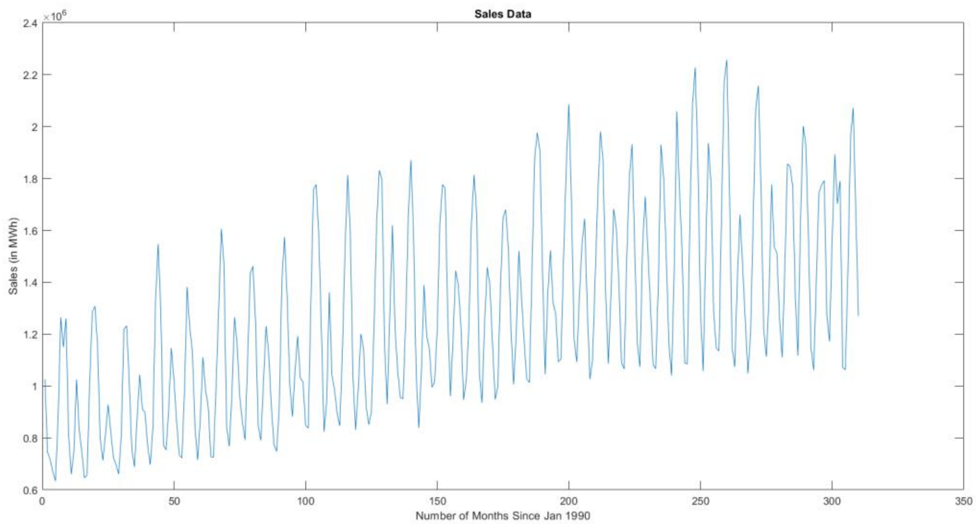
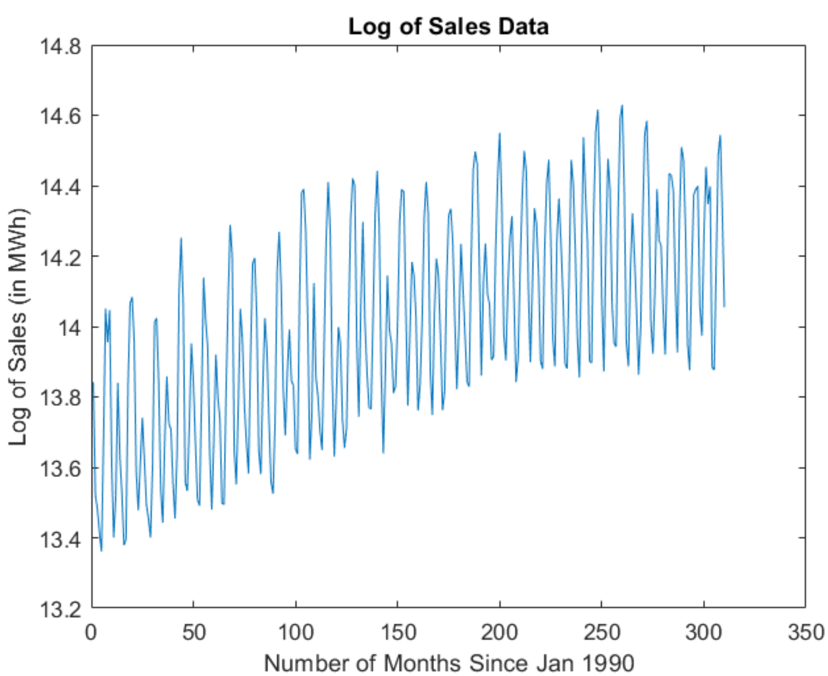
Task 2:

Below is a plot of the regular sales data. It is evident that the variance is increasing with time. Therefore, even if we removed the trend and seasonality, we would not have a weakly stationary time series.

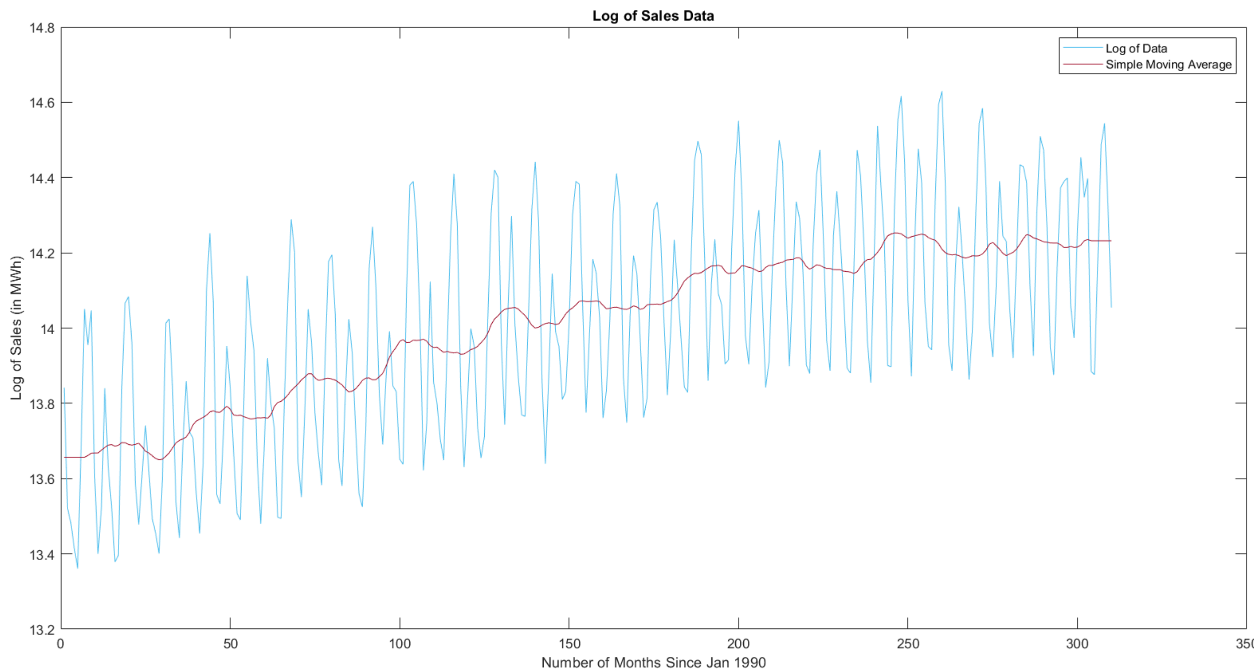


We plot the log of the data to check if the variance stays relatively constant:



This variance looks a lot better, and therefore we shall use it as the basis for our analysis.

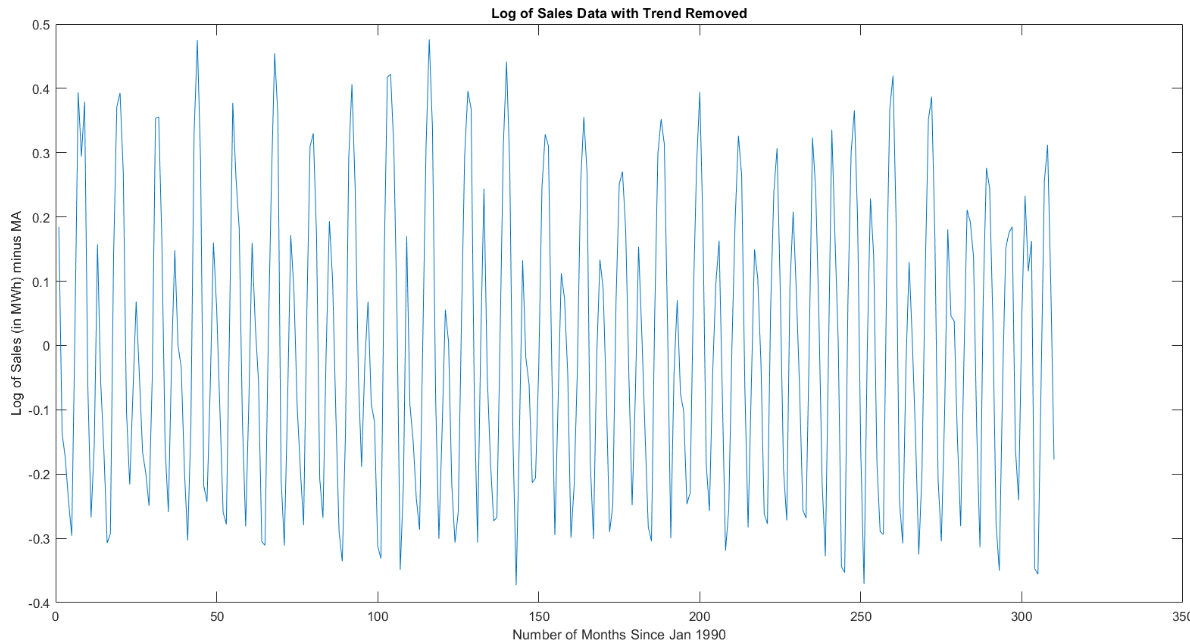
Task 3:



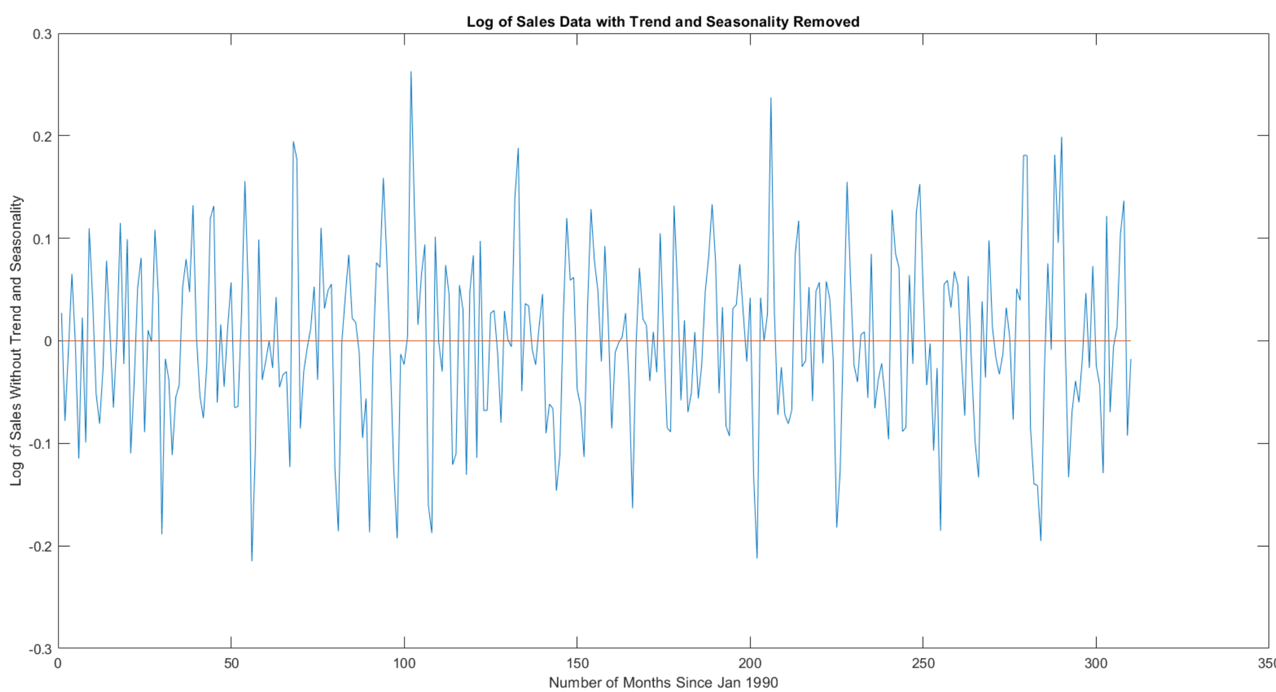
The red line is a 12- term simple moving average. It was calculated as the average of two asymmetric moving averages of length 11 (see code)

Task 4 AND Task 5

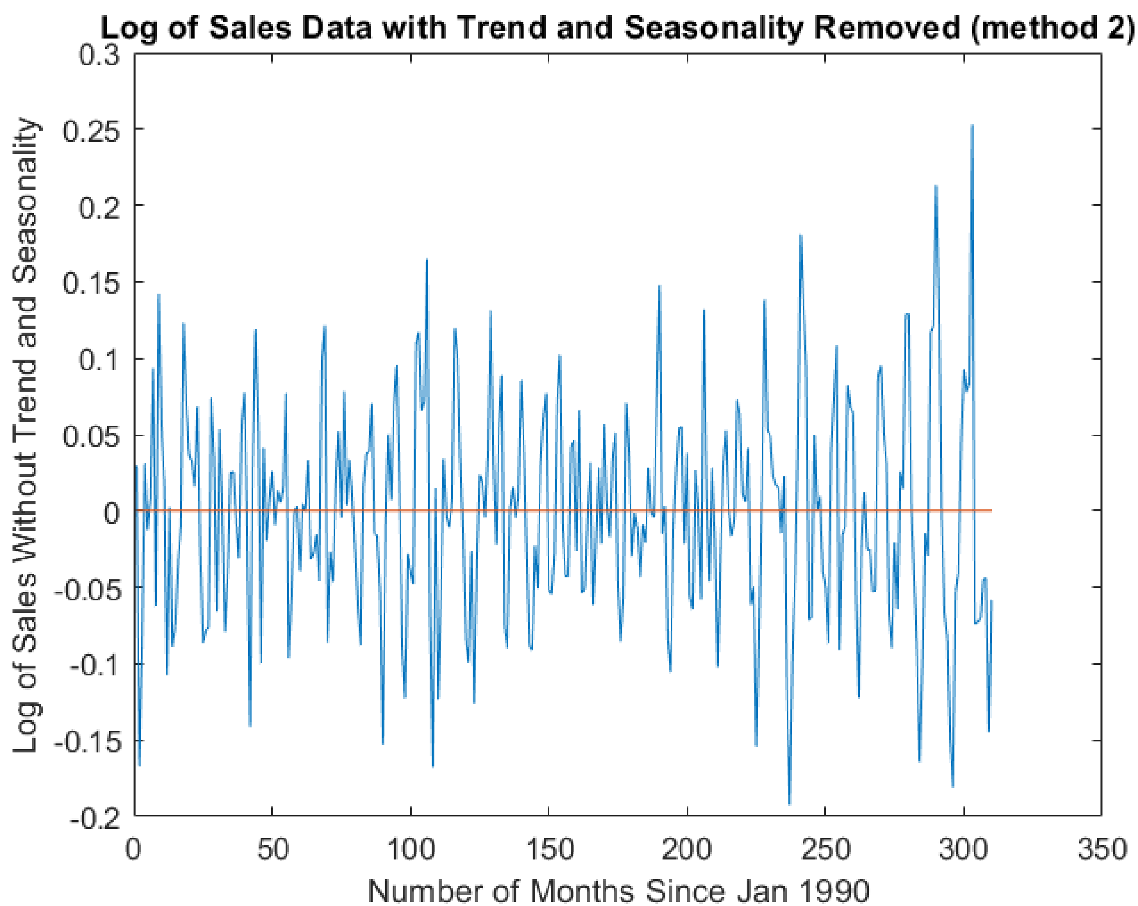
Below is a plot of the data with the trend removed:



Next, we remove the seasonality by using the equation seasonless (i) = trendless (i) – trendless (i – 12). Below is the plot:



This gives us a stationary time series, but does not give us the seasonal adjustments for our table. We therefore try a second method: take the average for each month and consider that to be the seasonal adjustment. Then subtract this average for the corresponding months from the original data to get the seasonless data. This procedure resulted in the following graph:



This appears to be stationary too, but looks a bit more wild compared to the previous method. The seasonal adjustments are:

Month	Seasonal Adjustment	Seasonal Adjustment (shifted so sum = 0)
January	14.1642	0.1550
Feb	14.0404	0.0312
March	13.9188	-0.0904

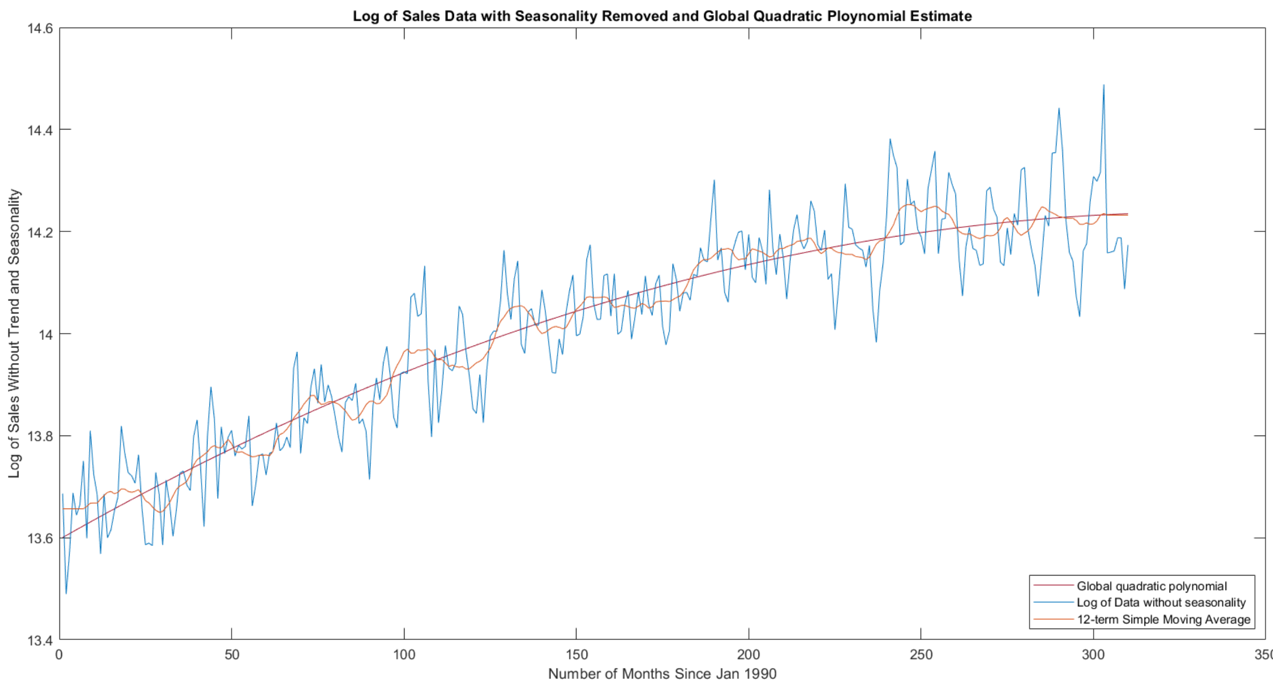
April	13.7361	-0.2731
May	13.7258	-0.2834
June	14.0267	0.0175
July	14.3095	0.3003
August	14.3655	0.3562
Sept	14.2461	0.2369
Oct	13.8902	-0.1190
Nov	13.7249	-0.2843
Dec	13.9622	-0.0470
Sum	168.11	0

Task 6:

We set M to be the nx3 Vandermonde matrix, and then calculate the coefficients of our quadratic using the fact that this matrix is full rank, i.e., coeff = (M'*M)^(-1) *(M'*Y), where M' is the transpose of M, and Y is our deseasonalized time series. The coeffiients we get are

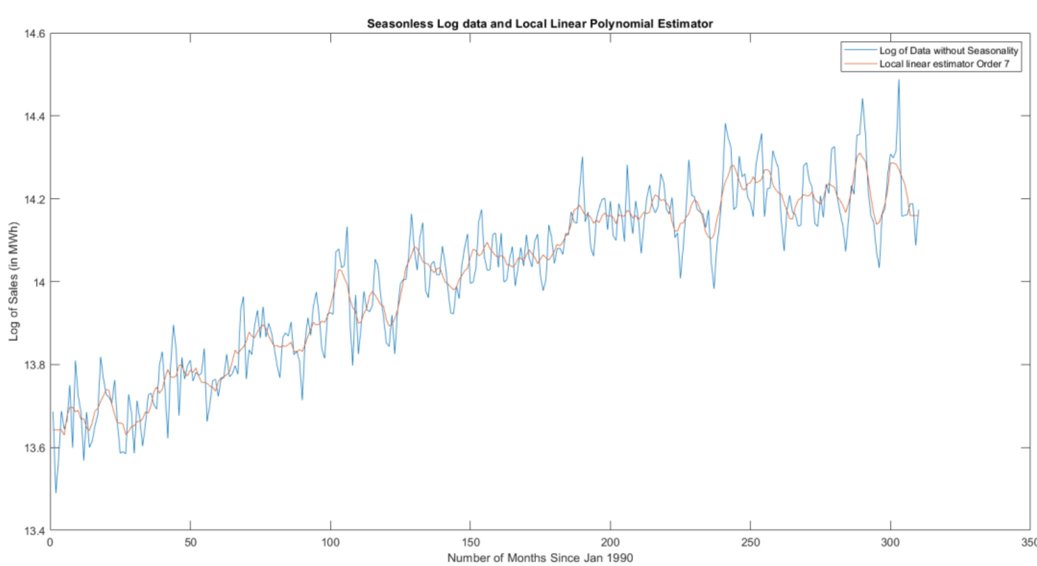
- alpha1 = 13.5966
- alpha2 = 0.0038
- alpha3 = -5.767e-06

We get the plot below:



The quadratic polynomial seems to be a smoothed version of the moving average, as it takes into account all data points, not just the points closest to the one in consideration. The global polynomial is therefore more useful in predicting future values.

Task 7



Farhad de Sousa 545 Data Assignment 1 Feb 27, 2023

We calculate the local polynomial estimator using the formula

$$[\beta_0, \beta_1] = (M^T * M)\text{^}(-1) * (M^T * X)$$

Where X = [Xt-3, Xt-2, …, Xt+2, Xt+3].

M = [1, -3; 1, -2; 1, -1; 1, 0; 1, 1; 1,2; 1,3]

We get $(M^T * M)^{-1} * (M^T) =$

0.1429	0.1429	0.1429	0.1429	0.1429	0.1429	0.1429
-0.1071	-0.0714	-0.0357	0	0.0357	0.0714	0.1071

Therefore we just end up taking a moving average of order 7.

Compared to the moving average of order 12, this one has more variability, and follows the data more closely.