Description

Midterm2020 is a dataset in the form of a CSV provided for the take-home portion of the Spring 2020 Midterm in MSDS 6373 Time Series.

Additional Realization

We do not know the origin of the Midterm2020 dataset other than it was provided for one portion of the midterm. Therefore, we cannot obtain another realization.

Condition 1

The Midterm2020 dataset appears to oscillate with some seasonality. The series trends up slightly, then declines during the last unit, which does appear to represent a year. It seems the series cannot be stationary because there is a level of dependency between the mean and time.

Condition 2

We can not accurately assess the homoscedasticity of the Midterm2020 dataset since the data is dependent on time (and therefore not stationary). Nonetheless, variance seems to be greater earlier in the measurements and smaller later in the measurements.

Condition 3 - The correlation between 𝑋t1 and 𝑋t2 depends only on 𝑡2 − 𝑡1. That is, the correlation between data points depends only on how far apart they are in time, not where they are in time.

Based on the first ACF chart (ACF of midterm2020), there appears to be a strong seasonal component represented in the sinusoidal degradation. Autocorrelation cycles are almost identical across similar volumes of lags, evenly spaced. In analyzing the ACFs of the first and second halves of the series, the autocorrelations seem to mirror themselves (when comparing the first half to the second half). Therefore, the data seems dependent on position in time, not just on the distance between each pair of points.

Conclusion

Because the three conditions of a stationary time series cannot be confirmed, we must conclude that this is not a stationary time series and that there is a dependency on time driving the position of each successive data point.