The Dota 2 International 10 Traveller's Guide

Franco Neil De Asis

June 1, 2020

1. Introduction

1.1 Background

The International is the premier esports tournament for the video game Dota 2, hosted and produced by the game's developer, Valve. It is being held annually starting from 2011 upto the present. The event has been attended and watched by fans across the globe.

The first iteration of the International was on August 1, 2011 and was held on Gamescom in cologne. The tournament was born due to re-allocation of marketing funds. Instead of doing traditional marketing through advertising, Valve opted to do a tournament instead, boasting a 1 million USD grand prize. Since then the tournament has been held every year with the last iteration being held in Shanghai, China.

The tournament boasts the largest prize pool in all of esports, giving out a total of 34 million USD in prizes in 2019. The prizepool is a result of a crowdfunding scheme called the battle pass. 25% of the purchases made from the battle pass goes into the prize pool, in which the community was able to raise more than 30 million to add to the prize pool.[1]

The International consists of 18 teams which came from 6 major regions which include: North America, South America, Southeast Asia, China, Europe and CIS. It is a spectacle that has been watched by millions of people. In 2019, the total peak viewers(excluding China) amounted to 1,968,497 and has an average viewership of 726,754. The venues that the tournament was held on has been consistently sold-out.[2]

The next iteration of the International will be held at the Ericsson Globe, Stockholm, Sweden. It has a seating capacity of 16,000 people which is expected to sellout.[3] However, on April 30, 2020, Valve announced that the event would be cancelled due to the current Covid-19 pandemic and is possible to return on 2021.[4]

1.2 Statement of the Problem

There is a lack of guidance regarding attending these kinds of Esport events. The only information mostly available is the place and venue of where these events happen. The attendee/tourist is left to its own mechanism on how they should plan and go about their trip.

Creating a traveller's guide would help the spectators plan and prepare their itinerary during their stay in Sweden. It would leverage data from Foursquare API to acquire information regarding Hotels, Food Venues and Pubs and process that data so that it can be easily consumed by the spectators.

What's the value of improving the esport experience?

It has been projected that esports are going to grow at unprecedented levels. According to a Goldman Sachs report[5], the monthly audience would reach 276 million by 2022. The audience is mostly young and digitally inclined, where 79% are under 35 yrs. old. In Dota 2, the audience is predominantly young to middle aged adults, hence the inclusion of Pubs. The total esport prize pool is projected to 413 million in 2022 and Dota 2 boast the highest of them all in 2019.

Improving the experience of offline events would also improve the outlook of esports.

1.3 Interest

The Dota2 fans, esport fans and tournament attendees would be interested in a traveller's guide that would involve the premier esport event(The International).

2. Data

2.1 Downloading Data

The data is acquired by running queries on the Foursquare API. The data that is collected involves a list of venues around 10 km of the Ericsson Globe. The venues that are of interest are Hotels, Food, and Pubs. The researcher only considered these 3 venues because the traveller travelled mainly to Sweden for a sporting event. Most of their time would be spent on the Ericsson Globe in which the event would last around 1 week at around 8 hrs per day. So the tournament attendee won't really have much time to do actions that a normal tourist would do. Which would make these venues: Hotels, Food, and Pubs, the most vital for the tournament attendee. The inclusion of pubs is due to the fact that Dota 2 is a community driven game with grassroots origins. Pubstomps have been getting more popular as a means of community gathering. So pubs would provide value for the attendee especially if there are community meetups after the games.

There would be 3 primary datasets representing each venue category(hotel, food, pubs). And the data that the researcher would be getting from the Foursquare API would look like this:

```
"venue": {
    "id": "4be69a77910020a1d075d414",
    "name": "Daisy's",
    "location": {
        "address": "Ulvsundav\u00e4gen 71",
        "lat": 59.340587099504525,
        "lng": 17.96113959880651,
        "labeledLatLngs": [
            {
                "label": "display",
                "lat": 59.340587099504525,
                "lng": 17.96113959880651
            }
        ],
        "distance": 8703,
        "postalCode": "167 32",
        "cc": "SE",
        "city": "Stockholm",
        "state": "Storstockholm",
        "country": "Sverige",
        "formattedAddress": [
            "Ulvsundav\u00e4gen 71",
            "167 32 Bromma",
            "Sverige"
        ]
```

Figure 1. JSON File sample of the data from Foursquare API

Primarily, the data acquired from the api would only be a list of venues around Ericsson Globe, without more specific details. The api also has a limitation of 50 venues per call, so there would be multiple calls in order to collect all the venues.

After having the multiple list of venues, it would now be combined into 3 main lists which correspond to the 3 venue categories (hotel, food, pub). Then, another round of calls to the api is going to be made in order to collect specific information of the venues. Each api call would correspond to each venue. So after a call is made, the additional information is then added to its corresponding dataset. This process is going to repeat until all the venues have received its details.

2.2 Data Processing and Exploration

The data have some unnecessary columns that have to be removed. 'Unnamed: 0' and 'Unnamed: 0.1' columns were removed. The 'postal_code' column was also removed due to the error of data collection which resulted in having the entire column getting null values.

After that, a preliminary exploration of the data was done using the describe() method. It was found out that certain columns have null values. Histograms were also done in order to view how the values are distributed.

The graphs for the hotel data show that the majority of the distance values are in the middle of 3000-6000 meters. The amount of venues that have likes less than 50 numbers the most. The rating is generally evenly distributed. The price_tier shows that most hotels are at the upper range in pricing however our data only has 14 values.

In the food data histogram, the distance column is more skewed to the 1000-6000 meters. The likes are similar to the hotel distribution with the difference that the range of the majority is increased to around ~100 which can be assumed that there are more likes overall for food venues. The rating values are at the upper tier of >7. The price_tier shows that the majority of the venues are at the mid to lower end of the pricing from 2 to 1.

The pub data histograms show that pubs are on the cheaper side with the majority of the venues having a price_tier of 1. The rating goes as low as <5 and somewhat evenly distributed. Most pubs have a distance of 2000-6000 meters. The likes are somewhat similar to the food and hotel where the majority has a value of <100.

After looking at the distribution, rows that have null values in the 'rating' and 'likes' column were dropped. The missing values in likes were then replaced by a value of zero.

The data was then plotted to a map to have a glimpse of the distribution of the venues in the map.
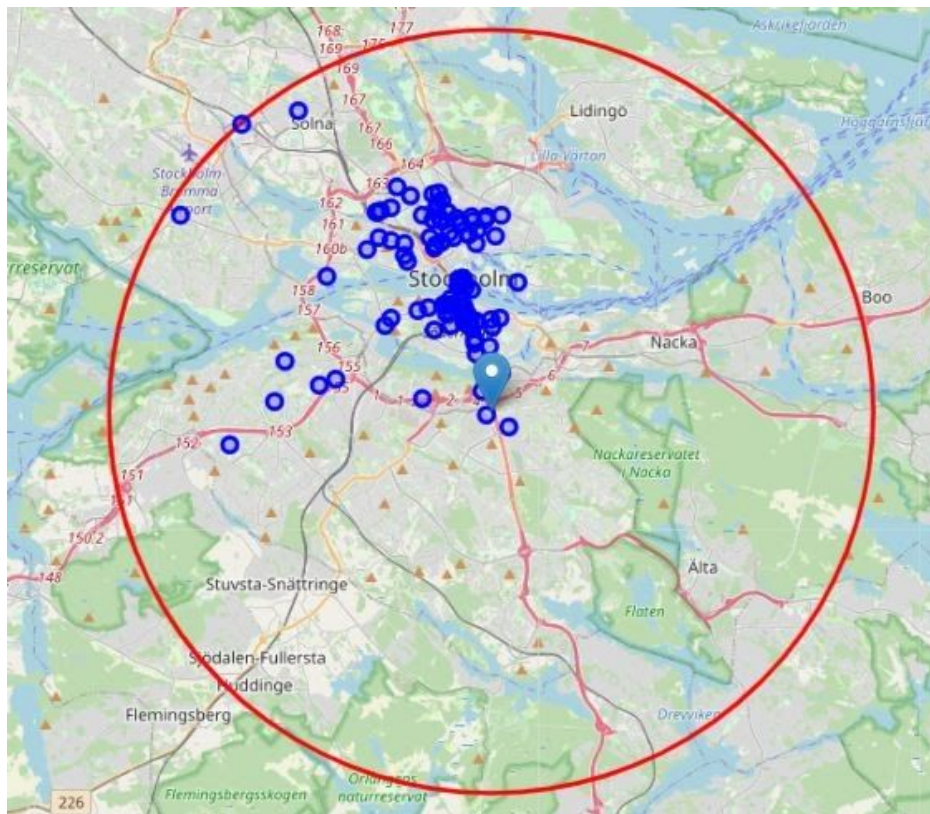

Figure 2. Sample distribution of the pub venues in the map

The final data ended up with 156 Hotels, 230 Food venues, and 102 Pubs.

3. Methodology

This project would try to determine which areas would be a good choice to stay in for the tourist/spectator. The methodology would be accounting all the hotel, food, and pub venues that are around 10 km of the Ericsson Globe.

In the first step, the researchers collected all the data necessary for the hotel, food and pub venues that are located around the Ericsson Globe. It has been made sure to include important parameters such as the geographic coordinates(latitude, longitude), distance from the Ericsson Globe(in meters), the rating of the venue, the number of likes, and the price tier.

In the second step, the researchers attempted to understand the density of the respective venue category around Stockholm. Using heatmaps to get a good view on the density and then cluster the venues according to their geographic coordinates to be able to show interesting clustering. Then the researchers tried to find patterns in the density and clustering of the venues and try to correlate these clusters to every other venue category.

Then in the third step, the researchers would now try to understand the grouping of the venues depending on the parameters: likes, rating, price, and distance.

Lastly, using both the density and parameter groupings, the researchers would give recommendations on which places the tourist can generally stay. They would try to give recommendations for tourists that want the highest quality hotel, food, and pub. Recommendations based on distance to the tournament venue(Ericsson Globe). And recommendations depending on their budget.

4. Analysis

4.1 Geographic Location Clustering and Analysis

This step involves grouping the data according to the geographic location. It would first show the heatmap and then find how the K-Means algorithm clusters the venues that would differentiate between regions. This step would try to know where and how dense a region and cluster is.

Only the latitude and longitude attributes are considered in this step. The coordinates are used to create a heatmap and then clustered using the K-Means algorithm. The heatmap would provide a basis on which areas are deemed to be dense or "hot" and then determine which clusters form these dense/hot areas. These clusters would then be grouped together to form a major cluster.

4.1.1 Hotels

Here is the result of the heatmap and clustering:
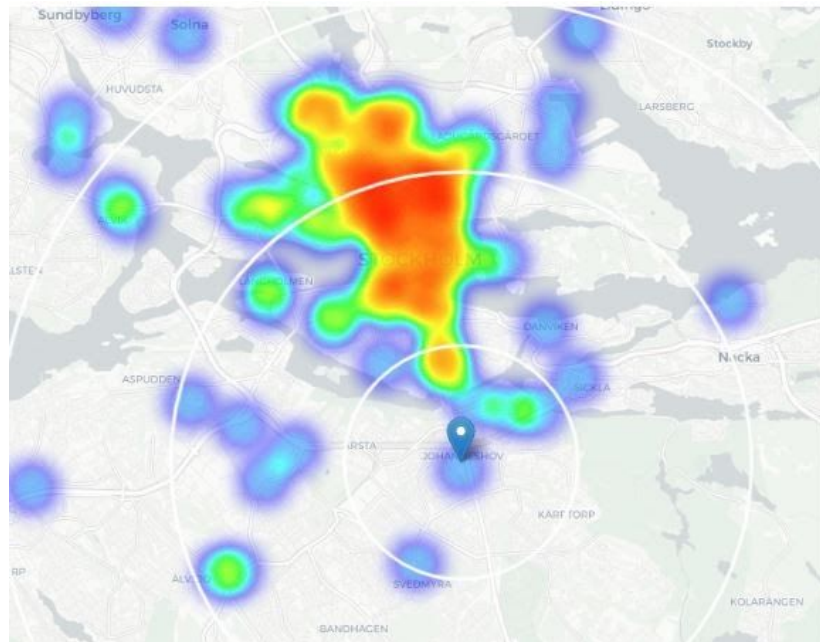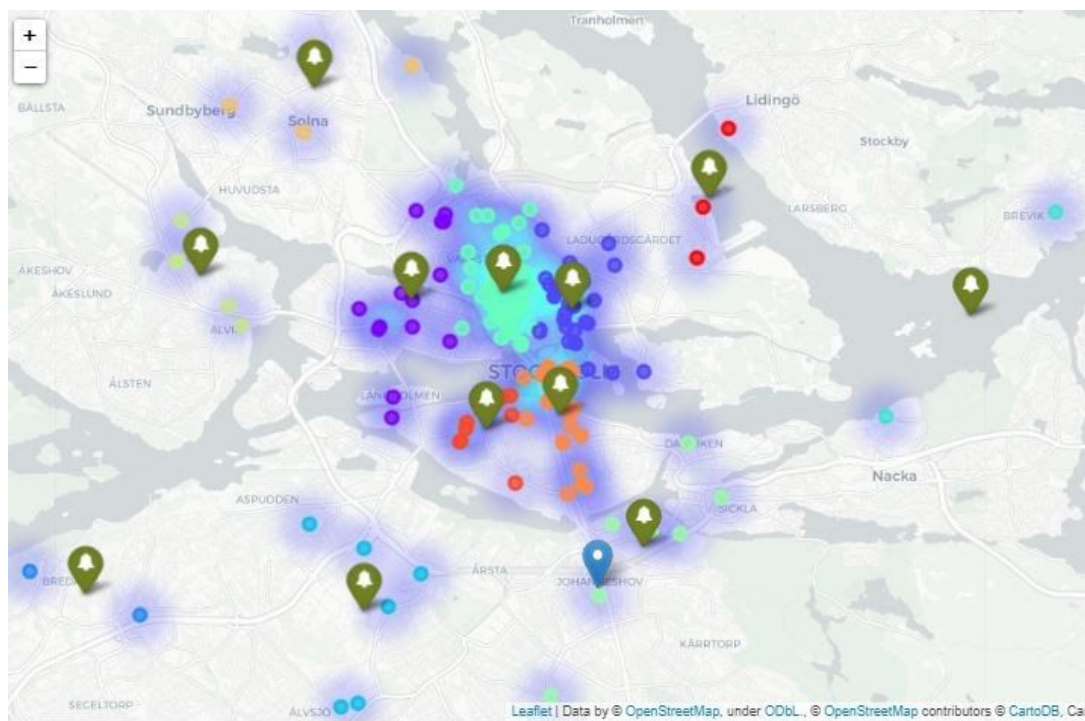

Figure 3. Hotel Data Heatmap

Figure 4. Hotel Data Clusters

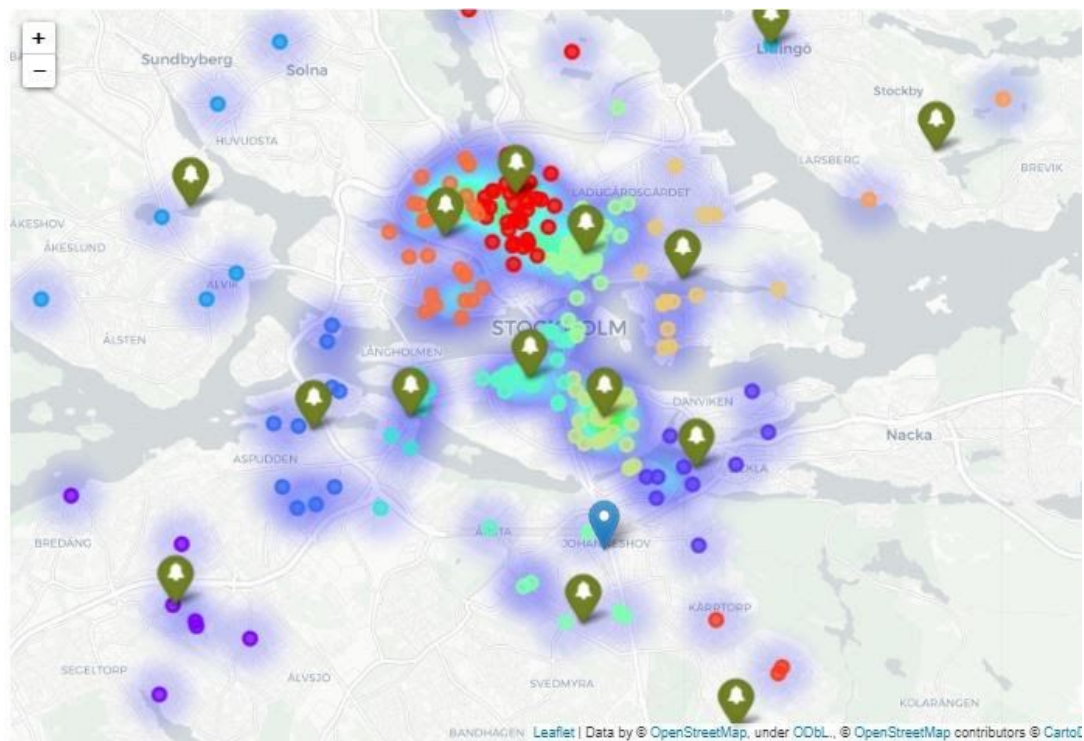**4.1.2 Food Venues**

Figure 5. Food Data Heatmap

Figure 6. Food Data Clusters

4.1.3 Pubs
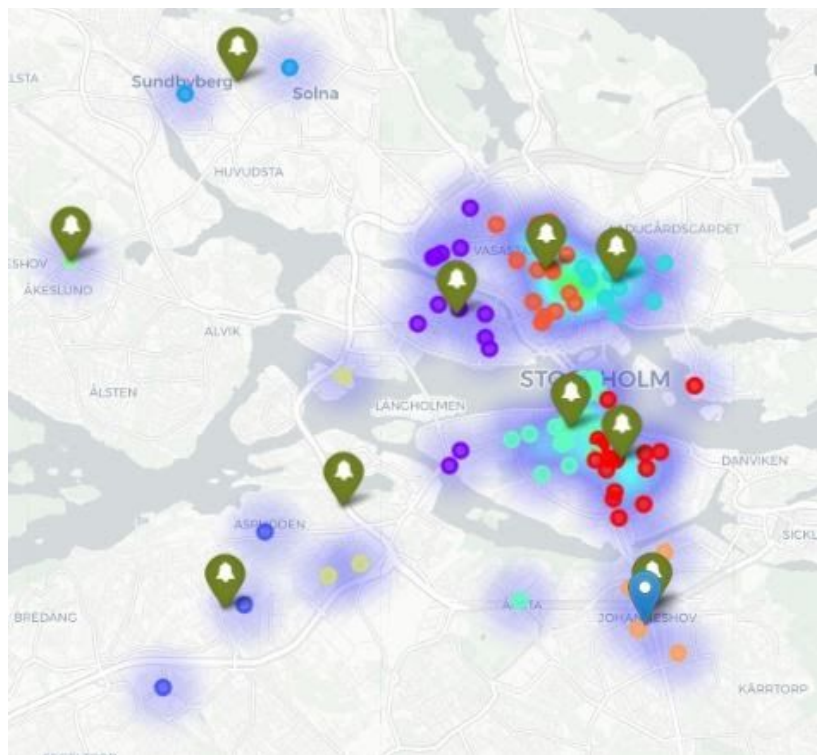

Figure 7. Pub Data Heatmap

Figure 8. Pub Data Clusters

4.1.4 Density Analysis

The researchers found out that in the three categories, the dense/hot area is located north of the Ericsson Globe. The clusters that comprised the area are:

Hotel	1, 2, 6, 10, 11
Food	0, 7, 9, 10, 13
Pub	0, 1, 4, 5, 9

Here are the statistics of the area:

```
Hotels
              lat  lng  category
Geo Clusters
1             14   14         14
2             30   30         30
6             52   52         52
10            23   23         23
11             8    8          8
Total: 127
-----------
Food Venues
              lat  lng  category
Geo Clusters
0             42   42         42
7             26   26         26
9             30   30         30
10            31   31         31
13            31   31         31
Total: 160
-----------
Pubs
              lat  lng  category
Geo Clusters
0             18   18         18
1             13   13         13
4             13   13         13
5             28   28         28
9             17   17         17
Total: 89
```

Figure 8. Dense Region Statistics

This region which we would now call the "dense region" has a total of *376* venues which represent **77.05%** of the total venues. Then the center point of the region was calculated by determining the mean of all coordinates of the venues in the dense region. The coordinates of the center point is [59.330048881331344, 18.063149202393614].

4.1.5 The 4 Districts: Norrmalm, Östermalm, Södermalm, Kungsholmen

The most dense region is located in 4 districts namely: Norrmalm, Ostermalm, Sodermalm, and Kungsholmen. Below are descriptions for the districts coming from some travel guides.

Norrmalm:

According to wikitravel.org, the district is well served by public transport. Most venues can also be reached by walking. The Stockholm central station is also located in this district along with the Cityterminalen (the bus terminal) and T-Centralen (the subway terminal).[6]

visitstockholm.com[7]:

> *"The City Center is the place for affordable shopping, especially for the young and trendy. Drottninggatan and the area between Sergels Torg and Kungsträdgården offer a mix of clothing chains and large-scale classic Stockholm department stores. Åhléns City, where Drottninggatan ends at Sergels Torg, is a popular department store with most of the best Swedish brands.*
>
> *There are several large chain stores in the nearby area and at Gallerian on Hamngatan, shopaholics can get their fill of various clothing chains and then stop for a bite at one of the many restaurants and cafés. NK (Nordiska kompaniet), on Hamngatan, is a national institution when it comes to fashion and design, and the nearby MOOD Stockholm is a new mall with shopping, food, and service.*
>
> *For a break from the hustle and bustle of the city center, you´ll find that Kungsträdgården is perfect for ice cream and people-watching; during spring thousands of pink cherry blossoms light up the park and when winter comes, the ice-skating rink is a very popular spot for both visitors and locals alike. Don't miss the newly refurbished Hötorgshallen food hall and the colorful market on the adjacent square. And if you're up for a good movie Norrmalm has quite a few movie theaters to choose from!"*

Östermalm:

visitstockholm.com[8]:

> *"Östermalm makes for an interesting contradiction. It's a posh residential neighborhood. But it also has the highest concentration of nightclubs in the city, found around Stureplan.*
>
> *Bibliotekstan is the city's most affluent shopping district, with exclusive Scandinavian and international brands. A slow saunter between the food stalls in the Östermalmshallen market hall is also mandatory.*

Take a walk along Strandvägen, and enjoy the magnificent buildings and wonderful seaside views."

Södermalm:

visitstockholm.com[9]:

"Södermalm is, after all, more than just trendy shops and cozy cafés. It's also an area with diverse architecture, popular city parks and great stories to tell."

Kungsholmen:

visitstockholm.com[10]:

"Kungsholmen is primarily a residential island, where you'll have a great opportunity to mix with the locals. Lots of lively restaurants, bars, and cafés are found here, especially along Hantverkargatan and Fleminggatan. Kungsholmen is home to the most prominent building of the Stockholm skyline, the City Hall. If you get the chance, don't miss climbing the tower to be amazed by the incredible view.

The walk along the tree-lined quay on Norrmälarstrand offers unparalleled views of Södermalm and boats sailing on the lake. If you continue past the City Hall you can actually stroll around the whole island along the water. It's a popular track for locals and is often either finished or started with a brunch in one of the many cafés in the area.

During summer, there are a couple of nice places to go for a swim on the island, for instance in the big green area of Rålambshovsparken. For a truly exotic experience, we recommend a swim from the rocks of Fredhällsbadet."

4.1.6 Generating the distance between the venues and the "Dense Region Center"

The researchers have first determined the distance from the dense region center to the Ericsson Globe. By drawing a straight line between the center of the dense region to the Ericsson Globe, the line covers 4.22km. Google maps determined that the center of the dense region corresponds to this address: **Klaratunneln, 111 52 Stockholm, Sweden.**

Using google maps in determining the travel time between the two points using public transportation, it was able to calculate that it takes 22mins from Klaratunneln, 111 52 Stockholm, Sweden to the Ericsson Globe. Travelling through a car is faster though.
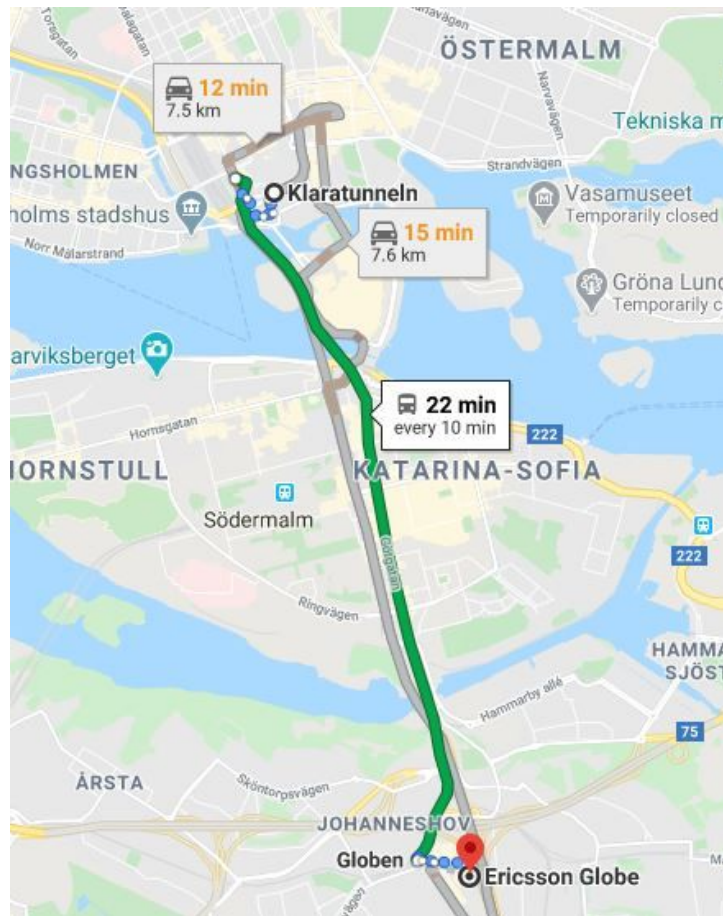
Here is the map:

Figure 9. Google Maps Travel estimation from the Dense Center to Ericsson Globe.

The researchers then proceeded to calculate the distance between the dense region center to each and every venue in the data. This new data would be used in the next processes.

4.2 Clustering Based on Ratings, Likes, Price, Distance to the Dense Region

The next step would be clustering the data based on other parameters aside from geographic location. The researchers would now cluster the data using other attributes of the venues, namely: ratings, likes, price and the newly acquired distance from the dense region center.

The parameters: ratings, likes, price, and distance to the dense region center, would be placed on a separate data frame where it would undergo standardization using the standard scalar algorithm. After that, the data would now be subjected to the K-Means algorithm to cluster the data.

4.2.1 Hotels

Here is the result of the clustering:

cluster_rel_to_dense	likes	rating	dist_to_dense_region	count
0	37.634615	7.598077	1.357068	52
1	314.200000	8.180000	0.539206	5
2	19.981481	5.798148	1.345565	54
3	154.285714	8.100000	1.092164	21
4	22.333333	6.645833	5.555754	24

Figure 10. Hotel Attribute Clustering Results

Cluster 1 [Most attractive hotels]:

The cluster shows that it has the highest average likes and rating while at the same time having the closest average distance to the dense region center. With these characteristics it could be assumed that this cluster represents the best venues. However, the number of venues in this cluster is only 5.

Cluster 4 [Least attractive hotels]:

The cluster is the opposite of cluster 1. Low end of likes and rating while having the highest distance to the dense center. These venues might possibly be the least attractive venues for tourists of the International 10.

Cluster 3 [Amazing Hotels with good characteristics]:

The cluster shows that it is the 2nd best cluster in terms of its attributes. High likes and rating but only second to cluster 1. It also has an average distance of 1km making the venues really attractive.

Cluster 0 [Decent Hotels]:

The cluster has good likes and above average ratings and the distance is really good too. It is above average in the least. The cluster has the 2nd highest venues at 52.

Cluster 2 [Average Hotels]:

It has a number of likes while having an average rating. It has acceptable characteristics but might not be a primary choice for venues. The cluster is the most numerous at 54.

4.2.2 Food Venues

Here are the results:

cluster_rel_to_dense	likes	rating	price_tier	dist_to_dense_region	count
0	50.243902	8.102439	1.536585	2.038127	82
1	38.000000	8.497143	1.314286	6.083909	35
2	481.315789	8.800000	2.105263	1.436900	19
3	104.043478	8.552174	3.173913	1.791393	23
4	95.521127	8.849296	1.633803	1.827704	71

Figure 11. Food Venue Attribute Clustering Results

Cluster 2 [Best Food Venues]:

The cluster has it all. Highest likes, 2nd highest rating, lowest distance to the dense center while maintaining a mid price tier. Based on the cluster and data from the foursquare api, these venues should be the best places to eat in.

Cluster 3 [Amazing Pricey Food]:

The cluster is very likeable, high rating, close in proximity but it has the highest price tier in all of the clusters. If it is compared to cluster 2, it can be assumed that people liked food venues more that have affordable pricing. If the tourist wants an amazing experience without minding the price, then these venues should provide for their needs.

Cluster 4 [Amazing Affordable Food]:

Arguably this cluster is better than cluster 3 because of its characteristics. In comparison to cluster 3 it is very likeable while only having a difference of +8.522351 in likes, highest rating, only having an average distance of ~1.8 km while being really affordable at a price tier of ~1.6. If tourists are looking for an amazing accessible dining experience while being friendly to the wallet, then these venues are the place to go.

Cluster 0 [Decent Food]:

The cluster has a decent amount of likes while still enjoying a high rating. The price tier is also at the low end and an average distance of 2 km ain't that bad.

Cluster 1 [Decent Cheap Food but requires more travel]:

If the tourists are looking for cheap good food and want to travel further than the dense region, then these venues would meet their expectations.

All the food venues are rated very highly which means that based on foursquare data, the food that is being offered in Stockholm is deemed to be highly desirable.

4.2.3 Pubs

Here are the results:

cluster_rel_to_dense	likes	rating	price_tier	dist_to_dense_region	count
0	14.842105	5.578947	1.394737	1.626495	38
1	383.000000	8.250000	1.750000	1.411780	4
2	50.647059	7.397059	1.147059	1.510102	34
3	136.866667	7.973333	2.600000	1.046439	15
4	14.363636	6.609091	1.272727	5.368291	11

Figure 12. Pub Attribute Clustering Results

Cluster 1 [Best Affordable Pubs]:

The cluster has the best likes and ratings while maintaining a price tier of <2.0. It is also close to the dense region center.

Cluster 3 [Pricey Good Pubs]:

It has a high amount of likes and is highly rated. It is also the cluster that is closest to the dense region center but the downside is that it represents the most expensive venues.

Cluster 2 [Cheap Good Pubs]:

Lowest price tier with a decent amount of likes and maintaining a high rating of > 7.0. It is also in close proximity to the dense region center. If the tourist is looking for a cheap yet decent pub, then these venues would be able to fit the bill.

Cluster 0 [Average Pubs]:

It has the lower end of likes and has the lowest average rating. The redeeming attributes are that the cluster is relatively cheap and somewhat accessible.

Cluster 4 [Average yet Far Pubs]:

Lowest amount of likes with a decent rating. It is fairly cheap too but the tourists need to get out of their way to reach the venues.

The Pubs are generally on the cheaper end of the price tier, so the venues are mostly good for a community night out.

4.2.4 Aggregate Analysis

The clustering of the data using the *likes, rating, price_tier, and distance_to_dense_region* attributes produced some *remarkable patterns*.

A primary assumption is made in the case of hotels. The hotel data doesn't have the price_tier data however its other characteristics exhibit the same patterns of food and pub data. So it is assumed that the hotels would have the same characteristics of price_tier to those of the food and pub data.

4.2.4.1 Best Clusters: Hotel [Cluster 1], Food [Cluster 2], Pub [Cluster 1]

```
-------------------------------------------------------------
Hotel [Cluster 1]
likes                     314.200000
rating                      8.180000
dist_to_dense_region        0.539206
Name: 1, dtype: float64
-------------------------------------------------------------
Food [Cluster 2]
likes                     481.315789
rating                      8.800000
price_tier                  2.105263
dist_to_dense_region        1.436900
Name: 2, dtype: float64
-------------------------------------------------------------
Pub [Cluster 1]
likes                     383.00000
rating                      8.25000
price_tier                  1.75000
dist_to_dense_region        1.41178
Name: 1, dtype: float64
-------------------------------------------------------------
```

Figure 13. Characteristics of the Best Cluster

They all have the similar characteristics of best likes, best rating, affordable price, and close proximity to the dense center.
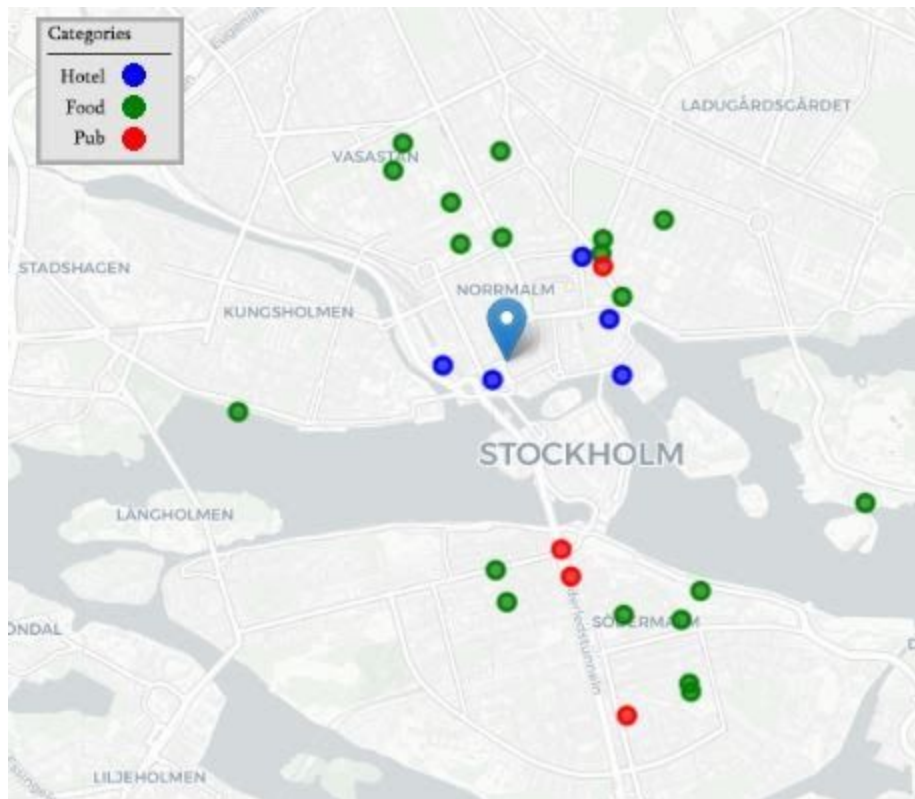
Here is the cluster in the map:

Figure 14. Map of the Best Clusters

As shown above the best venues are located in close proximity to each other and at the same time at the close proximity of the dense region center. They are located mostly in Norrmalm and Södermalm districts.

4.2.4.2 Fringe Clusters: Hotel [Cluster 4], Food [Cluster 1], Pub [Cluster 4]

```
---------------------------------------------------------
Hotel [Cluster 4]
likes                    22.333333
rating                    6.645833
dist_to_dense_region      5.555754
Name: 4, dtype: float64
---------------------------------------------------------
Food [Cluster 1]
likes                    38.000000
rating                    8.497143
price_tier                1.314286
dist_to_dense_region      6.083909
Name: 1, dtype: float64
---------------------------------------------------------
Pub [Cluster 4]
likes                    14.363636
rating                    6.609091
price_tier                1.272727
dist_to_dense_region      5.368291
Name: 4, dtype: float64
---------------------------------------------------------
```

Figure 15. Characteristics of the Fringe Clusters

The greatest characteristic that made these clusters least desirable is the distance from the dense region. The clusters also have low likes, mostly average to above average rating, but they have cheap pricing.
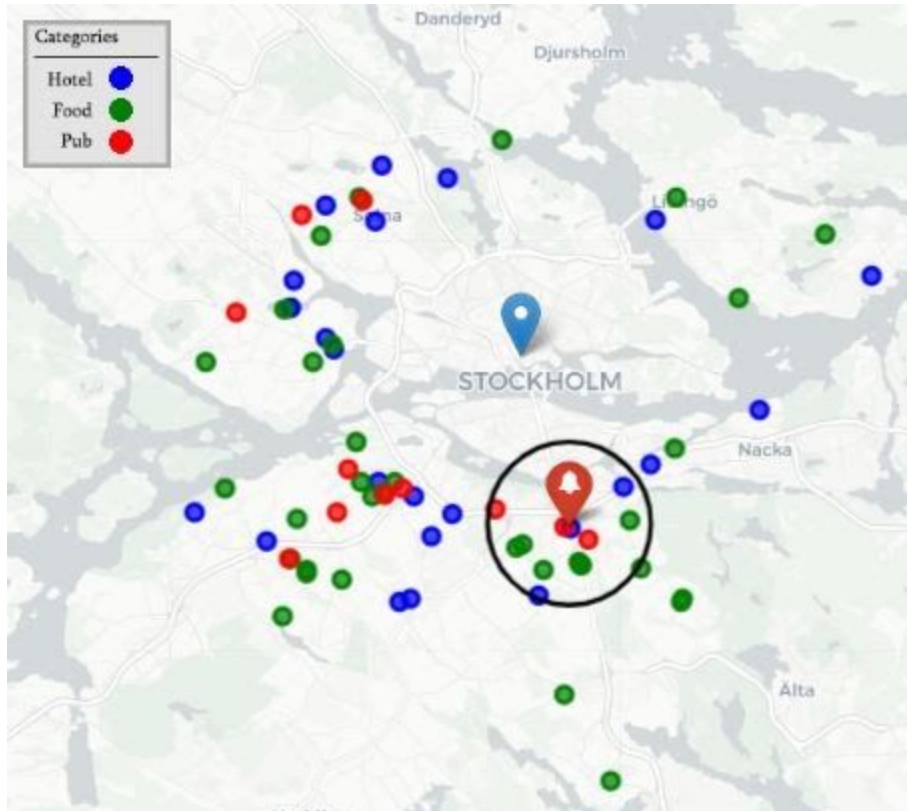
Map:

Figure 16. Map of the Fringe Clusters

However, as shown in the map. These clusters have a region which could probably make it really *desirable* for the International 10 tourist because of its **proximity(within 2 km) of the Ericsson Globe**. It would make the event area really accessible.

4.2.4.3 Amazing Clusters: Hotel[Cluster 3], Food[Cluster 3], Pub[Cluster 3]

```
------------------------------------------------------------
Hotel [Cluster 3]
likes                    154.285714
rating                     8.100000
dist_to_dense_region       1.092164
Name: 3, dtype: float64
------------------------------------------------------------
Food [Cluster 3]
likes                    104.043478
rating                     8.552174
price_tier                 3.173913
dist_to_dense_region       1.791393
Name: 3, dtype: float64
------------------------------------------------------------
Pub [Cluster 3]
likes                    136.866667
rating                     7.973333
price_tier                 2.600000
dist_to_dense_region       1.046439
Name: 3, dtype: float64
------------------------------------------------------------
```

Figure 17. Amazing Clusters Characteristics

These venues in these clusters have a glaring characteristic: they are the most expensive. They have amazing attributes in likes, rating and distance but they are more costly. These clusters are only 2nd to the best clusters and the difference between these 2 groups is that the *best clusters* have more affordable pricing, mostly in mid range. The fewer likes and ratings of the *amazing clusters* could possibly be attributed for having a higher price tier.
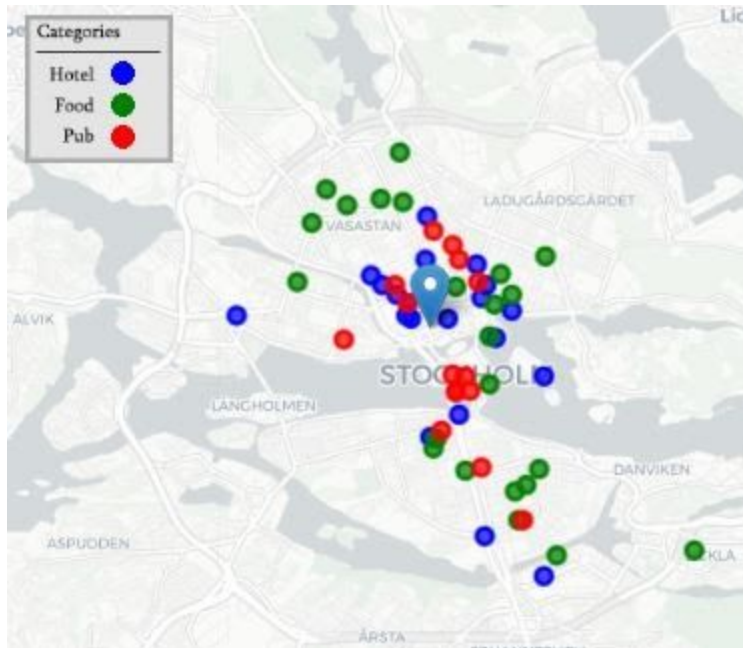
Map:

Figure 18. Amazing Clusters Map

4.2.4.4 Highest Amount of Venues(Most Numerous) Clusters: Hotel [Cluster 0 & 2], Food [Cluster 0 & 4], Pub [Cluster 0 & 2]

```
-----------------------------------------------------------
Hotel [Cluster 0]
likes                 37.634615
rating                 7.598077
dist_to_dense_region   1.357068
Name: 0, dtype: float64
-----------------------------------------------------------
Food [Cluster 0]
likes                 50.243902
rating                 8.102439
price_tier             1.536585
dist_to_dense_region   2.038127
Name: 0, dtype: float64
-----------------------------------------------------------
Pub [Cluster 0]
likes                 14.842105
rating                 5.578947
price_tier             1.394737
dist_to_dense_region   1.626495
Name: 0, dtype: float64
-----------------------------------------------------------
-----------------------------------------------------------
Hotel [Cluster 2]
likes                 19.981481
rating                 5.798148
dist_to_dense_region   1.345565
Name: 2, dtype: float64
-----------------------------------------------------------
Food [Cluster 4]
likes                 95.521127
rating                 8.849296          Hotel
price_tier             1.633803          Cluster 0: 52
dist_to_dense_region   1.827704          Cluster 2: 54
Name: 4, dtype: float64                  -----------------
-----------------------------------------------------------
                                         Food
Pub [Cluster 2]                          Cluster 0: 82
likes                 50.647059          Cluster 4: 71
rating                 7.397059          -----------------
price_tier             1.147059          Pub
dist_to_dense_region   1.510102          Cluster 0: 38
Name: 2, dtype: float64                  Cluster 2: 34
-----------------------------------------------------------
```

Figure 19. Most Numerous Clusters Characteristics

These clusters have low to mid amounts of likes and ratings. They are also affordable and close to the dense region center.
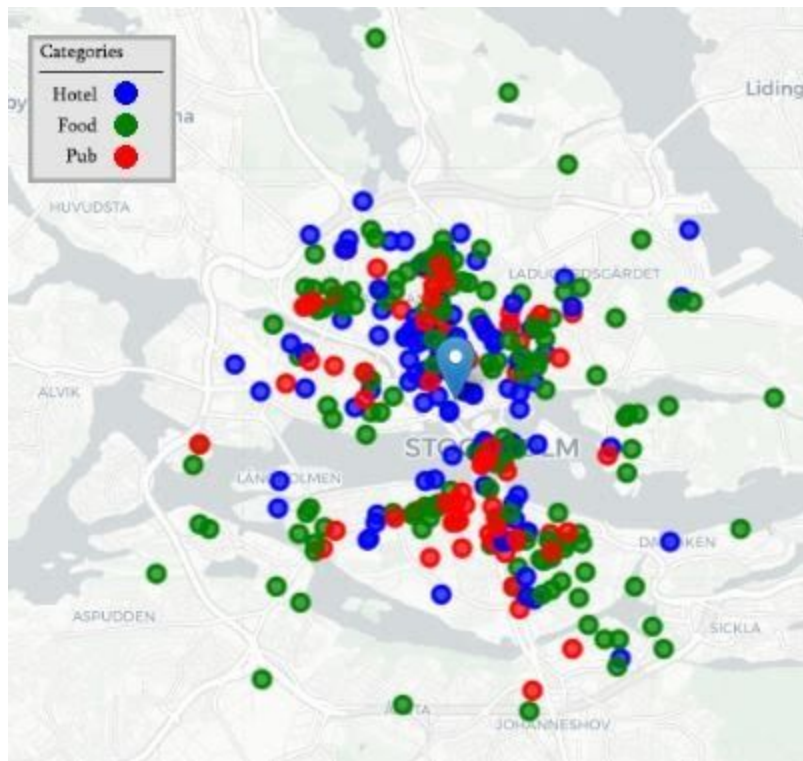
Map:

Figure 20. Most Numerous Clusters Map

These clusters are concentrated in the dense region. If the tourist just wants decent eats without any much preferences, then these venues can easily be found all over the place in the dense region.

4.3 Analysis on the Payment Column

An auxiliary analysis is done on the payment schemes of these venues to add recommendations for the tourists on how they can carry their money and pay for services.

```
Total hotel venues that accepts credit cards: 140
Total hotel venues that doesn't have data on payment: 16

Total food venues that accepts credit cards: 199
Total food venues that doesn't have data on payment: 30
Total food venues that only accepts cash: 1

Total pub venues that accepts credit cards: 87
Total pub venues that doesn't have data on payment: 14
Total food venues that only accepts cash: 1
```

```
            Hotels: 89.74%
            Foods: 86.52%
            Pubs: 85.29%
            Total: 87.30%
```

Figure 21. Payment Method Statistics

It is assumed that the venues that don't have data on payment, only accept cash. The venues have a high acceptance rate for credit cards of 87.30% which would allow more flexibility in payment options for the tourists.

5. Results and Discussion

Our endeavor has determined that Stockholm offers a decent amount and varied choices of Hotel, Food, and Pub venues for the International 10 attendees. The data was acquired through the Foursquare API and the data pertains to the venues that are within 10km of the Ericsson Globe. After processing the raw data, this effort attained 156 Hotels, 230 Food venues, and 102 Pubs.

The primary step done was to create a heatmap for each of the venues to identify the hotspots in Stockholm. Upon further observation, it was found out that the hotels, food venues, and pubs are concentrated in a region north of the Ericsson Globe. Clustering based on the geographical coordinates was done in order to identify which venues are comprising this "hot" region above the Ericsson Globe. All the pertaining clusters were visually identified and combined together. The mean of the coordinates of these concerned venues was then calculated. The mean represents the focal point or the center of the "Dense Region." This focal point represents the area that is the most accessible in the region.

The dense region comprises 4 districts, namely: Kungsholmen, Södermalm, Norrmalm and Östermalm. These 4 districts represent the Stockholm City Center. The dense region represents 77.05% of the total venues.

After getting the center point(59.330048881331344, 18.063149202393614), the distances from this point to all the venues are then calculated. It is found out that the distance of the straight line between this point and the Ericsson Globe is 4.22 km. By using google maps, the address(Klaratunneln, 111 52 Stockholm, Sweden) of this point was determined and an estimation of the travel time between this center to the Ericsson Globe was calculated. The time of travel averaged to 22 mins when using public transportation and the travel time ranges from 12 to 17 mins if travelling by car. This new data would help the tourist plan for their commute.

Now that the distances between the center and the venues have been calculated, another set of clustering is going to be performed. This new set of clustering involves parameters such as likes, rating, price tier, and the acquired data of distance to the dense region center. Each category(hotel, food, pub) was clustered separately. The results of the clusters were able to identify remarkable features and patterns. And these patterns that are present in each of the categories have been found out to be similar to each other which prompted an aggregate analysis.

The aggregate analysis that was done has an assumption that was made in which the behaviour of the price tier in the hotel is assumed to be the same as what is found in food

and pubs. The assumption was made because the hotels lacked the necessary price tier data.

The results of the aggregate analysis identified 4 major patterns for the data. These characteristics would also serve as recommendations for the tournament attendee. The first pattern pertains to the venues that are deemed to have the best qualities. It has the highest likes, rating, and close proximity to the center. At the same time, the venues have an affordable pricing, right in the middle. However, take note that these clusters have the fewest number of venues.

The second pattern pertains to the fringe venues that are located at the outer rim of the dense region. They have at least average in likes and rating but they are mostly the cheapest venues. One area needs specific attention though which is the area around the Ericsson Globe. There are venues in this pattern that are in close proximity with the event area. Because of its distance, these venues in this area can be deemed to be a premium location.

The third pattern represents the venues that are amazing but in exchange they are a bit pricey. These venues have high likes and ratings. The distance to the center is also really good. However, it represents the most expensive venues. These venues would probably be lower in demand and should be good for tourists that can afford the price.

The last pattern identifies the venues that are about average but accounts for the most in numbers. So if the tourist/attendee just wants a decent place to stay, eat, and drink, there are plenty of venues to choose from. This pattern represents the fact that Stockholm City Center has an abundance of venues that the attendees can go to.

Lastly, an auxiliary analysis is made regarding the payment options of the venues. It was found out that 87.3 % of the venues offer credit card payments and some accept NFC(near-field communication) payments. The venues in Stockholm have a lot of flexibility in terms of payment.

6. Conclusion

The aim of the project is to attempt to find travel recommendations for the International 10 attendee. Because the traveller isn't a normal tourist, its needs are also different. The chunk of the stay is going to be used in watching the tournament. So the venues that were considered were Hotel, Food venues, and Pubs. The inclusion of pubs is due to the video game, Dota 2, is community-driven and there have been events called pubstomps. Pubs would be an amazing place for a community gathering. The project has explored, processed and clustered the venue data and was able to come up with a couple of recommendations. One is the general recommendation to find a place in the Stockholm City Center(the 4 districts). The project was also able to give an estimate of the travel time from the city center to the Ericsson Globe which could help the attendee plan their day. The project was also able to give recommendations on which places the traveller can go depending on their preferences.

However, one thing to point out is that the data is from Foursquare API and it has been processed. The data is only a representation of the venues and does not paint the complete picture because it doesn't have the complete list of venues. But the project should be able to give the traveller a good enough picture on Stockholm. The project is not an end-all but rather a supplement that would help the attendee take less time to plan and more time to enjoy their stay.

References

[1] https://en.wikipedia.org/wiki/The_International_(Dota_2)

[2] https://escharts.com/tournaments/dota2/international-2019

[3] https://en.wikipedia.org/wiki/Ericsson_Globe

[4] http://blog.dota2.com/2020/04/the-international-2/

[5] https://www.goldmansachs.com/insights/pages/infographics/e-sports/index.html?cid=sch-pd-bing-esportshub-searchad-201810-----&mkwid=9b1dvEvt

[6] https://wikitravel.org/en/Stockholm/Norrmalm

[7] https://www.visitstockholm.com/guides/citynorrmalm/

[8] https://www.visitstockholm.com/guides/district-guide-ostermalm/

[9] https://www.visitstockholm.com/guides/see-the-sights-of-sodermalm/

[10] https://www.visitstockholm.com/guides/district-guide-kungsholmen/