Analysis with SQL

In [1]:
```
import sqlite3
import pandas as pd
```

In [3]:
```
conn = sqlite3.connect('pizzeria.db')
cursor = conn.cursor()
```

In [7]:
```
sales_df = pd.read_csv('sales_transactions.csv')
menu_df = pd.read_csv('menu_data.csv')
inventory_df = pd.read_csv('inventory_ingredients.csv')
reviews_df = pd.read_csv('customer_reviews.csv')
```

In [9]:
```
sales_df.to_sql('sales', conn, if_exists='replace', index=False)
menu_df.to_sql('menu', conn, if_exists='replace', index=False)
inventory_df.to_sql('inventory', conn, if_exists='replace', index=False)
reviews_df.to_sql('reviews', conn, if_exists='replace', index=False)
```

Out[9]: 5

In [11]:
```
conn.commit()
conn.close()
```

Total Sales by Item (Revenue per Menu Item)

In [27]:
```
conn = sqlite3.connect('pizzeria.db')
query = '''
SELECT m.Item_Name, SUM(s.Quantity * s.Price) AS Total_Revenue
FROM sales s
JOIN menu m ON s.Item_ID = m.Item_ID
GROUP BY m.Item_Name
ORDER BY Total_Revenue DESC;
'''
sales_by_item = pd.read_sql(query, conn)
sales_by_item.to_csv('sales_by_item.csv', index=False)  # index=False to avoid w
conn.close()

# Show the result
print(sales_by_item)
```

```
      Item_Name  Total_Revenue
0     Bolognese         777660
1     Carbonara         738188
2   Amatriciana         712287
3        Vegana         707508
4       Diavola         658980
5    Margherita         617632
6      Tiramisu         352818
7      Focaccia         287940
8      Espresso         230556
9     Coca-Cola         194100
10        Water          96635
```

Total Monthly Sales

In [29]:
```
conn = sqlite3.connect('pizzeria.db')
query = '''
SELECT strftime('%Y-%m', Date) AS Month, SUM(Quantity * Price) AS Total_Revenue
FROM sales
GROUP BY Month
ORDER BY Month;
'''
monthly_sales = pd.read_sql(query, conn)
monthly_sales.to_csv('monthly_sales.csv', index=False)  # index=False to avoid w

conn.close()

# Show the result
print(monthly_sales)
```

```
     Month  Total_Revenue
0  2023-01         473922
1  2023-02         469155
2  2023-03         438959
3  2023-04         436755
4  2023-05         452168
5  2023-06         399540
6  2023-07         408647
7  2023-08         503898
8  2023-09         479537
9  2023-10         447678
10 2023-11         444101
11 2023-12         419944
```

Most Used Ingredients

In [31]:
```
conn = sqlite3.connect('pizzeria.db')
query = '''
SELECT i.Ingredient, SUM(s.Quantity) AS Total_Usage
FROM sales s
JOIN inventory i ON s.Item_ID = i.Item_ID
GROUP BY i.Ingredient
ORDER BY Total_Usage DESC;
'''
inventory_usage = pd.read_sql(query, conn)
inventory_usage.to_csv('inventory_usage.csv', index=False)  # index=False to avo

conn.close()

# Show the result
print(inventory_usage)
```

```
      Ingredient  Total_Usage
0   Tomato Sauce        41589
1          Pasta        25127
2         Coffee        16642
3       Pancetta        16577
4         Cheese        16371
5    Ground Beef         8550
6   Vegan Cheese         8419
7      Mushrooms         8419
8   Bell Peppers         8419
9     Mascarpone         8397
10         Eggs         8328
11    Coca-Cola         8268
12        Basil         8267
13        Water         8239
14     Rosemary         8164
15    Olive Oil         8164
16        Flour         8164
17       Salami         8104
```

Review Rating Distribution

In [25]:
```
conn = sqlite3.connect('pizzeria.db')
query = '''
SELECT rating, COUNT(rating) AS Rating_Count
FROM reviews
GROUP BY rating
ORDER BY rating DESC;
'''
rating_distribution = pd.read_sql(query, conn)
conn.close()

# Show the result
print(rating_distribution)
```

```
  rating  Rating_Count
0      5             5
```

In [ ]:
```

```