

Milestone 1: Business Level Analysis Report

Analysis of Arizona Businesses using Yelp Dataset

1. Objective

- Analyze Yelp business data in Arizona to identify trends in business performance influenced by reviews, ratings, and geographical location.
- Highlight top-rated businesses and customer engagement in key categories.

Dataset Details:

- A refined dataset focused on businesses located in Arizona.

Tools Used:

- Apache Spark for efficient dataset querying and filtering.
- Python for data visualization.

2. Analysis and Findings

Query 1: Arizona Cities with Most Businesses

Objective:

- Identify cities in Arizona with the highest number of businesses.

Findings:

- Tucson leads with 9,249 businesses, followed by Oro Valley (286) and Marana (102).
- Smaller towns like Green Valley and Sahuarita also appear in the top 5.

Implication:

- Tucson serves as a central hub for businesses in Arizona.

Query 2: Average Business Ratings by Category

Objective:

- Evaluate average ratings of businesses across various categories.

Findings:

- Categories such as 'Diagnostic Services', 'Axe Throwing' and 'Jewelry Repair' have perfect 5.0 ratings.

Implication:

- Niche categories tend to excel in customer satisfaction.

Query 3: Highest Rated Restaurants by Zip Code

Objective:

- Identify zip codes with the highest-rated restaurants in Arizona.

Findings:

- Zip code 85701 stands out with an average rating of 3.96 and 22,901 reviews.
- Other top-performing zip codes include 85716 and 85719.

Implication:

- Centrally located zip codes generally achieve better ratings.

Query 4: Cities with Most Highly Rated Restaurants**Objective:**

- Find cities with the most restaurants rated 4 stars or higher.

Findings:

- Tucson tops the list with 986 highly-rated restaurants, while smaller cities like Oro Valley (31) and Marana (14) also contribute.

Implication:

- Urban centers tend to cluster more high-rated restaurants.

Query 5: Positive Review Analysis by Zip Code**Objective:**

- Identify zip codes with the highest count of positive reviews.

Findings:

- Zip code 85719 leads with 14,162 positive reviews, followed by 85705 (11,811).
- Zip codes 85711 and 85701 also rank high.

Implication:

- Specific areas show strong customer satisfaction, reflected in positive reviews.

3. Conclusion

Tucson dominates Arizona's Yelp dataset with the highest business count, top-rated restaurants, and most positive reviews. Specialty categories like 'Diagnostic Services' and 'Axe Throwing' achieve exceptional customer approval. Key areas like zip code 85719 emerge as hot spots for highly-rated dining experiences and positive feedback.

4. Recommendations**For Businesses:**

- Focus on Tucson and its surrounding regions for growth opportunities.

For Researchers:

- Explore the role of specific business attributes in driving customer satisfaction.