**Thomas Martin**

**CS-340-11221-M01 Client/Server Development 2024**

**8-1 Assignment: Data Aggregation Pipeline**

**Southern New Hampshire University**

**June 26, 2024**

Using the mongoimport tool, create the database “companies” by loading the documents found in the “companies.json” file into the “research” collection. This file is located in the “/usr/local/datasets/” directory in Apporto. Verify your load by issuing the following queries

A screenshot of a computer

Description automatically generated

db.research.find({"name" : "AdventNet"})

A screenshot of a computer program

Description automatically generated

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

A screenshot of a computer

Description automatically generated

Perform the following tasks using MongoDB queries:

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

A screenshot of a computer

Description automatically generated

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

Description automatically generated

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.

db.research.aggregate

Step 1: Match documents with offices in the US

"offices.country\_code": "USA"

Step 2: Unwind the offices array

$unwind: "$offices"

Step 3: Match only offices in the US

$match: {

"offices.country\_code": "USA"

Step 4: Group by state and count offices

$group: {

\_id: "$offices.state",

totalOffices: { $sum: 1 }

Step 5: Project the output format

$project: {

\_id: 0,

state: "$\_id",

totalOffices: 1

A screenshot of a computer

Description automatically generated