XN PROJECT: PROJECT ROADMAP (GROUP)

Sean McLean

Paula Andrea Romero Melo

Gauri Udaysing Shinde

Theodore R. Smith

College of Professional Studies, Northeastern University

PROJECT ROADMAP

Introduction

This project utilizes a comprehensive Amazon dataset to uncover correlations and enhance brand performance through various analytical methods. Preliminary findings indicate a possible link between sales quantities and marketing spending, necessitating further investigation. In the coming weeks, our team will systematically examine these relationships and apply advanced data analysis and machine learning techniques to improve brand positioning and sales on Amazon. Key milestones include assignments focused on different aspects of the XN Project, such as evaluating pros and cons, developing project roadmaps, conducting detailed interviews, and preparing final presentations. These efforts aim to deliver thorough analysis, insightful visualizations, and actionable recommendations, ultimately driving significant improvements in digital sales and marketing strategies for our client.

Analytical Approach

There are several possible analyses and machine learning techniques that can be employed to optimize brand positioning and performance on Amazon given the provided data and context:

Exploratory Data Analysis

With the provided Amazon dataset, conduct a comprehensive exploratory data analysis (EDA) using various approaches. Begin with descriptive statistics to calculate summary statistics and plot distributions for key metrics like impressions, clicks, cart adds, and purchases. Use data visualization techniques such as time series analysis, bar charts, and scatter plots to explore trends and relationships. Conduct correlation analysis to identify how variables are related and visualize these connections. Analyze the conversion funnel to understand drop-off rates and calculate conversion rates at different stages. Assess brand performance by examining share

metrics and comparing prices at different stages. Explore the impact of different shipping speeds on user behavior, and conduct segmented analysis based on these speeds. Investigate search query performance by analyzing query scores and identifying top-performing queries. Finally, compare performance metrics for brand-related and non-brand-related searches, and perform competitive analysis if multiple brands are included in the dataset.

To implement exploratory data analysis (EDA), start with data cleaning to ensure the dataset is accurate and consistent, addressing any missing or inconsistent values. Next, perform data transformation to create new derived variables as needed. Utilize visualization tools such as ggplot2 in R, to create insightful visualizations. Conduct statistical analysis using appropriate libraries to compute correlations, summary statistics, and perform hypothesis testing. Finally, compile and present findings in a report or interactive dashboard.

Business Intelligence (BI) Analysis

Descriptive Analytics:

Trend Analysis: Track changes in metrics such as search query volumes, impressions, clicks, cart adds, and purchases over time to identify trends and seasonality.

Performance Metrics: Calculate KPIs such as click-through rates (CTR), conversion rates, and brand share percentages for various stages (impressions, clicks, cart adds, purchases).

Comparative Analysis:

Brand vs. Non-Brand Comparison: Compare the performance of branded search terms (e.g., "Stonewall Kitchen Blueberry Jam") versus non-branded terms (e.g., "blueberry jam").

Competitor Benchmarking: Analyze how a brand performs relative to competitors in terms of search query scores, impressions, clicks, and purchases.

Visualizations:

Dashboards: Create interactive dashboards using tools like Tableau, Power BI, or Looker to visualize the data and track performance metrics.

Heatmaps: Use heatmaps to visualize the distribution of clicks, cart adds, and purchases across different shipping speeds, prices, and other variables.

Machine Learning Analysis

Predictive Analytics:

Sales Forecasting: Use historical data to forecast future sales volumes and identify high-performing search queries.

Behavior Prediction: Predict the likelihood of a search query leading to clicks, cart adds, and purchases.

Clustering:

Customer Segmentation: Use clustering algorithms to segment customers based on their behavior (e.g., search queries, clicks, cart adds, purchases).

Product Segmentation: Group similar products based on attributes like search query volumes, impressions, and conversion rates.

Classification:

Search Query Classification: Classify search queries as branded or non-branded and predict their potential performance.

Conversion Likelihood Classification: Classify search queries based on their likelihood to lead to conversions (purchases).

Regression Analysis:

Price Sensitivity Analysis: Use regression models to understand how changes in price affect clicks, cart adds, and purchases.

Ad Performance Analysis: Model the relationship between advertising spend and outcomes such as impressions, clicks, and purchases.

Natural Language Processing (NLP):

Sentiment Analysis: Analyze the sentiment of search queries and customer reviews to gauge customer perception and its impact on performance.

Search Term Optimization: Use NLP to optimize search terms and improve relevance for better search engine performance on Amazon.

Specific Models

Brand Influence Model:

Ranking Analysis: Develop a model to quantify the influence of a brand on search query performance by comparing branded versus non-branded term rankings.

Bidding Strategy: Formulate a bidding strategy based on the brand influence score, adjusting bids to optimize ad-spend efficiency.

Advertising Effectiveness Model:

Ad Spend Optimization: Create a model to optimize advertising spend by analyzing the impact of ad impressions on clicks and sales.

Return on Ad Spend (ROAS): Calculate and optimize ROAS by adjusting bidding strategies based on predicted outcomes.

Implementation Steps

Data Preparation: Clean and preprocess the data to handle missing values, normalize numerical features, and encode categorical variables. Split the data into training and test sets for machine learning models.

Feature Engineering: Create new features such as time-based indicators (e.g., day of the week, seasonality), interaction terms (e.g., price × shipping speed), and aggregated metrics (e.g., average click rate).

Model Training and Evaluation: Train various machine learning models (e.g., linear regression, decision trees, random forests, gradient boosting machines) and evaluate their performance using metrics such as RMSE, MAE, and accuracy.

Deployment: Deploy the best-performing models into a production environment where they can provide real-time insights and recommendations.

Monitoring and Optimization: Continuously monitor the performance of the deployed models and adjust as necessary to ensure they remain accurate and effective.

These are initial ideas for the analysis, and the feasibility of each idea will be evaluated to determine the most effective approaches. Through a deliberate process, the team will assess each concept's potential impact on optimizing brand positioning and performance on Amazon. Through careful evaluation, we aim to identify the most viable and valuable strategies for leveraging the available data to enhance digital sales and marketing efforts. The selected methodologies will be rigorously tested and refined to ensure they meet the project's goals and the sponsor's expectations.

Milestones

1. XN Project: Pros & Cons
   Date: April 22, 2024

Summary: Analyze the advantages and disadvantages of the XN Project.
2. XN Project: Digging Deeper (Group)
    Date: April 29, 2024
    Summary: This group assignment requires a deeper exploration of the XN Project.
3. XN Group Project: Interview Guide-Sponsor Project-Interview (Group)
    Date: May 6, 2024
    Summary: Create an interview guide for the sponsor project interview.
4. XN Project: Project Roadmap (Group)
    Date: May 20, 2024
    Summary: Focus on developing a project roadmap for the XN Project.
5. XN Project: Mid-Term Presentation (Group)
    Date: May 27, 2024
    Summary: Present the mid-term progress of the XN Project.
6. XN Project: Project Scope Document (Group)
    Date: May 27, 2024
    Summary: Create a project scope document for the XN Project.
7. XN Group Project: Project Draft (Group)
    Date: June 17, 2024
    Summary: Submit a draft of the XN Project.
8. XN Project: Final Draft (Group)
    Date: June 24, 2024
    Summary: The final draft of the XN Project is due.
9. Group XN Project: Presentation Slide Deck (Group)
    Date: June 30, 2024
    Summary: Prepare and submit a presentation slide deck for the XN Project.
10. Group XN Project: Project Deliverable and Presentation
    Date: June 30, 2024
    Summary: Submit the final project deliverable and giving a presentation for the XN Project.
11. XN Project: Individual Contribution
    Date: June 30, 2024
    Summary: Each individual details their contribution to the XN Project.

Job Assignments

Sean McLean: EDA and Analysis

Paula Andrea Romero Melo: Timeline management and communication lead

Gauri Udaysing Shinde: Research and Analysis

Theodore R. Smith: Synthesis of report and presentation

Key Risks and Strategies

Security and Privacy of Data

Managing enormous volumes of Amazon data carries serious dangers to data security and privacy, particularly if it contains sensitive customer information. It is imperative to use strong encryption protocols for data in transit and at rest in order to reduce these dangers. Strict access controls minimize the possibility of unwanted access by guaranteeing that only authorized individuals can access sensitive data.

Consistency and Quality of Data

The dependability of the Price Optimization Model and the Brand Influence Model can be compromised by inconsistent or poor-quality data, which can result in erroneous models and misguided methods. Strong data validation and cleaning procedures must be put in place to reduce this risk and guarantee high-quality data from the start.

Combining with Current Systems

Technical and operational difficulties may arise when integrating new data models with the systems and procedures of the business. It is essential to create precise integration plans and schedules, along with thorough testing processes to find and fix any problems early, in order to reduce these risks. The smooth transfer of data between systems can be facilitated by using API-based integrations, reducing interruptions. The company team will be able to successfully use the new models and incorporate them into their workflows if they receive training and support.

Measure(s) of Success

Success is measured by realizing a positively effective change. In a real consulting analysis, there would be an expected minimum rate of return. That could be a dollar figure (at a minimum the change should cover the consulting fee) or it could be a percentage increase in

profit. In this case, a measure of success is one of three things. 1. An increase in sales. 2. An increase in profit. 3. An increase in market share for a category.

Those are the quantifiable measures of success.   The other measures of success come later. Does the client call our company back for further consultation? Do they recommend us to other companies? Does this project generate additional clients?

Presentation Method

The presentation will be a standalone brief. The slides will speak for themselves. There will be enough information to fully comprehend the decisions and recommendations. There will also be clear contact information if anyone has any questions or wants to dive deeper into a subject.

References

Champigny, A. (2024, May 18). *6 Ways to Measure Your Consulting Firms Performance*.

Retrieved from deltek.com: https://www.deltek.com/en/blog/consulting-metrics-

associated-with-success

guidingmetrics.com. (2024, May 18). *The Management Consulting Industry's 10 Most Critical

Metrics*. Retrieved from guidingmetrics.com:

https://guidingmetrics.com/content/management-consulting-industry-metrics-kpis/