

INTRODUCTION

We are a marketing company located in Dubai, UAE. We took the exclusive rights to use Dubai metro stations to advertise our customers.

PROBLEM

As our focus group is local business owners and venues, we need to define our target customers based on their location near every metro station.

Not all the metro stations areas are similar to each other's when we check its business hot areas, as some areas have more hotels, other areas have shopping malls or parks ... etc.

Some metro stations are busier than others, so this should be considered in the pricing of the advertise especially in the peak hours of the station.

AS an expansion of our business areas as a company, in co-operation with Foursquare.com , we will start building a recommendation system based on venues visited in the station's area.

ASSUMPTION

We will use some open datasets and APIs to collect data about metro stations and the nearby venues and targeted customers.

Moreover, we could collect some useful data about the traffic transaction of the metro stations, thanks to Dubaipulse.com.

Once we have our data ready, we will start categorizing the venues to reduce the number of business areas categories.

Then, we will build some statistics on the available data to satisfy our business needs and help the decision makers to create the right strategy.

Finally, we will start building our user recommendation system using foursquare.com APIs.

DATA SOURCES

Finding open datasets now becomes easier than before, we used one dataset provided by Dubai government and another one found on Github.com then we used a public API from Foursquare.com.

1. DUBAI NEIGHBORHOODS DATASET

ABOUT DATASET

A dataset which contains neighborhood name along with its latitude and longitude.

SOURCE

[Github.com](#)

SAMPLE RESPONSE

	name	categories	lat	lng
0	The Greens الروضة	Neighborhood	25.092828	55.170851
1	Nad Al Sheba 2 د الشبا 2	Neighborhood	25.160512	55.346228
2	Burj Al Arab Open Beach	Beach	25.147411	55.193854
3	Arabian Ranches 2	Neighborhood	25.035080	55.267104
4	Nadd AlHammar Park	Park	25.201799	55.380516

2. DUBAI METRO STATIONS DATASET

ABOUT DATASET

A dataset which contains metro stations name and metro line along with its latitude and longitude.

SOURCE

[Dubaipulse.gov.ae](#)

SAMPLE RESPONSE

	station	line_name	longitude	latitude
0	World Trade Centre Metro Station	Red Metro Line	55.285061	25.224829
1	Emirates Towers Metro Station	Red Metro Line	55.279821	25.217214
2	Financial Centre Metro Station	Red Metro Line	55.275587	25.211030
3	Burj Khalifa/ Dubai Mall Metro Station	Red Metro Line	55.269518	25.201401
4	Business Bay Metro Station	Red Metro Line	55.260419	25.191275

3.NEARBY VENUES API

ABOUT API

Returns a list of recommended venues near the current location.

SOURCE

[Foursquare.com](https://foursquare.com)

SAMPLE RESPONSE

```
{
  "meta": {
    "code": 200,
    "requestId": "5ac51ef86a607143de8eg5cb"
  },
  "response": {
    "warning": {
      "text": "There aren't a lot of results near you. Try something more general, reset your filters, or expand the search area."
    },
    "suggestedRadius": 600,
    "headerLocation": "Lower East Side",
    "headerFullLocation": "Lower East Side, New York",
    "headerLocationGranularity": "neighborhood",
    "totalResults": 230,
    "suggestedBounds": {
      "ne": {
        "lat": 40.724216906965616,
        "lng": -73.9896507407283
      },
      "sw": {
        "lat": 40.71252743651852,
        "lng": -74.012484375
      }
    }
  }
}
```

```

        "lat":40.72151724718017,
        "lng":-73.98693222860872
    }
},
"groups": [
{
    "type":"Recommended Places",
    "name":"recommended",
    "items": [
    {
        "reasons":{
            "count":0,
            "items": [
            {
                "summary":"This spot is popular",
                "type":"general",
                "reasonName":"globalInteractionReason"
            }
            ]
        },
        "venue": {
            "id":"49b6e8d2f964a52016531fe3",
            "name":"Russ & Daughters",
            "location": {
                "address":"179 E Houston St",
                "crossStreet":"btwn Allen & Orchard St",
                "lat":40.72286707707289,
                "lng":-73.98829148466851,
                "labeledLatLngs": [
                {
                    "label":"display",
                    "lat":40.72286707707289,
                    "lng":-73.98829148466851
                }
                ],
                "distance":130,
                "postalCode":"10002",
                "cc":"US",
                "city":"New York",
                "state":"NY",
                "country":"United States",
                "formattedAddress": [
                    "179 E Houston St (btwn Allen & Orchard St"
                )
            }
        }
    }
}
]
}

```

```
        "New York, NY 10002",
        "United States"
    ],
},
"categories": [
{
    "id":"4bf58dd8d48988d1f5941735",
    "name":"Gourmet Shop",
    "pluralName":"Gourmet Shops",
    "shortName":"Gourmet",
    "icon":{
        "prefix":"https://ss3.4sqi.net/img/categories_v2/shops/food_gourmet_",
        "suffix":".png"
    },
    "primary":true
},
],
"popularityByGeo":0.9999983845502491,
"venuePage":{
    "id":"77298563"
}
}
]
}
}
```

4. VENUES' LIKES API

ABOUT API

Get Users Who Liked a Venue

SOURCE

Foursquare.com

SAMPLE RESPONSE

```
{  
  "meta": {  
    "code": 200,  
    "requestId": "5854c1f8351e3d5887252266"  
  },  
  "notifications": [  
    {  
      "item": {  
        "unreadCount": 13  
      },  
      "type": "notificationTray"  
    }  
  ],  
  "response": {  
    "like": false,  
    "likes": {  
      "count": 28,  
      "groups": [  
        {  
          "count": 28,  
          "items": [  
  
            ],  
          "type": "others"  
        }  
      ],  
      "summary": "28 Likes"  
    }  
  }  
}
```

5.MONTHLY TRIPS NUMBERS PER METRO STATION DATASET

ABOUT DATASET

Total number of the monthly trips per metro station
SOURCE

Dubaipulse.gov.ae

(One dataset per month, total number 9 datasets)

SAMPLE RESPONSE

	year	month	metro_station	trips
0	2019	Aug	Abu Baker Al Siddique Metro Station	316825
1	2019	Aug	Abu Hail Metro Station	230371
2	2019	Aug	ADCB Metro Station	489146
3	2019	Aug	Airport Terminal 1 Metro Station	231888
4	2019	Aug	Airport Terminal 3 Metro Station	144990

6.A RANDOM WORKING DAY ALL STATIONS CHECK-IN AND CHECK-OUT TRANSACTIONS DATASET

ABOUT DATASET

All stations transactions in 27th March 2018 (Normal working day)
SOURCE

Dubaipulse.gov.ae

SAMPLE RESPONSE

	txn_type	txn_date	txn_time	station
0	Check in	2018-03-27	23:05:19	Burj Khalifa/ Dubai Mall Metro Station
1	Check in	2018-03-27	21:32:43	Burj Khalifa/ Dubai Mall Metro Station
2	Check in	2018-03-27	19:28:22	Burj Khalifa/ Dubai Mall Metro Station
3	Check in	2018-03-27	22:26:32	Burj Khalifa/ Dubai Mall Metro Station
4	Check in	2018-03-27	14:06:16	Burj Khalifa/ Dubai Mall Metro Station

ANALYSIS AND FINDINGS

First of all, we listed the metro stations that our analysis will be based on

- World Trade Centre Metro Station
- Emirates Towers Metro Station
- Financial Centre Metro Station
- Burj Khalifa/ Dubai Mall Metro Station
- Business Bay Metro Station
- Noor Bank Metro Station
- First Abu Dhabi Bank Metro Station
- Mall of the Emirates Metro Station
- Sharaf DG Metro Station
- Dubai Internet City Metro Station
- Nakheel Metro Station
- DAMAC Properties Metro Station
- DMCC Metro Station
- Nakheel Harbour & Tower Metro Station
- Ibn Battuta Metro Station
- Energy Metro Station
- Danube Metro Station
- UAE Exchange Metro Station
- Etisalat Metro Station
- Al Qusais Metro Station
- Dubai Airport Free Zone Metro Station

- Al Nahda Metro Station
- Stadium Metro Station
- Al Qiyadah Metro Station
- Abu Hail Metro Station
- Abu Baker Al Siddique Metro Station
- Salah Al Din Metro Station
- Union (Green Line) Metro Station
- Baniyas Square Metro Station
- Palm Deira Metro Station
- Al Ras Metro Station
- Al Ghubaiba Metro Station
- Al Fahidi Metro Station
- BurJuman Metro Station
- Oud Metha Metro Station
- Dubai Healthcare City Metro Station
- Al Jadaf Metro Station
- Creek Metro Station
- Rashidya Depot
- Rashidiya Metro Station
- Emirates Metro Station
- Airport Terminal 3 Metro Station
- Airport Terminal 1 Metro Station
- GGICO Metro Station
- Deira City Centre Metro Station
- Al Rigga Metro Station
- Union Metro Station
- ADCB Metro Station
- Al Jafiliya Metro Station

NEARBY VENUES

Using the Explore API from foursquare, we could retrieve most of the nearby venues to the metro stations.

	Station	Venue ID	Venue Name	Category	Lat	Long.
0	World Trade Centre Metro Station	5c18d290da7080002c9a6e83	Opa Dubai	Greek Restaurant	25.226190	55.284894
1	World Trade Centre Metro Station	4d3170065c2db60c9518b86c	Conrad Dubai	Hotel	25.226089	55.283571

2	World Trade Centre Metro Station	4b0587eaf964a52090a622e3	Fairmont Dubai	Hotel	25.226206	55.284714
3	World Trade Centre Metro Station	564e368a498e0884971bfdc8	Bagatelle Dubai	French Restaurant	25.226147	55.284663

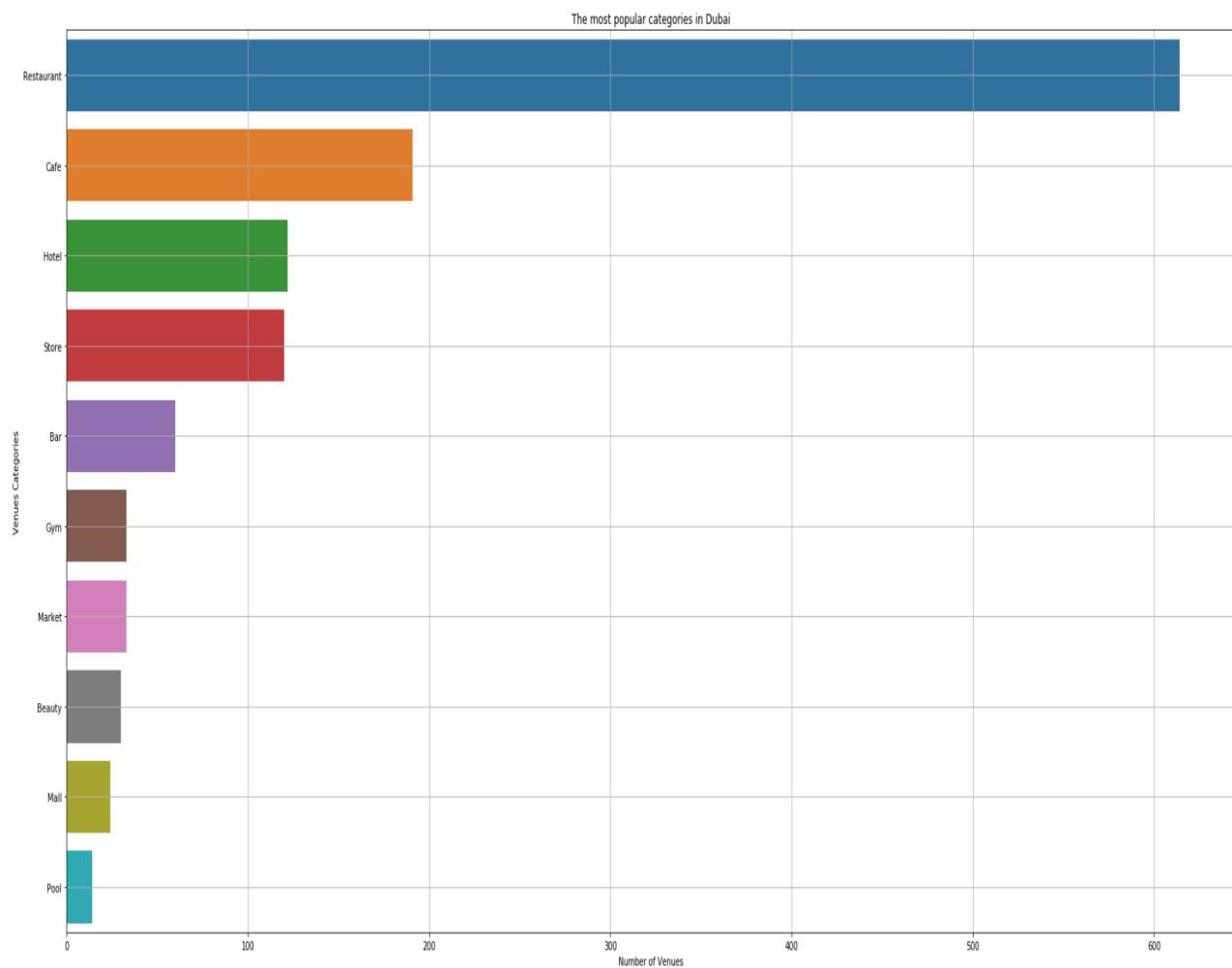
CATEGORIZING

In order to reduce the number of categories to help in the analysis process, we started to create a main category for the venues' categories so both of Greek Restaurant and French Restaurant will be “Restaurant”

	Station	Venue ID	Venue Name	Lat	Long	Category	Main Category
876	UAE Exchange Metro Station	5624b50d498e469eb77ddad9	Taqado Mexican Kitchen	24.980252	55.093491	Mexican Restaurant	Restaurant
857	UAE Exchange Metro Station	563095d6498ea09d5435fd44	Joga	24.980279	55.093656	Salad Place	Restaurant
1235	UAE Exchange Metro Station	536b469d498e308b4748fe87	1762	24.980698	55.093846	Deli / Bodega	Restaurant
258	UAE Exchange Metro Station	4e10288c8877ac23b655c6a4	Oregano	24.981092	55.093726	Italian Restaurant	Restaurant

338	UAE Exchange Metro Station	4e77053fe4cd32543800227e	Canteen L'Oreal ME	24.981446	55.092854	Restaurant	Restaurant
-----	----------------------------	--------------------------	--------------------	-----------	-----------	------------	------------

By grouping the Categories we found that the most popular business areas in Dubai are the following



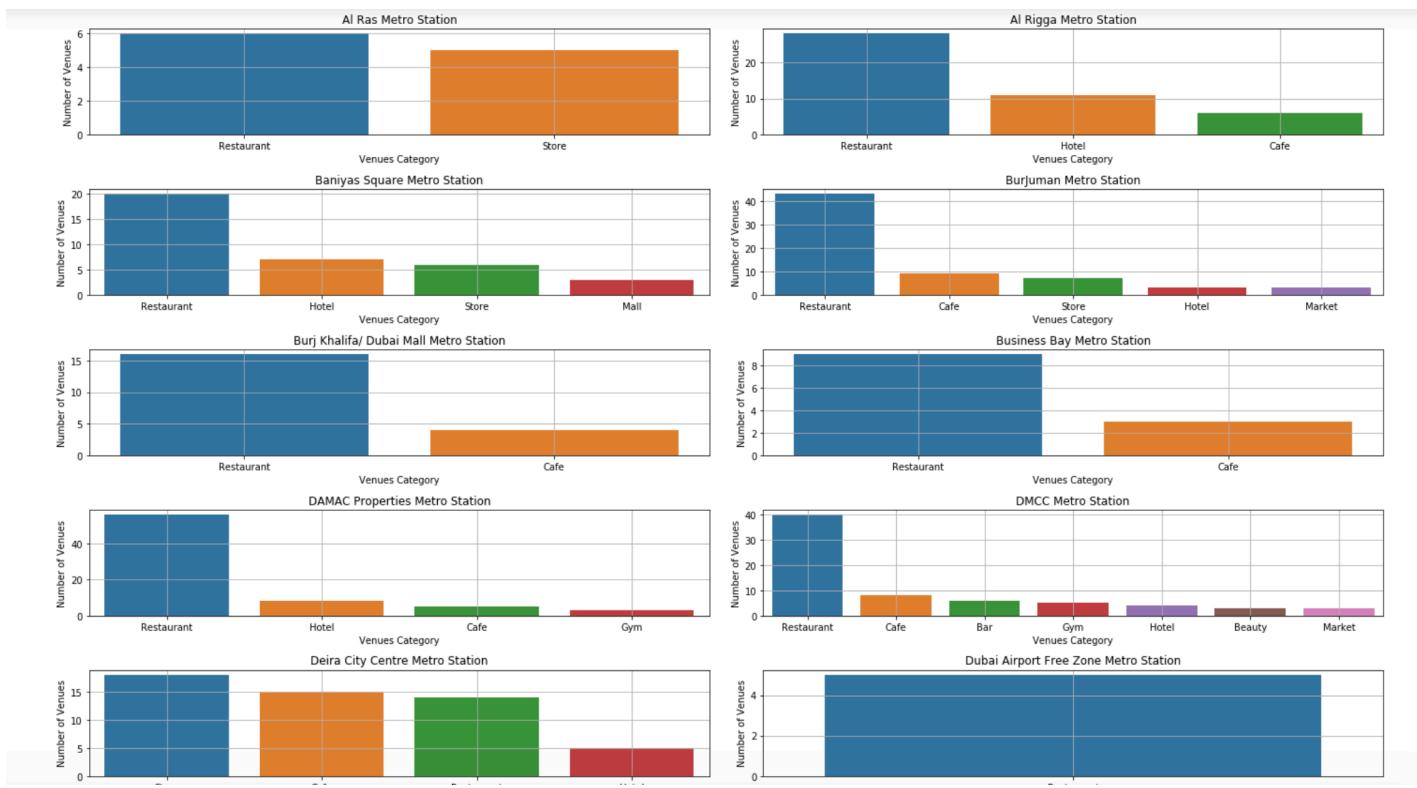
	Category	Number of Venues
0	Restaurant	614
1	Cafe	191
2	Hotel	122
3	Store	120
4	Bar	60
5	Gym	33
6	Market	33
7	Beauty	30
8	Mall	24
9	Pool	14

HOT BUSINESS AREAS

Restaurants are coming in the first place then Cafes, Hotel and General Stores. But is this popularity are shared among all the metro stations areas?

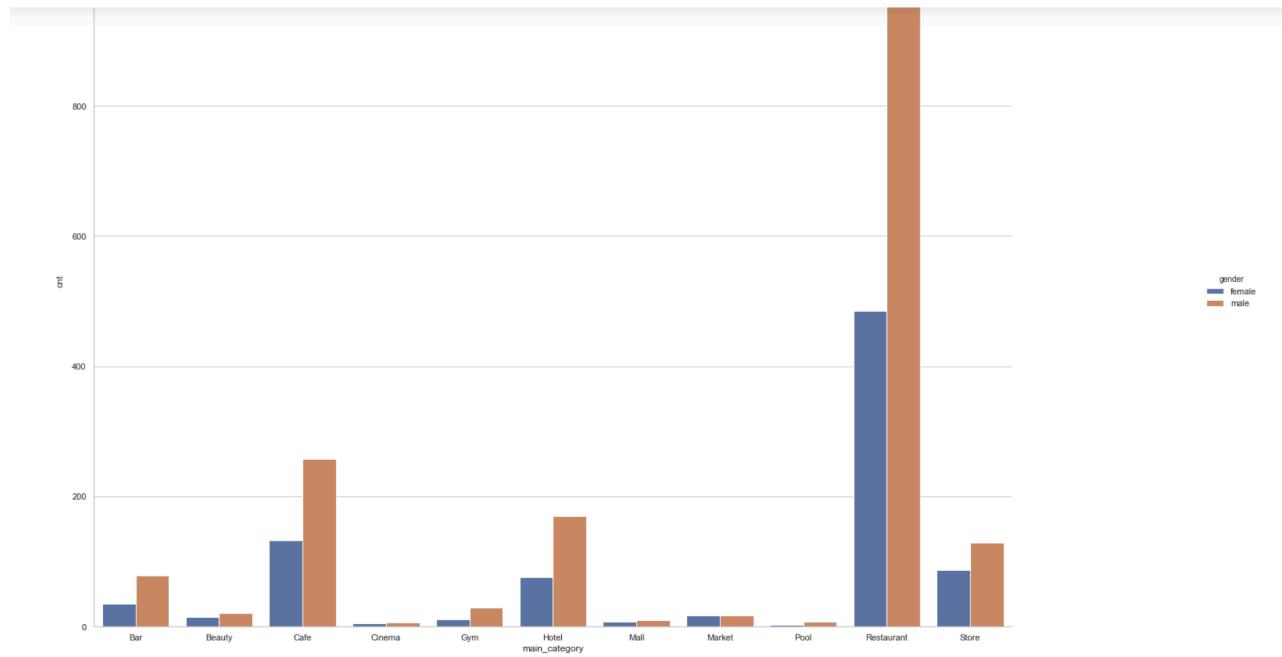
To have more insight to the metro station popular or hot business areas, we grouped both of Station and Main Category data and found the following sample:

	Station	Main Category	Number of Venues
0	ADCB Metro Station	Restaurant	27
1	ADCB Metro Station	Hotel	4
2	ADCB Metro Station	Cafe	3
3	Abu Baker Al Siddique Metro Station	Restaurant	11
4	Abu Baker Al Siddique Metro Station	Hotel	4



GENDER BASED CATEGORIZING

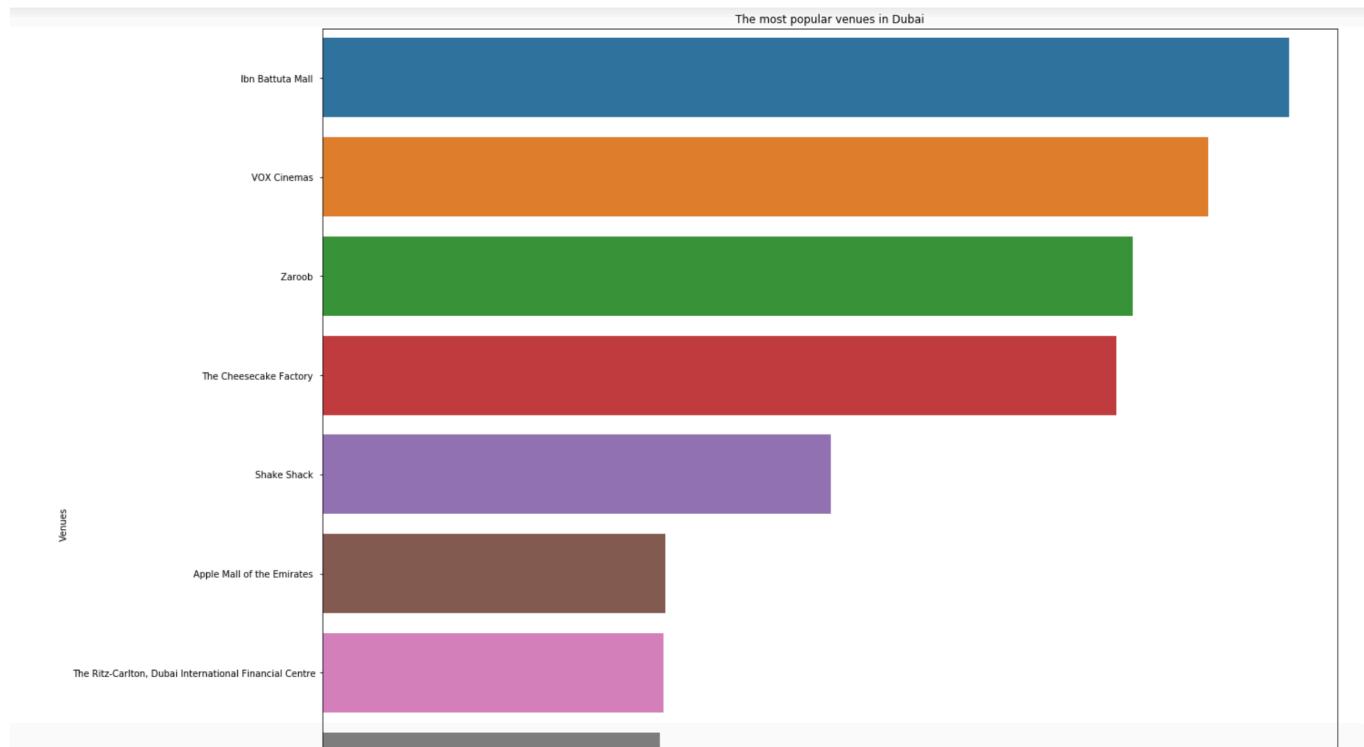
One more useful classification, checking if there some business categories are more preferred to a gender more than the other?



We found that venues like Restaurants, Cafes and Hotels have ore check-ins for males than females, but venues like Markets, Stores and Beauty related venues are almost have the same popularity for both of the genders.

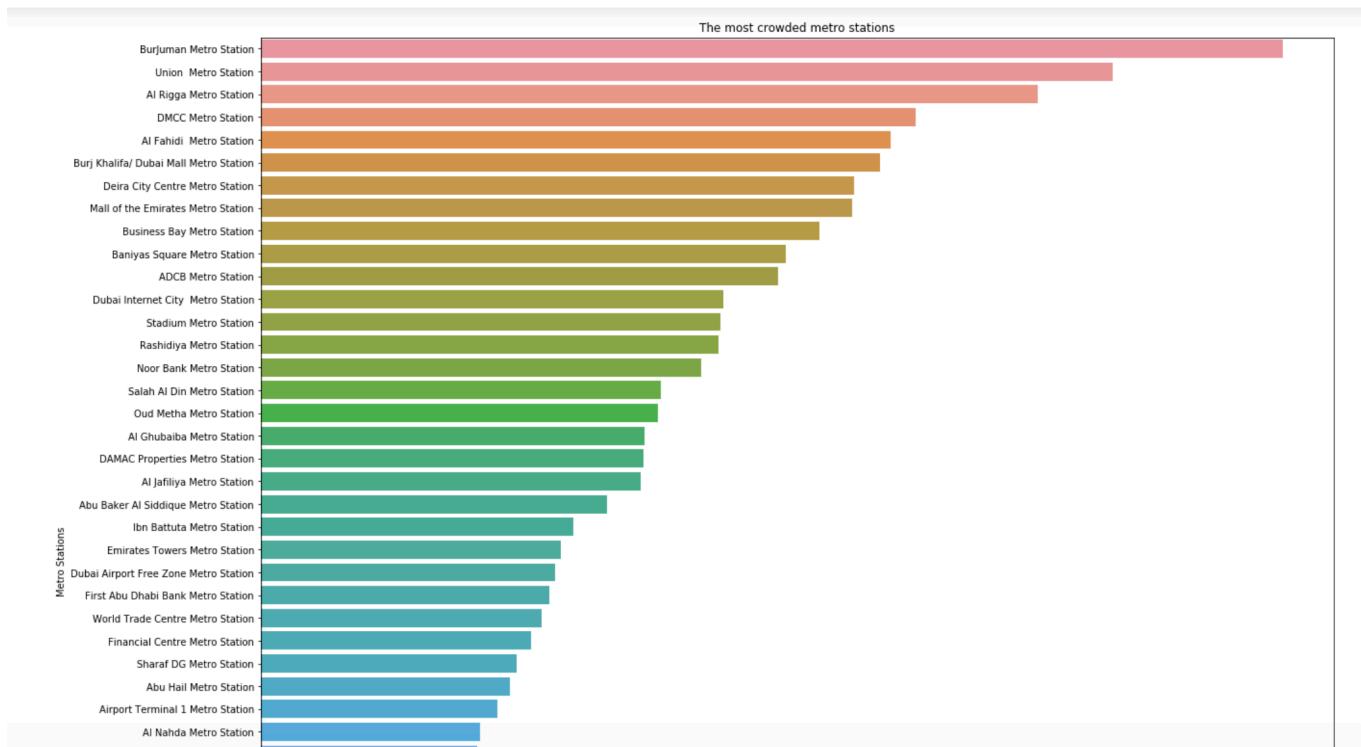
GOLD CLASS CUSTOMERS

In order to check the Gold class venues based on their popularity, we did the following statistic based on foursquare.com Likes API



STATIONS' TRIPS AND CLASSIFICATION

Picking the number of monthly transaction, we found that the most crowded metro stations can be as following



So based on the previous figure, we can categorize the metro stations from A to D based on the number of transactions.

Station	Number of Transactions from Jan 2019 to Aug 2019	Category
ADCB Metro Station	527789	B
Abu Baker Al Siddique Metro Station	353043	C
Abu Hail Metro Station	254428	D

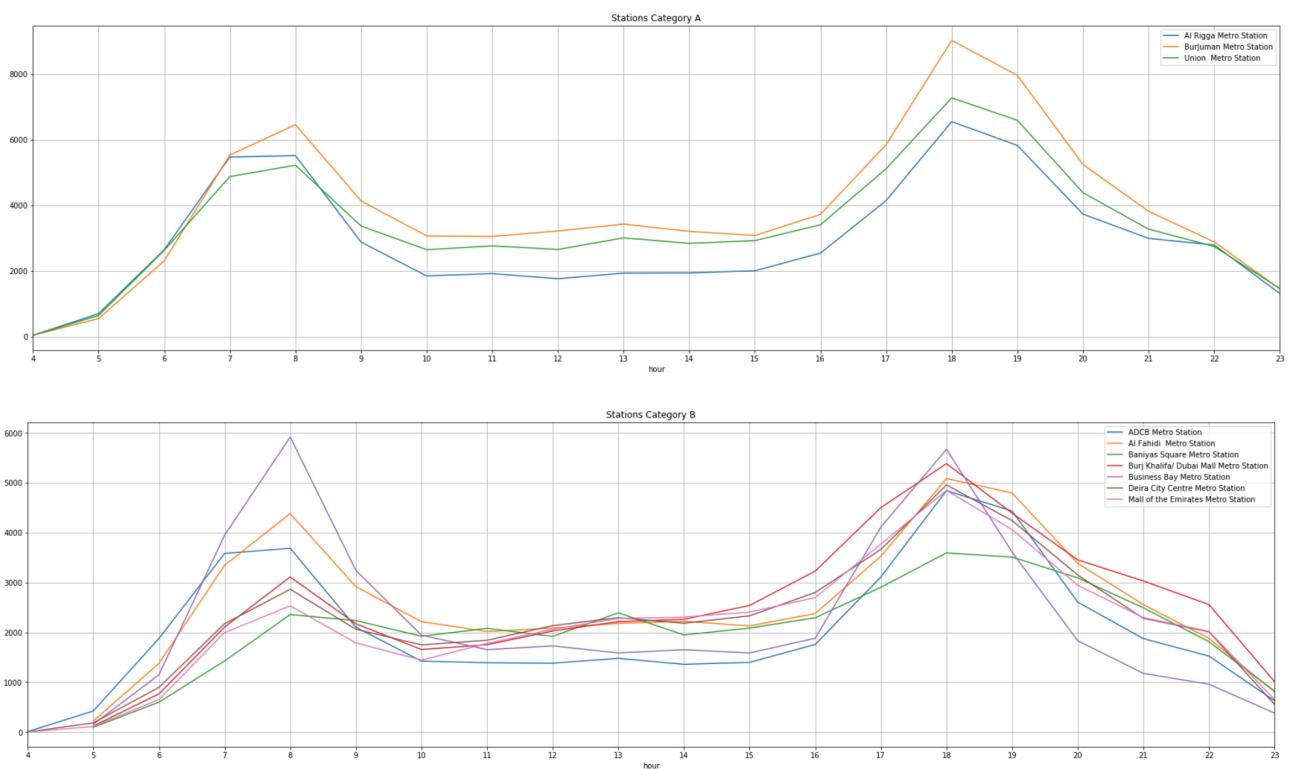
Station	Number of Transactions from Jan 2019 to Aug 2019	Category
Airport Terminal 1 Metro Station	241753	D
Airport Terminal 3 Metro Station	151413	D

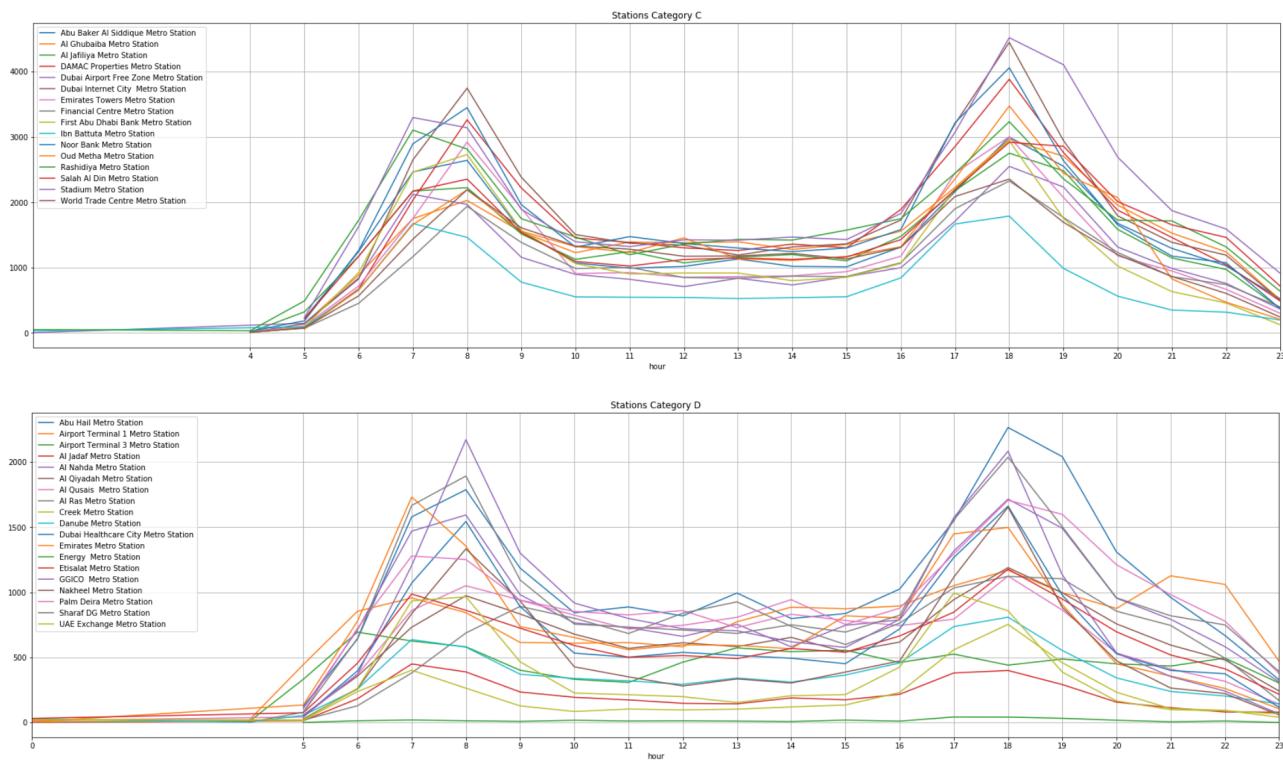
In [66]:

PEAK HOURS

Advertising price can't be the same comparing the peak hours with normal ones, so we need to calculate the peak hours of the metro stations in order to set the pricing for the digital boards' rental.

Picking a random working day's (27th March 2018) check-in and check-out transactions, we were able to create the following charts.

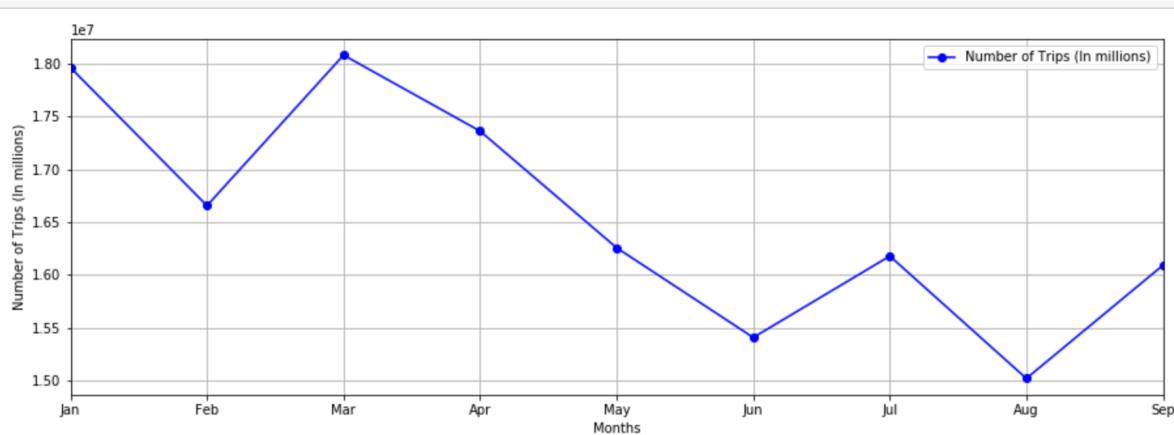




The peak hours for most of the stations are between 7:00 am: 8:00 am and 5:00 pm: 7:00 pm, so this could be considered in the pricing process.

HIGH/LOW SEASON

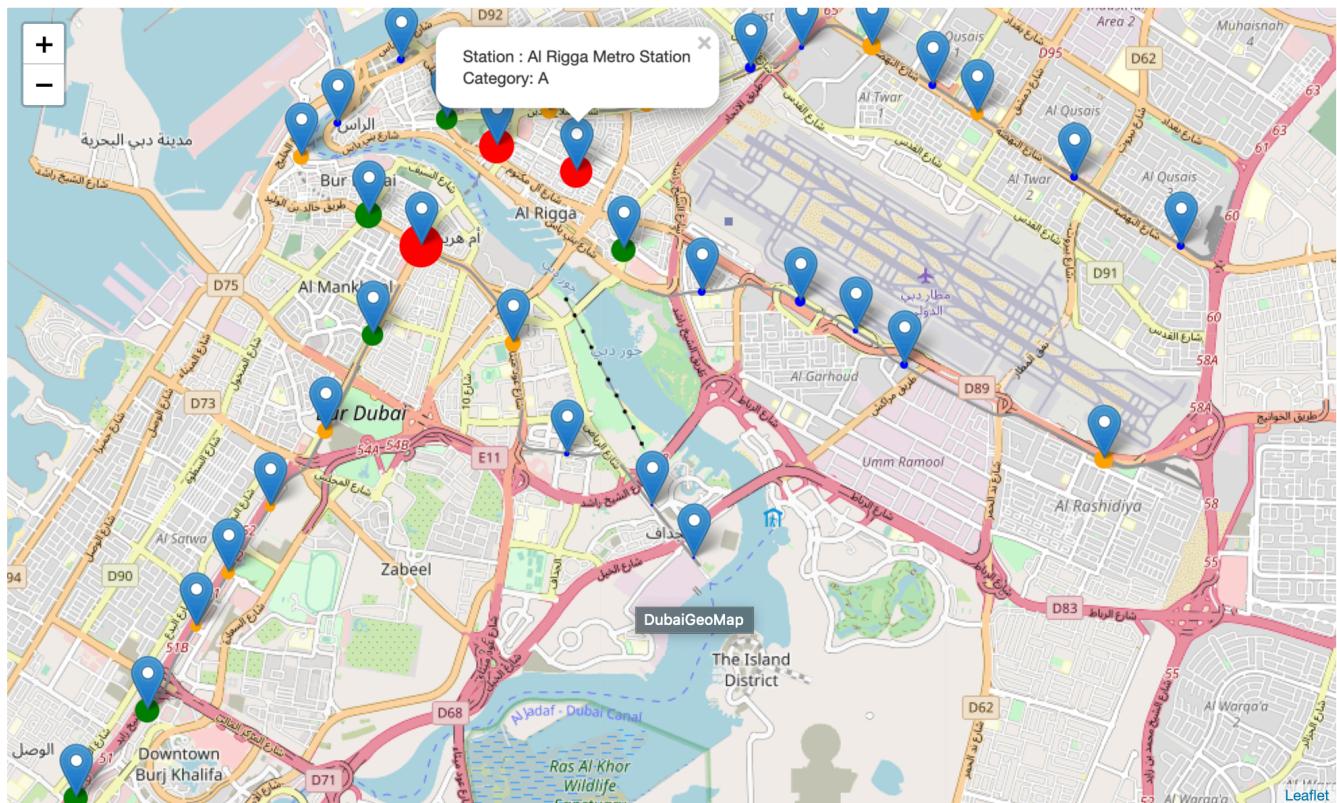
The following chart shows the number of transactions per month, it shows that summer months has less traffic than winter (Transactions are in Million)



GEOGRAPHICAL DISTRIBUTION

One more information will be an added value in order to see the whole picture, the geographical distribution of the stations' categories (A, B, C and D).

The following map show the geographical distribution of the stations' categories.



CONCLUSION

Stations like Al Rigga, Burjuman and Union metro station can be classified as Category A as they are the most visited station among the other stations and this will be considered in the advertising pricing process.

Number of passengers reaching its minimum in July and summer months, so we may reduce the price for advertising in those months.

The peak hours for most of the stations are between 7:00 am: 8:00 am and 5:00 pm: 7:00 pm, these will be the attractive hours for the customers that need their ad be viewed by the largest number of passengers.

Restaurants, hotels, stores and cafes have the largest number of venues around Dubai.

Venues like Ibn Battuta mall, Zaroob, Vox cinemas and cheesecake factory are the most liked venues based on foursquare.com API's data, so those could be good targeted customers.

RECOMMENDATION SYSTEM

As part of the future business areas the company want to invest in, we need to build a recommendation system based on foursquare users data to start suggesting venues to the user based on venues he visited.

We will user the Content Based methodology to build the recommendation system

System will start recommending venues based on other users' liked venues and the user liked the venue which our user currently liked.

Of course, the venue's category will be considered as the system will not recommend a hotel to a user that liked a junk food restaurant.

Before starting the suggestion process, the system will calculate the weight of every venue based on its popularity among other venues in the same area.

In our example the system suggesting venues to users who liked Burger King near ADCB Merto Station.

The system will filter the users who liked the same venue

	user_id		venue_id	main_category	venue_weight
1625	74154281		4c5b2a136bbdc9b650866c63	Restaurant	2
1626	80794171		4c5b2a136bbdc9b650866c63	Restaurant	2
1627	47751455		4c5b2a136bbdc9b650866c63	Restaurant	2

System will start recommending the popular venues which liked by the filtered users in the same category, sorted by popularity.

	venue_id	venue_name	main_category	station	venue_weight
0	4da1dd1e7aee548156fcfc9fe	Golden Fork	Restaurant	Al Ghubaiba Metro Station	4
1	4bdb148d3904a5937aa6489e	Saravana Bhavan	Restaurant	Al Ghubaiba Metro Station	4
2	5299aff011d2bb626ff08759	Mahesh Lunch Home	Restaurant	ADCB Metro Station	4
3	4b2cfa89f964a520aecb24e3	Blue Barjeel	Restaurant	Al Ras Metro Station	3
4	4bc99ce8937ca593f406a692	Pizza Hut	Restaurant	BurJuman Metro Station	3
5	4c5b2a136bbdc9b650866c63	Burger King	Restaurant	ADCB Metro Station	2
8	4d3058d9f986a093a7a5fb66	Biriani Esfahan Iranian Restaurant (بریانی اصفهان)	Restaurant	Union (Green Line) Metro Station	2
9	51decd8a498e66e7b8088ec6	Chatori Gali	Restaurant	Oud Metha Metro Station	2