Data Analytics

Today's agenda

- **Project recap**
- **Problem**
- **The Analytics team**
- **Process**
- **Insights & Summary**

Project Recap

“Social Buzz” is a fast growing technology unicorn that need to adapt quickly to it's global scale. Accenture has begun a 3 month POC focusing on these tasks:

- An audit of Social Buzz's big data practice
- Recommendations for a successful IPO
- Analysis to find Social Buzz's top 5 most popular categories of content

Problem

- The client has reached a massive scale within recent years and does not have the resources internally to handle it.
- Social Buzz receives over 100000 posts per day which amounts 36,500,000 posts every year, of which this all unstructured data making it very hard to make a sense.
- Identify the requirements that need to be delivered for this project.
- Merging of sample data set tables
- An analysis of their content categories that highlights the top 5 categories with the largest aggregate popularity

The Analytics team

Andrew Fleming

Chief Technology Architect

Marcus Rompton

Senior Principal

Swamesh Lotlikar

Data Analyst

Process

1. Data Understanding
2. Data Cleaning
3. Data Modelling
4. Data Analysis
5. Uncover Insights

Data Understanding

Content Table:
ID: Unique ID of the content that was uploaded (automatically generated).
User ID: Unique ID of a user that exists in the User table.
Type: A string detailing the type of content that was uploaded.
Category: A string detailing the category that this content is relevant to.
URL: Link to the location where this content is stored.

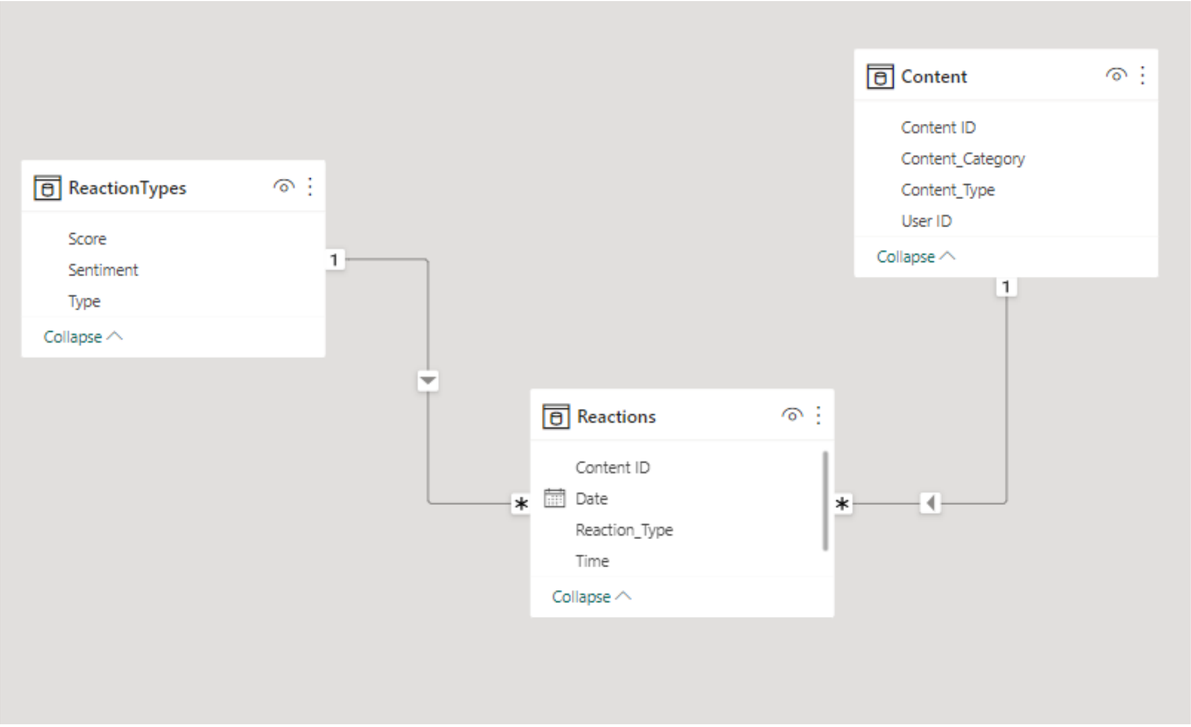
Reaction Table :
Content ID: Unique ID of a piece of content that was uploaded.
User ID: Unique ID of a user that exists in the User table who reacted to this piece of content.
Type: A string detailing the type of reaction this user gave.
Datetime: The date and time of this reaction.

ReactionTypes Table :
Type: A string detailing the type of reaction this user gave.
Sentiment: A string detailing whether this type of reaction is considered as positive, negative or neutral.
Score: This is a number calculated by Social Buzz that quantifies how “popular” each reaction is. A reaction type with a higher score should be considered as a more popular reaction.

Data Cleaning and Modelling

- Corrected irregular datatypes and converted date column to date format.
- Removed rows with NULL Reaction Types and unnecessary columns.
- Added Time of the Day column to analyze most active time.
- Corrected Content categories that had duplicate values like “animals” and Animals to a standard format ANIMALS.

- Used DAX in Power BI to form measures where AVERAGEX(),RELATED() and SUMX() function were used.



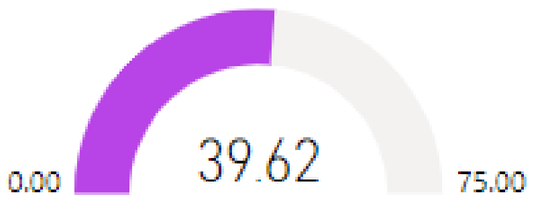
Initially used VLOOKUP() in Excel and Power Query Merge to combine all datasets in to the base (reactions). However, data modelling was preferred.

Insights

Unique Categories

16

Average Score



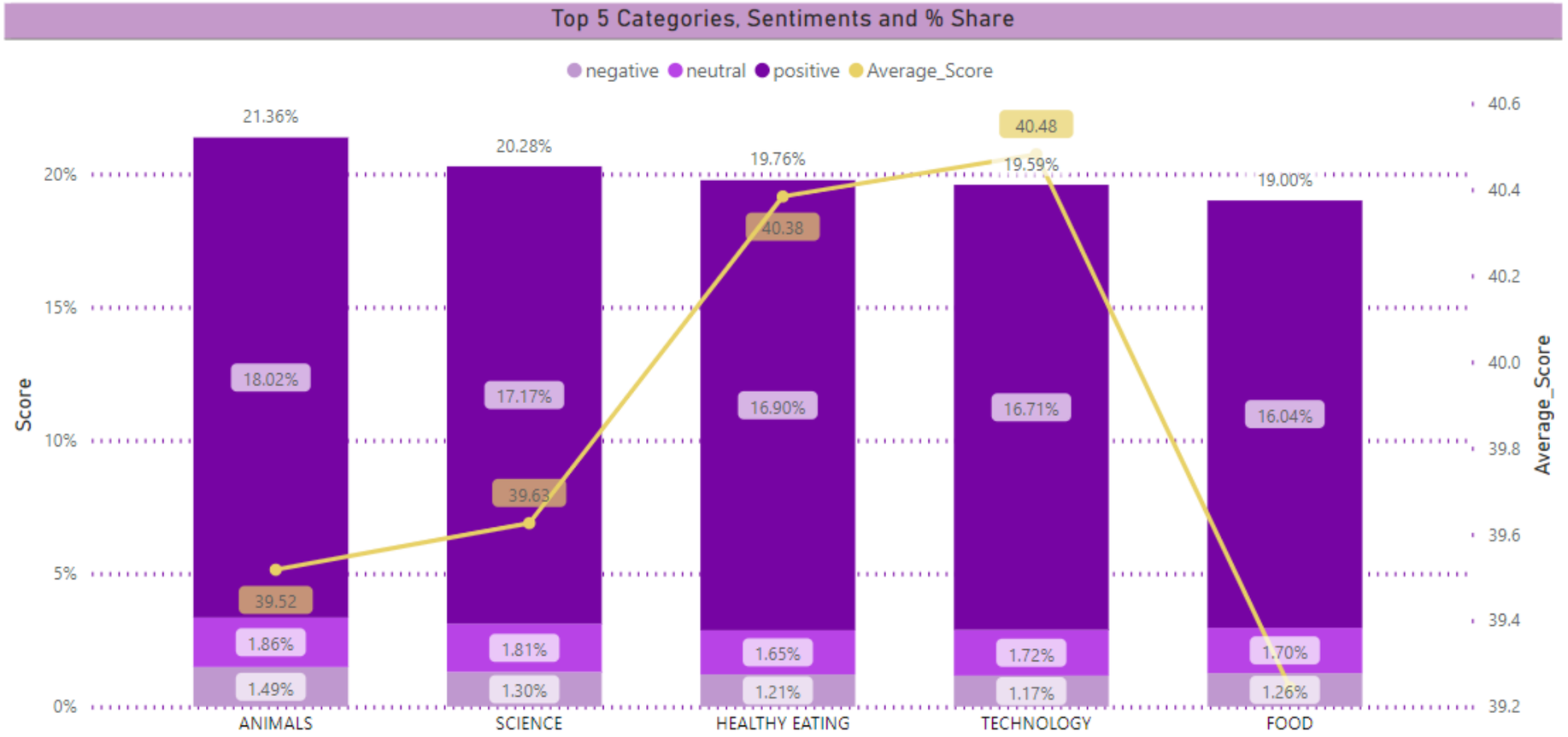
Total Reactions

24.57K

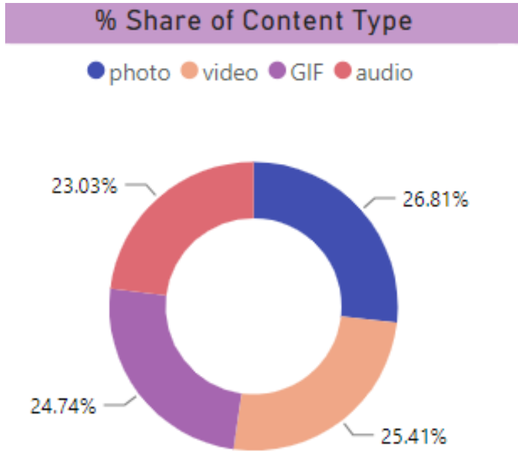
ANIMALS

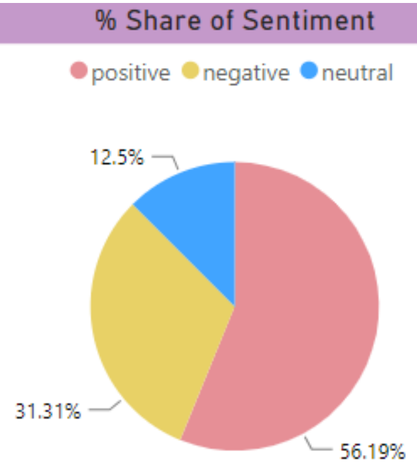
1897

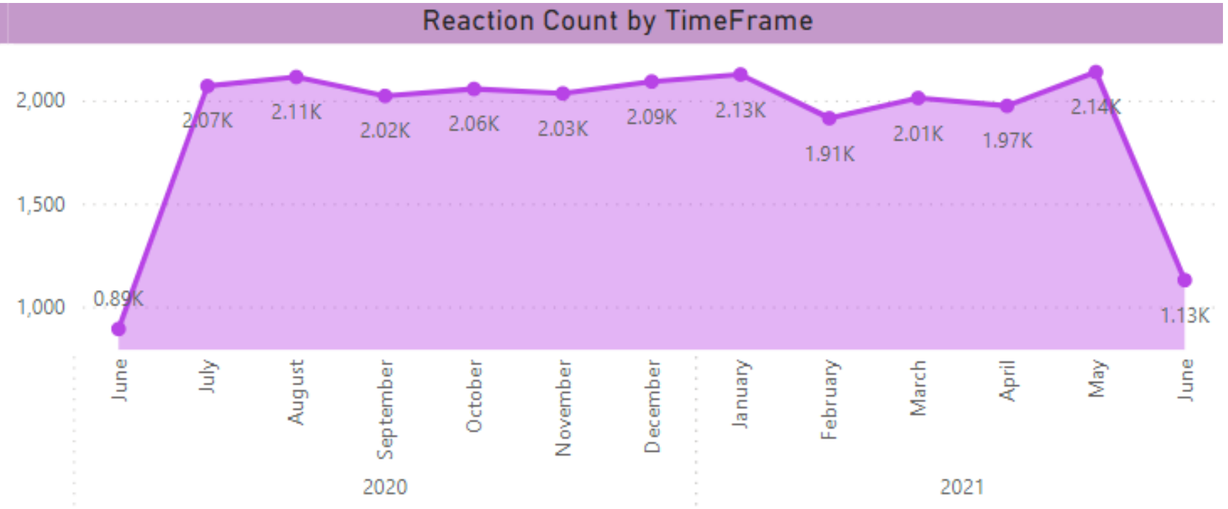
Highest Reaction Count



- The top 5 performing categories are Animals, Science, Healthy Eating, Technology and Food.
- Animals account for 21.36% share amongst all the categories.
- Amongst the top 5. the average score for Technology is highest at 40.48
- The proportion of positive:neutal:negative sentiments is consistent for the 5 categories where Positive is the major sentiment.







TimeOfDay	Reactions
Afternoon	3995
Evening	5079
Morning	12387
Night	3112

- The 4 content types are consistently divided between 23 % to 27%.
- All reactions have major positive sentiment at 56.19%
- People tend to be the most active during Morning and May 2021 saw the highest reactions received. However, on a larger scale, we can see a similar reaction counts from July 2020 till May 2021.

Summary

Animals, Science & Health Rule: Focus on real-world content like animals & science, and create campaigns with health brands to match user interest.
.

Tech is Booming : With rise in technology it is no surprise to see contents related to technology right up with the top categories. It show users enjoy your technology contents. I would recommend collaborating with some of the world's tech giants as this would definitely make user engagement skyrocket.

Morning Star : People seem to be the most active during Morning, strategize our campaigns and marketing ideas accordingly to gain the most user engagement,

Thank you!

ANY QUESTIONS?

Contact

Email : lotlikar.swamesh@gmail.com

Linkedin : [Linkedin]