**Team - City Crimes**

* Kevin Grzyb
* Mamtha Sharma
* Amy Peters

**Project Technical Description**

**Extraction**

Original data sources were found below and the data was originally formatted in CSV.

* Data.world – NYC crime data 2014
* Data.lacity.org – LA crime data 2014
* Data.cityofchicago.org – CHI crime data 2014

**Transformation**

We took the following steps to transform our dataframes:

* Date – transformed date and remove timestamp
* Truncated data frame to relevant columns
* Added column for City
* Renamed columns
* Reformatted Primary Crime Category
* Re-ordered columns
* Saved to csv
* Import into pandas to join data

**Load**

We took the following steps to load the transformed dataframes into a Postgres database

* Concatenated the data from 3 cities
* Created a new database called ETL-project in Postgres
* Created a table called city\_data to receive the concatenated crime data
* Connected to Postgres server through SQLAlchemy
* Tested the connection
* Uploaded the data to the table in Postgres using “to\_sql”
* Verified data accuracy by running queries