# **ETL Project**

Zack Snyder, Yin Cai, Tricia Palomino

**Data Cleanup & Analysis**

* The sources of data:
  + [Zillow Housing Price](https://www.kaggle.com/zhenyufan/zillow-housing-price)
  + [US Zip Code Latitude and Longitude](o%09https:/public.opendatasoft.com/explore/dataset/us-zip-code-latitude-and-longitude/export/)
* The type of transformation needed for this data (cleaning, joining, filtering, aggregating, etc):

o Cleaning: Text to columns

o Filters: Most recent year (2018), Exclude Time Zone,

o Joins: Zillow (base) with left join of Lat/Long dataset

* The type of final production database to load the data into (relational or non-relational).

o SQL: Zip = zip

* The final tables or collections that will be used in the zillow\_housing database. This is the final table we use after joining all the tables/DataFrames.