

Time Series Modelling of Boston Neighbourhoods

By Tom Ribaroff

Our Task

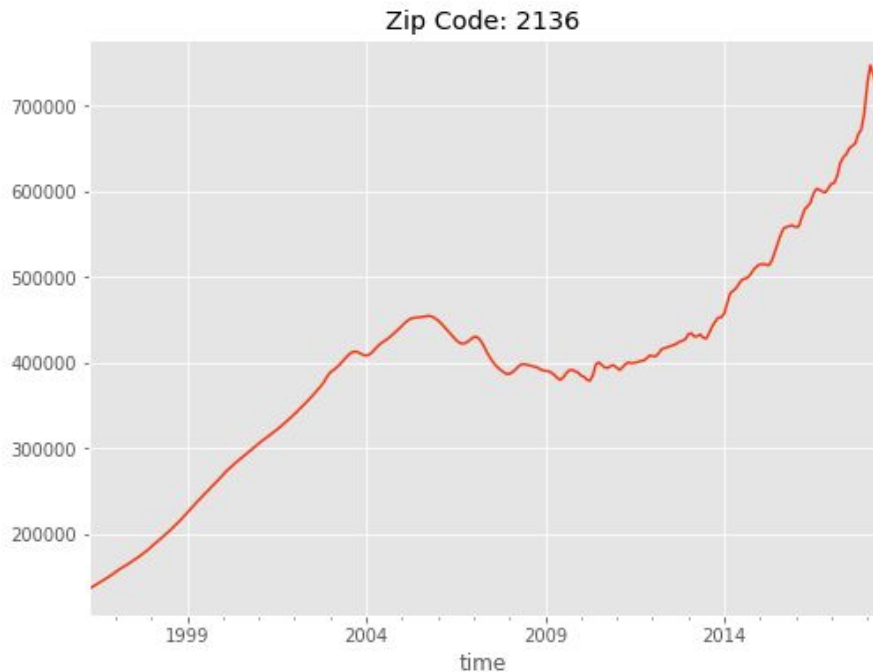
We have been given a very large Dataset, scraped from the website Zillow, on housing prices all across America. Our task is to answer the question:

“What are the top 5 zipcodes to invest in?”

The dataset is massive, and too large for a small laptop to handle, so we filtered our data to just contain that of the **city of Boston**. So we will attempt to find the best 5 zip codes in Boston to invest in!

| | RegionName | City | State | Metro | CountyName | time | value |
|---------|------------|-----------|-------|-------------|-------------------|------------|----------|
| 5850 | 1001 | Agawam | MA | Springfield | Hampden | 1996-04-01 | 113100.0 |
| 20573 | 1001 | Agawam | MA | Springfield | Hampden | 1996-05-01 | 112800.0 |
| 35296 | 1001 | Agawam | MA | Springfield | Hampden | 1996-06-01 | 112600.0 |
| 50019 | 1001 | Agawam | MA | Springfield | Hampden | 1996-07-01 | 112300.0 |
| 64742 | 1001 | Agawam | MA | Springfield | Hampden | 1996-08-01 | 112100.0 |
| ... | ... | ... | ... | ... | ... | ... | ... |
| 3835273 | 99901 | Ketchikan | AK | Ketchikan | Ketchikan Gateway | 2017-12-01 | 291900.0 |
| 3849996 | 99901 | Ketchikan | AK | Ketchikan | Ketchikan Gateway | 2018-01-01 | 294200.0 |
| 3864719 | 99901 | Ketchikan | AK | Ketchikan | Ketchikan Gateway | 2018-02-01 | 297500.0 |
| 3879442 | 99901 | Ketchikan | AK | Ketchikan | Ketchikan Gateway | 2018-03-01 | 302100.0 |
| 3894165 | 99901 | Ketchikan | AK | Ketchikan | Ketchikan Gateway | 2018-04-01 | 305100.0 |

Our definition of “The Best”



How do we define “best” in terms of investment? We could consider:

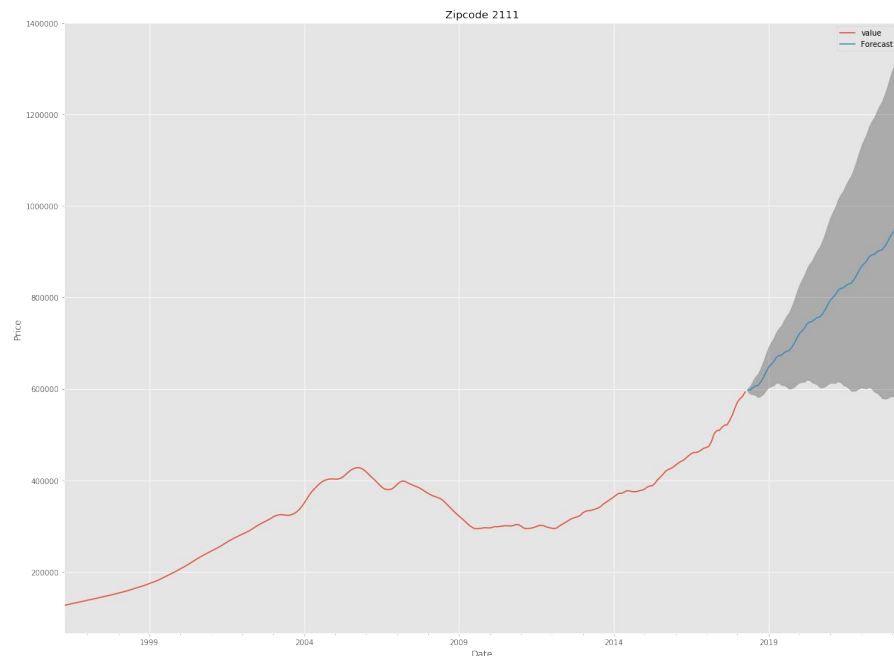
- Maximum gross return on our investment?
- Percentage rise of expected price
- Likelihood of losing money
- What time scale should we be looking at (5/10/20 years)?

All of them are sensible approaches, but for simplicity, this the one I am going for:

“The highest percentage rise in house price after 5 years, with a 95% assurance that our house price won't drop more than 10% in the first two years”

The outline of our method

- 1 - We use computer modelling (SARIMA) to model each of the 17 Neighbourhoods as accurately as possible
- 2 - Using each best possible model, it forecasts 5 years into the future
- 3 - We store each Neighbourhood's expected percentage price rise in 5 years, but only if it passes the test to make sure it isn't a risky investment
- 4 - We rank each stored Neighbourhood by percentage increase of price after 5 years, and return the top 5!



Our results

Great! We have our top 5 neighbourhoods.

Note: Although our top Zip Code has the highest percentage increase by 3%, the lower bound/upper bound on expected price is pretty wide! Perhaps we should consider the other neighbourhoods as better, because even if the returns are smaller, we are much more confident that these zip codes will be a moderate success as an investment.

| | Region | Final_known_price | 5_year_expected_price | 5_year_lower_bound_price | 5_year_upper_bound_price | 5_year_expected_pc_increase |
|---|--------|-------------------|-----------------------|--------------------------|--------------------------|-----------------------------|
| 0 | 2111 | 593600 | 961249 | 582642 | 1339856 | 61 |
| 1 | 2124 | 543900 | 862863 | 627755 | 1097971 | 58 |
| 2 | 2127 | 807900 | 1278992 | 1021348 | 1536636 | 58 |
| 3 | 2119 | 512600 | 797617 | 535603 | 1059631 | 55 |
| 4 | 2126 | 419400 | 651901 | 469589 | 834213 | 55 |

Conclusions

We have our final top recommendation in my opinion:

Invest in Zip Code 2127!

Further, this tool can be used to help anyone planning to invest in other neighbourhoods:

If you want to invest in a specific neighbourhood for a particular reason (say because it is close to a school, family member etc) then we can give you a good estimate for the likely return on a property.

We can also help to choose between two zipcodes that you like, or stop someone from investing in a potentially disastrous zip code!

