

Regional Sales Analysis – Exploratory Data Analysis (EDA)

CODECADEMY DATA ANALYTICS WITH AI BOOTCAMP -2025

Danny Chacon



Approach

- •Collect and review the *Regional Sales Dataset* (2014–2018).
- •Explore dataset structure: Sales Orders, Customers, Products, Regions, Budgets.
- •Perform data cleaning and preparation for analysis.
- •Conduct Exploratory Data Analysis (EDA) to uncover trends and insights.
- •Special focus: Compare 2017 Budgets vs. actual sales (Line Total) to evaluate performance.
- •Identify over- and under-performing products.
- •Provide strategic recommendations based on findings.

Data overview

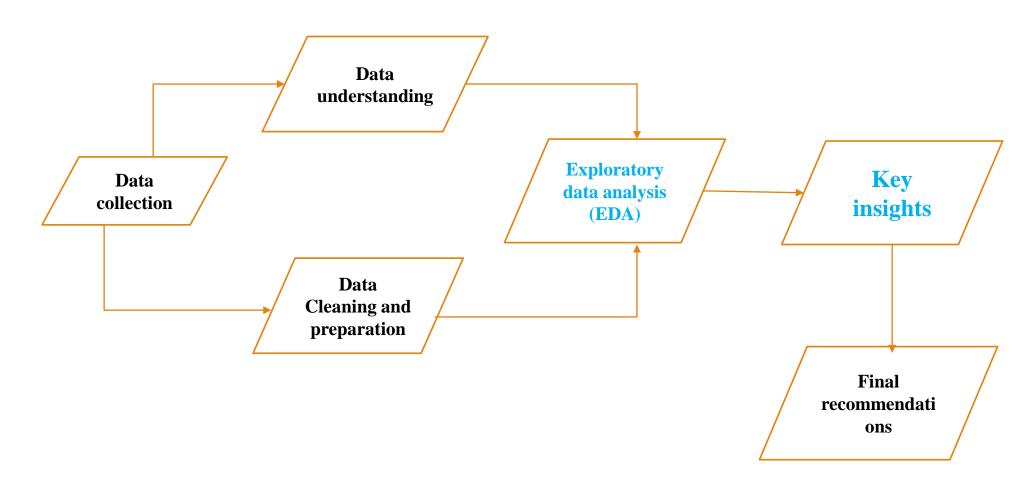


- Dataset covers **2014–2018 sales transactions**.
- Sales Orders table: **61,404 rows** × **12 columns**.
- Key fields: Order Number, Order Date, Channel, Product, Line Total, Profit.
- Other tables: Customers, Products, Regions, State Regions, 2017 Budgets.

	OrderNumber	OrderDate	Customer Name Index	Channel	Currency Code	Warehouse Code	Delivery Region Index	Product Description Index	Order Quantity	Unit Price	Line Total	Total Unit Cost
0	SO - 000225	2014-01- 01	126	Wholesale	USD	AXW291	364	27	6	2499.1	14994.6	1824.343
1	SO - 0003378	2014-01- 01	96	Distributor	USD	AXW291	488	20	11	2351.7	25868.7	1269.918
2	SO - 0005126	2014-01- 01	8	Wholesale	USD	AXW291	155	26	6	978.2	5869.2	684.740
3	SO - 0005614	2014-01- 01	42	Export	USD	AXW291	473	7	7	2338.3	16368.1	1028.852
4	SO - 0005781	2014-01- 01	73	Wholesale	USD	AXW291	256	8	8	2291.4	18331.2	1260.270
orders.shape ✓ 0.0s												
	0.0s 104, 12)											Python

Project workflow





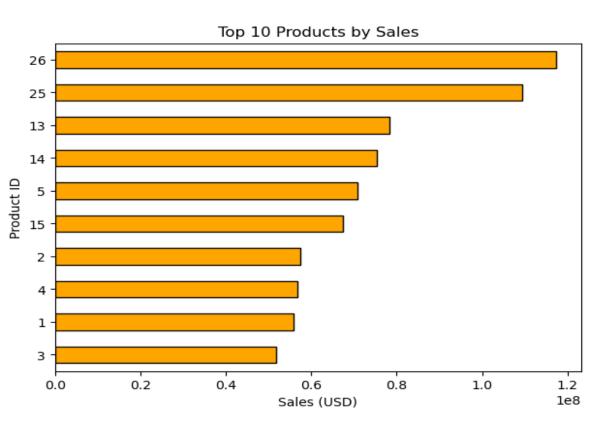


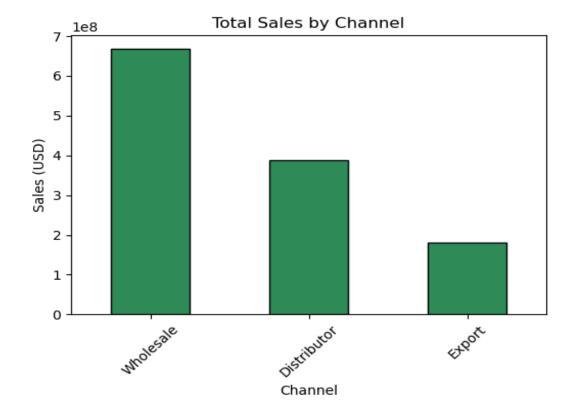
Key Insights

- •Sales remained stable between 2014–2016 (~298M USD/year) but dropped sharply in 2018 (~48M).
- •A small group of products drives the majority of revenue (top 3 products >40%).
- •One channel dominates sales, creating dependency risk.
- •California and Illinois lead in sales (>100M USD), while most states contribute marginally.
- •In 2017, several products exceeded budget forecasts, but some underperformed significantly



Graphs







Recommendations

- •Diversify channels to reduce dependency on a single distribution strategy.
- •Focus investment on high-performing products while reassessing low-selling ones.
- •Expand operations in top states (California, Illinois, Florida) and explore growth opportunities in underperforming regions.
- •Improve forecasting accuracy for budgets to better align expectations with reality.
- •Investigate the 2018 sales drop to confirm whether it is due to incomplete data or real market decline.