Chase

CS340

Sunday, December 14, 2024

Professor Wilson

* 1. Assignment: Data Aggregation Pipeline

1. Using the mongoimport tool, **create the database** “companies”, db.research.find({"name" : "AdventNet"})

db.research.find({"founded\_year" : 1996},{"name" : 1}).limit(10)

A screenshot of a computer

Description automatically generated

A screenshot of a computer program

Description automatically generated

1. Perform the following tasks **using MongoDB queries**

List only the first 20 names of companies founded after the year 2010, ordered alphabetically.

List only the first 20 names of companies with offices in either California or Texas, ordered by the number of employees and sorted largest to smallest.

A screenshot of a computer screen

Description automatically generated

A screenshot of a computer screen

Description automatically generated

1. Design and implement a MongoDB aggregation pipeline to show the total number of offices by state for all companies that have offices in the United States.

A screenshot of a computer

Description automatically generated

1. Design and implement a MongoDB aggregation pipeline to show the total number of offices by state for all companies that have offices in the United States. Be sure that you account for the fact that some companies have offices in several states. Explain your aggregation pipeline.

I used the following stages: First, I applied unwind to create individual documents for each office a company has. Next, I used match to filter out companies without an office in the USA. Then, I applied group to organize the documents by state code and used sum to calculate the total number of offices for each state. Lastly, I used sort with a value of -1 to order the results in descending order, displaying states with the most offices first.