Visualizing Chipotle's 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
import matplotlib.pyplot as plt
from collections import Counter

# set this so the graphs open internally
%matplotlib inline
```

Step 2. Import the dataset from this address.

Step 3. Assign it to a variable called chipo.

```
chipo = pd.read_csv('https://raw.githubusercontent.com/thieu1995/csv-files/main/data/pandas/chipotle.tsv', sep='\t')
```

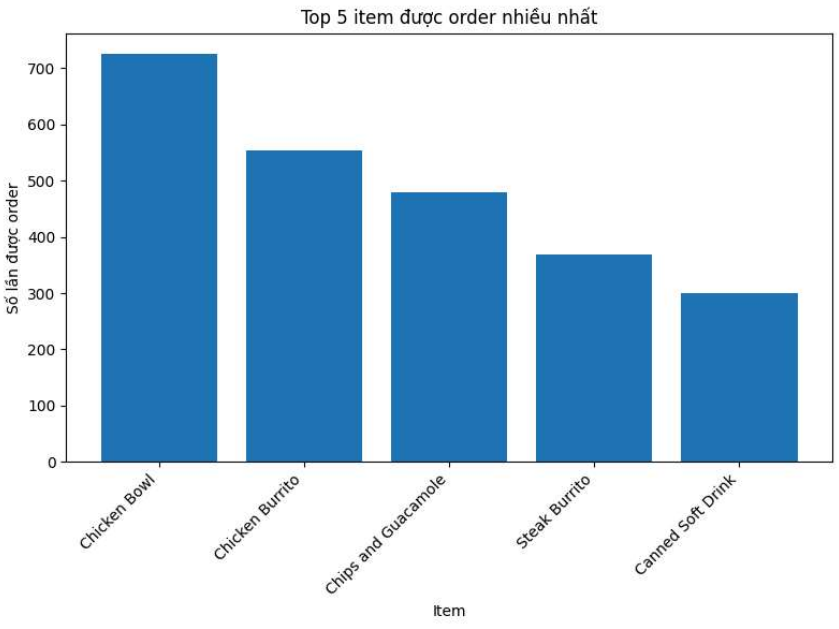
Step 4. See the first 10 entries

```
chipo.head(10)
```

	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
5	3	1	Chicken Bowl	[Fresh Tomato Salsa (Mild), [Rice, Cheese, Sou...	$10.98
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
9	5	1	Steak Burrito	[Fresh Tomato Salsa, [Rice, Black Beans, Pinto	$9.25

Step 5. Create a histogram of the top 5 items bought

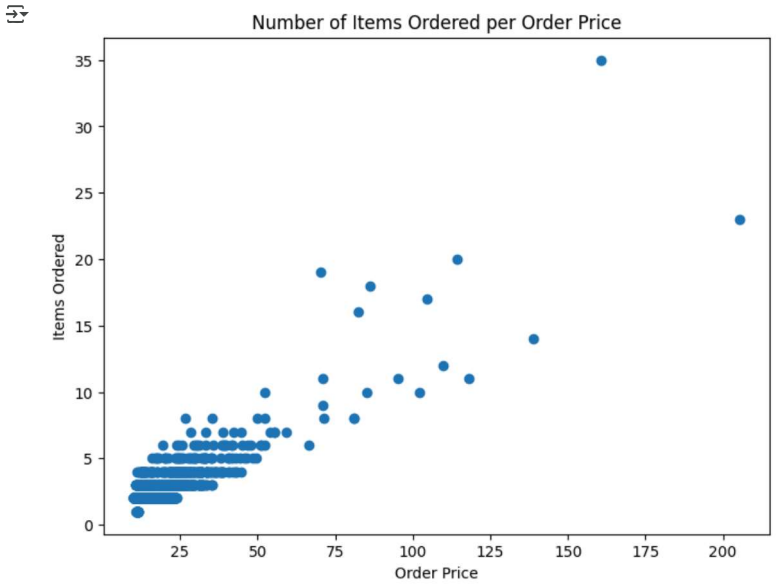
```
item_counts = chipo['item_name'].value_counts()
top5_items = item_counts.head(5)
plt.figure(figsize=(8, 6))
plt.bar(top5_items.index, top5_items.values)
plt.xlabel("Item")
plt.ylabel("Số lần được order")
plt.title("Top 5 item được order nhiều nhất")
plt.xticks(rotation=45, ha="right")
plt.tight_layout()
plt.show()
```



Step 6. Create a scatterplot with the number of items orderered per order price

Hint: Price should be in the X-axis and Items ordered in the Y-axis

```
chipo['item_price'] = chipo['item_price'].str.replace('$', '').astype(float)
order_items = chipo.groupby('order_id')['quantity'].sum()
order_prices = chipo.groupby('order_id')['item_price'].sum()
plt.figure(figsize=(8, 6))
plt.scatter(order_prices, order_items)
plt.xlabel("Order Price")
plt.ylabel("Items Ordered")
plt.title("Number of Items Ordered per Order Price")
plt.show()
```



Step 7. BONUS: Create a question and a graph to answer your own question.

Bắt đầu lập trình hoặc tạo mã bằng trí tuệ nhân tạo (AI).