Time Series Final Report

Forecasting Zillow Price

Group Members

Vivian Chu
Ding Ding
Max Calehuff
Dillon Quan

Problem Description

Data Set Description

The modified Zillow dataset recorded the following metrics on California housing:

1. Sale price (median) from January 2004 - December 2015 (monthly)
2. Mortgage rate (median) from January 2004 - August 2017 (monthly)
3. Unemployment rate from January 2004 - August 2017 (monthly)
4. Rental price (median) from January 2010 - August 2017 (monthly)

Column Name	Non-null Values	Data Type
Date	164 non-null	object
MedianSoldPrice_AllHomes.California	144 non-null	float64
MedianMortageRate	164 non-null	float64
UnemploymentRate	164 non-null	float64
MedianRentalPrice_AllHomes.California	92 non-null	float64

Total Number of Columns: 5
Total Number of Rows: 164

First five rows:

	Date	MedianSoldPrice_AllHomes.California	MedianMortageRate	UnemploymentRate	MedianRentalPrice_AllHomes.California
0	2004-01-31	326752.55	5.02	7.9	NaN
1	2004-02-29	329501.50	4.94	7.8	NaN
2	2004-03-31	340125.45	4.74	7.8	NaN
3	2004-04-30	355329.50	5.16	7.5	NaN
4	2004-05-31	367818.15	5.64	7.3	NaN

Forecasting Goals

We are interested in building a model that can accurately predict the median sale and rental prices of housing in California between January 2016, and August 2017. If we can build a model that can accurately predict prices in this time period, we could potentially deploy the model on a broader scale to aid purchasers in choosing the best time to sign a mortgage or a lease.

Methodology

Data Cleaning

1. We first changed the Date to datetime objects, sorted the data set by Date in descending order from least to most recent, and set the Date as the index.
2. We then tructuated the names of the columns associated with price:
    a. MedianSoldPrice_AllHomes.California changed to MedianPrice
    b. MedianRentalPrice_AllHomes.California changed to MedianRentalPrice
3. Next, we performed some minor exploration, and discovered a significant amount of NaN values in the MedianPrice (12.2%) and MedianRentalPrice (43.9%) columns. We replaced these missing fields with the median value of their respective columns.

Models

1. SARIMA (seasonal ARIMA) on Median Price
2. Exponential Smoothing on MedianPrice
    a. ETS(add/add)
    b. ETS(add/mul)
3. SARIMAX (seasonal ARIMA with exogenous variables) on Mortgage Rate
4. VAR (vector autoregression)
    a. VAR on Median Mortgage Rate
    b. VAR on Unemployed Rate Model
    c. VAR on Median Rental Price Model
    d. VAR on Unemployment rate and Median Rental Price
    e. VAR on All Features

Results

SARIMA Results Summary

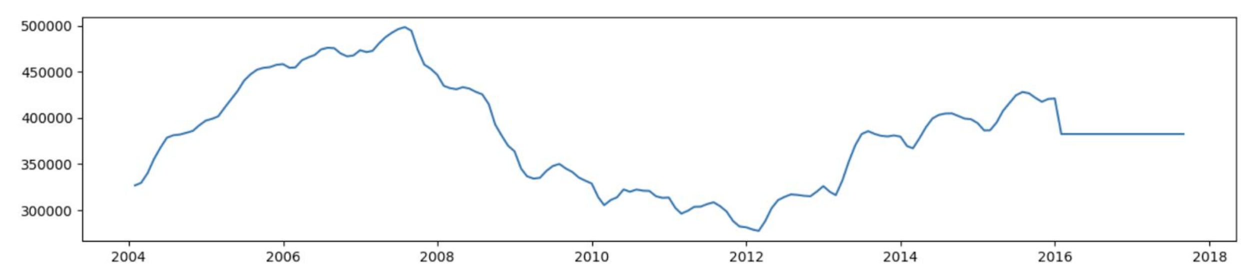
We started with a SARIMA model because the graph of the MedianPrice clearly showed both trend and seasonal components. The MedianPrice steadily increased, decreased, and increased again over a period of several years. There is also a notable small decline in the sale price of homes near the end of each year, which rises again after the winter months.

For our first run of SARIMA on the MedianPrice, we determined that d=2 for detrending and D=1 for removing the seasonality yielded the best ADF test results and the lowest RMSE. The optimal seasonality lag was m=18. Using auto_arima with AIC and BIC as our metrics, we determined a SARIMAX(1, 2, 1)x(1, 1, 1, 18) model to be the best fit. The AIC and BIC scores were AIC=2339.16, BIC=2335.41. The RMSE for this model was 34356.35.

Steps

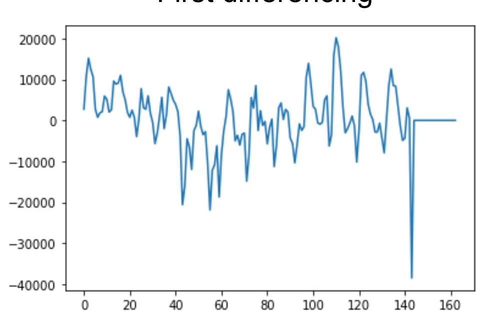
1. First, we observed both trend and seasonality in MedianPrice.

General Trend of MedianPrice

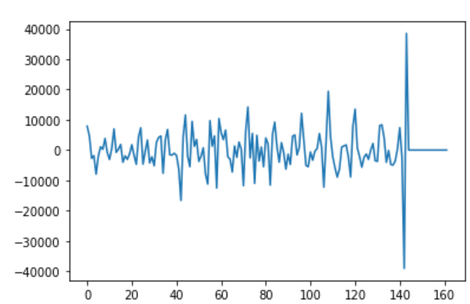


2. We differenced twice to remove the trend component.

First differencing



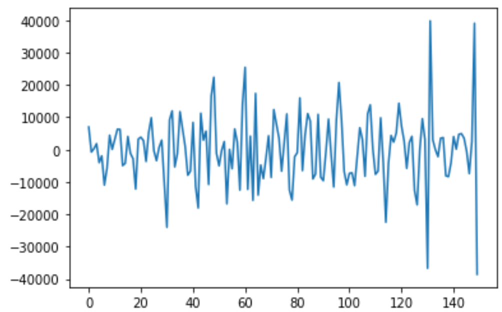
Second differencing



Next obtained p-value = 0.000012 from Dickey-Fuller on the second differencing, which is less than 0.05 and therefore acceptable.

3. We differenced once to remove the seasonal component. Assumed lag = 18.

First Differencing



4. We split the data into test and training splits based on the 80-20 rules, and fit auto_arima with d = 2 with seasonality m = 18.

   Model Parameters

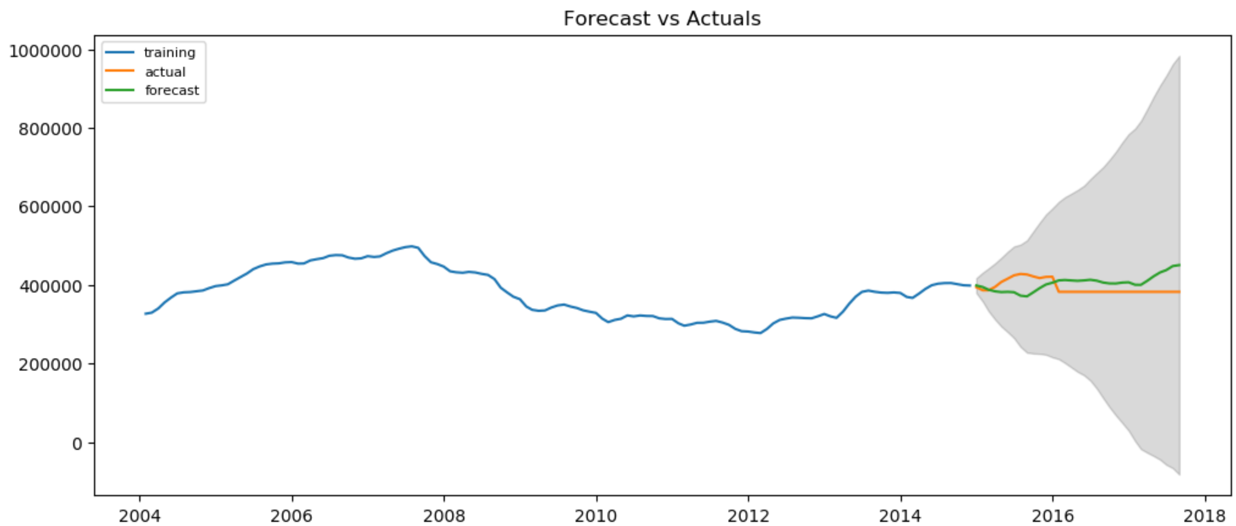
   model = pm.auto_arima(train, start_p=1, start_q=1,
   max_p=3, max_q=3, m=18,
   start_P=0, seasonal=True,
   d=2, D=1, trace=True,
   error_action='ignore',
   suppress_warnings=True,
   stepwise=True)

   Summary Table

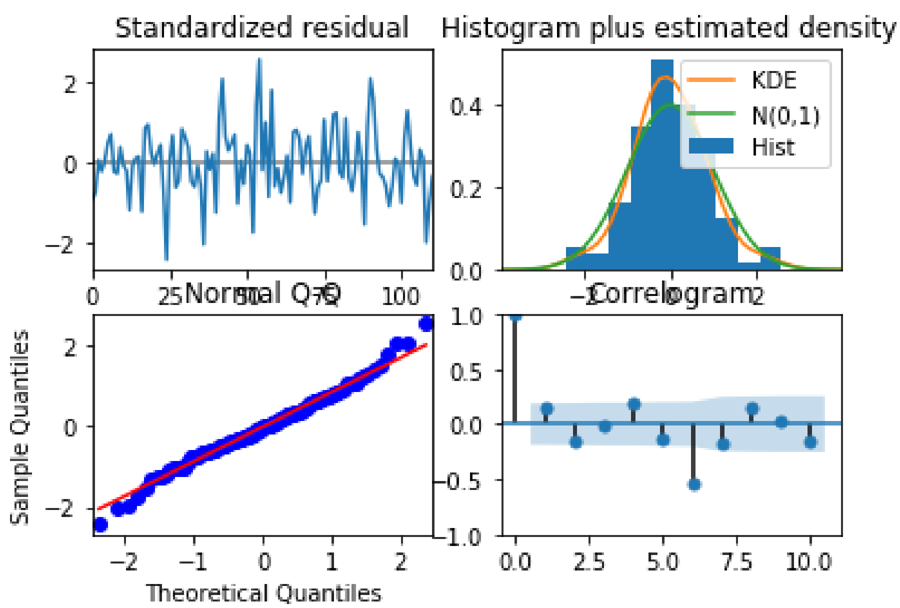
```
                                     Statespace Model Results
==========================================================================================
Dep. Variable:                                  y   No. Observations:                  131
Model:             SARIMAX(1, 2, 1)x(0, 2, 1, 18)   Log Likelihood                -999.179
Date:                            Wed, 11 Dec 2019   AIC                           2008.358
Time:                                    11:08:28   BIC                           2021.021
Sample:                                         0   HQIC                          2013.471
                                            - 131
Covariance Type:                              opg
==============================================================================
                 coef    std err          z      P>|z|      [0.025      0.975]
------------------------------------------------------------------------------
intercept    102.1136    765.480      0.133      0.894   -1398.200    1602.428
ar.L1         -0.9999      0.248     -4.038      0.000      -1.485      -0.515
ma.L1          0.9986      0.212      4.701      0.000       0.582       1.415
ma.S.L18      -0.9943      0.222     -4.480      0.000      -1.429      -0.559
sigma2      1.049e+08      0.000   3.66e+11      0.000    1.05e+08    1.05e+08
===================================================================================
Ljung-Box (Q):                      295.60   Jarque-Bera (JB):                 1.02
Prob(Q):                              0.00   Prob(JB):                         0.60
Heteroskedasticity (H):               0.89   Skew:                             0.14
Prob(H) (two-sided):                  0.75   Kurtosis:                         2.56
===================================================================================
```

Based on AIC and BIC, the best model is SARIMAX(1, 2, 1)x(0, 2, 1, 18). The AIC and BIC scores were AIC=2339.16, BIC=2335.41. The RMSE for this model was 34356.35.

5. We then fit this model with the training set and forecast using the testing set.



6. We then checked the residual diagnostics.



The diagnostics demonstrate White Noise behaviors: zero-mean, constant variance and non-correlation. Normality is also demonstrated.

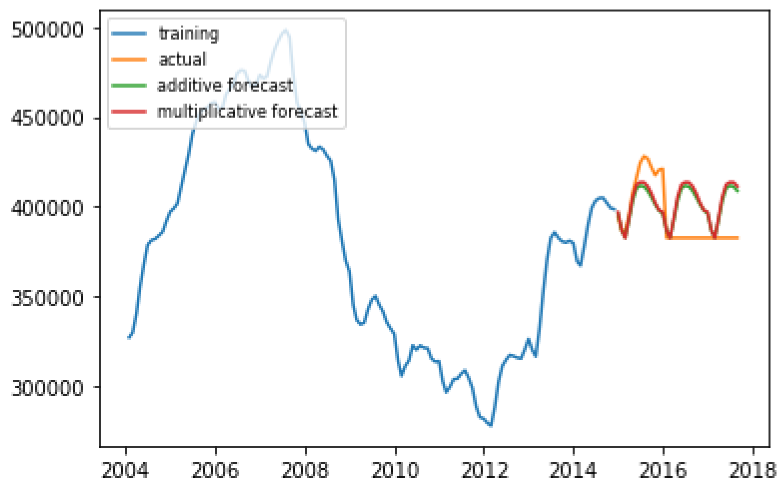
Exponential Smoothing Results Summary

Using exponential smoothing, the best model chosen from the RMSE evaluation is additive for trend and additive for seasonality with a RMSE of 18358.22. The RMSE for additive trend and multiplicative seasonality was 19220.75.

Steps

1. We fit two models, ETS(add/add), ETS(add/mul) with seasonal_periods = 12.

2. We then drew the forecasting plot.



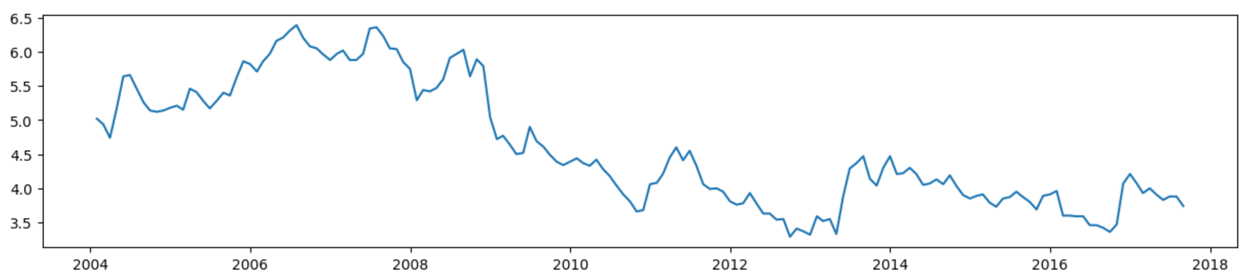
SARIMAX(MortgageRate) Results Summary

For this model, we introduced MortgageRate as an additional feature to predict MedianPrice. We first started with a SARIMA model because the graph of the MortgageRate clearly showed both trend and seasonal components. MortgageRate steadily increased, decreased, and increased again over a period of several years. There is also a notable small decline in MortgageRate near the end of each year, which rises again after the winter months.

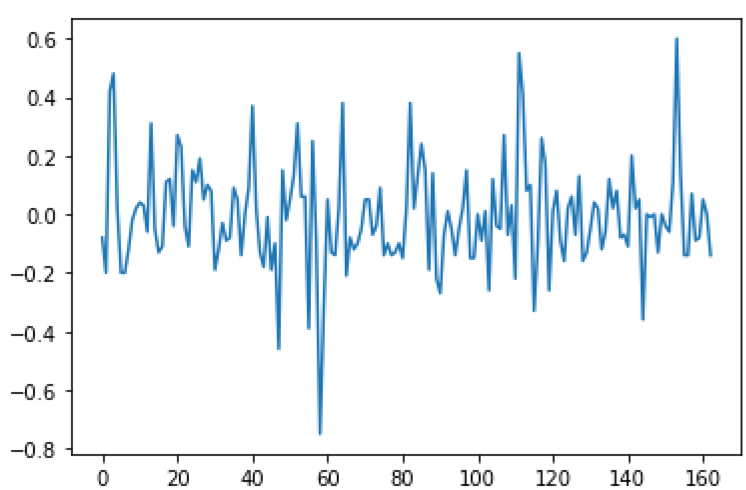
We determined that d=1 for trend yielded the best ADF test results and the lowest RMSE. No seasonality lag was required.

Using auto_arima with AIC and BIC as our metrics, we determined a SARIMAX(1, 1, 1) model to be the best fit. The AIC and BIC scores were AIC=2655.46, BIC=2666.93. The RMSE for this model was 19150.73.

1. First, we observed both trend and seasonality in MortgageRate.



2. We differenced once to eliminate the trend component.



Dickey-Fuller Test p-value = 1.268969e-08.

3. No seasonal differencing was necessary according to our ADF test.

4. We split the data into test and training splits based on the 80-20 rules, and fit auto_arima with d = 1 and no seasonality.

   Parameters

   ```
   model = pm.auto_arima(train, exogenous=exog_train, start_p=1, start_q=1,
                         max_p=3, max_q=3,
                         start_P=0, seasonal=False,
                         d=1,  trace=True,
                         error_action='ignore',
                         suppress_warnings=True,
                         stepwise=True)
   ```

   Summary

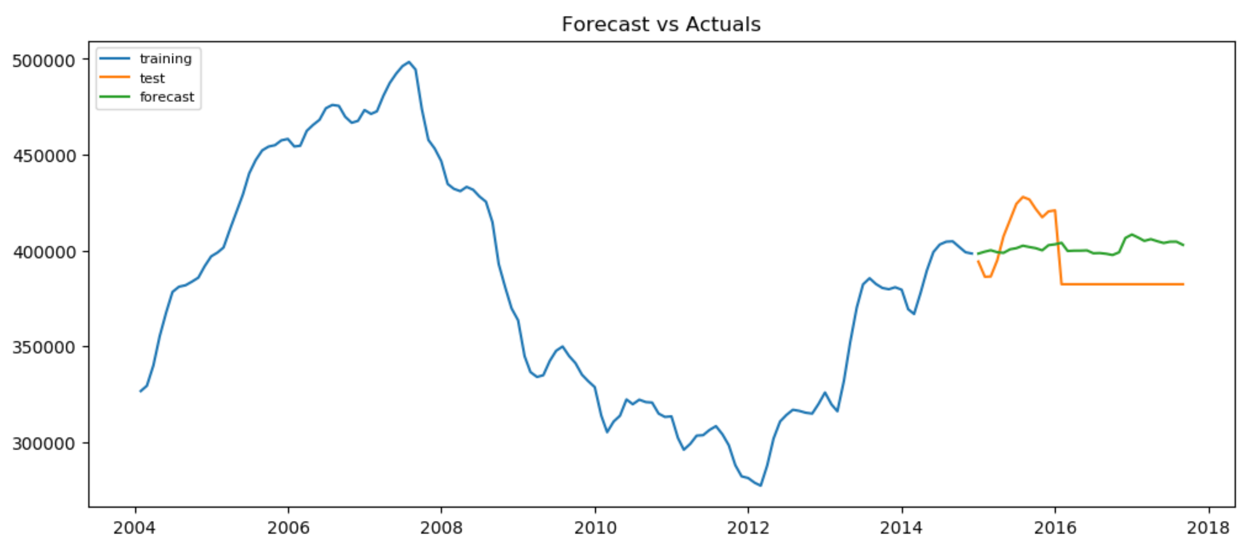
```
                             ARIMA Model Results
==============================================================================
Dep. Variable:                    D.y   No. Observations:                  130
Model:                 ARIMA(1, 1, 1)   Log Likelihood               -1296.457
Method:                       css-mle   S.D. of innovations           5171.713
Date:                Wed, 11 Dec 2019   AIC                           2602.915
Time:                        11:16:20   BIC                           2617.253
Sample:                             1   HQIC                          2608.741

==============================================================================
                 coef    std err          z      P>|z|      [0.025      0.975]
------------------------------------------------------------------------------
const       4892.7793   6276.945      0.779      0.437   -7409.807    1.72e+04
x1          -885.6816   1280.950     -0.691      0.491   -3396.298    1624.935
ar.L1.D.y      0.5003      0.102      4.903      0.000       0.300       0.700
ma.L1.D.y      0.3500      0.109      3.221      0.002       0.137       0.563
                                    Roots
=============================================================================
                  Real          Imaginary           Modulus         Frequency
-----------------------------------------------------------------------------
AR.1            1.9988           +0.0000j            1.9988            0.0000
MA.1           -2.8569           +0.0000j            2.8569            0.5000
-----------------------------------------------------------------------------
```

5. We then fit a SARIMAX(1,1,1) model with train and exog_train.

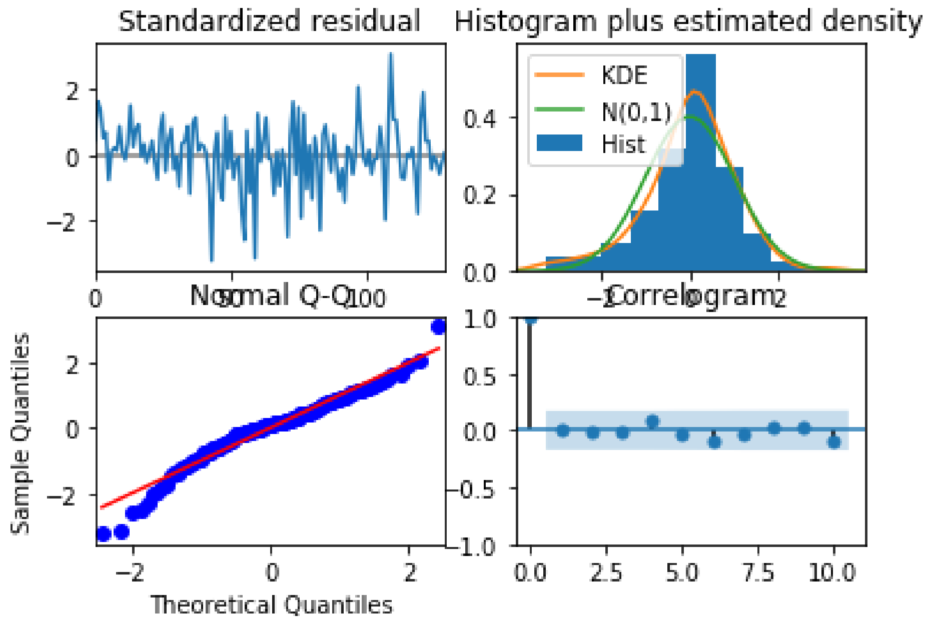
Summary

```
                               Statespace Model Results
==============================================================================
Dep. Variable:            MedianPrice   No. Observations:                  131
Model:               SARIMAX(1, 1, 1)   Log Likelihood               -1323.732
Date:                Wed, 11 Dec 2019   AIC                           2655.464
Time:                        11:16:22   BIC                           2666.934
Sample:                    01-31-2004   HQIC                          2660.125
                         - 11-30-2014
Covariance Type:                  opg
==============================================================================
                 coef    std err          z      P>|z|      [0.025      0.975]
------------------------------------------------------------------------------
x1          1.232e+04   2733.706      4.506      0.000    6960.419    1.77e+04
ar.L1          0.9151      0.058     15.874      0.000       0.802       1.028
ma.L1         -0.8762      0.067    -13.137      0.000      -1.007      -0.745
sigma2      4.088e+07      0.053   7.71e+08      0.000    4.09e+07    4.09e+07
===================================================================================
Ljung-Box (Q):                      252.36   Jarque-Bera (JB):                 9.17
Prob(Q):                              0.00   Prob(JB):                         0.01
Heteroskedasticity (H):               3.34   Skew:                             0.14
Prob(H) (two-sided):                  0.00   Kurtosis:                         4.27
===================================================================================
```

6. We then fit this model with the training set and forecast using the testing set.



7. We then checked the residual diagnostics.



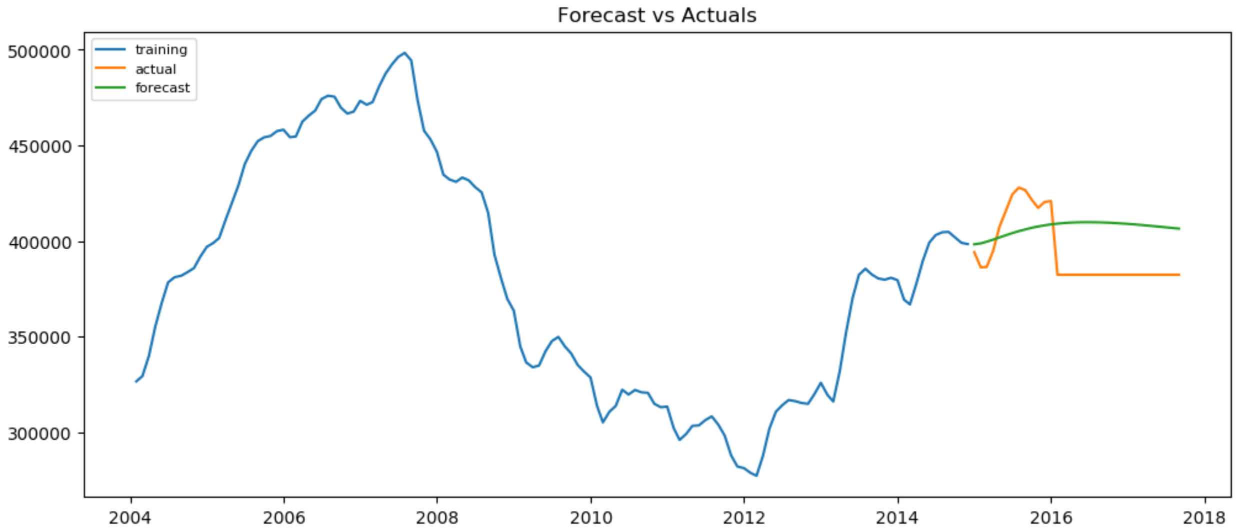
The diagnostics demonstrate White Noise behaviors: zero-mean, constant variance and non-correlation. Normality is also demonstrated.

VAR Results Summary

1. VAR (Mortgage Rate)

Using the feature median mortgage rate as an endogenous feature to predict median price, the RMSE was computed to be 22329.18 with p=2.

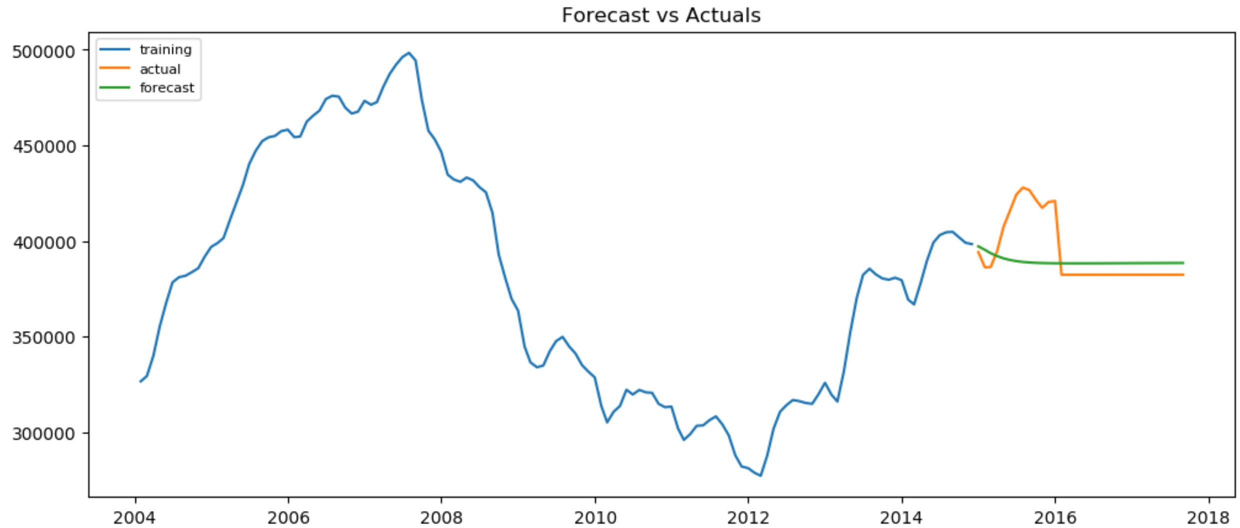
```
  Summary of Regression Results
==================================
Model:                         VAR
Method:                        OLS
Date:           Wed, 11, Dec, 2019
Time:                     11:18:51
--------------------------------------------------------------------
No. of Equations:         2.00000    BIC:                    13.9991
Nobs:                     129.000    HQIC:                   13.8675
Log likelihood:          -1244.73    FPE:                    962718.
AIC:                      13.7774    Det(Omega_mle):         892214.
--------------------------------------------------------------------
Results for equation MedianMortageRate
=========================================================================================
                          coefficient       std. error           t-stat            prob
-----------------------------------------------------------------------------------------
const                       -0.021004         0.098269           -0.214           0.831
L1.MedianMortageRate         1.041507         0.091833           11.341           0.000
L1.MedianPrice               0.000005         0.000002            2.084           0.037
L2.MedianMortageRate        -0.126380         0.091162           -1.386           0.166
L2.MedianPrice              -0.000004         0.000002           -1.548           0.122
=========================================================================================

Results for equation MedianPrice
=========================================================================================
                          coefficient       std. error           t-stat            prob
-----------------------------------------------------------------------------------------
const                     5046.023219      2941.929625            1.715           0.086
L1.MedianMortageRate       780.099785      2749.256266            0.284           0.777
L1.MedianPrice               1.666141         0.068942           24.167           0.000
L2.MedianMortageRate     -1281.873974      2729.176325           -0.470           0.639
L2.MedianPrice              -0.672617         0.071364           -9.425           0.000
=========================================================================================

Correlation matrix of residuals
                     MedianMortageRate  MedianPrice
MedianMortageRate             1.000000     0.217809
MedianPrice                   0.217809     1.000000
```


Forecast vs Actuals
training
actual
forecast
500000
450000
400000
350000
300000
2004
2006
2008
2010
2012
2014
2016
2018

2. VAR (Unemployed Rate)

Using the feature of unemployment rate as an endogenous feature to predict median price, the RMSE was computed to be 17390.46 with p=2.

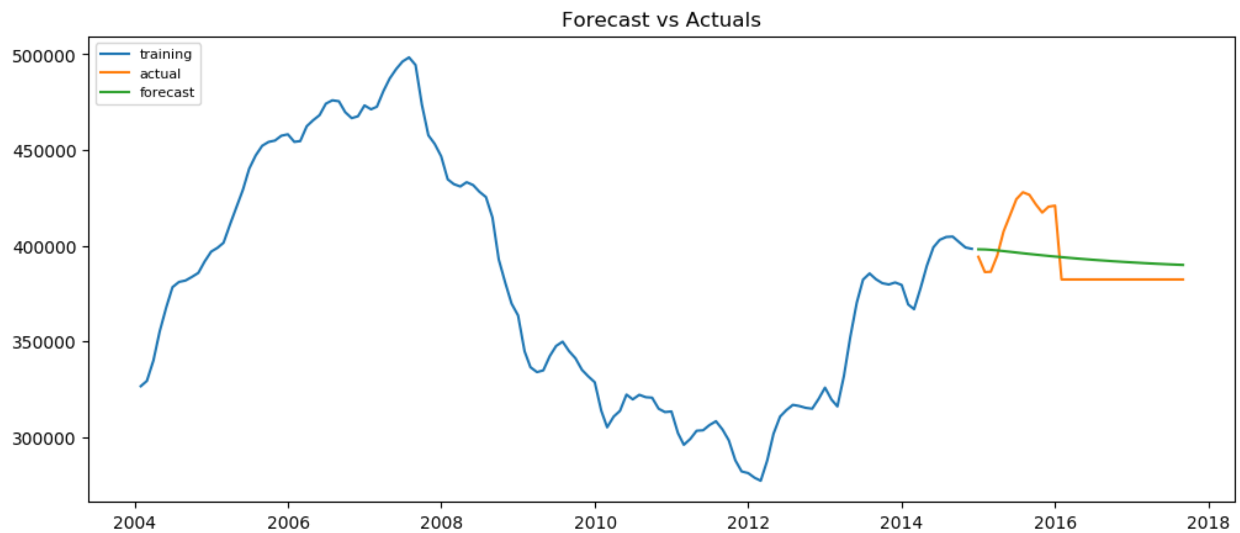
```
  Summary of Regression Results   
==================================
Model:                         VAR
Method:                        OLS
Date:           Wed, 11, Dec, 2019
Time:                     11:21:42
--------------------------------------------------------------------
No. of Equations:         2.00000    BIC:                    16.1327
Nobs:                     128.000    HQIC:                   15.9475
Log likelihood:          -1361.78    FPE:                7.42943e+06
AIC:                      15.8207    Det(Omega_mle):     6.67895e+06
--------------------------------------------------------------------
Results for equation UnemploymentRate
=======================================================================================
                         coefficient       std. error           t-stat            prob
---------------------------------------------------------------------------------------
const                       3.967600         0.989520            4.010           0.000
L1.UnemploymentRate         0.640480         0.090069            7.111           0.000
L1.MedianPrice              0.000005         0.000008            0.643           0.520
L2.UnemploymentRate        -0.121786         0.106489           -1.144           0.253
L2.MedianPrice             -0.000021         0.000015           -1.349           0.177
L3.UnemploymentRate         0.107360         0.091033            1.179           0.238
L3.MedianPrice              0.000012         0.000008            1.375           0.169
=======================================================================================

Results for equation MedianPrice
=======================================================================================
                         coefficient       std. error           t-stat            prob
---------------------------------------------------------------------------------------
const                  -14182.148118     10360.955059           -1.369           0.171
L1.UnemploymentRate       709.035749       943.083412            0.752           0.452
L1.MedianPrice              1.802578         0.088135           20.452           0.000
L2.UnemploymentRate       804.626088      1115.009805            0.722           0.471
L2.MedianPrice             -1.019221         0.159543           -6.388           0.000
L3.UnemploymentRate       173.487860       953.174878            0.182           0.856
L3.MedianPrice              0.223085         0.088287            2.527           0.012
=======================================================================================

Correlation matrix of residuals
                    UnemploymentRate  MedianPrice
UnemploymentRate            1.000000    -0.103813
MedianPrice                -0.103813     1.000000
```


Forecast vs Actuals
training
actual
forecast
500000
450000
400000
350000
300000
2004
2006
2008
2010
2012
2014
2016
2018

3. VAR (Median Rental Price)

Using the feature of rate as an endogenous feature to predict median price, the RMSE was computed to be 15423.91 with p=2.

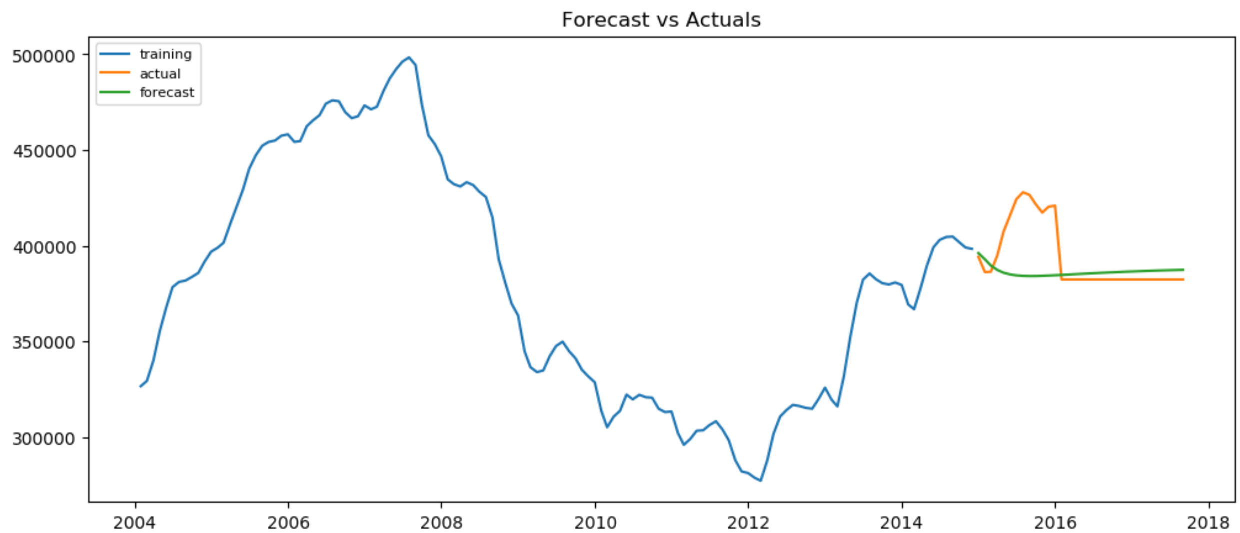
```
  Summary of Regression Results
==================================
Model:                         VAR
Method:                        OLS
Date:           Wed, 11, Dec, 2019
Time:                     11:22:05
--------------------------------------------------------------------
No. of Equations:         2.00000    BIC:                    25.5800
Nobs:                     128.000    HQIC:                   25.3948
Log likelihood:          -1966.40    FPE:                9.41577e+10
AIC:                      25.2680    Det(Omega_mle):     8.46464e+10
--------------------------------------------------------------------
Results for equation MedianRentalPrice
=========================================================================================
                          coefficient       std. error           t-stat              prob
-----------------------------------------------------------------------------------------
const                       64.913265        63.664866            1.020             0.308
L1.MedianRentalPrice         0.810301         0.090550            8.949             0.000
L1.MedianPrice               0.000344         0.000933            0.368             0.713
L2.MedianRentalPrice         0.240874         0.114621            2.101             0.036
L2.MedianPrice              -0.000465         0.001702           -0.273             0.785
L3.MedianRentalPrice        -0.100882         0.090188           -1.119             0.263
L3.MedianPrice               0.000219         0.000934            0.234             0.815
=========================================================================================

Results for equation MedianPrice
=========================================================================================
                          coefficient       std. error           t-stat              prob
-----------------------------------------------------------------------------------------
const                    10549.994884      5970.335881            1.767             0.077
L1.MedianRentalPrice         0.930753         8.491581            0.110             0.913
L1.MedianPrice               1.811016         0.087477           20.703             0.000
L2.MedianRentalPrice         6.702343        10.748906            0.624             0.533
L2.MedianPrice              -1.034985         0.159586           -6.485             0.000
L3.MedianRentalPrice       -11.081181         8.457627           -1.310             0.190
L3.MedianPrice               0.215195         0.087573            2.457             0.014
=========================================================================================

Correlation matrix of residuals
                     MedianRentalPrice  MedianPrice
MedianRentalPrice             1.000000    -0.070086
MedianPrice                  -0.070086     1.000000
```


Forecast vs Actuals
training
actual
forecast
500000
450000
400000
350000
300000
2004
2006
2008
2010
2012
2014
2016
2018

4. VAR (Unemployment rate and Median Rental Price)

Using the feature of rate as an endogenous feature to predict median price, the RMSE was computed to be 19263.65 with p=3.

```
  Summary of Regression Results
==================================
Model:                         VAR
Method:                        OLS
Date:           Tue, 10, Dec, 2019
Time:                     23:28:08
--------------------------------------------------------------------
No. of Equations:         3.00000    BIC:                    24.5524
Nobs:                     128.000    HQIC:                   24.1555
Log likelihood:          -2043.45    FPE:                2.36094e+10
AIC:                      23.8840    Det(Omega_mle):     1.88399e+10
--------------------------------------------------------------------
Results for equation UnemploymentRate
=======================================================================================
                         coefficient       std. error           t-stat            prob
---------------------------------------------------------------------------------------
const                       4.308374         1.016944            4.237           0.000
L1.UnemploymentRate         0.605508         0.091528            6.616           0.000
L1.MedianRentalPrice        0.001148         0.000837            1.372           0.170
L1.MedianPrice              0.000008         0.000009            0.970           0.332
L2.UnemploymentRate        -0.140065         0.106573           -1.314           0.189
L2.MedianRentalPrice        0.000160         0.001022            0.157           0.876
L2.MedianPrice             -0.000022         0.000015           -1.455           0.146
L3.UnemploymentRate         0.048865         0.095817            0.510           0.610
L3.MedianRentalPrice       -0.000849         0.000811           -1.047           0.295
L3.MedianPrice              0.000009         0.000009            1.048           0.295
=======================================================================================


Results for equation MedianRentalPrice
=======================================================================================
                         coefficient       std. error           t-stat            prob
---------------------------------------------------------------------------------------
const                    -189.087773       112.153249           -1.686           0.092
L1.UnemploymentRate        16.376584        10.094150            1.622           0.105
L1.MedianRentalPrice        0.743884         0.092306            8.059           0.000
L1.MedianPrice             -0.000052         0.000939           -0.056           0.956
L2.UnemploymentRate         7.478008        11.753393            0.636           0.525
L2.MedianRentalPrice        0.231547         0.112720            2.054           0.040
L2.MedianPrice             -0.000292         0.001676           -0.174           0.862
L3.UnemploymentRate         7.937336        10.567095            0.751           0.453
L3.MedianRentalPrice       -0.068504         0.089398           -0.766           0.444
L3.MedianPrice              0.000757         0.000946            0.800           0.424
=======================================================================================

Results for equation MedianPrice
=======================================================================================
                         coefficient       std. error           t-stat            prob
---------------------------------------------------------------------------------------
const                 -12486.437952     10543.955185           -1.184           0.236
L1.UnemploymentRate     1082.653859       948.989560            1.141           0.254
L1.MedianRentalPrice      -5.342086         8.678006           -0.616           0.538
L1.MedianPrice             1.769153         0.088249           20.047           0.000
L2.UnemploymentRate     1020.185341      1104.981308            0.923           0.356
L2.MedianRentalPrice       5.973705        10.597283            0.564           0.573
L2.MedianPrice            -1.013191         0.157607           -6.429           0.000
L3.UnemploymentRate      804.224998       993.452935            0.810           0.418
L3.MedianRentalPrice      -8.117699         8.404664           -0.966           0.334
L3.MedianPrice             0.263939         0.088949            2.967           0.003
=======================================================================================

Correlation matrix of residuals
                     UnemploymentRate  MedianRentalPrice  MedianPrice
UnemploymentRate             1.000000           0.066751    -0.084340
MedianRentalPrice            0.066751           1.000000    -0.137841
MedianPrice                 -0.084340          -0.137841     1.000000
```

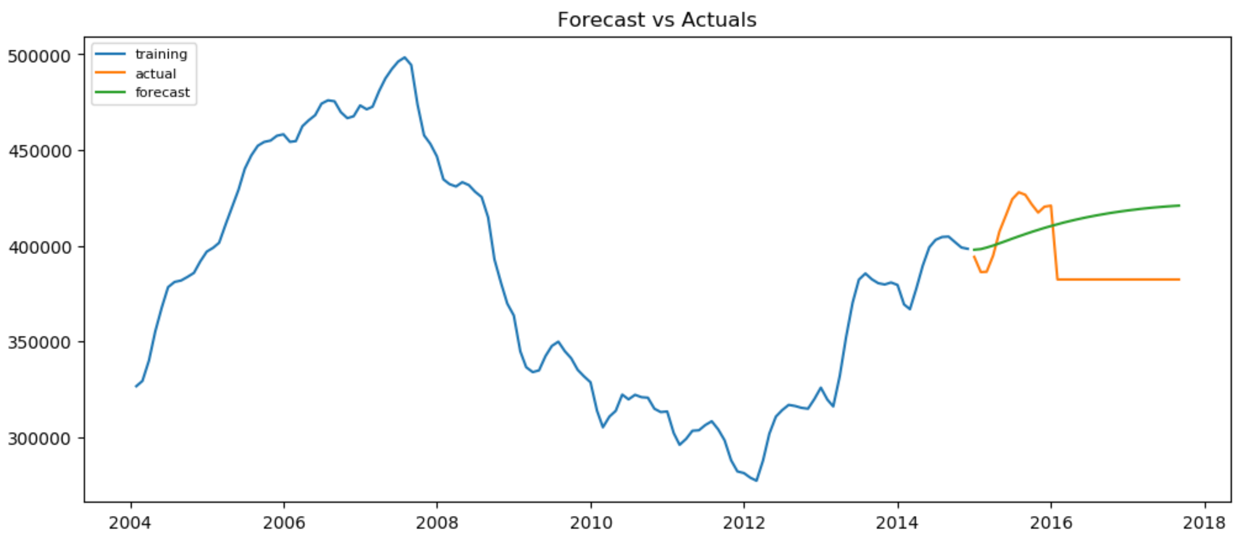


Forecast vs Actuals
training
actual
forecast
500000
450000
400000
350000
300000
2004
2006
2008
2010
2012
2014
2016
2018

5. VAR (All Features)

Using the feature of rate as an endogenous feature to predict median price, the RMSE was computed to be 28470.47 with p=4.

```
  Summary of Regression Results
==================================
Model:                         VAR
Method:                        OLS
Date:           Tue, 10, Dec, 2019
Time:                     23:28:08
--------------------------------------------------------------------
No. of Equations:         4.00000    BIC:                    21.3454
Nobs:                     130.000    HQIC:                   21.0835
Log likelihood:          -2076.63    FPE:                1.19860e+09
AIC:                      20.9043    Det(Omega_mle):     1.03065e+09
--------------------------------------------------------------------
Results for equation MedianMortageRate
=======================================================================================
                          coefficient       std. error           t-stat            prob
---------------------------------------------------------------------------------------
const                       -0.267206         0.309395           -0.864           0.388
L1.MedianMortageRate         0.901033         0.037537           24.004           0.000
L1.MedianPrice               0.000002         0.000001            2.777           0.005
L1.MedianRentalPrice         0.000009         0.000106            0.089           0.929
L1.UnemploymentRate          0.020337         0.026989            0.754           0.451
=======================================================================================

Results for equation MedianPrice
=======================================================================================
                          coefficient       std. error           t-stat            prob
---------------------------------------------------------------------------------------
const                   -16547.423050     11576.353908           -1.429           0.153
L1.MedianMortageRate     -1793.048907      1404.497039           -1.277           0.202
L1.MedianPrice               1.050804         0.020353           51.630           0.000
L1.MedianRentalPrice        -9.976747         3.971683           -2.512           0.012
L1.UnemploymentRate       3824.911829      1009.807022            3.788           0.000
=======================================================================================

Results for equation MedianRentalPrice
=======================================================================================
                          coefficient       std. error           t-stat            prob
---------------------------------------------------------------------------------------
const                      -78.873573        92.370912           -0.854           0.393
L1.MedianMortageRate         3.992310        11.206868            0.356           0.722
L1.MedianPrice               0.000229         0.000162            1.408           0.159
L1.MedianRentalPrice         0.922195         0.031691           29.099           0.000
L1.UnemploymentRate         18.812985         8.057528            2.335           0.020
=======================================================================================

Results for equation UnemploymentRate
=======================================================================================
                          coefficient       std. error           t-stat            prob
---------------------------------------------------------------------------------------
const                        3.925163         0.828940            4.735           0.000
L1.MedianMortageRate         0.060170         0.100571            0.598           0.550
L1.MedianPrice              -0.000005         0.000001           -3.435           0.001
L1.MedianRentalPrice         0.000348         0.000284            1.224           0.221
L1.UnemploymentRate          0.568192         0.072309            7.858           0.000
=======================================================================================

Correlation matrix of residuals
                     MedianMortageRate  MedianPrice  MedianRentalPrice  UnemploymentRate
MedianMortageRate             1.000000     0.288388           0.000990         -0.021122
MedianPrice                   0.288388     1.000000          -0.101017          0.018170
MedianRentalPrice             0.000990    -0.101017           1.000000          0.053939
UnemploymentRate             -0.021122     0.018170           0.053939          1.000000
```


Forecast vs Actuals
training
actual
forecast
500000
450000
400000
350000
300000
2004
2006
2008
2010
2012
2014
2016
2018

Model Summary & Conclusion

Model	RMSE
SARIMA	34356.35
ETS (add/add)	18358.22
ETS (add/mul)	19220.75
SARIMAX	19150.73
VAR on Median Mortgage Rate	22329.19
VAR on Unemployed Rate Model	17390.69
VAR on Median Rental Price Model	**15423.91**
VAR on Unemployment rate and Median Rental Price	19263.65
VAR on All Features	28470.47

We determined VAR (Median Rental Price) was our best model, with an RMSE of 15423.91.

This is likely because VAR is optimized for multiple time series that influence each other.

Both Median Rental Price and Median Sale Price are subject to many of the same exogenous forces which determine the cost of housing, such as the supply of vacant units and the income of the people living in the area.

These variables also endogenously influence each other, since if rental prices get high enough that it becomes economical to purchase a home, the increased demand on home purchases will drive up home purchase prices. Similarly, increases in the price of housing can make it more economical to rent, which will drive up rent prices.

Forecasting Results

Prediction Table

Date	ARIMA	ETS(add/add)	ETS(add/mul)	SARIMAX(Mortgage Rate)	VAR(Mortgage Rate)	VAR(Unemployed Rate)	VAR(Median Rental Price)
2014-12-31	391436.63513036900	395992.81068571600	396820.8747176020	398440.0579562530	398363.04505754800	397275.2905564390	398208.8976857380
2015-01-31	374593.99521364	386499.05524032600	386710.7477230130	399464.4596137580	398852.5001731120	395485.11477457800	398130.6776710070
2015-02-28	350937.52674160200	382691.08160577900	382351.8712434190	400197.36791521000	399724.35369695800	393681.94685051600	397927.5960710930
2015-03-31	328657.4670212170	392555.71282353100	393646.18774268100	399164.40671930800	400799.62939571300	392171.59455740000	397640.5637221780
2015-04-30	308130.33594193500	403051.5864186010	405244.2962998200	398832.7580612030	401955.3386925180	391006.9307157570	397286.0857625200
2015-05-31	289764.4283758740	410019.9408788870	412204.66534809700	400683.8356716060	403110.7363177880	390154.1440781190	396907.7948823540
2015-06-30	268988.77944083700	411598.5583221230	413522.80626730400	401271.426968534	404214.55310386600	389553.8158090040	396524.50214921400
2015-07-31	238454.80783249100	411092.09087429400	413474.55377007200	402569.1597161480	405235.72299404400	389141.87735123100	396148.2095598790
2015-08-31	216818.98229275000	408684.8854163860	411256.2369553260	401869.44732161100	406156.9041562010	388863.7768135590	395782.4583637180
2015-09-30	210592.4287015930	405064.27273872200	407286.6699504690	401268.66477530400	406969.9886064580	388678.84390522900	395428.4142585650
2015-10-31	205387.64295960200	400996.0336460640	401982.1894663950	400152.9513985940	407672.9989858830	388558.1591907790	395085.8590294880
2015-11-30	197960.4083358520	397824.2566207290	398013.78882411300	402835.62586145300	408267.94921852400	388481.4802201470	394754.45941369000
2015-12-31	184405.41621656500	396029.78192192100	396916.80393467700	403282.40338762900	408759.37424153700	388434.83510487400	394433.89527001300
2016-01-31	168104.0698505950	386505.29787645600	386730.42505193100	404081.7225048290	409153.3238802750	388408.6612378670	394123.9619028320
2016-02-29	150950.62036698800	382692.1356821320	382355.9669804730	399814.9374693960	409456.6784800790	388396.41466819800	393824.5129509180
2016-03-31	130647.83031248500	392555.89080554600	393647.0754659730	399968.5260685170	409676.68737502100	388393.6076292890	393535.4331109770
2016-04-30	110600.09196832500	403051.61647107000	405244.4886949590	399985.8948449740	409820.6614812460	388397.15953490000	393256.61027221700

2016-05-31	93222.25691163660	410019.9459532820	412204.7065480120	400114.51794217400	409895.7722937390	388404.9545078420	392987.9254989390
2016-06-30	69033.96148455320	411598.5591789400	413522.81496870600	398630.83396445500	409908.9241517860	388415.53843960700	392729.24846872400
2016-07-31	34207.964374855000	411092.0910189690	413474.5556017330	398738.5496492650	409866.67677427800	388427.912151346	392480.4373900820
2016-08-31	-7929.420511092050	408684.88544081400	411256.2373388700	398344.38746094400	409775.20210905600	388441.3898236700	392241.3398952010
2016-09-30	-48288.139447656600	405064.27274284700	407286.670030436	397695.4910781530	409640.26443427000	388455.5016363240	392011.7943305690
2016-10-31	-86439.7318021373	400996.03364676000	401982.1894830110	399133.0637987730	409467.21604997700	388469.926680407	391791.630895521
2016-11-30	-121943.0286496130	397824.25662084700	398013.78882757700	406599.63872272300	409261.00326053100	388484.4467140240	391580.67264005700
2016-12-31	-159864.53896964800	396029.7819219410	396916.8039354050	408393.3446138830	409026.17898930600	388498.9142660700	391378.7363235110
2017-01-31	-208376.8461212520	386505.29787645900	386730.42505208100	406732.0350163710	408766.9195070860	388513.2306248040	391185.63318823100
2017-02-28	-246741.59190473500	382692.1356821330	382355.96698050400	405065.3558019830	408487.04354685300	388527.3306475860	391001.1696731690
2017-03-31	-267736.94524152500	392555.89080554600	393647.07546597900	405980.62360984300	408190.0326270140	388541.17227970600	390825.1480851830
2017-04-30	-287337.49368354800	403051.61647107000	405244.48869496000	404920.452991702	407879.05178616000	388554.729323301	390657.3672344130
2017-05-31	-309157.09130508800	410019.9459532820	412204.706548012	403979.35109806300	407556.97019659100	388567.9864491960	390497.62303680400
2017-06-30	-337530.9347557370	411598.5591789400	413522.8149687060	404635.8734258120	407226.38130667300	388580.9357567120	390345.7090847290
2017-07-31	-368946.6380812870	411092.0910189690	413474.55560173300	404673.0302724310	406889.6222884980	388593.5744016900	390201.4171863970
2017-08-31	-400715.2785146090	408684.88544081400	411256.2373388700	402982.4594728200	406548.792654411	388605.9029615880	390064.5378747060

Responsibility Table

Group Members	Vivian Chu	Ding Ding	Max Calehuff	Dillon Quan
Proportion of Work	25%	25%	25%	25%
List of Work	Initial check for the data(summary tables and graphs) Discussion of problems and chosen methods Data Cleaning and Modeling Forecasting and analyzing the forecasting results Discussion of model selection Results	Initial check for the data(summary tables and graphs) Discussion of problems and chosen methods Final Discussion Write the Report	Initial check for the data(summary tables and graphs) Discussion of problems and chosen methods Final Discussion Write the Report	Initial check for the data(summary tables and graphs) Discussion of problems and chosen methods Model Fitting Forecasting and analyzing the forecasting results Final Discussion Prediction Table