

Sales Order Analysis: Data Analytics Project

Project Overview

This project delivers a comprehensive analysis of sales performance, transforming raw transactional data into high-level business intelligence. By integrating Python for data cleaning, SQL for complex business querying, and Power BI for visualization, the project identifies revenue-generating trends, regional performance gaps, and Year-over-Year (YoY) growth metrics to support data-driven decision-making.

Dataset Summary

The project utilizes a comprehensive sales dataset (orders.csv) containing 9,994 records with 16 attributes.

- Order Attributes: Order ID, Order Date, Ship Mode, and Segment.
- Geographic Data: Country, City, State, and Region.
- Product Details: Category (Furniture, Office Supplies, Technology), Sub-Category & Product ID
- Financial Metrics: Cost Price, List Price, Quantity, and Discount Percent.
- Calculated Fields: Sale Price, Profit, and Revenue.

These fields enable analysis of revenue trends, profitability, product performance, and seasonal sales behavior.

Tools & Technologies

- Python (Pandas, NumPy, Matplotlib/Seaborn): Data loading, cleaning, and EDA
- SQL (PostgreSQL / MySQL / SQL Server): Business analysis and advanced queries
- Power BI: Interactive dashboards and KPI reporting
- Jupyter Notebook: Analysis environment
- Gamma: Used for creating AI-powered slides presentation (PPT)

Project steps & Workflow

1. Exploratory Data Analysis (Python)

Using the Sales_Order_Analysis_.ipynb notebook, the raw data underwent a rigorous ETL process:

- Data Ingestion: Loaded raw CSV data using Pandas

- Data Cleaning: Standardized column names to snake case (lowercase with underscores), handled null values, and converted date strings to datetime objects to optimized data types.
- Feature Engineering: Calculated new metrics such as Discount Amount, Sale Price, and Profit for each transaction.

2. SQL-Based Business Analysis

The cleaned data migrated to a SQL environment to solve critical business problems:

- Revenue Leaders: Top 10 products contributing to the highest revenue.
- Regional Performance: Top 5 selling products within each specific geographic region.
- Time-Series Analysis: Month-over-Month (MoM) growth comparison between 2022 and 2023.
- Category Leaders: Determined which month had the highest sales for each product category.
- Profit Growth: Identified sub-categories with the highest profit growth in 2023 compared to 2022.

Here are some answers to specific business questions, including:

1. Total Revenue by Category

category	total_revenue
Technology	3934910.7
Furniture	3720742.4
Office Supplies	3423675.1

2. Top 10 Products by Revenue

product_id	total_revenue
TEC-CO-10004722	245056
OFF-BI-10000545	163777.7
TEC-MA-10002412	130406.4
FUR-CH-10002024	120090.7
TEC-PH-10001459	113041.9
TEC-CO-10001449	107388
OFF-BI-10003527	97082.9
TEC-MA-10000822	89622.3
FUR-BO-10002213	84014.8
TEC-MA-10001047	81549

3. Average Revenue by Shipping Mode

ship_mode	avg_revenue
Standard Class	1112.706105
First Class	1018.990832

4. Top 5 Products by Revenue (With Discounts Applied)

product_id	total_discounted_revenue
TEC-CO-10004722	245056
OFF-BI-10000545	163777.7
TEC-MA-10002412	130406.4
FUR-CH-10002024	120090.7
TEC-PH-10001459	113041.9

5. Top 3 Most Profitable Products per Category

category	product_id	total_profit
Furniture	FUR-CH-10002024	2246.2
Furniture	FUR-BO-10004834	1614.1
Furniture	FUR-TA-10000198	1229.7
Office Supplies	OFF-BI-10003527	3435.3
Office Supplies	OFF-BI-10001359	2080.2
Office Supplies	OFF-BI-10000545	1959
Technology	TEC-CO-10004722	5644
Technology	TEC-MA-10002412	3624.4
Technology	TEC-CO-10001449	2631.2

6. Top 5 highest profit generating products

product_id	total_profit
TEC-CO-10004722	5644
TEC-MA-10002412	3624.4
OFF-BI-10003527	3435.3
TEC-CO-10001449	2631.2
FUR-CH-10002024	2246.2

7. Month-over-Month Sales Comparison (2022 vs 2023)

order_month	sales_2022	sales_2023
1	94712.5	88632.6
2	90091	128124.2
3	80106	82512.3
4	95451.6	111568.6
5	79448.3	86447.9
6	94170.5	68976.5
7	78652.2	90563.8
8	104808	87733.6
9	79142.2	76658.6
10	118912.7	121061.5
11	84225.3	75432.8
12	95869.9	102556.1

3. Power BI Dashboard Overview

The Sales Performance Dashboard.pbix provides an interactive experience for stakeholders to explore the data:

- KPI Scorecard: Real-time tracking of Total Sales, Total Profit, and Quantity Sold.
- Category Drill-down: Comparison of performance across Furniture, Office Supplies, and Technology.
- Geographic Mapping: A visual representation of sales density across the United States.

Key Results & Recommendations

- High-Growth Sub-Categories: "Machines" exhibited the highest profit growth at 50.19%. Recommend increasing inventory and marketing spend for this category.
- Profit Optimization: Some categories like "Paper" showed a slight negative growth (-3.81%). A review of supplier costs or pricing strategies is recommended.
- Seasonal Strategy: Sales peaks consistently occur in specific months for different categories; marketing campaigns should be aligned with these historical peaks.