1. **SEARCH**
   1. Searched data.gov for data sets around Chicago Public Schools:
      1. [School Profiles](https://catalog.data.gov/dataset/chicago-public-schools-school-profile-information-sy1718)
      2. [School Progress Reports](https://catalog.data.gov/dataset/chicago-public-schools-school-progress-reports-sy1718)
      3. [School Locations](https://catalog.data.gov/dataset/chicago-public-schools-school-locations-sy1819)
   2. Included a data set around fast food restaurants to map using zip codes.
      1. [Fast Food Restaurants](https://data.world/datafiniti/fast-food-restaurants-across-america)
   3. Saved all files in .csv format to the data directory inside the project directory.
2. **EXTRACT / TRANSFORM - PYTHON / JUPYTER NOTEBOOK**
   1. Imported dependencies:
      1. pandas
      2. create\_engine from SQLAlchemy
   2. File one - school profiles:
      1. Read in the first file containing school profiles into a data frame
      2. Trimmed down the original data frame from 97 columns to 4 main columns and renamed to:
         1. school\_id
         2. finance\_id
         3. short\_name
         4. long\_name
   3. File two - school progress reports:
      1. Read in the second file containing school progress reports
      2. Trimmed down the original data frame from 168 columns to 3 main columns and renamed to:
         1. school\_id
         2. short\_name
         3. award
   4. File three - school locations:
      1. Read in the third file containing school locations
      2. Trimmed down the original data frame from 15 columns to 9 main columns and renamed to:
         1. school\_id
         2. network
         3. short\_name
         4. address
         5. zip
         6. grades
         7. community
         8. ward
         9. alderman
   5. File four - fast food
      1. Read in the fourth file containing national fast food information
      2. Filtered out any fast food restaurants outside the City of Chicago by specifying city must equal Chicago
      3. Cleaned up the existing list of restaurants by removing duplicates
      4. Trimmed down the cleaned data frame from 12 columns to 3 main columns and renamed to:
         1. name
         2. zip
         3. state
3. **LOAD - POSTGRESQL / PGADMIN**
   1. Created chicago\_public\_schools database in pgAdmin
   2. Using query tool, created 4 tables to match transformed data frames:
      1. school\_profiles (set school\_id as the PRIMARY KEY)
      2. school\_reports (set school\_id as the FOREIGN KEY)
      3. school\_locations (set school\_id as the FOREIGN KEY)
      4. fast\_food (not all zips match so keys wouldn’t work)
   3. Saved as .sql file in the project directory
   4. Using Jupyter Notebook:
      1. Set up connection to database
      2. Loaded data frames into the database
   5. Saved final Jupyter Notebook as .ipynb file in the project directory