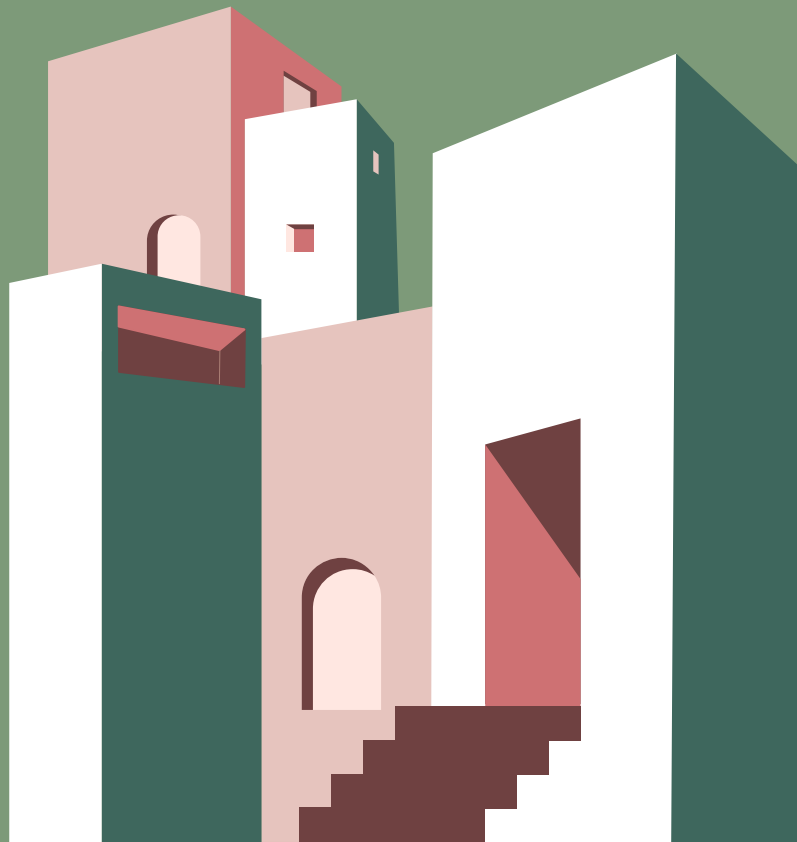
TIME SERIES ANALYSIS ON U.S. HOUSING PRICES

By: Meredith Lou



OBJECTIVES

01

UNDERSTAND THE DATA - HOUSING AND RECESSIONS

Source: St. Louis Federal Reserve
1980-2022, quarterly data
EDA: Plots, Decomposition, ACFs/PACFs and Smoothing

02

DETERMINE THE BEST FORECAST MODEL

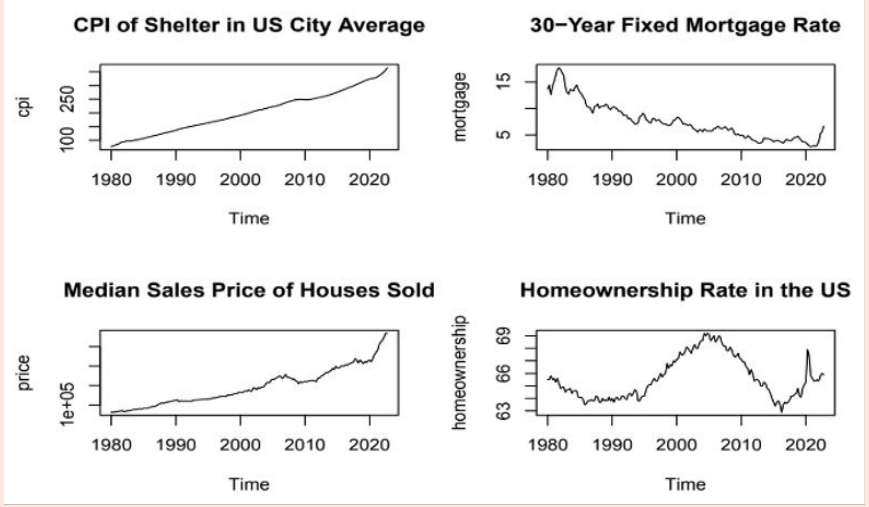
Considerations: Linear Regression, ARMA & ARIMA models

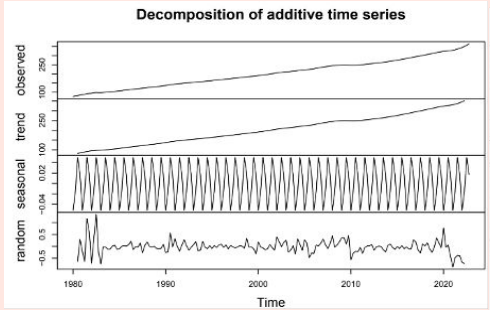
03

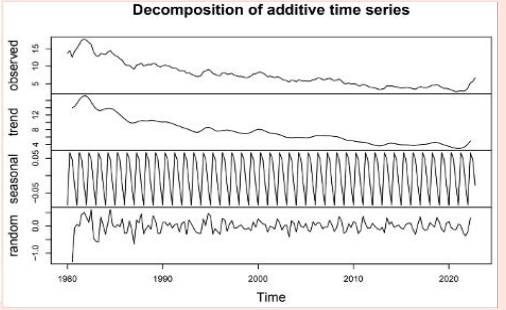
FORECAST MEDIAN HOUSING PRICES

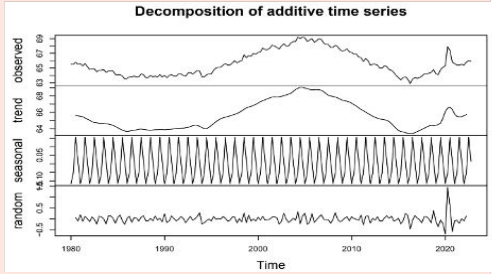
Use of an ARIMA model to forecast future housing prices

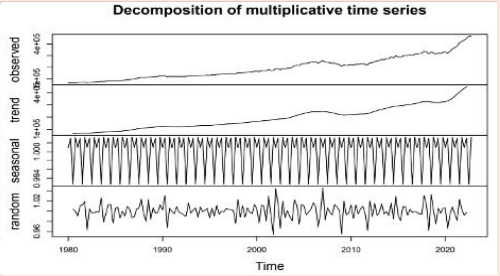
EXPLORATORY DATA ANALYSIS











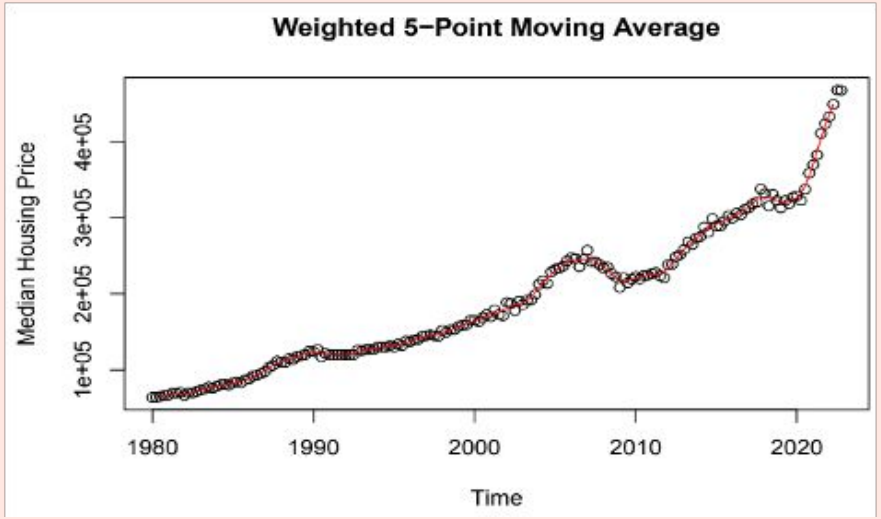
PLOTS OF RAW DATA

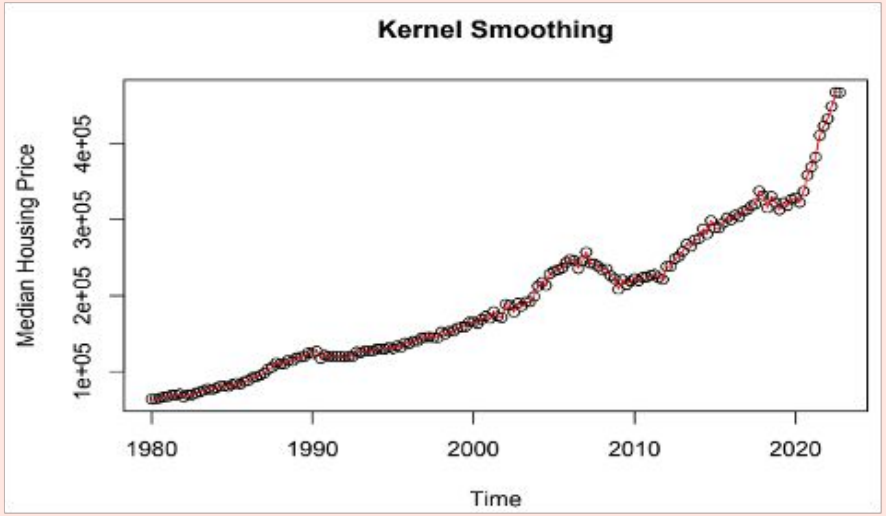
- Positive trend for CPI and Prices
- Negative trend for mortgage
- Fluctuations with homeownership rate
- No clear seasonality

DECOMPOSITION

- Confirms the trends and seasonality claims
- Little noise

EXPLORATORY DATA ANALYSIS



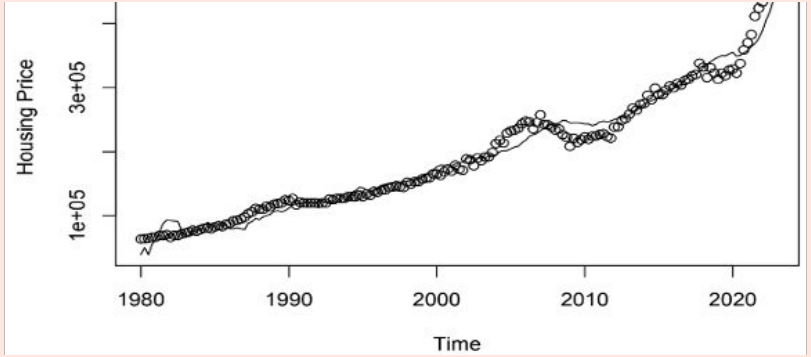


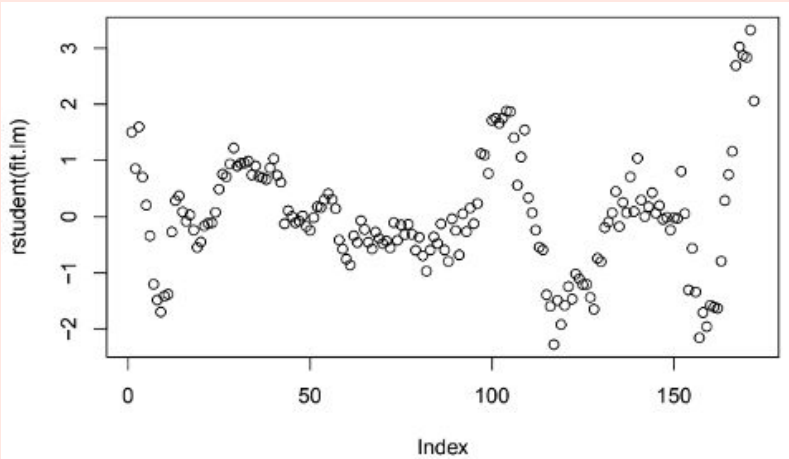
SMOOTHING

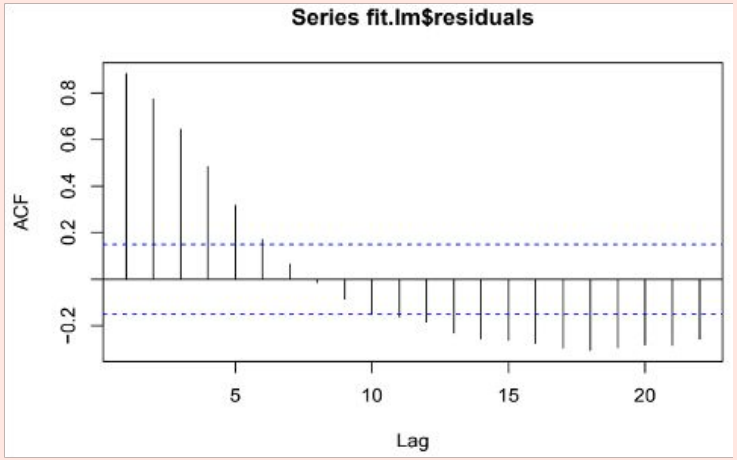
- Smoothed to study long-term trends and data structure
- 5-point moving average (left) and kernel smoothing (right)
  - Also tried smoothing splines and loess

LINEAR REGRESSION

- Price regressed on CPI, mortgage rates, and homeownership rates
- Good plot, but patterned residuals and a lot of autocorrelation

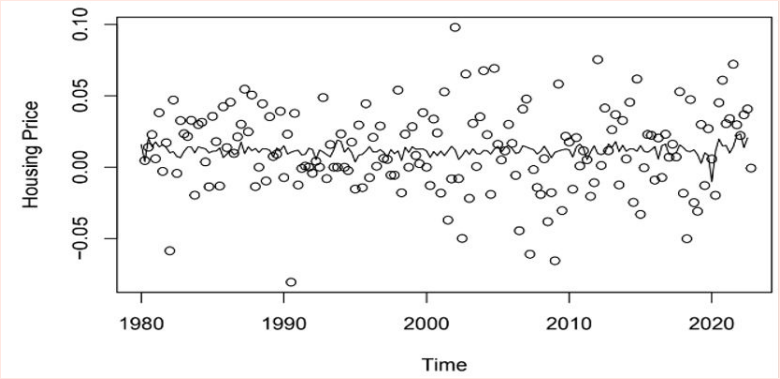


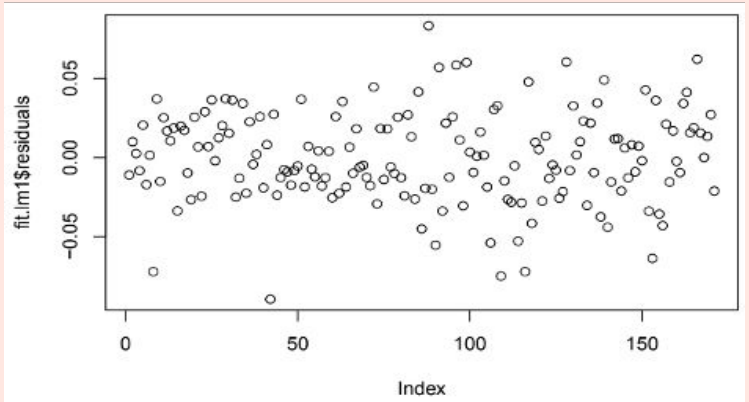


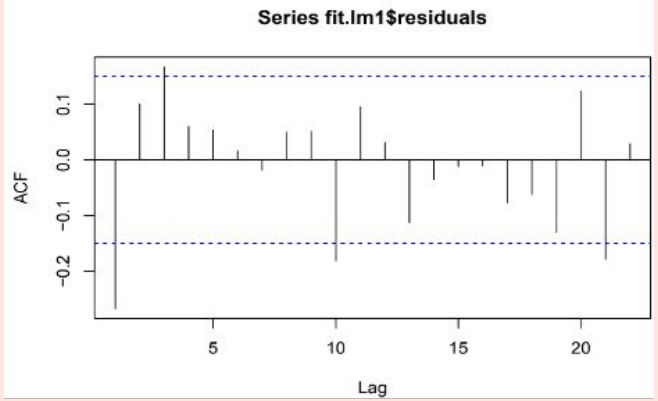


LINEAR MODEL OF THE DIFFERENCED LOGS

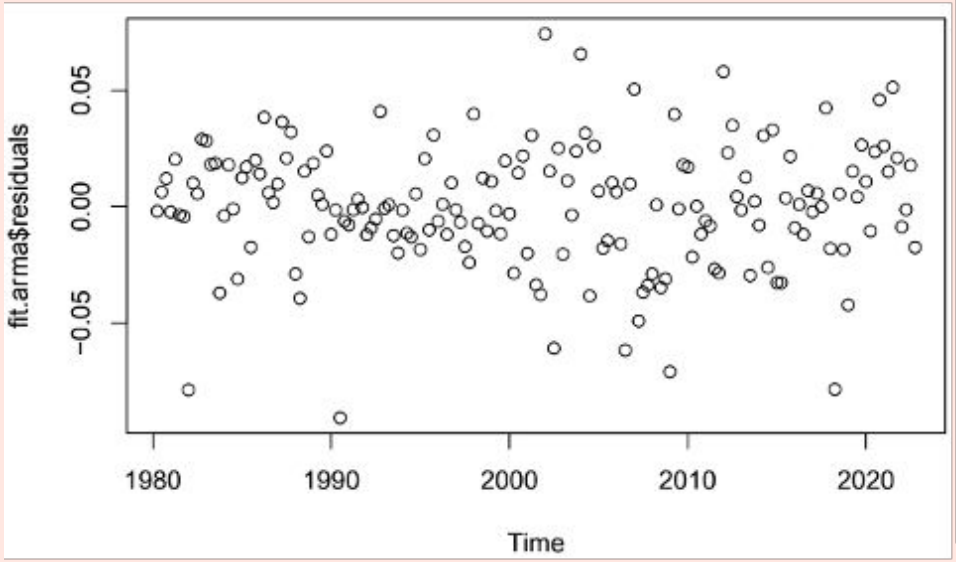
- Price regressed on CPI, mortgage rates, and homeownership rates
- Poor plot but good residual distribution and ACF

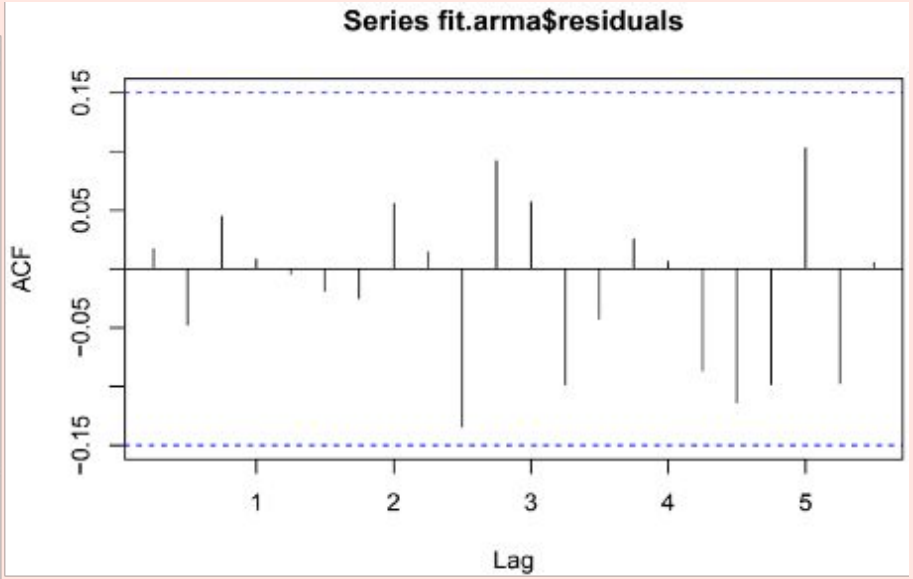






ARMA MODEL

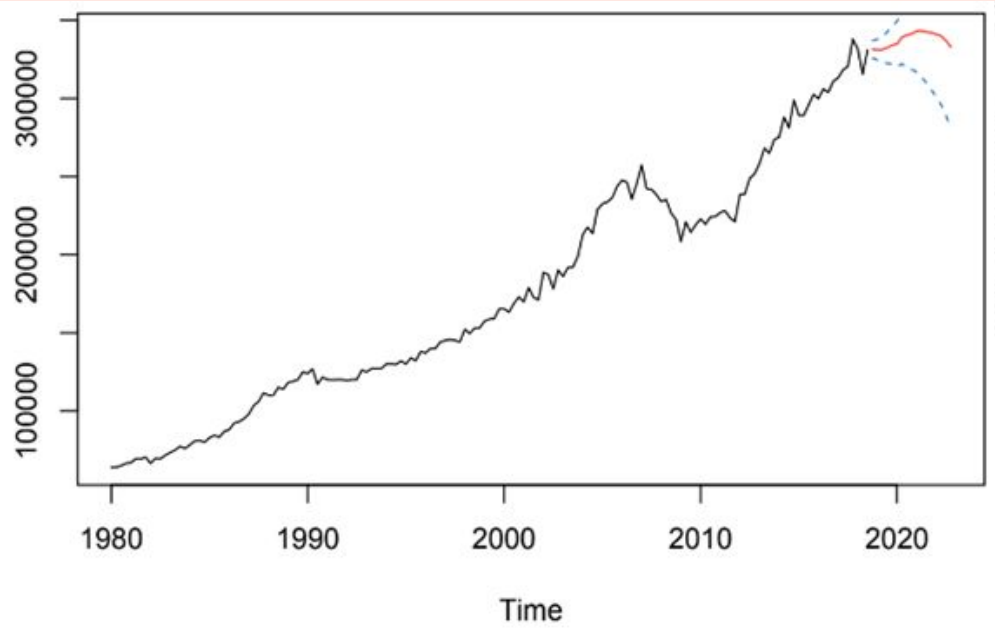




- No pattern in residual plot
- No autocorrelation in ACF
  - Confirmed by Box-Ljung test

- Better than the linear regressions
  - MAE of 0.01968 v. 0.02281
  - RMSE of 0.02625 v 0.02860

ARIMA MODEL FOR FORECASTING

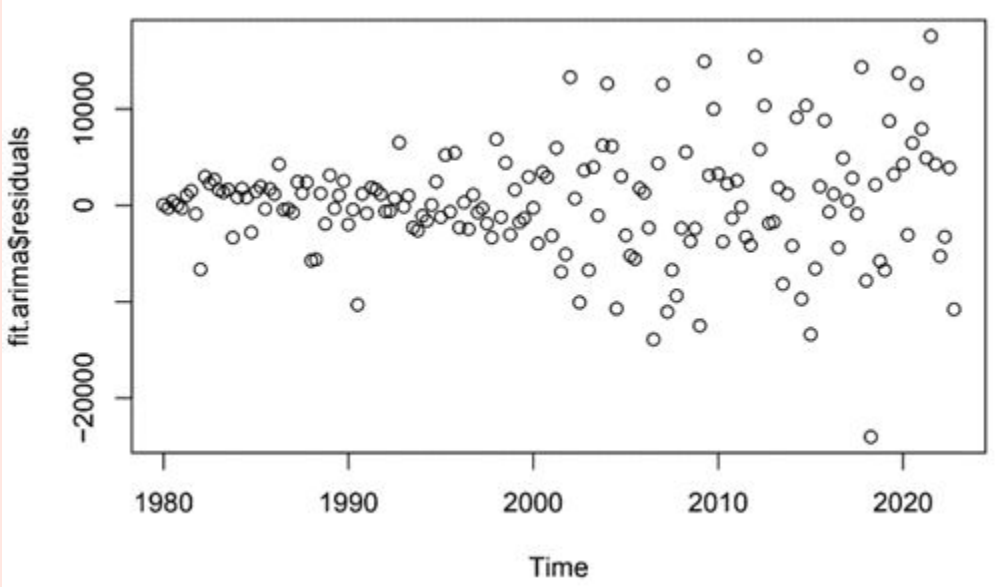


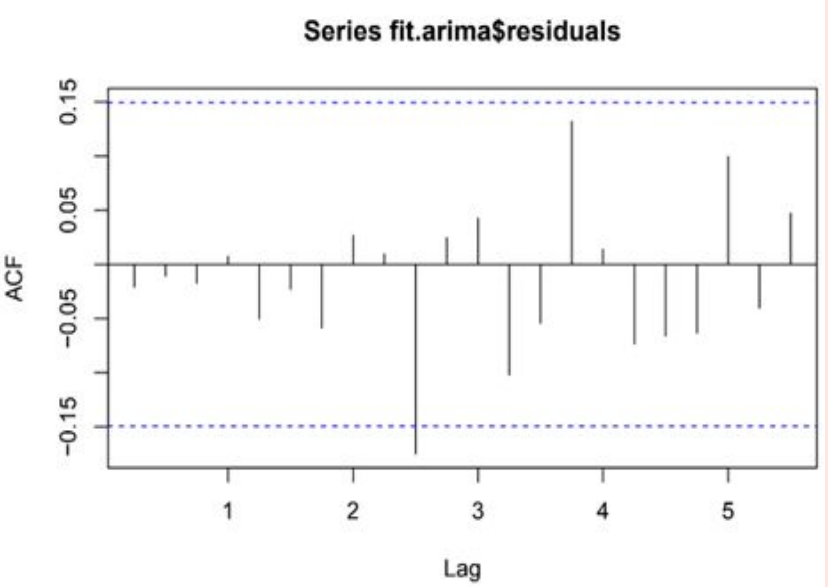

```
fit.arima.drop.pr$pred

##          Qtr1     Qtr2     Qtr3     Qtr4
## 2018                               331417.0
## 2019 331198.7 331172.2 332446.7 334031.7
## 2020 334768.8 339196.5 340426.0 341330.5
## 2021 343006.6 343029.4 342563.4 341912.3
## 2022 341039.3 339826.3 336834.7 332866.2
```

- Based on data set without last 10% of observations (17 points)
- Order (0, 3, 3)
- 29.41176% of the observed values were captured by the predicted values

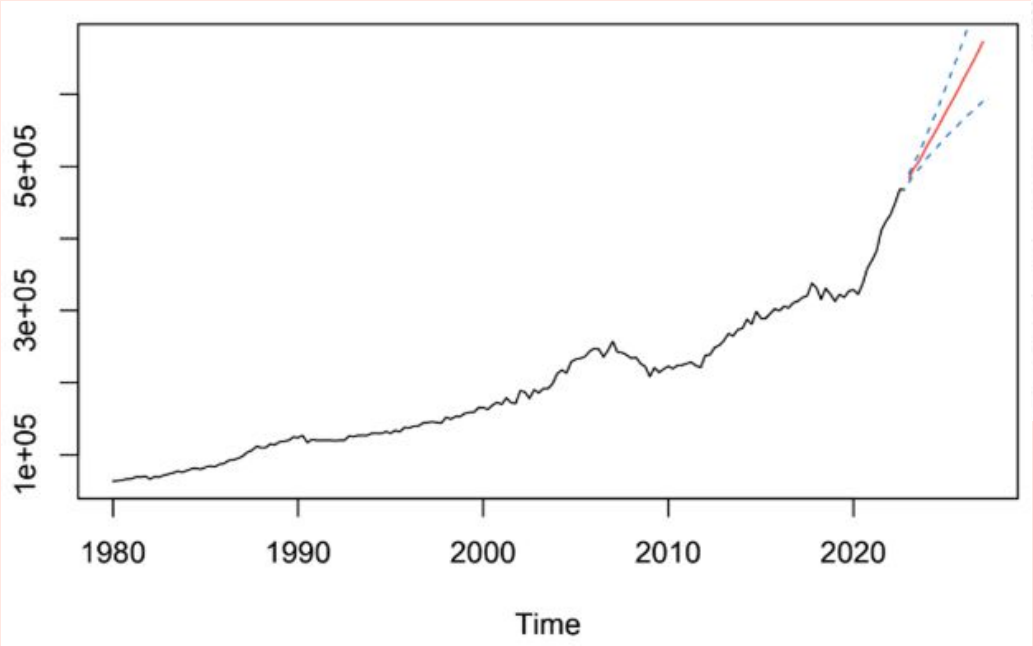
ARIMA MODEL (FULL DATA)





- Order (2, 3, 3)
- Fanning in residual plot - limitation
- Slight autocorrelation in ACF, but not much
    - Confirmed by Box-Ljung test

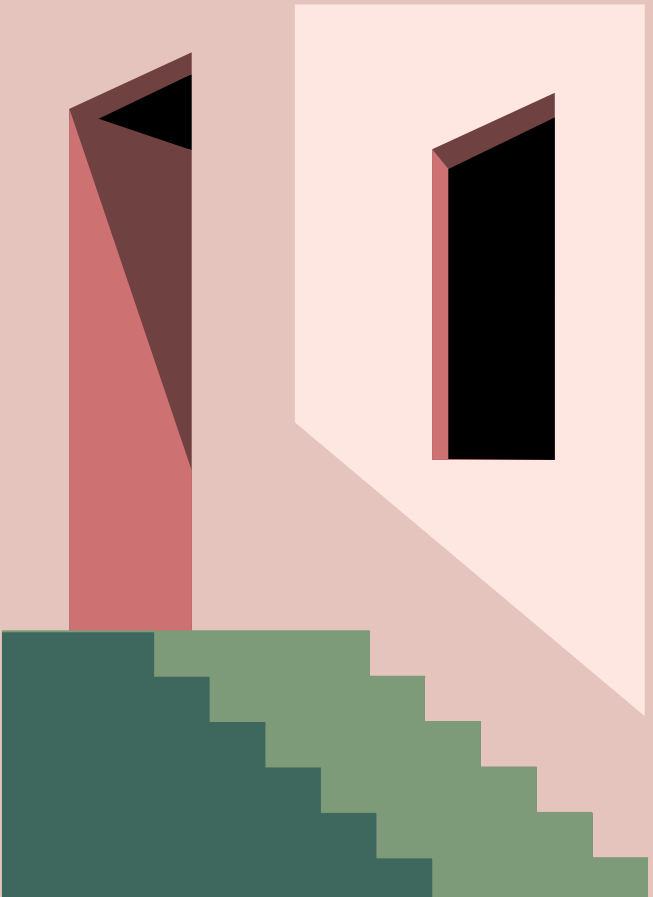
ARIMA MODEL FOR FORECASTING (FULL DATA)




```
fit.arima.pr$pred

##          Qtr1     Qtr2     Qtr3     Qtr4
## 2023 484001.4 496736.6 505438.0 517100.5
## 2024 529683.9 540551.1 551520.7 563609.0
## 2025 575908.2 587285.7 598738.1 611070.8
## 2026 623697.5 635450.6 647237.8 659903.0
## 2027 672875.6
```

- Forecast of the next four years
- Holt-Winters to forecast the separate predictors
  - Plugged values into ARIMA to forecast price



CONCLUSION

- Prices are going up, similar to pre-2008 recession
  - Could this be a bubble?
- Housing prices and recessions have a complex relationship
  - The lack of other economic regressors affect model accuracy
  - Current recessionary tensions are not as based on the housing market as those in 2008 were