


Discovery Challenge – Aggregate Queries – Tripadvisor Rankings



Hunter Betz



Right off the bat...

- Restaurants with a score of 4, 4.5, and 5 still ranked low
- # of restaurant reviews doesn't really matter
- Restaurants with similar #s of 4 and 5 stars ranked lower than those with higher #s of 5 stars
- Some exceptions exist however
- “Popular” restaurants (restaurants with >100 reviews) generally did not make the top 10
- Only 4 with >100 reviews made it to the top 10

```
SELECT
    restaurant,
    rank,
    COUNT(*) total,
    SUM(CASE
        WHEN review_stars = 5 THEN 1
        ELSE 0
    END) fiveCount,
    SUM(CASE
        WHEN review_stars = 4 THEN 1
        ELSE 0
    END) fourCount,
    SUM(CASE
        WHEN review_stars = 3 THEN 1
        ELSE 0
    END) threeCount,
    SUM(CASE
        WHEN review_stars = 2 THEN 1
        ELSE 0
    END) twoCount,
    SUM(CASE
        WHEN review_stars = 1 THEN 1
        ELSE 0
    END) oneCount
FROM
    tripadvisor
GROUP BY rank;
```

Positive reviews (4 & 5 stars) within a short time span generally lead to a high rank

```
SELECT
    restaurant,
    rank,
    COUNT(*) total,
    SUM(CASE
        WHEN
            rank <= 10
            AND (review_date LIKE '8/%/2015'
                OR review_date LIKE '9/%/2015'
                OR review_date LIKE '10/%/2015'
                OR review_date LIKE '11/%/2015')
        THEN
            1
        ELSE 0
    END) lastFourMonths
FROM
    tripadvisor
WHERE
    rank <= 10
GROUP BY restaurant
ORDER BY rank;
```

On average 54% of reviews were made between August and November for top 10 restaurants

```
SELECT
    restaurant,
    rank,
    COUNT(*) total,
    SUM(CASE
        WHEN
            rank >= 50
            AND (review_date LIKE '8/%/2015'
                OR review_date LIKE '9/%/2015'
                OR review_date LIKE '10/%/2015'
                OR review_date LIKE '11/%/2015')
        THEN
            1
        ELSE 0
    END) lastFourMonths
FROM
    tripadvisor
WHERE
    rank >= 50
GROUP BY restaurant
ORDER BY rank;
```

On average 33% of reviews were made between August and November for bottom 10 restaurants

Results:

| restaurant ▼ | rank | total | lastFourMonths |
|---------------------------------|------|-------|----------------|
| Ritz Restaurant | 1 | 144 | 84 |
| Prologue Restaurant & Bar | 2 | 98 | 98 |
| Pierogarnia Mandu | 3 | 260 | 109 |
| Familia Bistro | 4 | 45 | 37 |
| Wine Bar & Restaurant Literacka | 5 | 157 | 66 |
| Jozef K | 6 | 80 | 30 |
| Retro Cafe | 7 | 186 | 61 |
| Flisak 76 | 8 | 83 | 40 |
| Allora | 9 | 9 | 7 |
| Monte Carlo Pizza i Wino | 10 | 26 | 9 |

Restaurants such Allora, Prologue, and Familia Bistro must have opened relatively recently since a large portion of their reviews are between a 4 month span

| restaurant | rank | total | lastFourMonths |
|-----------------------------|------|-------|----------------|
| Bioway | 50 | 36 | 11 |
| Gdanski Bowke | 51 | 218 | 54 |
| Tawerna Mestwin | 52 | 62 | 21 |
| La Famiglia Pizza | 53 | 32 | 13 |
| Labeerynt Craft Beer Pub | 54 | 11 | 10 |
| Billy's American Restaurant | 55 | 234 | 56 |
| Cala Naprzod | 56 | 62 | 8 |
| Pyra Bar | 57 | 58 | 24 |
| Cafe Kamienica | 58 | 38 | 6 |
| Szafarnia 10 | 59 | 20 | 9 |
| W Starym Kadrze | 60 | 32 | 2 |

These restaurants were far less popular since they received fewer reviews in the same 4 month span as the top 10 restaurants

Interpretations:

- New restaurants, or restaurants that are experiencing a popularity boost rise up the ranks faster than other restaurants
- Additionally, the wider the gap between the # of 5 stars a restaurant receives and # of 4 or lower stars, the better
- Restaurants with # of 5 stars closer to the # of 4 stars tend to rank lower