

## Laboratorio 9 - Trumpsacciones

Profesores: Claudio Gutiérrez

Matías Toro

Auxiliares: Scarlett Plaza

Daniel Radrigán

Cristian Salazar

Fran Zautzik

En el servidor del curso, usted encontrará el esquema `trx_b` con datos relativos a la elección presidencial realizada en EE.UU en 2016. (Alerta de spoiler: Trump gana).

Cuando se trata de votaciones, la integridad de los datos es crucial. Hoy trabajaremos en agregar transacciones y restricciones a las siguientes tablas del esquema `trx_b`:

- `estado`(`nombre`, `voto_electoral`, `cierre`, `num_candidatos`)
- `condado`(`nombre`, `estado`, `reportado`)
- `candidato`(`nombre`, `partido`)
- `votosPorCondado`(`candidato`, `condado`, `estado`, `votos`)

La tabla `estado` contiene los nombres de los estados EE.UU., la cantidad de votos electorales que poseen, la hora de cierre de la votación y el número de candidatos que corren en él. La tabla `condado` contiene todos los condados de un estado, junto con la fracción de votos reportados hasta el momento (inicialmente 0). Pueden haber dos condados con el mismo nombre en diferentes estados. En `candidato` se registran los candidatos a la presidencia y sus respectivos partidos. Finalmente, en `votosPorCondado` está la cantidad de votos por cada candidato en cada condado (inicialmente 0).

En la mañana del día de la elección, las tablas comienzan con 0 votos. En la base de datos hay actualizaciones de los datos de la votación por cada hora. Las tablas `votosPorCondado1`, ..., `votosPorCondado9` tienen la misma estructura que `votosPorCondado` y representan la cantidad **total** de votos en ese momento (`votosPorCondadon` corresponde a las votaciones totales **hasta** la *n*ésima hora). Las tablas `condado1`, ..., `condado9` representan las fracciones de votos emitidos hasta cada hora de la elección, en cada condado, con la misma estructura que `condado`. Estas tablas con actualizaciones contienen una tupla por cada vez que cambia algo. Dado que son valores totales hasta la hora respectiva, para actualizar `votosPorCondado` y `condado`, es necesario reemplazar (no sumar) los valores a partir de las tablas de actualizaciones.

Al final del laboratorio, usted debe entregar un archivo `.txt` con las respuestas de las siguientes preguntas (incluyendo comandos SQL que ud. escriba o modifique).

**P1.** 10 PUNTOS Para practicar con las actualizaciones y restricciones, es necesario que copie las tablas que hay en la base de datos desde el esquema `trx_b` a `trx_p`. Para ahorrar tiempo, descargue el archivo `trx_carga.txt` desde Material Docente. ¡**NO** llegue y corra los comandos! Debe hacer las siguientes modificaciones a los comandos en `trx_carga.txt` antes. Las llaves primarias se definen como se indican anteriormente para las cuatro tablas correspondientes.

- Para cada tabla, reemplace “GGGGG” por un nombre único para su grupo.
- Añada una restricción de llave primaria a la tabla `[grupo]_estado` (ahora con el nombre de su grupo). Cree su tabla y copie los datos desde `estado` usando el comando `INSERT INTO ...`.

- (c) Agregue una llave primaria y una llave foránea (referenciando a `grupo_estado`) en `grupo_condado`. Agregue una condición que verifique que la fracción reportada sea un valor entre 0 y 1 (inclusivo).
  - (d) Agregue una llave primaria en `grupo_candidato`.
  - (e) Añada las restricciones de llave primaria y foránea relevantes en la tabla `grupo_votosPorCondado`. Cree y cargue los datos de su tabla.
- P2.** 10 PUNTOS Ahora queremos ir actualizando los datos sobre las votaciones en su copia de las tablas, comenzando con la primera hora de elección. Revise los contenidos de la tabla `votosPorCondado1`. Construya una consulta `UPDATE` en SQL para actualizar su tabla `grupo_votosPorCondado` para cada tupla en `votosPorCondado1` según corresponda (basada en la llave primaria). Ejecute su instrucción.
- P3.** 10 PUNTOS Haremos lo mismo con la fracción de votos reportados durante la primera hora. Revise los contenidos de la tabla `condado1`; sigue la misma idea que la tabla `votosPorCondado1`. Escriba las instrucciones para actualizar `grupo_condado` con los datos de `condado1`.
- P4.** 10 PUNTOS Piense en lo que pasaría si actualizamos `grupo_votosPorCondado` y luego intentamos actualizar `grupo_condado` y ocurriera un problema. ¿Los datos estarían bien? No, pues tendríamos votos contados en el condado, pero la fracción de datos escrutados seguiría siendo desactualizada. Por lo tanto, use una transacción para actualizar ambas tablas atómicamente, pero esta vez usando los datos de la segunda hora de votación.
- P5.** 10 PUNTOS Modifique su transacción para actualizar las horas siguientes (sí, puede seguir *copypasteando* hasta completar todas las horas; puede usar una transacción por hora).
- P6.** 10 PUNTOS Putin envía sus hackers a cargar datos maliciosos a su base de datos para lograr la victoria de Clinton.<sup>1</sup> Es decir, envía como actualizaciones las tablas `votosPorCondadoX` y `condadoX`. Lamentablemente para Vladimir (y para Hillary), los hackers creyeron que la columna `reportado` correspondía a un porcentaje y no a una fracción. Cree una transacción para cargar `votosPorCondadoX` y `condadoX` en sus tablas atómicamente. Vea cuidadosamente la salida de Postgres. Realice una consulta para verificar el condado y candidato que trataron de hackear. ¿Putin lo logró?

---

<sup>1</sup>... imagínese que estamos en un universo alternativo.