# ▾ Chipotle

Browse through the data & understand the data

What are the number of entries in the data set

Print the name of all the column header in the dataset

How is the data set indexed

Which Item was the most ordered item

For the most ordered find quantity of items ordered

What was the most ordered item based on choose description column

How many items were ordered in total

Convert the item price into float

Create a lambada function and change the type of item price

What was the total revenue for the period of the dataset

How many orders were made in the period

What was the average revenue per order

How many different items are sold ( or on the menu)

```
1 #tips
2 #1 describne info
3 #2 shape
4 #4 df index
5 #5 filter count
6
```

loading dataset

```
1 import numpy as np
2 import pandas as pd
3 chip = pd.read_csv("Chipotle.tsv",sep='\t')
4 chip
```

| | 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 |

## ▾ Chipotle

Browse through the data & understand the data

Sour Cream, Cheese...

```
1 chip.describe()
```

| | order_id | quantity |
|---|---|---|
| count | 4622.000000 | 4622.000000 |
| mean | 927.254868 | 1.075725 |
| std | 528.890796 | 0.410186 |
| min | 1.000000 | 1.000000 |
| 25% | 477.250000 | 1.000000 |
| 50% | 926.000000 | 1.000000 |
| 75% | 1393.000000 | 1.000000 |
| max | 1834.000000 | 15.000000 |

```
1 chip.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 4622 entries, 0 to 4621
Data columns (total 5 columns):
 #   Column              Non-Null Count  Dtype
---  ------              --------------  -----
 0   order_id            4622 non-null   int64
 1   quantity            4622 non-null   int64
 2   item_name           4622 non-null   object
 3   choice_description  3376 non-null   object
 4   item_price          4622 non-null   object
dtypes: int64(2), object(3)
memory usage: 180.7+ KB
```

```
1 chip.shape[1]
```

5

## What are the number of entries in the data set

```
1 chip.shape
```

```
(4622, 5)
```

## Print the name of all the column header in the dataset

```
1 chip.columns
```

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

## How is the data set indexed

```
1 chip.index
```

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

## Which Item was the most ordered item

```
1    chip.item_name.describe()
```

```
count             4622
unique              50
top       Chicken Bowl
freq               726
Name: item_name, dtype: object
```

```
1 #sir method
2 mostordered = chip.groupby('item_name')['quantity'].count()
3 mostordered
```

```
item_name
6 Pack Soft Drink              54
Barbacoa Bowl                  66
Barbacoa Burrito               91
Barbacoa Crispy Tacos          11
Barbacoa Salad Bowl            10
Barbacoa Soft Tacos            25
Bottled Water                 162
Bowl                            2
Burrito                         6
Canned Soda                   104
Canned Soft Drink             301
Carnitas Bowl                  68
```

```
Carnitas Burrito                              59
Carnitas Crispy Tacos                          7
Carnitas Salad                                 1
Carnitas Salad Bowl                            6
Carnitas Soft Tacos                           40
Chicken Bowl                                 726
Chicken Burrito                              553
Chicken Crispy Tacos                          47
Chicken Salad                                  9
Chicken Salad Bowl                           110
Chicken Soft Tacos                           115
Chips                                        211
Chips and Fresh Tomato Salsa                 110
Chips and Guacamole                          479
Chips and Mild Fresh Tomato Salsa              1
Chips and Roasted Chili Corn Salsa            22
Chips and Roasted Chili-Corn Salsa            18
Chips and Tomatillo Green Chili Salsa         43
Chips and Tomatillo Red Chili Salsa           48
Chips and Tomatillo-Green Chili Salsa         31
Chips and Tomatillo-Red Chili Salsa           20
Crispy Tacos                                   2
Izze                                          20
Nantucket Nectar                              27
Salad                                          2
Side of Chips                                101
Steak Bowl                                   211
Steak Burrito                                368
Steak Crispy Tacos                            35
Steak Salad                                    4
Steak Salad Bowl                              29
Steak Soft Tacos                              55
Veggie Bowl                                   85
Veggie Burrito                                95
Veggie Crispy Tacos                            1
Veggie Salad                                   6
Veggie Salad Bowl                             18
Veggie Soft Tacos                              7
Name: quantity, dtype: int64
```

```
1 mostordered.sort_values(ascending=False)
```

```
item_name
Chicken Bowl                                 726
Chicken Burrito                              553
Chips and Guacamole                          479
Steak Burrito                                368
Canned Soft Drink                            301
Steak Bowl                                   211
Chips                                        211
Bottled Water                                162
Chicken Soft Tacos                           115
Chicken Salad Bowl                           110
Chips and Fresh Tomato Salsa                 110
Canned Soda                                  104
Side of Chips                                101
Veggie Burrito                                95
Barbacoa Burrito                              91
Veggie Bowl                                   85
```

```
Carnitas Bowl                                   68
Barbacoa Bowl                                   66
Carnitas Burrito                                59
Steak Soft Tacos                                55
6 Pack Soft Drink                               54
Chips and Tomatillo Red Chili Salsa             48
Chicken Crispy Tacos                            47
Chips and Tomatillo Green Chili Salsa           43
Carnitas Soft Tacos                             40
Steak Crispy Tacos                              35
Chips and Tomatillo-Green Chili Salsa           31
Steak Salad Bowl                                29
Nantucket Nectar                                27
Barbacoa Soft Tacos                             25
Chips and Roasted Chili Corn Salsa              22
Chips and Tomatillo-Red Chili Salsa             20
Izze                                            20
Veggie Salad Bowl                               18
Chips and Roasted Chili-Corn Salsa              18
Barbacoa Crispy Tacos                           11
Barbacoa Salad Bowl                             10
Chicken Salad                                    9
Carnitas Crispy Tacos                            7
Veggie Soft Tacos                                7
Burrito                                          6
Veggie Salad                                     6
Carnitas Salad Bowl                              6
Steak Salad                                      4
Bowl                                             2
Salad                                            2
Crispy Tacos                                     2
Chips and Mild Fresh Tomato Salsa                1
Carnitas Salad                                   1
Veggie Crispy Tacos                              1
Name: quantity, dtype: int64
```

## For the most ordered find quantity of items ordered

```
1   chip.item_name.describe()
```

```
count              4622
unique               50
top        Chicken Bowl
freq                726
Name: item_name, dtype: object
```

## Cleansed data by dropping null value

```
1 cleanchip = chip.dropna()
2 cleanchip
```

| | order_id | quantity | item_name | choice_description | item_price |
|---|---|---|---|---|---|
| **1** | 1 | 1 | Izze | [Clementine] | $3.39 |
| **2** | 1 | 1 | Nantucket Nectar | [Apple] | $3.39 |
| **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 |
| **7** | 4 | 1 | Steak Burrito | [Tomatillo Red Chili Salsa, [Fajita Vegetables... | $11.75 |
| **...** | ... | ... | ... | ... | ... |
| **4617** | 1833 | 1 | Steak Burrito | [Fresh Tomato Salsa, [Rice, Black Beans, Sour ... | $11.75 |

## What was the most ordered item based on choose description column

```
1 cleanchip.item_name.describe() #WRONG
```

```
count             3376
unique              38
top       Chicken Bowl
freq               726
Name: item_name, dtype: object
```

```
1 chip.choice_description
```

```
0                                               NaN
1                                      [Clementine]
2                                           [Apple]
3                                               NaN
4       [Tomatillo-Red Chili Salsa (Hot), [Black Beans...
                            ...
4617    [Fresh Tomato Salsa, [Rice, Black Beans, Sour ...
4618    [Fresh Tomato Salsa, [Rice, Sour Cream, Cheese...
4619    [Fresh Tomato Salsa, [Fajita Vegetables, Pinto...
4620    [Fresh Tomato Salsa, [Fajita Vegetables, Lettu...
4621    [Fresh Tomato Salsa, [Fajita Vegetables, Pinto...
Name: choice_description, Length: 4622, dtype: object
```

```
1 #sir method
2 choice  = chip.groupby("choice_description")['quantity'].sum()
3 choice.sort_values(ascending=False)
```

```
choice_description
[Diet Coke]
[Coke]
[Sprite]
[Fresh Tomato Salsa, [Rice, Black Beans, Cheese, Sour Cream, Lettuce]]
[Fresh Tomato Salsa, [Rice, Black Beans, Cheese, Sour Cream]]
```

```
[Roasted Chili Corn Salsa, [Fajita Vegetables, Rice, Pinto Beans, Guacamole, Lettuce]
[Roasted Chili Corn Salsa, [Fajita Vegetables, Rice, Sour Cream, Lettuce]]
[Roasted Chili Corn Salsa, [Fajita Vegetables, Sour Cream, Lettuce, Guacamole]]
[Roasted Chili Corn Salsa, [Guacamole, Sour Cream, Rice, Fajita Vegetables, Lettuce]
[[Tomatillo-Red Chili Salsa (Hot), Tomatillo-Green Chili Salsa (Medium)], [Rice, Pint
Name: quantity, Length: 1043, dtype: int64
```

## How many items were ordered in total

```
1 chip.item_name.describe() #wrong
```

```
count              4622
unique               50
top        Chicken Bowl
freq                726
Name: item_name, dtype: object
```

```
1 chip.quantity.describe() #wrong
```

```
count    4622.000000
mean        1.075725
std         0.410186
min         1.000000
25%         1.000000
50%         1.000000
75%         1.000000
max        15.000000
Name: quantity, dtype: float64
```

```
1 #SIR METHOD
2 sumtotal = chip.quantity.sum()
3 sumtotal
```

```
4972
```

## Convert the item price into float

```
1 chip.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 4622 entries, 0 to 4621
Data columns (total 5 columns):
 #   Column              Non-Null Count  Dtype
---  ------              --------------  -----
 0   order_id            4622 non-null   int64
 1   quantity            4622 non-null   int64
 2   item_name           4622 non-null   object
 3   choice_description  3376 non-null   object
 4   item_price          4622 non-null   object
dtypes: int64(2), object(3)
memory usage: 180.7+ KB
```

```
1 chip[['item_price']].info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 4622 entries, 0 to 4621
Data columns (total 1 columns):
 #   Column      Non-Null Count  Dtype
---  ------      --------------  -----
 0   item_price  4622 non-null   object
dtypes: object(1)
memory usage: 36.2+ KB
```

```
1 chip.item_price.dtype # means object
```

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

```
1 chip.head()
```

| | 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 |

```
1 conv = lambda x : float(x[1:])
2 #it is a function havr to apply to
3 conv
```

```
<function __main__.<lambda>>
```

```
1 chip['conv_price'] = chip.item_price.apply(conv)
```

```
1 chip.head()
```

| | order_id | quantity | item_name | choice_description | item_price | conv_price |
|---|---|---|---|---|---|---|
| **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 | NaN | $2.39 | 2.39 |

```
1 chip.conv_price.dtype
```

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

Create a lambada function and change the type of item price

```
1 #covered above
```

What was the total revenue for the period of the dataset

```
1 chip.conv_price.sum()
```

```
34500.16
```

how many orders were made in the period

```
1 chip['order_id'].unique()
```

```
1834
```

```
1 chip['order_id'].unique().max()
```

```
1834
```

Look for sir method

```
1
```

What was the average revenue per order

```
1   ordamt = chip.groupby("order_id")['conv_price'].sum()
2   ordamt
```

```
order_id
1        11.56
2        16.98
3        12.67
4        21.00
5        13.70
         ...
1830     23.00
1831     12.90
1832     13.20
1833     23.50
1834     28.75
Name: conv_price, Length: 1834, dtype: float64
```

```
1   ordamt.mean()
```

```
18.81142857142869
```

How many different items are sold ( or on the menu)

```
1  chip.item_name.describe()
```

```
count              4622
unique               50
top        Chicken Bowl
freq                726
Name: item_name, dtype: object
```