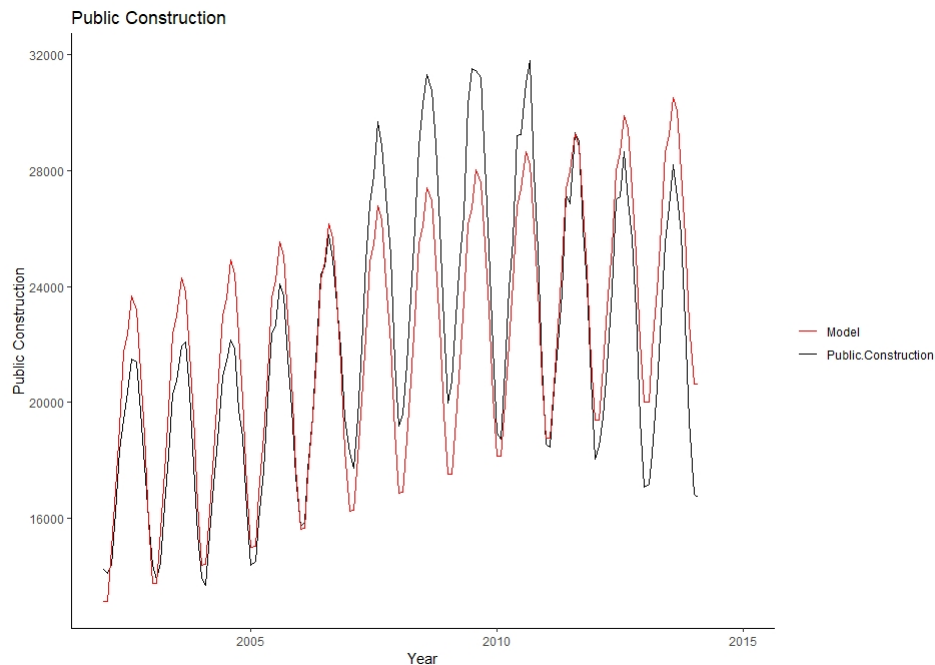


1 Public Construction Linear model with 12 month average lag

For some reason my R Markdown is very buggy, so I just took the code and pasted it below with the plot at the beginning.



```
library(tidyverse)
```

```
#read in the data
myData<-read.csv('ConstructionTimeSeriesDataV2.csv',
  red↪ stringsAsFactors=FALSE)
#change the format so we can use the dates
myData$Month.Year<-paste0("1-",myData$Month.Year)
myData['Month.Year']<-dmy(myData$Month.Year)
#look at the data
plot(myData$Month.Year,c(myData$Public.Construction), type="l",
  red↪ ylim=c(0, 85000))
lines(myData$Month.Year,myData$Private.Construction)
#Lets make a model. First one is a regress
#least squares regression
lmPublic<- lm(Public.Construction~Month, data = myData)
model<-c(Intercept,Slope)
lmPublic
intercept<-lmPublic$coefficients[1]
slope<-lmPublic$coefficients[2]
```

```
#now we have the model - put the data through it
myData$yhat<-18275.93+52.01*myData$Month
#seperate the seasonality difference
myData$R1<-myData$Public.Construction-myData$yhat
lines(myData$Month.Year,myData$yhat)
lines(myData$Month.Year,myData$R1)
#Now need to account for seasonality - lets use the 12 month lag
  red↵→ average
#best lag has the highest correlation

#I explored the lag even though it seems to be a 12 month has
  red↵→ highest correlation
#I cheated and used the lag() function in R
laggedTS <- data.frame(matrix(ncol = 24, nrow = 146))
for (i in 1:24){
  laggedTS[,i]<-lag(myData$R1,i)
}

# determine the best correlation (its at 12)
myCorrelation<-c(1:24)
for (i in 1:24){
myCorrelation[i]<-cor(myData$Public.Construction[(i+1):146],
  red↵→ laggedTS[(i+1):146,i])
}

#Get the average value of the data at every 12th Month.
#I used the mod() function to get months 1-12. Then I did some
  red↵→ DPLYR work to get the mean for each month
myData$Mod<-myData$Month%%12
averages<-myData %>% group_by(Mod) %>% summarize(avg=mean(R1))
averages<-rep(averages$avg,20)
averages<-averages[-1] #remove the first value as it is a
  red↵→ december
averages<-averages[1:146] #remove the extra data from the tail
myData<-myData[,1:6]
myData<-mutate(myData, S=averages)#add the seasonal component
  red↵→ into the data frame
myData<-mutate(myData, fit=S+yhat) #complete the model

#LEts plot it

ggplot(data=myData)+geom_line(aes(x=Month.Year, y=Public.
  red↵→ Construction,color="Public.Construction"))+geom_line(
  red↵→ aes(x=myData$Month.Year, y=myData$fit, color="Model"))
  red↵→ +
```

```
xlab("Year")+ylab("Public_Construction")+theme(legend.position=
  red↪ "bottom")+
scale_colour_manual(name='', values=c('Public.Construction'='
  red↪ black', 'Model'='red'))+
scale_x_date(limits = as.Date(c("2002-01-01","2015-01-01")))+
  red↪ ggtitle("Public_Construction")+
theme_classic()
```