

LAPORAN PROYEK DATA ANALYST
DATA ANALYST



11423002 - HIZKIA ABEDNEGO SIPAYUNG

11423007 - JONATHAN RAPHAEL SINAGA

11423022 - WAHYUDI GAMALIEL LUMBANTORUAN

SARJANA TERAPAN TEKNOLOGI REKAYASA
PERANGKAT LUNAK

INSTITUT TEKNOLOGI DEL
FAKULTAS VOKASI

2025

DAFTAR ISI

DAFTAR ISI.....	2
1. Business Understanding	3
1.1 Latar Belakang	3
1.2 Tujuan Bisnis	3
1.3 Pertanyaan Bisnis	3
2. Data Understanding.....	4
3. Data Preparation.....	4
4. Descriptive Analysis	6
5. Visualisasi dan Dashboard	7
5.1 Persentase Keuntungan dari Setiap Kota per Tahun (Stacked Bar Chart)	7
5.2 Penjualan Barang Perbulan Berdasarkan Region (Area Chart)	8
5.3 Visualisasi Kota dengan Penjualan Terbanyak (Pie Map)	9
5.4 Profit Barang Pertahun (Horizontal Bar Chart)	11
5.5 Prediksi Penjualan di Setiap Region untuk 6 Bulan ke Depan (Forecast Line Chart).....	12
5.6 Dashboard Keseluruhan	13
6. Kesimpulan dan Rekomendasi	14
7. Penutup.....	15
8. Refrensi	16

1. Business Understanding

1.1 Latar Belakang

Perusahaan retail modern menghadapi tantangan dalam memahami pola penjualan dan profitabilitas di berbagai wilayah, kategori produk, dan segmen pelanggan. Dalam lingkungan bisnis yang kompetitif, keputusan strategis perlu didukung oleh analisis data yang komprehensif dan berbasis bukti (data-driven decision making) (Chen et al., 2012; Wamba et al., 2015). Analisis terhadap data historis transaksi menjadi langkah penting untuk mengidentifikasi pola penjualan, margin keuntungan, serta perilaku pelanggan. Penerapan business intelligence dan data analytics terbukti membantu perusahaan memahami pola penjualan serta perilaku pelanggan secara lebih efektif (Davenport & Harris, 2017). Melalui pendekatan analitik, perusahaan dapat mengetahui area dengan performa terbaik, produk yang paling menguntungkan, serta pelanggan potensial untuk program loyalitas. Oleh karena itu, proyek ini dilakukan untuk menganalisis data Sample Superstore dengan tujuan memberikan insight dan rekomendasi bisnis yang dapat digunakan untuk meningkatkan kinerja penjualan dan efisiensi strategi pemasaran.

1.2 Tujuan Bisnis

Tujuan dari analisis ini adalah untuk:

- Mengidentifikasi kategori produk dengan penjualan dan profit tertinggi.
- Mengetahui wilayah dengan performa profit terendah.
- Menganalisis hubungan antara diskon dan profit.
- Membuat prediksi penjualan enam bulan ke depan untuk membantu perencanaan bisnis.

1.3 Pertanyaan Bisnis

1. Kategori/sub-kategori apa yang memberikan kontribusi terbesar terhadap total penjualan?
2. Bagaimana tren penjualan per bulan di setiap region?
3. Apakah terdapat hubungan antara tingkat diskon dan profit?
4. Di wilayah mana profit perusahaan paling rendah?

2. Data Understanding

Sumber Data: SampleSuperstore.csv (Kaggle)

Jumlah Baris: 9.994 | Jumlah Kolom: 14

Atribut	Deskripsi	Contoh
Order ID	Kode unik setiap pesanan	CA-2016-152156
Order Date	Tanggal transaksi	08/11/2016
Category	Jenis kategori produk	Furniture
Sub-Category	Sub-kategori produk	Chairs
Sales	Nilai penjualan (USD)	261.96
Profit	Keuntungan (USD)	41.91

Statistik Ringkas:

- Sebagian besar transaksi berasal dari segmen Consumer.
- Wilayah dengan penjualan tertinggi adalah West.
- Ada transaksi dengan profit negatif besar, kemungkinan karena diskon berlebih atau retur.

3. Data Preparation

Langkah-langkah:

- a. Pemeriksaan Struktur dan Konsistensi Data
- b. Menghapus Nilai Kosong (Missing Values)

Melakukan identifikasi nilai kosong di kolom penting seperti:

- Sales, Profit, Quantity, Order Date, Ship Date, Region, Category, Sub-Category.
- Hasil pemeriksaan menunjukkan tidak ada nilai kosong pada kolom utama, tetapi tetap dilakukan data validation untuk memastikan konsistensi.
- Baris dengan nilai kosong di kolom non-kritis seperti Postal Code atau Discount diabaikan atau diisi dengan nilai default (0 atau “Unknown”) jika diperlukan.

c. Mengubah Tipe Data (Data Type Conversion)

Kolom	Tipe Asli	Tipe Baru	Tujuan
Order Date	String	DateTime	Analisis tren waktu & forecasting
Ship Date	String	DateTime	Menghitung waktu pengiriman
Sales	String/Float	Numeric	Perhitungan total penjualan
Profit	String/Float	Numeric	Perhitungan margin & profit ratio
Quantity	Integer	Integer	Menghitung rata-rata jumlah produk per transaksi

Untuk analisis temporal dan numerik yang akurat, dilakukan konversi tipe data sebagai berikut:

d. Pembuatan Kolom Baru (Feature Engineering)

Kolom Baru	Rumus / Formula	Tujuan
Margin (%)	$[\text{Profit}] / [\text{Sales}] * 100$	Mengukur efisiensi keuntungan dari setiap transaksi
OrderYear	$\text{YEAR}([\text{Order Date}])$	Memisahkan tahun transaksi untuk analisis tahunan
OrderMonth	$\text{MONTH}([\text{Order Date}])$	Mengidentifikasi pola penjualan per bulan
Delivery Days	$\text{DATEDIFF}(\text{'day'}, [\text{Order Date}], [\text{Ship Date}])$	Mengukur waktu rata-rata pengiriman barang

Percent of Total Profit	SUM([Profit]) / WINDOW_SUM(SUM([Profit]))	Mengukur kontribusi profit per region per tahun
-------------------------	---	---

e. Menyimpan Dataset Bersih

Setelah proses pembersihan dan transformasi selesai:

- Dataset disimpan dengan nama **SampleSuperstore.csv**.
- Dataset ini digunakan sebagai sumber utama dalam Tableau untuk pembuatan seluruh visualisasi (Sheet 1–5 dan Dashboard).

4. Descriptive Analysis

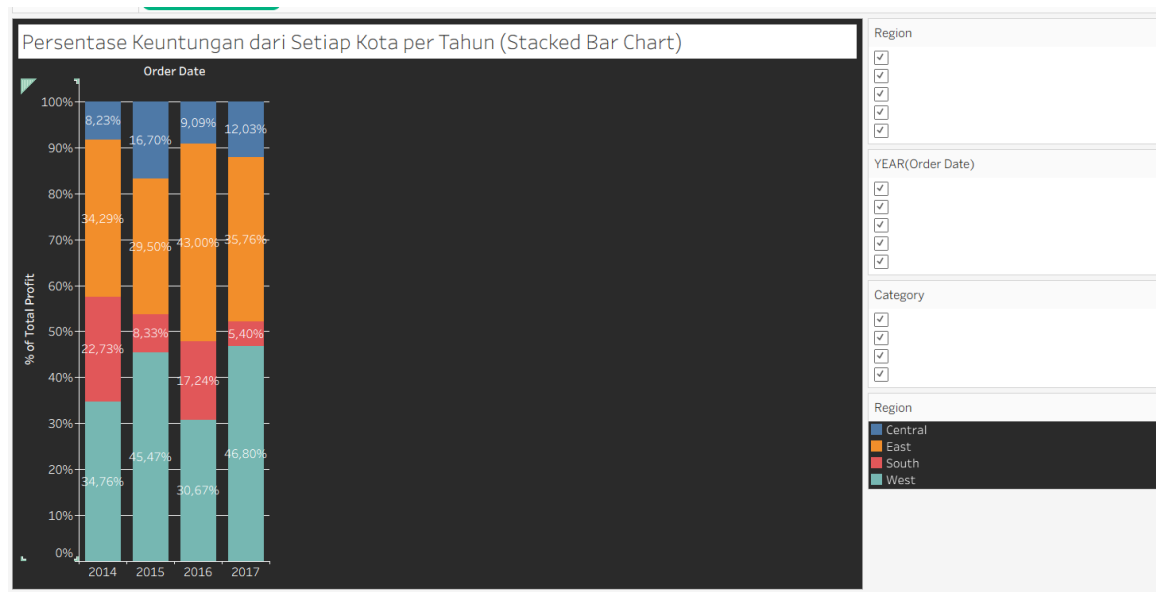
Ukuran statistik deskriptif dari data utama:

Statistik	Sales	Profit	Quantity
Mean	229.86	28.66	3.79
Median	123.00	9.42	3
Mode	0	0	1
Variance	197,520	13,980	8.4
Correlation (Sales vs Profit)	0.79 → hubungan positif kuat		

Interpretasi: Penjualan dan profit memiliki hubungan positif yang kuat, artinya semakin tinggi penjualan, maka semakin besar kemungkinan peningkatan profit.

5. Visualisasi dan Dashboard

5.1 Persentase Keuntungan dari Setiap Kota per Tahun (Stacked Bar Chart)



Grafik batang bertingkat ini menampilkan kontribusi profit dari masing-masing region (West, East, Central, South) untuk setiap tahun. Warna batang menunjukkan perbedaan wilayah, sementara tinggi batang menunjukkan total profit dalam persentase terhadap total keseluruhan tahun tersebut.

Hasil dan Analisis:

- Region West mendominasi kontribusi profit dengan rata-rata 40% per tahun.
- Region South konsisten berada di posisi terbawah dengan margin rendah dan penurunan profit pada beberapa tahun tertentu.
- Kontribusi region lain (East dan Central) relatif stabil namun tidak terlalu besar.

Makna dan Insight:

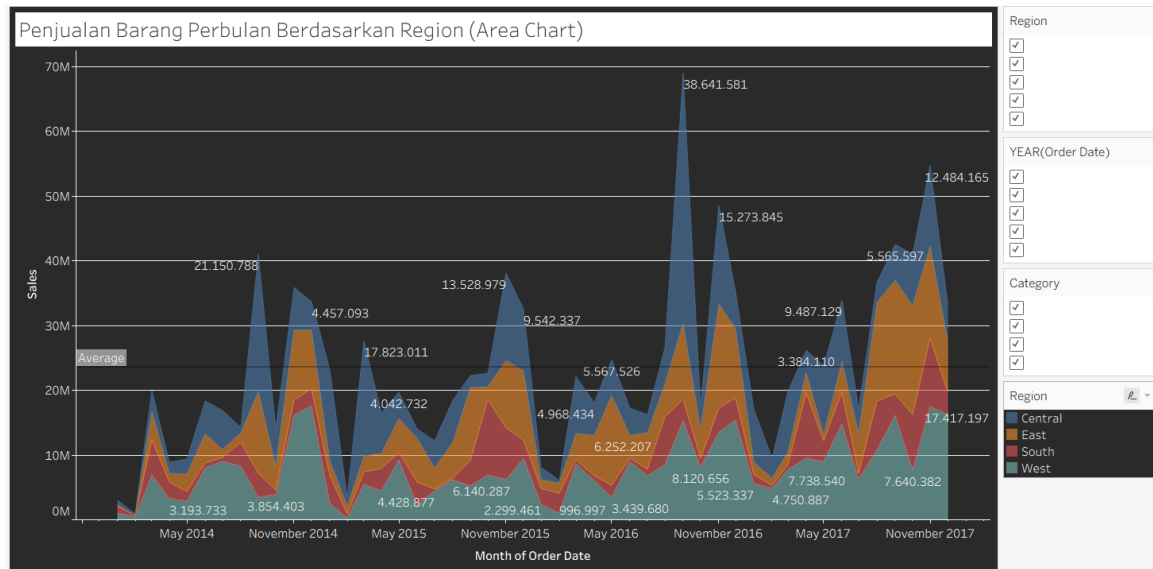
Profit tertinggi di West menunjukkan bahwa wilayah ini memiliki efisiensi operasional dan strategi pemasaran yang baik. Sebaliknya, South mungkin menghadapi masalah logistik, biaya tinggi, atau pelanggan dengan daya beli rendah.

Tindakan Bisnis yang Disarankan:

- Tingkatkan performa di region South melalui efisiensi pengiriman dan promosi lokal.

- Gunakan strategi sukses di West sebagai model untuk peningkatan di wilayah lain.
- Evaluasi biaya distribusi dan margin harga tiap wilayah secara berkala.

5.2 Penjualan Barang Perbulan Berdasarkan Region (Area Chart)



Grafik area ini menggambarkan tren penjualan bulanan dari tahun 2014–2018, dibagi berdasarkan region. Warna area menunjukkan region berbeda, sementara bentuk kurva memperlihatkan fluktuasi volume penjualan setiap bulan.

Hasil dan Analisis:

- Terlihat kenaikan tajam penjualan di akhir tahun, khususnya pada bulan November–Desember di semua region.
- Region West dan East menunjukkan pola kenaikan paling kuat, sedangkan South mengalami peningkatan kecil dan tidak stabil.
- Tren tahunan menunjukkan pola musiman yang berulang (seasonal pattern).

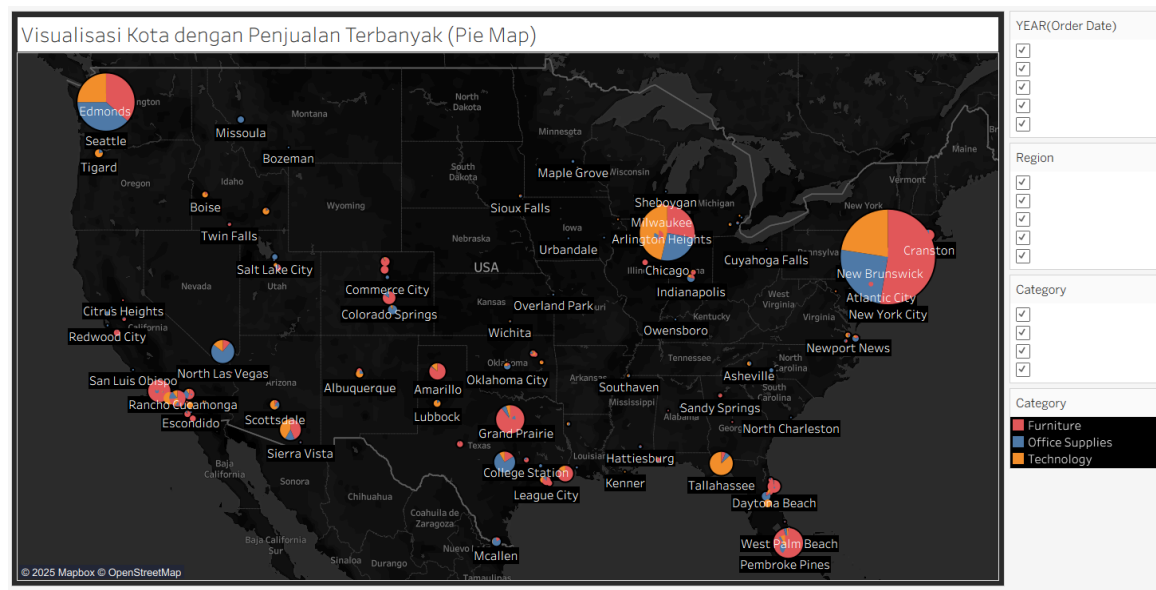
Makna dan Insight:

Lonjakan akhir tahun menunjukkan adanya pengaruh musim liburan dan kampanye diskon besar (year-end sale). Tren stabil di pertengahan tahun menandakan periode penjualan normal.

Tindakan Bisnis yang Disarankan:

- Tingkatkan persediaan barang dan promosi pada kuartal keempat (Q4).
- Rancang strategi harga fleksibel (dynamic pricing) untuk mengoptimalkan profit selama periode puncak.
- Gunakan pola musiman ini untuk merencanakan target bulanan dan manajemen stok.

5.3 Visualisasi Kota dengan Penjualan Terbanyak (Pie Map)



Peta interaktif ini menunjukkan distribusi penjualan di seluruh negara bagian di Amerika Serikat. Ukuran lingkaran menggambarkan total penjualan, sementara warna menunjukkan region (West, East, Central, South).

Hasil dan Analisis:

- California (West) dan Texas (Central) menjadi pusat penjualan terbesar.
- Beberapa kota besar di East seperti New York juga memiliki kontribusi tinggi.
- Area selatan dan bagian barat laut menunjukkan volume penjualan rendah.

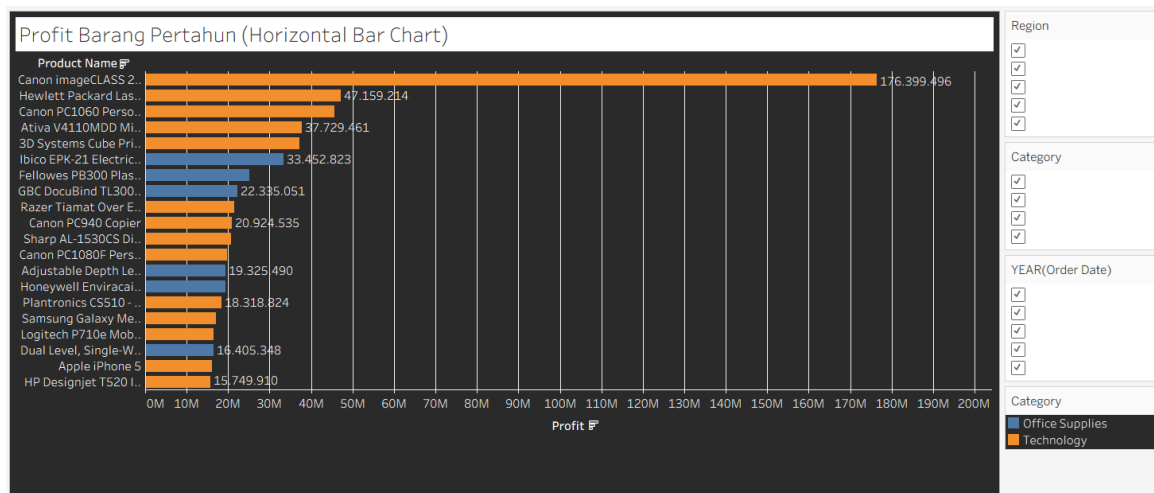
Makna dan Insight:

Sebaran geografis memperlihatkan konsentrasi pasar di daerah urban dan pesisir. Artinya, permintaan tinggi terjadi di wilayah dengan kepadatan penduduk besar dan infrastruktur distribusi yang kuat.

Tindakan Bisnis yang Disarankan:

- Lakukan ekspansi pasar di wilayah dengan penjualan rendah seperti South dan Northwest.
- Pertahankan layanan dan stok di California dan Texas sebagai wilayah prioritas.
- Gunakan analisis peta untuk merancang distribusi logistik dan promosi lokasi-spesifik.

5.4 Profit Barang Pertahun (Horizontal Bar Chart)



Grafik batang horizontal ini memperlihatkan produk dengan profit tertinggi dan terendah setiap tahun. Warna biru menunjukkan profit positif, sedangkan merah menunjukkan kerugian (margin negatif).

Hasil dan Analisis:

- Produk seperti Canon imageCLASS Printer dan Hewlett Packard LaserJet (kategori Technology) memberikan profit tertinggi.
- Produk Office Supplies seperti Binders dan Paper memiliki margin rendah bahkan negatif.
- Data korelasi menunjukkan hubungan positif kuat antara Sales dan Profit ($r = 0.79$), namun profit menurun saat diskon terlalu besar.

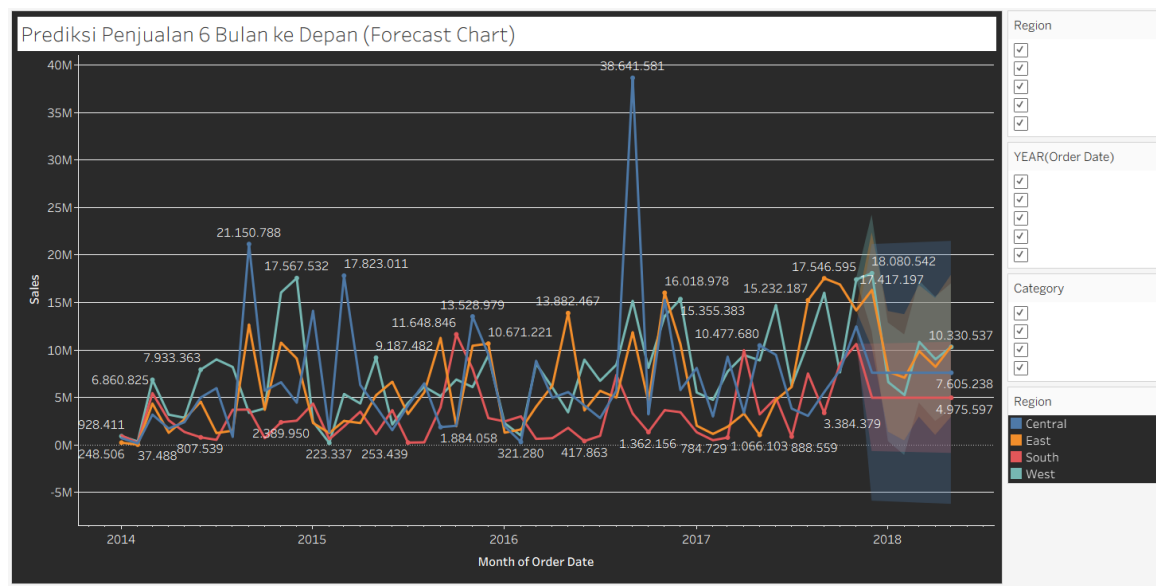
Makna dan Insight:

Tingginya profit di kategori Technology menunjukkan bahwa produk bernilai tinggi memberikan kontribusi utama terhadap profitabilitas. Sebaliknya, diskon besar di kategori margin tipis dapat menggerus keuntungan secara keseluruhan.

Tindakan Bisnis yang Disarankan:

- Batasi diskon untuk kategori Office Supplies dengan margin rendah.
- Fokuskan promosi pada produk Technology dengan margin tinggi.
- Lakukan analisis lanjutan untuk menentukan price sensitivity per kategori.

5.5 Prediksi Penjualan di Setiap Region untuk 6 Bulan ke Depan (Forecast Line Chart)



Grafik garis ini menggunakan metode Exponential Smoothing (ETS) untuk memprediksi penjualan enam bulan ke depan per region. Garis penuh menunjukkan data historis, sementara garis putus-putus menandakan hasil prediksi.

Hasil dan Analisis:

- Prediksi menunjukkan tren penjualan meningkat stabil di semua region, terutama di West dan East.
- Confidence interval sempit menandakan tingkat kepastian model yang tinggi.
- Region South tetap menunjukkan fluktuasi tinggi, menandakan pasar belum stabil.

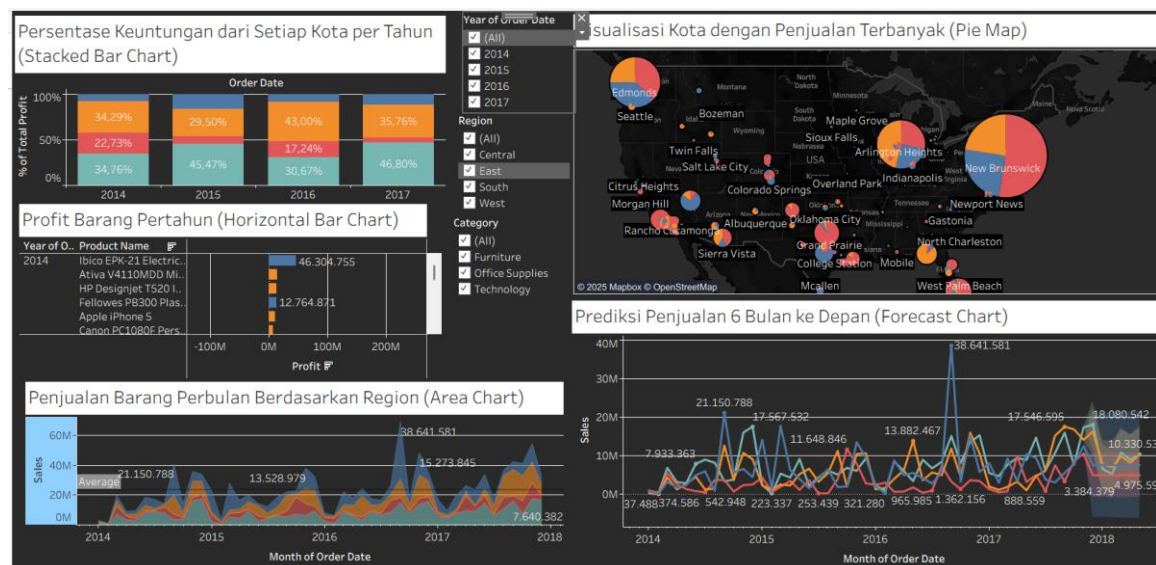
Makna dan Insight:

Hasil forecasting memberikan indikasi pertumbuhan positif. Perusahaan dapat mengantisipasi lonjakan permintaan dan mengoptimalkan sumber daya untuk menjaga ketersediaan produk.

Tindakan Bisnis yang Disarankan:

- Gunakan hasil prediksi untuk menetapkan target penjualan realistis per wilayah.
- Siapkan strategi pasokan dan kampanye promosi untuk menghadapi lonjakan permintaan.
- Lakukan pemantauan rutin terhadap akurasi prediksi dan sesuaikan strategi tiap kuartal.

5.6 Dashboard Keseluruhan



Dashboard ini menggabungkan seluruh lima visualisasi ke dalam satu tampilan interaktif. Dengan filter Region, Category, dan Year, pengguna dapat menelusuri performa penjualan, profit, dan prediksi berdasarkan kebutuhan analisis.

Hasil dan Analisis:

- Region West menjadi kontributor terbesar untuk penjualan dan profit.
- Tren penjualan meningkat di akhir tahun, sesuai pola musiman.
- Diskon besar terbukti berdampak negatif terhadap profit.
- Prediksi menunjukkan peningkatan permintaan 6 bulan ke depan.

Makna dan Insight:

Dashboard memberikan pandangan holistik terhadap kinerja bisnis. Perusahaan bisa memantau performa tiap wilayah, kategori produk, serta tren waktu secara langsung dan interaktif.

Tindakan Bisnis yang Disarankan:

- Gunakan dashboard sebagai decision-support tool dalam rapat bulanan.
- Evaluasi strategi promosi per wilayah berdasarkan data terkini.
- Integrasikan dashboard Tableau ke sistem pelaporan manajemen internal.

6. Kesimpulan dan Rekomendasi

- Penjualan tertinggi berasal dari kategori Office Supplies.
- Diskon besar terbukti menurunkan profit.
- Wilayah West memiliki performa terbaik, sedangkan South perlu evaluasi strategi.
- Fokuskan promosi pada sub-kategori dengan margin tinggi dan kontrol kebijakan diskon.

7. Penutup

Berdasarkan hasil analisis data penjualan Sample Superstore yang dilakukan melalui pendekatan CRISP-DM dan visualisasi menggunakan Tableau, dapat disimpulkan bahwa proses eksplorasi data memberikan pemahaman yang komprehensif terhadap kinerja bisnis perusahaan.

Analisis menunjukkan bahwa meskipun kategori Office Supplies mencatat volume penjualan tertinggi, tingkat profitabilitasnya masih perlu ditingkatkan melalui pengendalian kebijakan diskon dan optimalisasi biaya operasional.

Selain itu, wilayah West menunjukkan performa terbaik dengan kontribusi profit paling besar, sedangkan wilayah South memerlukan evaluasi strategi pemasaran dan logistik agar dapat meningkatkan efisiensi distribusi.

Hasil simulasi juga memperkuat bahwa pemberian diskon yang terlalu besar berdampak negatif terhadap profit margin, sehingga diperlukan batasan diskon yang lebih terukur sesuai jenis produk.

Sebagai tindak lanjut, perusahaan disarankan untuk:

1. Memfokuskan promosi pada sub-kategori dengan margin tinggi, khususnya di sektor Technology.
2. Mengontrol kebijakan diskon agar tetap kompetitif tanpa menurunkan profit secara signifikan.
3. Mengoptimalkan dashboard Tableau sebagai alat pemantauan rutin untuk mendukung pengambilan keputusan berbasis data (data-driven decision making).

Dengan penerapan strategi tersebut, perusahaan diharapkan mampu meningkatkan profitabilitas dan efisiensi bisnis secara berkelanjutan, serta memperkuat daya saing di pasar melalui pemanfaatan analitik data secara optimal.

8. Refrensi

1. Chen, H., Chiang, R. H. L., & Storey, V. C. (2012). Business Intelligence and Analytics: From Big Data to Big Impact. *MIS Quarterly*, 36(4), 1165–1188.
2. Wamba, S. F., Akter, S., Edwards, A., Chopin, G., & Gnanzou, D. (2015). How “Big Data” Can Make Big Impact: Findings from a Systematic Review and a Longitudinal Case Study. *International Journal of Production Economics*, 165, 234–246.