ACCENTURE PROJECT

Analysis of sample data sets with visualizations to understand the popularity of different content categories

You’ve identified *Reaction*, *Content*, and *Reaction Types* as our relevant data sets.

To clarify why you made this selection:

* The brief carefully it states that the client wanted to see “An analysis of their content categories showing the top 5 categories with the largest popularity”.
* As explained in the data model, popularity is quantified by the “Score” given to each reaction type.
* We therefore need data showing the content ID, category, content type, reaction type, and reaction score.
* So, to figure out popularity, we’ll have to add up which content categories have the largest score.

Importantly, before you start creating your data charts, you should plan your presentation structure. This will ensure your presentation answers the right questions.

* Here is a template that we would use at Accenture to create a presentation. You need to **download this template and populate slides 2-6.**

For each slide, think about:

* Agenda - What will your presentation cover?
* Project Recap - What are the key points from the brief?
* Problem - What is the problem that you answer in this presentation?
* The Analytics team - Who is on your team?
  + As a reminder from the earlier task - this includes: Andrew Fleming (Chief Technical Architect), Marcus Rompton (Senior Principle), and yourself!
* Process - How did you complete your analysis?

Analyze the background information provided to understand the requirements

Identify the right datasets required to answer the client’s business question

Clean the datasets and merge them to prepare the data for analysis

model and analyze data to create valuable insights for the client

Determine the answer to the client’s business question and present the findings to the client

There is a lot of information here and it’s easy to get lost in the data. So, to make sure you are using the right data to answer the business questions you’ll follow these steps:

1. Requirements gathering
2. Data cleaning
3. Data modelling

* Understand the client and business problem at hand.
* Identify the requirements that need to be delivered for this project.
* Identify which tasks you should focus on as a Data Analyst.
* **Big Data** refers to large, diverse sets of information that grow at ever-increasing rates. It includes the volume of information, the velocity or speed at which it is created and collected, and the variety or scope of the data points being covered (known as the "three v's" of big data). Source: [Investopedia](https://www.investopedia.com/financial-edge/0312/ipos-for-beginners.aspx).
* An **IPO** stands for Initial Public Offering - this happens when a private company becomes public by selling its shares on a stock exchange. Source: [Investopedia](https://www.investopedia.com/terms/b/big-data.asp).

Audit of big data practice, recommendations for IPO, analysis of popular content

**My task as data analysts**

Analysis of sample data sets with visualizations to understand the popularity of different content categories