

Decisions:

Decisions we made:

1. "session\_year", "start year", "start minute", "start quarter", and "end year" are not included in the database we built since it is not very useful in analyzing the questions we raised.
2. "timeID" and "limitID" are added as primary keys in times and limits tables.
3. All sub-tables are separated from the original recreational courses table due to a purpose of redundancy removal. Therefore, during the cleaning process, all the redundant(same) tuples are removed.
4. In the original data(XLSX type), comma is used to separate days in variable Days\_of\_the\_Weeks, such as "M, W". However, when we convert the XLSX file to a csv file, all columns are also separated with comma, which induces a problem that sql recognizes different days of the week as separate columns. Therefore, we modify the variable Days\_of\_the\_Weeks in the XLSX file by separating days with a space(" ").

Cleaning steps:

All data in the original table is not null and clearly labeled. Therefore, not a lot of cleaning processes are undertaken. However, during the process of loading, some schema mistakes are found.

1. For the programs table, we originally define the constraint that subsection have seven possibilities, including 'Child/Youth', 'Adult', 'Early Child', 'Child', 'Youth', 'Older Adult', 'All Ages', but we found by loading the data that there is another possibility, which is 'Youth/Adult' in the table. So, we add Youth/Adult in the constraint.
2. We found that we mis-spelled 'regular' in the schema.