CNIT 372 Final Project

Jordan Jollief (jjollief@purdue.edu)

Cassie Kim (kim4128@purdue.edu)

Dillon Wiggington (wiggingt@purdue.edu)

Motivation

We as a group were motivated to choose the option of YouTube as a creator. This is because none of us are creators on YouTube, but we all are consumers of YouTube, meaning we know what we like as a consumer and can tell what video will be good before we watch it, since we are the audience. We chose YouTube as a creator because we wanted to know the kind of data that creators would look at when they are uploading a video or thinking about uploading their next video. We wanted to know how a creator would think when presented data and how they would choose to upload videos, depending on the audience's reaction to their video.

Database Description

Tables:

- creator_data
    o YouTube Analysis dataset
- trending_data
    o YouTube trending dataset
- Comments
    o Comment csv file within YouTube Statistics dataset

Columns:

- creator_data: videotitle, videoviews, likes, no_of_comments, videolink, creatorname, uploaddate, etc.
- trending_data: videoid, title, views, likes, uploaddate, trendingdate, videolink, etc.
- comments: videoid, commentid, comment_text, likes, etc.

Questions

1. How can you create a cursor to display the video_id, the amount of likes and the title of the video from the trending_data table?
    a. This will help with using cursors, but it will also help creators see the correlation between likes and title of the video.

SQL

```sql
declare
    v_id trending_data.video_id%type;
    v_likes trending_data.likes%type;
    v_name trending_data.creatorname%type;
    cursor v_video is
        select video_id, likes, creatorname from trending_data;
begin
    open v_video;
    loop
    fetch v_video into v_id, v_likes, v_name;
        exit when v_video%notfound;
        dbms_output.put_line(v_id || ' ' || v_likes || ' ' || v_name);
    end loop;
    close v_video;
end;
```

Results

```
YwJyis4yOEA 20393 Carli Bybel
4OCcbUZuTUQ 359609 TREASURE (트레저)
NkE0AMGzpJY 1599289 MrBeast
JFm7YDVlqnI 1006096 DrakeVEVO
QcTwbXLMm9c 293238 charli d'amelio
dYZlrfarCIs 25720 Champions League on CBS Sports
vwnyP6979xE 11925 Champions League on CBS Sports
qDFjsMFojzs 147308 jeffreestar
ozWNPkhqi8o 93530 The Boys
7mH-ug3Dgo8 1793 CBS Sports HQ
0C80BSgjb8M 334693 YoungBoy Never Broke Again
m1t1ooj_Em0 79106 Simply Nailogical
4fSGeKU5Mvw 652 MMA Fight Nation
jWQNqlZ47Ec 40995 HSM
SCUXdRb5abU 166131 Apex Legends
FbM1yi4mMMc 1705178 MrBeast
6BEfZCH1nN8 33123 Official Ed and Lorraine Warren Channel
xE2f745DICI 956 SHOWTIME Sports
4rjNpA941y4 8446 UFC - Ultimate Fighting Championship
b8pSdMHicYw 953 kenanK TV
0opZqh_TprM 213877 Lyrical Lemonade
YNyFAro1E5M 293549 PoloGVEVO
fqT81qdPpOw 146350 P2istheName
oBOKCRuo2y4 10696 C&C Life
_rJYc_k-w84 6818 billschannel
```

2. What language is used the most out of all videos?
   a. This will help creators understand what the majority of the audience speaks and what language is the most used on YouTube.

SQL

```
create or replace procedure LanguageCount
    (most_used_language out varchar2) as
begin
    select language into most_used_language
    from (
        select language, count(*) as language_count
        from creator_data
        group by language
        order by language_count desc
    )
    where rownum = 1;
end;

declare
    most_used_language varchar2(50);
begin
    LanguageCount(most_used_language);
    dbms_output.put_line('The most used language is: ' || most_used_language);
end;
```

Results

The most used language is: English

3. How can you create a trigger that lets a user know when you are going to delete something from the 'comments' table?
   a. This will help users understand when a row is being deleted from the table, and also helps us use triggers in the correct way.

SQL

```
create or replace trigger comment_trigger
    before delete on comments
    for each row
begin
    dbms_output.put_line('Trigger fired before deleting a row from COMMENTS');
end;
```

Results

Trigger fired before deleting a row from COMMENTS

4. Does having more than 3 hashtags bring more subscribers than average?
   a. This helps creators understand the correlation between hashtag amounts and the subscription rate across videos.

SQL

```
WITH AverageSubscribers AS (
    SELECT AVG(TotalChannelSubcribers) AS AvgSubscribers
    FROM YoutubeAnalysis
)

SELECT
    CreatorName,
    Hashtags,
    TotalChannelSubcribers,
    CASE
        WHEN y.TotalChannelSubcribers > x.AvgSubscribers THEN 'Above Average'
        ELSE 'Below or Equal Average'
    END AS SubscriberComparison
FROM YoutubeAnalysis y
CROSS JOIN AverageSubscribers x
WHERE y.Hashtags > 3;
```

Results

```
CREATORNAME HASHTAGS TOTALCHANNELSUBSCRIBERS SUBSCRIBERCOMPARISON
----------------------------------------------------------------
Etl Qa Labs         4      7800     Below or Equal Average
100 Years Of Health 7      25500    Below or Equal Average
World Of Hemant 11         34       Below or Equal Average
Yidircars           6      551000   Below or Equal Average
Watchgecko          4      43500    Below or Equal Average
Vidyut Jammwal  6          1370000  Below or Equal Average
Namanh Kapur        5      105000   Below or Equal Average
Entertainment Live  6      15600    Below or Equal Average
Dji Tutorials       4      421000   Below or Equal Average
Ganesh Prasad       5      2250000  Below or Equal Average
Great Learning  4          694000   Below or Equal Average
Ganesh Prasad       5      2250000  Below or Equal Average
NBC News            4      7060000  Below or Equal Average
Thu Vu Data Analytics 6    52600    Below or Equal Average
Ganesh Prasad       4      2250000  Below or Equal Average
Dr Vivek Bindra 10         19800000        Above Average
Ganesh Prasad       8      2250000  Below or Equal Average
Ganesh Prasad       9      2250000  Below or Equal Average
Ganesh Prasad       5      2250000  Below or Equal Average
Ganesh Prasad       5      2250000  Below or Equal Average
Chef Ranveer        10     538000   Below or Equal Average
Luke Barouse        4      243000   Below or Equal Average
```

5. What is the average length of a video in seconds, and how does it affect the views?
   a. This question helps creators understand how long they should be making their videos and how much of an effect the duration has on the views.

SQL

```sql
create or replace procedure duration as
begin
    for dur in (
        select avg(duration_sec) as avg_in_sec, videoviews
        from creator_data
        group by videoviews
        order by videoviews)
    loop
        dbms_output.put_line('Average Duration In Seconds: ' || dur.avg_in_sec);
        dbms_output.put_line('Views: ' || dur.videoviews);
        dbms_output.put_line('----------------------------------------');
    end loop;
end;

exec duration;
```

Results

```
Average Duration In Seconds: 1429
Views: 745314
------------------------------------
Average Duration In Seconds: 1475
Views: 909640
------------------------------------
Average Duration In Seconds: 456
Views: 1049688
------------------------------------
Average Duration In Seconds: 603
Views: 1118970
------------------------------------
```

6. What is the relationship between channel views and likes on a video where the title is longer than 30 characters?
    a. This helps creators understand what makes an eye catching title whether it is long or short.

SQL

```
create or replace procedure title as
begin
    for v_title in (
        select videoviews, likes, videotitle
        from creator_data
        where length(videotitle) > 30
        order by likes)
    loop
        dbms_output.put_line('Video Title: ' || v_title.videotitle);
        dbms_output.put_line('Views: ' || v_title.videoviews);
        dbms_output.put_line('Likes: ' || v_title.likes);
        dbms_output.put_line('-------------------------------------');
    end loop;
end;
```

Results

```
Video Title: Unnoticable things when Storing information in Excel | YourExcelGuy
Views: 100
Likes: 6
-------------------------------------
Video Title: Byjus | Difference Inside vs Direct vs Centre Sale | BDA sale department 2022 #bda #byjus
Views: 766
Likes: 13
-------------------------------------
Video Title: Dreams Rush | Ayush Shrivastava
Views: 1200
Likes: 25
-------------------------------------
Video Title: Top countries with proven natural gas reserves
Views: 9965
Likes: 188
-------------------------------------
Video Title: Are Stress & Worry Hurting Your Sleep? Here's How to Get More Sleep
Views: 12288
Likes: 369
-------------------------------------
Video Title: This Is Why Red Bull Keeps Winning Races
Views: 169135
Likes: 688
-------------------------------------
Video Title: DISCIPLINE YOUR BRAIN | Wake Up Positive | Morning Motivational Speeches
Views: 62367
Likes: 771
-------------------------------------
```

7. What is the correlation between a video's upload date and the date it begins to trend?
   a. This will help creators understand that videos do not trend automatically and will also help creators see the dates that popular videos trend on.

SQL

```
create or replace procedure VideoDate as
begin
    for video_date in (
        select videotitle, publishedon, trending_date
        from trending_data)
    loop
        dbms_output.put_line('Video Title: ' || video_date.videotitle);
        dbms_output.put_line('Published On: ' || video_date.publishedon);
        dbms_output.put_line('Trending Date: ' || video_date.trending_date);
        dbms_output.put_line('-----------------------------------');
    end loop;
end;


exec videodate;
```

Results

```
-------------------------------------
Video Title: Apex Legends | Stories from the Outlands â€" â€œThe Endorsementâ€¿
Published On: 2020-08-11T17:00:10Z
Trending Date: 2020-08-12T00:00:00Z
-------------------------------------
Video Title: I left youtube for a month and THIS is what happened.
Published On: 2020-08-11T16:34:06Z
Trending Date: 2020-08-12T00:00:00Z
-------------------------------------
Video Title: Ultimate DIY Home Movie Theater for The LaBrant Family!
Published On: 2020-08-11T15:10:05Z
Trending Date: 2020-08-12T00:00:00Z
-------------------------------------
Video Title: I Haven't Been Honest About My Injury.. Here's THE TRUTH
Published On: 2020-08-11T20:00:04Z
Trending Date: 2020-08-12T00:00:00Z
-------------------------------------
Video Title: OUR FIRST FAMILY INTRO!!
Published On: 2020-08-12T00:17:41Z
Trending Date: 2020-08-12T00:00:00Z
-------------------------------------
Video Title: CGP Grey was WRONG
Published On: 2020-08-11T17:15:11Z
Trending Date: 2020-08-12T00:00:00Z
-------------------------------------
```

8. What comment has the most likes and what does the comment say?
    a. This helps creators engage with their community, allowing them to understand what similar experiences their audience had.

SQL

```sql
create or replace procedure most_liked_comment is
    v_comment varchar2(1024);
    v_likes number(38,0);
begin
    select comment_text, likes
    into v_comment, v_likes
    from comments
    order by likes desc
    fetch first row only;

    if v_comment is not null then
        dbms_output.put_line('Most Liked Comment: ' || v_comment);
        dbms_output.put_line('Likes: ' || v_likes);
    end if;
end;

exec most_liked_comment;
```

Results

```
Most Liked Comment: Like I said in the video, subscribe if you haven't already and you could win $10,000!
Likes: 891372
```

9. How can you find the creator's name, channel views, and channel subscribers for channels with the most number of weekly uploads?

SQL

```sql
CREATE OR REPLACE PROCEDURE PopularEngagement

AS

c_title CREATOR_DATA.VideoTitle%type;
c_views CREATOR_DATA.VideoViews%type;
c_likes CREATOR_DATA.NoOfLikes%type;
c_comments CREATOR_DATA.NoOfComments%type;

CURSOR c_popularengagement IS
    SELECT  VideoTitle, VideoViews, NoOfLikes, NoOfComments
    FROM CREATOR_DATA
    WHERE NoOfLikes > 1000000
    ORDER BY NoOfLikes DESC;

BEGIN

    OPEN c_popularengagement;
    LOOP
    FETCH c_popularengagement into c_title, c_views, c_likes, c_comments;
        EXIT WHEN c_popularengagement%notfound;
        DBMS_OUTPUT.PUT_LINE(c_title || ' ' || c_views || ' ' || c_likes || ' ' || c_comments);
    END LOOP;
    CLOSE c_popularengagement;
END;
```

Results

```
Ed Sheeran - Perfect 3278291072 18000000 476983
World's Most Dangerous Escape Room! 138831703 4000000 144901
Coke Studio Season 8| Tajdar-e-Haram| Atif Aslam 408785205 3300000 197501
Alone (Official Music Video) 2317708089 2300000 1000000
Fabiolous Escape 2 153649156 2100000 33768
I Attended Spider-Man : No Way Home Premiere | Ashish Chanchlani | LA vlog 21175732 2000000 49215
Camila Cabello - Bam Bam (Official Music Video) ft. Ed Sheeran 124158710 1700000 36550
School PTM 36559485 1600000 40241
Taylor Swift ft. Chris Stapleton - I Bet You Think About Me (Taylor's Version) 45332993 1500000 46497
Ava Max - Who's Laughing Now 141644942 1300000 39000
Childhood Dreams | Aakash Gupta | Stand-up Comedy | Crowd Work 28494371 1200000 12192
I opened free 5 star hotels for poor 11234302 1100000 78759
```

10. For videos with over 1,000,000 views, which video had the least amount of likes and how many comments were on the video?
    a. This question helps creators understand why a video with so many views might not have gotten the most viewer activity.

SQL

```sql
create or replace procedure one_million
    (least_liked_video out varchar2) as
begin
    select videotitle into least_liked_video
    from (
        select videotitle, likes, videoviews, comment_count
        from creator_data
        where videoviews > 1000000
        group by videotitle, likes, videoviews, comment_count
        having likes = (select min(likes) from creator_data)
    )
    where rownum = 1;
end;

declare
    least_liked_video varchar2(128);
begin
    one_million(least_liked_video);
    dbms_output.put_line('The video that has over 1,000,000 views and the least amount of likes is: ' || least_liked_video);
end;
```

Results

The video that has over 1,000,000 views and the least number of likes is: Pokémon Journeys

Team

Jordan Jollief – Senior in CIT

- Contributions
    o Milestones 1,2,3, SQL questions, set up GitHub, report, presentation slides.

Cassie Kim – Senior in CIT

- Contributions
    o Milestones 1,2,3, SQL questions, added to GitHub, report, presentation slides.

Dillon Wiggington – Senior in CIT

- Contributions
    o Milestones 1,2,3, SQL questions, added to GitHub, report, presentation slides.