Joel

In MongoDB, when querying data, the program will scan through all the documents to find a match to the data that is being requested. Depending on the size of the collection, this could take a while. This is where indexing can make a difference. Instead of scanning through every document, the index can limit the query to a specific group of documents which will decrease the amount of processing required.

On MongoDB’s website, the process of indexing is described as data structures “that store a small portion of the collection’s data set in an easy to traverse form.” [[1]](https://cyberactive.bellevue.edu/webapps/discussionboard/do/message?action=list_messages&course_id=_510116_1&nav=discussion_board&conf_id=_717709_1&forum_id=_2359938_1&message_id=_43852722_1" \l "_ftn1" \o ") This process allows MongoDB to associate certain query data with these documents in the index. MongoDB already automatically creates an index for the document ids, which can be found under the Indexes section.

A simple example of indexing could be a database of employees. If there are queries that constantly attempt to pull data by a certain field, like maybe job status, an index could be created for records with that field. This would allow MongoDB to already have a list of the data compiled so the next query will allow it to search through that list instead of through all the documents.

Another example of when indexing should be used could be a compound index, which could include multiple fields to increase query efficiency. This could be a great option for more complex queries that are searching for multiple points of data, like maybe a subscription service that needs to find users in a specific geography, with a specific membership option. Without indexing, the query could potentially be searching through millions of records per country, and millions per membership type.

The process of creating an index is fairly simple. There are several methods available online, depending on which programming language or method is being used to query the data. For pymongo, I found the following website which has a lot of useful information and describes how to create an index: https://www.analyticsvidhya.com/blog/2020/09/mongodb-indexes-pymongo-tutorial/