# Analyzing the best location for a new restaurant

Lakshmi Narasaiah

IBM Data Science Certificate Capstone

# Data

To attempt a solution at this problem, we'll be using the Foursquare API to fetch venue information and analyze different neighborhoods.

To fulfill the requirements of the business problem stated, we'll specifically be looking for two types of information:

- Data on the neighborhoods in general, specifically:
  - The categories for each venue in each Manhattan neighborhood
  - Popular days and hours for venues that serve our target customer (nightlife venues)
  - A popularity score for each of the target venues, calculated by:
    - The total number of tips
    - The like/dislike volume and ratio
- Data on venues that might compete with ours, specifically:
  - The relative frequency of all late night restaurants
  - The price range of late night restaurants in each neighborhood
  - The cuisine type of late night restaurants in each neighborhood

# Data (fetching popular venues in a neighborhood)

```
In [15]: uniques = venues['Category'].value_counts()
         uniques
```

```
Out[15]: Chinese              9
         Bubble Tea           6
         Cocktail             4
         American             4
         Dim Sum              4
         Vietnamese           4
         Bakery               3
         Noodles              3
         Hotpot               3
         Salon / Barbershop   3
         Optical              2
         Bar                  2
         Spa                  2
         Coffee Shop          2
         Malay                2
         Asian                2
         Dumplings            2
         Ice Cream            2
         Sandwiches           2
         Boutique             1
         Cosmetics            1
```

# Data (Fetching the popular hours for a venue)

In [11]:
```
VENUE_ID = venues.loc[0, 'venue.id']
url = 'https://api.foursquare.com/v2/venues/{}?&client_id={}&client_secret={}&v={}'.format(
    VENUE_ID,
    CLIENT_ID,
    CLIENT_SECRET,
    VERSION,
    )

details = requests.get(url).json()
popularTimes = details['response']['venue']['popular']
popularTimes
```

Out[11]:
```
{'status': 'Likely open',
 'richStatus': {'entities': [], 'text': 'Likely open'},
 'isOpen': True,
 'isLocalHoliday': False,
 'timeframes': [{'days': 'Today',
   'includesToday': True,
   'open': [{'renderedTime': '8:00 AM–3:00 PM'},
    {'renderedTime': '6:00 PM–8:00 PM'}],
   'segments': []},
  {'days': 'Mon',
   'open': [{'renderedTime': '7:00 AM–8:00 AM'},
    {'renderedTime': '11:00 AM–Noon'},
    {'renderedTime': '3:00 PM–9:00 PM'}],
   'segments': []},
  {'days': 'Tue',
   'open': [{'renderedTime': '7:00 AM–9:00 AM'},
    {'renderedTime': '11:00 AM–Noon'},
    {'renderedTime': '5:00 PM–8:00 PM'}],
   'segments': []},
```

# Data (Fetching venue attributes)

```
VENUE_ID = '4db3374590a0843f295fb69b'
url = 'https://api.foursquare.com/v2/venues/{}?&client_id={}&client_secret={}&v={}'.format(
    VENUE_ID,
    CLIENT_ID,
    CLIENT_SECRET,
    VERSION,
    )
request2 = requests.get(url).json()
attributes = request2['response']['venue']['attributes']
attributes
```

```
{'groups': [{'type': 'price',
   'name': 'Price',
   'summary': '$',
   'count': 1,
   'items': [{'displayName': 'Price', 'displayValue': '$', 'priceTier': 1}]},
  {'type': 'reservations',
   'name': 'Reservations',
   'count': 3,
   'items': [{'displayName': 'Reservations', 'displayValue': 'No'}]},
  {'type': 'payments',
   'name': 'Credit Cards',
   'summary': 'No Credit Cards',
   'count': 7,
   'items': [{'displayName': 'Credit Cards', 'displayValue': 'No'}]},
  {'type': 'outdoorSeating',
   'name': 'Outdoor Seating',
   'count': 1,
   'items': [{'displayName': 'Outdoor Seating', 'displayValue': 'No'}]},
  {'type': 'serves',
   'name': 'Menus',
   'summary': 'Dinner, Lunch & more',
   'count': 8,
```

# Data (fetching number of tips)

In [29]:
```
VENUE_ID = '4db3374590a0843f295fb69b'
url = 'https://api.foursquare.com/v2/venues/{}?&client_id={}&client_secret={}&v={}'.format(
    VENUE_ID,
    CLIENT_ID,
    CLIENT_SECRET,
    VERSION,
    )
request2 = requests.get(url).json()
attributes = request2['response']['venue']['tips']
attributes
```

Out[29]:
```
{'count': 172,
 'groups': [{'type': 'others',
   'name': 'All tips',
   'count': 172,
   'items': [{'id': '4df167a3b0fb807158b979f8',
     'createdAt': 1307666339,
     'text': 'Big tray chicken. Make sure you ask them to add hand-pulled noodles to it.',
     'type': 'user',
     'canonicalUrl': 'https://foursquare.com/item/4df167a3b0fb807158b979f8',
     'lang': 'en',
```

# Data (methodology)

After fetching all the required data, I'll be using clustering to analyze neighborhoods based on the following information:

- Popular hours and days for venues in the area
- Popularity scores for venues that complement our restaurant (nightlife venues)
- The relative frequency of restaurants that have similar price range, popular hours, and cuisine to ours

The end goal will be to find a cluster of neighborhoods that seem like a good fit for our specific restaurant. The ideal neighborhood will have:

- Popular nightlife venues with a lot of foot traffic
- A low number of restaurants that may compete with ours