In [12]:
```
# pip install pandas
```

In [14]:
```
import pandas as pd
import matplotlib.pyplot as plt
import seaborn as sns
```

In [18]:
```
path = 'chipotle.tsv'
chipo = pd.read_csv(path, sep = '\t')
```

In [19]:
```
chipo.head()
```

Out[19]:

	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 [20]:
```
chipo.shape
```

Out[20]:
```
(4622, 5)
```

In [21]:
```
chipo.columns
```

Out[21]:
```
Index(['order_id', 'quantity', 'item_name', 'choice_description',
       'item_price'],
      dtype='object')
```

In [22]:
```
chipo.index
```

Out[22]:
```
RangeIndex(start=0, stop=4622, step=1)
```

In [28]:
```
categorical_columns = chipo.select_dtypes(include='object').columns
numerical_columns = chipo.select_dtypes(exclude='object').columns
print(categorical_columns)
print(numerical_columns)
```

```
Index(['item_name', 'choice_description', 'item_price'], dtype='obje
ct')
Index(['order_id', 'quantity'], dtype='object')
```

In [29]:
```
#OR
chipo.dtypes
```

Out[29]:
```
order_id               int64
quantity               int64
item_name             object
choice_description    object
item_price            object
dtype: object
```

In [32]:
```
chipo['item_price'].head()
```

Out[32]:
```
0     $2.39
1     $3.39
2     $3.39
3     $2.39
4    $16.98
Name: item_price, dtype: object
```

In [33]:
```
chipo['item_price'] = chipo['item_price'].str.replace('$', '').asty
```

In [35]:
```
#checking
chipo['item_price'].dtype
```

Out[35]:
```
dtype('float64')
```

In [37]:
```
chipo.describe()
```

Out[37]:

	order_id	quantity	item_price
count	4622.000000	4622.000000	4622.000000
mean	927.254868	1.075725	7.464336
std	528.890796	0.410186	4.245557
min	1.000000	1.000000	1.090000
25%	477.250000	1.000000	3.390000
50%	926.000000	1.000000	8.750000
75%	1393.000000	1.000000	9.250000
max	1834.000000	15.000000	44.250000

In [42]:
```
#Sorting and Filtering Data
chipo[chipo.item_price > 10.00]['order_id'].count()
```

Out[42]:
```
1130
```

In [43]:
```
#OR
chipo[chipo.item_price > 10.00]['order_id'].size
```

Out[43]:
```
1130
```

! size measures total entries, while count() measures valid (non-null) entries

! size returns a single integer (total elements), while count() returns a series (per-column counts in a DataFrame) or a single integer (for a Series)

In [48]:
```
#A simple way to do that is to get the data frame with only two col
prices = chipo[['item_name', 'item_price']]
prices
```

Out[48]:

	item_name	item_price
0	Chips and Fresh Tomato Salsa	2.39
1	Izze	3.39
2	Nantucket Nectar	3.39
3	Chips and Tomatillo-Green Chili Salsa	2.39
4	Chicken Bowl	16.98
...	...	...
4617	Steak Burrito	11.75
4618	Steak Burrito	11.75
4619	Chicken Salad Bowl	11.25
4620	Chicken Salad Bowl	8.75
4621	Chicken Salad Bowl	8.75

4622 rows × 2 columns

In [49]:
```
#SORT
chipo.item_name.sort_values()
```

Out[49]:
```
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 [55]:
```
#OR
chipo.sort_values(by = 'item_name')
```

Out[55]:

	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

In [61]:
```
#most expensive item
chipo.sort_values(by = 'item_price', ascending=False)['quantity'].h
```

Out[61]:
```
3598    15
Name: quantity, dtype: int64
```

In [64]:
```
#without index
chipo.sort_values(by = "item_price", ascending=False)['quantity'].h
```

Out[64]:
```
15
```

In [67]:
```
#How many times the meal was ordered
print(chipo[chipo.item_name == "Veggie Salad Bowl"]['quantity'].sum
```

```
18
```

In [74]:
```
#How many times did someone order more than one Canned Soda
chipo[(chipo.item_name == "Canned Soda") & (chipo['quantity'] > 1)]
```

Out[74]: 20

In [78]:
```
#Unique products
unique_prod = chipo["item_name"].nunique()
print(unique_prod)
```

```
50
```

In [80]:
```
#Total revenue
total_revenue = (chipo['item_price'] * chipo['quantity']).sum()
print(total_revenue)
```

```
39237.02
```

In [82]:
```
#AVG
average_price = chipo['item_price'].mean()
print(round(average_price, 2))
```

```
7.46
```

In [87]:
```
#Total orders
total_orders = chipo['order_id'].nunique()
print(total_orders)
```

```
1834
```

In [88]:
```
#Total quantity
print(chipo['quantity'].sum())
```

```
4972
```

In [89]:
```
#Highest AVG price
chipo.groupby('item_name')['item_price'].mean().idxmax()
```

Out[89]: 'Bowl'

In [93]:
```
#How many items include “Chicken” in their name
chicken = chipo[chipo['item_name'].str.contains('Chicken')]['item_n
print(chicken)
```

```
1560
```

In [95]:
```
chipo.groupby('item_name')['quantity'].sum().idxmax()
```

Out[95]: 'Chicken Bowl'

In [98]:
```
# What was the most ordered item in the choice_description column

chipo.groupby('choice_description')['quantity'].sum().idxmax()
```

Out[98]: '[Diet Coke]'

In [106…
```
# What is the average revenue amount per order

(chipo['item_price'] * chipo['quantity']).mean()
```

Out[106… 8.48918649935093

In [108…
```
#Which product has the highest total quantity sold
chipo.groupby('item_name')['quantity'].sum().idxmax()
```

Out[108… 'Chicken Bowl'

In [109…
```
# calculate the average price of each product.
average_price = chipo.groupby('item_name')['item_price'].mean()
print(round(average_price, 2))
```

```
item_name
6 Pack Soft Drink                         6.61
Barbacoa Bowl                            10.19
Barbacoa Burrito                          9.83
Barbacoa Crispy Tacos                    10.93
Barbacoa Salad Bowl                      10.64
Barbacoa Soft Tacos                      10.02
Bottled Water                             1.87
Bowl                                     14.80
Burrito                                   7.40
Canned Soda                               1.32
Canned Soft Drink                         1.46
Carnitas Bowl                            10.83
Carnitas Burrito                         10.13
Carnitas Crispy Tacos                    11.14
Carnitas Salad                            8.99
Carnitas Salad Bowl                      11.06
Carnitas Soft Tacos                       9.40
Chicken Bowl                             10.11
Chicken Burrito                          10.08
Chicken Crispy Tacos                     10.05
Chicken Salad                             9.01
Chicken Salad Bowl                       11.17
Chicken Soft Tacos                        9.64
Chips                                     2.34
Chips and Fresh Tomato Salsa              3.29
Chips and Guacamole                       4.60
Chips and Mild Fresh Tomato Salsa         3.00
Chips and Roasted Chili Corn Salsa        3.08
Chips and Roasted Chili-Corn Salsa        2.39
Chips and Tomatillo Green Chili Salsa     3.09
Chips and Tomatillo Red Chili Salsa       3.07
Chips and Tomatillo-Green Chili Salsa     2.54
Chips and Tomatillo-Red Chili Salsa       2.99
Crispy Tacos                              7.40
Izze                                      3.39
Nantucket Nectar                          3.64
Salad                                     7.40
Side of Chips                             1.84
Steak Bowl                               10.71
Steak Burrito                            10.47
Steak Crispy Tacos                       10.21
Steak Salad                               8.91
Steak Salad Bowl                         11.85
Steak Soft Tacos                          9.75
Veggie Bowl                              10.21
Veggie Burrito                            9.84
Veggie Crispy Tacos                       8.49
Veggie Salad                              8.49
Veggie Salad Bowl                        10.14
Veggie Soft Tacos                        10.57
Name: item_price, dtype: float64
```

In [ ]:

In [112…

```
# Which choice appears most often?
```

```
chipo.groupby('choice_description').sum().sort_values(by = 'quantit
```

Out [112…

choice_description	order_id	quantity	item_name	item_price
[Diet Coke]	123455	159	Canned SodaCanned SodaCanned Soda6 Pack Soft D...	326.71