Shamaas Hussain   
MongoDB Trip Advisor Mini-Project   
  
MONGO DB QUERIES:

1. **Find all reviews for a hotel ‘Desert Rose Resort’**

db.reviews.find({"HotelInfo.Name":"Desert Rose Resort"})

OUTPUT: (Only showing some, result was larger 300 pages+)

/\* 1 \*/

{

"\_id" : ObjectId("5a1d7c469a807fe556269c76"),

"Reviews" : [

{

"Ratings" : {

"Service" : 5.0,

"Cleanliness" : "4",

"Overall" : 5.0,

"Value" : 5.0,

"Location" : "4",

"Rooms" : "4"

},

"AuthorLocation" : "",

"Title" : "“Best Deal for large families.”",

"Author" : "pb1bicdic",

"ReviewID" : "UR127021251",

"Content" : "So I found out about the hotel from searching this site. I found the reviews to be accurate and honest. When we check in, not five minutes into the hotel concierge gave us great info on free shows on the strip and the best way to maneuver from strip to downtown. If I could change anything it would be the hours of contenital breakfast. That time went so fast and the line so long that I only wanted to grab juice and leave. However, the staff was very accommodating and friendly that more than made up for it.",

"Date" : ISODate("2012-03-31T04:00:00.000Z")

},

{

"Ratings" : {

"Service" : 2.0,

"Cleanliness" : "2",

"Overall" : 2.0,

"Value" : 2.0,

"Sleep Quality" : "2",

"Rooms" : "2",

"Location" : "5"

},

"AuthorLocation" : "Sacramento, California",

"Title" : "“Not what I expected”",

"Author" : "Nancy I",

"ReviewID" : "UR126985135",

"Content" : "I am an elite owner with SHELL Vacations Club and this is one of ours. Usually SHELL properties are all GREAT, but this one leaves much to be desired! It was old and run-down, noisy, and had \"QUESTIONABLE\" employees and other guests! Not at all the nice atmosphere I expected or am used to with SHELL. It is a nice location...close to the strip...walking distance, but the place itself is almost a dive!",

"Date" : ISODate("2012-03-30T04:00:00.000Z")

},

{

"Ratings" : {

"Service" : 5.0,

"Cleanliness" : "5",

"Overall" : 5.0,

"Value" : 5.0,

"Sleep Quality" : "5",

"Rooms" : "5",

"Location" : "5"

},

"AuthorLocation" : "Milton, Canada",

"Title" : "“A nice piece of quiet away from the strip!”",

"Author" : "John E",

"ReviewID" : "UR126940047",

"Content" : "Before heading out to Vegas,for the first time,I checked reviews for this resort and it was rated in the top 10 in Vegas(Tripadvisor).My wife and myself chose this resort over others also because it had a kitchen,so we wouldn't have to buy food all the time,and because it was comprised of a number of 3 story buildings instead of overcrowded,smokey,noise filled towers.It was only a block from the strip and within walking distance to everything.We arrived in Vegas late but the 2 ladies at the reception desk were very helpful and without asking put us in a better,closer room to the front entrance.The service during our whole stay remained this way.There was no noise,everything,including the room, was very clean and breakfast was included everyday.Staff was very friendly and always willing to help.Don't get me wrong,we went to Vegas to see all the lights and glitz and glamor and we explored the strip thouroghly,but at night it was nice to come back to a place that felt more like home instead of always having to walk through a casino to find your elevator tucked away in a back corner.There was also a nice pool area(which we didn't use because it was a little cool)and if we ever go back to Vegas we will definitely be staying here again and have already recommended it to friends.",

"Date" : ISODate("2012-03-29T04:00:00.000Z")

},

{

"Ratings" : {

"Service" : 5.0,

"Cleanliness" : "5",

"Overall" : 4.0,

"Value" : 5.0,

"Sleep Quality" : "5",

"Rooms" : "5",

"Location" : "5"

},

"AuthorLocation" : "Singapore",

"Title" : "“Great value”",

"Author" : "staycation79",

"ReviewID" : "UR126931566",

"Content" : "The rooms are spacious, well equipped with lots of utensils, but we couldn't find a pair scissors. We used a pair which we brought along. It's a walk away from the noise and smoke of the casinos and we could understand why alot of the previous reviews appreciated that. Security is high, key cards not only open your room door but also the external doors to the block you are staying in. We did a bit of shopping at Las Vegas Town Square and appreciated the fact that Taxi drivers didn't have to go through the stop start traffic to drop us off at our hotel. Our room faced the road, but it didn't bother us as we slept really soundly. There's a free PC terminal with a printer to print your tour bookings and other coupons. We thought that was a nice touch. The breakfast buffet line was long, but it didn't bother us, since we would rather try the breakfast selections on the strip.",

1. Number of ratings for each Hotel. Sort the results.

Query:

db.reviews.aggregate([{ "$project": {"ratingCount": {$size: "$Reviews"},"\_id":0, "HotelInfo.Name":1, "HotelInfo.HotelID":1}}, {$group: {"\_id": "$HotelInfo.Name","noOfRatingForEachHotel": {$sum:"$ratingCount"}}}, {$sort: {"noOfRatingForEachHotel":-1}}]);  
  
Output:   
  
/\* 1 \*/

{

"\_id" : null,

"noOfRatingForEachHotel" : 145117

}

/\* 2 \*/

{

"\_id" : "Bellagio Las Vegas",

"noOfRatingForEachHotel" : 9714

}

/\* 3 \*/

{

"\_id" : "MGM Grand Hotel and Casino",

"noOfRatingForEachHotel" : 8883

}

/\* 4 \*/

{

"\_id" : "Venetian Resort Hotel Casino",

"noOfRatingForEachHotel" : 7418

}

/\* 5 \*/

{

"\_id" : "Hotel Pennsylvania New York",

"noOfRatingForEachHotel" : 6261

}

/\* 6 \*/

{

"\_id" : "Planet Hollywood Resort & Casino",

"noOfRatingForEachHotel" : 6144

}

/\* 7 \*/

{

"\_id" : "The Mirage Hotel & Casino",

"noOfRatingForEachHotel" : 5941

}

/\* 8 \*/

{

"\_id" : "Treasure Island - TI Hotel & Casino",

"noOfRatingForEachHotel" : 5940

}

/\* 9 \*/

{

"\_id" : "Wynn Las Vegas",

"noOfRatingForEachHotel" : 5580

}

1. Average overall ratings for each Hotel. Sort the results.

db.reviews.aggregate([{$unwind: "$Reviews"}, {$group: {"\_id": "$HotelInfo.Name", "avgOverAllRatingForEachHotel": {$avg: "$Reviews.Ratings.Overall"}}}, {$sort: {"avgOverAllRatingForEachHotel":-1}}]);

/\* 1 \*/

{

"\_id" : "Hotel Aquino",

"avgOverAllRatingForEachHotel" : 5.0

}

/\* 2 \*/

{

"\_id" : "Hostal Rio de Castro",

"avgOverAllRatingForEachHotel" : 5.0

}

/\* 3 \*/

{

"\_id" : "Residenza Grandi Vedute",

"avgOverAllRatingForEachHotel" : 5.0

}

/\* 4 \*/

{

"\_id" : "Conte House I",

"avgOverAllRatingForEachHotel" : 5.0

}

/\* 5 \*/

{

"\_id" : "The Mansion at MGM Grand",

"avgOverAllRatingForEachHotel" : 5.0

}

/\* 6 \*/

{

"\_id" : "Eterna Roma",

"avgOverAllRatingForEachHotel" : 5.0

}

/\* 7 \*/

{

"\_id" : "Residence Inn Indianapolis Fishers",

"avgOverAllRatingForEachHotel" : 5.0

}

/\* 8 \*/

{

"\_id" : "Centro Pellegrini",

"avgOverAllRatingForEachHotel" : 5.0

}

/\* 9 \*/

{

"\_id" : "Hotel Igel",

"avgOverAllRatingForEachHotel" : 5.0

}

1. Show hotels with number of 5.0 Overall ratings that they received.

db.reviews.aggregate([{$unwind: "$Reviews"}, {$match: {"Reviews.Ratings.Overall": 5}},{$group: {\_id: "$HotelInfo.Name", reviewsCount: {"$sum": 1}}},{$sort: {"reviewsCount": -1}}]);

/\* 1 \*/

{

"\_id" : null,

"reviewsCount" : 59923.0

}

/\* 2 \*/

{

"\_id" : "Bellagio Las Vegas",

"reviewsCount" : 5486.0

}

/\* 3 \*/

{

"\_id" : "Venetian Resort Hotel Casino",

"reviewsCount" : 3916.0

}

/\* 4 \*/

{

"\_id" : "Signature at MGM Grand",

"reviewsCount" : 3742.0

}

/\* 5 \*/

{

"\_id" : "Floridays Resort Orlando",

"reviewsCount" : 3742.0

}

/\* 6 \*/

{

"\_id" : "Wynn Las Vegas",

"reviewsCount" : 3370.0

}

/\* 7 \*/

{

"\_id" : "MGM Grand Hotel and Casino",

"reviewsCount" : 2670.0

}

/\* 8 \*/

{

"\_id" : "ARIA Resort & Casino",

"reviewsCount" : 2440.0

}

/\* 9 \*/

{

"\_id" : "Lake Buena Vista Resort Village & Spa",

"reviewsCount" : 2325.0

}

/\* 10 \*/

{

"\_id" : "The Royal Playa del Carmen",

"reviewsCount" : 2319.0

}

/\* 11 \*/

{

"\_id" : "The Mirage Hotel & Casino",

"reviewsCount" : 2260.0

}

/\* 12 \*/

{

"\_id" : "The Palazzo Resort Hotel Casino",

"reviewsCount" : 2207.0

}

/\* 13 \*/

{

"\_id" : "W New York - Times Square",

"reviewsCount" : 2193.0

}

/\* 14 \*/

{

"\_id" : "Hard Rock Hotel at Universal Orlando",

"reviewsCount" : 2169.0

}

/\* 15 \*/

{

"\_id" : "The Florida Hotel and Conference Center",

"reviewsCount" : 2168.0

}

/\* 16 \*/

{

"\_id" : "Encore At Wynn Las Vegas",

"reviewsCount" : 2117.0

}

/\* 17 \*/

{

"\_id" : "Sofitel New York",

"reviewsCount" : 2054.0

}

/\* 18 \*/

{

"\_id" : "Desert Rose Resort",

"reviewsCount" : 1963.0

}

/\* 19 \*/

{

"\_id" : "Tropicana Las Vegas",

"reviewsCount" : 1923.0

}

/\* 20 \*/

{

"\_id" : "Eventi - a Kimpton Hotel",

"reviewsCount" : 1901.0

}

/\* 21 \*/

{

"\_id" : "Secrets Maroma Beach Riviera Cancun",

"reviewsCount" : 1863.0

}

/\* 22 \*/

{

"\_id" : "Radisson Hotel Orlando - Lake Buena Vista",

"reviewsCount" : 1838.0

}

1. Number of Ratings given out per month/week

db.reviews.aggregate([ {$unwind: "$Reviews"}, {$project:{ month: {$month: "$Reviews.Date"}, day: {$dayOfWeek: "$Reviews.Date"} }}, {$group:{ \_id : {month: "$month", day: "$day"}, total:{$sum :1}}}])

/\* 1 \*/

{

"\_id" : {

"month" : 6,

"day" : 7

},

"total" : 12993.0

}

/\* 2 \*/

{

"\_id" : {

"month" : 5,

"day" : 7

},

"total" : 12159.0

}

/\* 3 \*/

{

"\_id" : {

"month" : 12,

"day" : 1

},

"total" : 13068.0

}

/\* 4 \*/

{

"\_id" : {

"month" : 12,

"day" : 2

},

"total" : 16960.0

}

/\* 5 \*/

{

"\_id" : {

"month" : 4,

"day" : 7

},

"total" : 15076.0

}

/\* 6 \*/

{

"\_id" : {

"month" : 11,

"day" : 7

},

"total" : 13933.0

}

/\* 7 \*/

{

"\_id" : {

"month" : 2,

"day" : 7

},

"total" : 14033.0

}

OTHER QUERIES THAT TAKE INTO ACCOUNT MONTH/DAY/YEAR (Allow 150 second + to run)   
  
 db.reviews.aggregate([{$unwind:"$Reviews"},{"$project":{"month":{$cond:["$Reviews.Date",{"$month":"$Reviews.Date"},-1]},"year":{$cond:["$Reviews.Date",{"$year":"$Reviews.Date" }, -1]},"Reviews":1}},{ "$group": {"\_id": {month: "$month",year: "$year"}, 'Reviews': {'$push': '$Reviews'}}},{$project: {"ratingCount": {$size: "$Reviews"}, "\_id": 1}}, {$sort: {"\_id.year":-1}}],{allowDiskUse:true});

db.reviews.aggregate([{$unwind:"$Reviews"},{"$project":{"dayOfWeek":{$cond:["$Reviews.Date",{"$dayOfWeek":"$Reviews.Date"},-1]}, "month":{$cond:["$Reviews.Date",{"$month":"$Reviews.Date"},-1]},"year":{$cond:["$Reviews.Date",{"$year":"$Reviews.Date" }, -1]},"Reviews":1}},{ "$group": {"\_id": {dayOfWeek: "$dayOfWeek", month: "$month",year: "$year"}, 'Reviews': {'$push': '$Reviews'}}},{$project: {"ratingCount": {$size: "$Reviews"}, "\_id": 1}}, {$sort: {"\_id.year":-1}}],{allowDiskUse:true});

1. Number of reviews per author

db.reviews.aggregate([{$unwind: "$Reviews"},{$group: {"\_id": "$Reviews.Author", perAuthorReviews: {$sum:1}}}, {$sort: {perAuthorReviews: -1}}], {allowDiskUse: true}).pretty();

/\* 1 \*/

{

"\_id" : "A TripAdvisor Member",

"perAuthorReviews" : 75503.0

}

/\* 2 \*/

{

"\_id" : "lass=",

"perAuthorReviews" : 17432.0

}

/\* 3 \*/

{

"\_id" : "Posted by an Accorhotels.com traveler",

"perAuthorReviews" : 4656.0

}

/\* 4 \*/

{

"\_id" : "Posted by an Easytobook.com traveler",

"perAuthorReviews" : 262.0

}

/\* 5 \*/

{

"\_id" : "Posted by a hotelsgrandparis.com traveler",

"perAuthorReviews" : 178.0

}

/\* 6 \*/

{

"\_id" : "David S",

"perAuthorReviews" : 97.0

}

/\* 7 \*/

{

"\_id" : "David B",

"perAuthorReviews" : 90.0

}

/\* 8 \*/

{

"\_id" : "David M",

"perAuthorReviews" : 89.0

}

/\* 9 \*/

{

"\_id" : "John C",

"perAuthorReviews" : 85.0

}

/\* 10 \*/

{

"\_id" : "John S",

"perAuthorReviews" : 84.0

}

/\* 11 \*/

{

"\_id" : "David H",

"perAuthorReviews" : 80.0

}

/\* 12 \*/

{

"\_id" : "John B",

"perAuthorReviews" : 74.0

}

/\* 13 \*/

{

"\_id" : "ITA\_And\_RE\_a",

"perAuthorReviews" : 69.0

}

/\* 14 \*/

{

"\_id" : "John M",

"perAuthorReviews" : 68.0

}

1. Find the number of hotels having better average overall rating than "Desert Rose Resort" – Should not be a static query

db.reviews.aggregate([{$project: {"HotelInfo.Name":1, "Reviews.Ratings.Overall":1}}, {$unwind: "$Reviews"}, {$group: {"\_id": "$HotelInfo.Name",allHotels: { $push: { \_id: "$HotelInfo.Name", avgOverAllRatingForEachHotel: {$avg: "$Reviews.Ratings.Overall"}}}}}, {$limit:5}], {allowDiskUse:true});

-I couldn’t finish this query. I pasted my attempt at it.   
  
Extra Queries for more findings:

1. Find hotels that have average 5.0 overall rating.

db.reviews.aggregate([{$unwind: "$Reviews"}, {$group: {"\_id": "$HotelInfo.Name", "avgOverallRating": {$avg: "$Reviews.Ratings.Overall"}}},{$project:{"avgOverallRating":1, "hasAVG": {$eq:["$avgOverallRating", 5.0]}}},{ "$match": { "hasAVG": true } } , {$sort: {"\_id":-1}}]);

/\* 1 \*/

{

"\_id" : "The Mansion at MGM Grand",

"avgOverallRating" : 5.0,

"hasAVG" : true

}

/\* 2 \*/

{

"\_id" : "Residenza Grandi Vedute",

"avgOverallRating" : 5.0,

"hasAVG" : true

}

/\* 3 \*/

{

"\_id" : "Residenza Ca Felice",

"avgOverallRating" : 5.0,

"hasAVG" : true

}

/\* 4 \*/

{

"\_id" : "Residence Inn Indianapolis Fishers",

"avgOverallRating" : 5.0,

"hasAVG" : true

}

/\* 5 \*/

{

"\_id" : "Petit Prince",

"avgOverallRating" : 5.0,

"hasAVG" : true

}

/\* 6 \*/

{

"\_id" : "Marini Park Hotel",

"avgOverallRating" : 5.0,

"hasAVG" : true

}

/\* 7 \*/

{

"\_id" : "La Piazzetta di Roma",

"avgOverallRating" : 5.0,

"hasAVG" : true

}

/\* 8 \*/

{

"\_id" : "Hotel Konfidentiel",

"avgOverallRating" : 5.0,

"hasAVG" : true

}

/\* 9 \*/

{

"\_id" : "Hotel Igel",

"avgOverallRating" : 5.0,

"hasAVG" : true

}

/\* 10 \*/

{

"\_id" : "Hotel Aquino",

"avgOverallRating" : 5.0,

"hasAVG" : true

}

/\* 11 \*/

{

"\_id" : "Hostal Rio de Castro",

"avgOverallRating" : 5.0,

"hasAVG" : true

}

/\* 12 \*/

{

"\_id" : "Freestanza Accommodation",

"avgOverallRating" : 5.0,

"hasAVG" : true

}

/\* 13 \*/

{

"\_id" : "Eterna Roma",

"avgOverallRating" : 5.0,

"hasAVG" : true

}

/\* 14 \*/

{

"\_id" : "Conte House I",

"avgOverallRating" : 5.0,

"hasAVG" : true

}

/\* 15 \*/

{

"\_id" : "Centro Pellegrini",

"avgOverallRating" : 5.0,

"hasAVG" : true

}

/\* 16 \*/

{

"\_id" : "Central Park Suite",

"avgOverallRating" : 5.0,

"hasAVG" : true

}

Findings:

-There is an unusual amount of reviews from "A TripAdvisor Member." These appear to be from websites from verified users, but they do not have any profile associated with them.

-There also seems to be a lot of reviews that are automated from "Posted by an Accorhotels.com traveler" and "Posted by an Easytobook.com traveler"

After some research, this seems to be a problem that is asked by many travelers:

https://www.tripadvisor.com/ShowTopic-g1-i12105-k4601523-o10-Why\_is\_An\_AccorHotels\_com\_traveler\_on\_my\_tag-TripAdvisor\_Support.html

-David and John seem to a common name for users who posts a lot of reviews. These might be bots.

-A lot of reviews are for hotel name null.

-

Part 4: mySQL (RESULTS NOT SHOWN SINCE ONLY PUT 1-2 TUPLES IN EACH TABLE)

Database Tables:

Hotel [Name, HotelURL, Price, Address, HotelID, ImgURL]   
 Reviews [HotelID, RatingID, AuthorID, AuthorLocation, Title, ReviewID, Content, Date]

Author[AuthorID, author]

Ratings[RatingID, Service, Cleanliness, Overall, Value, Sleep Quality, Rooms, Location]

1. Find all reviews for a hotel ‘Desert Rose Resort’

SELECT r1.\*

FROM Reviews r1, Hotel h1

WHERE h1.Name = ‘Desert Rose Resort’ AND h1. HotelID = r1. HotelID;

1. Number of ratings for each Hotel. Sort the results

SELECT h1.Name, COUNT(r1.RatingID) as NumRating

FROM Reviews r1, Hotel h1  
WHERE h1.HotelID=r1.HotelID

GROUP BY h1.HotelID

ORDER BY NumRating DESC

1. Average overall ratings for each Hotel. Sort the results

SELECT AVG(ra1.Overall) AS Avg\_Overall\_Rating

FROM Reviews r1, Hotel h1, Ratings ra1

WHERE h1.HotelID=r1.HotelID AND r1.RatingID=ra1.RatingID

GROUP BY h1.HotelID

ORDER BY Avg\_Overall\_Rating DESC;

1. Show hotels with number of 5.0 Overall ratings that they recieved.

SELECT h1.Name, COUNT(ra1.Overall)

FROM Reviews r1, Hotel h1, Ratings ra1

WHERE h1.HotelID=r1.HotelID AND r1.RatingID=ra1.RatingID AND ra1.Overall=5.0;

GROUP BY h1.Name

1. Number of Ratings given out per month/day of week

SELECT COUNT (r1.RatingID), MONTH(r1.date), DAY(r1.date)

FROM Reviews r1, Hotel h1, Ratings ra1  
WHERE h1.HotelID=r1.HotelID AND r1.RatingID=ra1.RatingID

GROUP BY MONTH(r1.date), DAY(r1.date)

1. Number of reviews per author

SELECT a1.Author, COUNT(r1.ReviewID)

FROM Reviews r1, Hotel h1, Author a1  
WHERE h1.HotelID=r1.HotelID AND r1.AuthorID = a1.AuthorID

GROUP BY a1.Author

1. Find the number of hotels having better average overall rating than "Desert Rose Resort" – Should not be a static query

SET @Checking = (SELECT AVG(r2.rating)

​​​​FROM Reviews r, Hotel h, Ratings r1

​​​​INNER JOIN Hotel ON h.HotelID=r.HotelID

AND

INNER JOIN Reviews ON r.RatingID = r1.RatingID

WHERE h.Name = 'Desert Rose Resort '

​​​​GROUP BY h.HotelID);

SELECT h.Name, AVG(r2.rating)

FROM Hotel h, Reviews r, Ratings r2

WHERE h.HotelID = r.HotelID AND r.RatingID = r2.RatingID

GROUP BY r.HotelID

HAVING AVG(r2.rating) > @Checking

­­­­­