

# Data Analysis

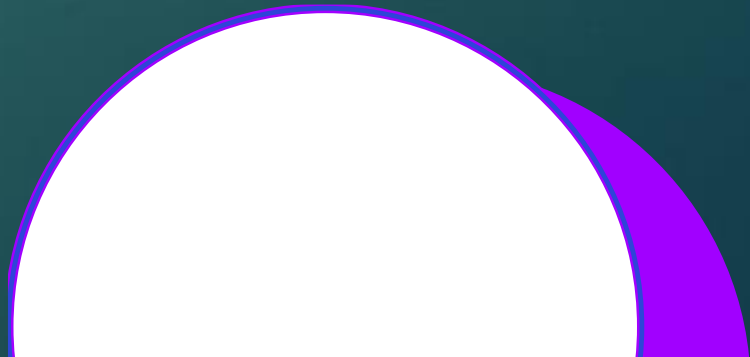
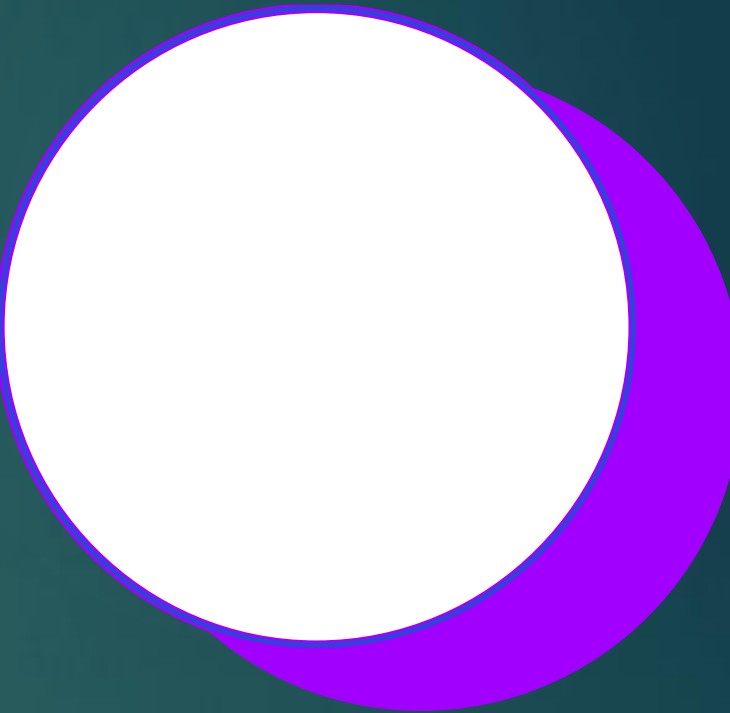


Hello and welcome, my name is 'Vishal Kumar' and today I will be presenting to you the results of the Data Analytics task.



# Today's agenda

Project recap  
Problem  
The Analytics team  
Process  
Insights  
Summary



## Today's agenda will be as follows:-

1. We will recap the overall project to give a high level understanding of the business problem we're tackling and the specific requirements.
2. We will dive into the specific problem that we, the **Data Analytics team**, have been focusing on and will give some background as to why this is such a big problem.
3. After introducing the problem, I will go over the team responsible from our side in tackling this task.
4. I will then go over the high-level process that we followed to complete this task, so that you have complete clarity in how we tackle these kinds of tasks.
5. Finally, I will go over the all important results and I will present them as a series of insights and visualizatio's from our analysis.

To wrap up, I will summarize and open for any questions.

# 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



To kick things off let me recap this engagement.

- We, **Accenture** have embarked on a **3 month pilot with Social Buzz to focus on 3 main tasks**, aligned with some of the biggest challenges that you're currently facing.
- Social Buzz has reached huge scale in recent years to become recognized as a global unicorn company. We are here to help you manage this scale and to guide you in the right direction.
- Firstly, we will be doing an audit of your big data practice and sharing best practices and industry expertise. Secondly we will be guiding you through a successful IPO, of which we have deep expertise and knowledge of within our team. And finally, we have conducted an analysis of your data to find insights regarding your top 5 most popular categories of content.

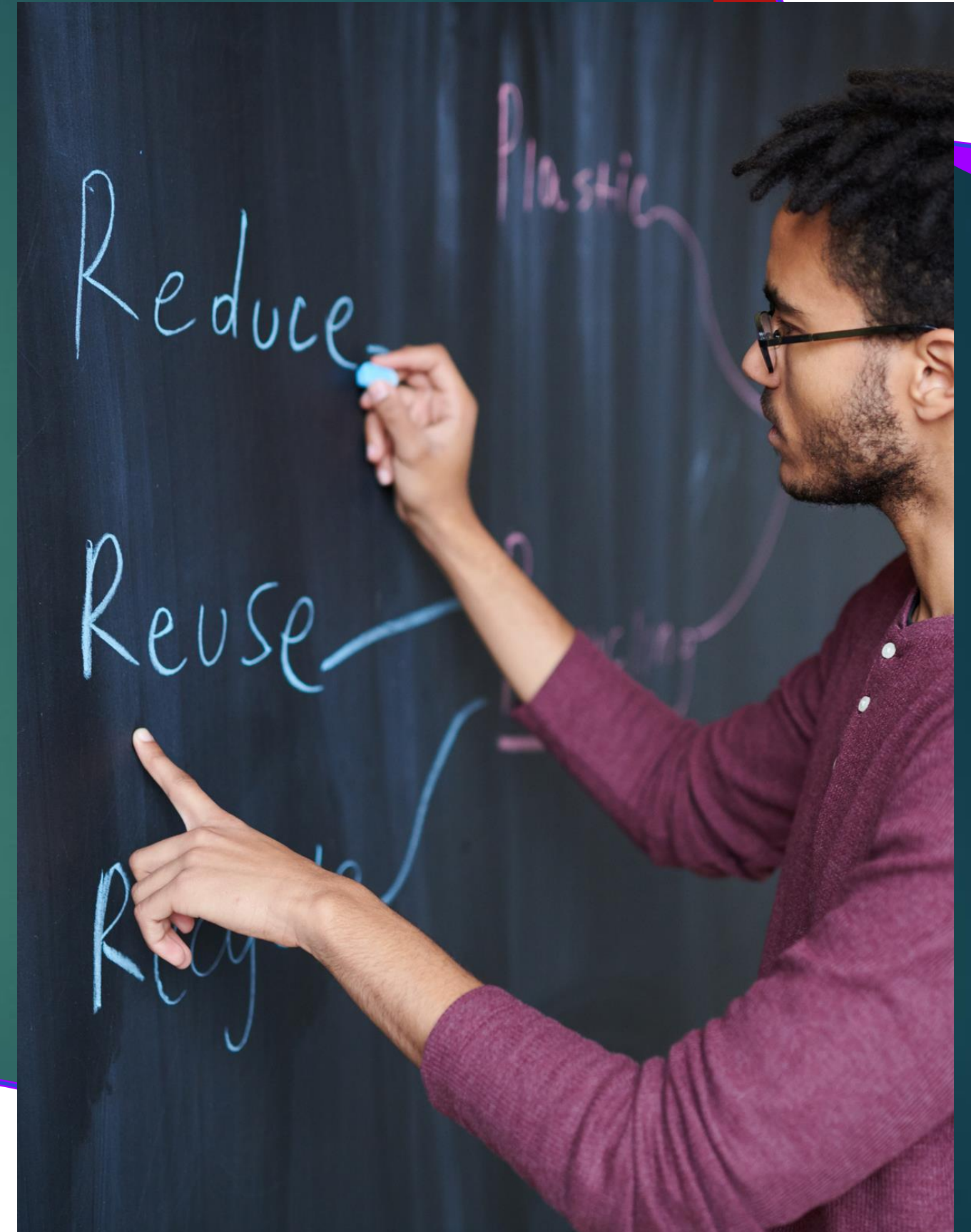
# Problem

Over 100000 posts per day

36,500,000 pieces of content  
per year!

But how to capitalize on it when there is so much?

Analysis to find Social Buzz's top 5 most popular  
categories of content

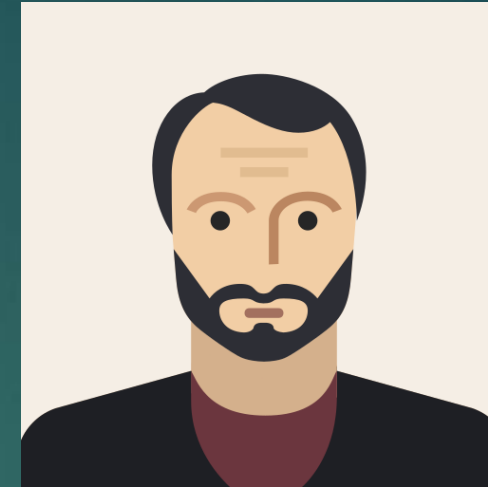




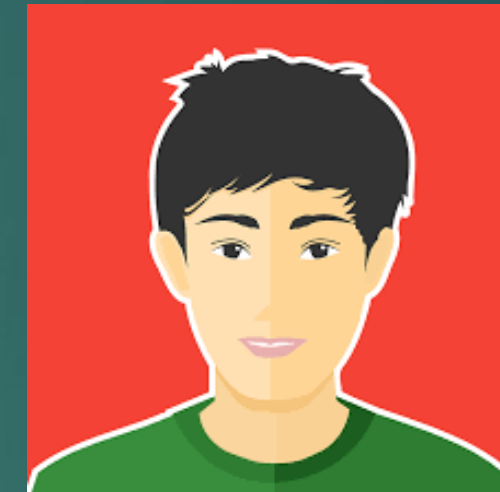
- Focusing on the last point that I mentioned there, this is what the Data Analytics team has been specifically focused on.
- Clearly with such grand scale, this comes with a lot of data and with such vast amounts of data comes challenges.
- To give a background on how much data you've been creating: - You told us that your platform receives over 100000 posts per day which amounts to 36 500 000 posts every year, of which, this is all unstructured data making it very hard to make sense of.
- In this day and age, content is king. Just look at some of the biggest platforms in the world, for example YouTube, Facebook and Netflix... they are all content businesses...
- But how to capitalize on it when there is so much?
- It's not just all about harvesting as much content as possible... The real value is in understanding and crunching this content to gain a deeper understanding of your audience and to therefore provide a more personalized and enjoyable experience.
- And this is where our data analytics expertise comes in, with the insights that we've uncovered from this task, we can show you exactly how to take analytics to production at scale.



# The Analytics team



ANDREW FLEMING  
Chief Technology  
Architect



MARCUS ROMPTON  
Senior Principal



MY SELF  
Data Analyst

- 
- Talking about experience, we have a large data analytics practice at Accenture but we had a team of 3 people primarily focusing on this task. Andrew Fleming is our Chief Technical Architect and his expertise really helped to guide the team to produce high quality analysis.
  - Marcus Rompton, a senior principal has worked with the worlds biggest clients on solving their data problems and was heavily involved in the data engineering side of this project.
  - And finally myself, **Vishal Kumar** , who was solely responsible for taking leadership guidance and delivering high quality insights from the raw datasets and turning these into business decisions.

# Process

1

Data Understanding

2

Data Cleaning

3

Data Modelling

4

Data Analysis

5

Uncover Insights

So, how did we tackle this problem?

Well we approached it in 5 steps:

1. Data understanding - the key to success on any data project is to understand the data in detail. So we took the time to understand the data model and domain of your business.
2. Data cleaning - after understanding your business, we then cleaned the available datasets and thought about what an ideal dataset should look like for this problem.
3. Data modelling - After ensuring the data was clean for analysis, we needed to process and model this data into a dataset that can precisely answer the business questions and produce the results needed.
4. Data analysis - With our new dataset, we used our analytical expertise to uncover insights from this dataset and to produce visualizations to describe the insights.
5. And finally we used these insights to unlock business decisions and to make recommendations on next steps.



# Insights



16

UNIQUE  
CATEGORIES



1897


REACTIONS TO "ANIMAL"  
POSTS

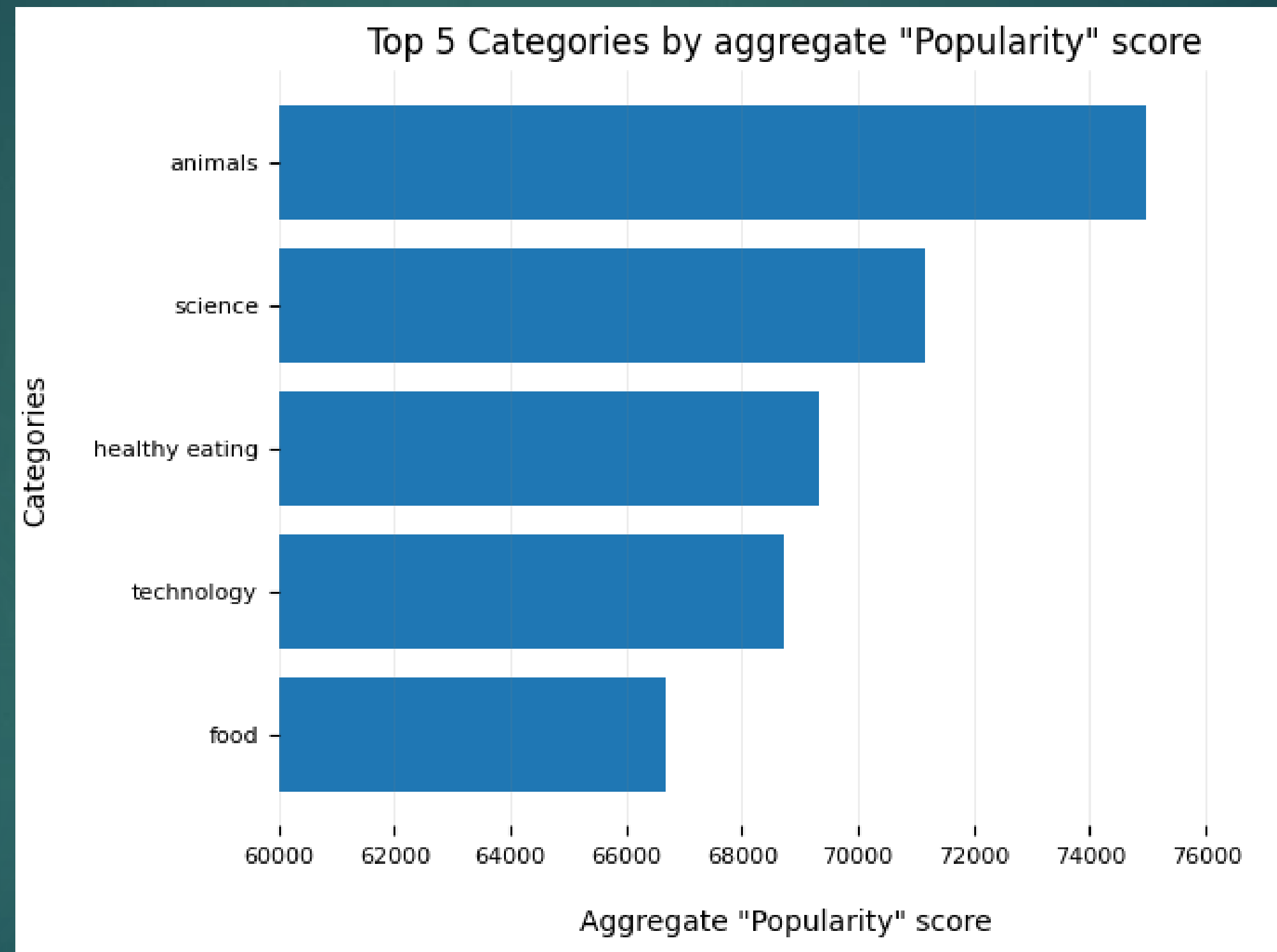



JANUARY

MONTH WITH  
MOST POSTS



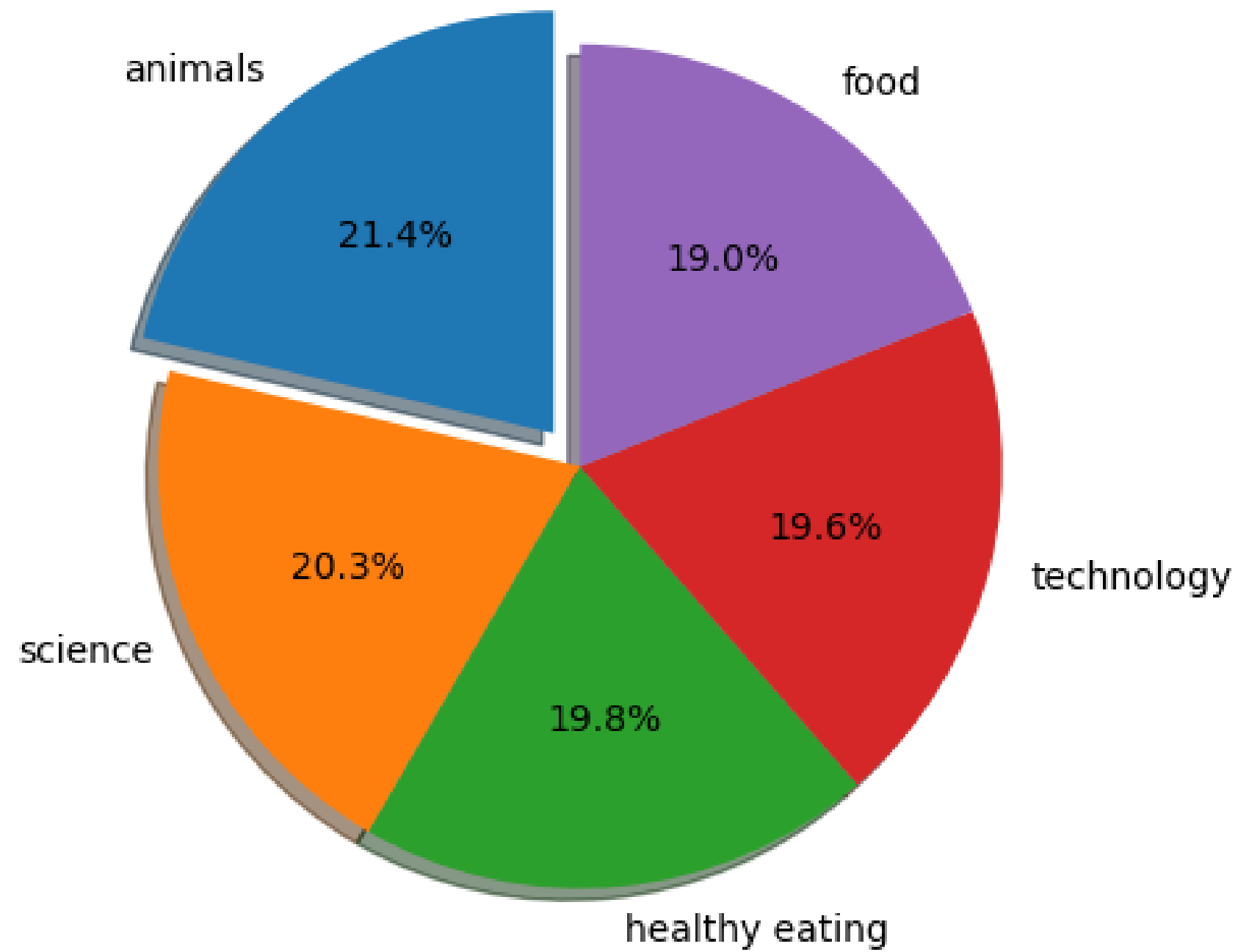
- 
- From your data we found that you had a total of 16 unique categories of posts across your sample dataset. This includes things such as Food, Science and Animals.
  - As well as this, there were 1897 reactions from just the animal category alone! People obviously really like animals!
  - And also the most common month for users to post within was January. This aligns with seasonal trends of social media users that feel the need to reconnect with people after calendar events such as Christmas.
  - But now, onto the main question... which is... what were the top 5 most popular categories of posts?




- 
- From our analysis, you can see that the top 5 most popular categories of posts were animals, science, healthy eating, technology and food in descending order.
  - Animals had an aggregate popularity score of around 74965. It is very interesting to see both food and healthy eating within the top 5, it really shows that food is a highly engaging content category. Healthy eating ranks slightly higher than food, so perhaps your user base may be skewed towards healthy eaters and health-conscious people.
  - Finally, its also interesting to see science and technology too. This may suggest that people enjoy consuming factual content and snippets of content that they can learn something from.



Popularity percentage share from top 5 categories



- 
- Additionally, you can see from this chart the % split of popularity between the top 5 categories. There is not much difference between the share of each category, however, the difference between the 1st most popular, animals and the 2nd most popular, science, is the largest gap equal to 1.1%.
  - In business terms, this could suggest that the most popular category, animals, is tailing away from the rest of the categories and may continue to get more and more popular. To avoid an issue where 1 content category consumes the entire platform, it will be important for you to ensure that any algorithms used to govern the content on the platform gives a fair balance to the content categories.

# Summary



## ANALYSIS

Animals and science are the two most popular categories of content, showing that people enjoy "real-life" and "factual" content the most.

## INSIGHT

Food is a common theme with the top 5 categories with "Healthy Eating" ranking the highest. This may give an indication to the audience within your user base. You could use this insight to create a campaign and work with healthy eating brands to boost user engagement.

## NEXT STEPS

This ad-hoc analysis is insightful, but it's time to take this analysis into large scale production for real-time understanding of your business. We can show you how to do this.



So to summarize:

- We tackled this task and found the top 5 most popular categories as asked, but we also went one step further.
- We found that animals and science are the two most popular categories, suggesting that users like "real-life" and "factual" content
- We also found that food was a common theme amongst popular content and the most popular food category was healthy eating. This could be a signal to show the types of people that are using your platform, and you could use this insight to boost engagement even further. For example, you could run a campaign with content focused on this category or work with healthy eating brands to promote content.
- As much as this analysis was insightful, we are ready to take it to the next stage and we have the expertise within Accenture to help you realize these kinds of insights in production across your organization and in real-time. We would love to help you with this.





Thank you!

ANY QUESTIONS?