



SQL-Mongo Project – Spatial Data of US Wildfires

BUAN 6320

Group Members:
Sneha Santosh Bhagwat
Divya Deepak Pai
Astha Bansal
Rajas Prasannakumar Thakur

Group #: 6

Activity	Sneha	Divya	Astha	Rajas
Prepared Data Model and Created Physical DB	X	X	X	X
Loaded Data into Database	X	X	X	X
Wrote SQL Queries				
Prepared Mongo Database				
Loaded data into Mongo DB				
Wrote Mongo Queries				
Prepared Report				
Reviewed Report				

Contents

Data Model	6
Assumptions/Notes about Data set.....	6
Data Entities and Relationships	7
Entity-Relationship Diagram	9
Physical Database	10
Assumptions/Notes About Data Set	10
Screen shot of Physical Database objects.....	11
Data in the Database.....	21
SQL Queries.....	22
Query 1.....	22
Question.....	22
Notes/Comments About SQL Query and Results (Include # of Rows in Result).....	22
Translation	22
Screen Shot of SQL Query and Results.....	23
Query 2.....	24
Question.....	24
Notes/Comments About SQL Query and Results (Include # of Rows in Result).....	24
Translation	24
Screen Shot of SQL Query and Results.....	25
Query 3.....	26
Question.....	26
Notes/Comments About SQL Query and Results (Include # of Rows in Result).....	26
Translation	26
Screen Shot of SQL Query and Results.....	27
Query 4.....	29
Question.....	29
Notes/Comments About SQL Query and Results (Include # of Rows in Result).....	29
Translation	29
Screen Shot of SQL Query and Results.....	29
Query 5.....	30
Question.....	30
Notes/Comments About SQL Query and Results (Include # of Rows in Result).....	30

Translation	30
Screen Shot of SQL Query and Results.....	31
Query 6.....	32
Question.....	32
Notes/Comments About SQL Query and Results (Include # of Rows in Result).....	32
Translation	32
Screen Shot of SQL Query and Results.....	32
Data Review for MongoDB.....	33
Assumptions/Notes About Data Collections, Attributes and Relationships between Collections	33
Physical Mongo Database	34
Assumptions/Notes About Data Set	34
Screen shot of Physical Database objects (Database, Collections and Attributes)	34
Data in the Database.....	34
MongoDB Queries/Code	35
Query 1.....	35
Question.....	35
Notes/Comments About MongoDB Query/Code and Results (Include # of Documents in Result) ...	35
Translation	35
Screen Shot of MongoDB Query/Code and Results.....	35
Query 2.....	36
Question.....	36
Notes/Comments About MongoDB Query/Code and Results (Include # of Documents in Result) ...	36
Translation	36
Screen Shot of MongoDB Query/Code and Results.....	36
Query 3.....	37
Question.....	37
Notes/Comments About MongoDB Query/Code and Results (Include # of Documents in Result) ...	37
Translation	37
Screen Shot of MongoDB Query/Code and Results.....	37
Query 4.....	38
Question.....	38
Notes/Comments About MongoDB Query/Code and Results (Include # of Documents in Result) ...	38
Translation	38

Screen Shot of MongoDB Query/Code and Results.....	38
Query 5.....	39
Question.....	39
Notes/Comments About MongoDB Query/Code and Results (Include # of Documents in Result) ...	39
Translation	39
Screen Shot of MongoDB Query/Code and Results.....	39
Query 6.....	40
Question.....	40
Notes/Comments About MongoDB Query/Code and Results (Include # of Documents in Result) ...	40
Translation	40
Screen Shot of MongoDB Query/Code and Results.....	40

Data Model

Assumptions/Notes about Data set

- We are assuming that there is only one NWCG unit for a combination of code and state in NWCG_Unit_Location table
- After looking at the data, we have included following tables in the data model:
 1. Fires
 2. Nwcg_unitidactive_20170109
 3. Idx_fires_shape_node
 4. Idx_fires_shape_parent
 5. Idx_fires_shape_rowid

All the other tables are excluded from our data model.

- There is identifying one-to-one relationship between Discovery_Data and Fires (In order to eliminate transitivity). Hence, the foreign key (FOD_ID) is also the primary key in Discovery_Data table.
- There is identifying one-to-one relationship between Controlled_Data and Fires. Hence, the foreign key (FOD_ID) is also the primary key in Controlled_Data table.
- From the given data, an ICS number is assigned to more than one fire. Hence, we have assumed there is one-to-many relationship between ICS and Fires. Also, as per the given dataset, some fires have not been assigned an ICS_209_Incident_Number.
- From the given data, a FIPS_CODE is assigned to more than one fire. Hence, we have assumed there is one-to-many relationship between FIPS and Fires. Also, as per the given dataset, some fires have not been assigned an FIPS_CODE.
- From the given data, an MTBS ID is assigned to more than one fire. Hence, we have assumed there is one-to-many relationship between MTBS and Fires. Also, as per the given dataset, some fires have not been assigned an MTBS ID.
- We have included parent_node of idx_Fires_Shape_parent table in idx_Fires_Shape_node table since all the other columns in both the tables were same and it was having one to one relationship.

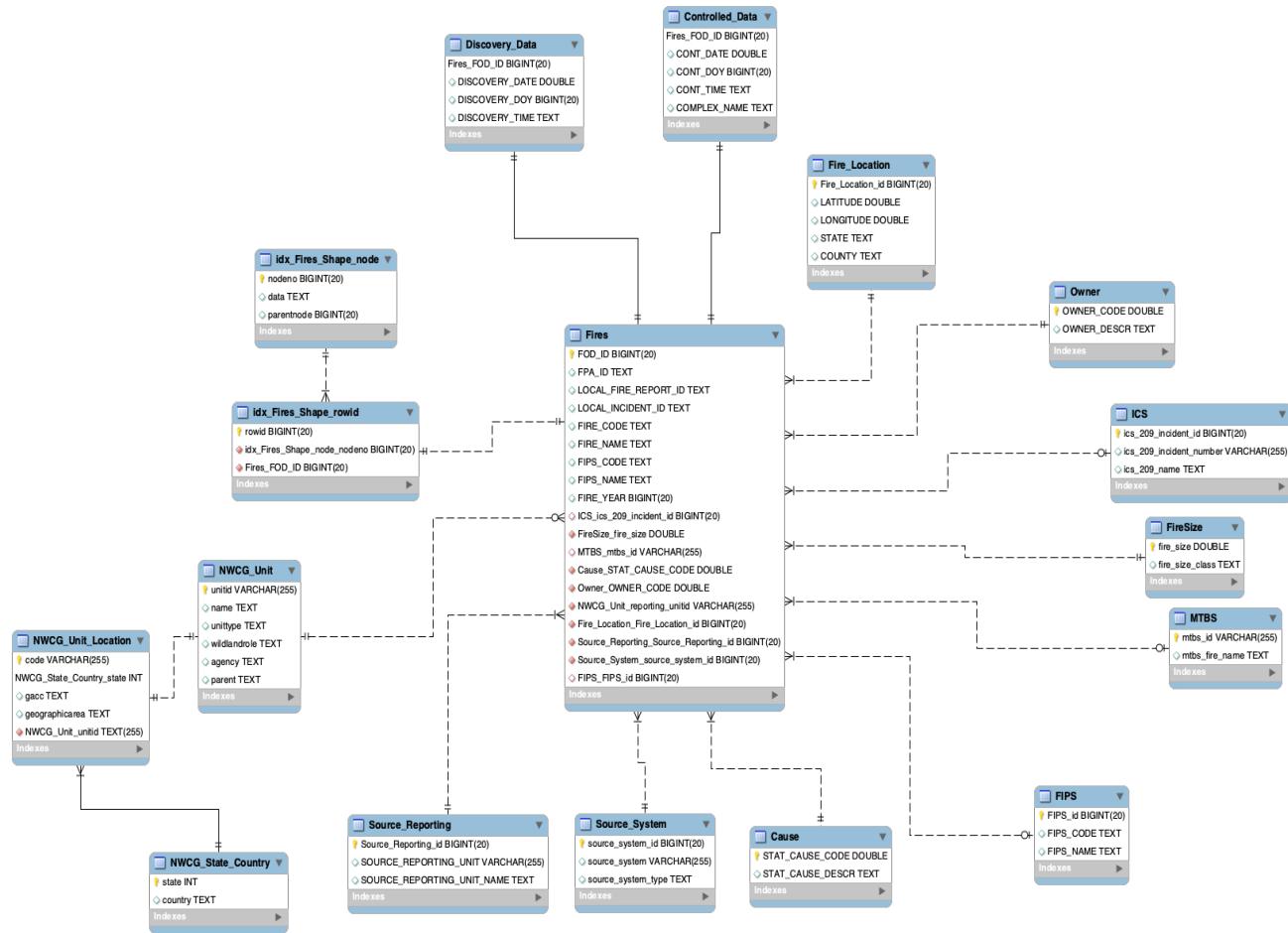
Data Entities and Relationships:

- A Fire has one and only one Discovery_Data. A Discovery_Data is associated with one and only one Fire.
- A Fire has one and only one Controlled_Data. A Controlled_Data is associated with one and only one Fire.
- A Fire has one and only one Fire_Location. A Fire_Location can be location of one to many Fires.
- A Fire location is managed by one and only one Owner. An Owner can manage one to many Fire locations.
- A Fire has one and only one Fire_Location. Fire_Location can have one to many Fires.
- A Fire can have zero to one Incident Identifier (ICS). An ICS can be assigned to one to many Fires.
- A Fire has one and only one Fire_Size. A Fire_Size can be associated with one to many Fires.
- A Fire can have zero to one MTBS Incident Identifier. A MTBS Incident Identifier can be associated with one to many Fires.
- A Fire can have zero to one FIPS Code. A FIPS Code can be associated with one to many Fires.
- A Fire can have one and only one Cause. A Cause can be related to one to many Fires.
- A Fire can have one and only one Source_System. A Source_system is related to one to many Fires.
- A Fire can be reported by one and only one Source_Reporting unit. A Source_Reporting unit can report one to many Fires.
- A Fire can be reported by one and only one NWCG_Unit. A NWCG_Unit can report zero to many Fires.
- A NWCG_Unit has one and only one NWCG_Unit_Location. A NWCG_Unit_Location has one and only one NWCG_Unit.
- A NWCG_Unit_Location has one and only one NWCG_State_Country. A NWCG_State_Country has one to many NWCG_Unit_Location.
- A Fire can have one and only one idx_Fires_Shape_rowid. A idx_Fires_Shape_rowid is associated with one and only one Fire.
- A idx_Fires_Shape_node can be associated with one to many idx_Fires_Shape_rowid. An idx_Fires_Shape_rowid is associated with one and only one idx_Fires_Shape_node.

Reasons why the data model is in 3NF:

- After performing Normalization, we can see that each table in the dataset has a **single value** in each cell. Also, while loading data in to the tables, we have loaded all the DISTINCT records. Hence, there are **unique rows** in each table. Thus, the **data model is in 1NF**.
- We have ensured that each non-key column is completely dependent on all the primary keys of the table. Thus, the data model **does not have partial dependency**. Thus, **it is in 2NF**.
- In order to remove transitive dependencies, the following actions were undertaken:
 1. In Discovery_Data table, we can get date information using 'DISCOVERY_DOY' and 'FIRE_YEAR' together, the issue of transitive dependency was arising as the non-key column ('DISCOVERY_DATE') was dependent on other non-key column ('DISCOVERY_DOY' and 'FIRE_YEAR'). So, we included 'FIRE_YEAR' column in Fires table to avoid transitive dependency.
 2. Since OWNER_DESCR was dependent on OWNER_CODE (Non-key), we separated them from Fires table into a new table called Owner and made OWNER_CODE the primary key of that table.
 3. Since FIRE_SIZE_CLASS was dependent on FIRE_SIZE (Non-key), we separated them from Fires table into a new table called FireSize and made fire_size the primary key of that table.
 4. Since MTBS_FIRE_NAME was dependent on MTBS_ID (Non-key), we separated them from Fires table into a new table called MTBS and made mtbs_id the primary key of that table.
 5. Since STAT_CAUSE_DESCR was dependent on STAT_CAUSE_CODE(Non-key), we separated them from Fires table into a new table called Cause and made STAT_CAUSE_CODE the primary key of that table.
 6. Since country was dependent on state (Non-key), we separated them from NWCG table into a new table called NWCG_State_Country and made state the primary key of that table.
 - Thus, the data model has **no transitive dependency**. Hence, **it is in 3NF**.
- In this way, the entire data model now follows good database and normalization principles and is normalized to the best of our understanding.

Entity-Relationship Diagram



Physical Database

Assumptions/Notes About Data Set

Assumptions made about data such as empty fields, sparse data, bad data, etc.

- From the given data, we observed that ICS_209_Incident_Number is not unique for every ICS_209_NAME, and it also contains Null values. Hence, we introduced a new column i.e. ics_209_incident_id as Primary Key in ICS table, to enforce Unique and Not Null constraints of Primary Key Integrity.
- From the given data, we observed that FIPS_CODE is not unique for every FIPS_NAME, and it also contains Null values. Hence, we introduced a new column i.e. FIPS_id as Primary Key in FIPS table, to enforce Unique and Not Null constraints of Primary Key Integrity.
- From the given data, we observed that SOURCE_REPORTING_UNIT is not unique for every SOURCE_REPORTING_UNIT_NAME. Hence, we introduced a new column i.e. Source_Reported_id as Primary Key in Source_Reported table, to enforce Unique constraint of Primary Key Integrity.
- From the given data, we observed that SOURCE_SYSTEM can have different SOURCE_SYSTEM_TYPE. Hence, we introduced a new column i.e. Source_System_id as Primary Key in Source_System table, to enforce Unique constraint of Primary Key Integrity.
- From the given data, there were Null values present in MTBS_ID and corresponding MTBS_NAME was null as well. Hence, in MTBS table, to incorporate Null values and to maintain Not Null integrity of primary key (mtbs_id), we changed Null value row of mtbs_id to '0' and kept corresponding mtbs_name as Null.

Screen shot of Physical Database objects

Fires

1 • `select * from Project.Fires;`

2

100% 1:2

Result Grid Filter Rows: Search Export/Import: Fetch rows:

FOD_ID	FPA_ID	LOCAL_FIRE_REPORT_ID	LOCAL_INCIDENT_ID	FIRE_CODE	FIRE_NAME	FIRE_YEAR	FireSize_fire_size	MTBS_mtbs_id	Cause_STAT_CAUSE_CODE	Owner_OW
1	FS-1418826	1	PNF-47	BJ8K	FOUNTAIN	2005	0.1	HULL	9	5
2	FS-1418827	13		AAC0	PIGEON	2004	0.25	HULL	1	5
3	FS-1418835	27	021	A32W	SLACK	2004	0.1	HULL	5	13
4	FS-1418845	43	6	HULL	DEER	2004	0.1	HULL	1	5
5	FS-1418847	44	7	HULL	STEVENOT	2004	0.1	HULL	1	5
6	FS-1418849	54	8	HULL	HIDDEN	2004	0.1	HULL	1	5
7	FS-1418851	58	9	HULL	FORK	2004	0.1	HULL	1	5
8	FS-1418854	3	02	BK5X	SLATE	2005	0.8	HULL	5	13
9	FS-1418856	5	03	BLPQ	SHASTA	2005	1	HULL	5	13
10	FS-1418859	61	10	HULL	TANGLEFOOT	2004	0.1	HULL	1	5
11	FS-1418861	64	11	HULL	FORK #2	2004	0.1	HULL	1	5
12	FS-1418863	71	14	HULL	MOKO #2	2004	0.1	HULL	1	5
13	FS-1418865	81	26	HULL	WIGHTS	2004	0.1	HULL	9	5
14	FS-1418869	99	21	HULL	TIGER	2004	6	HULL	4	14
15	FS-1418874	102	28	HULL	BROWN	2004	0.2	HULL	1	5
16	FS-1418876	103	22	HULL	SOPIAGO	2004	0.1	HULL	1	5
17	FS-1418878	109	23	HULL	POWER	2004	16823	CA38502120...	2	5
18	FS-1418881	111	35	BHA3	FREDS	2004	7700	CA38787120...	2	5
19	FS-1418884	115	22	HULL	HARTLESS	2004	0.1	HULL	5	13
20	FS-1418887	8	NM-LNF-008	AVPA	SPRING	2004	0.1	HULL	1	5
21	FS-1418888	12	NM-LNF-012	A49Y	ANGRY	2004	0.1	HULL	1	13
22	FS-1418893	13	NM-LNF-013	A5DG	DARK	2004	0.1	HULL	1	5
23	FS-1418894	16	NM-LNF-016	AVP4	CHURCH #2	2004	0.1	HULL	1	5

Fires 1

Action Output

Time	Action	Response	Duration / F
43 03:14:29	select * from Project.Fires LIMIT 0, 50000	50000 row(s) returned	0.0015 sec /

1 • `select count(*) from Project.Fires;`

2

100% 16:1

Result Grid Filter Rows: Search Export:

count(*)
1880465

Cause

1 • `SELECT * FROM Project.Cause;`

100% 1:1

Result Grid Filter Rows: Search Export/Import:

STAT_CAUSE_CODE	STAT_CAUSE_DESCR
1	Lightning
2	Equipment Use
3	Smoking
4	Campfire
5	Debris Burning
6	Railroad
7	Arson
8	Children
9	Miscellaneous
10	Fireworks
11	Powerline
12	Structure
13	Missing/Undefined
HULL	

Cause 1

Action Output

Time	Action	Response	Duration / F
46 03:16:16	SELECT * FROM Project.Cause LIMIT 0, 50000	13 row(s) returned	0.00032 sec

Controlled_Data

```
1 •  SELECT * FROM Project.Controlled_Data;
```

Result Grid Filter Rows: Search Edit: Export/Import: Fetch rows:

Fires_FOD_ID	CONT_DATE	CONT_DOY	CONT_TIME	COMPLEX_NAME
3	2453156.5	152	2024	NULL
4	2453189.5	185	1400	NULL
5	2453189.5	185	1200	NULL
6	2453187.5	183	1600	NULL
7	2453188.5	184	1400	NULL
8	2453437.5	67	1600	NULL
9	2453444.5	74	1700	NULL
10	2453188.5	184	1800	NULL
11	2453189.5	185	1900	NULL
12	2453189.5	185	2000	NULL
13	2453251.5	247	1630	NULL
14	2453276.5	272	1800	NULL
15	2453281.5	277	1500	NULL
16	2453281.5	277	1400	NULL
17	2453299.5	295	1000	NULL
18	2453295.5	291	1800	NULL
19	2453330.5	326	1630	NULL
20	2453160.5	156	1400	NULL
21	2453175.5	171	1940	NULL
22	2453178.5	174	1100	NULL
23	2453182.5	178	1300	NULL
24	2453188.5	184	1110	NULL
25	2453192.5	188	1450	NULL

Controlled_Data 1 Action Output

Time	Action	Response	Duration / F
03:16:37	SELECT * FROM Project.Controlled_Data LIMIT 0, 50000	50000 row(s) returned	0.0084 sec

```
1 •  SELECT count(*) FROM Project.Controlled_Data;
```

Result Grid Filter Rows: Search Export:

count(*)
1880465

Discovery_Data

```
1 •  SELECT * FROM Project.Discovery_Data;
```

Result Grid Filter Rows: Search Edit: Export/Import: Fetch rows:

Fires_FOD_ID	DISCOVERY_DATE	DISCOVERY_DOY	DISCOVERY_TIME
3	2453156.5	152	1921
4	2453184.5	180	1600
5	2453184.5	180	1600
6	2453186.5	182	1800
7	2453187.5	183	1800
8	2453437.5	67	1300
9	2453444.5	74	1200
10	2453187.5	183	1800
11	2453188.5	184	1830
12	2453188.5	184	1730
13	2453251.5	247	1600
14	2453276.5	272	1200
15	2453281.5	277	0630
16	2453281.5	277	0930
17	2453284.5	280	1415
18	2453291.5	287	1618
19	2453329.5	325	2000
20	2453160.5	156	1030
21	2453175.5	171	1530
22	2453177.5	173	1440
23	2453181.5	177	1910
24	2453187.5	183	1250
25	2453187.5	183	1712

Discovery_Data 1 Action Output

Time	Action	Response	Duration / F
03:17:52	SELECT * FROM Project.Discovery_Data LIMIT 0, 50000	50000 row(s) returned	0.011 sec / C

```
1 •  SELECT count(*) FROM Project.Discovery_Data;
2
3
4
5
```

100% 1:6

Result Grid Filter Rows: Search Export:

count(*)
1880465

FIPS

```
1 •  SELECT * FROM Project.FIPS;
```

100% 1:1

Result Grid Filter Rows: Search Edit: Export/Import:

FIPS_id	FIPS_CODE	FIPS_NAME
1	063	Plumas
2	061	Placer
3	017	El Dorado
4	003	Alpine
5	005	Amador
6	HULL	HULL
7	027	Lincoln
8	017	Deschutes
9	021	Buncombe
10	113	Macon
11	027	Caldwell
12	011	Avery
13	009	Converse
14	069	Wheeler
15	037	Lake
16	033	Sheridan
17	027	Grays Harbor
18	053	Lincoln
19	089	Sanders
20	049	Utah
21	005	Coconino
22	019	Johnson
23	061	Union

FIPS 1

Action Output

Time	Action	Response	Duration / F
03:18:33	SELECT * FROM Project.FIPS LIMIT 0, 50000	2694 row(s) returned	0.015 sec / C

Fire_Location

```
1 •  SELECT * FROM Project.Fire_Location;
```

100% 1:7

Result Grid Filter Rows: Search Edit: Export/Import: Fetch rows:

Fire_Location_id	LATITUDE	LONGITUDE	STATE	COUNTY
1	40.03694444	-121.00583333	CA	63
2	38.93305556	-120.40444444	CA	61
3	38.98416667	-120.73555556	CA	17
4	38.55916667	-119.91333333	CA	3
5	38.55916667	-119.93055556	CA	3
6	38.63527778	-120.10361111	CA	5
7	38.68833333	-120.15333333	CA	17
8	40.96805556	-122.43388889	CA	HULL
9	41.23361111	-122.28333333	CA	HULL
10	38.54833333	-120.14916667	CA	5
11	38.69166667	-120.15972222	CA	17
12	38.5275	-120.10611111	CA	5
13	38.78666667	-120.19333333	CA	17
14	38.43333333	-120.51	CA	5
15	38.67583333	-120.27972222	CA	17
16	38.56416667	-120.54222222	CA	17
17	38.52333333	-120.21166667	CA	5
18	38.78	-120.26	CA	17
19	38.945	-120.41166667	CA	17
20	33.44083333	-105.72055556	NM	27
21	33.30722222	-105.62861111	NM	27
22	33.44444444	-105.76805556	NM	27
23	33.55944444	-105.76611111	NM	27

Fire_Location 1

Action Output

Time	Action	Response	Duration / F
52 03:18:49	SELECT * FROM Project.Fire_Location LIMIT 0, 50000	50000 row(s) returned	0.0065 sec

```
1 •  SELECT count(*) FROM Project.Fire_Location;
```

100% 1:6

Result Grid Filter Rows: Search Export:

count(*)
1585927

FireSize

```
1 •  SELECT * FROM Project.FireSize;
```

100% 1:1

Result Grid Filter Rows: Search Edit: Export/Import: Fetch rows:

fire_size	fire_size_class
0.0001	A
0.0002	A
0.00022	A
0.00034	A
0.0004	A
0.000459	A
0.0008	A
0.0009	A
0.001	A
0.00138	A
0.00159	A
0.0016	A
0.002	A
0.0025	A
0.0027	A
0.003	A
0.004	A
0.005	A
0.006	A
0.0062	A
0.0068	A
0.007	A
0.008	A

FireSize 1

Action Output

Time	Action	Response	Duration / F
54 03:19:56	SELECT * FROM Project.FireSize LIMIT 0, 50000	13602 row(s) returned	0.0077 sec

ICS

1 • `SELECT * FROM Project.ICS;`

100% 1:1

Result Grid Filter Rows: Search Edit: Export/Import:

ics_209_incident_id	ics_209_incident_number	ics_209_name
3	CA-ENF-18044	FREDS
4	NC-NCS-050201401	Austin Creek
5	CA-LPF-1353	CHIMINEAS
6	AZ-TNF-105	THREE FIRE COMPLEX
7	CA-SQF-1857	NINE
8	AZ-TNF-136	Oak
9	AZ-TNF-139	Greenback
10	AZ-TNF-140	SALOME
11	AZ-TNF-?	Two Bar
12	NM-SNF-027	Mesa Camino
13	LA-KIF-050004	Box
14	MT-CNF-010	SAWMILL GULCH
15	MT-KNF-075	CAMP 32
16	LA-KIF-050004	East Taylor
17	LA-KIF-050012	Mothers Day
18	MT-CNF1-045	PAGET
19	LA-KIF-	Coochie Brake
20	CA-SQF-2233	WISHON
21	NM-GNF-01	East Fork
22	MO-MTF-000097	WOLF PEN HOLLOW
23	KS-PSF-903	Brother
24	AZ-ASF-163	Line
25	CO-PSF-875	Carizo Complex

ICS 1 Apply

Action Output

Time	Action	Response	Duration / F
55 03:20:10	SELECT * FROM Project.ICS LIMIT 0, 50000	23314 row(s) returned	0.0022 sec

MTBS

1 • `SELECT * FROM Project.MTBS;`

100% 1:1

Result Grid Filter Rows: Search Edit: Export/Import:

mtbs_id	mtbs_fire_name
AK5683615710019920517	YANTARNI
AK5759815753720050529	PILOT POINT
AK5761615232120150828	TWIN CREEKS
AK5884115754320150622	COPENHAGEN CREEK
AK588415640120150621	PAULS CREEK
AK5890415740520150704	COPENHAGEN
AK5899315855220121102	SNAKE RIVER
AK5903315365920030428	KAMISHAK BAY
AK5925816024020150622	GECIAK LAKE
AK5955415836020150711	LITTLE KOKWOK
AK5960815825420150621	KOKWOK RIVER
AK5975315141520050429	TRACY AVE
AK5978615491320030523	ILIAMNA
AK6000115759420130728	NUYAKUK RIVER
AK6002415119720070619	CARIBOU HILLS
AK6011915081019940830	WINDY POINT
AK6012815061920040814	GLACIER CREEK
AK6015015550020130619	NIKUBUMA LAKE
AK6015115098020050711	FOX CREEK WFU
AK6025715983220150620	LITTLE SWIFT CREEK
AK6027615391520130627	CURRANT CREEK
AK6030315782720150623	KING SALMON RIVER 1
AK6037615081420090630	SHANTA CREEK

MTBS 1 Apply

Action Output

Time	Action	Response	Duration / F
60 03:22:27	SELECT * FROM Project.MTBS LIMIT 0, 50000	10482 row(s) returned	0.0066 sec

Owner

```
1 • SELECT * FROM Project.Owner;
```

100% 1:1

Result Grid Filter Rows: Search Edit: Export/Import:

OWNER_CODE	OWNER_DESCR
2	BIA
3	NPS
4	FWS
5	USFS
6	OTHER FEDERAL
7	STATE
8	PRIVATE
9	TRIBAL
10	BOR
11	COUNTY
12	MUNICIPAL/LOC...
13	STATE OR PRIV...
14	MISSING/NOT S...
15	UNDEFINIED FE...
16	UNDEF

Owner 1

Action Output

Time	Action	Response	Duration / F
64 03:22:55	SELECT * FROM Project.Owner LIMIT 0, 50000	16 row(s) returned	0.0035 sec

Source_Reporting

```
1 • SELECT * FROM Project.Source_Reporting;
```

100% 1:1

Result Grid Filter Rows: Search Edit: Export/Import:

Source_Reporting_id	SOURCE_REPORTING_UNIT	SOURCE_REPORTING_UNIT_NAME
3	0514	Shasta-Trinity National Forest
4	0308	Lincoln National Forest
5	0601	Deschutes National Forest
6	0811	National Forests in North Carolina
7	0206	Medicine Bow-Routt National Forest
8	0614	Umatilla National Forest
9	0212	Pike San Isabel National Forest
10	0607	Ochoco National Forest
11	0602	Fremont-Winema National Forest
12	0202	Bighorn National Forest
13	0114	Kootenai National Forest
14	0418	Uinta National Forest
15	0307	Kaibab National Forest
16	0510	Six Rivers National Forest
17	0305	Coronado National Forest
18	0501	Angeles National Forest
19	0203	Black Hills National Forest
20	0310	Santa Fe National Forest
21	0417	Humboldt-Toiyabe National Forest
22	0302	Carson National Forest
23	0214	Shoshone National Forest
24	0502	Cleveland National Forest
25	0809	Ouachita National Forest

Source_Reporting 1

Action Output

Time	Action	Response	Duration / F
65 03:23:01	SELECT * FROM Project.Source_Reporting LIMIT 0, 50000	6646 row(s) returned	0.0084 sec

Source_System

```
1 • | SELECT * FROM Project.Source_System;
```

100% 1:1

Result Grid Filter Rows: Search Edit: Export/Import:

source_system_id	source_system	source_system_type
3	FWS-FMIS	FED
4	FA-NFIRS	NONFED
5	ST-NASF	NONFED
6	ST-AZAZS	NONFED
7	ST-MOMOS	NONFED
8	IA-AKACC	INTERAGCY
9	ST-MTMTS	NONFED
10	ST-SCSCS	NONFED
11	ST-COCOS	NONFED
12	ST-MEMES	NONFED
13	ST-ORORS	NONFED
14	ST-MSMSS	NONFED
15	ST-NMMNMS	NONFED
16	ST-SDSDS	NONFED
17	ST-UTUTS	NONFED
18	ST-WAWAS	NONFED
19	ST-WYYWS	NONFED
20	ST-OKOKS	NONFED
21	ST-WIWIS	NONFED
22	ST-MIMIS	NONFED
23	ST-KYKYS	NONFED
24	ST-ALALS	NONFED
25	ST-ARARS	NONFED

Source_System 1

Action Output

Time	Action	Response	Duration / F
66 03:23:08	SELECT * FROM Project.Source_System LIMIT 0, 50000	39 row(s) returned	0.0041 sec /

NWCG_Unit

```
1 • | SELECT * FROM Project.NWCG_Unit;
```

100% 1:1

Result Grid Filter Rows: Search Edit: Export/Import:

unitid	name	unitype	wildlandrole	agency	parent
CAMBCIFC	Canadian Interagency Forest Fire Centre	Interagency	Dispatch/Coordination Center	NULL	NULL
CAMBMBN	Manitoba Wildfire Program	International Country Subdivision(State/Province)	Incident Host Geographic	NULL	NULL
CANBNBN	New Brunswick Forest Fire Management Section	International Country Subdivision(State/Province)	Incident Host Geographic	NULL	NULL
CANLNLN	Newfoundland and Labrador Fire Management...	International Country Subdivision(State/Province)	Incident Host Geographic	NULL	NULL
CANSNSN	Nova Scotia Forest Protection	International Country Subdivision(State/Province)	Incident Host Geographic	NULL	NULL
CANTNTN	Northwest Territories Forest Management Division	International Country Subdivision(State/Province)	Incident Host Geographic	NULL	NULL
CAONONNN	Ontario Aviation	International Country Subdivision(State/Province)	Incident Host Geographic	NULL	NULL
CAPEPEN	Prince Edward Island Forests	International Country Subdivision(State/Province)	Incident Host Geographic	NULL	NULL
CAQCPCN	Parks Canada National Fire Management Division	International Country Subdivision(State/Province)	Incident Host Geographic	NULL	NULL
CAQQCCN	Quebec SOPFEU	International Country Subdivision(State/Province)	Incident Host Geographic	NULL	NULL
CASKSKN	Saskatchewan Wildfire Management	International Country Subdivision(State/Province)	Incident Host Geographic	NULL	NULL
CAYTYTN	Yukon Wildland Fire Management	International Country Subdivision(State/Province)	Incident Host Geographic	NULL	NULL
UM76NVR	Navassa Island National Wildlife Refuge	US Federal	Incident Host Geographic	FWS	NULL
USAKACC	Alaska Interagency Coordination Center	Interagency	Dispatch/Coordination Center	NULL	NULL
USAKADD	Anchorage Field Office	US Federal	Incident Host Geographic	BLM	NULL
USAKAFFMX	Anchorage Fire Department	US County/Local	Incident Host Geographic	C&L	NULL
USAKAHN	Ahtna Regional Native Corporation	Non-Government	Resource Provider Only	ANC	NULL
USAKAKA	Alaska Regional Office	US Federal	Incident Host Geographic	BIA	NULL
USAKAKD	Alaska Fire Service Resources	US Federal	Incident Host Geographic	BLM	NULL
USAKAKK	Alaska Fire Service Interagency Cache	Interagency	Fire/Radio Caches	NULL	NULL
USAKAKS	Central Office	US State	Incident Host Geographic	DVF	NULL
USAKALN	Aleut Regional Native Corporation	Non-Government	Resource Provider Only	ANC	NULL
USAKAMMMX	Anchorage Municipality	US County/Local	Incident Host Geographic	C&L	NULL

NWCG_Unit 1

Action Output

Time	Action	Response	Duration / F
62 03:22:44	SELECT * FROM Project.NWCG_Unit LIMIT 0, 50000	5867 row(s) returned	0.0082 sec ,

NWCG_Unit_Location

```
1 • SELECT * FROM Project.NWCG_Unit_Location;
```

100% 1:1

Result Grid Filter Rows: Search Edit: Export/Import:

code	NWCG_State_Country_state	gacc	geographicarea	NWCG_Unit_unitid
1AX	ID	USUTGBC	GB	USID1AX
1BCN	ID	USUTGBC	GB	USID1BCN
1BIN	ID	USUTGBC	GB	USID1BIN
1BN	ID	USUTGBC	GB	USID1BN
1BPN	ID	USUTGBC	GB	USID1BPN
1FN	ID	USUTGBC	GB	USID1FN
1GX	ID	USUTGBC	GB	USID1GX
1JN	ID	USUTGBC	GB	USID1JN
1MN	ID	USUTGBC	GB	USID1MN
1NAS	CA	USCAON...	CA	USCA1NAS
1ON	ID	USUTGBC	GB	USID1ON
1PX	ID	USUTGBC	GB	USID1PX
1TN	ID	USUTGBC	GB	USID1TN
1TPN	ID	USUTGBC	GB	USID1TPN
1VX	ID	USUTGBC	GB	USID1VX
1WX	ID	USUTGBC	GB	USID1WX
2BN	ID	USUTGBC	GB	USID2BN
2CX	ID	USUTGBC	GB	USID2CX
2FN	ID	USUTGBC	GB	USID2FN
2LLN	ID	USUTGBC	GB	USID2LLN
2PN	ID	USUTGBC	GB	USID2PN
3CN	ID	USUTGBC	GB	USID3CN
4BN	ID	USUTGBC	GB	USID4BN

NWCG_Unit_Location 1 Apply

Action Output

Time	Action	Response	Duration / F
63 03:22:49	SELECT * FROM Project.NWCG_Unit_Location LIMIT 0, 50000	5867 row(s) returned	0.0056 sec

NWCG_State_Country

```
1 • SELECT * FROM Project.NWCG_State_Country;
```

100% 1:1

Result Grid Filter Rows: Search Edit: Export/Import:

state	country
AK	US
AL	US
AR	US
AS	US
AZ	US
BC	CA
CA	US
CO	US
CT	US
DC	US
DE	US
FL	US
GA	US
GU	US
HI	US
IA	US
ID	US
IL	US
IN	US
KS	US
KY	US
LA	US
MA	US

NWCG_State_Country 1 Apply

Action Output

Time	Action	Response	Duration / F
67 03:34:59	SELECT * FROM Project.NWCG_State_Country LIMIT 0, 50000	68 row(s) returned	0.015 sec

Idx_Fires_Shape_node

```
1 • | SELECT * FROM Project.Idx_Fires_Shape_node;
```

100% 1:1

Result Grid Filter Rows: Search Edit: Export/Import: Fetch rows:

nodeno	data	parentnode
3	<0>	2235
4	<0>	56494
5	<0>	63763
6	<0>	81
7	<0>	63241
8	<0>	1719
9	<0>	42692
10	<0>	4842
11	<0>	81
12	<0>	765
13	<0>	42
14	<0>	817
15	<0>	2150
16	<0>	14129
17	<0>	8873
18	<0>	12536
19	<0>	7872
20	<0>	2727
21	<0>	12835
22	<0>	10041
23	<0>	911
24	<0>	63022
25	<0>	10224
idx_Fires_Shape_node	1	

Action Output

Time	Action	Response	Duration / F
03:20:52	SELECT * FROM Project.Idx_Fires_Shape_node LIMIT 0, 50000	50000 row(s) returned	0.0075 sec

```
1 • | SELECT count(*)| FROM Project.Idx_Fires_Shape_node;
```

100% 16:1

Result Grid Filter Rows: Search Export:

count(*)
72263

Idx_Fires_Shape_rowid

```
1 •  SELECT * FROM Project.idx_Fires_Shape_rowid;
```

100% 1:1

Result Grid Filter Rows: Search Edit: Export/Import: Fetch rows:

rowid	idx_Fires_Shape_node_nodeno	Fires_FOD_ID
3	6313	3
4	5777	4
5	5777	5
6	4515	6
7	111	7
8	5786	8
9	43366	9
10	1584	10
11	3812	11
12	1584	12
13	2210	13
14	41959	14
15	40999	15
16	12044	16
17	1820	17
18	6654	18
19	1273	19
20	3021	20
21	2553	21
22	1040	22
23	5185	23
24	7653	24
25	8940	25

idx_Fires_Shape_rowid 1 Apply

Action Output

Time	Action	Response	Duration / F
58 03:21:15	SELECT * FROM Project.idx_Fires_Shape_rowid LIMIT 0, 50000	50000 row(s) returned	0.011 sec / C

```
1 •  SELECT count(*) FROM Project.idx_Fires_Shape_rowid;
```

100% 1:6

Result Grid Filter Rows: Search Export:

count(*)
1880465

Data in the Database

Table Name	Primary Key	Foreign Key	# of Rows in Table
Fires	FOD_ID	1. ICS_ics_209_incident_id 2. FireSize_fire_size 3. MTBS_mtbs_id 4. Cause_STAT_CAUSE_CODE 5. Source_Reportin_g_ Source_Reportin_g_id 6. Source_System_source_system_id 7. Owner_OWNER_CODE 8. Fire_Location_Fire_Location_id 9. NWCG_Unit_reporting_unitid 10. FIPS_FIPS_id	1880465
Discovery_Data	Fires_FOD_ID	Fires_FOD_ID	1880465
Controlled_Data	Fires_FOD_ID	Fires_FOD_ID	1880465
Fire_Location	Fire_Location_id		1585927
Owner	OWNER_CODE		16
ICS	ics_209_incident_id		23314
FireSize	fire_size		13602
MTBS	mtbs_id		10482
Cause	STAT_CAUSE_CODE		13
Source_Reportin_g	Source_Reportin_g_id		6646
Source_System	source_system_id		39
FIPS	FIPS_id		2694
NWCG_Unit	unitid		5867
NWCG_Unit_Location	code NWCG_State_Country_state	NWCG_Unit_unitid NWCG_State_Country_state	5867
NWCG_State_Country	state		68
Idx_Fires_Shape_rowid	rowid	Idx_Fires_Shape_node_nodeno Fires_FOD_ID	1880465
Idx_Fires_Shape_node	nodeno		72263

SQL Queries

Query 1

Question

Q 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 best chance to win a share of the fund?

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

Here the beverage company would most likely give a share of the fund to the state which has highest count of fires caused due to debris burning – so that the state can use the fund to clean its debris and reduce the number of fires caused due to it in future.

Translation

Select State Name from Fire_Location table, Count(FOD_ID from Fires table) having cause debris burning from Fire_Location table
joined with Fires table on Fire_Location_id in Fires_Location table
joined with Cause table on stat_cause_code in Fires table
where STAT_CAUSE_DESCR is Debris
grouped by State Name
Sorted in descending order of number of fires
Limiting the number of rows to 1

Clean Up

Select State Name from Fire_Location, Count(FOD_ID from Fires) from Fire_Location
join Fires on Fire_Location_id in Fires_Location
join Cause on stat_cause_code in Fires
where STAT_CAUSE_DESCR like '%Debris%'
group by State Name
order by 2 desc
limit 1

Screen Shot of SQL Query and Results

```
5 •  select fl.state, count(fod_id) as 'Number of Fires having cause Debris Burning'  
6   from Fire_Location fl join  
7   Fires f on fl.Fire_Location_id = f.Fire_Location_Fire_Location_id join  
8   Cause c on c.stat_cause_code = f.Cause_STAT_CAUSE_CODE  
9   where c.STAT_CAUSE_DESCR like '%Debris%'  
10  group by fl.state  
11  order by 2 desc  
12  limit 1;  
13  
14
```

100% 1:16

Result Grid Filter Rows: Search Export: Fetch rows:

state	Number of Fires having cause Debris Burning
GA	88199

Result 5

Action Output

Time	Action	Response	Duration / Fetch Time
04:28:41	select fl.state, count(fod_id) as 'Number of Fires having cause Debris Burning' from Fire_Lo...	1 row(s) returned	9.666 sec / 0.00000...

Result Grid Form Editor Field Types Query Stats

Read Only

The state of Georgia(GA) had the most no. of fires(88199) caused due to Debris Burning. Hence, Georgia has the best chance to win.

Query 2

Question

Q 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 most appropriate.

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

Here we need to find top 3 forests where most fires are caused by children. That is, children are the cause of Fires. Such forests would need a rule of children being accompanied by an adult to prevent fires caused by them in future.

Assumption: We have considered NWCG_Reportng_Unit field for our results.

Translation

Select number of fires, Forest name, cause of fire from NWCG_Unit table
joined with Fires table on unitid in NWCG_Unit table
joined with Cause table on stat_cause_code in Fires table
where cause of fire was Children
grouped by Forest Name, Cause
sorted in descending order of number of fires
limiting the number of rows to 3

Clean Up

Select count(Name from NWCG_Unit), Forest name, Cause from NWCG_Unit
join Fires on unitid in NWCG_Unit
join Cause on stat_cause_code in Fires
where cause like '% Children%'
group by Forest Name, Cause
order by Number of Fires desc
Limit 3

Screen Shot of SQL Query and Results

```
9 •   select count(nu.name) as 'Number of Fires', nu.name as Forest, c.STAT_CAUSE_DESCR as Cause  
10    from NWCG_Unit nu  
11   join Fires f on nu.unitid = f.NWCG_Unit_reporting_unitid  
12   join Cause c on f.Cause_STAT_CAUSE_CODE = c.STAT_CAUSE_CODE  
13  where c.STAT_CAUSE_DESCR like '%Children%'  
14  group by nu.name, c.STAT_CAUSE_DESCR  
15  order by 1 desc  
16  limit 3;  
17
```

The screenshot shows a database query results interface. At the top, there is a code editor window displaying the SQL query. Below it is a results grid showing the output of the query. The results grid has three columns: 'Number of Fires', 'Forest', and 'Cause'. The data shows three rows: North Carolina Forest Service (7537 fires, Children cause), Georgia Forestry Commission (7190 fires, Children cause), and Florida Forest Service (4345 fires, Children cause). To the right of the results grid is a vertical toolbar with icons for different database functions: Result Grid, Form Editor, Field Types, and Query Plans.

Number of Fires	Forest	Cause
7537	North Carolina Forest Service	Children
7190	Georgia Forestry Commission	Children
4345	Florida Forest Service	Children

Action Output

Time	Action	Response	Duration / Fetch Time
39 19:56:09	select count(nu.name) as 'Number of Fires', nu.name as Forest, c.STAT_CAUSE_DESCR as C...	3 row(s) returned	0.525 sec / 0.00000...

The 3 forests: North Carolina Forest, Georgia Forest, Florida Forest had the most fires caused due to Children. Hence, these are the 3 forests where the ban would be the most appropriate.

Query 3

Question

Q 3) One advocacy group says Nature and not human actions 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)

As per the NWCG Report, Debris burn escapes are one of the leading specific causes of wildland fires. Windblown embers or fire creeping from the controlled burn area into un-cleared vegetation are the primary ignition mechanisms of the wildfire. Hence, we make 2 assumptions about wildfires when cause is debris.

Assumption 1: If we consider debris burning as both man-made and natural cause of fire, then we don't include it in any group. Hence, here we cannot support the claim of advocacy group.

Translation

```
Select number of fires due to Man-made Cause,  
select number of fires due to Natural cause from fires table where Cause_stat_code is in  
select stat_cause_code from Cause table where Cause is Lightning or Debris  
from Fires table  
where cause_stat_cause_code is not in  
select stat_cause_code from Cause table  
where cause is Lightning, Miscellaneous, Missing/Undefined or Debris
```

Clean Up

```
Select count(*) as 'Man-made Cause',  
select count(*) from Fires where Cause_stat_cause_code in (  
select stat_cause_code from Cause from Fires where stat_cause_descr = '%Lightning%' or  
stat_cause_descr = '%Debris%' )  
where cause_stat_cause_code not in (  
select stat_cause_code from Cause  
where stat_cause_descr = '%Lightning%' or stat_cause_descr = '%Miscellaneous%' or  
stat_cause_descr = '%Missing/Undefined%' or `stat_cause_descr = '%Debris%'` )
```

Screen Shot of SQL Query and Results

The screenshot shows a SQL query editor interface. The query itself is a multi-step process involving three tables: Fires, Cause, and Cause. It counts fires by cause and then filters those counts based on specific cause descriptions. The results are displayed in a grid, showing two rows: 'Man-made Cause' with a count of 682441 and 'Natural Cause' with a count of 707496. Below the results, a timeline shows the execution of the query, indicating it took 0.392 seconds.

```
47 • select count(*) as 'Man-made Cause',
48   (select count(*)
49     from Fires as f
50   where f.Cause_stat_code in (select c.stat_code
51                               from Cause c
52                               where c.stat_descr like '%Lightning%' or
53                                   c.stat_descr like '%Debris%')) as 'Natural Cause'
54
55   from Fires as f1
56   where f1.Cause_stat_code not in (select c1.stat_code
57                                     from Cause c1
58                                     where c1.stat_descr like '%Lightning%' or
59                                         c1.stat_descr like '%Miscellaneous%' or
60                                         c1.stat_descr like '%Missing/Undefined%' or
61                                         c1.stat_descr like '%Debris%');
61
```

Result Grid	Filter Rows:	Export:
Man-made Cause Natural Cause	682441	707496

Action Output

Time	Action	Response	Duration / Fetch Time
04:59:45	select count(*) as 'Man-made Cause', (select count(*) from Fires as f where f.Cause_stat_code in (select c.stat_code from Cause c where c.stat_descr like '%Lightning%' or c.stat_descr like '%Debris%')) as 'Natural Cause'	1 row(s) returned	0.392 sec / 0.000007...

The above query does support the statement claimed that Nature and not human actions is to blame for most wildfires.

Assumption 2: We consider debris burning as Natural cause. Now we can support the claim of advocacy group that nature and not human actions is to blame for most wildfires.

Translation

Select count of fires due to Man-made Cause,

Select count of Natural Cause Fires from fires table where Cause_stat_code is in

select stat_code from Cause table where Cause is Lightning

from Fires table

where cause_stat_code is not in

select stat_code from Cause table

where cause is Lightning, Miscellaneous, or Missing/Undefined

Clean Up

Select count(*) as 'Man-made Cause',

Select count(*) from Fires where Cause_stat_code in (

select stat_code from Cause from Fires where stat_descr = '%Lightning%')

where cause_stat_code not in (

Select stat_code from Cause

where stat_descr = '%Lightning%' or stat_descr = '%Miscellaneous%' or

stat_descr = '%Missing/Undefined%')

Screen Shot of SQL Query and Results

```
47 •   select count(*) as 'Man-made Cause',
48   (select count(*)
49     from Fires as f
50     where f.Cause_stat_cause_code in (select c.stat_cause_code
51                                         from Cause c
52                                         where c.stat_cause_descr like '%Lightning%')) as 'Natural Cause'
53
54   from Fires as f1
55   where f1.Cause_stat_cause_code not in (select c1.stat_cause_code
56                                         from Cause c1
57                                         where c1.stat_cause_descr like '%Lightning%' or
58                                         c1.stat_cause_descr like '%Miscellaneous%' or
59                                         c1.stat_cause_descr like '%Missing/Undefined%');
60
```

Result Grid Filter Rows: Search Export:

Result Grid Form Editor

Man-made Cause	Natural Cause
1111469	278468

Action Output

Time	Action	Response	Duration / Fetch Time
90 04:58:09	select count(*) as 'Man-made Cause', (select count(*) from Fires as f where f.Cause_stat_cause_code in (select c.stat_cause_code from Cause c where c.stat_cause_descr like '%Lightning%')) as 'Natural Cause'	1 row(s) returned	0.337 sec / 0.000009...

The above query does not support the statement claimed that Nature and not human actions is to blame for most wildfires.

Query 4

Question

Q 6) What were the forests that had more than one fire that lasted more than two days?

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

We have considered NWCG_Unit table which contains a list of Forest names.

Translation

Select distinct Forest Name from NWCG_Unit table

joined with Fires table on NWCG_Unit_reporting_unitid in NWCG_Unit table

joined with Controlled_Data table on FOD_ID in Fires table

joined with Discovery_Data table on FOD_ID in Fires table

where fire lasted for more than two days

grouped by Forest Name

having occurrence of more than one fire given by FOD_ID in Fires table

Clean Up

Select distinct Name from NWCG_Unit

Join Fires on NWCG_Unit_reporting_unitid in NWCG_Unit

Join Controlled_Data on FOD_ID in Fires

Join Discovery_Data on FOD_ID in Fires

where datediff(CONT_DATE in Controlled_Data, DISCOVERY_DATE in Discovery_Data) > 2 AND

FOD_ID in Fires > 1

Group by Name in NWCG_Unit

Screen Shot of SQL Query and Results

```
28 •  select distinct nu.name as Forest
29   from NWCG_Unit nu
30   join Fires f on nu.unitid = f.NWCG_Unit_reporting_unitid
31   join Controlled_Data as cd on cd.Fires_FOD_ID = f.FOD_ID
32   join Discovery_Data as dd on dd.Fires_FOD_ID = f.FOD_ID
33   where datediff(cd.CONT_DATE, dd.DISCOVERY_DATE) > 2
34   group by nu.name
35   having count(f.FOD_ID) > 1;
36
```

Result Grid Filter Rows: Search Export:

Forest
Alaska Regional Office
Arctic National Wildlife Refuge
Alaska State Office
Bering Land Bridge National Preserve
Chugach National Forest
Valdez-Copper River Area Forestry
Delta Area Forestry
Denali National Park & Preserve
Fairbanks Area Forestry
Alaska Fire Service - Galena Zone
Gates of the Arctic National Park &...
Haines Area Forestry
Innoko National Wildlife Refuge
Katmai National Park & Preserve
Kanuti National Wildlife Refuge

Result 9 Read Only

Action Output ◊

Time	Action	Response	Duration / Fetch Time
20 12:49:20	select distinct nu.name as Forest from NWCG_Unit nu join Fires f on nu.unitid = f.NWCG_Unit_reporting_...	587 row(s) returned	62.136 sec / 0.00014...

Query 5

Question

Q 7) Which state had more fires in the second half of a calendar year than the first half of the calendar year?

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

We have shown total number of fires in first half of calendar year as well as those in second half for side by side comparison.

Translation

```
Select state2, count1, count2 from (
  Select state from Fire_Location table, count(FOD_ID from Fires table)
  Where Discovery_Date from Discovery_Data table is greater than 6 AND
    Fire_Location_ID from Fire_Location table is equal to Fire_Location_ID from Fires table AND
    FOD_ID from Fires table is equal to FOD_ID from Discovery_Data table
  Grouped by state from Fire_Location table )
Joined with
  Select state from Fire_Location table, count(FOD_ID from Fires table) from
    Fire_Location table, Fires table, Discovery_Data table
  Where Discovery_Date from Discovery_Data table is less than or equal to 6 AND
    Fire_Location_ID from Fire_Location table is equal to Fire_Location_ID from Fires table AND
    FOD_ID from Fires table is equal to FOD_ID from Discovery_Data table
  Grouped by state from Fire_Location table ON state1 equals state2
```

Clean Up

```
Select state2, count1, count2 from (
  Select state from Fire_Location, count(FOD_ID from Fires)
  Where Discovery_Date from Discovery_Data > 6 AND
    Fire_Location_ID from Fire_Location = Fire_Location_ID from Fires AND
    FOD_ID from Fires = FOD_ID from Discovery_Data
  Group by state from Fire_Location )
Joined with
  Select state from Fire_Location, count(FOD_ID from Fires) from
    Fire_Location , Fires , Discovery_Data
  Where Discovery_Date from Discovery_Data <= 6 AND
    Fire_Location_ID from Fire_Location = Fire_Location_ID from Fires AND
    FOD_ID from Fires = FOD_ID from Discovery_Data
  Group by state from Fire_Location ON s1.state1 = s2.state2
```

Screen Shot of SQL Query and Results

```
118
119 • select s2.state2 as State, s1.count1 'Total Fires in First half of year', s2.count2 as 'Total Fires in Second half of year'
120   from (select fl.state as state2, count(f.FOD_ID) as count2
121     from Fire_Location fl, Fires f, Discovery_Data dd
122    where month(dd.Discovery_Date) > 6 and fl.Fire_Location_id = f.Fire_Location_fire_location_id and
123      f.FOD_ID = dd.Fires_fod_id
124    group by fl.state) s2 join (select fl.state as state1, count(f.FOD_ID) as count1
125      from Fire_Location fl, Fires f, Discovery_Data dd
126    where month(dd.Discovery_Date) <= 6 and fl.Fire_Location_id = f.Fire_Location_fire_location_id and
127      f.FOD_ID = dd.Fires_fod_id
128    group by fl.state) s1 on s1.state1 = s2.state2
129  where s2.count2 > s1.count1;
130
```

100% 13:132 1 error found

Result Grid Filter Rows: Search Export: Result Grid

State	Total Fires in First half of year	Total Fires in Second half of year
RI	219	260

Result 26 Read Only

Action Output

Time	Action	Response	Duration / Fetch Time
06:53:57	select s2.state2 as State, s1.count1 'Total Fires in First half of year', s2.count2 as 'Total Fires in Second half of...	1 row(s) returned	42.526 sec / 0.00001...

The state of Rhode Island(RI) had more fires in the second half of the year than in the first half.

Query 6

Question

Q 8) Which forest had the most number of fires?

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

We have considered NWCG_Unit and NWCG_Unit_reporting tables for our results.

Translation

Select count(Name from NWCG_Unit table), Forest Name from NWCG_Unit table

Joined with Fires table on unitid in NWCG_Unit table

Grouped by Forest name

Sorted in descending order of Number of Fires

Limiting the number of records to 1

Clean Up

Select count(Name from NWCG_Unit), Forest Name from NWCG_Unit

Join Fires on unitid in NWCG_Unit

Group by name in NWCG_Unit

Order by 1 desc

Limit 1

Screen Shot of SQL Query and Results

The screenshot shows a SQL query results interface. At the top, there is a code editor window containing the following SQL query:

```
3 • select count(nu.name) as 'Number of Fires', nu.name as Forest
4   from NWCG_Unit nu join Fires f on nu.unitid = f.NWCG_Unit_reporting_unitid
5   group by nu.name
6   order by 1 desc
7   limit 1;
```

Below the code editor is a result grid titled "Result Grid". The grid has two columns: "Number of Fires" and "Forest". There is one row of data: "167123" under "Number of Fires" and "Georgia Forestry Commission" under "Forest".

On the right side of the interface, there is a vertical toolbar with icons for "Result Grid", "Form Editor", "Field Types", "Query Stats", and "Execution Plan".

At the bottom of the interface, there is a status bar with the text "Result 13" and a "Read Only" button. Below the status bar, there is a table titled "Action Output" with columns "Time", "Action", "Response", and "Duration / Fetch Time". The table contains one row of data: "40 19:57:44 select count(nu.name) as 'Number of Fires', nu.name as Forest from NWCG_Unit nu join Fire... 1 row(s) returned 2.393 sec / 0.00000...".

The Georgia Forest had the most number of fires(167123).

Data Review for MongoDB

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

We have considered two primary collections – Fires and NWCG_Unit.

Physical Mongo Database

Assumptions/Notes About Data Set

We have considered two collections – Fires and NWCG_Unit as the primary collections.

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

The screenshot displays the MongoDB Compass interface with two databases open:

- Project.fires**: Contains 1.9 million documents across 10 collections. The "fires" collection is selected, showing a table with columns: FPA_ID, SOURCE_SYSTEM_TYPE, SOURCE_SYSTEM, NWCG_REPORTING_AGENCY, and NWCG_REPORTING_UNIT_ID. The data shows various FPA IDs and reporting agencies like USCAENF and USCAFNF.
- Project.nwcfg**: Contains 5.9k documents across 10 collections. The "nwcfg" collection is selected, showing a table with columns: UnitId, GeographicArea, Gacc, WildlandRole, and UnitType. The data shows various unit IDs and types like USAKCEKX and US County/Local.

```
To enable free monitoring, run the following command: db.enableFreeMonitoring()
To permanently disable this reminder, run the following command: db.disableFreeMonitoring()
---

> show dbs
Project 0.287GB
admin 0.000GB
config 0.000GB
local 0.177GB
testemp 0.000GB
> use Project
switched to db Project
> show collections
fires
nwcg
> db.fires.find().pretty();
```



```
"SOURCE_SYSTEM" : "FS-FIRESTAT",
"NWCG_REPORTING_AGENCY" : "FS",
"NWCG_REPORTING_UNIT_ID" : "USMMLNF",
"NWCG_REPORTING_UNIT_NAME" : "Lincoln National Forest",
"SOURCE_REPORTING_UNIT" : "0308",
"SOURCE_REPORTING_UNIT_NAME" : "Lincoln National Forest",
"LOCAL_FIRE_REPORT_ID" : "13",
"LOCAL INCIDENT_ID" : "NM-LNF-013",
"Fire_CODE" : "A5DG",
"Fire_NAME" : "DARK",
"ICS_209_INCIDENT_NUMBER" : "",
"ICS_209_NAME" : "",
"MTBS_ID" : "",
"MTBS_FIRE_NAME" : "",
"COMPLEX_NAME" : "",
"Fire_YEAR" : 2004,
"DISCOVERY_DATE" : 2453177.5,
"DISCOVERY_DOY" : 173,
"DISCOVERY_TIME" : "1440",
"STAT_CAUSE_CODE" : 1,
"STAT_CAUSE_DESCR" : "Lightning",
"CONT_DATE" : 2453178.5,
"CONT_DAY" : 174,
"CONT_TIME" : "1100",
"Fire_SIZE" : 0.1,
"Fire_SIZE_CLASS" : "A",
"Latitude" : 33.44444444,
"Longitude" : -105.7680556,
"Owner_CODE" : 5,
"Owner_DESCR" : "USFS",
"State" : "NM",
"County" : "27",
"FIPS_CODE" : "027",
"FIPS_NAME" : "Lincoln",
"Shape" : "<>"}
```

```
Type "it" for more
```



```
"SOURCE_SYSTEM" : "FS-FIRESTAT",
"NWCG_REPORTING_AGENCY" : "FS",
"NWCG_REPORTING_UNIT_ID" : "USNCNCF",
"NWCG_REPORTING_UNIT_NAME" : "National Forests in North Carolina",
"SOURCE_REPORTING_UNIT" : "0811",
"SOURCE_REPORTING_UNIT_NAME" : "National Forests in North Carolina",
"LOCAL_FIRE_REPORT_ID" : "2",
"LOCAL INCIDENT ID" : "002",
"Fire_CODE" : "RJ66",
"Fire_NAME" : "HOWARD GAP",
"ICS_209_INCIDENT_NUMBER" : "",
"ICS_209_NAME" : "",
"MTBS_ID" : "",
"MTBS_FIRE_NAME" : "",
"COMPLEX_NAME" : "",
"Fire_YEAR" : 2005,
"DISCOVERY_DATE" : 2453397.5,
"DISCOVERY_DOW" : 27,
"DISCOVERY_TIME" : "2200",
"STAT_CAUSE_CODE" : 7,
"STAT_CAUSE_DESCR" : "Arson",
"CONT_DATE" : 2453398.5,
"CONT_DOW" : 28,
"CONT_TIME" : "0300",
"Fire_SIZE" : 50.3,
"Fire_SIZE_CLASS" : "C",
"Latitude" : 35.00027778,
"Longitude" : -83.35111111,
"OWNER_CODE" : 5,
"OWNER_DESCR" : "USFS",
"STATE" : "NC",
"COUNTY" : "113",
"FIPS_CODE" : "113",
"FIPS_NAME" : "Macon",
"Shape" : "<>"

}
Type "it" for more
> db.nwcg.find().pretty();
```

```
Name" : "Cordova Volunteer Fire Department"
}
{
  "_id" : ObjectId("5ded4b55d66c05fac9ce040f"),
  "row_names" : "19",
  "OBJECTID" : 19,
  "UnitId" : "USAKCRCX",
  "GeographicArea" : "AK",
  "Gacc" : "USAKACC",
  "WildlandRole" : "Incident Host Geographic",
  "UnitType" : "US County/Local",
  "Department" : "AK",
  "Agency" : "C8L",
  "Parent" : "",
  "Country" : "US",
  "State" : "AK",
  "Code" : "CRCX",
  "Name" : "Crystal Creek Volunteer Fire Department"
}
{
  "_id" : ObjectId("5ded4b55d66c05fac9ce0410"),
  "row_names" : "20",
  "OBJECTID" : 20,
  "UnitId" : "USAKCRS",
  "GeographicArea" : "AK",
  "Gacc" : "USAKACC",
  "WildlandRole" : "Incident Host Geographic",
  "UnitType" : "US State",
  "Department" : "AK",
  "Agency" : "DVF",
  "Parent" : "",
  "Country" : "US",
  "State" : "AK",
  "Code" : "CRS",
  "Name" : "Valdez-Copper River Area Forestry"
}
Type "it" for more
>
```

```
Type "it" for more
> db.nwcg.find().count();
5867
> db.fires.find().count();
1880465
>
```

Data in the Database

Collection Name	Relationships with Other Collections (if any)	# of Documents in Collection
Fires		1880465
NWCG		5867

MongoDB Queries/Code

Query 1

Question

Q 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 best chance to win a share of the fund?

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

Here the beverage company would most likely give a share of the fund to the state which has highest count of fires caused due to debris burning – so that the state can use the fund to clean its debris and reduce the number of fires caused due to it in future.

Translation

Find STAT_CAUSE_DESCR matching to ‘Debris Burning’ from Fires Collection where

Grouped by State field

Sorted in descending order of fire count

Limiting the number of documents to 1

Screen Shot of MongoDB Query/Code and Results

```
> db.Fires.aggregate([
...   {
...     $match:{"STAT_CAUSE_DESCR":"Debris Burning"}
...   },
...   {
...     $group:{
...       _id:"$STATE",
...       fire_count:{$sum:1}
...     }
...   },
...   {
...     $sort:{fire_count: -1}
...   },
...   {
...     $limit:1
...   }
[... ])
{ "_id" : "GA", "fire_count" : 88199 }
```

The state of Georgia(GA) had the most no. of fires(88199) caused due to Debris Burning. Hence, Georgia has the best chance to win.

Query 2

Question

Q 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 most appropriate.

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

Here we need to find top 3 forests where most fires are caused by children. That is, children are the cause of Fires. Such forests would need a rule of children being accompanied by an adult to prevent fires caused by them in future.

Assumption: We have considered NWCG_Reportng_Unit field for our results.

Translation

Find STAT_CAUSE_DESCR matching to 'Children' from Fires Collection where

Grouped by NWCG_Reportng_Unit_Name field

Sorted in descending order of fire count

Limiting the number of documents to 3

Screen Shot of MongoDB Query/Code and Results

```
>
> db.Fires.aggregate([
...   {
...     $match:{"STAT_CAUSE_DESCR":"Children"}
...   },
...   {
...     $group:{
...       _id:"$NWCG_REPORTING_UNIT_NAME",
...       fire_count:{$sum:1}
...     }
...   },
...   {
...     $sort:{fire_count: -1}
...   },
...   {
...     $limit:3
...   },
...   {
...     $project:{
...       _id: 1,
...       fire_count: 1,
...     }
...   }
... ])
{ "_id" : "North Carolina Forest Service", "fire_count" : 7537 }
{ "_id" : "Georgia Forestry Commission", "fire_count" : 7190 }
{ "_id" : "Florida Forest Service", "fire_count" : 4345 }
```

The 3 forests: North Carolina Forest, Georgia Forest, Florida Forest had the most fires caused due to Children. Hence, these are the 3 forests where the ban would be the most appropriate.

Query 3

Question

Q 8) Which forest had the most number of fires?

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

We have considered NWCG_Reportng_Unit_Name for our results.

Translation

Project NWCG_Reportng_Unit_Name from Fires Collection where

Fire count is maximum

Sorted in descending order of fire count field

Limiting the number of documents to 1

Screen Shot of MongoDB Query/Code and Results

```
> db.Fires.aggregate([
...   {
...     $group:{
...       _id:"$NWCG_REPORTING_UNIT_NAME",
...       fire_count:{$sum:1}
...     }
...   },
...   {
...     $sort:{fire_count: -1}
...   },
...   {
...     $limit:1
...   },
...   {
...     $project:{
...       _id: 1,
...       fire_count: 1,
...     }
...   }
... ])
{ "_id" : "Georgia Forestry Commission", "fire_count" : 167123 }
> █
```

The Georgia Forest had the most number of fires(167123).

Query 4

Question

Q3)

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

Translation

Screen Shot of MongoDB Query/Code and Results

```
> db.Fires.aggregate([
...   {
...     $project: {
...       "STAT_CAUSE_DESCR": 1,
...       "Cause": {
...         $switch: {
...           branches: [
...             {case: { $eq : ["$STAT_CAUSE_DESCR", "Lightning"]}, then: "Natural Causes"},
...             {case: { $eq : ["$STAT_CAUSE_DESCR", "Debris Burning"]}, then: "Natural Causes"},
...             {case: { $and: [ { $ne : ["$STAT_CAUSE_DESCR", "Miscellaneous"]},
...                           { $ne : ["$STAT_CAUSE_DESCR", "Missing/Undefined"]},
...                           { $ne : ["$STAT_CAUSE_DESCR", "Debris Burning"]},
...                           { $ne : ["$STAT_CAUSE_DESCR", "Lightning"]} ]}, then: "Man-Made Causes"}
...           ],
...           default: 0
...         }
...       }
...     },
...     {
...       $group: { _id: "$Cause", count: { $sum: 1} }
...     }
...   ])
{ "_id" : "Man-Made Causes", "count" : 682441 }
{ "_id" : 0, "count" : 490528 }
{ "_id" : "Natural Causes", "count" : 707496 }
> █

> db.Fires.aggregate([
...   {
...     $project: {
...       "STAT_CAUSE_DESCR": 1,
...       "Cause": {
...         $switch: {
...           branches: [
...             {case: { $eq : ["$STAT_CAUSE_DESCR", "Lightning"]}, then: "Natural Causes"},
...             {case: { $and: [ { $ne : ["$STAT_CAUSE_DESCR", "Miscellaneous"]},
...                           { $ne : ["$STAT_CAUSE_DESCR", "Missing/Undefined"]},
...                           { $ne : ["$STAT_CAUSE_DESCR", "Lightning"]} ]}, then: "Man-Made Causes"}
...           ],
...           default: 0
...         }
...       }
...     },
...     {
...       $group: { _id: "$Cause", count: { $sum: 1} }
...     }
...   ])
{ "_id" : 0, "count" : 490528 }
{ "_id" : "Natural Causes", "count" : 278468 }
{ "_id" : "Man-Made Causes", "count" : 1111469 }
> █
```

Query 5

Question

Q6)

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

Translation

Screen Shot of MongoDB Query/Code and Results

```
> db.Fires.aggregate([
...   {
...     $project:{
...       "id":1,
...       "NWCG_REPORTING_UNIT_NAME":1,
...       "days":{$gt:[{$subtract:[ "$CONT_DOY", "$DISCOVERY_DOY"]},2]}
...     }
...   },
...   {
...     $match:{ "days":true}
...   },
...   {
...     $group:{
...       "_id": "$NWCG_REPORTING_UNIT_NAME",
...       units:{ "$sum":1}
...     }
...   },
...   {
...     $project:{
...       "_id":1,
...       "units":{$gt:[ "units",2]}
...     }
...   },
...   {
...     "$match":{ "units":true}
...   },
...   {
...     $sort:{ "NWCG_REPORTING_UNIT_NAME":1}
...   }
... ])
{
  "_id" : "Kenai National Wildlife Refuge", "units" : true
}
{
  "_id" : "Papago Agency", "units" : true
}
{
  "_id" : "Amador-El Dorado Unit", "units" : true
}
{
  "_id" : "Shoshone National Forest", "units" : true
}
{
  "_id" : "Felsenthal National Wildlife Refuge", "units" : true
}
{
  "_id" : "Yukon Delta National Wildlife Refuge", "units" : true
}
{
  "_id" : "Manti-Lasal National Forest", "units" : true
}
{
  "_id" : "Michigan Department of Natural Resources", "units" : true
}
{
  "_id" : "Lassen-Modoc Unit", "units" : true
}
{
  "_id" : "Riverside Unit", "units" : true
}
{
  "_id" : "Bitter Lake National Wildlife Refuge", "units" : true
}
{
  "_id" : "Saguaro National Park", "units" : true
}
{
  "_id" : "Louisiana Office of Forestry", "units" : true
}
{
  "_id" : "Craters of the Moon National Monument", "units" : true
}
{
  "_id" : "Wallowa Unit", "units" : true
}
{
  "_id" : "Rogue River-Siskiyou National Forest", "units" : true
}
{
  "_id" : "Yukon Flats National Wildlife Refuge", "units" : true
}
{
  "_id" : "Wind River Agency", "units" : true
}
{
  "_id" : "Big Thicket National Park", "units" : true
}
{
  "_id" : "Sonoma-Lake Napa Unit", "units" : true
}
Type "it" for more
> |||
```

Query 6

Question

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

Translation

Screen Shot of MongoDB Query/Code and Results