MongoDB Project DB submission

# **Group 4**

## Rishab Katta

## Tuheena Singh

## Mithila Nigudkar

We used the Publicly available Yelp dataset. We divided the dataset into 5 collections – Business, Checkin, Review, Tip and Photos(Photos.chunks and Photos.files). We used Bash to load the databases and Javascript to remove unused (Movie and Restaurant) data from the collections. The code for our project can be found at - <https://gitlab.com/rishabkatta>

Here’s some sample data from the Mongo Client after Loading the Database.

**All collections overview:**

> show collections

Business

Checkin

Photos.chunks

Photos.files

Review

Tip

**Business Collection :**

> db.Business.find().limit(10).pretty();

{

"\_id" : ObjectId("60735637e6ae54f2a4e7ecbb"),

"business\_id" : "oaepsyvc0J17qwi8cfrOWg",

"name" : "Great Clips",

"address" : "2566 Enterprise Rd",

"city" : "Orange City",

"state" : "FL",

"postal\_code" : "32763",

"latitude" : 28.9144823,

"longitude" : -81.2959787,

"stars" : 3,

"review\_count" : 8,

"is\_open" : 1,

"attributes" : {

"RestaurantsPriceRange2" : "1",

"BusinessAcceptsCreditCards" : "True",

"GoodForKids" : "True",

"ByAppointmentOnly" : "False"

},

"categories" : "Beauty & Spas, Hair Salons",

"hours" : null

}

{

"\_id" : ObjectId("60735637e6ae54f2a4e7ecbc"),

"business\_id" : "PE9uqAjdw0E4-8mjGl3wVA",

"name" : "Crossfit Terminus",

"address" : "1046 Memorial Dr SE",

"city" : "Atlanta",

"state" : "GA",

"postal\_code" : "30316",

"latitude" : 33.7470274,

"longitude" : -84.3534244,

"stars" : 4,

"review\_count" : 14,

"is\_open" : 1,

"attributes" : {

"GoodForKids" : "False",

"BusinessParking" : "{'garage': False, 'street': False, 'validated': False, 'lot': False, 'valet': False}",

"BusinessAcceptsCreditCards" : "True"

},

"categories" : "Gyms, Active Life, Interval Training Gyms, Fitness & Instruction",

"hours" : {

"Monday" : "16:0-19:0",

"Tuesday" : "16:0-19:0",

"Wednesday" : "16:0-19:0",

"Thursday" : "16:0-19:0",

"Friday" : "16:0-19:0",

"Saturday" : "9:0-11:0"

}

}

{

"\_id" : ObjectId("60735637e6ae54f2a4e7ecbd"),

"business\_id" : "bvN78flM8NLprQ1a1y5dRg",

"name" : "The Reclaimory",

"address" : "4720 Hawthorne Ave",

"city" : "Portland",

"state" : "OR",

"postal\_code" : "97214",

"latitude" : 45.5119069956,

"longitude" : -122.6136928797,

"stars" : 4.5,

"review\_count" : 13,

"is\_open" : 1,

"attributes" : {

"BusinessAcceptsCreditCards" : "True",

"RestaurantsPriceRange2" : "2",

"ByAppointmentOnly" : "False",

"BikeParking" : "False",

"BusinessParking" : "{'garage': False, 'street': True, 'validated': False, 'lot': False, 'valet': False}"

},

"categories" : "Antiques, Fashion, Used, Vintage & Consignment, Shopping, Furniture Stores, Home & Garden",

"hours" : {

"Thursday" : "11:0-18:0",

"Friday" : "11:0-18:0",

"Saturday" : "11:0-18:0",

"Sunday" : "11:0-18:0"

}

}

**Checkin collection:**

> db.Checkin.find().limit(5).pretty();

{

"\_id" : ObjectId("6073563ea3a7db545e367776"),

"business\_id" : "--0r8K\_AQ4FZfLsX3ZYRDA",

"date" : "2017-09-03 17:13:59"

}

{

"\_id" : ObjectId("6073563ea3a7db545e367777"),

"business\_id" : "--0zrn43LEaB4jUWTQH\_Bg",

"date" : "2010-10-08 22:21:20, 2010-11-01 21:29:14, 2010-12-23 22:55:45, 2011-04-08 17:14:59, 2011-04-11 21:28:45, 2011-04-26 16:42:25, 2011-05-20 19:30:57, 2011-05-24 20:02:21, 2011-08-29 19:01:31"

}

{

"\_id" : ObjectId("6073563ea3a7db545e367778"),

"business\_id" : "--2aF9NhXnNVpDV0KS3xBQ",

"date" : "2014-11-03 16:35:35, 2015-01-30 18:16:03, 2015-03-16 18:45:30, 2016-08-25 15:42:01, 2017-03-27 20:32:57, 2018-02-12 23:13:56, 2019-04-22 19:34:48, 2020-12-29 16:22:00"

}

{

"\_id" : ObjectId("6073563ea3a7db545e367779"),

"business\_id" : "--2mEJ63SC\_8\_08\_jGgVIg",

"date" : "2010-12-15 17:10:46, 2013-12-28 00:27:54, 2015-10-18 00:43:55, 2016-06-11 19:56:11"

}

**Review collection:**

> db.Review.find().limit(5).pretty();

{

"\_id" : ObjectId("6073565f268a064ef6803c92"),

"review\_id" : "8bFej1QE5LXp4O05qjGqXA",

"user\_id" : "YoVfDbnISlW0f7abNQACIg",

"business\_id" : "RA4V8pr014UyUbDvI-LW2A",

"stars" : 4,

"useful" : 1,

"funny" : 0,

"cool" : 0,

"text" : "This store is pretty good. Not as great as Walmart (or my preferred, Milford Target), but closer and in a easier area to get to. \nThe store itself is pretty clean and organized, the staff are friendly (most of the time), and BEST of all is the Self Checkout this store has! \nGreat clearance sections throughout, and great prices on everything in the store, in general (they pricematch too!). \nChristian, Debbie, Jen and Hanna are all very friendly, helpful, sensitive to all customer needs. Definitely one of the better Target locations in the area, and they do a GREAT job assisting customers for being such a busy store. Located directly in the Framingham Mall on Cochituate Rd / Route 30. 4 stars.",

"date" : "2015-07-03 20:38:25"

}

{

"\_id" : ObjectId("6073565f268a064ef6803c93"),

"review\_id" : "2l\_TDrQ7p-5tANOyiOlkLQ",

"user\_id" : "LWUnzwK0ILquLLZcHHE1Mw",

"business\_id" : "mD-A9KOWADXvfrZfwDs-jw",

"stars" : 4,

"useful" : 1,

"funny" : 0,

"cool" : 0,

"text" : "I am definitely a fan of Sports Authority. This particular location has a good check in deal. We came here near Christmas time to buy some presents and we had a good experience. The staff members were very friendly and they helped us find what we were looking for. We got some golf stuff, two pairs of shoes, a tennis racket bag, and some bicycle accessories. The store was clean and well organized. They have everything from apparel to basketballs and everything in between. Good spot to buy athletic merchandise.",

"date" : "2012-05-28 15:00:47"

}

{

"\_id" : ObjectId("6073565f268a064ef6803c94"),

"review\_id" : "KKVFopqzcVfcubIBxmIjVA",

"user\_id" : "99RsBrARhhx60UnAC4yDoA",

"business\_id" : "EEHhKSxUvJkoPSzeGKkpVg",

"stars" : 5,

"useful" : 0,

"funny" : 0,

"cool" : 0,

"text" : "I work in the Pru and this is the most affordable and tasty place in the food court. deals where a meal is $5-$7 and the chicken pesto is really good. I am not a chowda person but all there soups I have had are pretty damn good. Broccoli chicken is my favorite. Also, probably the most personable Food court staff I have ever had the pleasure of ordering from.",

"date" : "2014-05-07 18:10:21"

}

{

"\_id" : ObjectId("6073565f268a064ef6803c95"),

"review\_id" : "FdoBFTjXXMn4hVnJ59EtiQ",

"user\_id" : "eLAYHxHUutiXswy-CfeiUw",

"business\_id" : "WQFn1A7-UAA4JT5YWiop\_w",

"stars" : 1,

"useful" : 0,

"funny" : 0,

"cool" : 0,

"text" : "They NEVER seem to get our \norder correct, service is crappy, food is inconsistent and has gone down hill steadily in the last 6-9 months! WILL NEVER GO THERE AGAIN!",

"date" : "2017-09-08 23:26:10"

}

{

"\_id" : ObjectId("6073565f268a064ef6803c96"),

"review\_id" : "T5fAqjjFooT4V0OeZyuk1w",

"user\_id" : "SFQ1jcnGguO0LYWnbbftAA",

"business\_id" : "0AzLzHfOJgL7ROwhdww2ew",

"stars" : 2,

"useful" : 1,

"funny" : 1,

"cool" : 1,

"text" : "I've stayed at many Marriott and Renaissance Marriott's and this was a huge disappointment! The front desk and atrium is nice..there is a starbucks on site which is nice.\n\nThe rooms are run down and old. There is a flat screen but that is to be expected of a Renaissance.\n\nWe got this hotel via Priceline at a rate of $75/night...good deal for the price but this is not a true Renaissance.",

"date" : "2010-01-08 02:29:15"

}

**Tip Collection:**

> db.Tip.find().limit(5).pretty();

{

"\_id" : ObjectId("6073577d40be4423f96a7450"),

"user\_id" : "WCjg0jdHXMlwbqS9tZUx8Q",

"business\_id" : "ENwBByjpoa5Gg7tKgxqwLg",

"text" : "Carne asada chips...",

"date" : "2011-07-22 19:07:35",

"compliment\_count" : 0

}

{

"\_id" : ObjectId("6073577d40be4423f96a7451"),

"user\_id" : "fTsVDajAyDJ-YzsSdfXSDw",

"business\_id" : "oQyf1788YWsiDLupGva6sw",

"text" : "Cold cuts are the best",

"date" : "2015-06-09 14:35:57",

"compliment\_count" : 0

}

{

"\_id" : ObjectId("6073577d40be4423f96a7452"),

"user\_id" : "wDWoMG5N9oI4DJ-p7z8EBg",

"business\_id" : "XWFjKtRGZ9khRGtGg2ZvaA",

"text" : "1/2-price bowling & the \"Very\" Old Fashion are excellent, but the drink didn't help my bowling score!",

"date" : "2017-07-11 23:07:16",

"compliment\_count" : 0

}

{

"\_id" : ObjectId("6073577d40be4423f96a7453"),

"user\_id" : "tOFEtck\_Xnj7y0VzXtqCtA",

"business\_id" : "EUOuLlhkf1PBM9G9rdxBtQ",

"text" : "Most important 1 of the night.",

"date" : "2011-12-24 06:32:37",

"compliment\_count" : 0

}

{

"\_id" : ObjectId("6073577d40be4423f96a7454"),

"user\_id" : "g3HFkZgloxLQCvoJ5Zs5gg",

"business\_id" : "OQ2oHkcWA8KNC1Lsvj1SBA",

"text" : "Water Ave coffee served here--& @ L Jeans!",

"date" : "2018-03-22 15:28:09",

"compliment\_count" : 0

}

**Photos (GridFS)**

**Photos.files**

> db.Photos.files.find().limit(5).pretty();

{

"\_id" : ObjectId("6073b946452b1dc06440cef9"),

"length" : NumberLong(13307),

"chunkSize" : 261120,

"uploadDate" : ISODate("2021-04-12T03:06:46.915Z"),

"filename" : "./data/yelp\_photos/photos/-1MVd6ljY2JT6gCKnrc-uw.jpg",

"metadata" : {

}

}

{

"\_id" : ObjectId("6073b94642944a412f20d66b"),

"length" : NumberLong(43113),

"chunkSize" : 261120,

"uploadDate" : ISODate("2021-04-12T03:06:46.935Z"),

"filename" : "./data/yelp\_photos/photos/-1TUzO4XCmOeqi9eAGeDWA.jpg",

"metadata" : {

}

}

{

"\_id" : ObjectId("6073b9461135aeb4c4b88b25"),

"length" : NumberLong(217101),

"chunkSize" : 261120,

"uploadDate" : ISODate("2021-04-12T03:06:46.956Z"),

"filename" : "./data/yelp\_photos/photos/-4mowOKV3xGyfgQBh6ey\_g.jpg",

"metadata" : {

}

}

{

"\_id" : ObjectId("6073b9460060dad09ad5cc5e"),

"length" : NumberLong(26212),

"chunkSize" : 261120,

"uploadDate" : ISODate("2021-04-12T03:06:46.976Z"),

"filename" : "./data/yelp\_photos/photos/-IMbvKYt8ZxayUzfMgDVeA.jpg",

"metadata" : {

}

}

{

"\_id" : ObjectId("6073b946e3840215c3748cca"),

"length" : NumberLong(16993),

"chunkSize" : 261120,

"uploadDate" : ISODate("2021-04-12T03:06:46.994Z"),

"filename" : "./data/yelp\_photos/photos/-aQJUU104Sc9vuAV\_EtXrA.jpg",

"metadata" : {

}

}

**Photos.chunks**

> db.Photos.chunks.find().limit(5).pretty();

{

"\_id" : ObjectId("6073b946452b1dc06440cefa"),

"files\_id" : ObjectId("6073b946452b1dc06440cef9"),

"n" : 0,

"data" : BinData(0,"/9j/4AAQSkZ …

}

{

"\_id" : ObjectId("6073b94642944a412f20d66c"),

"files\_id" : ObjectId("6073b94642944a412f20d66b"),

"n" : 0,

"data" : BinData(0,"/9j/4AAQ…

}