Ethan Daugherty

SNHU

CS-340-T4528

8-1 Assignment: Data Aggregation Pipeline

04/24/2022

**Data Aggregation Pipeline**

1. 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:

Text

Description automatically generated

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

Text

Description automatically generated

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

Text

Description automatically generated

Provide **screenshots** of your statements and the results as evidence.

1. Perform the following tasks **using MongoDB queries**:
   1. List only the first 20 names of companies founded after the year 2010, ordered alphabetically.

Text

Description automatically generated

* 1. 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 with medium confidence

Provide **screenshots** of your statements and the results as evidence.

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

Text

Description automatically generated