ADS 506: Monthly Amtrak Ridership Forecasting Final Project

Team 3: Jimmy Nguyen, Aubrey Barrett, Luke Awino

11/15/2021

Contents

Libraries

```
library(astsa)
library(zoo)
library(xts)
library(pander)
library(tidyverse)
library(lubridate)
```

Data Set

Code:

```
# Load the data set from CSV file
df <- read.csv("/Coding//Monthly-Amtrak-Ridership-Forecasting/Data/Amtrak Ridership Data.csv")

# Rename columns
names(df)[1] <- 'Dates'
names(df[2]) <- 'Number_of_Passengers'


# First 12 months in 1991
head(df, n = 12) %>%
  pander(style = "grid", caption = "First 12 Months - 1991")
```

Table 1: First 12 Months - 1991

Dates	Number.of.Passengers
Jan-91	1708917
Feb-91	1620586
Mar-91	1972715
Apr-91	1811665
May-91	1974964
Jun-91	1862356
Jul-91	1939860
Aug-91	2013264
Sep-91	1595657
Oct-91	1724924
Nov-91	1675667
Dec-91	1813863

```r
# convert to time series object
df<- ts(data = df[,2], start = c(1991,1),
        end = c(2013,5), frequency = 12)

print("Starting Year and Month: ")
```

```
## [1] "Starting Year and Month: "
```

```r
start(df)
```

```
## [1] 1991    1
```

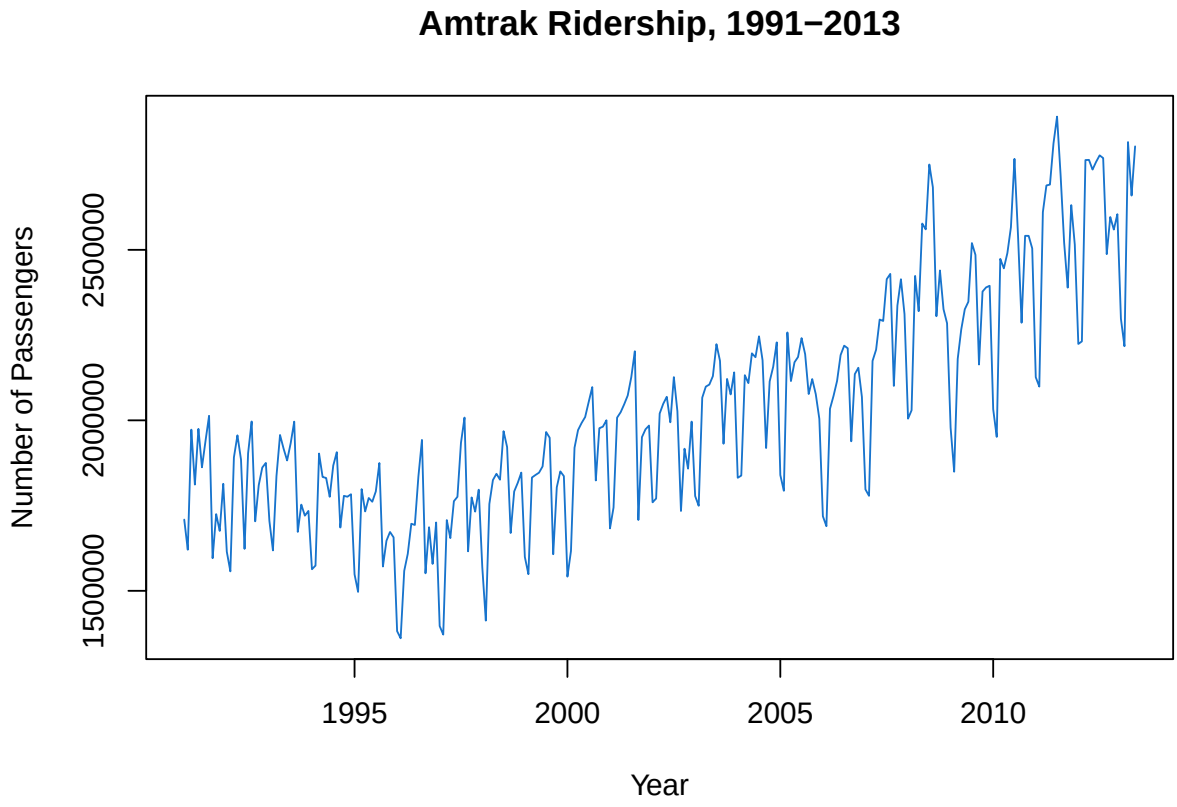
```r
print("Final Year and Month: ")
```

```
## [1] "Final Year and Month: "
```

```r
end(df)
```

```
## [1] 2013    5
```

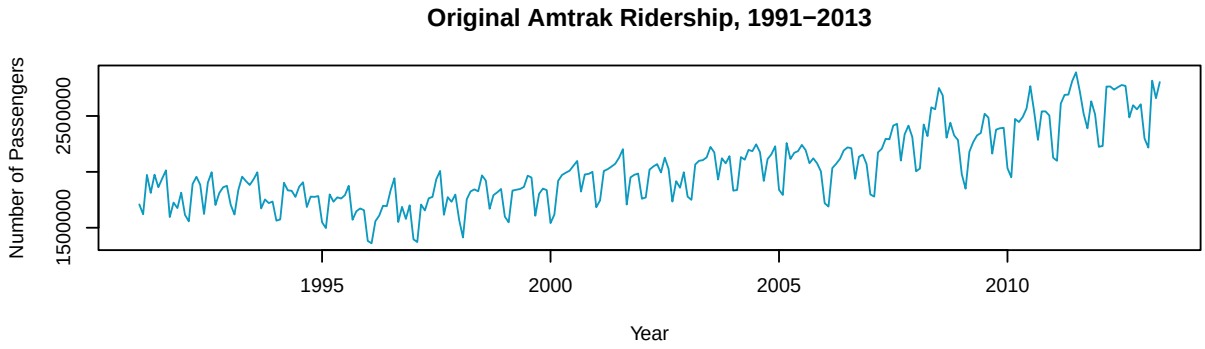
```r
# Make a quick time-series plot
ts.plot(df, xlab = "Year", ylab = "Number of Passengers",
     main = "Amtrak Ridership, 1991-2013", col = 4)
```

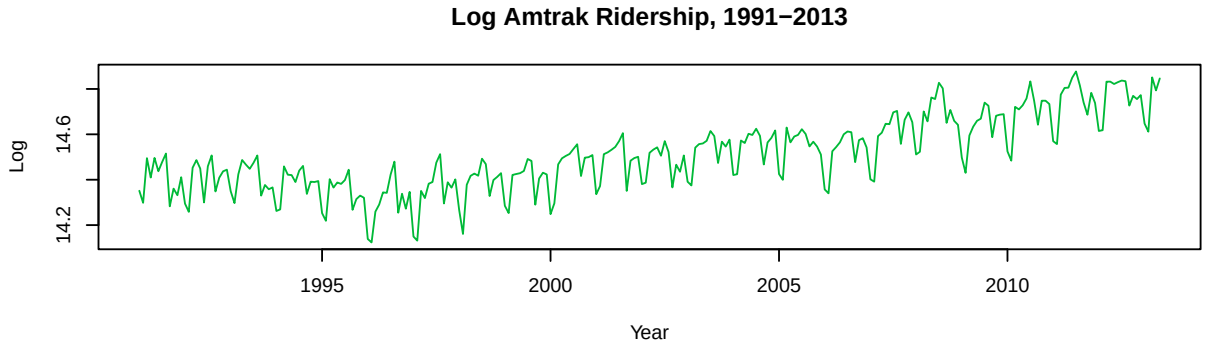


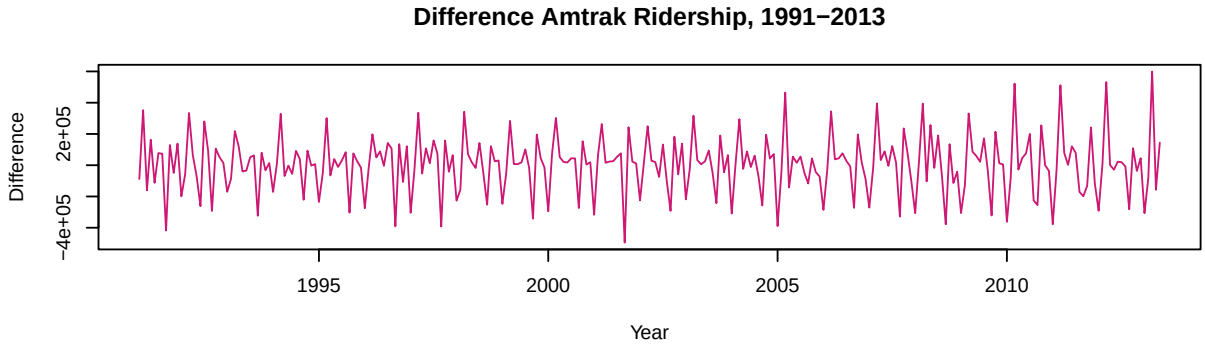
Data Transformations

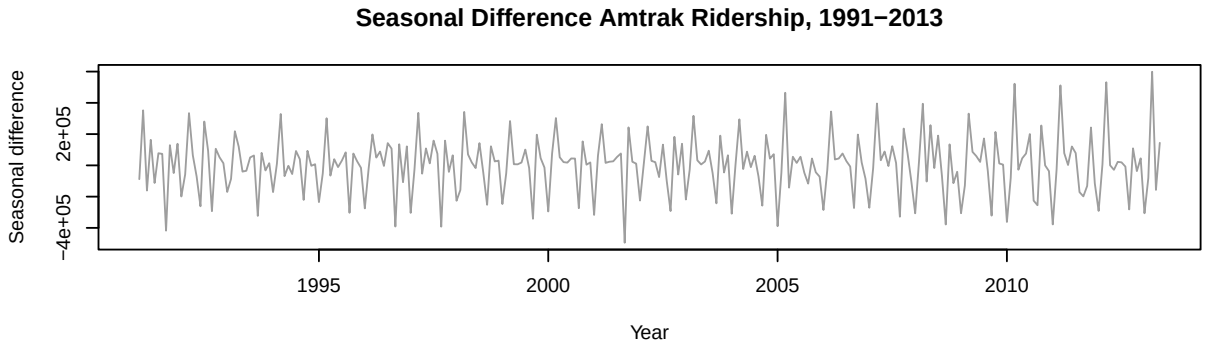
- We can use these transformations to remove or filter any possible trends in the data
- E.g., log transformations or difference, etc.

```
par(mfrow=c(4,1))

# Original
ts.plot(df, xlab = "Year", ylab = "Number of Passengers",
     main = "Original Amtrak Ridership, 1991-2013", col = 5)

# log
ts.plot(log(df), xlab = "Year", ylab = "Log",
     main = "Log Amtrak Ridership, 1991-2013", col = 3)

# difference
ts.plot(diff(df), xlab = "Year", ylab = "Difference",
     main = "Difference Amtrak Ridership, 1991-2013", col = 6)


# seasonal differencing
ts.plot(diff(df, s = 4), xlab = "Year", ylab = "Seasonal difference",
     main = "Seasonal Difference Amtrak Ridership, 1991-2013", col = 8)
```









Analysis:

- The log transformation can linearize a rapid growth trend and can also stabilize a series that exhibits

increasing variance. Although the main restriction is that the log transformation is only defined for positively valued time series.

- Difference transformations can remove linear trend.
- For time series with seasonal trends, seasonal differencing can remove those periodic patterns.
- For monthly or quarterly data, an appropriate value of s would be 12 or 4. Although there wasn't a difference when including the seasonal term in the bottom graph.