Summary of the London Case Study

Question: “Which borough of London has seen the greatest average increase in housing prices over the (approximately) two decades covered by the dataset?”

Answer: Hackney, with 6.198 times increase from 1998 to 2018.

Summary of procedure: I used pandas, numpy, and matplotlib in Python with a Jupyter Notebook to analyze the data. The dataset used was from: <https://data.london.gov.uk/download/uk-house-price-index/70ac0766-8902-4eb5-aab5-01951aaed773/UK%20House%20price%20index.xls>

I cleaned the data by transposing the data, dropping the first row, renaming “Unnamed: 0” and NaN columns, melting the dataframe, converting some column types, and dropping rows that do not contain boroughs. This data was quite messy at first, with several missing values and mismatched column counts, but these were fixed by removing the rows that were not boroughs. A year column was created separate from the date, as the year is the focus of the study. A ratio function was created to find the ratio between 2018 average price of a London borough and its average price in 1998. So, a 6.198 times ratio represents the average price of a Hackney borough in 2018 is 6.198 times its price in 1998. The top 20 ratios are represented on a chart and pictorially on the next page:

