**DSC 275/475: Time Series Analysis and Forecasting (Fall 2019)**

**HW #1**

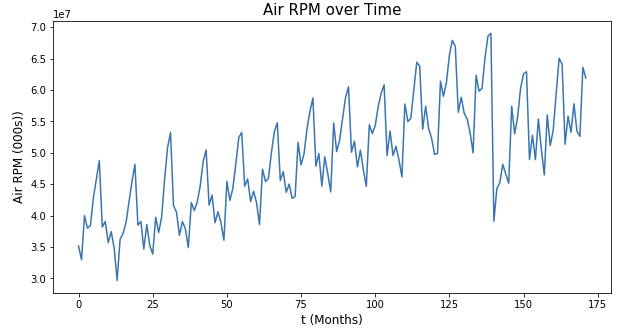
1. The Bureau of Transportation Statistics (BTS) conducted a study to evaluate the impact of Sept 11 attacks (9/11) on U.S. air transportation. The purpose of this study is to provide a greater understanding of the passenger travel behavior patterns of persons travelling by air before and after the event. In order to assess the impact of September 11, BTS took the following approach: Using data before September 11, it forecasted future data (under the assumption of no terrorist attack). Then, BTS compared the forecasted series with the actual data to assess the impact of the event.

The data is available in the file: *BTS\_Air\_Rail\_Vehicle\_Miles.xls*.

1. Is the goal of this study descriptive or predictive?

**Predictive**

1. Create a time series plot of the data, i.e. a plot yt 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 the first, second, and third month starting from 1990-01-01.**

**t = 1 refers to the time period from 1990-01-01 to 1990-02-01**

1. What are the values for y1, y­2 and y3 in the time series?

**35153577, 32965187, 39993913**

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

**Air: Seasonality + Trend + Noise**

**Vehicle: Seasonality + Trend**

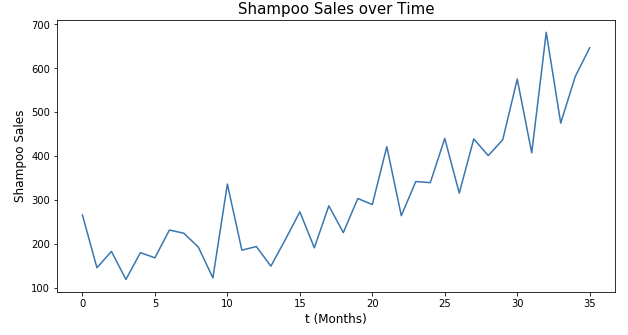
1. For the Rail data set, describe the trend, i.e. how does the trend vary across the time series?

**The rail pm decreases from t = 0 to roughly around t = 85, and start increasing afterwards**

1. ­Forecasting Shampoo Sales: The file *ShampooSales.xls* contains data on the monthly

sales of a certain shampoo over a 3 year period.

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



1. Which of the four components (level, trend, seasonality, noise) are present in this

series?

**Trend + Noise**

1. Do you expect to see seasonality in sales of shampoo?

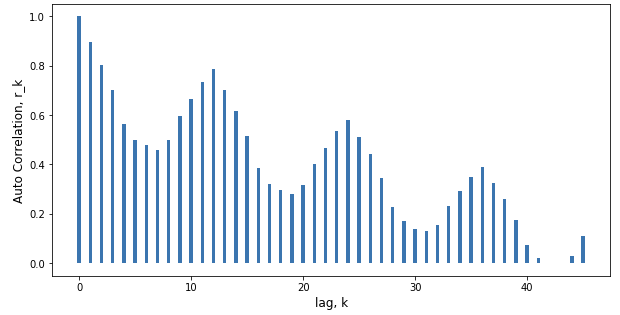
**Maybe, it seems to have similar cycle every 10 months**

1. The file, *Beverages\_Shipment\_2019.xlsx*, contains the US beverage product shipments data.
2. Find the sample autocorrelation function for this data set.

**[1.0, 0.8956332157588593, 0.8038648003815599, 0.7032492771028038, 0.5657227424115161, 0.4984403056090543, 0.47715372925402666, 0.4600033246976522, 0.499006858672507, 0.5970544150956527, 0.6662173124751137, 0.7336315475414809, 0.7876450914208581, 0.7017275345101591, 0.6174834802605921, 0.513363217336216, 0.38592596814601415, 0.32000739855701466, 0.29441875043022336, 0.28215168304349286, 0.31588261080792773, 0.40365546376252376, 0.46702748663277927, 0.533622334387032, 0.5820431880870733, 0.5119904096114375, 0.4412014793958346, 0.3438611388258802, 0.22933840555772286, 0.17107000090424815, 0.1398004215459315, 0.12943130808254413, 0.15656378253010417, 0.2324813705891291, 0.2907975351344827, 0.3503580878803203, 0.38907680228010244, 0.3250875310384353, 0.25979882110732694, 0.17344187559679336, 0.07305950786644617, 0.02013606285045337, 0.0013859916701638982,**

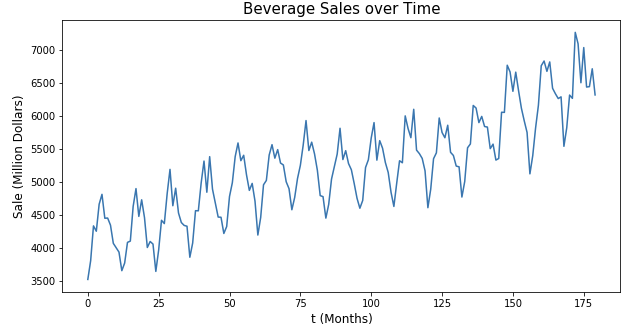
**-0.0011792319747295126, 0.029852497334172865, 0.10946774355892162]**

1. Is the time series stationary or nonstationary?



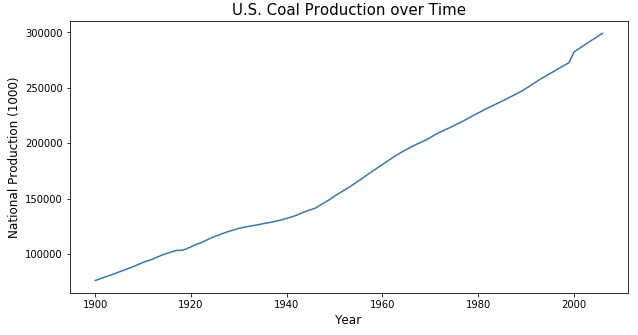
**The sample ACF is very persistent and also shows strong auto correlation even ta large lag values, which are the characteristic of a nonstationary time series**

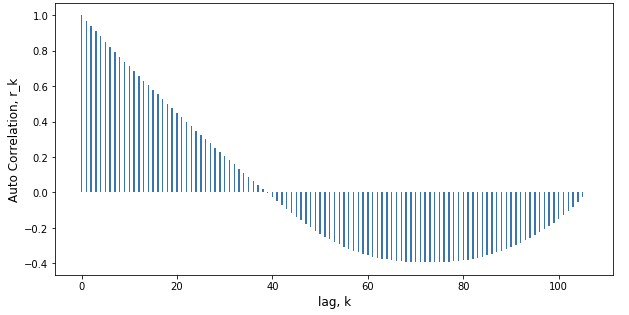
1. Comment on the seasonality of the time series.



**The beverage sale tends to have a cycle for every 24 months (1 year) and the overall trend is increasing**

1. Data on US coal production are given in *Coal\_Production\_US.xls*.
2. Plot the coal production data and the sample autocorrelation function.

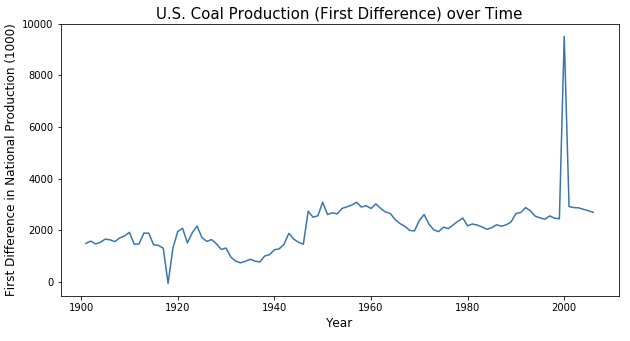


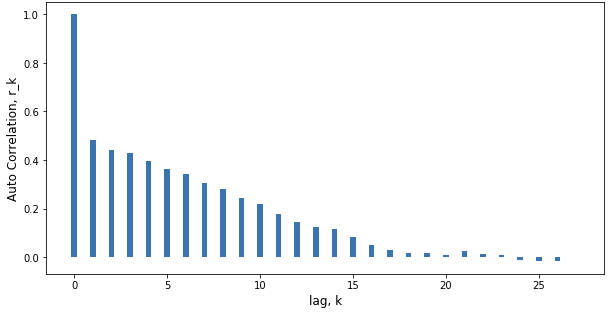


1. Is the time series stationary or non-stationary ?

**This time series is non-stationary**

1. Plot the first difference of the time series and compute the sample autocorrelation function of the first differences.





1. What impact has differencing had on the time series?

**After taking the first difference, the time series becomes stationary**