In [1]:

```
import pandas as pd
```

In [2]:

```
url = 'https://raw.githubusercontent.com/justmarkham/DAT8/master/data/chipotle.tsv'
chipo = pd.read_csv(url, sep = '\t')
```

In [3]:

```
chipo.head(10)
```

Out[3]:

	order_id	quantity	item_name	choice_description	item_price
0	1	1	Chips and Fresh Tomato Salsa	NaN	$2.39
1	1	1	Izze	[Clementine]	$3.39
2	1	1	Nantucket Nectar	[Apple]	$3.39
3	1	1	Chips and Tomatillo-Green Chili Salsa	NaN	$2.39
4	2	2	Chicken Bowl	[Tomatillo-Red Chili Salsa (Hot), [Black Beans...	$16.98
5	3	1	Chicken Bowl	[Fresh Tomato Salsa (Mild), [Rice, Cheese, Sou...	$10.98
6	3	1	Side of Chips	NaN	$1.69
7	4	1	Steak Burrito	[Tomatillo Red Chili Salsa, [Fajita Vegetables...	$11.75
8	4	1	Steak Soft Tacos	[Tomatillo Green Chili Salsa, [Pinto Beans, Ch...	$9.25
9	5	1	Steak Burrito	[Fresh Tomato Salsa, [Rice, Black Beans, Pinto...	$9.25

In [4]:

```
c = chipo.groupby('item_name').sum()
c = c.sort_values(['quantity'], ascending = False)
c.reset_index(inplace=True)
c.head(1)
#Q1
print("Most ordered item is",c.item_name[0])
#Q2
print("For the most-ordered item",c.quantity[0],"items were ordered")
```

```
Most ordered item is Chicken Bowl
For the most-ordered item 761 items were ordered
```

In [5]:

```
d = chipo.groupby('choice_description').sum()
d = d.sort_values(['quantity'], ascending = False)
d.reset_index(inplace=True)
d.head(1)
#Q3
print("The most ordered item in the choice_description is",d.choice_description[0])
```

```
The most ordered item in the choice_description is [Diet Coke]
```

In [6]:

```
#Q4
#print("Total",sum(chipo['quantity']),"quantities were ordered.")
print("Total",chipo.quantity.sum(),"quantities were ordered.")
```

```
Total 4972 quantities were ordered.
```

In [7]:

```
chipo.item_price.dtype
```

Out[7]:

```
dtype('O')
```

In [8]:

```
#that means item price is pandas object which is string
#convert to float
try:
    dollarizer = lambda x: float(x[1:-1])
    chipo.item_price = chipo.item_price.apply(dollarizer)

except:TypeError
```

In [9]:

```
#Q5
chipo.item_price.dtype
```

Out[9]:

```
dtype('float64')
```

In [10]:

```
chipo['revenue'] = chipo['quantity']*chipo['item_price']
chipo.head(5)
```

Out[10]:

	order_id	quantity	item_name	choice_description	item_price	revenue
0	1	1	Chips and Fresh Tomato Salsa	NaN	2.39	2.39
1	1	1	Izze	[Clementine]	3.39	3.39
2	1	1	Nantucket Nectar	[Apple]	3.39	3.39
3	1	1	Chips and Tomatillo-Green Chili Salsa	NaN	2.39	2.39
4	2	2	Chicken Bowl	[Tomatillo-Red Chili Salsa (Hot), [Black Beans...	16.98	33.96

In [11]:

```
#Q6
print("Total revenue for the period in the dataset is $", chipo.revenue.sum())
```

```
Total revenue for the period in the dataset is $ 39237.02
```

In [12]:

```
#Q7
print(chipo.order_id.value_counts().count(), "orders were made in the period")
```

```
1834 orders were made in the period
```

In [13]:

```
#Q8
print("The average revenue amount per order is",chipo.groupby(by=['order_id']).sum().mean()['revenue'
```

```
The average revenue amount per order is 21.394231188658654
```

In [14]:

```
#Q9
unique_item = chipo.item_name.unique()
```

In [15]:

```
print(len(unique_item), "different items were sold")
```

```
50 different items were sold
```

In [ ]:

```

```