

SALES ANALYSIS DASHBOARD

A robust teaching-learning platform dedicated for you to gain the industrial skills and knowledge by, making your learning journey more valuable.



Objectives of Project

The primary objective of the project, which consists of sales order records from an organization, is to provide comprehensive insights into various aspects of the company's sales performance and operations. This data is crucial for informed decision-making and strategic planning.

Introduction:

The dataset includes detailed information about sales transactions, including order dates, customer names, product details, sales amounts, profit margins, shipping modes, and regional data. By analyzing this data, the organization can achieve several key objectives:

1. Sales Performance Analysis:

- Evaluate total sales and profits over time to identify trends and patterns.
- Determine which products and categories contribute the most to sales and profit.
- Identify top-performing sales regions and cities.

2. Customer Insights:

- Analyze customer purchasing behavior to identify high-value customers.
- Understand customer preferences by examining the most purchased products and categories.

3. Operational Efficiency:

- Assess the impact of different shipping modes on sales and profit.
- Monitor the distribution of sales by various shipping methods to optimize logistics.

4. Discount and Profitability Analysis:

- Investigate the relationship between discount rates and profit margins.
- Identify optimal discount strategies that maximize profitability.

5. Product Performance:

- Identify top and bottom-performing products and categories to inform inventory and marketing strategies.
- Analyze sub-category performance to uncover opportunities for product development or discontinuation.

6. Regional and Market Analysis:

- Examine regional sales and profit data to identify market opportunities and challenges.
- Use geographic data to support regional marketing and sales strategies.

Datasets information:

This dataset contains information about orders, including details such as the order date, ship date, ship mode, customer information, product category, sales, quantity, discount, and profit. Here are some specific columns and their descriptions:

Row ID: Unique identifier for each row in the dataset.

Order ID: Unique identifier for each order.

Order Date: The date when the order was placed.

Ship Date: The date when the order was shipped.

Ship Mode: The mode of shipment (e.g., Standard Class, Second Class, First Class).

Customer ID: Unique identifier for each customer.

Customer Name: Name of the customer.

Segment: Customer segment, such as Consumer, Corporate, or Home Office.

Country/Region: Country or region where the order was placed.

City: City where the order was placed.

State: State where the order was placed.

Postal Code: Postal code of the city.

Region: Region in the United States (e.g., West, East).

Product ID: Unique identifier for each product.

Category: Product category (e.g., Office Supplies).

Sub-Category: Product sub-category (e.g., Art, Storage).

Product Name: Name of the product.

Sales: Sales amount for the order.

Quantity: Quantity of items ordered.

Discount: Discount applied to the order.

Profit: Profit earned from the order.

Task 1: Basic Data Exploration and Visualization

Objective: Gain a general understanding of the data, basic trends, and simple aggregates. This task focuses on high-level insights using simple charts and aggregations.*

1. **What is the total record count in the dataset?**

Goal: Verify the size of the dataset and ensure completeness.

Approach: Use a simple card visual in Power BI to display the total count of records.

2. **What is the average value for each numerical column?**

Goal: Understand the general magnitude of each numerical field.

Approach: Use card visuals or table visualizations to display averages of each column.

3. **What is the distribution of categorical values in key columns?**

Goal: Examine how values are distributed in each categorical column to identify any major categories.

Approach: Create bar charts for each categorical field to show the count of records per category.

4. **How many unique values exist for each categorical column?**

Goal: Determine the diversity of values in each categorical field.

Approach: Use distinct count measures for each categorical column in Power BI.

5. **What are the top 10 most common values in the main categorical columns?**

Goal: Identify the most frequent values in each major category.

Approach: Create top N filters in bar or column charts to show the top 10 values.

6. **What is the frequency of missing values in each column?**

Goal: Assess data quality by finding missing data in each column.

Approach: Create calculated columns that count missing values and visualize this count with a table or chart.

7. **What are the maximum and minimum values for each numerical column?**

Goal: Identify the range of values in each numerical field.

Approach: Display maximum and minimum values in cards or tables.

8. **What are the basic trends over time, if there is a time-based column?**

Goal: Get an initial sense of time trends, such as changes in volume over time.

Approach: Use a line chart with a date or timestamp column on the x-axis

9. **Are there any obvious outliers in numerical columns?**

Goal: Spot unusually high or low values that may require further investigation.

Approach: Use box plots to show distributions and potential outliers.

10. **What is the overall distribution (histogram) of each numerical column?**

Goal: Identify the shape of distributions (normal, skewed, etc.).

Approach: Use histograms to visualize the frequency of values in intervals.

Task 2: Comparative Insights

Objective: Conduct comparisons and identify patterns between different columns, considering more intricate visualizations.*

1. How do categorical columns correlate with each other?

Goal: Identify possible relationships or dependencies between categories.

Approach: Use a matrix or a scatter plot matrix to cross-tabulate categories.

2. What is the relationship between two key numerical columns?

Goal: Determine if there's any linear or non-linear relationship between two specific numerical fields.

Approach: Use a scatter plot to visualize the correlation and add a trend line.

3. What are the top 5 highest and lowest performing categories based on an aggregate measure?

Goal: Rank categories based on an aggregate measure (e.g., sum or average of a key metric).

Approach: Create a ranking measure and apply a filter for the top 5 and bottom 5 categories.

4. How does the distribution of a numerical column differ across various categorical values?

Goal: Examine if certain categories have higher or lower averages for a particular metric.

Approach: Use box plots or clustered bar charts to display the distribution per category.

5. What are the trends over time for each major category?

Goal: Compare time-based changes for different categorical values.

Approach: Create a line chart with categorical filters or use small multiples to show trends by category.

6. **How does the average of a numerical column change across different groups?**

Goal: Assess the variation in average values across different segments.

Approach: Use a column chart with a category on the x-axis and an average measure.

7. **What are the key contributing factors for any noticeable trends over time?**

Goal: Identify specific categories or groups that contribute to upward or downward trends.

Approach: Use decomposition trees or drill-through options on line charts.

8. **How many records fall into specific ranges (e.g., using a numerical column in bands)?**

Goal: Understand how values are distributed in predefined ranges.

Approach: Create a calculated column to classify data into bins and use a histogram.

9. **What is the impact of missing values on overall trends?**

Goal: See if missing data skews the results significantly.

Approach: Filter out rows with missing data and compare trends with and without them.

10. **How do aggregate values compare between two selected time periods?**

Goal: Assess whether there's been a change between two timeframes.

Approach: Use a comparison bar chart with filters for each time period or DAX calculations for year-over-year comparisons.

Task 3: Predictive Insights

Objective: Conduct complex analyses, combining multiple dimensions, and consider predictive modeling insights. Approach explanations provided for advanced analysis where necessary.*

1. **What are the top factors contributing to a specific outcome in the dataset?**

Goal: Identify key drivers of an outcome metric (e.g., sales volume or customer churn).

Approach: Use the Key Influencers visual to automatically find correlations and key drivers.

2. **How does a key metric vary over different categories and subcategories?**

Goal: Perform a drill-down analysis across multiple hierarchical categories.

Approach: Use hierarchical charts or drill-down features on bar or matrix charts.

3. **What is the seasonality or periodic pattern in time-based data?**

Goal: Identify repeating patterns on a monthly, weekly, or daily basis.

Approach: Use decomposition analysis with time as the main axis to isolate seasonal trends.

4. **What is the likelihood of a particular outcome based on categorical or numerical values?**

Goal: Predict the likelihood of an event based on past data patterns.

Approach: Use the Key Influencers visual or integrate Power BI with Azure Machine Learning for predictions.

5. How can we predict future trends based on historical data?

Goal: Forecast future values for key metrics.

Approach: Use forecasting features in line charts and configure forecast intervals.

6. How do two or more key metrics interact over time?

Goal: Analyze correlations between multiple metrics with respect to time.

Approach: Use combo charts or multiple lines on the same chart to track interactions over time.

7. What insights can we gain from outlier detection in a numerical column?

Goal: Identify anomalies that may indicate issues or opportunities.

Approach: Use clustering in scatter plots to identify points that deviate from the norm.

8. What is the percentage contribution of each category to an overall total and its variance over time?

Goal: Track the change in contribution of each category.

Approach: Use waterfall charts for changes over time and pie or donut charts for contribution breakdowns.

9. How do complex hierarchical relationships affect the outcome metric?

Goal: Examine multi-level relationships within the dataset.

Approach: Use the decomposition tree or custom hierarchies in matrix visuals.

10. How does the presence of one variable affect the predictive power of another in relation to an outcome?

Goal: Uncover interaction effects or dependencies between variables.

Approach: Use Power BI's custom visualizations with conditional formatting or incorporate machine learning models for interaction effects analysis.

Super-Advanced Level Analysis

This task required the knowledge of Machine learning Algorithms so its for those who have ML knowledge otherwise its **optional** part

This task involves more complex analysis and advanced insights, potentially requiring statistical methods or machine learning.

1. **Predictive Churn Modeling:**

- a. Can we build a predictive model to identify customers at high risk of churning? What are the top features driving churn?

2. **Customer Lifetime Value Prediction:**

- a. Can we predict the lifetime value of customers based on their demographics and purchasing behavior?

3. **Product/Service Recommendation System:**

- a. Using historical data, can we create a recommendation system to suggest products or services to users based on their purchase history or similar user profiles?

4. ****Customer Segmentation Using Clustering**:**

- a. Apply clustering algorithms to segment customers based on multiple variables (e.g., demographics, purchasing behavior, satisfaction).
What are the key segments, and what characteristics define each?

5. ****Anomaly Detection in Sales Trends**:**

- a. Can we identify anomalies or unusual spikes in sales data? What factors could explain these anomalies?

6. ****Sentiment Trend Analysis Over Time**:**

- a. Analyze sentiment in customer reviews over time to detect shifts in customer satisfaction or identify periods of increased negative feedback.

7. ****Cohort Analysis by Acquisition Date**:**

- a. Analyze customer cohorts based on acquisition date. How do different cohorts perform over time in terms of retention, spending, and satisfaction?

8. **Price Elasticity Analysis:**

- a. How sensitive are customers to changes in product prices? Can we estimate the effect of price changes on demand?

9. **Lifetime Engagement by Product Category:**

- a. Which product categories keep customers engaged the longest, and what factors contribute to prolonged engagement?

10. **Advanced Customer Satisfaction Model:**

- a. Build a regression or classification model to identify factors contributing most significantly to high or low customer satisfaction.

DOCUMENTATION

After Completion of the projects you have to Create one Docx file in that you have to Make Report of above Projects and include some Key Factors like

1. Introduction
2. Methodology
3. Requirement Analysis
4. Other Parameters depending upon the Projects
5. All Visualization like All Charts which is there in the Dashboards
6. Insights from the Charts as well as Dashboards

After creating the Reports for above project upload that docx file or pdf file in Assignment links

Submission Guidelines

Format: PowerPoint or PDF

Length: 1-20 slides.

Sections: Introduction, Key Findings, Actionable, Methodologies , Approaches, Insights, Conclusions

Tools and Technologies :

POWER BI O/R TABLEAU

Deadline:

Submit your report and presentation within 21 Days from the day you will start.

