Description of Queries

Weston Ortiz, Hans Weeks

April 16, 2017

Contents

1 Data Retrieval Queries

1. How many amendments were introduced before 2008 in each branch?

```
select type, count(type) from Amendment where
introduced_at < '2008' group by type;
```

2. Which legislators were born before the outset of WWI?

```
select distinct bioguide_id, `First Name` , `Last Name`, party
from
Legislator natural join Term where birthday < '1914-07-28';
```

3. Find all Legislators that shared a birthday

```
select L1.birthday, L1.bioguide_id, L1.`First Name`, L1.`Last Name`,
  L2.bioguide_id, L2.`First Name`, L2.`Last Name`
from Legislator L1, Legislator L2
where L1.birthday = L2.birthday and L1.bioguide_id != L2.bioguide_id;
```

4. Find all votes that have passed

```
select * from Vote where result = 'PASSED';
```

5. How many legislators have there been for each party (using only a legislators most recent term to decide their party)

```
select party, count(party) as count
from Term join (select bioguide_id, max(start) as ms
from Term GROUP BY bioguide_id) as latest
ON Term.start = latest.ms and Term.bioguide_id = latest.bioguide_id
GROUP BY party ORDER BY count DESC;
```

6. For a Bill find all congress members and how they voted on the latest vote for that bill

```
select `bioguide_id`, `First Name`, `Last Name`, how_voted
from (select * from Legislator natural join Legislator_Vote) as LV
     join Vote ON Vote.id = LV.Vote_id
where Vote_id = (select id from Vote where Bill_id =  'hr899-113'
      and number = (select max(number) from Vote where Bill_id = 'hr899-113'));
```

7. Find all bills sponsored by a legislator from NM and the name of the legislator that sponsored it.

```
select `Last Name`, Bill_id
from ((select bioguide_id from Term where state ='NM') as t1
     natural join Legislator) natural join Sponsor;
```

8. Union the shared fields from amendment and bill, represents all pieces of legislation in the database:

```
select *
from (select id, type, status, introduced_at as introduction_date,
      congress, number
      from Amendment) as t1
     union
     (select id, type, status, introduction_date,
     congress, number from Bill) limit 10;
```

9. For a given vote count the number of yes votes for each party

```
select party, count(party) as count
from (select distinct bioguide_id, party
      from Legislator_Vote natural join Term
      where Vote_id = 'h136-115.2017' and how_voted='Yea') as yes
GROUP BY party;
```

10. List the members of congress for congress 115 and order by state

```
select distinct `First Name`, `Last Name`, state
from Legislator natural join Term
where start >= (select begin from Congress where id = 115)
      and end <= (select end from Congress where id = 115) ORDER BY state
```

11. Find all subjects for Congress 114 (would be useful for navigation in the web interface)

```
select distinct subject from Subject join Bill ON Subject.Bill_id = Bill.id
where congress = 114;
```

12. Count the number of terms a member of congress has served (only if we have a record of them voting in our data)

```
select bioguide_id, `First Name`, `Last Name`, Terms
from (select bioguide_id, count(bioguide_id) as Terms
      from Term GROUP BY bioguide_id) as tm natural join Legislator
where bioguide_id in (select bioguide_id from Legislator_Vote)
ORDER BY Terms DESC;
```

13. Find all Legislators that share a last name

```
select L1.bioguide_id, L1.`First Name`, L1.`Last Name`,
       L2.bioguide_id, L2.`First Name`, L2.`Last Name`
from Legislator L1, Legislator L2
where L1.`Last Name` = L2.`Last Name`
      and L1.bioguide_id != L2.bioguide_id;
```

14. Find all the ways that have been voted

```
select distinct how_voted from Legislator_Vote;
```

2 Modification Queries

1. If a Legislator does not have a full name update it with their wikipedia id (which is a full name) only if their wikipedia id is not null as well

```
update Legislator
set official_full_name=wikipedia_id
where official_full_name IS NULL and wikipedia_id IS NOT NULL;
```

2. Change a legislators end date for a specific Term (like if they have been removed from office, in our case Tom Marino is likely to be appointed to another position soon)

```
update Term
set end = '2017-05-01'
where bioguide_id = (select bioguide_id from Legislator
     where `First Name` = 'Tom' and `Last Name` = 'Marino')
     and end > '2017-05-01';
```

3. Delete all Legislators , Sponsor, and Terms for those legislator if they never participated in a vote in our data

```
delete from Term
where bioguide_id
      not in (select bioguide_id from Legislator_Vote);

delete from Sponsor
where Legislator_id
      not in (select bioguide_id from Legislator_Vote);

delete from Legislator
where bioguide_id
      not in (select bioguide_id from Legislator_Vote);
```

4. Delete all votes that we do not have a record of Legislators voting on it

```
delete from Vote
where id
      not in (select Vote_id as id from Legislator_Vote);
```

5. Add a new Legislator (a special election has occured and Ron Estes needs to be added).

```
insert into Legislator (bioguide_id, ‘Last Name‘, ‘First Name‘,
                        birthday, gender, ‘wikipedia_id‘,
                        govtrack_id, official_full_name)
Value
('E000298', 'Estes', 'Ron', '1956-07-19', 'M', NULL, 412735, NULL);
```

6. Add a new term for that legislator (each legislator should have a corresponding Term)

```
insert into Term
VALUE ('E000298', '2017-04-27', '2019-01-03',
       'rep', 'KS', NULL, 4, 'Republican', 'h');
```

3 Useful Indexes

1. The query where we find all legislators that share a birthday was quite slow. We increased the speed of this dramatically by creating an index on their birthday.

   ```
   create index bday on Legislator (birthday);
   ```

   From our index log file using this index the query went from 12.58 sec to 0.07 sec

2. When finding all pairs of legislators with the same last name speed was imporved by adding a index on ‘Last Name‘ column

   ```
   create index lastName on Legislator (‘Last Name‘);
   ```

   From our index log file the query went from 6.97 sec to 0.17 sec

3. When counting unique ways in which legislators have voted the query speed was increased by added an index on the corresponding column:

   ```
   create index voteindex on Legislator_Vote(how_voted);
   ```

   From our index log file the query went from 1.43 sec to 0.27 sec

4 Modifying data from the last project

- For Legislator: First Name, Last name were the same due to an script typo.
- For Term: Party is Unknown for all members.
- For bill and amendments there are 'None' values instead of NULL.

All of this was fixed in our github repo and in our database: `https://github.com/hhweeks/congress_db/`

Term and Legislator were updated by correcting the script, Bill and Amendments were corrected by performing a sed replace on 'None' to NULL in the correct sql import files.