Laboratory Exercise - 2

Objective

- This case study provides an opportunity to demonstrate our ability to combine data sets and produce meaningful analysis. Specifically, we would like to provide a decision maker with more than just data – we want to provide insights, understanding, and wisdom. This exercise allows the student an opportunity to demonstrate progress (or mastery) of learning objectives 1, 2, 3, 4, and 5:

1) **O**btain data and understand data structures and data elements.
2) **S**crub data using scripting methods, to include debugging, for data manipulation in R and other tools.
3) **E**xplore data using essential qualitative analysis techniques including descriptive statistics.
4) **M**odel relationships between data using the appropriate analytical methodologies matched to the information and the needs of clients and users.
5) **IN**terpret the data, model, analysis, and findings. Communicate the results in a meaningful way.

Instructions

- The research question is can we predict which three zip codes provide the best investment opportunity for the Syracuse Real Estate Investment Trust (SREIT)?
- Using the base data available from Zillow (files.zillowstatic.com/research/public/Zip/Zip_Zhvi_SingleFamilyResidence.csv)
    - Review the data – clean as appropriate
    - Provide an initial data analysis to include (but not limited to):
        - Develop time series plots for the following Arkansas metro areas:
            - Hot Springs, Little Rock, Fayetteville, Searcy
            - Present all values from 1997 to present
            - Average at the metro area level
- Using data from Zillow:
    - Develop model(s) for forecasting average median housing value by zip code for 2018
    - Use the historical data from 1997 through 2017 as your training data
    - Integrate data from other sources (think Bureau of Labor Statistics and Census data) to improve upon your base model(s)
- Answer the following questions:
    - What technique/algorithm/decision process did you use to down sample? (BONUS FOR NOT DOWN SAMPLING)
    - What three zip codes provide the best investment opportunity for the SREIT?
    - Why?
- Bonus: Develop a geographic visualization that in your view best depicts the data and recommendations:
    - By state

          ○   Median housing for Dec (state average)

Additional Instructions

- Don't forget what you learned in your previous courses; do your own work, document any assistance, use comments for clarity.
- Use python to conduct your analysis and produce your graphics

Submission Items

- Report with graphics
- Supporting notebook for the report with final data set