

Metrics for Classifying Businesses as Popular and Successful

Popularity Metric

The popularity metric is based on the idea that customer foot traffic (numCheckins) and satisfaction (reviewRating) are indicators for how popular a business is. It combines the total number of check-ins and average review ratings to gauge popularity. The formula applied is:

Popularity Score = $(0.8 \times \text{Total Check-Ins}) + (0.2 \times \text{Average Review Rating})$

This weighted approach shows the importance of both attracting customers and maintaining quality service, with a bias towards pure customer volume.

Data Items Utilized:

- Total Check-Ins (numCheckins from the Business table): Represents the volume of customer visits for each business.
- Average Review Rating (reviewRating from the Business table): Reflects average customer satisfaction.

The weights have a bias towards total number of checkins as I believe this is a more important factor for measuring pure popularity. The query used to calculate this metric is:

```
-- Popularity metric (PopularityScore)
SELECT
    Category.CategoryName,
    Business.Name AS BusinessName,
    SUM(Business.numCheckins) AS TotalCheckIns,
    AVG(Business.reviewRating) AS AvgReviewRating,
    (0.8 * SUM(Business.numCheckins) + 0.2 * AVG(Business.reviewRating)) AS
PopularityScore
FROM
    Business
JOIN
    Category ON Business.CategoryID = Category.CategoryID
GROUP BY
    Category.CategoryName, Business.Name
ORDER BY
    Category.CategoryName, PopularityScore DESC;
```

This query aggregates check-ins and review ratings for each business. It ranks businesses within their respective categories based on their popularity score.

Success Metric

The success metric is based on longevity and consistent customer satisfaction. It combines the operational period (measured from the earliest to the most recent review) with the average review rating:

Success Score = Longevity + Average Rating

Data Items Utilized:

- Longevity: The number of years between the first and last review dates for each business.
- Average Rating (AvgRating): The average star rating from customer reviews.

This formula shows that businesses that have been in business over many years, and have decent customer reviews, are deemed successful.

The corresponding SQL query is:

```
-- Success metric (SuccessScore)
SELECT
    Category.CategoryName,
    Business.Name,
    DATE_PART('year', MAX(Review.Date)) - DATE_PART('year', MIN(Review.Date)) AS
LongevityYears,
    AVG(Review.Stars) AS AvgRating,
    (0.7 *(DATE_PART('year', MAX(Review.Date)) - DATE_PART('year', MIN(Review.Date)))) +
0.3 * AVG(Review.Stars) AS SuccessScore
FROM
    Business
LEFT JOIN Review ON Business.BusinessID = Review.BusinessID
JOIN Category ON Business.CategoryID = Category.CategoryID
GROUP BY
    Business.Name, Category.CategoryName
ORDER BY
    Category.CategoryName, SuccessScore DESC;
```

This query extracts the operational span (longevity) and average ratings to calculate the success score, It then ranks businesses within their respective categories based on their success score. The weights have a bias towards longevity as I believe that is the main factor for successfulness.