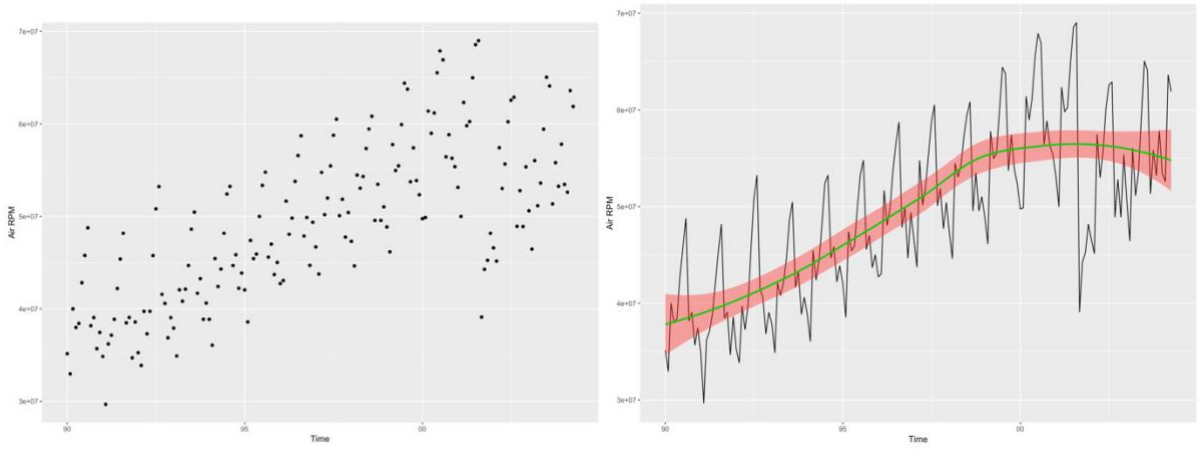
# DSC 275/475: Time Series Analysis and Forecasting (Fall 2019)
# HW #1
# Chunlei Zhou

1. **The Bureau of Transportation Statistics (BTS) conducted a study to evaluate the impact of Sept 11 attacks (9/11) on U.S. air transportation. The purpose of this study is to provide a greater understanding of the passenger travel behavior patterns of persons travelling by air before and after the event.**

a) **Is the goal of this study descriptive or predictive?**

The goal of this study is descriptive.

b) **Create a time series plot of the data, i.e. a plot $y_t$ versus *t, where t=1,2,3 …* What would t=1, 2, 3 refer to in the time series? Which time period does t=1 refer to?**




t=1, 2, 3 refer to three continuous time periods from 1990-01-01 to 1990-01-31, from 1990-02-01 to 1990-02-28, and from 1990-03-01 to 1990-03-31 respectively.

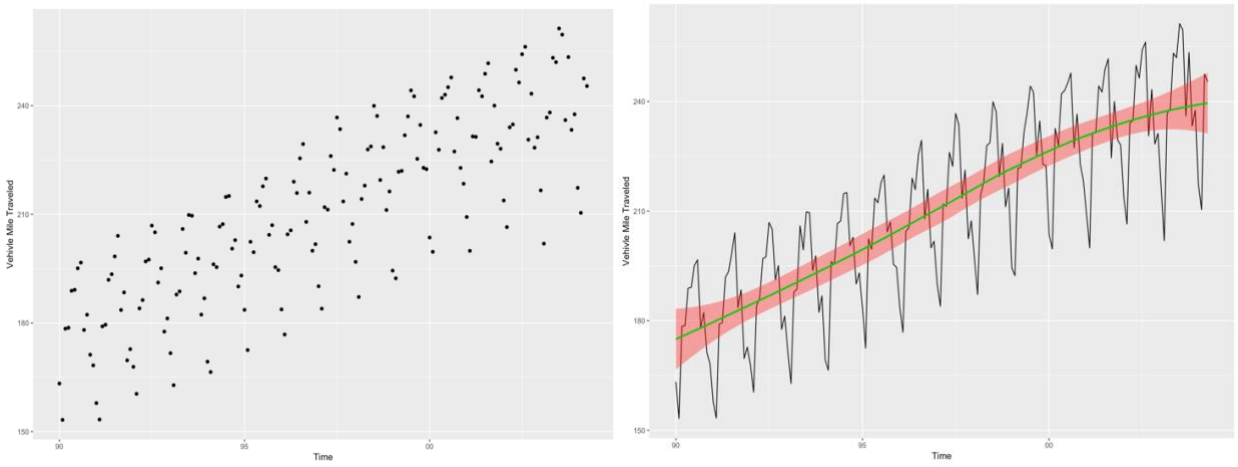
c) **What are the values for $y_1$, $y_2$ and $y_3$ in the time series?**

The values for y in different time periods is the revenue passenger mile contributed by passengers pay to travel by air. Air RPM = number of paying passenger * Miles per flight.

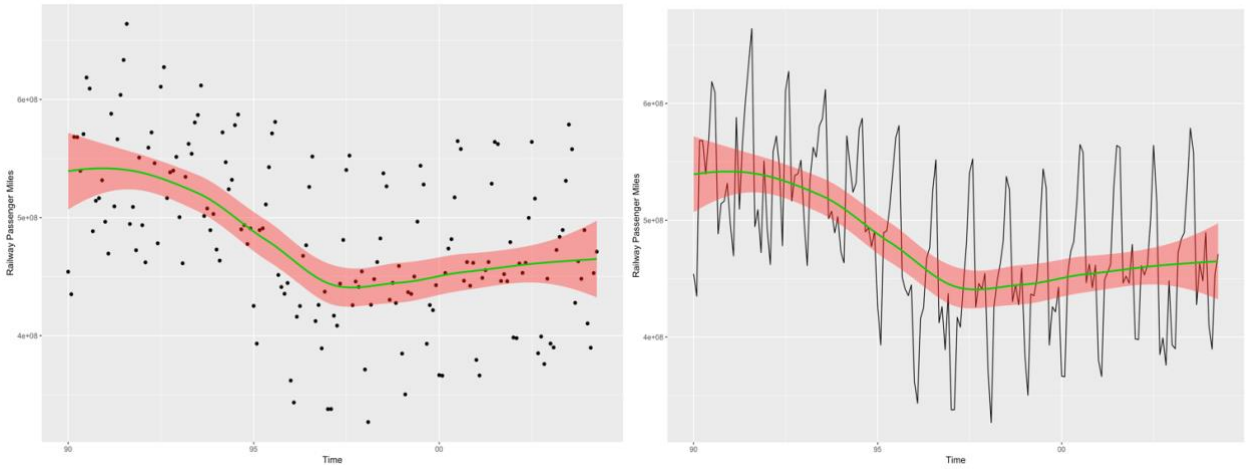
2. **In addition to air travel data, two additional time series are also provided in the same data file – Rail and Vehicle travel.**

a) **Which of these components appear in the *Air* and *Vehicle* time series: i) Level; ii) Seasonality; iii) Trend; iv) Noise. List for each data set.**

Based on the plot in question 1 b), level, seasonality, trend, and noise appear in the Air. Based on the plot below, level, seasonality, trend, and noise appear in the vehicle.



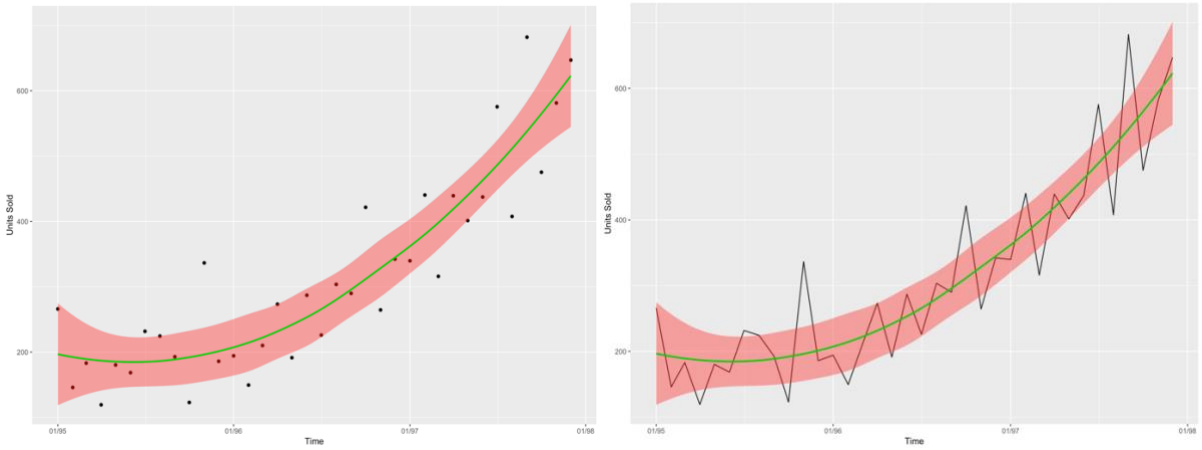

**b) For the Rail data set, describe the trend, i.e. how does the trend vary across the time series?**




Total miles traveled by railway decreased fast from 1990 to 1997. It increased very slowly and slightly from 1997 to 2004. The second half is relatively flatter than the first half.

**3. Forecasting Shampoo Sales: monthly sales of a certain shampoo over a 3-year period.**

**a) Create a time series plot of the data. Label the axes, and units.**




**b) Which of the four components (level, trend, seasonality, noise) are present in this series?**

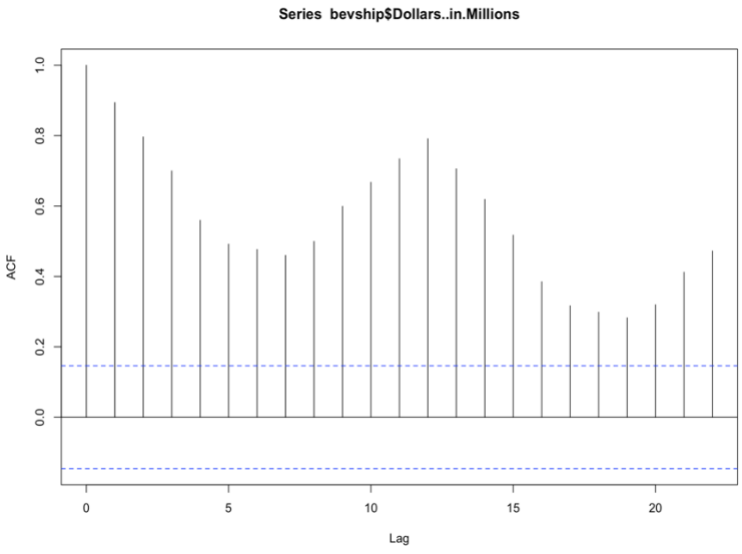
Trend, level, and noise present in this series.

**c) Do you expect to see seasonality in sales of shampoo?**

No, I don't, because the demand of shampoo does not depend on season. Almost everyone needs to take a shower or clean his/her hair in a constant frequency.

**4. The file, *Beverages_Shipment_2019.xlsx*, contains the US beverage product shipments data.**

a) Find the sample autocorrelation function for this data set.



**b) Is the time series stationary or nonstationary?**

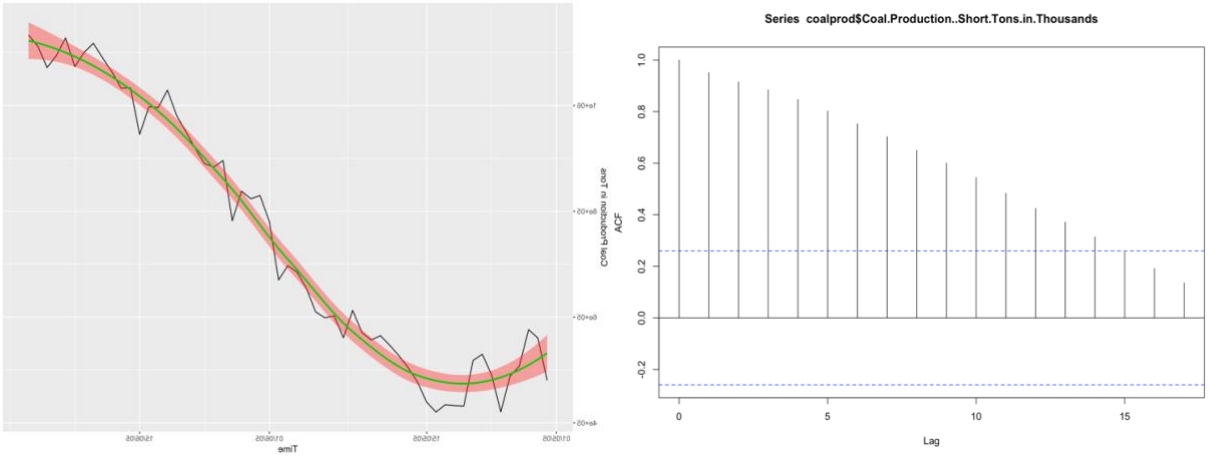
It is nonstationary.

**c) Comment on the seasonality of the time series.**

The time series is season. One season is about 12 time periods.

**5. Data on US coal production are given in *Coal_Production_US.xls*.**

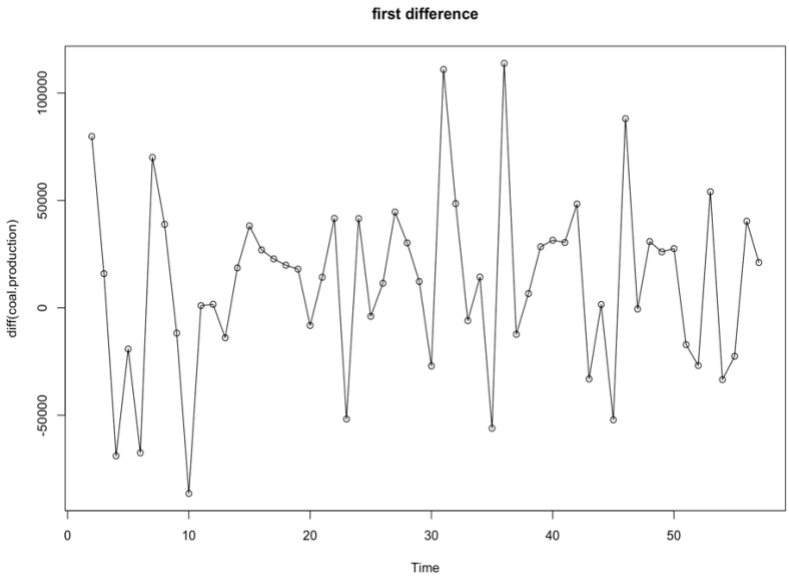
**a) Plot the coal production data and the sample autocorrelation function.**




**b) Is the time series stationary or non-stationary?**

It is non-stationary.

**c) Plot the first difference of the time series and compute the sample autocorrelation function of the first differences.**



sample autocorrelation function of the first differences, by lag

```
Autocorrelations of series ‘coal.production’, by lag

    0     1     2     3     4     5     6     7     8     9    10    11
1.000 0.952 0.915 0.884 0.847 0.803 0.752 0.703 0.650 0.600 0.544 0.483
   12    13    14    15    16    17
0.424 0.371 0.314 0.256 0.192 0.136
```

**d) What impact has differencing had on the time series?**

It removes the trend of the time series.