NoSQL Homework

In this exercise, you practice the interaction with a NoSQL Document database. For this, a dataset is provided, on which the given queries should be performed. The exercise consists of two parts: an on-premise exercise with **MongoDB** and a cloud-based exercises with **Azure Cosmos DB**. The tasks in this exercise are not hard themselves. The challenge lies in overcoming the issues of setting up a Document datastore on-premise and in the cloud.

Your report should include the **queries** and **commands** that you wrote, as well as the **obtained results**, together with **screenshots**, documenting your **setup/configuration process** and the **cloud migration process**. Please submit everything in a **single PDF-Document** through Moodle.

1 PART 1 – MONGODB (5 POINTS)

In the first part of the exercise, you have to setup a MondoDB Document datastore on your local machine. You can use the provided virtual machine to setup MongoDB if you want to avoid installing it on your local machine. However, you have to setup MongoDB there manually as well. Then you will work with MongoDB using a provided dataset (import, query, update).

1.1 PREREQUISITES

Before starting with this exercise, read through the **MongoDB slides** provided in the Moodle course. The slides explain the basic concepts of MongoDB.

1.2 SETUP MONGODB

To complete the first part of thze exercise, you have to install a MongoDB Server on your local machine (i.e., "on-premise").

- Download and install the `MongoDB Community Server` for on-premise installation
  - You do not have to install MongoDB as a service, if you only use it for this exercise, but that's up to you
  - You can choose to install `MondoDB Compass` (GUI for importing and querying data)
- Additionally, download and install `MonboDB Database Tools` (required for `mongodump`)
- After the installation, you have to run your MongoDB Server
- You can find the tools at https://www.mongodb.com/
  - Remember to select the appropriate platform when downloading
  - Instructions for installation and running can be found at: https://docs.mongodb.com/manual/administration/install-community/

1.3 IMPORT DATA

- Download the exercise dataset from the Moodle course (`dataset.zip`) and extract it
- Choose a new database
- Import each file into a collection
  - Hint: Can be done using MongoDB Compass

1.4 QUERY DATA

Execute the following queries on your local MongoDB. You can either use the command line client or the MonboDBCompass "MongoSH Beta". In your report, include the statements, as well as the outputs.

a) Select all Austrian cities (`countryID = 15`)!
b) Select all Austrian cities in ascending order!
c) Select all Austrian cities in descending order!
d) Select the number of Austrian cities that are included in the dataset!
e) Select solely the city names of all Austrian cities!
f) Select all countries, which exhibit a population between 15 and 20 millions of people!

1.5 UPDATE DATA

Update the data according to the following instructions on your local MongoDB. You can either use the command line client or the MonboDBCompass "MongoSH Beta". In your report, include the statements, as well as the outputs.

a) Increase the population of Austria (`countryID = 15`) by 3 persons.
b) Decrease the population of Austria by 3 persons.

2 PART 2 – AZURE COSMOS DB (5 POINTS)

In the second part, you will move to the cloud, using a Microsoft Azure Cosmos DB. You will migrate the data from your local MongoDB to the cloud storage and query it from there.

2.1 SETUP COSMOS DB

Create an Azure Cosmos DB Account using the Azure Portal.

- Choose a reasonable configuration (Subscription, Resource Group, Account Name, Location, can be chosen freely)
- For API, select "Azure Cosmos DB for MongoDB API"
- Hint: Create a new resource group for the exercise

2.2 MIGRATE DATA

To migrate the data from your on-premise MongoDB Server to the cloud, several steps are required. The data must first be exported from MongoDB and then prepared on an Azure storage account. An Azure Database Migration Service resource is needed to execute the migration process.

2.2.1 Export Data from MongoDB

- Use the tool `mongodump` to create a bson dump of your database with all the collections
- For details on `mongodump`, see
  https://docs.mongodb.com/database-tools/mongodump/

2.2.2 Upload all the files to a Storage Account

- Create a Storage Account in the Resource Group
- Select the Storage Account in the Portal and choose "Blob service" → "Containers"
- Create a new container (public access level must be "Container") and select it
- Upload all files created by the dump to a new folder in the container

- Copy the URL you see, when choosing "Properties" in the container view

2.2.3   Azure Database Migration Service (offline)

To migrate data to Azure, a new "Azure Database Migration Service" resource must be created. Search for the service name in the search bar on top of the portal. Follow the tutorial below, to create the resource and migrate your data

https://docs.microsoft.com/en-us/azure/dms/tutorial-mongodb-cosmos-db

- The "Prerequisites" step is already completed
- Skip the "Register the Microsoft.DataMigration resource provider" step
- Complete the "Create an instance" and "Create a migration project" steps
- Complete the "Specify source details" step
  - Choose "Data from Azure storage"
  - The "storage account connection string" is the URL you copied above in 2.2.2
- Complete the "Specify target details" step
  - Select your Cosmos DB name, the connection string is set automatically
- Complete the "Map to target databases", "Run the migration", "Monitor the migration" and "Verify data in Cosmos DB" steps

2.3   QUERY/UPDATE DATA

Use the "Data Explorer" of Cosmos DB to execute your queries in a Mongo Shell. To test if your CosmosDB works and contains all the required data, execute the following queries:

- Two queries of your choice from section 1.4
- One query of your choice from section 1.5