**Marketing Analytics**

**Adidas USA Sales**



**Project Proposal**

**Name: Sai Anudeep Janapareddy Student ID: 0836698**

|  |  |
| --- | --- |
| **Table of Contents** | **Page No.** |
| Introduction and Motivation | 1 |
| Dataset | 2 |
| Market Problem | 4 |
| Problem Approach | 5 |
| Tools | 6 |
| Timeline | 6 |

**Introduction and Motivation:**

**Why did you choose this company or topic? What interests you the most?**

Adidas, a globally recognized sportswear brand, maintains a robust position in the North American market. Headquartered in Herzogenaurach, Germany, Adidas AG was established by Adolf Dassler, a German artisan. Initially, the Dassler brothers, Rudolf and Adolf, jointly founded Gebrüder Dassler Schuhfabrik (Dassler Brothers Shoe Factory). Both excelled in the footwear industry, but subsequently pursued separate paths, with Rudolf founding Puma and Adolf founding Adidas.

This dataset furnishes valuable insights into the brand's sales performance in the years 2020 and 2021 across diverse retailers, rendering it particularly compelling from a marketing standpoint. The rationale behind selecting this dataset is its potential to yield valuable insights into market trends, and operational strategies within the apparel and sportswear sector.

**Dataset:**

**Dataset Link:**

<https://data.world/npbn/adidas-sales-data/workspace/file?filename=Adidas+Sales+Datasets.xlsx>

**Data source and its industry.**

The dataset was acquired from Adidas AG corporation and delineates the sales of products across its retail establishments in the United States.

**Details about the features of the dataset.**

The Adidas sales dataset contains several features that provide detailed information about each sale. Here's a breakdown of the key features:

Retailer Information

* **Retailer**: The name of the retailer.
* **Retailer ID**: A unique identifier for the retailer.

Date and Location

* **Invoice Date**: The date of each sale, ranging from January 2020 to December 2021
* **Region**: The geographical region.
* **State**: The state where the sale occurred.
* **City**: The city of the sale.

Product Details

* **Product**: Categorizes the product.

Sales Metrics

* **Price per Unit**: The price of each item sold.
* **Units Sold**: The quantity of items sold in each transaction.
* **Total Sales**: The total revenue generated from the sale (Price per Unit × Units Sold)
* **Operating Profit**: The profit earned from the sale
* **Operating Margin**: The profit margin as a percentage (e.g., 35%)

Sales Channel

* **Sales Method**: The method of sale.

**What this dataset is about such as Sales, Marketing, Budget, and HRM?**

The dataset primarily focuses on sales of Adidas products through various retail outlets, providing detailed information on sales metrics such as price per unit, units sold, total sales, operating profit, and operating margin. The data covers sales from January 2020 to December 2021. Sales are exclusively through various stores listed as outlets, online or in-store for different transactions. The dataset allows for analysis of daily sales performance, profitability of different product categories, and seasonal trends in sales and pricing. However, it does not include data on marketing strategies, budgeting processes, or human resource management; its focus is solely on sales performance and basic product categorization.

**What are some of the features, you are interested in?**

Several interesting features could be valuable for analysis:

Sales Performance Metrics

1. **Total Sales**: This metric gives a direct measure of revenue generated from each transaction
2. **Units Sold**: The quantity of items sold in each transaction, which can help identify popular products or high-volume sales days
3. **Operating Profit**: This shows the actual profit earned from each sale, allowing for profitability analysis
4. **Operating Margin**: Expressed as a decimal, this indicates the profit percentage for each sale

Product Information

1. **Product**: The dataset distinguishes between Apparel, Street Footwear, and Athletic Footwear, allowing for category-specific analysis
2. **Gender Type**: Products are categorized as Men's or Women's, enabling gender-based sales analysis

Temporal Data

1. **Invoice Date**: This allows for time-based analysis, including daily sales trends, weekly patterns, and seasonal variations.

Pricing Information

1. **Price per Unit**: This feature can be used to analyze pricing strategies and their impact on sales and profitability

Location Data

1. **City and State**: This data is useful in comparing sales across multiple locations.

Sales Channel

1. **Sales Method**: The data can be used to find variations in sales channels.

**What is the spread of the dataset?**

The dataset encompasses the timeframe spanning from January 2020 to December 2021, rendering it robust for conducting a comprehensive 2-year trend analysis due to its substantial 9649 entries.

**How Old and How Long?**

The data, although a few years old, still provides valuable insights into seasonal sales patterns, product performance, and pricing strategies for Adidas products in the market. The dataset covers a period of approximately 2 years, starting from January 2020 and ending in December 2021.

**Market Problem:**

Seasonal Demand Fluctuations and Inventory Management. The data reveals significant fluctuations in sales volumes and product demand across different months and regions particularly as the holiday season approaches. This presents a challenge for inventory management and pricing strategies.

Key Observations:

1. **Increasing Demand**: There's a noticeable increase in units sold as we move from October to December, especially for certain product categories.
2. **Price Variations**: Prices tend to increase for many products towards the end of the year, likely in response to higher demand.
3. **Category Disparities**: Street Footwear consistently outperforms other categories in terms of units sold and total sales.
4. **Gender Differences**: There are notable differences in sales patterns between men's and women's products.

The main challenge is balancing inventory levels and pricing strategies to:

1. Meet the increased demand during peak seasons without overstocking.
2. Maintain profitability during slower periods.
3. Optimize the product mix to capitalize on high-performing categories while supporting underperforming ones.
4. Address the disparity in sales between men's and women's products.

This problem requires a sophisticated approach to demand forecasting, inventory management, and dynamic pricing to maximize sales and profitability throughout the year while minimizing the risks of stockouts or excess inventory.

**Problem Approach/Solving the problem:**

Based on the Adidas sales dataset, here's a comprehensive problem-solving approach focusing on data cleaning, exploratory data analysis, visualization, analysis models, and conclusion:

1. Data Cleaning

* Remove duplicate entries, if any.
* Ensure consistency in date formats and convert to a standard datetime format.
* Check for and handle any missing values.
* Standardize text fields (e.g., product categories, gender types).

2. Exploratory Data Analysis (EDA)

* Calculate basic statistics for numerical columns (mean, median, standard deviation).
* Analyze the distribution of sales across product categories and gender types.
* Identify trends in sales volume and revenue over time.
* Examine the relationship between price and units sold.
* Investigate seasonality patterns in the data.

3. Data Visualization

* Create time series plots for total sales and units sold.
* Use bar charts to compare performance across product categories and gender types.
* Develop heatmaps to visualize sales patterns across days of the week and months.
* Create scatter plots to explore relationships between variables (e.g., price vs. units sold).
* Use box plots to show the distribution of operating margins across categories.

4. Analysis Models

Time Series Analysis

* Conduct seasonal decomposition to understand underlying patterns.

Regression Analysis

* Use multiple linear regression to identify factors influencing total sales.
* Implement logistic regression to predict high-performing days or products.

Clustering

* Apply K-means clustering to group similar products or sales days.

Machine Learning Models

* Develop a random forest model to predict sales volume based on various features.
* Use gradient boosting for sales forecasting and feature importance analysis.

5. Conclusion and Recommendations

* Summarize key findings from the analysis.
* Identify top-performing products and categories.
* Highlight seasonal trends and their impact on sales.
* Provide actionable insights for inventory management and pricing strategies.
* Suggest areas for further investigation or data collection.

By following this approach, we can gain deep insights into sales patterns, optimize inventory management, and develop data-driven strategies to improve the overall performance of Adidas products.

**Tools:**

Python and its libraries:

* NumPy: For effective data array management and numerical computations.
* Pandas: For exploratory analysis, data cleaning, and manipulation. The dataset will be analyzed, filtered, and organized using it.
* Seaborn: Developing intricate and educational plots to reveal trends and patterns in data visualization.
* Scikit-Learn: For applying predictive analytics and segmentation models.

Jupyter Notebook:

To plan, carry out, and record the complete analysis procedure. It facilitates interactive data exploration, which facilitates result visualization and interpretation.

Tableau:

To generate interactive reports, dashboards, and sophisticated visualizations. This can make it easier to convey the insights, particularly to stakeholders who require a clear, visual representation of the data.

**Timeline:**

Estimated Timeline of achieving the result

Week 1: Data Cleaning and Initial EDA

* Clean the dataset, addressing any inconsistencies or errors.
* Perform initial exploratory data analysis.
* Create basic summary statistics and initial visualizations.

Week 2: Advanced EDA and Visualization

* Conduct in-depth exploratory data analysis.
* Develop comprehensive visualizations (time series plots, bar charts, heatmaps).
* Identify key trends and patterns in the data.

Week 3: Time Series Analysis

* Implement time series decomposition to understand seasonal patterns.
* Evaluate and refine time series models.

Week 4: Regression and Machine Learning Models

* Develop multiple linear regression models to identify sales drivers.
* Implement logistic regression for high-performance prediction.
* Create and train random forest and gradient boosting models.

Week 5: Advanced Analytics and Clustering

* Perform price elasticity analysis.
* Apply K-means clustering to group similar products or sales days.

Week 6: Model Evaluation and Refinement

* Evaluate all models and refine as necessary.
* Conduct cross-validation and sensitivity analysis.
* Prepare final versions of all analyses and models.

Week 7: Conclusions and Recommendations

* Synthesize findings from all analyses.
* Develop actionable recommendations for inventory management and pricing strategies.
* Prepare final report and presentation materials.

**Total Estimated Time: 7 Weeks**

This plan allows for a comprehensive analysis of the Adidas sales data, progressing from basic cleaning and exploration to advanced modelling and strategic recommendations. Each week builds upon the previous, ensuring a thorough and methodical approach to solving the identified market problems.