CSC343 Introduction to Databases
Project Phase 2

November 19, 2021

Sarah Mansoor 1004183251 Sumaya Hassen 1000968679

---

1 Design Decisions

1.1 Questions

We are adding to question 3 from phase 1 to make it more specific.

3 Do boards that have a negative progress in credit accumulation by end of grade 11also have a negative progress in four year graduation rate? What about five year graduation rate? Do the results vary by board type?

1.2 Integrity Constraints

We are adding/changing these integrity constraints.

- CreditAccumulation[bID] $\subseteq$ Board[bID]
  - We are adding this integrity constraint to specify that CreditAccumulation has foreign key bID from Board.
- Graduation[bID] $\subseteq$ Board[bID]
  - We are adding this integrity constraint to specify that Graduation has foreign key bID from Board.
- $\sigma_{eResults<0 \vee eResults>1}\text{EQAO} = \varnothing$
  - We are adding this integrity constraint due to a mistake in this integrity constraint from phase 1:
    - $\sigma_{eResults<0 \wedge eResults>1}\text{EQAO} = \varnothing$
  - We made the mistake of saying AND when it should have been OR. The reason being that eResults cannot be less than 0 AND greater than 1 at the same time so this constraint would mean nothing.
- $\sigma_{oResults<0 \vee oResults>1}\text{OSSLT} = \varnothing$
  - We are adding this integrity constraint due to a mistake in this integrity constraint from phase 1:
    - $\sigma_{oResults<0 \wedge oResults>1}\text{OSSLT} = \varnothing$
  - We made the mistake of saying AND when it should have been OR. The reason being that oResults cannot be less than 0 AND greater than 1 at the same time so this constraint would mean nothing.
- $\sigma_{fourRate<0 \vee fourRate>1}$ Graduation $= \varnothing$
  - We are adding this integrity constraint due to a mistake in this integrity constraint from phase 1:
    - $\sigma_{fourRate<0 \wedge fourRate>1}$ Graduation $= \varnothing$
  - We made the mistake of saying AND when it should have been OR. The reason being that fourRate cannot be less than 0 AND greater than 1 at the same time so this constraint would mean nothing.
- $\sigma_{fiveRate<0 \vee fiveRate>1}$ Graduation $= \varnothing$

- We are adding this integrity constraint due to a mistake in this integrity constraint from phase 1:
  - $\sigma_{fiveRate<0 \wedge fiveRate>1}$ Graduation $= \varnothing$
- We made the mistake of saying AND when it should have been OR. The reason being that fiveRate cannot be less than 0 AND greater than 1 at the same time so this constraint would mean nothing.

2 Cleaning Process

- We created a new .csv file for each table based on the schema we had designed in phase 1. We decided to do this since our data set does not have too many rows so this could be easily done.
- Once we did this, we imported the files into postgreSQL along with our schema.
- We used COPY to import the .csv files.
- We had some errors doing so, since the NOT NULL constraint was violated due to some NULLS the data set. So we went into each .csv file to manually remove the rows with any NULL values. There were 2 of them in every data set.
- There were no other constraints we found that seemed to be violated in the data sets when importing the data in postgreSQL.
- Since our dataset only consisted of 72 rows, we did not have any trouble with overloading the database server and did not need to check for any additional constraints or violations when removing the 2 rows in each table with NULL values.