Analysis of Social buzz content

Today's agenda

Project recap
Problem
The Analytics team
Process
Insights
Summary

Project Recap

- **An audit of their big data practice**
- **Recommendations for a successful IPO**
- **An analysis of their content categories that highlights the top 5 categories with the largest aggregate popularity**

Problem

- The Project involved taking the Social buzz dataset, consisting of the contents posted by the users, the different reactions made by other users on the posts, and a score associated with each category.
- The goal of this dataset was to analyze the dataset, convert it into a clean dataset, and then return the best 5 most popular categories on the social buzz platform.

The Analytics team

Andrew Fleming (Chief Technical Architect)

Marcus Rompton (Senior Principle)

Muneeb Hassan (Data Analyst)

Process

1. Download and analyse the datasets
2. Clean the dataset by removing Unnecessary columns and empty rows.
3. Used Python to make sure the values of a column follow the same data format.
4. Used SQL queries to group the contents into categories and summing the total score of each category.
5. Ordered the categories into descending order by their score and returned the top 5 categories into a csv file.

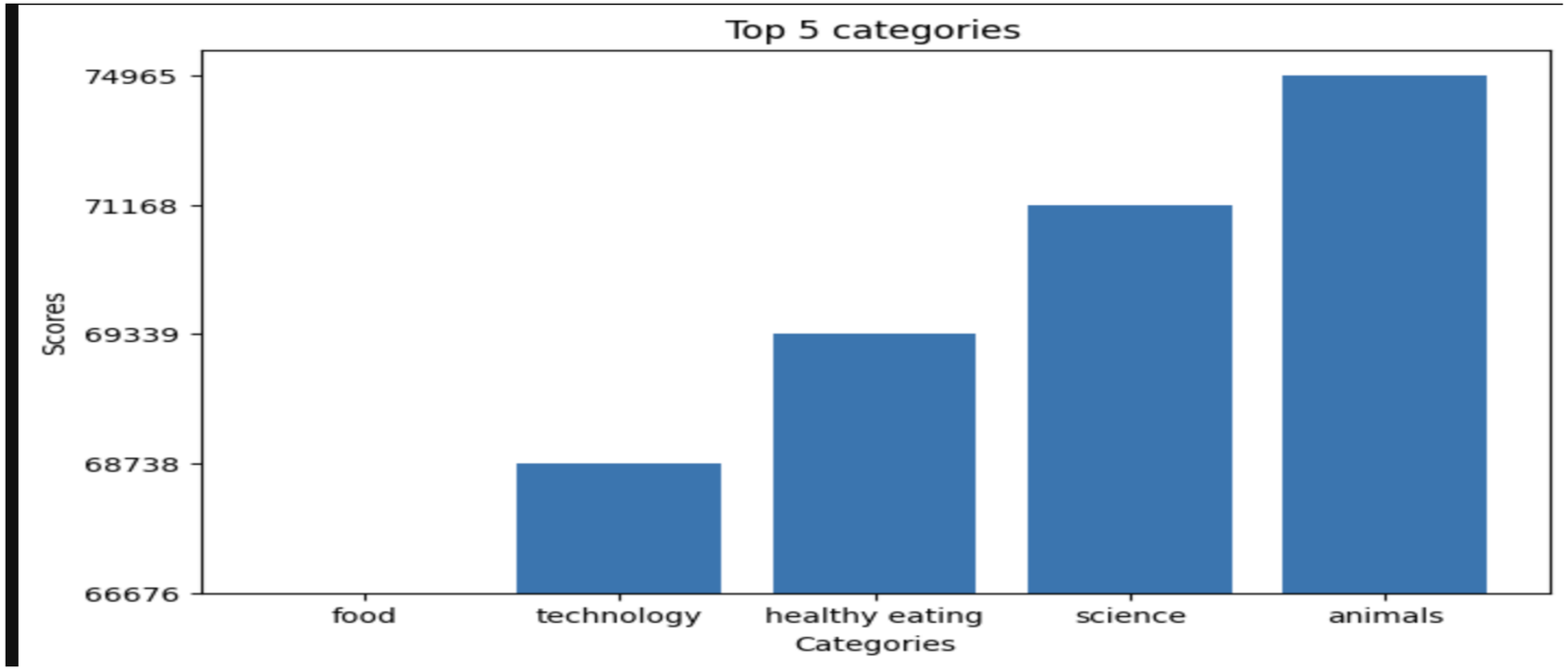
Insights

The most category was the animal category with a total score of 74965

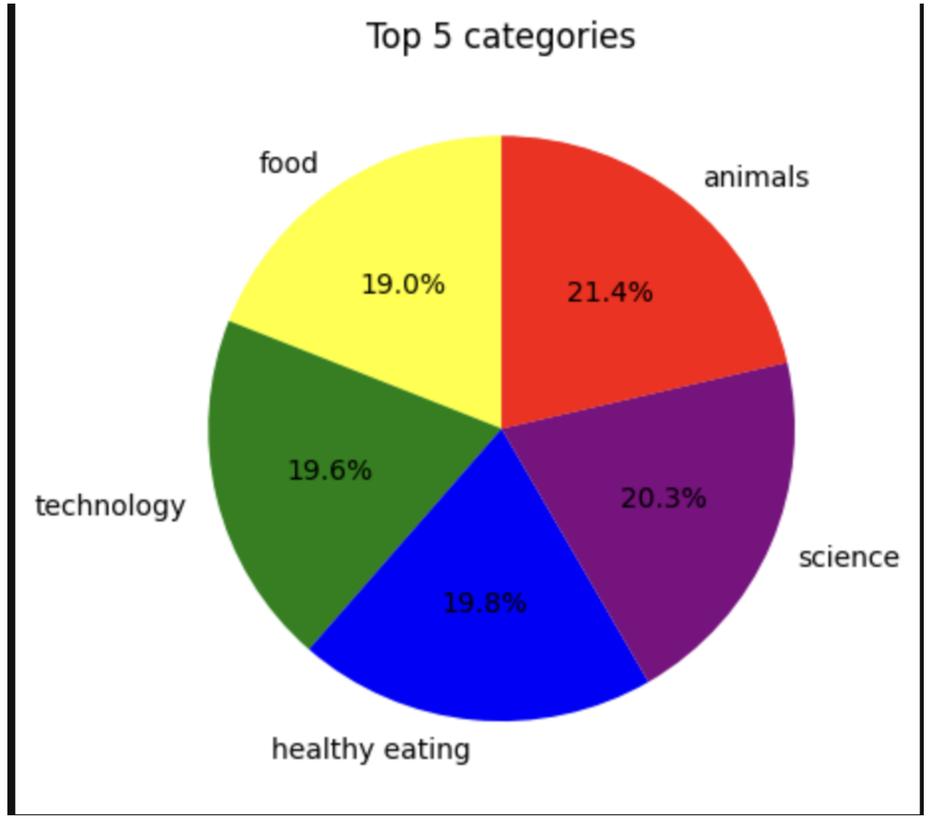
A total of 72 posts were made about the animals category.

Total of 1897 reactions were made on the animal category post, average almost 27 reactions per post.

BAR GRAPH

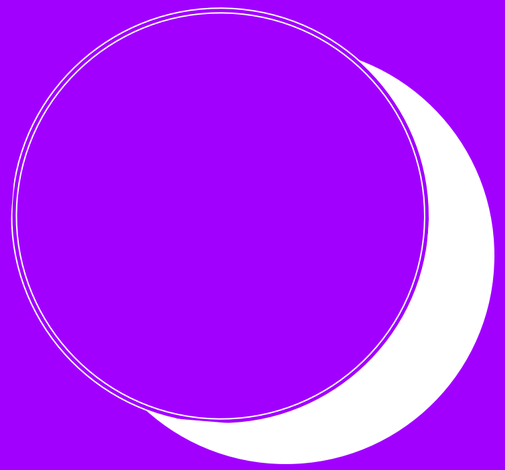

Top 5 categories
74965
71168
69339
68738
66676
Scores
food
technology
healthy eating
science
animals
Categories

PIE CHART


Top 5 categories
food
animals
19.0%
21.4%
technology
19.6%
20.3%
science
19.8%
healthy eating

Summary

- **Analysed and cleaned the dataset**
- **Performed queries to generate the top 5 categories**
- **Provided insights and graphs regarding the dataset**



Thank you!

ANY QUESTIONS?