

## Public Libraries Assignment

**Query 1:** “What is the most common library name in the 2018 data-set?”

Your answer should be a tuple with two columns, in this order: libname, count.

```
select LIBNAME as libname, count(libname)
from lib2018
group by libname
order by count(libname) desc
limit 1;
```

```
shimmy@TABLET-7SLA7USR: x + v
ND | 2216377 | 2162189 | 2201730 | 2.5061638922406876 | -1.7959059466873777
ID | 8179077 | 8029503 | 8597955 | 1.8628052072463265 | -6.6114791249779744
DC | 3632539 | 3593201 | 3930763 | 1.0947898545057735 | -8.5876965871511460
NH | 7045010 | 7028800 | 7236567 | 0.2306225813794673 | -2.8710713242895423
ME | 6746380 | 6731768 | 6811441 | 0.2170603621515180 | -1.1696937549631568
UT | 15326963 | 15295494 | 16096911 | 0.2057403311066645 | -4.9787005718053607
DE | 4122181 | 4117904 | 4125899 | 0.1038635189164196 | -0.1937759504049905
(10 rows)

libraries=# select * from lib2016 where LIBNAME='M';
libraries=# select stabr, fscskey, libname, address, city, zip, cnty, phone from lib2016 where libname='M';
stabr | fscskey | libname | address | city | zip | cnty | phone
-----+-----+-----+-----+-----+-----+-----+-----
(0 rows)

libraries=# select stabr, fscskey, libname, address, city, zip, cnty, phone from lib2017 where libname='M';
stabr | fscskey | libname | address | city | zip | cnty | phone
-----+-----+-----+-----+-----+-----+-----+-----
(0 rows)

libraries=# select stabr, fscskey, libname, address, city, zip, cnty, phone from lib2018 where libname='M';
stabr | fscskey | libname | address | city | zip | cnty | phone
-----+-----+-----+-----+-----+-----+-----+-----
(0 rows)

libraries=# select stabr, fscskey, libname, address, city, zip, cnty, phone from lib2018 where stabr='M';
stabr | fscskey | libname | address | city | zip | cnty | phone
-----+-----+-----+-----+-----+-----+-----+-----
(0 rows)

libraries=# select LIBNAME as libname, count(libname)
from lib2018
group by libname
order by count(libname) desc
limit 1;
      libname      | count
-----+-----
OXFORD PUBLIC LIBRARY |    10
(1 row)

libraries=#
```

## Query 2: “Which state has the most libraries?”

Your result set should contain seven tuples, in descending order of “number of libraries”. Each tuple should contain (and be labeled), in this order: state, number of libraries

State is represented by the 2-letter code. Use the 2018 dataset.

```
select STABR as state, count(*) as "number of libraries"
from lib2018
group by state
order by "number of libraries" desc
limit 7;
```

```
shimmy@TABLET-75LA7USR: . x + v
group by libname
order by count(libname) desc
limit 1;
-----+-----
libname | count
-----+-----
OXFORD PUBLIC LIBRARY | 10
(1 row)

libraries=# select STABR as state, count(*) as "number of libraries"
from lib2018 -- might need to join
group by state
order by "number of libraries" desc
limit 7;
state | number of libraries
-----+-----
NY | 756
IL | 623
TX | 560
IA | 544
PA | 451
MI | 398
WI | 381
(7 rows)

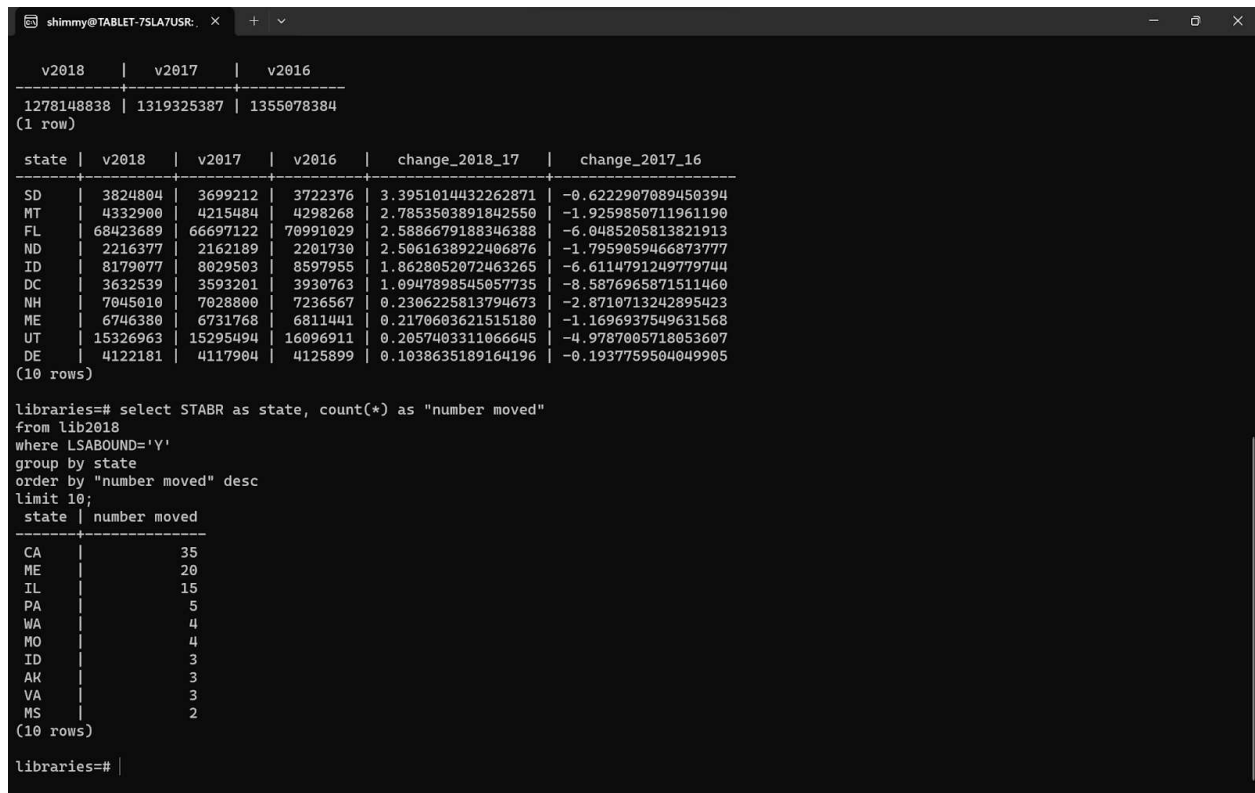
libraries=# select STABR as state, count(*) as "number of libraries"
from lib2018
group by state
order by "number of libraries" desc
limit 7;
state | number of libraries
-----+-----
NY | 756
IL | 623
TX | 560
IA | 544
PA | 451
MI | 398
WI | 381
(7 rows)

libraries=# |
```

**Query 3:** “For each state, in 2018, how many libraries changed their address in any way?”

Your result set should contain, in descending order of “number libraries moved in that state in 2018”. Each tuple should contain (and be labeled as) in this order: state, number moved.

```
select STABR as state, count(*) as "number moved"
from lib2018
where LSABOUND='Y'
group by state
order by "number moved" desc
limit 10;
```



The screenshot shows a terminal window with a dark background. It displays the results of a SQL query. The first table shows data for 2018, 2017, and 2016, along with change values. The second table shows the top 10 states by the number of libraries that moved in 2018.

	v2018	v2017	v2016		
1278148838	1319325387	1355078384			
(1 row)					

state	v2018	v2017	v2016	change_2018_17	change_2017_16
SD	3824804	3699212	3722376	3.3951014432262871	-0.6222907089450394
MT	4332900	4215484	4298268	2.7853503891842550	-1.9259850711961190
FL	68423689	66697122	70991029	2.5886670188346388	-6.0485205813821913
ND	2216377	2162189	2201730	2.5061638922406876	-1.7959059466873777
ID	8179077	8029503	8597955	1.8628052072463265	-6.6114791249779744
DC	3632539	3593201	3930763	1.0947898545057735	-8.5876965871511460
NH	7045010	7028800	7236567	0.2306225813794673	-2.8710713242895423
ME	6746380	6731768	6811441	0.2170603621515180	-1.1696937549631568
UT	15326963	15295494	16096911	0.2057403311066645	-4.9787005718053607
DE	4122181	4117904	4125899	0.1038635189164196	-0.1937759504049905
(10 rows)					

```
libraries=# select STABR as state, count(*) as "number moved"
from lib2018
where LSABOUND='Y'
group by state
order by "number moved" desc
limit 10;
```

state	number moved
CA	35
ME	20
IL	15
PA	5
WA	4
MO	4
ID	3
AK	3
VA	3
MS	2
(10 rows)	

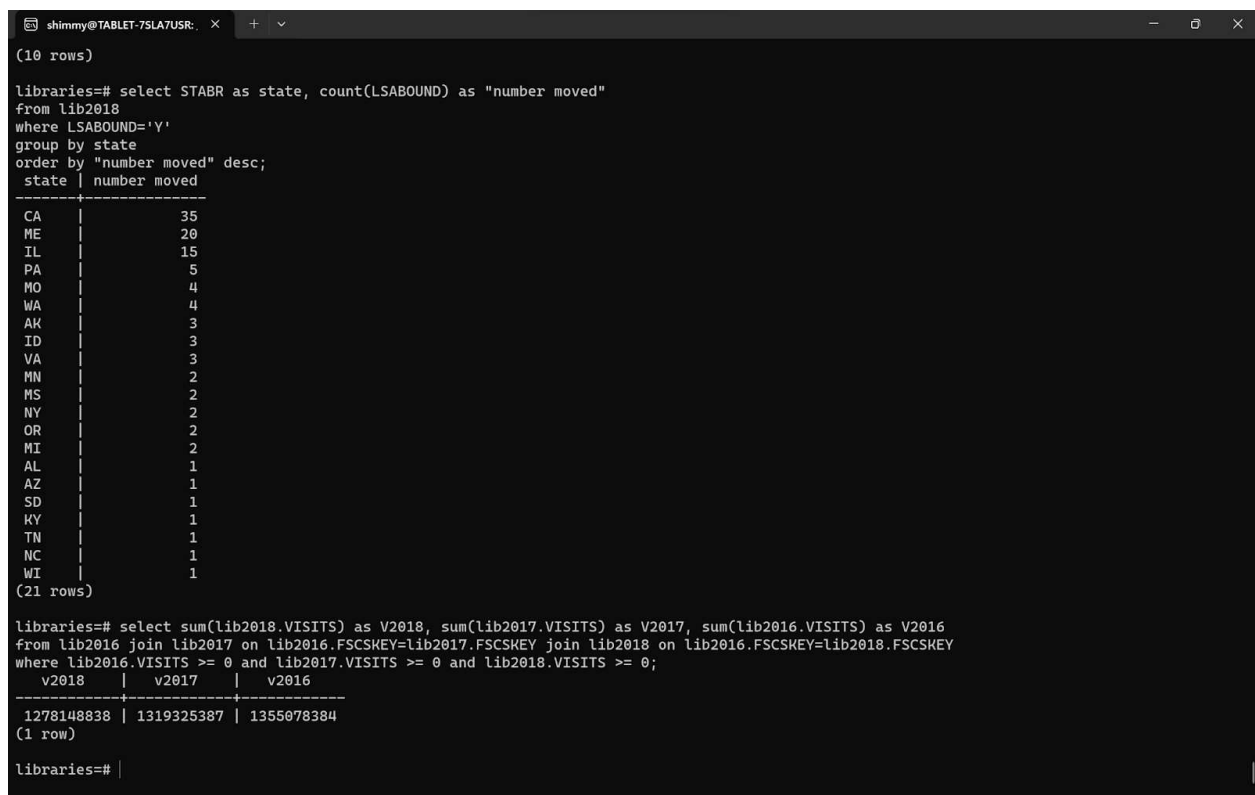
```
libraries=#
```

**Query 4:** “Return the number of visits to libraries in 2018, 2017, and 2016.”

The query is intended to get a sense of whether library usage has increased, decreased, or stayed roughly the same over this time period. Therefore: you need to do an “apples-to-apples” comparison, such that only libraries that were open in each of these years are used in the computation. Be sure to only include “valid survey responses/data” (see above) in the computation. Note: you’re aggregating data across the United States as a whole.

The result set should be a single tuple, containing (and labelled), in this order: V2018, V2017, V2016.

```
select sum(lib2018.VISITS) as V2018, sum(lib2017.VISITS) as
V2017, sum(lib2016.VISITS) as V2016
from lib2016 join lib2017 on lib2016.FSCSKEY=lib2017.FSCSKEY
join lib2018 on lib2016.FSCSKEY=lib2018.FSCSKEY
where lib2016.VISITS >= 0 and lib2017.VISITS >= 0 and
lib2018.VISITS >= 0;
```



The screenshot shows a terminal window with a dark background. The title bar indicates the user is 'shimmy' on a machine named 'TABLET-75LA7USR'. The terminal displays two SQL queries and their results. The first query counts the number of libraries moved by state, showing results for 21 states. The second query calculates the sum of visits for the years 2018, 2017, and 2016, returning a single row with three values.

```
(10 rows)

libraries=# select STABR as state, count(LSABOUND) as "number moved"
from lib2018
where LSABOUND='Y'
group by state
order by "number moved" desc;
state | number moved
-----+-----
CA    |          35
ME    |          20
IL    |          15
PA    |           5
MO    |           4
WA    |           4
AK    |           3
ID    |           3
VA    |           3
MN    |           2
MS    |           2
NY    |           2
OR    |           2
MI    |           2
AL    |           1
AZ    |           1
SD    |           1
KY    |           1
TN    |           1
NC    |           1
WI    |           1
(21 rows)

libraries=# select sum(lib2018.VISITS) as V2018, sum(lib2017.VISITS) as V2017, sum(lib2016.VISITS) as V2016
from lib2016 join lib2017 on lib2016.FSCSKEY=lib2017.FSCSKEY join lib2018 on lib2016.FSCSKEY=lib2018.FSCSKEY
where lib2016.VISITS >= 0 and lib2017.VISITS >= 0 and lib2018.VISITS >= 0;
v2018 | v2017 | v2016
-----+-----+-----
1278148838 | 1319325387 | 1355078384
(1 row)

libraries=#
```

5. “Do a by-state analysis of the above query (details below).”

Instead of reporting data aggregated across the United States (the previous query), you’re aggregating data by state. You’ll report the “raw number of visits” for each of the three years. In addition: you’ll report (by state), the trends of visits: specifically, the percentage change (whether positive or negative) in 2018 relative to 2017, and the percentage change in 2017 relative to 2016.

Your answer should be a tuple containing (and labeled) in this order: (a) state (b) V2018 (c) V2017 (d) V2016 (e) CHANGE\_2018\_17 (f) CHANGE\_2017\_16

Report only the first ten tuples when the result set is ordered by descending “percentage change from 2018 to 2017” values.

Note: this query uses the previous queries’ semantics with respect to e.g., library usage.

Suggestion: be careful about integer division!

```
select state, V2018, V2017, V2016, round((100.0 * V2018 / V2017)
- 100, 1) as CHANGE_2018_17, round((100.0 * V2017 / V2016) -
100, 1) as CHANGE_2017_16
from (select lib2016.STABR as state, sum(lib2018.VISITS) as
V2018, sum(lib2017.VISITS) as V2017, sum(lib2016.VISITS) as
V2016
      from lib2016 join lib2017 on
lib2016.FSCSKEY=lib2017.FSCSKEY join lib2018 on
lib2016.FSCSKEY=lib2018.FSCSKEY
      where lib2016.VISITS >= 0 and lib2017.VISITS >= 0 and
lib2018.VISITS >= 0
      group by state) as visits_table
order by CHANGE_2018_17 desc
limit 10;
```

```
shimmy@TABLET-75LA7USR: x + v
-----
1278148838 | 1319325387 | 1355078384
(1 row)

state | v2018 | v2017 | v2016 | change_2018_17 | change_2017_16
-----+-----+-----+-----+-----+-----
SD | 3824804 | 3699212 | 3722376 | 3.4 | -0.6
MT | 4332900 | 4215484 | 4298268 | 2.8 | -1.9
FL | 68423689 | 66697122 | 70991029 | 2.6 | -6.0
ND | 2216377 | 2162189 | 2201730 | 2.5 | -1.8
ID | 8179077 | 8029503 | 8597955 | 1.9 | -6.6
DC | 3632539 | 3593201 | 3930763 | 1.1 | -8.6
ME | 6746380 | 6731768 | 6811441 | 0.2 | -1.2
NH | 7045010 | 7028800 | 7236567 | 0.2 | -2.9
UT | 15326963 | 15295494 | 16096911 | 0.2 | -5.0
DE | 4122181 | 4117904 | 4125899 | 0.1 | -0.2
(10 rows)

libraries=# select state, V2018, V2017, V2016, round((100.0 * V2018 / V2017) - 100, 1) as CHANGE_2018_17,
round((100.0 * V2017 / V2016) - 100, 1) as CHANGE_2017_16
from (select lib2016.STABR as state, sum(lib2018.VISITS) as V2018, sum(lib2017.VISITS) as V2017, sum(lib2016.VISITS) as V2016
from lib2016 join lib2017 on lib2016.FSCSKEY=lib2017.FSCSKEY join lib2018 on lib2016.FSCSKEY=lib2018.FSCSKEY
where lib2016.VISITS >= 0 and lib2017.VISITS >= 0 and lib2018.VISITS >= 0
group by state) as visits_table
order by CHANGE_2018_17 desc
limit 10;
state | v2018 | v2017 | v2016 | change_2018_17 | change_2017_16
-----+-----+-----+-----+-----+-----
SD | 3824804 | 3699212 | 3722376 | 3.4 | -0.6
MT | 4332900 | 4215484 | 4298268 | 2.8 | -1.9
FL | 68423689 | 66697122 | 70991029 | 2.6 | -6.0
ND | 2216377 | 2162189 | 2201730 | 2.5 | -1.8
ID | 8179077 | 8029503 | 8597955 | 1.9 | -6.6
DC | 3632539 | 3593201 | 3930763 | 1.1 | -8.6
ME | 6746380 | 6731768 | 6811441 | 0.2 | -1.2
NH | 7045010 | 7028800 | 7236567 | 0.2 | -2.9
UT | 15326963 | 15295494 | 16096911 | 0.2 | -5.0
DE | 4122181 | 4117904 | 4125899 | 0.1 | -0.2
(10 rows)

libraries=# |
```