

The report for 4140 Assignment 4

1. Data cleaning

We delete all rows whose BILL_ID is empty. We also change the value of MONTH from string to int so that data of year, month, and day can be reached faster.

2. Query explanation

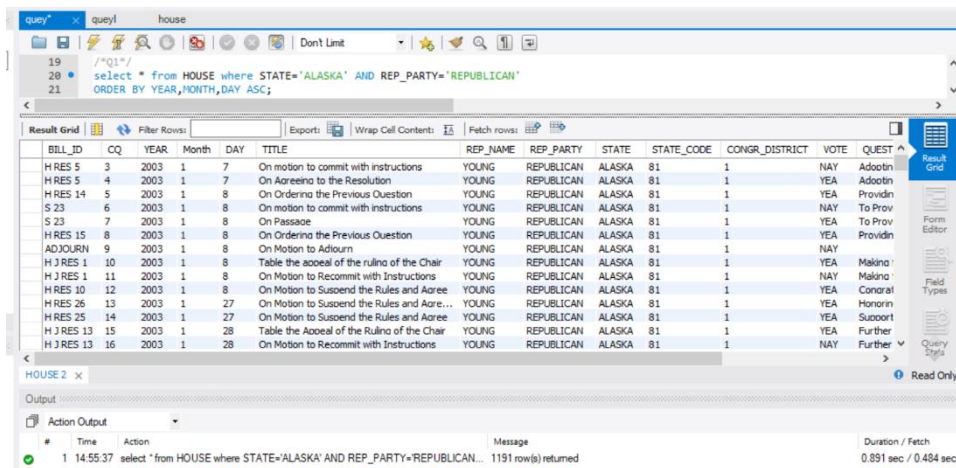
Query 1:

```
Select * from HOUSE where STATE='ALASKA' AND REP_PARTY='REPUBLICAN'  
ORDER BY YEAR, MONTH, DAY ASC;
```

Explanation:

This query selects rows whose representatives are from Alaska and their party belong to republican. The result is ordered by year, month, and day. In the local system, it takes 0.891 seconds to execute while in the cloud-based system, it takes only 3054 milliseconds. The cloud-based system has more execution time.

Local system:



BILL_ID	CQ	YEAR	Month	DAY	TITLE	REP_NAME	REP_PARTY	STATE	STATE_CODE	CONGR_DISTRICT	VOTE	QUEST
H RES 5	3	2003	1	7	On motion to commit with instructions	YOUNG	REPUBLICAN	ALASKA	81	1	NAY	Adoptin
H RES 5	4	2003	1	7	On Agreeing to the Resolution	YOUNG	REPUBLICAN	ALASKA	81	1	YEA	Adoptin
H RES 14	5	2003	1	8	On Ordering the Previous Question	YOUNG	REPUBLICAN	ALASKA	81	1	YEA	Providin
S 23	6	2003	1	8	On motion to commit with instructions	YOUNG	REPUBLICAN	ALASKA	81	1	NAY	To Prov
S 23	7	2003	1	8	On Passage	YOUNG	REPUBLICAN	ALASKA	81	1	YEA	To Prov
H RES 15	8	2003	1	8	On Ordering the Previous Question	YOUNG	REPUBLICAN	ALASKA	81	1	YEA	Providin
ADJOURN	9	2003	1	8	On Motion to Adjourn	YOUNG	REPUBLICAN	ALASKA	81	1	NAY	
H J RES 1	10	2003	1	8	Table the appeal of the ruling of the Chair	YOUNG	REPUBLICAN	ALASKA	81	1	YEA	Makino
H J RES 1	11	2003	1	8	On Motion to Recommit with Instructions	YOUNG	REPUBLICAN	ALASKA	81	1	NAY	Makino
H RES 10	12	2003	1	8	On Motion to Suspend the Rules and Agree	YOUNG	REPUBLICAN	ALASKA	81	1	YEA	Conoral
H RES 26	13	2003	1	27	On Motion to Suspend the Rules and Agree	YOUNG	REPUBLICAN	ALASKA	81	1	YEA	Honorin
H RES 25	14	2003	1	27	On Motion to Suspend the Rules and Agree	YOUNG	REPUBLICAN	ALASKA	81	1	YEA	Support
H J RES 13	15	2003	1	28	Table the Appeal of the Ruling of the Chair	YOUNG	REPUBLICAN	ALASKA	81	1	YEA	Further
H J RES 13	16	2003	1	28	On Motion to Recommit with Instructions	YOUNG	REPUBLICAN	ALASKA	81	1	NAY	Further

Cloud-based system:

Congratulations. Your connection to dashDB is successful.

Rows from the DASH6769.HOUSE table query1

BILL_ID	CQ	YEAR	MONTH	DAY	TITLE	REP_NAME	REP_PARTY	STATE	STATE CODE	CONGR DISTRICT	VOTE	QUESTION
H RES 5	3	2003	1	7	On motion to commit with instructions	YOUNG	REPUBLICAN	ALASKA	81	1	NAY	Adopting Rules for the One Hundred Eighth Congress
H RES 5	4	2003	1	7	On Agreeing to the Resolution	YOUNG	REPUBLICAN	ALASKA	81	1	YEA	Adopting Rules for the One Hundred Eighth Congress
												To Provide for a 5-Month Extension of the Temporary Extended Unemployment Compensation Act of 2002 and for a Transition
S 23	6	2003	1	8	On motion to commit with	YOUNG	REPUBLICAN	ALASKA	81	1	NAY	

GEThttps://assignment4db4140.mybluemix.netParamsSendSave

AuthorizationHeadersBodyPre-request ScriptTestsCookiesCode

TYPE

Inherit auth from parent

The authorization header will be automatically generated when you send the request. [Learn more about authorization](#)

This request is not inheriting any authorization helper at the moment. Save it in a collection to use the parent's authorization helper.

BodyCookies (1)Headers (9)Test Results

Status: 200 OKTime: 3054 msSize: 330.51 KB

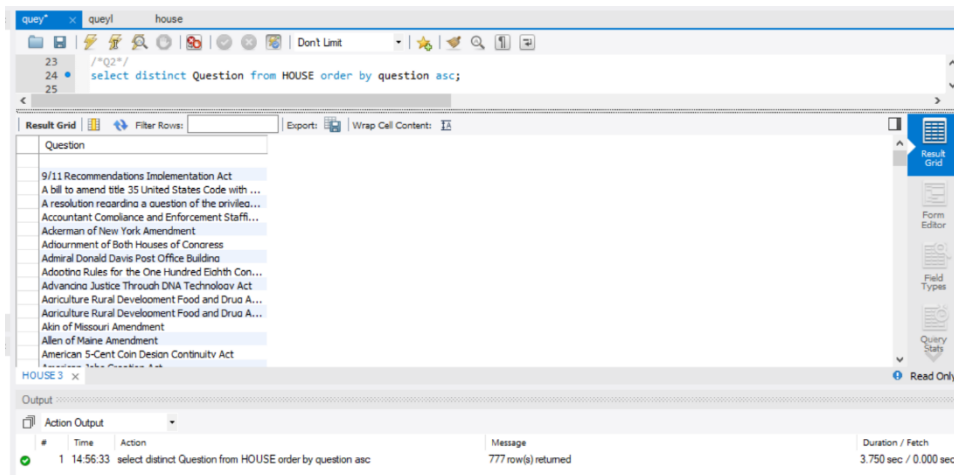
Query 2:

```
select distinct Question from HOUSE order by question asc;
```

Explanation:

This query lists distinct questions that are asked and are ordered in the ascending order. In the local system, the execution time is 3.750 seconds. However, in the cloud-based system, it takes less time: 477 milliseconds.

Local system:



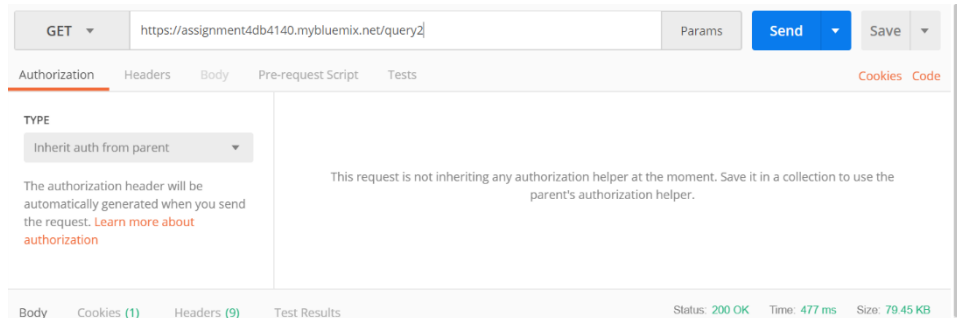
Cloud-based system:

Congratulations. Your connection to dashDB is successful.

Rows from the DASH6769.HOUSE table query2

QUESTION

9/11 Recommendations Implementation Act
A bill to amend title 35 United States Code with respect to patent fees and for other purposes
A resolution regarding a question of the privileges of the House regarding Roll Call 669.
Accountant Compliance and Enforcement Staffing Act



Query 3:

```
select REP_NAME, count(REP_NAME) as TimesVotingForNay from HOUSE where  
STATE='CALIFOR' and VOTE='NAY' GROUP BY REP_NAME;
```

Explanation:

This query counts how many times that representatives who are from California vote for "Nay". The execution time in local system is 0.797 seconds. The cloud-based system takes 570 milliseconds to execute the query. The cloud-based system runs faster than the local one.

Local system:

query1 query1

```
/*Q3*/
select REP_NAME, count(REP_NAME) as TimesVotingForNay from HOUSE where STATE='CALIFOR' and VOTE='NAY' GROUP BY REP_NAME;
```

REP_NAME	TimesVotingForNay
BACA	370
BECERRA	374
BERMAN	345
BONO	356
CALVERT	390
CAPPS	367
CARDOZA	371
COX	332
CUNNINGHAM	365
DAVIS	363
DOOLEY	358
DOOLITTLE	371
DREIER	393
ESHOO	355
FARR	387
FILNER	382

Output

#	Time	Action	Message	Duration / Fetch
1	15:00:41	select REP_NAME, count(REP_NAME) as TimesVotingForNay from HOUSE wher...	51 row(s) returned	0.797 sec / 0.000 sec

Cloud-based system:

Congratulations. Your connection to dashDB is successful.

Rows from the DASH6769.HOUSE table query3

REP_NAME	TIMESVOTINGFORNAY
COX	332
SCHIFF	377
HONDA	375
MILLENDEN	327
OSE	392

GET https://assignment4db4140.mybluemix.net/query3 Params Send Save

Authorization Headers Body Pre-request Script Tests Cookies Code

TYPE

Inherit auth from parent

The authorization header will be automatically generated when you send the request. [Learn more about authorization](#)

This request is not inheriting any authorization helper at the moment. Save it in a collection to use the parent's authorization helper.

Body Cookies (1) Headers (9) Test Results Status: 200 OK Time: 570 ms Size: 2.48 KB

Query 4:

select REP_NAME, COUNT(REP_NAME) as numberOfVotes FROM HOUSE group by REP_NAME ORDER BY COUNT(REP_NAME) asc limit 5;

Local system:

query1 query1

```
/*Q4*/
select REP_NAME, COUNT(REP_NAME) as numberOfVotes FROM HOUSE group by REP_NAME ORDER BY COUNT(REP_NAME) asc limit 5 ;
```

REP_NAME	numberOfVotes
MCCOLLUM	1191
BERKLEY	1191
MEEKS	1191
HAYES	1191
DELAY	1191

Output

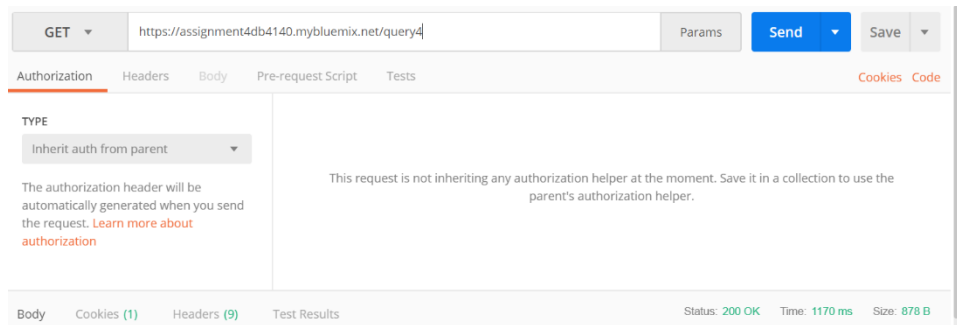
#	Time	Action	Message	Duration / Fetch
1	15:01:43	select REP_NAME, COUNT(REP_NAME) as numberOfVotes FROM HOUSE grou...	5 row(s) returned	1.609 sec / 0.000 sec

Cloud-based system:

Congratulations. Your connection to dashDB is successful.

Rows from the DASH6769.HOUSE table query4

REP_NAME	NUMBEROFVOTES
BURNS	1217
PRICE	1217
LEACH	1217
GILCHRES	1217
FRANK	1217



Explanation:

The query selects five representatives' name who vote the least times from all representatives. The query also orders all the results in the ascending order. The execution time in local system is 1.609 seconds and the cloud-based system takes 1170 milliseconds to execute the query. The cloud-based system runs faster than the local one.

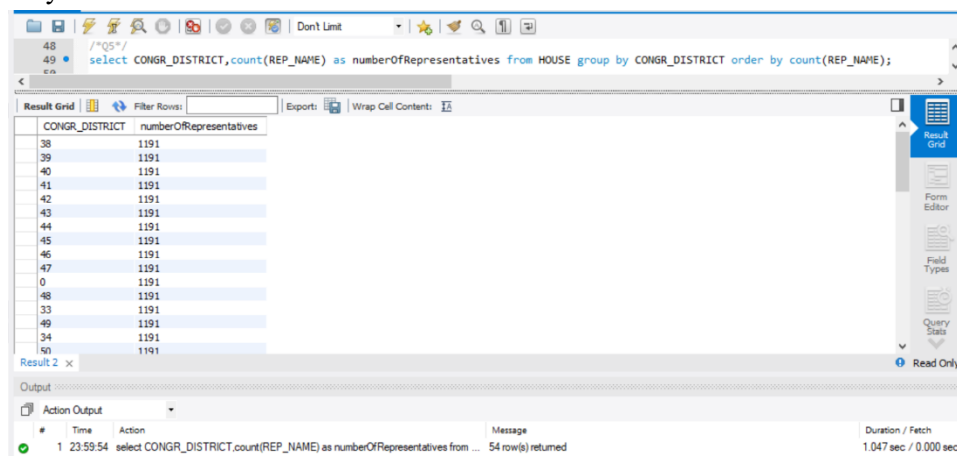
Query 5:

```
select CONGR_DISTRICT, count(REP_NAME) as numberOfRepresentatives from  
HOUSE group by CONGR_DISTRICT order by count(REP_NAME);
```

Explanation:

This query displays the statistics results of each district. It counts the quantity of representatives in each district. The results are ordered in ascending order of the count. In the local system, the execution time is 1.047 seconds while it is 1014 milliseconds in the cloud-based system.

Local system:

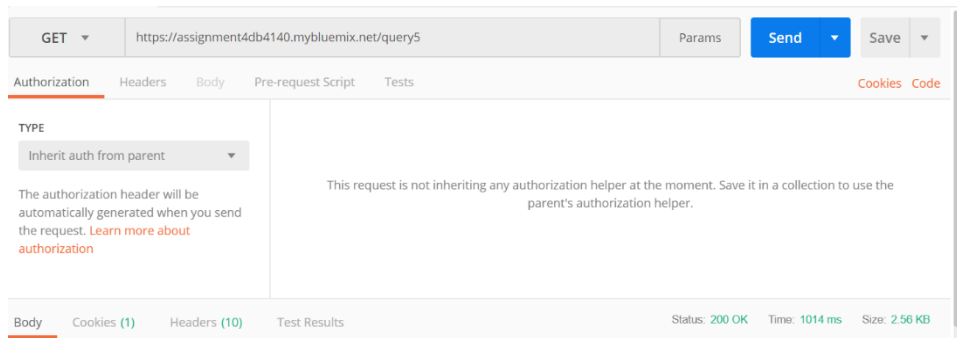


Cloud-based system:

Congratulations. Your connection to dashDB is successful.

Rows from the DASH6769.HOUSE table query5

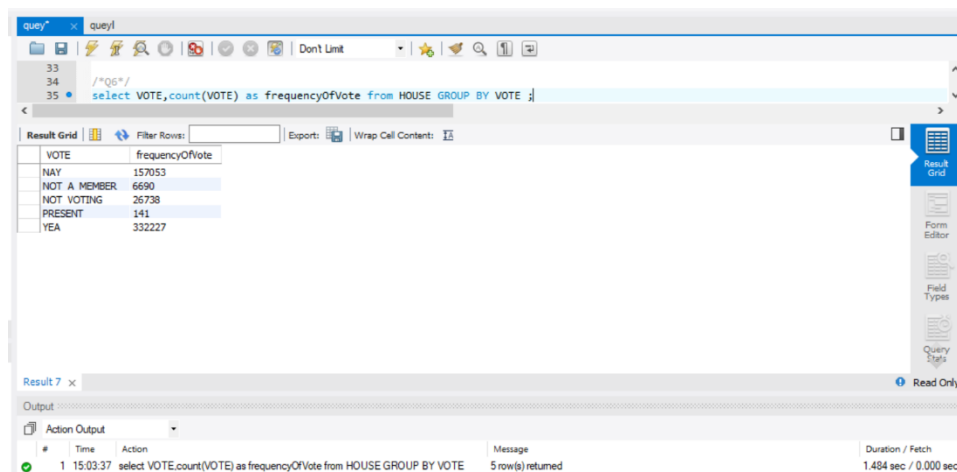
CONGR_DISTRICT	NUMBEROFREPRESENTATIVES
53	1217
36	1217
45	1217
50	1217
37	1217
47	1217



Query 6:

select VOTE, count(VOTE) as frequencyOfVote from HOUSE GROUP BY VOTE ;

Local system:

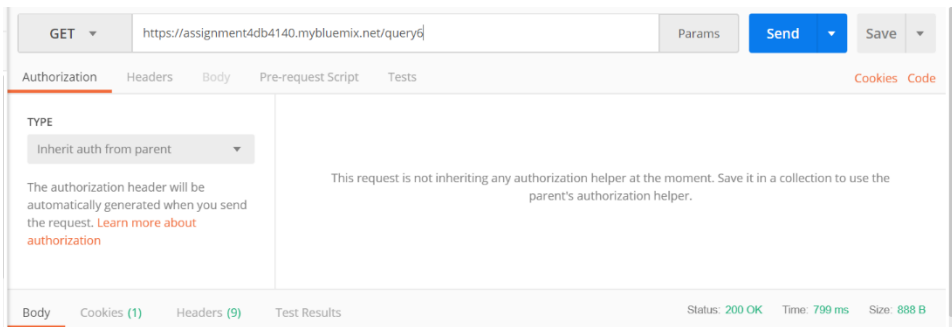


Cloud-based system:

Congratulations. Your connection to dashDB is successful.

Rows from the DASH6769.HOUSE table query6

VOTE	FREQUENCYOFVOTE
YEA	341119
NOT_A_MEMBER	6853
PRESENT	156
NOT_VOTING	27469
NAY	158666



Explanation:

The query 6 counts numbers of votes in the "HOUSE" table and group the result according to the type of votes. The execution time in local system is 1.484 seconds. The cloud-based system takes 799 milliseconds to execute the query. The cloud-based system runs faster than the local one.

Query 7:

SELECT state , count(state) as totalNumberOfVotes from HOUSE group by state;

Local system:

state	totalNumberOfVotes
ALABAMA	8337
ALASKA	1191
ARIZONA	9528
ARKANSAS	4764
CALIFOR	63123
COLORAD	8337
CONNECT	5955
DELAWARE	1191
FLORIDA	29775
GEORGIA	15483
HAWAII	2382
IDAHO	2382
ILLINOI	22629
INDIANA	10719
IOWA	5955
KANSAS	4764

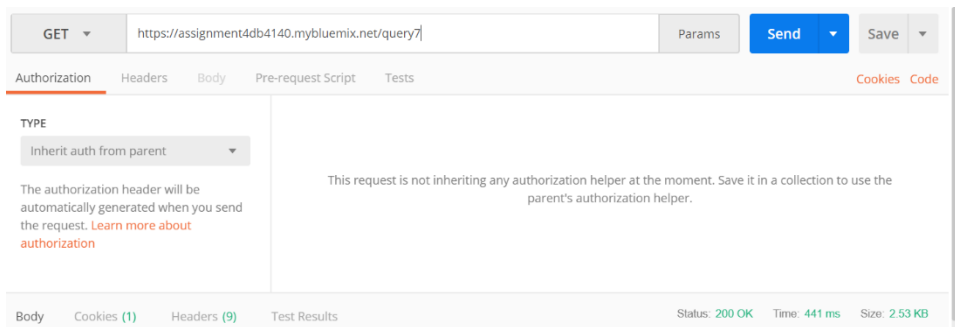
Output: 1 15:04:15 SELECT state ,count(state) as totalNumberOfVotes from HOUSE group by state 50 row(s) returned 1.828 sec / 0.000 sec

Cloud-based system:

Congratulations. Your connection to dashDB is successful.

Rows from the DASH6769.HOUSE table query7

STATE	TOTALNUMBEROFVOTES
CALIFOR	64501
NEW MEX	3651
OKLAHOM	6085
USA	1217
NEW HAM	2434
PENNSYL	23123
IDAHO	2434
ARIZONA	9736
NEW YOR	35293
VERMONT	1217



Explanation:

The query 7 counts numbers of votes in terms of the 'state' in the "HOUSE" table. The results are grouped by the 'state'. The execution time in local system is 1.828 seconds. The cloud-based system takes 441 milliseconds to execute the query. The cloud-based system runs faster than the local one.

Time	Local	Cloud		Faster
Q1	0.891 seconds	3054 milliseconds	3.054 seconds	Local
Q2	3.750 seconds	477 milliseconds	0.477seconds	Cloud
Q3	0.797 seconds	570milliseconds	0.57 seconds	Cloud
Q4	1.609 seconds.	1170 milliseconds	1.170 seconds	Cloud
Q5	1.047 seconds	1014milliseconds	1.014 seconds	Cloud
Q6	1.484 seconds.	799 milliseconds	0.799 seconds	Cloud
Q7	1.828 seconds	441 milliseconds	0.441 seconds	Cloud

By comparing these seven queries' executions times in local and cloud-based system, we can see that overall, the cloud-based system takes less time to get results. It is apparent that the local system is faster when the executed query is not complex. It may be because the speed of the cloud-based system is influenced by the internet speed since it needs to time to transmit data. However, when queries become more and more complicated, the cloud-based system shows its benefits. The reason may be because that the cloud-based system may has better hardware than the local system.

3. NodeJS application directory

F:\CSCI4140\A4\dash