CS450: Final Documentation

Emily Risley, Nicholas Grogg, Timothy Kudryn

04-28-2020

**Overview**

**Project design**

**Languages**

**Roles**

**GUI**

The GUI is made using tkinter. To run the GUI, you must have the following python packages installed:

tkinter

mysql.connector

sys

numpy

pandas

plotly

Install the packages in a Anaconda environment with a name of your choice. To start your GUI, use command "python Launch.py" when in the python folder.

**Statistics**

The Statistics Page allows the user to interact with numerical data in the database. The user must fill out several fields: state (must be long form name, capitalized. ex. Alabama. Can also be “All”, which provides an aggregate statistic), date (in form YYYY-MM-DD), and data type (cases, deaths, rate of death (deaths/cases)).

**Interactive Plots**

The Plot Page in the GUI will have 3 buttons: Cloropleth, Time Series, and Plot Report. Plot Report is covered in more depth in a later section.

*Choropleth*

Choropleth takes the user to a new page. On this new page, the user can enter a date (form YYYY-MM-DD) and select a data type (cases, deaths, rate of death (deaths/cases)). After hitting plot, the GUI will attempt to open a webpage in the user's default browser and display the generated graph. The graph is a United States choropleth, and shows higher counts/rates as more yellow and lower counts/rates as more purple. The plot will also be saved as choropleth.png in the GUI's home directory.

**Plot Report**

**Backend**

**Web Server**

For a web server, a web hosted CentOS 7 Linux server was used. This configuration was NOT a LAMP stack, as Apache and PHP were not used. Access was done via a user, password, IP and database combination by the front end. It was also possible to log into the server for maintenance using an SSH key, as password logins were disabled. CentOS 7 was chosen for it’s stability and extensive documentation.

**Database**

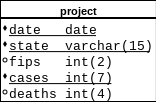
This database was a MariaDB database using the InnoDB engine. This engine was chosen as it’s an atomic engine, so transactions are either entirely successful or fail completely. The root user is not used for remote connections, so a special remote user was configured specially for the CS450 database.

**Database Values**

Most of these values are somewhat obvious, however the FIPS (Federal Information Processing Standards) codes are not. These are unique numeric codes assigned by NIST (National Institute of Standards and Technology) assigned to states and counties. For this project only the two digits used for states were given in the dataset, with the three digit county FIPS numbers not included.

|  |  |
| --- | --- |
| **Column** | **Description** |
| Date | When the data was recorded |
| State | What state is the data from |
| FIPS | Two digit code used to identify a state |
| Cases | The number of recorded cases |
| Deaths | The number of confirmed deaths |

**ER Diagram**



Underlined values are primary keys, these make up a composite primary key

Hollow symbols (Circles) are allowed to be NULL

Solid symbols (Diamonds) are not allowed to be NULL