**Exploring Three Cities & Their Hotel Rankings**

Akash Navneeth, David Diaz, Jayni Desai, & Spencer Wu

Naveen Jindal School of Management, The University of Texas at Dallas

ITSS 4300: Database Fundamentals

Professor Thiru Pandian

# Content of the Dataset

**Link to the Dataset (Kaggle):**

https://www.kaggle.com/datasets/thedevastator/sentiment-analyses-of-city-hotels

Booking.com is a travel booking website that allows its page visitors to book hotel rooms, flights, car rentals, and so much more all in one convenient place. The specific data that the team explored has five different comma-separated values (CSV) files called *hotel*, *location*, *rating_info*, *amenities*, and *quality*. Each of these tables contains data about three cities – New York City, Los Angeles, & Orlando – and how each of these cities' hotels have been rated based on consumer or visitor feedback that was posted on the Booking.com website and was compiled onto Kaggle, then the team did some data cleaning to ensure that we had only the relevant data necessary from Kaggle.

Descriptions of each of the tables highlight what data each table contains to further explain how the team came to examine this dataset:

- The hotel.csv file contains data about the hotel names and the number of reviews each of those hotels received while also including a hotel ID number for easy identification.
- The location.csv file contains data of which city, state, and zip code each of the hotels are located in.
- The rating_info.csv file contains data of the average rating (on a scale of 1 to 10) of how consumers felt about certain aspects of the hotel that they stayed in.
- The amenities.csv file contains data of what additional amenities are available for visitors of each hotel.
- The quality.csv file contains data of the relative sentiment (as average) of how consumers felt regarding their hotel experience, either positive or negative.

# Business/Project Objective

The exploration of the ratings of hotels in New York City, Los Angeles, and Orlando is to help Booking.com determine which hotels most people tend to visit as well as what amenities or hotel characteristics people look for (through number of visits to a particular hotel). These findings could help Booking.com find ways to attract more consumer/hotel visitors to certain hotels that do not garner enough traction by figuring out what kinds of offers the travel company could offer to increase the hotel visits.

Also based on this data from Booking.com and Kaggle, the team can figure out how much of an influence the rating of hotels really has on future hotel visitors. This can be accomplished by examining the overall quality rating that these hotels received based on the reviews.

# Project Insights

These are the questions that the team has come across to help Booking.com take advantage of the data to potentially increase website traffic by highlighting certain hotels with offers to potential customers/hotel visitors. To ensure that this is possible, below there are ten insights that could improve the Booking.com website's goals of increasing bookings for certain hotels.

The ten insights that the team is looking into is as follows:

1. What hotels are the 50 most reviewed on the platform?
2. Is there a particular city that has more hotels in the 50 most reviewed hotels?
3. Is there a particular city that has more hotels in the 50 least reviewed hotels?
4. What is the total number of hotels in each city to have at least 2000 reviews on the platform?
5. What are the ratings for the hotel characteristics for the 10 most rated hotels?
6. What amenities are offered at the hotels that are in the 10 most rated hotels?

7. What are the ratings for the hotel characteristics for the 10 least rated hotels?
8. What amenities are offered at the hotels that are in the 10 least rated hotels?
9. Does the number of reviews influence the overall score for the 50 most rated hotels?
10. What hotels are in the top 50 when it comes to overall positive quality?

## Target Audience

Our target audience is going to be the management of Booking.com because the insights that the team will gather from this exploration of the dataset pertains to how the business could grow. The growth of Booking.com stems from its clients – the hotels – so getting to help the hotels grow will in turn increase growth for the Booking.com website.

## Problem Statement

The problem that is being faced at the moment is how hotel bookings are being influenced by ratings shared on Booking.com. What this report looks to accomplish is discover the hotel attributes and reviews that attract the most customers. This would allow Booking.com to market and emphasize certain hotels that offer these experiences and will in turn increase the consumer base of Booking.com users.

# Physical Model

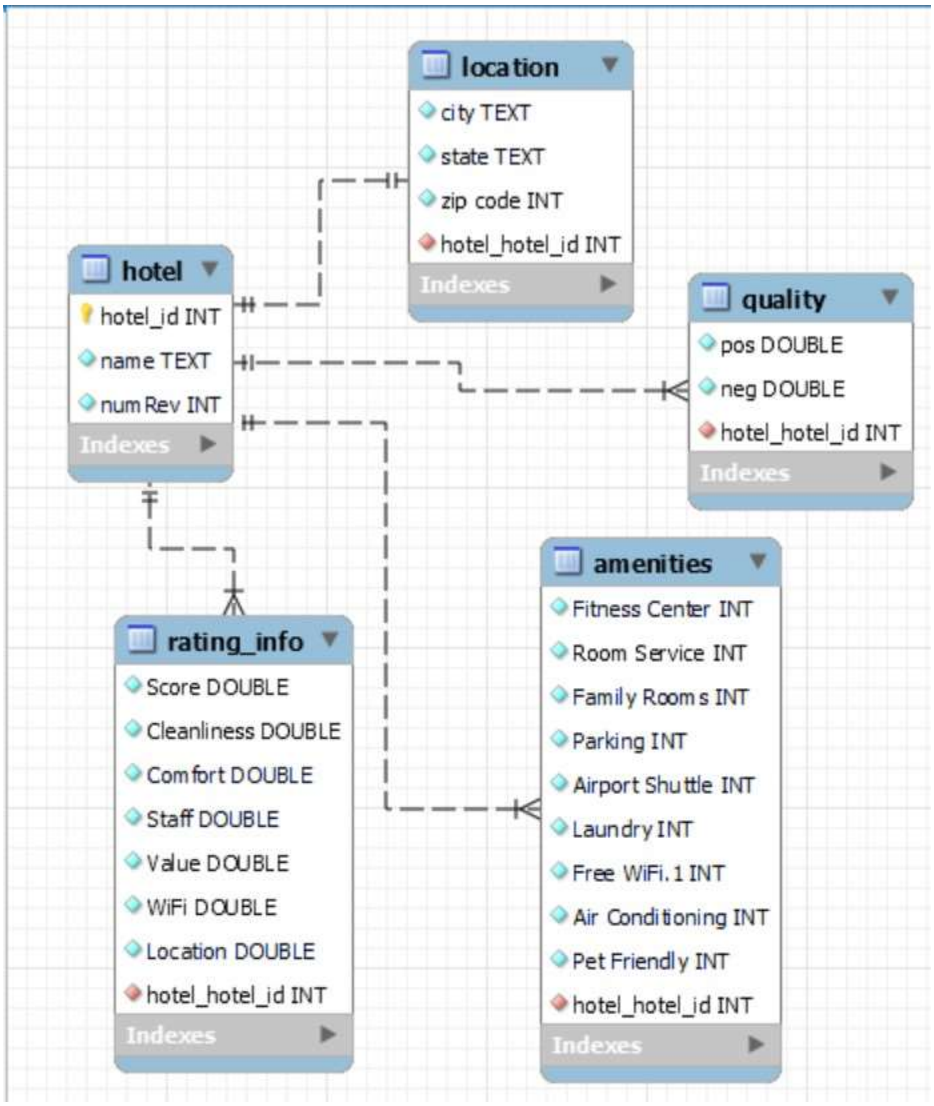
```
-- MySQL Workbench Forward Engineering


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_F
OR_DIVISION_BY_ZERO,NO_ENGINE_SUBSTITUTION';


-- -----------------------------------------------------

-- Schema mydb

-- -----------------------------------------------------


-- -----------------------------------------------------

-- Schema mydb

-- -----------------------------------------------------

CREATE SCHEMA IF NOT EXISTS `mydb` DEFAULT CHARACTER SET utf8mb3 ;

USE `mydb` ;


-- -----------------------------------------------------

-- Table `mydb`.`hotel`

-- -----------------------------------------------------

CREATE TABLE IF NOT EXISTS `mydb`.`hotel` (

`hotel_id` INT NOT NULL,

`name` TEXT NOT NULL,

`numRev` INT NOT NULL,

PRIMARY KEY (`hotel_id`))

ENGINE = InnoDB

DEFAULT CHARACTER SET = utf8mb3;
```

```
-- -----------------------------------------------------
-- Table `mydb`.`amenities`
-- -----------------------------------------------------
CREATE TABLE IF NOT EXISTS `mydb`.`amenities` (
`Fitness Center` INT NOT NULL,
`Room Service` INT NOT NULL,
`Family Rooms` INT NOT NULL,
`Parking` INT NOT NULL,
`Airport Shuttle` INT NOT NULL,
`Laundry` INT NOT NULL,
`Free WiFi.1` INT NOT NULL,
`Air Conditioning` INT NOT NULL,
`Pet Friendly` INT NOT NULL,
`hotel_hotel_id` INT NOT NULL,
INDEX `fk_amenities_hotel1_idx` (`hotel_hotel_id` ASC) VISIBLE,
CONSTRAINT `fk_amenities_hotel1`
FOREIGN KEY (`hotel_hotel_id`)
REFERENCES `mydb`.`hotel` (`hotel_id`)
ON DELETE NO ACTION
ON UPDATE NO ACTION)
ENGINE = InnoDB
DEFAULT CHARACTER SET = utf8mb3;


-- -----------------------------------------------------
-- Table `mydb`.`location`
-- -----------------------------------------------------
CREATE TABLE IF NOT EXISTS `mydb`.`location` (
`city` TEXT NOT NULL,
`state` TEXT NOT NULL,
`zip code` INT NOT NULL,
`hotel_hotel_id` INT NOT NULL,
```

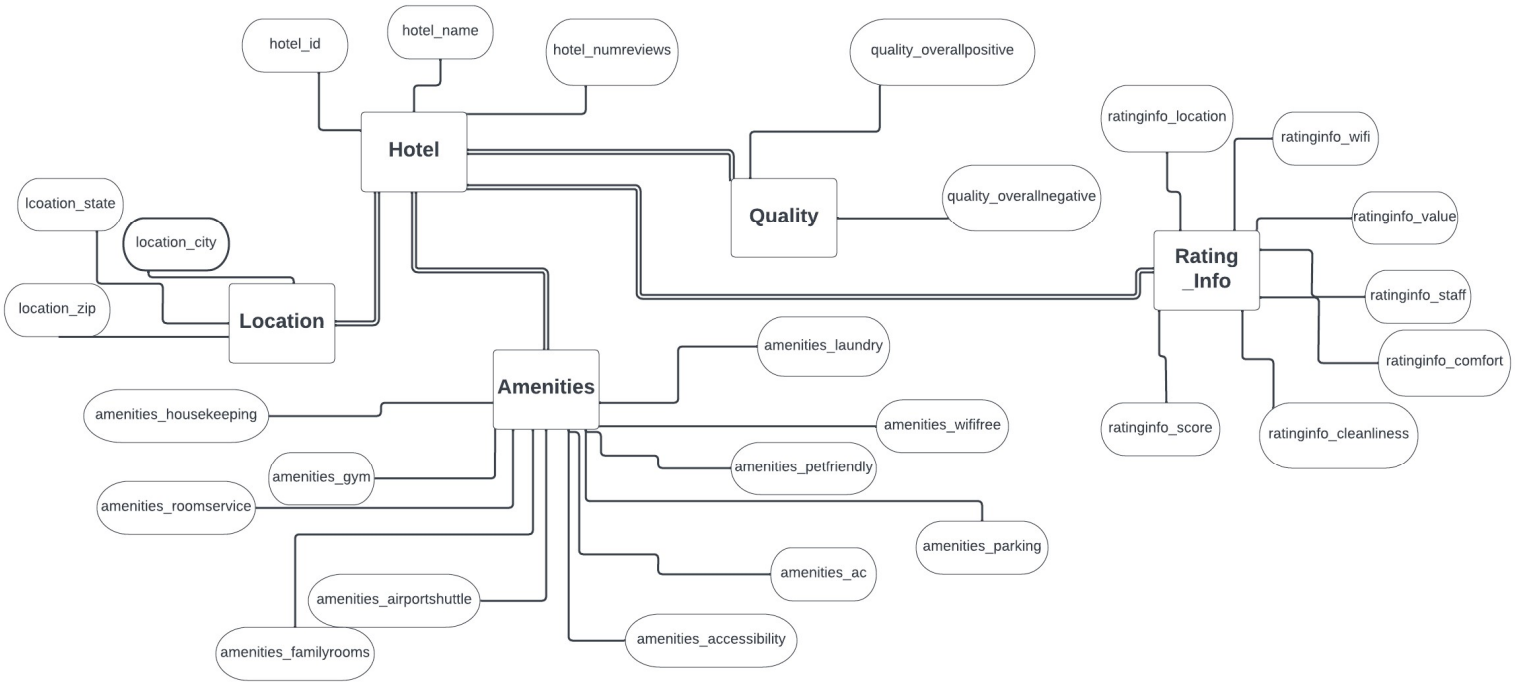
```
INDEX `fk_location_hotel_idx` (`hotel_hotel_id` ASC) VISIBLE,
CONSTRAINT `fk_location_hotel`
FOREIGN KEY (`hotel_hotel_id`)
REFERENCES `mydb`.`hotel` (`hotel_id`)
ON DELETE NO ACTION
ON UPDATE NO ACTION)
ENGINE = InnoDB
DEFAULT CHARACTER SET = utf8mb3;

-- -----------------------------------------------------
-- Table `mydb`.`quality`
-- -----------------------------------------------------
CREATE TABLE IF NOT EXISTS `mydb`.`quality` (
`pos` DOUBLE NOT NULL,
`neg` DOUBLE NOT NULL,
`hotel_hotel_id` INT NOT NULL,
INDEX `fk_quality_hotel1_idx` (`hotel_hotel_id` ASC) VISIBLE,
CONSTRAINT `fk_quality_hotel1`
FOREIGN KEY (`hotel_hotel_id`)
REFERENCES `mydb`.`hotel` (`hotel_id`)
ON DELETE NO ACTION
ON UPDATE NO ACTION)
ENGINE = InnoDB
DEFAULT CHARACTER SET = utf8mb3;

-- -----------------------------------------------------
-- Table `mydb`.`rating_info`
-- -----------------------------------------------------
CREATE TABLE IF NOT EXISTS `mydb`.`rating_info` (
`Score` DOUBLE NOT NULL,
`Cleanliness` DOUBLE NOT NULL,
```

```
`Comfort` DOUBLE NOT NULL,
`Staff` DOUBLE NOT NULL,
`Value` DOUBLE NOT NULL,
`WiFi` DOUBLE NOT NULL,
`Location` DOUBLE NOT NULL,
`hotel_hotel_id` INT NOT NULL,
INDEX `fk_rating_info_hotel1_idx` (`hotel_hotel_id` ASC) VISIBLE,
CONSTRAINT `fk_rating_info_hotel1`
FOREIGN KEY (`hotel_hotel_id`)
REFERENCES `mydb`.`hotel` (`hotel_id`)
ON DELETE NO ACTION
ON UPDATE NO ACTION)
ENGINE = InnoDB
DEFAULT CHARACTER SET = utf8mb3;

SET SQL_MODE=@OLD_SQL_MODE;
SET FOREIGN_KEY_CHECKS=@OLD_FOREIGN_KEY_CHECKS;
SET UNIQUE_CHECKS=@OLD_UNIQUE_CHECKS;
```

# Logical Model/ER Diagram

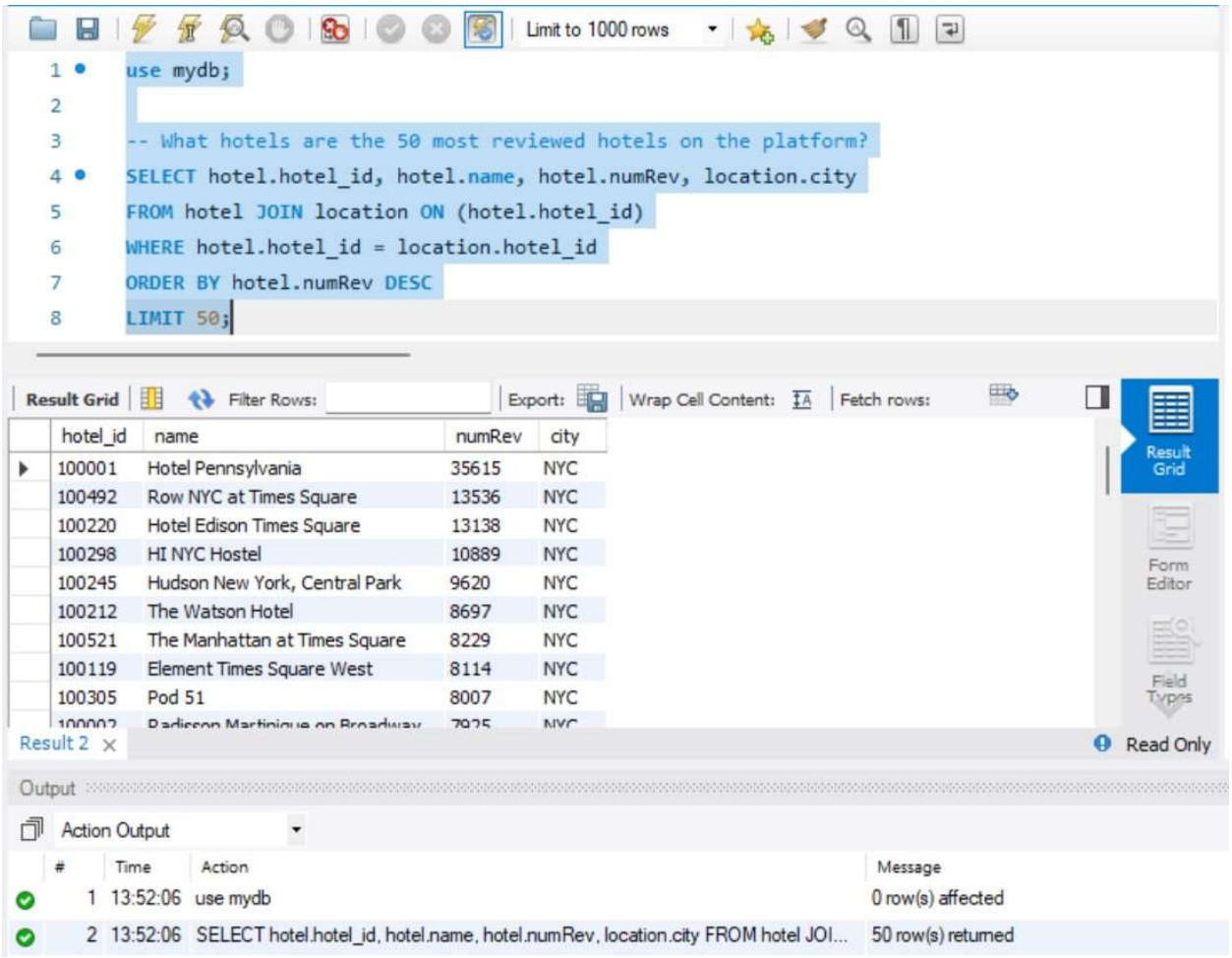


# Conceptual Model



# Project Insights Analysis

**Insight 1 – What hotels are the 50 most reviewed on the platform?**



```
use mydb;

-- What hotels are the 50 most reviewed hotels on the platform?
SELECT hotel.hotel_id, hotel.name, hotel.numRev, location.city
FROM hotel JOIN location ON (hotel.hotel_id)
WHERE hotel.hotel_id = location.hotel_id
ORDER BY hotel.numRev DESC
LIMIT 50;
```

Result Grid

| hotel_id | name | numRev | city |
|---|---|---|---|
| 100001 | Hotel Pennsylvania | 35615 | NYC |
| 100492 | Row NYC at Times Square | 13536 | NYC |
| 100220 | Hotel Edison Times Square | 13138 | NYC |
| 100298 | HI NYC Hostel | 10889 | NYC |
| 100245 | Hudson New York, Central Park | 9620 | NYC |
| 100212 | The Watson Hotel | 8697 | NYC |
| 100521 | The Manhattan at Times Square | 8229 | NYC |
| 100119 | Element Times Square West | 8114 | NYC |
| 100305 | Pod 51 | 8007 | NYC |

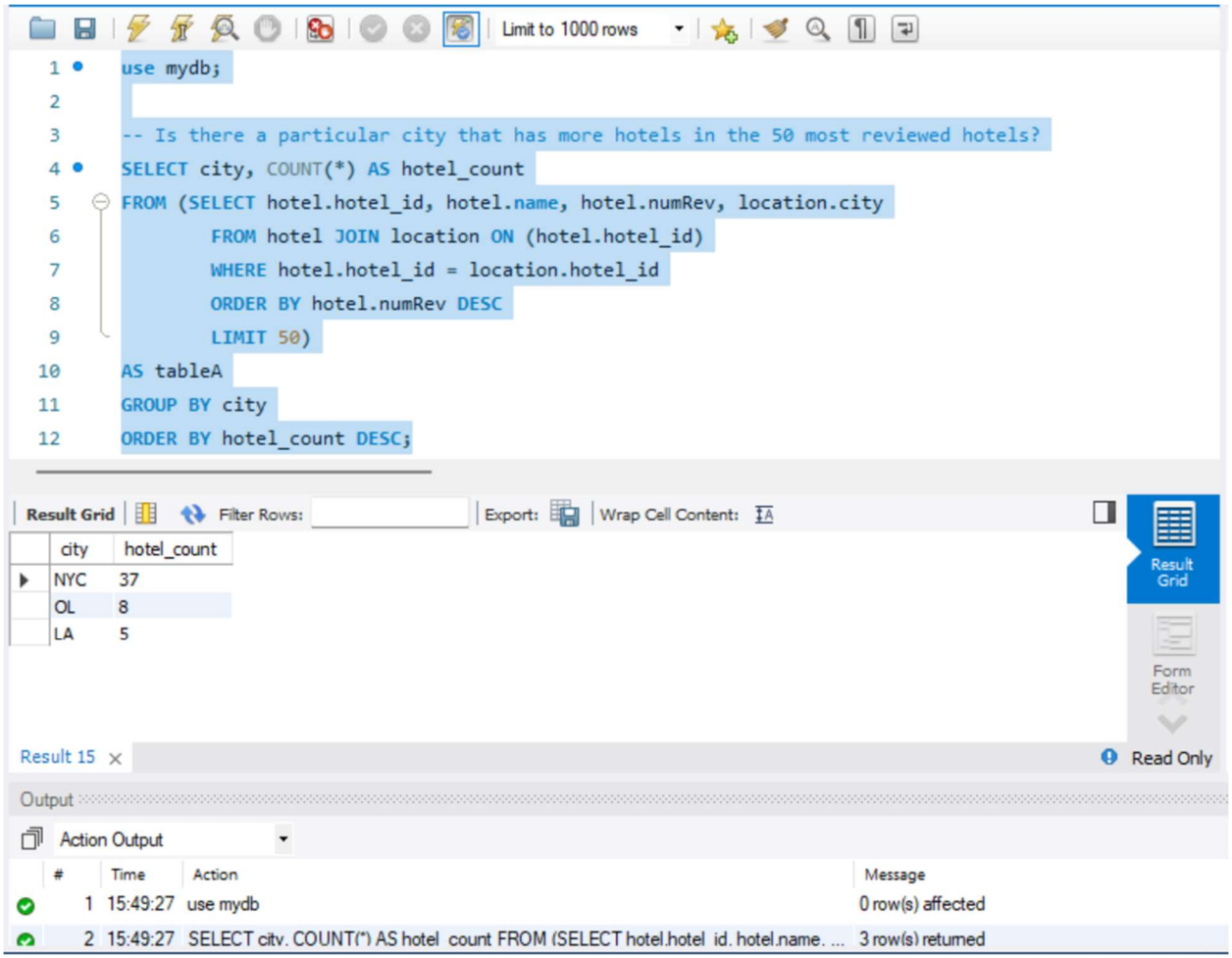
Output

Action Output

| # | Time | Action | Message |
|---|---|---|---|
| 1 | 13:52:06 | use mydb | 0 row(s) affected |
| 2 | 13:52:06 | SELECT hotel.hotel_id, hotel.name, hotel.numRev, location.city FROM hotel JOI... | 50 row(s) returned |

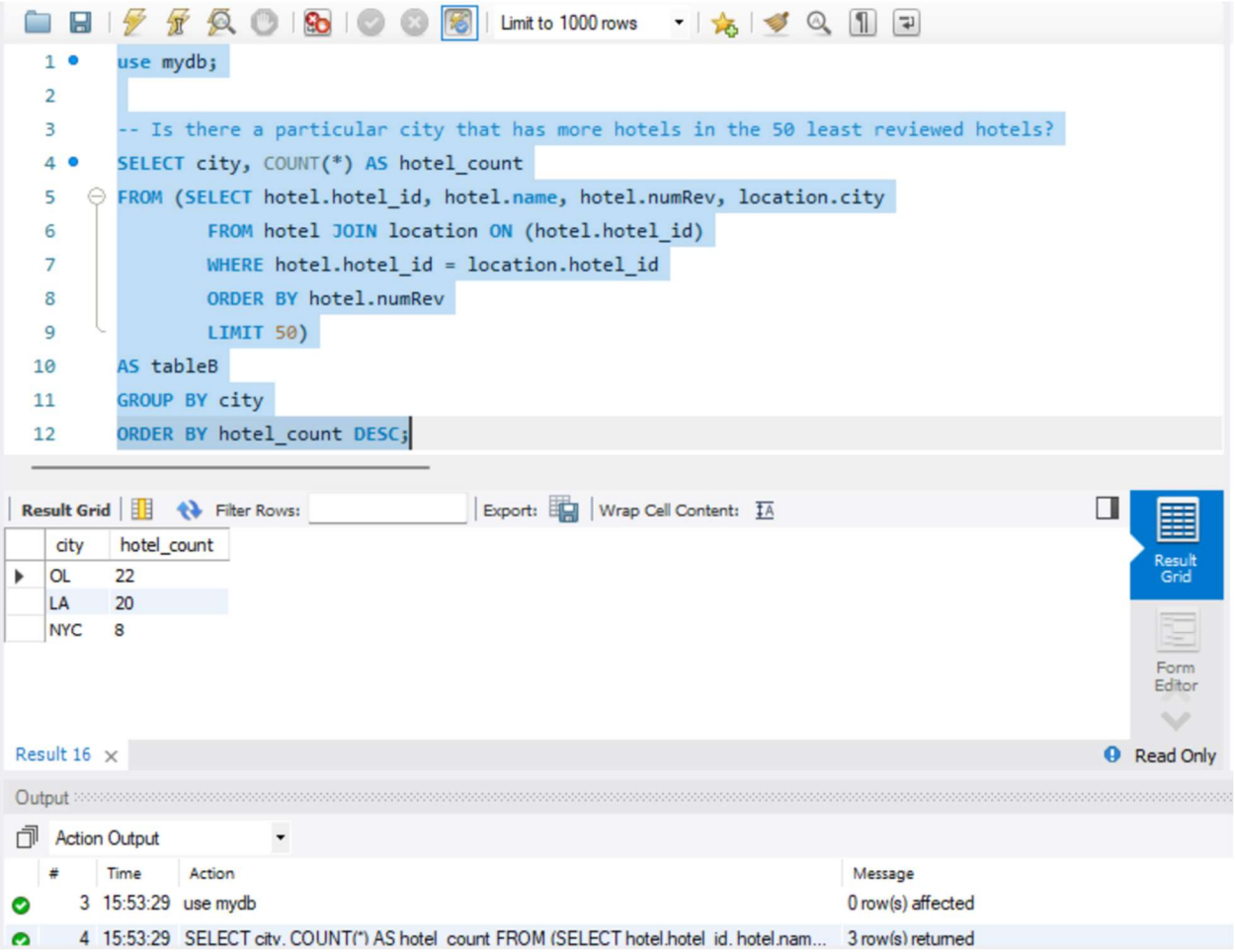
Analysis of Insight – From this insight, it looks like Hotel Pennsylvania is the most-rated hotel on the platform with an astounding 35,615 reviews. The next highest number of reviews is Row NYC at Times Square with 13,536 reviews. This indicates that Hotel Pennsylvania seems to be the most popular hotel in the dataset since there is a huge disparity in the number of reviews between the number of reviews between this hotel and the subsequent hotel in the list. From this initial insight, there is something to keep a look out for what makes Hotel Pennsylvania so popular for reviews. So, there will need to be some additional insights to find out what makes that hotel special.

**Insight 2 – Is there a particular city that has more hotels in the 50 most reviewed hotels?**




Analysis of Insight – Based on this insight, it looks like New York City, NY is dominating the list of the 50 most reviewed hotels on the platform, So, what this means for the Booking.com platform is that the management team needs to figure out ways to promote the hotels in the other two cities – Los Angeles, CA and Orlando, FL – because those hotels are not getting as many reviews maybe due to lower density of visitors or travelers to those destinations. But to be sure it's just the density of the population of the visitors and not other factors that the hotels are able to influence, this is going to require more insights to understand this disparity.

**Insight 3 – Is there a particular city that has more hotels in the 50 least reviewed hotels?**




Analysis of Insight – Based on this insight, it looks like Orlando, FL and Los Angeles are almost equally dominant in the list of the 50 least reviewed hotels on the platform. What this means for the Booking.com platform is that the management team needs to figure out ways to promote the hotels in the other two cities – Los Angeles, CA and Orlando, FL – because those hotels are not getting as many reviews but there is still something unclear about why these two cities are suffering from lower number of reviews for the hotels – especially when each of those hotels is getting less than ten total reviews on the platform.

## Insight 4 – What is the total number of hotels in each city to have at least 2000 reviews on the platform?

```
use mydb;

-- What is the total number of hotels in each city to have at least 2000 reivews on the platform?
SELECT city, COUNT(*) AS total_hotels
FROM location JOIN hotel ON (hotel.hotel_id)
WHERE hotel.hotel_id = location.hotel_id AND hotel.numRev >= 2000
GROUP BY location.city
ORDER BY total_hotels DESC;
```

Result Grid | Filter Rows: | Export: | Wrap Cell Content:

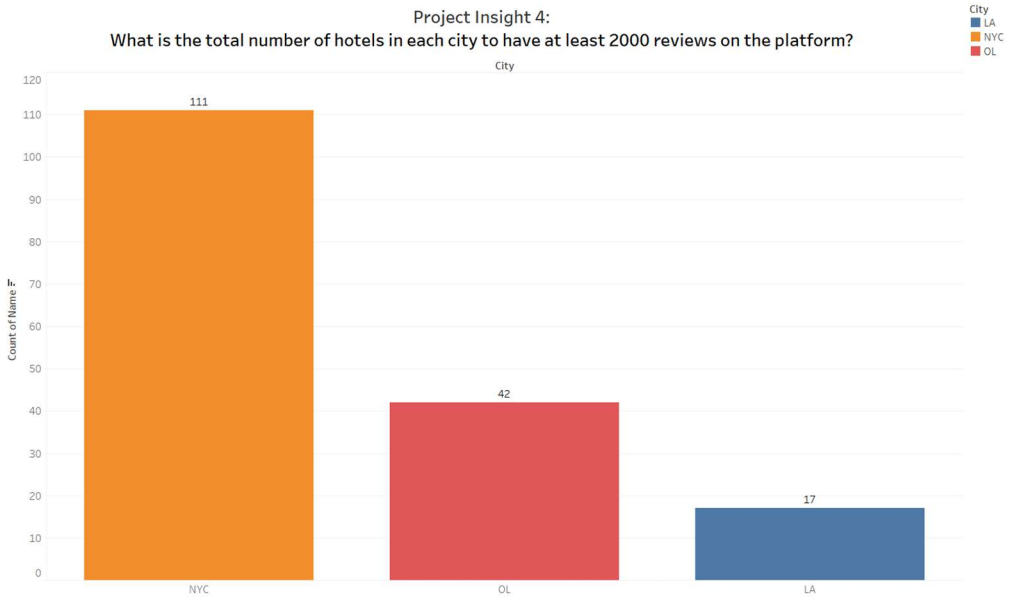
| city | total_hotels |
|---|---|
| NYC | 111 |
| OL | 42 |
| LA | 17 |

Result 18 | Read Only

Output

Action Output

| # | Time | Action | Message |
|---|---|---|---|
| 3 | 16:08:32 | use mydb | 0 row(s) affected |
| 4 | 16:08:32 | SELECT city, COUNT(*) AS total_hotels FROM location JOIN hotel ON (hotel.hotel... | 3 row(s) returned |



Analysis of Insight – Based on this insight, it looks like New York City dominates the list for the most hotels with more than 2000 reviews for each hotel. This is of no surprise since when the team examined the 50 most-rated hotels, NYC was dominating that list as well. Moreover, from this insight, it indicated that there is something that hotels in NYC are doing that allows them to get more reviews overall compared to hotels in the other cities of Los Angeles and Orlando.

## Insight 5 – What are the ratings for the hotel characteristics for the 10 most rated hotels?

```
use mydb;

-- What are the ratings for the hotel characteristics for the 10 most rated hotels?
SELECT hotel.name, hotel.numRev, rating_info.Score, rating_info.Cleanliness,
rating_info.Comfort, rating_info.Staff, rating_info.Value, rating_info.WiFi,
rating_info.Location
FROM  rating_info JOIN hotel ON (hotel.hotel_id)
WHERE hotel.hotel_id = rating_info.hotel_id
ORDER BY hotel.numRev DESC
LIMIT 10;
```

Result Grid

| name | numRev | Score | Cleanliness | Comfort | Staff | Value | WiFi | Location |
|---|---|---|---|---|---|---|---|---|
| Hotel Pennsylvania | 35615 | 5.6 | 5 | 4.9 | 5.6 | 5.2 | 6.3 | 8.3 |
| Row NYC at Times Square | 13536 | 7.1 | 7 | 6.8 | 7.1 | 6.4 | 7.6 | 8.8 |
| Hotel Edison Times Square | 13138 | 7.5 | 7.7 | 7.2 | 7.7 | 6.8 | 4.9 | 9.2 |
| HI NYC Hostel | 10889 | 8.5 | 8.6 | 8.1 | 8.7 | 8.2 | 8.9 | 8.5 |
| Hudson New York, Central Park | 9620 | 7.2 | 7.4 | 6.5 | 7.5 | 6.3 | 7.8 | 8.7 |
| The Watson Hotel | 8697 | 7.8 | 7.8 | 7.7 | 7.8 | 7.5 | 7.7 | 8.5 |
| The Manhattan at Times Square | 8229 | 6.9 | 6.4 | 6.6 | 7.3 | 6.1 | 8.9 | 9 |

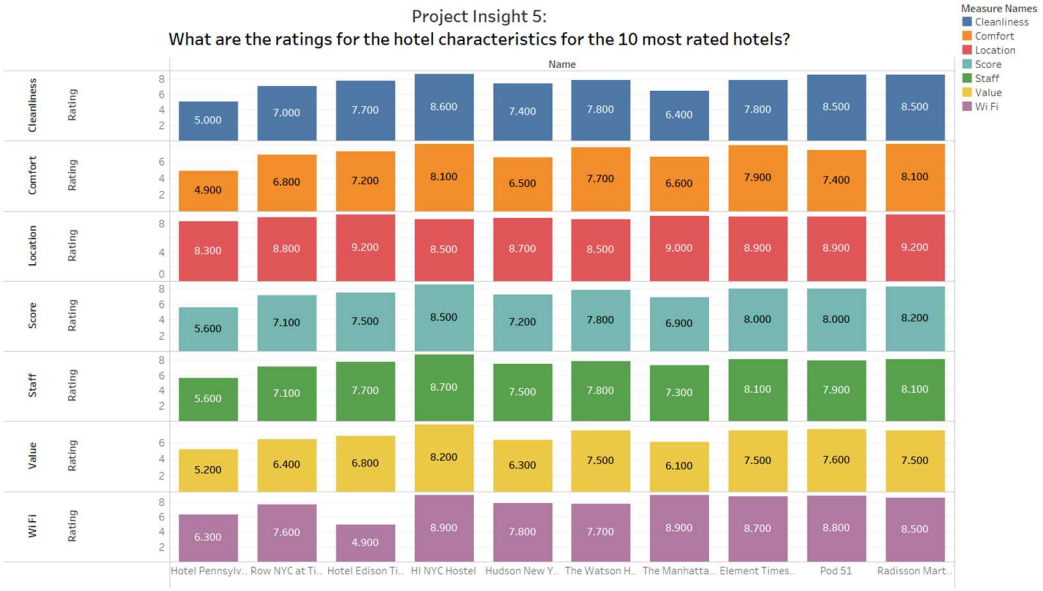
Result 10

Read Only

Output

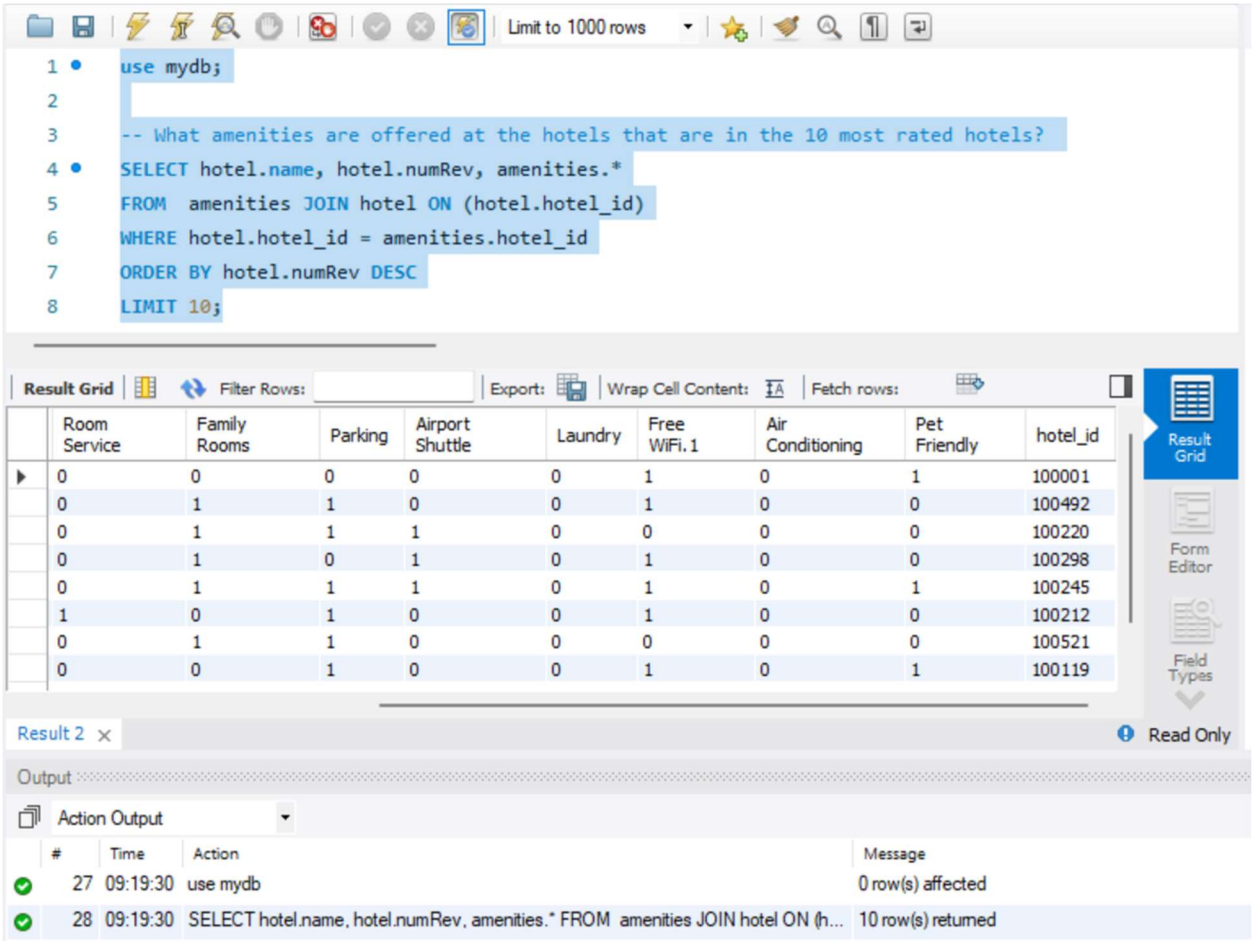
Action Output

| # | Time | Action | Message |
|---|---|---|---|
| 25 | 09:15:50 | use mydb | 0 row(s) affected |
| 26 | 09:15:50 | SELECT hotel.name, hotel.numRev, rating_info.Score, rating_info.Cleanliness, ratin... | 10 row(s) returned |



Analysis of Insight – Based on this insight, it looks like a majority (9 out of the 10 hotels listed) have ratings for each of the hotel rating criteria  of at least 6.0. However, it looks like Hotel Pennsylvania (the first hotel on the list with the most reviews), has very poor ratings but has the most reviews overall, which indicates that more reviews a hotel has, it's likely to have lower ratings. So, based on this insight, it is clear that the total number of reviews a hotel has on the Booking.com platform does not indicate that the hotel is of good quality – but it does indicate that the hotel has a lot of visitors who stay in that hotel.

**Insight 6 – What amenities are offered at the hotels that are in the 10 most rated hotels?**



```
use mydb;

-- What amenities are offered at the hotels that are in the 10 most rated hotels?
SELECT hotel.name, hotel.numRev, amenities.*
FROM  amenities JOIN hotel ON (hotel.hotel_id)
WHERE hotel.hotel_id = amenities.hotel_id
ORDER BY hotel.numRev DESC
LIMIT 10;
```

Result Grid

| Room Service | Family Rooms | Parking | Airport Shuttle | Laundry | Free WiFi. 1 | Air Conditioning | Pet Friendly | hotel_id |
|---|---|---|---|---|---|---|---|---|
| 0 | 0 | 0 | 0 | 0 | 1 | 0 | 1 | 100001 |
| 0 | 1 | 1 | 0 | 0 | 1 | 0 | 0 | 100492 |
| 0 | 1 | 1 | 1 | 0 | 0 | 0 | 0 | 100220 |
| 0 | 1 | 0 | 1 | 0 | 1 | 0 | 0 | 100298 |
| 0 | 1 | 1 | 1 | 0 | 1 | 0 | 1 | 100245 |
| 1 | 0 | 1 | 0 | 0 | 1 | 0 | 0 | 100212 |
| 0 | 1 | 1 | 0 | 0 | 0 | 0 | 0 | 100521 |
| 0 | 0 | 1 | 0 | 0 | 1 | 0 | 1 | 100119 |

Result 2

Output

Action Output

| # | Time | Action | Message |
|---|---|---|---|
| 27 | 09:19:30 | use mydb | 0 row(s) affected |
| 28 | 09:19:30 | SELECT hotel.name, hotel.numRev, amenities.* FROM amenities JOIN hotel ON (h... | 10 row(s) returned |

Analysis of Insight – Based on this insight, it looks like the most common amenities offered by the 10 most-rated hotels, the fitness center, free parking, and free Wi-Fi. The other amenities are sparsely offered between all of those 10 hotels. This indicates that out of the amenities, people look for fitness centers, free parking spaces at hotels, and free Wi-Fi services, so that they can enjoy their stay – this is because these services are so common in the 10 most-rated hotels. So, Booking.com needs to work with other hotels to offer these common amenities to draw more customers to other hotels that may not be doing so well.

## Insight 7 – What are the ratings for the hotel characteristics for the 10 least rated hotels?

```
-- What are the ratings for the hotel characteristics for the 10 least rated hotels?
SELECT hotel.name, hotel.numRev, rating_info.Score, rating_info.Cleanliness,
rating_info.Comfort, rating_info.Staff, rating_info.Value, rating_info.WiFi,
rating_info.Location
FROM  rating_info JOIN hotel ON (hotel.hotel_id)
WHERE hotel.hotel_id = rating_info.hotel_id
ORDER BY hotel.numRev
LIMIT 10;
```

Result Grid | Filter Rows: | Export: | Wrap Cell Content: | Fetch rows:

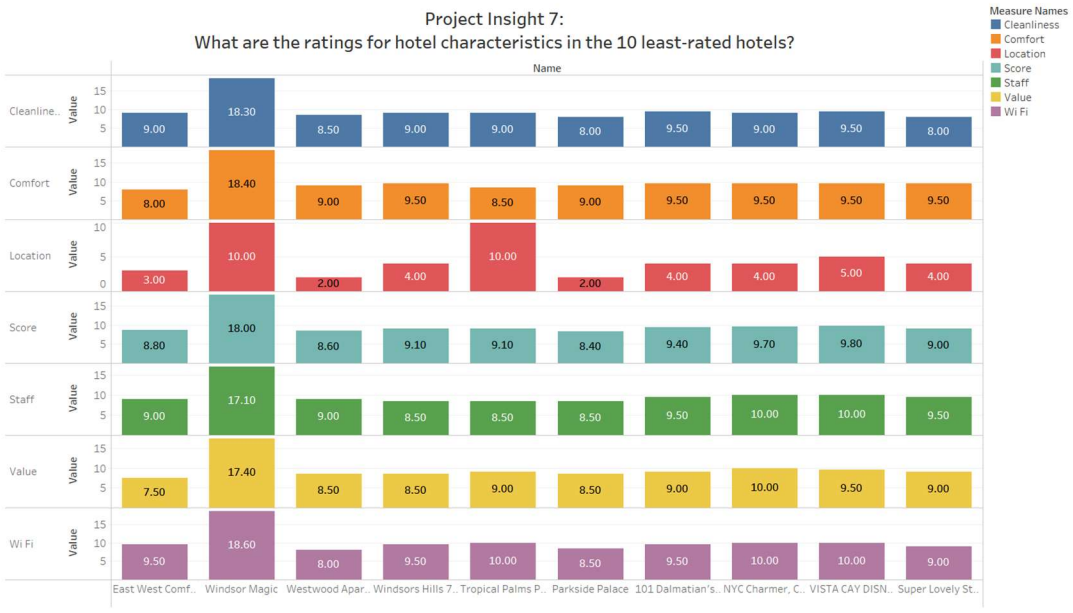
| name | numRev | Score | Cleanliness | Comfort | Staff | Value | WiFi | Location |
|---|---|---|---|---|---|---|---|---|
| East West Comfort: Downtown LA Inspiring Apa... | 5 | 8.8 | 9 | 8 | 9 | 7.5 | 9.5 | 3 |
| Windsor Magic | 5 | 8.9 | 9 | 9.5 | 8 | 8.5 | 9.5 | 2 |
| Westwood Apartment 409W | 5 | 8.6 | 8.5 | 9 | 9 | 8.5 | 8 | 2 |
| Windsors Hills 7749 | 5 | 9.1 | 9 | 9.5 | 8.5 | 8.5 | 9.5 | 4 |
| Tropical Palms Premium Loft Cottage 33 | 5 | 9.1 | 9 | 8.5 | 8.5 | 9 | 10 | 10 |
| Parkside Palace | 5 | 8.4 | 8 | 9 | 8.5 | 8.5 | 8.5 | 2 |
| 101 Dalmatian's Den | 5 | 9.4 | 9.5 | 9.5 | 9.5 | 9 | 9.5 | 4 |
| NYC Charmer, Central Park/Lincoln Center | 5 | 9.7 | 9 | 9.5 | 10 | 10 | 10 | 4 |
| VISTA CAY DISNEY PARADISE 6 | 5 | 9.8 | 9.5 | 9.5 | 10 | 9.5 | 10 | 5 |
| Super Lovely Studio Apartment | 5 | 9 | 8 | 9.5 | 9.5 | 9 | 9 | 4 |

Result 11 | Read Only

Output

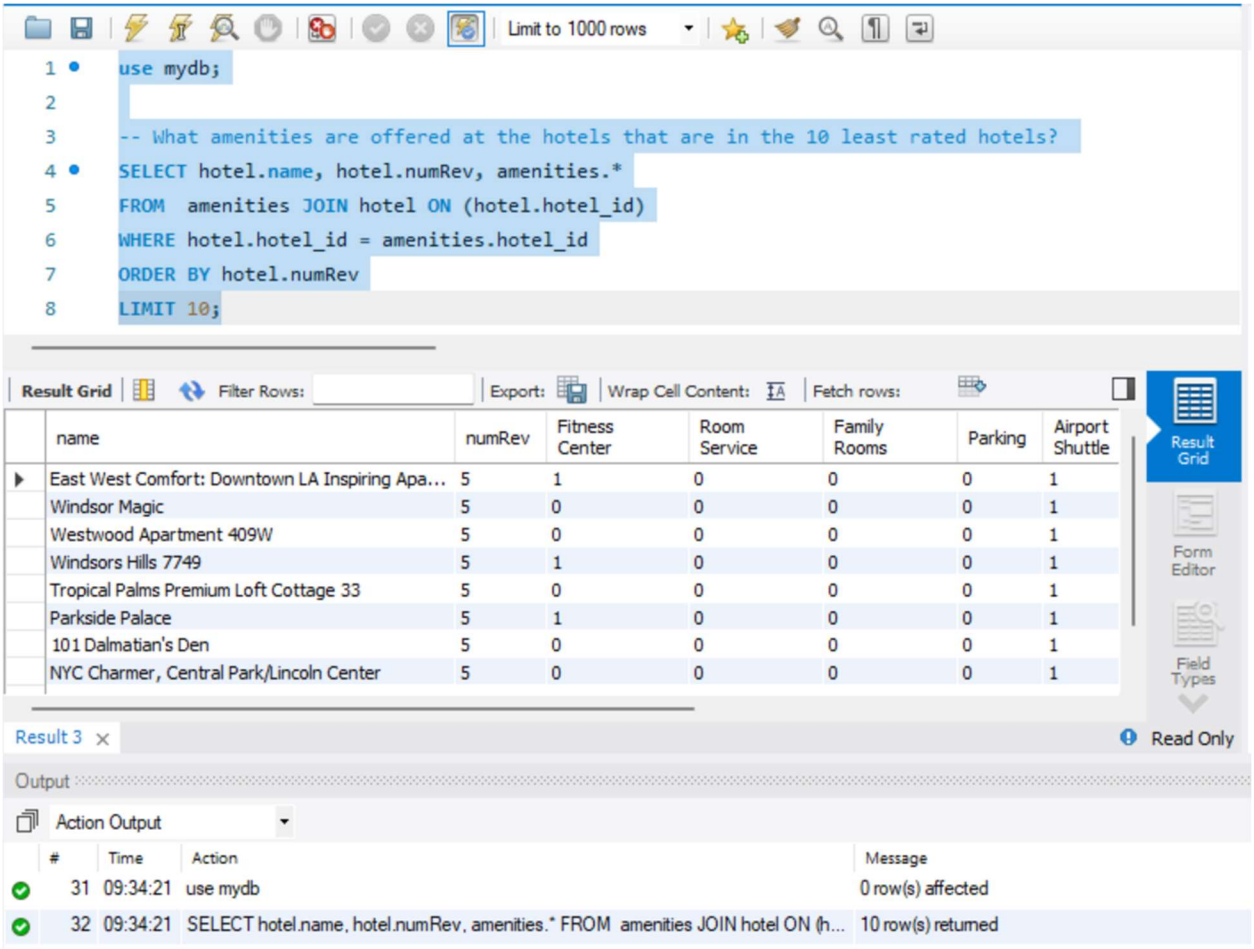
Action Output

| # | Time | Action | Message |
|---|---|---|---|
| 29 | 09:23:00 | use mydb | 0 row(s) affected |
| 30 | 09:23:00 | SELECT hotel.name, hotel.numRev, rating_info.Score, rating_info.Cleanliness, ratin... | 10 row(s) returned |



Analysis of Insight – Based on this insight, it highlights the issue with few reviews for hotels. The issue is that few reviews can skew the perceptions of the hotels themselves, because the 10 least rated hotels listed above have ratings above 8.0. So, why are these hotels reviewed the least but have the highest ratings? Well, that answer lies in how many bookings these hotels might be getting through the Booking.com platform. So, Booking.com will need to come up with ways to increase traffic on booking these hotels on the platform.

**Insight 8 - What amenities are offered at the hotels that are in the 10 least rated hotels?**



```
use mydb;

-- What amenities are offered at the hotels that are in the 10 least rated hotels?
SELECT hotel.name, hotel.numRev, amenities.*
FROM  amenities JOIN hotel ON (hotel.hotel_id)
WHERE hotel.hotel_id = amenities.hotel_id
ORDER BY hotel.numRev
LIMIT 10;
```

Result Grid

| name | numRev | Fitness Center | Room Service | Family Rooms | Parking | Airport Shuttle |
|---|---|---|---|---|---|---|
| East West Comfort: Downtown LA Inspiring Apa... | 5 | 1 | 0 | 0 | 0 | 1 |
| Windsor Magic | 5 | 0 | 0 | 0 | 0 | 1 |
| Westwood Apartment 409W | 5 | 0 | 0 | 0 | 0 | 1 |
| Windsors Hills 7749 | 5 | 1 | 0 | 0 | 0 | 1 |
| Tropical Palms Premium Loft Cottage 33 | 5 | 0 | 0 | 0 | 0 | 1 |
| Parkside Palace | 5 | 1 | 0 | 0 | 0 | 1 |
| 101 Dalmatian's Den | 5 | 0 | 0 | 0 | 0 | 1 |
| NYC Charmer, Central Park/Lincoln Center | 5 | 0 | 0 | 0 | 0 | 1 |

Result 3 | Read Only

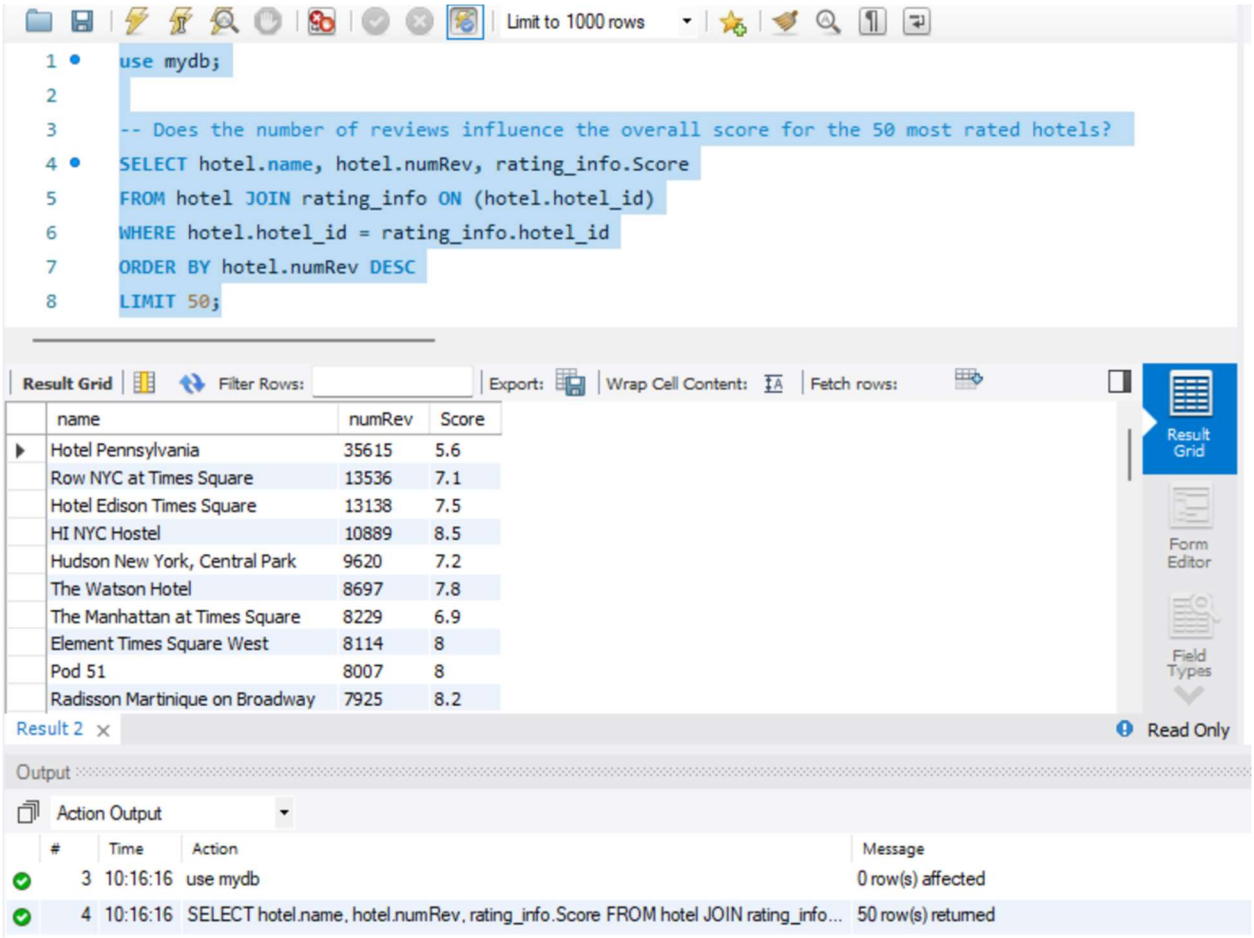
Output

Action Output

| # | Time | Action | Message |
|---|---|---|---|
| 31 | 09:34:21 | use mydb | 0 row(s) affected |
| 32 | 09:34:21 | SELECT hotel.name, hotel.numRev, amenities.* FROM amenities JOIN hotel ON (h... | 10 row(s) returned |

Analysis of Insight – Based on this insight, it looks like the most common amenities that the 10 least-rated hotels offer its guests include an airport shuttle service and a laundry room option. In comparison to Insight 6 (which looked into the amenities offered in the 10 most-rated hotels), the least-rated hotels are not offering some of the key amenities for guests like free Wi-Fi and free parking. This indicates that people refrain from going to hotels that don't offer as many of the 'given' amenities that people expect when they stay as a guest at a hotel. So, Booking.com can work with these hotels to help bring in more guests by creating offers for potential guests and slowly bringing together some of the 'given' amenities of a hotel stay.

**Insight 9 – Does the number of reviews influence the overall score for the 50 most rated hotels?**



```
use mydb;

-- Does the number of reviews influence the overall score for the 50 most rated hotels?
SELECT hotel.name, hotel.numRev, rating_info.Score
FROM hotel JOIN rating_info ON (hotel.hotel_id)
WHERE hotel.hotel_id = rating_info.hotel_id
ORDER BY hotel.numRev DESC
LIMIT 50;
```

Result Grid

| name | numRev | Score |
|---|---|---|
| Hotel Pennsylvania | 35615 | 5.6 |
| Row NYC at Times Square | 13536 | 7.1 |
| Hotel Edison Times Square | 13138 | 7.5 |
| HI NYC Hostel | 10889 | 8.5 |
| Hudson New York, Central Park | 9620 | 7.2 |
| The Watson Hotel | 8697 | 7.8 |
| The Manhattan at Times Square | 8229 | 6.9 |
| Element Times Square West | 8114 | 8 |
| Pod 51 | 8007 | 8 |
| Radisson Martinique on Broadway | 7925 | 8.2 |

Result 2 | Read Only

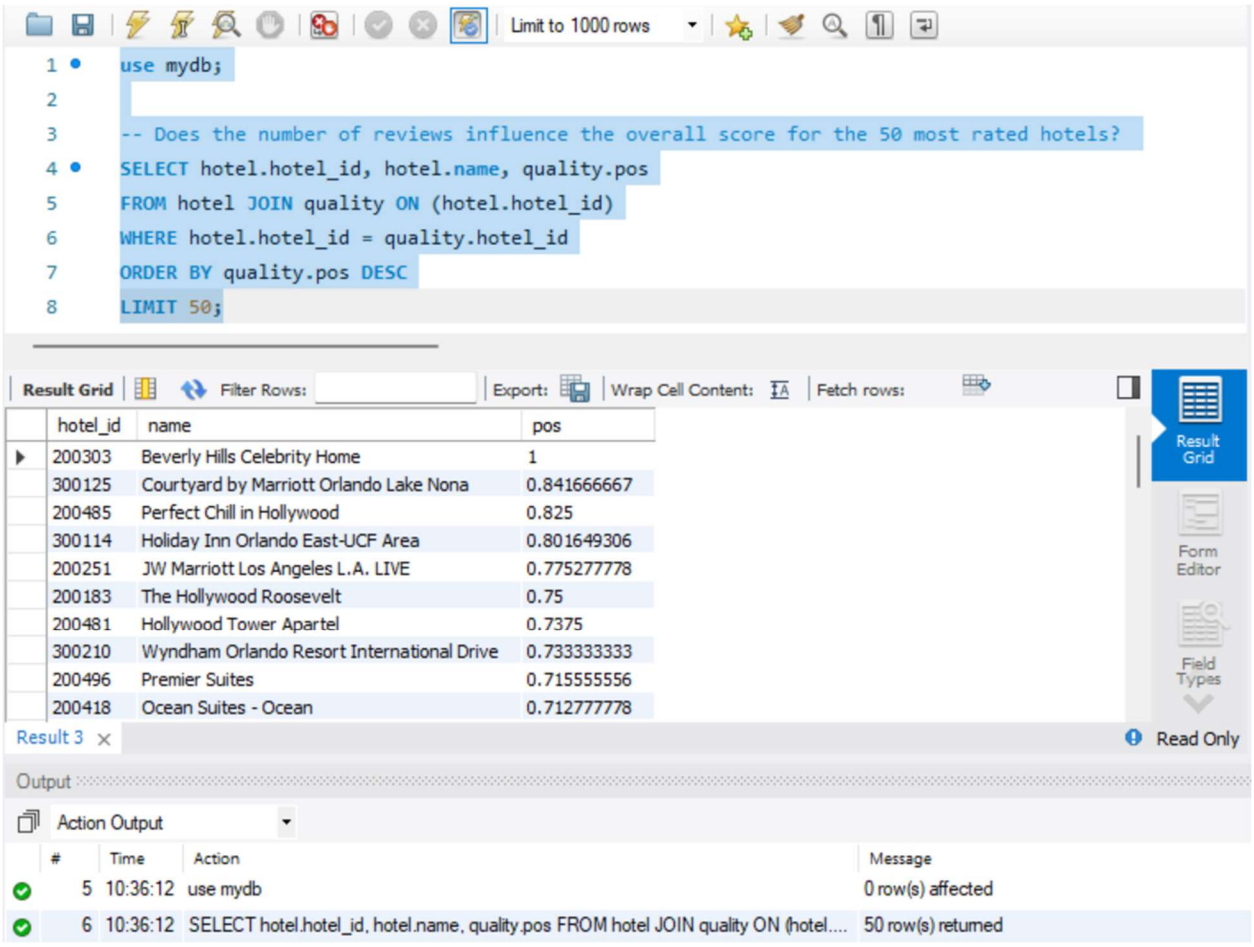
Output

Action Output

| # | Time | Action | Message |
|---|---|---|---|
| 3 | 10:16:16 | use mydb | 0 row(s) affected |
| 4 | 10:16:16 | SELECT hotel.name, hotel.numRev, rating_info.Score FROM hotel JOIN rating_info... | 50 row(s) returned |

Analysis of Insight – Based on this insight, it looks like there is a correlation between the number of reviews and the overall score of the hotel of the 50 most-rated hotels. As the number of reviews increased, it shows that the overall score has generally decreased. So, of course, with the overall score shows that there are some neutral or negative reviews that bring down the overall score of hotels that are in top positions of the 50 most-rated hotels. But to know for sure, this will require one more insight looking into the general positive and negative reviews of the hotels.

**Insight 10 – What hotels are in the top 50 when it comes to overall positive quality?**



```
use mydb;

-- Does the number of reviews influence the overall score for the 50 most rated hotels?
SELECT hotel.hotel_id, hotel.name, quality.pos
FROM hotel JOIN quality ON (hotel.hotel_id)
WHERE hotel.hotel_id = quality.hotel_id
ORDER BY quality.pos DESC
LIMIT 50;
```

Result Grid

| hotel_id | name | pos |
|---|---|---|
| 200303 | Beverly Hills Celebrity Home | 1 |
| 300125 | Courtyard by Marriott Orlando Lake Nona | 0.841666667 |
| 200485 | Perfect Chill in Hollywood | 0.825 |
| 300114 | Holiday Inn Orlando East-UCF Area | 0.801649306 |
| 200251 | JW Marriott Los Angeles L.A. LIVE | 0.775277778 |
| 200183 | The Hollywood Roosevelt | 0.75 |
| 200481 | Hollywood Tower Apartel | 0.7375 |
| 300210 | Wyndham Orlando Resort International Drive | 0.733333333 |
| 200496 | Premier Suites | 0.715555556 |
| 200418 | Ocean Suites - Ocean | 0.712777778 |

Result 3 | Read Only

Output

Action Output

| # | Time | Action | Message |
|---|---|---|---|
| 5 | 10:36:12 | use mydb | 0 row(s) affected |
| 6 | 10:36:12 | SELECT hotel.hotel_id, hotel.name, quality.pos FROM hotel JOIN quality ON (hotel.... | 50 row(s) returned |

Analysis of Insight – Based on this insight, it looks like the hotels listed are not the same as the 50 most-rated hotels – which is of no surprise because the most-rated hotels had overall score of nowhere near the 10 out of 10 rating. As such, the 50 hotels with the highest positive ratings out of their respective total reviews/ratings showcase that the hotels have moderate traffic/bookings because these hotels were not in the most-rated or least-rated lists that were gathered from previous insights/queries.

# Conclusion

Booking.com needs to do the following to help improve hotels and increase the likelihood the platform gets hotel reviews for future analysis:

1. Introduce new offers to potential hotel customers to visit certain hotels to help hotels increase bookings
2. Introduce additional offers or discounts after a hotel stay if a customer completes a review on the Booking.com platform
3. Provide information to hotels that use Booking.com in these cities with information on what people look for when booking a hotel at these destinations (i.e., what hotel characteristics or hotel amenities)