

## ✓ 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

```
import pandas as pd
```

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

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

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

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

```
# Chuyển item_price về float
chipo['item_price'] = chipo['item_price'].str.replace('$', '').astype(float)

# Lọc sản phẩm có giá > 10
products_over_10 = chipo[chipo['item_price'] > 10]
print("Products over $10.00:", products_over_10.shape[0])
```

```
↳ Products over $10.00: 1130
```

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

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

```
# Giá mỗi item: lấy giá trung bình theo item_name
item_prices = chipo.groupby('item_name')['item_price'].mean().reset_index()
print(item_prices.head())
```

```
↳
```

	item_name	item_price
0	6 Pack Soft Drink	6.610185
1	Barbacoa Bowl	10.187273
2	Barbacoa Burrito	9.832418
3	Barbacoa Crispy Tacos	10.928182
4	Barbacoa Salad Bowl	10.640000

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

```
sorted_items = item_prices.sort_values(by='item_name')
print(sorted_items.head())
```

```
↳
```

	item_name	item_price
0	6 Pack Soft Drink	6.610185
1	Barbacoa Bowl	10.187273
2	Barbacoa Burrito	9.832418
3	Barbacoa Crispy Tacos	10.928182
4	Barbacoa Salad Bowl	10.640000

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

```
most_expensive = chipo.sort_values(by='item_price', ascending=False).iloc[0]
print("Most expensive item:", most_expensive['item_name'])
print("Quantity ordered:", most_expensive['quantity'])
```

↻ Most expensive item: Chips and Fresh Tomato Salsa  
Quantity ordered: 15

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

```
veggie_salad_count = chipo[chipo['item_name'] == 'Veggie Salad Bowl'].shape[0]  
print("Veggie Salad Bowl ordered:", veggie_salad_count)
```

↻ Veggie Salad Bowl ordered: 18

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

```
canned_soda_orders = chipo[(chipo['item_name'] == 'Canned Soda') & (chipo['quantity'] > 1)]  
print("Canned Soda ordered more than once:", canned_soda_orders.shape[0])
```

↻ Canned Soda ordered more than once: 20

