Data Dive: Unraveling Insights from Social Buzz's Content Landscape
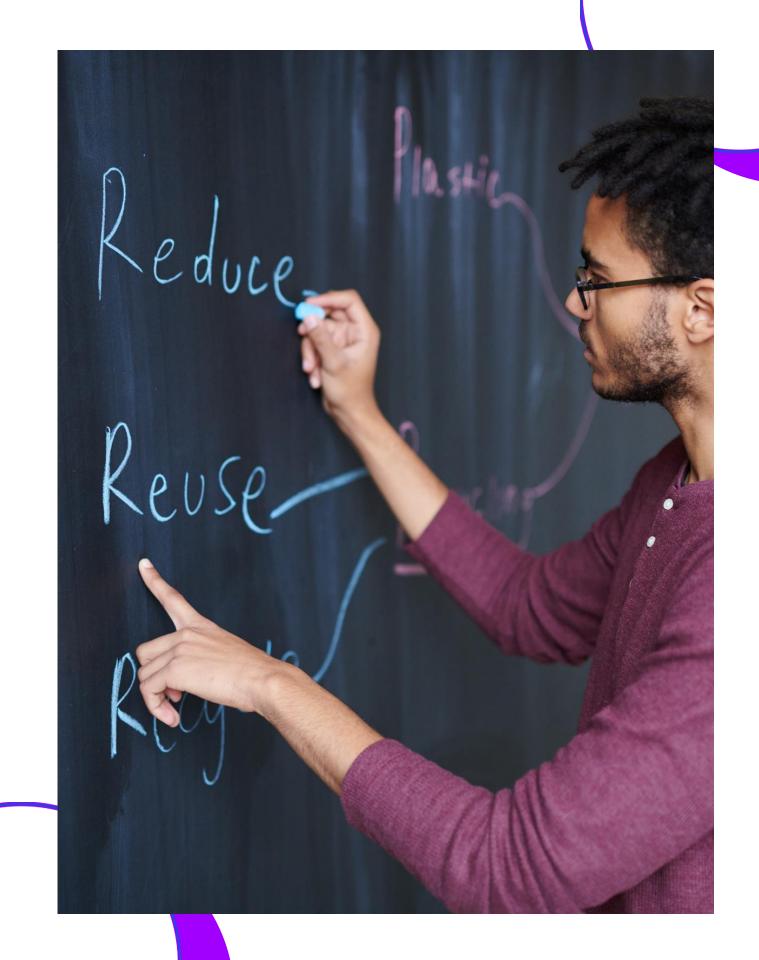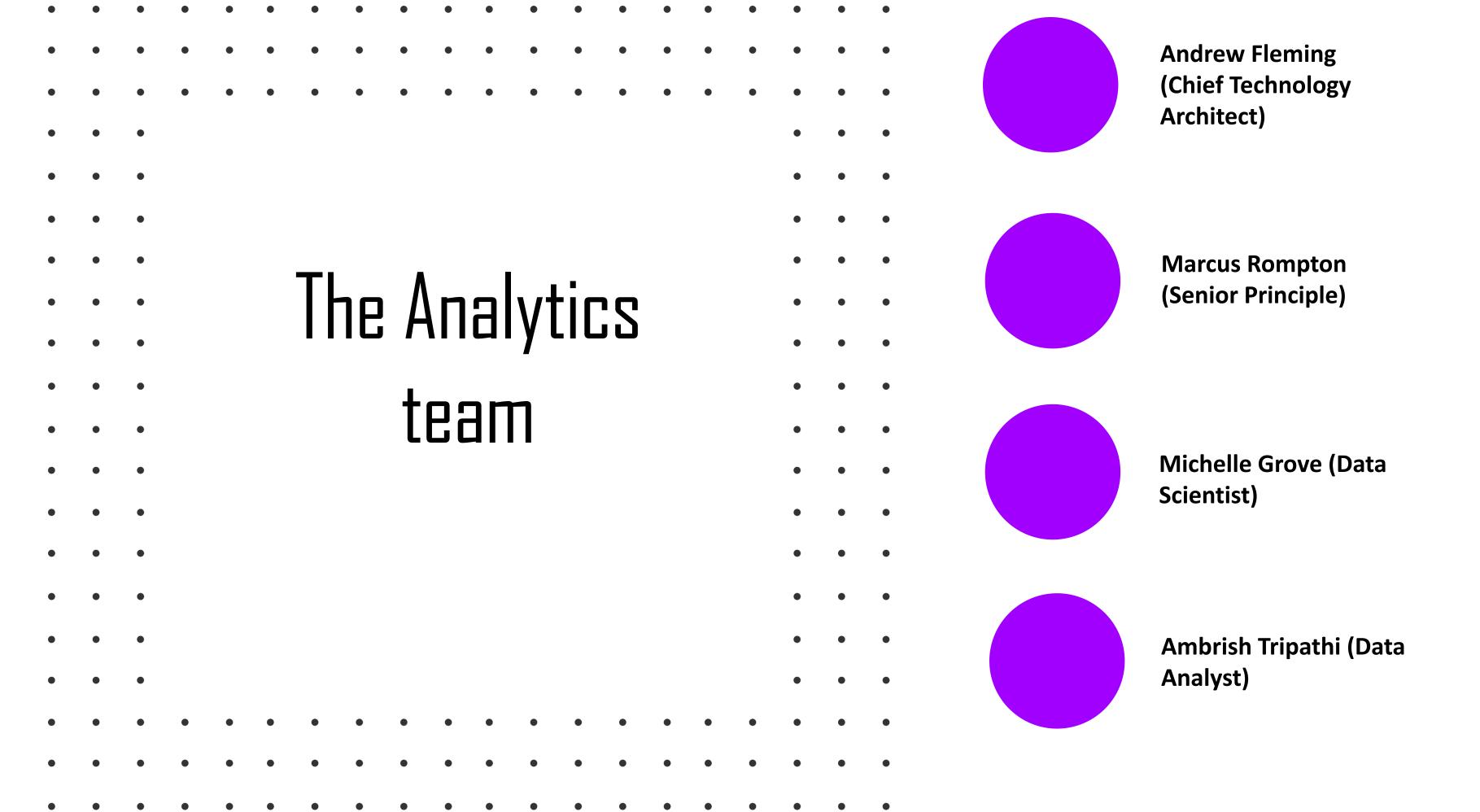
Today's agenda

Project Recap

Social Buzz, founded by ex-engineers, is a content-centric platform with over 500 million monthly users, emphasizing anonymity and diverse content reactions beyond likes. Scaling rapidly, they seek external expertise for an imminent IPO. The tsaks to be performed are:

- An Audit of Social Buzz's Big data practice.
- Recommendations for successful IPO.
- Analysis of content categories that highlights Top 5 categories with largest aggregate popularity.

Problem

- **Rapid growth strains resources.**
- **Managing massive unstructured data.**
- **Lack of experience in managing big data.**
- **Need for efficient data center and technology audit.**
- **Identifying weak spots in technology.**

- **How can we optimize resource allocation for scalability?**
- **What strategies/tools enhance unstructured data processing?**
- **What are the best practices from large corporations for big data?**
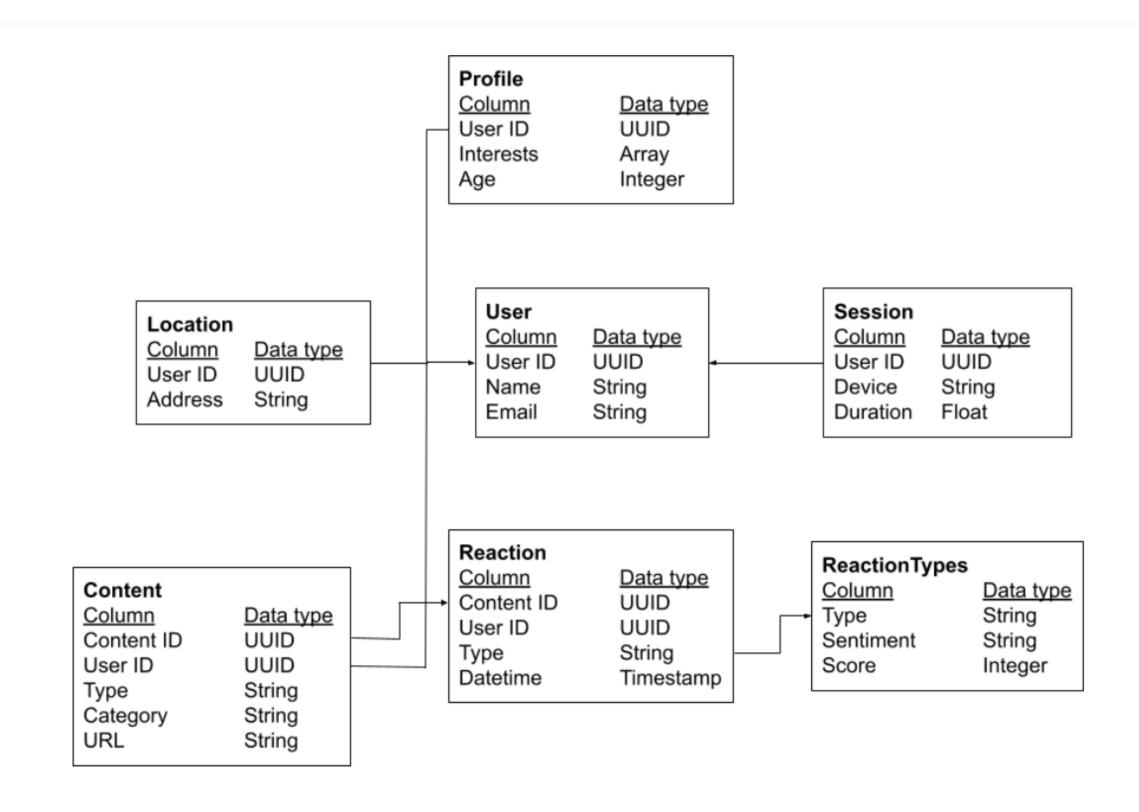- **How can we improve technology architecture for efficiency?**

The Analytics team

Andrew Fleming (Chief Technology Architect)

Marcus Rompton (Senior Principle)

Michelle Grove (Data Scientist)

Ambrish Tripathi (Data Analyst)

Data Model

According to the problems we need to solve we have used the required Dataset.

- **Reactions**
- **Content**
- **ReactionType**

Profile

Column	Data type
User ID	UUID
Interests	Array
Age	Integer

Location

Column	Data type
User ID	UUID
Address	String

User

Column	Data type
User ID	UUID
Name	String
Email	String

Session

Column	Data type
User ID	UUID
Device	String
Duration	Float

Content

Column	Data type
Content ID	UUID
User ID	UUID
Type	String
Category	String
URL	String

Reaction

Column	Data type
Content ID	UUID
User ID	UUID
Type	String
Datetime	Timestamp

ReactionTypes

Column	Data type
Type	String
Sentiment	String
Score	Integer

Process

1 Identified Business Problems and understood the requirements.

2 Collected the relevant datasets required to solve the problem. Cleaned the data using excel and removed unnecessary data.

3 Created the relationship between datasets and created model using unique and common keys.

4 Performed Analysis and Cisualized data on Power BI

5 Extracted Insights

Insights

Top 5 Categories:
- **Animals**
- **Science**
  Healthy Eating
- **Technology**
- **Food**
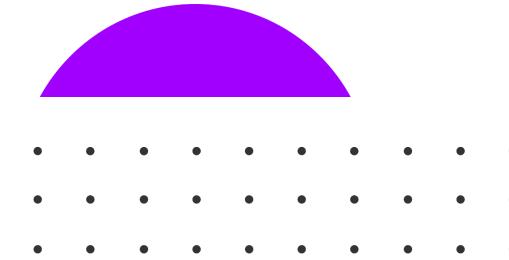
There are total **16** Unique categories

There **1895** reactions for **Animal** which is the most popular category.
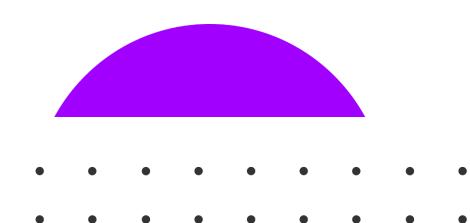
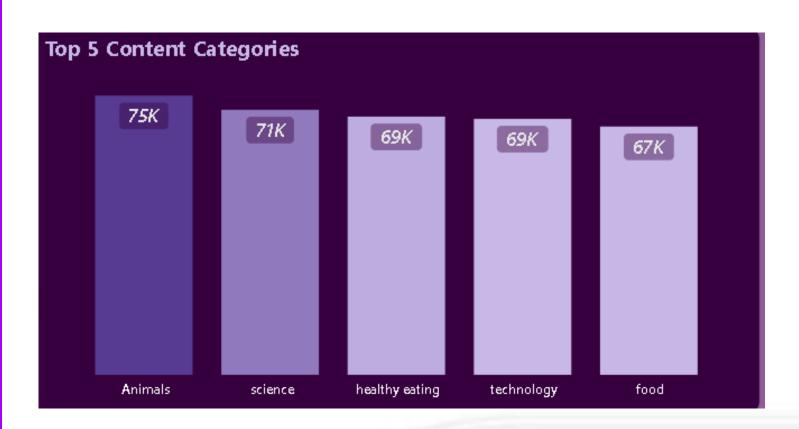**January** was the month with most posts. **781** posts were posted.

Sentiment analysis:
- **56.19% Positive** Reactions on Posts.
- **31.31%** and **12.5%** were **Negative** and **Neutral**.

January, August and **December** were the months with highest engament.

Bar Chart:
Top 5 Categories

Recommendation: Allocate resources for creating more content in the top-performing categories.
Suggestions: Conduct audience surveys to identify specific subtopics or trends within these categories for more targeted content.
Based on the preferences we can push such subplots to increase more engagement.

25K
Total Reactions

16
No. of Unique Catego...

KPIs:

- **Number of Reactions**
- **Number of Unique Categories**

Suggestions:
- Develop guides or tutorials to educate users on the available content categories.
- Establish community feedback platforms for users to suggest and vote on new content categories.

Solutions:
- Create in-app notifications or dedicated sections to showcase lesser-known categories.
- Implement user-generated category suggestions with a voting mechanism.
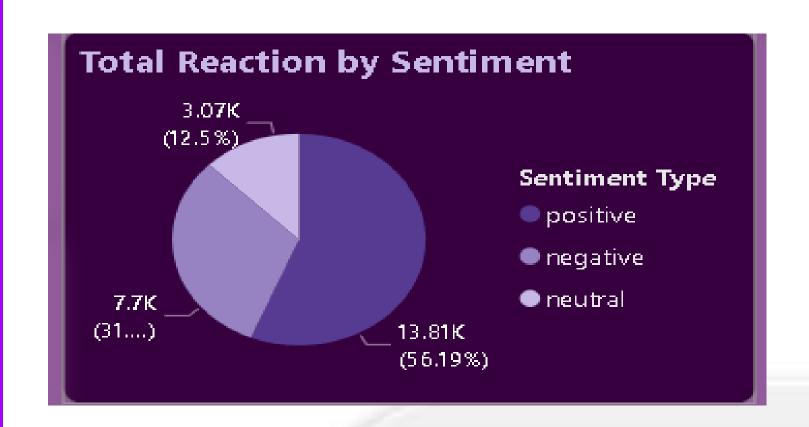
Total Posts by Month

Line Chart:
Total Posts By Month

Suggestions:
- Collaborate with influencers or run targeted marketing campaigns to maximize platform visibility
- Introduce content creation challenges to stimulate posting during slower months.

Solution:
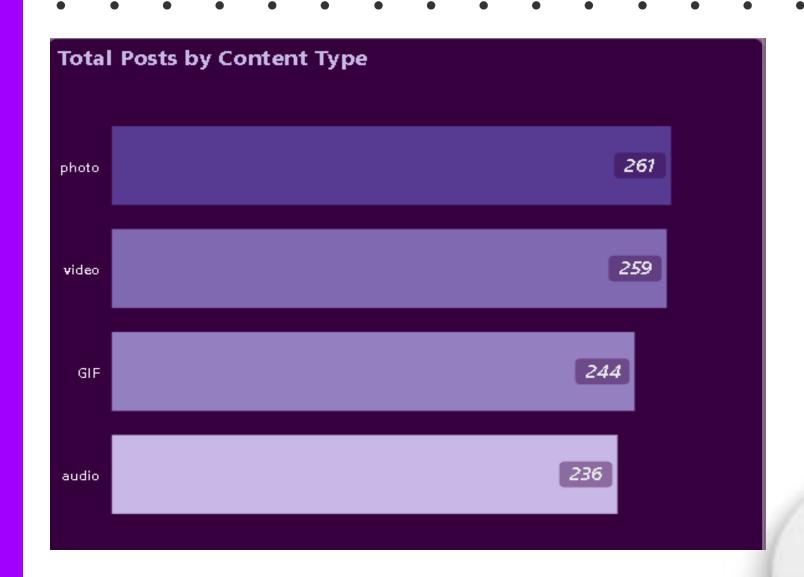Design campaigns that inspire users to create and share content with specific themes.

Total Reaction by Sentiment

3.07K
(12.5%)

Sentiment Type
- positive
- negative
- neutral

7.7K
(31....)

13.81K
(56.19%)

Pie Chart:
Sentiment Analysis – Reactions By Sentiment type

Suggestions:
- Invest in advanced sentiment analysis algorithms.
- Implement a content moderation system based on sentiment analysis.

Solution and Action:
- Collaborate with data scientists to improve the platform's ability to accurately analyze and categorize sentiments.
- Flag or review content with consistently negative sentiments for additional scrutiny.

Total Posts by Content Type

Content Type	Posts
photo	261
video	259
GIF	244
audio	236

Horizontal Bar Chart:
Content type Distribution – posts By
Content Type

Suggestion:
Develop an analytics dashboard for content creators to track the performance of their content types.

Action:
Provide insights into the reach and engagement of different content types to guide creators.

Total Reactions by Month

2126 · 1914 · 2012 · 1974 · 2138 · 2021 · 2070 · 2114 · 2022 · 2056 · 2034 · 2092

JAN · FEB · MAR · APR · MAY · JUN · JUL · AUG · SEP · OCT · NOV · DEC
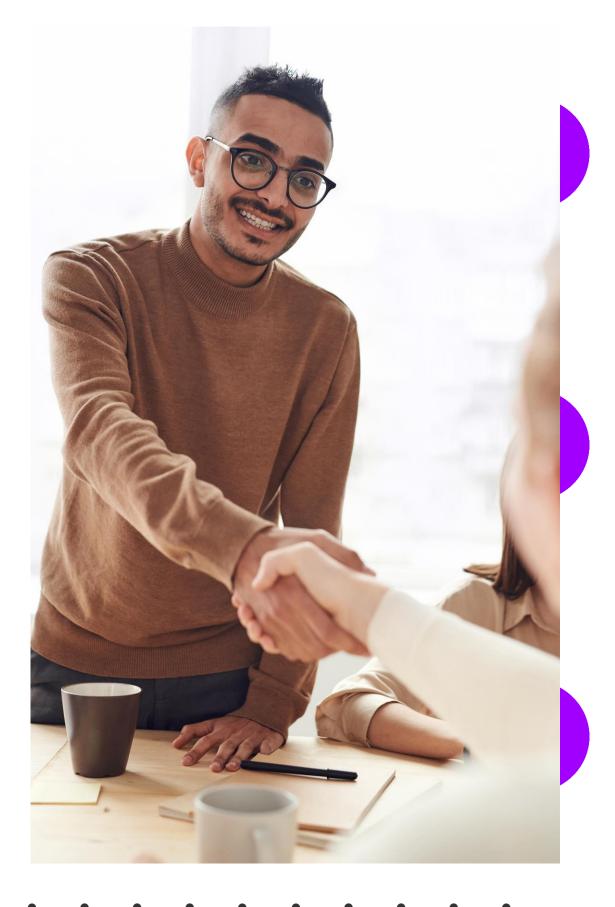
Line Chart:
Total Reactions By Month

Suggestion:
Utilize push notifications strategically.

Action:
Send notifications during specific months to encourage users to share content, react, or participate in community activities.

Summary

Social Buzz's audit revealed challenges in resource allocation, unstructured data management, and technology efficiency. Insights include top 5 categories (Animals, Science, Healthy Eating, Technology, Food), 16 unique categories, and peak engagement in January. Sentiment analysis showed 56.19% positive reactions. Next steps involve refining technology architecture and implementing best practices. The IPO preparation is on track with identified strengths and areas for improvement. The Analytics team successfully navigated the complexities, extracting actionable insights for enhanced scalability.

Analysis:
Identified business problems.
Conducted data cleaning and modeling.
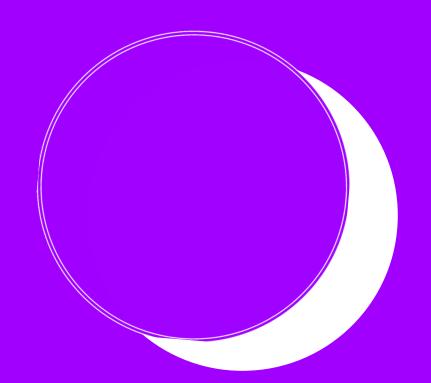Utilized Power BI for data visualization.

Insights:
Top 5 categories: Animals, Science, Healthy Eating, Technology, Food.
16 unique categories.
Highest engagement in January.

Next Steps:
Refine technology architecture.
Implement big data best practices.
Prepare for IPO based on identified strengths and improvements.

Thank you!

ANY QUESTIONS?