

User Requirements Doc: Top UK Youtubers 2024 Dashboard

Objective

To discover the top performing UK Youtubers to form marketing collaborations with throughout the year 2024.

Problems identified

- Sharon (the Head of Marketing) has found finding the top YouTube channels in the UK to run marketing campaigns with difficult.
- She has performed online research but constantly bumps into overly complicated and conflicting insights
- She has also held calls with different third-party providers, but they are all expensive options for underwhelming results
- The BI reporting team lack the bandwidth to assist her with this assignment

Target audience

- Primary – Sharon (Head of Marketing)
- Secondary - Marketing team members (who will be involved in running campaigns with the YouTubers)

Use cases

1. Identify the top Youtubers to run campaigns with

User story

As the Head of Marketing, I want to identify the top YouTubers in the UK based on subscriber count, videos uploaded and views accumulated, so that I can decide on which channels would be best to run marketing campaigns with to generate a good ROI.

Acceptance criteria

The dashboard should

- List the top YouTube channels by subscribers, videos and views
- Display key metrics (channel name, subscribers, videos, views, engagement ratios)
- Be user-friendly and easy to filter/sort
- Use the most recent data possible

2. Analyze the potential for marketing campaigns with Youtubers

User story

As the Head of Marketing, I want to analyze the potential for successful campaigns with the top Youtubers so that I can maximize the ROI

Acceptance criteria

The solution should

- Recommend YouTube channels best suited for different campaigns types (e.g. product placement, sponsored video series, influencer marketing)
- Consider reach, engagement and potential revenue based on estimated conversion rates
- Clearly explain the recommendations with data-driven justifications

Success criteria

Sharon can

- Easily identify the top performing YouTube channels based on the key metrics mentioned above
- Assess the potential for successful campaigns with top Youtubers based on reach, engagement and potential revenue
- Make informed decisions on the ideal collaborations to advance with based on recommendations

This allows Sharon to achieve a good ROI and build relationships with YouTubers for future collaborations, which leads to recognition within the company.

Information needed

Sharon needs the top Youtubers in the UK, and the key metrics needed include:

- Subscriber count
- Videos uploaded
- Views
- Average views
- Subscriber engagement ratio
- Views per subscriber

Data needed

The dataset to produce the information we need should include the following fields

- Channel name (string)
- Total subscribers (integer)
- Total videos uploaded (integer)
- Total views (integer)

Example of Query ran to obtain the required data field:

<pre>/* Data Cleaning steps: 1. Remove unnecessary column by selecting the one we need 2. Extract theyoutube channel names from the firts column 3. Rename the column names. */ SELECT NOMBRE, total_subscribers, total_views, total_videos FROM dbo.top_uk_youtuber_2024</pre>	<pre>-- extracting channel name/clean - using CHARINDEX and SUBSTRING function CREATE VIEW view_uk_youtuber_2024 AS SELECT CAST(SUBSTRING(NOMBRE, 1, CHARINDEX('@', NOMBRE) -1) AS VARCHAR(100)) AS channel_name, total_subscribers, total_views, total_videos FROM dbo.top_uk_youtuber_2024</pre>
---	--

Data quality checks

We need to add measures in place to confirm the dataset contains the data required without any issues – here are some of the data quality checks we need to conduct:

- Row count check
- Column count check
- Data type check
- Duplicate check

Example of Query ran to test the quality data. In order to make sure data is complete and accurate:

<pre>/* Testing data: 1. The data needs to be 100 records Youtube Channel (row count test) 2. The data need 4 fields (column test) 3. The channel column must be string format and the other column must be numeric data type (data types check). 4. Each record must be unique in the dataset (duplicate count check). Row count – 100 column count – 4 Data types - channel name = varchar - total_subscribers = int - total_views = int - total_videos = int duplicate_count = 0 */ -- 1. row count test SELECT COUNT(*) FROM view_uk_youtuber_2024;</pre>	<pre>-- 2. column count - using information schema SELECT COUNT(*) AS columns_count FROM INFORMATION_SCHEMA.COLUMNS WHERE TABLE_NAME = 'view_uk_youtuber_2024' -- 3. Data types check SELECT COLUMN_NAME, DATA_TYPE FROM INFORMATION_SCHEMA.COLUMNS WHERE TABLE_NAME = 'view_uk_youtuber_2024'; -- 4. Duplicate Test SELECT channel_name, COUNT(*) AS duplicate_count FROM view_uk_youtuber_2024 GROUP BY channel_name HAVING COUNT(*) > 1 -- TEST PASS!! COMPLETED (4/4) SELECT * FROM view_uk_youtuber_2024</pre>
---	--

Additional requirements

- Document the solution and include the data sources, transformation processes and walk through on analysis conclusions
- Make source code and docs available on GitHub
- Ensure the solution is reproducible and maintainable so that it can support future updates

Power Bi Dashboard: A snapshot of dashboard that has been created

