The Best Zip Codes For Your Investment

Caleb Elgut - September 2020

Introduction

- Data Source: Zillow
  - Includes: ID, **Zip Code**, City, County, State, Metropolitan Area, **Size Rank** (Grade of Urbanization), **Value per month**
- Python Libraries:
  - Pandas, Matplotlib, NumPy, SkLearn, Math, Warnings, StatsModels
- GOAL:
  - Examine time series data related to real estate buying/selling to determine which zip codes are the most efficient to invest in based on the needs of our stakeholders who desire a rural environment.
    - Risk Assessment is key
    - Low Risk & High ROI

How Do We Select Our Rural Zip Codes?

Method of Filtering Zip Codes

1. 25% lowest values in Size Rank Column
2. Values 15% above and below the median value
3. Coefficient of Variation (*Unitized Risk*) cannot be higher than .60
4. From here we choose the zip codes with the top 5 ROI

	RegionName	ROI	CV
13753	48894	2.561947	0.190111
11357	56360	2.290870	0.206014
11605	40008	2.035417	0.206328
13206	49339	1.900000	0.191471
11414	27019	1.759754	0.189331

```
Zipcode : 48894
Location: Westphalia, MI

Zipcode : 56360
Location: Osakis, MN

Zipcode : 40008
Location: Bloomfield, KY

Zipcode : 49339
Location: Pierson, MI

Zipcode : 27019
Location: Germanton, NC
```

Time Series???

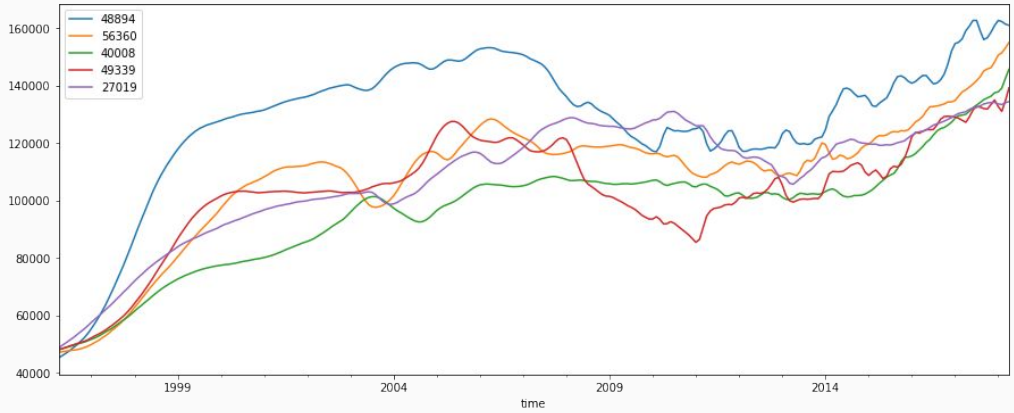
- Refers to datasets where the **progress of time** is an important dimension.
  - Stock Market Predictions
  - Climate Data
  - Yards Run by Week/Month/Year

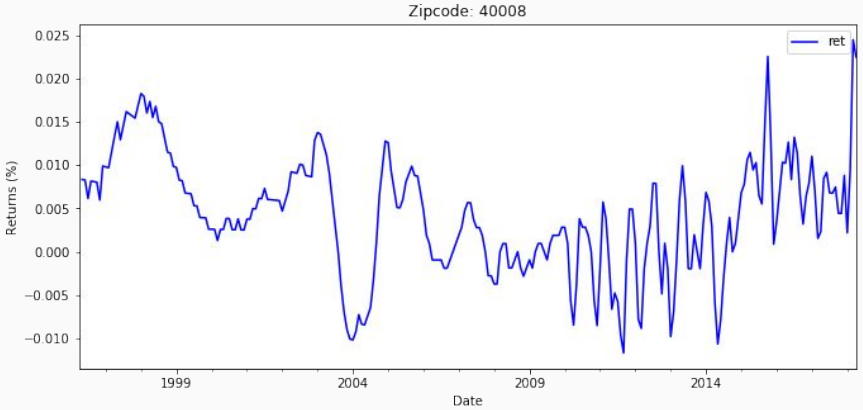
Steps For Time Series

1. Format: Long-Form
2. Plot & Test for Stationarity
    a. Stationarity: Statistical Properties Remain Constant Over Time
3. Get as Close to Stationarity as Possible
    a. Subtract The Rolling Mean (at times using weights)
    b. Differencing
4. Examine Autocorrelation & Partial Autocorrelation
5. Examine Effectiveness using Diagnostics
6. Forecast with SARIMA

Initial Time Series Analysis

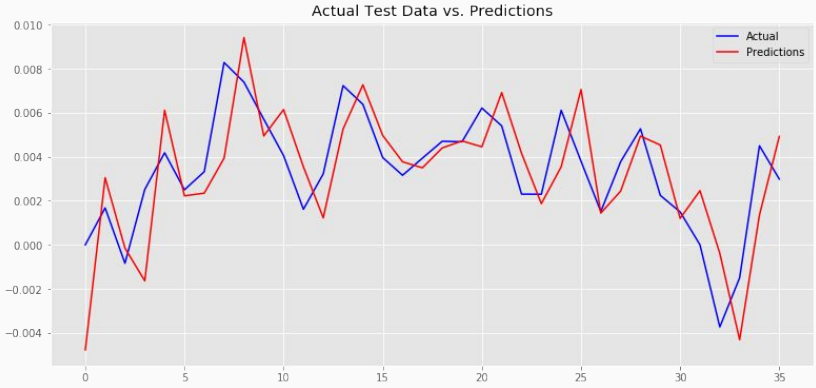
- Transformed from Yearly to monthly
  - Easier to examine changes

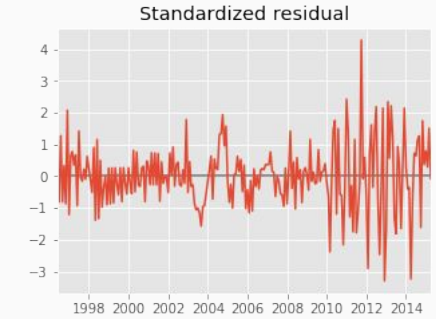


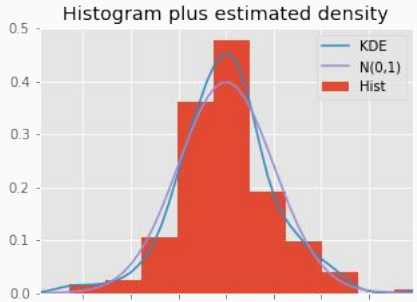


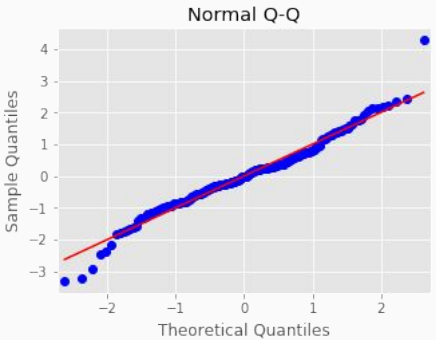
SARIMA RESULTS

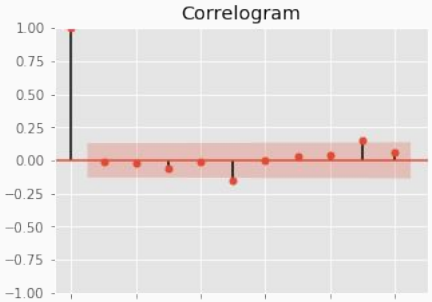
- We gathered the best SARIMA parameters from each zip code so that our model could harness as much of the data's signal as possible.











Results & Analysis

- Focus: ROI on 10-Year Investment
- Invest in:
  - Bloomfield, KY (40008) → 10.47 ROI
  - Osakis, MN (56360) → 5.75 ROI
  - Pierson, MI (49339) → 1.82 ROI

	Zip_code	1-year	3-year	5-year	10-year
2	40008	0.278668	1.072976	2.360854	10.477094
1	56360	0.208379	0.764886	1.577687	5.751005
3	49339	0.153610	0.404825	0.710747	1.822667
4	27019	0.042173	0.131907	0.229376	0.516576
0	48894	-0.041020	-0.073356	-0.079900	-0.081814

Future Plans?

- More Details:
  - Population
  - Schools
  - Prevalence of Parks
- Unexpected Events?
  - How to forecast taking emergencies into account?

Thank You!