DATA SALES PROJECT NUPAT TECHNOLOGIES

IMPORT PANDAS AND NUMPY

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

READ CSV FILE

In [5]:
```
sales_data = pd.read_csv('salesdata.csv')
sales_data
```

Out[5]:

	order_id	quantity	item_name	choice_description	item_price
0	1	1	Izze	[Clementine]	$3.39
1	1	1	Nantucket Nectar	[Apple]	$3.39
2	2	2	Chicken Bowl	[Tomatillo-Red Chili Salsa (Hot), [Black Beans...	$16.98
3	3	1	Chicken Bowl	[Fresh Tomato Salsa (Mild), [Rice, Cheese, Sou...	$10.98
4	4	1	Steak Burrito	[Tomatillo Red Chili Salsa, [Fajita Vegetables...	$11.75
...	...	...	...	...	...
195	115	1	Steak Burrito	[Roasted Chili Corn Salsa (Medium), [Black Bea...	$8.99
196	116	1	Steak Soft Tacos	[Fresh Tomato Salsa, [Rice, Cheese, Lettuce]]	$9.25
197	117	1	Barbacoa Soft Tacos	[Tomatillo-Red Chili Salsa (Hot), [Pinto Beans...	$8.99
198	118	1	Chicken Burrito	[Tomatillo-Green Chili Salsa (Medium), [Pinto ...	$8.49
199	119	1	Chicken Burrito	[[Fresh Tomato Salsa (Mild), Roasted Chili Cor...	$8.49

200 rows × 5 columns

Use replace attribute to remove the $ sign and convert the column to float type data before filtering.

In [10]:
```
sales_data['item_price'] = sales_data['item_price'].replace('$', '')
sales_data['item_price'] = sales_data['item_price'].astype(float)
```

In [11]:
```
sales_data
```

Out[11]:

	order_id	quantity	item_name	choice_description	item_price
0	1	1	Izze	[Clementine]	3.39
1	1	1	Nantucket Nectar	[Apple]	3.39
2	2	2	Chicken Bowl	[Tomatillo-Red Chili Salsa (Hot), [Black Beans...	16.98
3	3	1	Chicken Bowl	[Fresh Tomato Salsa (Mild), [Rice, Cheese, Sou...	10.98
4	4	1	Steak Burrito	[Tomatillo Red Chili Salsa, [Fajita Vegetables...	11.75
...	...	...	...	...	...
195	115	1	Steak Burrito	[Roasted Chili Corn Salsa (Medium), [Black Bea...	8.99
196	116	1	Steak Soft Tacos	[Fresh Tomato Salsa, [Rice, Cheese, Lettuce]]	9.25
197	117	1	Barbacoa Soft Tacos	[Tomatillo-Red Chili Salsa (Hot), [Pinto Beans...	8.99
198	118	1	Chicken Burrito	[Tomatillo-Green Chili Salsa (Medium), [Pinto ...	8.49
199	119	1	Chicken Burrito	[[Fresh Tomato Salsa (Mild), Roasted Chili Cor...	8.49

200 rows × 5 columns

How many products cost more than $10.00?

In [16]:
```
np.count_nonzero(sales_data.item_price > 10.00)
```

Out[16]:
```
60
```

How many products cost less than $10.00?

In [18]:
```
np.count_nonzero(sales_data.item_price < 10.00)
```

Out[18]:
```
140
```

What is the price of each item?

Creating a Data frame with two columns 'item_name' and 'item_price' using Min of 'item_price' so we can get the price for only one quantity of a particular item.

In [20]:
```
product_cost = sales_data.groupby(by ='item_name')[['item_price']].min().reset_index()
product_cost
```

Out[20]:

	item_name	item_price
0	Barbacoa Bowl	8.99
1	Barbacoa Burrito	8.99
2	Barbacoa Crispy Tacos	9.25
3	Barbacoa Soft Tacos	8.99
4	Canned Soda	1.09
5	Canned Soft Drink	1.25
6	Carnitas Bowl	8.99
7	Carnitas Burrito	8.99
8	Carnitas Soft Tacos	9.25
9	Chicken Bowl	8.49
10	Chicken Burrito	8.49
11	Chicken Crispy Tacos	8.75
12	Chicken Salad	10.98
13	Chicken Salad Bowl	8.75
14	Chicken Soft Tacos	8.75
15	Izze	3.39
16	Nantucket Nectar	3.39
17	Steak Bowl	8.99
18	Steak Burrito	8.99
19	Steak Crispy Tacos	9.25
20	Steak Soft Tacos	8.99
21	Veggie Bowl	8.49
22	Veggie Burrito	8.49
23	Veggie Salad Bowl	11.25

Sort the products by the name of the item.

In [21]:
```
sales_data.sort_values(by = "item_name",ascending=True).reset_index(drop=True)
```

Out[21]:

	order_id	quantity	item_name	choice_description	item_price
0	27	1	Barbacoa Bowl	[Roasted Chili Corn Salsa, [Fajita Vegetables,...	11.75
1	61	1	Barbacoa Bowl	[Tomatillo Red Chili Salsa, [Fajita Vegetables...	11.75
2	56	1	Barbacoa Bowl	[Tomatillo Red Chili Salsa, [Rice, Pinto Beans...	9.25
3	19	1	Barbacoa Bowl	[Roasted Chili Corn Salsa, [Fajita Vegetables,...	11.75
4	110	1	Barbacoa Bowl	[Tomatillo Red Chili Salsa, [Rice, Cheese, Let...	9.25
...	...	...	...	...	...
195	110	1	Veggie Bowl	[Roasted Chili Corn Salsa, [Rice, Pinto Beans,...	8.75
196	99	1	Veggie Burrito	[Tomatillo Red Chili Salsa, [Black Beans, Chee...	11.25
197	26	1	Veggie Burrito	[Tomatillo Red Chili Salsa, [Fajita Vegetables...	11.25
198	46	1	Veggie Burrito	[Fresh Tomato Salsa (Mild), [Black Beans, Rice...	8.49
199	83	1	Veggie Salad Bowl	[Fresh Tomato Salsa, [Fajita Vegetables, Rice,...	11.25

200 rows × 5 columns

What was the quantity of the most expensive item ordered?

In [23]:
```
quantity_exp_order = np.where(sales_data.item_price == sales_data.item_price.max(),sales_data.quantity,0)
for i in quantity_exp_order:
    if i!=0:
        print('The quantity of the most expensive item ordered is : ',i)
```

```
The quantity of the most expensive item ordered is :  2
The quantity of the most expensive item ordered is :  2
```

How many times was a Chicken Burrito ordered?

In [25]:
```
order_times = np.count_nonzero(sales_data.item_name == 'Chicken Burrito')
print('Chicken Burrito was ordered', order_times, 'times.' )
```

```
Chicken Burrito was ordered 39 times.
```

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

In [26]:
```
order_canned_soda = np.count_nonzero((sales_data.item_name == "Canned Soda") & (sales_data.quantity > 1))
print('Customers orderd more than one Canned Soda', order_canned_soda, 'times')
```

```
Customers orderd more than one Canned Soda 4 times
```

COMPLETED BY EMMANUEL ADEITAN, DATA SCIENCE AND MACHINE LEARNING INTERN, NUPAT TECHNOLOGIES.

In [ ]:
```
```