Project Report: Sales Analysis and Forecasting

1. Project Objectives and Insights

This project aims to analyze and forecast sales data to uncover trends, customer behaviors, and product performance. The insights gained will help guide business decisions in marketing, inventory, customer service, and revenue optimization.

Key Analysis Objectives

1. **Customer and Transaction Analysis**: Understand purchasing behaviors based on customer types and demographics.
2. **Product Demand Assessment**: Identify high-demand products to improve inventory management.
3. **Sales Trends by Time Period**: Determine high and low-performing months and quarters.
4. **Branch and Rating Distribution**: Analyze customer satisfaction across branches.
5. **Forecast Future Sales**: Use time-series analysis to predict future sales trends.

2. Data Extraction and Transformation

Data Extraction

The data contains various columns such as `Invoice ID`, `Branch`, `City`, `Customer_type`, `Gender`, `Product_line`, `Unit_price`, `Quantity`, `Tax_5%`, `Total`, `cogs`, `gross_margin_percentage`, `gross_income`, `Date`, `Time`, `Payment`, and `Rating`. SQL was used for extracting, cleaning, and transforming data for analysis.

SQL Queries and Insights

1. **Extract All Data**:
   - Query: `SELECT * FROM sales_data;`
   - **Insight**: Allows a complete view of the data for subsequent analysis.
2. **Average Purchase by Customer Type and Gender**:
   - Query to calculate the average purchase amount, total sales, and number of purchases for each customer type and gender.
   - **Insight**: Shows behavioral differences between customer groups to tailor marketing and loyalty initiatives.
3. **Average Quantity by Product Line**:
   - Query to identify high-demand products by setting a threshold on average quantity.

- o **Insight**: Products like Fashion Accessories have the lowest demand, while Electronic Accessories have the highest. This can guide inventory management.
4. **Branch and Rating Distribution**:
   - o Query to extract branch and rating data.
   - o **Insight**: Enables identification of the branch with the highest customer satisfaction ratings.
5. **Identify Highest Unit Price by Product Line**:
   - o Query to determine the product line with the highest unit price.
   - o **Insight**: Shows Fashion Accessories as the highest-priced category, which can inform pricing and marketing strategies.

Data Transformation

1. **Date and Time Standardization**:
   - o Converted `Date` and `Time` columns into a unified timestamp format.
   - o **Outcome**: Created a new column with standardized date-time data for consistent time-based analysis.

3. Data Cleaning and Preprocessing

1. **Missing Values**:
   - o Checked for missing values in critical columns like `Quantity`, `Unit_price`, and `Total`.
   - o **Outcome**: No missing values were found in key columns.
2. **Outliers in Rating**:
   - o Removed extreme ratings below 2 or above 9.5 to ensure data integrity.
   - o **Outcome**: 71 outlier rows were removed.

4. Sales Trend Analysis with Tableau Visualizations

Objectives

1. **Monthly and Quarterly Sales Trends**: Identify high-performing months and seasonal patterns.
2. **Product Line Performance**: Understand which products drive revenue and profit.
3. **Customer Demographics and Behavior**: Analyze purchasing patterns by customer type and gender.
4. **Revenue by City and Branch**: Compare performance across different locations.
5. **Payment Method Usage**: Determine the most preferred payment methods.
6. **Customer Satisfaction by Product Line**: Examine product lines with high and low customer satisfaction.

Visualizations and Insights

1. **Monthly and Quarterly Sales Trend (Bar Chart)**:
   - o **Insight**: January shows the highest sales, while February is the lowest. This helps in planning seasonal promotions.
2. **Revenue and Profit by Product Line (Bar Chart)**:
   - o **Insight**: Food and Beverages drive the highest revenue, while Health and Beauty is the lowest. This insight can guide decisions around product pricing and inventory.
3. **Customer Demographics (Stacked Bar Chart)**:
   - o **Insight**: Female members contribute more to revenue than other groups, suggesting targeted marketing opportunities.
4. **Payment Method Usage (Pie Chart)**:
   - o **Insight**: Cash is the most preferred payment method. This can guide efforts to promote digital payments.
5. **Customer Satisfaction by Product Line (Bar Chart)**:
   - o **Insight**: Fashion Accessories have the highest ratings, while Health and Beauty has the lowest, guiding quality improvement efforts.

5. Tableau Data Visualization Report

This Tableau Sales Dashboard provides a high-level and detailed view of customer behavior, product performance, payment preferences, and monthly revenue trends.

A. Customer Revenue by Gender

- Female customers generate slightly higher total revenue than male customers.
- The difference is not extreme, indicating balanced engagement across genders.

Insight:

Marketing strategies resonate well with both genders, but there may be opportunities to further personalize offers to maximize female customer lifetime value.

B. Product Line Performance (Total Revenue)

Revenue across product lines is relatively evenly distributed, which indicates a well-diversified product portfolio.

Top-performing product lines:

- ✓ Food and Beverages (highest revenue)

- ✓ Electronic Accessories
- ✓ Home and Lifestyle
- ✓ Lower but still strong contributors:
- ✓ Fashion Accessories
- ✓ Sports and Travel
- ✓ Health and Beauty

Insight:

No single product line overly dominates revenue, which reduces business risk. However, Food & Beverages clearly has the strongest customer pull.

C. Product Line Ratings

- Fashion Accessories and Food & Beverages receive stronger customer ratings.
- Health & Beauty and Sports & Travel show comparatively lighter engagement.

Insight:

High ratings indicate strong customer satisfaction and brand trust, especially in lifestyle and consumable products.

D. Monthly Sales Trend (Q1 2019)

- March records the highest revenue, followed closely by January.
- February shows a noticeable dip, likely due to fewer selling days or reduced customer activity.

Insight:

Sales performance exhibits short-term seasonality, even within a single quarter.

E. Payment Method Analysis

- E-wallets account for the largest share of revenue.
- Credit cards follow closely.
- Cash payments contribute the least.

Insight: Customers strongly prefer cashless payment options, reflecting modern consumer behavior and digital readiness.

6. Forecasting Future Sales with Time-Series Analysis in Python (Prophet Model)

Data Preprocessing

1. **Date Parsing**: Converted `Date` into datetime format.
2. **Aggregated Sales Data**: Summed daily sales for time-series analysis.
3. **Handled Missing Dates**: Filled missing dates to ensure continuous data.

Forecasting with Prophet Model

- **Forecasting Period**: Forecasted sales for the next 60 days.
- **Components**:
  - **Historical Data** (black dots): Actual sales from January to March 2019.
  - **Forecasted Sales** (blue line): Predicted sales for April to May 2019.
  - **Confidence Interval**: Range of uncertainty around forecasted sales.

Chart Interpretation

1. **Sales Trend**: Predicted stable sales initially, followed by a spike and sharp drop, possibly due to seasonal or one-time events.

7. Conclusion and Recommendations

Key Findings

1. Seasonal sales trends and high-demand months were identified.
2. Certain product lines, especially Food and Beverages, drive profitability.
3. Diverse payment preferences were observed, with cash being most popular.
4. Variations in customer satisfaction suggest areas for improvement in product quality.

Recommendations

1. **Boost Marketing in Peak Months**: Focus marketing efforts during high-demand periods.
2. **Promote Profitable Products**: Increase focus on high-revenue products and improve lower-margin ones.
3. **Enhance Loyalty Programs**: Encourage Normal customers to convert to Members.
4. **Expand in High-Revenue Locations**: Consider new branches in top-performing cities.

5. **Promote Digital Payment Options**.
6. **Improve Quality for Low-Rated Products**.
7. Increase promotional focus on Food & Beverages and Electronic Accessories.
8. Introduce bundles or cross-sell offers with Fashion Accessories to boost basket size.