SQL

SQL-Powered Insights for The Restaurant Industry

An Analysis of Top 250, Future 50, and Independence 100 Datasets

Malik Mohammad Junaid Majeed

Abstract

This report presents a structured SQL-based analysis of three major datasets: Top 250, Future 50, and Independence 100. The project focuses on data cleaning, query optimization, and extracting meaningful business insights. Key findings include identification of high-performing restaurants, growth trends, and regional distributions. The study demonstrates SQL as a powerful tool for in dustry benchmarking and strategic decision-making.

Introduction

The restaurant industry is highly dynamic, and data-driven analysis is essential for competitive advantage. Structured Query Language (SQL) enables analysts to clean, organize, and interpret datasets effectively. This report explores SQL methodologies applied to three datasets: Independence100 (independent restaurants), Future50 (emerging chains), and Top250 (established leaders). The analysis highlights sales performance, growth metrics, and operational efficiency.

Methodology

The analysis was conducted in MySQL. Initial steps included creating temporary cleaned tables for each dataset, ensuring removal of duplicates, NULL values, and invalid entries. Safe updates were di sabled to allow for data modifications. Columns such as YOY_Sales and YOY_Units were standardize d into numeric formats. Once cleaned, advanced queries involving aggregation, window functions, CTEs, and CASE statements were applied to generate insights.

Advanced SQL Concepts Applied

The project also applied advanced SQL techniques:

- Stored Procedures: For automating transaction insertion.
- Triggers: For automatic updates of balances after inserts.
- Window Functions: For ranking, rolling totals, and partitioned analysis.
- CTEs: For segmenting data such as 'High Check' vs 'Low Check'.
- CASE Statements: For categorizing restaurants by growth.
- UNION: For combining top sales performers across datasets.

SQL Queries and Insights:

Top Sales Performers in Independents

SQL:

```
SELECT Restaurant, City, State, SUM (Sales) AS TOTAL_SALES_RANKING
FROM Cleaned_independence100
GROUP BY Restaurant, City, State
ORDER BY TOTAL_SALES_RANKING DESC
LIMIT 10;
```

Result Grid | Filter Rows: | Export: | Wrap Cell Content: | Fetch rows:

Restaurant	City	State	TOTAL_SALES_RANKING
Carmine's (Times Square)	New York	N.Y.	39080335
The Boathouse Orlando	Orlando	Fla.	35218364
Old Ebbitt Grill	Washington	D.C.	29104017
LAVO Italian Restaurant & Nightclub	New York	N.Y.	26916180
Bryant Park Grill & Cafe	New York	N.Y.	26900000
Gibsons Bar & Steakhouse	Chicago	Ill.	25409952
Top of the World at the STRAT	Las Vegas	Nev.	25233543
Maple & Ash	Chicago	Ill.	24837595
Balthazar	New York	N.Y.	24547800
Smith & Wollensky	New York	N.Y.	24501000

Rolling Total of Meals Served (Window Function)

SQL:

```
SELECT *, SUM(`Meals Served`) OVER(ORDER BY `Rank` ) AS ROLLING_TOTAL
FROM independence100
ORDER BY `Rank`;
```

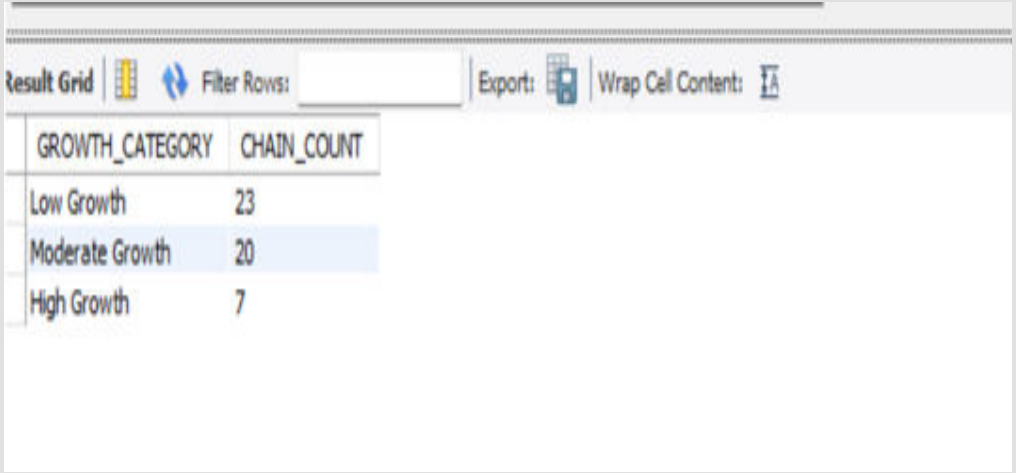
Result Grid | Filter Rows: | Export: | Wrap Cell Content:

Rank	Restaurant	Sales	Average Check	City	State	Meals Served	ROLLING_TOTAL
1	Carmine's (Times Square)	39080335	40	New York	N.Y.	469803	469803
2	The Boathouse Orlando	35218364	43	Orlando	Fla.	820819	1290622
3	Old Ebbitt Grill	29104017	33	Washington	D.C.	892830	2183452
4	LAVO Italian Restaurant & Nightclub	26916180	90	New York	N.Y.	198500	2381952
5	Bryant Park Grill & Cafe	26900000	62	New York	N.Y.	403000	2784952
6	Gibsons Bar & Steakhouse	25409952	80	Chicago	Ill.	348567	3133519
7	Top of the World at the STRAT	25233543	103	Las Vegas	Nev.	246054	3379573
8	Maple & Ash	24837595	99	Chicago	Ill.	210832	3590405
9	Balthazar	24547800	87	New York	N.Y.	519000	4109405
10	Smith & Wollensky	24501000	107	New York	N.Y.	257364	4366769
11	Angus Barn	24268160	75	Raleigh	N.C.	315000	4681769
12	Prime 112	23800000	135	Miami Beach	Fla.	206000	4887769
13	Joe's Seafood, Prime Steak & Sto...	23660000	86	Washington	D.C.	277850	5165619

Growth Segmentation in Future50

SQL:

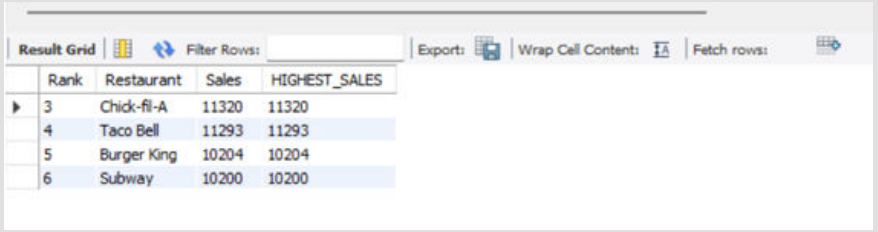
```
SELECT
CASE
   WHEN YOY_Sales > 50 THEN 'High Growth'
   WHEN YOY_Sales BETWEEN 20 AND 30 THEN 'Moderate Growth'
   ELSE 'Low Growth'
END AS GROWTH_CATEGORY,
COUNT(*) AS CHAIN_COUNT
FROM future50
GROUP BY GROWTH_CATEGORY
ORDER BY COUNT(*) DESC;
```



Restaurants Ranked 3rd–6th in Sales (Top250)

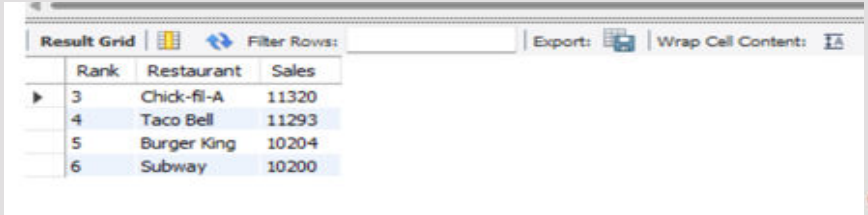
SQL:

```
SELECT `Rank`, Restaurant, Sales, MAX(Sales) AS HIGHEST_SALES
FROM Cleaned_top250
GROUP BY `Rank`, Restaurant, Sales
ORDER BY HIGHEST_SALES DESC
LIMIT 4 OFFSET 2;
```

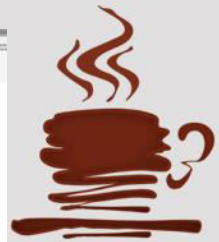


Rank	Restaurant	Sales	HIGHEST_SALES
3	Chick-fil-A	11320	11320
4	Taco Bell	11293	11293
5	Burger King	10204	10204
6	Subway	10200	10200

```
-- Alternative (Window Function)
SELECT `Rank`, Restaurant, Sales
FROM (
   SELECT *, DENSE_RANK() OVER(ORDER BY Sales DESC) AS rnk
   FROM Cleaned_top250
) AS RANKING
WHERE rnk > 2 AND rnk < 7;
```



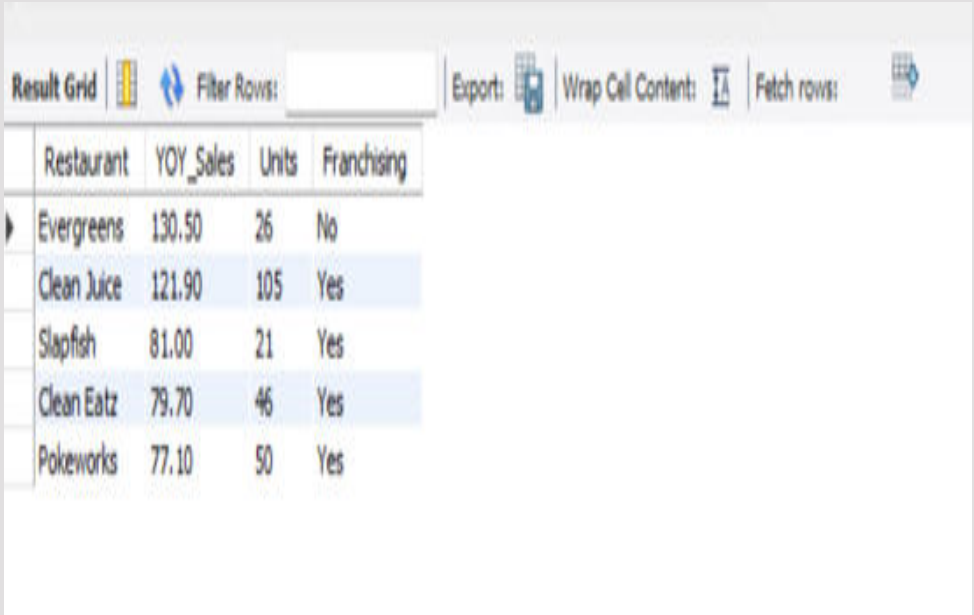
Rank	Restaurant	Sales
3	Chick-fil-A	11320
4	Taco Bell	11293
5	Burger King	10204
6	Subway	10200



Top 5 Emerging Chains by YOY Sales Growth (Future50)

SQL:

```
SELECT Restaurant, YOY_Sales, Units, Franchising
FROM Cleaned_future50
ORDER BY YOY_Sales DESC
LIMIT 5;
```



Restaurant	YOY_Sales	Units	Franchising
Evergreens	130.50	26	No
Clean Juice	121.90	105	Yes
Slapfish	81.00	21	Yes
Clean Eatz	79.70	46	Yes
Pokeworks	77.10	50	Yes

Geographic Distribution of Independent Restaurants

SQL:

```
SELECT DISTINCT State, COUNT(Restaurant) AS RESTAURANT_PER_STATE
FROM Cleaned_independence100
GROUP BY State
ORDER BY COUNT(Restaurant) DESC;
```

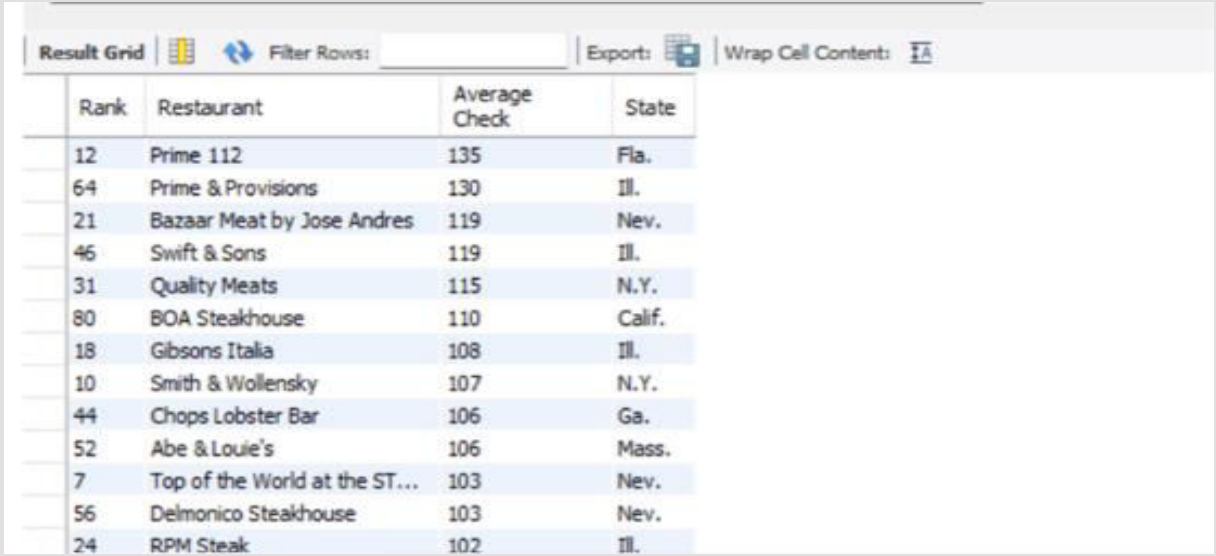
Result Grid | Filter Rows: | Export: | Wrap Cell Content:

State	RESTAURANT_PER_STATE
N.Y.	21
Ill.	18
Calif.	12
Nev.	11
Fla.	9
D.C.	9
Texas	3
Ind.	2
Ga.	2
Mich.	2
N.J.	2
Tenn.	2
N.C.	1

Outliers in Check Value

SQL:

```
SELECT `Rank`, Restaurant, `Average Check`, State
FROM Cleaned_independence100
WHERE `Average Check` > 100
ORDER BY `Average Check` DESC;
```



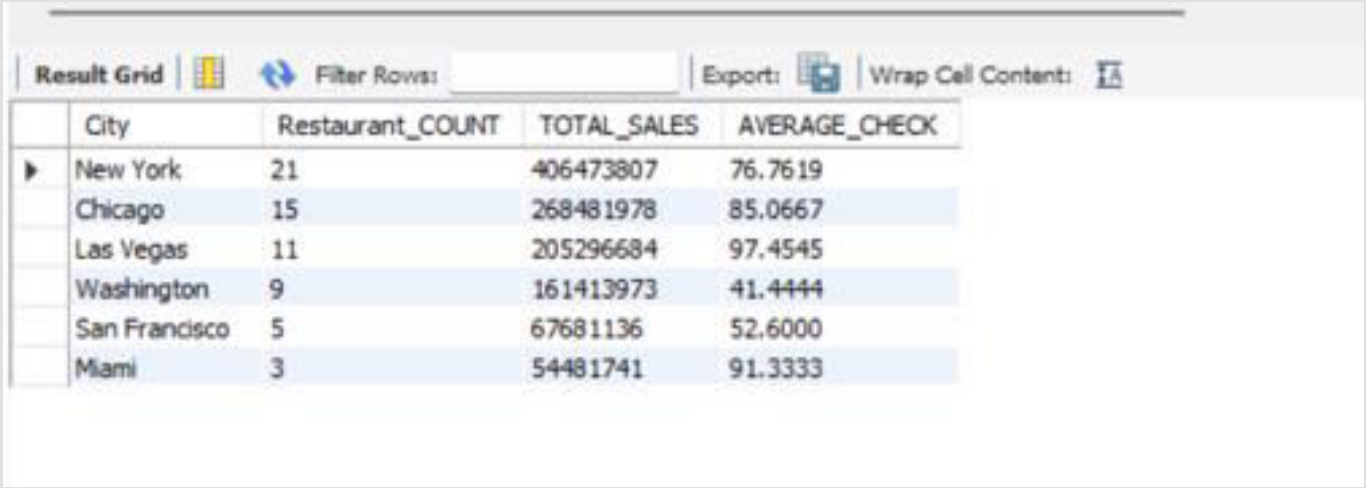
Result Grid | Filter Rows: | Export: | Wrap Cell Content:

Rank	Restaurant	Average Check	State
12	Prime 112	135	Fla.
64	Prime & Provisions	130	Ill.
21	Bazaar Meat by Jose Andres	119	Nev.
46	Swift & Sons	119	Ill.
31	Quality Meats	115	N.Y.
80	BOA Steakhouse	110	Calif.
18	Gibsons Italia	108	Ill.
10	Smith & Wollensky	107	N.Y.
44	Chops Lobster Bar	106	Ga.
52	Abe & Louie's	106	Mass.
7	Top of the World at the ST...	103	Nev.
56	Delmonico Steakhouse	103	Nev.
24	RPM Steak	102	Ill.

Sales by City (Independents)

SQL:

```
SELECT City, COUNT(Restaurant) AS Restaurant_COUNT, SUM(Sales) AS TOTAL_SALES, AVG(`Average Check`) AS
AVERAGE_CHECK
FROM Cleaned_independence100
GROUP BY City
HAVING Restaurant_COUNT > 2
ORDER BY TOTAL_SALES DESC;
```



Result Grid | Filter Rows: | Export: | Wrap Cell Content:

City	Restaurant_COUNT	TOTAL_SALES	AVERAGE_CHECK
New York	21	406473807	76.7619
Chicago	15	268481978	85.0667
Las Vegas	11	205296684	97.4545
Washington	9	161413973	41.4444
San Francisco	5	67681136	52.6000
Miami	3	54481741	91.3333

Results and Discussion

Analysis of the datasets revealed several key insights:

- Independent restaurants with high sales-to-meals ratios indicate operational efficiency.
- Franchised Future50 chains show higher average unit volumes compared to non-franchised.
- Certain states dominate the independent restaurant scene, highlighting regional market strengths.
- Top 250 segmentation analysis demonstrates competitive pressures in Quick Service categories.

Conclusion and Future Work

This project demonstrates the power of SQL for handling large-scale datasets in the restaurant industry. Through data cleaning, aggregation, and advanced analysis, meaningful insights were extracted. Future work can include integrating SQL with BI tools such as Power BI or Tableau for visualization, or extending the analysis with predictive modeling using Python and SQL integration.