

Ex1 - Filtering and Sorting Data

This time we are going to pull data directly from the internet. Special thanks to: <https://github.com/justmarkham> for sharing the dataset and materials.

Step 1. Import the necessary libraries

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

Step 2. Import the dataset from this [address](#).

```
In [2]: df = pd.read_table('.../DSML17_ASSIGNMENT/pandas/chipotle.tsv')
```

Step 3. Assign it to a variable called chipo.

```
In [3]: chipo = df
chipo
```

```
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
...
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

4622 rows x 5 columns

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

```
In [6]: chipo['price_float']=pd.to_numeric(chipo['item_price'].str.slice(1))
(chipo['price_float']>10).sum()

Out[6]: 1130

In [7]: chipo['price_float']=chipo.item_price.str.slice(1).astype(float)
(chipo['price_float']>10).sum()

Out[7]: 1130

In [6]: n=chipo.groupby('item_name').agg({'price_float':"max"})
(n['price_float']>10).sum()

Out[6]: 31
```

Step 5. What is the price of each item?

print a data frame with only two columns item_name and item_price

```
In [13]: chipo.groupby('item_name').agg('max')['price_float']

Out[13]:
```

item_name	price_float
6 Pack Soft Drink	12.98
Barbacoa Bowl	11.75
Barbacoa Burrito	11.75
Barbacoa Crispy Tacos	18.50
Barbacoa Salad Bowl	11.89
Barbacoa Soft Tacos	11.75
Bottled Water	15.00
Bowl	22.20
Burrito	7.40
Canned Soda	4.36
Canned Soft Drink	5.00
Carnitas Bowl	35.25
Carnitas Burrito	18.50
Carnitas Crispy Tacos	17.98
Carnitas Salad	8.99
Carnitas Salad Bowl	11.89
Carnitas Soft Tacos	11.75
Chicken Bowl	32.94
Chicken Burrito	35.00
Chicken Crispy Tacos	17.50
Chicken Salad	10.98
Chicken Salad Bowl	22.50
Chicken Soft Tacos	21.96
Chips	6.45
Chips and Fresh Tomato Salsa	44.25
Chips and Guacamole	17.80
Chips and Mild Fresh Tomato Salsa	3.00
Chips and Roasted Chili Corn Salsa	5.90
Chips and Roasted Chili-Corn Salsa	2.39
Chips and Tomatillo Green Chili Salsa	5.90
Chips and Tomatillo Red Chili Salsa	5.90
Chips and Tomatillo-Green Chili Salsa	4.78
Chips and Tomatillo-Red Chili Salsa	7.17
Crispy Tacos	7.40
Izze	3.39
Nantucket Nectar	6.78
Salad	7.40
Side of Chips	13.52
Steak Bowl	26.07
Steak Burrito	27.75
Steak Crispy Tacos	17.98
Steak Salad	8.99
Steak Salad Bowl	23.78
Steak Soft Tacos	18.50
Veggie Bowl	16.98
Veggie Burrito	33.75
Veggie Crispy Tacos	8.49
Veggie Salad	8.49
Veggie Salad Bowl	11.25
Veggie Soft Tacos	16.98

Name: price_float, dtype: float64

```
In [16]: chipo.groupby('item_name').agg({'price_float':"max"})

Out[16]:
```

price_float	
item_name	
6 Pack Soft Drink	12.98
Barbacoa Bowl	11.75
Barbacoa Burrito	11.75
Barbacoa Crispy Tacos	18.50
Barbacoa Salad Bowl	11.89
Barbacoa Soft Tacos	11.75
Bottled Water	15.00
Bowl	22.20
Burrito	7.40
Canned Soda	4.36
Canned Soft Drink	5.00
Carnitas Bowl	35.25
Carnitas Burrito	18.50
Carnitas Crispy Tacos	17.98
Carnitas Salad	8.99
Carnitas Salad Bowl	11.89
Carnitas Soft Tacos	11.75
Chicken Bowl	32.94
Chicken Burrito	35.00
Chicken Crispy Tacos	17.50
Chicken Salad	10.98
Chicken Salad Bowl	22.50
Chicken Soft Tacos	21.96
Chips	6.45
Chips and Fresh Tomato Salsa	44.25
Chips and Guacamole	17.80
Chips and Mild Fresh Tomato Salsa	3.00
Chips and Roasted Chili Corn Salsa	5.90
Chips and Roasted Chili-Corn Salsa	2.39
Chips and Tomatillo Green Chili Salsa	5.90
Chips and Tomatillo Red Chili Salsa	5.90
Chips and Tomatillo-Green Chili Salsa	4.78
Chips and Tomatillo-Red Chili Salsa	7.17
Crispy Tacos	7.40
Izze	3.39
Nantucket Nectar	6.78
Salad	7.40
Side of Chips	13.52
Steak Bowl	26.07
Steak Burrito	27.75
Steak Crispy Tacos	17.98
Steak Salad	8.99
Steak Salad Bowl	23.78
Steak Soft Tacos	18.50
Veggie Bowl	16.98
Veggie Burrito	33.75
Veggie Crispy Tacos	8.49
Veggie Salad	8.49
Veggie Salad Bowl	11.25
Veggie Soft Tacos	16.98

Step 6. Sort by the name of the item

```
In [19]: chipo.groupby('item_name').agg('max')

Out[19]:
```

order_id	quantity	choice_description	item_price	price_float
item_name				
6 Pack Soft Drink	1803	2	[Sprite]	\$6.49 12.98
Barbacoa Bowl	1825	1	[Tomatillo-Red Chili Salsa (Hot), Tomatillo-G...	\$9.25 11.75
Barbacoa Burrito	1827	1	[Tomatillo-Red Chili Salsa (Hot), Fresh Tomat...	\$9.25 11.75
Barbacoa Crispy Tacos	1330	2	[Roasted Chili Corn Salsa (Medium), Tomatillo...	\$9.25 18.50
Barbacoa Salad Bowl	1751	1	[Tomatillo Red Chili Salsa, [Black Beans, Chee...	\$9.39 11.89
Barbacoa Soft Tacos	1703	1	[Fresh Tomato Salsa (Mild), Roasted Chili Cor...	\$9.25 11.75
Bottled Water	1831	10	NaN	\$7.50 15.00
Bowl	279	3	[Braised Carnitas, Pinto Beans, [Sour Cream, C...	\$7.40 22.20
Burrito	313	1	[White Rice, Adobo-Marinated and Grilled Steak...	\$7.40 7.40
Canned Soda	1621	4	[Sprite]	\$4.36 4.36
Canned Soft Drink	1829	4	[Sprite]	\$5.00 5.00
Carnitas Bowl	1831	3	[Tomatillo-Red Chili Salsa (Hot), Fresh Tomat...	\$9.25 35.25
Carnitas Burrito	1821	2	[Tomatillo-Green Chili Salsa (Medium), Roaste...	\$9.25 18.50
Carnitas Crispy Tacos	1774	2	[Tomatillo-Red Chili Salsa (Hot), [Black Beans...	\$9.25 17.98
Carnitas Salad	1500	1	[Fresh Tomato Salsa (Mild), Roasted Chili Cor...	\$8.99 8.99
Carnitas Salad Bowl	1692	1	[Tomatillo Green Chili Salsa, [Rice, Pinto Bea...	\$9.39 11.89
Carnitas Soft Tacos	1811	1	[Tomatillo-Red Chili Salsa (Hot), [Cheese, Sou...	\$9.25 11.75
Chicken Bowl	1828	3	[Tomatillo-Red Chili Salsa (Hot), Tomatillo-G...	\$8.75 32.94
Chicken Burrito	1817	4	[Tomatillo-Red Chili Salsa (Hot), Tomatillo-G...	\$8.75 35.00
Chicken Crispy Tacos	1775	2	[Tomatillo-Green Chili Salsa (Medium), [Black ...	\$8.75 17.50
Chicken Salad	1598	1	[Tomatillo-Green Chili Salsa (Medium), Roaste...	\$8.49 10.98
Chicken Salad Bowl	1834	2	[Tomatillo Red Chili Salsa, [Rice, Fajita Vege...	\$8.75 22.50
Chicken Soft Tacos	1832	2	[Tomatillo-Red Chili Salsa (Hot), Roasted Chi...	\$8.75 21.96
Chips	1831	3	NaN	\$6.45 6.45
Chips and Fresh Tomato Salsa	1790	15	NaN	\$5.90 44.25
Chips and Guacamole	1832	4	NaN	\$8.90 17.80
Chips and Mild Fresh Tomato Salsa	279	1	NaN	\$3.00 3.00
Chips and Roasted Chili Corn Salsa	1513	2	NaN	\$5.90 5.90
Chips and Roasted Chili-Corn Salsa	1565	1	NaN	\$2.39 2.39
Chips and Tomatillo Green Chili Salsa	1787	2	NaN	\$5.90 5.90
Chips and Tomatillo Red Chili Salsa	1779	2	NaN	\$5.90 5.90
Chips and Tomatillo-Green Chili Salsa	1604	2	NaN	\$4.78 4.78
Chips and Tomatillo-Red Chili Salsa	1637	3	NaN	\$7.17 7.17
Crispy Tacos	217	1	[Adobo-Marinated and Grilled Steak]	\$7.40 7.40
Izze	1468	1	[Grapefruit]	\$3.39 3.39
Nantucket Nectar	1632	2	[Pomegranate Cherry]	\$6.78 6.78
Salad	575	1	[White Rice, Adobo-Marinated and Grilled Chick...	\$7.40 7.40
Side of Chips	1629	8	NaN	\$5.07 13.52
Steak Bowl	1796	3	[Tomatillo-Red Chili Salsa (Hot), Fresh Tomat...	\$9.25 26.07
Steak Burrito	1833	3	[Tomatillo-Green Chili Salsa (Medium), Tomati...	\$9.25 27.75
Steak Crispy Tacos	1810	2	[Roasted Chili Corn Salsa (Medium), Tomatillo...	\$9.25 17.98
Steak Salad	1406	1	[[Lettuce, Fajita Veggies]]	\$8.99 8.99
Steak Salad Bowl	1818	2	[Tomatillo Red Chili Salsa, [Rice, Black Beans...	\$9.39 23.78
Steak Soft Tacos	1808	2	[Tomatillo-Green Chili Salsa (Medium), Roaste...	\$9.25 18.50
Veggie Bowl	1780	2	[Tomatillo-Green Chili Salsa (Medium), Roaste...	\$8.75 16.98
Veggie Burrito	1830	3	[Tomatillo-Green Chili Salsa (Medium), Roaste...	\$8.75 33.75
Veggie Crispy Tacos	668	1	[Fresh Tomato Salsa (Mild), [Pinto Beans, Rice...	\$8.49 8.49
Veggie Salad	1635	1	[Tomatillo-Green Chili Salsa (Medium), Roaste...	\$8.49 8.49
Veggie Salad Bowl	1818	1	[Tomatillo Red Chili Salsa, [Fajita Vegetables...	\$8.75 11.25
Veggie Soft Tacos	1559	2	[Tomatillo Red Chili Salsa, [Fajita Vegetables...	\$8.75 16.98

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

```
In [162]: m = chipo['price_float'].max()
chipo[chipo['price_float']==m][['quantity','price_float']]

Out[162]:
```

quantity	price_float
3598	44.25

Step 8. How many times was a Veggie Salad Bowl ordered?

```
In [29]: chipo[chipo['item_name']=='Veggie Salad Bowl'].quantity.sum()

Out[29]: 18
```

Step 9. How many times did someone order more than one Canned Soda?

```
In [112]: chipo[chipo['item_name']=='Canned Soda'][chipo['quantity']>1].value_counts('quantity').sum()
C:\Users\shiva\AppData\Local\Temp\ipykernel_2516\4163584917.py:1: UserWarning: Boolean Series key will be reindexed to match DataFrame index.
  chipo[chipo['item_name']=='Canned Soda'][chipo['quantity']>1].value_counts('quantity').sum()

Out[112]: 20
```

Step 10. What is the order id of order with the highest total price of all items ordered in that order?

```
In [126]: m = chipo.groupby('order_id').sum()
m.max()

Out[126]:
```

quantity	price_float
35.00	205.25

dtype: float64

```
In [84]: m = chipo.groupby('order_id').agg({'price_float':'sum'})
m1 = m['price_float'].max()
m[m['price_float']==m1]

Out[84]:
```

price_float
order_id
926

205.25

Step 11. In How many orders, atleast one items is ordered multiple times?

```
In [4]: chipo.head()

Out[4]:
```

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

```
In [25]: a = chipo.groupby('order_id').agg('item_name').value_counts()
a[a.values>1].sum()

Out[25]: 695

In [ ]:
```

Step 12. Create a billing report (dataframe) which has columns as 'Order_Id', 'Total_Bill', 'Total Bill After Discount'. Restaurant owner has decided to give discount of 5% on orders of above \$50 total cost.

```
In [27]: df = chipo.groupby('order_id').agg({'price_float':'sum','order_id':'max'})
df['price_float'] = df['price_float'].apply(lambda x:x*0.95 if x>50 else x)
df['total_bill'] = df['price_float']
df['Total Bill After Discount'] = df['price_float']
df

Out[27]:
```

price_float	order_id	total_bill	Total Bill After Discount
order_id			
1	11.56	1	11.56
2	16.98	2	16.98
3	12.67	3	12.67
4	21.00	4	21.00
5	13.70	5	13.70
...
1830	23.00	1830	23.00
1831	12.90	1831	12.90
1832	13.20	1832	13.20
1833	23.50	1833	23.50
1834	28.75	1834	28.75

1834 rows x 4 columns

```
In [ ]:
```

```
In [ ]:
```