**World Quant University**

**Professor: Tiberiu Stoica**

**Econometrics**

Nikolas Lippmann Pareschi - [nikolaslippmann@gmail.com](mailto:nikolaslippmann@gmail.com)

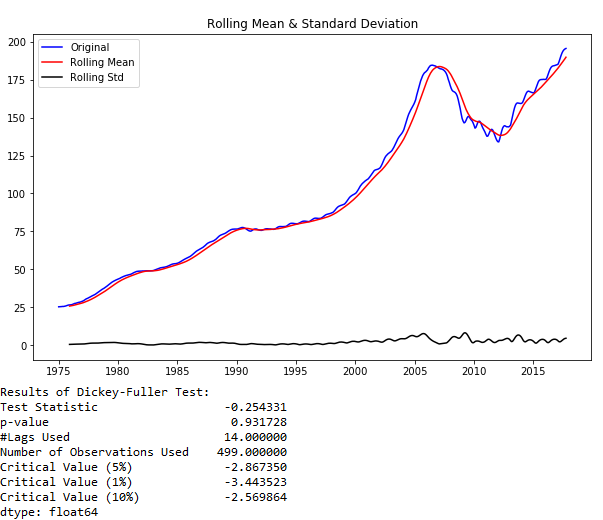
**Mini Project: Mini Project: Forecasting U.S. House Prices**

### 1: Implement the Augmented Dickey-Fuller Test for checking the existence of a unit root in Case-Shiller Index series.

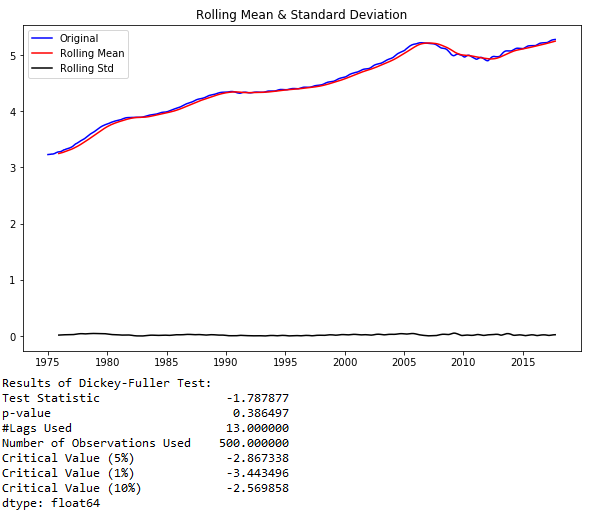
After reading the house data we defined a function to test for stationarity. The function was based on

<https://www.analyticsvidhya.com/blog/2016/02/time-series-forecasting-codes-python/>

We tested the original data for a unit root which implies in non-stationarity and we got results that clearly showed that the data was non-stationary (not rejected). A clearly no stable mean and standard deviation and a p-value superior than 0.05 and T-stats higher than the critical values.



After that we tried a trend elimination with log and we runned the test again. The results ere still far from stationarity, but better:



Our next step was to take the difference between the logs. That improved our results a lot:

### 

### Nevertheless we were yet not below a p-value of 0.05. We are not rejecting the unit root at that level. So we took once again the difference of our data (difference of difference of logs):

### 

### These results were a lot better. We rejected the unit root. The data seems more stationary. The rolling standard deviation is not stable but it is ok if we consider the subprime crisis in 2007-2008.

### 2 :Implement an ARIMA(p,d,q) model. Determine p, d, q using Box-Jenkins methodology. Discuss the results provided by Autocorrelation Function (ACF) and Partial Autocorrelation Function (PACF).

### We ploted the ACF and PACF in our best Dickey-Fuller double differentiated log data:

### 

### Considering:

### 

### <http://www.itl.nist.gov/div898/handbook/pmc/section4/pmc446.htm>

### And:

### 

### From Chapter 470 from NCSS Statistical Software (second picture):

### We believe a good ARIMA would be (1, 2, 0) because the partial autocorrelation is clearly finite, cutting quickly after some lags. So the model is (p, d, 0). The autocorrelation function intercalates between positive and negative number and we didn’t see major lags pointing out that p would be higher than 1. As we differentiated the data 2 times we have d = 2.

### 3: Forecast the future evolution of Case-Shiller Index using the ARMA model. Provide one month, two-month and three-month forecasts. Test the model using in-sample forecasts

### These are the predictions and the mean square error:

### 

### 

### 

### 4: Suggest exogenous variables that can be introduced which can improve the forecasts.

### 

House prices are correlated to interest rates. When it is cheap to borrow money the demand for houses usually improves. Also, money supply is correlated to house prices. Another variable is GDP, a hot economy will probably increase de demand for houses and their prices. Another indicator I believe is the number of people per square meter in a city for example. London, Singapore and Hong Kong have ultra-expensive houses. These cities have an elevated number of people per square meter.