Scenario

You are a financial data analyst at Chipotle and your manager has tasked you with analyzing the most recent sales numbers. She has provided the following set of questions she would like answered.

Thanks to Data in Motion - DataSet here -> https://d-i-motion.com/lessons/challenge-1-chipotle-sales/

In [ ]:
```
import pandas as pd

# Read the TSV file using pandas
df = pd.read_csv('chipotle.tsv', delimiter='\t')

# Display a preview of the DataFrame
print("Preview of the data:")
df.head()
```

```
Preview of the data:
```

Out[ ]:

	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 [ ]:
```python
# Group the data by the item_name column and sum the quantity for each item
item_counts = df.groupby('item_name')['quantity'].sum()
```

In [ ]:
```python
# Find the item with the highest quantity
most_ordered_item = item_counts.idxmax()

# Print the result
print(f"The most-ordered item is: {most_ordered_item}")
```

```
The most-ordered item is: Chicken Bowl
```

In [ ]:
```python
# Filter the DataFrame for the most-ordered item
most_ordered_item_df = df[df['item_name'] == most_ordered_item]

# Calculate the total quantity for the most-ordered item
total_ordered_quantity = most_ordered_item_df['quantity'].sum()

# Print the result
print(f"The most-ordered item is: {most_ordered_item}")
print(f"Total quantity ordered: {total_ordered_quantity}")
```

```
The most-ordered item is: Chicken Bowl
Total quantity ordered: 761
```

In [ ]:
```python
# Find the most ordered item in the choice_description column
most_ordered_item = df['choice_description'].value_counts().idxmax()

# Print the result
print(f"The most ordered item in the choice_description column is: {most_ordered_item}")
```

```
The most ordered item in the choice_description column is: [Diet Coke]
```

In [ ]:
```python
total_items_ordered = df['quantity'].sum()

# Print the result
print(f"The total number of items ordered is: {total_items_ordered}")
```

```
The total number of items ordered is: 4972
```

In [ ]:
```python
#Turn the item price into a float
df['item_price'] = df['item_price'].str.replace('$', '').astype(float)

# Print the updated DataFrame
df.head()
```

Out[ ]:

	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 [ ]:
```python
revenue = (df['item_price'] * df['quantity']).sum()
```

```
# Print the total revenue
print(f"Total revenue: {revenue}")
```

```
Total revenue: 39237.02
```

In [ ]:
```
num_orders = df['order_id'].nunique()

# Print the total number of orders
print(f"Total number of orders: {num_orders}")
```

```
Total number of orders: 1834
```

In [ ]:
```
average_revenue_per_order = df['item_price'].sum() / num_orders

# Round the average revenue per order to two decimal places
average_revenue_per_order = round(average_revenue_per_order, 2)

# Print the average revenue per order
print(f"Average revenue per order: ${average_revenue_per_order}")
```

```
Average revenue per order: $18.81
```

In [ ]:
```
num_different_items = len(df['item_name'].unique())

# Print the number of different items sold
print(f"Number of different items sold: {num_different_items}")
```

```
Number of different items sold: 50
```