



SQL-Mongo Project – Spatial Data of US Wildfires

BUAN 6320

Submitted By:
Rishi Bhatia

Contents

Data Model	4
Assumptions/Notes About Data Entities and Relationships	4
Entity-Relationship Diagram	4
Physical Database	5
Assumptions/Notes About Data Set	5
Screen shot of Physical Database objects.....	5
Data in the Database.....	7
SQL Queries.....	12
Query 1.....	12
Question.....	12
Notes/Comments About SQL Query and Results (Include # of Rows in Result).....	12
Translation	12
Screen Shot of SQL Query and Results.....	12
Query 2.....	13
Question.....	13
Notes/Comments About SQL Query and Results (Include # of Rows in Result).....	13
Translation	13
Screen Shot of SQL Query and Results.....	13
Query 3.....	14
Question.....	14
Notes/Comments About SQL Query and Results (Include # of Rows in Result).....	14
Translation	14
Screen Shot of SQL Query and Results.....	14
Query 4.....	15
Question.....	15
Notes/Comments About SQL Query and Results (Include # of Rows in Result).....	15
Translation	15
Screen Shot of SQL Query and Results.....	15
Query 5.....	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 6.....	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
Data Review for MongoDB.....	18
Assumptions/Notes About Data Collections, Attributes and Relationships between Collections	18
Physical Mongo Database	19
Assumptions/Notes About Data Set	19
Screen shot of Physical Database objects (Database, Collections and Attributes)	19
Data in the Database.....	24
MongoDB Queries/Code	25
Query 1.....	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 2.....	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 3.....	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 4.....	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

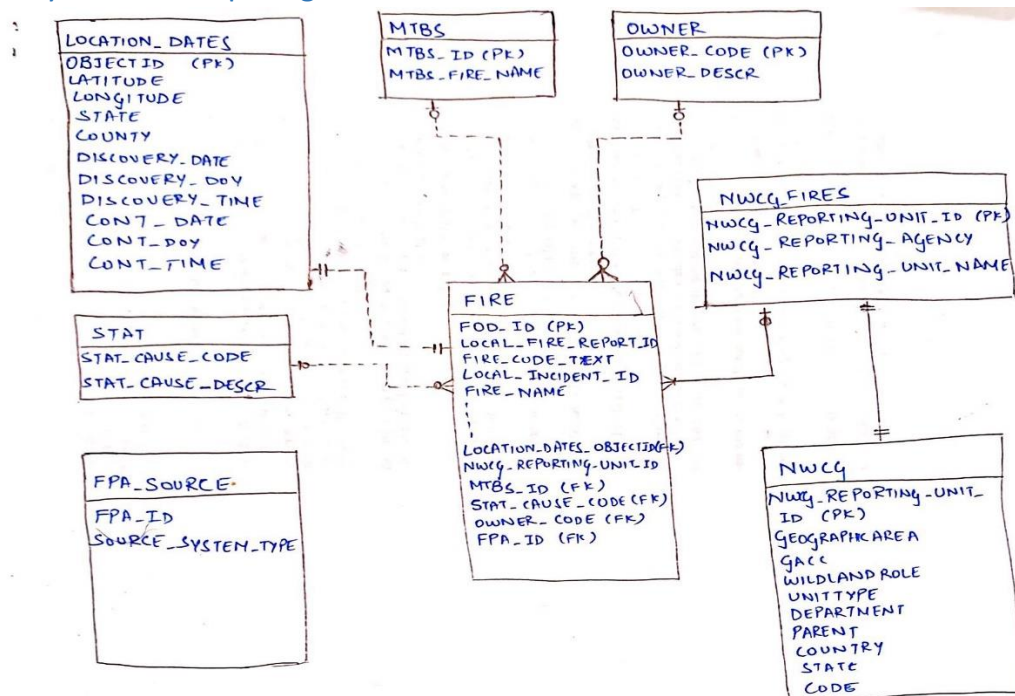
Query 5.....	29
Question.....	29
Notes/Comments About MongoDB Query/Code and Results (Include # of Documents in Result) ...	29
Translation	29
Screen Shot of MongoDB Query/Code and Results.....	29
Query 6.....	30
Question.....	30
Notes/Comments About MongoDB Query/Code and Results (Include # of Documents in Result) ...	30
Translation	30
Screen Shot of MongoDB Query/Code and Results.....	30

Data Model

Assumptions/Notes About Data Entities and Relationships

1. As FOD_ID is the global unique identifier and is non-null and unique, it is assumed to be the primary key of the main table. There are other foreign keys like STAT_CAUSE_ID, OWNER_CODE, MTBS_ID, NWCG_REPORTING_UNIT_ID, FPA_ID, OBJECTID which are the primary keys in other tables.
2. In Stat_Cause_Table, stat_cause_id is assumed to be the primary key as each key represent an individual stat_cause_descr, the stat_cause_id is unique and non_null and stat_cause_descr is directly dependant on the stat_cause_id.
3. In the Owner table, owner_code is assumed to be the primary key because it is non-null, unique and all the element of owner_descr depend directly on owner_code.
4. In the MTBS table, MTBS_ID directly depends on each unique MTBS_NAME. MTBS_ID is the primary key for this table because it is unique and non-null and every record of MTBS_Fire_Name directly depends on MTBS_ID.
5. The NWCG table is assumed to be a separate table because there are a lot of records in Unit_ID column, all of which are not used in the main table. NWCG table serves the purpose of all the records given by NWCG Reporting and all the columns from NWCG_FIRES, and NWCG are directly related to NWCG_REPORTING_Unit_ID. NWCG_REPORTING_Unit_ID, the primary key of the table is non-null and unique.
6. In the FPA_SOURCE table, FPA_ID is the primary key because all the other columns directly depend on the FPA_ID and the SOURCE_REPORTING_UNIT_ID cannot be a primary key because of its redundant values.

Entity-Relationship Diagram

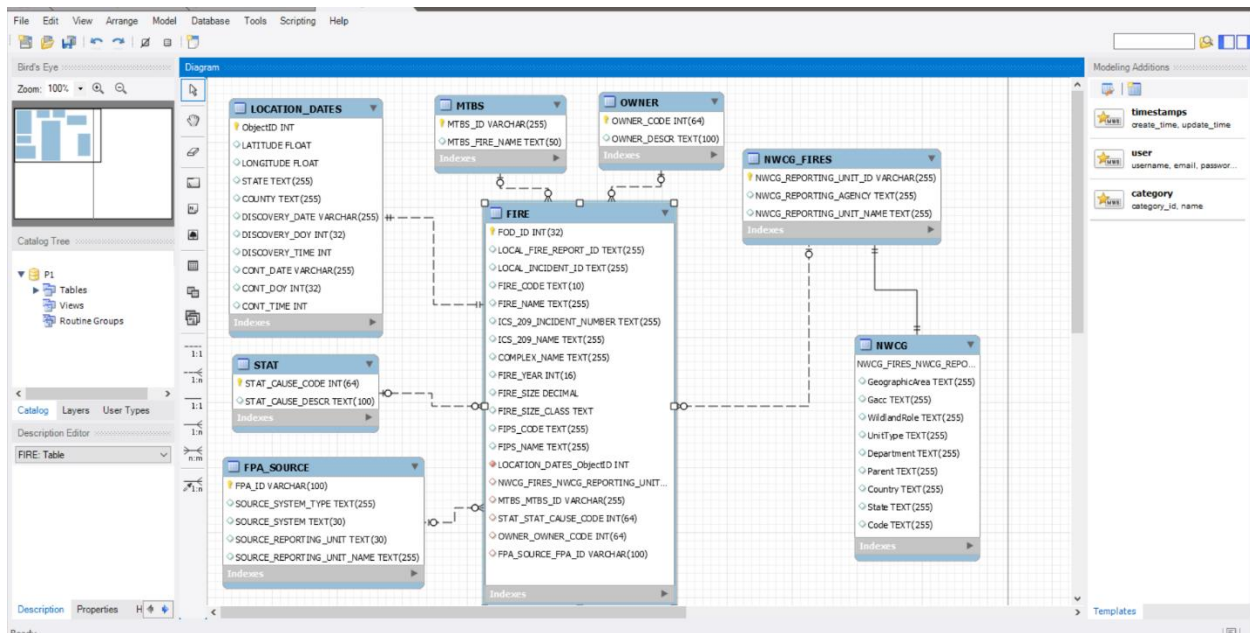


Physical Database

Assumptions/Notes About Data Set

1. Shape column was removed because there was no relation of the column with the primary key and the data type was not of standard data types in SQL.
2. Source_Reporting_Unit could be a primary key in the Source Code table with the four columns namely Source_Reporting_Unit, Source_Reporting_Unit_Name, Source_System_Type, Source_System. But while examining the data, there were few duplicated values in Source_Reporting_Unit with different matching values in other columns, due to which this table was separated but with FPA_ID which showed the relationship between all the columns of the table to the primary key, FPA_ID
3. FIPS_CODE could be the primary key for the FIPS table with the columns namely, FIPS_CODE, FIPS_NAME. But after cross-examining the data, there were some duplicated values in FIPS_CODE, which mapped to two different FIPS_NAME. Thus, FIPS table could not be separated.
4. NWCG and NWCG_FIRES need to be separated because they all are connected with a common primary key, i.e. NWCG_REPORTING_UNIT_ID.
5. All the state, county, longitude, latitude, and dates of discovery with contained can be placed in one table that is LOCATION_DATES where the primary key is the ObjectID.

Screen shot of Physical Database objects



Data in the Database

Table Name	Primary Key	Foreign Key	# of Rows in Table
FIRE	FOD_ID	LOCATION_DATE_OBJECTID NWCG_FIRES_NWCG_REPORTING_UNIT_ID MTBS_MTB_ID STAT_STAT_CAUSE_CODE OWNER_OWNER_CODE FPA_SOURCE_FPA_ID	1880456
FPA_SOURCE	FPA_ID	-	1880462
LOCATION_DATES	OBJECT_ID	-	1880456
MTBS	MTBS_ID	-	10481
NWCG	NWCG_REPORTING_UNIT_ID	NWCG_REPORTING_UNITID	5885
NWCG_FIRES	NWCG_REPORTING_UNIT_ID	-	5885
OWNER	OWNER_ID	-	16
STAT	STAT_CAUSE_CODE	-	13

The screenshot displays the SQL Developer environment. The left sidebar shows the 'SCHEMAS' tree with 'p3' expanded, listing tables like 'fire', 'fpa_source', 'location_dates', 'mtbs', 'nwcg', 'nwcg_fires', 'owner', 'stat', and 'views'. The main window shows a SQL script with two statements: 'use p3;' and 'select count(*) as Fire_Count from fire;'. The 'Result Grid' shows a single row with the value '1880456'. The bottom pane shows the 'Output' window with a log of actions and their durations.

SQL Script:

```

1 • use p3;
2 • select count(*) as Fire_Count from fire;

```

Result Grid:

Fire_Count
1880456

Output Log:

#	Time	Action	Message	Duration / Fetch
1	12:19:44	use p3	0 row(s) affected	0.000 sec
2	12:20:00	select * from fire	1880456 row(s) returned	0.015 sec / 5.594 sec
3	12:20:39	select count(*) from fire	1 row(s) returned	15.812 sec / 0.000 sec
4	12:21:04	select count(*) as Fire_Count from fire	1 row(s) returned	16.375 sec / 0.000 sec

File Edit View Query Database Server Tools Scripting Help

Navigator

SCHEMAS

Filter objects

- bowlingleagueexample
- employees
- entertainmentagencyexample
- p3
 - Tables
 - fire
 - fpa_source
 - location_dates
 - mtbs
 - nwcg
 - nwcg_fires
 - owner
 - stat
 - Views
 - Stored Procedures
 - Functions
- recipeseexample
- salesorderexample
- schoolscheduleexample
- test_db

Administration Schemas

Information

No object selected

SQL File p3

```

1 use p3;
2 select count(*) as FPA_SOURCE_COUNT from fpa_source;

```

Result Grid

FPA_SOURCE_COUNT
1880462

Output

Action Output

#	Time	Action	Message	Duration / Fetch
1	12:33:47	use p3	0 row(s) affected	0.000 sec
2	12:33:49	select count(*) as FPA_SOURCE_COUNT from fpa_source	Error Code: 2013. Lost connection to MySQL server during query	30.016 sec
3	12:36:04	SET GLOBAL interactive_timeout=60	0 row(s) affected	0.000 sec
4	12:36:08	select count(*) as FPA_SOURCE_COUNT from fpa_source	Error Code: 2013. Lost connection to MySQL server during query	30.000 sec
5	12:37:01	select count(*) as FPA_SOURCE_COUNT from fpa_source	1 row(s) returned	36.969 sec / 0.000 sec

SQLAdditions

Automatic context help is disabled. Use the toolbar to manually get help for the current caret position or to toggle automatic help.

Object Info Session

File Edit View Query Database Server Tools Scripting Help

Navigator

SCHEMAS

Filter objects

- bowlingleagueexample
- employees
- entertainmentagencyexample
- p3
 - Tables
 - fire
 - fpa_source
 - location_dates
 - mtbs
 - nwcg
 - nwcg_fires
 - owner
 - stat
 - Views
 - Stored Procedures
 - Functions
- recipeseexample
- salesorderexample
- schoolscheduleexample
- test_db

Administration Schemas

Information

No object selected

SQL File p3

```

1 use p3;
2 select count(*) as LOCATION_DATES_COUNT from location_dates;

```

Result Grid

LOCATION_DATES_COUNT
1880465

Output

Action Output

#	Time	Action	Message	Duration / Fetch
1	12:33:47	use p3	0 row(s) affected	0.000 sec
2	12:33:49	select count(*) as FPA_SOURCE_COUNT from fpa_source	Error Code: 2013. Lost connection to MySQL server during query	30.016 sec
3	12:36:04	SET GLOBAL interactive_timeout=60	0 row(s) affected	0.000 sec
4	12:36:08	select count(*) as FPA_SOURCE_COUNT from fpa_source	Error Code: 2013. Lost connection to MySQL server during query	30.000 sec
5	12:37:01	select count(*) as FPA_SOURCE_COUNT from fpa_source	1 row(s) returned	36.969 sec / 0.000 sec
6	12:38:32	select count(*) as LOCATION_DATES_COUNT from location_dates	1 row(s) returned	22.172 sec / 0.000 sec

SQLAdditions

Automatic context help is disabled. Use the toolbar to manually get help for the current caret position or to toggle automatic help.

Object Info Session

File Edit View Query Database Server Tools Scripting Help

Navigator

SCHEMAS

Filter objects

- bowlingleagueexample
- employees
- entertainmentagencyexample
- p3
 - Tables
 - fire
 - fpa_source
 - location_dates
 - mtbs
 - nwcg
 - nwcg_fires
 - owner
 - stat
 - Views
 - Stored Procedures
 - Functions
- recipiesexample
- salesordersexample
- schoolschedulngexample
- test_db

Administration Schemas

Information

No object selected

SQL File 0*

1 use p3;

2 select count(*) as MTBS_COUNT from mtbs;

Result Grid

	MTBS_COUNT
1	10481

Output

Action Output

#	Time	Action	Message	Duration / Fetch
2	12:33:49	select count(*) as FPA_SOURCE_COUNT from fpa_source	Error Code: 2013. Lost connection to MySQL server during query	30.016 sec
3	12:36:04	SET GLOBAL Interactive_timeout=60	0 row(s) affected	0.000 sec
4	12:36:08	select count(*) as FPA_SOURCE_COUNT from fpa_source	Error Code: 2013. Lost connection to MySQL server during query	30.000 sec
5	12:37:01	select count(*) as FPA_SOURCE_COUNT from fpa_source	1 row(s) returned	36.969 sec / 0.000 sec
6	12:38:32	select count(*) as LOCATION_DATES_COUNT from location_dates	1 row(s) returned	22.172 sec / 0.000 sec
7	12:39:39	select count(*) as MTBS_COUNT from mtbs	1 row(s) returned	1.265 sec / 0.000 sec

Object Info Session

SQLAdditions

Automatic context help is disabled. Use the toolbar to manually get help for the current caret position or to toggle automatic help.

File Edit View Query Database Server Tools Scripting Help

Navigator

SCHEMAS

Filter objects

- bowlingleagueexample
- employees
- entertainmentagencyexample
- p3
 - Tables
 - fire
 - fpa_source
 - location_dates
 - mtbs
 - nwcg
 - nwcg_fires
 - owner
 - stat
 - Views
 - Stored Procedures
 - Functions
- recipiesexample
- salesordersexample
- schoolschedulngexample
- test_db

Administration Schemas

Information

No object selected

SQL File 0*

1 use p3;

2 select count(*) as NWCG_COUNT from nwcg;

Result Grid

	NWCG_COUNT
1	5885

Output

Action Output

#	Time	Action	Message	Duration / Fetch
4	12:36:08	select count(*) as FPA_SOURCE_COUNT from fpa_source	Error Code: 2013. Lost connection to MySQL server during query	30.000 sec
5	12:37:01	select count(*) as FPA_SOURCE_COUNT from fpa_source	1 row(s) returned	36.969 sec / 0.000 sec
6	12:38:32	select count(*) as LOCATION_DATES_COUNT from location_dates	1 row(s) returned	22.172 sec / 0.000 sec
7	12:39:39	select count(*) as MTBS_COUNT from mtbs	1 row(s) returned	1.265 sec / 0.000 sec
8	12:39:56	select count(*) as NWCG_COUNT from nwcg	1 row(s) returned	0.000 sec / 0.000 sec
9	12:40:08	select count(*) as NWCG_COUNT from nwcg	1 row(s) returned	0.781 sec / 0.000 sec

Object Info Session

SQLAdditions

Automatic context help is disabled. Use the toolbar to manually get help for the current caret position or to toggle automatic help.

File Edit View Query Database Server Tools Scripting Help

Navigator

SCHEMAS

Filter objects

- bowlingleagueexample
- employees
- entertainmentagencyexample
- p3
 - Tables
 - fire
 - fpa_source
 - location_dates
 - mtbs
 - nwcg
 - nwcg_fires
 - owner
 - stat
 - Views
 - Stored Procedures
 - Functions
 - recipeseexample
 - salesorderexample
 - schoolsschedulingexample
 - test_db

Administration Schemas

Information

No object selected

SQL File 0*

```

1 use p3;
2 select count(*) as NWCG_FIRES_COUNT from nwcg_fires;

```

Result Grid

	NWCG_FIRES_COUNT
1	5885

Output

Action Output

#	Time	Action	Message	Duration / Fetch
7	12:39:39	select count(*) as MTBS_COUNT from mtbs	1 row(s) returned	1.265 sec / 0.000 sec
8	12:39:56	select count(*) as NWCG_COUNT from mtbs	1 row(s) returned	0.000 sec / 0.000 sec
9	12:40:08	select count(*) as NWCG_COUNT from nwcg	1 row(s) returned	0.781 sec / 0.000 sec
10	12:40:49	select count(*) as NWCG_FIRES_COUNT from nwcg_fires	1 row(s) returned	0.531 sec / 0.000 sec
11	12:40:57	select count(*) as NWCG_FIRES_COUNT from nwcg	1 row(s) returned	0.015 sec / 0.000 sec
12	12:41:05	select count(*) as NWCG_FIRES_COUNT from nwcg_fires	1 row(s) returned	0.000 sec / 0.000 sec

SQLAdditions

Automatic context help is disabled. Use the toolbar to manually get help for the current caret position or to toggle automatic help.

Read Only Context Help Snippets

Object Info Session

File Edit View Query Database Server Tools Scripting Help

Navigator

SCHEMAS

Filter objects

- bowlingleagueexample
- employees
- entertainmentagencyexample
- p3
 - Tables
 - fire
 - fpa_source
 - location_dates
 - mtbs
 - nwcg
 - nwcg_fires
 - owner
 - stat
 - Views
 - Stored Procedures
 - Functions
 - recipeseexample
 - salesorderexample
 - schoolsschedulingexample
 - test_db

Administration Schemas

Information

No object selected

SQL File 0*

```

1 use p3;
2 select count(*) as OWNER_COUNT from owner;

```

Result Grid

	OWNER_COUNT
1	16

Output

Action Output

#	Time	Action	Message	Duration / Fetch
8	12:39:56	select count(*) as NWCG_COUNT from mtbs	1 row(s) returned	0.000 sec / 0.000 sec
9	12:40:08	select count(*) as NWCG_COUNT from nwcg	1 row(s) returned	0.781 sec / 0.000 sec
10	12:40:49	select count(*) as NWCG_FIRES_COUNT from nwcg_fires	1 row(s) returned	0.531 sec / 0.000 sec
11	12:40:57	select count(*) as NWCG_FIRES_COUNT from nwcg	1 row(s) returned	0.015 sec / 0.000 sec
12	12:41:05	select count(*) as NWCG_FIRES_COUNT from nwcg_fires	1 row(s) returned	0.000 sec / 0.000 sec
13	12:41:24	select count(*) as OWNER_COUNT from owner	1 row(s) returned	0.016 sec / 0.000 sec

SQLAdditions

Automatic context help is disabled. Use the toolbar to manually get help for the current caret position or to toggle automatic help.

Read Only Context Help Snippets

Object Info Session

File Edit View Query Database Server Tools Scripting Help

SQL File 8*

```

1 use p3;
2 select count(*) as STAT_COUNT from stat;

```

Automatic context help is disabled. Use the toolbar to manually get help for the current caret position or to toggle automatic help.

Result Grid

STAT_COUNT
13

Result 10 x

Output

Action Output

#	Time	Action	Message	Duration / Fetch
9	12:40:08	select count(*) as NWCG_COUNT from nwcg	1 row(s) returned	0.781 sec / 0.000 sec
10	12:40:49	select count(*) as NWCG_FIRES_COUNT from nwcg_fires	1 row(s) returned	0.531 sec / 0.000 sec
11	12:40:57	select count(*) as NWCG_FIRES_COUNT from nwcg	1 row(s) returned	0.015 sec / 0.000 sec
12	12:41:05	select count(*) as NWCG_FIRES_COUNT from nwcg_fires	1 row(s) returned	0.000 sec / 0.000 sec
13	12:41:24	select count(*) as OWNER_COUNT from owner	1 row(s) returned	0.016 sec / 0.000 sec
14	12:41:43	select count(*) as STAT_COUNT from stat	1 row(s) returned	0.015 sec / 0.000 sec

Object Info Session

SQL Queries

Query 1

Question 1

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

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

The state which had the least number of fires with stat_cause_code = 5, or 'debris burning' will have the least chance to win a share of fund. The result will contain 1 row stating the state having the least chance along with the count of debris burning.

Assumption: The fund shall be allocated to the state most affected state which have high number of debris burning in order to eliminate emergency.

Translation

SELECT state, count of state from table location_dates JOIN table fire ON ObjectID where stat_cause_code is 5 group by state sort ascending limiting

Screen Shot of SQL Query and Results

The screenshot displays the SQL Server Enterprise Manager interface. The left pane shows the 'SCHEMAS' tree with 'p3' expanded, listing tables like 'fire', 'fpa_source', 'location_dates', 'mtbs', 'nwcg', 'nwcg_fires', 'owner', and 'stat'. The central pane shows the following SQL query:

```
1 use p3;
2
3 -- SELECT state, count of state from table location_dates JOIN table fire ON ObjectID where stat_cause_code is 5 group by
4 -- state sort ascending limiting 1.
5 -- SELECT state, count(state) from location_dates JOIN fire ON location_dates.ObjectID = fire.Location_Dates_ObjectID
6 -- where stat_cause_code = 5 group by state order by count limit 1.
7
8
9
10 select STATE, count(state) from location_dates JOIN fire ON location_dates.ObjectID = fire.LOCATION_DATES_OBJECTID
11 where fire.STATE_CAUSE_CODE = 5
12 group by location_dates.STATE
13 order by count(state) limit 1 ;
14
```

The right pane shows a message: "Automatic context help is disabled. Use the toolbar to manually get help for the current caret position or to toggle automatic help."

The bottom pane shows the 'Result Grid' with the following data:

STATE	count(state)
DC	2

The bottom pane also shows the 'Output' window with the following messages:

#	Time	Action	Message	Duration / Fetch
11	12:40:57	select count(*) as NWCg_FIRES_COUNT from nwcg	1 row(s) returned	0.015 sec / 0.000 sec
12	12:41:05	select count(*) as NWCg_FIRES_COUNT from nwcg_fires	1 row(s) returned	0.000 sec / 0.000 sec
13	12:41:24	select count(*) as OWNER_COUNT from owner	1 row(s) returned	0.016 sec / 0.000 sec

Query 2

Question 2

One of the reporting agencies suggested that children be banned from its forest unless there is one adult for every 3 children in a group visiting a forest. Name 3 forests where this would be the least appropriate.

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

The forests which prohibits children to go alone are assumed to be the forests where stat_cause_code = 8, which means children.

ASSUMPTION: The forests where stat cause is children but the count is minimum is assumed to be least appropriate for 1 adult per 3 children because minimum occurrences means less threat.

Translation

select nwcg reporting unit name, count of unit name from table nwcg_fires JOIN fire ON nwcg_reporting_unit_id where stat cause code is 8 and group by unit name and order by count with limit 3.

Screen Shot of SQL Query and Results

The screenshot displays the SQL Developer interface. The left pane shows the 'SCHEMAS' tree with 'p3' selected. The main pane shows the following SQL query:

```
1 use p3;
2 -- TRANSLATION
3 -- select nwcg reporting unit name, count of unit name from table nwcg_fires JOIN fire ON nwcg_reporting_unit_id
4 -- where stat cause code is 8 and group by unit name and order by count with limit 3.
5
6 -- CLEANUP
7 -- SELECT nwcg_reporting_unit_name, count(nwcg_reporting_unit_name) from nwcg_fires JOIN fire
8 -- ON nwcg_fires.NWCG_REPORTING_UNIT_ID = fire.NWCG_FIRES_NWCG_REPORTING_UNIT_ID WHERE stat_stat_cause_code = 8
9 -- group by NWCG_REPORTING_UNIT_NAME order by COUNT limit 3;
10 *
11 select distinct nwcg_fires.NWCG_REPORTING_UNIT_NAME, count(nwcg_fires.NWCG_REPORTING_UNIT_NAME) as Count
12 from nwcg_fires JOIN fire ON nwcg_fires.NWCG_REPORTING_UNIT_ID = fire.NWCG_FIRES_NWCG_REPORTING_UNIT_ID
13 where fire.STAT_STAT_CAUSE_CODE = 8
14 group by nwcg_fires.NWCG_REPORTING_UNIT_NAME
15 order by Count limit 3;
```

The 'Result Grid' shows the following data:

NWCG_REPORTING_UNIT_NAME	Count
Coville National Forest	1
El Malpais National Monument	1
Big Horn Canyon National Recreation Area	1

The 'Output' pane shows the execution log:

#	Time	Action	Message	Duration / Fetch
1	16:41:37	use p3	0 row(s) affected	0.000 sec
2	16:41:41	select distinct nwcg_fires.NWCG_REPORTING_UNIT_NAME, count(nwcg_fires.NWCG_REPORTING_UNIT_NAME) as Count from nwcg_fires JOIN fire ON nwcg_fires.NWCG_REPORTING_UNIT_ID = fire.NWCG_FIRES_NWCG_REPORTING_UNIT_ID where fire.STAT_STAT_CAUSE_CODE = 8 group by nwcg_fires.NWCG_REPORTING_UNIT_NAME order by Count limit 3;	3 row(s) returned	2.438 sec / 0.000 sec

Query 3

Question 3

One advocacy group says human actions and not Nature is to blame for most wildfires. Write a query that supports this statement.

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

It is assumed that out of 13 stat causes only lightning (1) and Undefined/Missing (13) are the natural caused reasons. This also leads to assumption that Miscellaneous is a human action.

The result gives a percentage of count of human actions caused events and as the result is 76.3255 we can confidently support the statement that 'Human actions and not Nature is to blame for most wildfires'.

Translation

SELECT percentage of count of stat_cause_code) from fire where stat_cause_code is (2,3,4,5,6,7,8,9,10,11,12).

Screen Shot of SQL Query and Results

The screenshot displays a SQL IDE interface with a query editor, a results grid, and an output pane. The query editor shows the following SQL code:

```
1 use p3;
2
3 -- SELECT percentage of count of stat_cause_code) from fire where stat_cause_code is (2,3,4,5,6,7,8,9,10,11,12)
4
5 -- SELECT count(stat_cause_code)/(select count(*) from fire)*100 from fire where stat_cause_code IN (2,3,4,5,6,7,8,9,10,11)
6
7 select count(stat_cause_code)/(select count(*) from fire)*100 as percentage_human_reason
8 from fire
9 where STAT_CAUSE_CODE IN(2,3,4,5,6,7,8,9,10,11,12);
10
```

The results grid shows a single row with the value 76.3255.

percentage_human_reason
76.3255

The output pane shows the execution log with the following messages:

#	Time	Action	Message	Duration / Fetch
2	16:41:41	select distinct nwcg_fires.NWCg_REPORTING_UNIT_NAME, count(nwcg_fires.NWCg_REPOR...	3 row(s) returned	2.438 sec / 0.000 sec
3	16:48:56	select count(stat_cause_code)/(select count(*) from fire)*100 as percentage_human_reason.f...	1 row(s) returned	6.250 sec / 0.000 sec

Query 4

Question 6

What were the forests that had no fires that lasted more than two days?

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

To find the forests where no fire lasted more than two days assumption is made that the forest name lies in NWCG_REPORTING_UNIT_NAME. The result comprises of 1640 forests in which $\text{cont_date} - \text{discovery_date} \leq 2$.

Translation

SELECT nwcg_reporting_unit_id, nwcg_reporting_unit_name from table fire JOIN table location_dates ON ObjectID JOIN nwcg_fires ON nwcg_reporting_unit_id WHERE $\text{cont_date} - \text{discovery_date} \leq 2$

Screen Shot of SQL Query and Results

The screenshot displays the SQL Server Enterprise Manager interface. The left pane shows the 'SCHEMAS' tree with 'p3' expanded, listing tables like 'fire', 'location_dates', and 'nwcg_fires'. The central pane shows the following SQL query:

```
1 use p3;
2
3 -- SELECT nwcg_reporting_unit_id, nwcg_reporting_unit_name from table fire JOIN table location_dates ON ObjectID
4 -- JOIN nwcg_fires ON nwcg_reporting_unit_id WHERE cont_date - discovery_date <=2
5
6 -- SELECT distinct(nwcg_reporting_unit_id), nwcg_reporting_unit_name from fire join location_dates
7 -- ON fire.location_dates_objectid = location_dates.objectid JOIN nwcg_fires
8 -- ON nwcg_fires.NWCG_REPORTING_UNIT_ID = fire.NWCG_FIRES_NWCG_REPORTING_UNIT_ID WHERE cont_date - discovery_date <=2
9
10
11 select distinct(nwcg_fires.NWCG_REPORTING_UNIT_ID) as UNIT_ID, (nwcg_fires.NWCG_REPORTING_UNIT_NAME) as Forest
12 from fire join location_dates
13 ON fire.location_dates_objectid = location_dates.objectid
14 JOIN nwcg_fires ON nwcg_fires.NWCG_REPORTING_UNIT_ID = fire.NWCG_FIRES_NWCG_REPORTING_UNIT_ID
15 where location_dates.cont_date - location_dates.discovery_date <= 2;
```

The right pane shows a message: "Automatic context help is disabled. Use the toolbar to manually get help for the current caret position or to toggle automatic help." Below the query, the 'Result Grid' shows the following data:

UNIT_ID	Forest
USCAPHF	Plumas National Forest
USCASHF	Eldorado National Forest
USCASHF	Shasta-Trinity National Forest
USNMLNF	Lincoln National Forest
USORDEF	Deschutes National Forest

The bottom pane shows the 'Output' window with the following message:

```
5 17:03:28 select distinct(nwcg_fires.NWCG_REPORTING_UNIT_ID) as UNIT_ID, (nwcg_fires.NWCG_REPORTING_UNIT_NAME) as Forest
6 17:03:43 select distinct(nwcg_fires.NWCG_REPORTING_UNIT_ID) as UNIT_ID, (nwcg_fires.NWCG_REPORTING_UNIT_NAME) as Forest
1640 row(s) returned
```

Query 5

Question 7

Which state had more fires in the first half of a calendar year than the second half of the calendar year?

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

To compare the count of fires in each state by first half of the year to second half of the year, we make two subqueries delivering the count of first half and second half. If the count of the first half is more than the second half, we print the state.

The result comprises of 38 such states out of 52 total states where fires were more in first half of the year.

Translation

SELECT state from (select state, count of state from location_dates where month of discovery_date less than equal to 6 group by state) JOIN (select state, count of state from location_dates where month of discovery_date more than 6 GROUP BY state) ON state where count_1 > count_2

Screen Shot of SQL Query and Results

The screenshot displays the SQL Developer interface. The left pane shows the 'SCHEMAS' tree with 'p3' expanded, listing tables like 'fire', 'fpa_source', 'location_dates', 'mtbs', 'nwcg', 'nwcg_fires', 'owner', 'stat', and 'stat'. The main editor shows the following SQL query:

```
1 use p3;
2 -- SELECT state from (select state, count of state from location_dates where month of discovery_date less than equal to 6
3 -- group by state) JOIN (select state, count of state from location_dates where month of discovery_date more than 6
4 -- GROUP BY state) ON state where count_1 > count_2
5
6 -- SELECT state from (select state, count(state) as count_1 from location_dates where month(discovery_date) <=6 GROUP BY state)
7 -- as a JOIN (select state, count(state) as count_2 from location_dates where month(discovery_date) >6 GROUP BY state) as b
8 -- ON a.state = b.state where count_1 > count_2
9
10 SELECT a.state from
11 (select state, count(state) as count_1 from location_dates
12 where (month(discovery_date) <=6)
13 GROUP BY state) as a JOIN (select state, count(state) as count_2 from location_dates
14 where (month(discovery_date) >6)
15 GROUP BY state) as b ON a.state = b.state
16 where count_1 > count_2 ;
```

The 'Result Grid' shows the following results:

state
AR
MI
TX
LA
KS

The 'Output' pane shows the execution log:

#	Time	Action	Message	Duration / Fetch
10	17:30:21	select count(stat_cause_code)/(select count(*) from fire)*100 as percentage_human_reason f...	1 row(s) returned	5.953 sec / 0.000 sec
11	17:33:04	SELECT a.state from (select state, count(state) as count_1 from location_dates where monthdisc...	38 row(s) returned	6.828 sec / 0.000 sec

Query 6

Question 8

Which forest had the least number of fires?

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

To find the forest with least number of fires, we assume that the forest names are given in the SOURCE_REPORTING_UNIT_NAME and also that while ordering it we assume the count ≤ 1 because in the main fire table, the least count of any forest reported is 1.

The result will publish the names of all the forests which have occurred 1 time in the mail table and they are thus considered to be the least in the count. There are 534 forest which only reported ≤ 1 fires.

Translation

select count, name from (select count of nwcg_reporting_unit_id, nwcg_reporting_unit_name from table fire JOIN table nwcg_fires ON nwcg_reporting_unit_id GROUP BY nwcg_reporting_unit_name ORDER BY count) WHERE Count ≤ 1

Screen Shot of SQL Query and Results

The screenshot displays the MySQL Workbench interface. The SQL Editor window contains the following query:

```
1 use p3;
2
3 -- SELECT count, name from (select count of nwcg_reporting_unit_id, nwcg_reporting_unit_name from table fire
4 -- JOIN table nwcg_fires ON nwcg_reporting_unit_id GROUP BY nwcg_reporting_unit_name ORDER BY count) WHERE Count <= 1
5
6 -- SELECT count, name FROM (SELECT count(nwcg_reporting_unit_id as COUNT, nwcg_reporting_unit_name as NAME FROM fire join nwcg
7 -- ON fire.NWCG_FIRES_NWCG_REPORTING_UNIT_ID = nwcg_fires.NWCG_REPORTING_UNIT_ID GROUP BY nwcg_reporting_unit_name
8 -- ORDER BY COUNT asc) mycount
9
10 SELECT COUNT, NAME
11 FROM (select count(fire.FPA_SOURCE_FPA_ID) as COUNT, fpa_source.SOURCE_REPORTING_UNIT_NAME as NAME
12 FROM fire join fpa_source
13 ON fire.FPA_SOURCE_FPA_ID = fpa_source.FPA_ID
14 GROUP BY NAME
15 ORDER BY COUNT asc) mycount
16 where count <= 1;
```

The Results window shows the following data:

COUNT	NAME
1	Alaska Interagency Coordination Center
1	Hinsdale County
1	Mineral County
1	Ouray County
1	Desecho National Wildlife Refuge

The Action Output window shows the execution details:

#	Time	Action	Message	Duration / Fetch
13	17.48.10	SELECT COUNT, NAME FROM (select count(fire.NWCG_FIRES_NWCG_REPORTING_UNIT_...	195 row(s) returned	4.453 sec / 0.000 sec
14	18.24.43	SELECT COUNT, NAME FROM (select count(fire.FPA_SOURCE_FPA_ID) as COUNT, fpa_sour...	534 row(s) returned	26.047 sec / 0.000 sec

Data Review for MongoDB

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

1. As FOD_ID is the global unique identifier and is non-null and unique, it is assumed to be the primary key of the main table. There are other foreign keys like STAT_CAUSE_ID, OWNER_CODE, MTBS_ID, NWCG_REPORTING_UNIT_ID, FPA_ID, OBJECTID which are the primary keys in other tables.
2. In Stat_Cause_Table, stat_cause_id is assumed to be the primary key as each key represent an individual stat_cause_descr, the stat_cause_id is unique and non_null and stat_cause_descr is directly dependant on the stat_cause_id.
3. In the Owner table, owner_code is assumed to be the primary key because it is non-null, unique and all the element of owner_descr depend directly on owner_code.
4. In the MTBS table, MTBS_ID directly depends on each unique MTBS_NAME. MTBS_ID is the primary key for this table because it is unique and non-null and every record of MTBS_Fire_Name directly depends on MTBS_ID.
5. The NWCG table is assumed to be a separate table because there are a lot of records in Unit_ID column, all of which are not used in the main table. NWCG table serves the purpose of all the records given by NWCG Reporting and all the columns from NWCG_FIRES, and NWCG are directly related to NWCG_REPORTING_Unit_ID. NWCG_REPORTING_Unit_ID, the primary key of the table is non-null and unique.
6. In the FPA_SOURCE table, FPA_ID is the primary key because all the other columns directly depend on the FPA_ID and the SOURCE_REPORTING_UNIT_ID cannot be a primary key because of its redundant values.
7. In the LOCATION_DATES table, OBJECTID is the primary key as all the other elements depends directly on the OBJECTID and OBJECTID is a non-null unique value column.

Physical Mongo Database

Assumptions/Notes About Data Set

To organize the data effectively we divided the Fires table into 8 tables and joined the data with NWCG_UnitIDActive_20170109 table. The division was done to organise the data in 3rd Normalized Form.

The assumptions were as follows:

1. Data is provided only within Fires and NWCG-UnitIDActive_20170109 tables.
2. The collections are distributed to the best effort of reducing redundancy.

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

MongoDB Compass Community - localhost:27017/p1

Connect View Help

My Cluster

7 DBS 14 COLLECTIONS

HOST
localhost:27017

CLUSTER
Standalone

EDITION
MongoDB 4.0.13 Community

Filter your data

> admin

> config

> local

▼ p1

fire

fpa_source

location

mtbs

nwcg

nwcg_fires

owner

stat

Collections

CREATE COLLECTION

Collection Name	Documents	Avg. Document Size	Total Document Size	Num. Indexes	Total Index Size	Properties
fire	1,880,465	495.9 B	932.6 MB	1	18.8 MB	
fpa_source	1,880,462	202.4 B	380.6 MB	1	18.8 MB	
location	1,880,465	247.7 B	465.8 MB	1	18.8 MB	
mtbs	10,481	88.4 B	926.2 KB	1	106.5 KB	
nwcg	5,885	257.8 B	1.5 MB	1	65.5 KB	
nwcg_fires	5,885	150.9 B	887.8 KB	1	65.5 KB	
owner	16	68.3 B	1.1 KB	1	16.4 KB	
stat	13	79.9 B	1.0 KB	1	16.4 KB	

MongoDB Compass Community - localhost:27017/p1.fire

Connect View Collection Help

My Cluster

7 DBS 14 COLLECTIONS

HOST
localhost:27017

CLUSTER
Standalone

EDITION
MongoDB 4.0.13 Community

Filter your data

- admin
- config
- local
- p1
 - fire
 - fpa_source
 - location
 - mtba
 - nwcg
 - nwcg_fire
 - owner
 - stat

p1.fire

Documents 1.9m TOTAL SIZE 889.4MB AVG. SIZE 496B INDEXES 1 TOTAL SIZE 18.0MB AVG. SIZE 18.0MB

Documents Aggregations Explain Plan Indexes

Filter OPTIONS FIND RESET

INSERT DOCUMENT VIEW LIST TABLE

Displaying documents 1 - 20 of 1880465

#	fire	_id ObjectId	FOO_ID Double	LOCAL_FIRE_REPORT_ID String	LOCAL_INCIDENT_ID String	FIRE_CODE String
1		5dec825f5e128328d892f996	1	"1"	"PWF-47"	"B3BK"
2		5dec825f5e128328d892f997	2	"13"	"13"	"AAC9"
3		5dec825f5e128328d892f998	3	"23"	"821"	"A32W"
4		5dec825f5e128328d892f999	4	"43"	"6"	""
5		5dec825f5e128328d892f99a	5	"44"	"7"	""
6		5dec825f5e128328d892f99b	6	"54"	"8"	""
7		5dec825f5e128328d892f99c	7	"58"	"9"	""
8		5dec825f5e128328d892f99d	8	"3"	"82"	"BKCK"
9		5dec825f5e128328d892f99e	9	"5"	"83"	"BLPQ"
10		5dec825f5e128328d892f99f	10	"61"	"18"	""
11		5dec825f5e128328d892f9a0	11	"64"	"11"	""
12		5dec825f5e128328d892f9a1	12	"71"	"14"	""

MongoDB Compass Community - localhost:27017/p1.fpa_source

Connect View Collection Help

My Cluster

7 DBS 14 COLLECTIONS

HOST
localhost:27017

CLUSTER
Standalone

EDITION
MongoDB 4.0.13 Community

Filter your data

- admin
- config
- local
- p1
 - fire
 - fpa_source
 - location
 - mtba
 - nwcg
 - nwcg_fire
 - owner
 - stat

p1.fpa_source

Documents 1.9m TOTAL SIZE 363.0MB AVG. SIZE 202B INDEXES 1 TOTAL SIZE 17.9MB AVG. SIZE 17.9MB

Documents Aggregations Explain Plan Indexes

Filter OPTIONS FIND RESET

INSERT DOCUMENT VIEW LIST TABLE

Displaying documents 1 - 20 of 1880462

#	fpa_source	_id ObjectId	FPA_ID String	SOURCE_SYSTEM_TYPE String	SOURCE_SYSTEM String	SOURCE_REPORTING_UNIT String
1		5dec913e929d8ce496db9c71	"2008CAIRS10092975"	"NONFED"	"ST-CACDF"	"CACZU"
2		5dec913e929d8ce496db9c72	"2008CAIRS10163484"	"NONFED"	"ST-CACDF"	"CACZU"
3		5dec913e929d8ce496db9c73	"2008CAIRS10222599"	"NONFED"	"ST-CACDF"	"CASCU"
4		5dec913e929d8ce496db9c74	"2008CAIRS10345987"	"NONFED"	"ST-CACDF"	"CALMU"
5		5dec913e929d8ce496db9c75	"2008CAIRS10419166"	"NONFED"	"ST-CACDF"	"CABTU"
6		5dec913e929d8ce496db9c76	"2008CAIRS10850097"	"NONFED"	"ST-CACDF"	"CASCU"
7		5dec913e929d8ce496db9c77	"2008CAIRS11104163"	"NONFED"	"ST-CACDF"	"CAMEU"
8		5dec913e929d8ce496db9c78	"2009CAIRS11104592"	"NONFED"	"ST-CACDF"	"CASLU"
9		5dec913e929d8ce496db9c79	"2009CAIRS11108635"	"NONFED"	"ST-CACDF"	"CATGU"
10		5dec913e929d8ce496db9c7a	"2009CAIRS11125800"	"NONFED"	"ST-CACDF"	"CASLU"
11		5dec913e929d8ce496db9c7b	"2009CAIRS11144527"	"NONFED"	"ST-CACDF"	"CAFKU"
12		5dec913e929d8ce496db9c7c	"2009CAIRS11144953"	"NONFED"	"ST-CACDF"	"CASBU"

MongoDB Compass Community - localhost:27017/p1.location

Connect View Collection Help

My Cluster

7 DBS 14 COLLECTIONS

HOST: localhost:27017

CLUSTER: Standalone

EDITION: MongoDB 4.0.13 Community

Filter your data

- admin
- config
- local
- p1
 - fire
 - fpa_source
 - location
 - mtbs
 - nwcg
 - nwcg_fire
 - owner
 - stat

p1.location

DOCUMENTS 1.9m TOTAL SIZE 444.2MB AVG. SIZE 248B INDEXES 1 TOTAL SIZE 17.9MB AVG. SIZE 17.9MB

Documents Aggregations Explain Plan Indexes

Filter OPTIONS FIND RESET

INSERT DOCUMENT VIEW LIST TABLE

Displaying documents 1 - 20 of N/A

#	location	_id ObjectID	ObjectID Double	LATITUDE Double	LONGITUDE Double	STATE String
1		5dec8f8f929d3ce496becae0	1	40.0369	-121.006	"CA"
2		5dec8f8f929d3ce496becae1	2	38.9331	-120.404	"CA"
3		5dec8f8f929d3ce496becae2	3	38.9842	-120.736	"CA"
4		5dec8f8f929d3ce496becae3	4	38.5592	-119.913	"CA"
5		5dec8f8f929d3ce496becae4	5	38.5592	-119.933	"CA"
6		5dec8f8f929d3ce496becae5	6	38.6353	-120.104	"CA"
7		5dec8f8f929d3ce496becae6	7	38.6883	-120.153	"CA"
8		5dec8f8f929d3ce496becae7	8	40.9681	-122.434	"CA"
9		5dec8f8f929d3ce496becae8	9	41.2336	-122.283	"CA"
10		5dec8f8f929d3ce496becae9	10	38.5483	-120.149	"CA"
11		5dec8f8f929d3ce496becaea	11	38.6917	-120.16	"CA"
12		5dec8f8f929d3ce496becaeb	12	38.5275	-120.186	"CA"

MongoDB Compass Community - localhost:27017/p1.mtbs

Connect View Collection Help

My Cluster

7 DBS 14 COLLECTIONS

HOST: localhost:27017

CLUSTER: Standalone

EDITION: MongoDB 4.0.13 Community

Filter your data

- admin
- config
- local
- p1
 - fire
 - fpa_source
 - location
 - mtbs
 - nwcg
 - nwcg_fire
 - owner
 - stat

p1.mtbs

DOCUMENTS 10.5k TOTAL SIZE 904.5KB AVG. SIZE 88B INDEXES 1 TOTAL SIZE 104.0KB AVG. SIZE 104.0KB

Documents Aggregations Explain Plan Indexes

Filter OPTIONS FIND RESET

INSERT DOCUMENT VIEW LIST TABLE

Displaying documents 1 - 20 of 10481

#	mtbs	_id ObjectID	MTBS_ID String	MTBS_FIRE_NAME String
1		5dec8f64929d3ce496bec1ef	"AFS-0307-10950614"	"CLEARWATER #2"
2		5dec8f64929d3ce496bec1f0	"AK5674215793820060522"	"PESHK RIVER"
3		5dec8f64929d3ce496bec1f1	"AK5683615710019920517"	"YAMTARDI"
4		5dec8f64929d3ce496bec1f2	"AK57508157577200950529"	"PILOT POINT"
5		5dec8f64929d3ce496bec1f3	"AK5761615232120150828"	"TWIN CREEKS"
6		5dec8f64929d3ce496bec1f4	"AK5884315754320150622"	"COPENHAGEN CREEK"
7		5dec8f64929d3ce496bec1f5	"AK5884415640120150621"	"PAULS CREEK"
8		5dec8f64929d3ce496bec1f6	"AK5890415740520150704"	"COPENHAGEN"
9		5dec8f64929d3ce496bec1f7	"AK5899315855220171102"	"SHAKE RIVER"
10		5dec8f64929d3ce496bec1f8	"AK5900315765920030428"	"KAMISHAK BAY"
11		5dec8f64929d3ce496bec1f9	"AK5925816024020150622"	"GECHIAK LAKE"
12		5dec8f64929d3ce496bec1fa	"AK5955415836020150711"	"LITTLE KODAK"

MongoDB Compass Community - localhost:27017/p1.nwgc

Connect View Collection Help

My Cluster

7 DBS 14 COLLECTIONS

HOST localhost:27017

CLUSTER Standalone

EDITION MongoDB 4.0.13 Community

Filter your data

admin

config

local

p1

fire

fpa_source

location

mtba

nwgc

nwgc_fire

owner

stat

p1.nwgc Documents

DOCUMENTS 5.9k TOTAL SIZE 1.4MB AVG. SIZE 258B INDEXES 1 TOTAL SIZE 64.0KB AVG. SIZE 64.0KB

Documents Aggregations Explain Plan Indexes

Filter OPTIONS FIND RESET

INSERT DOCUMENT VIEW LIST TABLE

Displaying documents 1 - 20 of 5885

#	mcg	_id ObjectId	MICRO_FIRES_MICRO_REPORTING_UNIT_	GeographicArea String	6acc String	WildlandHole String
1		5dec8f5f929d3ce496beaf2	"CAMBARI"	"IN"	"CAMBICFC"	"Dispatch/Coordination C
2		5dec8f5f929d3ce496beaf3	"CAMBARI"	"IN"	"CAMBICFC"	"Incident Host Geographi
3		5dec8f5f929d3ce496beaf4	"CAMBICFC"	"IN"	"CAMBICFC"	"Dispatch/Coordination C
4		5dec8f5f929d3ce496beaf5	"CAMBARI"	"IN"	"CAMBICFC"	"Incident Host Geographi
5		5dec8f5f929d3ce496beaf6	"CAMBARI"	"IN"	"CAMBICFC"	"Incident Host Geographi
6		5dec8f5f929d3ce496beaf7	"CAMBARI"	"IN"	"CAMBICFC"	"Incident Host Geographi
7		5dec8f5f929d3ce496beaf8	"CAMBARI"	"IN"	"CAMBICFC"	"Incident Host Geographi
8		5dec8f5f929d3ce496beaf9	"CAMBARI"	"IN"	"CAMBICFC"	"Incident Host Geographi
9		5dec8f5f929d3ce496beafa	"CAMBARI"	"IN"	"CAMBICFC"	"Incident Host Geographi
10		5dec8f5f929d3ce496beafb	"CAMBARI"	"IN"	"CAMBICFC"	"Incident Host Geographi
11		5dec8f5f929d3ce496beafc	"CAMBARI"	"IN"	"CAMBICFC"	"Incident Host Geographi
12		5dec8f5f929d3ce496beafd	"CAMBARI"	"IN"	"CAMBICFC"	"Incident Host Geographi

MongoDB Compass Community - localhost:27017/p1.nwgc_fire

Connect View Collection Help

My Cluster

7 DBS 14 COLLECTIONS

HOST localhost:27017

CLUSTER Standalone

EDITION MongoDB 4.0.13 Community

Filter your data

admin

config

local

p1

fire

fpa_source

location

mtba

nwgc

nwgc_fire

owner

stat

p1.nwgc_fire Documents

DOCUMENTS 5.9k TOTAL SIZE 867.0KB AVG. SIZE 151B INDEXES 1 TOTAL SIZE 64.0KB AVG. SIZE 64.0KB

Documents Aggregations Explain Plan Indexes

Filter OPTIONS FIND RESET

INSERT DOCUMENT VIEW LIST TABLE

Displaying documents 1 - 20 of 5885

#	mcg_fire	_id ObjectId	MICRO_REPORTING_UNIT_ID String	MICRO_REPORTING_AGENCY String	MICRO_REPORTING_UNIT_NAME String
1		5dec8f5f929d3ce496be93f5	"CAMBARI"	"	"Alberta Environment and Sustai
2		5dec8f5f929d3ce496be93f6	"CAMBARI"	"	"British Columbia Wildfire Mana
3		5dec8f5f929d3ce496be93f7	"CAMBICFC"	"	"Canadian Intermcg_reporting_a
4		5dec8f5f929d3ce496be93f8	"CAMBARI"	"	"Manitoba Wildfire Program"
5		5dec8f5f929d3ce496be93f9	"CAMBARI"	"	"New Brunswick Forest Fire Mana
6		5dec8f5f929d3ce496be93fa	"CAMBARI"	"	"Newfoundland and Labrador Fire
7		5dec8f5f929d3ce496be93fb	"CAMBARI"	"	"Nova Scotia Forest Protection"
8		5dec8f5f929d3ce496be93fc	"CAMBARI"	"	"Northwest Territories Forest M
9		5dec8f5f929d3ce496be93fd	"CAMBARI"	"	"Ontario Aviation"
10		5dec8f5f929d3ce496be93fe	"CAMBARI"	"	"Prince Edward Island Forests"
11		5dec8f5f929d3ce496be93ff	"CAMBARI"	"	"Parks Canada National Fire Man
12		5dec8f5f929d3ce496be9400	"CAMBARI"	"	"Quebec SORFEU"

MongoDB Compass Community - localhost:27017/p1.owner

Connect View Collection Help

My Cluster

- 7 DBS 14 COLLECTIONS
- HOST localhost:27017
- CLUSTER Standalone
- EDITION MongoDB 4.0.13 Community

Filter your data

- admin
- config
- local
- p1
 - fire
 - fpa_source
 - location
 - mtba
 - nwcg
 - nwcg_fire
 - owner
 - stat

p1.owner Documents

DOCUMENTS 16 TOTAL SIZE 1.1KB AVG. SIZE 68B INDEXES 1 TOTAL SIZE 16.0KB AVG. SIZE 16.0KB

Documents Aggregations Explain Plan Indexes

Filter OPTIONS FIND RESET

INSERT DOCUMENT VIEW LIST TABLE

Displaying documents 1 - 16 of 16

#	owner	_id ObjectId	OWNER_CODE Double	OWNER_DESCR String
1		5dec8fae929d3ce496be93e5	0	"FOREIGN"
2		5dec8fae929d3ce496be93e6	1	"BLM"
3		5dec8fae929d3ce496be93e7	2	"BLM"
4		5dec8fae929d3ce496be93e8	3	"NPS"
5		5dec8fae929d3ce496be93e9	4	"FWS"
6		5dec8fae929d3ce496be93ea	5	"USFS"
7		5dec8fae929d3ce496be93eb	6	"OTHER FEDERAL"
8		5dec8fae929d3ce496be93ec	7	"STATE"
9		5dec8fae929d3ce496be93ed	8	"PRIVATE"
10		5dec8fae929d3ce496be93ee	9	"TRIBAL"
11		5dec8fae929d3ce496be93ef	10	"BOR"
12		5dec8fae929d3ce496be93f0	11	"COUNTY"

MongoDB Compass Community - localhost:27017/p1.stat

Connect View Collection Help

My Cluster

- 7 DBS 14 COLLECTIONS
- HOST localhost:27017
- CLUSTER Standalone
- EDITION MongoDB 4.0.13 Community

Filter your data

- admin
- config
- local
- p1
 - fire
 - fpa_source
 - location
 - mtba
 - nwcg
 - nwcg_fire
 - owner
 - stat

p1.stat Documents

DOCUMENTS 13 TOTAL SIZE 1.0KB AVG. SIZE 80B INDEXES 1 TOTAL SIZE 16.0KB AVG. SIZE 16.0KB

Documents Aggregations Explain Plan Indexes

Filter OPTIONS FIND RESET

INSERT DOCUMENT VIEW LIST TABLE

Displaying documents 1 - 13 of 13

#	stat	_id ObjectId	STAT_CAUSE_CODE Double	STAT_CAUSE_DESCR String
1		5dec8fa7929d3ce496be93d8	1	"Lightning"
2		5dec8fa7929d3ce496be93d9	2	"Equipment Use"
3		5dec8fa7929d3ce496be93da	3	"Smoking"
4		5dec8fa7929d3ce496be93db	4	"Campfire"
5		5dec8fa7929d3ce496be93dc	5	"Debris Burning"
6		5dec8fa7929d3ce496be93dd	6	"Railroad"
7		5dec8fa7929d3ce496be93de	7	"Arson"
8		5dec8fa7929d3ce496be93df	8	"Children"
9		5dec8fa7929d3ce496be93e0	9	"Miscellaneous"
10		5dec8fa7929d3ce496be93e1	10	"Fireworks"
11		5dec8fa7929d3ce496be93e2	11	"Powerline"
12		5dec8fa7929d3ce496be93e3	12	"Structure"

Data in the Database

Collection Name	Relationships With Other Collections (if any)	# of Documents in Collection
FIRE	-	1880465
FPA_SOURCE	-	1880462
LOCATION	-	1880465
MTBS	-	10481
NWCG	-	5885
NWCG_FIRES	-	5885
OWNER	-	16
STAT	-	13

MongoDB Queries/Code

Query 1

Question 8

Which forest had the least number of fires?

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

To find the name of forests with the least number of fires, we made an assumption that FPA_SOURCE_REPORTING_UNIT_NAME has all the forest names that are called in the main file.

Also, we assumed that count <=1 will give all the forest names that have the least occurrences in the main fire table as there can be no forest name with count 0 because that will not let it be in the main table. We do this by creating an aggregation pipeline and using lookup, group, sort functions in it.

Translation

Group SOURCE_REPORTING_UNIT_NAME along with the count and order it in ascending order.

Screen Shot of MongoDB Query/Code and Results

The screenshot shows the MongoDB Compass interface for the 'p1.fpa_source' collection. The aggregation pipeline is defined as follows:

```
1 * field - The first field name.
2 *
3 *
4 *
5 *
6 {
7   _id: "$SOURCE_REPORTING_UNIT_NAME",
8   count: { $sum: 1 }
9 }
```

The results of the aggregation are displayed in a table:

_id	count
AFC Southeastern Region	239
AFC Ozark	235

A second stage, \$sort, is added to the pipeline. The output after the \$sort stage (Sample of 20 documents) is shown below:

```
1 *
2 * Provide any number of field/order pairs.
3 *
4 *
5 {
6   count: 1
7 }
```

The output shows that 'Black Hills National Forest' has the lowest count of 1.

Query 2

Question 1

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 state has the least chance to win a share of the fund?

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

The state which had the least number of fires with `stat_cause_code = 5`, or 'debris burning' will have the least chance to win a share of fund. The result will contain 1 row stating the state having the least chance along with the count of debris burning.

Assumption: The fund shall be allocated to the state most affected state which have high number of debris burning in order to eliminate emergency and thus neglecting the ones least affected.

We use the `$lookup` function to associate location table with fire table, and then match the `stat_stat_cause_code` and group by state name with it's count.

Translation

Join the tables fire and locations using `$lookup`, define parameters for the lookup. Then, match the `stat_stat_cause_code` for debris burning. After this, group, the output using the state name and also display the count of state for occurrences of `stat_stat_cause_code`.

Screen Shot of MongoDB Query/Code and Results

```
MongoDB shell version v4.0.13
connecting to: mongodb://127.0.0.1:27017/?sslapiServiceName=mongodb
Implicit session: session { "id" : UUID("08d2ae31-506e-4e17-bcb5-40253dbe68c3") }
MongoDB server version: 4.0.13
Server has startup warnings:
2019-12-08T19:43:22.084-0600 I CONTROL [initandlisten]
2019-12-08T19:43:22.084-0600 I CONTROL [initandlisten] ** WARNING: Access control is not enabled for the database.
2019-12-08T19:43:22.084-0600 I CONTROL [initandlisten] **           Read and write access to data and configuration is unrestricted.
2019-12-08T19:43:22.084-0600 I CONTROL [initandlisten]
---
Free Monitoring URL:
https://cloud.mongodb.com/freemonitoring/cluster/Q3ZETYEJULVAGVRJTLZ6NDMEJIG8DDCS
---
> use p1;
switched to db p1
> db.location.aggregate([
...{ $lookup: {from: "fire",    localField: "ObjectID", foreignField: "location_dates_ObjectID", as: "fire" }
...},
...{ $match: { ["stat_stat_cause_code": 5] }
...},
...{ $group: { statename: "$state", "count": { $sum: 1}}
...}
...]).sort("count": 1);
{"STATE" : DC, "count(state)" : 2}
>
```

Query 3

Question 2

One of the reporting agencies has suggested that children be banned from its forests unless there is one adult for every 3 children in a group visiting a forest. Name 3 forests where this would be the least appropriate.

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

The forests which prohibits children to go alone are assumed to be the forests where `stat_cause_code = 8`, which means children.

ASSUMPTION: The forests where `stat_cause` is children but the count is minimum is assumed to be least appropriate for 1 adult per 3 children because minimum occurrences means less threat to children

Translation

Create an aggregate pipeline using `$lookup` for `nwcg_fires` and `fire` tables. Then use `$match` to find the `stat_cause_code` for children; ie: 8. After this group the output, using `NWCG_REPORTING_UNIT_NAME` and it's count for `stat_cause_code`.

Screen Shot of MongoDB Query/Code and Results

```
MongoDB shell version v4.0.13
connecting to: mongodb://127.0.0.1:27017/?gssapiServiceName=mongodb
Implicit session: session { "id" : UUID("08d2ae31-506e-4e17-bcb5-40253dbe68c3") }
MongoDB server version: 4.0.13
Server has startup warnings:
2019-12-08T19:43:22.084-0600 I CONTROL [initandlisten]
2019-12-08T19:43:22.084-0600 I CONTROL [initandlisten] ** WARNING: Access control is not enabled for the database.
2019-12-08T19:43:22.084-0600 I CONTROL [initandlisten] **           Read and write access to data and configuration is unrestricted.
2019-12-08T19:43:22.084-0600 I CONTROL [initandlisten]
---
Free Monitoring URL:
https://cloud.mongodb.com/freemonitoring/cluster/Q3ZFYIEJULVAGVRJTLZ6MDMEJI6BDDCS
---
> use p1;
switched to db p1
> db.nwcg_fires.aggregate([
...{ $lookup:{from: "fire", localField: "NWCG_REPORTING_UNIT_ID", foreignField: "NWCG_FIRES_NWCG_REPORTING_UNIT_ID", as: "fire" }
...},
...{$match: {"stat_cause_code" : 8}
...},
...{ $group: {NWCG_REPORTING_UNIT_NAME: "$NWCG_REPORTING_UNIT_NAME", "count": {$sum: 1}}
...}
...}).sort("count": 1), {$limit: 3};
{"NWCG_REPORTING_UNIT_NAME" : Colville National Forest, "count" : 1}
{"NWCG_REPORTING_UNIT_NAME" : El Malpais National Monument, "count" : 1}
{"NWCG_REPORTING_UNIT_NAME" : Big Horn Canyon National Recreation Area, "count" : 1}
>
```

Query 4

Question 6

What were the forests that had no fires that lasted more than two days?

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

To find the forests where no fire lasted more than two days assumption is made that the forest name lies in NWCG_REPORTING_UNIT_NAME. The result comprises of 1640 forests in which `cont_date - discovery_date <= 2`.

Translation

This query also involves creating an aggregation pipeline, then using \$lookup for fire and nwcg_fires tables. We need to use a second \$lookup here for location and fire tables. Then match the difference of `cont_date` and `discovery_date <= 2` and group the output using `NWCG_REPORTING_UNIT_NAME`.

Screen Shot of MongoDB Query/Code and Results

```
MongoDB shell version v4.0.13
connecting to: mongodb://127.0.0.1:27017/?gssapiServiceName=mongodb
Implicit session: session { "id" : UUID("08d2ae31-506e-4e17-bcb5-40253dbe68c3") }
MongoDB server version: 4.0.13
Server has startup warnings:
2019-12-08T19:43:22.084-0600 I CONTROL [initandlisten]
2019-12-08T19:43:22.084-0600 I CONTROL [initandlisten] ** WARNING: Access control is not enabled for the database.
2019-12-08T19:43:22.084-0600 I CONTROL [initandlisten] **           Read and write access to data and configuration is unrestricted.
2019-12-08T19:43:22.084-0600 I CONTROL [initandlisten]
---
Free Monitoring URL:
https://cloud.mongodb.com/freemonitoring/cluster/Q3ZF1YE7JULVAGVRJ1L76WDM6J16BDDCS
---
> use p1;
switched to db p1
> db.fire.aggregate([
...{ $lookup: {from:"nwcg_fire", localField:"NWCG_FIRES_NWCG_REPORTING_UNIT_ID", foreignField:"NWCG_REPORTING_UNIT_ID", as:"nwcg_fire"}
... },
...{ $lookup: {from:"location", localField: "location_dates_objectid", foreignField: "objectid", as: "location" }
... },
...{ $match: { $expr: { $lte: [ { $subtract: [ "$CONT_DATE", "$DISCOVERY_DATE" ] }, 2 ] } } }
... },
...{ $group: { "_id": "$NWCG_REPORTING_UNIT_NAME" } }
...]);
{"UNIT_ID": "USCAPNF", "Forest": "Plumas National Forest"}
{"UNIT_ID": "USCAENF", "Forest": "Eldorado National Forest"}
{"UNIT_ID": "USCASHF", "Forest": "Shasta-Trinity National Forest"}
{"UNIT_ID": "USNMLNF", "Forest": "Lincoln National Forest"}
{"UNIT_ID": "USORDEF", "Forest": "Deschutes National Forest"}
{"UNIT_ID": "USNCNCF", "Forest": "National Forests in North Carolina"}
{"UNIT_ID": "USWYMBF", "Forest": "Medicine Bow National Forest"}
{"UNIT_ID": "USORUMF", "Forest": "Umatilla National Forest"}
{"UNIT_ID": "USCOPSF", "Forest": "Pike and San Isabel National Forest"}
{"UNIT_ID": "USOROCF", "Forest": "Ochoco National Forest"}
```

Query 5

Question 3

One advocacy group says human actions and not Nature is to blame for most wildfires. Write a query that supports this statement.

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

It is assumed that out of 13 stat causes only lightning (1) and Undefined/Missing (13) are the natural caused reasons. This also leads to assumption that Miscellaneous is a human action.

The result gives a percentage of count of human actions caused events and as the result is 76.3255 we can confidently support the statement that 'Human actions and not Nature is to blame for most wildfires'.

Translation

Use the db.collection.find() function to find the stat_cause_code that are relative to human causes and a percentage of that with respect to total number of records is found.

Screen Shot of MongoDB Query/Code and Results

```
MongoDB shell version v4.0.13
connecting to: mongodb://127.0.0.1:27017/?gssapiServiceName=mongodb
Implicit session: session { "id" : UUID("08d2ae31-506e-4e17-bcb5-40253dbe68c3") }
MongoDB server version: 4.0.13
Server has startup warnings:
2019-12-08T19:43:22.084-0600 I CONTROL [initandlisten]
2019-12-08T19:43:22.084-0600 I CONTROL [initandlisten] ** WARNING: Access control is not enabled for the database.
2019-12-08T19:43:22.084-0600 I CONTROL [initandlisten] **           Read and write access to data and configuration is unrestricted.
2019-12-08T19:43:22.084-0600 I CONTROL [initandlisten]
***
Free Monitoring URL:
https://cloud.mongodb.com/freemonitoring/cluster/Q3ZFTYEJULVAGVRJTLZ6WdHEJIG6DOCS
***
> use p1;
switched to db p1
> db.fire.find({
... "STAT_STAT_CAUSE_CODE": {"$in": [2,3,4,5,6,7,8,9,10,11,12]}
... },{
... "count(stat_stat_cause_code)/1880465*100 as percentage_human_reason": 1
... }
... });
{ "_id" : ObjectId("5dec825f5e128328d892f996") }
{ "_id" : ObjectId("5dec825f5e128328d892f998") }
{ "_id" : ObjectId("5dec825f5e128328d892f99d") }
{ "_id" : ObjectId("5dec825f5e128328d892f99e") }
{ "_id" : ObjectId("5dec825f5e128328d892f9a2") }
{ "_id" : ObjectId("5dec825f5e128328d892f9a3") }
{ "_id" : ObjectId("5dec825f5e128328d892f9a6") }
{ "_id" : ObjectId("5dec825f5e128328d892f9a7") }
{ "_id" : ObjectId("5dec825f5e128328d892f9a8") }
{ "_id" : ObjectId("5dec825f5e128328d892f9a9") }
{ "_id" : ObjectId("5dec825f5e128328d892f9b2") }
{ "_id" : ObjectId("5dec825f5e128328d892f9b3") }
{ "_id" : ObjectId("5dec825f5e128328d892f9b5") }
{ "_id" : ObjectId("5dec825f5e128328d892f9b9") }
{ "_id" : ObjectId("5dec825f5e128328d892f9ba") }
{ "_id" : ObjectId("5dec825f5e128328d892f9bb") }
{ "_id" : ObjectId("5dec825f5e128328d892f9bc") }
{ "_id" : ObjectId("5dec825f5e128328d892f9bd") }
{ "_id" : ObjectId("5dec825f5e128328d892f9be") }
```

Query 6

Question 7

Which state had more fires in the first half of a calendar year than the second half of the calendar year?

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

To compare the count of fires in each state by first half of the year to second half of the year, we make two subqueries delivering the count of first half and second half. If the count of the first half is more than the second half, we print the state.

The result comprises of 38 such states out of 52 total states where fires were more in first half of the year.

Translation

Using the `db.collection.group` function, we list out all the states where fires are more in the first half of a calendar year than the second half of the calendar year by defining a condition using `$gte` operator.

Screen Shot of MongoDB Query/Code and Results

```
MongoDB shell version v4.0.13
connecting to: mongodb://127.0.0.1:27017/?sslapiServiceName=mongodb
Implicit session: session { "id" : UUID("08d2ae31-506e-4e17-bcb5-40253dbe68c3") }
MongoDB server version: 4.0.13
Server has startup warnings:
2019-12-08T19:43:22.084-0600 I CONTROL [initandlisten]
2019-12-08T19:43:22.084-0600 I CONTROL [initandlisten] ** WARNING: Access control is not enabled for the database.
2019-12-08T19:43:22.084-0600 I CONTROL [initandlisten] **           Read and write access to data and configuration is unrestricted.
2019-12-08T19:43:22.084-0600 I CONTROL [initandlisten]
---
Free Monitoring URL:
https://cloud.mongodb.com/freemonitoring/cluster/Q3ZF1YEJULVAGVRJTLZ6MDWEJI6BDDCS
---
> use p1;
switched to db p1
> db.location.group([ { key: { state: 1 }, cond: { month["$CONT_DOY"]: {$gte: month["$DISCOVERY_DATE"]} }, initial: { } }
... ]);
{"state": "AR"}
{"state": "MI"}
{"state": "TX"}
{"state": "LA"}
{"state": "KS"}
{"state": "NE"}
{"state": "OK"}
{"state": "FL"}
{"state": "NC"}
{"state": "SC"}
{"state": "HI"}
{"state": "MO"}
{"state": "KY"}
{"state": "IN"}
{"state": "VA"}
{"state": "IL"}
```