In [1]:

```
import pandas as pd

df = pd.read_csv("tips.csv")
df.head()
```

Out[1]:

	total_bill	tip	sex	smoker	day	time	size
0	16.99	1.01	Female	No	Sun	Dinner	2
1	10.34	1.66	Male	No	Sun	Dinner	3
2	21.01	3.50	Male	No	Sun	Dinner	3
3	23.68	3.31	Male	No	Sun	Dinner	2
4	24.59	3.61	Female	No	Sun	Dinner	4

Data Understanding

- **total_bill:** total bill at restaurant
- **tip:** tip given to the waiter
- **sex:** gender of that person (who paid the bill)
- **smoker:** whether the person is smoker or non-smoker
- **day:** which day they came to restaurant
- **time:** which time they came to restaurant
- **size:** total no.of members came to restaurant

Dataset Understanding

In [3]:

```
df.shape
```

Out[3]:

```
(244, 7)
```

In [4]:

```
1 df.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 244 entries, 0 to 243
Data columns (total 7 columns):
 #   Column      Non-Null Count  Dtype  
---  ------      --------------  -----  
 0   total_bill  244 non-null    float64
 1   tip         244 non-null    float64
 2   sex         244 non-null    object 
 3   smoker      244 non-null    object 
 4   day         244 non-null    object 
 5   time        244 non-null    object 
 6   size        244 non-null    int64  
dtypes: float64(2), int64(1), object(4)
memory usage: 13.5+ KB
```

In [2]:

```
1 continous= ['total_bill','tip']
2 
3 discrete = ['sex','smoker','day','time','size']
4         # discrete_categorical = ['sex','smoker','day','time']
5         # discrete_count = ['size']
```