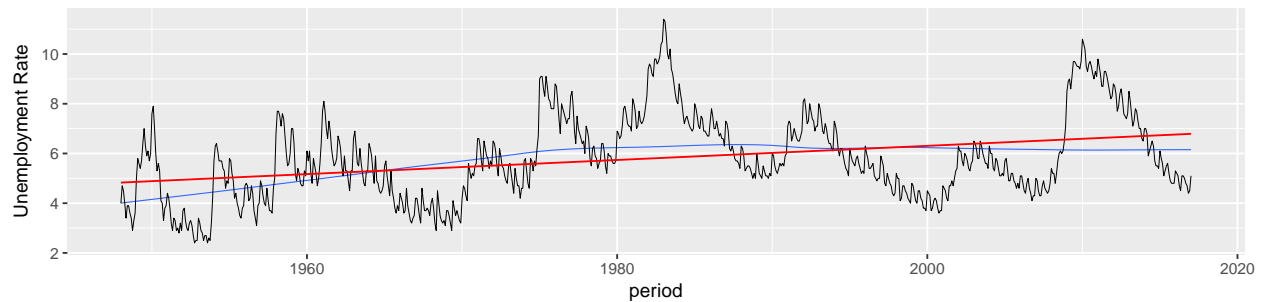


Project - 2

Dwipam, Krunal, Rakshesh

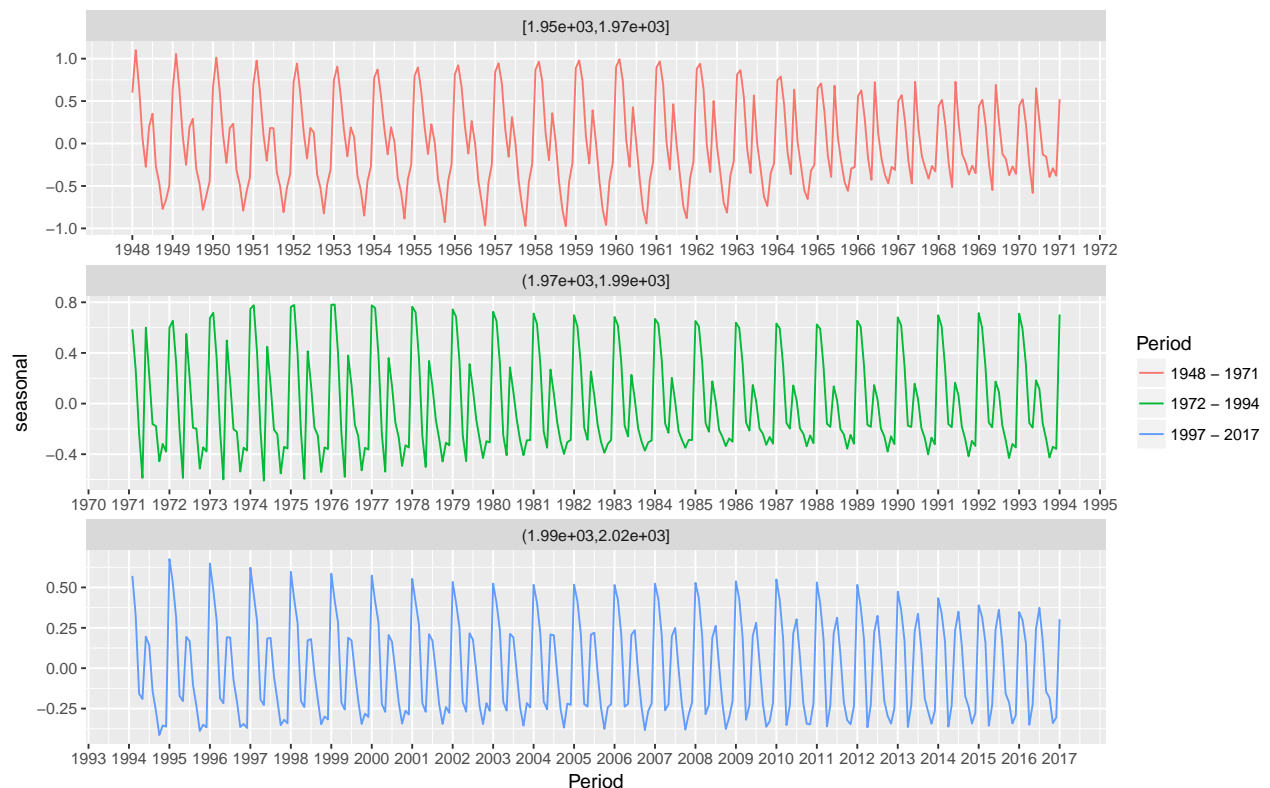
2/19/2017

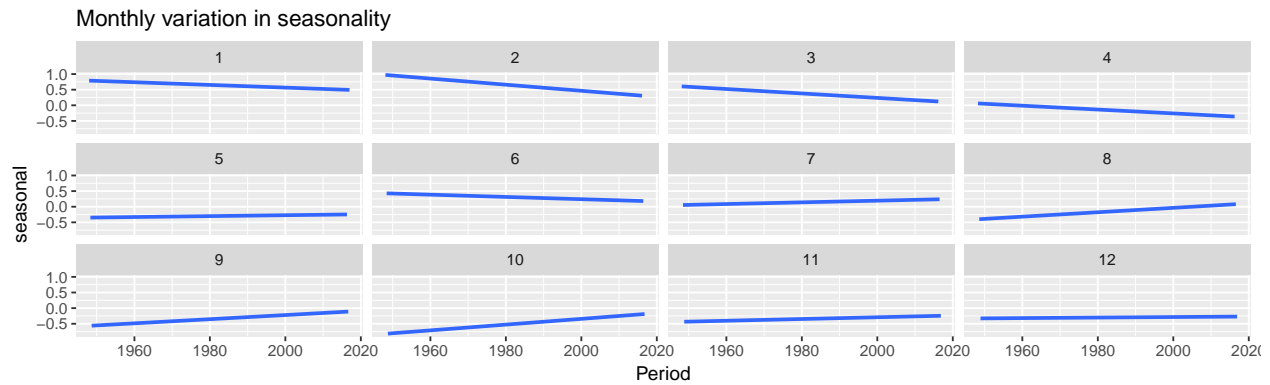
Question 1:



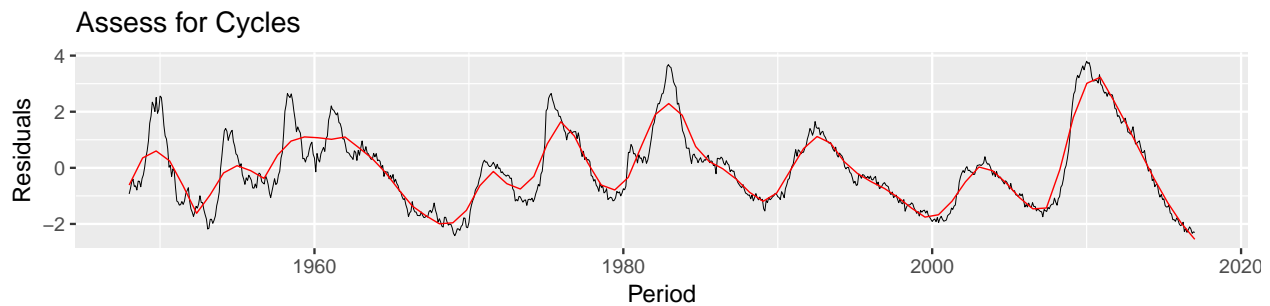
Clearly there seems to be some trend and some seasonality. Red line is the Linear fit, and blue line is the loess fit with degree = 1. It seems that there is an increasing trend over period of time for Linear fit. For Loess fit there was a small spike of unemployment during year 1975 - 1995. That's why there is a hump in the curve for Loess fit, but still comparing right and left tail, there is an increase in unemployment rate over all. This hump seems to be an irregularity over the period of time, we will stick with Linear model as it generalizes better. There seems to be a season of 12 months, where unemployment rate starts with low gradually increases and falls down for the next month and patterns changing every 3-4 months. However we don't know if there is seasonality. Let's fit trend and assess the residual to decompose the Time series into other components.

Seasonal plot per 23 months

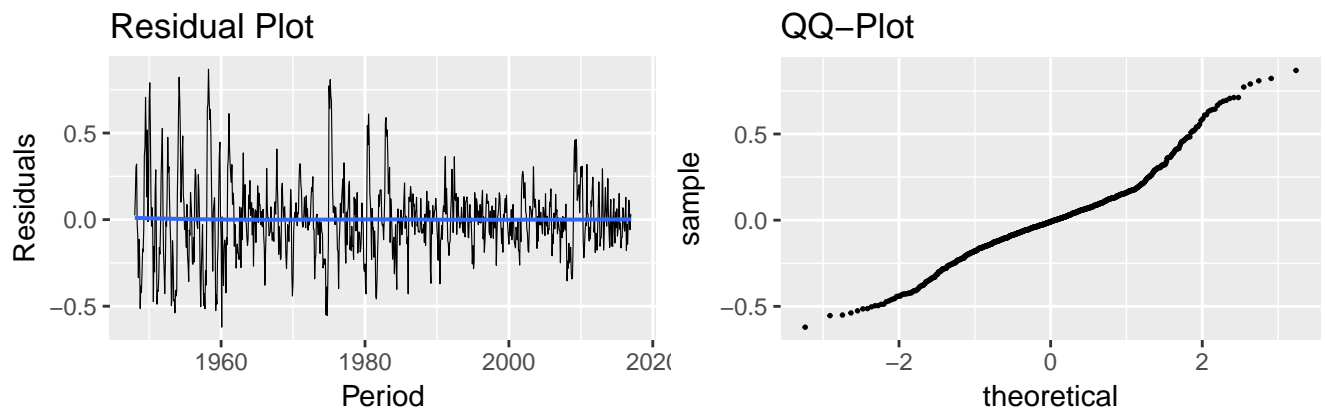




It seems that as the year increases unemployment rate during the second half of the around June increases. From 1995 to year 2000 there is repetitive pattern. For year 2014 - 2017 Both half of the year has equal highest unemployment rate. Looking at the monthly seasonal plot, it seems that there is decrease in unemployment rate over the period of 12 months. The slope of the line is decreasing from Month Jan - April while for months July, August, September, October increases and for november and december being somewhat normal. This also alligns with our pervious statement that unemployment rate is higher in first half of the year i.e. Jan - May. Now Let's assess the cycles and residuals from the model after fitting Trend and remainder.



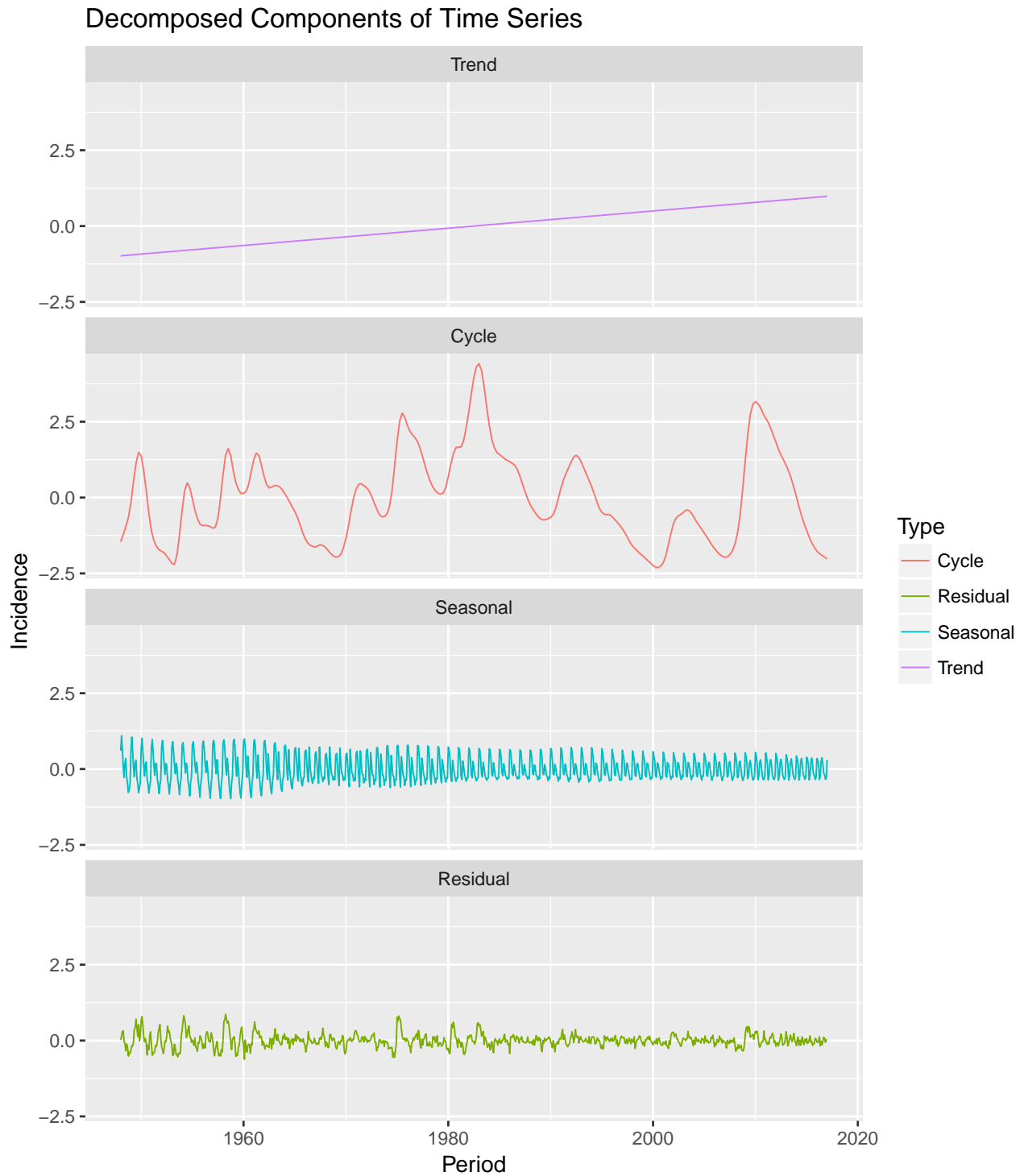
After removing, trend and seasonal component there seems to be 8-9 cycles, STL has already decomposed our time series for Cycles, hence lets just assess residuals after removing trend, seasonal and cycles.



Variance for each Time series components.

	Trend	Cycles	Seasonal	Residuals
1	0.3213793	2.140231	0.172514	0.04762451

Though, there is no visual pattern found in the residuals, the qqplot has small bends and hence it does not seems to be normal. Also the loess curve passes through 0.0 implying 0 mean. Cycles capture highest variance in the data. Lets plot Trend, Season, Cycle and residuals together.



Visually also it seems that, Cycles captures highest variance followed by Trend and then Seasonal. Hence our model is

$$y = \text{Trend} + \text{Seasonal} + \text{Cycle} + \text{Residual}(0, \text{variance})$$