

Pokemon community database Report

Yuhang Li 001442226 Yujin Xiang 001496297 Yijin Zhu 001498312

Instructed by : Prof. nikbearbrown

Abstract: In this project, we created this database for Pokemon Community which contained gamers, game developers, and game contents. We used the Twitter, Reddit, Instagram to gather data and establish connections between the different entities. What we want to acquire is to find the pokemon fans interested in the specific hashtags on 3 different social media. Finally we want to make a better interface for the query part in our database.

Key words: Pokemon, Game company, Gamer, Game content, SQL, Python

1 introduction

1.1 Background

Why is Pokemon?

Pokemon is also known as Pocket Monsters[b] in Japan, is a media franchise managed by The Pokemon Company, a Japanese consortium between Nintendo, Game Freak, and Creatures. The franchise copyright is shared by all three companies, but Nintendo is the sole owner of the trademark. The franchise was created by Satoshi Tajiri in 1995, and is centered on fictional creatures called "Pokemon", which humans, known as Pokemon Trainers, catch and train to battle each other for sport. The English slogan for the franchise is "Gotta Catch 'Em All". Works within the franchise are set in the Pokemon universe. And for us, Pokemon is a fantastic topic in our childhood because it builds a cartoon form in our minds. But there is a problem— where and how can we know more information about Pokemon? There are millions of fans of Pokemon in the world. If they can communicate or know the thought each other, it would be more convenient for Pokemon fans to

recognize some fresh information about Pokemon. So this is why my team wants to build this project.



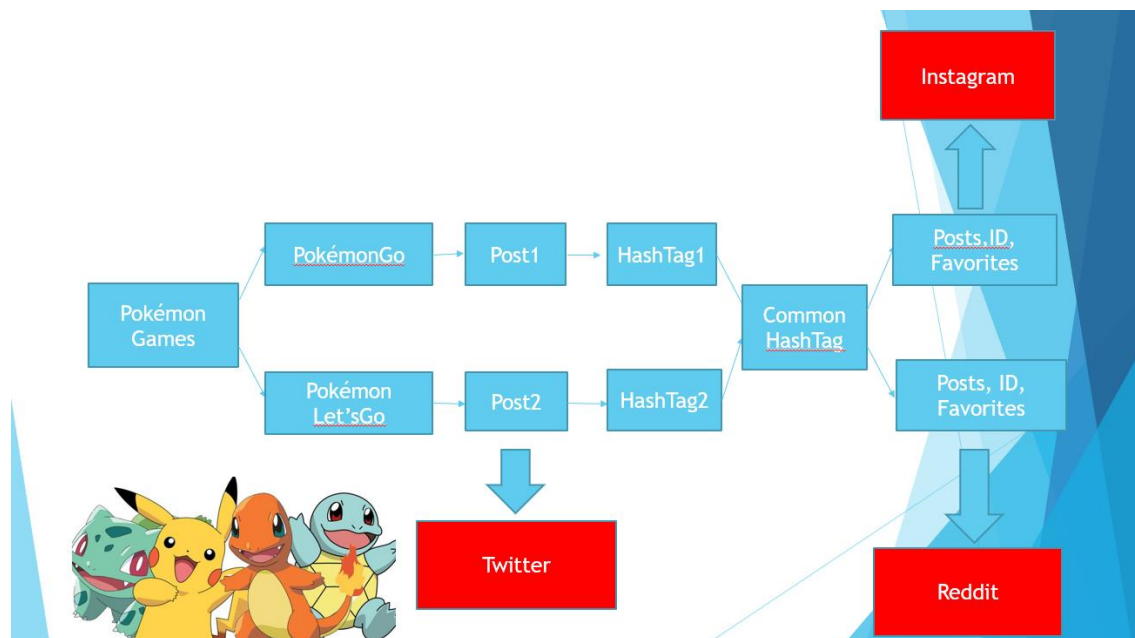
1.2 Introduction

The theme of our project is Pokemon domain. There are three roles in our Pokemon community(Game company, Game content, Gamer) and there are some relationships between each entity. For game company entity, it has attributes like name, Ins ID and post-on. For Game content entity, it has attributes like content name, game name and post-on. For Gamer entity, it has attributes like Reddit ID and post-on. For expanding the circle of communication, We use three social medias to scrape information from each Api. We get information about game company from Instagram and about game content from Twitter and about Gamer from Reddit. For the relationship, for example, a game company has may post in Instagram and each post has some hash-tags. Once we get the hash-tags from Instagram, We serve them as a game content entity and then we scrape all the posts that contain these hash-tags from Twitter. As for gamer, they would write some posts. If a post contains something about Pokemon games, in spite of social

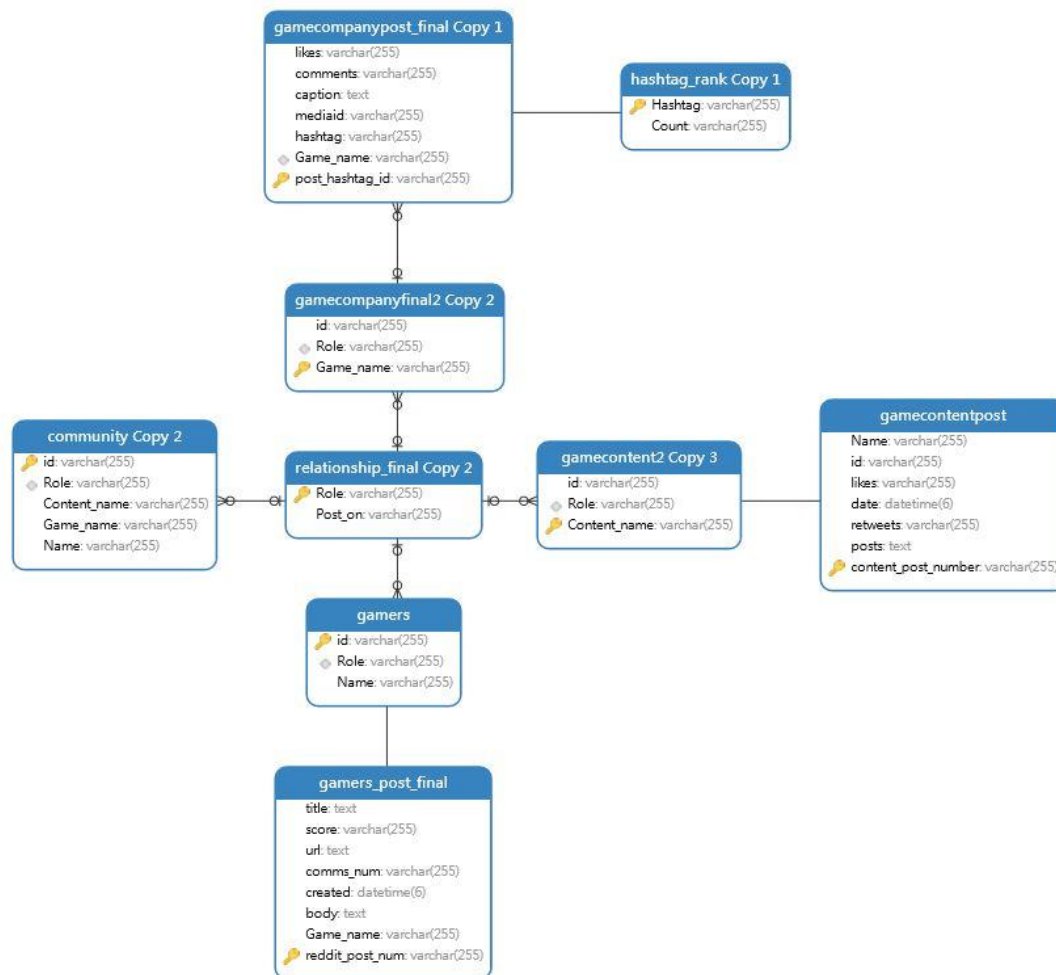
medias, the user who sends these posts is a gamer. So if they want to know more about Pokemon games, they can communicate each other.

1.3 Conceptual model

In this project, we want to build a Pokemon community database, there are 3 roles in this domain ---- Game company, Gamers and Game content. We choose 4 specific Pokemon game(PokemonGo, PokemonLet'sGo, Pokemon—SwordAndSheild and DetectivePikachu), then we looking for what Game company post and get hashtag table. After that , we choose 5 hashtag we interested (#Pikachu, #Eevee, #shiny, # PokemonGo, #swordandsheild) as Game contents and looking for what people talk about on Twitter. As for Gamer part, we use the 4 game name to looking for what gamers talk about on reddit.



1.4 ER-diagram



Entities: Community, Relationship, Game company, Game company Post, Game content, Game content post, Gamers, Gamers post, Hashtag.

Primary Key (PK): Community - id, Relationship-Role, Game company - Game_name, Game company Post - Post_Hashtag_id, Game content - Content-name, Game content post-Content_post_num, Gamers - id, Gamers post-reddit_post number, Hashtag- hashtag.

Foreign Key(FK): Community - Role(Relationship), Game company - Role(Relationship), Game company Post - Game name(Game company), Game content - Role(Relationship, Gamers - Role(Relationship), Gamers post, Hashtag.

Relationship:

One-to-one:

Gamers - Gamers post

One -to-many:

Relationship – Community/Game company/Game content/Gamers

Game company – Game company post;

Game content – Game content post;

Many-to-Many:

Game company post - Hashtag

1.5 Portfolio Link

All code, paper and output csv file is on GitHub.

The GitHub link of this project is :

https://github.com/CytusJuly/INFO6210_Pokemon_Final_Project

2 Collect data by using API

We acquire raw data from 3 different Social Media(Ins, Twitter & Reddit) by using 3 web API(Instaloader, Twitter API & PRAW).

2.1 Game company raw data

This part of data is written on Ins API. We output 4 dataframe for 4 different game and combine them into one table. For each game company , we collect data of caption, favorite, created and its hashtag.

Get Game Company Information Using Instaloader

We are looking for the pokemon company info on the instagram and finally chose 3 of them as our investigation objects: pokemon, pokemongoapp, detectivepikachu.

Import all the libraries needed

```
: import json
import oauth2 as oauth
import twitter
import urllib
import bs4
import numpy as np
import pandas as pd
import tweepy
import instaloader
import csv

L=instaloader.Instaloader()
profile=instaloader.Profile.from_username(L.context,'pokemon')
# Import the necessary libraries needed such like instaloader and tweepy.

posts=profile.get_posts()
name=[]
id=[]
likes=[]
comments=[]
caption=[]
hashtag=[]
mediaid=[]
i=0
for post in posts:
    if i <200:
        table_1=post.owner_username
        name.append(table_1)
        table_2=post.owner_id
        id.append(table_2)
        table_3=post.likes
        likes.append(table_3)
        table_4=post.comments
        comments.append(table_4)
        table_5=post.caption
        caption.append(table_5)
        table_6=post.caption_hashtags
        hashtag.append(table_6)
        table_7=post.mediaid
        mediaid.append(table_7)
        i +=1
# Get 200 recent posts on Twitter under the hashtag #pokemon.

user_info={'name':name,'id':id,'likes':likes,'comments':comments,'caption':caption,'hashtag':hashtag,'mediaid':mediaid}
df1=pd.DataFrame(user_info)
```

Now we get all the recent information posted under the account number: pokemon

df1							
	name	id	likes	comments	caption	hashtag	mediaid
0	pokemon	1711337410	32722	124	Coffee, clues, and #DetectivePikachu	[detectivepikachu]	2028634776084373191
1	pokemon	1711337410	33878	200	Transformation is an important tool in battle!	[tuesdaythoughts]	2028507254210276681
2	pokemon	1711337410	80382	957	Did you know there are only two known pure Wat...	[]	2027898995950596389
3	pokemon	1711337410	23754	91	The start of a new journey always begins with ...	[mondaymotivation]	2027826261308556109
4	pokemon	1711337410	61019	569	Eevee, stop. You're too cute! #DetectivePikachu	[detectivepikachu]	2027165846635379308
5	pokemon	1711337410	48210	610	About a quarter of dual-type Water-type Pokémo...	[]	2025061052806393720
6	pokemon	1711337410	69873	587	See how your favorite Pokémon measure up in th...	[]	2023605994382179189
7	pokemon	1711337410	32791	262	The TAG TEAM duos have one stellar Water-type ...	[pokemontcg]	2023521262898274310
8	pokemon	1711337410	69186	458	Honestly, this post is just an excuse to use t...	[detectivepikachu]	202279798594916003
9	pokemon	1711337410	164596	2566	Feeling cute, might evolve into one of eight d...	[]	2022090701356994465
10	pokemon	1711337410	68198	732	Nobody!\n\nRapidash:	[]	2021364674960015054

Now we use the concat instruction to get the total sheet.

```
df_gamecompany_raw = pd.concat([df1,df2,df3])
df_gamecompany_raw
# Concat all the 3 table into one.
```

	name	id	likes	comments	caption	mediaid	hashtag	Game_name
0	pokemon	1711337410	29343	117	Coffee, clues, and #DetectivePikachu	2028634776084373191	[detectivepikachu]	POKEMON
1	pokemon	1711337410	31036	178	Transformation is an important tool in battle!	2028507254210276681	[tuesdaythoughts]	POKEMON
2	pokemon	1711337410	79222	941	Did you know there are only two known pure Wat...	2027898995950596389	[]	POKEMON
3	pokemon	1711337410	23256	91	The start of a new journey always begins with ...	2027826261308556109	[mondaymotivation]	POKEMON
4	pokemon	1711337410	60643	565	Eevee, stop. You're too cute! #DetectivePikachu	2027165846635379308	[detectivepikachu]	POKEMON
5	pokemon	1711337410	48115	608	About a quarter of dual-type Water-type Pokémo...	2025061052806393720	[]	POKEMON
6	pokemon	1711337410	69809	588	See how your favorite Pokémon measure up in th...	2023605994382179189	[]	POKEMON
7	pokemon	1711337410	32736	261	The TAG TEAM duos have one stellar Water-type ...	2023521262898274310	[pokemontcg]	POKEMON

2.2 Hashtag and Game content

According to the previous table, we split the row with more than one hashtag into more rows. Then we count distinct hashtag and output a hashtag table. For every tags in this table, we choose 5 of them that we mostly interest as Game content in order to do the rest of project.

```
df1 = df1.drop('hashtag',axis=1).join(df1['hashtag'].str.split(', ',expand = True).stack().reset_index(level=1, drop=True).rename('# To divide the row which contained more than 1 hashtag into separate one(Normalize)').
```

df1

	name	id	likes	comments	caption	mediaid	hashtag
0	pokemon	1711337410	29343	117	Coffee, clues, and #DetectivePikachu	2028634776084373191	['detectivepikachu']
1	pokemon	1711337410	31036	178	Transformation is an important tool in battle!	2028507254210276681	['tuesdaythoughts']
2	pokemon	1711337410	79222	941	Did you know there are only two known pure Wat...	2027898995950596389	[]
3	pokemon	1711337410	23256	91	The start of a new journey always begins with ...	2027826261308556109	['mondaymotivation']
4	pokemon	1711337410	60643	565	Eevee, stop. You're too cute! #DetectivePikachu	2027165846635379308	['detectivepikachu']
5	pokemon	1711337410	48115	608	About a quarter of dual-type Water-type Pokémo...	2025061052806393720	[]
6	pokemon	1711337410	69809	588	See how your favorite Pokémon measure up in th...	2023605994382179189	[]
7	pokemon	1711337410	32736	261	The TAG TEAM duos have one stellar Water-type ...	2023521262898274310	['pokemontcg']
8	pokemon	1711337410	68955	458	Honestly, this post is just an excuse to use t...	2022797985984916003	['detectivepikachu']

Hashtag	Count
'belegendary']	10
'detectivepikachu']	10
'pokemoncenter']	11
'pokemonday']	18
'pokemongo']	5
'pokemonlivetweet']	6
'pokemonresolutions']	5
'pokemonultrasunmoon']	5
'teameevee'	10
'teampikachu'	10
'watchpokemon']	6
['detectivepikachu'	6
['detectivepikachu']	37
['detectivepikachumovie'	5
['eevee']	15
['firstdayofspring']	6
['gobattle']	5
['gosnapshot']	3
['happynewyear']	2
['mondaymotivation']	5
['mondaymotivation']	6
['nationalhatday']	2
['nationallazyday']	2
['pikachu']	2
['playpokemon']	20
['pokemoncenter']	13
['pokemonday']	45
['pokemondirect']	3
['pokemonduel']	27
['pokemongo']	10

id	Role	Content_name
2	GAMECONTENT	eevee
1	GAMECONTENT	pikachu
5	GAMECONTENT	pokemonletsgo
3	GAMECONTENT	shiny
4	GAMECONTENT	swordandshield

2.3 Game content raw data

This part of data is written on Twitter API. We output 5 dataframe for 5 different game content and combine them into one table. For each game content , we collect data of user id , name ,post, favorite, created and its hashtag.

Search all the key words on Twitter

Import all the libraries needed such like tweepy and numpy.

```
: import json
import oauth2 as oauth
import twitter
import urllib
import bs4
import numpy as np
import pandas as pd
import tweepy
# Import all the libraries needed.
```

Using auth to get access to the Twitter Developer platform.

```
: consumer_key = "kaXSMXPL5rJnf50AqXsHRaEr9"
consumer_secret = "GmSUPF3n6qQ15bh1cidhmtHn3BvFA0AlbtfnG2vGQKsT92aqXI"
access_token = "109935477811775360-9hsBFFfgVcBgI9eMqxq9g34oCZYTa"
access_token_secret = "AK24suXviBP0iGuVADydx5QTBMsttjq8yVFXj7ddsk0vd"
auth = tweepy.OAuthHandler(consumer_key, consumer_secret)
auth.set_access_token(access_token, access_token_secret)

api = tweepy.API(auth)
# Upload the auth to get access to the Developer Platform
```

Search all the information under the hashtag #pikachu.

```
df1 = pd.DataFrame(data=[tweet.user.screen_name for tweet in tweepy.Cursor(api.search,q='#pikachu').items(180)], columns=['Name'])
df1['id'] = np.array([tweet.id for tweet in tweepy.Cursor(api.search,q='#pikachu').items(180)])
df1['likes'] = np.array([tweet.favorite_count for tweet in tweepy.Cursor(api.search,q='#pikachu').items(180)])
df1['date'] = np.array([tweet.created_at for tweet in tweepy.Cursor(api.search,q='#pikachu').items(180)])
df1['retweets'] = np.array([tweet.retweet_count for tweet in tweepy.Cursor(api.search,q='#pikachu').items(180)])
df1['posts'] = np.array([tweet.text for tweet in tweepy.Cursor(api.search,q='#pikachu').items(180)])
df1
# To search all the contents such Like id, Likes, date under the hashtag #pikachu.
```

	Name	id	likes	date	retweets	posts
0	FPSheffield	1121425734516510720	0	2019-04-25 14:48:24	0	@FPSheffield @fpinternational NOW IN STORE!!!...
1	chocolat373love	1121425727088316416	0	2019-04-25 14:48:23	164	RT @oricon: 竹内涼真、ハリウッド俳優を英語でお出迎え 歌舞伎町で人生初カーベ ット...
2	chocolat373love	1121425707295338496	0	2019-04-25 14:48:18	37	RT @oricon: 【動画】竹内涼真、ハリウッド俳優を英語でお出迎え 歌舞伎町で人生 初カ...
3	TOONADOcom	1121425441774903304	0	2019-04-25 14:47:15	0	The Highly Anticipated DETECTIVE PIKACHU Gets ...
4	MieSorasanporoyo	1121424261376163841	0	2019-04-25 14:42:33	164	RT @oricon: 竹内涼真、ハリウッド俳優を英語でお出迎え 歌舞伎町で人生初カーベ ット...
5	min_zun	1121421258363686913	0	2019-04-25 14:30:37	2	RT @Buguwa: I'm trying a new technique :D\n\n#...

```
GameContentPostRaw = pd.concat([df1,df2,df3,df4,df5])
GameContentPostRaw
# Concat all the 5 tables into one.
```

	Name	id	likes	date	retweets	posts	Content_name
0	FPSheffield	1121425734516510720	0	2019-04-25 14:48:24	0	@FPSheffield @fpinternational NOW IN STORE!!!...	pikachu
1	chocolat373love	1121425727088316416	0	2019-04-25 14:48:23	164	RT @oricon: 竹内涼真、ハリウッド俳優を英語でお出迎え 歌舞伎町で人生初カーペット...	pikachu
2	chocolat373love	1121425707295338496	0	2019-04-25 14:48:18	37	RT @oricon: 【動画】竹内涼真、ハリウッド俳優を英語でお出迎え 歌舞伎町で人生初カ...	pikachu
3	TOONADOcom	1121425441774903304	0	2019-04-25 14:47:15	0	The Highly Anticipated DETECTIVE PIKACHU Gets ...	pikachu
4	MieSorasanporoy	1121424261376163841	0	2019-04-25 14:42:33	164	RT @oricon: 竹内涼真、ハリウッド俳優を英語でお出迎え 歌舞伎町で人生初カーペット...	pikachu
5	min_zun	1121421258363686913	0	2019-04-25 14:30:37	2	RT @Buguwa: I'm trying a new technique :D\n\n#...	pikachu
6	2019-04-25	...	RT @oricon: 竹内涼真、ハリウッド俳優を英語でお出迎え 歌舞伎町	...

2.4 Gamers raw data

This part of data is written on PRAW. We output 4 dataframe for 4 different game company and combine them into one table, in order to see what gamers talking about a specific game. For each game company, we collect data of user id , title ,score, url. Also, for each game name we acquire data from 5 different sort way (Hot, Top, New, Rising, Controversial) in order to acquire as more post as we can .

This code is try to acquire raw data about gamers playing different games under Pokemon domain. And we assume that whoever post on reddit should be a gamer from this specific game.

Also the normalization part of gamers' table are in the botton of this code.

Import library by using Reddit's API--- PRAW .

```
import praw
import pandas as pd
import datetime as dt
# Import all the libraries needed.
```

```
reddit = praw.Reddit(client_id='YUexek6uy1XnPQ', \
                     client_secret='DE6NwP7muXyP73LTvKPN9n0v1k8', \
                     user_agent='zhuzhuxia', \
                     username='jasonkk111', \
                     password='zyj92822')
# Choose all the items we want to get.
```

Log into subreddit "Pokemon Let's go"

Acquire Redditors posts under hot sort and output to dataframe

Acquire Redditors posts under top sort and output to dataframe

Acquire Redditors posts under new sort and output to dataframe

Acquire Redditors posts under rising sort and output to dataframe

Acquire Redditors posts under controversial sort and output to dataframe

Concat all 5 dataframe and generate table of POKEMON LET'S GO'S gamers.

```
: subreddit = reddit.subreddit('pokemonletsgo')

: hot_subreddit = subreddit.hot(limit=3000)
: top_subreddit = subreddit.top(limit=3000)
: rising_subreddit = subreddit.rising(limit=3000)
: controversial_subreddit = subreddit.controversial(limit=3000)
: new_subreddit = subreddit.new(limit=3000)

new_subreddit
controversial_subreddit
rising_subreddit
top_subreddit
hot_subreddit

: <praw.models.listing.generator.ListingGenerator at 0x26c939f9b38>
```

```
hot_dict = { "title":[],
             "score":[],
             "id":[],
             "name":[],
             "url":[],
             "comms_num": [],
             "created": [],
             "body":[]}

for submission in hot_subreddit:
    hot_dict["title"].append(submission.title)
    hot_dict["score"].append(submission.score)
    hot_dict["id"].append(submission.id)
    hot_dict["name"].append(submission.name)
    hot_dict["url"].append(submission.url)
    hot_dict["comms_num"].append(submission.num_comments)
    hot_dict["created"].append(submission.created)
    hot_dict["body"].append(submission.selftext)

tops_dict = { "title":[],
              "score":[],
              "id":[],
              "name":[],
              "url":[],
              "comms_num": [],
              "created": [],
              "body":[]}

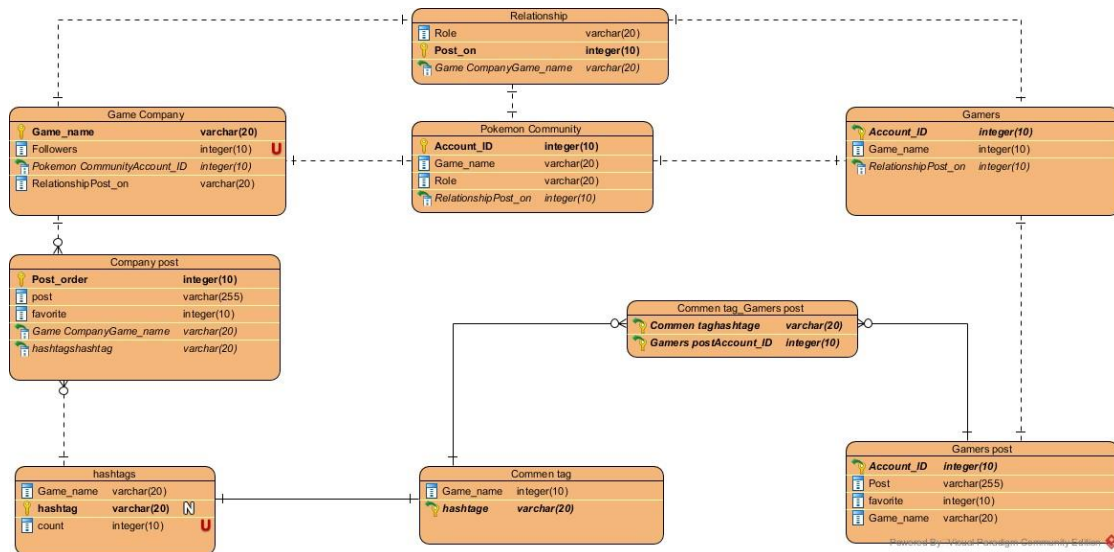
for submission in top_subreddit:
    tops_dict["title"].append(submission.title)
    tops_dict["score"].append(submission.score)
    tops_dict["id"].append(submission.id)
    tops_dict["name"].append(submission.name)
    tops_dict["url"].append(submission.url)
    tops_dict["comms_num"].append(submission.num_comments)
    tops_dict["created"].append(submission.created)
    tops_dict["body"].append(submission.selftext)
```

```
df_pokemonletsgo = pd.concat([controversial1_data,tops1_data,hot1_data,rising1_data,new1_data])
df_pokemonletsgo
```

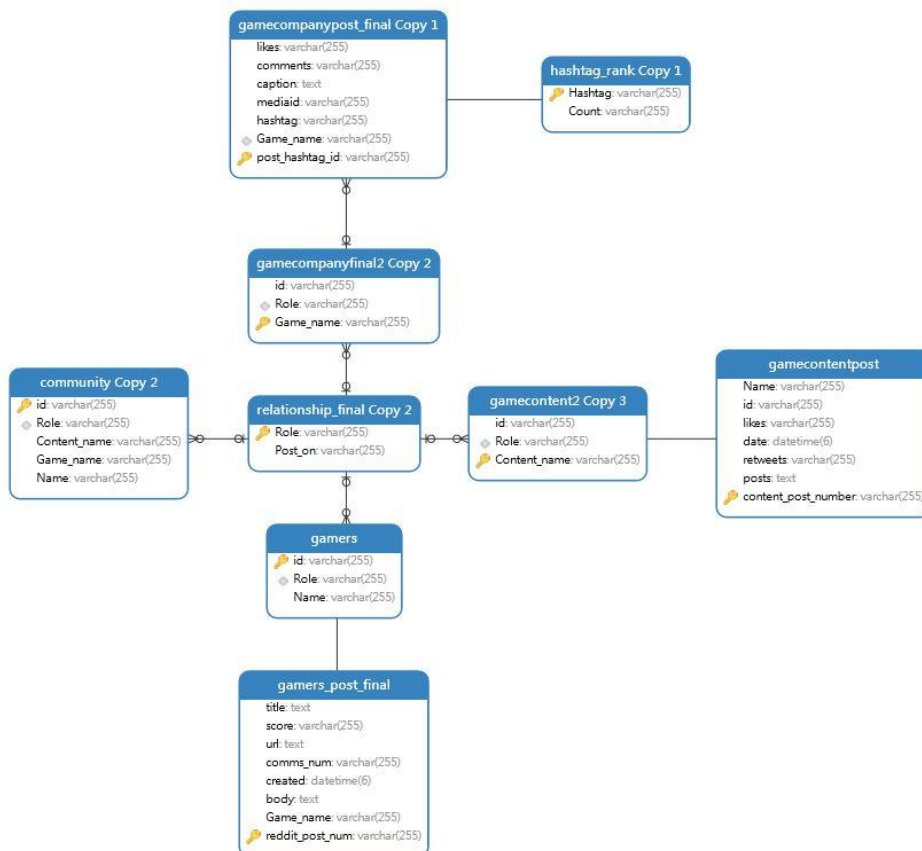
		title	score	id	name	url	comms_num	created	body
0		Am I the only one who prefers over world pokem...	6.0	9zu5ez	t3_9zu5ez	https://www.reddit.com/r/PokemonLetsGo/comment...	13.0	1.543023e+09	Idk i find it pretty cool especially when in t...
1		After a week of hunting I finally got it gimme...	0.0	b52wcy	t3_b52wcy	https://i.redd.it/k5fdtnvpe5o21.jpg	2.0	1.553469e+09	
2		The box art matches my tattoo	33.0	a00ee4	t3_a00ee4	https://i.redd.it/anwsepoveb021.jpg	15.0	1.543081e+09	
3		THERE IS A SECOND MASTER BALL	8.0	ablit9	t3_ablit9	https://i.redd.it/jbsottpodv721.jpg	12.0	1.546374e+09	
4		Chansey chaining really pays off (#41 in the c...	2.0	a6nake	t3_a6nake	https://i.redd.it/p32epkwzdl421.jpg	15.0	1.544946e+09	
5		Highest bonus multiplier?	2.0	a8e3im	t3_a8e3im	https://i.redd.it/11sbt6oyo521.jpg	1.0	1.545424e+09	
6		I'm totally Convinced: Total Caught Effects Sh...	0.0	ak4qb3	t3_ak4qb3	https://www.reddit.com/r/PokemonLetsGo/comment...	39.0	1.548535e+09	I realize this may be an unpopular opinion, bu...

3 Normalization

Before Normalization ER-D:



After Normalization ER-D:



This new database meets the requirement of 1NF, 2NF & 3NF:

First normal form (1NF)

- ✓ Each table has a primary key: minimal set of attributes which can uniquely identify a record
- ✓ The values in each column of a table are atomic (No multi-value attributes allowed).
- ✓ There are no repeating groups: two columns do not store similar information in the same table.

Second normal form (2NF)

- ✓ All requirements for 1st NF must be met.
- ✓ No partial dependencies.
- ✓ No calculated data

Third normal form (3NF)

- ✓ All requirements for 2nd NF must be met.
- ✓ Eliminate fields that do not directly depend on the primary key; that is no transitive dependencies.

Normalization part is contained in the ipynb files:

GameCompany_instagram.ipynb, Twitter2019.ipynb, reddit.ipynb.

```
df1['Game_name'] = 'POKEMON'
df2['Game_name'] = 'POKEMONGO'
df3['Game_name'] = 'DETECTIVE PIKACHU'
# Add a new column named Game_name and assign different values to it.
```

Now we use the concat instruction to get the total sheet.

```
df_gamecompany_raw = pd.concat([df1, df2, df3])
df_gamecompany_raw
# Concat all the 3 table into one.
```

	name	id	likes	comments	caption	mediaid	hashtag	Game_name
0	pokemon	1711337410	29343	117	Coffee, clues, and #DetectivePikachu	2028634776084373191	['detectivepikachu']	POKEMON
1	pokemon	1711337410	31036	178	Transformation is an important tool in battle...	2028507254210276681	['tuesdaythoughts']	POKEMON
2	pokemon	1711337410	79222	941	Did you know there are only two known pure Wat...	2027898995950596389	[]	POKEMON
3	pokemon	1711337410	23256	91	The start of a new journey always begins with ...	2027826261308556109	['mondaymotivation']	POKEMON
4	pokemon	1711337410	60643	565	Eevee, stop. You're too cute! #DetectivePikachu	2027165846635379308	['detectivepikachu']	POKEMON
5	pokemon	1711337410	48115	608	About a quarter of dual-type Water-type Pokémo...	2025061052806393720	[]	POKEMON
6	pokemon	1711337410	69809	588	See how your favorite Pokémon measure up in th...	2023605994382179189	[]	POKEMON
7	pokemon	1711337410	32736	261	The TAG TEAM duos have one stellar Water-type ...	2023521262898274310	['pokemontcg']	POKEMON

To delete the useless column 'name' and 'id'.

```
df_gamecompany_raw= df_gamecompany_raw.drop(['name', 'id'],axis=1)
df_gamecompany_raw
```

	likes	comments	caption	mediaid	hashtag	Game_name
0	29343	117	Coffee, clues, and #DetectivePikachu	2028634776084373191	['detectivepikachu']	POKEMON
1	31036	178	Transformation is an important tool in battle...	2028507254210276681	['tuesdaythoughts']	POKEMON
2	79222	941	Did you know there are only two known pure Wat...	2027898995950596389	[]	POKEMON
3	23256	91	The start of a new journey always begins with ...	2027826261308556109	['mondaymotivation']	POKEMON
4	60643	565	Eevee, stop. You're too cute! #DetectivePikachu	2027165846635379308	['detectivepikachu']	POKEMON
5	48115	608	About a quarter of dual-type Water-type Pokémo...	2025061052806393720	[]	POKEMON
6	69809	588	See how your favorite Pokémon measure up in th...	2023605994382179189	[]	POKEMON
7	32736	261	The TAG TEAM duos have one stellar Water-type ...	2023521262898274310	['pokemontcg']	POKEMON
8	68955	458	Honestly, this post is just an excuse to use t...	2022797985984916003	['detectivepikachu']	POKEMON
9	162907	2544	Feeling cute, might evolve into one of eight d...	2022090701356994465	[]	POKEMON
10	68087	729	Nobody:\r\n\r\nRapidash:	2021364674960015054	[]	POKEMON
11	51025	616	Old friends reunite when Brock and Misty visit...	2020646137274401351	['pokemontheseries']	POKEMON
12	21950	136	Celebrate spring with a bounty of log-in bonus...	2020623644757251709	['pokemonduel']	POKEMON
13	46553	1741	Fun fact: Water-type Pokémon make up 16% of th...	2019856478053044995	[]	POKEMON
14	147043	3654	Charizard knocking things over is a mood. 🦖\r\n...	2019163230761713938	['detectivepikachu']	POKEMON
15	27641	178	Not that we're counting... (We're definitely cou...	2019152970831489758	['detectivepikachu']	POKEMON
16	61958	412	Elementary, my dear (Detective) Pikachu! \r\nW...	2019146521577337017	['detectivepikachu']	POKEMON
17	44121	151	☀️ Put this ray of sunshine on your shelf, US ...	2018538621083942389	['pokemoncenter']	POKEMON
18	21004	121	April showers bring May flowers and Sprinkles?!	2018543723630026720	[]	POKEMON

```
GameCompany['Role']='GAMEDEVELOPER'
# Add a new row named Role and identify it as GAMEDEVELOPER.
```

GameCompany

	likes	comments	caption	mediaid	hashtag	Game_name	Role
0	29343	117	Coffee, clues, and #DetectivePikachu	2.028630e+18	['detectivepikachu']	POKEMON	GAMEDEVELOPER
1	31036	178	Transformation is an important tool in battle...	2.028510e+18	['tuesdaythoughts']	POKEMON	GAMEDEVELOPER
2	79222	941	Did you know there are only two known pure Wat...	2.027900e+18	[]	POKEMON	GAMEDEVELOPER
3	23256	91	The start of a new journey always begins with ...	2.027830e+18	['mondaymotivation']	POKEMON	GAMEDEVELOPER
4	60643	565	Eevee, stop. You're too cute! #DetectivePikachu	2.027170e+18	['detectivepikachu']	POKEMON	GAMEDEVELOPER
5	48115	608	About a quarter of dual-type Water-type Pokémo...	2.025060e+18	[]	POKEMON	GAMEDEVELOPER
6	69809	588	See how your favorite Pokémon measure up in th...	2.023610e+18	[]	POKEMON	GAMEDEVELOPER
7	32736	261	The TAG TEAM duos have one stellar Water-type ...	2.023520e+18	['pokemontcg']	POKEMON	GAMEDEVELOPER
8	68955	458	Honestly, this post is just an excuse to use t...	2.022800e+18	['detectivepikachu']	POKEMON	GAMEDEVELOPER
9	162907	2544	Feeling cute, might evolve into one of eight d...	2.022090e+18	[]	POKEMON	GAMEDEVELOPER
10	68087	729	Nobody:\r\n\r\n\r\n\r\nRapidash:	2.021360e+18	[]	POKEMON	GAMEDEVELOPER
11	51025	616	Old friends reunite when Brock and Misty visit...	2.020650e+18	['pokemontheseries']	POKEMON	GAMEDEVELOPER
12	21950	136	Celebrate spring with a bounty of log-in bonus...	2.020620e+18	['pokemonduel']	POKEMON	GAMEDEVELOPER

Normalization: Drop specific columns in order to meet 2NF and 3NF.

```
GAMERS= reddit_raw.drop(['title','score','url','comms_num','created','body','Game_name'],axis=1)
GAMERS
```

	id	name
0	1c9sok	t3_1c9sok
1	1mr9nt	t3_1mr9nt
2	5ad41g	t3_5ad41g
3	1c5w0o	t3_1c5w0o
4	493jab	t3_493jab
5	1eweur	t3_1eweur
6	1ccmhm	t3_1ccmhm
7	f7l14	t3_f7l14
8	17mock	t3_17mock
9	92u5al	t3_92u5al
10	1ow3as	t3_1ow3as
11	1hbin7	t3_1hbin7
12	1ioib1	t3_1ioib1
13	28ozi4	t3_28ozi4
14	4b4vr1	t3_4b4vr1

Normalization: create a new column named role for following normalization. And output final gamers table.

```
GAMERS['Role']='Gamers'
GAMERS
# We generate a new column named Role to identify the gamers, game_contents and game developer.
```

	id	name	Role
0	1c9sok	t3_1c9sok	Gamers
1	1mr9nt	t3_1mr9nt	Gamers
2	5ad41g	t3_5ad41g	Gamers
3	1c5w0o	t3_1c5w0o	Gamers
4	493jab	t3_493jab	Gamers
5	1eweur	t3_1eweur	Gamers
6	1ccmhm	t3_1ccmhm	Gamers
7	f7l14	t3_f7l14	Gamers
8	17mock	t3_17mock	Gamers
9	92u5al	t3_92u5al	Gamers
10	1ow3as	t3_1ow3as	Gamers
11	1hbin7	t3_1hbin7	Gamers

Normalization: drop columns for following normalization. And output final gamers' posts table.

```
GAMERS_POST= reddit_raw.drop(['id', 'name'],axis=1)
GAMERS_POST
# Drop the rows which are not needed: id, name.
```

	title	score	url	comms_num	created	body	Game_name
0	I bought a new shirt.	0.0	http://imgur.com/w6VCaXM	154.0	1.365867e+09		POKEMON
1	Here's the Ivysaur flower pot I just finished,...	540.0	http://imgur.com/fet4qQa	242.0	1.379649e+09		POKEMON
2	Keep this post at 666 upvotes and Lord helix w...	903.0	http://imgur.com/6lYdkOI	67.0	1.477930e+09		POKEMON
3	Would you play this if it exists?	600.0	http://imgur.com/TAtQVGP	593.0	1.365717e+09		POKEMON
4	I see your Pokemon collection and bring you mi...	43.0	http://i.imgur.com/noFE1iG.jpg	114.0	1.457203e+09		POKEMON
5	My new pokemon tattoo!!!!	23.0	http://imgur.com/7GAogUQ	25.0	1.369313e+09		POKEMON
6	Sinnoh starters according to my dad	17.0	http://imgur.com/sDJP5S6	31.0	1.365978e+09		POKEMON
7	A Pokemon Battle?	0.0	http://senorgif.memebase.com/2010/10/17/funny-...	14.0	1.295805e+09		POKEMON
8	HOLLY ARCEUS IT IS POSSIBLE!!!!!!!	0.0	http://i.imgur.com/VVDfhVX.jpg	13.0	1.359642e+09		POKEMON

4 Set up database by Navicat

In this part we input our csv.file in previous part into Navicat and set format and Null of each attribute and Primary Key and Foreign Key for each table. Establish relationship between tables.

The screenshot displays the Navicat Premium interface with two database windows open. The top window shows the 'gamecontentpost' table structure, and the bottom window shows the 'community' table structure.

gamecontentpost Table Structure:

Name	id	likes	date	retweets	posts	content_post_number
FPSheffield	112142573451651	0	2019-01-04 14:48:00.000000	0	@FPSheff	1
chocolat373love	112142572708831	0	2019-01-04 14:48:00.000000	164	RT @orice	2
chocolat373love	112142570729533	0	2019-01-04 14:48:00.000000	37	@orice	3
TOONADOcom	112142544177490	0	2019-01-04 14:47:00.000000	0	The Highl	4
MieSorasanporio	112142426137616	0	2019-01-04 14:42:00.000000	164	RT @orice	5
min_zun	112142125836368	0	2019-01-04 14:30:00.000000	2	RT @Bugi	6
r_k_y_261212	112142081210216	0	2019-01-04 14:28:00.000000	164	RT @orice	7
pemos2603	112142054110396	0	2019-01-04 14:27:00.000000	164	RT @orice	8
Tanatosup	112142050394243	0	2019-01-04 14:27:00.000000	164	RT @orice	9
maximumeast	112142041687289	0	2019-01-04 14:27:00.000000	164	RT @orice	10
aimiyarashi	112141931488665	0	2019-01-04 14:22:00.000000	164	RT @orice	11
eric_skavinsk	112141927074179	0	2019-01-04 14:22:00.000000	0	Projeto di	12
Rapun_jaki_1128	112141795130162	0	2019-01-04 14:17:00.000000	164	RT @orice	13
msdoxo	112141794129417	2	2019-01-04 14:17:00.000000	0	Pika pika	14
Takechan_050426	112141711786403	0	2019-01-04 14:14:00.000000	164	RT @orice	15
mrszuferino	112141638152896	0	2019-01-04 14:11:00.000000	0	@Warner	16
RINNAyuki	112141604454170	0	2019-01-04 14:09:00.000000	164	RT @orice	17
s2_satsuki_s2	112141564023598	0	2019-01-04 14:08:00.000000	164	RT @orice	18
zamas_megane	112141407537465	0	2019-01-04 14:02:00.000000	0	@meitani	19
shinju1999	112141405963765	0	2019-01-04 14:02:00.000000	164	RT @orice	20
nawamawara	112141377056619	0	2019-01-04 14:00:00.000000	164	RT @orice	21
Bentoandco	112141355262183	1	2019-01-04 14:00:00.000000	0	Can't wait	22
TheRealKungFoo	112141354175875	1	2019-01-04 13:59:00.000000	0	#deono i	23
potteriqatsi	112141262693064	0	2019-01-04 13:56:00.000000	2	RT @Bugi	24
imagicain	112141215469386	0	2019-01-04 13:54:00.000000	164	RT @orice	25
Bugwa	112141161982798	4	2019-01-04 13:52:00.000000	2	I'm trying	26
ProjectKroozie	112141126636141	0	2019-01-04 13:50:00.000000	0	In this sce	27
alex_ningin	112141093314074	0	2019-01-04 13:49:00.000000	2	RT @ZOT	28
pomoparade	112141081900511	0	2019-01-04 13:49:00.000000	164	RT @orice	29
1031_meat_623	112141069848181	0	2019-01-04 13:48:00.000000	164	RT @orice	30

community Table Structure:

id	Role	Content_name	Game_name	Name
9233				

The bottom window shows the 'community' table structure with columns: id, Role, Content_name, Game_name, and Name. The 'id' column is highlighted as the primary key.

5 Nine & Four question

These questions are from project requirement, we solve them by write queries on Navicat and output tables:

i. What are people saying about me (somebody)?

July | pokemon_project | 运行 | 停止 | 解释

```
1 SELECT
2   caption
3 FROM
4   gamecompanypost_final
5 WHERE
6   hashtag = "['pokemonday']"
7
8 /*What are people saying about me (somebody)? */
```

信息 | 结果 1 | 剖析 | 状态

caption
▶ < Think fast.< Be positive.< Follow your intuition.#TeamInstinct < Start recruiting, Team Instinct.
Trainers, Smeargle has arrived in Pokémon GO! 🐾 Looking to encounter Smeargle? If you spot Smeargle making
In celebration of the upcoming #PokemonDay, we're kicking off a new event, and we've invited some familiar Po
!! Attention !! Trainers will soon be able to change their Pokémon GO teams. < 🌟 It's time to start recruiting,
Put your research skills to work this weekend, Trainers! Clamperl will be featured in a three-hour Limited Researc
How are you celebrating #PokemonGOValentines? Show us by taking an AR photo with one of the many pink Po
#GOSnapshot 📸 update: Looking for an easier way to review your AR photos? With GO Snapshot, Trainers can e
Don't go Meltan hearts with these #PokemonGOValentines cards, Trainer! 🍷 🍷
Found a lot of Luvdisc, but need help finding some luv? 🍷 🍷 Here's some #PokemonGOValentines digital cards
Happy #ValentinesDay, Trainers! ❤️ Celebrate by sending a gift 📺, making a new friend, and catching some mor
It's time to spread the love, Trainers! #PokemonGOValentines is back! ❤️ Enjoy limited-time bonuses like Lure Mo
Attention, Trainers! A new feature called #GOSnapshot is coming to Pokémon GO! You can now take photos of A
Great news, Trainers! All day during your region's #PokemonGOCommunityDay, you can earn five Sinnoh Stones
For the first time, evolve Piloswine with a Sinnoh Stone during or up to one hour after Community Day to obtain
Since not every new Trainer has finished their quest to find Celebi yet, it's time to lend a hand. Make a new frien
We'd press A on this game inspired #DetectivePikachu art. Submit your art using #PikaCrew to be featured.
Mondays 🍷 #DetectivePikachu

ii. How viral are my posts?

```
1 SELECT
2   title,
3   score
4 FROM
5   gamers_post
6 WHERE
7   title LIKE "%shiny%"
8 ORDER BY
9   score DESC
10 /*How viral are my posts?*/
```

信息	结果 1	剖析	状态																																				
	<table><tr><th>title</th><th>score</th></tr><tr><td>► Shiny Gengar</td><td>21967</td></tr><tr><td>Shiny Gengar in the middle</td><td>15253</td></tr><tr><td>I carved Oddish a wooden s</td><td>14591</td></tr><tr><td>My brother and I got a shiny</td><td>13793</td></tr><tr><td>Shiny Drifloon!</td><td>11491</td></tr><tr><td>Shiny Drifloon!</td><td>11490</td></tr><tr><td>Shiny Drifloon!</td><td>11483</td></tr><tr><td>The reason why Shiny Mega</td><td>10975</td></tr><tr><td>First Gen 4 Shiny....</td><td>9422</td></tr><tr><td>Shiny Drifloon is my favorite</td><td>9053</td></tr><tr><td>My Shiny Team and I Just H</td><td>8970</td></tr><tr><td>Finally got a Shiny Gyarados</td><td>8782</td></tr><tr><td>My shiny Dragonair scarf! (x</td><td>8033</td></tr><tr><td>A couple months ago I com</td><td>7986</td></tr><tr><td>[OC] A pile of sleepy shiny s</td><td>7923</td></tr><tr><td>Trying to hatch a shiny stuffi</td><td>7633</td></tr><tr><td>A guy in my town is taking h</td><td>6876</td></tr></table>	title	score	► Shiny Gengar	21967	Shiny Gengar in the middle	15253	I carved Oddish a wooden s	14591	My brother and I got a shiny	13793	Shiny Drifloon!	11491	Shiny Drifloon!	11490	Shiny Drifloon!	11483	The reason why Shiny Mega	10975	First Gen 4 Shiny....	9422	Shiny Drifloon is my favorite	9053	My Shiny Team and I Just H	8970	Finally got a Shiny Gyarados	8782	My shiny Dragonair scarf! (x	8033	A couple months ago I com	7986	[OC] A pile of sleepy shiny s	7923	Trying to hatch a shiny stuffi	7633	A guy in my town is taking h	6876		
title	score																																						
► Shiny Gengar	21967																																						
Shiny Gengar in the middle	15253																																						
I carved Oddish a wooden s	14591																																						
My brother and I got a shiny	13793																																						
Shiny Drifloon!	11491																																						
Shiny Drifloon!	11490																																						
Shiny Drifloon!	11483																																						
The reason why Shiny Mega	10975																																						
First Gen 4 Shiny....	9422																																						
Shiny Drifloon is my favorite	9053																																						
My Shiny Team and I Just H	8970																																						
Finally got a Shiny Gyarados	8782																																						
My shiny Dragonair scarf! (x	8033																																						
A couple months ago I com	7986																																						
[OC] A pile of sleepy shiny s	7923																																						
Trying to hatch a shiny stuffi	7633																																						
A guy in my town is taking h	6876																																						

ii. What posts are likely to be interesting to me?

July	pokemon_project	运行	停止	解释
<pre>1 SELECT 2 id, 3 SUM(score) AS influence 4 FROM 5 gamers_post 6 WHERE 7 title LIKE "%pokemongo%" 8 GROUP BY 9 id 10 ORDER BY 11 influence DESC 12 /*How viral are my posts?*/</pre>				
信息	结果 1	剖析	状态	
	id	influence		
▶	4vv8kq	27281		
	4wyrq5	19456		
	4rrmpw	17925		
	4v18aq	16115		
	4vhbog	14226		
	4uung6	11226		
	505I15	10696		
	4nx50s	10301		
	4uny7t	9563		
	55i5cf	8448		
	807f9d	8033		
	4xkc6i	8015		
	4vhfwm	7235		
	5zpk2p	6894		
	4uwd7o	6742		
	bgzplm	6474		
	bgnm78	6194		

iv. What posts are like mine?

```
1 SELECT
2     title
3 FROM
4     gamers_post
5 WHERE
6     title LIKE "%shiny%"
7
8     /*What posts are like mine?*/
```

信息	结果 1	剖析	状态
----	------	----	----

title
LGBTQIA Pride Month Shiny Hunts
▶ Emulator Shiny
Did somebody say "Bad Shiny Sprite"?
Obligatory Shiny Post. Found this bad boy while button mas
Another one of the worst Shiny sprites
First shiny....god. damnit.
If were talking worst shiny sprites....
How I feel seeing all the shiny posts...
My First Shiny!
The only shiny I've ever caught. (First cake day post)
How I feel about all the "My first shiny" posts.
My first shiny :)
First shiny, karma, etc.
Surly I'll get a shiny
Shiny Mega Wailord
At least it isn't shiny...
Has anyone else bred a team for ORAS? Yes the Starly is shi

v. What users post like me?

```
1 SELECT DISTINCT
2   id
3 FROM
4   gamers_post
5 WHERE
6   title LIKE "%shiny%"
7
8   /*What posts are like mine?*/
```

信息	结果 1	剖析	状态
	id		
▶	8m3u23		
	1dr314		
	1ghrnv		
	1ahyp8		
	1ggylx		
	1cvzxq		
	1ghjao		
	16xjfy		
	15jf7k		
	14lr19		
	qs85u		
	1cz3v9		
	th2qw		
	3mflsm		
	1n8evv		
	u5o2t		
	2jbqma		

vi. Who should I be following?

```
1 SELECT DISTINCT
2   id
3 FROM
4   gamecontentpost
5 WHERE
6   posts LIKE "%letsgo%"
7 GROUP BY
8   id
9   /*Who should I be following?*/
```

信息	结果 1	剖析	状态
----	------	----	----

id

▶ 112113704534951

112113515843706

112140273095269

112066204833622

112053967475644

111814459568519

111814025833870

112140030464851

112139997058065

112139834146264

112139816353605

112139771534107

112139727766107

112139592840802

112139323883590

112139088897289

112138865673500

vii. What topics are trending in my domain?

```
1 |SELECT DISTINCT
2 |    hashtag,
3 |    COUNT( hashtag ) AS NUMBER
4 |FROM
5 |    gamecompanypost_final
6 |GROUP BY
7 |    hashtag
8 |ORDER BY
9 |    NUMBER DESC
10 | /*What topics are trending in my domain?*/
```

信息 结果 1 剖析 状态

hashtag	NUMBER
▶ []	101
['pokemonday']	45
['detectivepikachu']	37
['pokemonletsgo']	32
['pokemonduel']	27
['pokemontcg']	27
['pokemonholidaycheer']	22
['pokemongo']	22
['pokemonvalentine']	21
['playpokemon']	16
['pokemontheseries']	16
['pokemoncenter']	14
['pokemonswordshield']	12
['pokemontcg']	11
['pokemonday']	9
['pokemontheseries']	9
['playpokemon']	8

viii. What keywords/ hashtags should I add to my post?

```
1 SELECT DISTINCT
2   hashtag,
3   COUNT( hashtag ) AS NUMBER
4 FROM
5   gamecompanypost_final
6 WHERE
7   caption LIKE "%pokemongo%"
8 GROUP BY
9   hashtag
10 ORDER BY
11   NUMBER DESC
12
13 /*What keywords/ hashtags should I add to my post?*/
```

信息 结果 1 剖析 状态

hashtag	NUMBER
['pokemongo']	17
[]	15
['pokemonduel']	8
['pokemonday']	8
['pokemonletsgo']	8
['detectivepikachu']	5
['pokemonholidaycheer']	5
['playpokemon']	5
['pokemontcg']	4
['pokemoncenter']	4
['pokemonvalentine']	3
['pokemongo']	3
['pokemontheseries']	3
['pokemongocommunityday']	3
'pokemonletsgo']	2
'playpokemon']	2
['gobattle']	1

ix. Should I follow somebody back?

```
1 SELECT
2   id,
3   SUM( likes ) AS influence
4 FROM
5   gamecontentpost
6 WHERE
7   posts LIKE "%eevee%"
8 GROUP BY
9   id
10 ORDER BY
11   influence DESC
12 LIMIT 5
13
```

信息	结果 1	剖析	状态
----	------	----	----

id	influence
112137505186818	369
112115020174016	56
112137201003300	39
112116798902448	36
112130458011173	32

These for question we use both MySQL on Navicat and Python in jupyter notebook.

1. Domain tags (tags in your domain)

```
1 SELECT DISTINCT
2   hashtag
3 FROM
4   gamecompanypost_final
5   /*Domain tags (tags in your domain)*/
```

信息 结果 1 剖析 状态

hashtag
▶ ['detectivepikachu']
[]
['pokemongo']
['mondaymotivation']
'detectivepikachu']
['pokemontcg']
['pokemonresolutions']
['wednesdaywisdom']
'pokemonresolutions']
['pokemoncenter']
['pokemonduel']
['wednesdaywisdom']
['pokemontcg']
['nationalhatday']
['pokemontheseries']
['smashbrosultimate']
['happynewyear']

2.Synonyms (which tags are synonyms?)

```
: >>> from fuzzywuzzy import fuzz
>>> from fuzzywuzzy import process
```

C:\Users\cytus\Anaconda3\lib\site-packages\fuzzywuzzy\fuzz.py:11: UserWarning: Using slow pure-python SequenceMatcher. Install python-Levenshtein to remove this warning
warnings.warn('Using slow pure-python SequenceMatcher. Install python-Levenshtein to remove this warning')

```
: choices = ['pokemonday',
'detectivepikachu',
'pokemonletsgo',
'pokemontcg',
'pokemonduel',
'pokemontcg',
'pokemongo',
'pokemongo',
'pokemonholidaycheer',
'pokemonletsgogogo',
'playpokemon',
'teampikachu',
'smashbrosultimate',
'pikachu',
'happynewyear',
'nationallazyday'
]
process.extract('pikachu', choices, limit=5)
# Import all the hashtag we collected and to find the similarity of the word 'pikachu'.
```

```
: [('pikachu', 100),
('detectivepikachu', 90),
('teampikachu', 90),
('teampikachu', 90),
('detectivepikachu', 90)]
```

3.Mis-spellings (mis-spelled versions of words)

Cause the word 'pokemon' is a compound word which cannot be identified. In other words, there are no midspelling words in our database. So we pick some other words to do this part.

```
from spellchecker import SpellChecker

spell = SpellChecker()

# find those words that may be misspelled
misspelled = spell.unknown(['yes', 'encyclopeddia', 'dictionary', 'favorute', 'analysiz', 'woden', 'dog', 'forrrest'])

for word in misspelled:
    # Get the one `most likely` answer
    print(spell.correction(word))

    # Get a list of `likely` options
    print(spell.candidates(word))
```

```
# Cause the word 'pokemon' is a compound word which cannot be identified. In other words,
# there are no midspelling words in our database. So we pick some other words to do this part.
```

```
forrest
{'forrest'}
encyclopedia
{'encyclopedia'}
favorite
{'favorite'}
analysis
{'analysis'}
```

4.Semantic information (categories of tags)

```
1 SELECT DISTINCT
2   hashtag,
3   Game_name,
4   COUNT( hashtag ) AS number
5 FROM
6   gamecompanypost_final
7 GROUP BY
8   hashtag
9 ORDER BY
10  number DESC
11
12  /*IV. Semantic information (categories of tags)*/
```

信息	结果 1	剖析	状态
hashtag	Game_name	number	
▶ []	POKEMON	101	
['pokemonday']	POKEMONGO	45	
['detectivepikachu']	POKEMON	37	
['pokemonletsgo']	POKEMON	32	
['pokemonduel']	POKEMON	27	
['pokemontcg']	POKEMON	27	
['pokemonholidaycheer']	POKEMON	22	
['pokemongo']	POKEMON	22	
['pokemonvalentine']	POKEMONGO	21	
['playpokemon']	POKEMON	16	
['pokemontheseries']	POKEMON	16	
['pokemoncenter']	POKEMON	14	
['pokemonswordshield']	POKEMONGO	12	
['pokemontcg']	POKEMON	11	
'pokemonday']	POKEMONGO	9	
['pokemontheseries']	POKEMONGO	9	
'playpokemon']	POKEMON	8	

6 Use cases

We create 15 use cases on MySQL Navicat interface and output the results to prove our database's function.

6.1 Select all the contents which has the score higher than the AVG score from gamers_post.

```
1 SELECT * FROM gamers_post
2 WHERE score > (SELECT AVG(score) FROM gamers_post)
3 ORDER BY score
4
5 /*Select all the contents which has the score higher than the AVG score from gamers_post.*/
6
```

信息	结果 1	剖析	状态					
id	title	score	url	comms_num	created	body	Game_name	reddit_post_num
bdw5b7	Torchic Community Day Anr	2298	https://	216	5230-01-15	(Null)	POKEMON GO	7581
ap5cx9	My Pokeball decided to cry	2299	https://	186	0240-01-15	Guy lead	LETS GO PIKACHU/EEVEE	8886
bd3c5r	Almost all the evee	2302	https://	56	3536-01-15	This is my	POKEMON GO	7745
a6eciv	Hole in one with Ekins!	2305	https://	73	8402-01-15	(Null)	LETS GO PIKACHU/EEVEE	8885
a9rhj	Tecnology is incredible!	2348	https://	56	9312-01-15	(Null)	LETS GO PIKACHU/EEVEE	8884
bfcpha	This is why I bought this gar	2403	https://	100	3255-01-15	(Null)	LETS GO PIKACHU/EEVEE	8883
bdy734	How my girlfriend asked me	2500	https://	89	4435-01-15	(Null)	POKEMON	3691
bey714	When you want to level up a	2503	https://	29	6744-01-15	(Null)	POKEMON	3410
bey714	When you want to level up a	2504	https://	29	6744-01-15	(Null)	POKEMON	2558
bgz22w	Ryan Reynolds and his boi C	2560	https://	33	1350-01-15	(Null)	POKEMON	2005
bgz22w	Ryan Reynolds and his boi C	2570	https://	33	1350-01-15	(Null)	POKEMON	2864
bf0wyp	[OC] Just having a stroll with	2620	https://	32	6902-01-15	(Null)	POKEMON	3380
bf0wyp	[OC] Just having a stroll with	2621	https://	32	6902-01-15	(Null)	POKEMON	2548
bdw5b7	Torchic Community Day Anr	2298	https://	216	5230-01-15	(Null)	POKEMON	7581

6.2 To find the one who has fake favorites cause they have a high favorites but few comments.

```

1 SELECT
2   *,
3   ( likes - comments )
4 FROM
5   gamecompanypost_final
6 WHERE
7   game_name = 'POKEMON';
8 /*To find the one who has fake favorites cause they have a high favorites but few comments.*/

```

信息	结果 1	剖析	状态				
likes	comments	caption	mediaid	hashtag	Game_name	post_hashtag_id	(likes - comments)
29343	117	Coffee, clue	2028630000000000	['detectivepikachu']	POKEMON	1	29226
162907	2544	Feeling cute	2022090000000000	[]	POKEMON	10	160363
53639	375	☹️ Yes,	1972850000000000	['pokemongo']	POKEMON	100	53264
114222	543	#MondayM	1972200000000000	['mondaymotivation']	POKEMON	101	113679
114222	543	#MondayM	1972200000000000	['detectivepikachu']	POKEMON	102	113679
49459	185	Zeraora ap	1971290000000000	['pokemontcg']	POKEMON	103	49274
62823	962	The only thi	1970740000000000	[]	POKEMON	104	61861
38166	240	Stick with yc	1969210000000000	['pokemonresolutions']	POKEMON	105	37926
48882	439	#Wednesda	1968430000000000	['wednesdaywisdom']	POKEMON	106	48443
48882	439	#Wednesda	1968430000000000	['pokemonresolutions']	POKEMON	107	48443
112849	10935	Everyone's c	1967020000000000	[]	POKEMON	108	101914
111485	728	These cuties	1965600000000000	['pokemoncenter']	POKEMON	109	110757
68087	729	Nobody: Ra	2021360000000000	[]	POKEMON	11	67358
81475	1321	Which villai	1964200000000000	['pokemongo']	POKEMON	110	80154
42437	180	Pokémon Di	1964000000000000	['pokemonduel']	POKEMON	111	42257

6.3 To find the general post whose comments are more than the least one.

```
1 SELECT *
2 FROM gamecompanypost_final
3 WHERE comments > SOME
4 (SELECT comments
5  FROM gamecompanypost_final
6  WHERE game_name = 'POKEMON')
7
8 /*To find the general post whose comments are more than the least one.*/
```

likes	comments	caption	mediaid	hashtag	Game_name	post_hashtag_id
29343	117	Coffee, clue	202863000000000	['detectivepikachu']	POKEMON	1
162907	2544	Feeling cute	202209000000000	[]	POKEMON	10
53639	375	☺️ Yes,	197285000000000	['pokemongo']	POKEMON	100
114222	543	#MondayM	197220000000000	['mondaymotivation']	POKEMON	101
114222	543	#MondayM	197220000000000	['detectivepikachu']	POKEMON	102
49459	185	Zeraora app	197129000000000	['pokemontcg']	POKEMON	103
62823	962	The only thi	197074000000000	[]	POKEMON	104
38166	240	Stick with yc	196921000000000	['pokemonresolutions']	POKEMON	105
48882	439	#Wednesda	196843000000000	['wednesdaywisdom']	POKEMON	106
48882	439	#Wednesda	196843000000000	['pokemonresolutions']	POKEMON	107
112849	10935	Everyone's c	196702000000000	[]	POKEMON	108
111485	728	These cuties	196560000000000	['pokemoncenter']	POKEMON	109
68087	729	Nobody: Ra	202136000000000	[]	POKEMON	11
81475	1321	Which villai	196420000000000	['pokemongo']	POKEMON	110
42437	180	Pokémon D	196400000000000	['pokemonduel']	POKEMON	111
35183	261	Stay focus	196336000000000	['wednesdaywisdom']	POKEMON	112
101314	1259	🕵️ #Detecti	196128000000000	['detectivepikachu']	POKEMON	113

6.4 Write the subquery which has the highest comment from gamecompanypost.

```
1 SELECT *
2 FROM gamecompanypost_final
3 WHERE comments > ALL
4 (SELECT comments
5  FROM gamecompanypost_final
6  WHERE game_name = 'POKEMONGO')
7
8 /*Write the subquery which has the highest comment from gamecompanypost*/
```

likes	comments	caption	mediaid	hashtag	Game_name	post_hashtag_id
65095	999	A vast new	189149000000000	['pokemongo']	POKEMON	172

6.5 Select all the contents from the post sheet where the game_name likes PO.

```
1 SELECT *
2 FROM gamecompanypost_final
3 WHERE game_name like '%PO%'
4 GROUP BY game_name
5
6 /*Select all the contents from the post sheet where the game_name likes PO.*/
```

信息	结果 1	剖析	状态				
	likes	comments	caption	mediaid	hashtag	Game_name	post_hashtag_id
▶	29343	117	Coffee, clue	202863000000000	['detectivepikachu']	POKEMON	1
	18337	224	From its snc	202856000000000	['detectivepikachu']	POKEMONGO	227

6.6 Select the null body.

```
1 SELECT *
2 FROM gamers_post
3 WHERE body IS NULL;
4 /*Select the null body.*/
```

信息	结果 1	剖析	状态							
	id	title	score	url	comms_num	created	body	Game_name	reddit_post_num	
▶	1c9sok	I bought a new shirt.	0	http://i	154	8669-01-13	(Null)	POKEMON		1
	1mr9nt	Here's the Ivysaur flower po	540	http://i	242	6485-01-13	(Null)	POKEMON		2
	5ad41g	Keep this post at 666 upvote	903	http://i	67	9300-01-14	(Null)	POKEMON		3
	1c5w0o	Would you play this if it exis	600	http://i	593	7165-01-13	(Null)	POKEMON		4
	493jab	I see your Pokemon collectio	43	http://i	114	2025-01-14	(Null)	POKEMON		5
	1eweur	My new pokemon tattoo!!!!	23	http://i	25	3134-01-13	(Null)	POKEMON		6
	1ccmhm	Sinnoh starters according to	17	http://i	31	9778-01-13	(Null)	POKEMON		7
	f7l14	A Pokemon Battle?	0	http://s	14	8047-01-12	(Null)	POKEMON		8
	17mock	HOLLY ARCEUS IT IS POSSIB	0	http://i	13	6424-01-13	(Null)	POKEMON		9
	92u5al	TIL: The movement of a pok	3145	https://	251	8639-01-15	(Null)	POKEMON		10

6.7 Find the one whose name = POKEMON and score is less than 1000.

```
1 SELECT COUNT(comms_num)
2 FROM gamers_post
3 WHERE Game_name = 'POKEMON' AND score < 1000
4 /*Find the one whose name = POKEMON and score is less than 1000.*/
```

信息	结果 1	剖析	状态
COUNT(comms_num)			
▶	2590		

6.8 To calculate the MIN, MAX and AVG from the table gamers_post.

```
1 SELECT MIN(comms_num), MAX(comms_num), AVG(comms_num)
2 FROM gamers_post
3 /*To calculate the MIN, MAX and AVG from the table gamers_post.*/
```

信息	结果 1	剖析	状态
MIN(comms_num) MAX(comms_num) AVG(comms_num)			
▶	0	80553	111.5134

6.9 To find all the post where name is not null.

```
1 |SELECT DISTINCT(Game_name)
2 |FROM gamers_post
3 |WHERE Game_name IS NOT NULL
4 |/*To find all the post where name is not null.*/
```

信息	结果 1	剖析	状态
	Game_name		
▶	POKEMON		
	DETECTIVE PIKACHU		
	POKEMON GO		
	LETS GO PIKACHU/EEVEE		

6.10 To find the game name and comments number from the gamers_post where comments number is less than 100.

```
1 |SELECT Game_name, COUNT(comms_num)
2 |FROM gamers_post
3 |GROUP BY Game_name
4 |HAVING COUNT(comms_num)<100
5 |
6 |/*To find the game name and comments number from the gamers_post where comments number is less than 100.*/
```

信息	结果 1	剖析	状态
	Game_name	COUNT(comms_num)	
▶	POKEMON	3725	
	DETECTIVE PIKA	704	
	POKEMON GO	3452	
	(Null)	1	
	LETS GO PIKACH	3763	

6.11 Write a subquery where comms_num is larger than the smallest one with its name = POKEMON.

```

1 SELECT *
2 FROM gamers_post
3 WHERE comms_num > ANY
4 (SELECT comms_num
5  FROM gamers_post
6  WHERE Game_name = 'POKEMON')
7 /*Write a subquery where comms_num is larger than the smallest one with its name = POKEMON.*/
8
9

```

信息	结果 1	剖析	状态							
	id	title	score	url	comms_num	created	body	Game_name	reddit_u	
	1c9sok	I bought a new shirt.	0	http://imgur.com/w6VCaXM	154	8669-01-13	(Null)	POKEMON	1	
	1mr9nt	Here's the Ivysaur flower po	540	http://imgur.com/fet4qQa	242	6485-01-13	(Null)	POKEMON	2	
	5ad41g	Keep this post at 666 upvote	903	http://imgur.com/6lYdkOI	67	9300-01-14	(Null)	POKEMON	3	
	1c5w0o	Would you play this if it exis	600	http://imgur.com/TAtQVGP	593	7165-01-13	(Null)	POKEMON	4	
	493jab	I see your Pokemon collectio	43	http://i.imgur.com/noFE1G.jpg	114	2025-01-14	(Null)	POKEMON	5	
	1eweur	My new pokemon tattoo!!!!	23	http://imgur.com/7GAogUQ	25	3134-01-13	(Null)	POKEMON	6	
	1ccmhm	Sinnoh starters according to	17	http://imgur.com/sDJP5S6	31	9778-01-13	(Null)	POKEMON	7	
	f7l14	A Pokemon Battle?	0	http://senorgif.memebase.com/2010/10/17/fun	14	8047-01-12	(Null)	POKEMON	8	
	17mock	HOLLY ARCEUS IT IS POSSIB	0	http://i.imgur.com/VVDfhVX.jpg	13	6424-01-13	(Null)	POKEMON	9	
	92u5al	TIL: The movement of a pok	3145	https://v.redd.it/1a4jndbjhvc11	251	8639-01-15	(Null)	POKEMON	10	
	1ow3as	Regigas is in XY!	5	http://imgur.com/CM6QMPk	28	3457-01-13	(Null)	POKEMON	11	
	1hbin7	Good old days, I think we ha	254	http://i.imgur.com/EBeUZN3.jpg	53	5278-01-13	(Null)	POKEMON	12	

6.12 To find all the url where comms_num between 1 and 40.

```
1 |SELECT url
2 |FROM gamers_post
3 |WHERE comms_num BETWEEN 1 AND 40
4 |
5 |/*To find all the url where comms_num between 1 and 40.*/
```

信息 结果 1 剖析 状态

url

► <http://imgur.com/7GAogUQ>
<http://imgur.com/sDJP5S6>
<http://senorgif.memebase.com/2010/10/17/fun>
<http://i.imgur.com/VVDfhVX.jpg>
<http://imgur.com/CM6QMPk>
<http://i.imgur.com/M9gZeyJ.jpg>
<http://i.imgur.com/Yhynrhr.jpg>
<http://qkme.me/3tlquy>
<http://i.imgur.com/tNiVWzC.jpg>
<https://www.reddit.com/r/pokemon/comments>
<http://www.quickmeme.com/meme/3t15hm/>
<http://www.quickmeme.com/meme/353qv1/>
<http://cosplayspot.com/raichu-cosplay/pikachu>
<http://imgur.com/cJFO4gO>

6.13 Get all the hashtag where count is between 10 and 20.

```
1 SELECT COUNT(hashtag)
2 FROM hashtag_rank
3 WHERE Count BETWEEN 10 AND 20
4
5 /*Get all the hashtag where count is between 10 and 20.*/
```

信息	结果 1	剖析	状态
COUNT(hashtag)			
▶	20		

6.14 Using OR instruction to find all the contents whose name = POKEMON and POKEMONGO.

```
1 SELECT *
2 FROM gamers_post
3 WHERE Game_name = 'POKEMON' OR Game_name = 'POKEMON GO'
4
5 /*Using OR instruction to find all the contents whose name = POKEMON and POKEMONGO.*/
```

信息	结果 1	剖析	状态					
id	title	score	url	comms_num	created	body	Game_name	reddit_u
▶ 1c9sok	I bought a new shirt.	0	http://imgur.com/w6VCaXM	154	8669-01-13	(Null)	POKEMON	1
1mr9nt	Here's the Ivysaur flower po	540	http://imgur.com/fet4qQa	242	6485-01-13	(Null)	POKEMON	2
5ad41g	Keep this post at 666 upvote	903	http://imgur.com/6lYdkOI	67	9300-01-14	(Null)	POKEMON	3
1c5w0o	Would you play this if it exis	600	http://imgur.com/TAtQVGP	593	7165-01-13	(Null)	POKEMON	4
493jab	I see your Pokemon collecti	43	http://i.imgur.com/noFE1iG.jpg	114	2025-01-14	(Null)	POKEMON	5
1eweur	My new pokemon tattoo!!!!	23	http://imgur.com/7GAogUQ	25	3134-01-13	(Null)	POKEMON	6
1ccmhm	Sinnoh starters according to	17	http://imgur.com/sDJP5S6	31	9778-01-13	(Null)	POKEMON	7
f7l14	A Pokemon Battle?	0	http://senorgif.memebase.com/2010/10/17/fun	14	8047-01-12	(Null)	POKEMON	8
17mock	HOLLY ARCEUS IT IS POSSIB	0	http://i.imgur.com/VVDfhVX.jpg	13	6424-01-13	(Null)	POKEMON	9
92u5al	TIL: The movement of a pok	3145	https://v.redd.it/1a4jndbjhvc11	251	8639-01-15	(Null)	POKEMON	10
1ow3as	Regigas is in XY!	5	http://imgur.com/CM6QMPk	28	3457-01-13	(Null)	POKEMON	11
1hbin7	Good old days, I think we ha	254	http://i.imgur.com/EBeUZN3.jpg	53	5278-01-13	(Null)	POKEMON	12
1ioib1	I really like Espeon	7	http://imgur.com/KYgzLy4	66	2974-01-13	(Null)	POKEMON	13
20axi4	I haven't seen this posted b	15	http://i.imgur.com/M0s7eul.jpg	10	2176-02-14	(Null)	POKEMON	14

6.15 To find all the post whose name is not Game_name.

```
1 SELECT *
2 FROM gamers_post
3 WHERE NOT Game_name = 'POKEMON'
4
5 /*To find all the post whose name is not Game_name.*/
```

信息	结果 1	剖析	状态							
id	title	score	url	comms_num	created	body	Game_name	reddit_u		
azcoo0	Anyone else still a little won	1	https://www.reddit.com/r/detectivepikachu/cor	8	1964-01-15	Just by lo	DETECTIVE PIKA	3726		
bauoac	A closer look at the Detectiv	0	https://youtu.be/w4fZdbx0GrU	0	7347-01-15	(Null)	DETECTIVE PIKA	3727		
baud8y	DETECTIVE PIKACHU IN REA	2	https://youtu.be/33UnnVs164E	1	7331-01-15	(Null)	DETECTIVE PIKA	3728		
beqf3r	WATCH ME OPEN A HUGE \	5	https://youtu.be/qjCp4Fi2Eik	0	6198-01-15	(Null)	DETECTIVE PIKA	3729		
avsudl	When you see that your frier	11	https://i.redd.it/c42w6mj93cj21.jpg	0	3704-01-15	(Null)	DETECTIVE PIKA	3730		
axcedh	I have a MAJOR problem wi	11	https://youtu.be/GLWsQF2Rxqc	6	7336-01-15	(Null)	DETECTIVE PIKA	3731		
aremxj	The detective pikachu tease	0	https://youtu.be/rdSIA59rmgl	0	3608-01-15	(Null)	DETECTIVE PIKA	3732		
b7t2hk	Make it until morning - Dete	4	https://youtu.be/ev8XYlv8QIk	1	0679-01-15	(Null)	DETECTIVE PIKA	3733		
bhcq05	Im Screaming The Entire De	0	https://www.reddit.com/r/detectivepikachu/cor	3	2202-01-15	Hi! I am S	DETECTIVE PIKA	3734		
b9ip5i	HEY TIM!!! I DONT FLOAT!!!!	0	https://www.youtube.com/watch?v=AfffmHoxv	0	4135-01-15	(Null)	DETECTIVE PIKA	3735		
avev34	Doea mewtwo have a tail, o	8	https://i.redd.it/r152si6it4j21.jpg	7	2823-01-15	(Null)	DETECTIVE PIKA	3736		
9x47nv	Did it have to be Ryan Revn	20	https://www.reddit.com/r/detectivepikachu/cor	11	2301-01-15	It is great	DETECTIVE PIKA	3737		

7 View

We create 6 view based on MySQL Navicat according to previous queries, and output the results.

7.1 viewed based on case 3

```
1 SELECT
2   `gamecompanypost_final`.`likes` AS `likes`,
3   `gamecompanypost_final`.`comments` AS `comments`,
4   `gamecompanypost_final`.`caption` AS `caption`,
5   `gamecompanypost_final`.`mediaid` AS `mediaid`,
6   `gamecompanypost_final`.`hashtag` AS `hashtag`,
7   `gamecompanypost_final`.`Game_name` AS `Game_name`,
8   `gamecompanypost_final`.`post_hashtag_id` AS `post_hashtag_id`
9 FROM
10  `gamecompanypost_final`
11 WHERE
12  `gamecompanypost_final`.`comments` > ANY ( SELECT `gamecompanypost_final`.`comments` FROM `gamecompanypost_final` WHERE ( `gamecompanypost_final`.`Game_name` = 'POKEMON' ) )
```

likes	comments	caption	mediaid	hashtag	Game_name	post_hashtag_id
29343	117	Coffee, clue	2028630000000000	['detectivepikachu']	POKEMON	1
162907	2544	Feeling cute	2022090000000000	[]	POKEMON	10
53639	375	☺️ Yes,	1972850000000000	['pokemongo']	POKEMON	100
114222	543	#MondayM	1972200000000000	['mondaymotivation']	POKEMON	101
114222	543	#MondayM	1972200000000000	['detectivepikachu']	POKEMON	102
49459	185	Zeraora app	1971290000000000	['pokemontcg']	POKEMON	103
62823	962	The only thi	1970740000000000	[]	POKEMON	104
38166	240	Stick with yc	1969210000000000	['pokemonresolutions']	POKEMON	105
48882	439	#Wednesda	1968430000000000	['wednesdaywisdom']	POKEMON	106
48882	439	#Wednesda	1968430000000000	['pokemonresolutions']	POKEMON	107
112849	10935	Everyone's c	1967020000000000	[]	POKEMON	108
111485	728	These cuties	1965600000000000	['pokemoncenter']	POKEMON	109
68087	729	Nobody: Ra	2021360000000000	[]	POKEMON	11
81475	1321	Which villai	1964200000000000	['pokemongo']	POKEMON	110
42437	180	Pokémon D	1964000000000000	['pokemonduel']	POKEMON	111
35183	261	Stay focus	1963360000000000	['wednesdaywisdom']	POKEMON	112

7.2 view based on case 9

定义	高级	SQL 预览
1	SELECT DISTINCT	
2	`gamers_post`.`Game_name` AS `Game_name`	
3	FROM	
4	`gamers_post`	
5	WHERE	
6	(`gamers_post`.`Game_name` IS NOT NULL)	

信息	结果 1
Game_name	
▶ POKEMON	
DETECTIVE PIKA	
POKEMON GO	
LETS GO PIKACH	

7.3 View based on case 7

1	SELECT
2	count(`gamers_post`.`comms_num`) AS `COUNT(comms_num)`
3	FROM
4	`gamers_post`
5	WHERE
6	((`gamers_post`.`Game_name` = 'POKEMON') AND (`gamers_post`.`score` < 1000))

信息	结果 1
COUNT(comms_num)	
▶	2590

7.4 View based on case 6

```
1  SELECT
2    `gamers_post`.`id` AS `id`,
3    `gamers_post`.`title` AS `title`,
4    `gamers_post`.`score` AS `score`,
5    `gamers_post`.`url` AS `url`,
6    `gamers_post`.`comms_num` AS `comms_num`,
7    `gamers_post`.`created` AS `created`,
8    `gamers_post`.`body` AS `body`,
9    `gamers_post`.`Game_name` AS `Game_name`,
10   `gamers_post`.`reddit_post_num` AS `reddit_post_num`
11  FROM
12    `gamers_post`
13  WHERE
14    isnull( `gamers_post`.`body` )
```

信息	结果 1								
	id	title	score	url	comms_num	created	body	Game_name	reddit_post_num
▶	1c9sok	I bought	0	http://i	154	8669-01-13 00:00:00.00(Null)	POKEMON		1
	1mr9nt	Here's th	540	http://i	242	6485-01-13 16:00:00.00(Null)	POKEMON		2
	5ad41g	Keep thi	903	http://i	67	9300-01-14 00:00:00.00(Null)	POKEMON		3
	1c5w0o	Would y	600	http://i	593	7165-01-13 00:00:00.00(Null)	POKEMON		4
	493jab	I see you	43	http://i	114	2025-01-14 00:00:00.00(Null)	POKEMON		5
	1eweur	My new	23	http://i	25	3134-01-13 00:00:00.00(Null)	POKEMON		6
	1ccmhm	Sinnoh s	17	http://i	31	9778-01-13 00:00:00.00(Null)	POKEMON		7
	f7l14	A Poken	0	http://s	14	8047-01-12 00:00:00.00(Null)	POKEMON		8
	17mock	HOLLY A	0	http://i	13	6424-01-13 00:00:00.00(Null)	POKEMON		9
	92u5al	TIL: The	3145	https://	251	8639-01-15 00:00:00.00(Null)	POKEMON		10
	1ow3as	Regigas	5	http://i	28	3457-01-13 16:00:00.00(Null)	POKEMON		11
	1hbin7	Good ol	254	http://i	53	5278-01-13 00:00:00.00(Null)	POKEMON		12
	1ioib1	I really li	7	http://i	66	2974-01-13 01:00:00.00(Null)	POKEMON		13
	28ozi4	I haven't	15	http://i	18	3176-03-14 00:00:00.00(Null)	POKEMON		14
	4b4vr1	The girlf	17	http://i	30	4231-01-14 00:00:00.00(Null)	POKEMON		15

7.5 View based on case

```
1 SELECT
2   `gamers_post`.`id` AS `id`,
3   `gamers_post`.`title` AS `title`,
4   `gamers_post`.`score` AS `score`,
5   `gamers_post`.`url` AS `url`,
6   `gamers_post`.`comms_num` AS `comms_num`,
7   `gamers_post`.`created` AS `created`,
8   `gamers_post`.`body` AS `body`,
9   `gamers_post`.`Game_name` AS `Game_name`,
10  `gamers_post`.`reddit_post_num` AS `reddit_post_num`
11 FROM
12   `gamers_post`
13 WHERE
14   ( `gamers_post`.`score` > ( SELECT avg( `gamers_post`.`score` ) FROM `gamers_post` ) )
15 ORDER BY
16   `gamers_post`.`score`
```

信息	结果 1								
	id	title	score	url	comms_num	created	body	Game_name	reddit_post_num
▶	bdw5b7	Torchic	2298	https://	216	5230-01-15 00:00:00.000000	(Null)	POKEMON GO	7581
	ap5cx9	My Poke	2299	https://	186	0240-01-15 22:00:00.000000	Guy leadi	LETS GO PIKACH	8886
	bd3c5r	Almost a	2302	https://	56	3536-01-15 00:00:00.000000	This is my	POKEMON GO	7745
	a6eclv	Hole in	2305	https://	73	8402-01-15 00:00:00.000000	(Null)	LETS GO PIKACH	8885
	a9rhhj	Tecnolo	2348	https://	56	9312-01-15 00:00:00.000000	(Null)	LETS GO PIKACH	8884
	bfcpha	This is w	2403	https://	100	3255-01-15 00:00:00.000000	(Null)	LETS GO PIKACH	8883
	bdy734	How my	2500	https://	89	4435-01-15 23:00:00.000000	(Null)	POKEMON	3691
	bey714	When yc	2503	https://	29	6744-01-15 00:00:00.000000	(Null)	POKEMON	3410
	bey714	When yc	2504	https://	29	6744-01-15 00:00:00.000000	(Null)	POKEMON	2558
	bgz22w	Ryan Re	2560	https://	33	1350-01-15 00:00:00.000000	(Null)	POKEMON	2005
	bgz22w	Ryan Re	2570	https://	33	1350-01-15 00:00:00.000000	(Null)	POKEMON	2864
	bf0wyp	[OC] Jus	2620	https://	32	6902-01-15 00:00:00.000000	(Null)	POKEMON	3380
	bf0wyp	[OC] Jus	2621	https://	32	6902-01-15 00:00:00.000000	(Null)	POKEMON	2548
	3mflsm	Surfiv I'll	2728	http://i	786	2463-01-14 00:00:00.000000	(Null)	POKEMON	863

7.6 View based on case 15

```
1 SELECT
2   `gamers_post`.`id` AS `id`,
3   `gamers_post`.`title` AS `title`,
4   `gamers_post`.`score` AS `score`,
5   `gamers_post`.`url` AS `url`,
6   `gamers_post`.`comms_num` AS `comms_num`,
7   `gamers_post`.`created` AS `created`,
8   `gamers_post`.`body` AS `body`,
9   `gamers_post`.`Game_name` AS `Game_name`,
10  `gamers_post`.`reddit_post_num` AS `reddit_post_num`
11 FROM
12   `gamers_post`
13 WHERE
14   ( `gamers_post`.`Game_name` <> 'POKEMON' )
```

信息	结果 1								
	id	title	score	url	comms_num	created	body	Game_name	reddit_post_num
▶	azcoo0	Anyone	1	https://	8	1964-01-15 00:00:00.000	Just by lo	DETECTIVE PIKA	3726
	bauoac	A closer	0	https://	0	7347-01-15 00:00:00.000	(Null)	DETECTIVE PIKA	3727
	baud8y	DETECTI	2	https://	1	7331-01-15 00:00:00.000	(Null)	DETECTIVE PIKA	3728
	beqf3r	WATCH	5	https://	0	6198-01-15 00:00:00.000	(Null)	DETECTIVE PIKA	3729
	avsudl	When yc	11	https://	0	3704-01-15 00:00:00.000	(Null)	DETECTIVE PIKA	3730
	axcedh	I have a	11	https://	6	7336-01-15 00:00:00.000	(Null)	DETECTIVE PIKA	3731
	aremxj	The dete	0	https://	0	3608-01-15 00:00:00.000	(Null)	DETECTIVE PIKA	3732
	b7t2hk	Make it	4	https://	1	0679-01-15 00:00:00.000	(Null)	DETECTIVE PIKA	3733
	bhcq05	Im Screa	0	https://	3	2202-01-15 00:00:00.000	Hi! I am S	DETECTIVE PIKA	3734
	b9ip5i	HEY TIM	0	https://	0	4135-01-15 00:00:00.000	(Null)	DETECTIVE PIKA	3735
	avev34	Doea m	8	https://	7	2823-01-15 00:00:00.000	(Null)	DETECTIVE PIKA	3736
	9x47qy	Did it ha	20	https://	11	2301-01-15 09:00:00.000	It is great	DETECTIVE PIKA	3737
	bb0pdf	Detectiv	5	https://	1	7659-01-15 00:00:00.000	Do you th	DETECTIVE PIKA	3738
	b0x0xf	What are	5	https://	5	4752-01-15 07:00:00.000	Just said	DETECTIVE PIKA	3739

8 Function

We create 6 functions on MySQL Navicat, and prove them useful.

8.1 Using the primary key to search the related id

```
1 CREATE DEFINER=`root`@`localhost` FUNCTION `get_id`(i int) RETURNS varchar(255) CHARSET utf8mb4
2     READS SQL DATA
3     DETERMINISTIC
4 BEGIN
5     #Routine body goes here...
6     declare result varchar(255);
7     select id into result from gamers_post
8     where reddit_post_num=i;
9     RETURN result;
10 END
```

输入参数

i 1

☐ 原始数据模:

确定 取消

`get_id`(@i)

► 1c9sok

8.2 Using the primary key to search the related name and hashtag

```
1 CREATE DEFINER=`root`@`localhost` FUNCTION `get_hashtag`(i int) RETURNS varchar(255) CHARSET utf8mb4
2     READS SQL DATA
3     DETERMINISTIC
4 BEGIN
5     #Routine body goes here...
6     declare result varchar(255);
7     select hashtag into result from gamecompanypost_final
8     where post_hashtag_id = i;
9     RETURN result;
10 END
```

输入参数

i 15

☐ 原始数据模:

确定 取消

`get_hashtag`(@i)

▶ ['detectivepikachu']

8.3 Get the hashtag rank

```
1 CREATE DEFINER=`root`@`localhost` PROCEDURE `F1`()
2 BEGIN
3     SELECT * FROM hashtag_rank;#Routine body goes here...
4
5 END
```

Hashtag	Count
'belegendary'	10
'detectivepikachu'	10
'pokemoncenter'	11
'pokemonday'	18
'pokemongo'	5
'pokemonlivetweet'	6
'pokemonresolutions'	5
'pokemonultrasunmoo	5
'teameevee'	10
'teampikachu'	10
'watchpokemon'	6
['detectivepikachu'	6
['detectivepikachu'	37
['detectivepikachumov	5
['eevee']	15
['firstdayofspring']	6
['gobattle'	5
['gosnapshot']	3
['happynewyear']	2
['mondaymotivation'	5
['mondaymotivation']	6
['nationalhatday']	2
['nationallazyday']	2
['pikachu']	2
['playpokemon']	20

8.4 Get the hashtag rank

```
1 CREATE DEFINER=`root`@`localhost` PROCEDURE `F2`()
2 BEGIN
3     SELECT DISTINCT id
4     FROM gamers_post
5     WHERE title LIKE "%shiny%"
6 ;
7 END
```

id
8m3u23
1dr314
1ghrnv
1ahyp8
1ggylx
1cvzxq
1ghjao
16xjfy
15jf7k
14lr19
qs85u
1cz3v9
th2qw
3mflsm
1n8evv
u5o2t
2jbqma
atwcj0
5opehl
ain0nn

8.5 Get all the information with a game name Pokemon.

```

1 CREATE DEFINER=`root`@`localhost` PROCEDURE `F3`()
2 BEGIN
3     SELECT * , (likes-comments) FROM gamecompanypost_final
4     WHERE game_name = 'POKEMON';
5
6
7 END

```

likes	comments	caption	mediaid	hashtag	Game_name	post_hashtag_id	(likes-comments)
29343	117	Coffee, clue	2028630000000000	['detectivepikachu']	POKEMON	1	29226
162907	2544	Feeling cute	2022090000000000	[]	POKEMON	10	160363
53639	375	☺️ Yes,	1972850000000000	['pokemongo']	POKEMON	100	53264
114222	543	#MondayM	1972200000000000	['mondaymotivation']	POKEMON	101	113679
114222	543	#MondayM	1972200000000000	['detectivepikachu']	POKEMON	102	113679
49459	185	Zeraora app	1971290000000000	['pokemontcg']	POKEMON	103	49274
62823	962	The only thi	1970740000000000	[]	POKEMON	104	61861
38166	240	Stick with yc	1969210000000000	['pokemonresolutions']	POKEMON	105	37926
48882	439	#Wednesda	1968430000000000	['wednesdaywisdom']	POKEMON	106	48443
48882	439	#Wednesda	1968430000000000	['pokemonresolutions']	POKEMON	107	48443
112849	10935	Everyone's c	1967020000000000	[]	POKEMON	108	101914
111485	728	These cuties	1965600000000000	['pokemoncenter']	POKEMON	109	110757
68087	729	Nobody: Ra	2021360000000000	[]	POKEMON	11	67358
81475	1321	Which villai	1964200000000000	['pokemongo']	POKEMON	110	80154
42437	180	Pokémon Du	1964000000000000	['pokemonduel']	POKEMON	111	42257
35183	261	Stay focus	1963360000000000	['wednesdaywisdom']	POKEMON	112	34922
101314	1259	🕵️ #Detecti	1961280000000000	['detectivepikachu']	POKEMON	113	100055

8.6 Get all the information with 'eevee' in their posts.

```
|CREATE DEFINER=`root`@`localhost` PROCEDURE `F4`()
|BEGIN
|    SELECT
|        id,
|        SUM( likes ) AS influence
|    FROM
|        gamecontentpost
|    WHERE
|        posts LIKE "%eevee%"
|    GROUP BY
|        id
|    ORDER BY
|        influence DESC
|    LIMIT 5
-;
END
```

id	influence
▶ 112137505186818	369
112115020174016	56
112137201003300	39
112116798902448	36
112130458011173	32

9 License

LICENSE:

MIT License

Copyright (c) 2019 Yijin Zhu

Permission is hereby granted, free of charge, to any person obtaining a copy of this software and associated documentation files (the "Software"), to deal in the Software without restriction, including without limitation the rights to use, copy, modify, merge, publish, distribute, sublicense, and/or sell copies of the Software, and to permit persons to whom the Software is furnished to do so, subject to the following conditions:

The above copyright notice and this permission notice shall be included in all copies or substantial portions of the Software.

THE SOFTWARE IS PROVIDED "AS IS", WITHOUT WARRANTY OF ANY KIND, EXPRESS OR IMPLIED, INCLUDING BUT NOT LIMITED TO THE WARRANTIES OF MERCHANTABILITY, FITNESS FOR A PARTICULAR PURPOSE AND NONINFRINGEMENT. IN NO EVENT SHALL THE AUTHORS OR COPYRIGHT HOLDERS BE LIABLE FOR ANY CLAIM, DAMAGES OR OTHER LIABILITY, WHETHER IN AN ACTION OF CONTRACT, TORT OR OTHERWISE, ARISING FROM, OUT OF OR IN CON

10 Reference

1. <https://stackoverflow.com/questions/10383044/fuzzy-string-comparison>
2. https://www.tutorialspoint.com/python/python_spelling_check.htm