In [1]:

```
import numpy as np
import matplotlib as mpl
import matplotlib.pyplot as plt
import pandas as pd
```

In [6]:

```
url = 'https://raw.githubusercontent.com/justmarkham/DAT8/master/data/chipotle.tsv'
```

In [7]:

```
# read the above file
```

In [14]:

```
df = pd.read_csv(url, sep = '\t')
```

In [9]:

```
#print first 5 and last 7 records
```

In [15]:

```
df.head(5)
```

Out[15]:

	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

In [16]:

```
df.tail(7)
```

Out[16]:

	order_id	quantity	item_name	choice_description	item_price
4615	1832	1	Chicken Soft Tacos	[Fresh Tomato Salsa, [Rice, Cheese, Sour Cream]]	$8.75
4616	1832	1	Chips and Guacamole	NaN	$4.45
4617	1833	1	Steak Burrito	[Fresh Tomato Salsa, [Rice, Black Beans, Sour ...	$11.75
4618	1833	1	Steak Burrito	[Fresh Tomato Salsa, [Rice, Sour Cream, Cheese...	$11.75
4619	1834	1	Chicken Salad Bowl	[Fresh Tomato Salsa, [Fajita Vegetables, Pinto...	$11.25
4620	1834	1	Chicken Salad Bowl	[Fresh Tomato Salsa, [Fajita Vegetables, Lettu...	$8.75
4621	1834	1	Chicken Salad Bowl	[Fresh Tomato Salsa, [Fajita Vegetables, Pinto...	$8.75

In [12]:

```
# print total records and type of variables
```

In [17]:

```
df.info()#

# OR

df.shape[0]
# 4622 observations
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 4622 entries, 0 to 4621
Data columns (total 5 columns):
order_id              4622 non-null int64
quantity              4622 non-null int64
item_name             4622 non-null object
choice_description    3376 non-null object
item_price            4622 non-null object
dtypes: int64(2), object(3)
memory usage: 180.6+ KB
```

Out[17]:

```
4622
```

In [18]:

```
#Print the name of all the columns.
```

In [19]:

```
df.columns
```

Out[19]:

```
Index(['order_id', 'quantity', 'item_name', 'choice_description',
       'item_price'],
      dtype='object')
```

In [20]:

```
#How is the dataset indexed?
```

In [21]:

```
df.index
```

Out[21]:

```
RangeIndex(start=0, stop=4622, step=1)
```

In [22]:

```
#Which was the most ordered item? and How many items were ordered?
```

In [23]:

```
c = df.groupby('item_name')
c = c.sum()
c = c.sort_values(['quantity'], ascending=False)
c.head(1)
```

Out[23]:

	order_id	quantity
item_name		
Chicken Bowl	713926	761

In [24]:

```
#What was the most ordered item in the choice_description column?
```

In [25]:

```
c = df.groupby('choice_description').sum()
c = c.sort_values(['quantity'], ascending=False)
c.head(1)
```

Out[25]:

	order_id	quantity
choice_description		
[Diet Coke]	123455	159

In [26]:

```
#Turn the item price into a float
```

In [27]:

```
dollar = lambda x: float(x[1:-1])
df.item_price = df.item_price.apply(dollar)
```

In [28]:

```
#How much was the revenue for the period in the dataset?
```

In [30]:

```
revenue = (df['quantity']* df['item_price']).sum()

print('Revenue was: $' + str(np.round(revenue,2)))
```

```
Revenue was: $39237.02
```

In [31]:

```
#print a data frame with only two columns item_name and item_price
```

In [32]:

```
# delete the duplicates in item_name and quantity
filtered = df.drop_duplicates(['item_name','quantity'])

# select only the products with quantity equals to 1
one_prod = filtered[filtered.quantity == 1]

# select only the item_name and item_price columns
price_per_item = one_prod[['item_name', 'item_price']]

# sort the values from the most to less expensive
price_per_item.sort_values(by = "item_price", ascending = False)
```

Out[32]:

	item_name	item_price
606	Steak Salad Bowl	11.89
1229	Barbacoa Salad Bowl	11.89
1132	Carnitas Salad Bowl	11.89
7	Steak Burrito	11.75
168	Barbacoa Crispy Tacos	11.75
39	Barbacoa Bowl	11.75
738	Veggie Soft Tacos	11.25
186	Veggie Salad Bowl	11.25
62	Veggie Bowl	11.25
57	Veggie Burrito	11.25
250	Chicken Salad	10.98
5	Chicken Bowl	10.98
8	Steak Soft Tacos	9.25
554	Carnitas Crispy Tacos	9.25
237	Carnitas Soft Tacos	9.25
56	Barbacoa Soft Tacos	9.25
92	Steak Crispy Tacos	9.25
664	Steak Salad	8.99
54	Steak Bowl	8.99
3750	Carnitas Salad	8.99
21	Barbacoa Burrito	8.99
27	Carnitas Burrito	8.99
33	Carnitas Bowl	8.99
11	Chicken Crispy Tacos	8.75
12	Chicken Soft Tacos	8.75
44	Chicken Salad Bowl	8.75
1653	Veggie Crispy Tacos	8.49
16	Chicken Burrito	8.49
1694	Veggie Salad	8.49
1414	Salad	7.40
510	Burrito	7.40
520	Crispy Tacos	7.40
673	Bowl	7.40

	item_name	item_price
298	6 Pack Soft Drink	6.49
10	Chips and Guacamole	4.45
1	Izze	3.39
2	Nantucket Nectar	3.39
674	Chips and Mild Fresh Tomato Salsa	3.00
111	Chips and Tomatillo Red Chili Salsa	2.95
233	Chips and Roasted Chili Corn Salsa	2.95
38	Chips and Tomatillo Green Chili Salsa	2.95
3	Chips and Tomatillo-Green Chili Salsa	2.39
300	Chips and Tomatillo-Red Chili Salsa	2.39
191	Chips and Roasted Chili-Corn Salsa	2.39
0	Chips and Fresh Tomato Salsa	2.39
40	Chips	2.15
6	Side of Chips	1.69
263	Canned Soft Drink	1.25
28	Canned Soda	1.09
34	Bottled Water	1.09

In [33]:

```
#What was the quantity of the most expensive item ordered?
```

In [34]:

```
df.sort_values(by = "item_price", ascending = False).head(1)
```

Out[34]:

	order_id	quantity	item_name	choice_description	item_price
3598	1443	15	Chips and Fresh Tomato Salsa	NaN	44.25

In [35]:

```
#How many times were a Veggie Salad Bowl ordered?
```

In [36]:

```
df[df.item_name == "Veggie Salad Bowl"]
```

Out[36]:

	order_id	quantity	item_name	choice_description	item_price
186	83	1	Veggie Salad Bowl	[Fresh Tomato Salsa, [Fajita Vegetables, Rice,...	11.25
295	128	1	Veggie Salad Bowl	[Fresh Tomato Salsa, [Fajita Vegetables, Lettu...	11.25
455	195	1	Veggie Salad Bowl	[Fresh Tomato Salsa, [Fajita Vegetables, Rice,...	11.25
496	207	1	Veggie Salad Bowl	[Fresh Tomato Salsa, [Rice, Lettuce, Guacamole...	11.25
960	394	1	Veggie Salad Bowl	[Fresh Tomato Salsa, [Fajita Vegetables, Lettu...	8.75
1316	536	1	Veggie Salad Bowl	[Fresh Tomato Salsa, [Fajita Vegetables, Rice,...	8.75
1884	760	1	Veggie Salad Bowl	[Fresh Tomato Salsa, [Fajita Vegetables, Rice,...	11.25
2156	869	1	Veggie Salad Bowl	[Tomatillo Red Chili Salsa, [Fajita Vegetables...	11.25
2223	896	1	Veggie Salad Bowl	[Roasted Chili Corn Salsa, Fajita Vegetables]	8.75
2269	913	1	Veggie Salad Bowl	[Fresh Tomato Salsa, [Fajita Vegetables, Rice,...	8.75
2683	1066	1	Veggie Salad Bowl	[Roasted Chili Corn Salsa, [Fajita Vegetables,...	8.75
3223	1289	1	Veggie Salad Bowl	[Tomatillo Red Chili Salsa, [Fajita Vegetables...	11.25
3293	1321	1	Veggie Salad Bowl	[Fresh Tomato Salsa, [Rice, Black Beans, Chees...	8.75
4109	1646	1	Veggie Salad Bowl	[Tomatillo Red Chili Salsa, [Fajita Vegetables...	11.25
4201	1677	1	Veggie Salad Bowl	[Fresh Tomato Salsa, [Fajita Vegetables, Black...	11.25
4261	1700	1	Veggie Salad Bowl	[Fresh Tomato Salsa, [Fajita Vegetables, Rice,...	11.25
4541	1805	1	Veggie Salad Bowl	[Tomatillo Green Chili Salsa, [Fajita Vegetabl...	8.75
4573	1818	1	Veggie Salad Bowl	[Fresh Tomato Salsa, [Fajita Vegetables, Pinto...	8.75

In [ ]:

```

```