FINAL REPORT

Team Members: İrem Uslu 170709043

Hatice Özgül Bilici 170709023

Reşit Ökkeş Çiçek 170709029

The data we found gives detailed information about the Railroad crossing accident in the USA. These accidents have many different attributes such as the date of their occurrence, the weather conditions in which the accident occurred, the number of dead and injured. This dataset can be used for a large number of research such as accident prediction, examining the locations of accidents, extracting cause-effect relationships to predict accidents.Also,the dataset cover accidents from 1st January 1975 to 28 February 2021.

2) We made changes in our database according to Phase 1 and added new tables. We created 2 new tables named Incident2Value and IncidentTypeName by removing some data from the Incident table. We added the data we extracted from the Incidents table to these tables. We transferred the data such as temperature, car number, train speed, user age, injured, killed,number of loco

from incidents to the IncidentTypeName table. We linked the Inciden2Value table with the incidents table. We also linked the Incident2Value table with the incidenttypename table. Thus, our table has become compatible with 3NF rules.

3) We first downloaded our data as a .csv file from https://www.kaggle.com/yogidsba/us-highway-railgrade-crossing-accident. Then, we transferred the data in the csv file in php language to phpmyadmin as a single table. The code first creates a single table by taking the attributes from the first row in the csv. Then it transfers the data in each row to this table. Afterwards, we wrote new code, pulled the data from the table, and transferred all the data to the tables in our er diagram with the help of if statements. In our data, there were 23 rows belonging to the same accident and report number. Because of Date, Accident Number, Report Number are primary keys, we did not transfer these 23 rows belonging to the same primary key.

4) Our data has 239464 rows. So, we couldn’t use mySQL Workbench. We have a Centos 7 server, we used it with PhpMyAdmin and PHP.

5)

**View incidentsWithRailRoad :** It shows incidents details with Rail Road Company, Highway User, Highway Position, Visibility Situation.

**View incidentsWithStates:** It shows incident’s number, report key, date, time and county, state informations. For example we can use this when we want see a state’s incidents.

**View incidentsWithTypes:** It shows incident number, report key, incident details like temperature, injured numbers or death numbers. For each detail like temperature, injured, killed it has one row.

**PROCEDURE HighestTrainSpeedKilled:** Selects the number of the incident with the highest train speed. Then it displays the total number of deaths in incidents at this speed.

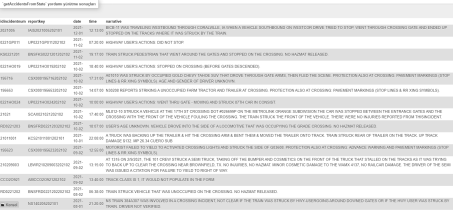
**PROCEDURE LowestTrainSpeedKilled:** Selects the number of the incident with the lowest train speed. Then it displays the total number of deaths in incidents at this speed.

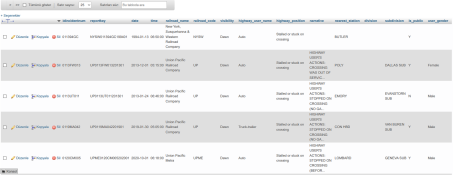
**PROCEDURE getAccidentsFromState:** Shows the last 20 accidents of the entered state.

**PROCEDURE getMostState:** It gives the latest accident information in the state with the most incidents.

6)

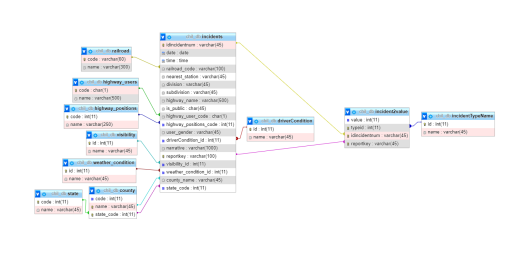
Our favorite procedure **getAccidentsFromState’s output:**

****Our favorite view **incidentsWithRailRoad**’s output:



7) On the return of our Phase 1 report, we learned from our teacher that our table is not 3NF-compliant and that some data could be stored more efficiently. We have made the necessary corrections.

8)

9)

