Understanding Success in Local Business Mobile & Web Apps

Research Proposal

By

Christopher Prasetya Mulya - 49209794
Jacob Rosen - 28837649
Jade Yu - 67606236
Natalie Crystal Coutinho - 66731928
Pranav Kumar Mahesh - 39703434

Master of Data Science

The University of British Columbia

British Columbia, Canada

May 13, 2024.

Introduction:

Craver is an intelligent engagement platform that enables restaurants to connect with their customers through mobile apps, self-serve kiosks and online web ordering. It helps restaurants increase the lifetime value of each customer and improve restaurant profitability. The Craver team has provided five open-ended questions. The primary goal of this project is to derive valuable and actionable insights through big data analysis of the merchant and the restaurant operation data with Google BigQuery. Analyzing customer behavior data can help businesses better understand the customers' needs and identify pain points or areas for improvement. By addressing these issues, Craver can enhance the overall customer experience and build stronger relationships with their customers (Verhoef et al., 2015).

The ultimate purpose is to increase the profit of Craver customers. To achieve this purpose, we will leverage Google BigQuery (Tigani & Naidu, 2014) to analyze vast amounts of customer data to identify patterns and segment customers based on their behavior. This enables personalized marketing efforts tailored to specific customer groups, leading to higher engagement and conversion rates (Wedel & Kannan, 2016).

Aims and Objectives:

The aim of this project is to provide actionable insights for Craver, focusing on enhancing customer engagement and optimizing merchant success within their ecosystem.

Key objectives to be accomplished include:

1. **Consumer Behaviour & Growth Analysis for Jules/Lovebird:** Investigate the impact of loyalty program usage on guest behavior and analyze customer growth patterns for specific large merchants.
   - (Jules/Lovebird) Looking deeper into the loyalty program to see how earning and using credit (stars) influences guests' behavior when it comes to ordering in person/over the phone vs door dash vs craver
   - (Jules/Lovebird) Comparing the growth of customer behavior in terms of their evolution as a Jules and/or Love Bird customer (i.e. if they were just a customer of Jules and then became one at Lovebird too and vice versa)
   - (Jules/Lovebird) How we can increase ticket values based on recurring trends
     - E.g. complementary items/pairings?
   - (Jules/Lovebird) Customer patterns/growth before and after the installation of Craver
   - (Jules/Lovebird) Other basic recurring trends on a consumer basis (i.e. popularity of items, what sells based on time of year, etc.)
2. **Understanding Customer Engagement and Retention:** Differentiate between sticky and non-sticky guests and identify factors influencing customer repurchase behavior.

3. **Impact of Restaurant Type on Merchant Success:** Evaluate how different restaurant types (coffee shop, bakery, quick-service restaurant, and food truck) influence merchant success metrics.
4. **Characteristics of High-Revenue Customers:** Identify common features among high-revenue customers to enhance revenue generation strategies.
5. **Cart Abandonment Analysis and Purchase Trends:** Investigate cart abandonment rates and purchase trends to optimize the ordering process and improve customer satisfaction.

Dataset:

The dataset used for this project is provided by Craver on their Google BigQuery platform. The dataset contains 103 tables representing the entire backend of Craver's operations. The data were collected between October 2017 and April 2024. Our project will focus primarily on the companies, locations, users, and orders tables. Each table has over 20 columns represented by a variety of data types including timestamps, integers, and strings. The companies and locations tables provide information about each merchant and their collection of restaurant locations. The users table provides information about the consumers using the app to make purchases from merchants. The orders table provides information about each individual order placed by a user. There are roughly 10 million orders, 1 million users, 613 companies, and 1,729 locations in the dataset.

Data Preparation Steps:

1. **Data Integration:** Merge or join relevant tables to combine information from multiple sources into a single dataset.
2. **Feature Engineering:** Create new features or derive additional insights from existing data to enhance analysis.
3. **Data Transformation:** Perform transformations to make the data suitable for analysis.
4. **Data Validation:** Validate the prepared dataset to ensure its quality and integrity before analysis.

Methodology:

Objective 1: Consumer Behaviour & Growth Analysis for Jules/Lovebird

To comprehensively understand guest behavior and purchasing patterns at Jules/ Lovebird, we will conduct analyses across multiple dimensions. Firstly, we will delve into the influence of loyalty programs on guest return behavior by examining tables Users, Orders, Companies, and Reward_Tiers. Utilizing the 'orders' table and user IDs, we will analyze the distribution of orders based on delivery options, discerning any potential influence of delivery method on loyalty program participation. This is done by the GROUP BY clause to count the number of orders per user based on the delivery option registered under the loyalty program. Secondly, our approach involves defining item sets and conducting item pairing analysis using the 'orders' and ‘order_items’ table to identify product pairs based on their IDs that are frequently purchased together. This analysis will provide insights into guest preferences and purchasing behavior. Lastly, we will explore recurring trends in guest behavior using data from ‘users’ and ‘orders’ tables. This will include monthly sales analysis to identify sales trends over time taking the ‘placed_at’ column as month and calculating the order frequency on a month on month basis. Similarly, conducting a monthly product analysis to pinpoint products frequently purchased within each month. Additionally, we will delve into channel analysis to understand the impact of different channels on guest behavior and revenue generation and calculate their lifetime value.

Besides, in order to look into the customer growth pattern of Jules/Lovebird, we will mainly use tables with information about orders, users, companies and payment details. While our focus is Craver users in this project, to better provide some insights to large merchants like Jules/Lovebird, we will analyze the behavior of customers ordering from all platforms. Our approach centers around analyzing the customer growth of the two merchants and the behavioral patterns of customers after their first purchase. First of all, we will look into the number and distribution of customers of Jules and Lovebird and check whether there is an overlap. This part will be illustrated through data visualization. Second, we will conduct a Customer Growth Analysis of both Jules and Lovebirds. We will utilize the table ‘orders’ and extract information about when customers make their first purchase at Jules and Lovebird and how long it takes them to become a customer of Lovebird/Jules as well. We will also explore the possible reasons behind their such behavior. This analysis will be performed by inner join ‘orders’ with ‘companies’ and other possibly related tables, through examination of columns like ‘updated_at’ and possible factor analysis.

Furthermore, to understand the impact of Craver, we will compare the customer growth pattern of Craver merchants before and after the installation of Craver by joining tables including ‘square_payments’, ‘orders’ and so on. We will define a SQL function to distinguish orders placed through Craver or other platforms by examining columns like ‘payment_detail’.

Objective 2: Understanding Customer Engagement and Retention

To understand the differentiation between sticky and non-sticky users, we will primarily rely on dataset tables containing information on orders, users, and coupons usage. Our approach involves defining and analyzing the behavior of these distinct user types. Initially, we will conduct a Frequency of Orders Analysis by examining the ‘orders’ table and grouping data based on user IDs to calculate the frequency of orders for each user. This analysis will be performed using a GROUP BY query in SQL to count the number of orders made by each user. Additionally, we will perform a Time Intervals Analysis to analyze the time intervals between orders (‘placed_at’ column) for each user, calculating the average time between consecutive orders to identify any patterns or differences in order frequency over time between sticky and non-sticky users. Furthermore, we will explore the impact of coupons on user stickiness by examining data from the ‘order_coupon’ table and analyzing coupon details ('title' column) to assess their influence on user behavior.

Objective 3: Impact of Restaurant Type on Merchant Success

The companies, locations, orders, and users tables will be the primary tables required to explore the success of merchants based on the restaurant type. To begin, each restaurant location will need to be classified as a restaurant type. There will be five classifications — coffee shop, bakery, quick-service restaurant, food truck, and other. We classify the location because some companies have a variety of restaurant types under their umbrella (i.e. a company may have a coffee shop and a food truck). Without going through the entire dataset and manually entering a location type, we can use a few name-based rules to classify locations. Coffee shops will have the words ‘cafe’, ‘coffee’, ‘bean’, or ‘brew’ in the name. Food trucks are labeled as such in the location name field. Bakeries will have the word bakery or bread in it and others will include juice shops and retail. All other restaurants will be considered fast-service restaurants. This is a slightly crude way of classifying the restaurants, but given the sheer number of locations, it offers the most efficient way possible. SQL’s LIKE filter will be used to create these rules.

After classifying each location, we can explore different metrics of success on the groupings and determine the significance of the results with statistical tests. Some of the metrics to explore include the number of order items per order, the total number of orders, the total revenue from app purchases, the number of unique users, and the number of sticky users. These metrics can be averaged over the month and the number of restaurants in each grouping. To calculate these metrics for each location, the locations table can be joined with the users or orders tables. The resulting table can then be grouped by restaurant type. We may consider visualizing the distributions as well.

Objective 4: Characteristics of High-Revenue Customers

To identify the common features or behaviors among the upper quartile customers, we will focus on the tables containing information on orders, users and companies. Initially, we'll employ SQL’s SUM function to compute total revenue per merchant which is present in the calculated_total column from the orders table. Subsequently, we’ll divide the data into quartiles based on their revenue and filter to spotlight merchants within the upper quartile, gauged by the percentage of revenue derived from Craver transactions. These merchants will be known as successful high revenue merchants. The success is defined based on how much of the revenue generated is through Craver. We will separate the lower quartile and the upper quartile customers into two tables to compare them and identify any qualitative or quantitative features that are capable of differentiating these two types of customers. We will also track the frequency of orders made by the high-revenue customers to assess their loyalty and engagement.

Objective 5: Cart Abandonment Analysis and Purchase Trends

While our focus is Craver merchants equipped with all kinds of payment integration services(i.e. Square, Stripe, Toast and so on), we will first conduct relevant analysis on those with Square payment service given the availability of related tables(e.g. Table ‘square_payments’).

To identify the trends correlated with cart abandonment and completion, we will utilize the 'orders' table and columns like 'placed_at'. The null value of this column suggests cart abandonment of customers. First we will calculate the cart abandonment rate. Additionally, we will look into factors contributing to purchase abandonment including time, basket size, merchant type, intended items and so on.

Deliverables, Statement of Work, and Timeline:

Our collaboration with Craver involves conducting thorough analysis using BigQuery to derive actionable insights into consumer behavior and merchant success factors. With access to Craver's BigQuery database and support from management, our roles encompass data analysis, project management, and communication, ensuring alignment with stakeholder objectives.

To navigate the project efficiently, we've outlined a timeline from May 1, 2024 to June 20, 2024, with key milestones including a mid-project presentation on May 27, 2024 and the final presentation on June 25, 2024. We propose fortnight meetings with Michael Jones to provide progress updates and ensure project alignment. Our weekly goals will be structured to meet the evolving needs of the project, ensuring consistent progress towards our objectives.

Our team, comprising individuals skilled in data analysis, statistical analysis, proficiency in data manipulation tools, project management, leadership, communication, problem-solving, business acumen, market research, understanding of consumer behavior, data visualization, database management, and familiarity with BigQuery and related technologies, includes Christopher Mulya, Jacob Rosen, Jade Yu, Natalie Coutinho and Pranav Kumar Mahesh.

Craver has prioritized five key insights for our focus, each of which aligns with specific scope questions. Additionally, each team member will be responsible for part of the scope questions. For Objective 1, Natalie Coutinho is assigned part a, c, and e , while Jade Yu oversees parts b and d of Objective 1, as well as Objective 5. Christopher Mulya, Jacob Rosen and Pranav Kumar Mahesh will be responsible for Objectives 2, 3 and 4 respectively. Meeting minutes will be rotated among team members to ensure thorough documentation. While each team member has assigned responsibilities, collaboration and support among team members are encouraged to address any challenges encountered across different questions.

TASK	SUBTASK	TASK OWNER	PCT OF TASK COMPLETE	WEEK 1					WEEK 2					WEEK 3					WEEK 4				
				M	T	W	TH	F	M	T	W	TH	F	M	T	W	TH	F	M	T	W	TH	F
Project Conception and Initiation	Project Kickoff Meeting	All	100%																				
	Develop Project Plan and Objectives		100%																				
	Proposal for MDS - Draft		85%																				
	Proposal Approval by Craver		0%																				
	Proposal for MDS - Final		50%																				
Data Preparation	EDA	All	30%																				
Conducting Analyses	Key Objective 1	Natalie / Jade	0%																				
	Key Objective 2	Christopher	0%																				
	Key Objective 3	Jacob	0%																				
	Key Objective 4	Pranav	0%																				
	Key Objective 5	Jade	0%																				
Deliverables Preparation	Dashboard Development	All	0%																				
	Final Report		0%																				
Presentation	Prepare Midterm / Final Slide Deck	All	0%																				
	Presentation Midterm / Final		0%																				

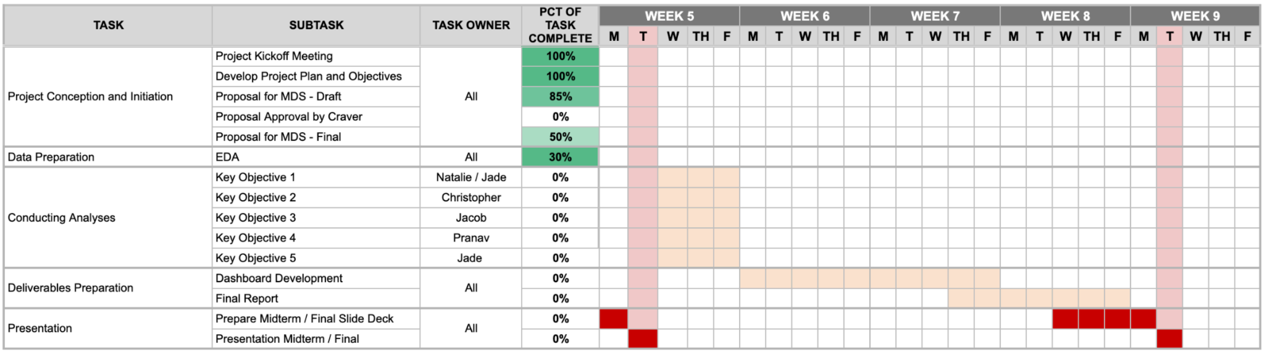

Figure 1. Capstone Project Timeline (Gantt Chart)

For deliverables, our team will focus on creating an interactive dashboard utilizing Looker, providing real-time insights into consumer behavior and merchant success metrics. This dashboard will serve as a dynamic tool for stakeholders to visualize and explore key findings from our analysis.

Additionally, we will produce a comprehensive project report documenting our methodology, findings, and recommendations. This report will offer a detailed overview of our analysis process, including insights derived from the data, actionable recommendations for Craver and its merchants, and potential areas for further exploration. It will serve as a valuable resource for stakeholders to understand the implications of our findings and guide future decision-making efforts.

Conclusion:

In conclusion, our research proposal outlines a comprehensive plan to analyze consumer behavior and merchant success within Craver's ecosystem. By leveraging Craver's extensive dataset and utilizing advanced data analysis techniques, we aim to derive actionable insights that will drive improvements in customer engagement and profitability for Craver and its merchants.

Through a systematic approach encompassing objectives such as consumer behavior analysis, customer engagement understanding, and evaluation of merchant success factors, we will uncover valuable insights to guide strategic decision-making. Our proposed methodology includes data preparation steps, detailed analyses, and statistical tests to ensure robust findings.

Furthermore, our proposed timeline and deliverables demonstrate our commitment to delivering tangible results within a specified timeframe. The creation of an interactive dashboard using Looker and a comprehensive project report will provide stakeholders with valuable tools and insights to optimize business strategies and enhance performance.

Moreover, by providing these essential tools and insights, we aim to empower Craver and its stakeholders to make informed decisions, optimize strategies, and propel the success of their businesses to new heights.

References:

Tigani, J., & Naidu, S. (2014). *Google bigquery analytics*. John Wiley & Sons.

Verhoef, P. C., Kannan, P. K., & Inman, J. J. (2015). From multi-channel retailing to Omni-channel retailing. *Journal of Retailing*, *91*(2), 174–181. https://doi.org/10.1016/j.jretai.2015.02.005

Wedel, M., & Kannan, P. K. (2016). Marketing analytics for data-rich environments. *Journal of Marketing*, *80*(6), 97–121. https://doi.org/10.1509/jm.15.0413