This project involved several modules and many techniques that I had no previous experience with. The requirement of having three different data sources was definitely one of the biggest challenges for me. I have worked primarily with flat file sources, and only used a handful of APIs. The web table source was something I had never seen, and web scraping in general was a new topic for me. It took quite a bit of research to understand the HTML structure and how to parse it effectively.

Once I managed to get the data sets imported, most of the transformation/cleaning steps that I performed were fairly straightforward. The biggest task overall was definitely cleaning up strings—both in the column names and in the county name data. Since I was planning to use the county names as the key for joining the data sets, I knew it would be important to keep my formatting completely consistent. This project was a great refresher on string methods; it had been a long time since I used most of them. I also gained a lot of useful experience using the “apply()” method to apply various functions to my data frames.

Another big challenge for me was the database portion of the project. I did not have any experience working with SQL or databases, so this was a lot of new material to absorb. The modules I was using (pandas, sqlite3) have convenient methods for accomplishing most of the tasks I needed to accomplish, but I still need to spend much more time studying the underlying concepts. I enjoyed finally learning some SQL, though some of the more advanced topics from the textbook are still difficult for me to understand. Completing this project has given me enough knowledge to continue practicing working with my own databases, and I look forward to exploring SQL and database manipulation in more depth.

Finally, the visualizations portion of the project lead me to research a number of new features/options of Matplotlib. I have used this module quite a bit, but I typically only produce basic plots with default settings. This was a great opportunity to explore the module’s capabilities in more depth, and get some practice using more advanced plotting and layout manipulation.

Overall, I thought this project, and the course as a whole, was very effective. Some courses I have taken used proprietary modules and functions, and I really appreciated that this course’s content always focused on using common, industry accepted modules. I believe this makes the skills and knowledge we have learned in this course far more relevant to the real-world applications we can expect in the professional world. Thank you for all of your feedback along the way!