DATA 624 PROJECT 1

Project 1

Samriti Malhotra

June 25, 2020

Table of Contents

## Series S02

### Import the necessary libraries

knitr::opts\_chunk$set(echo = TRUE)  
  
library(fpp2)

## Warning: package 'fpp2' was built under R version 3.5.3

## Loading required package: ggplot2

## Loading required package: forecast

## Warning: package 'forecast' was built under R version 3.5.3

## Loading required package: fma

## Warning: package 'fma' was built under R version 3.5.3

## Loading required package: expsmooth

## Warning: package 'expsmooth' was built under R version 3.5.3

library(dplyr)

##   
## 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

library(imputeTS)

## Warning: package 'imputeTS' was built under R version 3.5.3

library(urca)

## Warning: package 'urca' was built under R version 3.5.3

library(MLmetrics)

## Warning: package 'MLmetrics' was built under R version 3.5.3

##   
## Attaching package: 'MLmetrics'

## The following object is masked from 'package:base':  
##   
## Recall

### Load the dataset

data\_project <- readxl::read\_excel("./project1data/Data Set for class.xls")  
head(data\_project)

## # A tibble: 6 x 7  
## SeriesInd group Var01 Var02 Var03 Var05 Var07  
## <dbl> <chr> <dbl> <dbl> <dbl> <dbl> <dbl>  
## 1 40669 S03 30.6 123432400 30.3 30.5 30.6  
## 2 40669 S02 10.3 60855800 10.0 10.2 10.3  
## 3 40669 S01 26.6 10369300 25.9 26.2 26.0  
## 4 40669 S06 27.5 39335700 26.8 27.0 27.3  
## 5 40669 S05 69.3 27809100 68.2 68.7 69.2  
## 6 40669 S04 17.2 16587400 16.9 16.9 17.1

### Series 02 Subset

S02 <- subset(data\_project, group == 'S02', select = c(SeriesInd, Var02, Var03))  
head(S02)

## # A tibble: 6 x 3  
## SeriesInd Var02 Var03  
## <dbl> <dbl> <dbl>  
## 1 40669 60855800 10.0  
## 2 40670 215620200 10.4  
## 3 40671 200070600 11.1  
## 4 40672 130201700 11.3  
## 5 40673 130463000 11.5  
## 6 40676 170626200 11.8

### Exploratory Analysis of variables in Series S02

predictobs <- 1623:1762  
S2 <- ts(S02[-predictobs, 2:3])  
  
S2v1 <- ts(S02[-predictobs,2])  
S2v2 <- ts(S02[-predictobs,3])  
  
  
summary(S2v1)

## Var02   
## Min. : 7128800   
## 1st Qu.: 27880300   
## Median : 39767500   
## Mean : 50633098   
## 3rd Qu.: 59050900   
## Max. :480879500

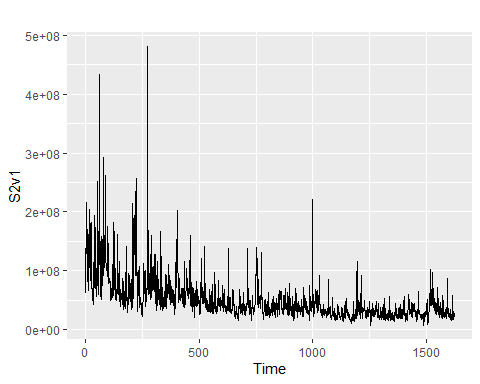
summary(S2v2)

## Var03   
## Min. : 8.82   
## 1st Qu.:11.82   
## Median :13.76   
## Mean :13.68   
## 3rd Qu.:15.52   
## Max. :38.28   
## NA's :4

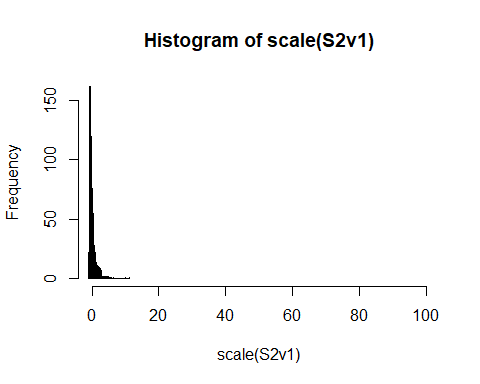
Var02 has no missing values and Var03 has 4 missing values.

### Exploratory Analysis of Var02 in series 2

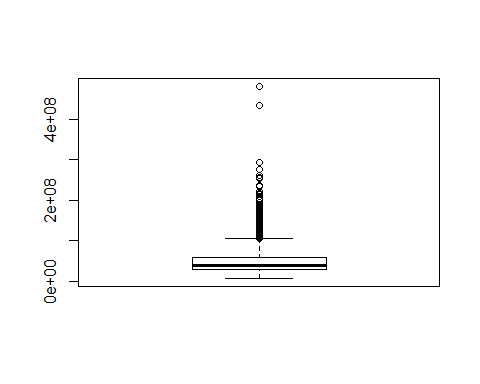
autoplot(S2v1)



hist(scale(S2v1),br=100,xlim=c(0,100) )

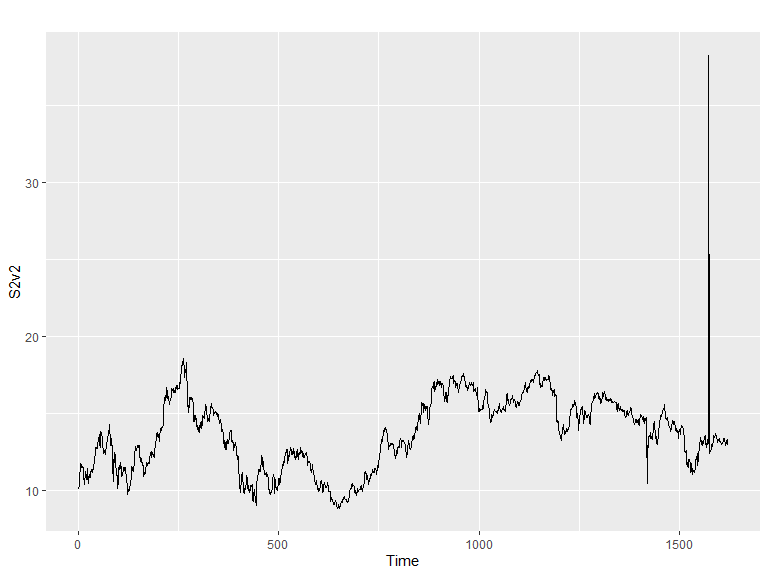


boxplot(S2v1)

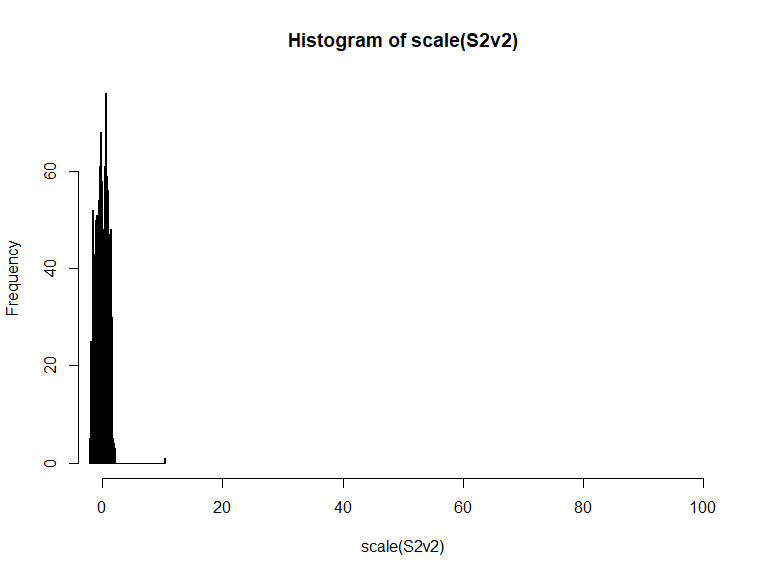
 Looking at above plots fixed distribution assumption do not holds, as histogram is not be bell-shaped, and the normal probability plot is not be linear. it is right skewed with many outliers. May be after supressing outliers distribution plot will improve.

### Exploratory Analysis of Var03 in series 2

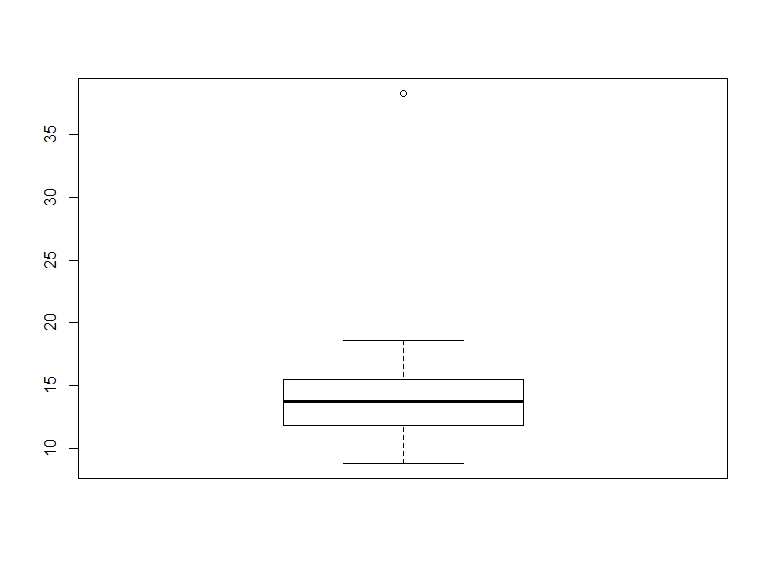
# interpreting NA values in our variable,  
S2v2 <- na\_interpolation(S2v2)   
  
  
autoplot(S2v2)



hist(scale(S2v2),br=100,xlim=c(0,100) )



boxplot(S2v2)



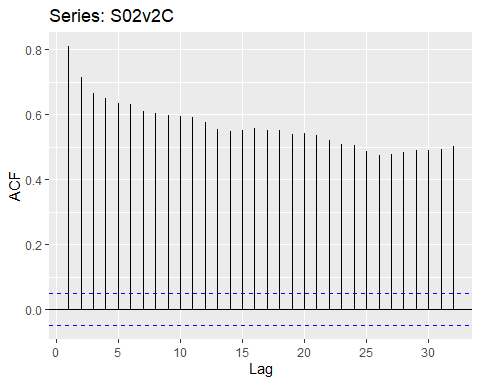
Looking at above plots fixed distribution assumption holds good for V03, as histogram is bell-shaped, and the normal probability plot is linear. It has one outliers.

### Oulier for var02

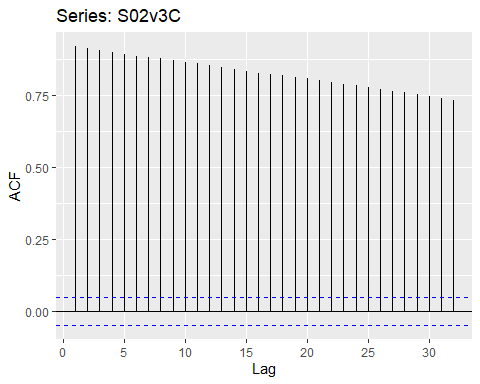
S02v2C <- tsclean(S2v1)  
  
S02v3C <- S2v2

### ACF of Var02 and Var03

ggAcf(S02v2C)



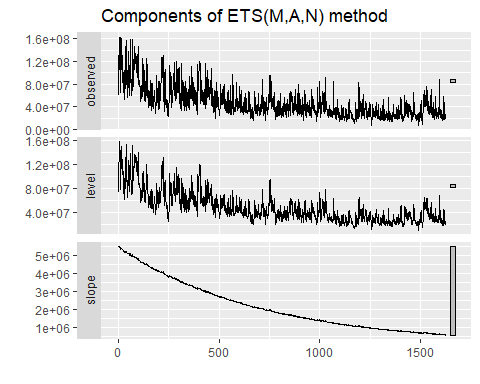
ggAcf(S02v3C)

 ACF plots for var02 & var03 shows Trend with insignificant seasonality.

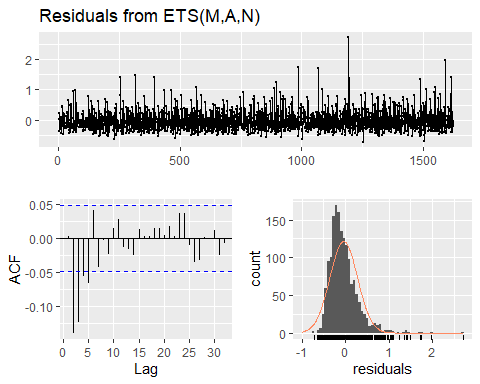
## Applying Models

### ETS model to Var02

fit\_ets <- ets(S02v2C)  
autoplot(fit\_ets)

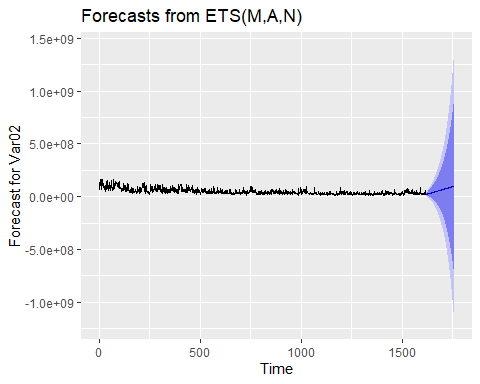


checkresiduals(fit\_ets)



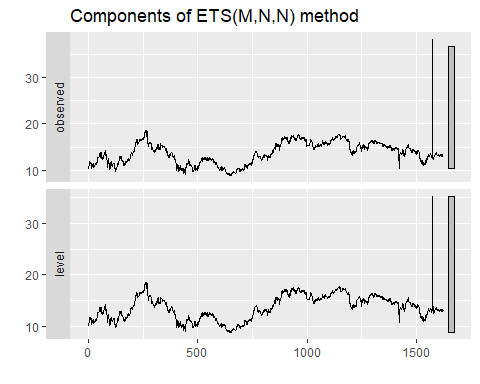
##   
## Ljung-Box test  
##   
## data: Residuals from ETS(M,A,N)  
## Q\* = 75.32, df = 6, p-value = 3.297e-14  
##   
## Model df: 4. Total lags used: 10

fit\_ets %>% forecast(h=140) %>%  
 autoplot() +  
 ylab("Forecast for Var02")

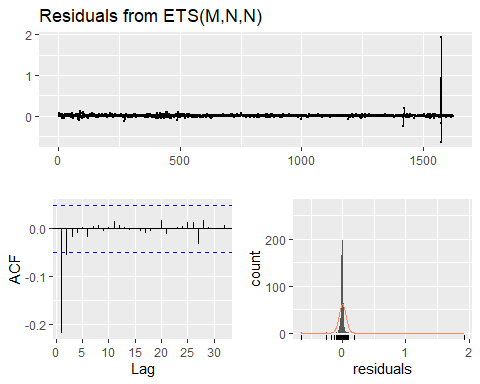


### ETS model to Var03

#Var 03  
fit\_ets\_V03 <- ets(S02v3C)  
autoplot(fit\_ets\_V03)

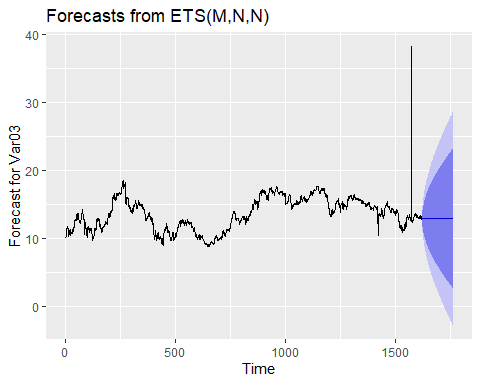


checkresiduals(fit\_ets\_V03)



##   
## Ljung-Box test  
##   
## data: Residuals from ETS(M,N,N)  
## Q\* = 83.799, df = 8, p-value = 8.327e-15  
##   
## Model df: 2. Total lags used: 10

fit\_ets\_V03 %>% forecast(h=140) %>%  
 autoplot() +  
 ylab("Forecast for Var03")

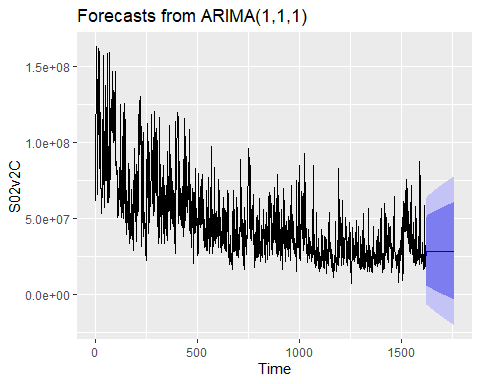


### Applying ARIMA Model to Var02.

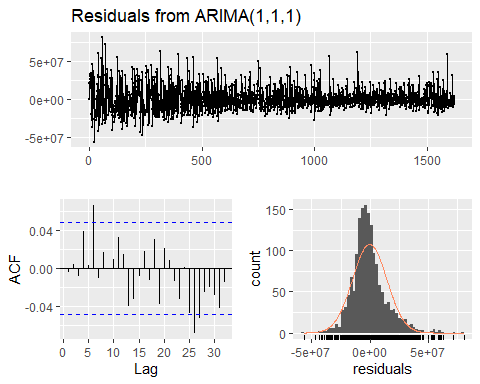
arima\_fit\_v02 <- auto.arima(S02v2C)  
  
summary(arima\_fit\_v02)

## Series: S02v2C   
## ARIMA(1,1,1)   
##   
## Coefficients:  
## ar1 ma1  
## 0.5075 -0.9502  
## s.e. 0.0283 0.0128  
##   
## sigma^2 estimated as 2.166e+14: log likelihood=-29053.65  
## AIC=58113.31 AICc=58113.32 BIC=58129.48  
##   
## Training set error measures:  
## ME RMSE MAE MPE MAPE MASE  
## Training set -421309.4 14704115 10777336 -8.965164 24.71449 0.9189302  
## ACF1  
## Training set -0.004111647

autoplot(forecast(arima\_fit\_v02, h=140))



checkresiduals(arima\_fit\_v02)



##   
## Ljung-Box test  
##   
## data: Residuals from ARIMA(1,1,1)  
## Q\* = 10.552, df = 8, p-value = 0.2284  
##   
## Model df: 2. Total lags used: 10

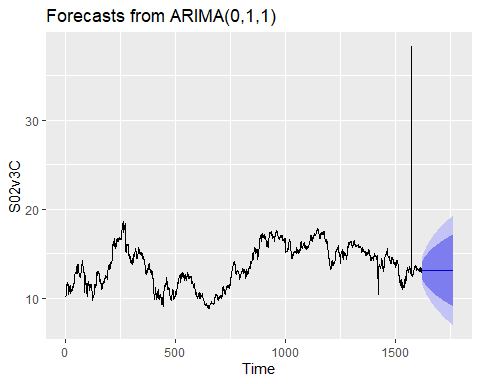
The auto.arima results with ARIMA(1,0,1) model with no drift ,

### Arima model for Var 03

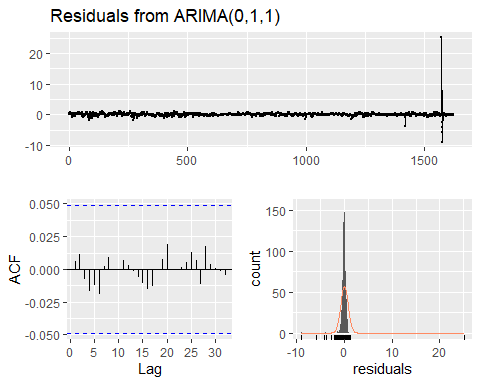
arima\_fit\_v03 <- auto.arima(S02v3C)  
  
summary(arima\_fit\_v03)

## Series: S02v3C   
## ARIMA(0,1,1)   
##   
## Coefficients:  
## ma1  
## -0.6687  
## s.e. 0.0191  
##   
## sigma^2 estimated as 0.6075: log likelihood=-1895.94  
## AIC=3795.87 AICc=3795.88 BIC=3806.66  
##   
## Training set error measures:  
## ME RMSE MAE MPE MAPE MASE  
## Training set 0.005200808 0.7789411 0.2901101 -0.102729 2.13786 1.383264  
## ACF1  
## Training set 0.006273859

autoplot(forecast(arima\_fit\_v03, h=140))



checkresiduals(arima\_fit\_v03)



##   
## Ljung-Box test  
##   
## data: Residuals from ARIMA(0,1,1)  
## Q\* = 1.8347, df = 9, p-value = 0.9938  
##   
## Model df: 1. Total lags used: 10

A portmanteau test returns a large p-value 0.9938, also suggesting that the residuals are white noise. The ACF plot of the residuals from the ARIMA(0,1,1) model shows that all autocorrelations are within the threshold limits, indicating that the residuals are behaving like white noise

### MAPE Calculation:

print(paste0("Accuracy for Var 02"))

## [1] "Accuracy for Var 02"

print(paste0("MAPE for S02 Var02 using ETS model ::: ", MLmetrics::MAPE(fit\_ets$fitted,S02v2C)))

## [1] "MAPE for S02 Var02 using ETS model ::: 0.279104207788701"

print(paste0("MAPE for S02 Var02 using Auto ARIMA model ::: ", MLmetrics::MAPE(arima\_fit\_v02$fitted,S02v2C)))

## [1] "MAPE for S02 Var02 using Auto ARIMA model ::: 0.247144896628619"

print(paste0("Accuracy for Var 03"))

## [1] "Accuracy for Var 03"

print(paste0("MAPE for S02 Var03 using ETS model ::: ", MLmetrics::MAPE(fit\_ets\_V03$fitted,S02v3C)))

## [1] "MAPE for S02 Var03 using ETS model ::: 0.0156105029139132"

print(paste0("MAPE for S02 Var03 using Auto ARIMA model ::: ", MLmetrics::MAPE(arima\_fit\_v03$fitted,S02v3C)))

## [1] "MAPE for S02 Var03 using Auto ARIMA model ::: 0.0213786043674216"

Looking at MAPE we are using ARIMA for forcast of Var03.

Also looking at the residuals for both models variables is having constant variance and normal distrubution and also residuals are uncorrelated with nearly zero mean.The mean of the residuals is close to zero and there is no significant correlation in the residuals series.

### Writing forcast of V03 to csv

fc <- forecast(arima\_fit\_v03, h=140)  
fc$mean<-fc$mean  
fc$upper<-fc$upper  
fc$lower<-fc$lower  
fc$x<-fc$x  
  
fc

## Point Forecast Lo 80 Hi 80 Lo 95 Hi 95  
## 1623 13.07344 12.074573 14.07231 11.545803 14.60108  
## 1624 13.07344 12.021187 14.12570 11.464157 14.68273  
## 1625 13.07344 11.970382 14.17650 11.386457 14.76043  
## 1626 13.07344 11.921816 14.22507 11.312182 14.83470  
## 1627 13.07344 11.875217 14.27167 11.240915 14.90597  
## 1628 13.07344 11.830363 14.31652 11.172317 14.97457  
## 1629 13.07344 11.787073 14.35981 11.106110 15.04077  
## 1630 13.07344 11.745192 14.40169 11.042060 15.10482  
## 1631 13.07344 11.704593 14.44229 10.979968 15.16692  
## 1632 13.07344 11.665163 14.48172 10.919666 15.22722  
## 1633 13.07344 11.626808 14.52008 10.861006 15.28588  
## 1634 13.07344 11.589444 14.55744 10.803863 15.34302  
## 1635 13.07344 11.552997 14.59389 10.748123 15.39876  
## 1636 13.07344 11.517404 14.62948 10.693688 15.45320  
## 1637 13.07344 11.482608 14.66428 10.640471 15.50641  
## 1638 13.07344 11.448556 14.69833 10.588393 15.55849  
## 1639 13.07344 11.415203 14.73168 10.537385 15.60950  
## 1640 13.07344 11.382508 14.76438 10.487382 15.65950  
## 1641 13.07344 11.350434 14.79645 10.438328 15.70856  
## 1642 13.07344 11.318945 14.82794 10.390171 15.75671  
## 1643 13.07344 11.288012 14.85887 10.342863 15.80402  
## 1644 13.07344 11.257606 14.88928 10.296360 15.85052  
## 1645 13.07344 11.227700 14.91918 10.250624 15.89626  
## 1646 13.07344 11.198272 14.94861 10.205617 15.94127  
## 1647 13.07344 11.169298 14.97759 10.161305 15.98558  
## 1648 13.07344 11.140759 15.00612 10.117658 16.02923  
## 1649 13.07344 11.112635 15.03425 10.074646 16.07224  
## 1650 13.07344 11.084908 15.06198 10.032242 16.11464  
## 1651 13.07344 11.057563 15.08932 9.990421 16.15646  
## 1652 13.07344 11.030584 15.11630 9.949160 16.19772  
## 1653 13.07344 11.003957 15.14293 9.908437 16.23845  
## 1654 13.07344 10.977668 15.16922 9.868232 16.27865  
## 1655 13.07344 10.951704 15.19518 9.828524 16.31836  
## 1656 13.07344 10.926055 15.22083 9.789296 16.35759  
## 1657 13.07344 10.900708 15.24618 9.750532 16.39635  
## 1658 13.07344 10.875653 15.27123 9.712215 16.43467  
## 1659 13.07344 10.850881 15.29600 9.674329 16.47255  
## 1660 13.07344 10.826382 15.32050 9.636861 16.51002  
## 1661 13.07344 10.802148 15.34474 9.599797 16.54709  
## 1662 13.07344 10.778169 15.36871 9.563125 16.58376  
## 1663 13.07344 10.754438 15.39245 9.526832 16.62005  
## 1664 13.07344 10.730947 15.41594 9.490906 16.65598  
## 1665 13.07344 10.707690 15.43919 9.455337 16.69155  
## 1666 13.07344 10.684659 15.46222 9.420114 16.72677  
## 1667 13.07344 10.661848 15.48504 9.385228 16.76166  
## 1668 13.07344 10.639251 15.50763 9.350669 16.79622  
## 1669 13.07344 10.616862 15.53002 9.316427 16.83046  
## 1670 13.07344 10.594675 15.55221 9.282495 16.86439  
## 1671 13.07344 10.572685 15.57420 9.248864 16.89802  
## 1672 13.07344 10.550886 15.59600 9.215526 16.93136  
## 1673 13.07344 10.529274 15.61761 9.182473 16.96441  
## 1674 13.07344 10.507845 15.63904 9.149699 16.99718  
## 1675 13.07344 10.486592 15.66029 9.117197 17.02969  
## 1676 13.07344 10.465513 15.68137 9.084959 17.06192  
## 1677 13.07344 10.444603 15.70228 9.052980 17.09390  
## 1678 13.07344 10.423858 15.72303 9.021253 17.12563  
## 1679 13.07344 10.403274 15.74361 8.989773 17.15711  
## 1680 13.07344 10.382848 15.76404 8.958534 17.18835  
## 1681 13.07344 10.362575 15.78431 8.927529 17.21935  
## 1682 13.07344 10.342453 15.80443 8.896756 17.25013  
## 1683 13.07344 10.322479 15.82441 8.866207 17.28068  
## 1684 13.07344 10.302648 15.84424 8.835878 17.31101  
## 1685 13.07344 10.282958 15.86393 8.805765 17.34112  
## 1686 13.07344 10.263406 15.88348 8.775863 17.37102  
## 1687 13.07344 10.243989 15.90289 8.746167 17.40072  
## 1688 13.07344 10.224705 15.92218 8.716674 17.43021  
## 1689 13.07344 10.205550 15.94133 8.687380 17.45950  
## 1690 13.07344 10.186522 15.96036 8.658279 17.48860  
## 1691 13.07344 10.167619 15.97926 8.629369 17.51751  
## 1692 13.07344 10.148838 15.99805 8.600646 17.54624  
## 1693 13.07344 10.130177 16.01671 8.572107 17.57478  
## 1694 13.07344 10.111633 16.03525 8.543747 17.60314  
## 1695 13.07344 10.093205 16.05368 8.515563 17.63132  
## 1696 13.07344 10.074890 16.07199 8.487553 17.65933  
## 1697 13.07344 10.056687 16.09020 8.459713 17.68717  
## 1698 13.07344 10.038592 16.10829 8.432040 17.71484  
## 1699 13.07344 10.020605 16.12628 8.404530 17.74235  
## 1700 13.07344 10.002723 16.14416 8.377182 17.76970  
## 1701 13.07344 9.984944 16.16194 8.349993 17.79689  
## 1702 13.07344 9.967268 16.17962 8.322959 17.82393  
## 1703 13.07344 9.949691 16.19719 8.296077 17.85081  
## 1704 13.07344 9.932213 16.21467 8.269347 17.87754  
## 1705 13.07344 9.914831 16.23205 8.242764 17.90412  
## 1706 13.07344 9.897545 16.24934 8.216327 17.93056  
## 1707 13.07344 9.880352 16.26653 8.190033 17.95685  
## 1708 13.07344 9.863251 16.28363 8.163879 17.98300  
## 1709 13.07344 9.846241 16.30064 8.137864 18.00902  
## 1710 13.07344 9.829320 16.31756 8.111986 18.03490  
## 1711 13.07344 9.812487 16.33440 8.086242 18.06064  
## 1712 13.07344 9.795740 16.35114 8.060630 18.08625  
## 1713 13.07344 9.779079 16.36780 8.035148 18.11174  
## 1714 13.07344 9.762501 16.38438 8.009795 18.13709  
## 1715 13.07344 9.746006 16.40088 7.984568 18.16232  
## 1716 13.07344 9.729592 16.41729 7.959465 18.18742  
## 1717 13.07344 9.713259 16.43363 7.934485 18.21240  
## 1718 13.07344 9.697004 16.44988 7.909626 18.23726  
## 1719 13.07344 9.680827 16.46606 7.884886 18.26200  
## 1720 13.07344 9.664727 16.48216 7.860263 18.28662  
## 1721 13.07344 9.648703 16.49818 7.835756 18.31113  
## 1722 13.07344 9.632753 16.51413 7.811363 18.33552  
## 1723 13.07344 9.616877 16.53001 7.787083 18.35980  
## 1724 13.07344 9.601074 16.54581 7.762913 18.38397  
## 1725 13.07344 9.585342 16.56154 7.738854 18.40803  
## 1726 13.07344 9.569681 16.57720 7.714902 18.43198  
## 1727 13.07344 9.554089 16.59279 7.691057 18.45583  
## 1728 13.07344 9.538567 16.60832 7.667317 18.47957  
## 1729 13.07344 9.523112 16.62377 7.643681 18.50320  
## 1730 13.07344 9.507724 16.63916 7.620147 18.52674  
## 1731 13.07344 9.492402 16.65448 7.596714 18.55017  
## 1732 13.07344 9.477146 16.66974 7.573381 18.57350  
## 1733 13.07344 9.461954 16.68493 7.550147 18.59674  
## 1734 13.07344 9.446825 16.70006 7.527010 18.61987  
## 1735 13.07344 9.431760 16.71512 7.503969 18.64291  
## 1736 13.07344 9.416756 16.73013 7.481024 18.66586  
## 1737 13.07344 9.401814 16.74507 7.458172 18.68871  
## 1738 13.07344 9.386932 16.75995 7.435412 18.71147  
## 1739 13.07344 9.372111 16.77477 7.412744 18.73414  
## 1740 13.07344 9.357348 16.78954 7.390167 18.75672  
## 1741 13.07344 9.342644 16.80424 7.367678 18.77921  
## 1742 13.07344 9.327997 16.81889 7.345278 18.80161  
## 1743 13.07344 9.313408 16.83348 7.322966 18.82392  
## 1744 13.07344 9.298875 16.84801 7.300739 18.84614  
## 1745 13.07344 9.284397 16.86249 7.278598 18.86829  
## 1746 13.07344 9.269975 16.87691 7.256541 18.89034  
## 1747 13.07344 9.255607 16.89128 7.234568 18.91232  
## 1748 13.07344 9.241293 16.90559 7.212676 18.93421  
## 1749 13.07344 9.227033 16.91985 7.190867 18.95602  
## 1750 13.07344 9.212825 16.93406 7.169138 18.97775  
## 1751 13.07344 9.198669 16.94821 7.147488 18.99940  
## 1752 13.07344 9.184565 16.96232 7.125917 19.02097  
## 1753 13.07344 9.170511 16.97637 7.104425 19.04246  
## 1754 13.07344 9.156508 16.99038 7.083009 19.06387  
## 1755 13.07344 9.142555 17.00433 7.061670 19.08521  
## 1756 13.07344 9.128652 17.01823 7.040406 19.10648  
## 1757 13.07344 9.114797 17.03209 7.019217 19.12767  
## 1758 13.07344 9.100990 17.04589 6.998101 19.14878  
## 1759 13.07344 9.087231 17.05965 6.977059 19.16982  
## 1760 13.07344 9.073520 17.07336 6.956089 19.19079  
## 1761 13.07344 9.059856 17.08703 6.935191 19.21169  
## 1762 13.07344 9.046237 17.10065 6.914364 19.23252

write.csv(fc,"s02v03.csv")