Capstone project- Food Business in Chittagong

Introduction:

Chittagong ,known as “Chattrogram” is a major coastal city and financial centre in southeastern Bangladesh. The city has a population of more than 2.5 million ,while the metropolitan area had a population of 4,009,423 in 2011. making it the second largest in the country. It is the capital of an eponymous district and division. The city is located on the banks of the Karnaphuli river between the Chittagong hill tracts and the Bay of Bengal. Modern Chittagong is Bangladesh's second most significant urban center after Dhaka.

Food business is one of the most popular businesses in this city. Lots of restaurants and food corners are available here. With a population over 4 million and large number of restaurants makes it crucial to make choices where and which type of restaurant should be a successful business.

Restaurants are also a favourite pastime and hangout place for teenagers and students as the city has a lack of playgrounds and parks. So students are the most common consumers of these food businesses. This city is also popular for its legendary “mezbani” food.

But in this project we will mainly focus on the consumers who are students. Foods that are popular to the students and find the best place to start a food business.

Problem:

To find the answers of the following questions

1. **What type of restaurant is most popular in the city among the students?**
2. **What kind of restaurant can be a potential for business ?**
3. **Which area has the highest density of educational institutions ?**
4. **Which is the best place to start a food business**

Data Section:

Chittagong city’s demographic shows various amounts of restaurants. These restaurants are of different kinds. Certain types of restaurants populate certain areas.

- Places in the city having restaurants with their latitude and longitude
    - Data source: Foursquare API
    - Description : Using this data we will find the places that has restaurants and the density of the restaurants with types, example: cafe, biriyani , chinese etc.

- Places in the city having educational institutions
    - Data source: Foursquare API
    - Description : Using this data we will find the places that have educational institutions which will help to visualize the density of the student population .
      Example: schools, colleges and universities and their locations

Methodology:

1. We will get all the locations of educational institutions in the city by using foursquare API.First we get all the data of the colleges in the city.

```
In [218]: search_query = 'College'
          radius = 7000
          print(search_query + ' .... OK!')

          College .... OK!

In [219]: url = 'https://api.foursquare.com/v2/venues/search?client_id={}&client_secret={}&ll={},{}&v={}&query={}&radius={}&limit=
          url

Out[219]: 'https://api.foursquare.com/v2/venues/search?client_id=3FML0TLCYPYUGJ0FT5I0VCMFIEZ0TWJVI2AP4OK0TFJJ4CFJ&client_secret
          =XUGMCCREGBAXMYE1CJ3QRAZ0QDMLDDJQMFT3RHZSSFWYQVNR&ll=22.3307998,91.8412863&v=20180604&query=College&radius=7000&limit
          =200'

In [220]: results = requests.get(url).json()
          results
```

```
In [222]: dataframe = json_normalize(venues)
          dataframe
```

	id	name	categories	referralId	hasPerk	location.lat	location.lng	location.labeledLatLngs	location.distance	location.cc
0	4d3977d081258cfa062c9b5f	Chittagong Medical College Boys Hostel	[{'id': '4d954b06a243a5684965b473', 'name': 'R...	v-1587910011	False	22.356672	91.830798	[{'label': 'display', 'lat': 22.35667211535404...	3075	BD
1	515bc883e4b045a6e11e631e	Chittagong Medical College - Lower lecture gal...	[{'id': '4bf58dd8d48988d1b3941735', 'name': 'M...	v-1587910011	False	22.361135	91.830655	[{'label': 'display', 'lat': 22.36113520413447...	3549	BD
2	54680b61498e9e5bff89864e	Chittagong Maa O Shishu Hospital Medical College	[{'id': '4bf58dd8d48988d196941735', 'name': 'H...	v-1587910011	False	22.322726	91.806307	[{'label': 'display', 'lat': 22.32272571735471...	3712	BD
3	4f5606d0e4b088020c2ca4ea	College Lab	[{'id': '4bf58dd8d48988d1a5941735',	v-1587910011	False	22.341265	91.829796	[{'label': 'display', 'lat':	1660	BD

Then we get all the data of universities and schools in the same way.

```
In [228]: venues = results['response']['venues']
          dataframe2 = json_normalize(venues)
          dataframe2
```

	id	name	categories	referralId	hasPerk	location.address	location.lat	location.lng	location.labeledLatLngs
0	52c152e2498ea03ae1a35d47	Premier University	[{'id': '4bf58dd8d48988d1ae941735', 'name': 'U...	v-1587910030	False	Probortok Circle	22.335182	91.840813	[{'label': 'display', 'lat': 22.33518218994140...
1	511748dbe4b073158cd980ee	Islamia University College	[{'id': '4bf58dd8d48988d1af941735', 'name': 'C...	v-1587910030	False	Darogahat Road	22.321500	91.831588	[{'label': 'display', 'lat': 22.32149968761159...
2	526e51af11d2978e57351a36	Chittagong Independent University	[{'id': '4bf58dd8d48988d1ae941735', 'name': 'U...	v-1587910030	False	Jamal Khan	22.339855	91.826705	[{'label': 'display', 'lat': 22.339855, 'lng':...
3	50a49bc1e4b0f5b95fd1e4d5	independent university of bangladesh	[{'id': '4bf58dd8d48988d1a3941735', 'name': 'C...	v-1587910030	False	Jamal.khan	22.347139	91.833542	[{'label': 'display', 'lat': 22.34713935852050...
4	522cfe737e48fb57037b8b3d	International Islamic University Chittagong	[{'id': '4bf58dd8d48988d1ae941735', 'name': 'U...	v-1587910030	False	Chawbazar	22.355057	91.839437	[{'label': 'display', 'lat': 22.35505699355748...

Dropping unwanted values

```
In [231]: venues = results['response']['venues']
          dataframe3 = json_normalize(venues)
          dataframe3
```

	id	name	categories	referralId	hasPerk	location.lat	location.lng	location.labeledLatLngs	location.distance	location.cc
0	5285a255498eb56ea39d2c8f	Govt Muslim High School Jame Masjid	[{'id': '4bf58dd8d48988d138941735', 'name': 'M...	v-1587910041	False	22.335663	91.836736	[{'label': 'display', 'lat': 22.3356632660779...	715	BD
1	50ab1596e4b0c3da5996d90a	Aparnacharan Girls High School	[{'id': '4bf58dd8d48988d1ab941735', 'name': 'S...	v-1587910041	False	22.336693	91.831848	[{'label': 'display', 'lat': 22.33669281005859...	1172	BD
2	4f54cfc8e4b03be3658be846	GEMS English Medium School	[{'id': '4f4533814b9074f6e4fb0106', 'name': 'M...	v-1587910041	False	22.342064	91.827383	[{'label': 'display', 'lat': 22.34206444952175...	1903	BD
3	4f539d26e4b0bf6b5e6dbf02	Red Rose Kinder Garten School	[{'id': '4f4533814b9074f6e4fb0107', 'name': 'N...	v-1587910041	False	22.345003	91.827383	[{'label': 'display', 'lat': 22.34500299141766...	2132	BD
4	4ee3ba7f61aff5a3411b901f	Saint Placid's High School	[{'id': '4bf58dd8d48988d13d941735', 'name': 'H...	v-1587910041	False	22.332908	91.839003	[{'label': 'display', 'lat': 22.33290751556591...	332	BD

2. We will use foursquare API to find the venues of all the restaurants and cafes within Chittagong City.

```
In [235]: venues = results['response']['venues']
          dataframe4 = json_normalize(venues)
          dataframe4
```

	id	name	categories	referralId	hasPerk	location.lat	location.lng	location.labeledLatLngs	location.distance	location.cc
0	5a472e15646e381ec6d9d733	Hotel ABP Restaurant & Biriyani House	[{'id': '4bf58dd8d48988d1c4941735', 'name': 'R...	v-1587910068	False	22.333973	91.840130	[{'label': 'display', 'lat': 22.333973, 'lng':...	372	BD
1	58fcbea0ba4a64209cad6f89	LA GONDOLA RESTAURANT	[{'id': '4bf58dd8d48988d10f941735', 'name': 'I...	v-1587910068	False	22.329506	91.836430	[{'label': 'display', 'lat': 22.329506, 'lng':...	520	BD
2	506ae721e4b0e2923556bbce	Diamond Restaurant, Biponi Bitaan	[{'id': '4bf58dd8d48988d1c7941735', 'name': 'S...	v-1587910068	False	22.334524	91.832164	[{'label': 'display', 'lat': 22.33452448058991...	1026	BD
3	5d9ada2a88d7280007c171ea	Shri Gobindo Vegetarian Restaurant	[{'id': '4bf58dd8d48988d1d3941735', 'name': 'V...	v-1587910068	False	22.338545	91.836960	[{'label': 'display', 'lat': 22.338545, 'lng':...	970	BD
4	4f1c2854e4b0817d3d6d1de6	Tava Restaurant and Lounge	[{'id': '4bf58dd8d48988d142941735', 'name': 'A...	v-1587910068	False	22.362013	91.811987	[{'label': 'display', 'lat': 22.36201311104485...	4601	BD

```
In [242]: venues = results['response']['venues']
          dataframe5 = json_normalize(venues)
          dataframe5
```

	id	name	categories	referralId	hasPerk	location.address	location.crossStreet	location.lat	location.lng	location.labeledLat
0	510faa98e4b00af1c7ace70a	Cafe Ali Restaurant & Biriyani House	[{'id': '4bf58dd8d48988d16d941735', 'name': 'C...	v-1587910072	False	Tayeb Bhaban, Bayazid Bostami Rd.	Bayazid Bostami Rd.	22.366851	91.821646	[{'label': 'display' 22.36685106819
1	514bffb8e4b00b3e00e29dcf	Brost Cafe	[{'id': '4bf58dd8d48988d16e941735', 'name': 'F...	v-1587910072	False	Highway Plaza, Lalkhan Bazar, Ctg.	NaN	22.338400	91.850800	[{'label': 'display' 22.3384, 'lng
2	5406d07f498e87b5e7634afe	Café Milano	[{'id': '4bf58dd8d48988d16d941735', 'name': 'C...	v-1587910072	False	Premier University (Opposite of CSCR) Academic...	O.R. Nizam Road, Probartak,Chittagong.	22.359852	91.827817	[{'label': 'display' 22.35985239740
3	533c2d41498ebf53586324f4	Café Shawarma	[{'id': '4bf58dd8d48988d16d941735', 'name': 'C...	v-1587910072	False	NaN	NaN	22.346953	91.818542	[{'label': 'display' 22.34695269594

Dropping unwanted values

3. We will fix unwanted values from the data and filter it to make it usable and visualize current data.

Dropping values that doesnt indicate institution.

```
In [223]: dataframe = dataframe.drop([0,1,4,8,12,15,18,19,23,25,26,27,28,32])
```

```
In [224]: dataframe
```

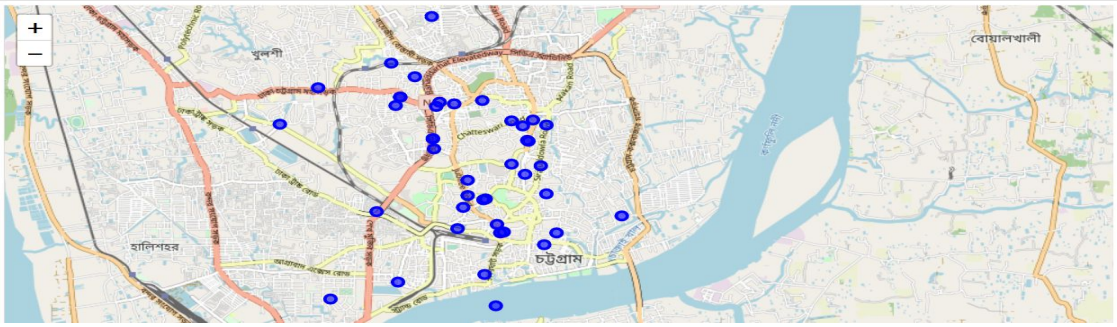
29	4ee503ede30005f8b867ea44	Dewanhat City Corporation College	[{'id': '4bf58dd8d48988d198941735', 'name': 'C...	v-1587910011	False	22.339130	91.813446	[{'label': 'display', 'lat': 22.33913040161133...	3012
30	515320f4e4b063e1c3b8f464	Chittagong Medical College	[{'id': '4bf58dd8d48988d198941735', 'name': 'C...	v-1587910011	False	22.359991	91.829536	[{'label': 'display', 'lat': 22.35999123453215...	3467
31	538f0ecc498e6cb5c9cb24df	Marine city Medical college	[{'id': '4bf58dd8d48988d1b3941735', 'name': 'M...	v-1587910011	False	22.359589	91.823015	[{'label': 'display', 'lat': 22.359589, 'lng':...	3716
33	5135e419e4b04e2c7b29338d	Chittagong Govt. Women College Playground	[{'id': '4bf58dd8d48988d1b8941735', 'name': 'C...	v-1587910011	False	22.364388	91.819351	[{'label': 'display', 'lat': 22.36438751220703...	4368
34	506578d3e4b0166bad932d1c	Presidency Int'l School & College	[{'id': '4bf58dd8d48988d198941735', 'name': 'C...	v-1587910011	False	22.375637	91.821894	[{'label': 'display', 'lat': 22.37563669432329...	5375

Now we will filter the food data to make it usable

In [245]:
```
filtered_columns4 = ['name', 'categories'] + [col for col in food_data.columns if col.startswith('location.')] + ['id']
food_filtered = food_data.loc[:, filtered_columns4]

# function that extracts the category of the venue
def get_category_type(row):
    try:
        categories_list = row['categories']
    except:
        categories_list = row['venue.categories']

    if len(categories_list) == 0:
        return None
    else:
        return categories_list[0]['name']

# filter the category for each row
food_filtered['categories'] = food_filtered.apply(get_category_type, axis=1)

# clean column names by keeping only last term
food_filtered.columns = [column.split('.')[-1] for column in food_filtered.columns]
food_filtered = pd.concat([food_filtered,type_col],axis =1)
food_filtered
```

In [251]:
```
edu_map = folium.Map(location=[latitude, longitude], zoom_start=13) # generate map centred around Chittagong city

for lat, lng, label in zip(edu_filtered.lat, edu_filtered.lng, edu_filtered.categories):
    folium.features.CircleMarker(
        [lat, lng],
        radius=5,
        color='blue',
        popup=label,
        fill = True,
        fill_color='blue',
        fill_opacity=0.6
    ).add_to(edu_map)

# display map
edu_map
```

Out[251]:

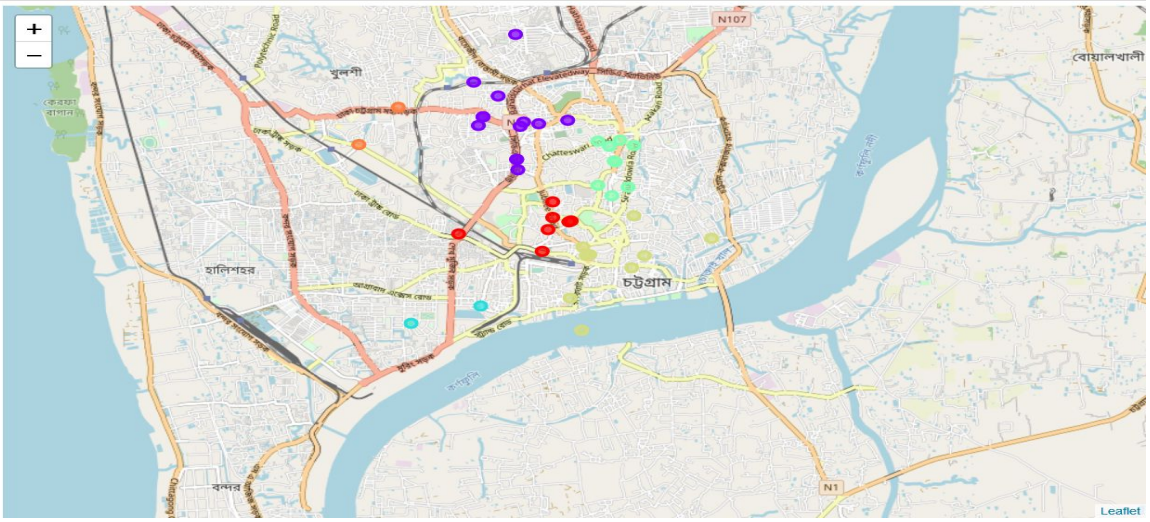


4. Using K Means clustering we will find the cluster where the educational institution is mostly dense.

In [254]:
```
kclusters = 7

edu_clustering = edu_filtered[['lat','lng']]

# run k-means clustering
edu_kmeans = KMeans(n_clusters=kclusters, random_state=0).fit(edu_clustering)

# check cluster labels generated for each row in the dataframe
edu_kmeans.labels_[0:10]
```

Out[254]:
```
array([3, 0, 2, 5, 1, 5, 5, 5, 0, 4])
```

Out[258]:



We can find that cluster 1 is the most dense.

```
In [259]: edu_filtered['Labels'].value_counts()
```

```
Out[259]: 1    13
          5     9
          4     9
          0     7
          6     2
          3     2
          2     1
          Name: Labels, dtype: int64
```

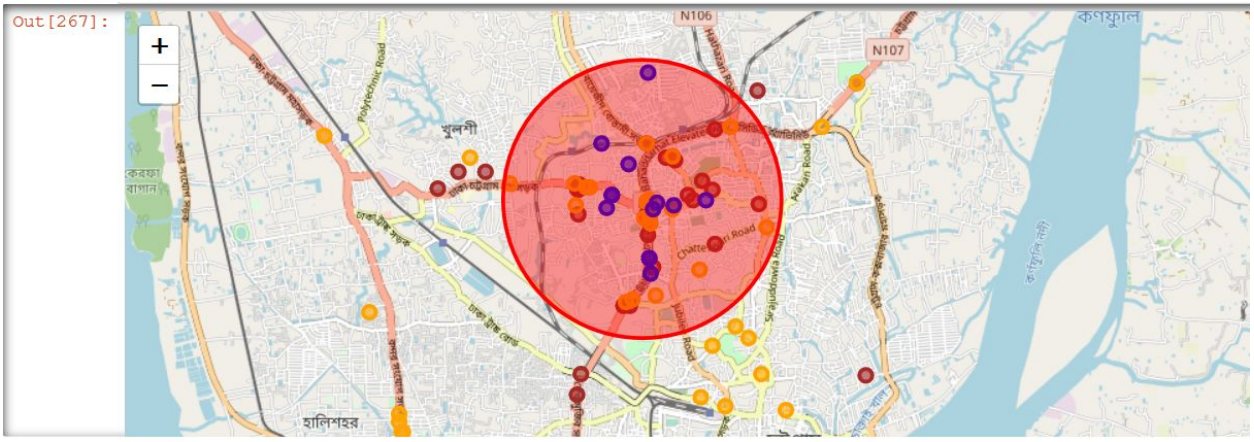
We found the area of cluster 1.

```
In [260]: inist =edu_filtered.groupby(['Labels','address'])['name']
          inist.first()
```

```
Out[260]: Labels  address
          0       Alpana Bhaban, Anandabag                                  Red Rose Kinder Garten School
                  Enayed Bazar Women College                                              College Lab
                  Ice Factory Road                                        Chittagong Collegiate School
                  Jamal Khan                                        Chittagong Independent University
                  Jubilee Road                                             GEMS English Medium School
                  Moniruzzaman Islamabadi Road, Enayet Bazar               Enayet Bazar Women College
                  Port City Complex, Dewanhat                       Dewanhat City Corporation College
          1       51 Panchlaish R/A                                 Presidency Int'l School & College
                  Chittagong Medical College                               Chittagong Medical College
                  Dampara                                                        Premier University
                  Dampara, Wasa                          Chittagong Mohila Shomiti School and College (...
                  Residential Area                                       Marine city Medical college
                  Zakir Hossain Rd                                  Chittagong Nasirabad Girls School
                  Zakir Hossain Rd, Chittagong                        Omargani MES University College
                  Zakir Hossain Road                               Ispahani Public School and College
          2       Chandgaon                              Chattagram International Medical college & Hos...
```

5. From the clusters we will locate all the food businesses and their type in the area.

Out[267]:



In [285]:
```
zonal_food_places = food_filtered[get_distance(food_filtered.lat,food_filtered.lng,latmean,lngmean)<=1.9]
```

In [300]:
```
zonal_food_places.reset_index(drop=True)
```

Out[300]:

	name	categories	lat	lng	labeledLatLngs	distance	cc	city	state	country	formattedAddress	address	crossStre
0	Tava Restaurant and Lounge	Asian Restaurant	22.362013	91.811987	[{'label': 'display', 'lat': 22.36201311104485...	4601	BD	NaN	NaN	বাংলাদেশ	[বাংলাদেশ]	NaN	N
1	Mazetto restaurant (Radisson Blu Hotel)	Italian Restaurant	22.348166	91.822903	[{'label': 'display', 'lat': 22.348166, 'lng':...	2705	BD	NaN	NaN	বাংলাদেশ	[বাংলাদেশ]	NaN	N
2	Karnafuly Chinese Restaurant	Chinese Restaurant	22.351328	91.828787	[{'label': 'display', 'lat': 22.351328, 'lng':...	2622	BD	NaN	NaN	বাংলাদেশ	[বাংলাদেশ]	NaN	N
3	Pitstop Restaurant	Thai Restaurant	22.347718	91.819828	[{'label': 'display', 'lat': 22.34771772832590...	2903	BD	NaN	NaN	বাংলাদেশ	[বাংলাদেশ]	NaN	N
4	Bangaliana Restaurant	Restaurant	22.347370	91.818924	[{'label': 'display', 'lat': 22.34737, 'lng': ...	2950	BD	NaN	NaN	বাংলাদেশ	[বাংলাদেশ]	NaN	N

6. Finally based on the type of popular food business in our chosen cluster we will answer the questions

In [297]:
```
pie_data = pd.DataFrame(zonal_food_places.type.value_counts().reset_index().values,columns=["Type","Count"])
pie_data = pie_data.sort_index(axis=0,ascending=True)
pie_data
```

Out[297]:

	Type	Count
0	Cafe	25
1	Restaurant	19

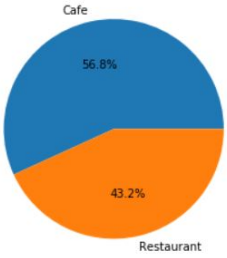
Result:

Now we can answer the questions of the problem:

1. We can say Cafe is the most popular type of restaurant among the students.

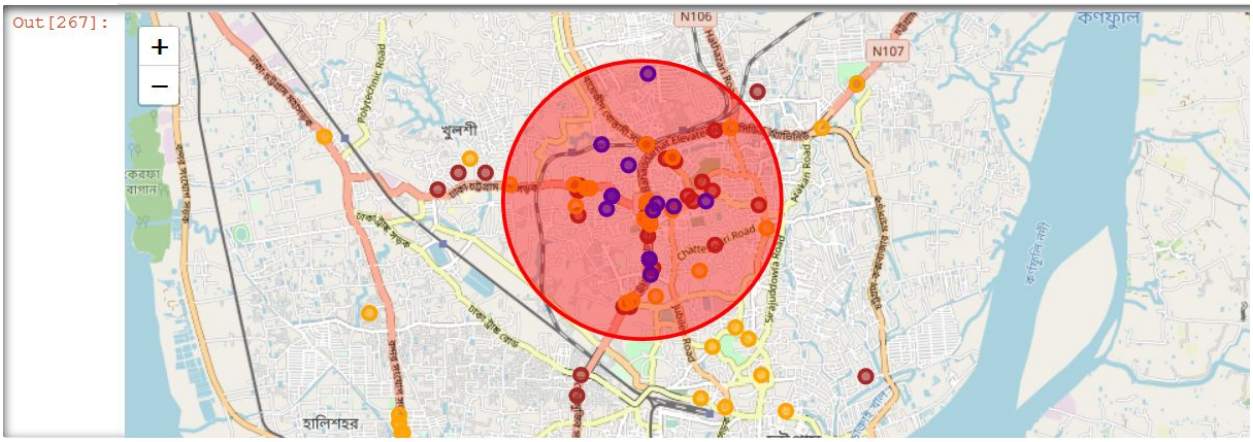
This pie chart shows Cafe type food business is the most popular

```
In [299]: my_labels = pie_data.Type
          plt.pie(pie_data.Count,labels=my_labels,autopct='%1.1f%%')
          plt.axis('equal')
          plt.show()
```



2. Because of the lack of good restaurants,different types of restaurant like chinese,italian,indian can be a potential in the city
3. Dampara , Zakir Hossain Rd has the most dense area with educational institutions
4. O.R Nizam road is the best place to open a food business.

Out[267]:



Discussion:

During this capstone project, the main problem was getting suitable data. There is basically no adequate data about this topic. So the project had to be completely based on Foursquare API information. Some of the data are also outdated so it may get hard to get the true assumptions.

Conclusion:

From the analysis we can conclude , if anyone wants to start a new food business, going with the cafe can provide the highest number of consumers. Also the location provided will also work as a factor in this subject.