Assignment – 11

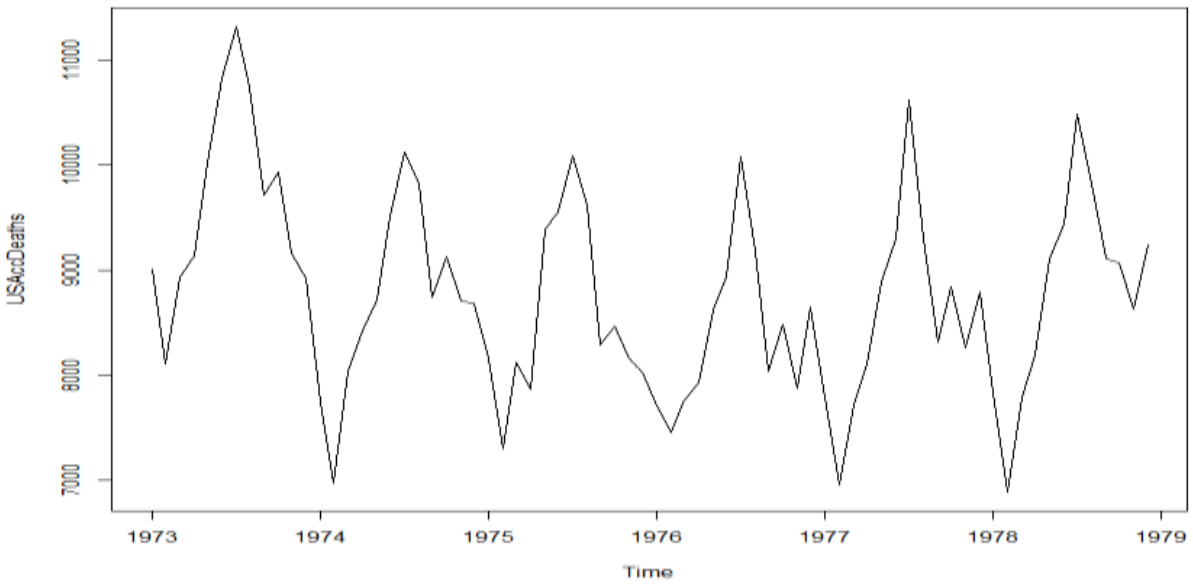
Data: USAccDeaths

Description: Monthly totals of accidental deaths in the USA

Data Type: Time series.

Time Interval: 1973 – 1978

Plot: Frequency of monthly accidents.

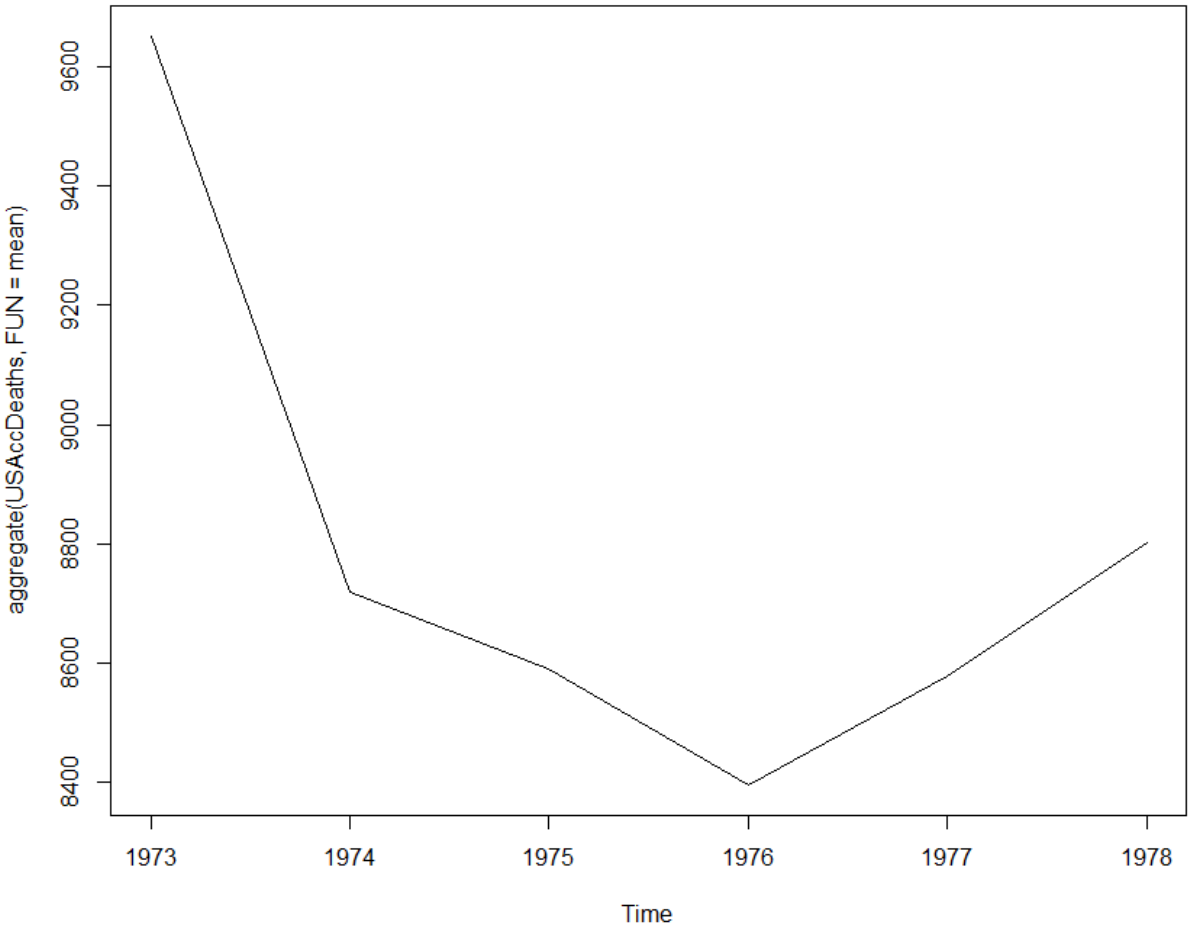


Plot Description: Figure shows Frequency of accidental deaths in corresponding months from Jan-Dec from Year, 1973-1978. We can see the trends where the number of accidents goes up primarily in month of June and gradually decrease as it reaches December.

We can also see that the numbers of accidents are high in the year 1973 reduced in 1974 and again is increasing after 1976.

We can see this clearly in the next Plot.

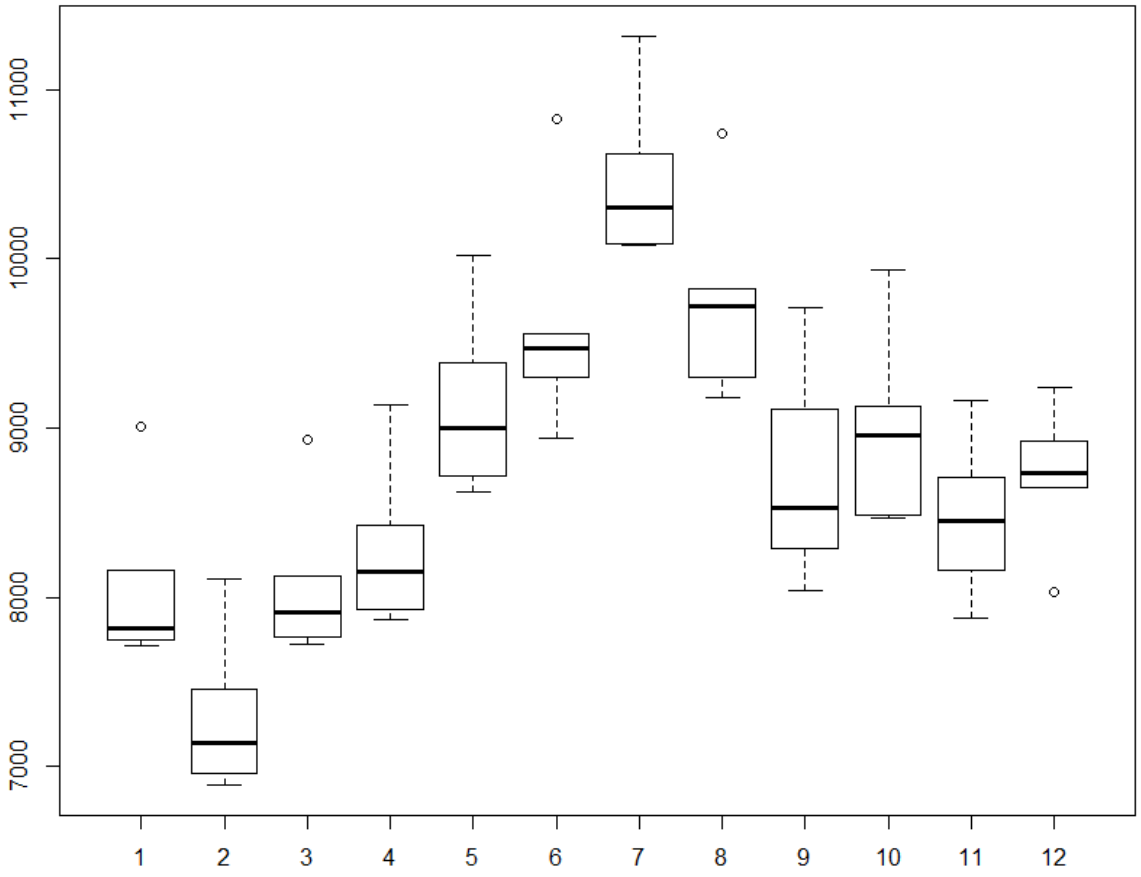
Plot: Average Deaths every Year



Plot Description: We use average deaths every year to plot this graph. From the graph it is clear that the number of accidental deaths has reduced considerably from 1973 to 1976 and after 1976 the average numbers of deaths are again increasing.

If we look into the past, 1973 is the time USA was hit by recession; people lost their jobs. People using vehicles reduced as the fuel prices increases and less people could afford vehicles. As the number of people driving vehicles decreased number of accidents decreased. After 1975 there was a recession recovery period for USA. And after 1976 again people started using vehicles leading to more accidents. And this graph will go on increasing as the number of people using vehicle is increasing and so will the accidents in proportion.

Plot : Box Plot of accidents every month

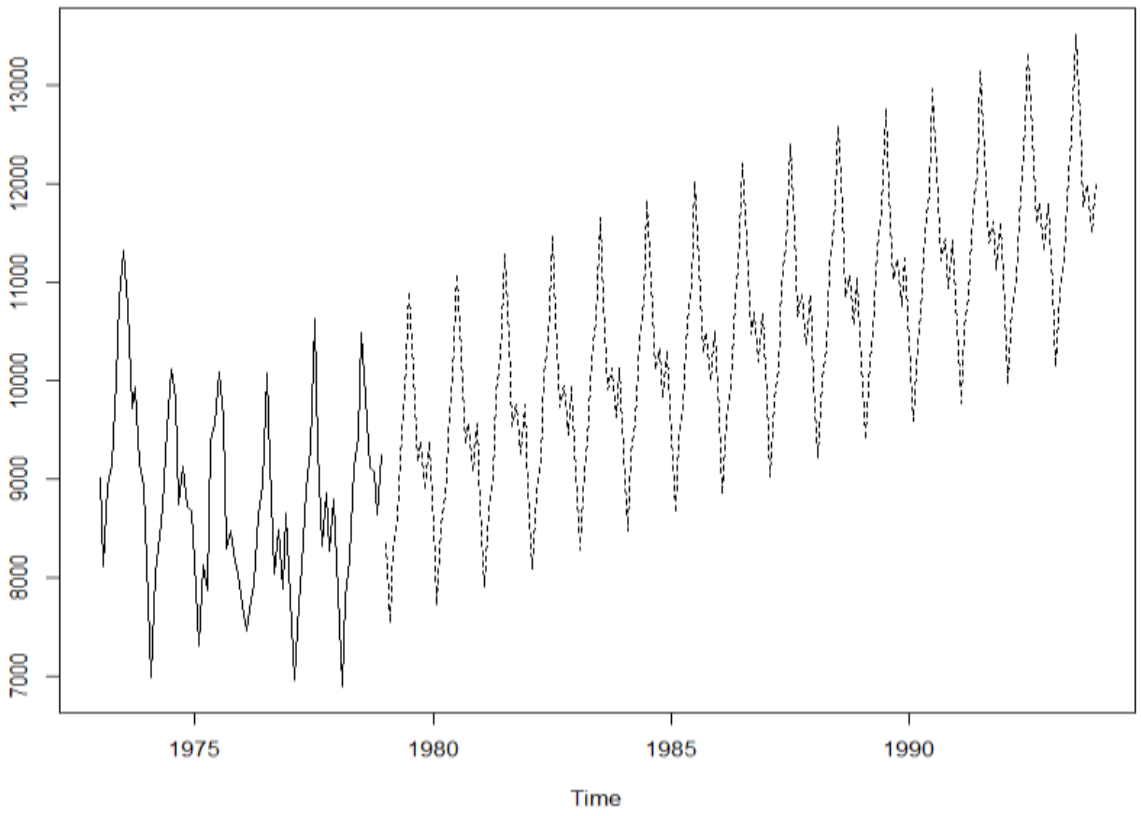


Plot Description: Figure above shows the box plot which shows minimum, maximum and median number of deaths in each month.

We can clearly see that the numbers of deaths are more in the months of June, July and August. **This time is summer when many people get out of their house, travel and hence increasing the number of vehicles on the road and eventually increasing the number of accidents.**

→From the data we can say that the number of accidents will go on increasing every year as the number of people driving will go on increasing and it will be maximum during summer time using similar hypothesis.

Plot: Time Series Plot of next 15 Years from 1978



Plot Description: Figure above shows the trend in which the number of Death will vary corresponding to e previous year. As we predicted that the numbers of Deaths will increase using simple observation, ARIMA model predicted the same using previous knowledge.

ARIMA Model: (0, 1, 1)

Description: As we can see from the second plot of Average deaths in every Year that the mean is not standardized, therefore we use MA and I to standardize it perform analysis and then predict. We took AR=0 as the output is not dependent on the weighted sum of past. If taken 1 it will try to reduce the output (reducing error) which is not desired here.

```
Coefficients:
          ma1     sma1
      -0.4303  -0.5528
s.e.   0.1228   0.1784

sigma^2 estimated as 99347:  log likelihood = -425.44,  aic = 856.88
```

Code :

```
data("USAccDeaths")
View(USAccDeaths)
class(USAccDeaths)
plot(USAccDeaths)

abline(reg=lm(USAccDeaths~time(USAccDeaths)))
cycle(USAccDeaths)


plot(aggregate(USAccDeaths, FUN=mean))
boxplot(USAccDeaths~cycle(USAccDeaths))

(fit = arima(USAccDeaths, c(0,1,1),seasonal=list(order=c(0,1,1), period=12)))
pred=predict(fit,n.ahead = 15*12)
ts.plot(USAccDeaths,pred$pred, lty=c(1,2))
```