# das-exercises-filtering-sorting-1

August 10, 2024

## 1 Filtering and Sorting Data

### 1.0.1 Step 1. Import the necessary libraries

[64]:
```
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
import seaborn as sns
```

### 1.0.2 Step 2. Import the dataset *chipotle.tsv*

[65]:
```
chipo=pd.read_csv("chipotle.tsv", sep='\t')
```

### 1.0.3 Step 3. Assign it to a variable called chipo.

[66]:
```
chipo=df
```

[67]:
```
df
```

[67]:

| | order_id | quantity | item_name | choice_description | item_price |
|---|---|---|---|---|---|
| 0 | 1 | 1 | Chips and Fresh Tomato Salsa | NaN | 2 |
| 1 | 1 | 1 | Izze | [Clementine] | 3 |
| 2 | 1 | 1 | Nantucket Nectar | [Apple] | 3 |
| 3 | 1 | 1 | Chips and Tomatillo-Green Chili Salsa | NaN | 2 |
| 4 | 2 | 2 | Chicken Bowl | | |
| … | … | … | … | | |
| 4617 | 1833 | 1 | Steak Burrito | | |
| 4618 | 1833 | 1 | Steak Burrito | | |
| 4619 | 1834 | 1 | Chicken Salad Bowl | | |
| 4620 | 1834 | 1 | Chicken Salad Bowl | | |
| 4621 | 1834 | 1 | Chicken Salad Bowl | | |

```
4     [Tomatillo-Red Chili Salsa (Hot), [Black Beans…          16
…                                                   …         …
4617  [Fresh Tomato Salsa, [Rice, Black Beans, Sour …          11
4618  [Fresh Tomato Salsa, [Rice, Sour Cream, Cheese…          11
4619  [Fresh Tomato Salsa, [Fajita Vegetables, Pinto…          11
4620  [Fresh Tomato Salsa, [Fajita Vegetables, Lettu…           8
4621  [Fresh Tomato Salsa, [Fajita Vegetables, Pinto…           8

[4622 rows x 5 columns]
```

### 1.0.4 Step 4. How many products cost more than $10.00?

Use `str` attribute to remove the $ sign and convert the column to proper numeric type data before filtering.

[72]:
```
df['item_price']=df['item_price'].apply(lambda x: x.replace("$"," "))
```

[70]:
```
df
```

[70]:
```
      order_id  quantity                              item_name  \
0            1         1           Chips and Fresh Tomato Salsa
1            1         1                                   Izze
2            1         1                       Nantucket Nectar
3            1         1  Chips and Tomatillo-Green Chili Salsa
4            2         2                           Chicken Bowl
…          …       …                                       …
4617      1833         1                          Steak Burrito
4618      1833         1                          Steak Burrito
4619      1834         1                     Chicken Salad Bowl
4620      1834         1                     Chicken Salad Bowl
4621      1834         1                     Chicken Salad Bowl

                                  choice_description  item_price
0                                                NaN           2
1                                      [Clementine]           3
2                                           [Apple]           3
3                                              NaN           2
4     [Tomatillo-Red Chili Salsa (Hot), [Black Beans…          16
…                                                   …         …
4617  [Fresh Tomato Salsa, [Rice, Black Beans, Sour …          11
4618  [Fresh Tomato Salsa, [Rice, Sour Cream, Cheese…          11
4619  [Fresh Tomato Salsa, [Fajita Vegetables, Pinto…          11
4620  [Fresh Tomato Salsa, [Fajita Vegetables, Lettu…           8
4621  [Fresh Tomato Salsa, [Fajita Vegetables, Pinto…           8

[4622 rows x 5 columns]
```

[71]:
```
df['item_price']=df['item_price'].astype(int)
```

[54]:
```
df[df['item_price']>10]
```

[54]:
```
      order_id  quantity           item_name  \
4            2         2        Chicken Bowl
7            4         1       Steak Burrito
13           7         1        Chicken Bowl
39          19         1       Barbacoa Bowl
42          20         1        Chicken Bowl
…          …        …                  …
4610      1830         1       Steak Burrito
4611      1830         1      Veggie Burrito
4617      1833         1       Steak Burrito
4618      1833         1       Steak Burrito
4619      1834         1  Chicken Salad Bowl

                                  choice_description  item_price
4     [Tomatillo-Red Chili Salsa (Hot), [Black Beans…          16
7     [Tomatillo Red Chili Salsa, [Fajita Vegetables…          11
13    [Fresh Tomato Salsa, [Fajita Vegetables, Rice,…          11
39    [Roasted Chili Corn Salsa, [Fajita Vegetables,…          11
42    [Roasted Chili Corn Salsa, [Rice, Black Beans,…          11
…                                                  …         …
4610  [Fresh Tomato Salsa, [Rice, Sour Cream, Cheese…          11
4611  [Tomatillo Green Chili Salsa, [Rice, Fajita Ve…          11
4617  [Fresh Tomato Salsa, [Rice, Black Beans, Sour …          11
4618  [Fresh Tomato Salsa, [Rice, Sour Cream, Cheese…          11
4619  [Fresh Tomato Salsa, [Fajita Vegetables, Pinto…          11

[1025 rows x 5 columns]
```

### 1.0.5 Step 5. What is the price of each item?

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

[55]:
```
df[['item_name',"item_price"]]
```

[55]:
```
                              item_name  item_price
0          Chips and Fresh Tomato Salsa           2
1                                  Izze           3
2                      Nantucket Nectar           3
3  Chips and Tomatillo-Green Chili Salsa           2
4                          Chicken Bowl          16
…                                    …         …
4617                     Steak Burrito          11
4618                     Steak Burrito          11
```

```
4619                    Chicken Salad Bowl          11
4620                    Chicken Salad Bowl           8
4621                    Chicken Salad Bowl           8

[4622 rows x 2 columns]
```

### 1.0.6 Step 6. Sort by the name of the item

[56]:
```
df.sort_values(by="item_name")
```

[56]:
```
      order_id  quantity          item_name  \
3389      1360         2  6 Pack Soft Drink
341        148         1  6 Pack Soft Drink
1849       749         1  6 Pack Soft Drink
1860       754         1  6 Pack Soft Drink
2713      1076         1  6 Pack Soft Drink
…          …          …                 …
2384       948         1  Veggie Soft Tacos
781        322         1  Veggie Soft Tacos
2851      1132         1  Veggie Soft Tacos
1699       688         1  Veggie Soft Tacos
1395       567         1  Veggie Soft Tacos

                                     choice_description  item_price
3389                                        [Diet Coke]          12
341                                         [Diet Coke]           6
1849                                            [Coke]           6
1860                                        [Diet Coke]           6
2713                                            [Coke]           6
…                                                  …           …
2384  [Roasted Chili Corn Salsa, [Fajita Vegetables,…           8
781   [Fresh Tomato Salsa, [Black Beans, Cheese, Sou…           8
2851  [Roasted Chili Corn Salsa (Medium), [Black Bea…           8
1699  [Fresh Tomato Salsa, [Fajita Vegetables, Rice,…          11
1395  [Fresh Tomato Salsa (Mild), [Pinto Beans, Rice…           8

[4622 rows x 5 columns]
```

[57]:
```
df.reset_index()
```

[57]:
```
      index  order_id  quantity                              item_name  \
0         0         1         1           Chips and Fresh Tomato Salsa
1         1         1         1                                   Izze
2         2         1         1                       Nantucket Nectar
3         3         1         1  Chips and Tomatillo-Green Chili Salsa
4         4         2         2                           Chicken Bowl
…     …        …         …                                      …
```

```
4617   4617      1833         1                          Steak Burrito
4618   4618      1833         1                          Steak Burrito
4619   4619      1834         1                     Chicken Salad Bowl
4620   4620      1834         1                     Chicken Salad Bowl
4621   4621      1834         1                     Chicken Salad Bowl

                                     choice_description  item_price
0                                                   NaN           2
1                                           [Clementine]           3
2                                                [Apple]           3
3                                                   NaN           2
4     [Tomatillo-Red Chili Salsa (Hot), [Black Beans…          16
…                                                   …           …
4617  [Fresh Tomato Salsa, [Rice, Black Beans, Sour …          11
4618  [Fresh Tomato Salsa, [Rice, Sour Cream, Cheese…          11
4619  [Fresh Tomato Salsa, [Fajita Vegetables, Pinto…          11
4620  [Fresh Tomato Salsa, [Fajita Vegetables, Lettu…           8
4621  [Fresh Tomato Salsa, [Fajita Vegetables, Pinto…           8

[4622 rows x 6 columns]
```

### 1.0.7 Step 7. What was the quantity of the most expensive item ordered?

[78]:
```
df['item_price'].unique()
```

[78]:
```
array([ 2,  3, 16, 10,  1, 11,  9,  4,  8, 22, 17,  5, 23,  6, 32,  7, 18,
       35, 27, 26, 21, 13, 12, 44, 33, 15])
```

[79]:
```
a=df[df['item_price'] == df['item_price'].max()]
```

[80]:
```
a
```

[80]:
```
      order_id  quantity                    item_name choice_description  \
3598      1443        15  Chips and Fresh Tomato Salsa                NaN

      item_price
3598          44
```

[81]:
```
a['quantity']
```

[81]:
```
3598    15
Name: quantity, dtype: int64
```

### 1.0.8 Step 8. How many times were a Veggie Salad Bowl ordered?

[83]:
```
df['item_name'].value_counts()
```

[83]:
```
Chicken Bowl                            726
Chicken Burrito                         553
Chips and Guacamole                     479
Steak Burrito                           368
Canned Soft Drink                       301
Chips                                   211
Steak Bowl                              211
Bottled Water                           162
Chicken Soft Tacos                      115
Chips and Fresh Tomato Salsa            110
Chicken Salad Bowl                      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
Chips and Roasted Chili-Corn Salsa       18
Veggie Salad Bowl                        18
Barbacoa Crispy Tacos                    11
Barbacoa Salad Bowl                      10
Chicken Salad                             9
Veggie Soft Tacos                         7
Carnitas Crispy Tacos                     7
Burrito                                   6
Carnitas Salad Bowl                       6
Veggie Salad                              6
```

```
Steak Salad                               4
Crispy Tacos                              2
Bowl                                      2
Salad                                     2
Chips and Mild Fresh Tomato Salsa         1
Veggie Crispy Tacos                       1
Carnitas Salad                            1
Name: item_name, dtype: int64
```

[85]:
```
df[df.item_name == "Veggie Salad Bowl"].quantity.sum()
```

[85]:
```
18
```

### 1.0.9 Step 9. How many times people orderd more than one Canned Soda?

[100]:
```
a=df[ df['item_name'] == "Canned Soda"]
```

[105]:
```
a
```

[105]:

| | order_id | quantity | item_name | choice_description | item_price |
|---|---|---|---|---|---|
| 18 | 9 | 2 | Canned Soda | [Sprite] | 2 |
| 28 | 14 | 1 | Canned Soda | [Dr. Pepper] | 1 |
| 51 | 23 | 2 | Canned Soda | [Mountain Dew] | 2 |
| 53 | 24 | 1 | Canned Soda | [Sprite] | 1 |
| 107 | 47 | 1 | Canned Soda | [Dr. Pepper] | 1 |
| … | … | … | … | … | … |
| 3897 | 1562 | 1 | Canned Soda | [Mountain Dew] | 1 |
| 3926 | 1575 | 1 | Canned Soda | [Dr. Pepper] | 1 |
| 3936 | 1578 | 1 | Canned Soda | [Diet Dr. Pepper] | 1 |
| 4008 | 1604 | 1 | Canned Soda | [Diet Coke] | 1 |
| 4051 | 1621 | 1 | Canned Soda | [Sprite] | 1 |

[104 rows x 5 columns]

[103]:
```
a[a['quantity']>1]
```

[103]:

| | order_id | quantity | item_name | choice_description | item_price |
|---|---|---|---|---|---|
| 18 | 9 | 2 | Canned Soda | [Sprite] | 2 |
| 51 | 23 | 2 | Canned Soda | [Mountain Dew] | 2 |
| 162 | 73 | 2 | Canned Soda | [Diet Coke] | 2 |
| 171 | 76 | 2 | Canned Soda | [Diet Dr. Pepper] | 2 |
| 350 | 150 | 2 | Canned Soda | [Diet Coke] | 2 |
| 352 | 151 | 2 | Canned Soda | [Coca Cola] | 2 |
| 698 | 287 | 2 | Canned Soda | [Coca Cola] | 2 |
| 700 | 288 | 2 | Canned Soda | [Coca Cola] | 2 |
| 909 | 376 | 2 | Canned Soda | [Mountain Dew] | 2 |
| 1091 | 450 | 2 | Canned Soda | [Dr. Pepper] | 2 |

| | | | | | |
|---|---|---|---|---|---|
| 1092 | 450 | 2 | Canned Soda | [Coca Cola] | 2 |
| 1944 | 787 | 2 | Canned Soda | [Dr. Pepper] | 2 |
| 2135 | 859 | 2 | Canned Soda | [Diet Coke] | 2 |
| 2235 | 901 | 4 | Canned Soda | [Sprite] | 4 |
| 2905 | 1156 | 2 | Canned Soda | [Coca Cola] | 2 |
| 2906 | 1156 | 2 | Canned Soda | [Sprite] | 2 |
| 3152 | 1258 | 2 | Canned Soda | [Dr. Pepper] | 2 |
| 3364 | 1349 | 2 | Canned Soda | [Coca Cola] | 2 |
| 3592 | 1440 | 2 | Canned Soda | [Diet Coke] | 2 |
| 3866 | 1550 | 2 | Canned Soda | [Mountain Dew] | 2 |