Oulin Yu, Tasnim Iqbal,
Shrabana Hazra, David Stifler

Project 1: Product Diversification

Executive Summary

The owner of a 50-seat café near the University of Nebraska, Omaha, is considering expanding the menu to include alcoholic beverages. Alongside this, they are exploring extending business hours from the current 6 AM - 3 PM to midnight, aiming to attract a broader market and increase patronage.

Objectives

By conducting comprehensive market research using surveys and analyses of secondary data, we evaluate the feasibility of the expansion. A summary of our findings follows below while more complete research outcomes are presented in the appendices of the market research proposal. In particular, we address the following points:

- Size, demand and saturation of a potential market for the business expansion.
- Potential increases in revenues and expenses associated with this expansion.
- Overall viability and profitability of undertaking this expansion.

Values Added

- Better customer satisfaction by catering to diverse needs and creating a campus social hub
- Attracting new customer segments and increase existing customer loyalty
- Maximizing space utilization and increase revenue streams

Target Market

Our target market comprises UNO students over 21, around 7000 people. While we expect to capture some of the faculty, staff, and local population as well, we do not believe that these groups are our target market and do not anticipate a statistically significant increase in patronage from it.

Competition

Total bars near or on campus: 15 (8 within 10 min walk) with average patrons of 100 - 150 per night.

Risks

- Regulatory compliance issues related to alcohol sales
- Competition with existing bars and cafes
- Revenue depends on the time of the academic year

Opportunities

- Attracting a large customer base of 7000 potential customer, hence increased revenue
- Distinguishing the cafe as a unique third space offering both coffee and alcohol, catering to a niche market

Results

Survey results from 200 people indicate that 60% of students go out often to socialize in the evening or late night. Nearly all of these students also stated that they drink alcohol during

social events. 90% of the students believe that location is important, and the majority prefers the campus. Furthermore, 45% of respondents indicated that they believed current offerings on or near campus could stand to be expanded. Finally, our revenue model indicates that in the optimistic case, up to 300 drinks can be sold each day, generating up to $500 of additional net profit. However, it should be noted that these numbers act as an upper bound estimate, and the true revenue can be strongly affected by the academic calendar.

Conclusion

Our current offerings and business hours do not capture this market segment, and therefore we believe there is a large potential customer base to be served by the proposed business expansion. The expansion can meet the diverse needs of students, making the cafe a unique destination.

Market Research Proposal

We propose conducting comprehensive market research to evaluate the feasibility of extending the café's operating hours from 3 PM to midnight and introducing alcohol to the menu. This research will provide data-driven insights into market demand, potential customer base, and competitive landscape, enabling us to make an informed decision.

Objectives

The objectives are summarized below:

- Market Demand
  - Conduct a survey among students and staff to assess their interest in purchasing alcohol during the proposed extended operating hours on campus.
  - Analyze social media platforms and online reviews of nearby establishments that sell alcohol to gain insights into current customer preferences and trends.
- Market Size
  - Research the potential customer base by analyzing the number of students, staff, and local residents on campus and in the surrounding areas.
  - Examine demographic data to identify the target age group and assess their spending habits, focusing on those most likely to frequent the café during the proposed extended hours.
- Market Saturation
  - Investigate the presence of nearby establishments that sell alcohol to evaluate the level of competition.
  - To determine if there is room for our shop to enter the market, conduct a competitive analysis to assess the strengths, weaknesses, and market share of existing establishments selling alcohol in the area.

Data Collection Methods

We will gather data through two methods: conducting *online surveys with questionnaires* that were distributed to UNO students and staff, as well as *analyzing secondary data from industry reports and market studies*. These approaches will provide both direct customer insights and broader market trends to inform our analysis.

Research Outcomes

The research outcomes are summarized and categorized into three parts:

1. The survey questionnaire along with the answers collected from 200 people from UNO.
2. Saturation research regarding different bars within 15 minutes walking from the UNO campus. This is collected from various resources such as reviews and Google Maps
3. A revenue model that is based on secondary data found on the Internet regarding UNO students and general drinking habits among university students in North America.

Each of these outcomes are appended below.

Conclusion

We believe this research will provide valuable insights, enabling us to make an informed decision regarding our expansion plans. These findings will guide our strategy, ensuring it aligns with market demand and opportunities. From these, we present the executive summary.

Appendix A: Survey results

AI-generated data including the location and demographic behavior of 200 people response:

Survey Respondents: 200
Student Year:

- First year (freshman): 42
- Second year (sophomore): 34
- Third year (junior): 29
- Fourth year (senior): 24
- Graduate student: 21

Employee Position:

- Tenured or tenure-track faculty: 11
- Non-tenure-track faculty: 9
- Academic support staff: 10
- Non-academic staff: 11
- None of the above: 4

Social Activities:

- Go dancing: 49
- Listen to music or see a play: 72
- Play sports or participate in group exercise: 79
- Make art or play music: 38
- Go to bars, cafes, or restaurants: 92

Preferred Locations for Social Activities:

- Clubs: 31
- Bars: 59
- Restaurants: 82
- Cafes: 68
- None of the above: 21

Importance of Location:

- Very important: 99
- Somewhat important: 81
- Not important: 20

Frequency of Alcohol Consumption:

- Almost every time: 39
- Some of the time: 81
- Rarely: 52
- Never: 28

Number of People Typically Met With:

- None - I go out solo: 21
- 1-3 people: 88
- 3-6 people: 73
- 6 or more people: 18

Preference for Social Event Locations:

- Close to campus: 109
- Far off campus: 39
- Either/both: 52

Preferred Time for Socializing:

- Morning: 11
- Mid-day: 19
- Afternoon: 51
- Evening: 81
- Late night: 38

Typical Spending on Social Activities:

- Less than $10: 21
- $10-30: 59
- $30-50: 68
- $50-100: 32
- More than $100: 20

Frequency of Nonalcoholic Specialty Drink Consumption:

- Almost every time: 48
- Some of the time: 83
- Rarely: 39
- Never: 30

Rating of Social Places Near Campus:

- Very good - plenty of places and variety: 39

- Pretty good - enough places and variety: 71
- So-so - it would be nice for more places to open: 58
- Pretty bad - only a few places and/or they're very similar: 19
- Very bad - almost nothing interesting and almost no variety: 13

Typical Number of Drinks at a Bar:

- 1: 49
- 2: 58
- 3: 41
- 4: 21
- 5+: 12
- N/A: 19

Time Spent at a Bar:

- 1 hour: 19
- 2 hours: 51
- 3 hours: 78
- More than 3 hours: 33
- N/A: 19

Average Times Going Out to a Bar:

- 1 to 2 times: 48
- 3 to 4 times: 79
- More than 5 times: 41
- N/A: 32

Appendix B: Saturation Research

The saturation research presented below is collected from various sources (see references) with bars that are within 15 minutes walking from the UNO campus.

Example Secondary Market Survey Data for Bars
1. Number of Bars in the Area:

- Total Bars: 15
- Bars within a 15-minute walk from campus: 8

2. Average Number of Patrons per Night:

- Bar 1: 120 patrons
- Bar 2: 150 patrons
- Bar 3: 100 patrons
- Bar 4: 130 patrons
- Bar 5: 140 patrons

3. Crowd Levels:

- Always Crowded: 3 bars
- Sometimes Crowded: 3 bars
- Rarely Crowded: 2 bar

4. Types of Patrons:

- Students: 60%
- Faculty/Staff: 20%
- Local Residents: 20%

5. Average Spending per Patron:

- Less than $10: 10%
- $10-30: 50%
- $30-50: 30%
- More than $50: 10%

6. Availability of Coffee and Alcohol:

- Bars Offering Both: 3
- Bars Offering Only Alcohol: 2
- Bars Offering Only Coffee: 0

7. Operating Hours:

- Bars Open Until Midnight: 3
- Bars Open Until 10 PM: 2
- Bars Open 24 Hours: 0

References:

1. https://www.visitomaha.com/blog/post/top-10-places-to-grab-a-drink-in-omaha-according-to-yelp/
2. https://www.djsdugout.com/
3. https://spiritworldwine.com/
4. https://lcc.nebraska.gov/rules-and-regulations

Appendix C: Revenue Model

In this section, we will explore how much revenue the expansion of the business would generate. This, in turn, would give the net profit.

It should be noted that at the early stage of this business expansion, it is reasonable to consider only simple drinks such as beer, seltzer and wine. Around 40% of students in North America stated that these simple drinks are their preferred drink [1]. While cocktails do sound appealing, they tend to be more expensive for students, and serving these requires more qualification. Importantly, for the owner, it requires a larger initial investment, rearranging the cafe space, and hiring bartenders. Consequently, we assume only simple drinks would be served at this first stage of the expansion.

In such a simplified model, the revenue is directly proportional to the number of drinks that can be sold by the cafe. We can estimate the number of drinks served with two approaches: top-down or bottom-up methods.

Top-down method:
UNO has ~15000 students and roughly ~7000 of them are above the legal age. Male and female students drink on average 4 or 9 drinks per week, respectively [2]. Given that the UNO student population has a male to female ratio of ~45:55 [3], the average number of drinks per week per student is 6.
As a result, if 5% of the students come to the cafe regularly (to be confirmed with the survey), then *on average 300 drinks would be served daily*.

Bottom-up method:
As with many other cafes found on or near the campus, we are bounded by the number of available seats. From this we can estimate an upper bound to the number of drinks served. Assuming that each client consumes on average 1 drink per 45 minutes, this translates to *no more than 600 drinks* served during the additional 9 hours of operation. However, the business rarely runs at maximum capacity, and in the most optimistic case, this number should be halved. As such, a reasonable upper bound is *300 drinks per day*.

Revenue:
Both the top-down and bottom-up approaches point to a magic number of 300 drinks served per day. The markup is generally around 200-300% [4], and given the campus setting, it is safe to assume a 200% markup, or 50% profit. Looking at the average pricing near ONU, a typical beer selling for 7$ generates 3.5$ of profit. This yields an upper bound to the revenue of around 1000$.

Additional cost:
Assuming 1 server per 25 guests, at least 2 servers are needed, plus a person in the back, and perhaps a bouncer later at night, we can expect a minimum of 4 additional employees with 10 hours shift (additional time for setting things up and cleaning). With a minimum wage of 12$/h, the total cost for employees is ~500$. Further expenses are however expected.

Net profit:
Combining revenue and additional cost for the expansion of the business, the net profit for the owner is 500$ per day. However, there are some caveats to be considered. First, this number is an upper bound *and* assumes the business is doing well. How the competition (attendance, promotion, etc.) plays out with other bars in the area is unknown, but can be perhaps clarified by the survey. Moreover, bars, especially one on the campus, are strongly affected by the academic calendar.

Towards more detailed estimates:
We can use the survey results and additional resources [5] to have a more detailed calculation. However, the above text gives a good idea of a first order approximation to the potential of such business expansion.

References
[1] https://www.barandrestaurant.com/people/survey-reveals-what-college-students-want-drink
[2] https://www.scramsystems.com/blog/2014/02/infographic-much-college-students-drink/

[3] https://www.usnews.com/best-colleges/university-of-nebraska-omaha-2554/student-life
[4] https://home.binwise.com/blog/beer-pricing-for-bars
[5] https://www.bestcolleges.com/research/alcohol-use-in-college-statistics/

Oulin Yu, Tasnim Iqbal,
Shrabana Hazra, David Stifler

Project 2: Website Conversion

Executive Summary

This company is a direct-to-customer online shopping site that specializes in exclusive and high performance kitchen tools, especially knives. It differentiates itself in the market because of its curation and expertise and relationships with manufacturers.

The goal of this report is to study why clients often put items into the cart without checking out through survey interviews.

Finding Users

The interviewees were selected randomly on the website, and prompted with a survey for them to complete. In this sampling, we expected the majority are people who add items but do not check out while still including few users who do purchase and can tell us what pushed them over the edge. The goal of the interview is to understand what parameters dictate the checkouts.

Interview Design

To gain insights into why users abandon their carts, we needed to design an interview that focuses on the user experience, pain points, and barriers during the shopping process. Here is a proposed interview design, including sample questions and expected answers:

Structure

1. **Introduction**: Brief explanation of the survey purpose, estimated time to complete (around 5-7 minutes), and assurance of data privacy.
2. **Demographic Questions**: Understand user profiles (age, gender, device usage).
3. **Behavioral Questions**: Assess users' shopping habits and experiences on the website.
4. **Experience and Feedback Questions**: Identify specific issues related to cart abandonment.
5. **Closing Questions**: Open-ended feedback and suggestions.

The questions and responses are found in the appendix at the end of this document.

Interview Outcome

Here are the **key findings** categorized by common themes identified during interviews. The findings reflect typical user feedback when analyzing cart abandonment issues on e-commerce websites.

- **User Experience Barriers**:
    - **Complicated Checkout Process**: Users are frustrated with the long, multi-step checkout process.
    - **Mandatory Account Creation:** Users expressed annoyance with mandatory account creation before making a purchase.

    - **Navigation Issues**: Users struggle with finding items or information about them, leading to cart abandonment.
    - **Overwhelming Selection:** Users described difficulties with choosing the right product for their needs, and wished they had a way to filter based on their use case to find exactly the right knife for them.
- **Trust and Credibility Issues**:
    - **Concerns About Website Security:** Some users were wary about entering their payment information due to concerns about the website's security. They felt unsure about the site's legitimacy or data protection measures.
    - **Lack of Product Reviews and Information**: Users are hesitant to purchase products without sufficient customer reviews.
- **Pricing and Payment Concerns**:
    - **High and Unexpected Shipping costs:** A significant number of users mentioned that they abandoned their cart when they saw unexpectedly high shipping fees. Many users added items to their cart to check the total cost, including shipping, and left when it was too high
    - **Comparison Shopping**: Many users compare prices on other websites before completing a purchase, often finding better deals elsewhere.
    - **Lack of Visible Discounts and Promo codes**: Users expect to see discounts or promotional offers during checkout but often abandon the cart when they don't find them.
    - **Limited Payment Options:** Users expressed frustration with the limited number of payment options available. Some preferred using digital wallets (e.g., Apple Pay, Google Pay, or PayPal), but the website only supported credit cards.

note: these responses are conjectural, as this project is meant to simulate the process rather than reflect real-world information or circumstances

Personae

The survey allowed the identification of three user segments that are illustrated by three different personae:

1. John is a professional chef who is seeking to buy a new knife for the next couple of years. He represents one user segment who would shop at a company that specializes in premium knives. John is already very knowledgeable about knives. It is clear to him that a knife from this company would be one of his top picks, but before pulling the trigger he would like to compare this option with a few other companies.

John

Job Title

Head Chef

Age

35 to 44 years

Highest Level of Education

Associate degree (e.g. AA, A

Industry

Food & Beverage

Organization Size

1-10 employees

Goals or Objectives

Working towards his first Michelin Star
Revenue for the Restaurant

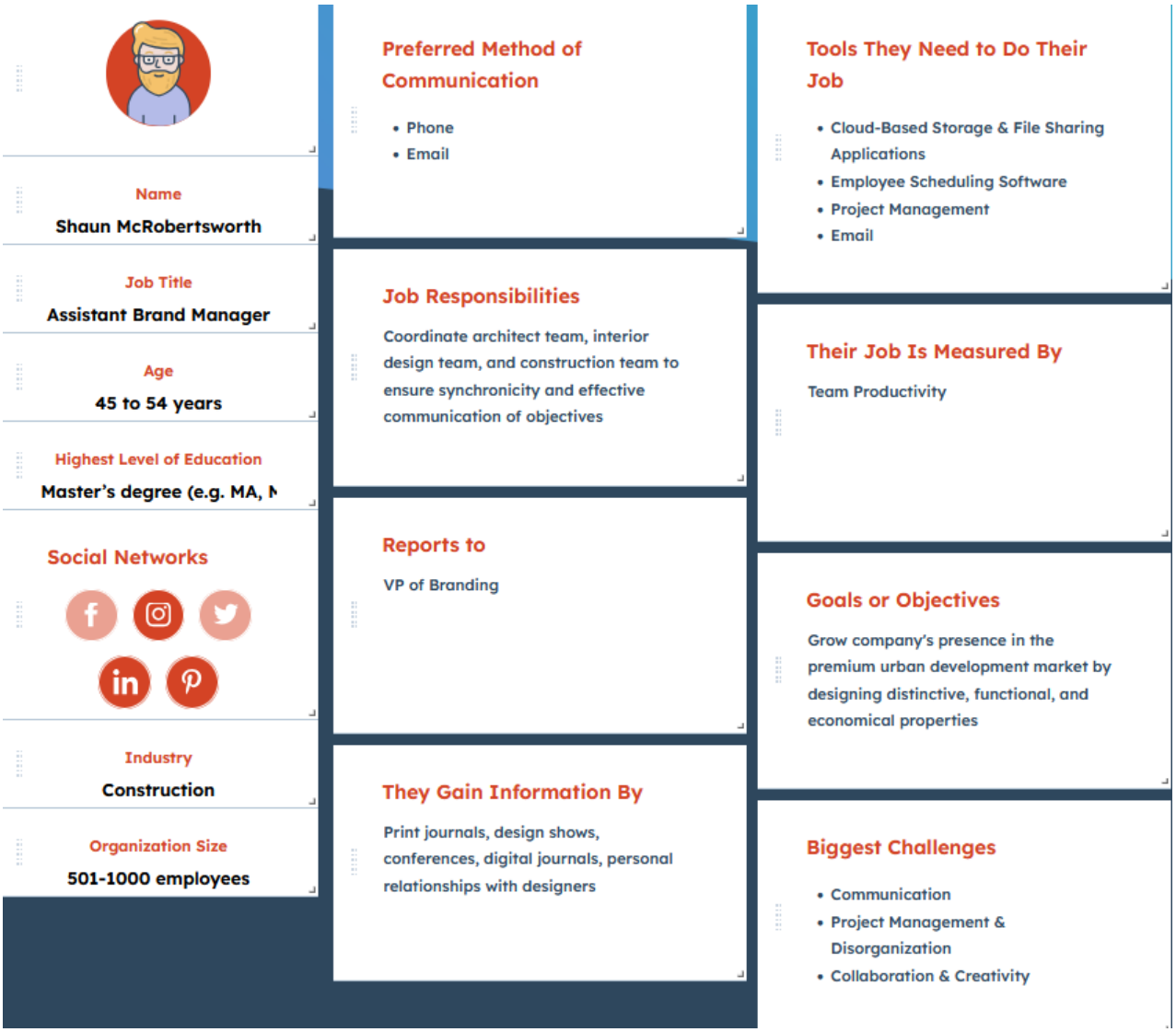
Job Responsibilities

People management
Quality control
Creation of new recipes

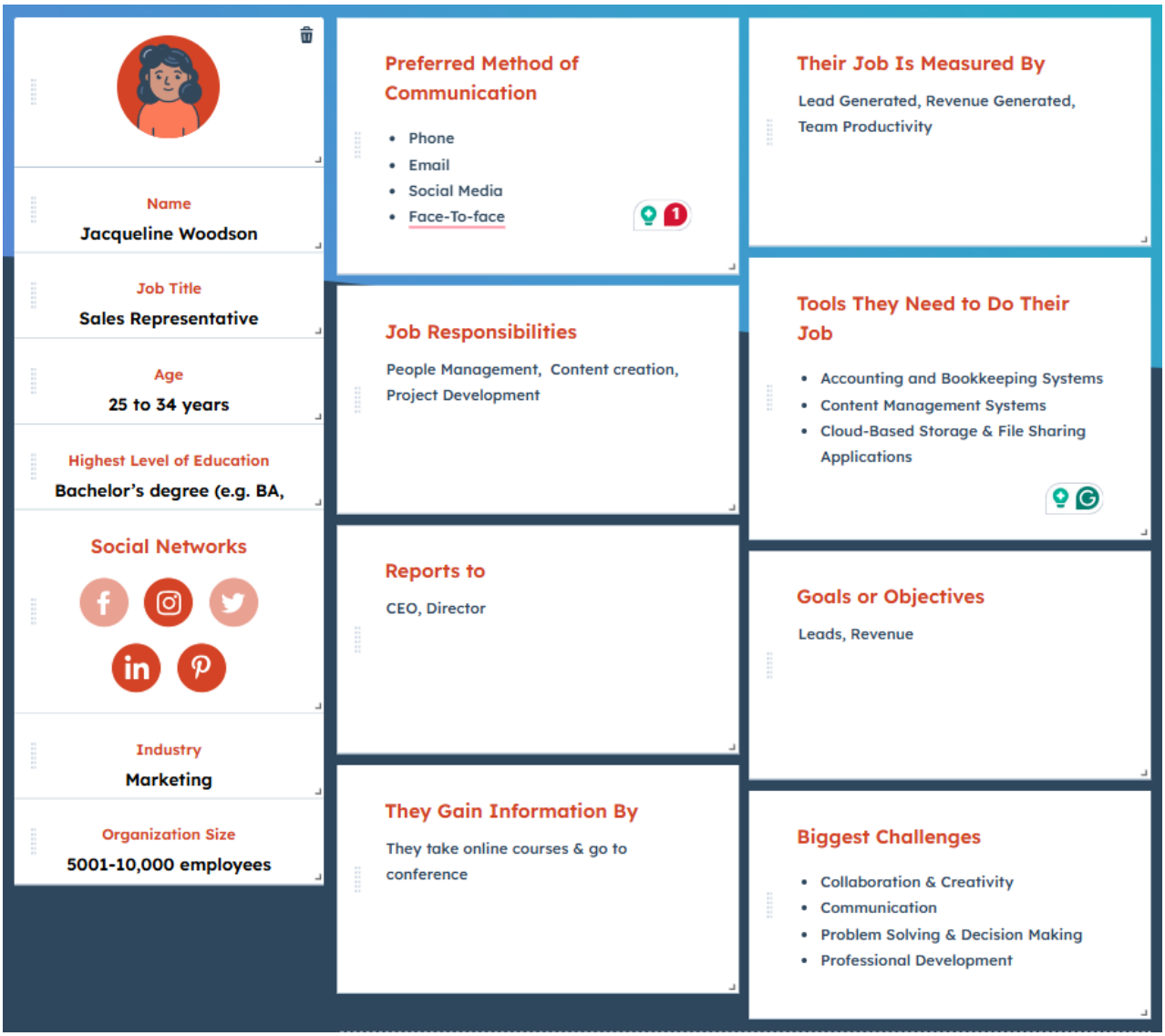
Biggest Challenges

Time coordination
Smooth management of people

2. Shaun is an amateur enthusiast / hobbyist cook. As a member of the upper middle class, he wants “the best” but is very judicious about spending money, and follows the maxim “buy once, cry once”.



3. Jacqueline, a young Sales Representative in marketing, values high-quality kitchen tools like these. She expects a smooth shopping experience with clear product info, secure checkout, and visible discounts. However, unexpected costs, limited payment options, or security concerns may lead her to abandon her cart.



Storyboarding

We have created three storyboards based on the above personae:

https://miro.com/app/board/uXjVLHFMpGQ=/

Conclusion

This UX research project revealed four key barriers that prevent users from completing their purchases on the e-commerce website, leading to high cart abandonment rates. The main issues identified were

1. a complex and lengthy checkout process
2. unexpected additional costs (such as shipping fees)

3. poor mobile usability, and
4. overwhelming selection, absent tools to help select the right product for users' needs.

To address these problems, the company should focus on streamlining the shopping process in four different but equally important ways. These include:

1. offering guest checkout - users should be able to purchase without having to sign up or become members
2. enhancing pricing transparency - users shouldn't feel as though they are surprised by additional costs between putting an item in their cart and checking out
3. optimizing the mobile shopping experience - the site should be easily navigable on all devices, since many people shop using phones and tablets rather than laptops or desktops
4. guiding users through the selection process - even the most experienced customers could be having their first interaction with this company and its products

Additionally, implementing user-friendly features like a "Save for Later" option and expanding payment methods could further reduce friction in the purchasing process. Implementing these changes will not only reduce cart abandonment but also foster greater user confidence and satisfaction, ultimately boosting conversion rates and customer loyalty.

Appendix: Interview questions

1. Demographic Questions

These questions help us categorize the user base and understand different user behaviors.

- **Q1**: What is your age group?
  - **Options**:
    - 18-24
    - 25-34
    - 35-44
    - 45-54
    - 55+
  - **Expected Answer Distribution**: Varies, with a focus on 25-44 (main shopping age groups).
- **Q2**: What device do you usually use to shop online?
  - **Options**:
    - Mobile Phone
    - Desktop/Laptop
    - Tablet
  - **Expected Answer**: Majority use mobile phones, followed by desktops/laptops. This finding has correlation with the previous survey question where it was noted that the younger crowd tend to use mainly mobile phones.
- **Q3**: How often do you shop online?
  - **Options**:
    - Weekly
    - Monthly
    - A few times a year
    - Rarely

- **Expected Answer**: Majority are likely monthly or weekly shoppers.

2. Behavioral Questions

Understanding the typical user journey and behavior on the website.

- **Q4**: Have you ever added items to your cart on our website but decided not to complete the purchase?
  - **Options**:
    - Yes
    - No
  - **Expected Answer**: Majority will answer "Yes," as the focus is on cart abandonment.
- **Q5**: What are the main reasons you add items to your cart without buying them? (Select all that apply)
  - **Options**:
    - Checking the total cost with shipping
    - Saving items for later
    - Comparing prices with other websites
    - Not ready to purchase
    - Difficulty in completing the checkout process
  - **Expected Answer**: High responses for "Comparing prices," "Checking total cost," and "Saving items for later."

3. Experience and Feedback Questions

Diving deeper into specific pain points during the checkout process.

- **Q6**: How would you rate your experience with the checkout process on our website?
  - **Options**:
    - Very easy
    - Easy
    - Neutral
    - Difficult
    - Very difficult
  - **Expected Answer**: Mixed, but a notable percentage may choose "Difficult" or "Neutral," indicating areas for improvement.
- **Q7**: What factors most influenced your decision not to complete the purchase? (Select up to 3)
  - **Options**:
    - High shipping costs
    - Long checkout process
    - Mandatory account creation
    - Lack of payment options
    - Lack of trust in the website
    - Product availability or delivery time
  - **Expected Answer**: High responses for "High shipping costs," "Long checkout process," and "Mandatory account creation."
- **Q8**: Did you encounter any issues with the website while adding items to your cart or during checkout?
  - **Options**:

        - Yes
        - No
    - **Follow-up (If Yes)**: What issues did you experience? (Open-ended)
    - **Expected Answer**: Many users may report issues like slow loading times, unclear buttons, or difficulty navigating.
- **Q9**: How important are the following factors when deciding to complete a purchase? (Rate on a scale of 1-5)
    - **Factors**:
        - Total cost (including shipping)
        - Ease of checkout
        - Availability of discounts/coupons
        - Trust in the website's security
        - Availability of reviews and product information
    - **Expected Answer**: High importance given to "Total cost," "Ease of checkout," and "Trust in website security."

4. Open-Ended Feedback

Gather qualitative insights and user suggestions.

- **Q10**: What would make you more likely to complete your purchase on our website in the future? (Open-ended)
    - **Expected Answers**:
        - More visible discounts and promotions
        - Simplified checkout process
        - Better clarity on shipping costs upfront
        - Option for guest checkout without account creation
- **Q11**: Is there anything you would change about our website to improve your shopping experience? (Open-ended)
    - **Expected Answers**:
        - Improve site speed on mobile devices
        - More detailed product descriptions and reviews
        - A way to filter the product selection based on intended application
        - Clearer information on return policies

Oulin Yu, Tasnim Iqbal,
Shrabana Hazra, David Stifler

Project 3: Customer Satisfaction

Proposal for Customer Satisfaction Survey

To: Jane Smith, Chief Executive Office,
Date: 12/08/2024

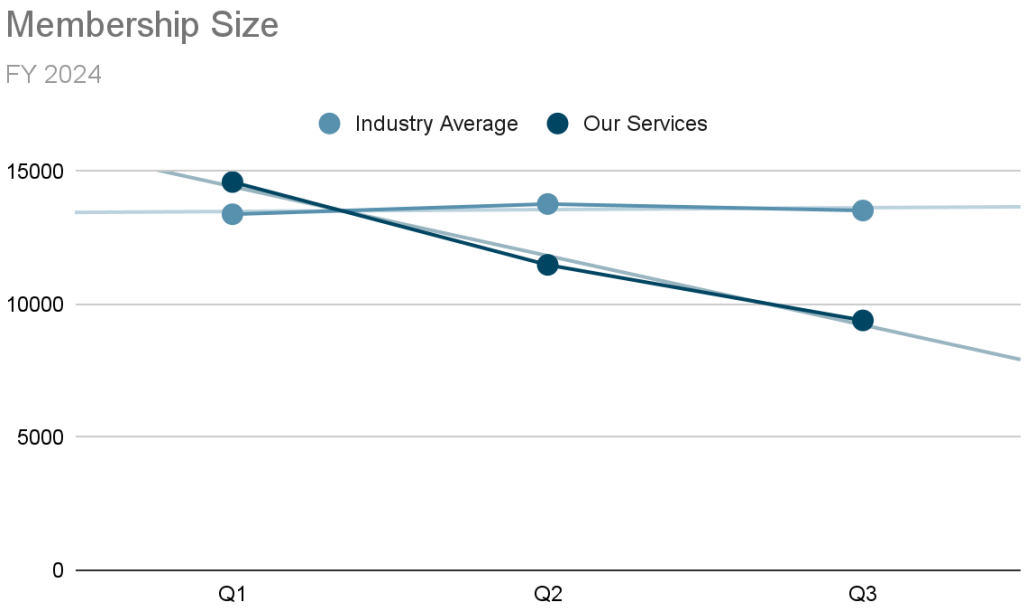
Subject:

Proposal for Customer Satisfaction Survey to Address Membership Churn

---

Overview:

Our makeup company has experienced a steady decline in membership subscriptions over the past three business quarters. As shown here, total membership has **decreased almost 36%** from Q1 to Q3 in FY 2024:



This decrease comes at a time when growth in the specialist makeup **industry as a whole** is generally flat, as the chart shows. The reasons for this **churn** are unknown, but we are confident they include some combination of the following:
Internal Factors

- customer satisfaction with product and/or service
- fashionability of products
- pricing model for memberships

External Factors

- inflation
- changing trends in fashion
- competition from other companies

To address this issue, we propose conducting a comprehensive customer satisfaction survey. The goal of this survey is to identify the key factors contributing to membership churn and to gather actionable insights that will help us improve our services and retain our members.

Survey Objectives:

1. Understand Demographics: Identify the age and income ranges of our members to better tailor our services.
2. Assess Customer Satisfaction: Measure satisfaction levels with various membership benefits and identify areas for improvement.
3. Evaluate Pricing Sensitivity: Determine how our members perceive the value of our membership pricing and the impact of potential discounts.
4. Identify Churn Drivers: Understand the primary reasons why members cancel their subscriptions and what features could retain them.

Proposed Survey Design:

1. Demographics:
    - Age Range: 18–25, 26–35, 36–50, 51+
    - Monthly Income Range: <$3,000, $3,001–$5,000, $5,001–$8,000, $8,000+
2. Customer Satisfaction:
    - Satisfaction with membership benefits (scale of 1 to 5)
    - Valued membership features (e.g., discounts, free shipping, exclusive product access, personal consultations)
    - Suggestions for improving membership experience
3. Pricing Sensitivity:
    - Perception of membership pricing (reasonable or not)
    - Impact of a 10% discount on membership reconsideration
4. Churn Drivers:
    - Reasons for canceling membership (e.g., pricing, service quality, lack of use, external financial pressures)
    - Features that could retain membership (e.g., more exclusive products, better rewards, lower pricing options, flexible payment plans)

Expected Insights:

1. Primary Churn Drivers: We expect to identify the main reasons for membership cancellations, with a particular focus on pricing and customer satisfaction.
2. Customer Satisfaction Levels: Insights into how satisfied our members are with their current benefits and what improvements they desire.
3. Valued Features: Understanding which membership features are most important to our members, allowing us to prioritize enhancements.
4. Pricing Sensitivity: Gauging how sensitive our members are to pricing changes and the potential impact of discounts on membership retention.
5. Retention Strategies: Identifying actionable strategies to reduce churn, such as adjusting pricing models, enhancing benefits, and improving customer service.

Conclusion:

By conducting this survey, we aim to gain a deeper understanding of our members' needs and preferences. The insights gathered will enable us to make data-driven decisions to enhance our membership offerings, improve customer satisfaction, and ultimately reduce churn.
We seek your approval to proceed with this survey and would appreciate any feedback or suggestions you may have.
Thank you for your consideration.

Market Research Team

AB Testing

To test whether the decline in memberships is due to the relative high cost, an AB testing with dependent groups was used. Each user was randomly (i.e. 50% chance) offered a 20% off discount while browsing the company's website for various services. The revenue generated for each user was recorded and classified as either control (no discount offered) or variant (discount offered).

For quantitative analysis, the recorded data are first cleaned to generate the summary statistics. The data are then visualized and, finally, hypothesis testing is performed. Below is the detailed quantitative analysis performed on the data obtained (Kaggle), and the Python code is found in the appendices.

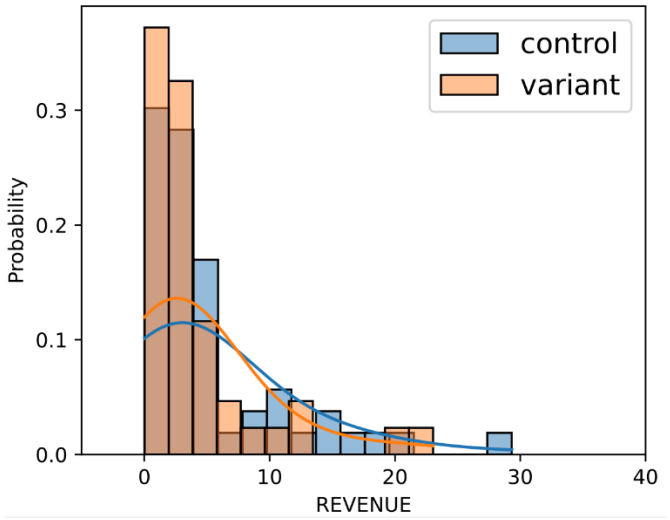
Data cleaning:
Among the 10000 collected entries, we first clean the data to increase the accuracy of our analysis. Specifically, the following operations were performed:

- Duplicating entries were removed
- Some data with same user ID yet belonging to different variant groups is inconsistent and hence removed
- An outlier was removed

Summary statistics:
The summary statistics of the cleaned data are shown below. It is noted that most users simply browse and do not make any purchase.

	Count	**Median**	**Mean**	**STD**
Control group	2413	0.0	0.113780	1.156753
Variant group	2407	0.0	0.074499	0.854413

It is inconclusive when one compares the means of the two groups. To gain further insight, we visualize the situation below.



Hypothesis Testing:

Even with the above figure, it is unclear which group is more likely to make a subscription. In addition, the histograms do not look normally distributed, which is confirmed with the normality tests (Shapiro-Wilk):

	Shapiro-Wilk test statistics	**p-value**
Control group	0.7513	0.0000
Variant group	0.7053	0.0000

As expected, for the standard choice of $\alpha = 0.05$, we reject the normality for both groups since $p < 0.05$. Thus, we can compare the means through the Mann–Whitney U test where the null hypothesis is the two groups generate the same revenue:

Mann-Whitney U test statistics	**p-value**
1249	0.4218

Importantly, in this test, since $p > 0.05$, it means that we cannot reject the null hypothesis, consistent with the histogram above.

Conclusion:

The data collected showed that there is no significant difference in the revenue when a 20% discount on subscription fees were randomly offered to different users. This contradicts the data collected from our survey analysis which showed most responses said that the pricing was not reasonable. Such discrepancy may be attributed to the fact that the discount offered during the AB testing was not aggressive enough. Further analysis, such as other criteria or more discount offers, is needed to unravel the exact reason for the decline in the membership rate.

Oulin Yu, Tasnim Iqbal,
Shrabana Hazra, David Stifler

Conclusion

In this document, we first present an overview of the problem and a survey proposal to the C-suite. An AB testing proposal is then addressed, followed by a full analysis of the data from Kaggle.

Finally, to facilitate the communication of the survey and AB testing outcomes, here is a powerpoint summarizing all the results:
https://docs.google.com/presentation/d/1hg5FiTZLHZJbfJT9WG0Io0scrMk-MYW4jNk88rAYprE/edit?usp=sharing

Appendix: Python code for the AB Testing Analysis

```
import pandas as pd
import seaborn as sns
import matplotlib.pyplot as plt
import numpy as np
from scipy.stats import shapiro, mannwhitneyu


# import and cleaning
df = pd.read_csv('AB_Test_Results.csv') # load csv
df = df.drop_duplicates() # remove duplicate items
# remove duplicate where a user belongs to multiple variant
duplicates = (df.groupby('USER_ID')['VARIANT_NAME'].nunique() == 1)
duplicates = duplicates[duplicates]
df = df[df['USER_ID'].isin(duplicates.index)] #
df = df[df['REVENUE'] < 100] # remove outlier
# print stats
print(df.groupby('VARIANT_NAME').agg({'USER_ID': 'nunique',
                                      'REVENUE': ['sum', 'mean', 'median',
'count']}))


# plot
f, axes = plt.subplots(1, figsize=(5,4))
plt.rcParams.update({'font.size': 14})
n = 2
variant = df.loc[(df['VARIANT_NAME'] == 'variant') & (df['REVENUE'] > 0), 'REVENUE']
control = df.loc[(df['VARIANT_NAME'] == 'control') & (df['REVENUE'] > 0), 'REVENUE']
sns.histplot(control, stat="probability", kde = True, kde_kws={"bw_adjust":n},
label='control', binwidth=n)
sns.histplot(variant, stat="probability", kde = True, kde_kws={"bw_adjust":n},
label='variant', binwidth=n)
```

```
plt.xlim([-5, 40])
plt.legend()
plt.subplots_adjust(hspace = 0.3)
plt.savefig('output.pdf')
plt.clf()


# normality tests
test_stat,p_value = shapiro(control)
print('Test Stat=%.4f,p-value=%.4f'%(test_stat,p_value))
test_stat,p_value = shapiro(variant)
print('Test Stat=%.4f,p-value=%.4f'%(test_stat,p_value))


# Non-parametric test
test_stat,pvalue = mannwhitneyu(control,variant)
print('Test stat = %.4f,p-value=%.4f'%(test_stat,pvalue))
```