# **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):
  + Cleaning: Text to columns
  + Filters: Most recent year (2003), Exclude Time Zone,
  + Joins: Zillow (base) with left join of Lat/Long dataset
  + Aggregate: 12 mo. Average by zip
* The type of final production database to load the data into (relational or non-relational).
  + SQL: zip = zit
* The final tables or collections that will be used in the production database.