Challenge #1: Chipotle Sales

by Eboni Brown

Which was the most-ordered item and how many items were ordered?

```
most_ordered = chipo[['item_name','quantity']]
most_ordered = most_ordered.groupby(by='item_name')['quantity'].sum().sort_values(ascending=False).reset_index()
most_ordered.head()
```

	item_name	quantity
0	Chicken Bowl	761
1	Chicken Burrito	591
2	Chips and Guacamole	506
3	Steak Burrito	386
4	Canned Soft Drink	351

761 Chicken Bowls ordered!

What was the most ordered item in the choice_description column?

```
most_ordered_choice = chipo[['choice_description','quantity']]
most_ordered_choice = most_ordered_choice.groupby(by='choice_description')['quantity'].sum().sort_values(ascending=False).reset_index()
most_ordered_choice.head()
```

	choice_description	quantity
0	[Diet Coke]	159
1	[Coke]	143
2	[Sprite]	89
3	[Fresh Tomato Salsa, [Rice, Black Beans, Chees...	49
4	[Fresh Tomato Salsa, [Rice, Black Beans, Chees...	42

Diet Coke was the most ordered choice!

How many items were ordered in total?

```
total = chipo['quantity'].sum()
print(f'There were {total} items ordered in total.')
```

```
There were 4972 items ordered in total.
```

Summed the quantity ordered per item to find total number of items ordered

Turn the item price into a float

Before

```
[8] chipo['item_price'].head()
```

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

```
[9] #remove '$'
    chipo['item_price'] = chipo['item_price'].str.replace('$', '', regex=False)
    #change data type
    chipo['item_price'] = chipo['item_price'].astype('float')
```

After

```
chipo['item_price'].head()
```

```
0     2.39
1     3.39
2     3.39
3     2.39
4    16.98
Name: item_price, dtype: float64
```

used .replace() to remove the '$' so that the item_price column can be converted to a float and aggregations can be performed on it.

How much was the revenue for the period in the dataset?

```
revenue = chipo[['item_name','item_price','quantity']].groupby(by='item_name').agg({'quantity': 'sum', 'item_price': 'first'}).reset_index()
```

```
def total_revenue(df):
  total_revenue = 0
  for index, col in df.iterrows():
    quantity = col['quantity']
    item_price = col['item_price']
    revenue = quantity * item_price
    total_revenue += revenue
  return total_revenue

print(f'Total revenue is ${total_revenue(revenue)}')
```

```
Total revenue is $40361.88
```

I used a function to calculate total revenue which was $40361.88 for the duration of this data set

What is the average revenue amount per order?

```
avg_revenue = total_revenue(revenue) / num_orders
print(f'The average revenue per order is ${round(avg_revenue,2)}')
```

```
The average revenue per order is $8.73
```

I used previous calculations to find average revenue per order, about $8.73 per order.

How many different items are sold?

```
diff_items_count = most_ordered.groupby(by='item_name')['quantity'].sum().sort_values(ascending=False).reset_index()
num_diff_items = diff_items_count['item_name'].count()
print(f'There are {num_diff_items} different items sold at Chipotle.')
```

```
There are 50 different items sold at Chipotle.
```

This dataset shows 50 different items sold by Chipotle.

Thank you!