

```
In [1]: import pandas as pd
import numpy as np
```

```
In [2]: xls = pd.ExcelFile('yelp.xlsx')
```

```
In [5]: df = xls.parse('yelp_data')
```

```
In [6]: df.head()
```

```
Out[6]:
```

	name	category_0	category_1	take_out	review_count	stars	city_id	state_id
0	China Sea Chinese Restaurant	Restaurants	Chinese	True	11	2.5	1	1
1	Discount Tire Center	Tires	Automotive	False	24	4.5	1	1
2	Frankfurters	Restaurants	Hot Dogs	True	3	4.5	1	1
3	Fred Dietz Floral	Shopping	Flowers & Gifts	False	6	4.0	1	1
4	Kuhn's Market	Food	Grocery	False	8	3.5	1	1

```
In [12]: df.apply(np.min)
```

```
Out[12]: name          1-Day Paint & Body Centers
category_0          Active Life
category_1          Accountants
take_out            False
review_count         3
stars              1.0
city_id             1
state_id            1
dtype: object
```

```
In [13]: df.apply(np.max)
```

```
Out[13]: name          e.b. Pepper
category_0    Women's Clothing
category_1          Yoga
take_out      True
review_count   1305
stars         5.0
city_id       13
state_id      2
dtype: object
```

```
In [15]: df.name.count()
```

```
Out[15]: 600
```

```
In [16]:
```

```
df.category_0.count()
```

Out[16]: 600

```
In [17]: df.count()
```

```
Out[17]: name          600
category_0      600
category_1      600
take_out        600
review_count    600
stars           600
city_id         600
state_id        600
dtype: int64
```

```
In [18]: df.stars.max()
```

Out[18]: 5.0

```
In [19]: df.review_count.max()
```

Out[19]: 1305

```
In [20]: df.stars.min()
```

Out[20]: 1.0

```
In [21]: df.review_count.min()
```

Out[21]: 3

```
In [22]: df.stars.mean()
```

Out[22]: 3.495

```
In [23]: df.review_count.mean()
```

Out[23]: 33.77166666666667

```
In [24]: df.mean()
```

```
Out[24]: take_out      0.323333
review_count  33.771667
stars         3.495000
city_id       9.193333
state_id      1.500000
dtype: float64
```

```
In [25]:
```

```
df.median()
```

```
Out[25]: take_out      0.0
review_count  10.0
stars         3.5
city_id       10.5
state_id       1.5
dtype: float64
```

```
In [26]: df.stars.mode()
```

```
Out[26]: 0    3.5
dtype: float64
```

```
In [27]: df.review_count.mode()
```

```
Out[27]: 0    3
dtype: int64
```

```
In [28]: df.mode()
```

```
Out[28]:
```

	name	category_0	category_1	take_out	review_count	stars	city_id	state_id
0	Giant Eagle	Shopping	Pizza	False	3.0	3.5	8.0	1
1	NaN	NaN	NaN	NaN	NaN	NaN	NaN	2

```
In [30]: df.stars.std()
```

```
Out[30]: 0.9555959428597502
```

```
In [31]: df.review_count.std()
```

```
Out[31]: 86.9018954316009
```

```
In [32]: df.std()
```

```
Out[32]: take_out      0.468139
review_count  86.901895
stars         0.955596
city_id       2.997933
state_id       0.500417
dtype: float64
```

```
In [33]: df.describe()
```

```
Out[33]:
```

	review_count	stars	city_id	state_id
count	600.000000	600.000000	600.000000	600.000000
mean	33.771667	3.495000	9.193333	1.500000
std	86.901895	0.955596	2.997933	0.500417

	review_count	stars	city_id	state_id
<b>min</b>	3.000000	1.000000	1.000000	1.000000
<b>25%</b>	5.000000	3.000000	8.000000	1.000000
<b>50%</b>	10.000000	3.500000	10.500000	1.500000
<b>75%</b>	25.250000	4.000000	12.000000	2.000000
<b>max</b>	1305.000000	5.000000	13.000000	2.000000

In [34]:

df.info()

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 600 entries, 0 to 599
Data columns (total 8 columns):
#   Column          Non-Null Count  Dtype
---  -
0   name             600 non-null   object
1   category_0       600 non-null   object
2   category_1       600 non-null   object
3   take_out         600 non-null   bool
4   review_count     600 non-null   int64
5   stars            600 non-null   float64
6   city_id          600 non-null   int64
7   state_id         600 non-null   int64
dtypes: bool(1), float64(1), int64(3), object(3)
memory usage: 33.5+ KB
```

In [35]:

df.stars.sum()

Out[35]: 2097.0

In [37]:

df.review\_count.sum()

Out[37]: 20263

In [38]:

df.sum()

```
Out[38]: name           China Sea Chinese RestaurantDiscount Tire Cent...
category_0  RestaurantsTiresRestaurantsShoppingFoodFoodRes...
category_1  ChineseAutomotiveHot DogsFlowers & GiftsGrocer...
take_out    194
review_count 20263
stars        2097.0
city_id      5516
state_id     900
dtype: object
```

In [ ]: