



# Time Series Analysis of Proportion of Issued License Plates in Shanghai

Carrie Yan



# Introduction

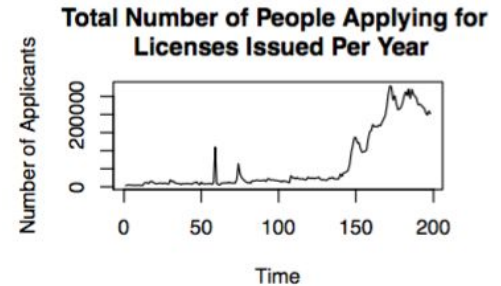
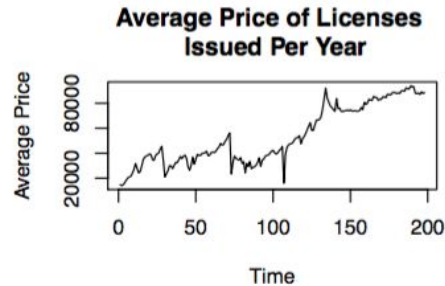
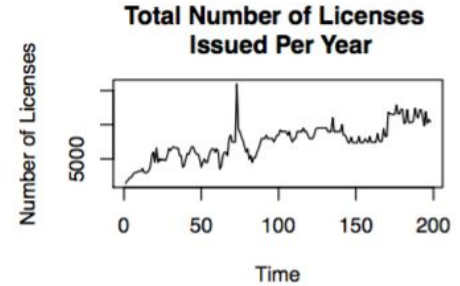
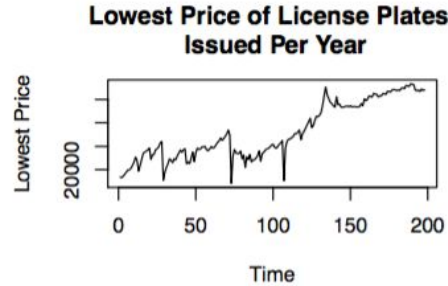


- Monthly data from January of 2002 to 2018
- Sell a limited number of license plates to fossil-fuel car buyers
- Forecast the monthly proportion of license plates issued to the number of applicants up until the year 2020
- License plate in Shanghai is referred to as "the most expensive piece of metal in the world"
- Average price is about \$13,000
- Government's attempt to combat Shanghai's air pollution problem

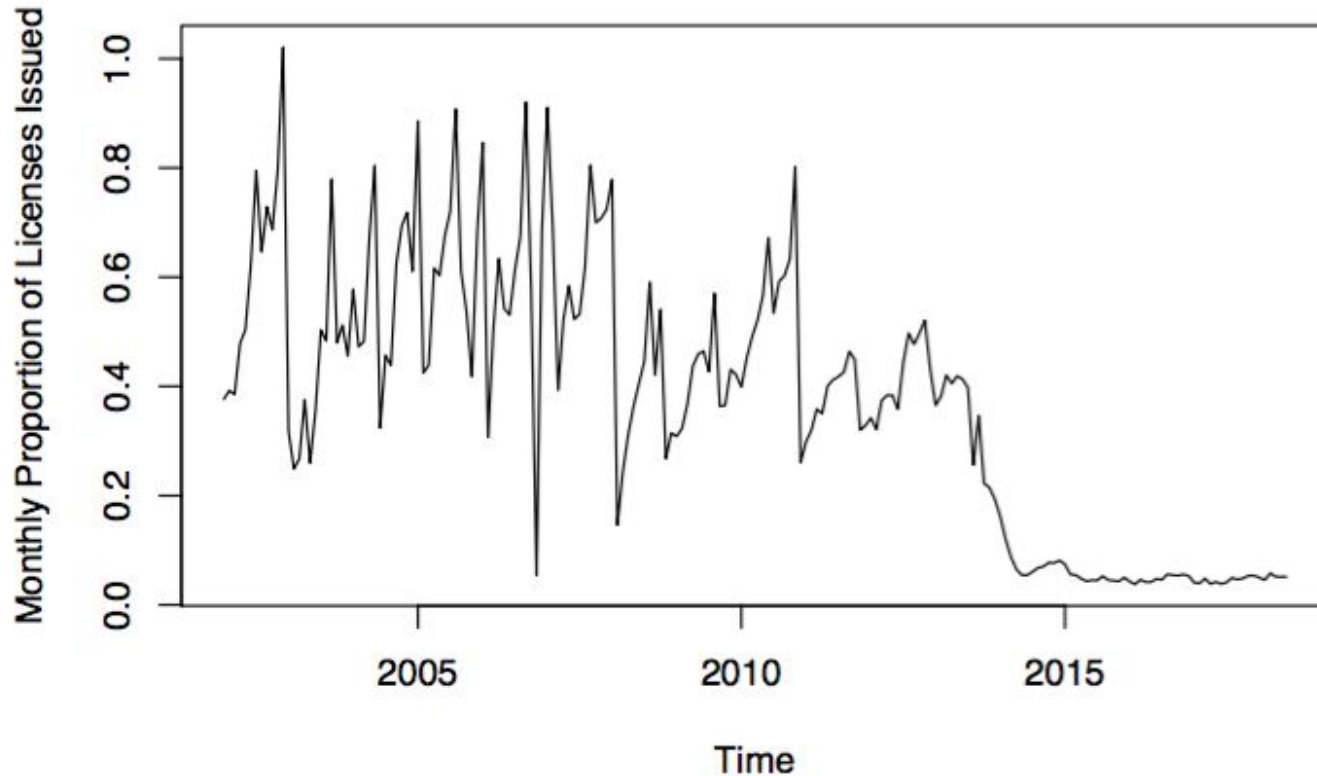
# Initial Analysis

We convert the data into a time series and plot each of the four variables:

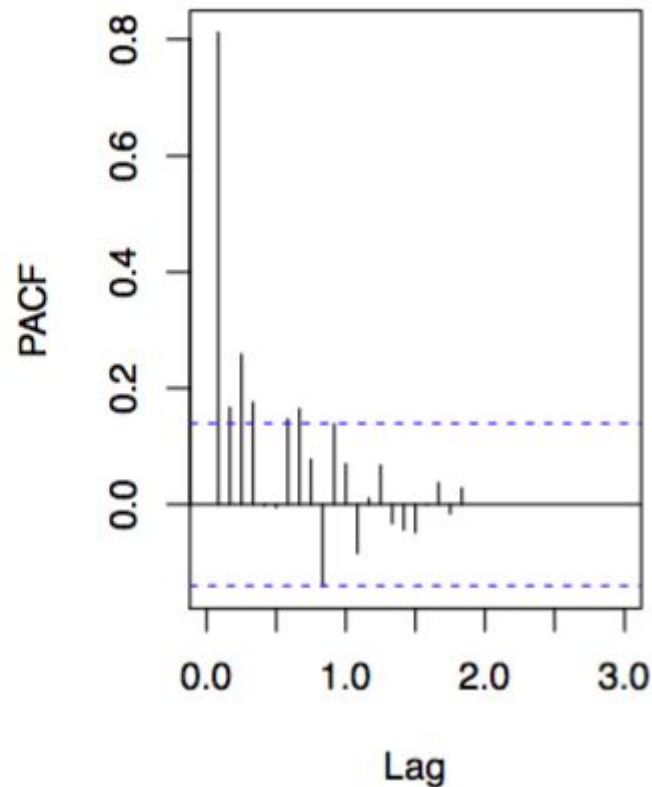
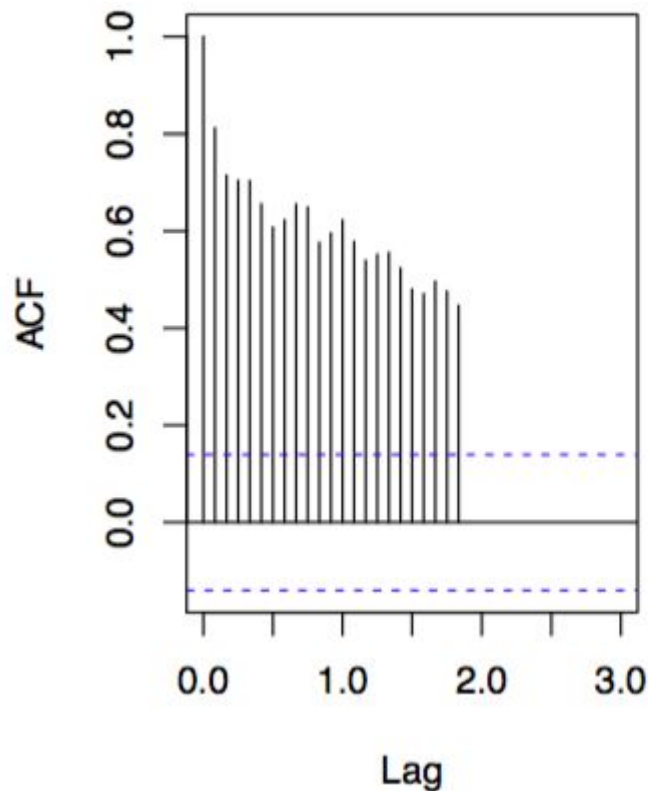
1. Lowest price
2. Total number of license plates issued
3. Average price
4. Total number of applicants



# Monthly Proportion of Licenses Issued in Shanghai (2002-2018)



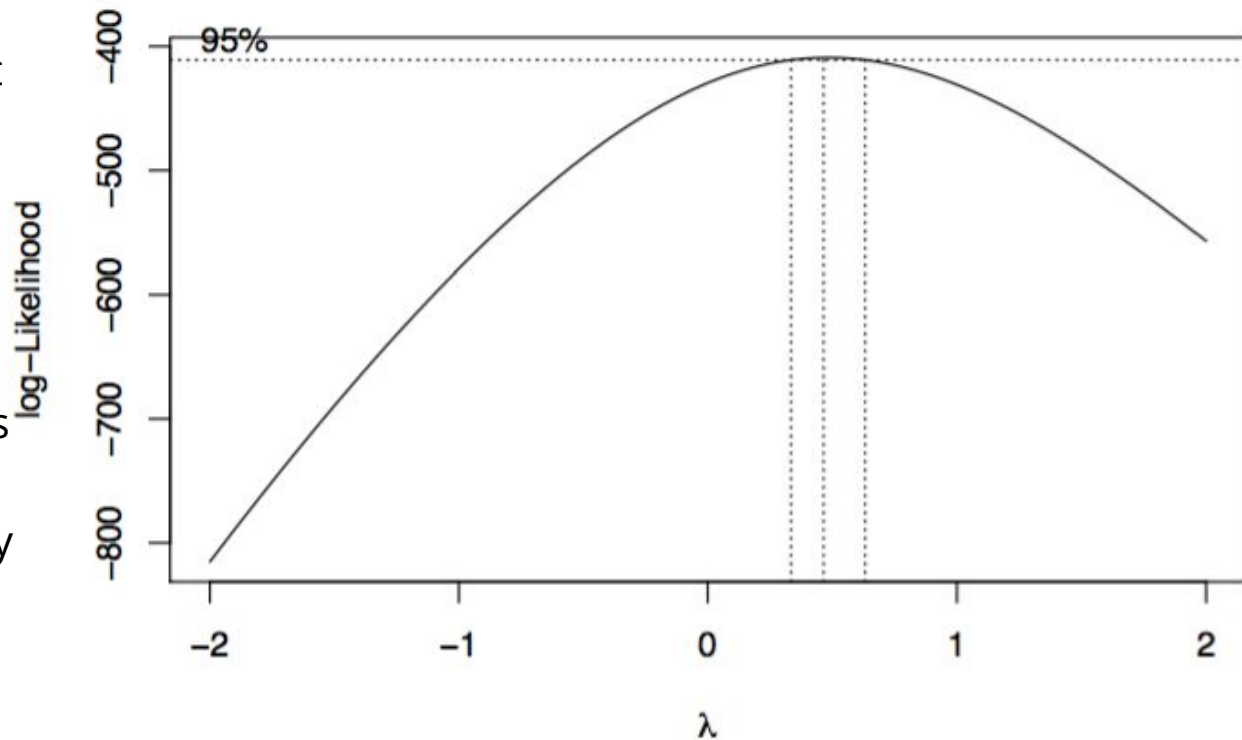
# ACF and PACF of Proportion of Shanghai-Issued License Plates



# Transformations

# Box-Cox

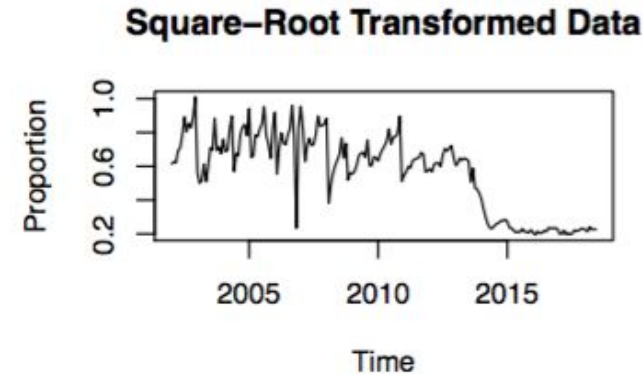
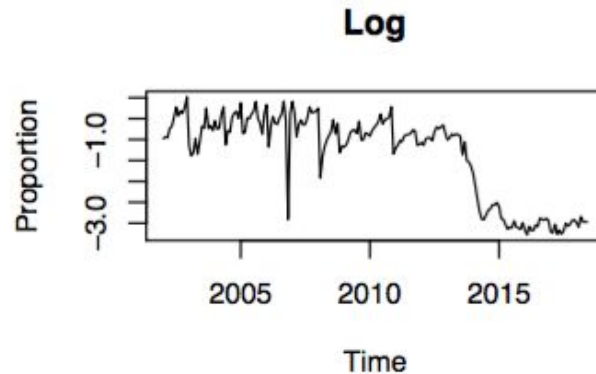
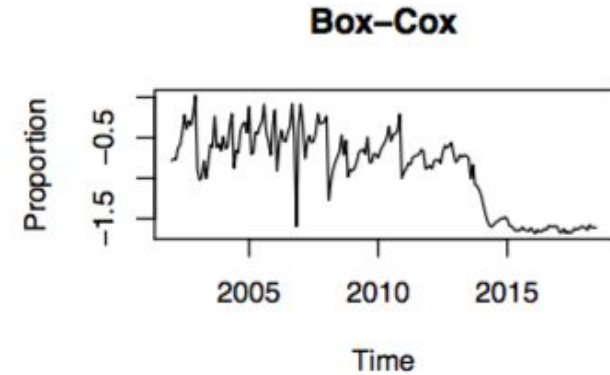
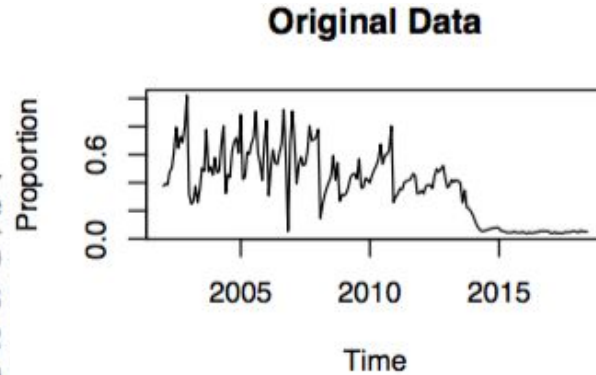
- Initial time series was not stationary and due to heteroscedasticity, violated our constant error of variance assumption
- Variance of error changes over time
- $\lambda = 0.46$  which is relatively close to 0.5
- Indicates square-root transformation performs best



# Transformations

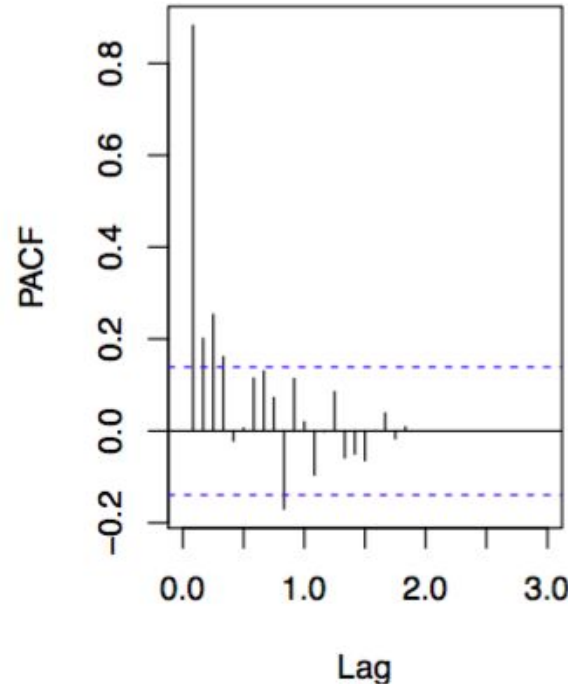
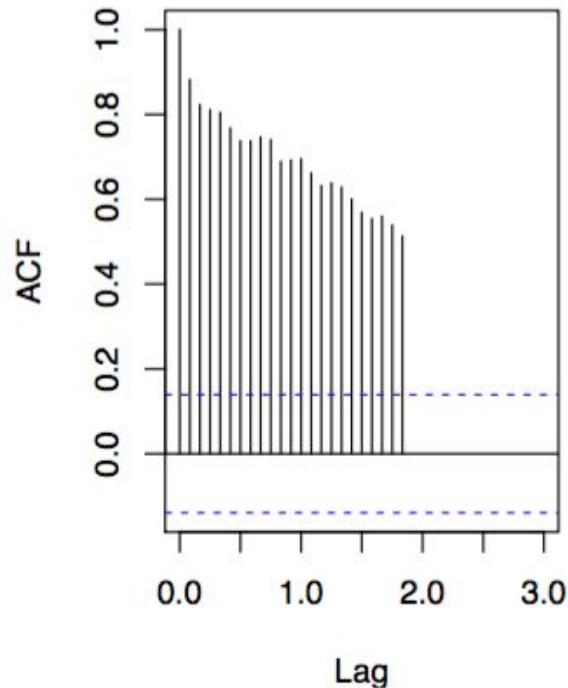
	Variance
Original Time Series	0.0609
Box-Cox Transformation	0.2393
Log Transformation	1.0672
Square-Root Transformation	0.0540

- Find the variances of each transformation to determine the best fit for our model
- Square-root transformation gives the smallest variance value





# ACF and PACF of Square-Root Transformed Time Series



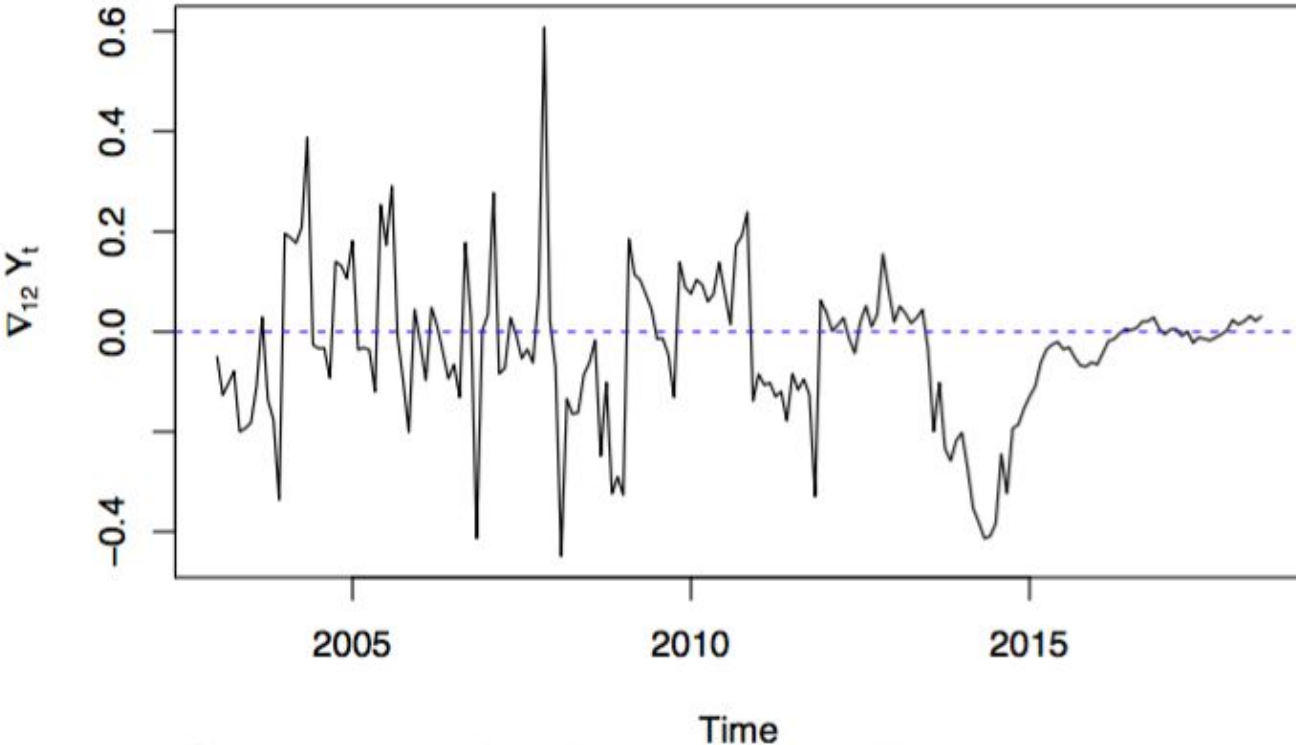
- ACF is still tailing off
- PACF cuts off around lag 0.8
- Supports our initial assumption that the series follows an AR(p) model



# Differencing to Remove Seasonality and Trend



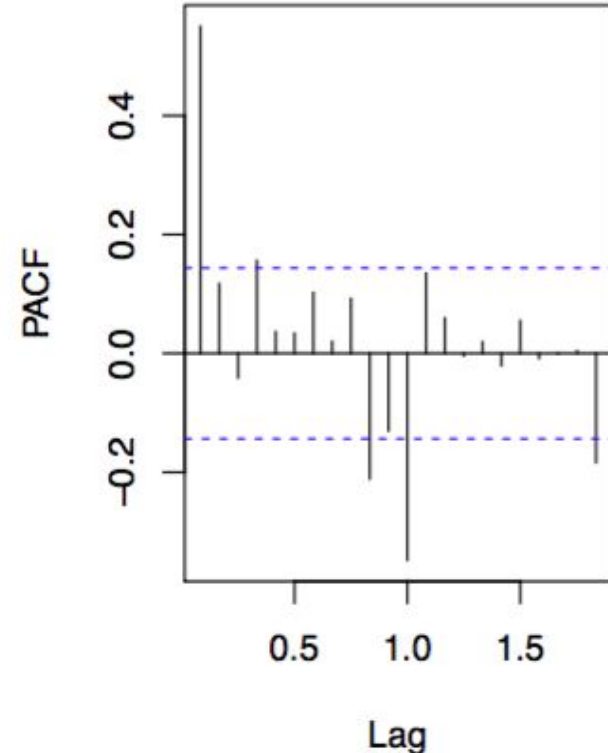
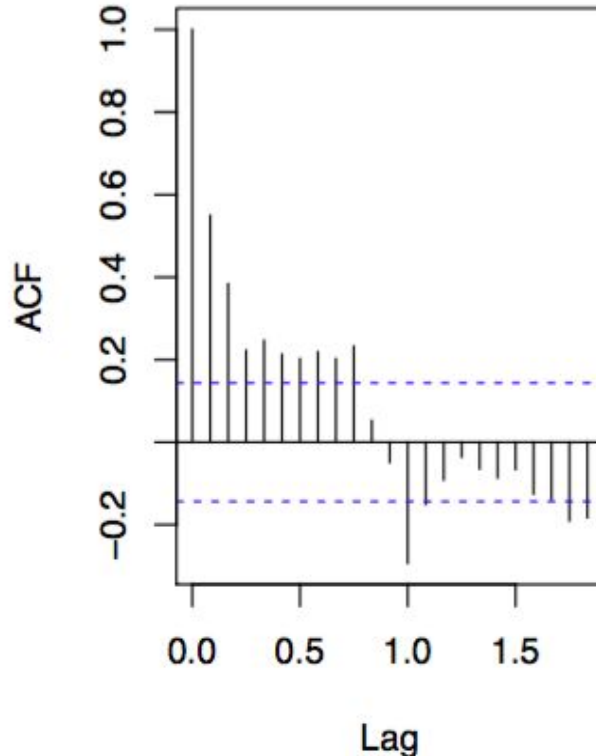
# Removing seasonality



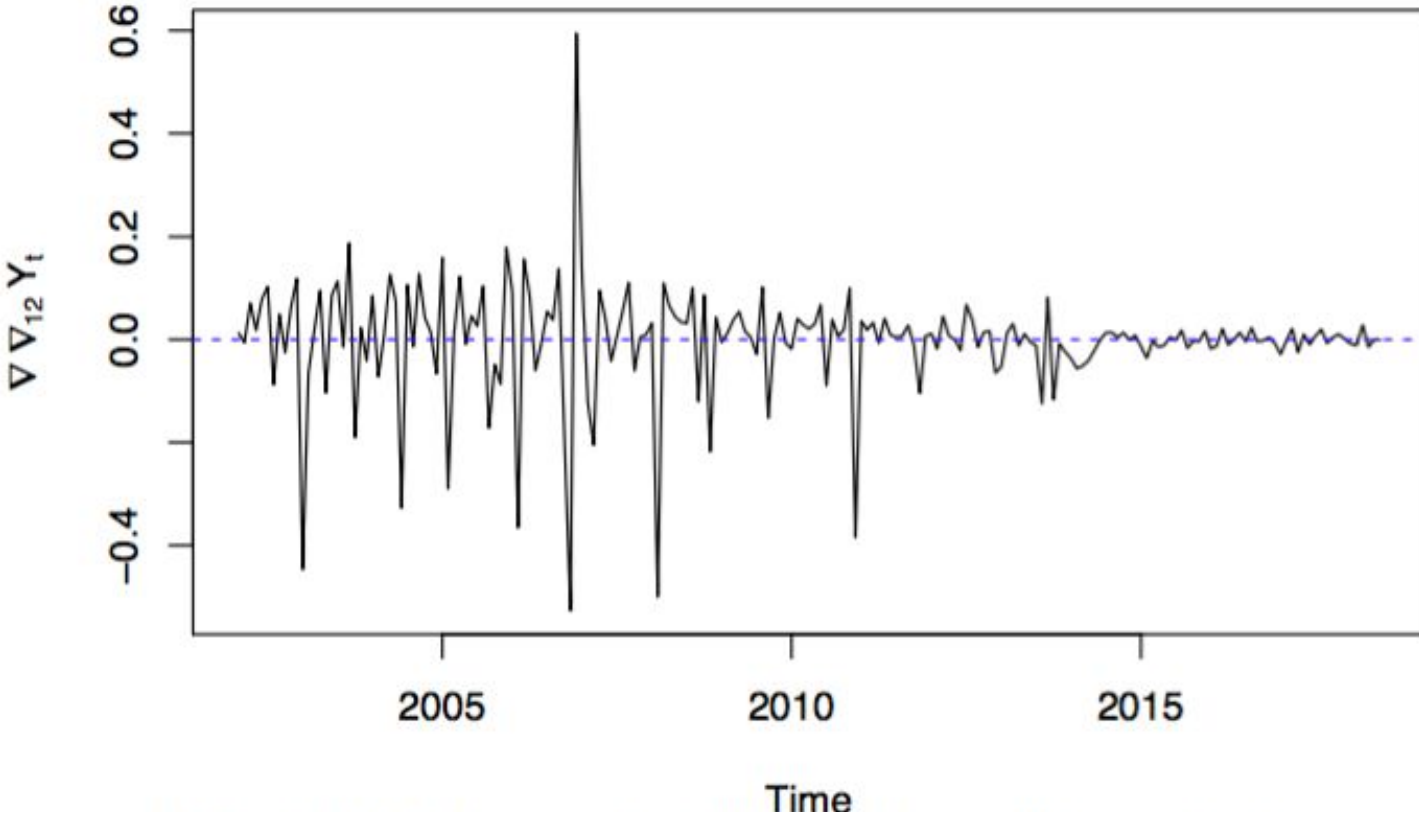
- Data still does not look stationary after applying square-root transformation, so we apply differencing
- Difference once at lag 12 to remove seasonality
- De-seasonalized data fluctuates around the mean = 0 line

# ACF and PACF plots of de-seasonalized data

- ACF slowly begins to decay
- PACF oscillates between the bounds

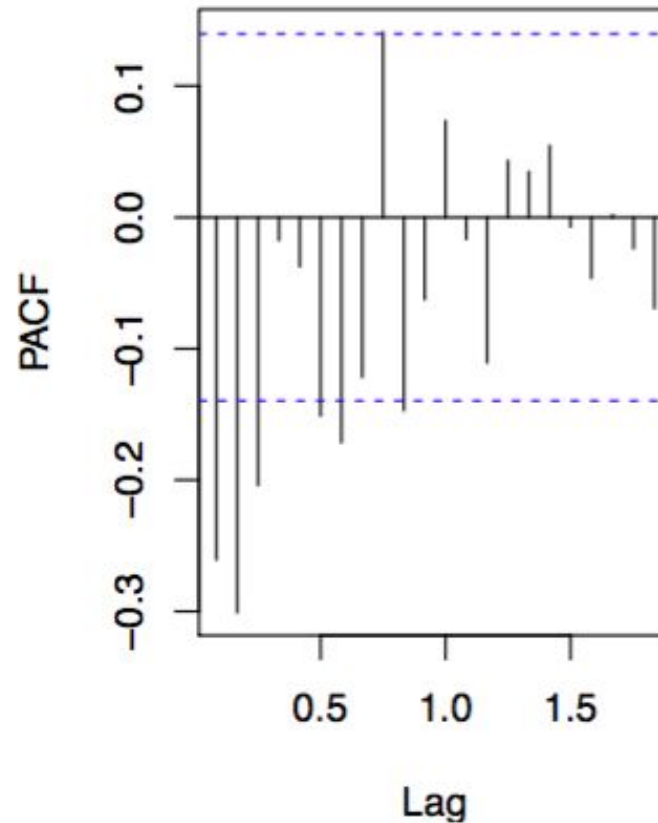
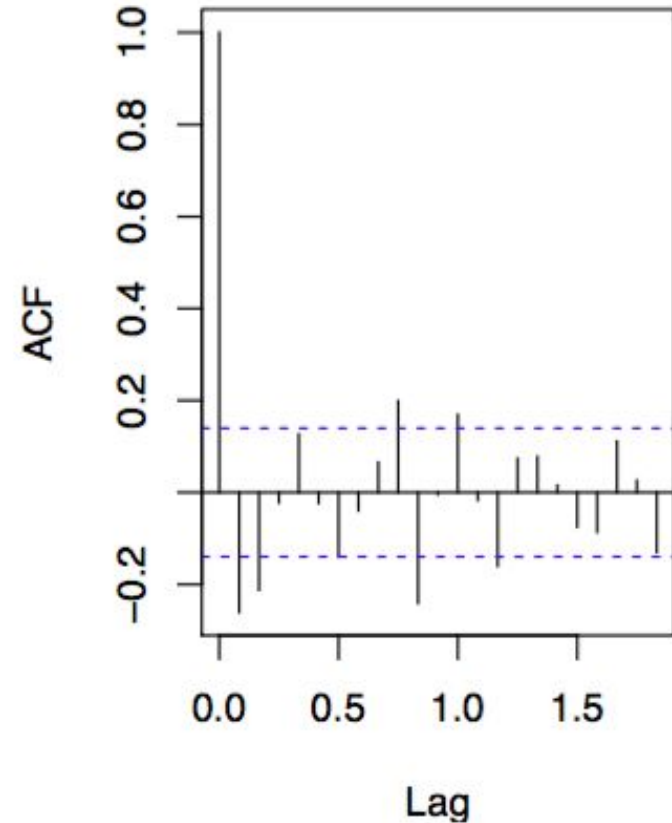


# Removing trend and seasonality



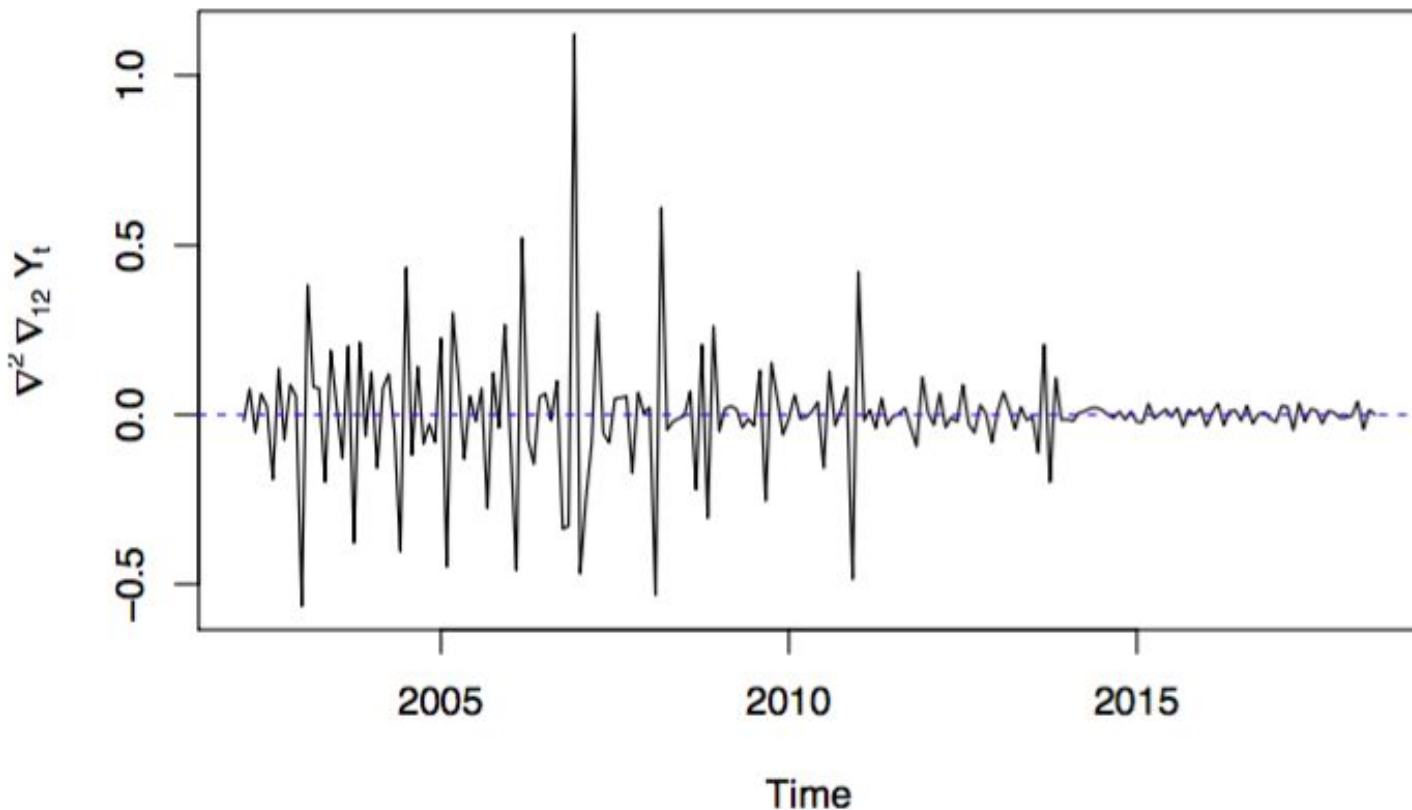
- Difference again at lag 1 to remove trend
- Plot fluctuates very closely around the mean = 0 line
- Stationary

# ACF and PACF of de-trended and de-seasonalized data



- ACF oscillates between the bounds
- PACF seems to cut off around lag 0.1

# Overdifferencing



- Variance increases from 0.012 to 0.031 when we difference a 2nd time at lag 1
- Indicates overdifferencing
- Only difference at lag 1 once

# Parameter Estimation using Yule-Walker

- Preliminary estimation using Yule-Walker gives an AR(10) model

Call:

```
ar(x = shanghai_prop.diff1, method = "yule-walker")
```

Coefficients:

1	2	3	4	5	6	7	8
-0.4189	-0.4095	-0.2802	-0.1567	-0.2030	-0.2471	-0.2010	-0.1161
9	10						
0.0761	-0.1469						

Order selected 10     $\sigma^2$  estimated as    0.009353



# Fitting an ARMA process

- Use `auto.arima()` function to find the estimated model is a ARIMA(1,0,1)

```
Series: shanghai_prop.diff1  
ARIMA(1,0,1) with zero mean
```

```
Coefficients:
```

	ar1	ma1
	0.3560	-0.8184
s.e.	0.0962	0.0570

```
sigma^2 estimated as 0.009847:  log likelihood=176.33  
AIC=-346.65    AICc=-346.53    BIC=-336.8
```

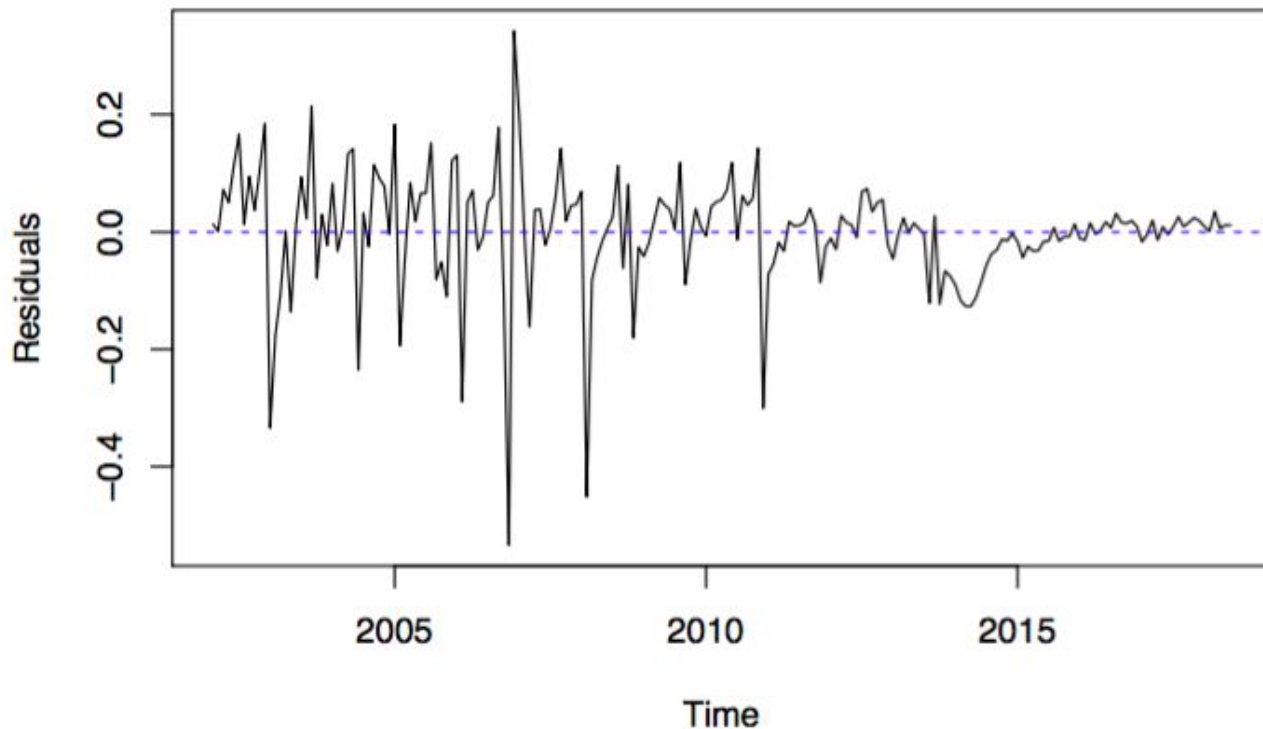
# Checking for best model fit

- We run further AIC tests and find our best model using estimated orders of  $(p,q)$
- Test each of the possible  $\text{ARMA}(p,q)$  parameter values to see which process gives us the smallest value of AIC using a for-loop
- $\text{ARMA}(1,1)$  gives us the lowest AIC value of -346.0987, so it is the best model for our data

	AIC
AR(1)	-317.5312
MA(1)	-336.1856
ARMA(1,1)	-346.0987

# Plotting Residuals of ARMA(1,1)

- Residuals seem to oscillate about the line at error 0



# Diagnostic Checking of Residuals

## Shapiro-Wilk normality test

```
data: err  
W = 0.86508, p-value = 3.169e-12
```

## Box-Ljung test

```
data: err  
X-squared = 0.080816, df = 1, p-value = 0.7762
```

## Box-Pierce test

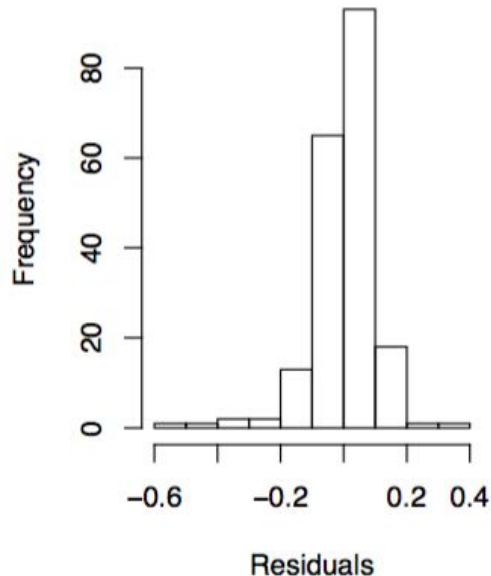
```
data: err  
X-squared = 0.079598, df = 1, p-value = 0.7778
```

- Check for normality of errors, if the residuals are not heteroskedastic and have constant variance, and if the residuals are serially correlated
- ARMA(1,1) passes the other 2 tests but does not pass the Shapiro-Wilk test of normality of errors
- Ljung-Box test (✓)
  - constant variance
  - residuals are not heteroskedastic
  - residuals are random
- Box-Pierce test (✓)
  - residuals are serially correlated

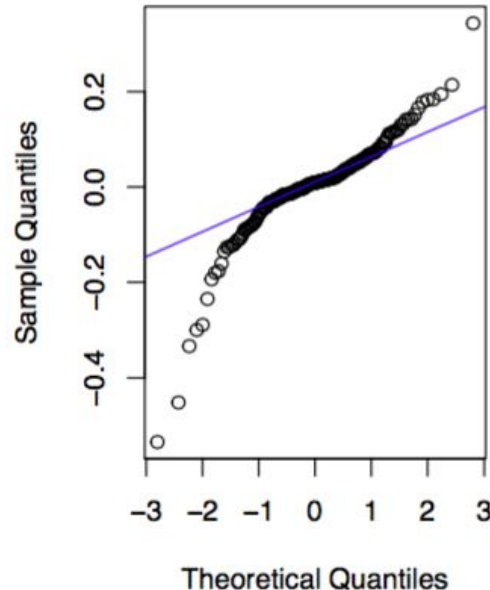
# Histogram and QQ-Plot

- Histogram shows data is normally distributed
- QQ-plot shows errors follow diagonal line so they are normally distributed

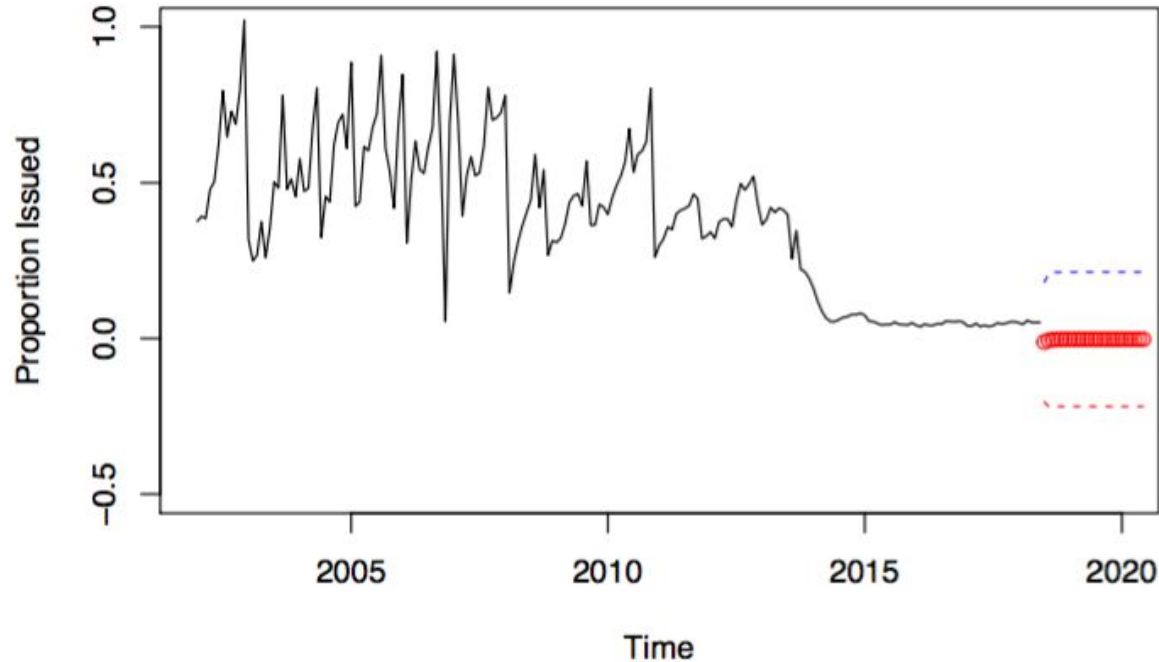
**Histogram of Residuals**



**Normal Q-Q Plot**



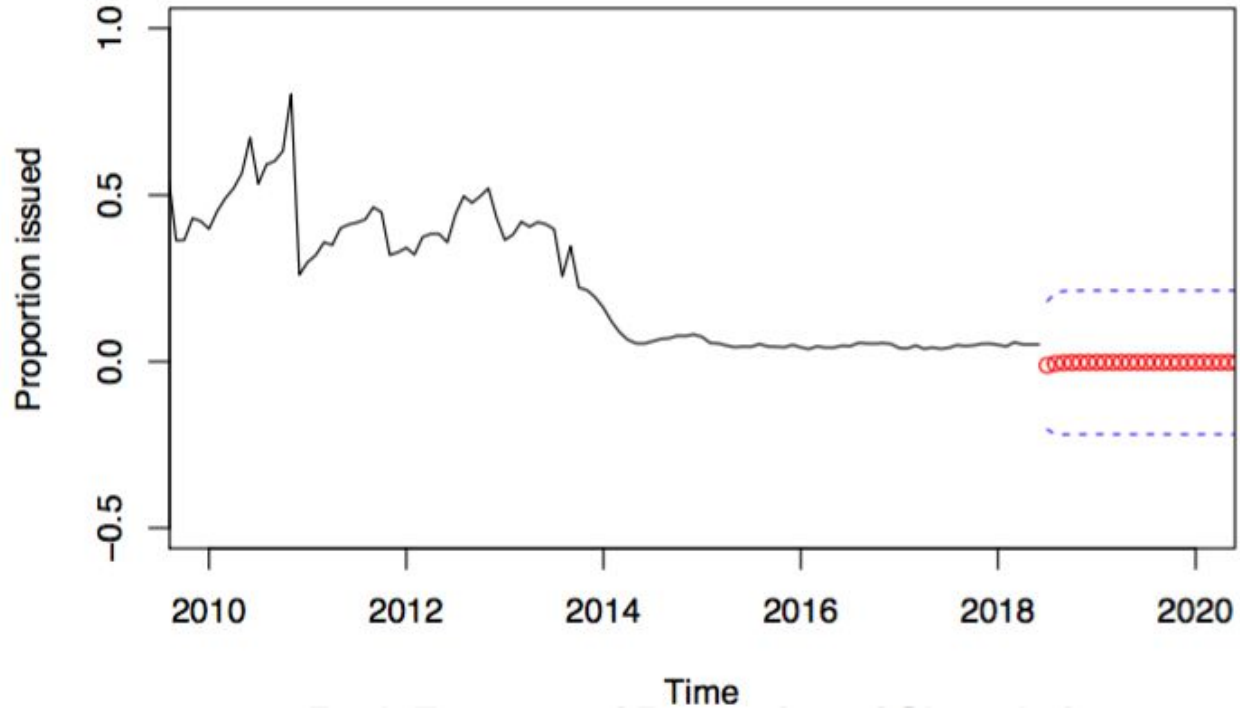
# Forecasting



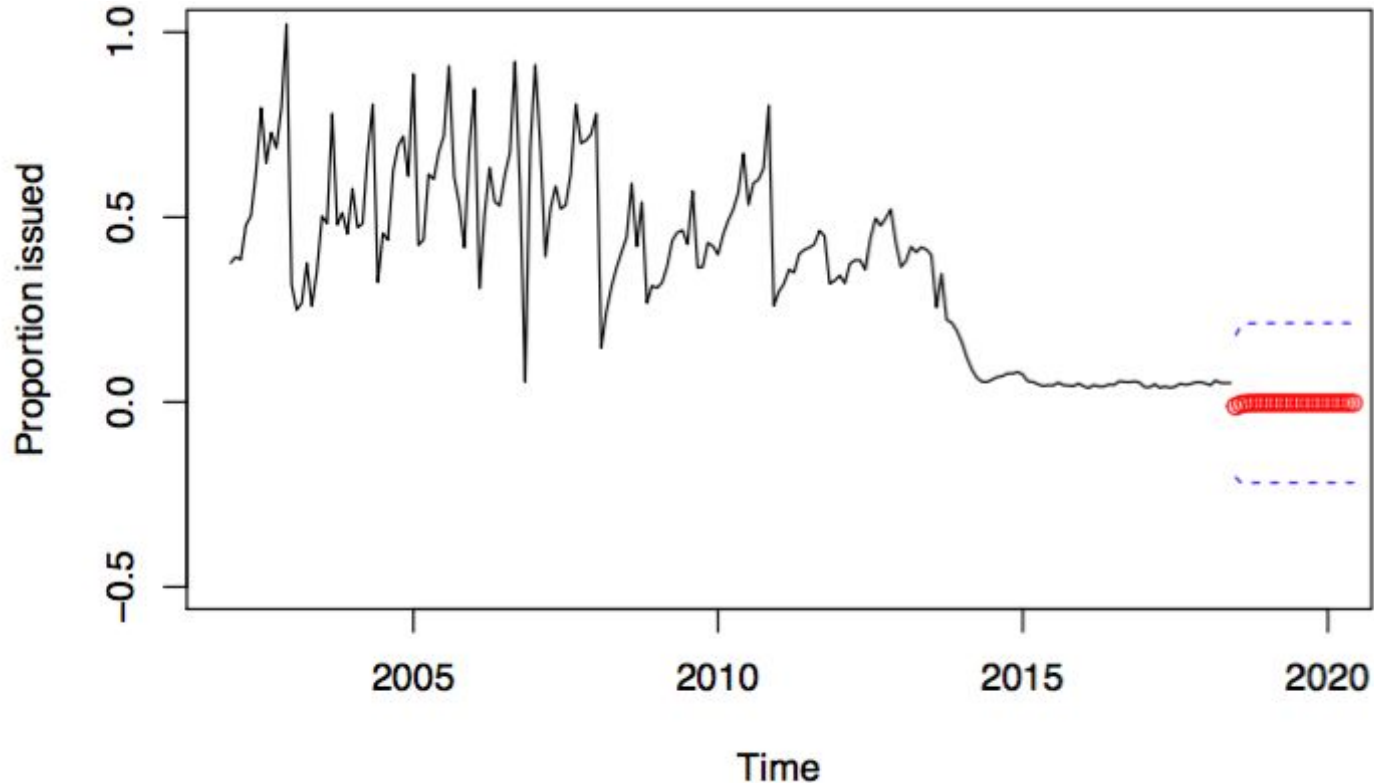
- We predict the proportion of license plates issued for the next 2 years using our ARMA(1,1) model
- Plot upper and lower bounds to calculate a 95% confidence interval for the predicted values

# Back-Transformation

- We obtain back forecasted values for the proportion for the next 2 years
- Our predicted values stay consistent for the next 2 years (2018-2020) so our proportion remains relatively consistent



# Back Forecast of Proportion of Shanghai-issued license plates from 2002-2020





# Conclusion

- Used monthly data to analyze the proportion of Shanghai-issued license plates per month to total number of applicants from 2002 to 2018
- Concluded ARMA(1,1) was the best model using Yule-Walker method, for-loop to compare AIC values, and auto.arima() function
- Forecasted values for the next 2 years (2018 to 2020)
- Predicted values show the proportion will stay relatively consistent as time goes on
- Consistent trend as number of license plates issued and number of applicants continue to fluctuate

# The Future of Shanghai

- If a certain number of people apply for a license plate per month, then the Shanghai government attempts to regulate the number of license plates by proportionally reducing the number available at auction
- Goal is to contain or reduce pollution, as forecasted for the next 2 years





Thank you!

