

MongoDB practical session test part

BSc data science for responsible business

EC Lyon. EM Lyon

March 2026, Előd Egyed-Zsigmond

Send the report to elod.egyed-zsigmond@insa-lyon.fr before 18h00 this evening!

This exercise supposes that you have completed the training part of the lab.

You will start by working on the `velov_geo` collection.

Provide the query and the answers for the following questions:

1. What are the distinct values of the `availability` field in the `velov_geo` collection ?
2. Provide the average number of available bikes in the commune that has the more `velov_stations`.
3. Collection comparison
 - a. Download the GeoJSON dataset with the actual velo'V station situation from the following page : <https://data.grandlyon.com/portail/fr/jeux-de-donnees/stations-velo-v-metropole-lyon-disponibilites-temps-reel/telechargements> . Import it in the same database, in the `velov2026` collection!
 - b. Count the number of `velov` stations in the new collection
 - c. Retrieve the list of communes (with no repetitions)
 - d. Provide a MongoDB query that verifies whether there are `velov` stations with their "`total_stands.capacity`" different from "`main_stands.capacity`". Provide the count of such stations.
 - e. We want to study the differences between the collections `velov2026` and `velov_geo`. Study the data and propose mongodb solutions to list the stations that are in common, and the stations that are only present in one of the two collections, not using their geographic coordinates. Optimize your queries as much as possible. Check your results!
Describe and explain your approach, provide the mongodb queries and illustrate the results with meaningful and easy to understand examples. Provide statistics as well (number of stations in the different categories).
4. Data enrichment. Look for downloadable data in JSON format on the website <https://data.grandlyon.com/> (for example Touristic interest points : <https://data.grandlyon.com/portail/fr/jeux-de-donnees/points-interet-touristiques-metropole-lyon-v2/telechargements> but you have veloV charging station, parks, bike lanes, as well) that can complement the information of Vélo'V stations.

- a) Provide the URL that allows you to download the data and explain how you think it can complement the MongoDB collection of Vélo’V stations.
 - b) Download the data and import it into MongoDB, simplify it to have more than one document. Describe your procedure.
 - c) Provide at least 3 MongoDB queries that combine your new collection with the `velov_geo` collection using the `$lookup` operator:
 - c.1 first query that involves address comparison (`properties.address`, `properties.nom`, ...) check the fields)
 - c.2 second query that implies geographic coordinate comparison. Create indexes if necessary.
 - c.3 for the additional query use your imagination
- Describe your method. Provide: query, results and explanation. Describe your method. Provide: query, results and explanation.

