Group #: Mon4pm\_Group4

Sujeeth Shetty

Sai Harshavardhan Bachina

Srisha Balaji

Siddharth Hareendran

**SQL-Mongo Project – Spatial Data of US Wildfires**

BUAN 6320

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Activity** | **Sujeeth Shetty** | **Sai Harshavardhan Bachina** | **Srisha Balaji** | **Siddharth Hareendran** |
| Prepared Data Model and Created Physical DB | x | x | x | x |
| Loaded Data into Database | x | x | x | x |
| Wrote SQL Queries | x | x | x | x |
| Prepared Mongo Database | x | x | x | x |
| Loaded data into Mongo DB | x | x | x | x |
| Wrote Mongo Queries | x | x | x | x |
| Prepared Report | x | x | x | x |
| Reviewed Report | x | x | x | x |

Contents

Data Model 5

Assumptions/Notes About Data Entities and Relationships 5

Entity-Relationship Diagram 6

Physical Database 7

Assumptions/Notes About Data Set 7

Screen shot of Physical Database objects 7

Data in the Database 15

SQL Queries 16

Query 1 16

Question 16

Notes/Comments About SQL Query and Results (Include # of Rows in Result) 16

Translation 16

Screen Shot of SQL Query and Results 16

Query 2 17

Question 17

Notes/Comments About SQL Query and Results (Include # of Rows in Result) 17

Translation 17

Screen Shot of SQL Query and Results 17

Query 3 18

Question 18

Notes/Comments About SQL Query and Results (Include # of Rows in Result) 18

Translation 18

Screen Shot of SQL Query and Results 18

Query 4 19

Question 19

Notes/Comments About SQL Query and Results (Include # of Rows in Result) 19

Translation 19

Screen Shot of SQL Query and Results 19

Query 5 20

Question 20

Notes/Comments About SQL Query and Results (Include # of Rows in Result) 20

Translation 20

Screen Shot of SQL Query and Results 20

Query 6 21

Question 21

Notes/Comments About SQL Query and Results (Include # of Rows in Result) 21

Translation 21

Screen Shot of SQL Query and Results 21

Data Review for MongoDB 22

Assumptions/Notes About Data Collections, Attributes and Relationships between Collections 22

Physical Mongo Database 23

Assumptions/Notes About Data Set 23

Screen shot of Physical Database objects (Database, Collections and Attributes) 23

Data in the Database 23

MongoDB Queries/Code 23

Query 1 24

Question 24

Notes/Comments About MongoDB Query/Code and Results (Include # of Documents in Result) 24

Translation 24

Screen Shot of MongoDB Query/Code and Results 24

Query 2 25

Question 25

Notes/Comments About MongoDB Query/Code and Results (Include # of Documents in Result) 25

Translation 25

Screen Shot of MongoDB Query/Code and Results 25

Query 3 26

Question 26

Notes/Comments About MongoDB Query/Code and Results (Include # of Documents in Result) 26

Translation 26

Screen Shot of MongoDB Query/Code and Results 26

Query 4 27

Question 27

Notes/Comments About MongoDB Query/Code and Results (Include # of Documents in Result) 27

Translation 27

Screen Shot of MongoDB Query/Code and Results 27

Query 5 28

Question 28

Notes/Comments About MongoDB Query/Code and Results (Include # of Documents in Result) 28

Translation 28

Screen Shot of MongoDB Query/Code and Results 28

Query 6 29

Question 29

Notes/Comments About MongoDB Query/Code and Results (Include # of Documents in Result) 29

Translation 29

Screen Shot of MongoDB Query/Code and Results 29

# Data Model

## Assumptions/Notes About Data Entities and Relationships

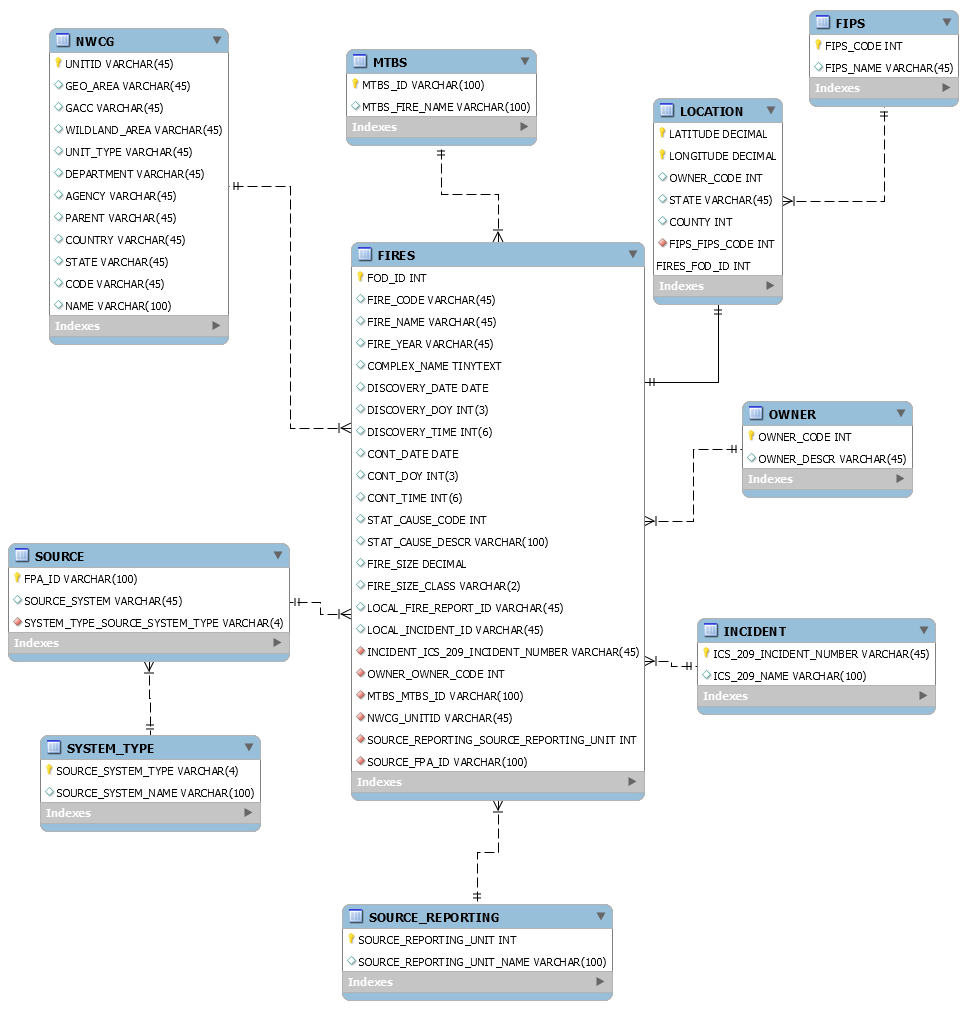
Include assumptions about data entities and their relationships with each other.

Include reasons why the data model is in 3NF.

Outliers:



## Entity-Relationship Diagram



# Physical Database

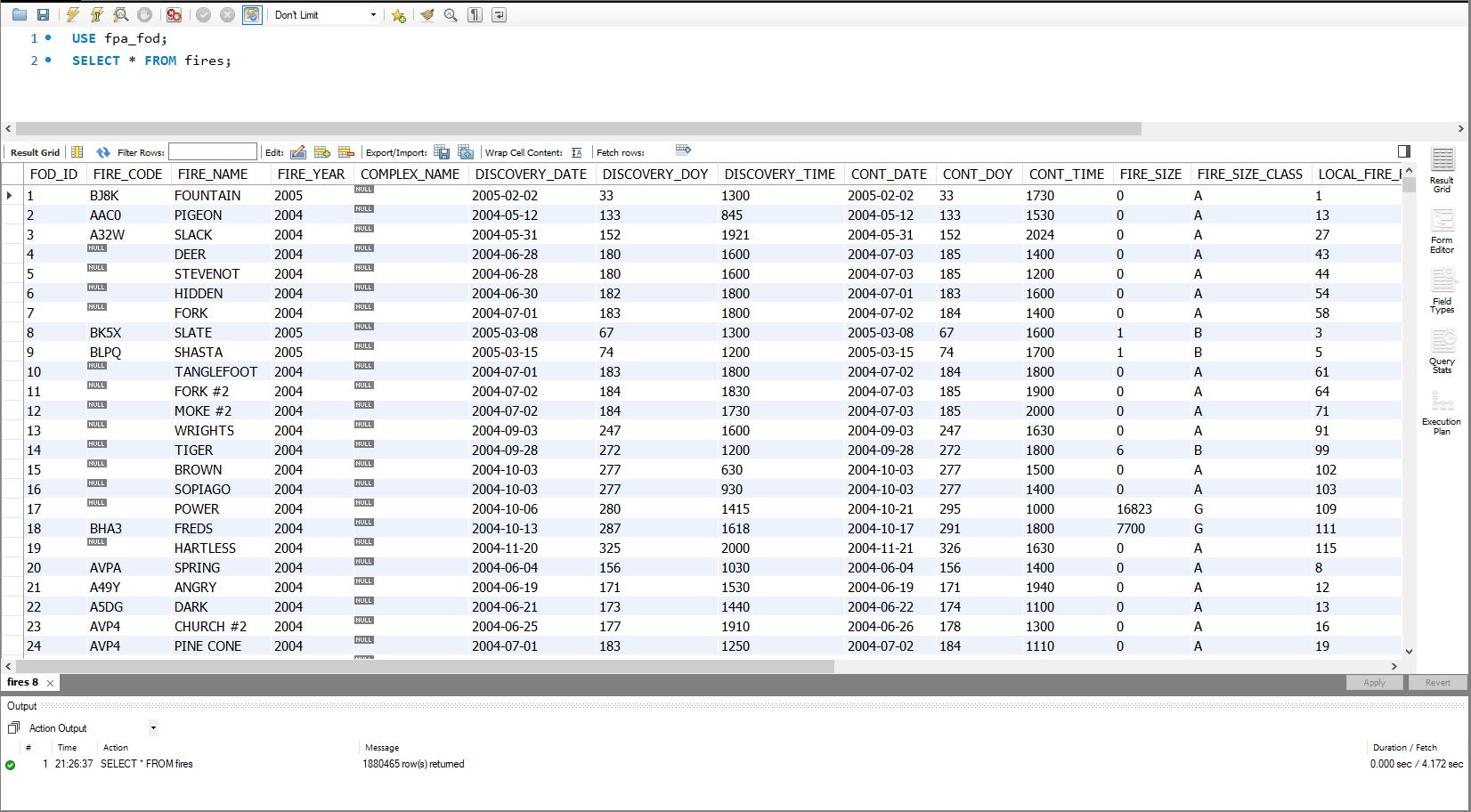
## Assumptions/Notes About Data Set

Include any assumptions made about data such as empty fields, sparse data, bad data, etc.

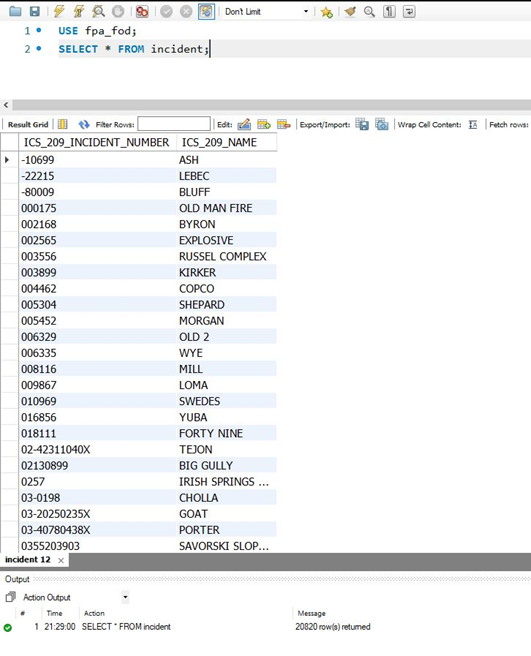
## Screen shot of Physical Database objects

**FIPS Table:**

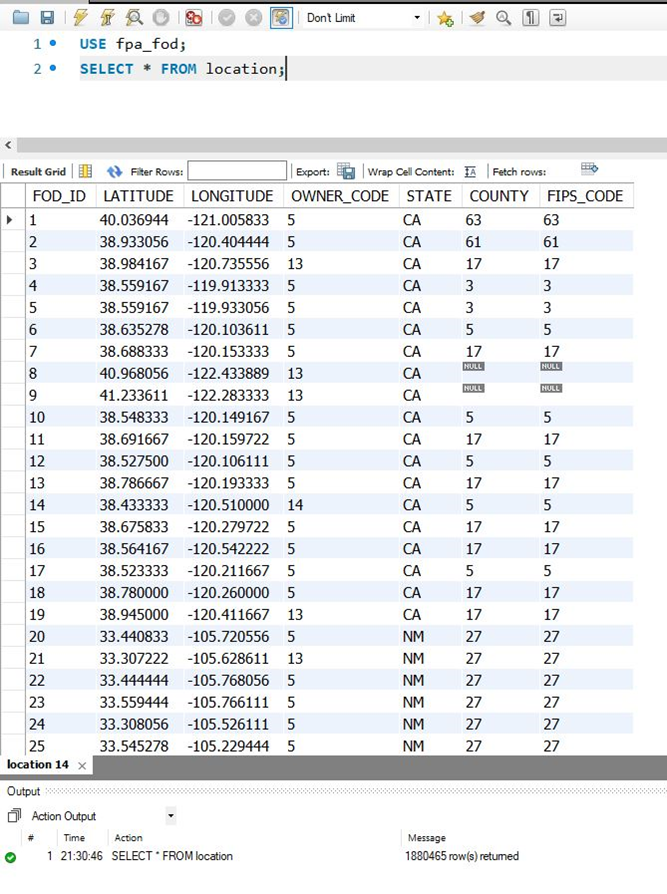


**FIRES Table:**

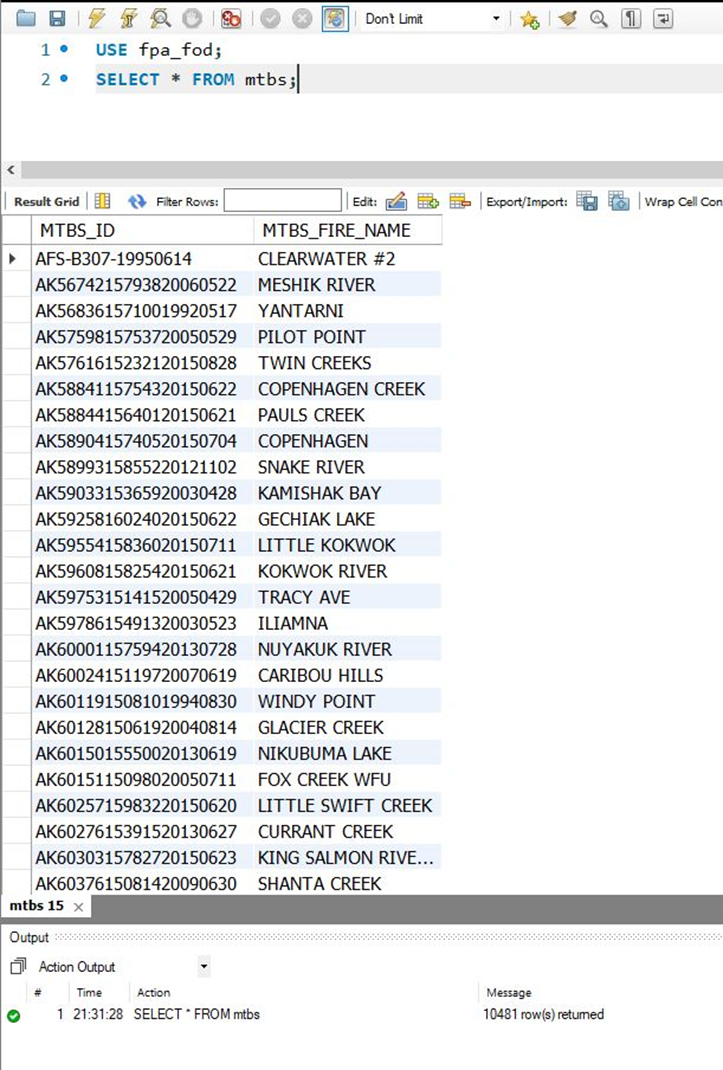
**INCIDENT Table:**

****

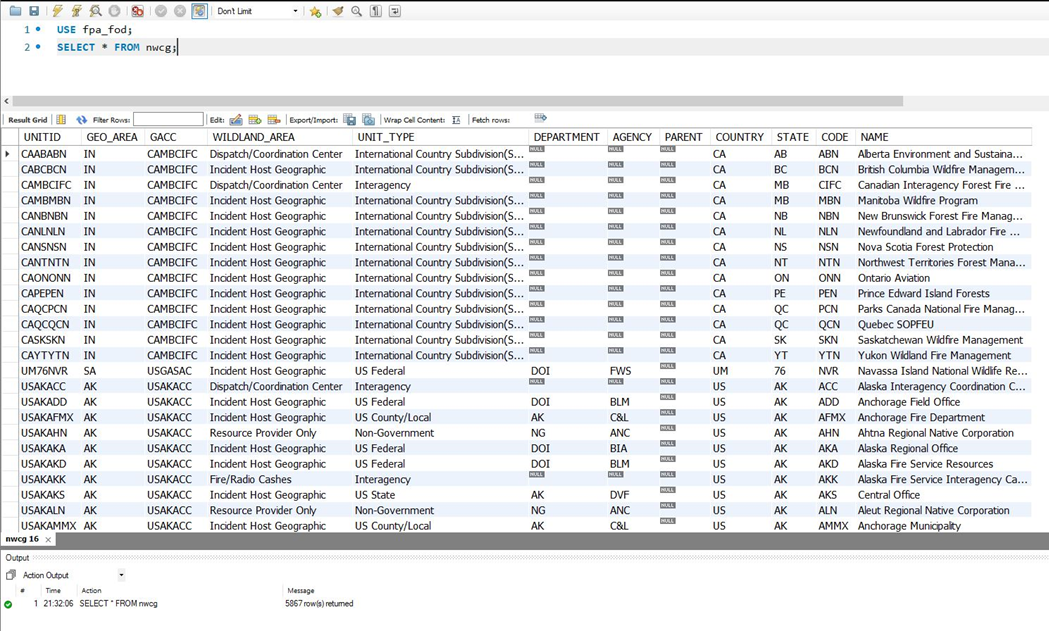
**Location Table:**

****

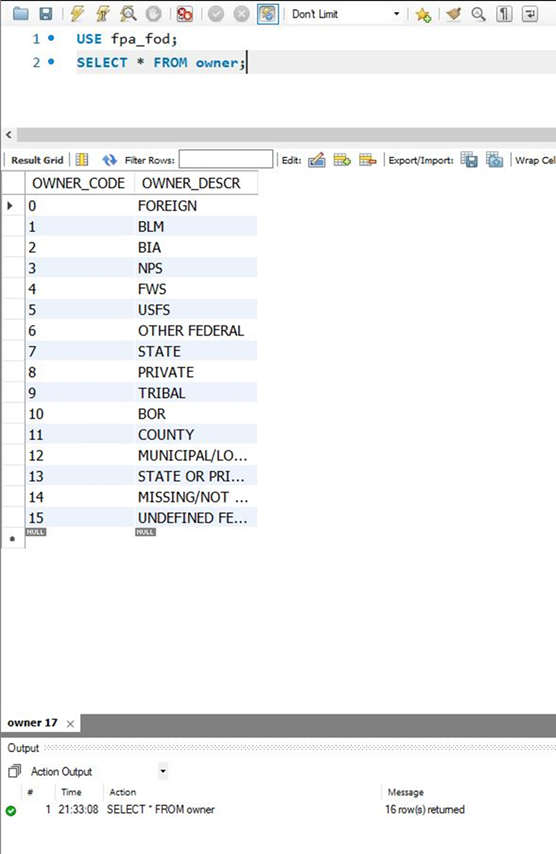
**MTBS Table:**

****

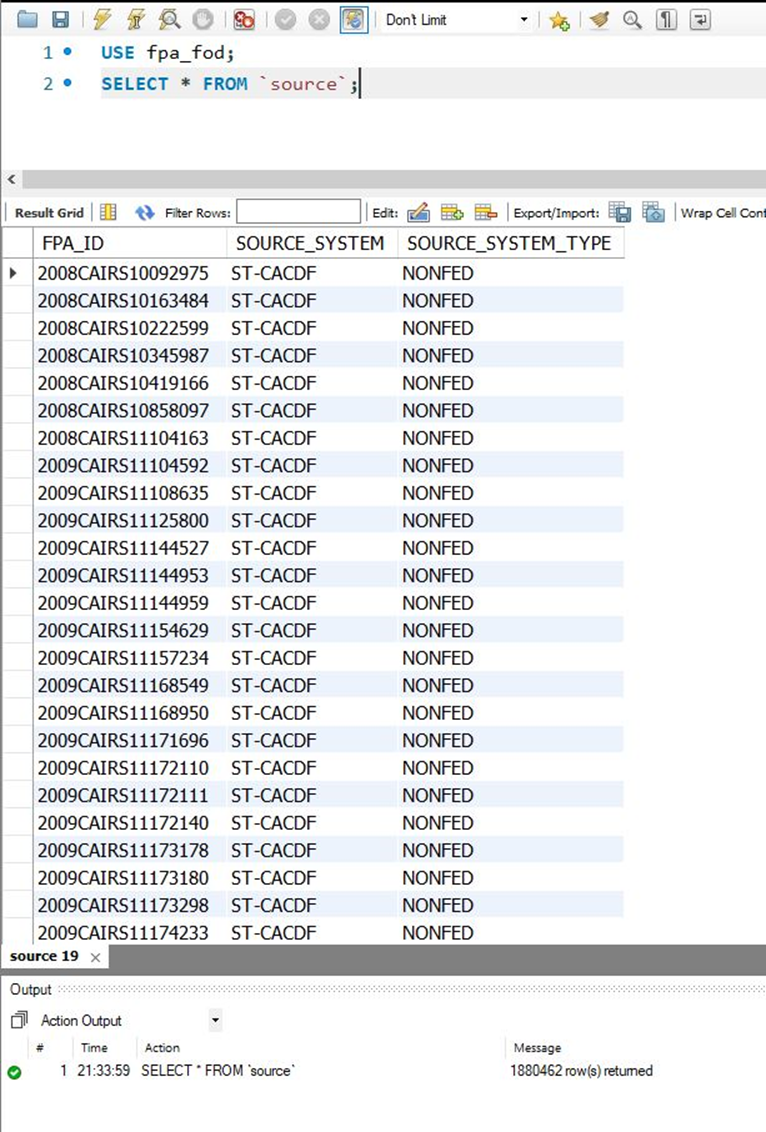
**NWCG Table:**

****

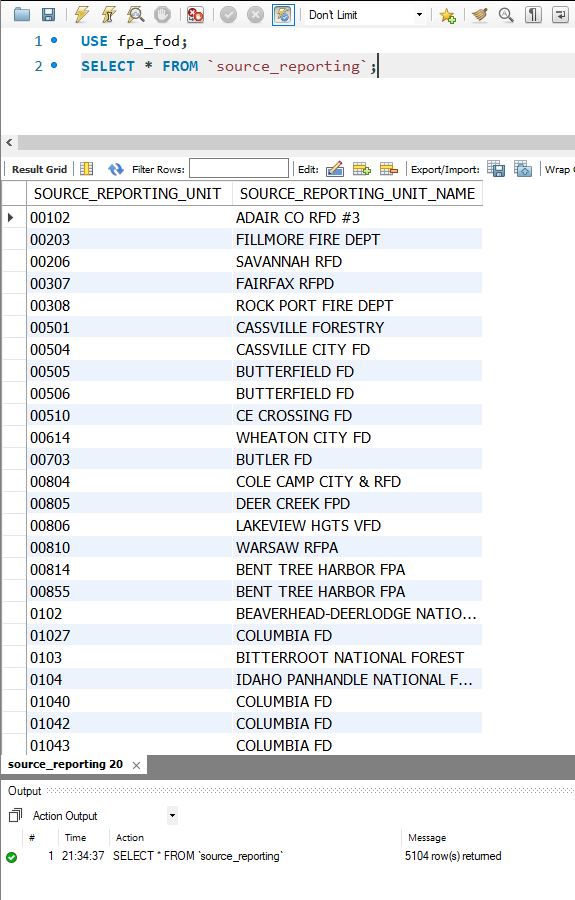
**OWNER Table:**

****

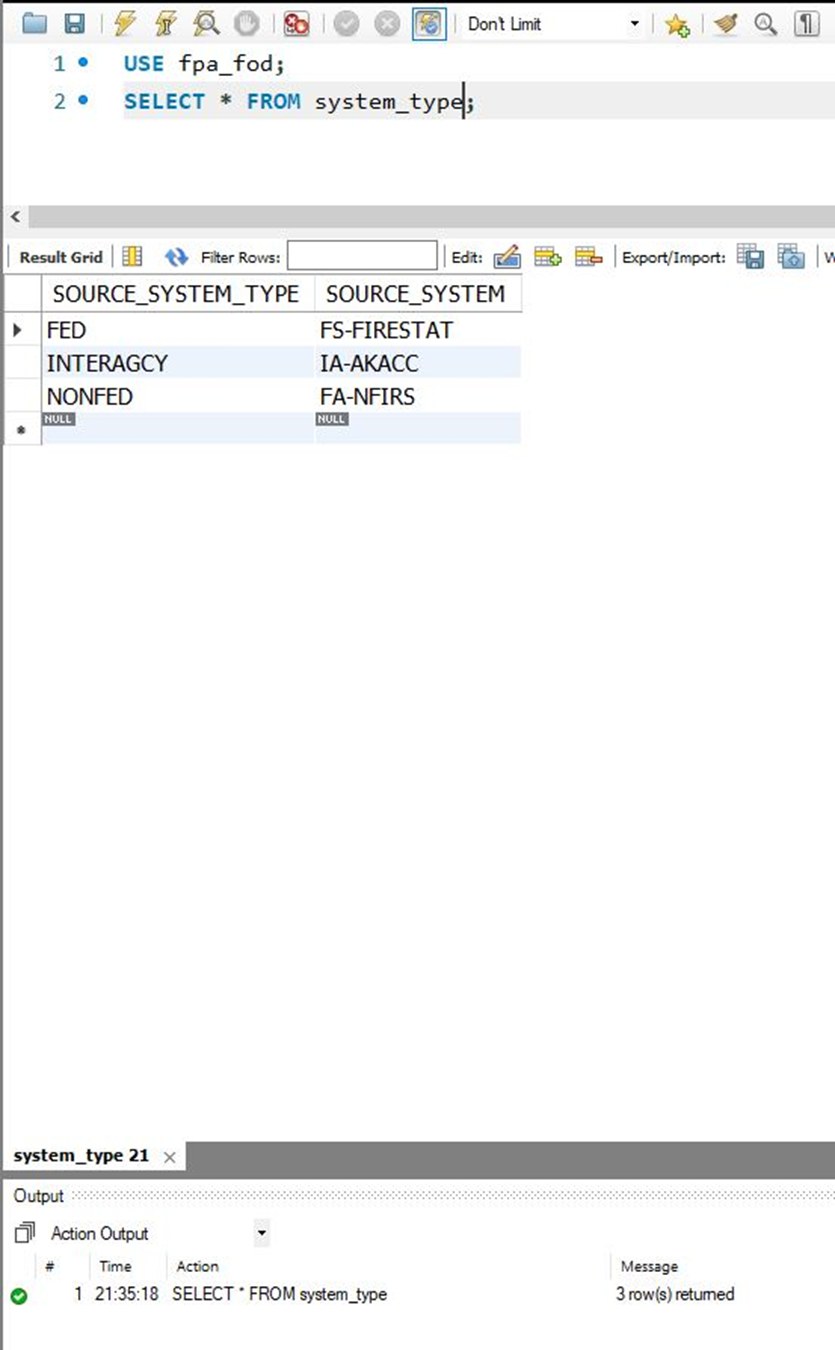
**Source Table:**

****

**SOURCE\_REPORTING Table:**



**SYSTEM\_TYPE Table:**

****

## 

## Data in the Database

|  |  |  |  |
| --- | --- | --- | --- |
| **Table Name** | **Primary Key** | **Foreign Key** | **# of Rows in Table** |
| FIRES | FOD\_ID | 1)INCIDENT\_ ICS\_209\_INCIDENT\_NUMBER  2)OWNER\_OWNER\_CODE  3)MTBS\_MTBS\_ID  4)NWCG\_UNITID  5)SOURCE\_REPORTING\_SOURCE\_REPORTING\_UNIT  6)SOURCE\_FPA\_ID | 1880465 |
| NWCG | UNITID | - | 5867 |
| MTBS | MTBS\_ID | - | 10481 |
| FIPS | FIPS\_CODE | - | 285 |
| SOURCE | FPA\_ID | SYSTEM\_TYPE\_SOURCE\_SYSTEM\_TYPE | 1880462 |
| SYSTEM\_TYPE | SOURCE\_SYSTEM\_TYPE | - | 3 |
| SOURCE\_REPORTING | SOURCE\_REPORTING\_UNIT | - | 5104 |
| INCIDENT | ICS\_209\_INCIDENT\_NUMBER | - | 20820 |
| OWNER | OWNER\_CODE | - | 16 |
| LOCATION | LATITUDE  LONGITUDE  (COMPOSITE KEY CONFIGURATION) | 1)FIPS\_FIPS\_CODE  2)FIRES\_FOD\_ID | 1880465 |

# SQL Queries

## Query 1

### Question

A leading beverage company has announced a billion-dollar fund for removing debris from forests, rivers and mountains in the US. All states are interested. Which 2 states have the least chance to win a share of the fund?

### Notes/Comments About SQL Query and Results (Include # of Rows in Result)

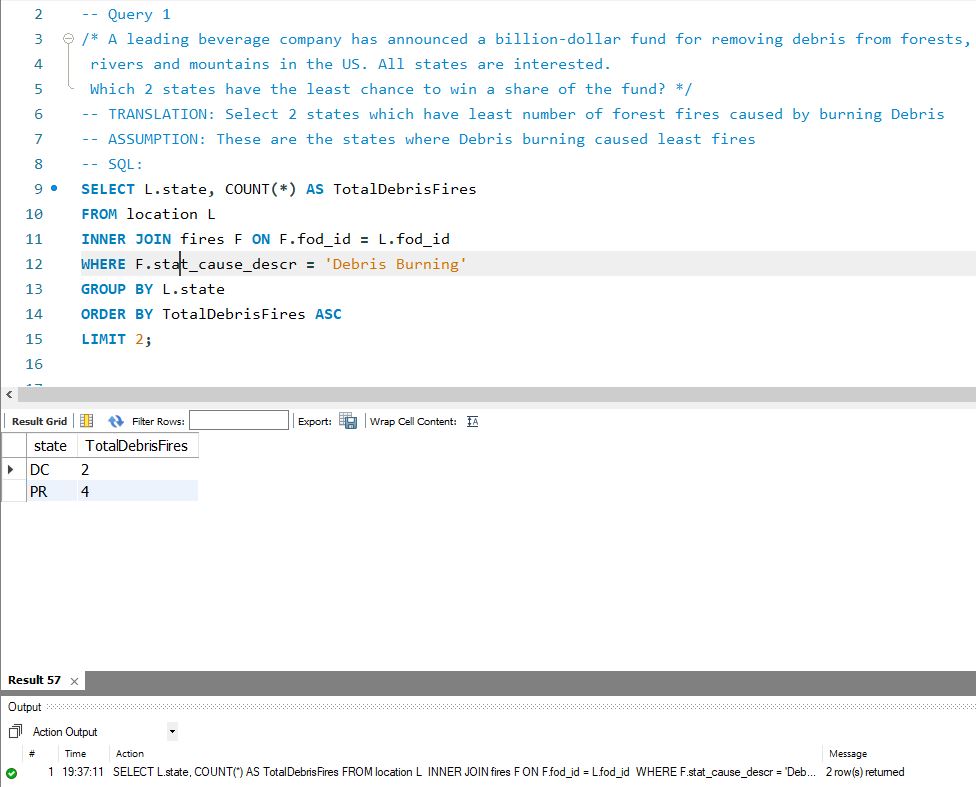
Assumption: These are the states where Debris burning caused least fires

Obtained 2 rows of Data.

### Translation

Select 2 states which have least number of forest fires caused by burning debris.

### Screen Shot of SQL Query and Results



## Query 2

### Question

One of the reporting agencies has suggested that children be banned from its forests unless there is one adult for every 4 children in a group visiting a forest. Name top 5 forests where this would be the least appropriate.

### Notes/Comments About SQL Query and Results (Include # of Rows in Result)

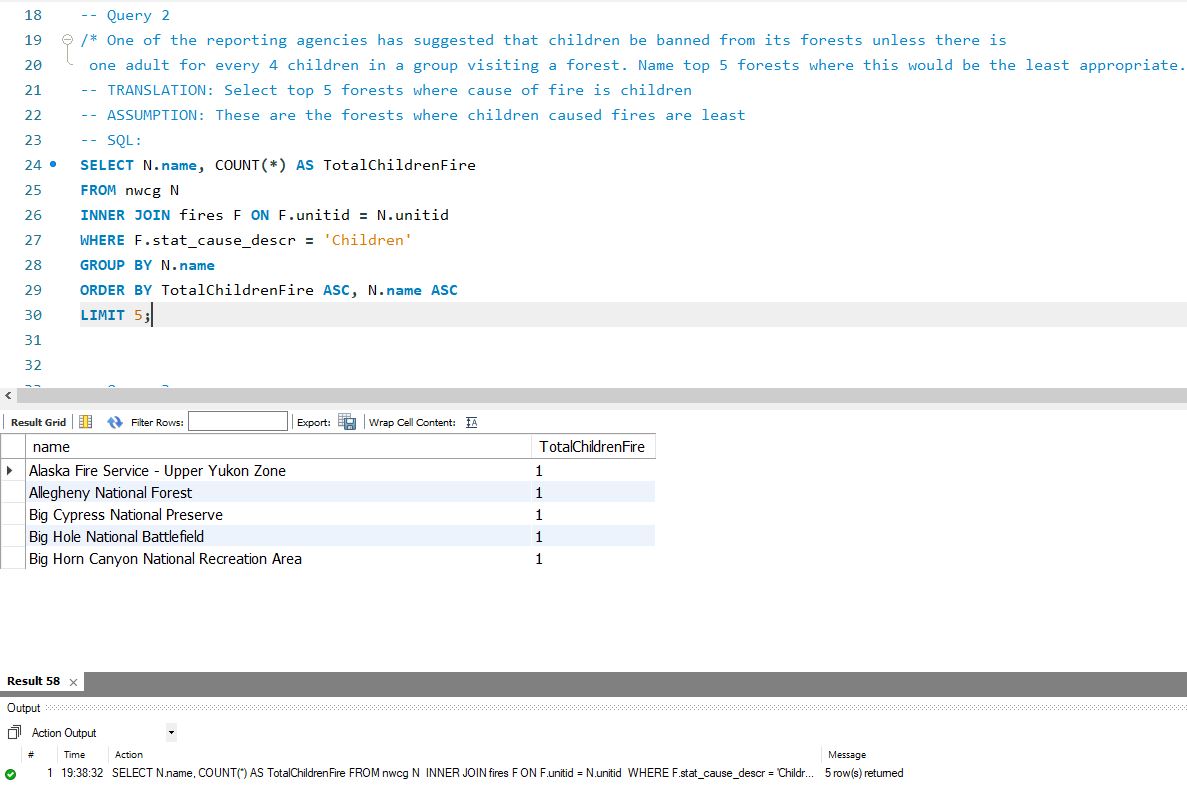
Assumption: These are the forests where children caused fires are the least

Obtained 5 rows of data.

### Translation

Select top 5 forests where cause of fire is children

### Screen Shot of SQL Query and Results



## Query 3

### Question

One advocacy group says human actions and nature are equally to blame for most wildfires. Write a query that can help determine the truth of this statement.

### Notes/Comments About SQL Query and Results (Include # of Rows in Result)

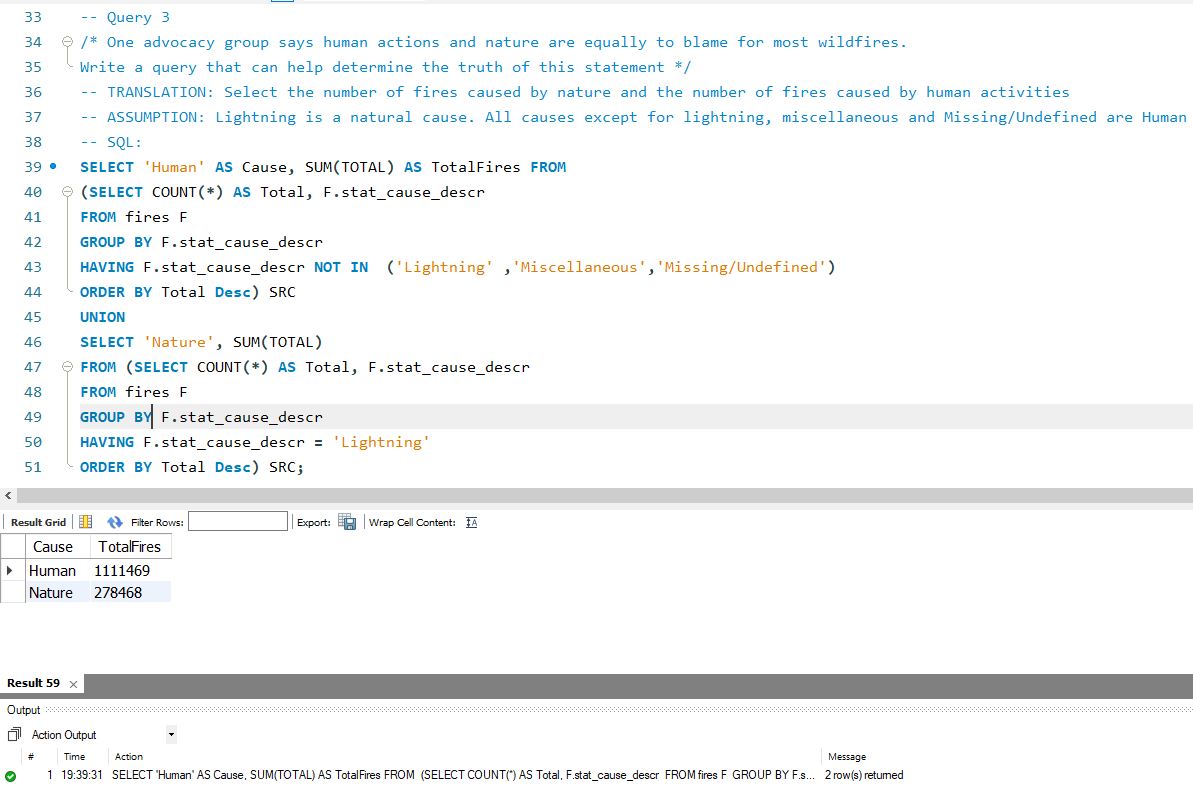
Assumption: Lightning is a natural cause. All causes except for lightning, miscellaneous and Missing/Undefined are Human.

Obtained 2 Rows of data with columns, cause and total fires

### Translation

Select the number of fires cause by nature and the number of fires caused by human activities

### Screen Shot of SQL Query and Results



## Query 4

### Question

What were the forests that had only one fire that lasted more than two days?

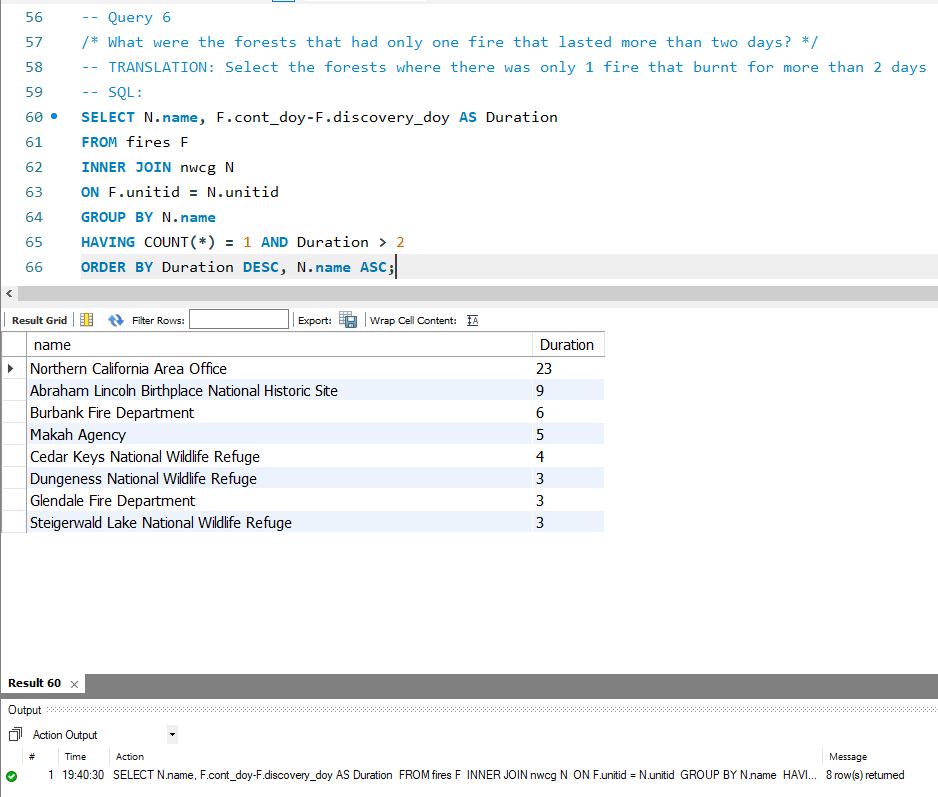
### Notes/Comments About SQL Query and Results (Include # of Rows in Result)

Obtained 8 rows of data.

### Translation

Select the forests where there was only 1 fire that lasted for more than two days.

### Screen Shot of SQL Query and Results



## Query 5

### Question

Which state had fires only in the second half of the calendar years?

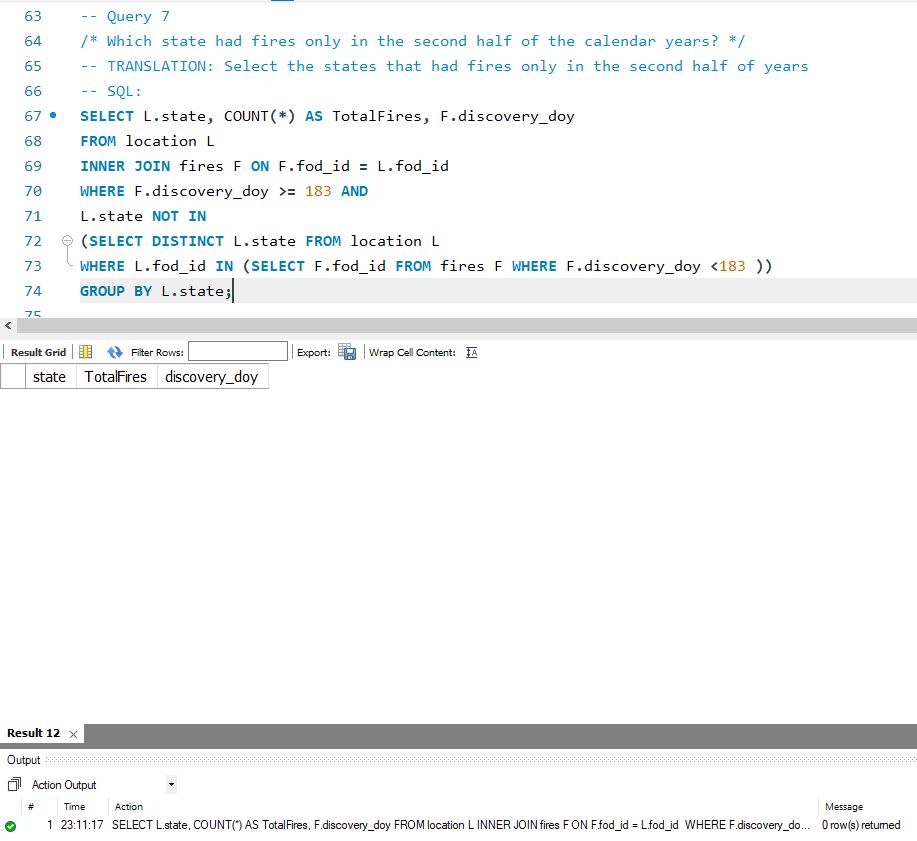
### Notes/Comments About SQL Query and Results (Include # of Rows in Result)

Returned Null. 0 records.

### Translation

Select the states that had fire only in the second half of the year

### Screen Shot of SQL Query and Results



## Query 6

### Question

Which forest had the number of fires equal to the average number of wild fires in the US?

### Notes/Comments About SQL Query and Results (Include # of Rows in Result)

Assumption: Average is the average number of wild fires per year in the US

Assumption: Since there was 0 forests with average number of fires, we set a range around the average.

Obtained 1 records.

### Translation

Select the forest that have fires equal to average number of fires per state in US

### Screen Shot of SQL Query and Results



# Data Review for MongoDB

## Assumptions/Notes About Data Collections, Attributes and Relationships between Collections

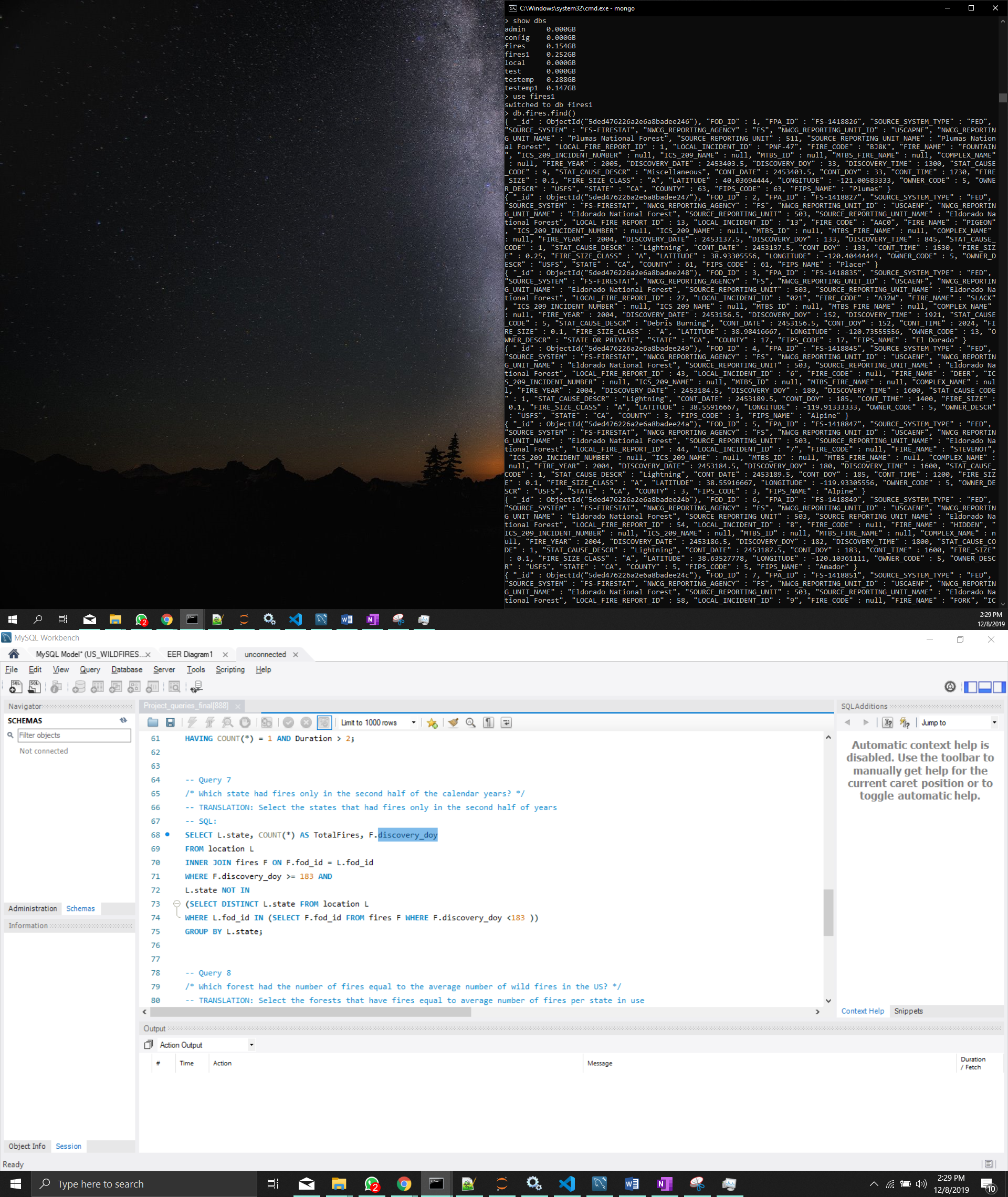
1. The database implemented here is a denormalized database. Since MongoDB is inherently unstructured and doesn’t require structured normalized database, we have implemented a denormalized database resulting in a single collection.
2. All queries implemented in this project is implemented on a single collection ‘fires’.

# Physical Mongo Database

## Assumptions/Notes About Data Set

All queries implemented in this project is implemented on a single collection ‘fires’.

## Screen shot of Physical Database objects (Database, Collections and Attributes)



## Data in the Database

|  |  |  |
| --- | --- | --- |
| **Collection Name** | **Relationshps With Other Collections (if any)** | **# of Documents in Collection** |
| fires | - | 1880465 |

# MongoDB Queries/Code

## Query 1

### Question

A leading beverage company has announced a billion-dollar fund for removing debris from forests, rivers and mountains in the US. All states are interested. Which 2 states have the least chance to win a share of the fund?

### Notes/Comments About MongoDB Query/Code and Results (Include # of Documents in Result)

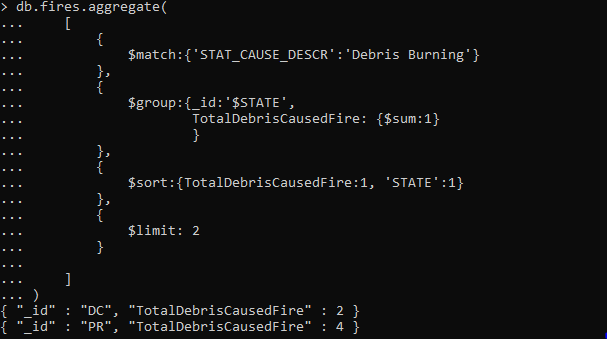
Assumption: These are the states where Debris burning caused least fires

Query returned 2 documents.

### Translation

Select 2 states which have least number of forest fires caused by burning debris.

### Screen Shot of MongoDB Query/Code and Results



## Query 2

### Question

One of the reporting agencies has suggested that children be banned from its forests unless there is one adult for every 4 children in a group visiting a forest. Name top 5 forests where this would be the least appropriate.

### Notes/Comments About MongoDB Query/Code and Results (Include # of Documents in Result)

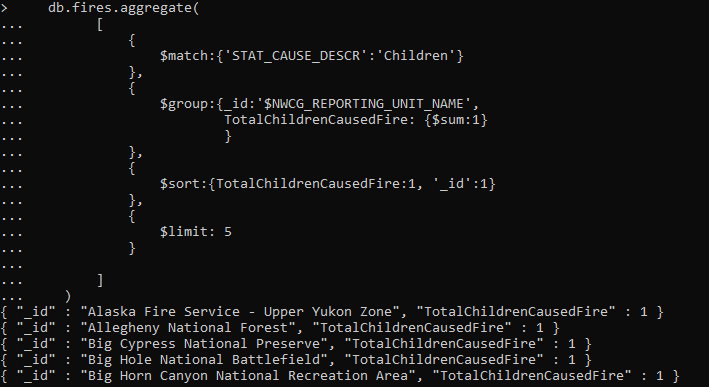
Assumption: These are the forests where children caused fires are least

Query returned 5 documents.

### Translation

Select top 5 forests where cause of fire is children

### Screen Shot of MongoDB Query/Code and Results



## Query 3

### Question

One advocacy group says human actions and nature are equally to blame for most wildfires. Write a query that can help determine the truth of this statement.

### Notes/Comments About MongoDB Query/Code and Results (Include # of Documents in Result)

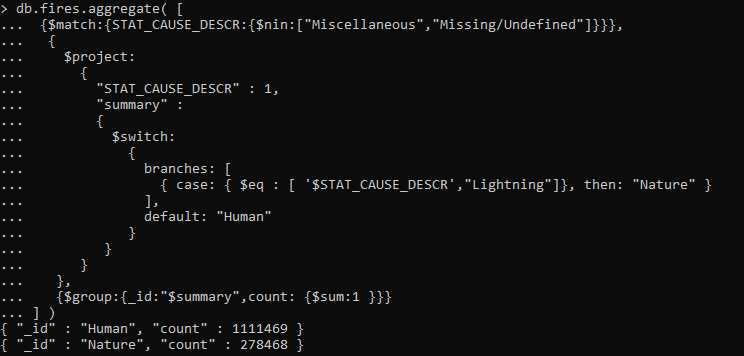
Assumption: Lightning is a natural cause. All causes except for lightning, miscellaneous and Missing/Undefined are Human

Query returned 2 documents

### Translation

Select the number of fires cause by nature and the number of fires caused by human activities

### Screen Shot of MongoDB Query/Code and Results



## Query 4

### Question

What were the forests that had only one fire that lasted more than two days?

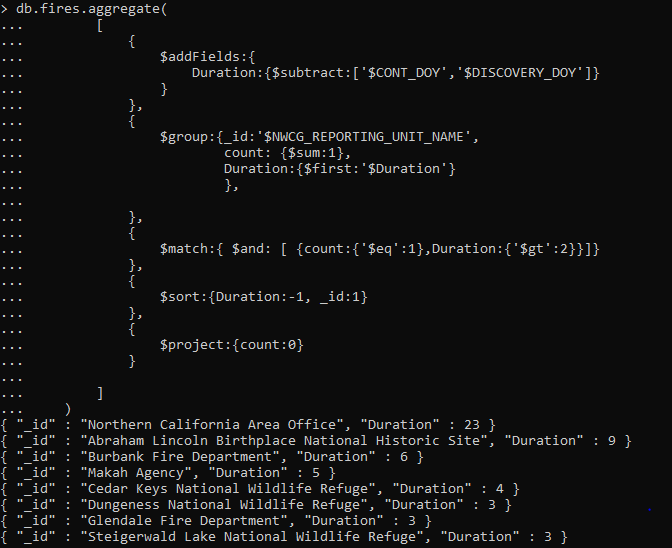
### Notes/Comments About MongoDB Query/Code and Results (Include # of Documents in Result)

Query retuned 8 documents

### Translation

Select the forests where there was only 1 fire that lasted for more than two days.

### Screen Shot of MongoDB Query/Code and Results



## Query 5

### Question

Which state had fires only in the second half of the calendar years?

### Notes/Comments About MongoDB Query/Code and Results (Include # of Documents in Result)

Query returned 0 documents

### Translation

Select the states that had fire only in the second half of the year

### Screen Shot of MongoDB Query/Code and Results



## Query 6

### Question

Which forest had the number of fires equal to the average number of wild fires in the US?

### Notes/Comments About MongoDB Query/Code and Results (Include # of Documents in Result)

Assumption: Average is the average number of wild fires per state in the US

Query returned 23 documents

### Translation

Select the forest that have fires equal to average number of fires per state in US

### Screen Shot of MongoDB Query/Code and Results

## 