Project Title
- ● USA Hospitals

Team Members
- ● Madeline Proctor
- ● Shapoor Mirzakwal

Project Description/Outline
- ● We are creating a data source of hospitals around the USA and identifying if they are a children's hospital or not.

Rough breakdown of tasks

https://github.com/maproctor95/hospitals_ETL.git

Collect Data	https://www.kaggle.com/datasets/digitalbro/usa-childrens-hospitals https://data.world/dhs/hospitals
Extract Data from sources	● CSV file from Kaggle ● CSV file from Data World
Clean and Format the data	Created a common column to merge on and then merged the data together Delete after combining two data sets: {'X', 'Y', 'ID', 'ZIP4', 'COUNTYFIPS', 'NAICS_CODE', 'NAICS_DESC', 'SOURCE', 'SOURCE_DAT', 'VAL_METHOD', 'VAL_DATE', 'ALT_NAME', 'ST_FIPS', 'OWNER', 'TTL_STAFF', 'DATE_CREAT', 'Beds', 'Alternative Name', 'Address', 'Address 1', 'Address 2', 'City', 'State', 'Zip', 'Hospital Website', 'Latitude', 'Longitude', 'Type', 'Wikipedia Source', 'Magnet Designation'}) Renamed:('Hospital': 'hospital_name', 'ADDRESS': 'address', 'CITY': 'city', 'STATE': 'state', 'ZIP': 'zip', 'TYPE': 'type', 'STATUS': 'status', 'POPULATION': 'population', 'COUNTY': 'county', 'COUNTRY': 'country', 'LATITUDE': 'latitude', 'LONGITUDE': 'longitude', 'WEBSITE': 'website', 'BEDS': 'beds', 'TRAUMA': 'trauma', 'HELIPAD': 'helipad', 'ADDRESS2': 'address2', 'TELEPHONE':'telephone', 'FID': 'fid', 'STATE_ID': 'state_id'})
Create a database with the data	Hospital database in PGAdmin
Write up a summary	In the ReadMe file in Jupyter notebook