

# HOLIDAY PLANNER DATABASE

(Rajendra Kumar Rajkumar, Monish Hirisave Raghu, Prudhvi Sekharamahanti)

## **1. Abstract**

To build a holiday planner site with corresponding database which will fulfill the necessities of a tourist planning for a vacation with information about tourist attractions, history, weather, hotels, transport services and social media trends of respective city of interest.

## **2. Introduction**

The project requires static data about tourist attractions, hotels, transport services to be obtained from their respective websites through web scraping and pre-defined datasets. It also requires dynamic data about weather and social media trends to be obtained from a weather website and social media websites through their respective APIs. All the above data is to be collected for popular cities of The United States of America.

The intention of the project is to offer one stop solution for all the basic information required by a tourist for his/her travel plan. Existing sites/databases are more confined to a specific information rather than providing everything required. This site/database will reduce the time and effort spent in searching data at multiple places. In addition, this site/database is integrated with social media trends which can give a wider view about the current happenings at respective cities.

Background Research: Background research included framing necessities of a tourist from various resources and verifying whether the existing holiday websites hold information regarding the same. Most of the website databases lacked one or the other necessary information. For example, a website which provided information about flights lacked information on tourist attractions and another website, which provided information about tourist attractions and hotels didn't have information about flights and latest weather.

## **3. Data Source Detailing:**

Data Sources: • Travel.usnews.com - Tourist attraction information.

A) kaggle.com – Flight information.

This dataset was obtained from kaggle and unnecessary columns were deleted using python. The dataset has 9 columns of information related to the flights like

- Origin city ID

- Departure city ID
- Arrival time
- Departure time
- Distance
- Stops in between
- Travel duration
- Flight ID
- Airline ID

#### B) Hotelscombined.com – Hotel information.

Information for the hotels was web scrapped from Hotelscombined.com using beautiful soups package in python. This had 8 columns which gave various parameters as shown below:-

- Hotel ID
- Hotel Name
- Hotel rating
- Hotel Type
- Price
- Review Count
- Landmark
- Address ID

#### C) Openweathermap.com – Updated weather information.

The data for weather is obtained from the above mentioned API. We have around 12 different values to check the weather .The data is fetched for 5 days from the day the user wants to retrieve .(for example if you check for May 1<sup>st</sup> it will fetch records till May 5<sup>th</sup>)

- Weather ID
- Temperature
- Max\_temperature
- Min\_temperature
- Pressure level
- Cloudiness
- Humidity
- Wind speed
- Rain volume
- Snow volume
- Weather description
- Date

D) Twitter – Social Media trends.

The data for the twitter is obtained from twitter API and these data are divided into three different tables for convenience.

- City ID
- Tweet ID
- Content
- Created Time
- Favourite count
- Retweet count
- User id
- Name
- Hashtag Id
- User ID
- Username
- User screenname
- Followers count

E) Reddit - Social Media trends.

Reddit data was obtained through the official API, and the dataset has more than 10 columns which were divided into two tables like reddit post and reddit user.

- City id
- Content
- Created time
- Upvote count
- Downvote count
- Post ID
- User ID
- Comments made
- Content Title
- Total Posts
- Total Comments
- Total Comments Score
- Total Post Score

F) Instagram - Social Media trends.

Instagram data was also obtained through the official API and divided into two tables like Instagram post and Instagram hashtag.

- Insta ID
- City ID
- Username
- Data created
- Content
- Likes
- Hashtag ID

G) Kaggle datasets – City/state information

The various states and city was obtained from kaggle and unnecessary columns and rows were deleted.

- City Id
- City Name
- City description
- State iD
- State Name
- State Description

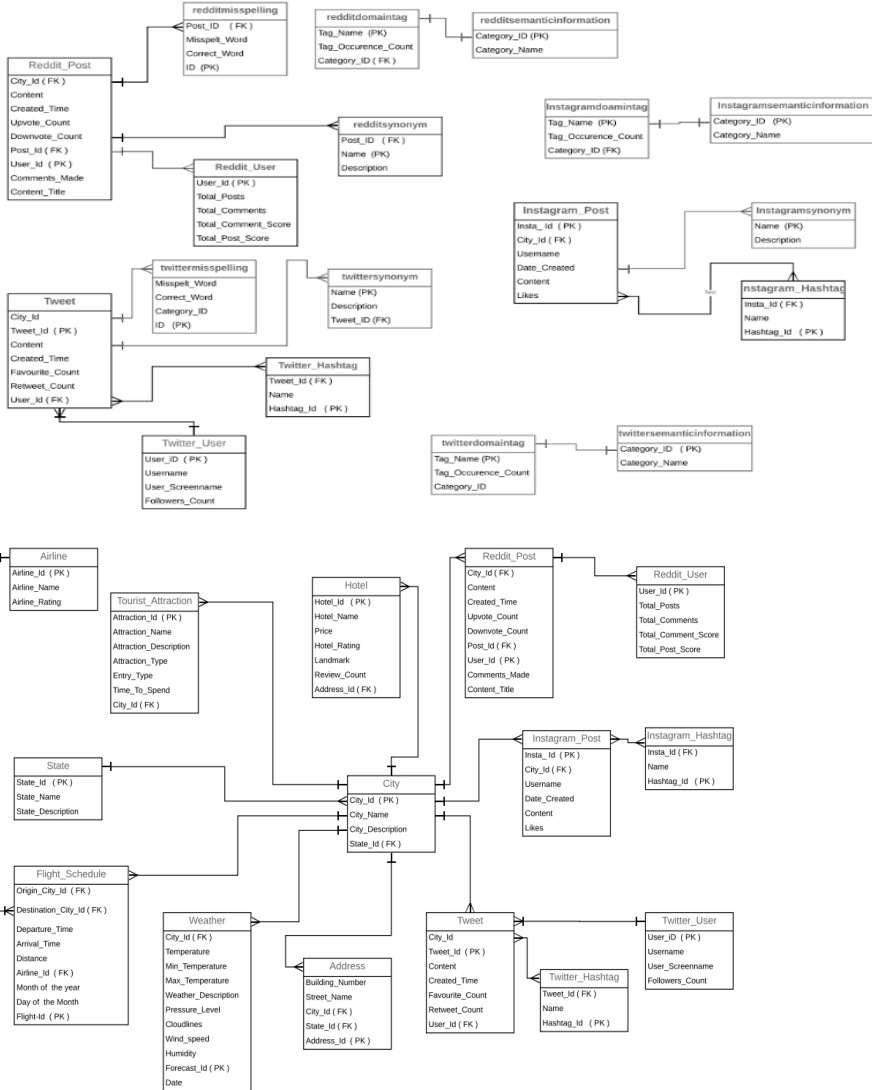
F) Tourist attraction

The dataset was taken from an authentic travel website. The dataset contains information about more 20 cities and their top attractions.

- Attraction Name
- Attraction Type
- Attraction description
- Entry type
- Time to S

## **4.ER diagram**

An entity relationship model or the entity-relationship(ER) diagram, is a graphical representation of entities and their relationships to each other, typically used in computing with regard to the organization of data within the databases or information systems.



we have used two diagrams to make it understandable and legible , as we had more than 25 tables to be shown.

## **5. Normalization**

Check that tables are in First Normal Form (1NF)

- Each table has a primary key: minimal set of attributes which can uniquely identify a record
  - Generated Primary keys for the tables which didn't had one.
- The values in each column of a table are atomic (No multi-value attributes allowed)
  - Cleaned the data so there we no atomic values in every table
- There are no repeating groups: two columns do not store similar information in the same table
  - Drop the tables which from the dataset which showed redundancy

Check that tables are in Second Normal Form (2NF)

- All requirements for 1st NF must be met.
  - Satisfied
- No partial dependencies.
  - While tidying the datasets cleaned the partial dependencies
- No calculated data
  - The variables depended upon other variables were removed

Check that tables are in Third Normal Form (3NF) and the final SQL

- All requirements for 2nd NF must be met.
  - Satisfied

**All the 26 tables in the model have a Primary Key (PK) and many non-key attributes and all these non key attributes are based on only primary key. All of them are in Third Normal Form**

## **Python Script For Dyanamic Data Collection(Weather):**

```
from instaloader import Instaloader, Profile
```

```
import xlrd
```

```
#Location of the excel file containg hotel links,city id and stateid
```

```
loc = ("C:/Users/prudh/Desktop/INFO 6210 DMDD/Project/hotels_urls.xlsx")
```

```
#opeing the workbook from the above mentioned location
```

```
wb = xlrd.open_workbook(loc)
```

```
#Fetching data from sheet 1
```

```
sheet = wb.sheet_by_index(0)
```

```
...
```

```
import pandas as ps
```

```
#Retriving Instagram API response
```

```
L =Instaloader()
```

```
#Declaring variables for Instagram data
```

```
user_name=[]
```

```
post_content=[]
```

```
post_likes=[]
```

```
date_created=[]
```

```
comments=[]
```

```
Insta_postid=[]
```

```
hashtags=[]
```

```
city_id=[]
```

```
count=100
```

```

posts_data_final=ps.DataFrame({ "City id":city_id,
                                "User Name":user_name,
                                "Post":post_content,
                                "DATE_CREATED":date_created,
                                "LIKES":post_likes,
                                "COMMENTS":comments,
                                "Post id":Insta_postid})

for i in range(sheet.nrows):
    city= sheet.cell_value(i,1)
    #Iterating over each post of user about a particular city
    for post in L.get_hashtag_posts(city):

        city_id.append(city)

        #User who posted the post
        user_name.append(post.owner_username)

        #Content of the post
        post_content.append(post.caption)

        #Number of likes
        post_likes.append(post.likes)

        #Date when the post has been created
        date_created.append(post.date)

        #Number of comments received for the post
        comments.append(post.comments)

```



```
#Hashtags of that post
tag=post.caption_hashtags
tag_string=' '.join(str(x) for x in tag)
hashtags.append(tag_string)
```

```
postid='ins'+ str(count)
Insta_postid.append(postid)
count=count+1
```

```
#Limiting the number of posts to 20 per user
if len(user_name)==10:
    break
```

```
#Appending all the attributes to a data frame
posts_data=ps.DataFrame({ "City id":city_id,
                           "User Name":user_name,
                           "Post":post_content,
                           "DATE_CREATED":date_created,
                           "LIKES":post_likes,
                           "COMMENTS":comments,
                           "Post id":Insta_postid,
                           })
posts_data_final=ps.concat([posts_data_final, posts_data])
user_name.clear()
post_content.clear()
post_likes.clear()
date_created.clear()
comments.clear()
```

```
Insta_postid.clear()
```

```
hashtags.clear()
```

```
city_id.clear()
```

C:\Users\prudh\Anaconda3\lib\site-packages\ipykernel\_launcher.py:67: FutureWarning: Sorting because non-concatenation axis is not aligned. A future version of pandas will change to not sort by default.

To accept the future behavior, pass 'sort=False'.

To retain the current behavior and silence the warning, pass 'sort=True'.

#Displaying the data frame

posts\_data\_final

Out[65]:

	COMMENTS	City id	DATE_CREATED	LIKES	Post	Post id	User Name	hashtags
0	0.0	boston	2019-04-26 15:13:59	0.0	They have the couch in the bay window as their...	ins100	chandrebo	misszsazsa misscoco sunseekers boston bostons ...
1	0.0	boston	2019-04-26 15:13:53	1.0	Thanks to all who tuned in to @concertwindow l...	ins101	lowlily1	vermont boston spring springinvtr springflowers...
2	0.0	boston	2019-04-26 15:13:39	1.0	Scene list for a new old project\n\nInitiated ...	ins102	nick_j_yeh	screenplay drama sütterlin handwriting product...
3	0.0	boston	2019-04-26 15:13:31	0.0	There's a fine line between being a Dulha and ...	ins103	jsjbespoke	jsjbespokestudio jadejasatyajeetbespokestudio ...

COMMENTS	City id	DATE_CREATED	LIKES	Post	Post id	User Name	hashtags
4 0.0	boston	2019-04-26 15:13:20	1.0	1966 in Iran. My Mother and Father. All the wa...	ins104	kambizmostofi	iran losangeles iranian longbeach sanfrancisco...
5 0.0	boston	2019-04-26 15:13:14	4.0	Still dance shot. Fabulous performance by my d...	ins105	ndelisle	dancer shesgottalent ballet contemporarydance ...
6 0.0	boston	2019-04-26 15:13:07	0.0	They say it is Spring. . . Boston weather hasn...	ins106	jodibean328	beaconhill boston acornstreet springinboston t...
7 0.0	boston	2019-04-26 15:13:03	2.0	Live And Let Live\n.\n.\n.\n.\n.\n.\n#orlando ...	ins107	bobmann29	orlando chicago atlanta gratefulheart peaceful...
8 0.0	boston	2019-04-26 15:12:51	0.0	Tulips are out at the garden! -\n.\n#boston #b...	ins108	paulksayed	boston bostonpublicgarden igersboston tulips f...
9 0.0	boston	2019-04-26 15:12:49	4.0	"Boston is an oasis in the desert, a place whe...	ins109	photography_ing03	boston massachusetts flowers cherryblossom lov...
0 0.0	newyorkcity	2019-04-26 15:14:29	0.0	We love winning 🏆🏆 DM now let's roll 👉 nashvill...	ins110	binary_medlyn	neworleans hongkong denmark boston denver saud...
1 0.0	newyorkcity	2019-04-26 15:14:14	0.0	@nas @tribeca #rapper #legend #hiphop #style #...	ins111	photographsbyj	rapper legend hiphop style mc newyorkcity phot...
2 0.0	newyorkcity	2019-04-26 15:14:03	3.0	🌟 Tag a friend 🙏\nFollow 👉 @Maludelos\n----- --...	ins112	maludelos	caniche latinamerica uruguay pitbullpuppies hu...
3 0.0	newyorkcity	2019-04-26 15:14:01	4.0	Hatz. #sony #sonya7iii #photography #streetpho...	ins113	badnickbadnick_photography	sony sonya7iii photography streetphotography a...

COMMENTS	City id	DATE_CREATED	LIKES	Post	Post id	User Name	hashtags
4 0.0	newyorkcity	2019-04-26 15:13:59	0.0	madrocktapas\nCheck @j_hollington125 's video/...	ins114	madrocktapas	hellskitchen spanishfood spanishbar pic fashio...
5 0.0	newyorkcity	2019-04-26 15:13:52	1.0	Dimitri from Paris. #legendary #dj #music #dis...	ins115	publicist_lifestyle	legendary dj music discohouse dance newyorkcit...
6 0.0	newyorkcity	2019-04-26 15:13:50	0.0	The Guggenheim Museum Atrium comparison with F...	ins116	james_mary_oconnor_architect	fultonstation guggenheim guggenheimmuseum fult...
7 0.0	newyorkcity	2019-04-26 15:13:47	0.0	We are partnering with @coveymusic and @super....	ins117	hindsighttees	hindsighttees tees apparel grunge clothing bra...
8 0.0	newyorkcity	2019-04-26 15:13:45	0.0	no day shall erase you from the moment of time...	ins118	qtpiemaomao	groundzero 911memorial newyorkcity
9 0.0	newyorkcity	2019-04-26 15:13:42	0.0	Volume Barbie Doll set ✨ Click The Link In My B...	ins119	queenminksbyneni_	eyelashextensiominks momentoffa eyebrows eyelas...
0 0.0	miami	2019-04-26 15:14:44	0.0	#miami #ftlauderdale #florida #atlanta #georgi...	ins120	matthios_mmr	miami ftlauderdale florida atlanta georgia hou...
1 0.0	miami	2019-04-26 15:14:42	0.0	Check out @3.3rica - \nShop one of a kind read...	ins121	rhondis_eyecandies	northbronx southbronx dyckman inwood washingto...
2 0.0	miami	2019-04-26 15:14:42	0.0	Glad to see you back!! The classic of @dstoryt...	ins122	dstorytellerflower	flowerbox flores flowers flowerlover flowersta...
3 0.0	miami	2019-04-26 15:14:42	0.0	Shop for prom dresses and accessories at JDO C...	ins123	jdo_couture_dress_shop	miami miamiwedding miamiweddingphotog rapher mi...
4 0.0	miami	2019-04-26 15:14:40	0.0	Foxcityla \n#popbeats #popstarroyalty #pophit ...	ins124	foxcityla	popbeats popstarroyalty pophit popbeatsforsale...

COMMENTS	City id	DATE_CREATED	LIKES	Post	Post id	User Name	hashtags
5 0.0	miami	2019-04-26 15:14:36	0.0	Delicioso Trifásico de chicharrón costilla y f...	ins125	bandejapaisa.restaurant	bandejapaisarestaurant colombianoseni west...
6 0.0	miami	2019-04-26 15:14:35	1.0	Shop for prom dresses and accessories at JDO C...	ins126	shopjdo	miami miamiwedding miamiweddingphotog rapher mi...
7 0.0	miami	2019-04-26 15:14:35	0.0	MIA REALTY CORP. NOW HIRING AGENTS 🏠 LEADS PROV...	ins127	mia_realty	commercialrealestate miami miamirealestate now...
8 0.0	miami	2019-04-26 15:14:35	1.0	Eliş Şile Bezi Sevda Bluz \\nKampanya !! İndiri...	ins128	elissilebezi	elissilebezi silemodacom beach beachwear germa...
9 0.0	miami	2019-04-26 15:14:35	0.0	Feliz Viernes mi gente, #Miami, #sol, #limochiva	ins129	gustavofranco11	miami sol limochiva
. ... .	...	...	...	...	...	...	...
0 0.0	sanfransisco	2019-04-26 15:14:14	0.0	@ericrhoads a reminder of yesterday's plein ai...	ins320	tubacschoolofflineartllc	pace19 pleinairconvention sanfransisco
1 0.0	sanfransisco	2019-04-26 15:13:05	8.0	🤔 Rep your sign 🌐🌐\\nstarcrossedcharms.com\\n\\n#...	ins321	starcrossedcharms	jewelrygram zodiac fashionista taurusgang taur...
2 0.0	sanfransisco	2019-04-26 14:56:57	7.0	\\n🌐FAB FRIDAY🌐 \\n \\nFetchin g myself a car wit...	ins322	doorwaysofchicago	sanfransisco classiccars classiccar classiccar...
3 1.0	sanfransisco	2019-04-26 14:54:09	1.0	Pick your poison, pick your ending...\\n.\\n#mel...	ins323	catastrophe27	melsdrivein jukebox pickyourpoison photography...
4 1.0	sanfransisco	2019-04-26 14:52:04	8.0	Sign up for summer cycling fun! Book your kids...	ins324	tripsforkidsmarin	tripsforkidsmarin tripsforkidsmarinstor y trail...

COMMENTS	City id	DATE_CREATED	LIKES	Post	Post id	User Name	hashtags
5 2.0	sanfransisco	2019-04-26 14:50:27	20.0	-----\n👉👉👉👉👉...	ins325	derin_n_demir	arizona ankara brasil phuket california canada...
6 3.0	sanfransisco	2019-04-26 14:45:55	7.0	Sharing our trip to San Fransisco on the blog ...	ins326	calliemoonfarmhouse	travelcalifornia sanfransisco sanfranciscofood...
7 4.0	sanfransisco	2019-04-26 14:44:15	40.0	Movement Appreciation Day 🌟\n\nI just implemen...	ins327	blitxcaroline	movementappreciationday gratefultomove
8 0.0	sanfransisco	2019-04-26 14:31:32	9.0	I love weekends almost as much as I love pizza...	ins328	whenieatpizza	whenieatpizza pizza pizzaislife food foodlover...
9 0.0	sanfransisco	2019-04-26 14:28:40	6.0	First 2 days in california have been incredibl...	ins329	essentiallymegz2213	travel wonerlust california sanfransisco napav...
0 0.0	sanantonio	2019-04-26 15:21:28	0.0	Another wall installed this morning for @ksatn...	ins330	foreverpetalsbyvee	foreverpetalsbyvee morethanjustpaper paperflow...
1 0.0	sanantonio	2019-04-26 15:21:24	0.0	A little detour on the way to Austin... #sanantonio...	ins331	gingermama1r	sanantonio riverwalksanantonio thealamo texas
2 0.0	sanantonio	2019-04-26 15:21:10	0.0	Breeze Weekend event #1. #Friday night: Rumtop...	ins332	nyamcashpromotion	1 friday breezeweekendhouston trinidadcarnival...
3 0.0	sanantonio	2019-04-26 15:21:04	0.0	[ 1 MÊS DE INTERCÂMBIO NOS ESTADOS UNIDOS ]\n\n...	ins333	intercambioeviagem	intercâmbionasférias eua timesquare disney go...
4 0.0	sanantonio	2019-04-26 15:20:26	0.0	MOLLETE! Bolillo tostado con frijoles refritos...	ins334	culebra30	igsanantonio sanantonio culebrasupermeatmarket...
5 1.0	sanantonio	2019-04-26 15:20:06	1.0	Shoes shoes 👠 - Michael by Michael Kors \nShoe...	ins335	cmsanantonio	

COMMENTS	City id	DATE_CREATED	LIKES	Post	Post id	User Name	hashtags
6 0.0	sanantonio	2019-04-26 15:19:54	1.0	Work graciously gave us the day off for Battle...	ins336	cinderfoodie	satx sanantonio tx texas 210 safood safoodpics...
7 1.0	sanantonio	2019-04-26 15:19:50	0.0	Mother's Day is right around the corner and th...	ins337	spaviaalamoranch	
8 0.0	sanantonio	2019-04-26 15:19:26	1.0	Love #SanAntonio 🌸🌺🌻🌼🌾🌿🌻🌺 Businesses, schools,...	ins338	hezkjones	sanantonio battleofflowersparade battleoftheal...
9 0.0	sanantonio	2019-04-26 15:19:25	1.0	Hoy viernes te ofrecemos estos deliciosos Moji...	ins339	elmuseodelpollo	mojitos drink mojitos2x1 mojitosfrozen mojitos...
0 0.0	orlando	2019-04-26 15:21:50	0.0	Mom deserves something special for this Mother...	ins340	wildaboutcleaning	wildaboutcleaning wildlifematters cleaning cle...
1 0.0	orlando	2019-04-26 15:21:48	1.0	Check out our IG exclusive weekend sale!!#weap...	ins341	premierepawn	weapons firearms shotgun fire 2ndamendment ins...
2 0.0	orlando	2019-04-26 15:21:46	1.0	#newyork #malaysia#london #toronto#miami #texas...	ins342	lawrence_ifey12	newyork malaysia london toronto miami texas fl...
3 0.0	orlando	2019-04-26 15:21:21	1.0	Transform your skin with Microneedling 😊 This n...	ins343	bridgetbachandtro mbo	bridgettrombo medicaaesthetician saloninx mi...
4 0.0	orlando	2019-04-26 15:21:15	0.0	🔴 TONIGHT!! 🔴 ****\nITS ABOUT TO GO DOWN 📉 !!!\n...	ins344	spasemarketing	funkfest kickoffparty orlando funkfestorlando ...
5 1.0	orlando	2019-04-26 15:21:13	3.0	Ya se siente el verano a la vuelta de la esqui...	ins345	rik.riv	realestate larosarealty realtorlife mindset bu...
6 0.0	orlando	2019-04-26 15:21:04	0.0	[ 1 MÊS DE INTERCÂMBIO NOS ESTADOS UNIDOS ]\n\...	ins346	intercambioeviagem	intercâmbionasférias eua timessquare disney go...

	COMM ENTS	City id	DATE_CR EATED	LIK ES	Post	Pos t id	User Name	hashtags
7	1.0	orlando	2019-04-26 15:20:59	3.0	Faded faded Faded bout to drop all my videos f...	ins347	juswav3	juswave independentartist star music linkinbio...
8	0.0	orlando	2019-04-26 15:20:43	0.0	TOMORROW Visit Us at SoFlo Taco 🌮 Battle & Cra...	ins348	theonlyhardhops	soflo hops spirits cocktails naturalingredient...
9	0.0	orlando	2019-04-26 15:20:12	0.0	Value yourself and others will too ❤️❤️ #modernm...	ins349	modernmermaidhair_ r_	modernmermaidhair whatimwearing hairstylist st...

250 rows × 8 columns

#Separating hashtags part to a seperate data frame for further data manipulations

```
insta_hashtag=posts_data_final[['Post id','hashtags']]
```

...

#Display of hashtags data frame

```
insta_hashtag
```

Out[62]:

	Post id	hashtags
0	ins100	misszsazsa misscoco sunseekers boston bostons ...
1	ins101	vermont boston spring springinvtt springflowers...
2	ins102	screenplay drama sütterlin handwriting product...
3	ins103	jsjbespokestudio jadejasatyajeetbespokestudio ...
4	ins104	iran losangeles iranian longbeach sanfrancisco...
5	ins105	dancer shesgottalent ballet contemporarydance ...
6	ins106	beaconhill boston acornstreet springinboston t...
7	ins107	orlando chicago atlanta gratefulheart peaceful...
8	ins108	boston bostonpublicgarden igersboston tulips f...
9	ins109	boston massachusetts flowers cherryblossom lov...



	Post id	hashtags
0	ins110	neworleans hongkong denmark boston denver saud...
1	ins111	rapper legend hiphop style mc newyorkcity phot...
2	ins112	caniche latinamerica uruguay pitbullpuppies hu...
3	ins113	sony sonya7iii photography streetphotography a...
4	ins114	hellskitchen spanishfood spanishbar pic fashio...
5	ins115	legendary dj music discohouse dance newyorkcit...
6	ins116	fultonstation guggenheim guggenheimmuseum fult...
7	ins117	hindsighttees tees apparel grunge clothing bra...
8	ins118	groundzero 911memorial newyorkcity
9	ins119	eyelashextensiominksmomentoffa eyebrows eyelas...
0	ins120	miami ftlauderdale florida atlanta georgia hou...
1	ins121	northbronx southbronx dyckman inwood washingto...
2	ins122	flowerbox flores flowers flowerlover flowersta...
3	ins123	miami miamiwedding miamiweddingphotographer mi...
4	ins124	popbeats popstarroyalty pophit popbeatsforsale...
5	ins125	bandejapaisarestaurant colombianosenmiami west...
6	ins126	miami miamiwedding miamiweddingphotographer mi...
7	ins127	commercialrealestate miami miamirealestate now...
8	ins128	elissilebezi silemodacom beach beachwear germa...
9	ins129	miami sol limochiva
...	...	...
0	ins320	pace19 pleinairconvention sanfransisco
1	ins321	jewelrygram zodiac fashionista taurusgang taur...
2	ins322	sanfransisco classiccars classiccar...
3	ins323	melsdrivein jukebox pickyourpoison photography...
4	ins324	tripsforkidsmarin tripsforkidsmarinstory trail...

	Post id	hashtags
5	ins325	arizona ankara brasil phuket california canada...
6	ins326	travelcalifornia sanfransisco sanfranciscofood...
7	ins327	movementappreciationday gratefultomove
8	ins328	whenieatpizza pizza pizzaislife food foodlover...
9	ins329	travel wonerlust california sanfransisco napav...
0	ins330	foreverpetalsbyvee morethanjustpaper paperflow...
1	ins331	sanantonio riverwalksanantonio thealamo texas
2	ins332	1 friday breezeweekendhouston trinidadcarnival...
3	ins333	intercâmbionasférias eua timesquare disney go...
4	ins334	igsanantonio sanantonio culebrasupermeatmarket...
5	ins335	
6	ins336	satx sanantonio tx texas 210 safood safoodpics...
7	ins337	
8	ins338	sanantonio battleofflowersparade battleoftheal...
9	ins339	mojitos drink mojitos2x1 mojitosfrozen mojitos...
0	ins340	wildaboutcleaning wildlifematters cleaning cle...
1	ins341	weapons firearms shotgun fire 2ndamendment ins...
2	ins342	newyork malaysia london toronto miami texas fl...
3	ins343	bridgettrombo medicalaesthetician saloninxs mi...
4	ins344	funkfest kickoffparty orlando funkfestorlando ...
5	ins345	realestate larosarealty realtorlife mindset bu...
6	ins346	intercâmbionasférias eua timesquare disney go...
7	ins347	juswave independentartist star music linkinbio...
8	ins348	soflo hops spirits cocktails naturalingredient...
9	ins349	modernmermaidhair whatimwearing hairstylist st...

250 rows × 2 columns

#Reshaping of hashtags data frame to store each hash tag in a seperate row

reshaped = \

(insta\_hashtag.set\_index(insta\_hashtag.columns.drop('hashtags',1).tolist())

.hashtags.str.split(' ', expand=True)

.stack()

.reset\_index()

.rename(columns={0:'hashtags'})

.loc[:, insta\_hashtag.columns]

)

#Converting attributes of hashtag data frame to seperate lists

postid = reshaped['Post id'].tolist()

hashtag=reshaped['hashtags'].tolist()

hashtagid=[]

count=1

#Autocreation of unique hashtag id for each hashtag

for i in postid:

    hashtagid.append('hash'+str(count))

    count=count+1

...

#Appending all the manipulated attributes to a new data frame

Instagram\_hashtags\_city=ps.DataFrame({"HashTag Id":hashtagid,

    "Post id":postid,

    "Hash Tag":hashtag})

#Display of consolidated final hashtags data frame

Instagram\_hashtags\_city

Out[51]:

	HashTag Id	Post id	Hash Tag
0	hash1	ins100	tupac
1	hash2	ins100	meagainsttheworld
2	hash3	ins100	cali
3	hash4	ins100	sample
4	hash5	ins100	mines
5	hash6	ins100	listen
6	hash7	ins100	weatcoast
7	hash8	ins100	atl
8	hash9	ins100	boston
9	hash10	ins100	newyork
10	hash11	ins100	indiana
11	hash12	ins100	texas
12	hash13	ins100	arizona
13	hash14	ins100	ohio
14	hash15	ins100	weed
15	hash16	ins100	420
16	hash17	ins100	producer
17	hash18	ins100	beats
18	hash19	ins100	smoke
19	hash20	ins100	writer
20	hash21	ins100	lyrics
21	hash22	ins100	music
22	hash23	ins100	new
23	hash24	ins101	nature
24	hash25	ins101	naturelovers
25	hash26	ins101	califonia

	HashTag Id	Post id	Hash Tag
26	hash27	ins101	europa
27	hash28	ins101	newportbeach
28	hash29	ins101	newyorkcity
29	hash30	ins101	newyork
...	...	...	...
4186	hash4187	ins348	grocery
4187	hash4188	ins349	orlandobloom
4188	hash4189	ins349	orlando
4189	hash4190	ins349	bloom
4190	hash4191	ins349	katyperry
4191	hash4192	ins349	korlando
4192	hash4193	ins349	actor
4193	hash4194	ins349	singer
4194	hash4195	ins349	fanpage
4195	hash4196	ins349	love
4196	hash4197	ins349	look
4197	hash4198	ins349	lordoftherings
4198	hash4199	ins349	lotr
4199	hash4200	ins349	will
4200	hash4201	ins349	legolas
4201	hash4202	ins349	potc
4202	hash4203	ins349	beauty
4203	hash4204	ins349	beautiful
4204	hash4205	ins349	role
4205	hash4206	ins349	movie
4206	hash4207	ins349	film

	HashTag Id	Post id	Hash Tag
4207	hash4208	ins349	celebrity
4208	hash4209	ins349	celebs
4209	hash4210	ins349	cute
4210	hash4211	ins349	gorgeous
4211	hash4212	ins349	hot
4212	hash4213	ins349	edit
4213	hash4214	ins349	photoedit
4214	hash4215	ins349	videoedit
4215	hash4216	ins349	fanpage

4216 rows × 3 columns

...

```
export_csv = posts_data_final.to_csv(r'C:\Users\prudh\Desktop\INFO 6210 DMDD\Project\CSV
files\Instagram_post_API_latest1.csv', index = None, header = True)
```

```
export_csv = Instagram_hashtags_city.to_csv(r'C:\Users\prudh\Desktop\INFO 6210 DMDD\Project\CSV
files\Instagram_hashtag_latest1.csv', index = None, header = True)
```

...

""CONCLUSIONS:

This code uses the list of cities, selected for the project and tries to find the popularity parameters related

to the same through Twitter API, stores the results in dataframe and exports it to a CSV file.

CONTRIBUTIONS:

RAJENDRA KUMAR RAJKUMAR - 40%

MONISH HIRISAVE RAGHU - 20%

PRUDHVI CHANDRA SEKHARAMAHANTI -40%

CITATIONS:

1. <https://www.geeksforgeeks.org>
2. [https://github.com/nikbearbrown/INFO\\_6210](https://github.com/nikbearbrown/INFO_6210)

The code regarding extraction of information from twitter was used from the above mentioned resources

Original writtten code - 60%

Code referenced from external sources(but modified suiting needs) - 40%

## **6.1 Social Media Use Cases**

**Use case 1:** What are people saying about me (somebody)?

#stored Procedure (1) :sp\_tweets\_about\_city()

#Description: Consists of city\_name as input parameter and returns the list of all tweets made on the

# with regard to the particular city.

DELIMITER \$\$

CREATE DEFINER=`root`@`localhost` PROCEDURE `sp\_tweets\_about\_city`(IN city\_name text)

READS SQL DATA

BEGIN

select t.Tweet\_ID, t.Content, tu.User\_Name, city.City\_Name

from holiday\_planner.city city, holiday\_planner.tweet t, holiday\_planner.twitteruser tu

where t.City\_ID = city.City\_ID and

city.City\_Name = city\_Name and t.User\_ID = tu.User\_ID;

END\$\$

DELIMITER ;

# Procedure call for sp\_tweets\_about\_city() with 'Boston' as input parameter

Call sp\_tweets\_about\_city ('Boston');

#drop procedure sp\_tweets\_about\_city;

Output:

Result Grid	Filter Rows:	Export:	Wrap Cell Content:
Tweet_ID	Content	User_Name	City_Name
1121771463911116800	@AOC I ask you AOC, "Should the Boston Bomber have the right to vote?" A simple yes or no, will tell a lot.	TOM LEFLER	Boston
1121771466326863873	Democrats Are Answering the Wrong Questions. <a href="https://t.co/vQe1bmuFgP">https://t.co/vQe1bmuFgP</a>	US President News	Boston
1121771475915223040	"Self-reflection is one of the biggest smallest things you can do." - @malialazu @Boston_CM #CMinclusive	Small Army	Boston

**Use case(2):** How viral are my posts??

#use case (2) :sp\_twitter\_virality\_of\_a\_city()

#Description: Consists of city\_name as input parameter and returns the list of all tweets made on the

# with regard to the particular city along with the retweet count and favourite count denoting the virality of the posts associated with the city

DELIMITER \$\$

CREATE DEFINER='root'@'localhost' PROCEDURE `sp\_twitter\_virality\_of\_a\_city`(IN city\_name text)

READS SQL DATA

BEGIN

select t.Tweet\_ID, t.Content, tu.User\_Name, t.Retweet\_Count, t.Favourite\_Count , city.City\_Name

from holiday\_planner.city city, holiday\_planner.tweet t, holiday\_planner.twitteruser tu

where t.City\_ID = city.City\_ID and

city.City\_Name = city\_Name and t.User\_ID = tu.User\_ID;

END\$\$;



DELIMITER ;

# Procedure call for sp\_tweets\_about\_city() with 'Boston' as input parameter

Call sp\_twitter\_virality\_of\_a\_city ('Chicago');

#drop procedure sp\_twitter\_virality\_of\_a\_city;

Output:

Result Grid		Filter Rows:	Export:	Wrap Cell Content:	
Tweet_ID	Content	User_Name	Retweet_Count	Favourite_Count	City_Name
1121771469170728961	RT @reallifemaggie: more BREAKING news for you, this one a little more close to my &lt;3 cuz Chicago connex: https://...	Anny Shaw	3	2	Chicago

**Use case(3):** What posts are likely to be interesting to me?

#use case (3) :sp\_interesting\_reddit\_posts()

#Description: Picks up the reddit posts associated with travel as the overall intention of the databse to support

#travellers seeking informmation

DELIMITER \$\$

CREATE DEFINER=`root`@`localhost` PROCEDURE `sp\_interesting\_reddit\_posts`()

READS SQL DATA

BEGIN

select city.City\_Name, p.Content\_Title as 'Heading', p.Content as 'Content', p.Post\_ID as 'Post ID'

from holiday\_planner.redditpost p, holiday\_planner.city city

where p.city\_id = city.City\_ID and

p.Content like '%travel%';

END\$\$;

DELIMITER ;

# Procedure call for sp\_interesting\_reddit\_posts with predefined 'trvael' search crriteria

Call sp\_interesting\_reddit\_posts ();

Output:

Result Grid	Filter Rows:	Export:	Wrap Cell Content:
City_Name	Heading	Content	Post ID
Miami	New Traveler Here, Quick Questions	New Traveler Here, Quick Questions	bdj29w
Seattle	Traveling to Seattle, then Vancouver, then Banff in June. Confused on rental cars...	Traveling to Seattle, then Vancouver, then Banff in June. Confused on rental cars...	bdthfs
Boston	Traveling solo to Iceland in the summer. I need your suggestions!	Traveling solo to Iceland in the summer. I need your suggestions!	be9f4g
Miami	Funds needed for 2 months travelling around the USA	Funds needed for 2 months travelling around the USA	bf1q01
Chicago	Looking for Romantic Midwest Vacation Ideas, Picky Travelers	Looking for Romantic Midwest Vacation Ideas, Picky Travelers	bfc8jw
Chicago	Traveling the US by train	Traveling the US by train	bfk09l
Los Angeles	Traveling Minors and Hotels	Traveling Minors and Hotels	bfxsy5
Los Angeles	what's the best city to travel to in California?	what's the best city to travel to in California?	bg50vg
Seattle	Travel suggestions for September: warm, sunny, beaches, hiking	Travel suggestions for September: warm, sunny, beaches, hiking	bgmn68
Los Angeles	Help: West Coast USA - 2nd Half May - 15 days travel	Help: West Coast USA - 2nd Half May - 15 days travel	bhc7c

**Use Case(4)** : What posts like me?

#use case (4) :sp\_similar\_instagram\_posts

#Description: Picks up the set of instagram posts which has hashtags similar to that user's interests confirming the

# similarity in interest

DELIMITER \$\$

```
CREATE DEFINER='root'@'localhost' PROCEDURE `sp_similar_instagram_posts`(IN place_to_visit text, IN favourite_shoes text,
```

```
IN favousite_music text, IN clothing text)
```

```
READS SQL DATA
```

```
BEGIN
```

```
select p.`Post id`, p.Post
```

```
from holiday_planner.instagramhashtag h, holiday_planner.instagrampost p
```

```
where h.`Post id` = p. `Post id` and
```

```
(h.`Hash Tag` like place_to_visit or h.`Hash Tag` like favourite_shoes or h.`Hash Tag` like favousite_music
```

or h.`Hash Tag` like clothing );

END\$\$

DELIMITER ;

# Procedure call for users\_with\_most\_retweeted\_tweets() with '2' as input parameter

Call sp\_similar\_instagram\_posts('sydney','sneakers','retro','menswear');

Output:

Result Grid			Filter Rows:	Export:	Wrap Cell Content:
	Post id	Post			
▶	ins159	Familialamezcladotcomlatinbillboardslasvegasurb...			
	ins162	WeresportingfirsttimehomebuyervibesthisFriday...			
	ins162	WeresportingfirsttimehomebuyervibesthisFriday...			
	ins163	ASandCatattheWashingtonDCZooCaughthimsta...			
	ins166	ÀesperadeAnaJuliafamilyfamiliafamilypicturesfa...			
	ins255	ILOVEthispiece			
	ins281	Whenyouthatthat soundjustrightfeelnopainresona...			
	ins344	TONIGHTITSABOUTTODOWNLionsPride123C...			

**Use Case (5):** What users post like me?

#use case (5) :sp\_similar\_instagram\_users

#Description: Picks up the set of instagram users who use hashtags similar to that user's interests confirming the

# similarity in interest

DELIMITER \$\$

CREATE DEFINER=`root`@`localhost` PROCEDURE `sp\_similar\_instagram\_users`(IN place\_to\_visit text, IN favourite\_shoes text,

IN favousite\_music text, IN clothing text)

READS SQL DATA

BEGIN

select p.`User Name`

from holiday\_planner.instagramhashtag h, holiday\_planner.instagrampost p

where h.`Post id` = p. `Post id` and

(h.`Hash Tag` like place\_to\_visit or h.`Hash Tag` like favourite\_shoes or h.`Hash Tag` like favousite\_music  
or h.`Hash Tag` like clothing );

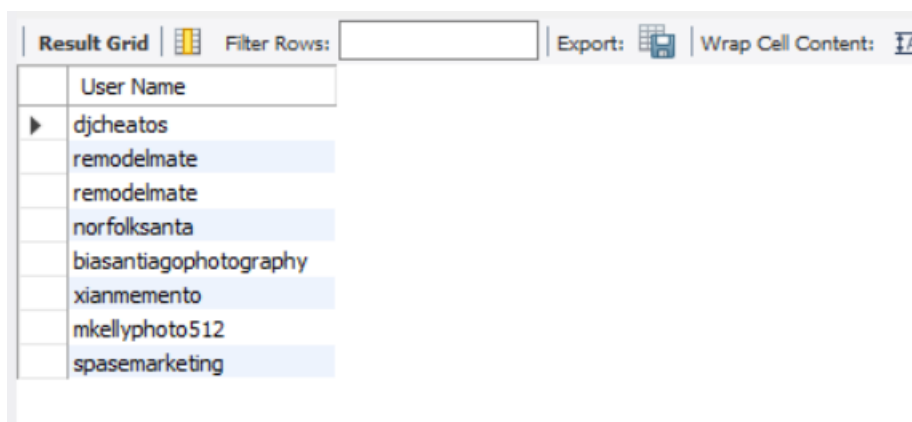
END\$\$

DELIMITER ;

# Procedure call for users\_with\_most\_retweeted\_tweets() with '2' as input parameter

Call sp\_similar\_instagram\_users('sydney','sneakers','retro','menswear');

Output:



The screenshot shows a database query result grid. At the top, there is a toolbar with 'Result Grid', 'Filter Rows' (with a search box), 'Export' (with a download icon), and 'Wrap Cell Content' (with a text icon). Below the toolbar is a table with one column labeled 'User Name'. The table contains the following rows:

User Name
djcheatos
remodelmate
remodelmate
norfolksanta
biasantiagophotography
xianmemento
mkellyphoto512
spasemarketing

**Use Case (6):** Who should I be following?

#use case (6) :sp\_users\_to\_follow()

#Description: Consists of number\_of\_users as input parameter and returns the list of twitters users(userID, username and user secreen name),

#where number of results is controlled by the input paramater

#Joins used: Tables 'tweet' and 'twitteruser' are joined using 'User\_ID'.

# Since the databse holds information about tweets tweeted with regard to cities. The user with most retweeted trave tweets must be

# a good content provider and is suitable person to be followed.

DELIMITER \$\$

CREATE DEFINER=`root`@`localhost` PROCEDURE `sp\_users\_to\_follow`(IN number\_of\_users INT)

READS SQL DATA

BEGIN

```
SELECT sum(t.Retweet_Count) as 'Retweet Count', u.User_ID, u.User_Name,u.User_Screen_Name
from holiday_planner.tweet t, holiday_planner.twitteruser u
where t.User_ID = u.User_ID
group by u.User_ID
order by sum(t.Retweet_Count) DESC
LIMIT number_of_users;
```

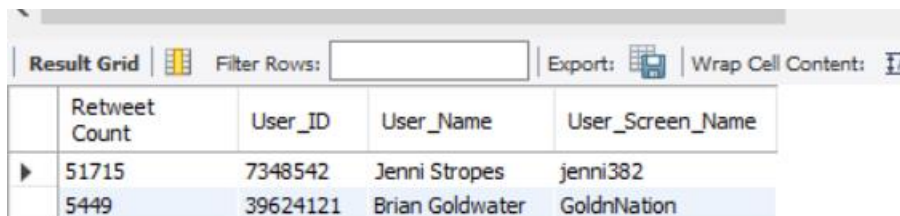
END\$\$

DELIMITER ;

# Procedure call for users\_with\_most\_retweeted\_tweets() with '2' as input parameter

Call sp\_users\_to\_follow (2);

Output:



The screenshot shows a SQL query result grid with the following data:

	Retweet Count	User_ID	User_Name	User_Screen_Name
▶	51715	7348542	Jenni Stropes	jenni382
	5449	39624121	Brian Goldwater	GoldnNation

**Use Case (7):** What topics are trending in my domain?

#use case (7) :sp\_city\_specific\_twitter\_trending\_topics()

#Description: Displays topics trending in an the city of interest of the user. The input parameters will be city name and

# the number of tpics the user is interested in"

DELIMITER \$\$

```
CREATE DEFINER=`root`@`localhost` PROCEDURE `sp_city_specific_twitter_trending_topics`(IN city_name text,
number_of_topics INT)
```

READS SQL DATA

BEGIN

```

select t.Trend_Topic as 'Trending Topics', t.Tweet_Volume as 'Tweet Volume', city.City_Name as 'City Name'
from holiday_planner.twittertrend t, holiday_planner.city city
where t.City_ID = city.City_ID and
city.city_Name like city_name
LIMIT number_of_topics;

END$$

DELIMITER ;

```

```

# Procedure call for users_with_most_retweeted_tweets() with '2' as input parameter
Call sp_city_specific_twitter_trending_topics('Seattle', 10);

```

Output:

Result Grid			
Filter Rows:		Export:	
Wrap Cell Content:			
	Trending Topics	Tweet Volume	City Name
▶	#NFLDraft	294766	Seattle
	Sweat	54978	Seattle
	Giants	300581	Seattle
	Packers	45453	Seattle
	Andre Dillard	12005	Seattle
	Haskins	132478	Seattle
	Daniel Jones	101774	Seattle
	Josh Allen	48083	Seattle
	Devin Bush	43360	Seattle
	Josh Jacobs	11832	Seattle

**Use Case(8)** : What keywords/ hashtags should I add to my post?

#use case (8) :sp\_hashtags\_to\_use()

#Description: Displays the hashtags which are popular in travel domain associated with cities and the user can select the number of

#hashtags he/she wants to see in the output, which is presented along with Retweet and Favourite count.

DELIMITER \$\$

CREATE DEFINER=`root`@`localhost` PROCEDURE `sp\_hashtags\_to\_use`(IN number\_of\_topics INT)

READS SQL DATA

BEGIN

select h.Name, (t.Retweet\_Count + t.Favourite\_Count) as 'Total Popularity of hashtag'

from holiday\_planner. tweet t, holiday\_planner. twitterhashtag h, holiday\_planner.twitteruser u

where t.User\_ID = u.User\_ID and h.Tweet\_ID = t.Tweet\_ID

group by h.Name

order by 'Total Popularity of hashtag' DESC

LIMIT number\_of\_topics ;

END\$\$

DELIMITER ;

# Procedure call for users\_with\_most\_retweeted\_tweets() with '2' as input parameter

Call sp\_hashtags\_to\_use(5);

Output:

Result Grid			Filter Rows:	Export:	Wrap Cell Content:
	Name	Total Popularity of hashtag			
▶	ShowandVern	0			
	Seattle	4			
	CMindusive	0			
	dirtythirty	0			
	TheCollege	1			

#Use Case (9): Should I follow somebody back?

**use case (9):**sp\_reddit\_users\_to\_follow

#Description: Picks up the set of reddit users who post asticles regarding travel, journey, hotels so that aspiring travellers

# follow them to get much inputs.

DELIMITER \$\$

```
CREATE DEFINER=`root`@`localhost` PROCEDURE `sp_reddit_users_to_follow`(IN number_of_users INT)
```

```
    READS SQL DATA
```

```
BEGIN
```

```
select p.User_ID, p.Content, p.Content_Title
```

```
from holiday_planner.redditpost p, holiday_planner.reddituser u
```

```
where p.User_ID = u.User_ID and
```

```
p.content like '%travel%'
```

```
or p.content like '%journey%'
```

```
or p.content like '%hotel%'
```

```
or p.content like '%restaurant%'
```

```
LIMIT number_of_users;
```

```
END$$
```

```
DELIMITER ;
```

# Procedure call for users\_with\_most\_retweeted\_tweets() with '2' as input parameter

Call sp\_reddit\_users\_to\_follow(5);

Output:

	User_ID	Content	Content_Title
►	Myardraug	Finding a hotel room in Las Vegas with a jacuzzi in the room?	Finding a hotel room in Las Vegas with a jacuzzi in the room?
	alirecalcitrant	Traveling Minors and Hotels	Traveling Minors and Hotels
	Myardraug	Finding a hotel room in Las Vegas with a jacuzzi in the room?	Finding a hotel room in Las Vegas with a jacuzzi in the room?
	AccFire99	New Traveler Here, Quick Questions	New Traveler Here, Quick Questions
	alirecalcitrant	Traveling Minors and Hotels	Traveling Minors and Hotels

## **6.2 Tables with syntactic and semantic information about tags:**



Social Tagging Tables created:

1. instagramdomaintag

Result Grid	Filter Rows:	Edit:
Tag_Name	Tag_Occurence_Count	Category_ID
amsterdam	2	touristplace
architecture	6	touristplacetype
art	12	touristplacetype
atlanta	22	touristplace
austin	13	touristplace
austria	3	touristplace
bangkok	3	touristplace
berlin	2	touristplace
beverlyhills	3	touristplace
boatlife	4	touristplacetype
boston	17	touristplace
brazil	3	touristplace
california	21	touristplace
charleston	9	touristplace
chicago	13	touristplace
colombia	3	touristplace
dallas	9	touristplace
dekonailou...	4	touristplace
denmark	3	touristplace
denver	19	touristplace

2. instagramsemanticinformation

Result Grid	Filter Rows:	
Category_ID	Category_Name	
Category_ID	Category_Name	
touristplace	Tourist Place	
touristplacetype	Tourist Place Type	
transportoption	Transport Option	
travellerrequirement	Traveller Requirement	
travellertype	Traveller Type	
travelpurpose	Travel Purpose	
NULL	NULL	

3. instagramsynonym

```

8 • select *
9   from instagramsynonym;

```

Result Grid | Filter Rows: | Ed


	Name	Description
▶	ATL	Hartsfield International Airport
	Atx	Austin Texas
	CCM	Canada Cycle & Motor Co. Ltd
	CMG	Chipotle Mexican Grill
	DMV	Department of Motor Vehides
	NBa	National Basket association
	Nfldraft	National football league draft
	nightlifeny	NewYork city nightlife
	Philly	Philadelphia
	Rnb	Rhythm and blues
	SWV	Sisters with Voices band
	TBT	Throwback Thursday
	Wanderlust	a strong desire to travel
*	NULL	NULL

#### 4. redditdomaintag

```

11 • select *
12   from redditdomaintag;

```

Result Grid | Filter Rows: | Edit: 

	Tag_Name	Tag_Occurence_Count	Category_ID
▶	Austin	2	touristplace
	Beach	2	touristplacetype
	flights	3	transportoption
	itinerary	2	travelplan
	Miami	2	touristplace
*	NULL	NULL	NULL



#### 5. redditmisspelling

```

15 • select *
16   from redditmisspelling;

```

<

Result Grid   Filter Rows:  Edit:

	Post_ID	Misspelt_Word	Correct_Word	ID
▶	bf1q01	travelling	traveling	1
*	NULL	NULL	NULL	NULL




#### 6. redditsemanticinformation

```

18 • select *
19   from redditsemanticinformation;

```

<

Result Grid   Filter Rows:  Export:  Wr

	Category_ID	Category_Name
▶	transportoption	Transport Option
	travelplan	Travel Plan
	touristplace	Tourist Place
	touristplacetype	Tourist Place Type



#### 7. redditsynonym

```

21 • select *
22   from redditsynonym;

```

<

Result Grid   Filter Rows:

	Post_ID	Name	Description
▶	bdou78	Bro	Brother
*	NULL	NULL	NULL

#### 8. twitterdomaintag

Result Grid			
		Filter Rows:	
	Tag_Name	Tag_Occurence_Count	Category_ID
▶	Antonio	2	touristplace
	Atlanta	3	touristplace
	Atlanta,	2	touristplace
	Boston	2	touristplace
	Chicago	3	touristplace
	Chicago,	4	touristplace
	D.C.	4	touristplace
	Dallas	4	touristplace
	Denver	4	touristplace
	Georgia,	2	touristplace
	Louis	2	touristplace
	Maine	2	touristplace
	Miami	4	touristplace
	Milan	2	touristplace
	Nashville	2	touristplace
	Natchez	2	touristplace
	Nevada	2	touristplace
	Orleans	2	touristplace
	Philadelphia	4	touristplace
	Seattle	2	touristplace

#### 9. twittermisspelling

Result Grid				
		Filter Rows:		
	Misspelt_Word	Correct_Word	Tweet_ID	ID
▶	ya	you	1121771468495491072	1
*	NULL	NULL	NULL	NULL

#### 10. twittersemanticinformation

Result Grid		
		Filter Rows:
	Category_ID	Category_Name
▶	Category_ID	Category_Name
	touristplace	Tourist Place
*	NULL	NULL

#### 11. twittersynonym

Result Grid	Filter Rows:	Edit:
Name	Description	Tweet_ID
D.C	District of Columbia	1121771304829505536
NULL	NULL	NULL

## 6.3 General Use Cases:

### Use case 1:

List of tourist attractions based on city and attraction type:

```
create procedure sp_attraction_city_type
(@cityid nvarchar(50),
@entrytype nvarchar(50))
as
begin
```

```
select c.city_name,ta.attraction_name,ta.attraction_type,ta.entry_type,ta.attraction_id,
c.state_id
from touristattraction ta
join city c
on ta.city_id=c.city_id
where c.city_name=@cityid and ta.entry_type=@entrytype;
end
```

output:

```
execute sp_attraction_city_type 'boston','free';
```

	city_name	attraction_name	attraction_type	entry_type	Time_To_Spend	attraction_id	state_id
1	Boston	Boston Common	Recreation	Free	1 to 2 hours	bostoncommon	massachusetts
2	Boston	Boston Public Garden	Recreation	Free	1 to 2 hours	bostonpublicgarden	massachusetts
3	Boston	Faneuil Hall Marketplace	shopping	Free	1 to 2 hours	faneuilhallmarketplace	massachusetts
4	Boston	Freedom Trail	Sughtseeing	Free	2 hours to half day	freedomtrail	massachusetts
5	Boston	Samuel Adams Brewery	Wineries/Breweries	Free	1 to 2 hours	samueladamsbrewery	massachusetts

### Use case 2:

Display weather of a city for a particular date

```
create procedure sp_weather_city_date
@cityname nvarchar(50),
@date datetime2(7)
```

```

as
begin
select
c.City_Name,w.Date,w.Temp_C_,w.Weather_Condition,w.Minimum_Temperature_C_,w.Maximum_Temperature
_C_,
w.Pressure_hPa_,w.Humidity___,w.Cloudiness___,w.Wind_miles_hr_
from weather w
join city c
on c.City_ID=w.City_id
where c.City_Name=@cityname and date=@date;
end

```

Output:

execute sp\_weather\_city\_date 'boston','2019-04-28 03:00:00'

Results		Messages								
	City_Name	Date	Temp_C_	Weather_Condition	Minimum_Temperature_C_	Maximum_Temperature_C_	Pressure_hPa_	Humidity___	Cloudiness___	Wind_miles_hr_
1	Boston	2019-04-28 03:00:00.0000000	5.55	broken clouds	5.55	5.55	1003.01	66	77	9.73095

### Use case 3:

Display days with harsh weather based on city input.

```

create procedure sp_city_harsh_weatner
@cityname nvarchar(50)
as
begin
select c.city_name, w.date,w.weather_condition, w.temp_c_
from weather w
join city c
on c.City_ID=w.City_id
where c.City_Name='boston' and w.Weather_Condition like '%rain%' or w.Weather_Condition like '%snow';
end

```

Output:

exec sp\_city\_harsh\_weatner ' boston '

	city_name	date	weather_condition	temp_c_
1	Boston	2019-04-27 00:00:00.0000000	light rain	13.79
2	Boston	2019-04-26 21:00:00.0000000	light rain	11.25
3	Boston	2019-04-26 18:00:00.0000000	light rain	10.05
4	Boston	2019-04-26 15:00:00.0000000	moderate rain	8.62
5	Boston	2019-04-30 21:00:00.0000000	light rain	13.06
6	Boston	2019-04-26 12:00:00.0000000	light rain	7.95
7	Boston	2019-04-29 00:00:00.0000000	light rain	6.58
8	Boston	2019-04-29 03:00:00.0000000	light rain	4.21
9	Boston	2019-04-30 15:00:00.0000000	light rain	6.05
10	Boston	2019-04-27 18:00:00.0000000	light rain	12.86
11	Boston	2019-04-30 18:00:00.0000000	light rain	9.64
12	Boston	2019-04-27 03:00:00.0000000	light rain	15.65
13	Boston	2019-04-27 06:00:00.0000000	light rain	14.86

#### Use case 4:

Display list of cheapest hotels for a city with approximate price/night.

```
create procedure sp_hotels_cheapest
```

```
@city nvarchar(50)
```

```
as
```

```
begin
```

```
select c.city_name,h.hotel_name,h.rating,h.total_reviewers,h.price_per_night,a.building_number,
a.street_name,h.landmark
```

```
from hotel h
```

```
join address a
```

```
on h.Address_Id=a.Address_Id
```

```
join city c
```

```
on a.City_Id = c.City_ID
```

```
where c.City_Name=@city
```

```
order by h.PRICE_PER_NIGHT;
```

```
end
```

Output:

```
exec sp_hotels_cheapest 'new york city'
```

	city_name	hotel_name	rating	total_reviewers	price_per_night	building_number	street_name	landmark
1	New york City	YOTEL New York	4.5	12,471 reviews	NULL	570	Tenth Avenue	0.5 miles to City center
2	New york City	Hotel Newton	4	1,720 reviews	128	2528	Broadway	2.7 miles to City center
3	New york City	The Watson Hotel	3.5	4,917 reviews	185	440	West th Street	0.8 miles to City center
4	New york City	Wyndham Garden Chinatown	4.5	1,722 reviews	203	93	Bowery	2.8 miles to City center
5	New york City	Crowne Plaza JFK Airport New York City	3.5	923 reviews	231	138	-th Ave	12 miles to City center
6	New york City	Springhill Suites New York Manhattan/Times Squa...	no rating available	No review available	269	338	West th Street	0.4 miles to City center
7	New york City	The Iroquois New York	4.5	4,406 reviews	352	49	West th Street	0.2 miles to City center
8	New york City	PUBLIC, an Ian Schrager hotel	4.5	702 reviews	365	215	Chrystie St	2.4 miles to City center
9	New york City	Gardens Suites Hotel by Affinia	4.5	2,174 reviews	403	215	E th St	1.3 miles to City center

#### Use case 5:

Display flight schedules for origin, destination, month and day input

```
create procedure sp_flight_schedule
@origincity nvarchar(50),
@destination nvarchar(50),
@month int,
@day int
as
begin

select a.airline_name, f.departure_time,f.arrival_time,f.distance,f.Air_Time,
f.Origin_City_ID,f.Destination_City_ID
from flightschedule f
join airline a
on a.airline_id=f.Airline_ID
join city c
on c.City_ID=f.Destination_City_ID
where f.Destination_City_ID=@destination and f.Month_Of_The_Year=@month and f.Day_Of_The_Month=@day
and f.Origin_City_ID=@origincity
end
```

Output:

Exec sp\_flight\_schedule 'boston', 'chicago',4,6

Results		Messages					
	airline_name	departure_time	arrival_time	distance	Air_Time	Origin_City_ID	Destination_City_ID
1	Southwest Airlines	2019-04-26 17:16:00.0000000	2019-04-26 18:57:00.0000000	861	123	boston	chicago
2	United Airlines	2019-04-26 09:00:00.0000000	2019-04-26 10:36:00.0000000	867	128	boston	chicago
3	American Airlines	2019-04-26 15:10:00.0000000	2019-04-26 16:50:00.0000000	867	130	boston	chicago
4	American Airlines	2019-04-26 13:55:00.0000000	2019-04-26 15:35:00.0000000	867	131	boston	chicago
5	American Airlines	2019-04-26 09:10:00.0000000	2019-04-26 10:45:00.0000000	867	131	boston	chicago
6	United Airlines	2019-04-26 10:37:00.0000000	2019-04-26 12:18:00.0000000	867	NULL	boston	chicago
7	American Airlines	2019-04-26 07:50:00.0000000	2019-04-26 09:25:00.0000000	867	131	boston	chicago
8	American Airlines	2019-04-26 12:50:00.0000000	2019-04-26 14:25:00.0000000	867	128	boston	chicago
9	American Airlines	2019-04-26 11:00:00.0000000	2019-04-26 12:35:00.0000000	867	132	boston	chicago
10	United Airlines	2019-04-26 07:03:00.0000000	2019-04-26 08:41:00.0000000	867	135	boston	chicago

#### **Use case 6:**

Display ratings of services offered by an airline.

```
create procedure sp_airline_rating
@airline nvarchar(50)
as
begin
```



```

select
airline_name,food_and_beverage_rating,Inflight_Entertainment_Rating,Seat_Comfort_Rating,Staff_Service_Ratin
g
,Value_For_Money_Rating
from airline
where airline_name=@airline
end

```

Output:

```
exec sp_airline_rating 'american airlines'
```

Results		Messages				
	airline_name	food_and_beverage_rating	Inflight_Entertainment_Rating	Seat_Comfort_Rating	Staff_Service_Rating	Value_For_Money_Rating
1	American Airlines	2	2	2	2	2

## Functions:

### Function 1:

Function to convert Celsius to Fahrenheit

```

CREATE FUNCTION dbo.temp (@temp float)
RETURNS float
AS
BEGIN
declare @tempF float
SET @tempF = ( @temp * (9/5))+32

RETURN @tempF
END

```

### Output:

```

select city_id,date,dbo.temp(temp_c_) as [Temperature F],Weather_Condition
from weather
where City_id='boston'

```

	city_id	date	Temperature F	Weather_Condition
1	boston	2019-04-26 00:00:00.0000000	39.18	broken clouds
2	boston	2019-04-27 03:00:00.0000000	47.65	light rain
3	boston	2019-04-27 06:00:00.0000000	46.86	light rain
4	boston	2019-04-27 09:00:00.0000000	42.55	light rain
5	boston	2019-04-27 12:00:00.0000000	42.05	overcast clouds
6	boston	2019-04-27 15:00:00.0000000	44.66	light rain
7	boston	2019-04-27 18:00:00.0000000	44.86	light rain
8	boston	2019-04-27 21:00:00.0000000	44.37	light rain
9	boston	2019-04-28 00:00:00.0000000	39.03	light rain
10	boston	2019-04-28 03:00:00.0000000	37.55	broken clouds
11	boston	2019-04-28 06:00:00.0000000	37.71	broken clouds
12	boston	2019-04-26 03:00:00.0000000	37.23	broken clouds

### **Function 2:**

Function to display Price in rupees and as well as into float data type.

```
CREATE FUNCTION dbo.price_conversion (@price float)
RETURNS float
AS
BEGIN
declare @rupees float
declare @p float
set @p = CONVERT(float,@price)
SET @rupees = @p * 69.85

RETURN @rupees
END
```

### **Output:**

```
select city_name,hotel_name,dbo.price_conversion(Price_per_night) as [Price in Rupees],
rating
from hotel h
join address a
on h.Address_Id=a.Address_Id
join city c
on a.City_Id=c.City_ID
where c.City_Name='new york city'
```

	city_name	hotel_name	Price in Rupees	rating
1	New york City	Crowne Plaza JFK Airport New York City	16135.35	3.5
2	New york City	PUBLIC, an Ian Schrager hotel	25495.25	4.5
3	New york City	YOTEL New York	NULL	4.5
4	New york City	Wyndham Garden Chinatown	14179.55	4.5
5	New york City	Springhill Suites New York Manhattan/Times Squa...	18789.65	no rating available
6	New york City	The Watson Hotel	12922.25	3.5
7	New york City	Gardens Suites Hotel by Affinia	28149.55	4.5
8	New york City	The Iroquois New York	24587.2	4.5
9	New york City	Hotel Newton	8940.8	4

### **Function 3:**

Auto-generation of email id for an airline.

```
alter function dbo.emailairline(@airline nvarchar(50),@airlineid nvarchar(50))
returns nvarchar(50)
as
begin
declare @emailid nvarchar(50)
declare @name nvarchar(50)
declare @aid nvarchar(50)

set @name = replace(@airline, ' ','_')
```

```

set @emailid = @name + '@airways.com'
return @emailid
end

```

#### **Output:**

	airline_name	airline_id	Email
1	Endeavor Air	9E	Endeavor_Air@airways.com
2	American Airlines	AA	American_Airlines@airways.com
3	Aloha	AQ	Aloha@airways.com
4	Alaska Airlines	AS	Alaska_Airlines@airways.com
5	JetBlue Airways	B6	JetBlue_Airways@airways.com
6	Continental Airlines	CO	Continental_Airlines@airways.com
7	Atlantic Coast Airlines	DH	Atlantic_Coast_Airlines@airways.com
8	Delta Air Lines	DL	Delta_Air_Lines@airways.com
9	Atlantic Southeast Airlines	EV	Atlantic_Southeast_Airlines@airways.com
10	Frontier	F9	Frontier@airways.com
11	AirTran Airways	FL	AirTran_Airways@airways.com
12	Hawaiian Airlines	HA	Hawaiian_Airlines@airways.com

#### **Function 4:**

```

create function dbo.attractioncountfree(@city nvarchar(50))
returns int
as
begin
declare @count int
set @count = (select count(attraction_id) from touristattraction where City_ID=@city )
return @count
end

```

#### **Output:**

```

select city_name,dbo.attractioncountfree(city_id) as {Number of tourist attractions}from
city
where city_name='boston'

```

	city_name	Number of tourist attractions
1	Boston	10

#### **Function 5:**

Calculate the number of Instagram posts for a city

```

create function dbo.insta_posts_count (@cityid nvarchar(50))
returns int

```

```

as
begin
declare @count int
set @count = (select count(post_id) from instagrampost where City_id=@cityid)
return @count
end

```

#### **Output:**

```

select city_name,dbo.insta_posts_count (city_id) as [Number of Instagram Posts]
from city where city_name='chicago'

```

	city_name	Number of Instagram Posts
1	Chicago	10

#### **Function 6:**

Display month name from month number.

```

Create function dbo.month_convert(@month int)
returns nvarchar(20)
as
begin
declare @m nvarchar(50)
set @m = (Select DateName( month , DateAdd( month , @month , -1 ) ))
return @m
end

```

#### **Output:**

```

Select month_of_the_year,dbo.month_convert(month_of_the_year) as Month
from flightschedule
where Origin_City_ID = 'boston' and Destination_City_ID='chicago'

```

	month_of_the_year	Month
1	5	May
2	6	June
3	4	April
4	5	May
5	4	April
6	5	May
7	5	May
8	5	May
9	5	May
10	4	April
11	6	June
12	6	June
13	6	June

## **Views:**

### View 1:

Displays city and state name

```
create view view1
as select city_name,State_Name
from city c join state s
on c.City_ID=s.State_ID;
```

### Output:

```
select * from view1
```

	city_name	State_Name
1	Virginia	Virginia
2	Montana	Montana

### View 2:

Display hashtags of a particular user

```
create view view2
as select user_name,hash_tag
from instagramhashtag ih
join instagrampost i
on i.Post_id = ih.Post_id
where User_Name='chandrebo'
```

### Output:

```
select * from view2
```

	user_name	hash_tag
1	chandrebo	tupac
2	chandrebo	newyork
3	chandrebo	indiana
4	chandrebo	texas
5	chandrebo	arizona
6	chandrebo	ohio
7	chandrebo	weed
8	chandrebo	420
9	chandrebo	producer
10	chandrebo	beats
11	chandrebo	smoke
12	chandrebo	weed

### View 3:

Display hotel name along with its complete address

```
alter view view3 as
```

```

select hotel_name,building_number,street_name,city_name,state_name
from hotel h join
address a
on a.Address_Id=h.Address_Id
join city c
on c.City_ID = a.City_Id
join state s on a.State_id=s.State_ID

```

#### Output:

	hotel_name	building_number	street_name	city_name	state_name
1	Crowne Plaza JFK Airport New York City	138	-th Ave	New york City	New York
2	PUBLIC, an Ian Schrager hotel	215	Chrystie St	New york City	New York
3	YOTEL New York	570	Tenth Avenue	New york City	New York
4	Wyndham Garden Chinatown	93	Bowery	New york City	New York
5	Springhill Suites New York Manhattan/Times Squa...	338	West th Street	New york City	New York
6	The Watson Hotel	440	West th Street	New york City	New York
7	Gardens Suites Hotel by Affinia	215	E th St	New york City	New York
8	The Iroquois New York	49	West th Street	New york City	New York
9	Hotel Newton	2528	Broadway	New york City	New York

#### View 4:

Display free attractions of a city

```
create view view4 as
```

```

select c.city_name,ta.attraction_name,ta.attraction_type,ta.entry_type
,ta.Time_To_Spend,ta.attraction_id,c.state_id
from city c
join touristattraction ta
on ta.city_id=c.city_id
where c.city_name='boston' and ta.entry_type='free'

```

#### Output:

	city_name	attraction_name	attraction_type	entry_type	Time_To_Spend	attraction_id	state_id
1	Boston	Boston Common	Recreation	Free	1 to 2 hours	bostoncommon	massachusetts
2	Boston	Boston Public Garden	Recreation	Free	1 to 2 hours	bostonpublicgarden	massachusetts
3	Boston	Faneuil Hall Marketplace	shopping	Free	1 to 2 hours	faneuilhallmarketplace	massachusetts
4	Boston	Freedom Trail	Sughtseeing	Free	2 hours to half day	freedomtrail	massachusetts
5	Boston	Samuel Adams Brewery	Wineries/Breweries	Free	1 to 2 hours	samueladamsbrewery	massachusetts

	city_name	attraction_name	attraction_type	entry_type	Time_To_Spend	attraction_id	state_id
1	Boston	Boston Common	Recreation	Free	1 to 2 hours	bostoncommon	massachusetts
2	Boston	Boston Public Garden	Recreation	Free	1 to 2 hours	bostonpublicgarden	massachusetts
3	Boston	Faneuil Hall Marketplace	shopping	Free	1 to 2 hours	faneuilhallmarketplace	massachusetts
4	Boston	Freedom Trail	Sughtseeing	Free	2 hours to half day	freedomtrail	massachusetts
5	Boston	Samuel Adams Brewery	Wineries/Breweries	Free	1 to 2 hours	samueladamsbrewery	massachusetts

## **9.Audit Validity**

The whole dataset had more information than required for the proposed conceptual database model. To fit in to the database schema, the dataset was reformatted. Some of the values were in '-', which were replaced by 0. A few of the columns have been renamed for the purpose of making it simple and understandable.

## **10. Audit Completeness**

The dataset is up to date, needs no more cleaning and matches the quality of the real world data. The data was cross verified with other sources.

## **11. Audit Consistence/Uniformity:**

The possible range of the dataset is covered from the new resultant data set. The data does not have any null values, limitations, negative values.

Also the values appear to remain constant with what is represented in the original source.

## **12.References & Citations**

- [https://github.com/nikbearbrown/INFO\\_6210](https://github.com/nikbearbrown/INFO_6210).
- YouTube videos: ▪ <https://www.youtube.com/watch?v=4UcqECQe5Kc>
- <https://www.youtube.com/watch?v=PWZKTWJ9bJE>
- SQL concepts: ▪ w3schools.com ▪ essentialsql.com.
- Kaggle.com
- Hotelscombined.com
- Openweathermap.com
- Twitter
- Reddit
- Instagram

## **13. License**

---

Copyright <2019> <Rajendra Kumar Rajkumar><Monish Hirisave Raghu> <Prudhvi Sekharamahanti>

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 CONNECTION WITH THE SOFTWARE OR THE USE OR OTHER DEALINGS IN THE SOFTWARE.