ITWS Database Final Project
Team members: Aayushi Baghel and Josef Di Pietrantonio

Data:

The following are descriptions of the two datasets we used for the projects:

1.) Offenses Known to Law Enforcement by Metropolitan and Nonmetropolitan Counties 2015.

This data set was retrieved from the FBI website using the following link:
https://ucr.fbi.gov/crime-in-the-u.s/2015/crime-in-the-u.s.-2015/resource-pages/downloads/download-printable-files .
The license information for first data set is provided here: https://www.usa.gov/government-works .
The data contained in the set include violent and property crime totals by state by county for the year 2015.
Pre-Processing of this data set included separating the state name from the metropolis indicator as a block entry.
Headers and footnotes were also removed.

2.) Census Survey County Estimates from 2010-2017

The data set was retrieved from the census website using the following link:
https://www2.census.gov/programs-surveys/popest/datasets/2010-2017/counties/asrh/cc-est2017-alldata.csv
with file details at https://www2.census.gov/programs-surveys/popest/technical-documentation/file-layouts/2010-2017/cc-est2017-alldata.pdf
The license information for second data set is provided here:
https://www.census.gov/about/policies/open-gov/open-data.html
The data contained in the set are census demographics of males, females, hispanics, and non-hispanics in the United States by state by county
Pre-Processing of data set included choosing the correct year code for 2015 and age code for all age groups.

What Files are in the Folder?

ProjectSchema.sql: defines the schema of the Database

Demographics_fill.py: This python file basically fills the database with the values of csv files

app.py: This contains flask framework and handles the routing between browser and python file

database_updation.py: this is the file that updates and selects from Database

login.html: a login page which asks for user id and password, which are then later used to access the database

queries.html: a html page to provide input queries

result.html: for displaying results in a browser

error.html: if user is not allowed to access data base

Note: queries.html, result.html, error.html are in the templates folder

roles.sql: A file defining assignment of roles to two different users(demo and demo_select)

requirements.txt: contains the two files that are needed to be pip installed(psycopg2 and flask)

Processing.sql: since the demographics data consisted of string 'county' after each county name we had to process it so that it matched with the crime table

How to build your application?

Our application was built using the python IDE PyCharm.
The following is how it was built.
1) user roles were created
2) Demographics_fill.py was executed
3) app.py was executed
4) login.html was opened and data entered
5) database_updation.py takes input for queries as well as user id and password from app.py and returns output via results.html

How to load the data into the application?

First we created roles(roles.sql) and granted privileges
We used python3 in command line to create the Project Schema.
We used python3 to load the data using the Demographics_fill.py file.
Then we made an UPDATE to the four demographics tables : MALE_Demos, FEMALE_Demos, Hispanic_Demos, and Non_Hispanic_Demos.
We updated the tables by removing the space and word ' County' after all of the county names in order to make table joins on state county combinations.

How to run and use the application to explore the data?

1.) Run app.py in the background.
2.) Run login.html to open authentication page
3.) select the query you want, according to privileges you have

For someone who does not have access to the database:
Upon entering user and password credentials they will be denied access and brought to an access denied html page.

For someone who has selection authentication:
Upon entering acceptable user and password credentials they will be taken to an html page that has radio buttons to choose from directed queries.
4.) Select the desired directed query of interest.
5.) Select from drop down list or enter necessary fields to conduct the query.
6.) Click submit to run the query and view the results.
7.) Exit out of the html and return to step 2

For someone with database insert AND update authentication:
Upon entering acceptable user and password credentials they will be able to perform insert as well as the update