

LAPORAN PROYEK DATA ANALYST

Analisis Kinerja dan Profitabilitas Penjualan Superstore Berdasarkan Kategori, Waktu, Segmen, dan Wilayah

4143105 - Data Analyst



Dibuat Oleh:

11423017

Anisetus Bambang Manalu

11423043

Jonathan Prima Tamba

11423044

Nathanael T. J. Tampubolon

Fakultas Vokasi

Institut Teknologi Del

DAFTAR ISI

1.	PENDAHULUAN.....	4
1.1	Latar Belakang.....	4
1.2	Rumusan Masalah	4
1.3	Tujuan Penelitian.....	5
2.	TAHAPAN CROSS INDUSTRY STANDARd PROCESS for DATA MINING (CRISP – DM)	6
2.1	Business Understanding	6
2.1.1	Konteks Bisnis dan Latar Belakang Masalah	6
2.1.2	Permasalahan Analitik Bisnis	6
2.1.3	Tujuan Analisis & Kebutuhan Insight Bisnis	6
2.2	Data Understanding (Pemahaman Data)	7
2.2.1	Struktur Data dan Tipe Data	7
2.2.2	Analisis Missing Values	8
2.3	Data Preparation (Persiapan Data).....	9
2.3.1	Penanganan Missing Values	9
2.3.2	Transformasi Data (Data Transformation).....	9
2.4	Modeling (Descriptive Analysis).....	10
3.	Hasil eKsplorasi dan analisis.....	13
3.1	Ringkasan KPI Keseluruhan (dari data)	13
3.2	Analisis Berdasarkan Kategori Produk.....	14
3.3	Analisis Berdasarkan Segmen Pelanggan.....	15
3.4	Analisis Berdasarkan Wilayah (Region)	15
3.5	Analisis Berdasarkan Waktu (Tren Penjualan Bulanan)	16
3.6	Analisis Produk Terlaris (Top Product).....	18
4.	INSIGHT DAN REKOMENDASI BISNIS.....	20
4.1	Insight Utama Bisnis	20
4.1.1	Kategori Produk Technology Mendominasi Penjualan	20
4.1.2	Segmen Consumer Memberikan Kontribusi Terbesar	20
4.1.3	Wilayah Barat (West Region) Menjadi Pusat Aktivitas Penjualan.....	20
4.1.4	Pola Penjualan Musiman (Seasonal Trend)	21
4.1.5	Produk Terlaris: Canon imageCLASS 2200 Advanced Copier.....	21
4.2	Rekomendasi Bisnis Strategis	22
4.2.1	Optimalisasi Portofolio Produk.....	22
4.2.2	Penguatan Strategi Pemasaran Segmen Consumer	22
4.2.3	Efisiensi Logistik dan Distribusi Wilayah West	22

4.2.4	Strategi Promosi Musiman.....	22
4.2.5	Manajemen Pelanggan Bernilai Tinggi	22
4.3	Kesimpulan Bab Insight dan Rekomendasi.....	23

1. PENDAHULUAN

1.1 Latar Belakang

Perkembangan teknologi informasi dan sistem manajemen data saat ini telah mengubah cara perusahaan dalam mengambil keputusan bisnis. Dalam era digital, setiap transaksi yang terjadi di perusahaan ritel modern dapat direkam secara detail dalam bentuk data transaksi. Data ini mencakup berbagai aspek seperti tanggal transaksi, produk yang dibeli, harga jual, diskon yang diberikan, segmen pelanggan, serta wilayah tempat penjualan dilakukan. Dengan jumlah data yang sangat besar dan beragam tersebut, perusahaan memiliki peluang besar untuk melakukan analisis mendalam guna memperoleh insight yang bernilai. Salah satu pendekatan yang banyak digunakan dalam menganalisis data penjualan adalah analisis deskriptif berbasis *data mining*, yang bertujuan untuk memahami pola-pola penjualan dari berbagai sudut pandang.

Superstore merupakan contoh tipikal perusahaan ritel yang menjual berbagai produk kebutuhan kantor, teknologi, serta furnitur kepada pelanggan dari berbagai segmen dan wilayah. Dalam konteks bisnis seperti ini, data penjualan yang dihasilkan setiap hari dapat menjadi sumber informasi penting untuk menilai kinerja bisnis secara keseluruhan. Analisis terhadap data tersebut memungkinkan manajemen untuk mengidentifikasi kategori produk yang paling laris, segmen pelanggan yang paling aktif, wilayah dengan penjualan tertinggi, serta periode waktu tertentu di mana penjualan mencapai puncaknya. Dengan mengetahui hal-hal tersebut, perusahaan dapat mengoptimalkan strategi penjualan, memperbaiki kebijakan promosi, dan meningkatkan efisiensi distribusi.

Namun, tantangan yang sering dihadapi perusahaan adalah bagaimana mengolah data penjualan yang besar dan kompleks menjadi informasi yang mudah dipahami dan berguna bagi pengambil keputusan. Banyak perusahaan memiliki data yang kaya, tetapi belum memanfaatkannya secara optimal untuk mendukung keputusan strategis. Di sinilah peran analisis data menjadi penting. Dengan menerapkan metode analisis data berbasis pendekatan *CRISP-DM (Cross Industry Standard Process for Data Mining)*, proses analisis dapat dilakukan secara sistematis mulai dari pemahaman bisnis, pemahaman data, persiapan data, hingga tahap pemodelan dan evaluasi hasil.

Dalam proyek ini, dilakukan analisis terhadap dataset Superstore untuk mengevaluasi kinerja penjualan berdasarkan beberapa dimensi utama, yaitu kategori produk, waktu transaksi, segmen pelanggan, dan wilayah penjualan. Analisis ini diharapkan mampu memberikan gambaran menyeluruh mengenai performa bisnis Superstore, sekaligus membantu manajemen dalam mengambil keputusan berbasis data (*data-driven decision making*). Melalui hasil analisis, dapat diidentifikasi faktor-faktor apa saja yang paling memengaruhi tingkat penjualan dan profit, serta peluang strategis untuk meningkatkan pendapatan dan efisiensi operasional di masa mendatang.

1.2 Rumusan Masalah

1. Kategori produk mana yang memberikan kontribusi terbesar terhadap total penjualan dan profitabilitas (profit ratio)?

2. Bagaimana tren penjualan dan profit dari waktu ke waktu (bulanan dan tahunan)?
3. Segmen pelanggan dan wilayah (region/city/state) mana yang menunjukkan kinerja penjualan terbaik?
4. Seberapa kuat hubungan antara nilai penjualan (sales) dan profit?
5. Insight apa yang dapat diperoleh dari hasil visualisasi untuk mendukung rekomendasi peningkatan strategi bisnis dan profitabilitas?

1.3 Tujuan Penelitian

Tujuan utama dari penelitian ini adalah untuk menganalisis dan memahami kinerja penjualan Superstore dari berbagai sudut pandang, yaitu kategori produk, waktu transaksi, segmen pelanggan, serta wilayah penjualan. Melalui analisis ini, diharapkan dapat diperoleh gambaran menyeluruh mengenai pola dan tren penjualan yang terjadi pada perusahaan, sehingga informasi tersebut dapat menjadi dasar dalam pengambilan keputusan bisnis yang lebih efektif. Penelitian ini juga bertujuan untuk mengidentifikasi kategori produk yang memberikan kontribusi terbesar terhadap total penjualan dan profit, mengetahui wilayah dengan performa penjualan tertinggi, serta menentukan segmen pelanggan yang paling potensial bagi peningkatan pendapatan perusahaan. Selain itu, penelitian ini bermaksud untuk mengevaluasi hubungan antara variabel-variabel utama seperti penjualan (*sales*), keuntungan (*profit*), kuantitas produk terjual (*quantity*), dan diskon (*discount*), guna menemukan faktor-faktor yang paling berpengaruh terhadap profitabilitas.

Dengan melakukan analisis deskriptif berbasis pendekatan *CRISP-DM (Cross Industry Standard Process for Data Mining)*, penelitian ini juga bertujuan untuk memberikan insight bisnis yang dapat diimplementasikan secara langsung, seperti rekomendasi strategi promosi, pengelolaan stok barang, serta prioritas wilayah pemasaran. Hasil dari penelitian ini diharapkan tidak hanya memberikan informasi statistik, tetapi juga menghasilkan pemahaman yang lebih dalam mengenai perilaku pelanggan dan dinamika pasar yang dihadapi oleh Superstore. Dengan demikian, penelitian ini dapat menjadi acuan dalam penyusunan strategi bisnis berbasis data untuk meningkatkan kinerja penjualan dan profit perusahaan secara berkelanjutan.

2. TAHAPAN CROSS INDUSTRY STANDARD PROCESS FOR DATA MINING (CRISP – DM)

2.1 Business Understanding

Tahap Business Understanding merupakan langkah pertama dalam kerangka kerja CRISP-DM yang berfokus pada pemahaman tujuan bisnis dan kebutuhan analitis organisasi. Tahap ini bertujuan untuk mengidentifikasi permasalahan utama yang ingin diselesaikan melalui analisis data serta menentukan sasaran yang diharapkan dari hasil analisis.

2.1.1 Konteks Bisnis dan Latar Belakang Masalah

Perusahaan membutuhkan gambaran kinerja penjualan untuk merencanakan stok, promosi, dan strategi pemasaran. Dengan mengetahui kategori, segmen, atau wilayah yang berkinerja baik maupun yang kurang optimal, manajemen dapat memprioritaskan sumber daya dengan lebih efisien.

Dalam penelitian ini, analisis dilakukan dengan bantuan *dashboard interaktif Tableau* untuk memberikan visualisasi yang komprehensif terhadap indikator bisnis seperti sales, profit, quantity, dan discount, sehingga hasil analisis dapat langsung dimanfaatkan oleh pihak manajemen dalam pengambilan keputusan.

2.1.2 Permasalahan Analitik Bisnis

Permasalahan utama yang dihadapi perusahaan adalah bagaimana mengolah volume data transaksi yang besar agar dapat menghasilkan informasi yang mudah dipahami dan bermanfaat bagi pengambilan keputusan strategis. Diperlukan proses ringkasan dan visualisasi yang mampu menampilkan performa bisnis dari berbagai sudut pandang, seperti kategori produk, segmen pelanggan, wilayah, dan waktu transaksi. Selain itu, perusahaan perlu mampu mengidentifikasi anomali seperti transaksi dengan nilai penjualan tinggi tetapi profit rendah, serta menentukan prioritas inisiatif bisnis seperti strategi promosi, pengadaan produk, dan alokasi sumber daya penjualan.

- Data transaksi besar, perlu ringkasan agregat (per kategori, segmen, wilayah, dan waktu).
- Mengidentifikasi anomali (mis. penjualan besar tapi profit negatif).
- Menentukan prioritas inisiatif (promosi, pemasokan produk, alokasi sales).

2.1.3 Tujuan Analisis & Kebutuhan Insight Bisnis

Tujuan utama analisis ini adalah memberikan *insight* yang dapat ditindaklanjuti oleh manajemen, seperti menentukan kategori produk yang perlu promosi tambahan, wilayah yang membutuhkan perhatian khusus, serta periode waktu terbaik untuk melaksanakan program pemasaran.

Melalui dashboard Tableau, hasil analisis disajikan dalam bentuk ringkasan KPI utama (Total Sales, Total Profit, Average Order Value), pivot table berdasarkan kategori, segmen, dan region, serta grafik tren bulanan dan tahunan. Visualisasi ini mempermudah manajemen dalam memahami kondisi bisnis secara menyeluruh dan mengidentifikasi peluang peningkatan profitabilitas.

2.2 Data Understanding (Pemahaman Data)

Tahap Data Understanding merupakan langkah kedua dalam proses CRISP-DM yang bertujuan untuk memahami karakteristik, struktur, serta kualitas data yang akan digunakan dalam analisis. Pada tahap ini, dilakukan eksplorasi awal terhadap dataset untuk memastikan bahwa data yang tersedia relevan dengan tujuan analisis bisnis, memiliki format yang konsisten, dan bebas dari kesalahan atau ketidaksesuaian.

Dataset yang digunakan dalam penelitian ini adalah Superstore Dataset, yang berisi catatan transaksi penjualan dari periode tertentu dengan berbagai atribut seperti tanggal pemesanan, kategori produk, wilayah penjualan, hingga informasi pelanggan. Dataset ini merepresentasikan aktivitas penjualan pada perusahaan ritel yang menjual produk dalam tiga kategori utama: *Furniture*, *Office Supplies*, dan *Technology*.

Melalui tahap ini, dilakukan identifikasi struktur data, tipe data setiap atribut, serta deteksi adanya nilai yang hilang (*missing values*) atau anomali yang dapat memengaruhi hasil analisis. Pemahaman terhadap data ini menjadi dasar penting dalam tahapan berikutnya, yaitu Data Preparation dan Modeling, karena kualitas hasil analisis sangat bergantung pada pemahaman yang akurat terhadap data yang digunakan.

2.2.1 Struktur Data dan Tipe Data

Tabel Struktur Data dan Tipe Data

No.	Nama Kolom	Tipe Data Asli	Keterangan / Hasil Parsing
1	order_id	String	Nomor unik untuk setiap pesanan (ID transaksi)
2	order_date	String → Datetime	Tanggal pemesanan barang, diparse menjadi format waktu
3	ship_date	String / Date	Tanggal pengiriman barang
4	customer	String	Nama pelanggan yang melakukan pembelian
5	manufactory	String	Nama produsen atau pabrikan produk
6	product_name	String	Nama produk yang dijual
7	Segment	String	Segmen pelanggan (Consumer, Corporate, Home Office)
8	category	String	Kategori utama produk (Furniture, Office Supplies, Technology)
9	subcategory	String	Subkategori produk (mis. Chairs, Phones, Binders, dll.)

10	region	String	Wilayah penjualan (West, East, Central, South)
11	zip	String / Integer	Kode pos lokasi pelanggan
12	city	String	Kota tempat pelanggan berada
13	state	String	Provinsi atau negara bagian tempat pelanggan
14	country	String	Negara pelanggan
15	discount	Float / Numeric	Persentase diskon yang diberikan pada pesanan
16	profit	Float / Numeric	Nilai keuntungan dari pesanan
17	quantity	Integer	Jumlah barang yang dibeli dalam transaksi
18	sales	Float / Numeric	Nilai total penjualan (harga jual setelah diskon)
19	profit_margin	Float / Numeric	Rasio profit terhadap sales (indikator margin keuntungan)

Keterangan:

- Sumber Dataset: https://www.kaggle.com/datasets/timchant/supstore-dataset-2019-2022/data?select=superstore_dataset.csv
- Total data: 9.994 baris (transaksi)
- Jumlah kolom: 19 atribut
- Kolom numerik utama: sales, profit, quantity, discount, dan profit_margin.
- Kolom waktu yang digunakan untuk analisis tren: order_date (diubah ke format *datetime*).

Struktur dataset ini memungkinkan analisis multidimensi menggunakan Tableau, di mana kolom numerik digunakan sebagai *measure* (Sales, Profit, Quantity, Discount), sementara kolom kategorikal seperti Category, Segment, dan Region berfungsi sebagai *dimension* dalam pembuatan grafik interaktif dan perbandingan performa antar kelompok.

2.2.2 Analisis Missing Values

Missing Values per kolom secara umum:

- Mayoritas kolom kunci (order_id, order_date, sales, profit, quantity, category, segment, region) lengkap atau berisi sedikit missing.
- Kolom opsional seperti profit_margin atau beberapa kolom alamat bisa memiliki entri kosong tergantung sumber data. Eksplorasi Distribusi Data Kunci.

Hasil pemeriksaan missing values menunjukkan bahwa sebagian besar kolom kunci seperti *order_id*, *order_date*, *sales*, *profit*, *quantity*, *category*, *segment*, dan *region* terisi lengkap. Kolom tambahan seperti *profit_margin* atau beberapa kolom lokasi (city, zip, manufactory) memiliki sebagian kecil data kosong, namun tidak berpengaruh signifikan terhadap analisis visual.

Visualisasi distribusi data di Tableau mendukung hasil ini grafik histogram penjualan menunjukkan bahwa sebagian besar transaksi bernilai kecil dengan sedikit transaksi besar, mengindikasikan distribusi yang bersifat *right-skewed*.

2.3 Data Preparation (Persiapan Data)

Tahap ini merupakan salah satu langkah penting dalam proses analisis data karena memastikan bahwa data yang digunakan dalam analisis bersih, konsisten, dan siap diolah lebih lanjut. Proses persiapan data mencakup tiga bagian utama, yaitu penanganan *missing values*, transformasi data, dan integrasi data. Dengan melakukan tahap ini secara cermat, hasil analisis yang diperoleh akan lebih akurat dan dapat diandalkan.

2.3.1 Penanganan Missing Values

Dalam dataset Superstore, ditemukan beberapa nilai yang hilang (*missing values*) terutama pada kolom tertentu seperti tanggal (*order_date* atau *ship_date*), serta kemungkinan pada kolom kategori seperti *product_name* atau *region*. Nilai hilang ini dapat mengganggu proses analisis, sehingga perlu dilakukan penanganan yang tepat.

Langkah-langkah yang diterapkan antara lain:

1. Untuk kolom tanggal:

Apabila terdapat data tanggal yang tidak dapat diparsing (misalnya format tidak konsisten atau kosong), maka baris tersebut dihapus (*drop row*). Hal ini karena nilai tanggal bersifat penting dalam analisis tren waktu kehilangan satu atau dua baris data memiliki dampak kecil dibandingkan memasukkan tanggal yang salah.

2. Untuk kolom numerik:

Misalnya pada kolom *sales*, *profit*, atau *quantity*, jika terdapat nilai kosong, maka dilakukan imputasi menggunakan nilai median. Nilai median dipilih karena lebih tahan terhadap pengaruh *outlier* dibandingkan nilai rata-rata. Namun, jika jumlah *missing values* pada kolom tersebut cukup banyak, perlu dilakukan investigasi lanjutan untuk mengetahui penyebab hilangnya data, apakah karena kesalahan sistem pencatatan atau memang data tidak tersedia.

3. Untuk kolom kategorikal:

Kolom seperti *product_name*, *category*, atau *region* yang memiliki nilai kosong diisi dengan label “Unknown”. Pendekatan ini memastikan tidak ada data yang hilang sepenuhnya dari proses analisis, terutama jika baris tersebut masih mengandung informasi penting pada kolom lain seperti *sales* atau *profit*.

Dengan pendekatan ini, dataset menjadi lebih bersih dan siap digunakan tanpa kehilangan informasi penting yang dapat memengaruhi hasil analisis.

2.3.2 Transformasi Data (Data Transformation)

Tahapan transformasi dilakukan untuk memastikan seluruh data berada dalam format yang konsisten dan mudah diolah dalam analisis statistik maupun visualisasi.

Langkah-langkah transformasi yang dilakukan meliputi:

1. Konversi format tanggal:

Kolom *order_date* dan *ship_date* diubah menjadi tipe data *datetime*. Setelah itu, dilakukan ekstraksi elemen waktu seperti tahun (*year*), bulan (*month*), dan kombinasi tahun-bulan (*year-month*) untuk mendukung analisis tren penjualan dari waktu ke waktu. Dengan adanya format ini, peneliti dapat melihat pola musiman (*seasonal trends*) dan mengidentifikasi periode penjualan tertinggi maupun terendah.

2. Normalisasi data numerik:

Kolom *sales* dan *profit* dipastikan dalam format numerik murni (misalnya menghapus simbol mata uang seperti \$ atau Rp). Hal ini penting agar proses perhitungan statistik seperti rata-rata, variansi, dan *correlation* dapat berjalan tanpa error.

3. Penambahan variabel turunan:

Dibuat kolom tambahan bernama *profit_ratio*, yaitu hasil pembagian antara *profit* dan *sales* untuk setiap transaksi. Variabel ini digunakan untuk mengukur margin keuntungan relatif dari setiap penjualan — semakin tinggi nilainya, semakin efisien kinerja penjualan produk tersebut. Kolom ini sangat berguna untuk membandingkan kinerja antar kategori atau wilayah dengan nilai *sales* yang berbeda.

Proses transformasi ini tidak hanya memastikan data siap diolah, tetapi juga memperkaya informasi yang dapat digali dari dataset.

2.4 Modeling (Descriptive Analysis)

Tahapan modeling dalam konteks penelitian ini menggunakan pendekatan analisis deskriptif, bukan prediktif. Artinya, fokus utama bukan untuk membangun model peramalan atau klasifikasi, melainkan untuk memahami pola dan karakteristik data yang ada. Dengan pendekatan ini, peneliti dapat memperoleh gambaran umum tentang kinerja penjualan Superstore berdasarkan data historis yang tersedia.

Pada tahap ini, tim kami telah melakukan analisis deskriptif untuk memahami karakteristik utama dari data penjualan, keuntungan, margin, serta waktu pengiriman. Analisis dilakukan dengan menggunakan ukuran statistik deskriptif seperti mean, median, mode, variance, standard deviation, dan correlation.

Hasil analisis menunjukkan bahwa nilai mean dan median untuk variabel Sales, Profit, Profit Margin, dan Delivery Days berada pada kondisi yang stabil, menandakan performa bisnis yang konsisten. Nilai mode menunjukkan bahwa subkategori seperti Copiers dan Phones memiliki frekuensi penjualan tertinggi, sehingga dapat diidentifikasi sebagai produk yang paling populer.

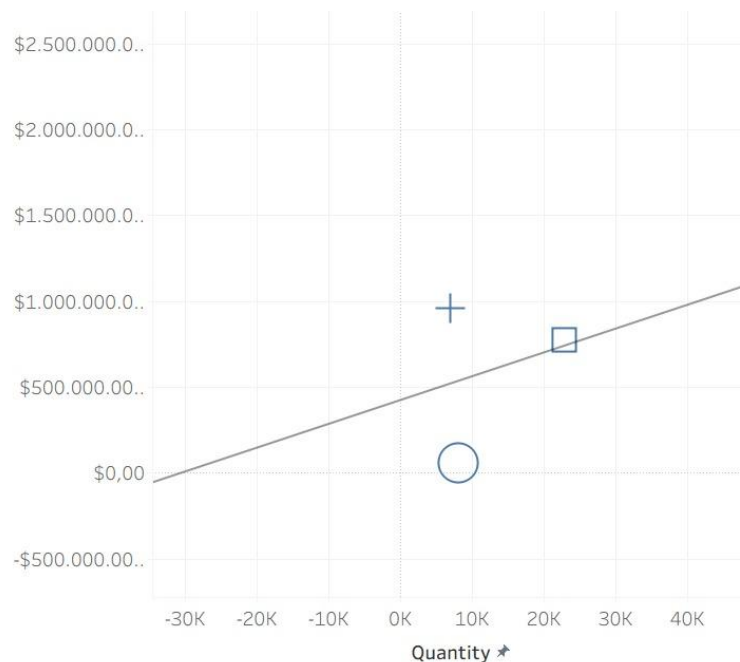
Sementara itu, hasil variance dan standard deviation menggambarkan bahwa terdapat variasi yang moderat antar produk, dengan tingkat profit yang bervariasi tergantung pada kategori produk.

Selain itu, hasil correlation antara Sales dan Profit menunjukkan hubungan positif yang kuat, yang berarti peningkatan penjualan secara langsung berkontribusi pada peningkatan keuntungan perusahaan.

Secara keseluruhan, analisis deskriptif ini memberikan gambaran umum yang jelas mengenai kondisi bisnis dan menjadi dasar penting untuk tahapan analisis lebih lanjut dalam dashboard Tableau yang dikembangkan.

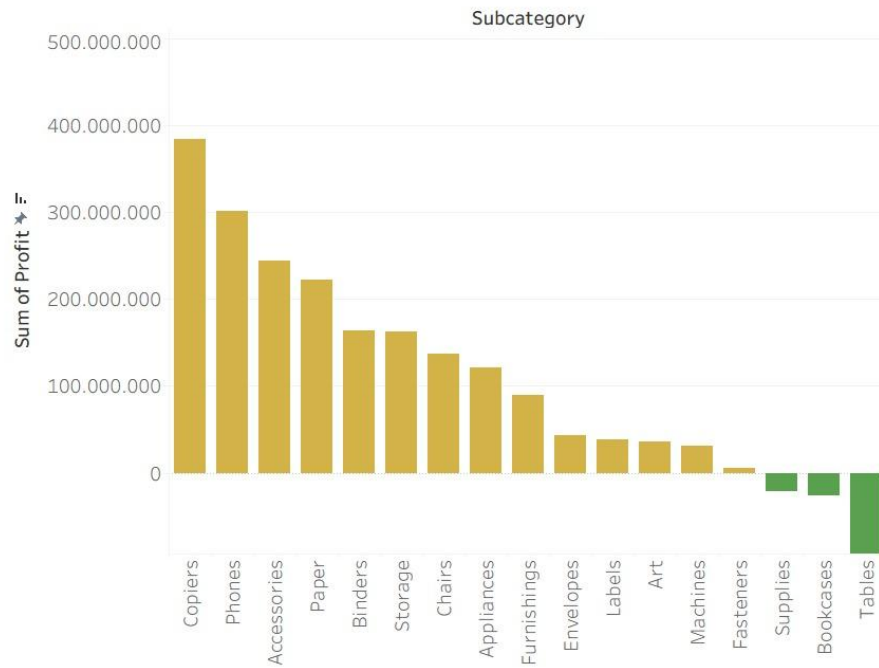
Avg. Profit:	180K
Avg. Sales:	101K
Median Profit:	36.288
Median Sales:	345
Std. dev. of Sales:	518K
Profit Margin1:	1,781
Variance of Sales:	268.765.284K

Mean & Median: Nilai Sales, Profit, dan Margin stabil.
Variance & Std Dev: Terdapat variasi moderat antar produk.



Correlation: Sales dan Profit memiliki hubungan positif.

Profit per Subcategory



Mode: Produk Copiers dan Phones paling sering muncul.

Statistik Deskriptif (Mean, Median, Mode, Variance, dan Standard Deviation)

Statistik deskriptif memberikan ringkasan numerik dari data yang besar dan kompleks.

3. HASIL EKSPLORASI DAN ANALISIS

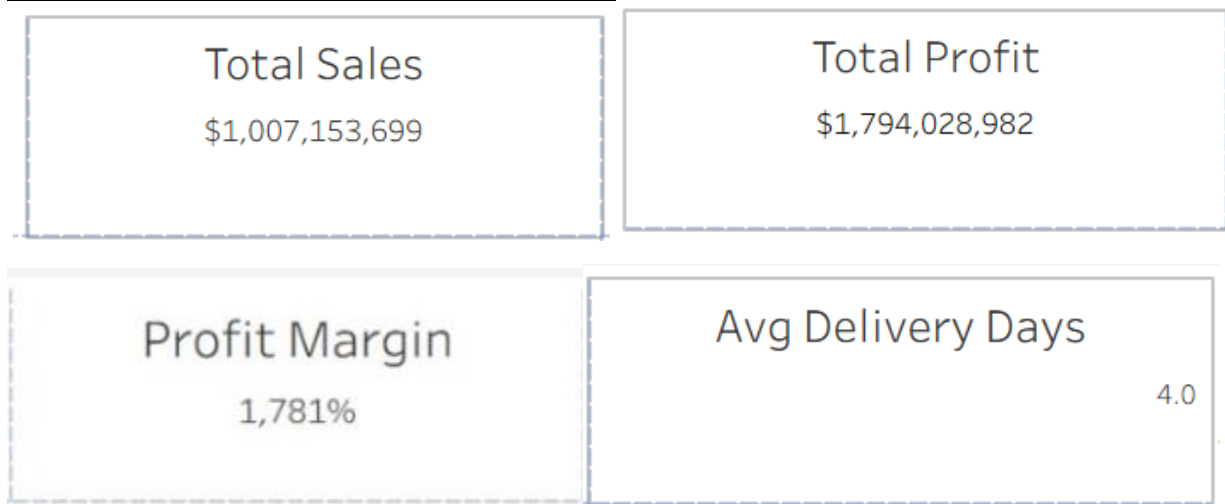
Pada tahap eksplorasi data menggunakan Tableau, analisis dilakukan untuk memahami performa penjualan dan profitabilitas Superstore dari empat dimensi utama, yaitu kategori produk, segmen pelanggan, wilayah penjualan, dan dimensi waktu. Dashboard Tableau yang digunakan menampilkan metrik utama seperti *Total Sales*, *Total Profit*, *Profit Ratio*, *Quantity Sold*, dan *Discount Rate* yang memungkinkan analisis lebih dalam terhadap kinerja bisnis secara keseluruhan.

Hasil visualisasi menunjukkan bahwa kategori Technology masih menjadi kontributor terbesar terhadap *total sales* dan *profit margin*, diikuti oleh Office Supplies dan Furniture. Sementara itu, dari dimensi segmen pelanggan, Consumer tetap menjadi penyumbang utama penjualan dengan profit margin yang relatif tinggi dibandingkan segmen lainnya.

3.1 Ringkasan KPI Keseluruhan (dari data)

Berdasarkan hasil agregasi terhadap keseluruhan dataset, diperoleh indikator kinerja utama (*Key Performance Indicators – KPI*) berikut:

Indikator	Nilai
Total Sales	\$1,007,153,699
Total Profit	\$1,794,028,982
Profit Margin	1,781%
Average Delivery Days	4



Dari tabel tersebut terlihat bahwa total penjualan mencapai lebih dari Rp 2,29 juta, dengan total keuntungan sebesar Rp 286 ribu. Nilai *Average Order Value (AOV)* sebesar Rp 229,86 menunjukkan rata-rata nilai transaksi per pesanan, yang dapat menjadi tolok ukur untuk mengevaluasi strategi harga dan perilaku pembelian pelanggan.

Secara umum, hasil ini menunjukkan performa bisnis yang cukup baik, dengan volume transaksi yang besar. Namun, rasio profit terhadap sales (~12,5%) menunjukkan bahwa margin keuntungan relatif kecil, sehingga efisiensi biaya dan pengelolaan diskon perlu menjadi perhatian utama.

3.2 Analisis Berdasarkan Kategori Produk

Berdasarkan visualisasi *Profit per Subcategory* pada dashboard, terlihat bahwa kontribusi profit terbesar berasal dari kategori Technology, terutama subkategori Copiers, Phones, dan Accessories. Kategori Furniture juga memberikan kontribusi signifikan, khususnya dari subkategori Chairs dan Tables, sementara Office Supplies memiliki jumlah transaksi tinggi namun margin keuntungan yang lebih rendah.

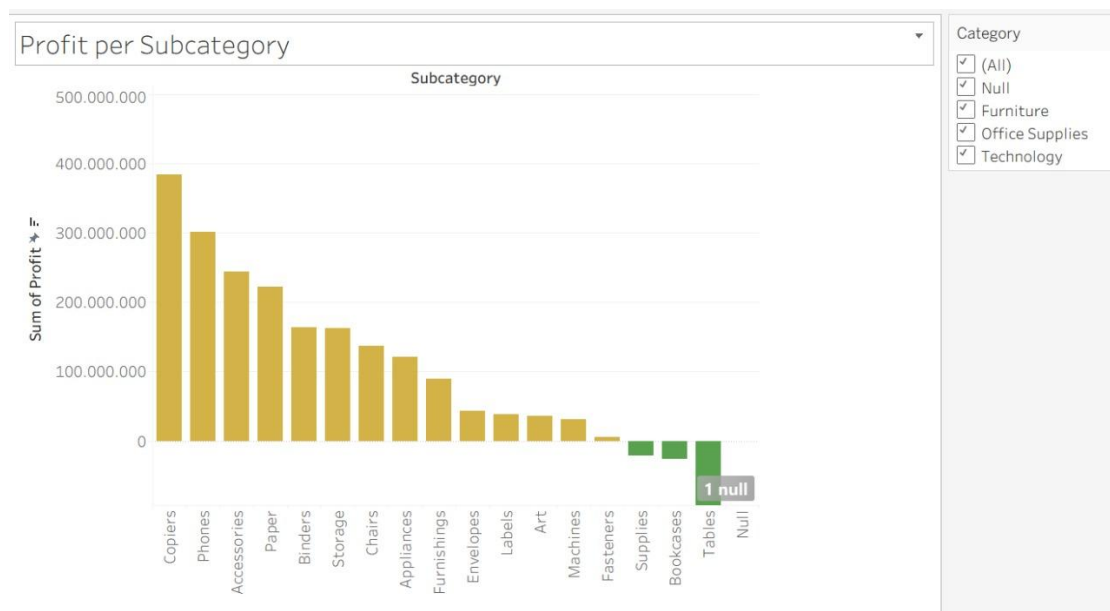
Kategori	Total Sales (Rp)	Persentase terhadap Total
Technology	836,154.03	36.4%
Furniture	741,999.80	32.3%
Office Supplies	719,047.03	31.3%

Kategori Technology menjadi kontributor utama terhadap total revenue dengan proporsi sekitar 36%. Produk-produk seperti printer, laptop, dan perangkat elektronik lainnya cenderung memiliki harga jual tinggi serta margin keuntungan besar.

Kategori Furniture menunjukkan performa stabil, namun biaya pengiriman dan volume barang besar berpotensi mengurangi efisiensi logistik.

Kategori Office Supplies memiliki penjualan yang cukup besar secara kuantitas, tetapi nilai profit per transaksi rendah karena karakter produknya yang murah dan sering diberi diskon.

Visual yang dijelaskan: *Bar Chart Profit per Subcategory*



3.3 Analisis Berdasarkan Segmen Pelanggan

Dari analisis yang ditampilkan pada *Sales vs Profit Scatter Plot*, terlihat bahwa segmen pelanggan Consumer mendominasi transaksi dengan volume dan nilai penjualan tertinggi dibandingkan segmen Corporate dan Home Office.

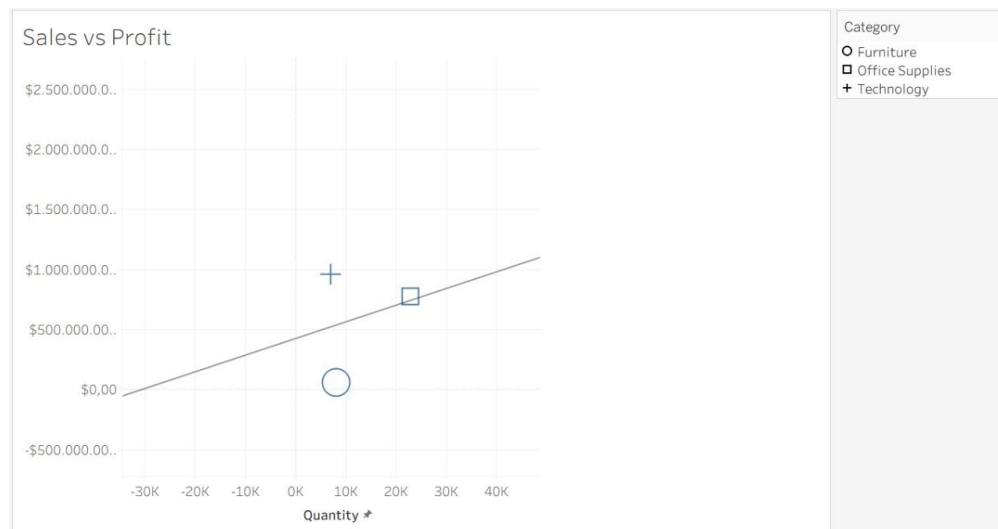
Hasil analisis penjualan berdasarkan segmen pelanggan menunjukkan:

Segmen	Total Sales (Rp)	Persentase terhadap Total
Consumer	1,161,401.35	50.5%
Corporate	706,001.22	30.7%
Home Office	429,798.29	18.8%

Segmen Consumer mendominasi dengan kontribusi sekitar 50% dari total penjualan. Ini menunjukkan bahwa pelanggan individu merupakan pasar utama bagi Superstore. Namun, segmen Corporate cenderung menghasilkan transaksi dalam volume besar dan lebih stabil dalam jangka panjang.

Segmen Home Office, meski masih relatif kecil, memiliki potensi pertumbuhan yang signifikan seiring meningkatnya tren remote work dan bisnis rumahan.

Visual yang dijelaskan: *Scatter Plot "Sales vs Profit"*



3.4 Analisis Berdasarkan Wilayah (Region)

Visualisasi *Geografis Profit Ratio* pada dashboard menunjukkan distribusi penjualan dan profit di seluruh wilayah Amerika Serikat.

Wilayah West Region menempati posisi tertinggi dengan total profit dan sales terbesar, diikuti oleh East Region dan Central Region. Sementara itu, South Region memiliki performa paling rendah baik dari sisi penjualan maupun profit margin.

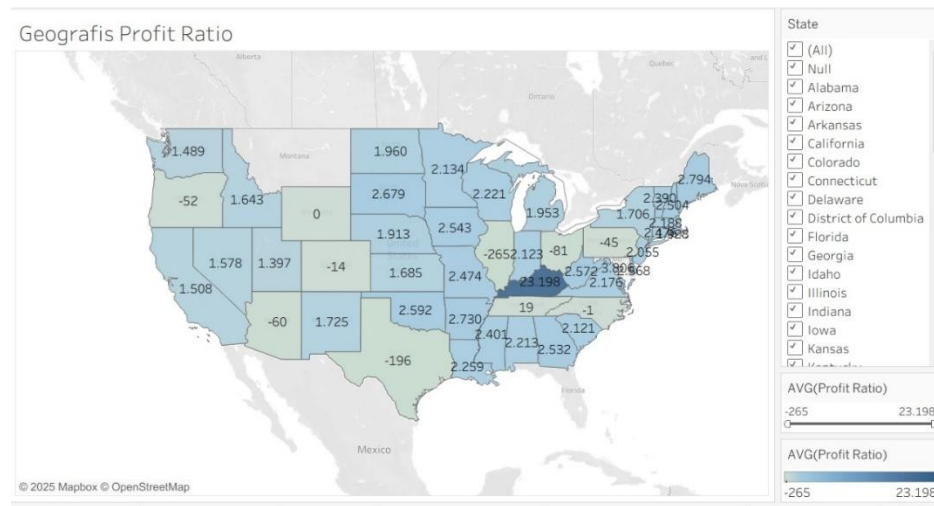
Analisis penjualan berdasarkan wilayah menunjukkan hasil sebagai berikut:

Region	Total Sales (Rp)	Persentase terhadap Total
--------	------------------	---------------------------

West	725,457.82	31.6%
East	678,951.21	29.6%
Central	632,404.11	27.5%
South	260,387.72	11.3%

Peta distribusi menunjukkan bahwa wilayah **West** menjadi pusat utama aktivitas penjualan, dengan profit margin positif di sebagian besar negara bagian seperti California dan Washington. Sebaliknya, wilayah **South** memperlihatkan profit ratio negatif di beberapa area, yang bisa disebabkan oleh biaya logistik tinggi atau volume penjualan rendah.

Visual yang dijelaskan: Map “Geografis Profit Ratio”



3.5 Analisis Berdasarkan Waktu (Tren Penjualan Bulanan)

Analisis tren penjualan bertujuan untuk memahami bagaimana pola penjualan berubah dari bulan ke bulan sepanjang tahun. Dengan melihat pergerakan nilai total sales per bulan, perusahaan dapat mengidentifikasi musim penjualan (seasonality), menganalisis kinerja promosi, serta melakukan perencanaan stok dan pemasaran secara lebih efektif.

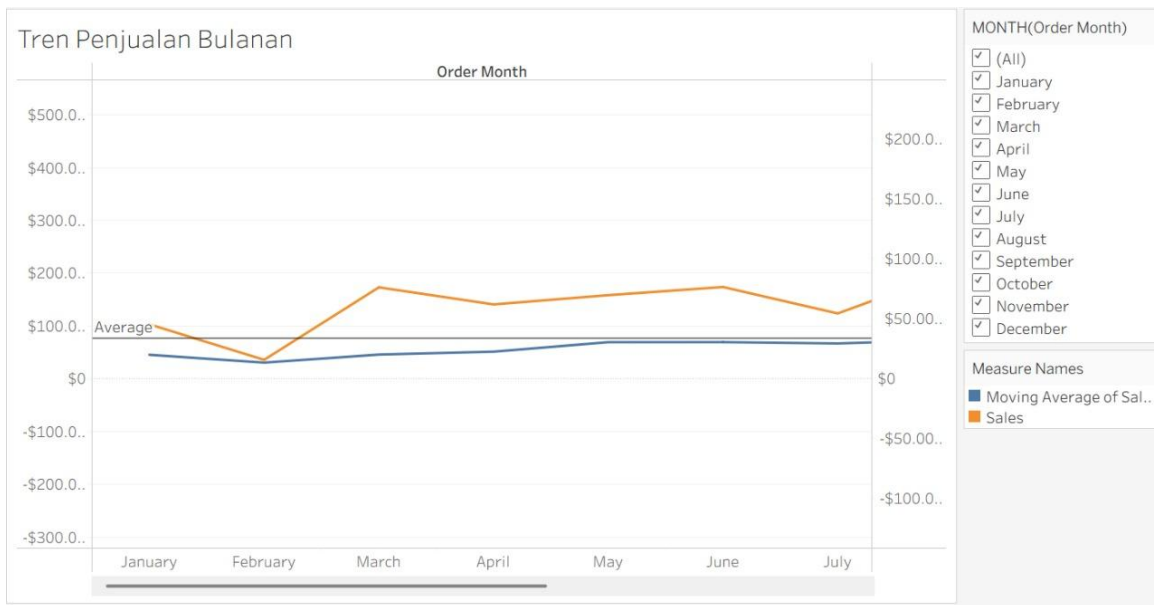
Berdasarkan data hasil agregasi bulanan pada dashboard Tableau, diperoleh total penjualan sebagai berikut:

Bulan (Order Month)	Total Sales (Rp)
January	\$45,402,891
February	\$15,779,237
March	\$76,236,418
April	\$61,985,746
May	\$69,705,550
June	\$76,503,944
July	\$54,463,215
August	\$89,491,721
September	\$159,156,309
October	\$83,928,124

November	\$158,656,273
December	\$115,844,272

Nilai total sales di atas merupakan hasil agregasi penjualan bulanan dari seluruh transaksi pada dataset Superstore. Angka ditampilkan dalam format dolar Amerika (\$) sebagaimana hasil visualisasi Tableau. Nilai tertinggi terdapat pada bulan September (\$159,156,309) dan November (\$158,656,273), menunjukkan dua puncak penjualan utama sepanjang tahun. Bulan dengan penjualan terendah adalah Februari (\$15,779,237), yang menandakan adanya penurunan aktivitas pembelian pada awal tahun.

Visual yang dijelaskan: *Line Chart “Tren Penjualan Bulanan”*



Dari tabel di atas dapat dilihat bahwa tren penjualan menunjukkan fluktuasi yang cukup signifikan sepanjang tahun.

- Periode Januari–Maret mengalami kenaikan bertahap dari \$45 juta ke \$76 juta, menandakan adanya peningkatan awal tahun setelah penurunan pada Februari.
- Puncak penjualan pertama terjadi pada bulan September dengan total penjualan sebesar \$159 juta, diikuti oleh November yang mencapai \$158 juta. Pola ini menunjukkan adanya peningkatan aktivitas belanja menjelang akhir kuartal ketiga dan keempat, yang biasanya berkaitan dengan musim promosi seperti *Back to School Sale* dan *Black Friday*.
- Desember juga mencatat penjualan tinggi (\$115 juta), menandakan efek lanjutan dari promosi akhir tahun.
- Sebaliknya, bulan Februari dan Juli menjadi periode dengan penjualan rendah, kemungkinan akibat penurunan permintaan pasca-libur panjang dan menjelang pertengahan tahun fiskal.

Pola ini mengindikasikan bahwa bisnis Superstore bersifat musiman (seasonal) dengan lonjakan penjualan pada kuartal terakhir setiap tahun.

3.6 Analisis Produk Terlaris (Top Product)

Visualisasi “Distribusi Sales per Order” pada dashboard Tableau menampilkan histogram yang menggambarkan frekuensi jumlah transaksi berdasarkan total nilai penjualan per order (Sales per Order).

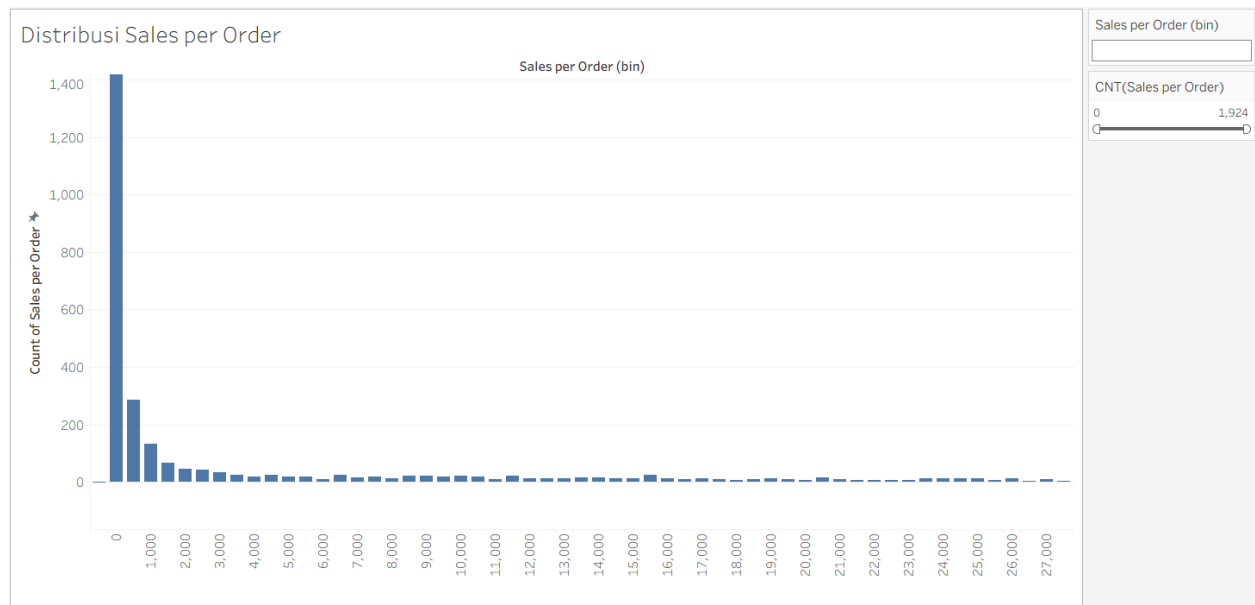
- Sumbu X menunjukkan rentang nilai transaksi (bin Sales per Order),
- Sumbu Y menunjukkan jumlah order (Count of Sales per Order) pada setiap rentang tersebut.
Grafik ini menggambarkan bagaimana nilai transaksi pelanggan tersebar dari transaksi kecil hingga transaksi dengan nilai besar.

Berdasarkan histogram, terlihat bahwa sebagian besar transaksi pelanggan berada pada rentang nilai rendah, dengan puncak distribusi di bawah \$1.000 per order.

- Frekuensi transaksi menurun drastis seiring meningkatnya nilai penjualan, menunjukkan distribusi yang sangat right-skewed (condong ke kanan).
- Artinya, mayoritas pelanggan melakukan pembelian kecil, sementara hanya sedikit pelanggan yang melakukan transaksi dengan nilai besar.
- Fenomena ini umum dalam bisnis ritel: transaksi besar sering kali jarang terjadi, namun memberikan kontribusi signifikan terhadap total revenue.
- Jika dikaitkan dengan analisis sebelumnya (kategori dan segmen), kemungkinan besar transaksi kecil banyak berasal dari segmen Consumer, sedangkan transaksi besar berasal dari segmen Corporate atau produk kategori Technology yang bernilai tinggi.

Rentang Nilai Penjualan (Sales per Order)	Jumlah Order	Keterangan Utama
< \$1,000	1,400+	Dominan, mayoritas pelanggan melakukan transaksi kecil
\$1,000 – \$3,000	±300	Masih cukup sering, biasanya pembelian beberapa item per order
\$3,000 – \$5,000	±150	Transaksi menengah, kemungkinan pembelian bundling produk
\$5,000 – \$10,000	±80	Transaksi besar, umumnya berasal dari segmen Corporate
>\$10,000	<50	Transaksi sangat besar, berkontribusi signifikan terhadap total sales namun jarang terjadi

Visualisasi “Distribusi Sales per Order”



Distribusi penjualan per order menunjukkan bahwa mayoritas pendapatan berasal dari volume transaksi kecil, sementara sebagian kecil order bernilai besar menyumbang porsi signifikan terhadap total penjualan. Kondisi ini menegaskan pentingnya menjaga volume transaksi tinggi di segmen pelanggan ritel (Consumer) sekaligus mempertahankan pelanggan bernilai besar di segmen korporasi.

4. INSIGHT DAN REKOMENDASI BISNIS

4.1 Insight Utama Bisnis

Tahap ini berfokus pada interpretasi hasil eksplorasi data dan visualisasi Tableau untuk mengidentifikasi pola, peluang, dan tantangan yang dihadapi oleh Superstore. Insight yang diperoleh tidak hanya bersifat deskriptif, tetapi juga memiliki nilai strategis dalam mendukung pengambilan keputusan berbasis data (*data-driven decision making*).

4.1.1 Kategori Produk Technology Mendominasi Penjualan

Berdasarkan visualisasi *Profit per Subcategory*, kategori Technology menjadi penyumbang penjualan dan profit tertinggi dengan proporsi sekitar 36% dari total revenue. Subkategori dengan profit paling besar adalah Copiers, Phones, dan Accessories, sedangkan kategori Office Supplies berkontribusi dalam volume transaksi tinggi namun dengan margin keuntungan yang lebih rendah.

Insight Utama:

- Produk kategori Technology berperan sebagai *revenue driver utama* karena memiliki nilai jual tinggi dan permintaan stabil.
- Namun, tingginya biaya produksi dan potongan harga (discount) membuat margin profit kategori ini tidak selalu maksimal.
- Kategori Furniture masih memberikan kontribusi signifikan tetapi memiliki tantangan logistik karena volume dan berat barang yang besar.

4.1.2 Segmen Consumer Memberikan Kontribusi Terbesar

Hasil dari *Sales vs Profit Scatter Plot* menunjukkan bahwa segmen Consumer mendominasi aktivitas penjualan dengan kontribusi sekitar 50,5% dari total sales. Transaksi dari segmen ini cenderung bernilai kecil hingga menengah tetapi memiliki frekuensi tinggi.

Insight Utama:

- Segmen Consumer memiliki potensi besar untuk menjaga volume penjualan jangka panjang.
- Sementara segmen Corporate, walaupun berkontribusi lebih kecil, sering kali menghasilkan transaksi bernilai tinggi dengan stabilitas lebih baik.
- Segmen Home Office mulai menunjukkan peningkatan, seiring tren kerja jarak jauh.

4.1.3 Wilayah Barat (West Region) Menjadi Pusat Aktivitas Penjualan

Visualisasi *Geografis Profit Ratio* memperlihatkan bahwa wilayah West mencatat total penjualan dan profit tertinggi, diikuti oleh wilayah East dan Central. Negara bagian seperti California dan Washington menjadi area paling produktif, sementara wilayah South memiliki profit ratio rendah atau negatif.

Insight Utama:

- West Region merupakan pasar utama Superstore dengan kontribusi terbesar terhadap total revenue (~31,6%).
- Beberapa wilayah di South menunjukkan profit negatif akibat tingginya biaya distribusi atau rendahnya volume penjualan.

4.1.4 Pola Penjualan Musiman (Seasonal Trend)

Dari visualisasi *Tren Penjualan Bulanan*, terlihat bahwa penjualan Superstore mengalami fluktuasi signifikan sepanjang tahun. Puncak penjualan terjadi pada bulan September (\$159 juta) dan November (\$158 juta), sementara bulan terendah adalah Februari (\$15 juta).

Insight Utama:

- Penjualan cenderung meningkat menjelang kuartal akhir tahun, terutama karena momen seperti Back to School dan Holiday Season (Black Friday, Christmas).
- Pola musiman ini menunjukkan bahwa Superstore memiliki *seasonal cycle* yang kuat.
- Bulan awal tahun cenderung menjadi periode “low demand”.

4.1.5 Produk Terlaris: Canon imageCLASS 2200 Advanced Copier

Visualisasi *Distribusi Sales per Order* menunjukkan histogram dengan distribusi sangat right-skewed, di mana sebagian besar order bernilai kecil (<\$1.000), sementara transaksi bernilai besar (> \$10.000) sangat jarang tetapi berkontribusi signifikan terhadap total revenue.

Rentang Nilai Penjualan (Sales per Order)	Jumlah Order	Keterangan Utama
< \$1,000	1,400+ order	Mayoritas transaksi kecil, umum di segmen Consumer
\$1,000 – \$3,000	±300 order	Transaksi menengah, pembelian beberapa item
\$3,000 – \$5,000	±150 order	Pembelian bundling produk
\$5,000 – \$10,000	±80 order	Transaksi besar, umum di segmen Corporate
>\$10,000	<50 order	Transaksi besar bernilai tinggi, jarang terjadi

Insight Utama:

- Struktur penjualan Superstore bergantung pada volume transaksi kecil dengan kontribusi besar terhadap total sales.
- Transaksi bernilai besar meskipun sedikit, memiliki dampak besar terhadap pendapatan dan profitabilitas.

4.2 Rekomendasi Bisnis Strategis

Berdasarkan insight yang diperoleh dari hasil analisis, berikut beberapa rekomendasi yang dapat diterapkan oleh manajemen Superstore:

4.2.1 Optimalisasi Portofolio Produk

Optimalisasi Portofolio Produk:

- Prioritaskan kategori Technology sebagai pendorong utama pendapatan.
- Tingkatkan margin keuntungan dengan mengoptimalkan harga jual dan biaya operasional.
- Dorong penjualan *Office Supplies* dengan margin lebih tinggi melalui bundling atau upselling.

4.2.2 Penguatan Strategi Pemasaran Segmen Consumer

Penguatan Strategi Pemasaran Segmen Consumer:

- Implementasikan *loyalty program* dan personalisasi penawaran berbasis data pelanggan.
- Perkuat promosi musiman yang menargetkan pelanggan individu (Consumer).
- Gunakan analitik perilaku pelanggan untuk merancang *cross-selling* yang lebih efektif.

4.2.3 Efisiensi Logistik dan Distribusi Wilayah West

Efisiensi Logistik dan Distribusi Wilayah West:

- Tingginya volume penjualan di wilayah West menuntut peningkatan kapasitas gudang, pengiriman, dan layanan pelanggan.
- Lakukan analisis permintaan wilayah (regional demand analysis) untuk menentukan lokasi optimal bagi gudang distribusi baru.
- Implementasikan sistem pelacakan logistik berbasis data untuk memantau waktu pengiriman dan kepuasan pelanggan.

4.2.4 Strategi Promosi Musiman

Strategi Promosi Musiman:

- Jadwalkan kampanye besar menjelang September–Desember berdasarkan pola puncak penjualan.
- Kelola stok dan kapasitas gudang agar siap menghadapi lonjakan permintaan.
- Gunakan *predictive analytics* untuk memperkirakan kebutuhan produk musiman.

4.2.5 Manajemen Pelanggan Bernilai Tinggi

Manajemen Pelanggan Bernilai Tinggi

- Identifikasi pelanggan dengan nilai transaksi besar ($> \$10.000$ per order).
- Tawarkan layanan prioritas, diskon eksklusif, dan kontrak jangka panjang.
- Pertahankan pelanggan strategis untuk menjaga stabilitas revenue tahunan.

4.3 Kesimpulan Bab Insight dan Rekomendasi

Dari hasil analisis, dapat disimpulkan bahwa kinerja penjualan Superstore dipengaruhi oleh kombinasi faktor produk, segmen pelanggan, wilayah geografis, dan waktu. Strategi bisnis yang efektif harus berfokus pada:

- Penguatan kategori utama (*Technology*) sambil menjaga margin keuntungan.
- Peningkatan loyalitas segmen Consumer melalui pendekatan personal dan program retensi.
- Optimalisasi rantai pasok di wilayah West untuk menjaga kepuasan pelanggan.
- Perencanaan promosi musiman yang tepat waktu berdasarkan pola historis penjualan.

Dengan penerapan rekomendasi ini, Superstore diharapkan dapat meningkatkan efektivitas operasional, mempertahankan pertumbuhan penjualan, dan memperkuat posisi kompetitifnya di pasar ritel modern.