Data Science sous R

Projects

2024

1 Customer Segmentation

1.1 Description

In the highly competitive market landscape, businesses must identify and understand the heterogeneous nature of their customer database. Customer segmentation addresses this challenge by recognizing the diversity of customers and categorizing them into distinct groups or segments based on shared characteristics. The primary objective is to move away from a one-size-fits-all approach and, instead, to adopt a more focused and personalized marketing strategy. By segmenting customers, a company can identify which groups are most profitable, which are underserved, which have high growth potential, and which might be at risk of churn. This strategic segmentation enables businesses to tailor their products, services, and marketing messages to meet each segment's unique needs and desires. The segmentation criteria can be based on various features such as demographic information, purchase history, online behavior, and customer life-cycle stage. By leveraging customer segmentation, companies streamline their marketing efforts and enhance their product offerings, customer service, and overall strategic positioning. This customer-centric approach aligns the company's objectives with customer needs and ensures that every marketing effort creates value for both the customer and the business.

1.2 Dataset

- https://www.kaggle.com/datasets/vetrirah/customer

2 Product Reviews Analysis

2.1 Description

Nowadays, product reviews are a goldmine of consumer feedback and are pivotal in shaping the reputation and the demand for products. Thus, the analyses of product reviews that go beyond merely computing the positive and negative ratings of product reviews can uncover meaningful patterns and nuanced opinions expressed in customer feedback. They can, for instance, reveal insights about customers' perceptions of a product's features, performance, and service experiences and even compare public sentiment toward competitors' products.

2.2 Dataset

- https://www.kaggle.com/datasets/snap/amazon-fine-food-reviews

3 Price Elasticity Estimation

3.1 Description

Many economists believe that U.S. gasoline consumption in the short run is mainly unresponsive to fluctuations in the retail price of gasoline. Whether this view is correct is not self-evident. The magnitude of the short-run price elasticity of gasoline demand is immediate for policymakers. Knowledge of this elasticity is also essential for gauging the macroeconomic effects of gasoline price fluctuations. Moreover, the magnitude of the price elasticity of gasoline demand plays a vital role in the debate about speculation in oil markets. Finally, price elasticity is essential in various industrial organizations' automobile market microeconomic models. The relationship between the retail price of gasoline and its consumption in the U.S. has significant implications for economic policy, market forecasting, and industry strategy. The standard hypothesis posits that U.S. gasoline consumption is relatively inelastic in the short run, meaning that price changes do not substantially affect the quantity consumed. This project aims to empirically investigate and estimate the short-run price elasticity of gasoline demand, challenging or affirming prevailing economic theories. The outcome

of such analysis is not merely academic; it has far-reaching consequences for policymakers who regulate fuel prices, economists who assess macroeconomic vulnerabilities, and stakeholders debating market dynamics and speculation impacts.

3.2 Dataset

- https://journaldata.zbw.eu/dataset/anticipation-tax-avoidance-and-the-price-elasticity-of-gasoline-demand

4 Integrating Marketing Mix Modeling and Brand Equity

4.1 Description

The purchase timing of households is usually modeled at the category level. However, many potential explanatory variables are observed at the brand level. To explain inter-purchase times, one has to either construct category-level measures of marketing efforts or integrate the model with a model for brand choice. This project aims to create marketing mix modeling (MMM) and brand equity models to quantify marketing activities' short- and long-term effects. Likewise, it seeks to untangle the complex interplay between different marketing variables and isolate the contribution of each to the bottom line. The end goal is to optimize the marketing strategy for better ROI, efficiency, and market penetration.

4.2 Dataset

- https://www.kaggle.com/datasets/datatattle/dt-mart-market-mix-modeling/data

5 Quantify Marketing Channel Effectiveness

5.1 Description

Measuring the impact of different advertising channels on demand through marketing mix modeling is critical for different business activities. It enables one to understand the effectiveness of advertising channels and allows for better allocation of marketing budgets, maximizing return on investment (ROI).

5.2 Dataset

- https://github.com/jamesrawlins1000/Market-mix-modelling-data

6 Sales Forecasting

6.1 Description

Sales forecasting is a strategic endeavor that is the backbone of a company's operational and financial planning. Many factors can change the sales of a set of products. If a company ignores them, it may lose money. Consequently, sales prediction is one of the most crucial plans for a company. Sales forecasting gives the company an idea of arranging stocks, calculating revenue, and deciding to make a new investment. Another advantage of knowing future sales is that achieving predetermined targets from the beginning of the season can positively affect stock prices and investors' perceptions. Also, not reaching the projected target could significantly damage stock prices, conversely. And it will be a big problem.

6.2 Datasets

- https://community.tableau.com/s/question/0D54T00000CWeX8SAL/sample-superstore-sales-excelxls
- https://www.kaggle.com/datasets/aslanahmedov/walmart-sales-forecast

7 Sales Trend and Seasonality Analysis

7.1 Description

Knowing customer behavior helps predict sales, and it plays an important role in the context of online retail. Businesses can tailor their marketing strategies, optimize product offerings, and enhance customer experiences by understanding purchasing patterns, preferences, and factors driving customer decisions. Analyzing these behaviors allows for accurate demand forecasting, better inventory management, and personalized promotions, ultimately leading to increased

customer satisfaction, higher conversion rates, and improved revenue. Insights into customer behavior also enable identifying high-value customers, helping businesses prioritize retention efforts and maximize customer lifetime value.

7.2 Dataset

- https://github.com/jamesrawlins1000/Online-Retail-Data-Set-

8 Pricing Strategy Analysis

Understanding price policies' short- and long-term effects on sales is crucial for developing effective pricing strategies that maximize revenue and customer satisfaction. Businesses can identify optimal price points that drive sales and encourage customer loyalty by analyzing how different pricing adjustments influence consumer demand over time. Additionally, insights into price elasticity help manage inventory more efficiently, ensuring that stock levels align with anticipated demand changes. This approach minimizes stockouts and overstock situations, leading to improved supply chain management and a more balanced relationship between sales growth and operational efficiency.

8.1 Dataset

- https://github.com/jamesrawlins1000/Online-Retail-Data-Set-

9 Evaluation Criteria

Each project will be evaluated based on the following elements:

1. **Data**: sources, wrangling, and management.
2. **Data analysis**: statistics, machine learning, business insight.
3. **Communication of the results**: summarization and visualization.
4. **Operationalization**: added value considering users' viewpoints.

10 Deliverables

1. An application or an R Notebook file demonstrating the project results.
2. A video describing your results and solution.
3. A report explaining what was done and how with a maximum of eight pages + appendix.
4. A 10-minute presentation of your project.

11 Groups

The groups were randomly generated using the code below, considering a list of students' names ordered by their surnames.

```
set.seed(42)
students <- 1:32
group.numbers <- sample(1:8, 8)
groups <- split(students, sample(rep(1:8, each=4)))
(named.groups <- setNames(groups, group.numbers))
```

Table 1: Member of the groups randomly generated

Group	Members
Customer Segmentation	CORBET Rayan DESMEDT Justine RAKOTONANAHARY Dorian TELLIER Pauline
Product Reviews Analysis	CHIKHI Silya NURSOO Ryan PICHEREAU Chloe ROUVEYROL Coralie
Price Elasticity Estimation	MANZANARES Nicolas VASSAUX Robin VERCHERE Laura MEHATS Arthur
MMM & Brand Equity	RIZET Lutece ASENCIO Victor CLEMENT Estelle HALLE Tom
Quality Marketing Channel Effectiveness	BLANCKE Ines CHAUVIN Lina DELOISON Bo SORNET Nathan
Sales Forecasting	ANDRE Clémence BOCQUEL Mathieu HOLLIER-LAROUSSE Nicolas ROBIN Alix
Sales Trend and Seasonality Analysis	BRUNET Maxence COMTE Noelie KREISS Paul THUILLER Jossua
Price Strategy Analysis	HUISSOUD Liam MACHEDA Andréa BLATTMANN Olivia PEZE Thomas