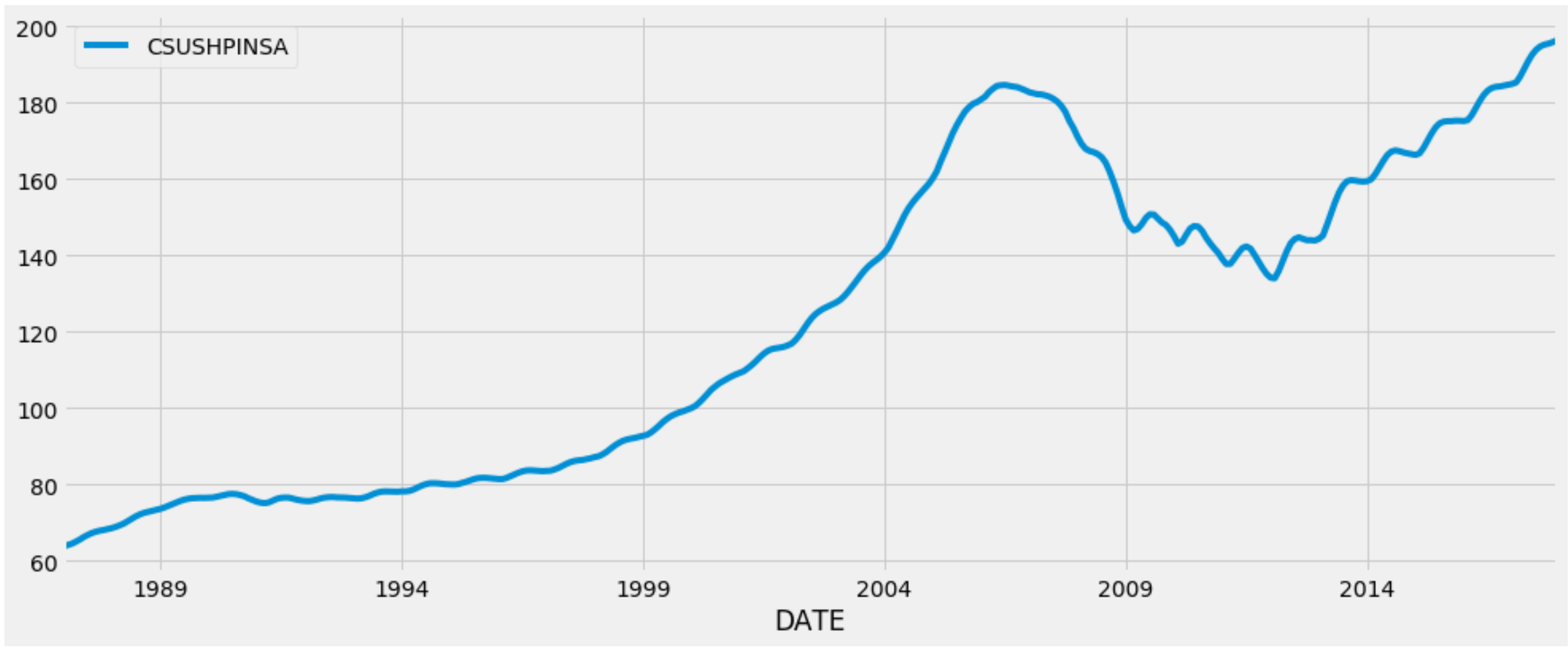
In [201]:

Removing all variables...

In [201]: runfile('C:/Users/Tawanda Vera/Downloads/3-Refine.py', wdir='C:/Users/Tawanda Vera/Downloads')



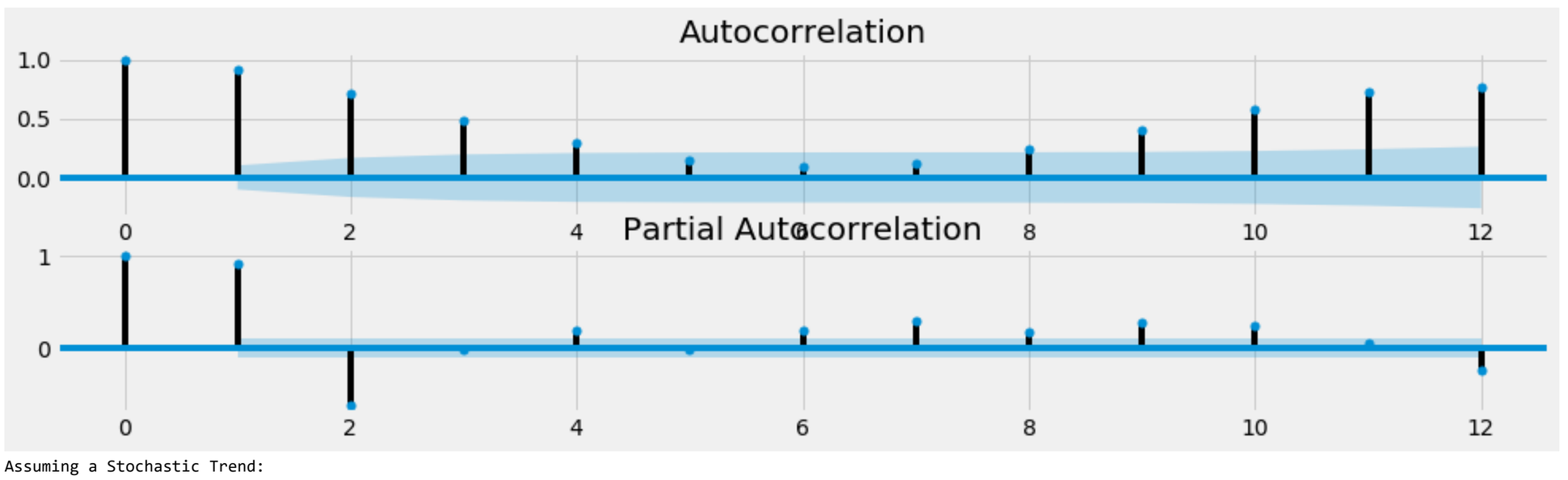
Results of the Augumented Dickey-Fuller Test:

Test statistic	-0.509918
p-value	0.890067
#Lags Used	13
Number of Observations Used	358
Critical Value	{'1%': -3.448748905151901, '5%': -2.8696473721...
Maximized information criterion	30.5212
Critical Value (1%)	-3.44875
Critical Value (5%)	-2.86965
Critical Value (10%)	-2.57109

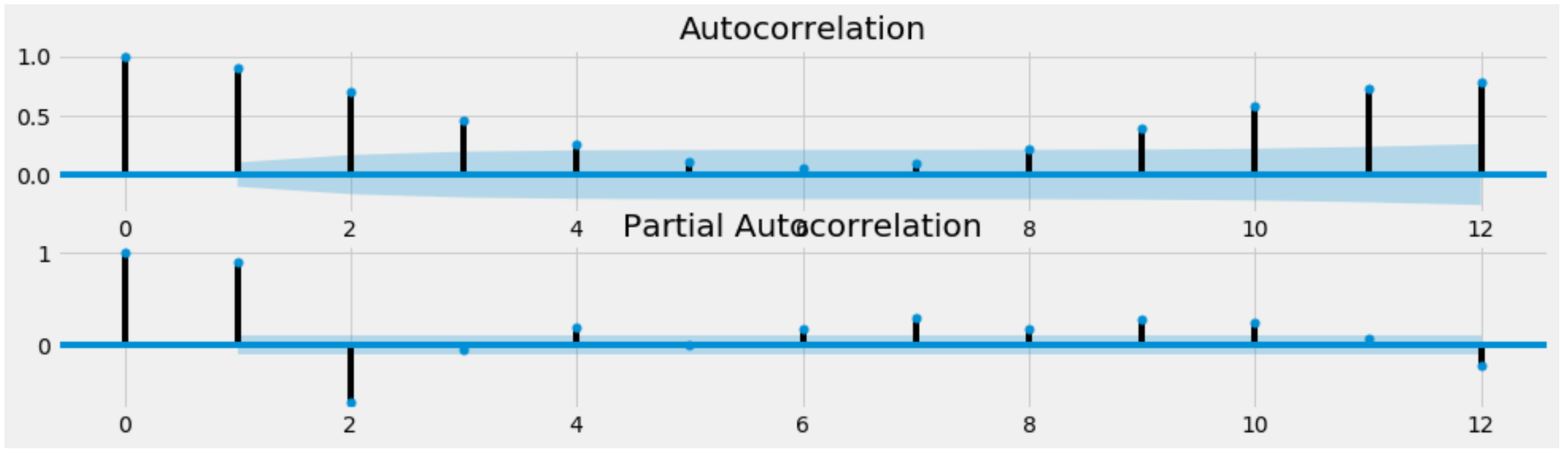
dtype: object

The pvalue = 0.97 is the higher as compared to the critical values of(1%)= -3.45, (5%) = -2.86 and (10%)= -2.57 Therefore we cannot reject the null hypothesis. As such, The U.S. National Home Price Index data presents a unit root therefore, it is nonstationary.

Assuming a Deterministic Trend:



Assuming a Stochastic Trend:



The graphs plotted, show that ACF is decreasing very very slowly. WHile, PACF cuts off at 2. Thus, we can conclude this to be a ARIMA(1,0,0) = AR(1) [The lag at which the PACF cuts off is the indicated number of AR terms]

Results for one month, two-month and three-month forecasts: One-month Forecast 116.385599
Two-month Forecast 72.692691
Three-month Forecast 47.568603
Four-month Forecast 32.462686
dtype: float64

We can add other indices related to home price Index to improve the prediction. These include the economic growth rate, the consumer price index, nominal wages as a percentage of GDP, the short-term interest rate, mortgage loans as a share of GDP and population in the 15-64 cohort as a percentage of GDP.

In [202]: