Submittes by: Muhammad Arham Adeel

Ex1 - Filtering and Sorting Data

Step 1. Import the necessary libraries

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

Step 2. Import the dataset from this address.

In [2]:
```
url = "https://raw.githubusercontent.com/justmarkham/DAT8/master/data/chipotle.tsv"
```

Step 3. Assign it to a variable called chipo.

In [11]:
```
chipo = pd.read_csv(url,sep="\t")
chipo.head()
```

Out[11]:

	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

In [12]:
```
chipo.item_price.dtype
```

Out[12]:
```
dtype('O')
```

In [13]:
```
remove_dollar_sign = lambda x: float(x[1:-1])
chipo.item_price = chipo.item_price.apply(remove_dollar_sign)
```

In [14]:
```
chipo.item_price.dtype
```

Out[14]:
```
dtype('float64')
```

In [15]:
```
chipo.head()
```

Out[15]:

	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 [16]:
```
chipo[chipo["item_price"] > 10]
```

Out[16]:

	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

Step 5. What is the price of each item?

print a data frame with only two columns item_name and item_price

In [22]:
```
each_itm_prc = chipo[['item_name', 'item_price','quantity']]
each_itm_prc.drop_duplicates(subset= ['item_name','quantity'], inplace = True)
each_itm_prc[each_itm_prc['quantity'] == 1].nunique()
```

```
C:\Users\eAgLe\AppData\Local\Temp\ipykernel_3096\2326408891.py:2: SettingWithCopyWarn
ing:
A value is trying to be set on a copy of a slice from a DataFrame

See the caveats in the documentation: https://pandas.pydata.org/pandas-docs/stable/us
er_guide/indexing.html#returning-a-view-versus-a-copy
  each_itm_prc.drop_duplicates(subset= ['item_name','quantity'], inplace = True)
```

Out[22]:
```
item_name     50
item_price    19
quantity       1
dtype: int64
```

In [ ]:

Step 6. Sort by the name of the item

In [23]:
```
chipo.sort_values(by = "item_name")
```

Out[23]:

	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 [24]:
```
 chipo.item_price.max()
```

Out[24]:
```
44.25
```

In [27]:
```
chipo[chipo['item_price'] == chipo.item_price.max()]
```

Out[27]:

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

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

In [30]:
```
chipo[chipo["item_name"] == "Veggie Salad Bowl"].sum()[1]
```

Out[30]:
```
18
```

In [31]:
```
#Second method
chipo['item_name'].value_counts()
```

Out[31]:
```
Chicken Bowl                                726
Chicken Burrito                             553
Chips and Guacamole                         479
Steak Burrito                               368
Canned Soft Drink                           301
Steak Bowl                                  211
Chips                                       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
Izze                                         20
Chips and Tomatillo-Red Chili Salsa          20
Veggie Salad Bowl                            18
Chips and Roasted Chili-Corn Salsa           18
Barbacoa Crispy Tacos                        11
Barbacoa Salad Bowl                          10
Chicken Salad                                 9
Veggie Soft Tacos                             7
Carnitas Crispy Tacos                         7
Veggie Salad                                  6
Carnitas Salad Bowl                           6
Burrito                                       6
Steak Salad                                   4
Crispy Tacos                                  2
Salad                                         2
Bowl                                          2
Chips and Mild Fresh Tomato Salsa             1
Veggie Crispy Tacos                           1
Carnitas Salad                                1
Name: item_name, dtype: int64
```

In [33]:
```
chipo['item_name'].value_counts().loc['Veggie Salad Bowl']
```

Out[33]:
```
18
```

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

In [34]:
```
more_than_one = chipo[(chipo["quantity"] > 1) & (chipo["item_name"] == "Canned Soda")]
more_than_one
```

Out[34]:

	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 [35]:
```
len(more_than_one)
```

Out[35]:
```
20
```