Time Series analysis on Zip codes

Intro

For this project I had a client that wanted to know what zip codes he should invest in. I got my dataset from Zillow which helped me get an idea what I wanted to do and how I wanted to do it. The dataset contained many columns that described the zip code. It gave the name of the city, the metro the city was located at and more.

Selection process for my top 5 zip codes

The first part of my project was to come up with a way to get my zip codes. I decided to use an investment strategy that was very safe and sound. My strategy was that I wanted zip codes that had stable prices and also had great returns. In order to get zip codes that fit my strategy my first step was to filter houses that had stable monthly returns. To do this, I decided to get the average standard deviation per month for each house. Then I decided to take the houses that had had less than the average standard deviation per month. This gave me my stable zip code prices. Next thing I wanted to do was to get the best returns out of these zip code. For this I decided to get the average monthly returns from my houses in the low standard deviation set. Lastly I picked my houses that gave me the best returns.

Modeling

After I picked my zip codes, I decided to run sarima models on them. I picked sarima models because I’m making an assumption that all zip codes have seasonal influences on price

After running zip codes through my first 5 models I got my basic vanilla models.

Improving my modeling

After I got my vanilla models, I decided that I wanted to improve them by using a grid search function on my sarima parameters. This helped me select my best parameters for my sarima models

Conclusion and future work

After running my models I decided that the best zip codes were 2 in Houston Texas, 1 in El Passo Texas, 1 in Pflugerville Texas and finally one in Nashville Texas. I would advise my client to invest in these zip codes because based on my models he could make a 50% return on investment if my client invest. Another thing I want to focus on is the modeling aspect of it since there was something slightly off on my models. I want to learn what is causing it and why. That is my project I hope you all liked it. Thanks for giving this a read