Demand Forecasting

IMSOHYUN

2018 년 9 월 11 일

data

```
library(dplyr) ; library(ggplot2)
```

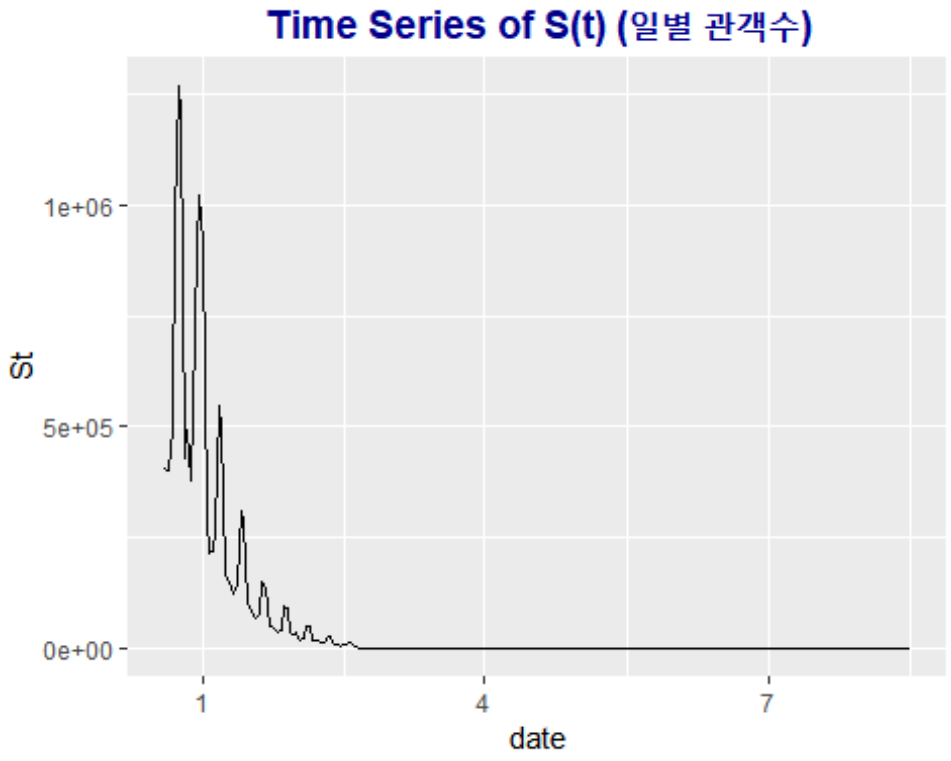
```
##
## Attaching package: 'dplyr'
```

```
## The following objects are masked from 'package:stats':
##
##     filter, lag
```

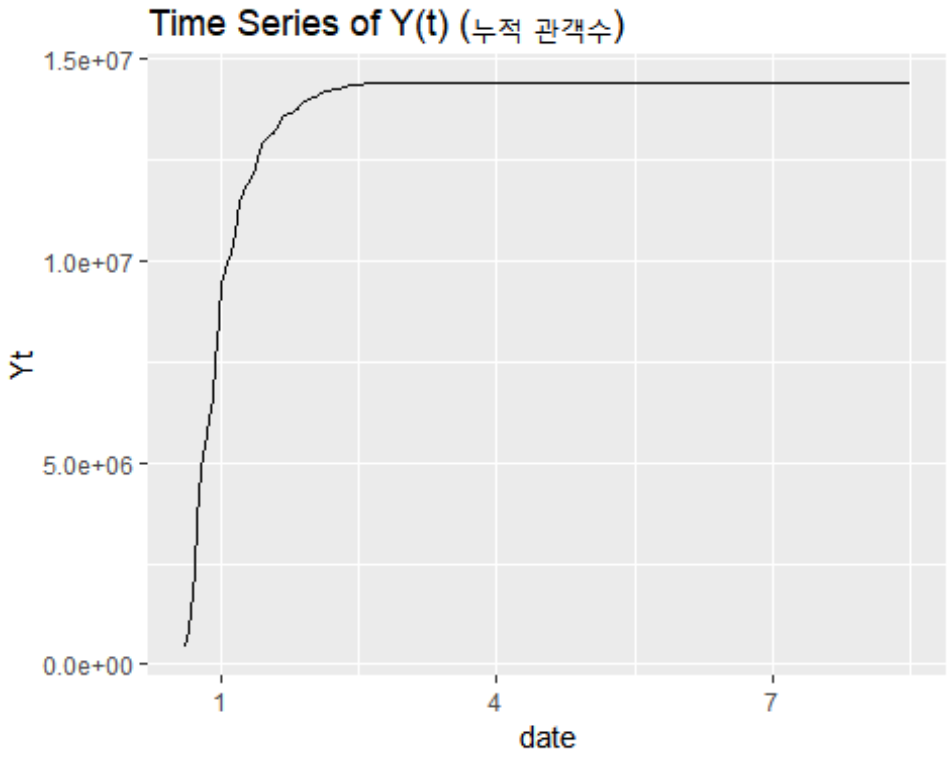
```
## The following objects are masked from 'package:base':
##
##     intersect, setdiff, setequal, union
```

```
# 신과 함께 - 죄와 벌
data1 = read.csv("AlongWithTheGods.csv",stringsAsFactors = F)
data1$date=as.Date(data1$date)
data1 = data1[-c(1:6),]

### Time Series plot
ggplot(data1,aes(date,St))+geom_line()+ggtitle("Time Series of S(t) (일별 관객수)")+
  theme(plot.title = element_text(color="darkblue", size=14, face="bold",hjust=0.5))
```




```
ggplot(data1,aes(date,Yt))+geom_line()+ggtitle("Time Series of Y(t) (누적 관객수)")
```



```
#ggplot(table1,aes(date,St))+geom_line()+ggtitle("Time Series of S(t) (일별 관객수)")
```

function

```
### Functions
ols_bass = function(data,n){
  ols = lm(St~Yt_1+I(Yt_1^2),data=data[1:n,])
  a = ols$coef[1] ; b = ols$coef[2] ; c = ols$coef[3]
  m = (-sqrt(b^2-4*a*c)-b)/(2*c) ; q=-m*c ; p=q-b
  total = data$Yt[length(data$Yt)]
  se = 100*(m-total)/total
  return(c(m,se))
}

ols_logis = function(data,n){
  ols = lm(St~-1+Yt_1+I(Yt_1^2),data=data[1:n,])
  a = ols$coef[1] ; b = ols$coef[2]
  q = a ; m = -q/b
  total = data$Yt[length(data$Yt)]
  se = 100*(m-total)/total
  return(c(m,se))
}

ols_gum = function(data,n){
  ols = lm(St~-1+Yt_1+I(Yt_1*log(Yt_1)),data=data[1:n,])
  a = ols$coef[1] ; b = ols$coef[2]
  q = -b ; m = exp(-a/b)
  total = data$Yt[length(data$Yt)]
  se = 100*(m-total)/total
  return(c(m,se))
}

ols_exp = function(data,n){
  ols = lm(St~Yt_1,data=data[1:n,])
  a = ols$coef[1] ; b = ols$coef[2]
  p = -b  ; m = a/p
  total = data$Yt[length(data$Yt)]
  se = 100*(m-total)/total
  return(c(m,se))
}
```

Find A Best Model

```
### Find A Best Model
holidays <- as.Date(c("2017-12-25", "2018-01-01", "2018-02-15", "2018-02-16", "2018-0
2-17", "2018-02-18",
                      "2018-03-01", "2018-05-05", "2018-05-07", "2018-05-22", "2018-0
6-06", "2018-06-13", "2018-08-15"))
table1 <- data.frame()
for (i in 1:length(data1$St)) {
  if (weekdays(data1$date[i]) %in% c("토요일", "일요일") | data1$date[i] %in% holiday
s) {
    table1 <- rbind(table1, data.frame(date = rep(data1[i,1], 2), St = rep(data1[i,2]
/2, 2))) }
  else { table1 <- rbind(table1, data1[i,1:2]) }
```

```
}

for (i in 2:length(table1$St)){
  table1$Yt[1] = table1$St[1]
  table1$Yt[i] = table1$St[i]+table1$Yt[i-1]
  table1$Yt_1[1] = NA
  table1$Yt_1[i] = table1$Yt[i-1]
  table1$t = seq(1,length(table1$St))
}
data.frame(n = c(7,14,28),
           Bass = c(ols_bass(table1,7)[2],ols_bass(table1,14)[2] ,ols_bass(table1,28)
[2]),
           Logistic = c(ols_logis(table1,7)[2],ols_logis(table1,14)[2],ols_logis(tabl
e1,28)[2]),
           Gumbel = c(ols_gum(table1,7)[2],ols_gum(table1,14)[2],ols_gum(table1,28)
[2]),
           Exponential = c(ols_exp(table1,7)[2],ols_exp(table1,14)[2],ols_exp(table1,
28)[2]))

##    n       Bass  Logistic    Gumbel Exponential
## 1  7        NaN -69.17581 -56.82009  -129.44485
## 2 14 -35.894488 -47.34650 -39.69862   772.41995
## 3 28  -8.562222 -17.44321 -12.75310    45.26231
```

Q-Q plot

```
total1=table1$Yt[length(table1$Yt)]
bass = function(data,n){
  ols = lm(St~Yt_1+I(Yt_1^2),data=data[1:n,])
  a = ols$coef[1] ; b = ols$coef[2] ; c = ols$coef[3]
  m = 11211221 ; q=-m*c ; p=q-b
  return(c(p,q))
}
p1=bass(table1,155)[1] ; q1=bass(table1,155)[2] ; k1=p1+q1 ; c1=q1/p1

qqtable1 = table1 %>% select(t,Yt) %>% filter(t<=94) %>%
  mutate(Ur=Yt/(total1+1),quan=(1/k1)*log((1+c1*Ur)/(1-Ur)))
ggplot(qqtable1,aes(quan,t))+geom_point()+geom_smooth(method="lm")+
  ggtitle("신과함께 :  Bass Q-Q plot")
```


신과함께 : Bass Q-Q plot
90
60
30
0
t
0
50
100
150
quan