DOCUMENTATION

SALES ANALYSIS USING PYTHON

**Skills:**

With this project I will be developing the following skills-

1. Python Programming
2. Data cleaning and preprocessing
3. Data visualization
4. Statistical analysis
5. Time-series analysis
6. Machine learning
7. Business acumen
8. Communication
9. Version control
10. Project Management

**Solved Business Problems:**

* **Identifying Sales Trends**

Understanding the overall trend in sales helps businesses adapt to market changes and make informed decisions on inventory, pricing, and marketing strategies.

* **Product Performance Analysis**

Businesses can identify their best-performing and underperforming products. This information helps in optimizing the product portfolio, marketing efforts, and investment allocation.

* **Pricing Strategy Optimization**

By analyzing sales data, businesses can evaluate the impact of different pricing strategies on sales volume and revenue. This information is crucial for setting competitive prices and maximizing profits.

* **Forecasting and Demand Planning**

Sales analysis aids in accurate forecasting, allowing businesses to plan production, staffing, and other resources based on anticipated demand.

**Identifying Sales Trend:**

Steps overview-

1. Load and explore sales data with Python.
2. Visualize time-series data to identify trends.
3. Apply statistical methods to smooth data and highlight trends.

**Dataset**

The dataset that we will be using for doing sales analysis is from the retail industry. It is historical sales data for 45 stores located at different locations and each store has some departments inside it. The store also runs some promotional markdown events throughout the year. These markdowns precede prominent holidays, the four largest of which are the Super Bowl, Labor Day, Thanksgiving, and Christmas. The weeks including these holidays are weighted five times higher in the evaluation than non-holiday weeks.

We have 3 csv files- Stores, Features and Sales.

*Stores*

Anonymized information about the 45 stores, indicating the type and size of store

*Features*

Contains additional data related to the store, department, and regional activity for the given dates.

* Store - the store number.
* Date - the week.
* Temperature - average temperature in the region.
* Fuel\_Price - cost of fuel in the region.
* MarkDown1-5 - anonymized data related to promotional markdowns. MarkDown data is only available after Nov 2011, and is not available for all stores all the time. Any missing value is marked with a NA.
* CPI - the consumer price index.
* Unemployment - the unemployment rate.
* IsHoliday - whether the week is a special holiday week.

*Sales*

Historical sales data, which covers 2010-02-05 to 2012-11-01. Within this tab you will find the following fields:

* Store - the store number.
* Dept - the department number.
* Date - the week.
* Weekly\_Sales - sales for the given department in the given store.
* IsHoliday - whether the week is a special holiday week.

**Data Exploration and Cleaning**

*Load the Data*

* Import the necessary Python libraries (Pandas, NumPy, Matplotlib, Seaborn).
* Load the three CSV files (Stores, Features, Sales) into Pandas DataFrames.

*Explore the Data*

* Use descriptive statistics and visualizations to understand the structure and content of each dataset.
* Check for missing values, outliers, and anomalies in the data.

*Merge Datasets*

* Merge the Sales and Features datasets on common columns (Store and Date).
* Explore the merged dataset to ensure consistency.

**Feature Engineering**

*Handle Missing Values*

* Decide on a strategy to handle missing values in the dataset, especially in Markdown columns.
* Impute missing values or consider dropping columns/rows if appropriate.

*Create Time Features*

* Extract additional time-related features from the 'Date' column (e.g., day of the week, month, year).
* Consider creating a separate column for holiday weeks.

**Exploratory Data Analysis (EDA)**

*Analyze Sales Trends*

* Visualize overall sales trends across all stores and departments.
* Explore sales patterns over time, considering holidays and non-holidays.

*Department-wise Analysis*

* Explore sales trends for specific departments.
* Identify high-performing and low-performing departments.

*Store-wise Analysis*

* Analyze sales performance for individual stores.
* Identify top-performing and underperforming stores.

**Holiday Impact Analysis**

*Evaluate Holiday Sales*

* Analyze the impact of holidays on weekly sales.
* Compare sales during holiday weeks to non-holiday weeks.

*Weighted Evaluation*

* Implement the weighted evaluation for weeks including Super Bowl, Labor Day, Thanksgiving, and Christmas.
* Adjust sales figures accordingly for more accurate analysis.

**Feature Correlation Analysis**

*Correlation Analysis*

* Examine correlations between features (e.g., temperature, fuel price, CPI, unemployment) and sales.
* Identify significant factors influencing sales.

**Recommendations**

*Focus on Increasing Overall Sales*

* Identify periods of consistent growth in overall sales and leverage marketing strategies or promotions during those times to maintain or increase sales momentum.

*Investigate Sales Declines*

* Investigate any noticeable declines in sales and assess factors contributing to these downturns. Implement targeted strategies to address and reverse declining trends.

*Optimize High-Performing Departments*

* Identify departments that consistently perform well and invest in optimizing and expanding product offerings within these departments. This could involve introducing new products or enhancing marketing efforts.

*Improve Low-Performing Departments*

* Analyze and understand the challenges faced by low-performing departments. Implement strategies such as promotions, product diversification, or targeted marketing to improve their performance.

*Learn from Top-Performing Stores*

* Analyze top-performing stores to understand the factors contributing to their success. Share best practices and successful strategies with other stores to drive overall improvement.

*Support Underperforming Stores*

* Identify underperforming stores and provide targeted support. This could involve additional training, local marketing efforts, or operational adjustments to enhance their performance.

*Optimize Holiday Sales Strategies*

* Evaluate the success of holiday sales strategies and promotions. Refine and optimize these strategies based on the analysis of past holiday sales performance to maximize revenue during peak seasons.

*Plan for Non-Holiday Periods*

* Develop strategies to maintain sales momentum during non-holiday periods. Consider introducing promotions or events to drive customer engagement during slower times.

*Adjust Resource Allocation*

* Based on the weighted evaluation, consider adjusting resource allocation, inventory planning, and marketing efforts specifically for weeks including Super Bowl, Labor Day, Thanksgiving, and Christmas.

*Plan Promotions Strategically*

* Plan and execute promotional markdown events strategically, aligning them with the weighted evaluation periods to capitalize on increased consumer activity.

*Optimize Pricing Strategies*

* Analyze the correlation between features like fuel price, CPI, and unemployment with sales. Optimize pricing strategies based on these correlations to ensure competitive pricing and maximize revenue.

*Weather-Responsive Marketing*

* Leverage temperature data to tailor marketing strategies. For example, promote seasonal products during weather-appropriate periods to enhance customer engagement.

**DOCUMENTATION END**