

## **Lab Exercise 2 - Syracuse Real Estate Investment Trust**

### **IST 718 - Big Data Analytics | Kelly Hwang**

#### **Introduction**

Property values have been propelled to record high numbers in some parts of the United States over the last decade. As a real estate investor, can it be predicted where the next opportunity for high returns will be? Using external studies and datasets of average single family home prices from all domestic zip codes, this forecast aims to suggest the top zip codes that Syracuse Real Estate Investment Trust should consider investing in.

Goals:

1. Develop model(s) for forecasting average median housing value by zip code for 2019
2. Predict top three zip codes for future investment opportunities
3. Plot time series data

#### **Analysis and Models**

##### **About the Data**

The following datasets were used in this analysis:

- Zillow public data set for Single Family Residence from 1996 - 2018

The following studies were used in conjunction for targeted downsampling:

- USA Today: [31 cities adding the most jobs as the US economy grows](#)

A variety of studies and datasets were considered for downsampling, including but not limited to unemployment rates, population size, population growth, and racial demographics. However, projected job growth opportunity an extremely important factor when considering property value growth, and the breakdown provided by USA Today provides a very thorough view into top job growth metro areas. The study compiled industry growth, employment / unemployment rates, and median household income to determine the fastest growing job markets, with some metro areas selected having double the national job growth rate. Over the past half decade, overall employment rose from 15.0 percent to 28.3 percent in the 31 cities on this list. Using this study, the following metro areas were marked in the existing dataset as top job growth opportunity metros:

- Albany, OR

- Atlanta-Sandy Springs-Roswell, GA
- Austin-Round Rock, TX
- Bend-Redmond, OR
- Boise City, ID
- Cape Coral-Fort Myers, FL
- Charlotte-Concord-Gastonia, NC
- Charlotte-Concord-Gastonia, SC
- Coeur d'Alene, ID
- Colorado Springs, CO
- Denver-Aurora-Lakewood, CO
- Elkhart-Goshen, IN
- Fayetteville-Springdale-Rogers, AR
- Fayetteville-Springdale-Rogers, MO
- Fort Collins, CO
- Gainesville, GA
- Greeley, CO
- Lake Charles, LA
- Lake Havasu City-Kingman, AZ
- Las Vegas-Henderson-Paradise, NV
- Medford, OR
- Nashville-Davidson--Murfreesboro--Franklin, TN
- North Port-Sarasota-Bradenton, FL
- Orlando-Kissimmee-Sanford, FL
- Phoenix-Mesa-Scottsdale, AZ
- Portland-Vancouver-Hillsboro, OR
- Portland-Vancouver-Hillsboro, WA
- Prescott, AZ
- Provo-Orem, UT
- Raleigh, NC
- Reno, NV
- St. George, UT
- Wenatchee, WA

Main attributes kept for analysis:

- RegionName - domestic zip code
- City, State, Metro had their own individual columns
- Every column is its own month/year from Apr 1996 to Jan 2019 with average single family home value per row

## Exploring the Data - Initial Analysis of Overall Data

Top 15 biggest states containing the most metro areas:

TX	56
OH	41
IN	39
NC	38
GA	34
PA	33
CA	31
MI	29
FL	28
IL	27
TN	27
NY	26
WI	25
MO	25
MN	23

Top 10 most expensive average values by state in 2018:  
state

DC	811585.648148
CA	757504.701577
HI	697839.919355
MA	463019.224620
CO	434521.063830
WA	411164.059305
NJ	402386.503984
NY	394115.227878
UT	341825.917431
MD	336392.650273

Top 10 most expensive average values by metro area in 2018:  
metro

San Jose-Sunnyvale-Santa Clara	1.637981e+06
San Francisco-Oakland-Hayward	1.328119e+06
Glenwood Springs	1.240951e+06
Summit Park	1.023750e+06
Los Angeles-Long Beach-Anaheim	9.309977e+05

Santa Maria-Santa Barbara	9.271787e+05
Vineyard Haven	9.043867e+05
Urban Honolulu	8.874222e+05
Edwards	8.765833e+05
Napa	8.424917e+05
Santa Cruz-Watsonville	8.024521e+05
San Diego-Carlsbad	7.727231e+05
Salinas	7.323408e+05
Kahului-Wailuku-Lahaina	7.264296e+05
Santa Rosa	7.262200e+05
Gardnerville Ranchos	6.718562e+05
Brunswick	6.599300e+05
Oxnard-Thousand Oaks-Ventura	6.562475e+05
San Luis Obispo-Paso Robles-Arroyo Grande	6.473116e+05
Hailey	6.222183e+05
Naples-Immokalee-Marco Island	6.074647e+05
New York-Newark-Jersey City	6.068555e+05
Key West	6.048500e+05
Ocean City	6.033236e+05
Kapaa	5.998125e+05
Seattle-Tacoma-Bellevue	5.722022e+05
Boulder	5.721327e+05
Boston-Cambridge-Newton	5.643083e+05
Bridgeport-Stamford-Norwalk	5.108698e+05
Barnstable Town	4.929771e+05

It's important to note that none of the top job growth metro areas are in the top 30 most expensive last year. This further justifies these metros as future potential investment opportunities, as the initial investment is not very costly.

### **Exploration of Arkansas Metro Areas**

The following metro areas were evaluated: Hot Springs, Little Rock, Fayetteville, Searcy

Top 3 metro areas in AK by average home value:  
metro

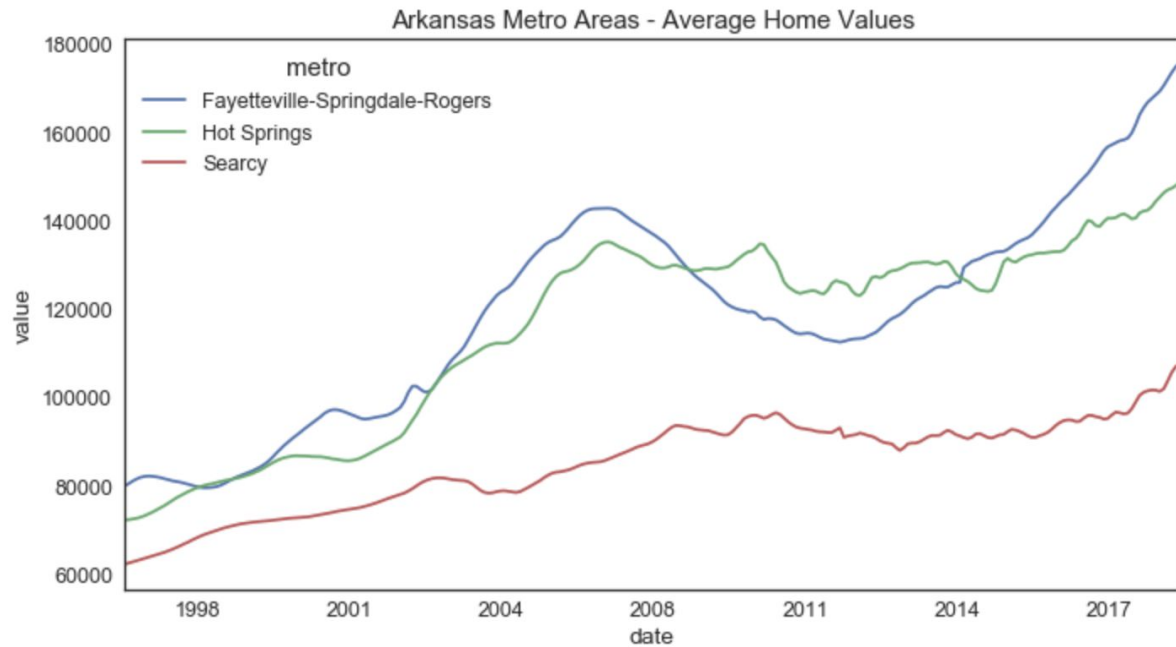
Fayetteville-Springdale-Rogers	119597.957839
Hot Springs	115423.962264
Searcy	85482.553191

Median value per year for Fayetteville:

year

1997	81050.0
1998	79050.0
1999	80750.0
2000	88050.0
2001	91800.0
2002	97250.0
2003	104000.0
2004	119000.0
2005	133500.0
2006	147050.0
2007	142900.0
2008	133700.0
2009	120500.0
2010	112150.0
2011	109050.0
2012	106500.0
2013	117550.0
2014	126000.0
2015	128750.0
2016	140000.0
2017	152850.0
2018	163600.0
2019	170500.0

Fayetteville-Springdale-Rogers is a metro area that was marked as high job growth opportunity.



### Exploration of Top Job Growth Metro Areas

Top job growth metro areas by state:

state

OR	5
CO	4
FL	4
AZ	3
WA	2
ID	2
TX	2
NV	2
NC	2
UT	2
GA	2
MO	1
IN	1
LA	1
SC	1
TN	1
AR	1

Count of zip codes by state:

state

CO 177

FL 172

GA 168

AZ 155

OR 116

NC 116

TN 87

NV 80

TX 75

ID 38

UT 30

WA 28

AR 21

SC 13

IN 11

LA 8

MO 3

Total zip codes for modeling used: 1,297

## **Models / Results**

Facebook's Prophet forecasting model was used to predict home values, using the data for these metro areas from 1997 - 2019 (Jan). All 1,297 zip codes were put through the Prophet model, and growth calculated for each. From this base model output calculation, the top forecasted growth areas were in Oregon and Florida. Below are the average margins of error for those regions:

- prediction off by +/- \$ 97,382.12 on average for zipcode 97210
- prediction off by +/- \$ 96,160.83 on average for zipcode 97201
- prediction off by +/- \$ 64,788.83 on average for zipcode 34236

Continuing use of the model, future dates were forecasted until Oct 2020, and with this new projection the top 3 highest growth predicted zip codes slightly changed, but were still in Oregon and Florida.

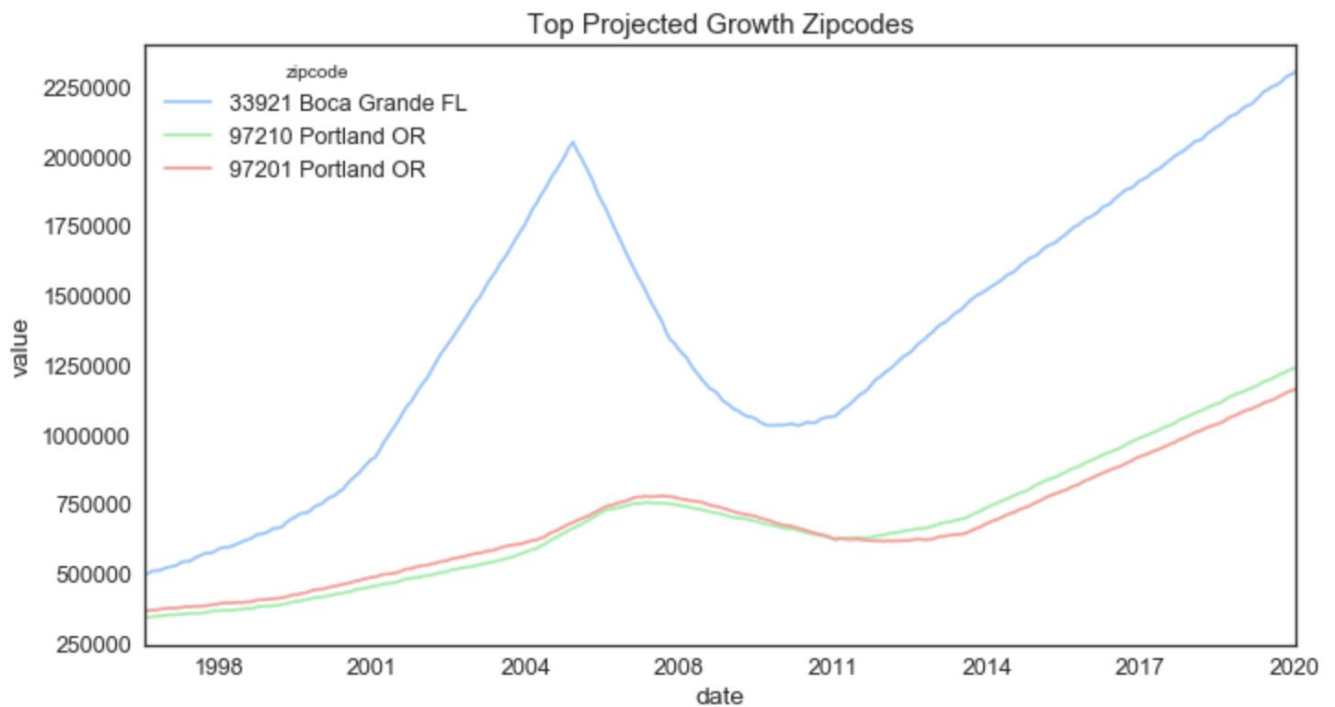
3 zip codes with largest growth:

33921

97210

97201

These zip codes are for the following cities:



## Conclusions

Using historical housing data from 1997 through 2017, the recommendation for investment opportunities would be to invest in Portland, OR and Boca Grande, FL as these are the zip codes with high job growth potential showing the greatest predicted home value growth rate. Portland's main growing industries are education and health services, with a rapidly growing population currently at 2.4 million (up 7.1 percent from 5 years ago). Boca Grande (in Florida's Cape Coral-Fort Myers metro area) has added 10,900 new jobs in the past half decade in the construction industry and continues to show signs of a growing population as well. Both of these areas are not yet at the top of the most expensive home values, so it would be beneficial to invest while they are still in the early growth stages.



## Resources

Code:

<https://docs.python.org/3/library/datetime.html#datetime.datetime.strptime>

<https://hackernoon.com/reshaping-data-in-python-fa27dda2ff77>

<https://dzone.com/articles/pandas-find-rows-where-columnfield-is-null>

<https://stackoverflow.com/questions/23836277/add-leading-zeros-to-strings-in-pandas-dataframe>

<https://stackoverflow.com/questions/42820696/using-prophet-package-to-predict-by-group-in-dataframe-in-r>

Data:

[https://files.zillowstatic.com/research/public/Zip/Zip\\_Zhvi\\_SingleFamilyResidence.csv](https://files.zillowstatic.com/research/public/Zip/Zip_Zhvi_SingleFamilyResidence.csv)

<https://www.usatoday.com/story/money/economy/2018/11/12/us-economy-grows-cities-adding-most-jobs/38319445/>