INTRODUCTION

This report analyzes user behaviors for airport shuttle businesses in Arizona (AZ) using the Yelp dataset. Ten queries, categorized as simple (user-centric insights) and complex (multi-dataset analysis), provide trends in reviews, ratings, and elite user interactions. Data from the business, review, user, and tip datasets were filtered for "Airport Shuttles" in AZ and processed with PySpark for scalability. These insights help identify user preferences and trends to support data-driven strategies for businesses in this category.

ANALYSIS OF QUERIES

1. **Query 1 (Top 10 Users with Most Shuttle Reviews)**:
   - *Purpose*: Identify the top 10 users with the most shuttle reviews and their average reviews across all businesses.
   - *Analysis*: Groups users by ID and name, calculates total and average reviews, orders by total reviews in descending order, and limits results to 10.
   - *Result*: Richard tops with 4 shuttle reviews and an average of 38.75 reviews across businesses. Other users, like D, Ess, and Kurt, each have 2 reviews, with averages between 35.5 and 80.0.

2. **Query 2 (Top 10 Users with Highest Average Ratings)**:
   - *Purpose*: Identify the top 10 users with the highest average ratings given to shuttle services.
   - *Analysis*: Groups users by ID and name, calculates their average star rating, orders results in descending order of average rating, and limits to 10.
   - *Result*: All top users, including Jason, Chris, and Liz, have a perfect average rating of 5.0, showcasing consistent high ratings across their reviews.

3. **Query 3 (Top 10 Users with Most Reviews by Rating)**:
   - *Purpose*: Identify the top 10 users with the highest total reviews, broken down by star ratings.
   - *Analysis*: Groups users by ID and name, counts the number of reviews for each star rating (1 to 5), calculates total reviews, and filters users with at least one review. Results are ordered by total reviews in descending order, limited to 10.
   - *Result*: Ceci and Nancy have 2 five-star reviews each, while other users like Becky, Kelly, and Tommy have only 1 review, all rated 5 stars.

4. **Query 4 (Users with Reviews Across Multiple Businesses)**:
   - *Purpose*: Identify users who have reviewed more than one distinct business.
   - *Analysis*: Groups users by ID and name, counts the distinct businesses they have reviewed, filters users who have reviewed more than one business, and orders the results by the number of distinct businesses in descending order.
   - *Result*: Richard has reviewed 3 distinct businesses, while others like Aisha, Isabella, D, and Marco have reviewed 2 distinct businesses each.

5. **Query 5 (Users with Most Recent Reviews)**:
   - *Purpose*: Identify users who have the most recent shuttle service reviews.
   - *Analysis*: Groups users by ID and name, selects the most recent review date for each user, orders the results by the latest review date, and limits to the top 10.
   - *Result*:
     - Jason has the most recent review, dated 2024-11-22.
     - Other users like Chris and Liz have recent reviews dated around 2024-11-20, showing a consistent flow of new reviews.

6. **Query 6 (Top 100 Businesses with Most Useful Votes)**:
   - *Purpose*: Identify the top 100 businesses with the most useful votes from user reviews.
   - *Analysis*: Joins business_reviews and user_reviews, filters businesses with more than 10 reviews, and aggregates data for useful, funny, and cool votes, ordered by total useful votes.
   - *Result*:
     - Arizona Shuttle Service leads with 15,498 useful votes and a 2.5 average rating, followed by Arizona Stagecoach with 6,820 votes.

- Other businesses like FastPark & Relax and 24hr Express have high ratings (4.5–5.0) with fewer useful votes.

7. **Query 7 (Business Reviews in Tucson and Phoenix)**:
   - *Purpose*: Identify businesses in Tucson and Phoenix with user reviews, highlighting the average ratings, total compliments, and user activity.
   - *Analysis*: Filters businesses in Tucson and Phoenix, where review count exceeds 5, and user review count exceeds 100. Computes average user ratings, total compliments, and tracks the first yelp date for each business. Results are ordered by business name.
   - *Result*:
     - "24hr Express" and "Limoport Transportation" stand out with perfect 5.0 business ratings, while businesses like "Arizona Shuttle Service" and "Yellow Cab" show significantly lower ratings (2.5 and 1.5, respectively).
     - Compliments vary widely, with "Arizona Shuttle Service" leading with 152,438 total compliments..

8. **Query 8 (User Tips for Businesses)**:
   - *Purpose*: Identify businesses with user tips and list the users associated with those tips.
   - *Analysis*: Joins the business_tips and user_tips tables on business_id, collecting the user names for each business. The results are limited to the top 5 businesses based on business name..
   - *Result*:
     - The businesses "King Pin Cab Company" and "24hr Express" have multiple tips from a small group of users, with "King Pin Cab Company" showing repeated names like Ceci.
     - "FastPark & Relax" and "Jeannie's Van & Taxi Service" have significantly more user names, indicating a broader range of contributors.

9. **Query 9 (User Names for Business Tips)**:
   - *Purpose*: Identify the top businesses with the most unique users who have provided tips, highlighting the number of users and the user names..
   - *Analysis*: Joins the business_tips and user_tips tables to collect distinct user names for each business. Groups the data by business name, counts the number of unique users per business, and orders the results by the number of users in descending order.
   - *Result*:
     - "FastPark & Relax" leads with 49 distinct users offering tips, followed by "Jeannie's Van & Taxi Service" and "Yellow Cab" with 36 users each.
     - Businesses like "Sky Harbor Express" have fewer unique users, with only 4 distinct contributors.

10. **Query 10 (Elite vs. Non-Elite Users' Reviews and Fans by Business)**:
    - *Purpose*: To compare review counts, user activity, and fan base between elite and non-elite users across businesses.
    - *Analysis*: The query calculates the number of elite and non-elite users, total reviews, average reviews per user, and total fans, grouped by business name.
    - *Result*:
      - FastPark & Relax leads with 343 elite users and 14,749 reviews.
      - Arizona Stagecoach has 1 elite user with a high average of 1,261 reviews.
      - King Pin Cab Company and Adobe Transportation have fewer elite users but still show notable review counts.
      - Elite users contribute significantly to the review activity across businesses.

CONCLUSION

This report analyzes user engagement with airport shuttle businesses in Arizona, revealing key trends in reviews, ratings, and fan activity. Elite users significantly contribute to businesses like FastPark & Relax, with high review counts and fan bases. Businesses such as Arizona Shuttle Service and Yellow Cab show lower ratings, suggesting areas for improvement. Companies with more active users, like Jeannie's Van & Taxi Service, display stronger engagement. Overall, businesses with higher elite user interactions may experience better customer retention, providing insights for strategic improvements.