

## **Milestone 2 - Individual Report- User Level Analysis**

### **Overview**

This report explores user behavior, business performance, and sentiment trends in the Health & Medical sector in Arizona using the Yelp dataset. The analysis integrates multiple datasets—**reviews**, **users**, **businesses**, **tips**, and **check-ins**—to derive actionable insights. It includes 10 queries, categorized into four **simple** and six **complex** analyses. Key focus areas include identifying top contributors, influential businesses, user sentiment, and engagement trends.

### **Main Results & Analysis**

#### **Simple Queries**

1. **Top Users by Review Count:**
  - Identified the most active users by total reviews.
  - Observation: High engagement users had consistent ratings above 4.0.
  - **Plot Insight:** A bar chart highlighted the top 10 users, showing a steep drop in review counts after the first few users.
2. **Average Reviews per User:**
  - The average number of reviews per user in the dataset was calculated.
  - Insight: This metric highlighted a moderate user engagement trend in the Health & Medical category.
3. **Users with the Most Useful Votes:**
  - Ranked users based on the total number of "useful" votes received.
  - Observation: Influential users frequently provide detailed, informative reviews.
  - **Plot Insight:** A bar chart revealed the top users, emphasizing the correlation between engagement and utility.
4. **Users with the Most Funny Votes:**
  - Identified users whose reviews are considered humorous by others.
  - Insight: Humor plays a secondary role in user engagement compared to utility.
  - **Plot Insight:** A bar chart illustrated user rankings, showing how the top humorous reviewers differed from those with useful votes.

#### **Complex Queries**

1. **Most Influential Users by Votes:**
  - Users were ranked based on "useful," "funny," and "cool" votes.
  - Insight: High vote counts often correlated with users who write extensive reviews and maintain high average ratings.

- **Plot Insight:** A bar chart visually ranked influential users, with "useful" votes dominating user influence metrics.
- 2. **Sentiment Analysis by User Review Count:**
  - Explored how sentiment scores vary with review activity.
  - Observation: Active users tended to have neutral to positive sentiment, indicating consistent satisfaction.
  - **Plot Insight:** A scatter plot revealed a positive correlation between review count and sentiment, with high-rating users clustering in the upper-right quadrant.
- 3. **Trends in User Sentiment Over Time:**
  - Analyzed yearly sentiment and rating trends.
  - Insight: Sentiment scores were stable over time, while the volume of reviews steadily increased, reflecting growing customer engagement.
  - **Plot Insight:** A line chart demonstrated stable sentiment trends and a steady rise in average ratings, emphasizing the sector's resilience.
- 4. **User Behavior by Business Attributes:**
  - Examined how attributes like WiFi availability and price range affect user sentiment.
  - Observation: Businesses offering free WiFi and moderate pricing (levels 2-3) generally received higher sentiments.
- 5. **User Engagement by City:**
  - Identified cities with the highest user engagement based on total reviews.
  - Insight: Urban areas like Phoenix and Tucson dominated engagement metrics due to population density and service availability.
  - **Plot Insight:** A horizontal bar chart highlighted Phoenix as the top city for user engagement, with engagement tapering off in smaller cities.
- 6. **User Influence by Review Length:**
  - Analyzed the correlation between review length and "useful" votes.
  - Observation: Longer reviews received more useful votes, indicating the value of detailed feedback.
  - **Plot Insight:** A scatter plot showed a strong positive correlation, with longer reviews clustering at higher "useful" vote counts.

## Conclusion

The analysis reveals that user engagement and business performance in the Health & Medical sector are significantly influenced by detailed reviews, business attributes, and geographic location. Active users contribute positively to community sentiment, while businesses with customer-friendly attributes, like free WiFi and moderate pricing, achieve higher satisfaction levels. Sentiment trends have remained consistent over the years, despite growing review volumes, underscoring the stability of service quality. These insights can guide businesses to enhance customer experiences and engagement strategies.