Objective: analysis of their content categories that highlights the top 5 categories with the

largest aggregate popularity

Tasks:

- Extraction of sample data sets using SQL

- Loading of sample data sets into Accenture sandbox database

- Merging of sample dataset tables

- focus on the analysis of sample datasets with visualizations to understand popularity of different content categories.

- Creation of an up-to-date big data best practices presentation

- Virtual session with Social Buzz team to present previous client success stories relevant to them

**NB:** popularity is quantified by the “Score” given to each reaction type.

- how many unique categories are there?

- how many reactions are there for most popular categories?

- what was the month with the most posts made?

Presentation:

- Summarise the 1st two quarters of 2021, and (3rd + 4th) quarters of 2020.

- from 1000 contents with 16 categories, 4 distinct content types, 25000 reactions, 15 distinct reaction types, and 3 types of sentiment.

- top 5 categories using reaction score ( 2020, 2021 ).

- pie chart by content ( categories ): for 2020, 2021. x

- pie chart by reaction ( top 5 categories and 11 others ): for 2020, 2021.

- bar graph(months with the most posts/ reactions)

- pie chart by reactions( sentiment ): 2020, 2021

Data analyst: responsible for analysis and translation of project requirements into insights. Data analysts sift through data and provide reports and visualizations to explain what insights the data is hiding.

Data scientist: produce predictions of future active user forecasts and developing algorithms for client scale. data scientist is someone who knows more programming than a statistician, and more statistics than a software engineer.