

TEAM ID: NM2023TMID07649

TEAM TOPIC: SUBSCRIBERS GALORE:EXPLORING THE WORLD'S TOP YOUTUBE CHANNELS

Data collection and extraction

Collect the Dataset:

Step 1:

The screenshot displays the Smart Internz web application interface. The browser address bar shows the URL: https://nme.smartinternz.com/Student/guided_project_info/6613#. The application has a dark sidebar on the left with navigation links: Home, Projects, and Support. The main content area is titled "Guided Project" and contains a vertical list of steps in a workflow:

- Specify The Business Problem
- Business Requirements
- Literature Survey (Student Will Write)
- Social Or Business Impact
- Data Collection & Extraction From Database
- Collect The Dataset** (highlighted in red)
- Uploading Data Into IBM Cognos
- Data Preparation
- Prepare The Data For Visualization
- Preparing A Data Module

The "Collect The Dataset" step is expanded, showing a white box with the text: "Please use the link to download the dataset: <https://www.kaggle.com/datasets/rajkumarpandey02/list-of-most-subscribed-youtube-channels-in-world>". The bottom of the screen shows a Windows taskbar with various application icons and a system tray indicating the time as 23:21 on 21-10-2023.

Step 2:

<https://www.kaggle.com/datasets/rajkumarpandey02/list-of-most-subscribed-youtube-channels-in-world>

The screenshot shows the Kaggle dataset page for 'List of most-subscribed YouTube channels in World'. The dataset is a CSV file (3.2 kB) with 67 rows and 9 columns. The columns are: # (Index Number), Rank, Name, Link, Brand channel, and Subscriber Primary Category. The dataset is titled '50 Most-Subscribed YouTube Channels in Worlds'. A bar chart shows the distribution of unique values for the Rank column, with 51 unique values. The dataset is available for download (1 kB) and has a version 1 (3.2 kB).

#	Rank	Name	Link	Brand channel	Subscriber Primary Category
0	1	T-Series	Link	Yes	238 Hindi[7][8] Music
1	2	Cocomelon	Link	Yes	155 English Education United States
2	3	Sony Entertainment Television India	Link	Yes	153 Hindi[9] Entertainmentn India
3	4	MrBeast	Link	No	137 English Entertainmentn United States
4	5	PewDiePie	Link	No	111 English Games Sweden

Step3:

The screenshot shows the Excel spreadsheet with the data from the Kaggle dataset. The data is organized into columns: Rank, Name, Link, Brand channel, Subscriber Primary Category, and Country. The data is sorted by Rank, with the top 19 channels listed.

Rank	Name	Link	Brand channel	Subscriber Primary Category	Country	
0	1	T-Series	Link	Yes	238 Hindi[7][8] Music	India
1	2	Cocomelon	Link	Yes	155 English Education	United States
2	3	Sony Entertainment Television India	Link	Yes	153 Hindi[9] Entertainmentn	India
3	4	MrBeast	Link	No	137 English Entertainmentn	United States
4	5	PewDiePie	Link	No	111 English Games	Sweden
5	6	Kids Diana	Link	Yes	109 English[10] Entertainmentn	Ukraine
6	7	Like Nasty	Link	No	105 English Entertainmentn	United States
7	8	Vlad and N Link	Link	No	94.9 English Entertainmentn	Russia
8	9	WWE	Link	Yes	93.8 English Sports	United States
9	10	Zee Music	Link	Yes	93.4 Hindi[13][Music	India
10	11	Blackpink	Link	Yes	84.8 Korean Music	South Korea
11	12	Goldmine	Link	Yes	83.3 Hindi Film	India
12	13	5-Minute Link	Link	Yes	79.2 English How-to	Cyprus[a]
13	14	Sony SAB	Link	Yes	78.2 Hindi Entertainmentn	India
14	15	BangtanTV	Link	No	73.9 Korean Music	South Korea
15	16	Justin Bie	Link	No	71.1 English Music	Canada
16	17	Hybe Labe	Link	Yes	69.5 Korean Music	South Korea
17	18	Canal Kon	Link	Yes	66.4 Portugues Music	Brazil

Uploading the data in IBM Cognos:

The screenshot displays the IBM Cognos Analytics 'Content' page. The browser address bar shows the URL: <https://ap1.ca.analytics.ibm.com/bi/?perspective=content>. The page title is 'Content'. Below the title, there are tabs for 'My content', 'Team content', 'Samples', and 'Favorites'. The 'My content' tab is active. On the right side of the page, there is a 'New' button with a plus icon. Below this, there is a table listing content items. The table has three columns: 'Name', 'Type', and 'Last Modified'. The items listed are:

Name	Type	Last Modified
<input type="checkbox"/> List of most-subscribed YouTube channels.csv CSV	Uploaded file	19/10/2023, 01:29
<input type="checkbox"/> pie chart	Exploration	19/10/2023, 02:38
<input type="checkbox"/> Subscribers dashboard	Dashboard	19/10/2023, 02:58
<input type="checkbox"/> subscribers story	Story	20/10/2023, 04:03

The right sidebar contains a 'New' button and a vertical list of icons for various content types and tools.