

# CNIT 372 Group Project

## Milestone 3

Megan Clecak  
Yasmine Graham  
Elliot Martin  
Marshall Bledsoe

## Problem Description:

Our group chose to research Option B: YouTube Data Analysis as a Content Creator. This data can be used to analyze genres, videos, and popular youtube channels within those genres. YouTube is a growing platform where individuals or groups can upload videos and create content for a specified audience. An average content creator (YouTuber) can make \$0.18 per view or \$18 per 1,000 views. One of the most popular YouTubers made an estimated \$54 million in 2021, with some of his videos reaching an upwards of 142 million views<sup>1</sup>. We can analyze popular content creators like MrBeast to measure the number of uploads, content category, and amount of views. We can also analyze the earnings of the content to gain insights into which content generates the most income. This information could help future YouTubers create successful channels that meet a target audience.

We are using the Global YouTube Statistics dataset from Kaggle. We plan on analyzing which categories are most popular, in which country they are most popular, and how much earnings are generated. Based on that information, a potential content creator will gain insightful information about what content is successful and in which country.

---

<sup>1</sup> Daugherty & Sineriz. "How Much Do Youtubers Make?"

# Questions:

## Question 1:

Which channel categories or niches are found in the dataset?

This question is important for a new or aspiring youtube content creator because they may want to look at what categories there are for youtube. Based on this query, you can also look at the frequency of counts per category, which may give some insight on how popular one category is over another. A new creator may also consider how popular this category is, indicating it may also be difficult to start a channel in that category because of how many youtubers are already creating content. For example, while broad, Entertainment may not be a good category to choose from because there are over 241 counts in this category.

The solution created for this question was a procedure that counts the number of categories from the YoutuberInfo table. The results given are the counts of how many youtubers create content in this category per the data set used from Kaggle. Using our prior knowledge of PL/SQL procedures, we created a query that will use a for loop to select the categories from each row and create a list of the different categories along with the count of the times it occurs. Since it is a procedure, we run the code by execution. The DBMS outputs the category and the count in a list.

### **Solution:**

```
CREATE OR REPLACE PROCEDURE GetCategoryCounts AS
BEGIN
    FOR category_rec IN (SELECT category, COUNT(*) AS category_count
                        FROM youtuberinfo
                        GROUP BY category
                        ORDER BY category_count DESC)
    LOOP
        DBMS_OUTPUT.PUT_LINE('Category: ' || category_rec.category);
        DBMS_OUTPUT.PUT_LINE('Count: ' || category_rec.category_count);
        DBMS_OUTPUT.PUT_LINE("");
    END LOOP;
END GetCategoryCounts;
/

EXEC GetCategoryCounts;
```

### **Results:**

```
Category: Entertainment
Count: 241
Category: Music
```

Count: 202  
Category: People & Blogs  
Count: 132  
Category: Gaming  
Count: 94  
Category: Comedy  
Count: 69  
Category: Film & Animation  
Count: 45  
Category: Education  
Count: 45  
Category: Howto & Style  
Count: 40  
Category:  
Count: 39  
Category: News & Politics  
Count: 26  
Category: Science & Technology  
Count: 17  
Category: Shows  
Count: 13  
Category: Sports  
Count: 11  
Category: Pets & Animals  
Count: 4  
Category: Trailers  
Count: 2  
Category: Autos & Vehicles  
Count: 2  
Category: Nonprofits & Activism  
Count: 2  
Category: Movies  
Count: 2  
Category: Travel & Events  
Count: 1

## Question 2:

For each country, what is the most (and least) popular genre or category for channel content?

This question is important because as a content creator, growth can be achieved much faster if you know what topic and/or genre is most popular in your area or country at large. The categories provide information regarding the preferences for the general viewer. Listing the categories for a specified country will allow a new content creator to specifically cater to a greater audience. While there are some common categories with multiple countries, it is still beneficial information to gauge and base the new youtube channel on those targeted audiences.

For this solution we created a procedure called FindPopularGenres with a loop to sort through the YoutuberInfo table and find the most and least popular genre based on those youtubers in the data set used from Kaggle. From our previous knowledge of PL/SQL Procedures, to create two variables, most\_popular\_category and least\_popular\_category, to store the data found from the table query. The program will loop through each distinct country to find the most and least popular genre. The solution also gives a DBMS output that displays the country and the two created variables in a column, separating each country's data with dashes.

### **Solution:**

```
CREATE OR REPLACE PROCEDURE FindPopularGenres AS
BEGIN
    FOR country_rec IN (SELECT DISTINCT country FROM YOUTUBERINFO)
    LOOP

        DECLARE
            most_popular_category VARCHAR2(255);
            least_popular_category VARCHAR2(255);
        BEGIN

            SELECT category
            INTO most_popular_category
            FROM YOUTUBERINFO
            WHERE Country = country_rec.Country
            ORDER BY uploads DESC
            FETCH FIRST 1 ROWS ONLY;

            SELECT category
            INTO least_popular_category
            FROM YOUTUBERINFO
            WHERE country = country_rec.country
            ORDER BY uploads ASC
            FETCH FIRST 1 ROWS ONLY;
```

```

        DBMS_OUTPUT.PUT_LINE('Country: ' || country_rec.Country);
        DBMS_OUTPUT.PUT_LINE('Most Popular Category: ' || most_popular_category);
        DBMS_OUTPUT.PUT_LINE('Least Popular Category: ' || least_popular_category);
        DBMS_OUTPUT.PUT_LINE('-----');
    EXCEPTION
        WHEN NO_DATA_FOUND THEN
            DBMS_OUTPUT.PUT_LINE('No data found for Country: ' || country_rec.Country);
            DBMS_OUTPUT.PUT_LINE('-----');
    END;
END LOOP;
END FindPopularGenres;
/

```

EXECUTE FINDPOPULARGENRES;

### Results:

Country: United States  
 Most Popular Category: Nonprofits & Activism  
 Least Popular Category: Film & Animation

-----  
 Country: Indonesia  
 Most Popular Category: News & Politics  
 Least Popular Category: Entertainment

-----  
 Country: Thailand  
 Most Popular Category: News & Politics  
 Least Popular Category: Music

-----  
 Country: Saudi Arabia  
 Most Popular Category: Gaming  
 Least Popular Category: Entertainment

-----  
 Country: Bangladesh  
 Most Popular Category: News & Politics  
 Least Popular Category: News & Politics

-----  
 Country: Finland  
 Most Popular Category: Gaming  
 Least Popular Category: Gaming

-----  
 Country: Brazil  
 Most Popular Category: Entertainment

Least Popular Category: Music

-----  
Country: Argentina

Most Popular Category: Gaming

Least Popular Category: Music

-----  
Country: South Korea

Most Popular Category: Entertainment

Least Popular Category: Music

-----  
Country: Malaysia

Most Popular Category: Film & Animation

Least Popular Category: Film & Animation

-----  
Country: Morocco

Most Popular Category: Music

Least Popular Category: Music

-----  
Country: Russia

Most Popular Category: Gaming

Least Popular Category: People & Blogs

-----  
Country: Spain

Most Popular Category: Gaming

Least Popular Category: Entertainment

-----  
Country: Iraq

Most Popular Category: Music

Least Popular Category: Comedy

-----  
Country: Japan

Most Popular Category: Entertainment

Least Popular Category: Entertainment

-----  
Country: Barbados

Most Popular Category: Entertainment

Least Popular Category: Entertainment

-----  
Country: Venezuela

Most Popular Category: People & Blogs

Least Popular Category: People & Blogs

-----  
Country: Netherlands

Most Popular Category: Music

Least Popular Category: Howto & Style

-----

Country: Singapore

Most Popular Category: Entertainment

Least Popular Category: People & Blogs

-----

Country: Italy

Most Popular Category: Entertainment

Least Popular Category: Entertainment

-----

Country: Egypt

Most Popular Category: Howto & Style

Least Popular Category: People & Blogs

-----

Country: Chile

Most Popular Category: Gaming

Least Popular Category: Gaming

-----

Country: Germany

Most Popular Category: Entertainment

Least Popular Category: Comedy

-----

Country: Latvia

Most Popular Category: Comedy

Least Popular Category: Comedy

-----

Country: France

Most Popular Category: Gaming

Least Popular Category: People & Blogs

-----

Country: Cuba

Most Popular Category: Gaming

Least Popular Category: Gaming

-----

Country: Peru

Most Popular Category: Entertainment

Least Popular Category: Entertainment

-----

Country: Samoa

Most Popular Category: Music

Least Popular Category: Music

-----

Country: Mexico

Most Popular Category: People & Blogs



Least Popular Category: Comedy

-----

No data found for Country:

-----

Country: Afghanistan

Most Popular Category: Music

Least Popular Category: Music

-----

Country: Ukraine

Most Popular Category: News & Politics

Least Popular Category: Entertainment

-----

Country: Vietnam

Most Popular Category: Music

Least Popular Category: Entertainment

-----

Country: El Salvador

Most Popular Category: Gaming

Least Popular Category: Gaming

-----

Country: Pakistan

Most Popular Category: News & Politics

Least Popular Category:

-----

Country: Australia

Most Popular Category: Gaming

Least Popular Category: Gaming

-----

Country: Kuwait

Most Popular Category: Gaming

Least Popular Category: Gaming

-----

Country: Philippines

Most Popular Category: News & Politics

Least Popular Category:

-----

Country: Sweden

Most Popular Category:

Least Popular Category: Gaming

-----

Country: Canada

Most Popular Category: Music

Least Popular Category: Gaming

-----

Country: China  
Most Popular Category: Howto & Style  
Least Popular Category: Howto & Style

-----

Country: Jordan  
Most Popular Category: Entertainment  
Least Popular Category: Music

-----

Country: Andorra  
Most Popular Category: Howto & Style  
Least Popular Category: Howto & Style

-----

Country: Ecuador  
Most Popular Category: People & Blogs  
Least Popular Category: Entertainment

-----

Country: United Arab Emirates  
Most Popular Category: People & Blogs  
Least Popular Category: Gaming

-----

Country: India  
Most Popular Category: People & Blogs  
Least Popular Category: Entertainment

-----

Country: United Kingdom  
Most Popular Category: News & Politics  
Least Popular Category: Howto & Style

-----

Country: Colombia  
Most Popular Category: Music  
Least Popular Category: Music

-----

Country: Turkey  
Most Popular Category: Music  
Least Popular Category: Entertainment

-----

Country: Switzerland  
Most Popular Category: Sports  
Least Popular Category: Sports

-----

### Question 3:

What are the top-ranked YouTubers for various countries, highlighting their respective categories or genres, that contribute to their popularity?

**\*\*Content creators from their respective country can use this data to see how they may be most successful on YouTube in their country.**

The question allows a potential content creator to understand a bit more about which categories are most popular in a country. Also, we can understand more based on the top ranked YouTubers. The new creator could find out who the top creators are and do further research such as watching their videos. Another unique feature of this question is that the content creator could specifically look for their country or target their audience in a country who may prefer the category that the new content creator makes.

We used a procedure called "topRanked" to list the YouTuber name, the country, and the channel type from the channel\_stats table. This data then retrieves the rows where the rank\_in\_country attribute is equal to 1. It then loops through the dataset until no more rows are detected, then the loop ends. The results then list the youtuber, their country, and their channel type in a comfortably formatted manner

#### **Solution:**

create or replace procedure topRanked

as

begin

    for current\_youtuber in (

        select channel\_stats.youtuber, youtuberinfo.country, channel\_type

        from channel\_stats

        inner join youtuberinfo

        on channel\_stats.youtuber = youtuberinfo.youtuber

        where rank\_in\_country = '1') loop

        dbms\_output.put\_line('Youtuber: ' || RPAD(current\_youtuber.youtuber, 50) || ' Country: ' ||

RPAD(current\_youtuber.country, 30) || ' Channel Type: ' ||

RPAD(current\_youtuber.channel\_type, 30));

    end loop;

end topRanked;

/

EXEC topRanked;

## Results:

Youtuber: يyyyyyy يyyyyyy يyyyyyy   toyorajanahtv	Country: Jordan	Channel Type: Music
Youtuber: Spinnin' Records	Country: Netherlands	Channel Type: Music
Youtuber: Bounce Patrol - Kids Songs	Country: Australia	Channel Type: Music
Youtuber: PANDA BOI	Country: Italy	Channel Type: Entertainment
Youtuber: Tsuriki Show	Country: Germany	Channel Type: Comedy
Youtuber: mmoshaya	Country: Saudi Arabia	Channel Type: Entertainment
Youtuber: SQUEEZIE	Country: France	Channel Type: Entertainment
Youtuber: Like Nastya VNM	Country: Vietnam	Channel Type: Entertainment
Youtuber: Les' Copaque Production	Country: Malaysia	Channel Type: Entertainment
Youtuber: يyyyyyy Liziqi	Country: China	Channel Type: Howto
Youtuber: يyyyyyy - Al-Remas	Country: Iraq	Channel Type: Music
Youtuber: LEGO	Country: Singapore	Channel Type: Entertainment
Youtuber: يyyyyyy يyyyyyy يyyyyyy 5 يyyyyyy	Country: Egypt	Channel Type: Howto
Youtuber: Doc Tops	Country: Ecuador	Channel Type: Entertainment
Youtuber: Saad Lamjarred   اّ اّ اّ اّ اّ اّ اّ اّ اّ اّ	Country: Morocco	Channel Type: Music
Youtuber: Drawblogs	Country: Peru	Channel Type: Entertainment
Youtuber: Avicii	Country: Sweden	Channel Type: Music
Youtuber: SlivkiShow	Country: Ukraine	Channel Type: Entertainment
Youtuber: dednahype	Country: Latvia	Channel Type: Comedy
Youtuber: FIFA	Country: Switzerland	Channel Type: Sports
Youtuber: T-Series	Country: India	Channel Type: Music
Youtuber: MrBeast	Country: United States	Channel Type: Entertainment
Youtuber: PewDiePie	Country: Japan	Channel Type: Entertainment
Youtuber: BLACKPINK	Country: South Korea	Channel Type: Music
Youtuber: Justin Bieber	Country: Canada	Channel Type: Music
Youtuber: Canal KondZilla	Country: Brazil	Channel Type: Music
Youtuber: El Reino Infantil	Country: Argentina	Channel Type: Music
Youtuber: Ed Sheeran	Country: United Kingdom	Channel Type: Music
Youtuber: JuegaGerman	Country: Chile	Channel Type: Games
Youtuber: Get Movies	Country: Russia	Channel Type: Film
Youtuber: Fernanfloo	Country: El Salvador	Channel Type: Games
Youtuber: HAR PAL GEO	Country: Pakistan	Channel Type: Entertainment
Youtuber: ABS-CBN Entertainment	Country: Philippines	Channel Type: Entertainment
Youtuber: Shakira	Country: Colombia	Channel Type: Howto
Youtuber: Rihanna	Country: Barbados	Channel Type: Music
Youtuber: shfa2 - اّ اّ اّ اّ	Country: United Arab Emirates	Channel Type: People
Youtuber: Mikecrack	Country: Spain	Channel Type: Games
Youtuber: WorkpointOfficial	Country: Thailand	Channel Type: Entertainment
Youtuber: PowerKids TV	Country: Channel Type: Film	
Youtuber: Jess No Limit	Country: Indonesia	Channel Type: Entertainment
Youtuber: BETER Bi اّ اّ اّ	Country: Turkey	Channel Type: Games
Youtuber: AboFlah	Country: Kuwait	Channel Type: Entertainment
Youtuber: Brawl Stars	Country: Finland	Channel Type: Games

## Question 4:

How do the average earnings of top YouTubers vary by country?

In order for a YouTuber to increase their income, they may want to emulate the content of the top earners on the site and, more importantly, the YouTuber's country.

This procedure finds the average earnings of the top n youtubers in each country. If the procedure is executed without the number parameter, it will show the top 3 YouTubers by default. One parameter is required, the country name. Based on this country name, the procedure will find a list of the top YouTubers from that country.

This procedure uses our knowledge of user defined variables, input parameters, subqueries, inner joins, loops, and basic arithmetic to compute a YouTubers ranking.

### **Solution:**

```
create or replace procedure top5EachCountry(country_choice in varchar2, numrows in number)
as
earnings number;
lowest number;
highest number;
begin
    for current_youtuber in(
        select youtuberinfo.youtuber, rank_in_country, youtuberinfo.country,
        ((channel_stats.lowest_yearly_earnings + channel_stats.highest_yearly_earnings)/2),
        channel_stats.lowest_yearly_earnings, channel_stats.highest_yearly_earnings
        from youtuberinfo
        inner join channel_stats
        on youtuberinfo.youtuber = channel_stats.youtuber
        where rank_in_country <= numrows+1 AND youtuberinfo.country is not null AND
        youtuberinfo.country = country_choice
        order by channel_stats.country, rank_in_country) loop
        lowest := current_youtuber.lowest_yearly_earnings;
        highest := current_youtuber.highest_yearly_earnings;
        earnings := (lowest + highest)/2;
        dbms_output.put_line('Youtuber: ' || RPAD(current_youtuber.youtuber, 50) || ' Country: ' ||
        RPAD(current_youtuber.country, 30) || ' Channel Rank: ' || RPAD(earnings, 30));
    end loop;

end top5EachCountry;
/
```

```
exec top5EachCountry('India', 5);
```

**Results:**

Youtuber: T-Series

Country: India

Channel Rank: 57600000

Youtuber: SET India

Country: India

Channel Rank: 46500000

Youtuber: Zee Music Company

Country: India

Channel Rank: 20500000

Youtuber: Sony SAB

Country: India

Channel Rank: 42300000

Youtuber: Zee TV

Country: India

Channel Rank: 43500000

## Question 5:

What is the average earnings for YouTubers per content category?

Understanding the average earnings per content category will help inform youtubers about which content categories are most profitable. This will likely drive certain categories to be really popular when they go viral. This will also show which categories may go under represented in the algorithm, and have room for growth.

This procedure estimates average earnings for each YouTube category by adding all the category's lowest yearly earnings and the category's highest yearly earnings then dividing that number by the number of YouTubers in that category. Since this information is in different tables an inner join statement was used. This procedure uses group by functions to perform the calculations on each category. The result is grouped by category and ordered by average earnings in descending order. This procedure uses a cursor and a loop to display the results in the DBMS Output window.

### **Solution:**

CREATE or REPLACE procedure CategoryAvgEarnings

as

CURSOR All\_Categories is

select Category, sum(Lowest\_yearly\_earnings)+sum(highest\_yearly\_earnings)/(2\*count(\*))

as AvgEarnings, count(\*) as NumInCategory

from Channel\_Stats STATS inner join YouTuberInfo INFO

on STATS.Youtuber = INFO.YouTuber

group by Category

order by AvgEarnings desc;

begin

DBMS\_OUTPUT.PUT(RPAD('Category', 25, ' '));

DBMS\_OUTPUT.PUT(LPAD('AvgEarnings', 15, ' '));

DBMS\_OUTPUT.PUT\_LINE(RPAD(' NumInCategory', 15, ' '));

DBMS\_OUTPUT.PUT(RPAD('-', 24, '-'));

DBMS\_OUTPUT.PUT(RPAD(' ', 16, '-'));

DBMS\_OUTPUT.PUT\_LINE(RPAD(' ', 15, '-'));

for current\_category in All\_Categories LOOP

DBMS\_OUTPUT.PUT(RPAD(nvl(current\_category.Category, 'N/A'), 20, ' '));

DBMS\_OUTPUT.PUT(LPAD(to\_char(current\_category.AvgEarnings,  
'\$999,999,999,999'),20,' '));

DBMS\_OUTPUT.PUT\_LINE(LPAD(current\_category.NumInCategory, 12, ' '));

end LOOP;

end CategoryAvgEarnings;

/

execute CategoryAvgEarnings;

**Results:**

Category	AvgEarnings	NumInCategory
Shows	\$12,955,381	13
N/A	\$7,402,073	39
Autos & Vehicles	\$6,994,875	2
Sports	\$5,109,023	11
Pets & Animals	\$5,091,696	4
Film & Animation	\$4,776,213	45
Education	\$4,421,939	45
Comedy	\$4,269,795	69
News & Politics	\$4,095,717	26
Entertainment	\$3,975,413	241
Music	\$3,529,024	202
People & Blogs	\$3,484,644	132
Movies	\$2,895,525	2
Nonprofits & Activism	\$2,496,375	2
Trailers	\$2,310,700	2
Gaming	\$1,727,601	94
Science & Technology	\$1,283,507	17
Howto & Style	\$1,230,536	40
Travel & Events	\$796,500	1



# Individual Contributions:

## Marshall

- Cleaned data set with google sheets and created 3 new data sheets
- Assisted creating 3 new tables for data set
- Created tables in SQL and imported data
- Created working database
- Wrote the procedure for question 5

## Megan

- Assisted creating 3 new tables for data set
- Assisted helping clean data set with google sheets
- Created tables in SQL and imported data
- Created "master database" and granted permissions to all group members
- Assisted creating the import data set into empty database script worksheet
- De-bugged and fixed query for question 3
- Completed descriptions/solution portions of question 1 & 2

## Elliot

- Fixed the Milestone 1 questions and descriptions
- Created PL/SQL solutions for questions 2 and 4
- Helped edit and finish question descriptions

## Yasmine

- Cleaned the data so that it would import into SQL seamlessly
  - Changed the "NAN" values to "NULL" so SQL could understand the value
- Imported data into SQL Developer
- Granted SELECT, UPDATE access to the teammates
- Created import data set into empty database script
- Updated Problem Description
- Wrote a query to solve question 1

# Citations

Daughtery, K. & Sineriz, M. (2023, September 12). *How Much do Youtubers Make?* Finance Buzz.

<https://financebuzz.com/how-much-youtubers-make#:~:text=The%20average%20YouTuber%20makes%20about,%2C%20up%2040%25%20from%202020>