CHIPOTLE SALES

Shahzad Azeez

AGENDA

1. Introduction
2. Challenging Questions
3. Summary

INTRODUCTION

You are a financial data analyst at Chipotle and your manager has tasked you with analyzing the most recent sales numbers.

CHALLENGING QUESTIONS

1. Which was the most-ordered item?
2. For the most-ordered item, how many items were ordered?
3. What was the most ordered item in the choice_description column?
4. How many items were ordered in total?
5. Turn the item price into a float
6. How much was the revenue for the period in the dataset?
7. How many orders were made in the period?
8. What is the average revenue amount per order?
9. How many different items are sold?

DATA SET

```
data.head()
```

	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

1.Which was the most-ordered item?

1. Which was the most-ordered item?

```
# Group the data by item and sum the quantities
item_counts = data.groupby('item_name')['quantity'].sum().reset_index()

# Find the most ordered item
most_ordered_item = item_counts.loc[item_counts['quantity'].idxmax()]

# Print the result
print("The most ordered item is:", most_ordered_item['item_name'])
```

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

2.For the most-ordered item, how many items were ordered?

2. For the most-ordered item, how many items were ordered?

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

# Find the most ordered item
most_ordered_item = item_counts.loc[item_counts['quantity'].idxmax()]

# Extract the item and its quantity
most_ordered_item_name = most_ordered_item['item_name']
most_ordered_item_quantity = most_ordered_item['quantity']

# Print the results
print(f"The quantity ordered for most ordered item is :  {most_ordered_item_quantity} ")
```

```
The quantity ordered for most ordered item is :  761
```

3. What was the most ordered item in the choice_description column?

3. What was the most ordered item in the choice_description column?

```
# Fill NaN values in 'choice_description' with a placeholder (e.g., 'Unknown')
data['choice_description'].fillna('Unknown', inplace=True)

# Group the data by 'choice_description' and sum the 'quantity' for each choice description
choice_counts = data.groupby('choice_description')['quantity'].sum().reset_index()

# Sort the results in descending order based on quantity
sorted_choice_counts = choice_counts.sort_values(by='quantity', ascending=False)

#Find the first most ordered choice description
first_most_ordered_choice = sorted_choice_counts.iloc[0]

#Find the first most ordered choice description quantity
first_most_ordered_choice_description = first_most_ordered_choice['choice_description']
first_most_ordered_choice_quantity = first_most_ordered_choice['quantity']


# Print the result
print(f"The_most_ordered_item_is '{first_most_ordered_choice_description}' with
      {first_most_ordered_choice_quantity} items ordered.")
```

```
The_most_ordered_item_is 'Unknown' with 1382 items ordered.
```

```
# Fill NaN values in 'choice_description' with a placeholder (e.g., 'Unknown')
data['choice_description'].fillna('Unknown', inplace=True)

# Group the data by 'choice_description' and sum the 'quantity' for each choice description
choice_counts = data.groupby('choice_description')['quantity'].sum().reset_index()

# Sort the results in descending order based on quantity
sorted_choice_counts = choice_counts.sort_values(by='quantity', ascending=False)

# Find the second most ordered choice description (skip the first, which is 'Unknown' or NaN)
second_most_ordered_choice = sorted_choice_counts.iloc[1]

# Extract the choice description and its quantity
second_most_ordered_choice_description = second_most_ordered_choice['choice_description']
second_most_ordered_choice_quantity = second_most_ordered_choice['quantity']
# Print the result
print(f"The second most ordered choice description is '{second_most_ordered_choice_description}' with
      {second_most_ordered_choice_quantity} items ordered.")
```

```
The second most ordered choice description is '[Diet Coke]' with 159 items ordered.
```

4. How many items were ordered in total?

4. How many items were ordered in total?

```
# Calculate the total number of items ordered
total_items_ordered = data['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
```

5. Turn the item price into a float

5. Turn the item price into a float

```
# Fill NaN values in 'choice_description' with a placeholder (e.g., 'Unknown')
data['choice_description'].fillna('Unknown', inplace=True)

#replace '$'
data['item_price'] = data['item_price'].replace('$', '').inplace=True

# Convert 'item price' column to float
data['item_price'] = data['item_price'].astype(float)

# Print the DataFrame to verify the conversion
print(data)
```

```
      order_id  quantity                              item_name  \
0            1         1           Chips and Fresh Tomato Salsa
1            1         1                                   Izze
2            1         1                       Nantucket Nectar
3            1         1  Chips and Tomatillo-Green Chili Salsa
4            2         2                           Chicken Bowl
...        ...       ...                                    ...
4617      1833         1                          Steak Burrito
4618      1833         1                          Steak Burrito
4619      1834         1                     Chicken Salad Bowl
4620      1834         1                     Chicken Salad Bowl
4621      1834         1                     Chicken Salad Bowl

                                choice_description  item_price
0                                          Unknown         1.0
1                                      [Clementine]         1.0
2                                          [Apple]         1.0
3                                          Unknown         1.0
4     [Tomatillo-Red Chili Salsa (Hot), [Black Beans...         1.0
```

```
data.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 4622 entries, 0 to 4621
Data columns (total 5 columns):
 #   Column              Non-Null Count  Dtype
---  ------              --------------  -----
 0   order_id            4622 non-null   int64
 1   quantity            4622 non-null   int64
 2   item_name           4622 non-null   object
 3   choice_description  4622 non-null   object
 4   item_price          4622 non-null   float64
dtypes: float64(1), int64(2), object(2)
memory usage: 180.7+ KB
```

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

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

```
# Calculate the revenue for each row and add it as a new column 'revenue'
data['revenue'] = data['quantity'] * data['item_price']

# Calculate the total revenue for the period
total_revenue = data['revenue'].sum()

# Print the total revenue
print(f"The total revenue for the period is: ${total_revenue:.2f}")
```

```
The total revenue for the period is: $4972.00
```

7. How many orders were made in the period?

7. How many orders were made in the period?

```
# Count the number of unique orders made in the period
total_orders = data['order_id'].nunique()

# Print the total number of orders
print(f"The total number of orders made in the period is: {total_orders}")
```

```
The total number of orders made in the period is: 1834
```

8. What is the average revenue amount per order?

8. What is the average revenue amount per order?

```
# Calculate the average revenue per order
average_revenue_per_order = total_revenue / total_orders

# Print the average revenue per order
print(f"The average revenue amount per order is: ${average_revenue_per_order:.2f}")
```

```
The average revenue amount per order is: $2.71
```

9. How many different items are sold?

9. How many different items are sold?

```
# Count the number of different items sold
different_items_sold = data['item_name'].nunique()

# Print the number of different items sold
print(f"The number of different items sold is: {different_items_sold}")
```

```
The number of different items sold is: 50
```

SUMMARY

- Most ordered item is chicken bowl, ordered quantity is 761.
- Most people chose the choice description “diet coke” where 159 items ordered in the period
- There is a revenue of $4972 for the period
- The Average revenue amount per order is $2.71

THANK YOU

Shahzad Azeez
shahzadazeez3242@gmail.com