Data Science and Visualization

By: Charleen Carr

Extracted housing data from Zillow on New Jersey, Tennessee, California, Utah and Texas.

Extracted data sets from realtor.com – used the following “RDC\_InventoryCoreMetrics\_Zip.csv”

Transformed all data in separate Jupyter Notebook files.

Loaded all data into separate schemas in MongoDB.

From MongoDB:

extracted both schemas

loaded into a pandas data frame

merged them together by zip code

then calculated the differences

Sources:

<https://www.zillow.com/>

<https://www.realtor.com/research/data/>