BAMA 517 - Data-Driven Marketing Accenture Data Analytics & Visualization Capstone Project

Submitted by:

Rahul Dhomane	**19801844**
Roshan Shetty	**13426960**
Daksh Arora	**68334820**
Sumaita Ahmed	**57330508**
Bhairavi Wadekar	**12491247**

Under the guidance of:
Prof. Daisy Ya

UBC SAUDER
SCHOOL OF BUSINESS

SAUDER SCHOOL OF BUSINESS

UNIVERSITY OF BRITISH COLUMBIA

December 2023

Social Buzz Overview: Founded in 2010 by two former engineers from a significant social media conglomerate, Social Buzz, headquartered in San Francisco, has rapidly grown to engage over 500 million users monthly. Focused on elevating content, the platform emphasizes user anonymity and tracks diverse reactions—over 100 ways—beyond conventional likes and comments. With 200 technical staff among its 250 employees, Social Buzz manages an immense volume of unstructured data daily. To tackle their scaling needs, we embark on a three-month project to prove our advisory firm's expertise by analyzing the top 5 content categories with the highest aggregate popularity, offering critical insights for their ongoing success in the dynamic realm of social media and content creation.

1. **Executive Summary**
   Established in 2010, Social Buzz has rapidly grown to attract over 500 million monthly active users. The platform offers diverse engagement with over 100 ways for users to interact, resulting in a daily creation of 100,000+ pieces of content. Social Buzz maintains a secure, personalized digital environment focused on user privacy and anonymity. This synthesis of dynamic engagement, content diversity, and a commitment to user privacy solidifies Social Buzz's prominent position in the social media landscape. In response to our comprehensive analysis of Social Buzz, we propose implementing a unique strategy termed "Buzz Craft." This innovative approach combines content and medium crafting to elevate user engagement, mitigate negative experiences, and foster a dynamic and positive user environment

2. **Introduction**
   Since 2010, SocialBuzz has thrived with 500M monthly users, 100+ ways to engage, and 100,000+ daily content pieces. Our mission: amplify user experiences. The primary objective of our analysis is to identify and propose a robust strategy for SocialBuzz to elevate user engagement levels and refine the user experience. In the dynamic digital industry, achieving this objective involves addressing multifaceted challenges, including adapting to changing user preferences, consistently upgrading content offerings, and enhancing content quality. As we delve into this objective, we focus on formulating actionable recommendations that align with SocialBuzz's commitment.

3. **Problem Statement**
   This analysis aims to deliver actionable insights in order to boost user engagement and enhance user experience for SocialBuzz. User engagement, measured by interaction frequency, and user experience, encompassing reactions to content, are pivotal for platform success. The central question is: *"How can SocialBuzz increase user engagement and improve user experience?"* This analysis will investigate current user engagement, explore user reactions, and propose strategies for personalized content and medium of content delivery. The goal is to equip SocialBuzz with data-driven insights to elevate user engagement and experience effectively.

4. **Methodology:**

   Data Collection:
   Data for this study were provided by Social Buzz and comprised three distinct datasets: User, Content, and Reaction.

- User Dataset:

- ID: A unique identifier automatically generated for each user.
        - Name: The full name of the user.
        - Email: The email address associated with each user.
- Content Dataset:
        - ID: A unique identifier automatically generated for each uploaded content.
        - User ID: A unique identifier of the user who uploaded the content, corresponding to the User dataset.
        - Type: A categorical string describing the type of content.
        - Category: A categorical string specifying the relevant category of the content.
        - URL: A link to the location where the content is stored.
- Reaction Dataset:
        - Content ID: A unique identifier linking each reaction to the content it pertains to.
        - User ID: A unique identifier of the user who reacted to the content, matching a user in the User dataset.
        - Type: A categorical string describing the type of reaction.
        - Datetime: The date and time when the reaction occurred.
- ReactionTypes Dataset:
        - Type: A categorical string specifying the type of reaction.
        - Sentiment: A categorical string categorizing reactions as positive, negative, or neutral.
        - Score: A numerical value calculated by Social Buzz, quantifying the popularity of each reaction type. Higher scores indicate more popular reactions.

Data cleaning and manipulation:

Data cleaning and manipulation steps were essential to ensure the data's quality and consistency, allowing for meaningful analysis of user-generated content and reactions within the Social Buzz platform. Several steps were taken to prepare the data for analysis:

- Column Removal: In the Content dataset, the 'URL' column was removed as it was deemed irrelevant to the analysis.
- Category Standardization: Categories in the Content dataset were standardized to ensure consistency. Variations of the same category were mapped to a single keyword. This standardization helped reduce data noise.
- Missing Value Handling: All datasets were checked for missing values, and no instances of missing data were found.
- Outliers: Outliers were not removed from the Reaction dataset. This decision was made because users may express a wide range of sentiments, and removing outliers would disregard valid and diverse reactions, which are essential for a comprehensive analysis.
- Data Integration: The User, Content, and Reaction datasets were joined based on the 'User ID' field, facilitating integrated analysis across these datasets.
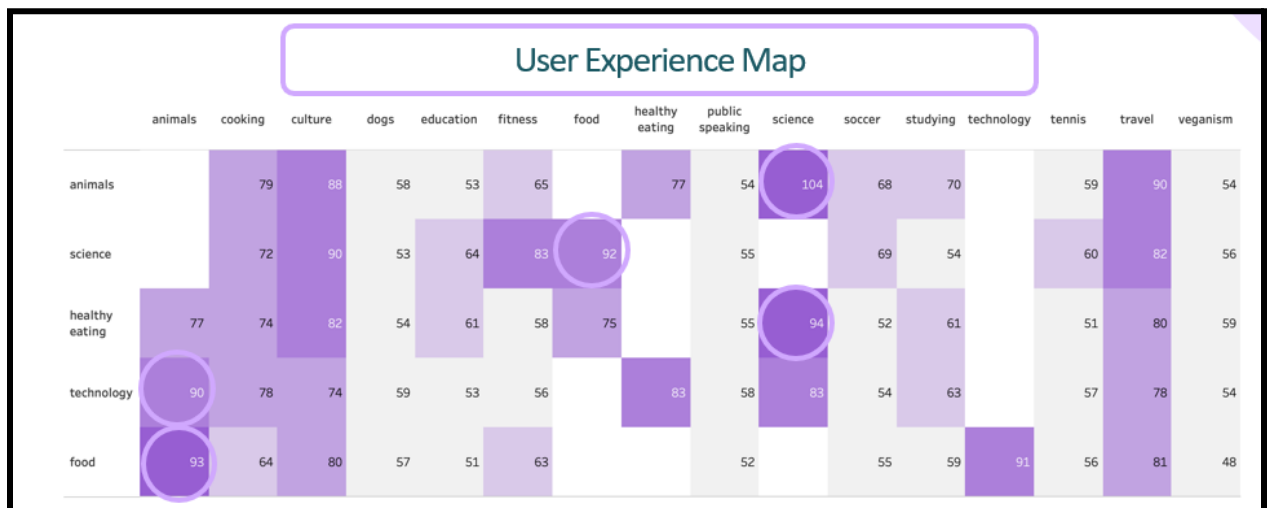
5. **Analysis and Recommend:**

The analysis utilized Python for data cleaning, Excel for manipulation, and Tableau for visualization, ensuring a seamless analytical process.

User Experience Map

Overview of the Analytical Framework:
A user experience map was generated in the form of a heatmap by categorizing users into two groups: those with above-average and below-average scores across various categories. Subsequently, we assessed the proportion of users who consistently provided above-average responses in multiple categories. The darker shades indicate a higher number of users with positive scores. For example, the cell at the intersection of "animals" for rows and "science" for columns is very dark and has the number 104, meaning a large number of users interested in animals also scored positively in the science category.

	animals	cooking	culture	dogs	education	fitness	food	healthy eating	public speaking	science	soccer	studying	technology	tennis	travel	veganism
animals		79	88	58	53	65		77	54	104	68	70		59	90	54
science		72	90	53	64	83	92		55		69	54		60	82	56
healthy eating	77	74	82	54	61	58	75		55	94	52	61		51	80	59
technology	90	78	74	59	53	56		83	58	83	54	63		57	78	54
food	93	64	80	57	51	63			52		55	59	91	56	81	48

Steps:
- Calculate the total score for each user for all the categories
- Calculate average of each category and tag number of users above the average and below the average as 1 and 0 respectively.
- Count the number of users who are interested in the top 5 and rest of the categories as well
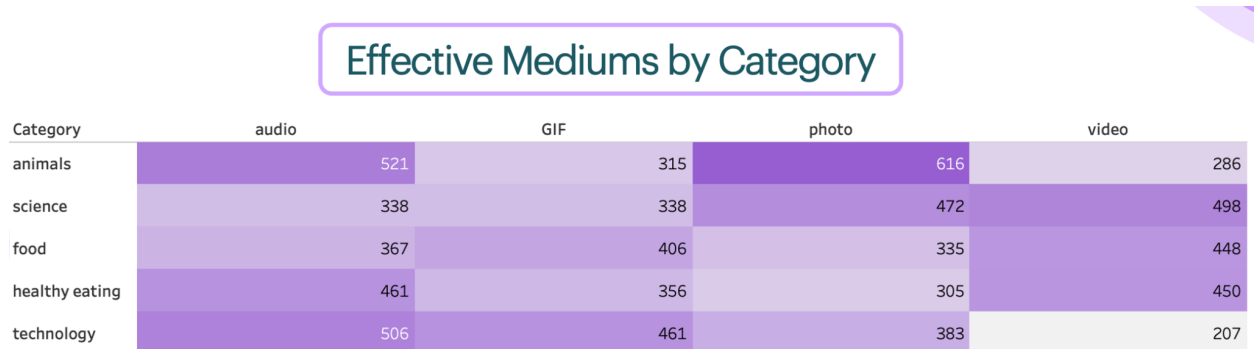
Insight and recommendation:
This visualization helps in quickly identifying which categories have the highest positive responses from users and how these responses correlate across different categories. Positive user reactions in one category consistently extend across diverse areas, revealing an interconnected user preference. Additionally, the user experience map highlights that followers of top categories also express interest in others. For example, animal enthusiasts often engage with science, and science followers show interest in food-related content.

Based on the insights gleaned from this visualization, we recommend that Social Buzz strategically optimize content promotions, curation, and user engagement strategies. Leveraging the observed patterns of interconnected user preferences, the platform can target promotions effectively, curate content based on related categories, encourage cross-promotions between content creators, and design engagement strategies that prompt users to explore a broader array of content. These measures will enhance the user experience and drive increased user engagement, contributing to sustained platform growth.

Effective Mediums by Category

Overview of the Analytical Framework:

The analysis focused on user preferences across categories, utilizing Tableau's count function. Specifically, we examined the number of times users engaged in each category and their preferred medium, laying the groundwork for our recommendation on "medium crafting."

Effective Mediums by Category

Category	audio	GIF	photo	video
animals	521	315	616	286
science	338	338	472	498
food	367	406	335	448
healthy eating	461	356	305	450
technology	506	461	383	207

Steps:
- Count the number of users who are in the top 5 categories and use audio, GIF, photo, video as medium

Insight and recommendation:
This heat map provides a comprehensive view of user engagement across various content categories, stratified by different mediums. It reveals intriguing insights, such as the animal category drawing the highest number of engagements predominantly through photos. In contrast, categories like technology draw the lowest number of engagement through videos.

Leveraging this knowledge, SocialBuzz can strategically target content categories via the most engaging mediums, thus enhancing its potential to attract and retain users. This sophisticated approach aligns user preferences with content delivery methods, ensuring a more personalized and engaging user experience, which can be pivotal in driving platform growth and user satisfaction.

6. **Implementation Plan for Social Buzz AB Testing:**

1. Development of Two Versions (A and B):

1.1 Feature/Content Element Details:
Define the feature under consideration – perhaps a new content presentation format or user interface modification. For example, in Version A, explores a redesigned content layout emphasizing images, while Version B maintains the current format. Clearly outline the changes made in each version.

1.2 Technical Implementation:

Ensure technical equivalence between Versions A and B. Guarantee that both versions offer a comparable user experience, maintain performance standards and are free of technical glitches. Technical implementation details should align with the seamless functionality of Social Buzz.

2. Random Assignment of Users:

2.1 User Segmentation:
Define criteria for user segmentation, considering factors like user activity and preferences. Utilize a random assignment algorithm to ensure fairness and comparability between the two user groups.

2.2 Testing Duration:
Determine the testing duration, considering Social Buzz's user activity patterns. Aim for a duration that captures daily and weekly variations, ensuring sufficient data collection for meaningful insights.

3. Metrics Analysis:

3.1 Key Performance Indicators (KPIs):
Identify KPIs specific to Social Buzz, such as engagement duration, content interaction rates, and user retention. These metrics will gauge the success of the tested feature modifications.

7. **Risks and Mitigation**
Risk 1: Changing User Preference
Adapting to evolving user preferences may lead to users migrating to platforms that better align with their current interests, resulting in a potential loss of user base for Social Buzz.

Mitigation:
Leverage data analytics to track user behaviour, content consumption patterns, and emerging trends. Implement machine learning algorithms to predict and adapt to shifting user preferences dynamically. Regularly solicit user feedback through surveys or engagement metrics to stay attuned to evolving interests.

Risk 2: Targeting the Wrong User Market
Different users possess distinct preferences for content types, formats, and topics. Incorrectly identifying or targeting the wrong user market may result in content that fails to resonate, leading to decreased engagement.

Mitigation:
Analyze content consumption patterns within potential user markets using data analytics. Utilize demographic information and user behaviour data to identify distinct audience segments. Customize content strategies to cater specifically to the preferences and interests of these identified audience segments, ensuring a more targeted and relevant user experience.

Risk 3: Lack of Content Diversity
Risk Description: A limited range of content on Social Buzz may lead to decreased user engagement, exclusion of potential users, and a competitive disadvantage against platforms offering a more diverse range of content.

Mitigation:
Regularly introduce new content categories aligned with user interests and emerging trends. Monitor user engagement metrics to identify underrepresented content areas and adjust the content strategy accordingly. Foster partnerships or collaborations to bring in diverse content creators and perspectives, enriching the overall content landscape on Social Buzz.

8. **Conclusion**
Our project analyzed top content categories, leading to the "Buzz Craft" strategy—integrating content and medium crafting for enhanced user experiences. The analytical process, utilizing Python, Excel, and Tableau, revealed interconnected user preferences and guided targeted promotions.

The AB testing plan outlines meticulous steps for feature development, user assignment, and results analysis. Proactive risk mitigation ensures adaptability to changing user preferences, precise targeting, and sustained content diversity.

Social Buzz is poised for continued success with data-driven strategies and user-centric experiences in the dynamic digital landscape. Our partnership aims to navigate risks and mitigations as mentioned above and set benchmarks for user engagement and content diversity excellence. Together, we look forward to elevating Social Buzz to new heights.

9. **References:**
Citations for data sources, tools, or methodologies used.