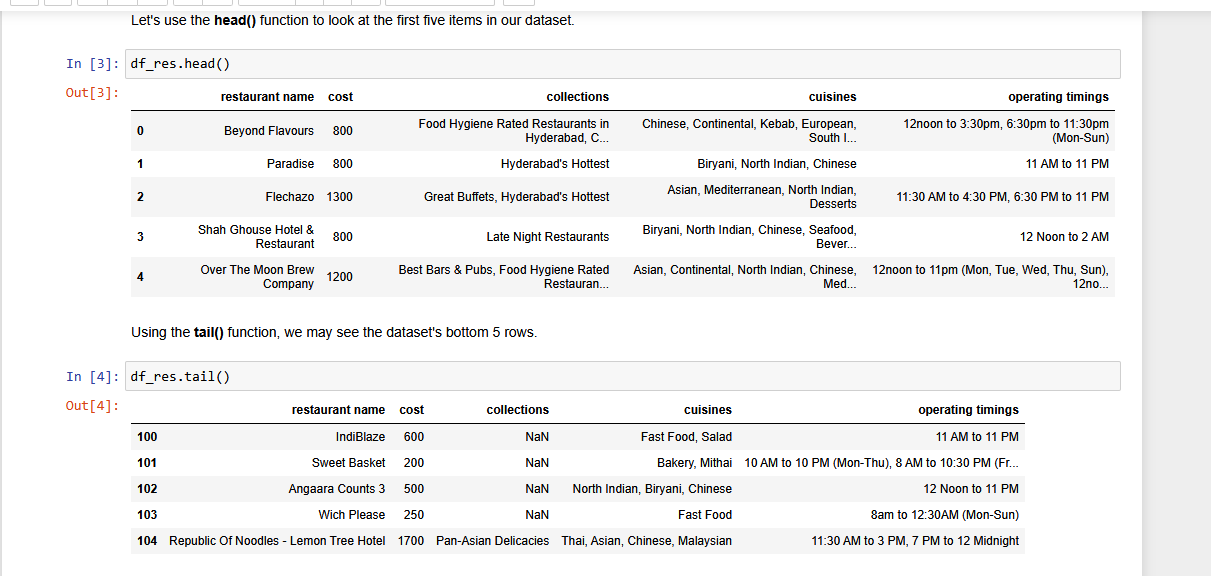
# Assessment Report

# **Demonstrate several NLP use cases and data insights to have a better grasp of customer feedback of restaurants**

## Objectives

1. You need to perform all the tasks from data preprocessing to exploratory data analysis On a given dataset
2. Perform sentiment analysis on a given dataset, Visualize the task to get a better overall understanding of people’s opinions.

#### **Download the dataset and read it into a pandas dataframe**



**Some Insights from the data**

* From this dataset, I have extracted some insights that will be used for customer feedback. They are as follows:
* Cost 500 has the most value count. It means most of the customers prefer the low cost cuisines
* 11 AM to 11 PM is the most visited timing for restaurants
* In that relationship data points are scattered between the range of 200 to 2000.
* It means the cost of the different cuisines is varying in between 200 to 2000.
* only 1 cost is too high which is greater than 2500 it relates I guess to Asian, Italian, continental cuisines
* Majority of the data points are between the value 0 to thousand and very few are greater than 1000.
* It means the cost of the restaurants are between 0 to 1000 mostly and very rare are soo expensive as they are exceeding the cost of 1000
* American cuisines are more costly than other cuisines
  + 10 Downing Street is much expensive amongst the rest of the restaurants cost=1900 and eat.fit is most less expensive restaurant cost = 500 but it also depends on the cuisines

Task no 2

# Data set: 2



Some Insights and outcomes from this data set

* Maximum number of value count in **5 star** rating and least number of values indicates to **floating** numbers of ratings. It means most of the customers rated 5 star because they liked the service of the restaurant. Whereas least number of customers have given floating numbers. Furthermore, 1 star of customers are more than 2 and 3 star ratings
* 5 star rating will be called as **positive** reviews although 1 and 2 stars will be called as **negative** reviews
* We use \*\*textblob\*\* function to analyze the polarity and the subjectivity of sentiments of the reviews of the customers
* The value of polarity can be between the -1 to 1.
* The reviews with the Negative polarity values shows the negative sentiments
* The reviews with the positive polarity values indicates the positive sentiments
* Subjectivity value can be rise between 0 and 1
* It quantifies the amount of personal opinion and the actual information contained in that text
* The higher the value of subjectivity contains the review of the personal opinion of that particular customer
* The value nears zero consists of the actual information of the review not the personal opinion of the particular customer
* This graph indicates that the highest number of customer-reviewed neutral sentiment polarity then most of them contain positive sentiment and few contain negative sentiments
* This 150th review contain positive sentiments due to the positive value of polarity also the personal opinion in review more than 50%.
* Like has maximum sentiment polarity than 5 star. The sentiment polarity has very diverse values according to their ratings. 1 has negative sentiments and they are decreasing monotonically
* These are mostly negative reviews or sentiments, we can say that it is having a negative polarity. If the polarity is less than zero it is containing negative sentiments
* If the polarity is greater than zero it is containing positive sentiments