



Week 3 :
Assignment 2 – Task 2

Assignment 2 : Data Cleaning (chipotle.tsv):

General Task Solve: 2

1. Quantity of each item?

Top 5 items available in store.

```
In [109]: top_items = data['standardized_item_name'].value_counts().head(5)           # To print top 5 items
          print(top_items)
```

```
Chicken Bowl      717
Chicken Burrito   546
Chips and Guacamole 474
Steak Burrito     365
Canned Soft Drink 290
Name: standardized_item_name, dtype: int64
```

2. Which is the most ordered item in choice_description?

```
In [116]: filtered_data = data[data['choice_description'] != 'None']                # As we replaced None at NULL places, we w
          # ill try filter None here
          choice_sum = filtered_data.groupby('choice_description').sum()           # Find sum of unique choice_description
          choice_sum = choice_sum.sort_values(['quantity'], ascending=False)      # Print in Ascending = False, meaning Desc
          # ending!
          choice_sum.head(5)
```

Out[116]:

	order_id	quantity	item_price
choice_description			
Diet Coke	123222	158	325.46
Coke	117334	135	278.79
Sprite	80426	89	133.93
Fresh Tomato Salsa, Rice, Black Beans, Cheese, Sour Cream, Lettuce	41581	48	464.26
Fresh Tomato Salsa, Rice, Black Beans, Cheese, Sour Cream	36041	42	407.25

Assignment 2 : Data Cleaning (chipotle.tsv):

General Task Solve: 2

3. What is total order count?

```
In [117]: data['quantity'].sum()           # Summing all the quantities
```

```
Out[117]: 4913
```

4. What is revenue?

```
In [ ]: data['item_price'].sum()          # Summing all the item_price
```

5. What are number of orders?

```
In [119]: len(data['order_id'].unique())  # Len is to get the length of the list consisting of unique order_id
```

```
Out[119]: 1834
```

6. What is the average revenue amount per order?

```
In [123]: data['revenue'] = data['quantity'] * data['item_price']      # For calculating revenue of each item, rev = qty * price
order_grouped = data.groupby(by=['order_id']).sum()                    # Group similar data type of order_id
order_grouped.mean()['revenue'].round(decimals=4)                      # After the data of every rev is stored in revenue list, find sum for every grouped data
```

```
Out[123]: 21.237
```

7. How many different items are sold?

```
In [125]: data['standardized_item_name'].value_counts().count()      # Counts uniquely available items
```

```
Out[125]: 50
```

Assignment 2 : Data Cleaning (chipotle.tsv):

General Task Solve: 2

8. What is revenue from each product?

```
In [129]: data['revenue'] = data['quantity'] * data['item_price']      # For calculating revenue of each item, rev = qty
          * price
          data.groupby('standardized_item_name')['revenue'].sum()    # Grouping items then revenue's sum to find revenue of each item
          # revenue of each item
```

```
Out[129]: standardized_item_name
6 Pack Soft Drink          369.93
Barbacoa Bowl             663.00
Barbacoa Burrito          885.60
Barbacoa Crispy Tacos     141.24
Barbacoa Salad Bowl       97.02
Barbacoa Soft Tacos       250.50
Bottled Water             628.90
Bowl                       74.00
Burrito                   44.40
Canned Soda              189.66
Canned Soft Drink        590.00
Carnitas Bowl            815.23
Carnitas Burrito         618.74
Carnitas Crispy Tacos    98.52
Carnitas Salad           8.99
Carnitas Salad Bowl      66.36
Carnitas Soft Tacos      376.00
Chicken Bowl            7979.84
Chicken Burrito         6407.70
Chicken Crispy Tacos     531.44
Chicken Salad           81.09
Chicken Salad Bowl     1476.59
Chicken Soft Tacos     1164.24
Chips                   574.05
Chips and Fresh Tomato Salsa 984.16
Chips and Guacamole     2443.97
Chips and Mild Fresh Tomato Salsa 3.00
Chips and Roasted Chili Corn Salsa 73.75
Chips and Roasted Chili-Corn Salsa 43.02
Chips and Tomatillo Green Chili Salsa 144.55
Chips and Tomatillo Red Chili Salsa 153.40
Chips and Tomatillo-Green Chili Salsa 88.43
Chips and Tomatillo-Red Chili Salsa 90.82
Crispy Tacos            14.80
Izze                    64.41
Nantucket Nectar        111.87
Salad                   14.80
Side of Chips           290.68
Steak Bowl              2477.89
Steak Bowl              2477.89
Steak Burrito           4225.43
Steak Crispy Tacos      378.12
Steak Salad             35.64
Steak Salad Bowl        386.05
Steak Soft Tacos        555.64
Veggie Bowl            910.91
Veggie Burrito          988.83
Veggie Crispy Tacos      8.49
Veggie Salad            50.94
Veggie Salad Bowl       182.52
Veggie Soft Tacos        93.52
Name: revenue, dtype: float64
```