DMDD Assignment 4

Project Objective:

The objective of our project is to design a database of YouTube and Twitter influencers catering to business marketing needs. The scope of our project includes records on the influencer's reach, type of content, and origin of the content, amongst other things. This sophisticated system can further be used to deduce and match which social media influencer is a good fit for advertising businesses based on many criteria.

Team Members:

1. Hamzah Mukadam (NUID: 002741426)
2. Shreya Maher (NUID: 002770227)
3. Veenadharini Shukla (NUID: 002704948)

We have Scrapped real time data from Twitter and YouTube and normalized our database for the Twitter and Youtube Tables as below:

1st Normal Form

Table Name: Youtube_Profile_Data

1. Eliminate duplicative columns from the same table.

    - No duplicative columns found

2. Create separate tables for each group of related data
    Table Name: Youtube_Profile_Data
      - The column type could be separated in a different table called category with primary key as id:

```
1      CREATE TABLE category
2    ⊝ AS (SELECT DISTINCT type
3          FROM youtube_profile_data);
4
5 •    ALTER TABLE category ADD id INT PRIMARY KEY AUTO_INCREMENT;
6
7 •    SELECT * FROM influencer_demo.category;
```

id	type
1	Gaming
2	Programming
3	Fashion
4	Makeup
5	Sports
6	News Channel
7	type
8	News Blogger
9	Fashion Blogger
10	Fitness
11	Movie Reviews
12	Food Blogger
13	Tech Channels
14	News
15	Food
16	Techology
17	Horror Stories
18	Beauty

category 2 ✕   Apply   Revert

3. Identify each row with a unique column(primary key)

   Each row is uniquely identified by column: handle

   ALTER TABLE youtube_profile_data ADD PRIMARY KEY(handle);

Table Name: profile_data_youtube_twitter

1. Eliminate duplicative columns from the same table.
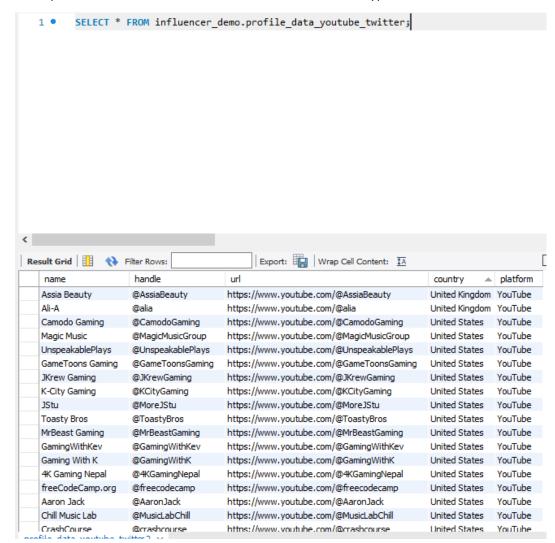
   -No duplicative columns found, all columns are of the same type

```
1 •    SELECT * FROM influencer_demo.profile_data_youtube_twitter;
```

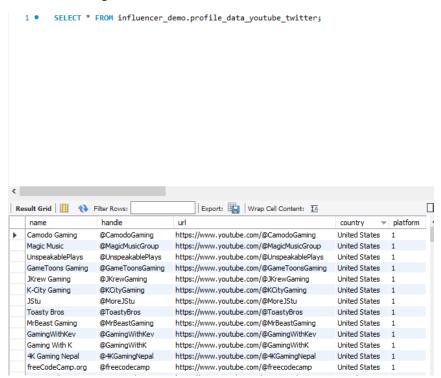| name	handle	url	country	platform
Assia Beauty	@AssiaBeauty	https://www.youtube.com/@AssiaBeauty	United Kingdom	YouTube
Ali-A	@alia	https://www.youtube.com/@alia	United Kingdom	YouTube
Camodo Gaming	@CamodoGaming	https://www.youtube.com/@CamodoGaming	United States	YouTube
Magic Music	@MagicMusicGroup	https://www.youtube.com/@MagicMusicGroup	United States	YouTube
UnspeakablePlays	@UnspeakablePlays	https://www.youtube.com/@UnspeakablePlays	United States	YouTube
GameToons Gaming	@GameToonsGaming	https://www.youtube.com/@GameToonsGaming	United States	YouTube
JKrew Gaming	@JKrewGaming	https://www.youtube.com/@JKrewGaming	United States	YouTube
K-City Gaming	@KCityGaming	https://www.youtube.com/@KCityGaming	United States	YouTube
JStu	@MoreJStu	https://www.youtube.com/@MoreJStu	United States	YouTube
Toasty Bros	@ToastyBros	https://www.youtube.com/@ToastyBros	United States	YouTube
MrBeast Gaming	@MrBeastGaming	https://www.youtube.com/@MrBeastGaming	United States	YouTube
GamingWithKev	@GamingWithKev	https://www.youtube.com/@GamingWithKev	United States	YouTube
Gaming With K	@GamingWithK	https://www.youtube.com/@GamingWithK	United States	YouTube
4K Gaming Nepal	@4KGamingNepal	https://www.youtube.com/@4KGamingNepal	United States	YouTube
freeCodeCamp.org	@freecodecamp	https://www.youtube.com/@freecodecamp	United States	YouTube
Aaron Jack	@AaronJack	https://www.youtube.com/@AaronJack	United States	YouTube
Chill Music Lab	@MusicLabChill	https://www.youtube.com/@MusicLabChill	United States	YouTube
CrashCourse	@crashcourse	https://www.youtube.com/@crashcourse	United States	YouTube

2. Create separate tables for each group of related data

   For type Platform since can be grouped we have created a new table and inserted grouped data

   CREATE TABLE `profile_platform` (

     `platform_key` int(11) DEFAULT NULL,

     `platform_type` varchar(30) DEFAULT NULL

)

```
1 •    INSERT INTO `influencer_demo`.`profile_platform`
2   ⊖ (`platform_key`,
3       `platform_type`)
4      VALUES
5   ⊖ (2,
6      'Twitter');
7 •    SELECT * FROM influencer_demo.profile_platform;
```

	platform_key	platform_type
▶	1	YouTube
	2	Twitter

Table after changes:

```
1 •    SELECT * FROM influencer_demo.profile_data_youtube_twitter;
```

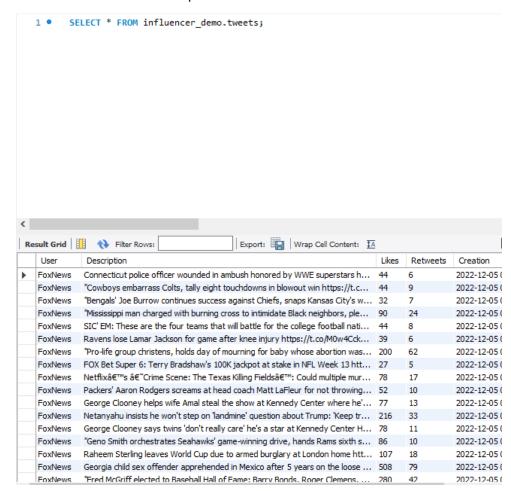| name	handle	url	country	platform
Camodo Gaming	@CamodoGaming	https://www.youtube.com/@CamodoGaming	United States	1
Magic Music	@MagicMusicGroup	https://www.youtube.com/@MagicMusicGroup	United States	1
UnspeakablePlays	@UnspeakablePlays	https://www.youtube.com/@UnspeakablePlays	United States	1
GameToons Gaming	@GameToonsGaming	https://www.youtube.com/@GameToonsGaming	United States	1
JKrew Gaming	@JKrewGaming	https://www.youtube.com/@JKrewGaming	United States	1
K-City Gaming	@KCityGaming	https://www.youtube.com/@KCityGaming	United States	1
JStu	@MoreJStu	https://www.youtube.com/@MoreJStu	United States	1
Toasty Bros	@ToastyBros	https://www.youtube.com/@ToastyBros	United States	1
MrBeast Gaming	@MrBeastGaming	https://www.youtube.com/@MrBeastGaming	United States	1
GamingWithKev	@GamingWithKev	https://www.youtube.com/@GamingWithKev	United States	1
Gaming With K	@GamingWithK	https://www.youtube.com/@GamingWithK	United States	1
4K Gaming Nepal	@4KGamingNepal	https://www.youtube.com/@4KGamingNepal	United States	1
freeCodeCamp.org	@freecodecamp	https://www.youtube.com/@freecodecamp	United States	1

3. Identify each row with a unique column(primary key) ALTER TABLE
   `influencer_demo`.`profile_platform`
   CHANGE COLUMN `platform_key` `platform_key` INT(11) NOT NULL ,
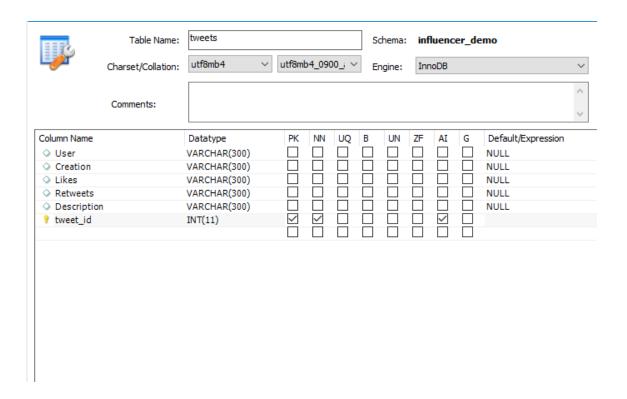   ADD PRIMARY KEY (`platform_key`);;

Table Name: tweets

1. Eliminate duplicative columns from the same table.

   The table does not consist duplicate columns

2. Create separate tables for each group of related data

   No separate tables needed as the table does not comprise of group data

3. Identify each row with a unique column(primary key)

   Since Tweets are multiple we cannot have user as primary key, hence we created a new primary
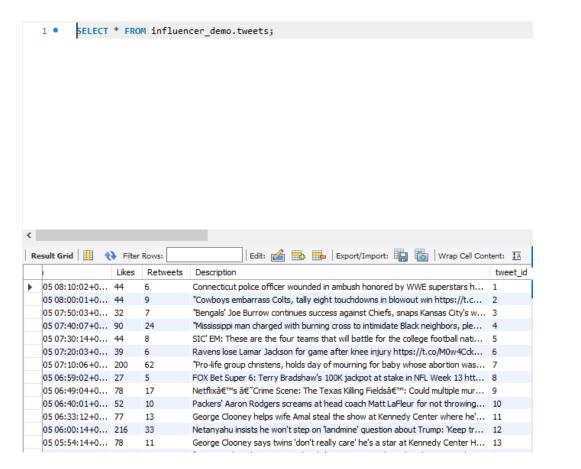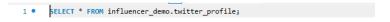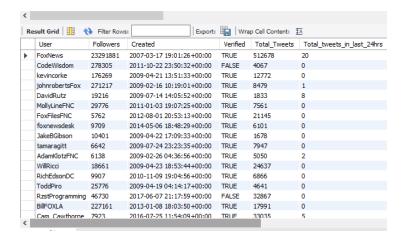
   key called tweet_id

| Table Name: | tweets | | Schema: | **influencer_demo** |
| Charset/Collation: | utf8mb4 | utf8mb4_0900_i | Engine: | InnoDB |

Comments:

Column Name	Datatype	PK	NN	UQ	B	UN	ZF	AI	G	Default/Expression
◇ User	VARCHAR(300)	☐	☐	☐	☐	☐	☐	☐	☐	NULL
◇ Creation	VARCHAR(300)	☐	☐	☐	☐	☐	☐	☐	☐	NULL
◇ Likes	VARCHAR(300)	☐	☐	☐	☐	☐	☐	☐	☐	NULL
◇ Retweets	VARCHAR(300)	☐	☐	☐	☐	☐	☐	☐	☐	NULL
◇ Description	VARCHAR(300)	☐	☐	☐	☐	☐	☐	☐	☐	NULL
🔑 tweet_id	INT(11)	☑	☑	☐	☐	☐	☐	☑	☐	
		☐	☐	☐	☐	☐	☐	☐	☐	

```
1 •    SELECT * FROM influencer_demo.tweets;
```

Result Grid | Filter Rows: | Edit: | Export/Import: | Wrap Cell Content:

	Likes	Retweets	Description	tweet_id
05 08:10:02+0...	44	6	Connecticut police officer wounded in ambush honored by WWE superstars h...	1
05 08:00:01+0...	44	9	"Cowboys embarrass Colts, tally eight touchdowns in blowout win https://t.c...	2
05 07:50:03+0...	32	7	"Bengals' Joe Burrow continues success against Chiefs, snaps Kansas City's w...	3
05 07:40:07+0...	90	24	"Mississippi man charged with burning cross to intimidate Black neighbors, ple...	4
05 07:30:14+0...	44	8	SIC' EM: These are the four teams that will battle for the college football nati...	5
05 07:20:03+0...	39	6	Ravens lose Lamar Jackson for game after knee injury https://t.co/M0w4Cck...	6
05 07:10:06+0...	200	62	"Pro-life group christens, holds day of mourning for baby whose abortion was...	7
05 06:59:02+0...	27	5	FOX Bet Super 6: Terry Bradshaw's 100K jackpot at stake in NFL Week 13 htt...	8
05 06:49:04+0...	78	17	Netflixâ€™s â€˜Crime Scene: The Texas Killing Fieldsâ€™: Could multiple mur...	9
05 06:40:01+0...	52	10	Packers' Aaron Rodgers screams at head coach Matt LaFleur for not throwing...	10
05 06:33:12+0...	77	13	George Clooney helps wife Amal steal the show at Kennedy Center where he'...	11
05 06:00:14+0...	216	33	Netanyahu insists he won't step on 'landmine' question about Trump: 'Keep tr...	12
05 05:54:14+0...	78	11	George Clooney says twins 'don't really care' he's a star at Kennedy Center H...	13

Table Name: twitter profile

1. Eliminate duplicative columns from the same table.

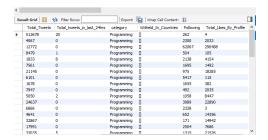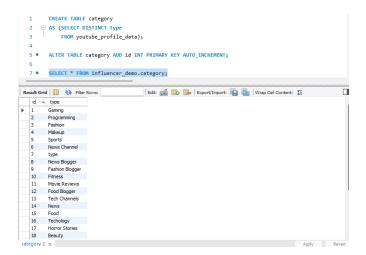   -No duplicative columns found, all columns are of the same type

   ```
   1 •   SELECT * FROM influencer_demo.twitter_profile;
   ```

User	Followers	Created	Verified	Total_Tweets	Total_tweets_in_last_24hrs
FoxNews	23291881	2007-03-17 19:01:26+00:00	TRUE	512678	20
CodeWisdom	278305	2011-10-22 23:50:32+00:00	FALSE	4067	0
kevincorke	176269	2009-04-21 13:51:33+00:00	TRUE	12772	0
johnrobertsFox	271217	2009-02-16 10:19:01+00:00	TRUE	8479	1
DavidRutz	19216	2009-07-14 14:05:52+00:00	TRUE	1833	8
MollyLineFNC	29776	2011-01-03 19:07:25+00:00	TRUE	7561	0
FoxFilesFNC	5762	2012-08-01 20:53:13+00:00	TRUE	21145	0
foxnewsdesk	9709	2014-05-06 18:48:29+00:00	TRUE	6101	0
JakeBGibson	10401	2009-04-22 17:09:33+00:00	TRUE	1678	0
tamaragitt	6642	2009-07-24 23:23:35+00:00	TRUE	7947	0
AdamKlotzFNC	6138	2009-02-26 04:36:56+00:00	TRUE	5050	2
WillRicci	18661	2009-04-23 18:53:44+00:00	TRUE	24637	0
RichEdsonDC	9907	2010-11-09 19:04:56+00:00	TRUE	6866	0
ToddPiro	25776	2009-04-19 04:14:17+00:00	TRUE	4641	0
RzstProgramming	46730	2017-06-07 21:17:59+00:00	FALSE	32867	0
BillFOXLA	227161	2013-01-08 18:03:50+00:00	TRUE	17991	0
Cam_Cawthorne	7923	2016-07-25 11:54:09+00:00	TRUE	33035	5

2. The column category could be separated in a different table called category with primary key as id:

   before:

   ```
   1 •   SELECT * FROM influencer_demo.twitter_profile;
   ```
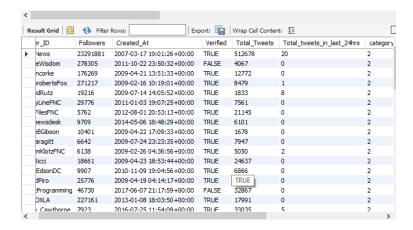
Total_Tweets	Total_tweets_in_last_24hrs	category	Witheld_In_Countries	Following	Total_Likes_By_Profile
512678	20	Programming	⬚	262	4
4067	0	Programming	⬚	2300	2033
12772	0	Programming	⬚	62067	290488
8479	1	Programming	⬚	504	105
1833	8	Programming	⬚	2138	4154
7561	0	Programming	⬚	1695	1492
21145	0	Programming	⬚	975	18285
6101	0	Programming	⬚	5417	115
1678	0	Programming	⬚	1655	382
7947	0	Programming	⬚	492	2035
5050	2	Programming	⬚	1058	8447
24637	0	Programming	⬚	3989	22890
6866	0	Programming	⬚	2328	3
4641	0	Programming	⬚	652	14356
32867	0	Programming	⬚	171	14942
17991	0	Programming	⬚	2504	7686
33035	5	Programming	⬚	1315	21526

   Creation of category as in table 2

```
1    CREATE TABLE category
2    AS (SELECT DISTINCT type
3        FROM youtube_profile_data);
4
5  ●  ALTER TABLE category ADD id INT PRIMARY KEY AUTO_INCREMENT;
6
7  ●  SELECT * FROM influencer_demo.category;
```

Result Grid | Filter Rows: | Edit: | Export/Import: | Wrap Cell Content: I_A

id	type
1	Gaming
2	Programming
3	Fashion
4	Makeup
5	Sports
6	News Channel
7	type
8	News Blogger
9	Fashion Blogger
10	Fitness
11	Movie Reviews
12	Food Blogger
13	Tech Channels
14	News
15	Food
16	Techology
17	Horror Stories
18	Beauty

category 2 ×   Apply   Revert

After updation based on table

```
1  ●  SELECT * FROM influencer_demo.twitter_profile;
```

Result Grid | Filter Rows: | Export: | Wrap Cell Content: I_A

er_ID	Followers	Created_At	Verified	Total_Tweets	Total_tweets_in_last_24hrs	category
News	23291881	2007-03-17 19:01:26+00:00	TRUE	512678	20	2
eWisdom	278305	2011-10-22 23:50:32+00:00	FALSE	4067	0	2
ncorke	176269	2009-04-21 13:51:33+00:00	TRUE	12772	0	2
rrobertsFox	271217	2009-02-16 10:19:01+00:00	TRUE	8479	1	2
idRutz	19216	2009-07-14 14:05:52+00:00	TRUE	1833	8	2
yLineFNC	29776	2011-01-03 19:07:25+00:00	TRUE	7561	0	2
FilesFNC	5762	2012-08-01 20:53:13+00:00	TRUE	21145	0	2
newsdesk	9709	2014-05-06 18:48:29+00:00	TRUE	6101	0	2
eBGibson	10401	2009-04-22 17:09:33+00:00	TRUE	1678	0	2
aragitt	6642	2009-07-24 23:23:35+00:00	TRUE	7947	0	2
mKlotzFNC	6138	2009-02-26 04:36:56+00:00	TRUE	5050	2	2
Ricci	18661	2009-04-23 18:53:44+00:00	TRUE	24637	0	2
EdsonDC	9907	2010-11-09 19:04:56+00:00	TRUE	6866	0	2
dPiro	25776	2009-04-19 04:14:17+00:00	TRUE	TRUE	0	2
Programming	46730	2017-06-07 21:17:59+00:00	FALSE	32867	0	2
OXLA	227161	2013-01-08 18:03:50+00:00	TRUE	17991	0	2
Cawthorne	7923	2016-07-25 11:54:09+00:00	TRUE	33035	5	2

3. Identify each row with a unique column(primary key)

ALTER TABLE `influencer_demo`.`twitter_profile`

CHANGE COLUMN `User_ID` `User_ID` VARCHAR(300) NOT NULL ,

ADD PRIMARY KEY (`User_ID`);

;

User_ID is primary key and uniquely identifies data in table

Tables Added in 1NF: Category, Profile platform

2nd Normal Form

Table Name: Youtube_Profile_Data

1. Meet all requirements of 1NF

    -We have demonstrated that all requirements of 1NF are met in the above screenshots.

2. Remove subsets of data that apply to multiple rows of a table and place them in separate tables.

    -We did not have duplicative columns of data to begin with, hence not required to separate the subsets of the data to a different table.
    For 'type' column we identified multiple rows of duplicate data hence separated that into a table 'category' with an additional auto-increment primary key 'id'.

3. Create relationships between these new tables and their predecessors through use of foreign keys:

    -Updated the type column in the Youtube_Profile_data table with the foreign key:

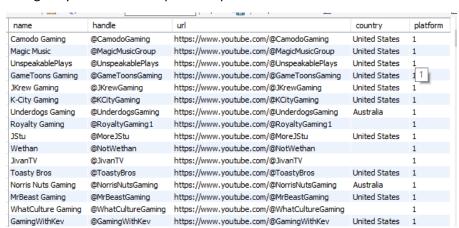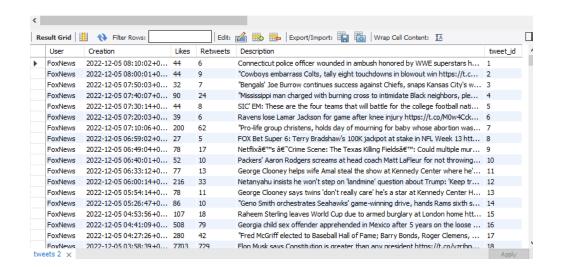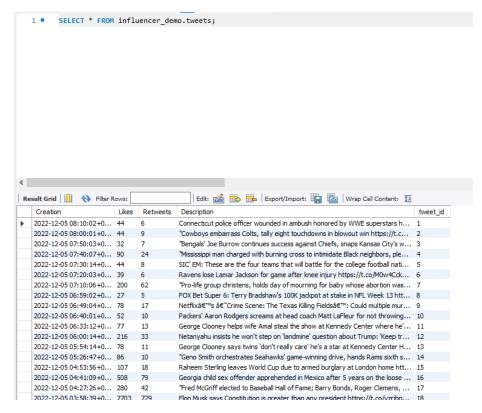| ï»¿handle	subscribers	otherLinks	joinedon	total views	type
TechyPathshala	606K	['https://www.youtube.com/channel/UC9Cowa...	2017-16-2	165,727,011	1
QAYZERGAMING	104K	['https://www.youtube.com/channel/UCfXoodv...	2018-27-11	50,137,826	1
RAHULGAMING920	1.6M	['https://www.youtube.com/redirect?event=ch...	2014-25-10	166,347,105	1
4KGamingNepal	1.04M	['https://www.youtube.com/redirect?event=ch...	2017-18-10	84,339,506	1
MiroHQ	24.7K	['https://www.youtube.com/redirect?event=ch...	2012-12-9	0	2
2withmosh	2.95M	['https://www.youtube.com/redirect?event=ch...	2014-6-10	142,329,938	2
freecodecamp	6.81M	['https://www.youtube.com/redirect?event=ch...	2014-16-12	465,384,906	2
AaronJack	338K	['https://www.youtube.com/redirect?event=ch...	2018-9-9	21,085,075	2
MusicLabChill	782K	['https://www.youtube.com/redirect?event=ch...	2019-5-7	93,156,733	2
TEDxTalks	37M	['https://www.youtube.com/redirect?event=ch...	2009-23-6	0	2
crashcourse	14.2M	['https://www.youtube.com/redirect?event=ch...	2006-19-5	1,750,944,033	2
namanhkapur	136K	['https://www.youtube.com/redirect?event=ch...	2012-14-9	7,507,520	2
CroatCode	513K	['https://www.youtube.com/redirect?event=ch...	2012-17-6	52,967,719	2

Table Name: profile_data_youtube_twitter

1. Meet all requirements of 1NF

   -We have demonstrated that all requirements of 1NF are met in the above screenshots.

2. Remove subsets of data that apply to multiple rows of a table and place them in separate tables.

   -We did not have duplicative columns of data to begin with, hence not required to separate the subsets of the data to a different table.

   For 'type' column we identified multiple rows of duplicate data hence separated that into a table 'category' with an additional auto-increment primary key 'id'.

3. Create relationships between these new tables and their predecessors through use of foreign keys:

   ALTER TABLE `influencer_demo`.`profile_platform`

   CHANGE COLUMN `platform_key` `platform_key` INT(11) NOT NULL ,

   ADD PRIMARY KEY (`platform_key`);

   ;

   Foreign key relation developed with platform

name	handle	url	country	platform
Camodo Gaming	@CamodoGaming	https://www.youtube.com/@CamodoGaming	United States	1
Magic Music	@MagicMusicGroup	https://www.youtube.com/@MagicMusicGroup	United States	1
UnspeakablePlays	@UnspeakablePlays	https://www.youtube.com/@UnspeakablePlays	United States	1
GameToons Gaming	@GameToonsGaming	https://www.youtube.com/@GameToonsGaming	United States	1
JKrew Gaming	@JKrewGaming	https://www.youtube.com/@JKrewGaming	United States	1
K-City Gaming	@KCityGaming	https://www.youtube.com/@KCityGaming	United States	1
Underdogs Gaming	@UnderdogsGaming	https://www.youtube.com/@UnderdogsGaming	Australia	1
Royalty Gaming	@RoyaltyGaming1	https://www.youtube.com/@RoyaltyGaming1		1
JStu	@MoreJStu	https://www.youtube.com/@MoreJStu	United States	1
Wethan	@NotWethan	https://www.youtube.com/@NotWethan		1
JivanTV	@JivanTV	https://www.youtube.com/@JivanTV		1
Toasty Bros	@ToastyBros	https://www.youtube.com/@ToastyBros	United States	1
Norris Nuts Gaming	@NorrisNutsGaming	https://www.youtube.com/@NorrisNutsGaming	Australia	1
MrBeast Gaming	@MrBeastGaming	https://www.youtube.com/@MrBeastGaming	United States	1
WhatCulture Gaming	@WhatCultureGaming	https://www.youtube.com/@WhatCultureGaming		1
GamingWithKev	@GamingWithKev	https://www.youtube.com/@GamingWithKev	United States	1

Table Name: tweets

1. Meet all requirements of 1NF

   -We have demonstrated that all requirements of 1NF are met in the above screenshots.

2. Remove subsets of data that apply to multiple rows of a table and place them in separate tables.

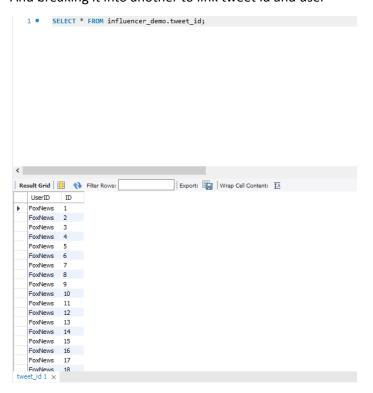   Breaking tweet_id and user

   Before:


```
1 •    SELECT * FROM influencer_demo.tweets;
```

User	Creation	Likes	Retweets	Description	tweet_id
FoxNews	2022-12-05 08:10:02+0...	44	6	Connecticut police officer wounded in ambush honored by WWE superstars h...	1
FoxNews	2022-12-05 08:00:01+0...	44	9	"Cowboys embarrass Colts, tally eight touchdowns in blowout win https://t.c...	2
FoxNews	2022-12-05 07:50:03+0...	32	7	"Bengals' Joe Burrow continues success against Chiefs, snaps Kansas City's w...	3
FoxNews	2022-12-05 07:40:07+0...	90	24	"Mississippi man charged with burning cross to intimidate Black neighbors, ple...	4
FoxNews	2022-12-05 07:30:14+0...	44	8	SIC' EM: These are the four teams that will battle for the college football nati...	5
FoxNews	2022-12-05 07:20:03+0...	39	6	Ravens lose Lamar Jackson for game after knee injury https://t.co/M0w4Cck...	6
FoxNews	2022-12-05 07:10:06+0...	200	62	"Pro-life group christens, holds day of mourning for baby whose abortion was...	7
FoxNews	2022-12-05 06:59:02+0...	27	5	FOX Bet Super 6: Terry Bradshaw's 100K jackpot at stake in NFL Week 13 htt...	8
FoxNews	2022-12-05 06:49:04+0...	78	17	Netflixâ€™s â€˜Crime Scene: The Texas Killing Fieldsâ€™: Could multiple mur...	9
FoxNews	2022-12-05 06:40:01+0...	52	10	Packers' Aaron Rodgers screams at head coach Matt LaFleur for not throwing...	10
FoxNews	2022-12-05 06:33:12+0...	77	13	George Clooney helps wife Amal steal the show at Kennedy Center where he'...	11
FoxNews	2022-12-05 06:00:14+0...	216	33	Netanyahu insists he won't step on 'landmine' question about Trump: 'Keep tr...	12
FoxNews	2022-12-05 05:54:14+0...	78	11	George Clooney says twins 'don't really care' he's a star at Kennedy Center H...	13
FoxNews	2022-12-05 05:26:47+0...	86	10	"Geno Smith orchestrates Seahawks' game-winning drive, hands Rams sixth s...	14
FoxNews	2022-12-05 04:53:56+0...	107	18	Raheem Sterling leaves World Cup due to armed burglary at London home htt...	15
FoxNews	2022-12-05 04:41:09+0...	508	79	Georgia child sex offender apprehended in Mexico after 5 years on the loose ...	16
FoxNews	2022-12-05 04:27:26+0...	280	42	"Fred McGriff elected to Baseball Hall of Fame; Barry Bonds, Roger Clemens, ...	17
FoxNews	2022-12-05 03:58:39+0...	7703	729	Elon Musk says Constitution is greater than any president https://t.co/yzribp...	18

Dropping user from table

```
1 •    SELECT * FROM influencer_demo.tweets;
```

Creation	Likes	Retweets	Description	tweet_id
2022-12-05 08:10:02+0...	44	6	Connecticut police officer wounded in ambush honored by WWE superstars h...	1
2022-12-05 08:00:01+0...	44	9	"Cowboys embarrass Colts, tally eight touchdowns in blowout win https://t.c...	2
2022-12-05 07:50:03+0...	32	7	"Bengals' Joe Burrow continues success against Chiefs, snaps Kansas City's w...	3
2022-12-05 07:40:07+0...	90	24	"Mississippi man charged with burning cross to intimidate Black neighbors, ple...	4
2022-12-05 07:30:14+0...	44	8	SIC' EM: These are the four teams that will battle for the college football nati...	5
2022-12-05 07:20:03+0...	39	6	Ravens lose Lamar Jackson for game after knee injury https://t.co/M0w4Cck...	6
2022-12-05 07:10:06+0...	200	62	"Pro-life group christens, holds day of mourning for baby whose abortion was...	7
2022-12-05 06:59:02+0...	27	5	FOX Bet Super 6: Terry Bradshaw's 100K jackpot at stake in NFL Week 13 htt...	8
2022-12-05 06:49:04+0...	78	17	Netflix's 'Crime Scene: The Texas Killing Fields': Could multiple mur...	9
2022-12-05 06:40:01+0...	52	10	Packers' Aaron Rodgers screams at head coach Matt LaFleur for not throwing...	10
2022-12-05 06:33:12+0...	77	13	George Clooney helps wife Amal steal the show at Kennedy Center where he'...	11
2022-12-05 06:00:14+0...	216	33	Netanyahu insists he won't step on 'landmine' question about Trump: 'Keep tr...	12
2022-12-05 05:54:14+0...	78	11	George Clooney says twins 'don't really care' he's a star at Kennedy Center H...	13
2022-12-05 05:26:47+0...	86	10	"Geno Smith orchestrates Seahawks' game-winning drive, hands Rams sixth s...	14
2022-12-05 04:53:56+0...	107	18	Raheem Sterling leaves World Cup due to armed burglary at London home htt...	15
2022-12-05 04:41:09+0...	508	79	Georgia child sex offender apprehended in Mexico after 5 years on the loose ...	16
2022-12-05 04:27:26+0...	280	42	"Fred McGriff elected to Baseball Hall of Fame; Barry Bonds, Roger Clemens, ...	17
2022-12-05 03:58:39+0...	7703	729	Elon Musk says Constitution is greater than any president https://t.co/yzribp...	18

And breaking it into another to link tweet id and user

```
1 •    SELECT * FROM influencer_demo.tweet_id;
```

UserID	ID
FoxNews	1
FoxNews	2
FoxNews	3
FoxNews	4
FoxNews	5
FoxNews	6
FoxNews	7
FoxNews	8
FoxNews	9
FoxNews	10
FoxNews	11
FoxNews	12
FoxNews	13
FoxNews	14
FoxNews	15
FoxNews	16
FoxNews	17
FoxNews	18

Table Name: twitter profile

1. Meet all requirements of 1NF

   -We have demonstrated that all requirements of 1NF are met in the above screenshots.

2. Remove subsets of data that apply to multiple rows of a table and place them in separate tables.

   -We did not have duplicative columns of data, hence not required to separate the subsets of the data to a different table.

3rd Normal Form

Table Name: Youtube_Profile_Data

1. Meet all requirements of 2NF

   -We have demonstrated that all requirements of 1NF are met in the above screenshots

2. Remove the columns that are not dependent upon the primary key

   - We do not have any columns dependent upon the primary key

Table Name: profile_data_youtube_twitter

1. Meet all requirements of 2NF

   -We have demonstrated that all requirements of 1NF are met in the above screenshots

2. Remove the columns that are not dependent upon the primary key

   - We do not have any columns dependent upon the primary key

Table Name: tweets

1. Meet all requirements of 2NF

   -We have demonstrated that all requirements of 1NF are met in the above screenshots

2. Remove the columns that are not dependent upon the primary key

   - We do not have any columns dependent upon the primary key

Table Name: twitter profile

1. Meet all requirements of 2NF

   -We have demonstrated that all requirements of 1NF are met in the above screenshots

2. Remove the columns that are not dependent upon the primary key

   - We do not have any columns dependent upon the primary key

Creating VIEWS for the SQL queries previously created for the Use Cases:

USE CASE 1: Name and handle of the YouTuber with the highest view count.

SQL:

create view HighestViewCount as

select profile_data.name, profile_data.handle, max(youtube_profile_data.total_views) as max_views

from profile_data join youtube_profile_data

on profile_data.handle = youtube_profile_data.handle

group by youtube_profile_data.handle

order by max(youtube_profile_data.total_views) desc

limit 1;

USE CASE 2: Country of the YouTuber with the highest view count.

SQL:

create view CountryofHighestViewCount as

select profile_data.name, profile_data.country, max(youtube_profile_data.total_views) as max_views

from

profile_data join youtube_profile_data

on profile_data.handle = youtube_profile_data.handle

group by youtube_profile_data.handle

order by max(youtube_profile_data.total_views) desc

limit 1;

USE CASE 3: Name and handle of the YouTuber with the highest view count based on the country.

SQL:

create view SlovakiaHighestViewCount as

select profile_data.name, profile_data.country, max(youtube_profile_data.total_views) as max_views

from

profile_data join youtube_profile_data

on profile_data.handle = youtube_profile_data.handle

where country = 'Slovakia'

group by youtube_profile_data.handle

order by max(youtube_profile_data.total_views) desc;

USE CASE 4: Name and handle of the YouTuber with the with the highest amount of subscribers.

SQL:

create view HighestNoOfSubscribers as

select profile_data.name, profile_data.handle, max(youtube_profile_data.subscribers) as subscribers

from

profile_data join youtube_profile_data

on profile_data.handle = youtube_profile_data.handle

group by youtube_profile_data.handle

order by max(youtube_profile_data.subscribers) desc

limit 1;

USE CASE 5: Country of the YouTuber with the highest amount of subscribers.

SQL:

create view HighestSubscriberYoutuberCountry as

select profile_data.name, profile_data.country, max(youtube_profile_data.subscribers) as subscribers

from

profile_data join youtube_profile_data

on profile_data.handle = youtube_profile_data.handle

group by youtube_profile_data.handle

order by max(youtube_profile_data.subscribers) desc

limit 1;

USE CASE 7: Name and handle of the YouTuber with the highest amount of subscribers based on a country.

SQL:

create view HighestSubscriberCountryYoutuber as

SELECT profile_data.name, profile_data.handle

FROM profile_data INNER JOIN youtube_profile_data

ON profile_data.handle = youtube_profile_data.handle

WHERE profile_data.country = "United States"

GROUP BY youtube_profile_data.total_views

HAVING MAX(youtube_profile_data.subscribers)

order by max(youtube_profile_data.total_views) desc

limit 1;

USE CASE 8: Handle of YouTuber with the highest views based on a country.

SQL:

create view HighestViewsCountryYoutuber as

SELECT youtube_profile_data.handle, youtube_profile_data.total_views

FROM youtube_profile_data INNER JOIN profile_data

ON youtube_profile_data.handle = profile_data.handle

WHERE profile_data.country ="United States"

GROUP BY youtube_profile_data.total_views

ORDER by youtube_profile_data.total_views DESC;

USE CASE 9: Handles of the YouTubers with the highest subscribers based on a country.

SQL:

create view HighestSubscriberCountryYoutuber as

SELECT influencers.youtube profile data.handle

FROM influencers.youtube profile data INNER JOIN influencers.profile_data

ON influencers.youtube profile data.handle = influencers.profile data.handle

GROUP BY influencer.youtube profile data.subscribers DESC HAVING country = "";

USE CASE 10: Join date of the recommended YouTuber based on country.

SQL:

create view CountryYoutuberJoinDate as

select profile_data.name, profile_data.handle, youtube_profile_data.joined_on from

profile_data join youtube_profile_data

on profile_data.handle = youtube_profile_data.handle

where country = 'United States'

group by youtube_profile_data.handle

;

USE CASE 11: Name and handle of the recommended YouTuber based on subscriber count.

SQL:

create view SubscriberCountYoutuber as

select profile_data.name, profile_data.handle from

profile_data join youtube_profile_data

on profile_data.handle = youtube_profile_data.handle

where youtube_profile_data.subscribers = '1.72M subscribers'

group by youtube_profile_data.handle;

USE CASE 12: Countries of the YouTubers sorted by type.

SQL:

create view TypeSortedYoutuberCountries as

select profile_data.name, profile_data.handle, youtube_profile_data.type from

profile_data join youtube_profile_data

on profile_data.handle = youtube_profile_data.handle

where youtube_profile_data.type = 'Programming'

group by youtube_profile_data.handle;

USE CASE 13: Types of the YouTubers sorted by countries.

SQL:

create view CountrySortedYoutuberTypes as

select profile_data.name, profile_data.handle, youtube_profile_data.type from

profile_data join youtube_profile_data

on profile_data.handle = youtube_profile_data.handle

where youtube_profile_data.type = 'Programming' and youtube_profile_data.country = 'United States'

group by youtube_profile_data.handle;

USE CASE 14: Total views of a YouTuber sorted by type and countries.

SQL:

create view Types&CountriesSorted_YoutuberTotalViews as

SELECT influencer.profile data.total views FROM influencer.profile data INNER JOIN influencer.youtuber

profile data ON influencer.profile_ data.handle = influencer.youtuber profile.handle GROUP BY type,

country;

USE CASE 15: URL of the recommended YouTuber based on subscribers.

SQL:

create view Subscriber_YoutuberURL as

SELECT influencer.profile data.url FROM influencer.profile data INNER JOIN influencer.youtuber profile

data ON influencer.profile_ data.handle = influencer.youtuber profile.handle WHERE subscribers = "";