**Extract Transform Load Project**

For this project, I scraped the transparent California website to obtain police salary for 2019. The website restricts visitors to scrape only 50 pages or maximum 2,500 records.

The outline below will illustrate how I extracted, transformed, and loaded my data.

**Extract:** I extracted the police salary records from the original data source from transparent California through web scraping. The data was formatted in html. Once I extracted the data, I saved the raw data into a csv file.

**Transform:** Following the web scraping, I transformed the raw data from the csv file by converting the data types from object to integer so that it can be easily loaded to a database because SQL is strict in its data type. I also replaced the text ‘Not provided’ to ‘0’ in order to have a consistent data type.

![Table

Description automatically generated]()

**Load:** After I completed the transformation process, I connected to my local database in pgAdmin and loaded the transformed data into the database.