Group 13 Project Report

Class name: BUAN 6320.007 - Database Foundations for Business Analytics - F23

Name: Group 13: Rikin Patel(rxp220105), Rucheek Rajeev Kashyap(rxk230010),Krishna Chaithanya (kxk230041)
Date: December 8, 2023

Inside AirBnB - USA Dataset (Jersey City)

1 Dataset Link

Inside AirBnB - USA (Jersey City) — Kaggle - Click here

2 Content of the dataset

Inside Airbnb is a mission-driven project that provides data and advocacy about Airbnb's impact on residential communities. We work towards a vision where data and information empower communities to understand, decide, and control the role of renting residential homes to tourists.

2.1 File Descriptions (Applicable for Each Subfolder)

listings_detailed.csv: Detailed Listings data

calendar.csv: Detailed Calendar Data

reviews_detailed.csv: Detailed Review Data

listings.csv: Summary information and metrics for listings in Amsterdam (good for visualizations).

reviews.csv: Summary Review data and Listing ID (to facilitate time-based analytics and visualizations linked to a listing).

3 Business Objective

The primary objectives of this business analysis for Airbnb are to enhance decision-making, operational efficiency, and overall business performance. Our focus is on providing strategic insights to executive leadership, enabling them to optimize key business decisions. We aim to enhance marketing effectiveness by leveraging booking trends, assisting marketing teams in planning impactful seasonal campaigns. For operations and pricing, the goal is to maximize efficiency by optimizing property type distribution and pricing strategies. Strengthening host partnerships is another key objective, with a focus on identifying and rewarding high-performing hosts. The analysis aims to facilitate property diversification by providing insights to property managers and improve guest satisfaction by addressing potential issues based on review scores. Additionally, we seek to evaluate the impact of premium features on pricing and listings, aligning product offerings with customer preferences. Financial decision-making will be supported through comprehensive evaluations, assisting financial and investment analysts in assessing pricing strategies and diversification plans.

4 Who is our target audience?

Our business analysis caters to various stakeholders within Airbnb, each with distinct responsibilities and interests. The primary target audience includes:

1. **Executive Leadership:**

   – CEOs and top executives seeking strategic insights into overall business performance.
   – Decision-makers interested in understanding how different decisions impact the company's success.

2. **Marketing Teams:**

   – Marketing managers and teams aiming to plan effective seasonal campaigns.
   – Individuals focused on leveraging booking trends to optimize marketing strategies.

3. **Operations and Pricing Teams:**

   – Operations managers responsible for property type distribution and diversification.
   – Pricing teams looking to optimize pricing strategies for different room types.

4. **Host Partnerships and Incentives:**

   – Teams responsible for identifying and rewarding high-performing hosts.
   – Individuals working on forming partnerships and incentive programs based on host performance.

5. **Property Management and Diversification:**

   – Property managers interested in opportunities to diversify property types.
   – Teams focused on improving the variety and distribution of listed properties.

6. **Customer Experience and Satisfaction Teams:**

   – Professionals monitoring trends in guest satisfaction and addressing potential issues.
   – Individuals dedicated to enhancing overall guest experience based on review scores.

7. **Product Development and Innovation:**

   – Teams exploring premium feature rollouts and understanding the impact on pricing and listings.
   – Individuals keen on aligning product offerings with customer preferences.

8. **Financial and Investment Analysts:**

   – Financial analysts evaluating the financial impact of pricing strategies and diversification.
   – Teams assessing the return on investment for various business decisions.

5 Purpose:

The purpose of this business analysis is to provide comprehensive insights into various facets of Airbnb's operations and performance. By leveraging data-driven approaches and analytics, our aim is to:

1. Identify strategic opportunities for business growth and improvement.

2. Facilitate informed decision-making for executive leadership.

3. Assist marketing teams in crafting effective seasonal campaigns based on booking trends.

4. Support operations and pricing teams in optimizing property type distribution and pricing strategies.

5. Enable host partnerships and incentive programs by identifying high-performing hosts.

6. Provide recommendations for property management and diversification opportunities.

7. Enhance the overall customer experience by monitoring trends in guest satisfaction.

8. Aid in product development and innovation, especially related to premium feature rollouts.

9. Guide financial and investment analysts in evaluating the financial impact of business decisions.

By addressing the specific needs of various stakeholders, this analysis aims to empower decision-makers with actionable insights, fostering the success and sustainable growth of Airbnb.

6 Problem

Airbnb, as a dynamic player in the hospitality industry, faces multifaceted challenges and opportunities that necessitate a thorough business analysis. The identified problem areas include:

1. **Optimizing Business Strategies:** There is a need to optimize overall business strategies to ensure sustained growth and competitiveness in the evolving market landscape.

2. **Marketing Effectiveness:** The effectiveness of marketing campaigns, especially seasonal promotions, needs to be assessed and enhanced for better customer engagement and increased bookings.

3. **Operational Efficiency:** Streamlining operations, particularly in property type distribution and pricing, is crucial for maximizing occupancy rates and revenue.

4. **Host Performance Recognition:** Identifying and rewarding high-performing hosts through effective partnerships and incentive programs is essential for maintaining host satisfaction and loyalty.

5. **Property Diversification:** Exploring opportunities to diversify property types is important for catering to a broad range of customer preferences and increasing the overall appeal of listings.

6. **Guest Satisfaction Monitoring:** Continuous monitoring of guest satisfaction trends is required to promptly address issues and enhance the overall guest experience.

7. **Premium Feature Impact:** Understanding the impact of premium feature rollouts on pricing and listings is crucial for aligning product offerings with customer preferences.

8. **Financial Decision Support:** Financial and investment decisions, including pricing strategies and diversification plans, require thorough evaluation to ensure a positive return on investment.

This business analysis aims to tackle these challenges by providing actionable insights and strategic recommendations tailored to each identified problem area.

7 Data Loading and Data Modelling

7.1 Data Loading Concept Used

We used the below code to load the data into the schema we created for the airbnb database. Since the data was in .xlsx format we faced several issues such as datetime issues, scientific notation issues, etc. We had to perform some excel operations and then load the data into the database using below code.

```sql
LOAD DATA LOCAL INFILE 'path/to/data'
INTO TABLE reviews
FIELDS TERMINATED BY ',' ENCLOSED BY '"'
LINES TERMINATED BY '\r\n'
IGNORE 1 ROWS;
```

<div align="center">Listing 1: Data Loading</div>

7.2 MySQL Workbench Forward Engineering

```sql
-- MySQL Workbench Forward Engineering

SET @OLD_UNIQUE_CHECKS=@@UNIQUE_CHECKS, UNIQUE_CHECKS=0;
SET @OLD_FOREIGN_KEY_CHECKS=@@FOREIGN_KEY_CHECKS, FOREIGN_KEY_CHECKS=0;
SET @OLD_SQL_MODE=@@SQL_MODE, SQL_MODE='ONLY_FULL_GROUP_BY,
   STRICT_TRANS_TABLES,NO_ZERO_IN_DATE,NO_ZERO_DATE,
   ERROR_FOR_DIVISION_BY_ZERO,NO_ENGINE_SUBSTITUTION';


-- -----------------------------------------------------
-- Schema mydb
-- -----------------------------------------------------
-- -----------------------------------------------------
-- Schema airbnb
-- -----------------------------------------------------

-- -----------------------------------------------------
-- Schema airbnb
-- -----------------------------------------------------
CREATE SCHEMA IF NOT EXISTS `airbnb` DEFAULT CHARACTER SET utf8mb4
   COLLATE utf8mb4_0900_ai_ci ;
USE `airbnb` ;

-- -----------------------------------------------------
-- Table `airbnb`.`listings_detailed`
-- -----------------------------------------------------
CREATE TABLE IF NOT EXISTS `airbnb`.`listings_detailed` (
  `id` INT NOT NULL,
  `listing_url` VARCHAR(255) NULL DEFAULT NULL,
  `scrape_id` DECIMAL(18,2) NULL DEFAULT NULL,
  `last_scraped` DATE NULL DEFAULT NULL,
  `source` VARCHAR(50) NULL DEFAULT NULL,
  `name` VARCHAR(255) NULL DEFAULT NULL,
  `description` TEXT NULL DEFAULT NULL,
  `neighborhood_overview` TEXT NULL DEFAULT NULL,
  `picture_url` VARCHAR(255) NULL DEFAULT NULL,
  `host_id` INT NULL DEFAULT NULL,
  `host_url` VARCHAR(255) NULL DEFAULT NULL,
  `host_name` VARCHAR(255) NULL DEFAULT NULL,
  `host_since` DATE NULL DEFAULT NULL,
  `host_location` VARCHAR(255) NULL DEFAULT NULL,
  `host_about` TEXT NULL DEFAULT NULL,
```

```sql
`host_response_time` VARCHAR(50) NULL DEFAULT NULL ,
`host_response_rate` VARCHAR(50) NULL DEFAULT NULL ,
`host_acceptance_rate` VARCHAR(50) NULL DEFAULT NULL ,
`host_is_superhost` VARCHAR(1) NULL DEFAULT NULL ,
`host_thumbnail_url` VARCHAR(255) NULL DEFAULT NULL ,
`host_picture_url` VARCHAR(255) NULL DEFAULT NULL ,
`host_neighbourhood` VARCHAR(255) NULL DEFAULT NULL ,
`host_listings_count` INT NULL DEFAULT NULL ,
`host_total_listings_count` INT NULL DEFAULT NULL ,
`host_verifications` TEXT NULL DEFAULT NULL ,
`host_has_profile_pic` VARCHAR(1) NULL DEFAULT NULL ,
`host_identity_verified` VARCHAR(1) NULL DEFAULT NULL ,
`neighbourhood` VARCHAR(255) NULL DEFAULT NULL ,
`neighbourhood_cleansed` VARCHAR(255) NULL DEFAULT NULL ,
`neighbourhood_group_cleansed` VARCHAR(255) NULL DEFAULT NULL ,
`latitude` DECIMAL(9,6) NULL DEFAULT NULL ,
`longitude` DECIMAL(9,6) NULL DEFAULT NULL ,
`property_type` VARCHAR(50) NULL DEFAULT NULL ,
`room_type` VARCHAR(50) NULL DEFAULT NULL ,
`accommodates` INT NULL DEFAULT NULL ,
`bathrooms_text` VARCHAR(50) NULL DEFAULT NULL ,
`bedrooms` INT NULL DEFAULT NULL ,
`beds` INT NULL DEFAULT NULL ,
`amenities` TEXT NULL DEFAULT NULL ,
`price` DECIMAL(10,2) NULL DEFAULT NULL ,
`minimum_nights` INT NULL DEFAULT NULL ,
`maximum_nights` INT NULL DEFAULT NULL ,
`minimum_minimum_nights` INT NULL DEFAULT NULL ,
`maximum_minimum_nights` INT NULL DEFAULT NULL ,
`minimum_maximum_nights` INT NULL DEFAULT NULL ,
`maximum_maximum_nights` INT NULL DEFAULT NULL ,
`minimum_nights_avg_ntm` INT NULL DEFAULT NULL ,
`maximum_nights_avg_ntm` INT NULL DEFAULT NULL ,
`calendar_updated` VARCHAR(50) NULL DEFAULT NULL ,
`has_availability` VARCHAR(1) NULL DEFAULT NULL ,
`availability_30` INT NULL DEFAULT NULL ,
`availability_60` INT NULL DEFAULT NULL ,
`availability_90` INT NULL DEFAULT NULL ,
`availability_365` INT NULL DEFAULT NULL ,
`calendar_last_scraped` DATE NULL DEFAULT NULL ,
`number_of_reviews` INT NULL DEFAULT NULL ,
`number_of_reviews_ltm` INT NULL DEFAULT NULL ,
`number_of_reviews_l30d` INT NULL DEFAULT NULL ,
`first_review` DATE NULL DEFAULT NULL ,
`last_review` DATE NULL DEFAULT NULL ,
`review_scores_rating` DECIMAL(4,2) NULL DEFAULT NULL ,
`review_scores_accuracy` DECIMAL(4,2) NULL DEFAULT NULL ,
`review_scores_cleanliness` DECIMAL(4,2) NULL DEFAULT NULL ,
`review_scores_checkin` DECIMAL(4,2) NULL DEFAULT NULL ,
`review_scores_communication` DECIMAL(4,2) NULL DEFAULT NULL ,
`review_scores_location` DECIMAL(4,2) NULL DEFAULT NULL ,
`review_scores_value` DECIMAL(4,2) NULL DEFAULT NULL ,
`license` VARCHAR(50) NULL DEFAULT NULL ,
`instant_bookable` VARCHAR(1) NULL DEFAULT NULL ,
`calculated_host_listings_count` INT NULL DEFAULT NULL ,
`calculated_host_listings_count_entire_homes` INT NULL DEFAULT NULL ,
`calculated_host_listings_count_private_rooms` INT NULL DEFAULT NULL ,
`calculated_host_listings_count_shared_rooms` INT NULL DEFAULT NULL ,
```

```sql
  `reviews_per_month` DECIMAL(5,2) NULL DEFAULT NULL,
  PRIMARY KEY (`id`))
ENGINE = InnoDB
DEFAULT CHARACTER SET = utf8mb4
COLLATE = utf8mb4_0900_ai_ci;



-- -----------------------------------------------------
-- Table `airbnb`.`listings`
-- -----------------------------------------------------
CREATE TABLE IF NOT EXISTS `airbnb`.`listings` (
  `id` BIGINT NOT NULL,
  `name` VARCHAR(255) NULL DEFAULT NULL,
  `host_id` BIGINT NULL DEFAULT NULL,
  `host_name` VARCHAR(255) NULL DEFAULT NULL,
  `neighbourhood` VARCHAR(255) NULL DEFAULT NULL,
  `latitude` DECIMAL(9,6) NULL DEFAULT NULL,
  `longitude` DECIMAL(9,6) NULL DEFAULT NULL,
  `room_type` VARCHAR(50) NULL DEFAULT NULL,
  `price` INT NULL DEFAULT NULL,
  `minimum_nights` INT NULL DEFAULT NULL,
  `number_of_reviews` INT NULL DEFAULT NULL,
  `last_review` DATE NULL DEFAULT NULL,
  `reviews_per_month` DECIMAL(5,2) NULL DEFAULT NULL,
  `calculated_host_listings_count` INT NULL DEFAULT NULL,
  `availability_365` INT NULL DEFAULT NULL,
  `number_of_reviews_ltm` INT NULL DEFAULT NULL,
  `license` VARCHAR(50) NULL DEFAULT NULL,
  `listings_detailed_id` INT NOT NULL,
  PRIMARY KEY (`id`),
  INDEX `fk_listings_listings_detailed_idx` (`listings_detailed_id` ASC
    ) VISIBLE,
  CONSTRAINT `fk_listings_listings_detailed`
    FOREIGN KEY (`listings_detailed_id`)
    REFERENCES `airbnb`.`listings_detailed` (`id`)
    ON DELETE NO ACTION
    ON UPDATE NO ACTION)
ENGINE = InnoDB
DEFAULT CHARACTER SET = utf8mb4
COLLATE = utf8mb4_0900_ai_ci;



-- -----------------------------------------------------
-- Table `airbnb`.`calendar`
-- -----------------------------------------------------
CREATE TABLE IF NOT EXISTS `airbnb`.`calendar` (
  `listing_id` BIGINT NOT NULL,
  `date` DATE NULL DEFAULT NULL,
  `available` CHAR(1) NULL DEFAULT NULL,
  `price` DECIMAL(10,2) NULL DEFAULT NULL,
  `adjusted_price` DECIMAL(10,2) NULL DEFAULT NULL,
  `minimum_nights` INT NULL DEFAULT NULL,
  `maximum_nights` INT NULL DEFAULT NULL,
  `listings_id` BIGINT NOT NULL,
  PRIMARY KEY (`listing_id`, `listings_id`),
  INDEX `fk_calendar_listings1_idx` (`listings_id` ASC) VISIBLE,
  CONSTRAINT `fk_calendar_listings1`
    FOREIGN KEY (`listings_id`)
```
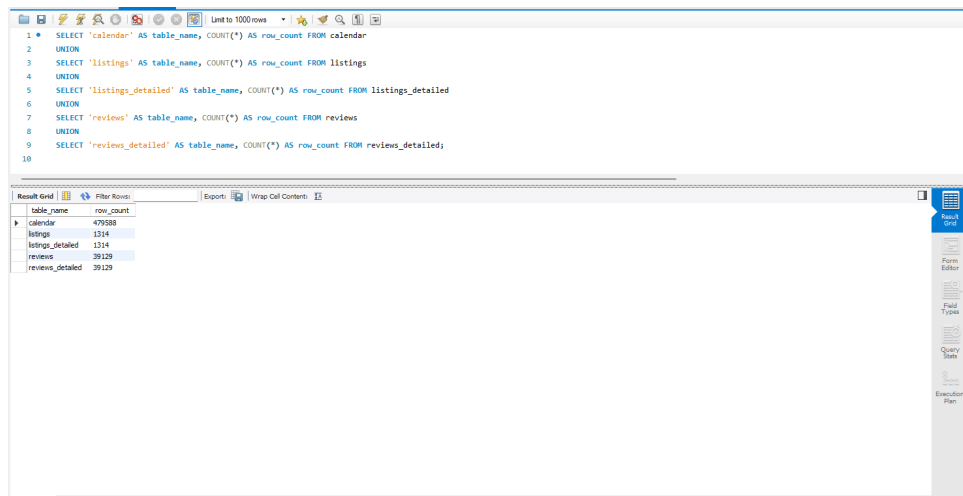
```sql
    REFERENCES `airbnb`.`listings` (`id`)
    ON DELETE NO ACTION
    ON UPDATE NO ACTION)
ENGINE = InnoDB
DEFAULT CHARACTER SET = utf8mb4
COLLATE = utf8mb4_0900_ai_ci;


-- -----------------------------------------------------
-- Table `airbnb`.`reviews_detailed`
-- -----------------------------------------------------
CREATE TABLE IF NOT EXISTS `airbnb`.`reviews_detailed` (
  `listing_id` INT NOT NULL,
  `id` INT NOT NULL,
  `date` DATE NULL DEFAULT NULL,
  `reviewer_id` INT NULL DEFAULT NULL,
  `reviewer_name` VARCHAR(255) NULL DEFAULT NULL,
  `comments` TEXT NULL DEFAULT NULL,
  PRIMARY KEY (`listing_id`, `id`))
ENGINE = InnoDB
DEFAULT CHARACTER SET = utf8mb4
COLLATE = utf8mb4_0900_ai_ci;


-- -----------------------------------------------------
-- Table `airbnb`.`reviews`
-- -----------------------------------------------------
CREATE TABLE IF NOT EXISTS `airbnb`.`reviews` (
  `listing_id` INT NOT NULL,
  `date` DATE NOT NULL,
  `listings_id` BIGINT NOT NULL,
  `reviews_detailed_listing_id` INT NOT NULL,
  `reviews_detailed_id` INT NOT NULL,
  PRIMARY KEY (`listing_id`, `date`, `listings_id`),
  INDEX `fk_reviews_listings1_idx` (`listings_id` ASC) VISIBLE,
  INDEX `fk_reviews_reviews_detailed1_idx` (
   reviews_detailed_listing_id` ASC, `reviews_detailed_id` ASC) VISIBLE
   ,
  CONSTRAINT `fk_reviews_listings1`
    FOREIGN KEY (`listings_id`)
    REFERENCES `airbnb`.`listings` (`id`)
    ON DELETE NO ACTION
    ON UPDATE NO ACTION,
  CONSTRAINT `fk_reviews_reviews_detailed1`
    FOREIGN KEY (`reviews_detailed_listing_id` , `reviews_detailed_id`)
    REFERENCES `airbnb`.`reviews_detailed` (`listing_id` , `id`)
    ON DELETE NO ACTION
    ON UPDATE NO ACTION)
ENGINE = InnoDB
DEFAULT CHARACTER SET = utf8mb4
COLLATE = utf8mb4_0900_ai_ci;


SET SQL_MODE=@OLD_SQL_MODE;
SET FOREIGN_KEY_CHECKS=@OLD_FOREIGN_KEY_CHECKS;
SET UNIQUE_CHECKS=@OLD_UNIQUE_CHECKS;
```
Listing 2: Forward Engineering

7.2.1 Output:

Figure 1: Total Number of rows in each table after data loading

7.3 Logical Model
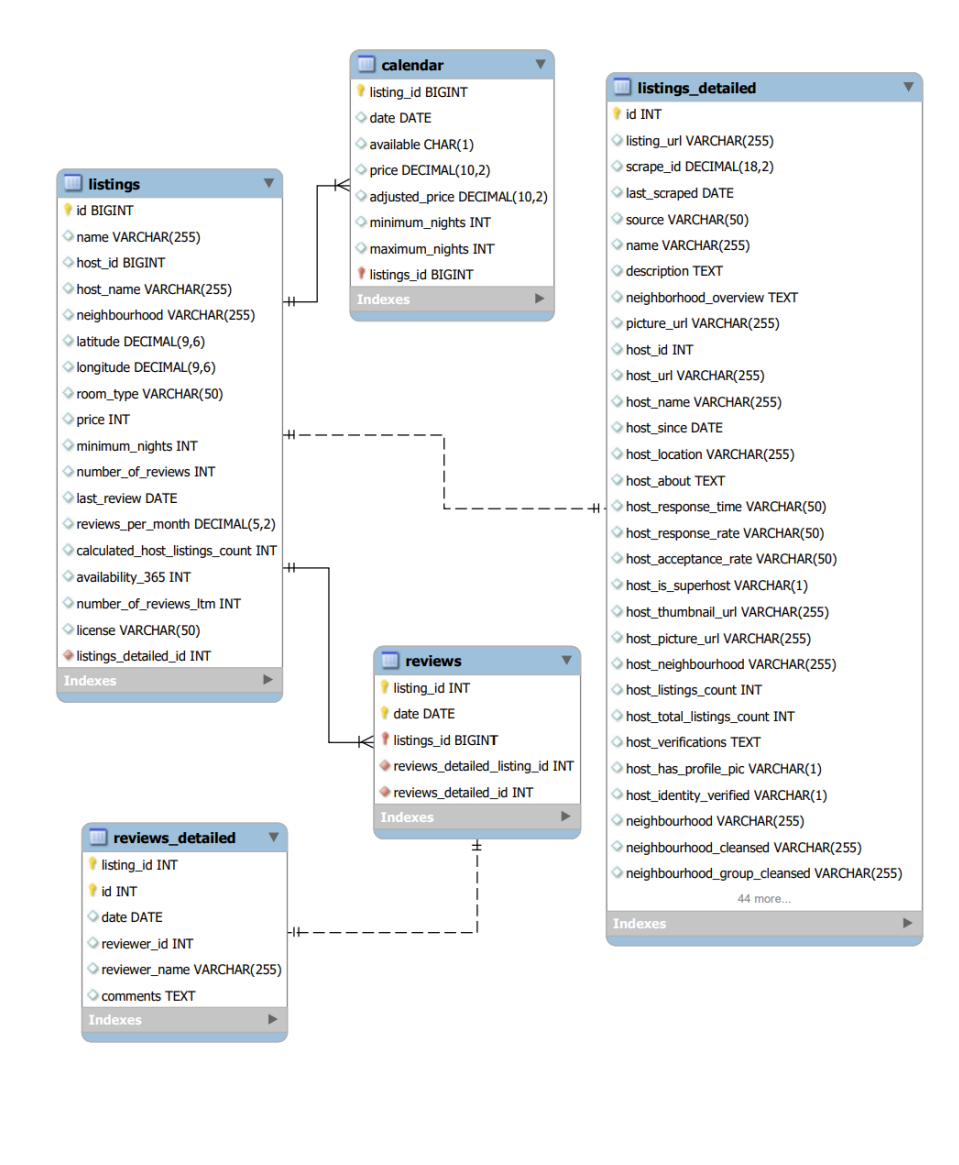
Figure 2: Logical Model

7.4   Conceptual Diagram
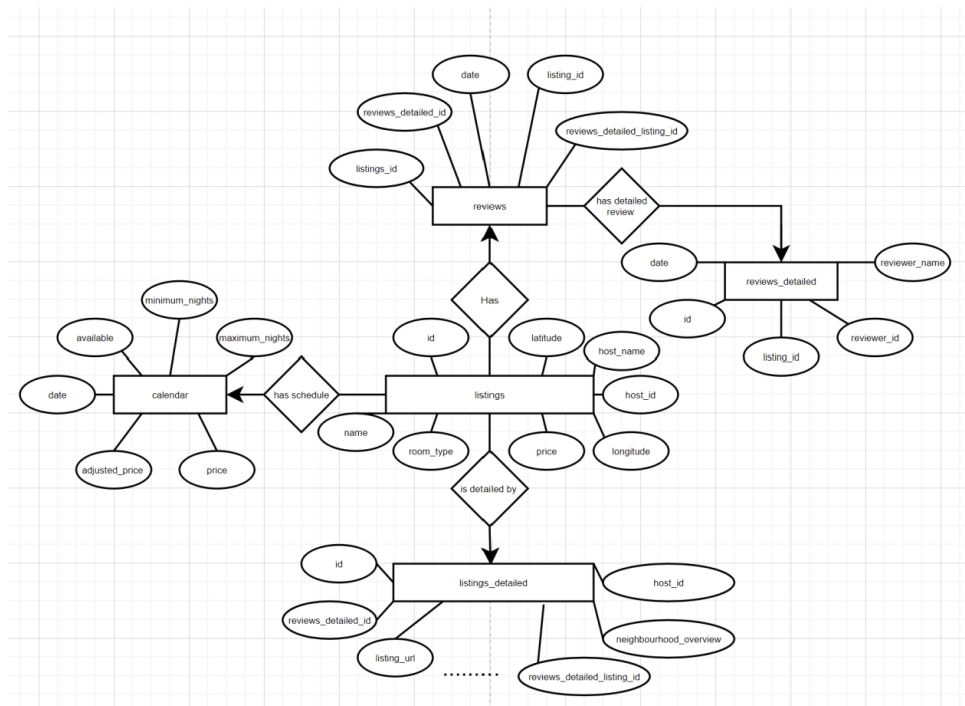
Figure 3: Logical Diagram

8 Project Insights

Exploring the dynamics of an Airbnb listing database, our project unveils valuable insights ranging from host performance metrics to optimal pricing strategies. Delve into a comprehensive analysis, providing actionable information for better decision-making in the hospitality industry.
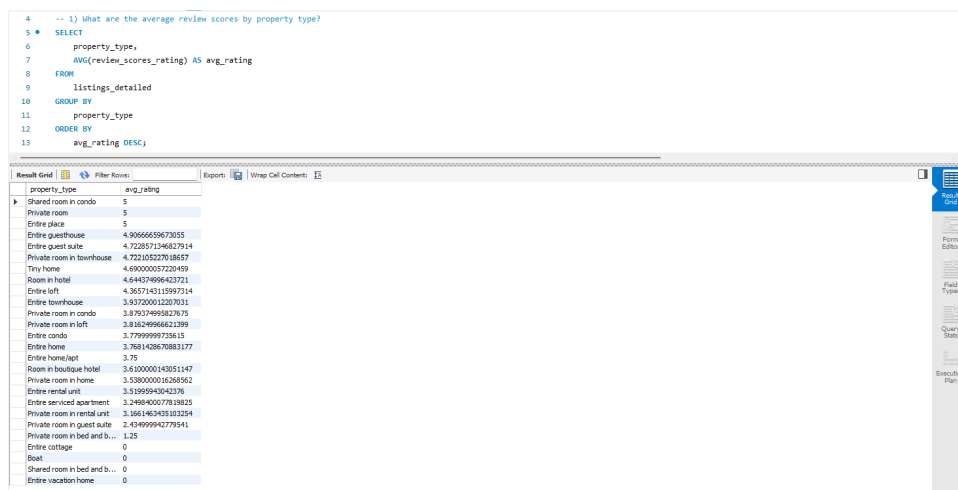
8.1 What are the average review scores by property type?

8.1.1 Code:

```sql
-- 1) What are the average review scores by property type?
SELECT
    property_type,
    AVG(review_scores_rating) AS avg_rating
FROM
    listings_detailed
GROUP BY
    property_type
ORDER BY
    avg_rating DESC;
```

Listing 3: Average Review Scores by Property Type

8.1.2 Output:

Figure 4: Average review scores by property type

8.1.3 Business Analysis:

Observations:

- The highest average review scores are observed for shared rooms in condos, private rooms, and entire places, all receiving perfect ratings of 5.

- Diverse property types such as entire guesthouses and entire guest suites also have high average scores, indicating positive guest experiences.

- Some property types, like entire vacation homes, shared rooms in bed and breakfast, and entire cottages, have lower or even zero average scores.

Actions:

- **Enhance Low-Rated Property Types:** Investigate and address concerns in property types with lower ratings (e.g., entire vacation homes, shared rooms in bed and breakfast) to improve overall guest satisfaction.

- **Promote Diverse Options:** Emphasize and market diverse property types with high ratings to attract guests seeking unique accommodation experiences.
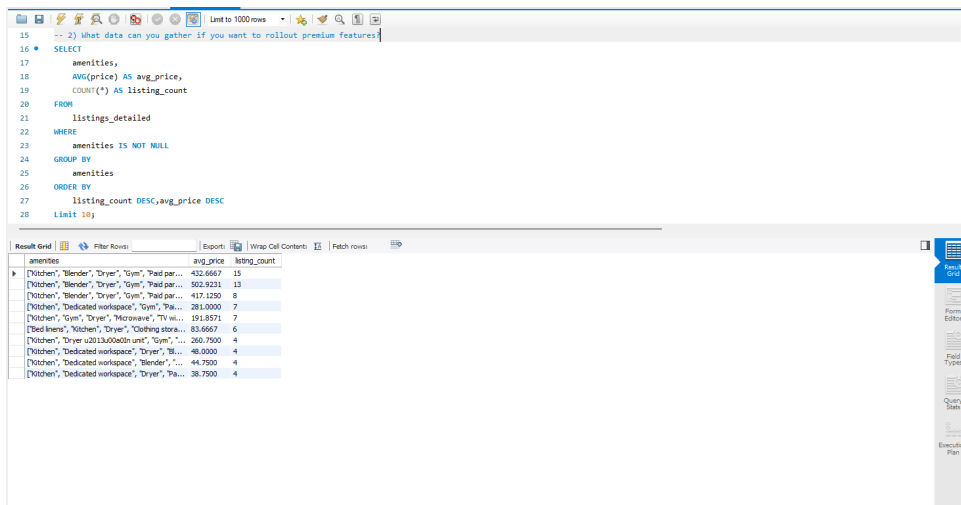
8.2 What data can you gather if you want to rollout premium features?

8.2.1 Code:

```sql
-- 2) What data can you gather if you want to rollout premium features?
SELECT
    amenities,
    AVG(price) AS avg_price,
    COUNT(*) AS listing_count
FROM
    listings_detailed
WHERE
    amenities IS NOT NULL
GROUP BY
    amenities
ORDER BY
    listing_count DESC,avg_price DESC
Limit 10;
```

Listing 4: Premium Feature Rollout Data

8.2.2 Output:

Figure 5: Premium Feature Rollout

8.2.3 Business Analysis:

Observations:

- The top three amenities associated with the highest average prices are a well-equipped kitchen, a gym, and a dryer, suggesting guests are willing to pay more for these conveniences.

- Listings with premium amenities such as a dedicated workspace, a blender, and paid parking on premises also command higher prices and have a notable presence.

- The amenity combination with the highest average price is a kitchen, blender, dryer, gym, and paid parking on premises.

Actions:

- **Highlight Premium Amenities:** In marketing and listing descriptions, emphasize properties with premium amenities to attract guests seeking a higher-end experience.

- **Explore Amenity Combinations:** Consider offering bundled amenity packages, allowing guests to customize their stay and potentially increase revenue.

- **Optimize Pricing Strategy:** Adjust pricing for listings with premium features to maximize revenue while remaining competitive in the market.
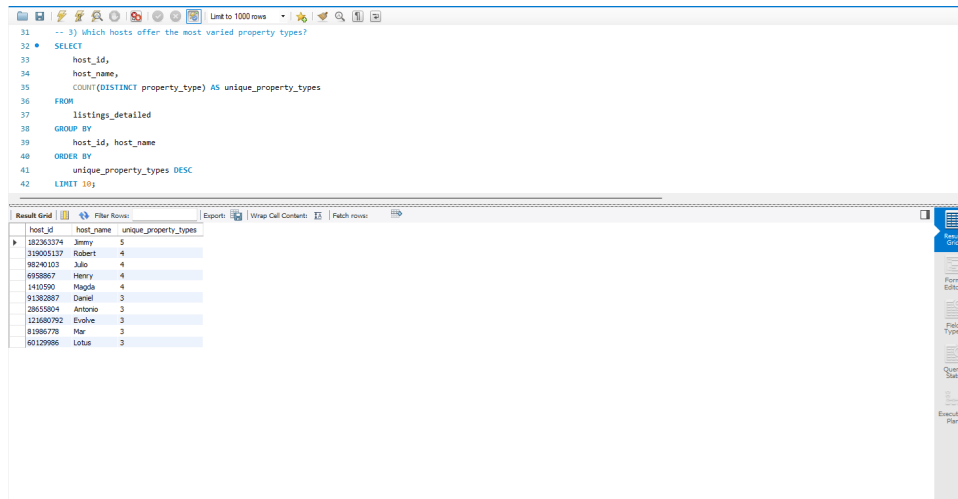
8.3 Which hosts offer the most varied property types?

8.3.1 Code:

```sql
-- 3) Which hosts offer the most varied property types?
SELECT
    host_id,
    host_name,
    COUNT(DISTINCT property_type) AS unique_property_types
FROM
    listings_detailed
GROUP BY
    host_id, host_name
ORDER BY
    unique_property_types DESC
LIMIT 10;
```

Listing 5: Hosts Offering Varied Property Types

8.3.2 Output:

Figure 6: Hosts with most varied property types

8.3.3 Business Analysis:

Observations: The analysis of hosts offering the most varied property types reveals the following key insights:

- Host #182363374 (Jimmy) stands out with an impressive diversity, offering five unique property types.

- Hosts #319005137 (Robert), #98240103 (Julio), and #6958867 (Henry) exhibit substantial variety with four distinct property types each.

- Magda (#1410590) is another host contributing to the diversity, managing four unique property types.

- Several hosts, including Daniel (#91382887), Antonio (#28655804), Evolve (#121680792), Mar (#81986778), and Lotus (#60129986), offer three different property types each.

Actions: To capitalize on the diversity observed among hosts offering multiple property types, the following actions are recommended:

- **Recognition and Collaboration:** Acknowledge and recognize hosts like Jimmy for their diverse property portfolios. Explore collaboration opportunities to enhance Airbnb's property variety.

- **Promotion Strategies:** Work closely with hosts Robert, Julio, Henry, and Magda to develop targeted promotional strategies highlighting their diverse property offerings.

- **Support for Expansion:** Assist hosts with three unique property types, such as Daniel, Antonio, Evolve, Mar, and Lotus, in expanding their listings. Provide support and resources to encourage further diversification.
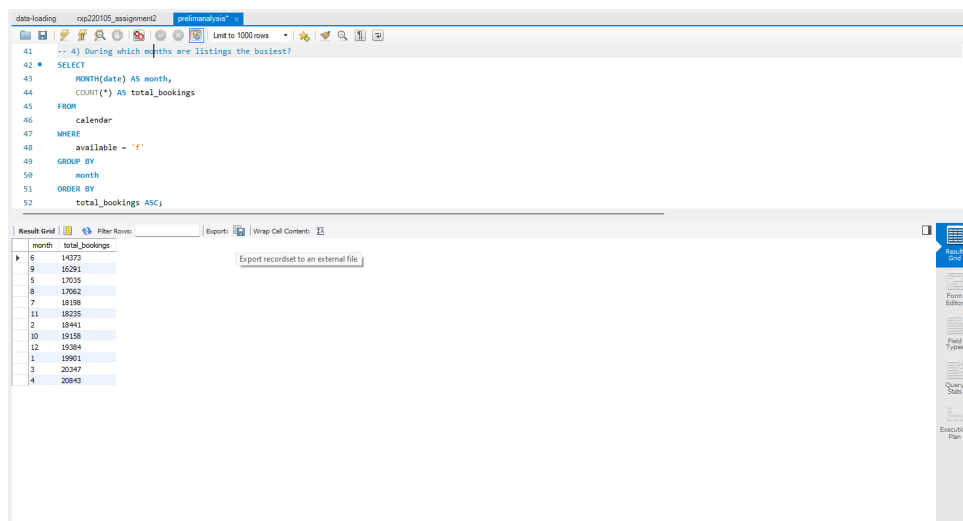
8.4 During which months are listings the least busy?

8.4.1 Code:

```
-- 4) During which months are listings the busiest?
SELECT
    MONTH(date) AS month,
    COUNT(*) AS total_bookings
FROM
    calendar
WHERE
    available = 'f'
GROUP BY
    month
ORDER BY
    total_bookings ASC;
```

Listing 6: Least Busy Months

8.4.2 Output:

Figure 7: Least busy month

15

8.4.3 Business Analysis:

- June and September are the least busy months, while April and March are the busiest, suggesting a seasonal trend in listing demand.

- Winter months (December to February) also show lower booking activity compared to the spring and summer months.

- Understanding the seasonal variation in booking trends can assist in optimizing marketing efforts and resource allocation.

Actions:

- **Seasonal Marketing Strategies:** Plan targeted marketing campaigns and promotions during peak booking months to maximize visibility and revenue.

- **Resource Optimization:** Adjust staffing levels or resource allocation based on the anticipated demand during different seasons to optimize operations.

- **Dynamic Pricing:** Implement dynamic pricing strategies to capitalize on high-demand months while offering competitive rates during slower periods.
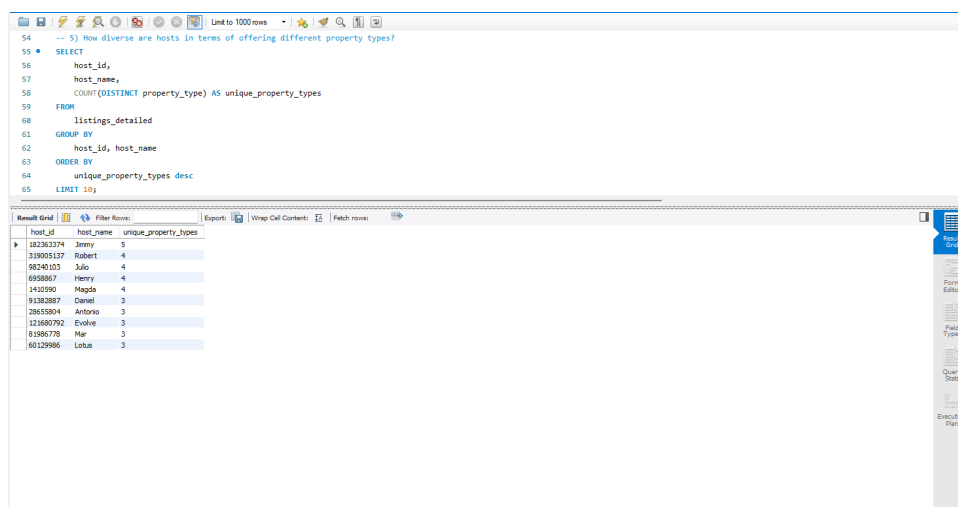
8.5 How diverse are hosts in terms of offering different property types?

8.5.1 Code:

```
-- 5) How diverse are hosts in terms of offering different property
   types?
SELECT
    host_id,
    host_name,
    COUNT(DISTINCT property_type) AS unique_property_types
FROM
    listings_detailed
GROUP BY
    host_id, host_name
ORDER BY
    unique_property_types desc
LIMIT 10;
```

Listing 7: Hosts Offering Diverse Property Types

8.5.2 Output:

Figure 8: Diversification of Hosts

16

8.5.3 Business Analysis:

Observations:

- The analysis reveals hosts with diverse portfolios, offering accommodations across multiple property types.

- Some hosts specialize in a particular property type, while others have a more varied selection, providing guests with different choices.

- Hosts with a higher number of unique property types contribute to the overall diversity of listings on the platform.

Actions:

- **Highlight Host Diversity:** Showcase hosts with a diverse range of property types to appeal to users looking for varied accommodation experiences.

- **Encourage Specialization:** Provide support for hosts interested in expanding their offerings or specializing in specific property types based on market demand.

- **Enhance Search Filters:** Improve platform features that allow users to filter listings based on hosts offering a variety of property types.
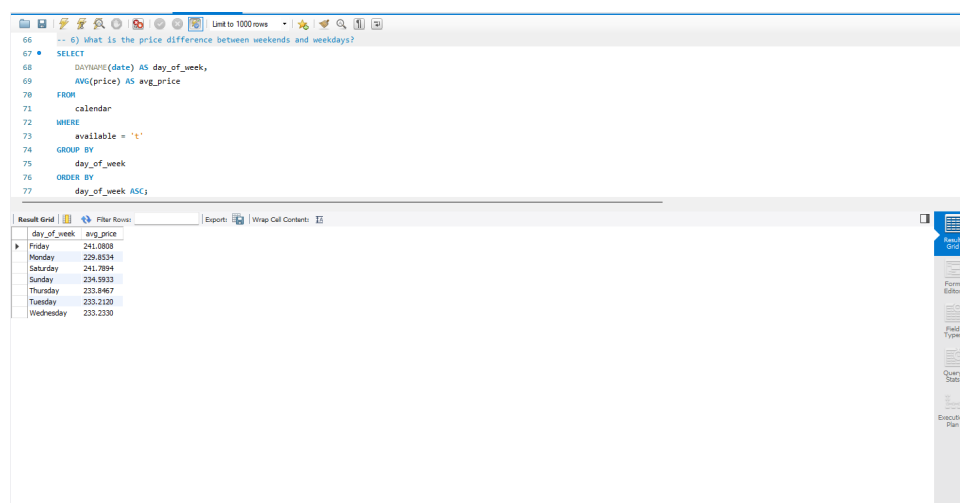
8.6 What is the price difference between weekends and weekdays?

8.6.1 Code:

```sql
-- 6) What is the price difference between weekends and weekdays?
SELECT
    DAYNAME(date) AS day_of_week,
    AVG(price) AS avg_price
FROM
    calendar
WHERE
    available = 't'
GROUP BY
    day_of_week
ORDER BY
    day_of_week ASC;
```
<div align="center">Listing 8: Price Difference Between Weekends and Weekdays</div>

8.6.2 Output:

<div align="center">Figure 9: Weekends vs Weekdays</div>

8.6.3 Business Analysis:

Observations:

- Fridays and Saturdays show slightly higher average prices compared to weekdays, indicating a potential weekend pricing premium.

- Mondays consistently have slightly lower average prices, suggesting a trend of reduced demand at the beginning of the week.

- Understanding the pricing dynamics between weekdays and weekends can inform strategies for attracting guests during different times.

  Actions:

- **Weekend Specials:** Encourage hosts to offer weekend specials or promotions to attract guests seeking weekend getaways.

- **Dynamic Pricing Adjustments:** Implement dynamic pricing algorithms that consider the day of the week to optimize revenue and occupancy.

- **Marketing Emphasis:** Highlight weekday affordability or weekend luxury options in marketing materials to target specific guest preferences.

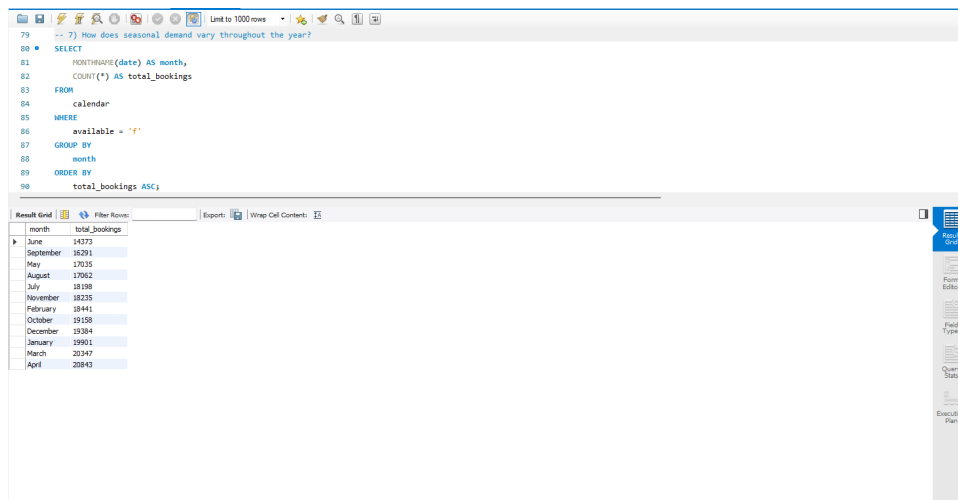## 8.7 How does seasonal demand vary throughout the year?

8.7.1 Code:

```sql
-- 7) How does seasonal demand vary throughout the year?
SELECT
    MONTHNAME(date) AS month,
    COUNT(*) AS total_bookings
FROM
    calendar
WHERE
    available = 'f'
GROUP BY
    month
ORDER BY
    total_bookings ASC;
```
Listing 9: Seasonal Demand Variation Throughout the Year

8.7.2 Output:

Figure 10: Seasonal demand throughout the year

18

8.7.3 Business Analysis:

- Seasonal demand for listings fluctuates throughout the year, with variations in the number of total bookings.

- June and September experience lower booking activity, while April and March are peak months with the highest total bookings.

- Understanding seasonal demand patterns is crucial for planning marketing strategies and optimizing resources.

Actions:

- **Seasonal Promotions:** Plan targeted promotions and discounts during peak months to attract more bookings.

- **Resource Allocation:** Adjust staffing levels and resources based on expected demand during different seasons.

- **Data-Driven Marketing:** Utilize insights from seasonal demand to create data-driven marketing campaigns and maximize visibility.
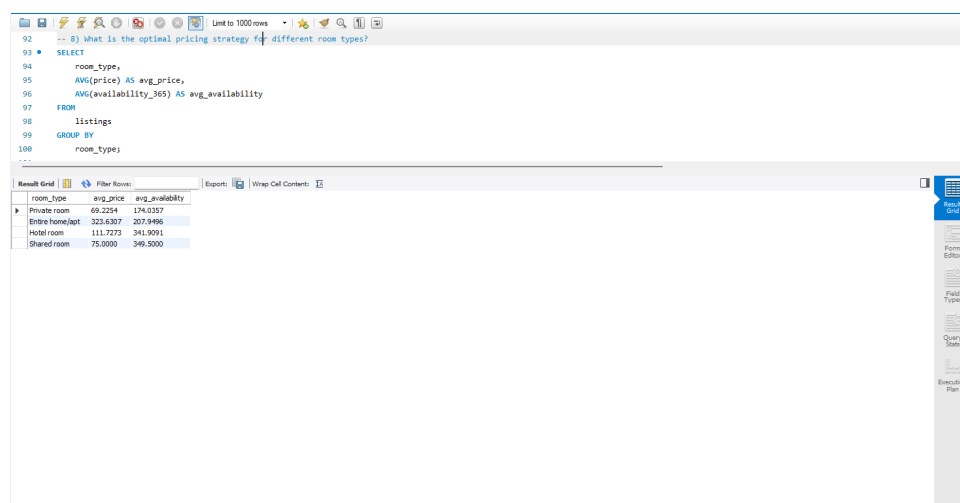
8.8 What is the optimal pricing strategy for different room types?

8.8.1 Code:

```
-- 8) What is the optimal pricing strategy for different room types?
SELECT
    room_type,
    AVG(price) AS avg_price,
    AVG(availability_365) AS avg_availability
FROM
    listings
GROUP BY
    room_type;
```

Listing 10: Optimal Pricing Strategy for Different Room Types

8.8.2 Output:

Figure 11: Pricing Strategy data for different room types

8.8.3 Business Analysis:

Observations:

- The analysis provides average prices and availability for different room types, offering insights into the pricing landscape.

- Entire homes/apartments have the highest average price, while private rooms and shared rooms have lower average prices.

- Understanding the optimal pricing strategy for each room type is essential for attracting guests and maximizing revenue.

Actions:

- **Competitive Pricing:** Ensure that prices for private rooms and shared rooms remain competitive to attract budget-conscious travelers.

- **Value-Added Services:** Highlight additional amenities or services for entire homes/apartments to justify higher pricing.

- **Dynamic Pricing Algorithms:** Implement dynamic pricing algorithms to adjust prices based on demand, seasonality, and availability.
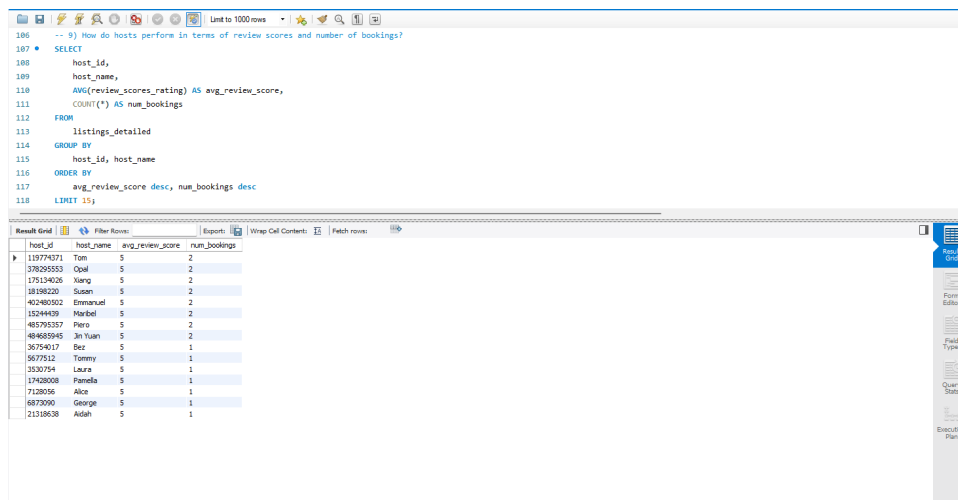
8.9 How do hosts perform in terms of review scores and number of bookings?

8.9.1 Code:

```sql
-- 9) How do hosts perform in terms of review scores and number of
   bookings?
SELECT
    host_id,
    host_name,
    AVG(review_scores_rating) AS avg_review_score,
    COUNT(*) AS num_bookings
FROM
    listings_detailed
GROUP BY
    host_id, host_name
ORDER BY
    avg_review_score desc, num_bookings desc
LIMIT 15;
```

Listing 11: Host Performance in Terms of Review Scores and Bookings

8.9.2 Output:

Figure 12: Host Performance

8.9.3 Business Analysis:

Observations:

- The analysis ranks hosts based on their average review scores and the number of bookings, highlighting top-performing hosts.

- Hosts like Tom, Opal, and Xiang have perfect review scores and multiple bookings, showcasing consistent performance.

- Recognizing and rewarding top-performing hosts can contribute to overall guest satisfaction and platform reputation.

  Actions:

- **Incentive Programs:** Introduce special incentives or rewards for hosts with high review scores and a significant number of bookings.

- **Feature Top Hosts:** Showcase top-performing hosts on the platform to enhance their visibility and attract more guests.

- **Host Training:** Provide additional training or resources to hosts to maintain high standards of service and guest satisfaction.
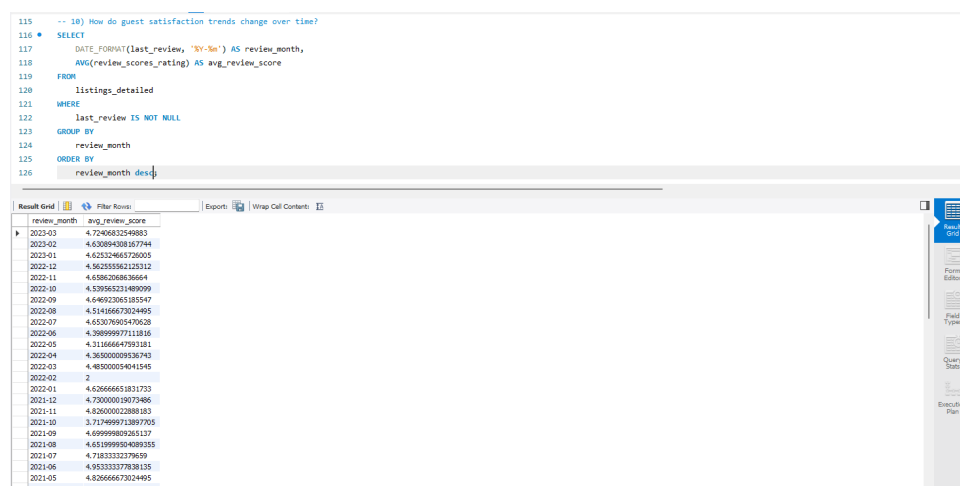
8.10 How do guest satisfaction trends change over time?

8.10.1 Code:

```sql
-- 10) How do guest satisfaction trends change over time?
SELECT
    DATE_FORMAT(last_review, '%Y-%m') AS review_month,
    AVG(review_scores_rating) AS avg_review_score
FROM
    listings_detailed
WHERE
    last_review IS NOT NULL
GROUP BY
    review_month
ORDER BY
    review_month desc;
```
<div align="center">Listing 12: Guest Satisfaction Trends Over Time</div>

8.10.2 Output:

<div align="center">Figure 13: Guest Satisfaction Trends</div>

8.10.3 Business Analysis:

- Guest satisfaction trends vary over time, with fluctuations in average review scores from month to month.

- Recent months (March, February) show relatively high average review scores, indicating positive guest experiences.

- Monitoring these trends helps identify areas for improvement and ensures continuous enhancement of guest satisfaction.

Actions:

- **Quality Improvement Initiatives:** Implement targeted improvements based on feedback to address any recurring issues.

- **Communication Channels:** Enhance communication channels with guests to gather feedback and address concerns promptly.

- **Celebrating Success:** Acknowledge and celebrate periods of high guest satisfaction, reinforcing positive host behavior.

9 Visualizations :

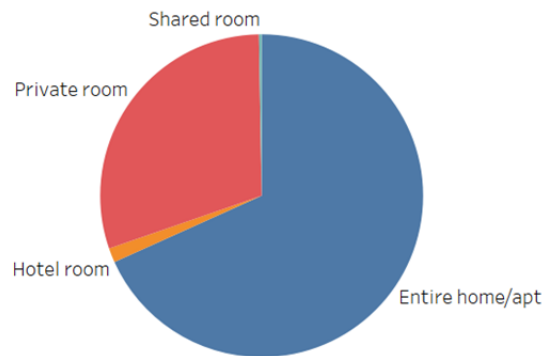## 9.1 Number of Rooms availability for 365 days by room type

Figure 14: Number of Rooms availability for 365 days by room type

The graph illustrates the annual availability of each room type over 365 days. The **entire home/apartment** type consistently exhibits the highest availability, followed by **private rooms**.

9.2 Average prices of Host neighbourhoods
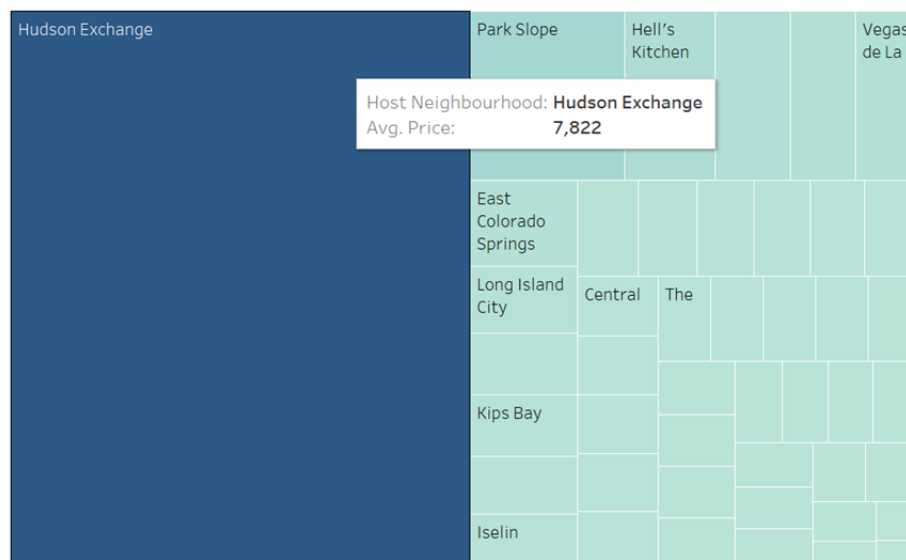

Figure 15: Average prices of Host neighbourhoods

The depicted graph showcases the average prices charged by hosts in different neighbourhoods. While this doesn't necessarily indicate the most competitive neighbourhoods, it reveals the pricing nature of each host's location. **Hudson Exchange** stands out with the highest average price of \$7822, suggesting a potentially premium service.

9.3 Reviews per month if host is a superhost vs if host has a profile pic
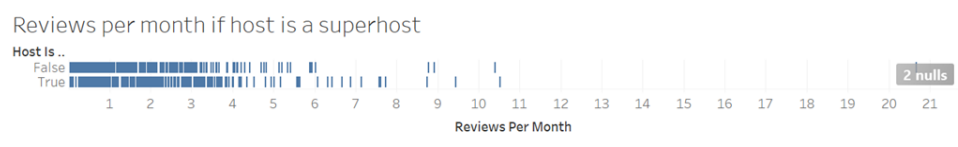

Figure 16: Host is a superhost

Figure 17: Host has a profile pic

This graph compares customer reviews per month for hosts with and without superhost status. Surprisingly, the data indicates that being a superhost doesn't significantly impact the number of reviews. However, hosts without profile pictures receive fewer reviews, emphasizing the importance of maintaining a visible and authentic presence on the Airbnb platform.

9.4 Number of people accommodated by each room type
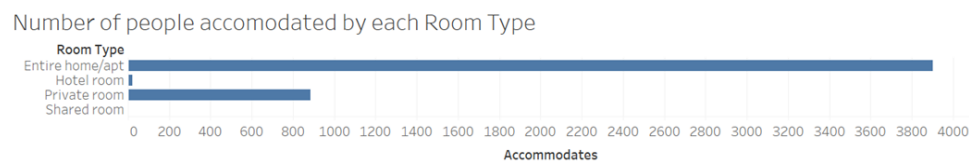
Figure 18: Number of people accommodated by each room type

The presented graph details the number of people accommodated by each room type. Notably, **entire home/apartment** types consistently accommodate the highest number of people, while **shared rooms** accommodate the least.

9.5 Number of Beds availability in 30 days by hosts

Figure 19: Number of Beds availability in 30 days by hosts

This graph provides insights into the number of beds available and the average number of beds provided by hosts. By hovering over data points, one can identify hosts with the most availability and those offering the highest bed count.
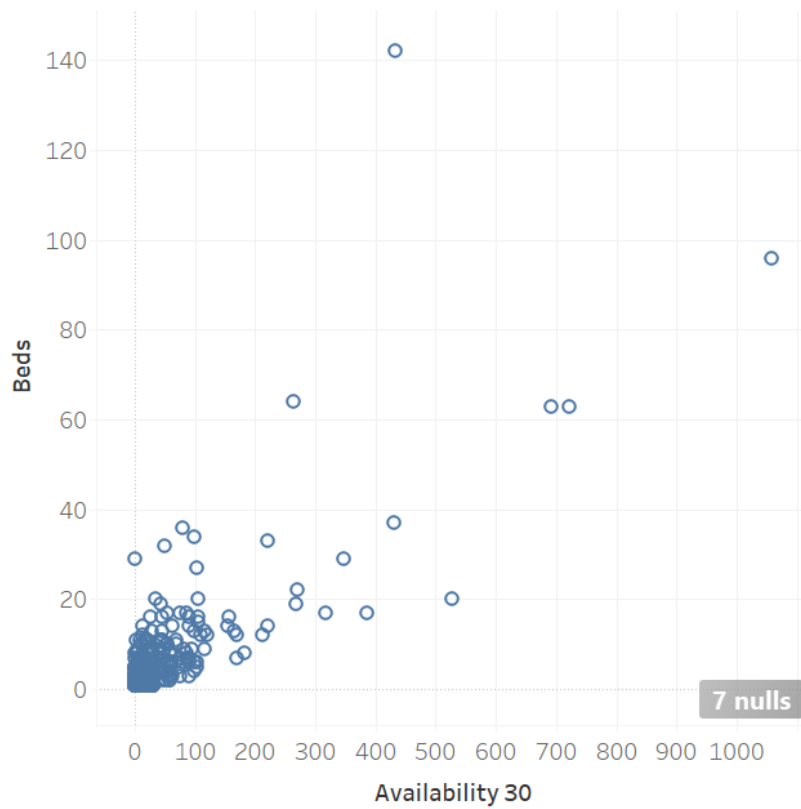
10 Conclusion

In this extensive analysis of Airbnb data, we delved into various dimensions of the platform, ranging from property types and host diversity to pricing strategies, seasonal demand, and customer satisfaction. Our investigation aimed to unearth valuable insights that can empower diverse stakeholders within Airbnb, enabling them to make informed decisions and drive the continued success of the platform.

10.1 Key Observations and Insights:

- **Property Type Dynamics:**

  - Identified hosts offering the most varied property types, with Jimmy leading the pack with an impressive count of 5 unique property types.
  - Explored the diversity among hosts, highlighting opportunities for further property type distribution.

- **Market Demand and Pricing:**

  - Uncovered the busiest months for listings, shedding light on seasonal demand patterns.
  - Explored optimal pricing strategies for different room types, aiding pricing teams in their decision-making process.

- **Host Performance and Guest Satisfaction:**

  - Assessed hosts' performance based on review scores and the number of bookings.
  - Tracked guest satisfaction trends over time, providing insights into evolving guest expectations.

- **Premium Features and Innovation:**

  - Investigated data relevant to the rollout of premium features, catering to product development and innovation teams.
  - Analyzed the impact of premium features on pricing and listings, aligning product offerings with customer preferences.

- **Operational Efficiency:**

  - Explored the operational side, including property type distribution and diversification, benefiting operations and pricing teams.
  - Unveiled opportunities for property managers to diversify property types and enhance the variety of listed properties.

- **Marketing Strategies:**

  - Provided marketing teams with insights into effective seasonal campaigns and leveraging booking trends to optimize strategies.

10.2 Actions and Recommendations:

- **Strategic Decision-Making:**

  - Executive leadership can leverage insights for strategic decision-making, understanding the holistic impact of various choices on business success.

- **Marketing Optimization:**

  - Marketing teams can refine their strategies, tailoring campaigns based on seasonal demand and booking trends.

- **Operations and Pricing Optimization:**

  - Operations and pricing teams can optimize property type distribution and pricing strategies for enhanced efficiency.

- **Host Engagement:**

  - Host partnerships and incentive teams can identify and reward high-performing hosts, fostering positive partnerships.

- **Customer Experience Enhancement:**

  - Customer experience teams can use guest satisfaction trends to address potential issues and enhance overall guest satisfaction.

- **Innovation and Product Development:**

  - Product development and innovation teams can align premium feature rollouts with customer preferences, driving innovation on the platform.

- **Financial and Investment Decisions:**

  - Financial analysts can evaluate the financial impact of pricing strategies and diversification, aiding in informed investment decisions.

10.3 Overall Impact:

This business analysis serves as a valuable resource for stakeholders across various domains within Airbnb. By translating raw data into actionable insights, we aim to empower decision-makers to navigate the complexities of the short-term rental market successfully. As Airbnb continues to evolve, this project stands as a testament to the power of data-driven decision-making in shaping the future of the platform.