Time Series Analysis on U.S. Citizen Air Travel to Canada

Leanne Lee

Introducion

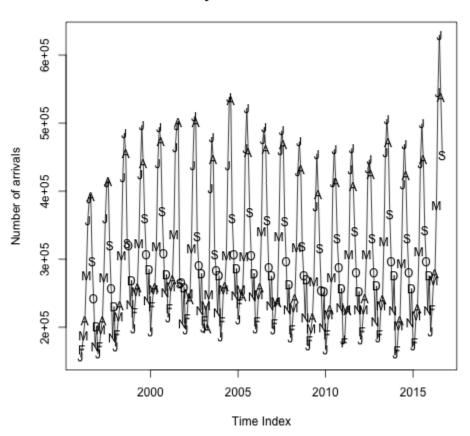
The scientific question motivating my work is "is there a seasonal trend for U.S.citizens to travel to Canada by air?" Canada is one of the closest country across the border of the United States. Many U.S. citizen consider Canada a great spot to travel during holiday. In the West Coast, we neighbor near Vancouver and Victoria, while the east coast has Toronto, Montreal and Quebec.

Dataset

The dataset comes from the U.S. Department of Commerce, National Travel & Tourism Office. [NTTO Monthly Tourism Statistics] (http://travel.trade.gov/research/monthly/departures/index.html). The dataset contain U.S. outbound travel by world regions from Jan 1996 to Sep 2016. There are 249 monthly data points throughout these twenty years.

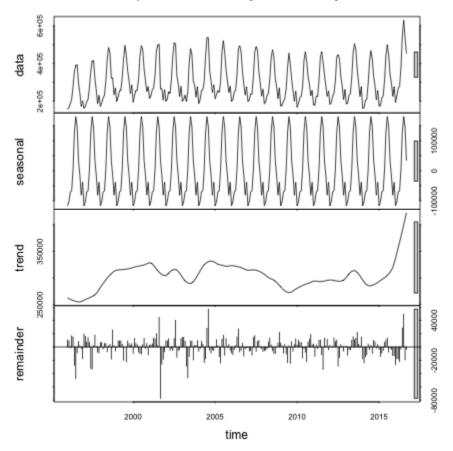
Exploratory Data Analysis

Monthly Air arrivals to Canada



First, I plot the time series for the monthly air arrivals from U.S. to Canada. By examining the time series, the plot looks seasonal with high peak in June, July and August. Students usually travel during their summer break and this is the main drive of the increase in air travel in these months. There is also a smaller peak in between of each cycle. The small peaks take place in December, which is during Christmas break and students tend to go travel more often. Because the graph doesn't get effect when the variation increase with the level of the series, so I didn't consider the log of the time series.

Seasonal Decomposition of Time Series by Loess for monthly Canada Travels



After the basic inspection on the time series, I ran the stl() command for a season decomposition of Time Series by Loess. I first inspect the data section. There is no sign of non-stationary, so I did not consider to take differencing. I consider the bar on the right hand side as one unit of variation. The bar on the seasonal panel is slightly bigger than the data panel, meaning the seasonal effect is larger than the variation in the data. In the trend panel, it has a larger variation box than the data and seasonal panels, meaning the variation attributed to the trend is much smaller than the seasonal component. It also indicated the trend is not dominating, but there's an upward trend at the end.