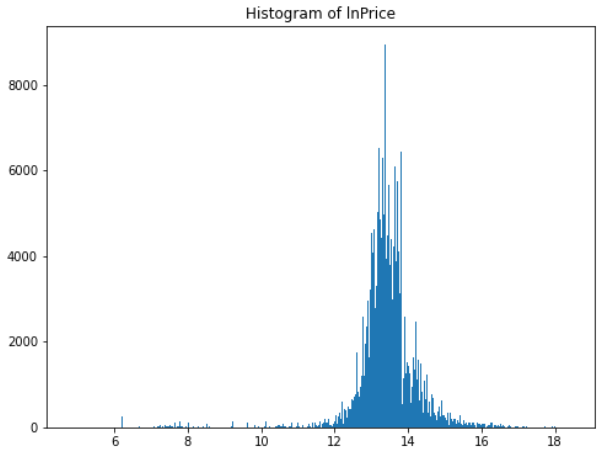
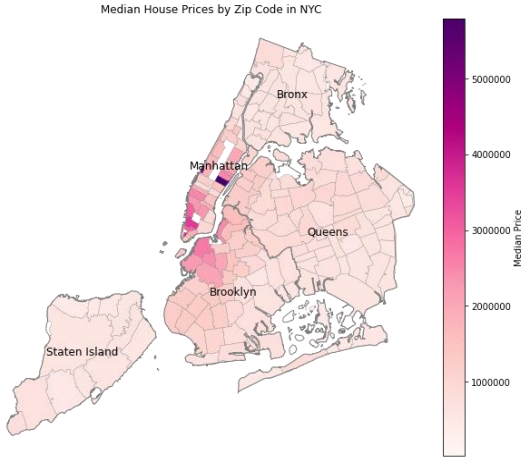
CS590 Project: Zillow Housing Price

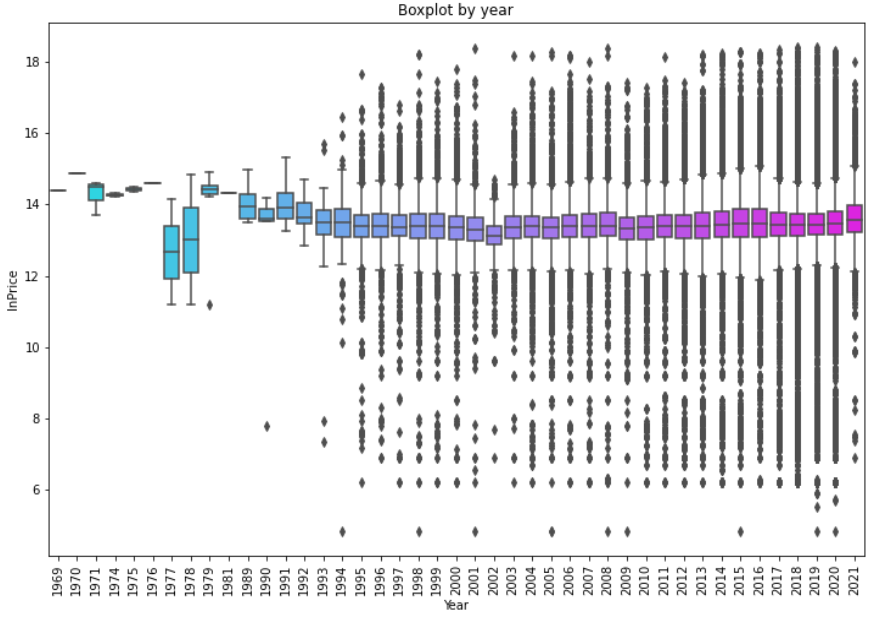
Minda Zhu (mz203), Oded Schlesinger (oas13), Uyanga Gansukh (ug19), Avinno Faruk (af330)

- **DATASET:** NY Housing data from Zillow API, Census Data, Income data from IRS.
- **PURPOSE:** Data-driven correlational study of the factors that drive the housing prices, with a retrospective focus on flipping.
- **RELEVANCE:**
  - Insights into the failure of Zillow's predictive model.
  - Insights into housing price mechanism for housing market participants.
- **UNEXPECTED FINDINGS:**
  - Housing price distribution does **NOT** match income distribution.
  - Houses in Lower Manhattan and Northern Brooklyn have shown extreme prices for which the markup in prices could **NOT** be explained by the increase in their intrinsic value in a short period.
  - Surprisingly, the most **expensive houses** tend to be those sold frequently and relisted on Zillow, resulting in a **flipping effect**. This can be explained by the fact that most of the real estate is traded for **investment** and **speculating** purposes.

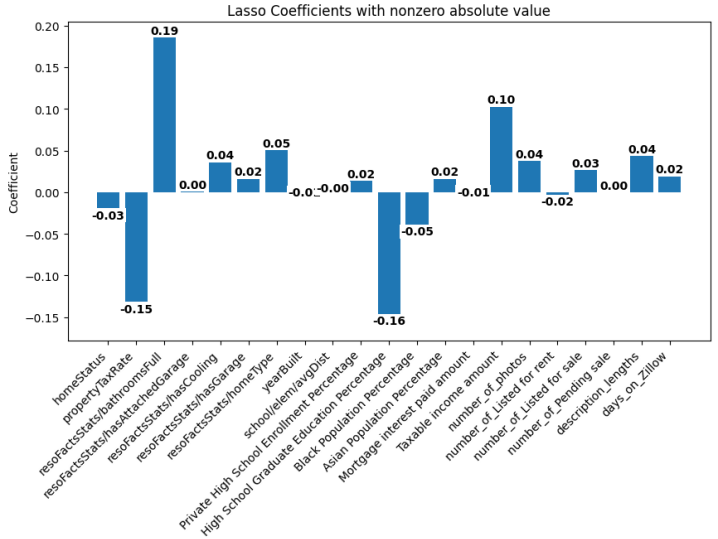
New York City housing price trend







Model: Lasso regression



Finding: Flipping effect