rp3.480.000** dan **Rp3.250.000**, yang meskipun lebih rendah, tetap memberikan kontribusi positif terhadap kinerja tim penjualan secara keseluruhan.

Jika dibandingkan dengan target yang telah ditetapkan, yaitu minimal **Rp2.500.000 per karyawan per bulan**, maka keempat karyawan tersebut berhasil melampaui target yang ditentukan. Tidak ada karyawan yang memiliki nilai penjualan di bawah ambang batas minimum, yang mengindikasikan bahwa secara individu, performa penjualan setiap anggota tim sudah berada pada tingkat yang baik dan sesuai dengan harapan.

Dari pencapaian ini, dapat disimpulkan bahwa **seluruh karyawan telah memenuhi target KPI** yang ditetapkan untuk periode tersebut. Bahkan, beberapa di antaranya menunjukkan performa yang sangat unggul, seperti Zildan dan Siti. Hal ini mencerminkan adanya efektivitas dalam sistem penjualan dan mungkin juga faktor seperti pengalaman, strategi pendekatan pelanggan, atau segmentasi yang tepat.

Sebagai rekomendasi, perusahaan dapat mempertimbangkan untuk memberikan **insentif atau penghargaan kinerja** kepada karyawan dengan pencapaian tertinggi sebagai bentuk motivasi dan pengakuan. Selain itu, strategi dan pendekatan yang digunakan oleh Zildan dan Siti dapat dianalisis lebih lanjut dan dijadikan best practice yang dapat diterapkan atau ditularkan kepada anggota tim lain untuk menjaga dan meningkatkan kinerja secara konsisten di periode berikutnya.

#### 4. KPI Total Penjualan: Prediksi Penjualan Berdasarkan Metode Pembayaran



Visualisasi ini disajikan dalam bentuk **grafik donat (donut chart)** untuk menggambarkan distribusi prediksi penjualan berdasarkan metode pembayaran yang digunakan oleh pelanggan. Masing-masing segmen diberi warna berbeda yang mewakili metode pembayaran tertentu, seperti biru untuk Transfer Bank, oranye untuk Tunai, ungu untuk Debit, dan hijau muda untuk QRIS. Di tengah grafik ditampilkan persentase kontribusi masing-masing metode terhadap total penjualan, yang memudahkan pembaca dalam mengidentifikasi proporsi serta preferensi pelanggan dalam memilih cara bertransaksi.

Dari grafik tersebut diperoleh beberapa insight penting. Transfer Bank merupakan metode pembayaran paling dominan dengan kontribusi sebesar **46%**, yang menunjukkan kecenderungan pelanggan dalam menggunakan metode digital yang dianggap aman dan praktis. Di posisi kedua, Tunai masih digunakan oleh sebagian besar pelanggan dengan persentase **22,9%**, mengindikasikan bahwa transaksi konvensional masih memiliki tempat, meskipun tren digital semakin menguat. Sementara itu, metode Debit dan QRIS memiliki proporsi yang hampir seimbang, yaitu masing-masing sebesar **15,6%** dan **15,5%**, yang mencerminkan adanya peningkatan penggunaan metode digital berbasis kartu dan QR code, khususnya di kalangan pelanggan yang melek teknologi.

Jika dibandingkan dengan target yang telah ditetapkan, maka hasil visualisasi menunjukkan beberapa pencapaian dan juga ketidaksesuaian. Untuk metode Transfer Bank, target adalah peningkatan  $\geq 15\%$ , dan dengan kontribusi sebesar 46%, dapat disimpulkan

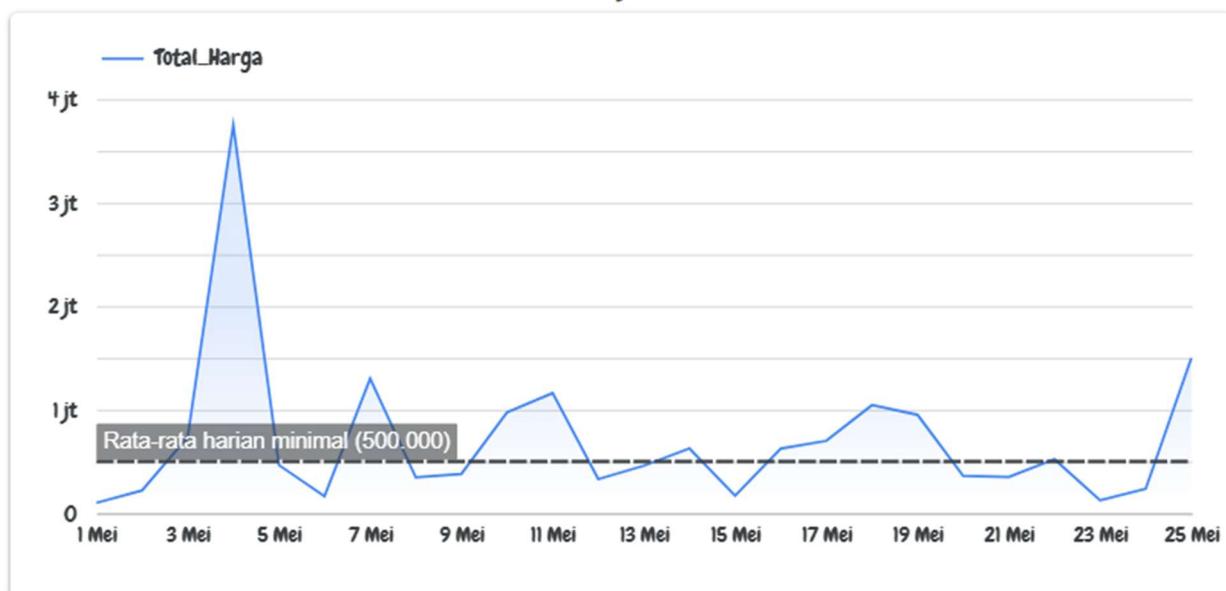
bahwa target tersebut **berhasil tercapai dengan sangat baik**. Metode Tunai memiliki target untuk stabil atau turun  $\leq 5\%$ , namun dengan kontribusi sebesar 22,9%, dapat disimpulkan bahwa **penggunaan tunai masih relatif tinggi dan belum memenuhi target pengurangan**. Untuk Debit, target maksimal adalah 35%, dan pencapaian 15,6% masih jauh di bawah ambang batas, sehingga masih sesuai target. Sedangkan untuk QRIS, targetnya adalah meningkat sebesar 10% dari bulan sebelumnya. Namun karena tidak tersedia data bulan sebelumnya untuk perbandingan, maka **belum dapat disimpulkan apakah target peningkatan tercapai atau tidak**, meskipun secara nilai persentase (15,5%) tergolong menjanjikan.

Dari analisis tersebut, dapat disimpulkan bahwa metode pembayaran digital seperti Transfer Bank, Debit, dan QRIS telah mulai mendominasi perilaku transaksi pelanggan, sementara penggunaan Tunai masih tinggi dan memerlukan perhatian lebih dalam upaya transformasi digital. Pencapaian target untuk metode Transfer Bank dan Debit sangat positif, sedangkan QRIS berpotensi meningkat, namun perlu data lanjutan untuk evaluasi lebih akurat.

Sebagai rekomendasi, perusahaan dapat mendorong penggunaan metode QRIS lebih agresif dengan memberikan **diskon khusus atau promo cashback** agar penggunaannya meningkat secara konsisten dari bulan ke bulan. Selain itu, edukasi pelanggan terkait kemudahan dan keamanan metode pembayaran digital juga perlu diperluas, terutama untuk mengurangi ketergantungan pada transaksi tunai. Transformasi ini akan sejalan dengan efisiensi operasional dan pengelolaan transaksi yang lebih modern dan transparan.

## 5. KPI Total Penjualan: Total Pendapatan Harian

**Total Pendapatan Harian**



Visualisasi ini disajikan dalam bentuk **grafik garis (line chart)** yang menampilkan tren pendapatan harian perusahaan selama periode 1 hingga 25 Mei 2025. Sumbu-X mewakili tanggal, sedangkan sumbu-Y menunjukkan nilai total pendapatan harian dalam rupiah. Garis berwarna biru menggambarkan fluktuasi pendapatan dari hari ke hari, dengan arsiran di bawah kurva yang menyoroti volume total pendapatan. Sebagai acuan, ditambahkan garis horizontal putus-putus berwarna hitam pada level Rp500.000 yang berfungsi sebagai indikator target rata-rata minimal pendapatan harian yang ditetapkan perusahaan.

Dari grafik tersebut, terlihat adanya **fluktuasi yang cukup tajam** pada pendapatan harian. Puncak tertinggi terjadi pada tanggal **4 Mei** dengan pendapatan mencapai **Rp3.750.000**, yang kemungkinan besar disebabkan oleh aktivitas promosi, penjualan massal, atau adanya pesanan besar dalam satu hari. Beberapa hari lain seperti **7, 10, 11, 18, 19, dan 25 Mei** juga menunjukkan performa penjualan yang cukup baik karena berhasil melampaui ambang target harian. Namun, pada mayoritas hari lainnya, pendapatan tercatat **berada di bawah target minimum Rp500.000**, bahkan ada yang sangat rendah seperti pada **1 Mei (Rp100.000)** dan **23 Mei (Rp125.000)**. Hal ini menandakan bahwa penjualan masih belum konsisten dan sangat dipengaruhi oleh faktor-faktor eksternal atau musiman.

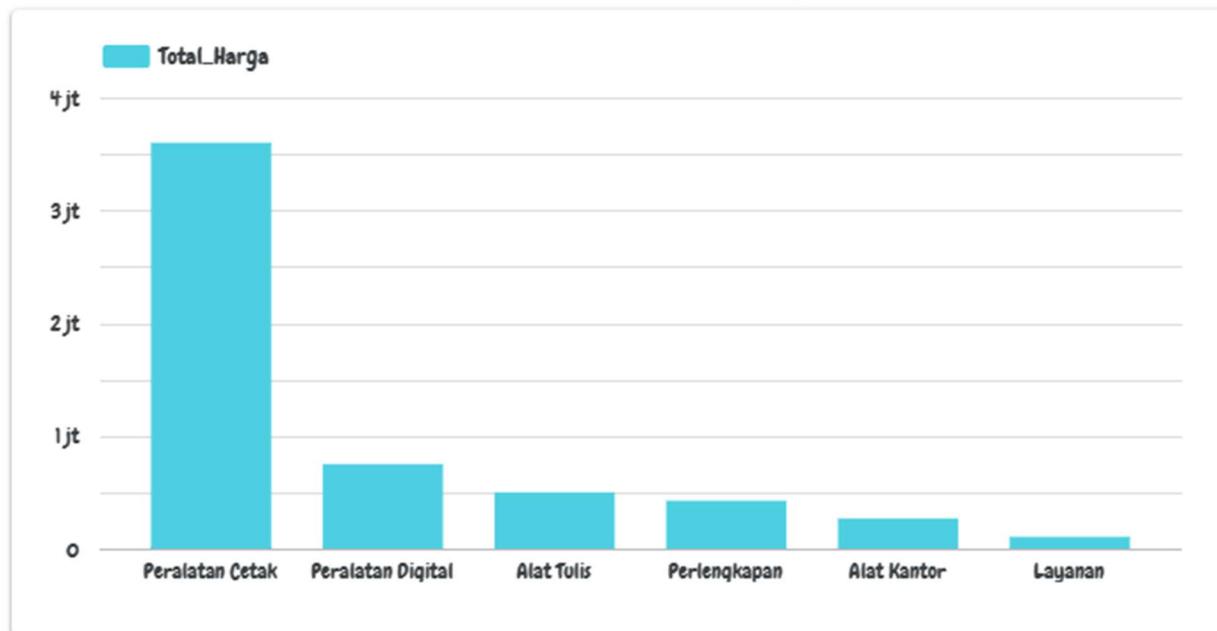
Berdasarkan perhitungan, total pendapatan selama 25 hari adalah **Rp18.744.000**, dengan rata-rata pendapatan harian sebesar **Rp749.760**. Secara angka rata-rata, KPI ini **berhasil melampaui target minimal Rp500.000 per hari**. Namun, jika dilihat dari distribusi hari per hari, hanya sekitar **9 dari 25 hari (36%)** yang berhasil mencapai atau melampaui target tersebut. Ini menunjukkan bahwa **rata-rata tinggi tersebut ditopang oleh beberapa hari lonjakan besar**, bukan karena konsistensi kinerja harian yang stabil.

Kesimpulannya, meskipun secara rata-rata target telah tercapai, **stabilitas dan keberlanjutan pencapaian target harian masih menjadi tantangan utama**. Penjualan yang terlalu bergantung pada momen tertentu dapat menimbulkan risiko terhadap arus kas dan perencanaan stok. Oleh karena itu, dibutuhkan strategi yang dapat meningkatkan konsistensi penjualan harian.

Sebagai rekomendasi, perusahaan perlu mempertimbangkan penerapan **strategi promosi yang merata sepanjang bulan**, seperti diskon terbatas harian, bundling produk, atau loyalty program. Selain itu, analisis terhadap hari-hari dengan penjualan tinggi dapat digunakan sebagai dasar untuk mengidentifikasi pola perilaku pelanggan, yang kemudian dioptimalkan untuk mendongkrak performa pada hari-hari dengan penjualan rendah. Pendekatan berbasis data seperti ini diharapkan dapat meningkatkan stabilitas dan memperkuat pencapaian target KPI secara berkelanjutan.

## 6. KPI Transaksi: Rata-Rata Nilai Transaksi per Kategori Produk

### Rata-rata Nilai Transaksi per Kategori Produk



Visualisasi ini menggunakan **diagram batang vertikal** untuk menggambarkan rata-rata nilai transaksi dari setiap kategori produk. Sumbu-X memuat nama kategori produk, sedangkan sumbu-Y menunjukkan nilai rata-rata transaksi dalam satuan rupiah. Setiap batang mewakili kategori produk tertentu, dengan tinggi batang mencerminkan seberapa besar rata-rata nilai transaksi yang terjadi dalam kategori tersebut. Kategori yang dianalisis meliputi: Peralatan Cetak, Peralatan Digital, Alat Tulis, Perlengkapan, Alat Kantor, dan Layanan.

Dari visualisasi, terlihat bahwa **Peralatan Cetak menempati posisi tertinggi**, dengan rata-rata transaksi mencapai **Rp3.600.000**, jauh melampaui kategori lainnya. Hal ini menunjukkan bahwa transaksi pada kategori ini bersifat besar, baik dari sisi harga satuan produk maupun kuantitas pembelian. **Peralatan Digital dan Alat Tulis** berada di kisaran menengah, masing-masing mencapai **Rp750.000** dan **Rp500.000**, yang menandakan bahwa produk-produk dalam kategori ini cenderung memiliki nilai ekonomis yang sedang hingga tinggi. Sementara itu, kategori seperti **Perlengkapan (Rp425.555)**, **Alat Kantor (Rp268.333)**, dan **Layanan (Rp104.636)** mencatatkan nilai transaksi yang lebih rendah, mengindikasikan produk/jasa yang terjangkau atau pembelian dalam jumlah kecil per transaksi.

Bila dibandingkan dengan target KPI yang telah ditetapkan, maka seluruh kategori **berhasil memenuhi atau bahkan melampaui target**. Misalnya, Alat Tulis memiliki target Rp100.000 dan tercapai sebesar Rp500.000. Demikian pula, Perlengkapan (target Rp75.000) dan Layanan (target Rp200.000), walaupun nilai Layanan lebih rendah dari target, yaitu Rp104.636, menandakan **Layanan menjadi satu-satunya kategori yang**

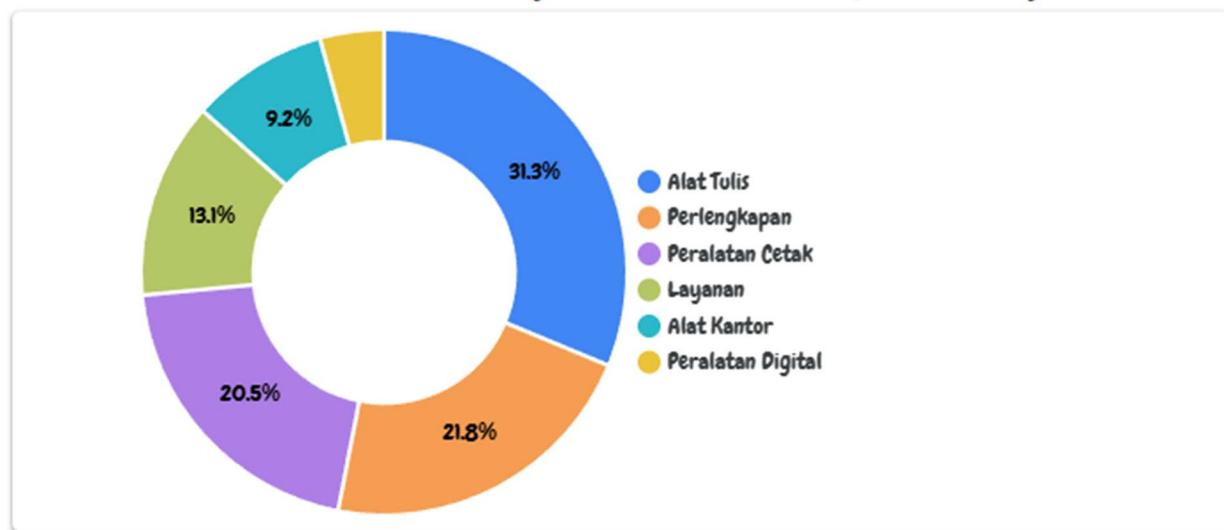
**belum mencapai target.** Sementara itu, kategori lainnya seperti Peralatan Cetak (target  $\geq$  Rp80.000) dan Peralatan Digital (target  $\geq$  Rp60.000) telah melampaui target dengan sangat signifikan.

Dari analisis ini dapat disimpulkan bahwa rata-rata nilai transaksi per kategori secara umum menunjukkan performa yang sangat baik, khususnya untuk produk fisik seperti alat tulis dan peralatan cetak. Hal ini mencerminkan bahwa pelanggan cenderung melakukan pembelian dengan nominal tinggi pada beberapa kategori tertentu, terutama pada produk dengan nilai dan fungsi khusus. Namun demikian, **kategori Layanan masih menjadi perhatian utama**, karena nilai rata-rata transaksinya belum mencapai target.

Sebagai rekomendasi, perusahaan dapat fokus meningkatkan nilai transaksi pada kategori Layanan dengan strategi seperti **menawarkan bundling jasa, meningkatkan kualitas layanan berbayar, atau mempromosikan paket langganan**. Selain itu, untuk kategori dengan performa tinggi, strategi **upselling** dan **cross-selling** dapat dioptimalkan guna mempertahankan dan meningkatkan nilai rata-rata transaksi. Pemahaman terhadap karakteristik masing-masing kategori juga dapat membantu dalam penyusunan strategi penjualan yang lebih efektif dan tersegmentasi.

## 7. KPI Transaksi: Persentase Kontribusi Kategori Produk Terhadap Total Penjualan

**Persentase Kontribusi Kategori Produk Terhadap Total Penjualan**



Visualisasi yang digunakan adalah **diagram donat**, yang menggambarkan sebaran kontribusi masing-masing kategori produk terhadap total penjualan. Setiap irisan pada diagram menunjukkan proporsi penjualan yang berasal dari kategori tertentu, dengan

besaran persentase yang tertera untuk memperjelas kontribusi relatif. Kategori yang divisualisasikan mencakup Alat Tulis, Perlengkapan, Peralatan Cetak, Layanan, Alat Kantor, dan Peralatan Digital.

Berdasarkan visualisasi, **kategori Alat Tulis memberikan kontribusi terbesar**, yaitu 31,3% dari total penjualan, mengindikasikan dominasi yang kuat dalam performa penjualan dan permintaan yang tinggi. **Perlengkapan (21,8%) dan Peralatan Cetak (20,5%)** juga mencatatkan kontribusi yang besar, yang menunjukkan bahwa produk-produk dalam kategori ini memiliki potensi penjualan yang baik, baik dari sisi volume maupun nilai. **Layanan berada di posisi tengah dengan kontribusi 13,1%**, sementara **Alat Kantor (9,2%) dan Peralatan Digital (4,3%)** berada di posisi terbawah.

Jika dibandingkan dengan target yang ditetapkan, dapat dilihat bahwa **Alat Tulis (target = 25%) dan Peralatan Cetak (target = 10%) berhasil melampaui target**, menunjukkan performa yang sangat baik. **Perlengkapan juga melebihi rentang target kontribusinya (10%–15%)**, yang menandakan pertumbuhan signifikan namun mungkin perlu pengelolaan ulang agar distribusi produk tetap seimbang. Sebaliknya, **kategori Layanan (target = 25%–35%) belum mencapai target, hanya sebesar 13,1%**, menunjukkan adanya kesenjangan antara harapan dan realisasi. **Alat Kantor (target = 20%) dan Peralatan Digital (target = 10%–13%) juga belum memenuhi target**, dengan kontribusi di bawah harapan.

Dari hasil ini dapat disimpulkan bahwa kontribusi penjualan masih terfokus pada kategori tertentu, terutama produk fisik seperti Alat Tulis dan Peralatan Cetak, sementara beberapa kategori lainnya tertinggal jauh dari target. Hal ini menunjukkan adanya ketidakseimbangan dalam portofolio penjualan produk, yang dapat berdampak pada ketahanan dan keberagaman sumber pendapatan jangka panjang.

Sebagai rekomendasi, perlu dilakukan upaya **peningkatan penjualan pada kategori yang belum mencapai target**, terutama Layanan, Alat Kantor, dan Peralatan Digital. Strategi seperti **bundling, promosi tematik, pelatihan staf penjualan**, serta **penyesuaian harga atau model layanan** dapat menjadi langkah yang tepat untuk mendorong pertumbuhan di kategori tersebut. Selain itu, karena beberapa kategori sudah melampaui target, maka perusahaan juga perlu **menjaga keseimbangan kontribusi antar kategori** agar struktur pendapatan lebih stabil dan tidak tergantung pada kategori tertentu saja.

Link Dashboard Penjualan Kelompok 7:

<https://lookerstudio.google.com/reporting/e97e8966-6c18-4f99-8cfa-91346040dcff>

## H. Kesimpulan

Proyek ini telah berhasil menunjukkan potensi besar dalam meningkatkan kapabilitas analitik bisnis Faeyza Printing melalui pembangunan sistem *Business Intelligence* (BI) yang terintegrasi. Dengan mengkonsolidasikan data krusial dari berbagai dimensi — mencakup penjualan, pelanggan, produk, dan karyawan — ke dalam sebuah *data warehouse* sentral melalui proses *Extract, Transform, and Load* (ETL) yang efisien, Faeyza Printing kini memiliki fondasi data yang kuat dan terstruktur. Integrasi ini merupakan langkah fundamental yang memungkinkan perancangan *dashboard Key Performance Indicator* (KPI) yang intuitif dan *real-time*, memberikan manajemen kemampuan untuk memantau kinerja bisnis secara holistik. Lebih jauh lagi, dengan penerapan analisis *data mining* pada data yang telah terintegrasi, perusahaan dapat menggali wawasan prediktif dan deskriptif yang mendalam, memungkinkan pengambilan keputusan yang lebih cepat, tepat, dan berbasis data. Manfaat utama yang diharapkan dari implementasi sistem BI ini sangat signifikan, meliputi peningkatan efisiensi operasional melalui identifikasi area perbaikan, optimalisasi arus kas dan penetapan target penjualan yang lebih realistik dan ambisius, serta pengembangan strategi pemasaran yang jauh lebih tertarget berkat pemahaman mendalam tentang segmentasi dan perilaku pelanggan. Dengan demikian, sistem BI ini tidak hanya berfungsi sebagai alat pelaporan, tetapi juga sebagai mesin pendorong pertumbuhan dan keunggulan kompetitif bagi **Faeyza Printing**.

## I. Pembagian Tugas Anggota Tim

Nama	Tugas
Prana Ichlasul Kautsar	Melakukan Pengambilan Data dengan Owner Objek Studi Kasus untuk analisis. Melakukan Rekayasa dan Modifikasi data agar lebih fit untuk analisis diolah Melakukan process Extract, Transform, Dan load di Pentaho Data Integration, berupa sorting, stream Lookup, Select Values dan Output xlsx dan sql
Achmad Vickram Ritonga	Mengerjakan Dashboard Penjualan untuk 5 KPI dan mendesain di Looker Studio Membuat kolom Calculate di Looker Studio untuk analisis KPI dan Chart dashboard
Gading Ainunzaky Widianto	Membuat Rekomendasi dari Hasil Perancangan Dashboard KPI serta Langkah lebih lanjut. Menyusun Laporan dan Presentasi
Nicodemus Benaya Gavia L. T.	Melakukan Data mining Untuk mendapatkan Insight dari data analisis Melakukan process datamining menggunakan Metode Random Forest dan Classification Melakukan Evaluasi Model dan hasil data mining di Vscode
Fredy Arya Hutama	Menentukan Objektif dari objek tugas besar Membuat Balance Scorecards dan menurunkannya menjadi KPI untuk standarisasi pengolahan data dan target analisis
M. Affan Febryan T.	Membuat perancangan star schema dari data dan dimensi yang ada Melakukan implementasi star schema di RDBMS dengan SQL.