

# 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
```

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

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

```
In [2]: url = 'https://raw.githubusercontent.com/justmarkham/DAT8/master/data/chipotle.tsv'
chipo = pd.read_csv(url, sep = '\t')
```

```
In [3]: chipo.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
4	2	2	Chicken Bowl	[Tomatillo-Red Chili Salsa (Hot), [Black Beans...	\$16.98

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

```
In [4]: chipo["item_price"].dtypes
```

```
Out[4]: dtype('O')
```

```
In [5]: chipo["item_price"] = chipo["item_price"].str.split("$").apply(lambda x: float(x[-1]))
```

```
In [6]: chipo[chipo["item_price"] > 10]
```

	order_id	quantity	item_name	choice_description	item_price
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
13	7	1	Chicken Bowl	[Fresh Tomato Salsa, [Fajita Vegetables, Rice,...	11.25
23	12	1	Chicken Burrito	[[Tomatillo-Green Chili Salsa (Medium), Tomati...	10.98
...	...	...	...	...	...
4610	1830	1	Steak Burrito	[Fresh Tomato Salsa, [Rice, Sour Cream, Cheese...	11.75
4611	1830	1	Veggie Burrito	[Tomatillo Green Chili Salsa, [Rice, Fajita Ve...	11.25
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

1130 rows × 5 columns

```
In [7]: (chipo["item_price"] > 10).sum()
```

```
Out[7]: 1130
```

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

print a data frame with only two columns item\_name and item\_price

```
In [8]: chipo.drop_duplicates(["item_name"])
```

	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
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
10	5	1	Chips and Guacamole	NaN	4.45
11	6	1	Chicken Crispy Tacos	[Roasted Chili Corn Salsa, [Fajita Vegetables,...	8.75
12	6	1	Chicken Soft Tacos	[Roasted Chili Corn Salsa, [Rice, Black Beans,...	8.75
16	8	1	Chicken Burrito	[Tomatillo-Green Chili Salsa (Medium), [Pinto ...	8.49
18	9	2	Canned Soda	[Sprite]	2.18
21	11	1	Barbacoa Burrito	[[Fresh Tomato Salsa (Mild), Tomatillo-Green C...	8.99
27	14	1	Carnitas Burrito	[[Tomatillo-Green Chili Salsa (Medium), Roaste...	8.99
33	17	1	Carnitas Bowl	[Tomatillo-Red Chili Salsa (Hot), [Black Beans...	8.99
34	17	1	Bottled Water	NaN	1.09
38	18	1	Chips and Tomatillo Green Chili Salsa	NaN	2.95
39	19	1	Barbacoa Bowl	[Roasted Chili Corn Salsa, [Fajita Vegetables,...	11.75
40	19	1	Chips	NaN	2.15
44	20	1	Chicken Salad Bowl	[Fresh Tomato Salsa, [Fajita Vegetables, Pinto...	8.75
54	25	1	Steak Bowl	[Fresh Tomato Salsa (Mild), [Black Beans, Rice...	8.99
56	26	1	Barbacoa Soft Tacos	[Fresh Tomato Salsa, [Fajita Vegetables, Black...	9.25
57	26	1	Veggie Burrito	[Tomatillo Red Chili Salsa, [Fajita Vegetables...	11.25
62	28	1	Veggie Bowl	[Fresh Tomato Salsa, [Fajita Vegetables, Rice,...	11.25
92	40	1	Steak Crispy Tacos	[Fresh Tomato Salsa, Sour Cream]	9.25
111	49	1	Chips and Tomatillo Red Chili Salsa	NaN	2.95
168	75	1	Barbacoa Crispy Tacos	[Tomatillo Red Chili Salsa, [Rice, Black Beans...	11.75
186	83	1	Veggie Salad Bowl	[Fresh Tomato Salsa, [Fajita Vegetables, Rice,...	11.25
191	85	1	Chips and Roasted Chili-Corn Salsa	NaN	2.39
233	102	1	Chips and Roasted Chili Corn Salsa	NaN	2.95
237	103	1	Carnitas Soft Tacos	[Tomatillo Green Chili Salsa, [Fajita Vegetabl...	9.25
250	109	1	Chicken Salad	[Roasted Chili Corn Salsa (Medium), [Black Bea...	10.98
263	114	1	Canned Soft Drink	[Coke]	1.25
281	123	2	Steak Salad Bowl	[Tomatillo Red Chili Salsa, [Black Beans, Chee...	23.78
298	129	1	6 Pack Soft Drink	[Sprite]	6.49
300	130	1	Chips and Tomatillo-Red Chili Salsa	NaN	2.39
445	193	3	Bowl	[Braised Carnitas, Pinto Beans, [Sour Cream, C...	22.20
510	214	1	Burrito	[Adobo-Marinated and Grilled Chicken, Pinto Be...	7.40
520	217	1	Crispy Tacos	[Adobo-Marinated and Grilled Steak]	7.40
554	230	1	Carnitas Crispy Tacos	[Roasted Chili Corn Salsa]	9.25
664	276	1	Steak Salad	[Tomatillo-Red Chili Salsa (Hot), [Black Beans...	8.99
674	279	1	Chips and Mild Fresh Tomato Salsa	NaN	3.00
738	304	1	Veggie Soft Tacos	[Tomatillo Red Chili Salsa, [Fajita Vegetables...	11.25
1132	468	1	Carnitas Salad Bowl	[Fresh Tomato Salsa, [Rice, Black Beans, Chees...	11.89
1229	501	1	Barbacoa Salad Bowl	[Fresh Tomato Salsa, [Rice, Fajita Vegetables,...	11.89
1414	575	1	Salad	[Brown Rice, Adobo-Marinated and Grilled Chick...	7.40
1653	668	1	Veggie Crispy Tacos	[Fresh Tomato Salsa (Mild), [Pinto Beans, Rice...	8.49
1694	686	1	Veggie Salad	[[Fresh Tomato Salsa (Mild), Roasted Chili Cor...	8.49
3750	1500	1	Carnitas Salad	[[Fresh Tomato Salsa (Mild), Roasted Chili Cor...	8.99

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

```
In [9]: chipo["item_name"].sort_values()
```

```
Out[9]: 3389    6 Pack Soft Drink
341      6 Pack Soft Drink
1849    6 Pack Soft Drink
1860    6 Pack Soft Drink
2713    6 Pack Soft Drink
...
2384    Veggie Soft Tacos
781     Veggie Soft Tacos
2851    Veggie Soft Tacos
1699    Veggie Soft Tacos
1395    Veggie Soft Tacos
Name: item_name, Length: 4622, dtype: object
```

```
In [10]: chipo.sort_values(by="item_name")
```

	order_id	quantity	item_name	choice_description	item_price
3389	1360	2	6 Pack Soft Drink	[Diet Coke]	12.98
341	148	1	6 Pack Soft Drink	[Diet Coke]	6.49
1849	749	1	6 Pack Soft Drink	[Coke]	6.49
1860	754	1	6 Pack Soft Drink	[Diet Coke]	6.49
2713	1076	1	6 Pack Soft Drink	[Coke]	6.49
...	...	...	...	...	...
2384	948	1	Veggie Soft Tacos	[Roasted Chili Corn Salsa, [Fajita Vegetables,...	8.75
781	322	1	Veggie Soft Tacos	[Fresh Tomato Salsa, [Black Beans, Cheese, Sou...	8.75
2851	1132	1	Veggie Soft Tacos	[Roasted Chili Corn Salsa (Medium), [Black Bea...	8.49
1699	688	1	Veggie Soft Tacos	[Fresh Tomato Salsa, [Fajita Vegetables, Rice,...	11.25
1395	567	1	Veggie Soft Tacos	[Fresh Tomato Salsa (Mild), [Pinto Beans, Rice...	8.49

4622 rows × 5 columns

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

```
In [11]: chipo.sort_values(by="item_price", ascending=False).head(1)
```

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

```
In [12]: chipo["item_price"].sort_values(ascending=False)
chipo.loc[3598]
```

```
Out[12]: order_id          1443
quantity          15
item_name      Chips and Fresh Tomato Salsa
choice_description      NaN
item_price        44.25
Name: 3598, dtype: object
```

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

```
In [13]: chipo[chipo["item_name"]=="Veggie Salad Bowl"].count()
```

```
Out[13]: order_id          18
quantity          18
item_name          18
choice_description  18
item_price          18
dtype: int64
```

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

```
In [14]: chipo[(chipo["item_name"]=="Canned Soda") & (chipo["quantity"] >1)]
```

	order_id	quantity	item_name	choice_description	item_price
18	9	2	Canned Soda	[Sprite]	2.18
51	23	2	Canned Soda	[Mountain Dew]	2.18
162	73	2	Canned Soda	[Diet Coke]	2.18
171	76	2	Canned Soda	[Diet Dr. Pepper]	2.18
350	150	2	Canned Soda	[Diet Coke]	2.18
352	151	2	Canned Soda	[Coca Cola]	2.18
698	287	2	Canned Soda	[Coca Cola]	2.18
700	288	2	Canned Soda	[Coca Cola]	2.18
909	376	2	Canned Soda	[Mountain Dew]	2.18
1091	450	2	Canned Soda	[Dr. Pepper]	2.18
1092	450	2	Canned Soda	[Coca Cola]	2.18
1944	787	2	Canned Soda	[Dr. Pepper]	2.18
2135	859	2	Canned Soda	[Diet Coke]	2.18
2235	901	4	Canned Soda	[Sprite]	4.36
2905	1156	2	Canned Soda	[Coca Cola]	2.18
2906	1156	2	Canned Soda	[Sprite]	2.18
3152	1258	2	Canned Soda	[Dr. Pepper]	2.18
3364	1349	2	Canned Soda	[Coca Cola]	2.18
3592	1440	2	Canned Soda	[Diet Coke]	2.18
3866	1550	2	Canned Soda	[Mountain Dew]	2.18

```
In [15]: chipo[(chipo["item_name"]=="Canned Soda") & (chipo["quantity"] >1)].count()
```

```
Out[15]: order_id          20
quantity          20
item_name          20
choice_description  20
item_price          20
dtype: int64
```