

changepoint analysis

Afiqah

2023-02-06

Workflow

1. Prepare R

1.1 Set working directory

1.2 Load packages

1.3 Load data

2. Analysis

2.1 Explore data

2.2 Prepare dataframe for variables

2.3 Changepoint analysis

2.3.1 Overall trend

2.3.2 Triage category

2.3.3 Discharge outcome

2.3.4 ED discharge diagnosis

Analysis

1. Prepare R

1.1 Set working directory

```
getwd()
```

```
## [1] "C:/Users/Afiqah Syamimi/Documents/USM/Research Project/Air pollution/Data air poluution"
```

1.2 Load packages

```
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(ggplot2)  
library(gtsummary)  
library(EnvCpt)
```

```
## Loading required package: changepoint
```

```
## Loading required package: zoo
```

```
##  
## Attaching package: 'zoo'
```

```
## The following objects are masked from 'package:base':  
##  
##   as.Date, as.Date.numeric
```

```
## Successfully loaded changepoint package version 2.2.4  
## See NEWS for details of changes.
```

```
## Loading required package: MASS
```

```
##  
## Attaching package: 'MASS'
```

```
## The following object is masked from 'package:gtsummary':  
##  
##   select
```

```
## The following object is masked from 'package:dplyr':  
##  
##   select
```

```
library(changepoint)
library(changepoint.influence)
library(ggpubr)
```

1.3 Load data

```
data <- read.csv('EDbyWeek.csv')
```

2. Analysis

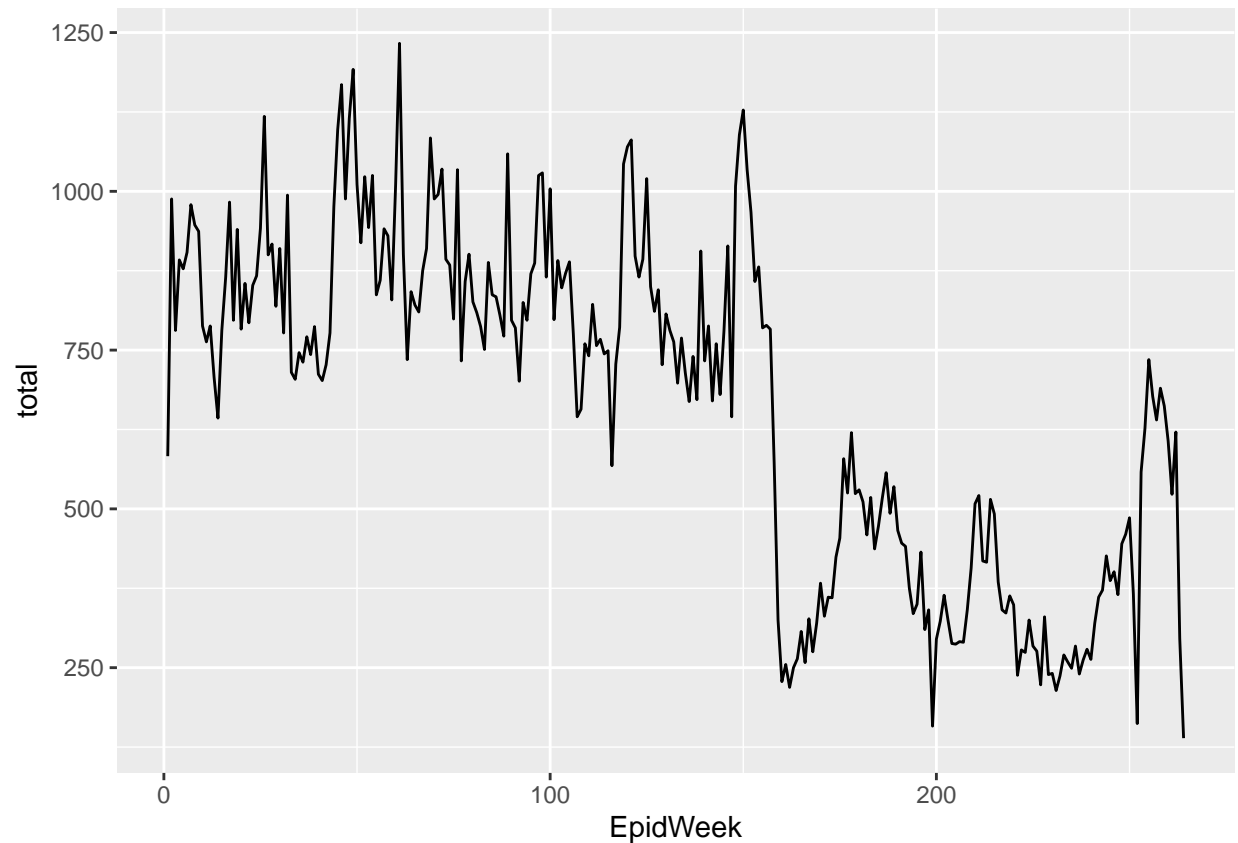
2.1 Explore data

```
summary(data)
```

```
##      EpidWeek      Period      total      age1
## Min.   : 1.00   Min.   :1.000   Min.   : 139.0   Min.   : 75.0
## 1st Qu.: 66.75   1st Qu.:1.000   1st Qu.: 413.8   1st Qu.:218.2
## Median :132.50   Median :1.000   Median : 737.5   Median :422.0
## Mean   :132.50   Mean   :1.451   Mean   : 665.7   Mean   :375.4
## 3rd Qu.:198.25   3rd Qu.:2.000   3rd Qu.: 865.5   3rd Qu.:496.2
## Max.   :264.00   Max.   :3.000   Max.   :1233.0   Max.   :765.0
##      age2      age3      green      green.avg
## Min.   : 35.0   Min.   : 21.00   Min.   : 98.0   Min.   : 14.00
## 1st Qu.:110.8   1st Qu.: 73.75   1st Qu.: 306.2   1st Qu.: 43.75
## Median :196.5   Median :112.00   Median : 592.0   Median : 84.57
## Mean   :182.2   Mean   :108.09   Mean   : 526.9   Mean   : 75.27
## 3rd Qu.:236.0   3rd Qu.:136.00   3rd Qu.: 685.5   3rd Qu.: 97.93
## Max.   :385.0   Max.   :351.00   Max.   :1004.0   Max.   :143.43
##      red      red.avg      yellow      yellow.avg
## Min.   : 4.00   Min.   :0.5714   Min.   : 22.00   Min.   : 3.143
## 1st Qu.:16.00   1st Qu.:2.2857   1st Qu.: 77.75   1st Qu.:11.107
## Median :21.00   Median :3.0000   Median :121.00   Median :17.286
## Mean   :22.13   Mean   :3.1613   Mean   :116.67   Mean   :16.667
## 3rd Qu.:26.00   3rd Qu.:3.7143   3rd Qu.:150.25   3rd Qu.:21.464
## Max.   :69.00   Max.   :9.8571   Max.   :325.00   Max.   :46.429
##      admission      admission.avg      death      death.avg
## Min.   : 31.0   Min.   : 4.429   Min.   :0.0   Min.   :0.00000
## 1st Qu.:102.0   1st Qu.:14.571   1st Qu.:0.0   1st Qu.:0.00000
## Median :129.0   Median :18.429   Median :0.0   Median :0.00000
## Mean   :123.5   Mean   :17.648   Mean   :0.5   Mean   :0.07143
## 3rd Qu.:149.0   3rd Qu.:21.286   3rd Qu.:1.0   3rd Qu.:0.14286
## Max.   :286.0   Max.   :40.857   Max.   :4.0   Max.   :0.57143
##      discharge      discharge.avg      allergic      assault
## Min.   : 108.0   Min.   : 15.43   Min.   : 0.000   Min.   : 0.000
## 1st Qu.: 309.8   1st Qu.: 44.25   1st Qu.: 4.000   1st Qu.: 2.750
## Median : 601.5   Median : 85.93   Median : 6.000   Median : 4.000
## Mean   : 541.6   Mean   : 77.38   Mean   : 6.288   Mean   : 4.807
## 3rd Qu.: 715.8   3rd Qu.:102.25   3rd Qu.: 9.000   3rd Qu.: 7.000
## Max.   :1074.0   Max.   :153.43   Max.   :16.000   Max.   :15.000
```

##	traumaburn	cns	derm	ent
##	Min. : 18.00	Min. : 4.00	Min. : 1.00	Min. : 0.00
##	1st Qu.: 67.75	1st Qu.:13.00	1st Qu.:10.00	1st Qu.:12.00
##	Median : 88.00	Median :19.00	Median :15.00	Median :15.00
##	Mean : 87.09	Mean :19.34	Mean :16.19	Mean :15.85
##	3rd Qu.:108.00	3rd Qu.:25.00	3rd Qu.:21.00	3rd Qu.:20.00
##	Max. :147.00	Max. :50.00	Max. :43.00	Max. :35.00
##	cd	gi	msp	neo
##	Min. : 19.00	Min. : 18.00	Min. : 0.000	Min. : 0.00
##	1st Qu.: 66.75	1st Qu.: 42.75	1st Qu.: 4.000	1st Qu.:27.75
##	Median :132.00	Median : 77.00	Median : 6.000	Median :37.50
##	Mean :129.07	Mean : 74.41	Mean : 6.261	Mean :35.39
##	3rd Qu.:167.00	3rd Qu.:100.25	3rd Qu.: 8.000	3rd Qu.:46.00
##	Max. :392.00	Max. :147.00	Max. :16.000	Max. :81.00
##	nephro	ong	omfs	ophthal
##	Min. : 0.000	Min. :0.000	Min. : 0.000	Min. : 0.000
##	1st Qu.: 4.000	1st Qu.:0.000	1st Qu.: 1.000	1st Qu.: 3.000
##	Median : 6.000	Median :1.000	Median : 3.000	Median : 6.000
##	Mean : 6.795	Mean :1.038	Mean : 3.019	Mean : 6.045
##	3rd Qu.: 9.000	3rd Qu.:2.000	3rd Qu.: 4.000	3rd Qu.: 8.000
##	Max. :18.000	Max. :5.000	Max. :12.000	Max. :16.000
##	others	poison	psy	respi
##	Min. : 3.00	Min. : 0.000	Min. :0.0000	Min. : 21.0
##	1st Qu.: 9.00	1st Qu.: 1.000	1st Qu.:0.0000	1st Qu.: 94.0
##	Median :12.00	Median : 2.000	Median :1.0000	Median :222.5
##	Mean :12.87	Mean : 2.424	Mean :0.8523	Mean :199.6
##	3rd Qu.:16.00	3rd Qu.: 3.000	3rd Qu.:1.0000	3rd Qu.:281.8
##	Max. :28.00	Max. :11.000	Max. :8.0000	Max. :466.0

```
ggplot(data, aes(x=EpidWeek, y=total)) + geom_line()
```



2.2 Prepare dataframe for variables

```
all <- c(data$total)
green <- c(data$green)
yellow <- c(data$yellow)
red <- c(data$red)
discharge <- c(data$discharge)
admission <- c(data$admission)
death <- c(data$death)
allergic <- c(data$allergic)
assault <- c(data$assault)
trauma <- c(data$traumaburn)
cns <- c(data$cns)
derm <- c(data$derm)
ent <- c(data$ent)
cd <- c(data$cd)
gi <- c(data$gi)
msp <- c(data$msp)
neo <- c(data$neo)
nephro <- c(data$nephro)
ong <- c(data$ong)
omfs <- c(data$omfs)
ophthal <- c(data$ophthal)
others <- c(data$others)
poison <- c(data$poison)
psy <- c(data$psy)
respi <- c(data$respi)
```

2.3 Changepoint analysis

Primary model

```
cpt.all <- envcpt(all)
```

2.3.1 Overall trend

```
## Fitting 12 models
```

```
##      |
```

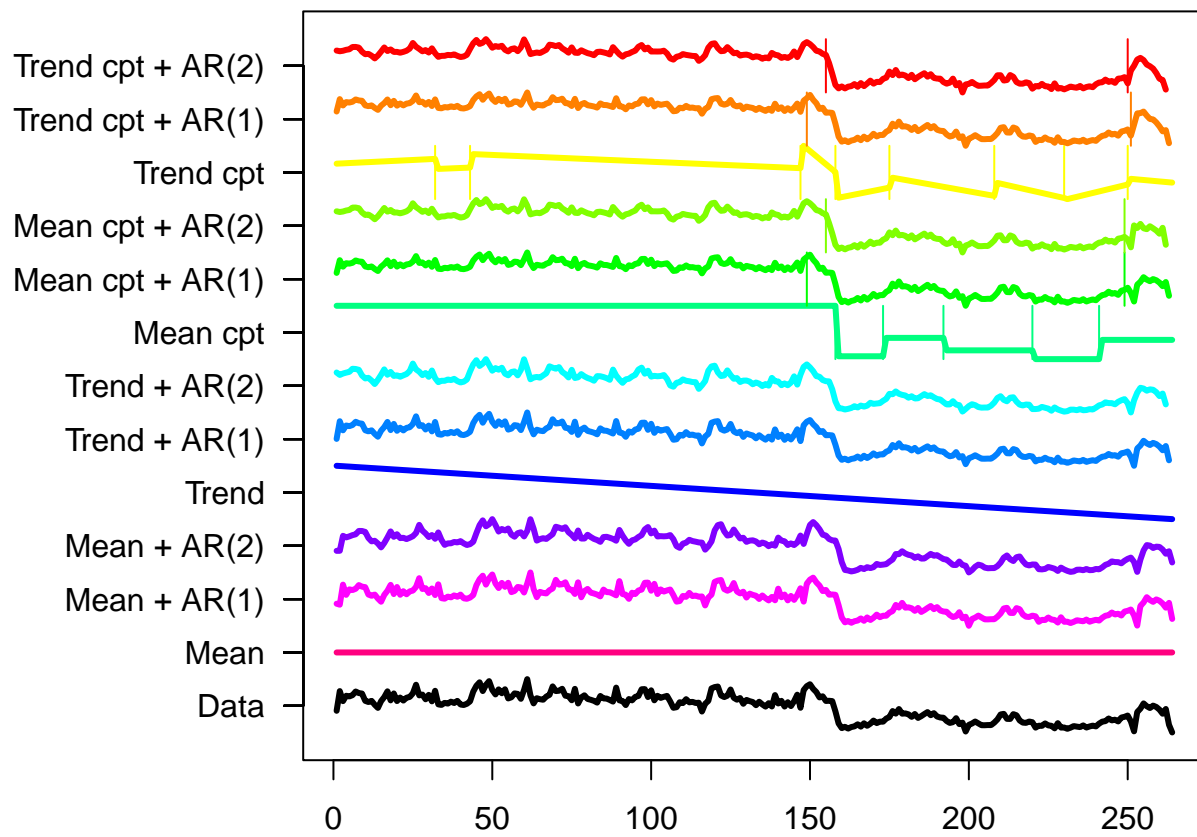
```
BIC(cpt.all) # returns the BIC for each model.
```

```
##      mean      meancpt      meanar1      meanar2  meanar1cpt  meanar2cpt
##  3704.424  3267.567  3239.187  3225.434   3170.135   3162.467
##      trend  trendcpt  trendar1  trendar2  trendar1cpt  trendar2cpt
##  3486.984  3223.276  3210.252  3181.588   3178.460   3169.639
```

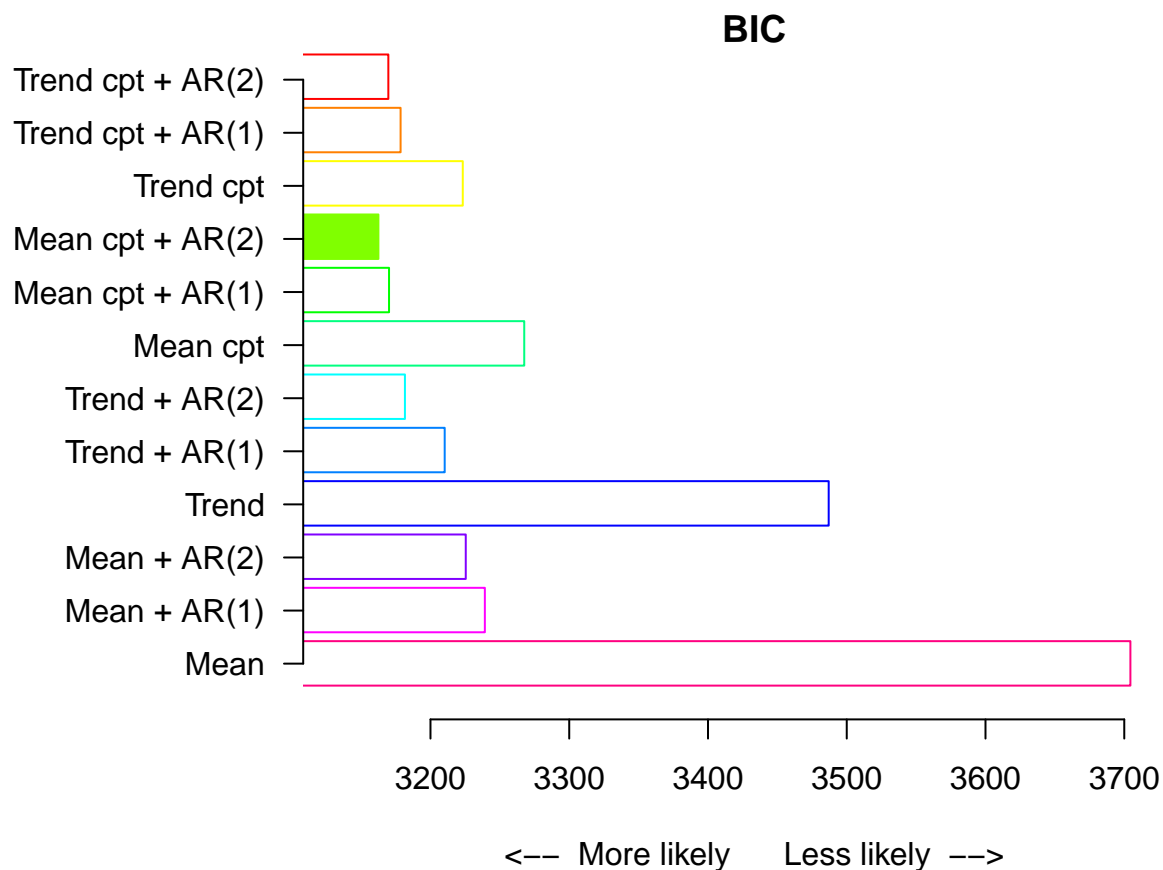
```
which.min(BIC(cpt.all))
```

```
## meanar2cpt
##          6
```

```
plot(cpt.all,type='fit') # plots the fits
```



```
plot(cpt.all,type="bic") # plots the bic values
```



```
### Selection of model based on best BIC
```

```
cpt.all$meanar2cpt
```

```
## Class 'cpt.reg' : Changepoint Regression Object
##      ~~~ : S4 class containing 12 slots with names
##           cpttype date version data.set method test.stat pen.type pen.value minseglen cpts ncpts
##
## Created on   : Sun Jan 29 03:56:54 2023
##
## summary(.) :
## -----
## Created Using changepoint version 2.2.4
## Changepoint type      : Change in regression
## Method of analysis    : PELT
## Test Statistic       : Normal
## Type of penalty       : MBIC with value, 33.41007
## Maximum no. of cpts   : Inf
## Changepoint Locations : 155 249
```

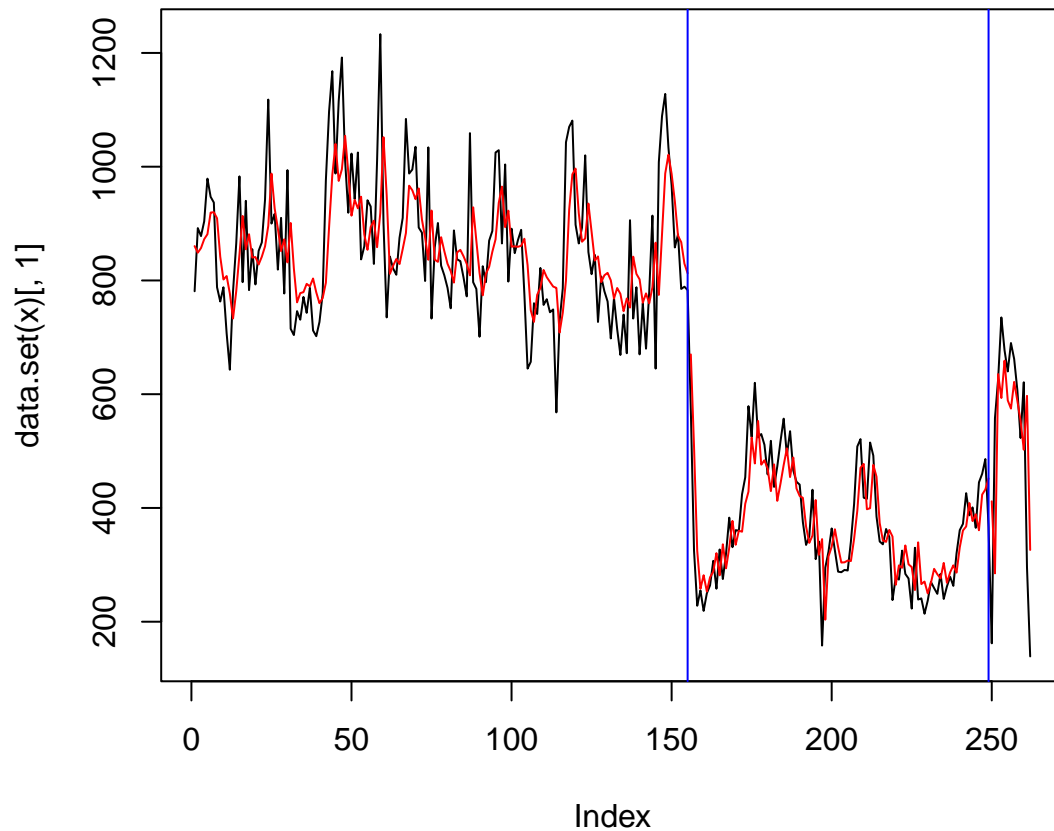
```
cpts(cpt.all$meanar2cpt)
```

```
## [1] 155 249
```



```
### Plot model
```

```
plot(cpt.all[[which.min(BIC(cpt.all))+1]])
abline(v=cpts(cpt.all$meanar2cpt), col='blue')
```



```
### Test for model fitness
```

```
cpt.all.fit = envcpt(all,models=c("meanar2cpt"))$meanar2cpt
```

```
## Fitting 1 models
```

```
## |
```

```
resid.all = all - rep(param.est(cpt.all.fit)$beta[,1],times=seg.len(cpt.all.fit)) -
  rep(param.est(cpt.all.fit)$beta[,2],times=seg.len(cpt.all.fit))*1:length(all)
```

```
## Warning in all - rep(param.est(cpt.all.fit)$beta[, 1], times =
## seg.len(cpt.all.fit)): longer object length is not a multiple of shorter object
## length
```

```
## Warning in rep(param.est(cpt.all.fit)$beta[, 2], times = seg.len(cpt.all.fit))
## * : longer object length is not a multiple of shorter object length
```

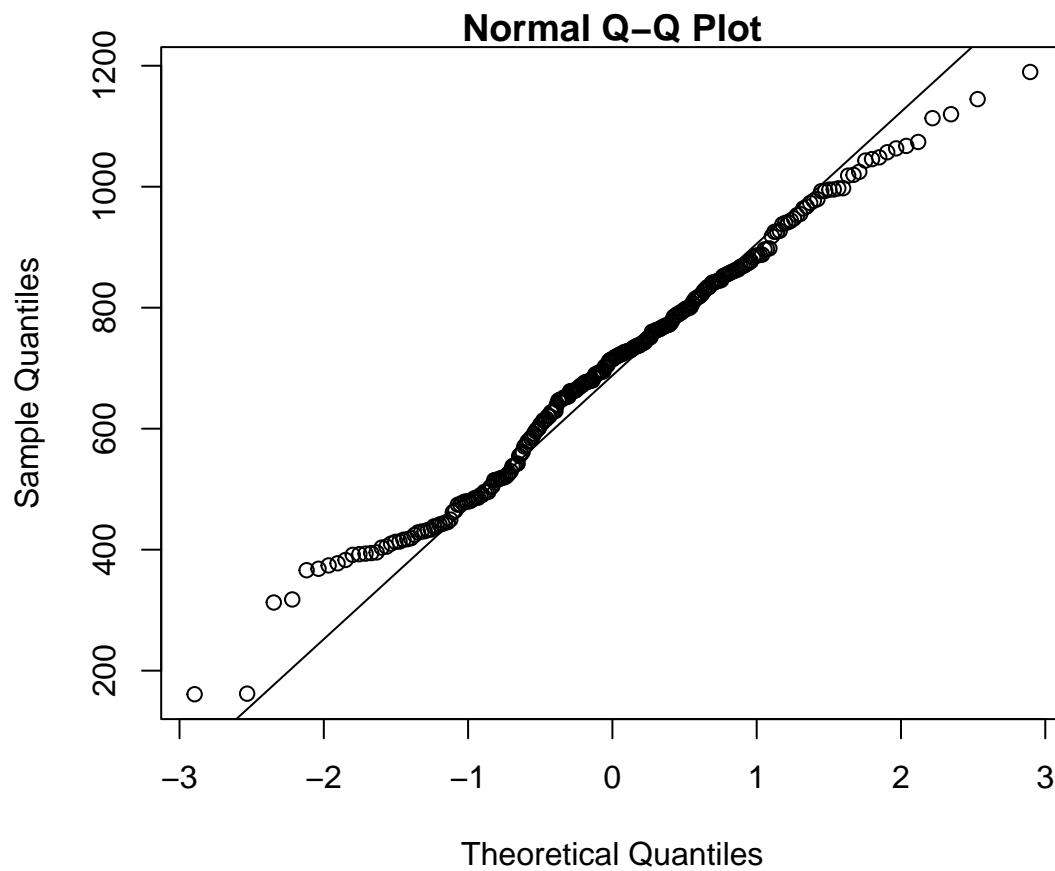
```
corrected.all = resid.all + param.est(cpt.all.fit)$beta[3,1] + param.est(cpt.all.fit)$beta[3,2]*1:length(
shapiro.test(corrected.all)
```

```
##
## Shapiro-Wilk normality test
##
## data:  corrected.all
## W = 0.99007, p-value = 0.06879
```

```
ks.test(resid.all, pnorm, mean = mean(corrected.all), sd = sd(corrected.all))
```

```
##
## Asymptotic one-sample Kolmogorov-Smirnov test
##
## data:  resid.all
## D = 0.58363, p-value < 2.2e-16
## alternative hypothesis: two-sided
```

```
qqnorm(corrected.all)
qqline(corrected.all)
```



2.3.2 Triage category

```
### Primary model
```

```
cpt.green <- envcpt(green)
```

2.3.2.1 Triage category: Green

```
## Fitting 12 models
```

```
## |
```

```
|
```

```
BIC(cpt.green) # returns the BIC for each model.
```

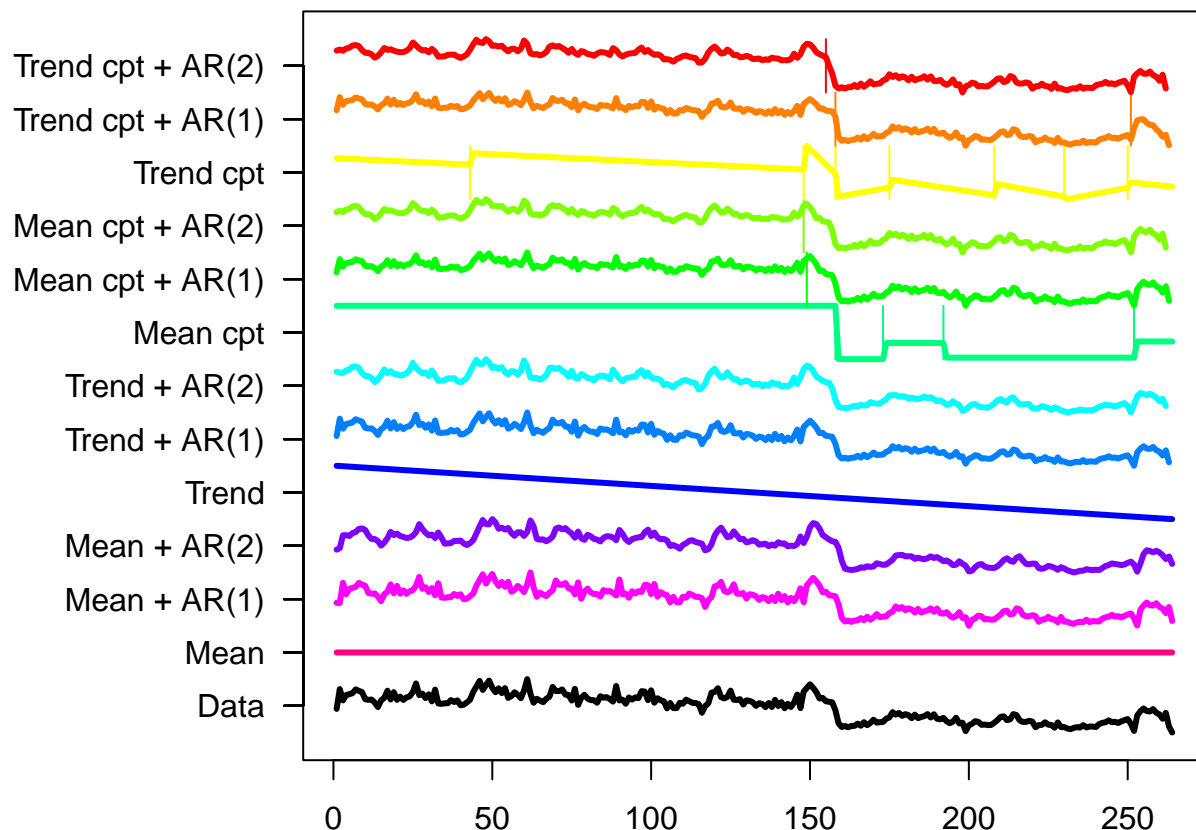
```
##      mean      meancpt      meanar1      meanar2      meanar1cpt      meanar2cpt
## 3608.674 3161.883 3135.582 3109.401 3065.212 3044.690
##      trend      trendcpt      trendar1      trendar2      trendar1cpt      trendar2cpt
## 3341.756 3100.902 3101.359 3064.254 3063.736 3087.412
```

```
which.min(BIC(cpt.green))
```

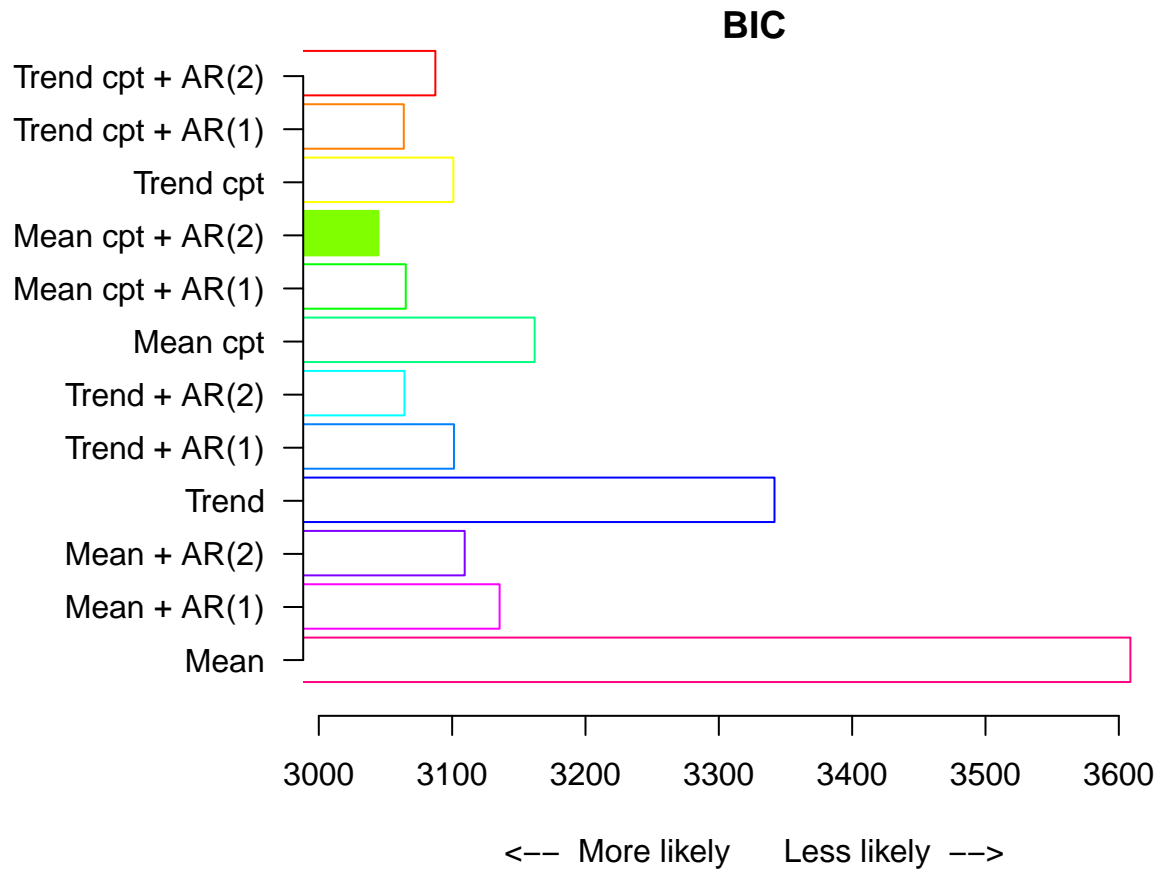
```
## meanar2cpt
```

```
##      6
```

```
plot(cpt.green,type='fit') # plots the fits
```



```
plot(cpt.green,type="bic") # plots the bic values
```



```
### Selection of model based on best BIC
```

```
cpt.green$meanar2cpt
```

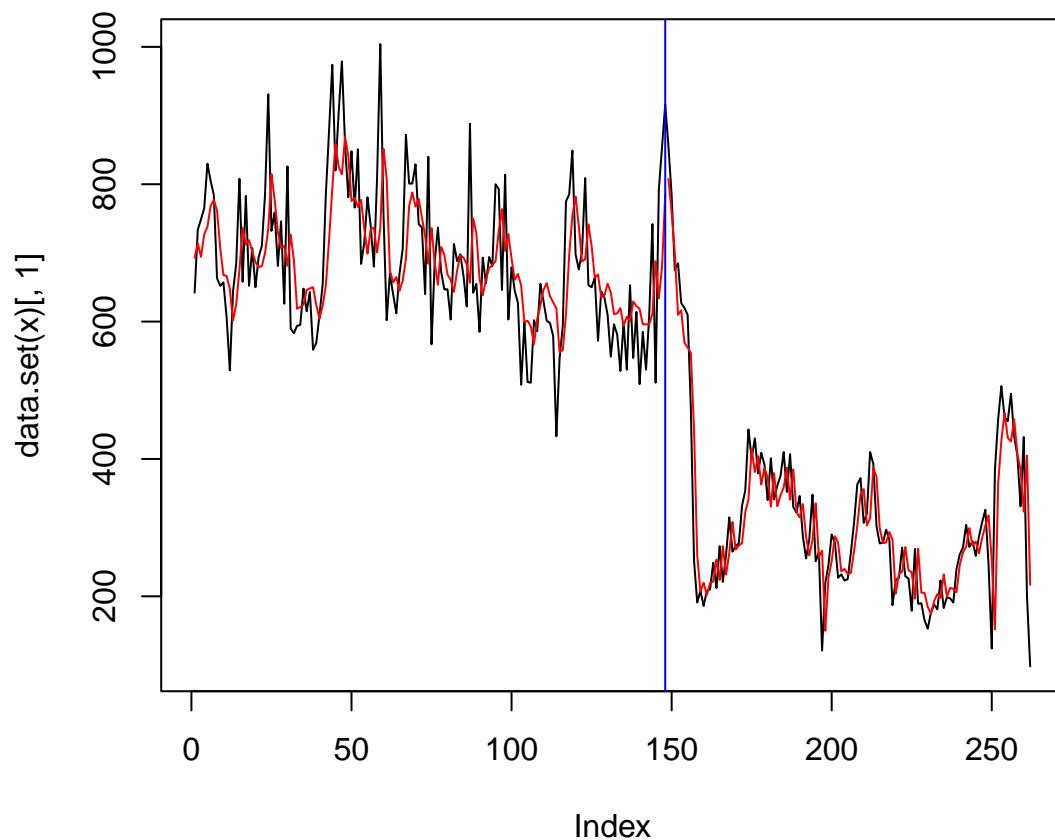
```
## Class 'cpt.reg' : Changepoint Regression Object
##      ~~      : S4 class containing 12 slots with names
##      cpttype date version data.set method test.stat pen.type pen.value minseglen cpts ncpts
##
## Created on   : Sun Jan 29 03:56:54 2023
##
## summary(.)  :
## -----
## Created Using changepoint version 2.2.4
## Changepoint type      : Change in regression
## Method of analysis    : PELT
## Test Statistic       : Normal
## Type of penalty       : MBIC with value, 33.41007
## Maximum no. of cpts   : Inf
## Changepoint Locations : 148
```

```
cpts(cpt.green$meanar2cpt)
```

```
## [1] 148
```

```
### Plot model
```

```
plot(cpt.green[[which.min(BIC(cpt.green))+1]])
abline(v=cpts(cpt.green$meanar2cpt), col='blue')
```



```
### Test for model fitness
```

```
cpt.green.fit = envcpt(all,models=c("meanar2cpt"))$meanar2cpt
```

```
## Fitting 1 models
```

```
## |
```

```
resid.green = green - rep(param.est(cpt.green.fit)$beta[,1],times=seg.len(cpt.green.fit)) -
  rep(param.est(cpt.green.fit)$beta[,2],times=seg.len(cpt.green.fit))*1:length(green)
```

```
## Warning in green - rep(param.est(cpt.green.fit)$beta[, 1], times =
## seg.len(cpt.green.fit)): longer object length is not a multiple of shorter
## object length
```

```
## Warning in rep(param.est(cpt.green.fit)$beta[, 2], times =
## seg.len(cpt.green.fit)) * : longer object length is not a multiple of shorter
## object length
```

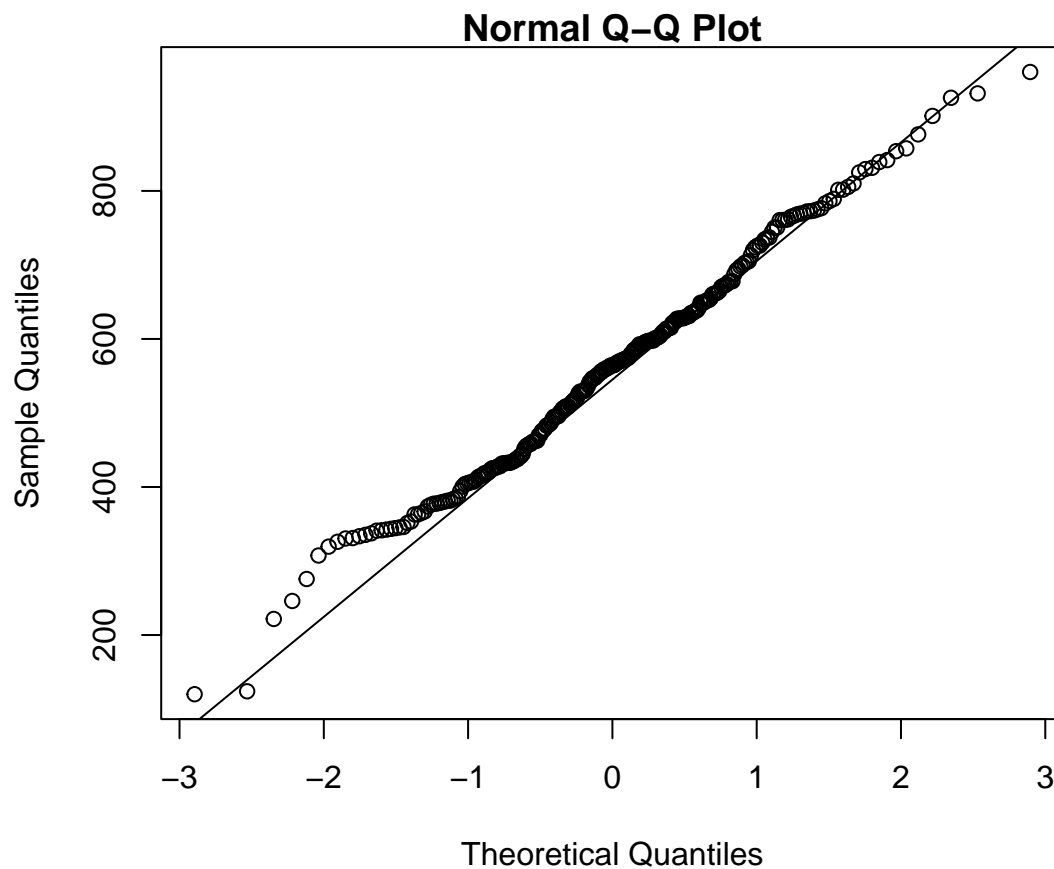
```
corrected.green = resid.green + param.est(cpt.green.fit)$beta[3,1] + param.est(cpt.green.fit)$beta[3,2]  
  
shapiro.test(corrected.green)
```

```
##  
## Shapiro-Wilk normality test  
##  
## data:  corrected.green  
## W = 0.99219, p-value = 0.1776
```

```
ks.test(resid.green, pnorm, mean = mean(corrected.green), sd = sd(corrected.green))
```

```
##  
## Asymptotic one-sample Kolmogorov-Smirnov test  
##  
## data:  resid.green  
## D = 0.68165, p-value < 2.2e-16  
## alternative hypothesis: two-sided
```

```
qqnorm(corrected.green)  
qqline(corrected.green)
```



```
### Primary model
```

```
cpt.yellow <- envcpt(yellow)
```

2.3.2.2 Triage category: Yellow

```
## Fitting 12 models
```

```
## |
```

```
|
```

```
BIC(cpt.yellow) # returns the BIC for each model.
```

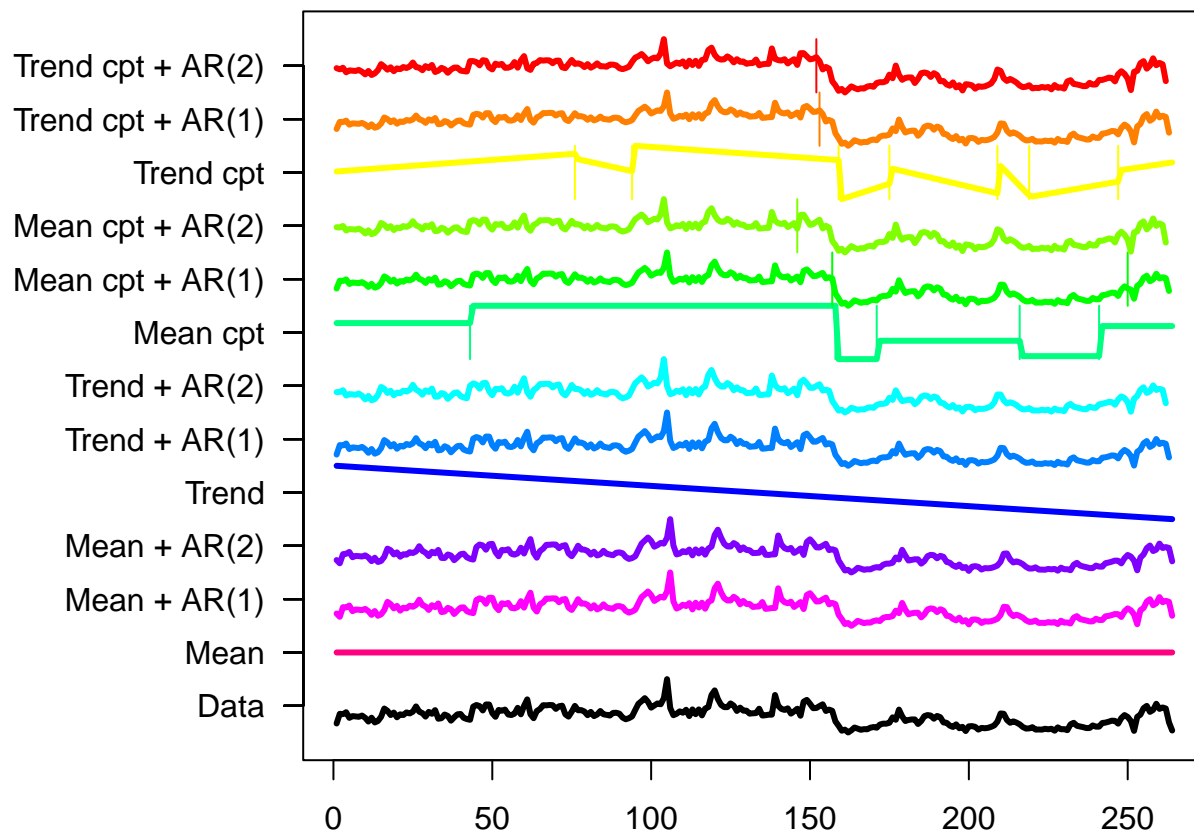
```
##      mean      meancpt      meanar1      meanar2      meanar1cpt      meanar2cpt
##  2821.266  2569.293  2505.815  2510.494  2467.136  2503.449
##      trend      trendcpt      trendar1      trendar2      trendar1cpt      trendar2cpt
##  2769.467  2534.810  2492.921  2487.668  2477.371  2477.080
```

```
which.min(BIC(cpt.yellow))
```

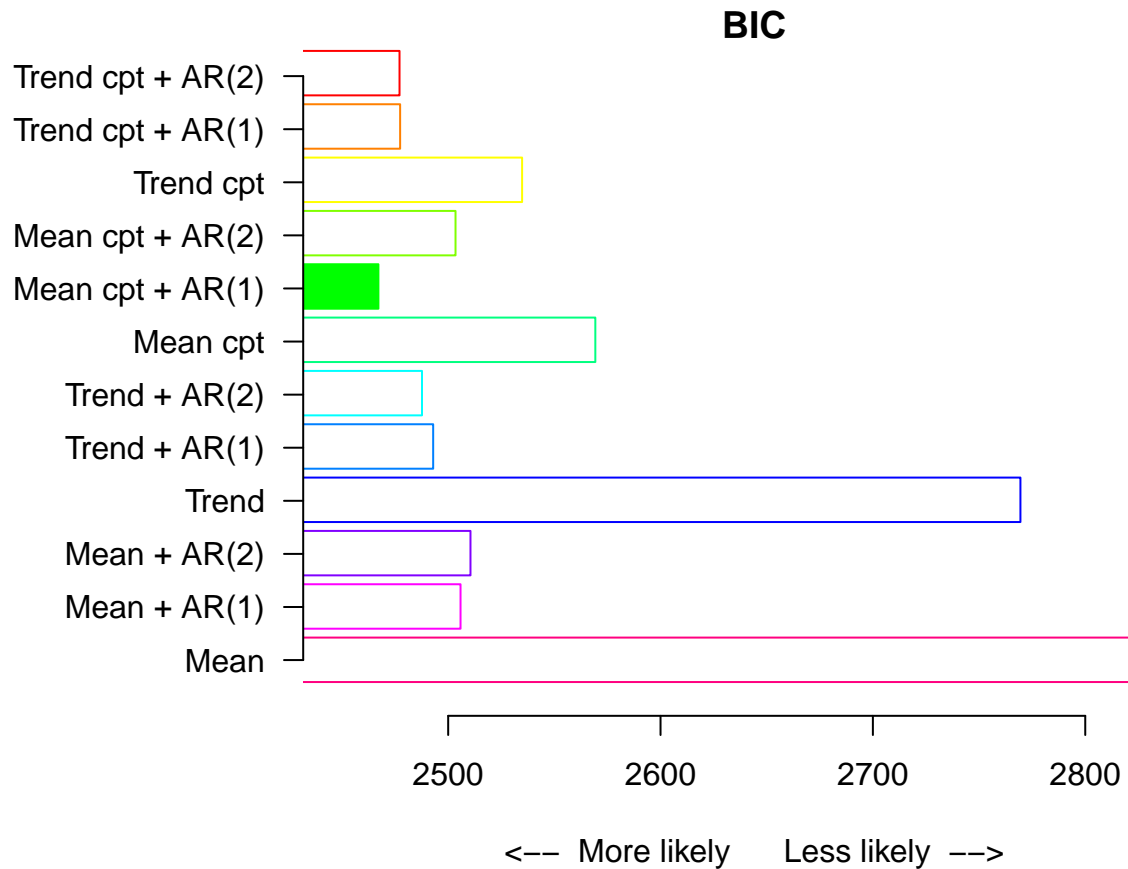
```
## meanar1cpt
```

```
##      5
```

```
plot(cpt.yellow,type='fit') # plots the fits
```



```
plot(cpt.yellow,type="bic") # plots the bic values
```



```
### Selection of model based on best BIC
```

```
cpt.yellow$meanar1cpt
```

```
## Class 'cpt.reg' : Changepoint Regression Object
##      ~~      : S4 class containing 12 slots with names
##      cpttype date version data.set method test.stat pen.type pen.value minseglen cpts ncpts
##
## Created on   : Sun Jan 29 03:56:54 2023
##
## summary(.)   :
## -----
## Created Using changepoint version 2.2.4
## Changepoint type      : Change in regression
## Method of analysis    : PELT
## Test Statistic       : Normal
## Type of penalty       : MBIC with value, 27.86077
## Maximum no. of cpts   : Inf
## Changepoint Locations : 157 250
```

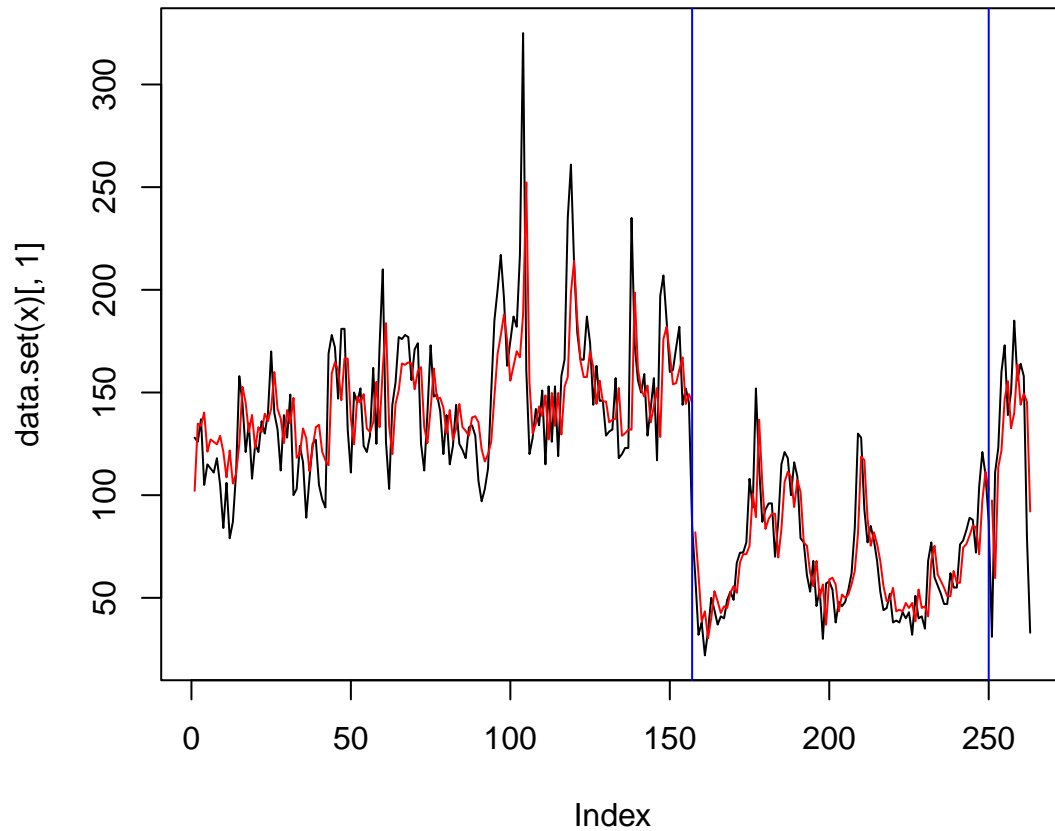
```
cpts(cpt.yellow$meanar1cpt)
```

```
## [1] 157 250
```



```
### Plot model
```

```
plot(cpt.yellow[[which.min(BIC(cpt.yellow))+1]])
abline(v=cpts(cpt.yellow$meanar1cpt), col='blue')
```



```
### Test for model fitness
```

```
cpt.yellow.fit = envcpt(yellow,models=c("meanar1cpt"))$meanar1cpt
```

```
## Fitting 1 models
```

```
## |
```

```
resid.yellow = yellow - rep(param.est(cpt.yellow.fit)$beta[,1],times=seg.len(cpt.yellow.fit)) -
  rep(param.est(cpt.yellow.fit)$beta[,2],times=seg.len(cpt.yellow.fit))*1:length(yellow)
```

```
## Warning in yellow - rep(param.est(cpt.yellow.fit)$beta[, 1], times =
## seg.len(cpt.yellow.fit)): longer object length is not a multiple of shorter
## object length
```

```
## Warning in rep(param.est(cpt.yellow.fit)$beta[, 2], times =
## seg.len(cpt.yellow.fit)) * : longer object length is not a multiple of shorter
## object length
```

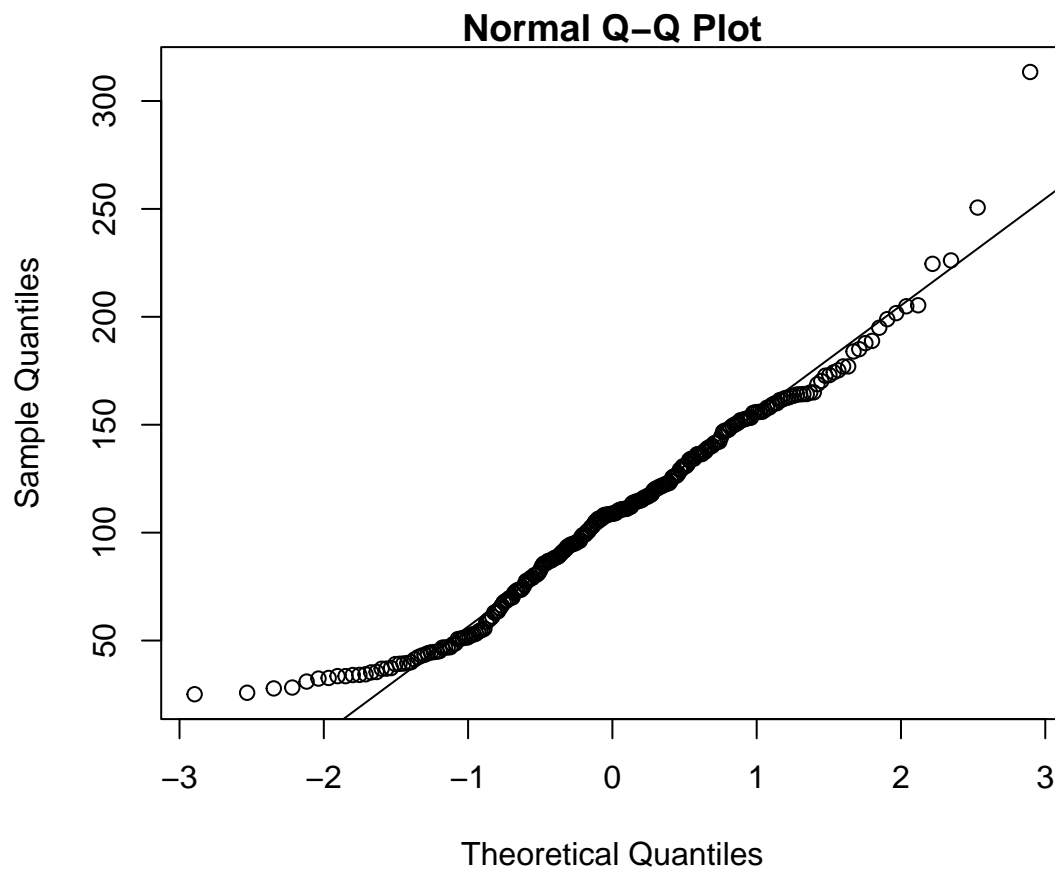
```
corrected.yellow = resid.yellow + param.est(cpt.yellow.fit)$beta[3,1] + param.est(cpt.yellow.fit)$beta[  
shapiro.test(corrected.yellow)
```

```
##  
##  Shapiro-Wilk normality test  
##  
## data:  corrected.yellow  
## W = 0.97209, p-value = 4.913e-05
```

```
ks.test(resid.yellow, pnorm, mean = mean(corrected.yellow), sd = sd(corrected.yellow))
```

```
##  
##  Asymptotic one-sample Kolmogorov-Smirnov test  
##  
## data:  resid.yellow  
## D = 0.69252, p-value < 2.2e-16  
## alternative hypothesis: two-sided
```

```
qqnorm(corrected.yellow)  
qqline(corrected.yellow)
```



```
### Primary model
```

```
cpt.red <- envcpt(red)
```

2.3.2.3 Triage category: Red

```
## Fitting 12 models
```

```
## |
```

```
BIC(cpt.red)
```

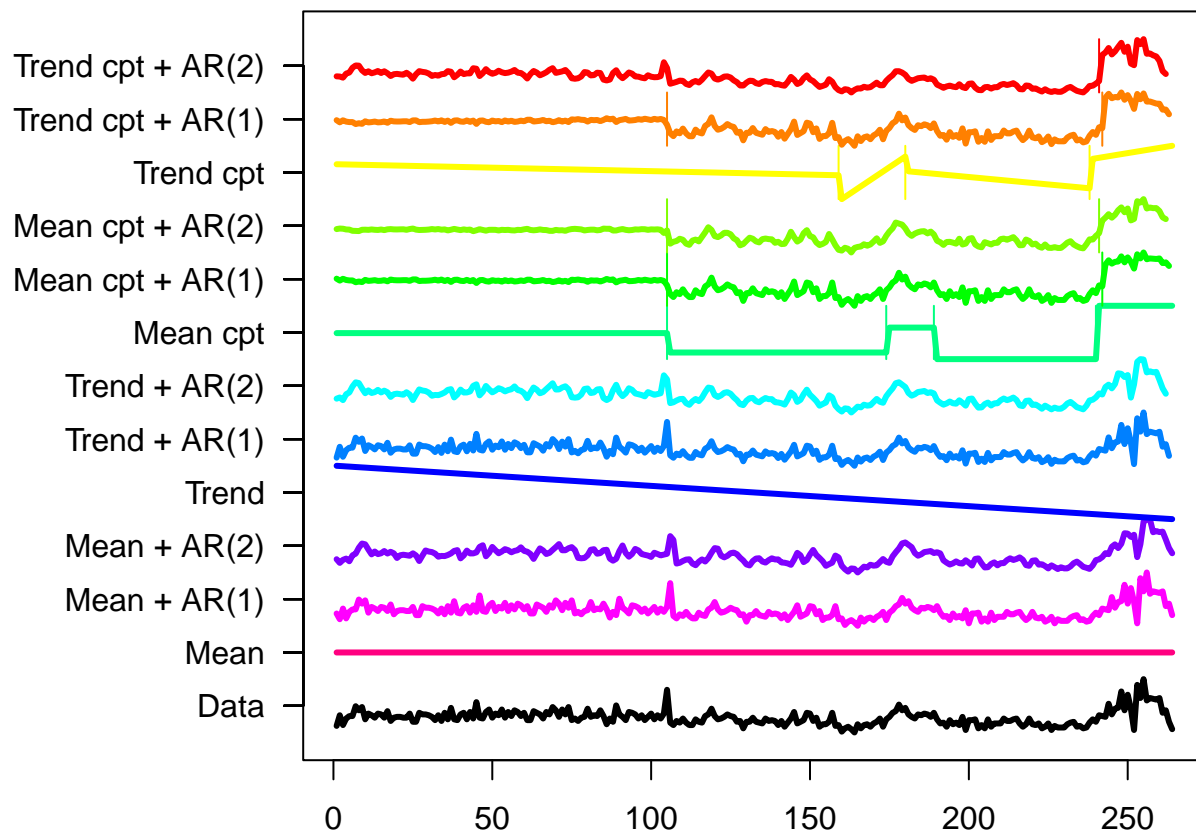
	mean	meancpt	meanar1	meanar2	meanar1cpt	meanar2cpt
##	1940.540	1763.376	1849.088	1813.302	1736.282	1739.086
	trend	trendcpt	trendar1	trendar2	trendar1cpt	trendar2cpt
##	1944.276	1776.489	1846.236	1803.744	1750.518	1790.381

```
which.min(BIC(cpt.red))
```

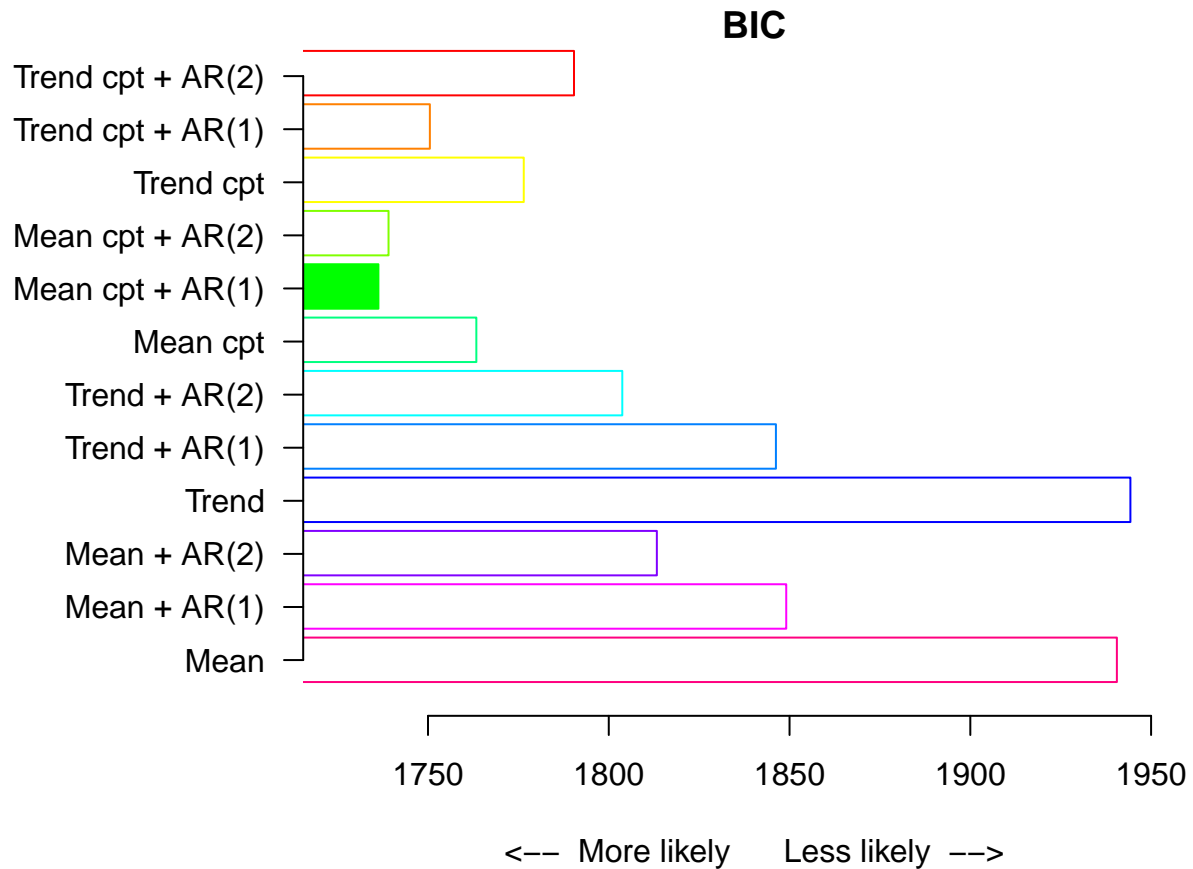
```
## meanar1cpt
```

```
## 5
```

```
plot(cpt.red,type='fit')
```



```
plot(cpt.red,type="bic")
```



```
### Selection of model based on best BIC
```

```
cpt.red$meanar1cpt
```

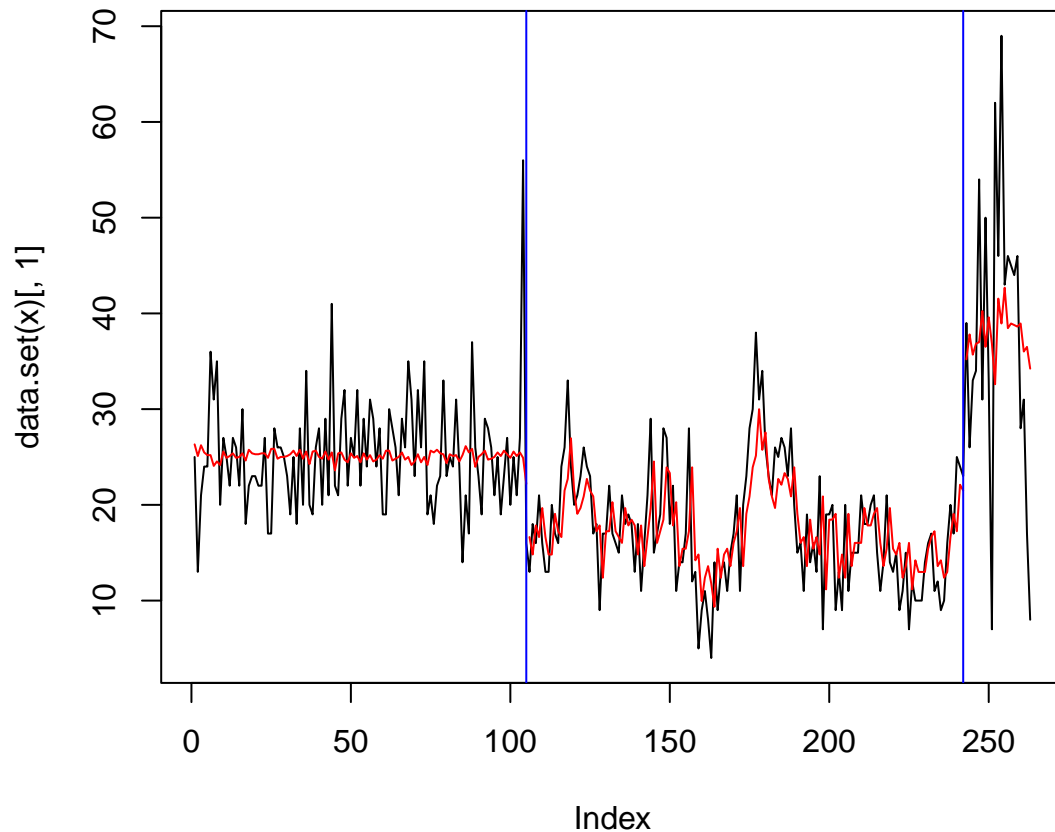
```
## Class 'cpt.reg' : Changepoint Regression Object
##      ~~      : S4 class containing 12 slots with names
##               cpttype date version data.set method test.stat pen.type pen.value minseglen cpts ncpts
##
## Created on   : Sun Jan 29 03:56:54 2023
##
## summary(.)   :
## -----
## Created Using changepoint version 2.2.4
## Changepoint type      : Change in regression
## Method of analysis    : PELT
## Test Statistic       : Normal
## Type of penalty       : MBIC with value, 27.86077
## Maximum no. of cpts   : Inf
## Changepoint Locations : 105 242
```

```
cpts(cpt.red$meanar1cpt)
```

```
## [1] 105 242
```

```
### Plot model
```

```
plot(cpt.red[[which.min(BIC(cpt.red))+1]]) + abline(v=cpts(cpt.red$meanar1cpt), col='blue')
```



```
## integer(0)
```

```
### Test for model fitness
```

```
cpt.red.fit = envcpt(red,models=c("meanar1cpt"))$meanar1cpt
```

```
## Fitting 1 models
```

```
## |
```

```
resid.red = red - rep(param.est(cpt.red.fit)$beta[,1],times=seg.len(cpt.red.fit)) -  
  rep(param.est(cpt.red.fit)$beta[,2],times=seg.len(cpt.red.fit))*1:length(red)
```

```
## Warning in red - rep(param.est(cpt.red.fit)$beta[, 1], times =  
## seg.len(cpt.red.fit)): longer object length is not a multiple of shorter object  
## length
```

```
## Warning in rep(param.est(cpt.red.fit)$beta[, 2], times = seg.len(cpt.red.fit))  
## * : longer object length is not a multiple of shorter object length
```

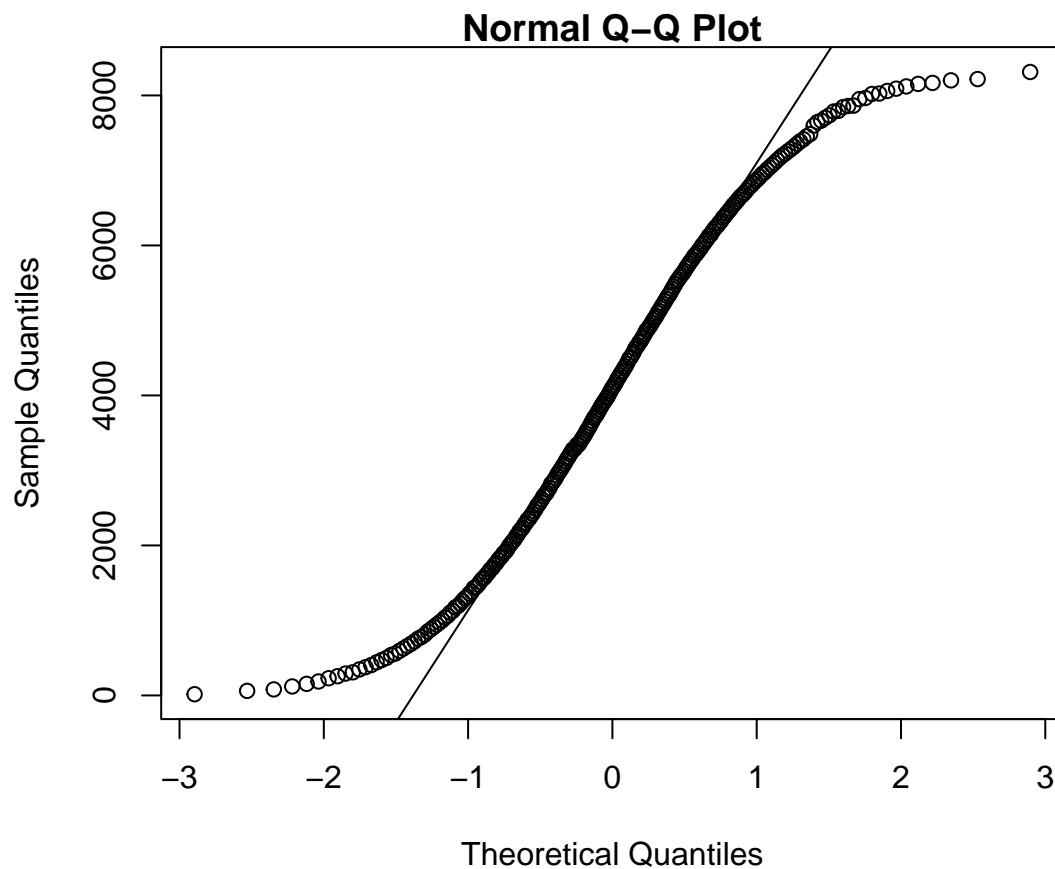
```
corrected.red = resid.red + param.est(cpt.red.fit)$beta[3,2] + param.est(cpt.red.fit)$beta[3,1]*1:length
shapiro.test(corrected.red)
```

```
##
## Shapiro-Wilk normality test
##
## data:  corrected.red
## W = 0.95829, p-value = 6.738e-07
```

```
ks.test(resid.red, pnorm, mean = mean(corrected.red), sd = sd(corrected.red))
```

```
##
## Asymptotic one-sample Kolmogorov-Smirnov test
##
## data:  resid.red
## D = 0.9578, p-value < 2.2e-16
## alternative hypothesis: two-sided
```

```
qqnorm(corrected.red)
qqline(corrected.red)
```



2.3.3 Visit outcome

```
### Primary model
```

```
cpt.discharge <- envcpt(discharge)
```

2.3.3.1 Visit outcome: Discharge

```
## Fitting 12 models
```

```
## |
```

```
BIC(cpt.discharge)
```

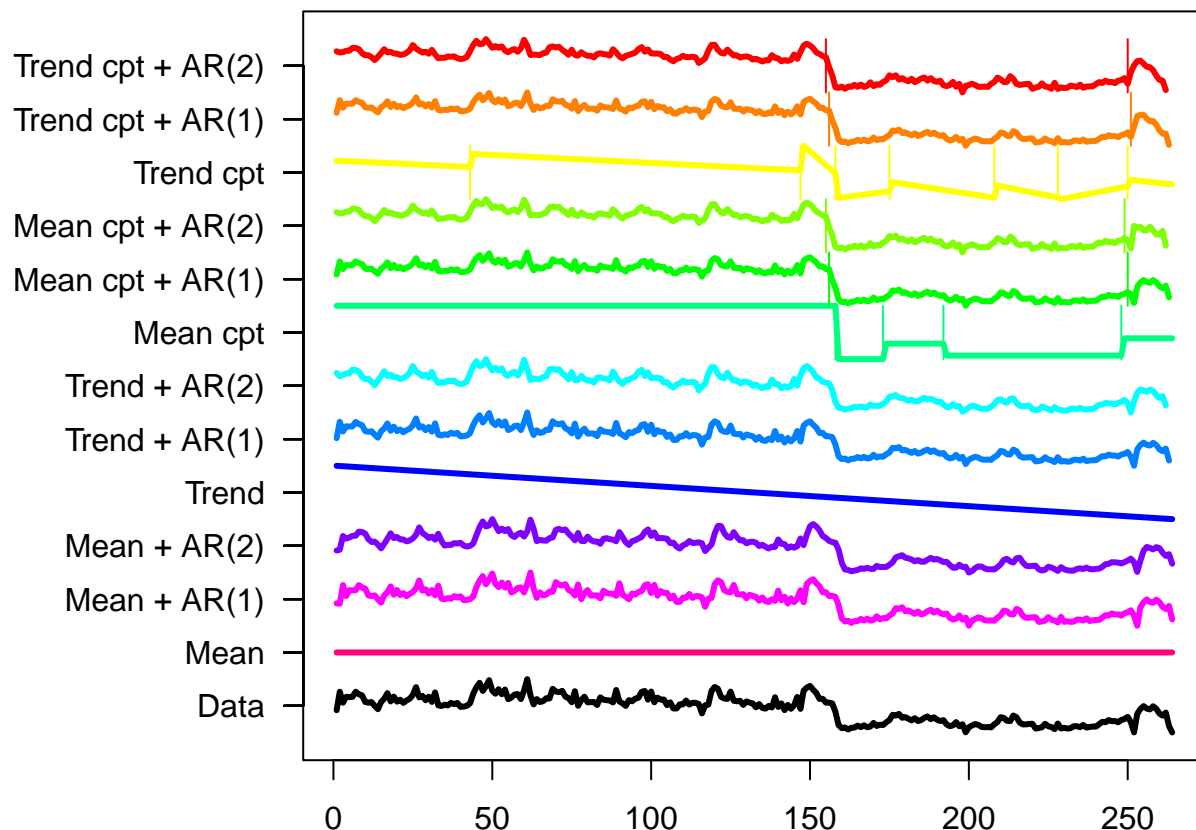
##	mean	meancpt	meanar1	meanar2	meanar1cpt	meanar2cpt
##	3642.000	3193.834	3162.529	3148.642	3085.994	3075.499
##	trend	trendcpt	trendar1	trendar2	trendar1cpt	trendar2cpt
##	3417.751	3142.384	3134.819	3106.721	3092.991	3082.544

```
which.min(BIC(cpt.discharge))
```

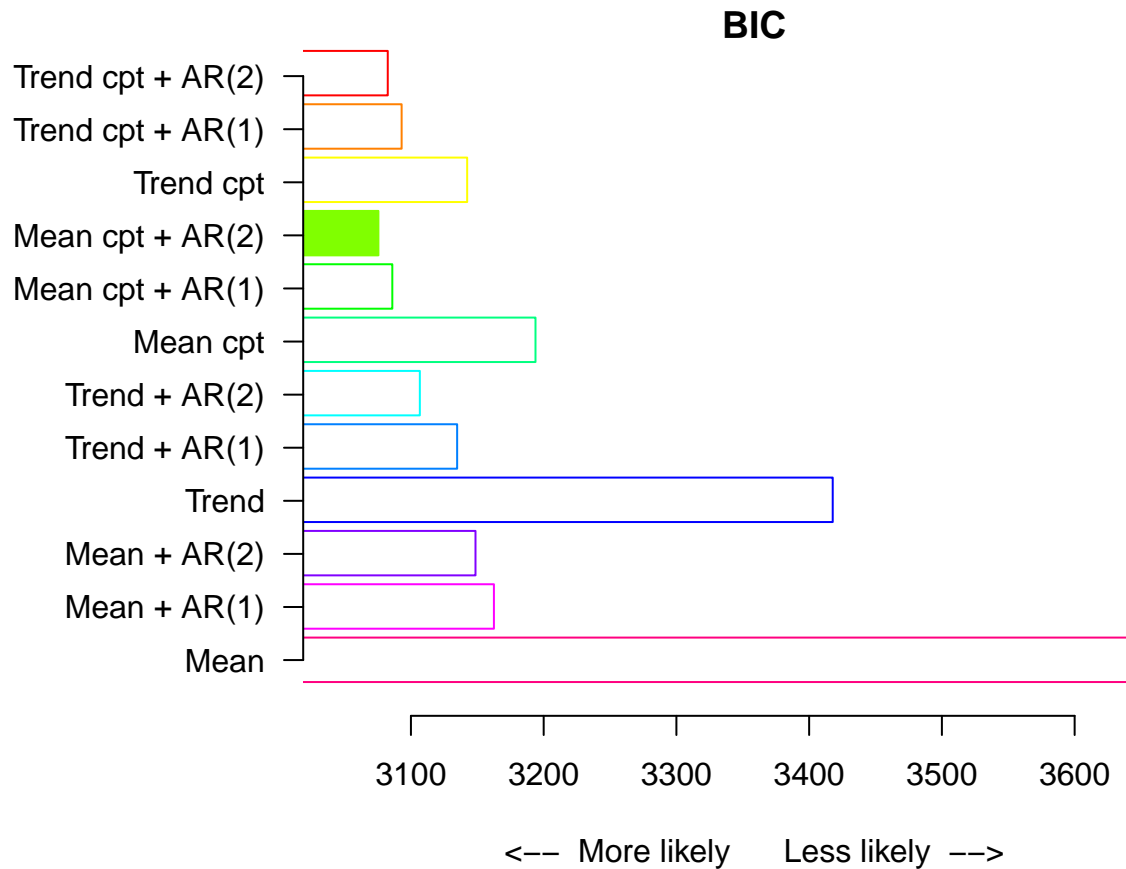
```
## meanar2cpt
```

```
## 6
```

```
plot(cpt.discharge,type='fit')
```



```
plot(cpt.discharge,type="bic")
```



```
### Selection of model based on best BIC
```

```
cpt.discharge$meanar2cpt
```

```
## Class 'cpt.reg' : Changepoint Regression Object
##      ~~      : S4 class containing 12 slots with names
##      cpttype date version data.set method test.stat pen.type pen.value minseglen cpts ncpts
##
## Created on   : Sun Jan 29 03:56:54 2023
##
## summary(.)  :
## -----
## Created Using changepoint version 2.2.4
## Changepoint type      : Change in regression
## Method of analysis     : PELT
## Test Statistic        : Normal
## Type of penalty       : MBIC with value, 33.41007
## Maximum no. of cpts   : Inf
## Changepoint Locations : 155 249
```

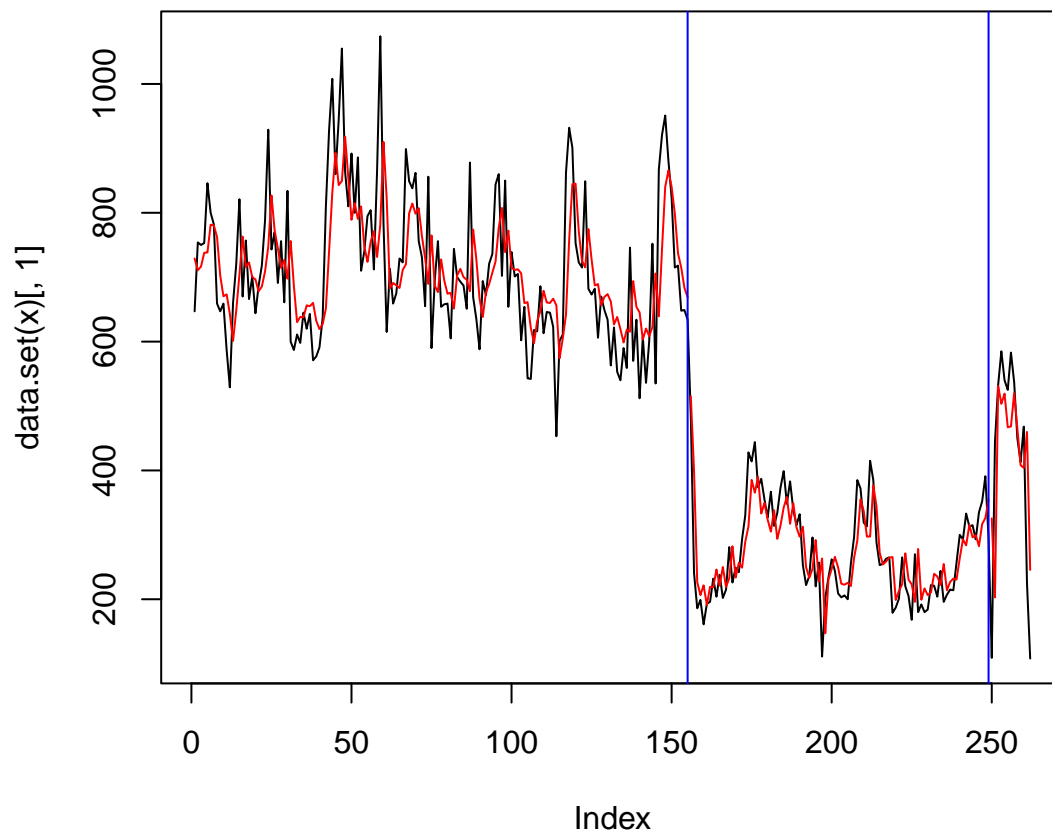
```
cpts(cpt.discharge$meanar2cpt)
```

```
## [1] 155 249
```



```
### Plot model
```

```
plot.discharge <- plot(cpt.discharge[[which.min(BIC(cpt.discharge))+1]]) + abline(v=cpts(cpt.discharge$
```



```
### Test for model fitness
```

```
cpt.discharge.fit = envcpt(discharge,models=c("meanar2cpt"))$meanar2cpt
```

```
## Fitting 1 models
```

```
## |
```

```
resid.discharge = discharge - rep(param.est(cpt.discharge.fit)$beta[,1],times=seg.len(cpt.discharge.fit
```

```
## Warning in discharge - rep(param.est(cpt.discharge.fit)$beta[, 1], times =  
## seg.len(cpt.discharge.fit)): longer object length is not a multiple of shorter  
## object length
```

```
## Warning in rep(param.est(cpt.discharge.fit)$beta[, 2], times =  
## seg.len(cpt.discharge.fit)) * : longer object length is not a multiple of  
## shorter object length
```

```
corrected.discharge = resid.discharge + param.est(cpt.discharge.fit)$beta[3,2] + param.est(cpt.discharge)

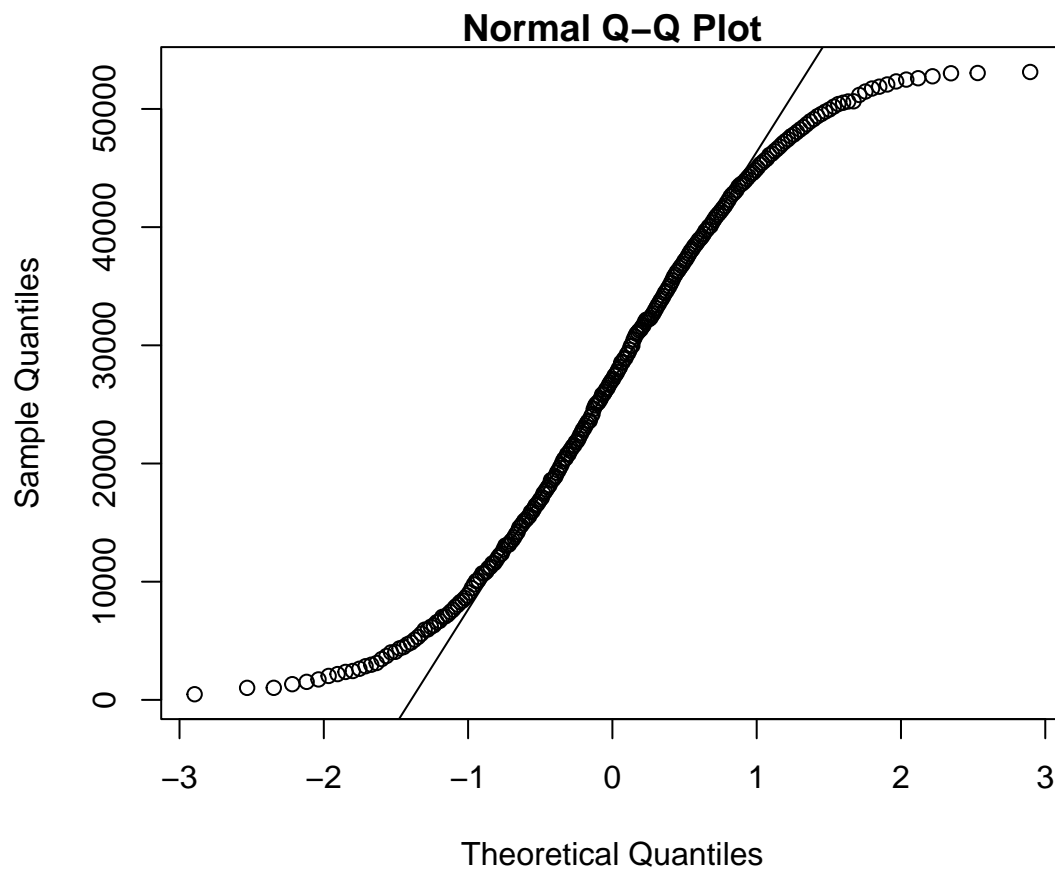
shapiro.test(corrected.discharge)
```

```
##
##  Shapiro-Wilk normality test
##
## data:  corrected.discharge
## W = 0.9556, p-value = 3.211e-07
```

```
ks.test(resid.discharge, pnorm, mean = mean(corrected.discharge), sd = sd(corrected.discharge))
```

```
##
##  Asymptotic one-sample Kolmogorov-Smirnov test
##
## data:  resid.discharge
## D = 0.95746, p-value < 2.2e-16
## alternative hypothesis: two-sided
```

```
qqnorm(corrected.discharge)
qqline(corrected.discharge)
```



```
### Primary model
```

```
cpt.admission <- envcpt(admission)
```

2.3.3.2 Visit outcome: Admission

```
## Fitting 12 models
```

```
## |
```

```
BIC(cpt.admission)
```

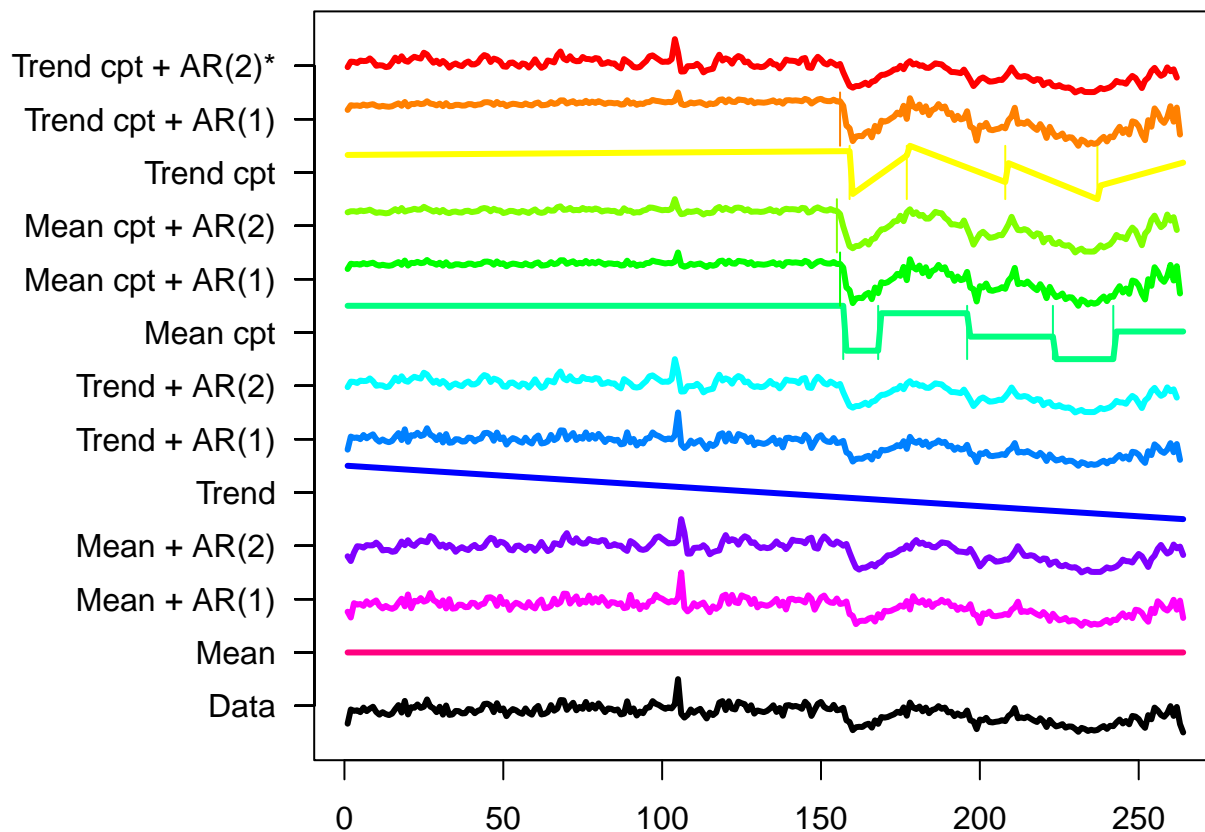
	mean	meancpt	meanar1	meanar2	meanar1cpt	meanar2cpt
##	2669.748	2450.924	2476.576	2451.812	2418.401	2410.359
	trend	trendcpt	trendar1	trendar2	trendar1cpt	trendar2cpt
##	2579.954	2438.783	2447.642	2421.020	2426.534	2454.476

```
which.min(BIC(cpt.admission))
```

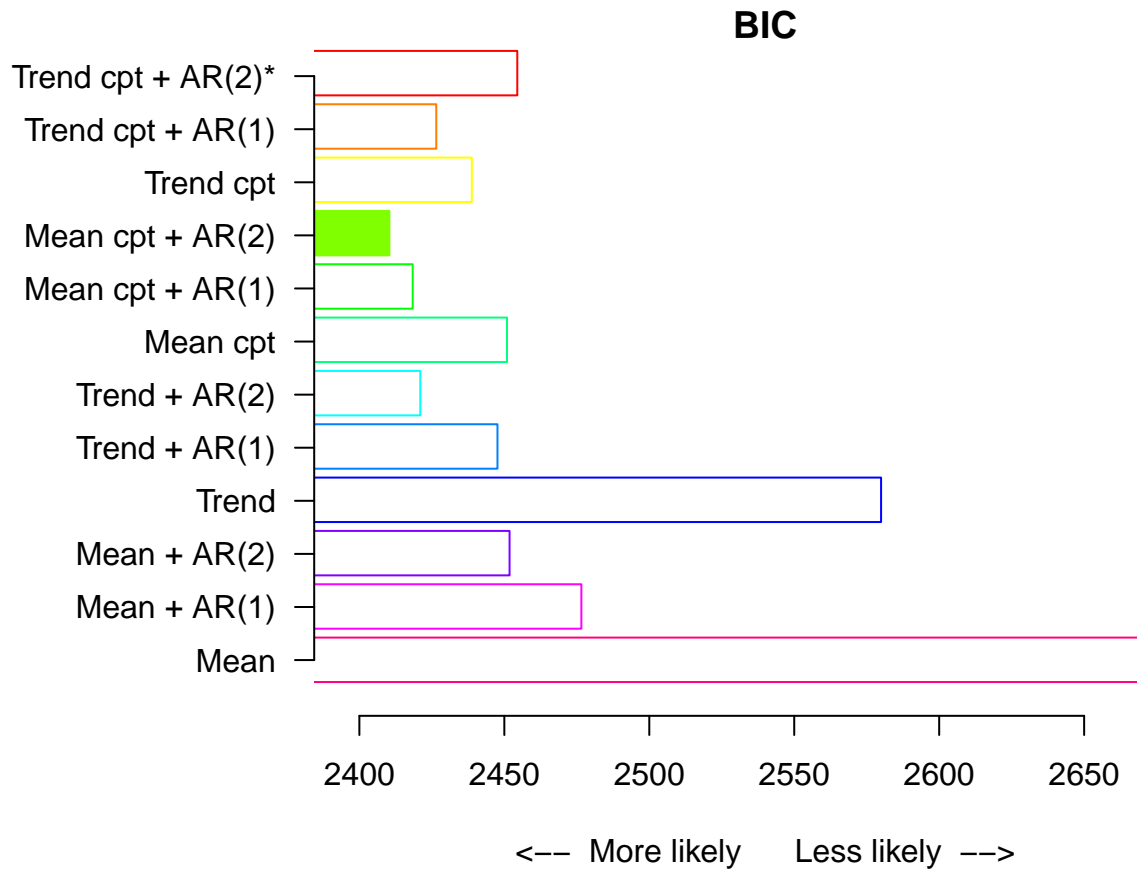
```
## meanar2cpt
```

```
## 6
```

```
plot(cpt.admission,type='fit')
```



```
plot(cpt.admission,type="bic")
```



```
### Selection of model based on best BIC
```

```
cpt.admission$meanar2cpt
```

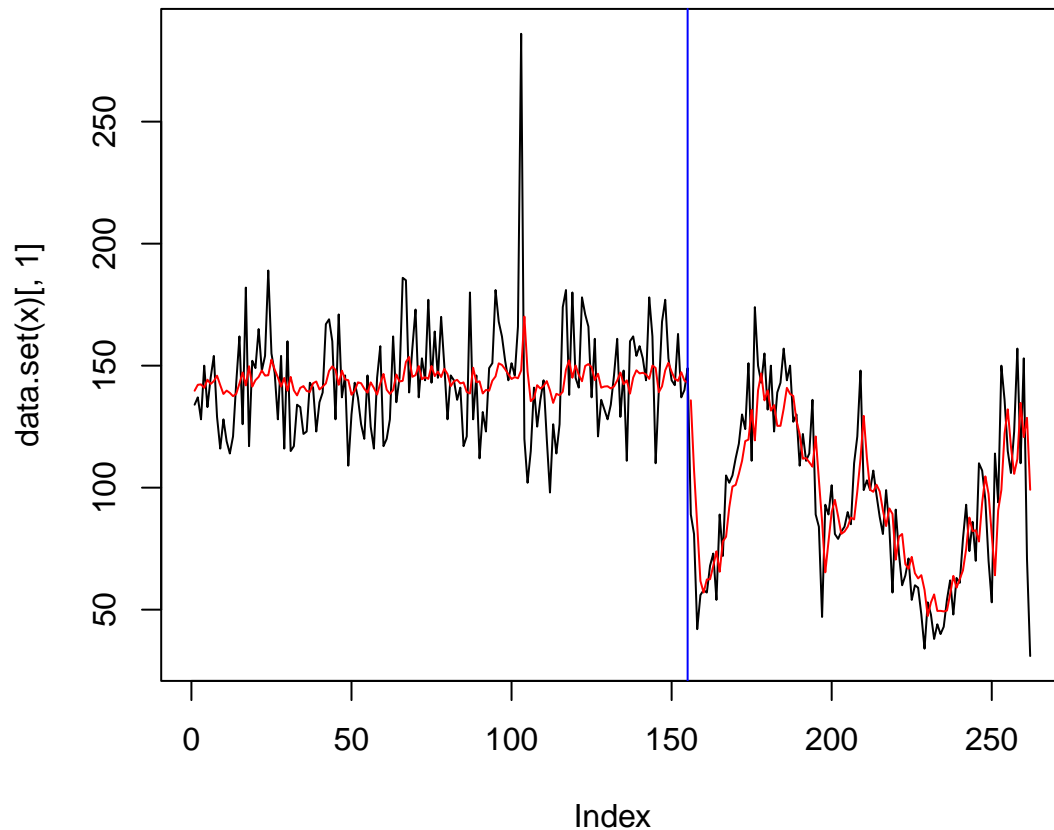
```
## Class 'cpt.reg' : Changepoint Regression Object
##      ~~      : S4 class containing 12 slots with names
##      cpttype date version data.set method test.stat pen.type pen.value minseglen cpts ncpts
##
## Created on   : Sun Jan 29 03:56:54 2023
##
## summary(.)   :
## -----
## Created Using changepoint version 2.2.4
## Changepoint type      : Change in regression
## Method of analysis    : PELT
## Test Statistic       : Normal
## Type of penalty       : MBIC with value, 33.41007
## Maximum no. of cpts   : Inf
## Changepoint Locations : 155
```

```
cpts(cpt.admission$meanar2cpt)
```

```
## [1] 155
```

```
### Plot model
```

```
plot(cpt.admission[[which.min(BIC(cpt.admission))+1]]) + abline(v=cpts(cpt.admission$meanar2cpt), col='blue')
```



```
## integer(0)
```

```
### Test for model fitness
```

```
cpt.admission.fit = envcpt(admission,models=c("meanar2cpt"))$meanar2cpt
```

```
## Fitting 1 models
```

```
## |
```

```
resid.admission = admission - (rep(param.est(cpt.admission.fit)$beta[,1],times=seg.len(cpt.admission.fit)))
```

```
## Warning in admission - (rep(param.est(cpt.admission.fit)$beta[, 1], times =  
## seg.len(cpt.admission.fit))): longer object length is not a multiple of shorter  
## object length
```

```
## Warning in rep(param.est(cpt.admission.fit)$beta[, 2], times =  
## seg.len(cpt.admission.fit)) * : longer object length is not a multiple of  
## shorter object length
```

```
corrected.admission = resid.admission + param.est(cpt.admission.fit)$beta[2,2] + param.est(cpt.admission)

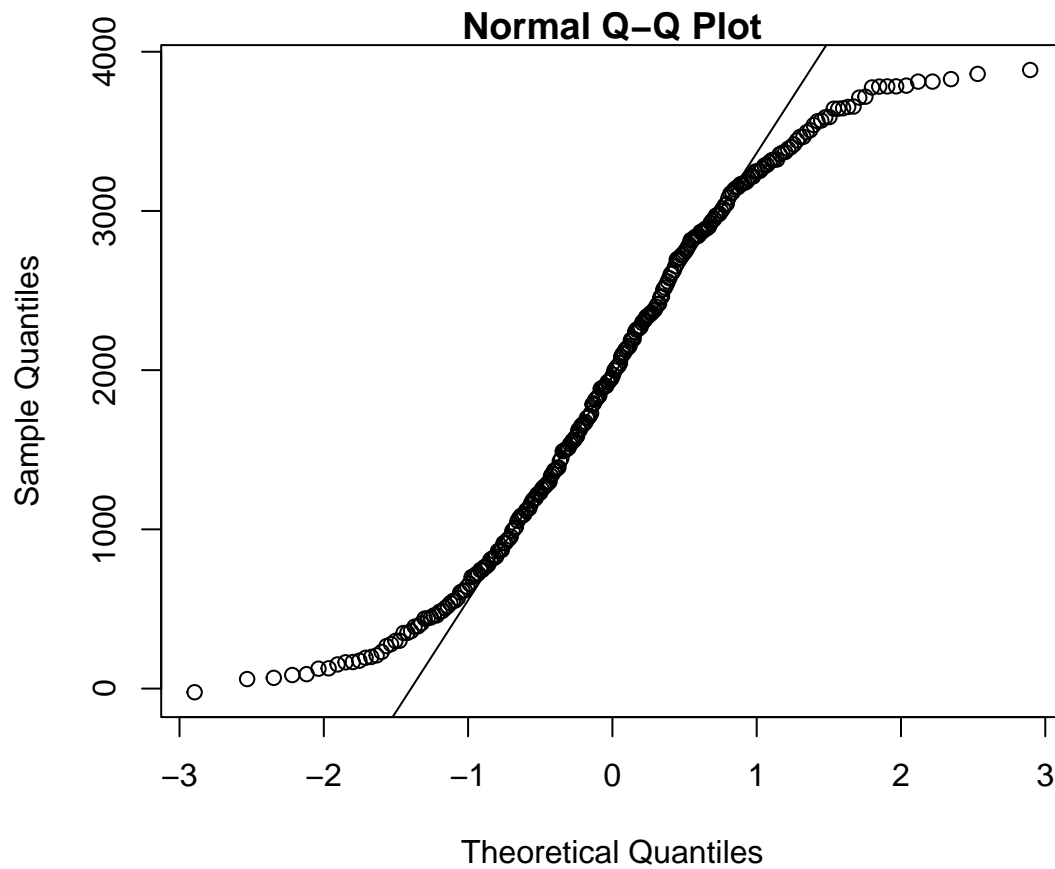
shapiro.test(corrected.admission)
```

```
##
##  Shapiro-Wilk normality test
##
## data:  corrected.admission
## W = 0.95534, p-value = 2.989e-07
```

```
ks.test(resid.admission, pnorm, mean = mean(corrected.admission), sd = sd(corrected.admission))
```

```
##
##  Asymptotic one-sample Kolmogorov-Smirnov test
##
## data:  resid.admission
## D = 0.95244, p-value < 2.2e-16
## alternative hypothesis: two-sided
```

```
qqnorm(corrected.admission)
qqline(corrected.admission)
```



```
### Primary model
```

```
cpt.death <- envcpt(death)
```

2.3.3.3 Visit outcome: Death

```
## Fitting 12 models
```

```
## |
```

```
|
```

```
BIC(cpt.death)
```

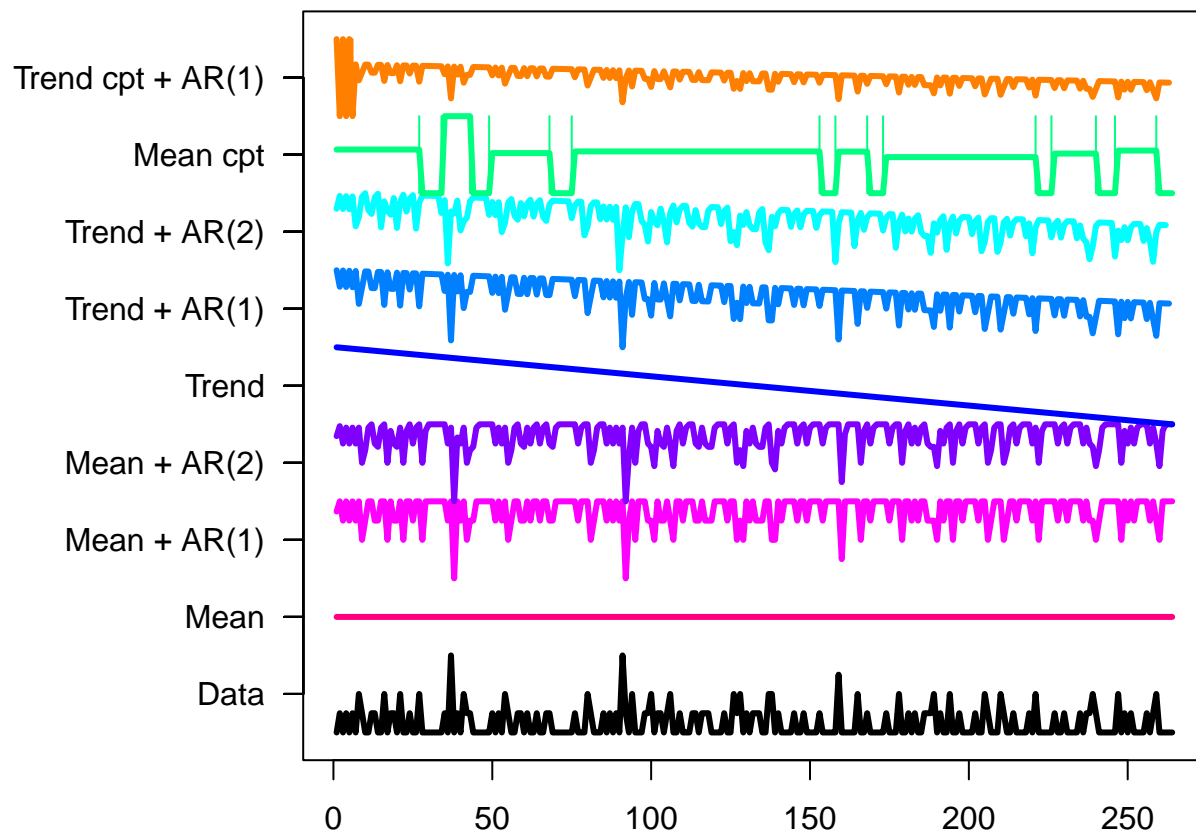
##	mean	meancpt	meanar1	meanar2	meanar1cpt	meanar2cpt
##	600.3316	-Inf	602.9307	608.4173	NaN	NaN
##	trend	trendcpt	trendar1	trendar2	trendar1cpt	trendar2cpt
##	604.3545	NaN	604.6745	608.6632	366.4741	NaN

```
which.min(BIC(cpt.death))
```

```
## meancpt
```

```
## 2
```

```
plot(cpt.death,type='fit')
```



```
### Selection of model based on best BIC
```

```
cpt.death$trendar1cpt
```

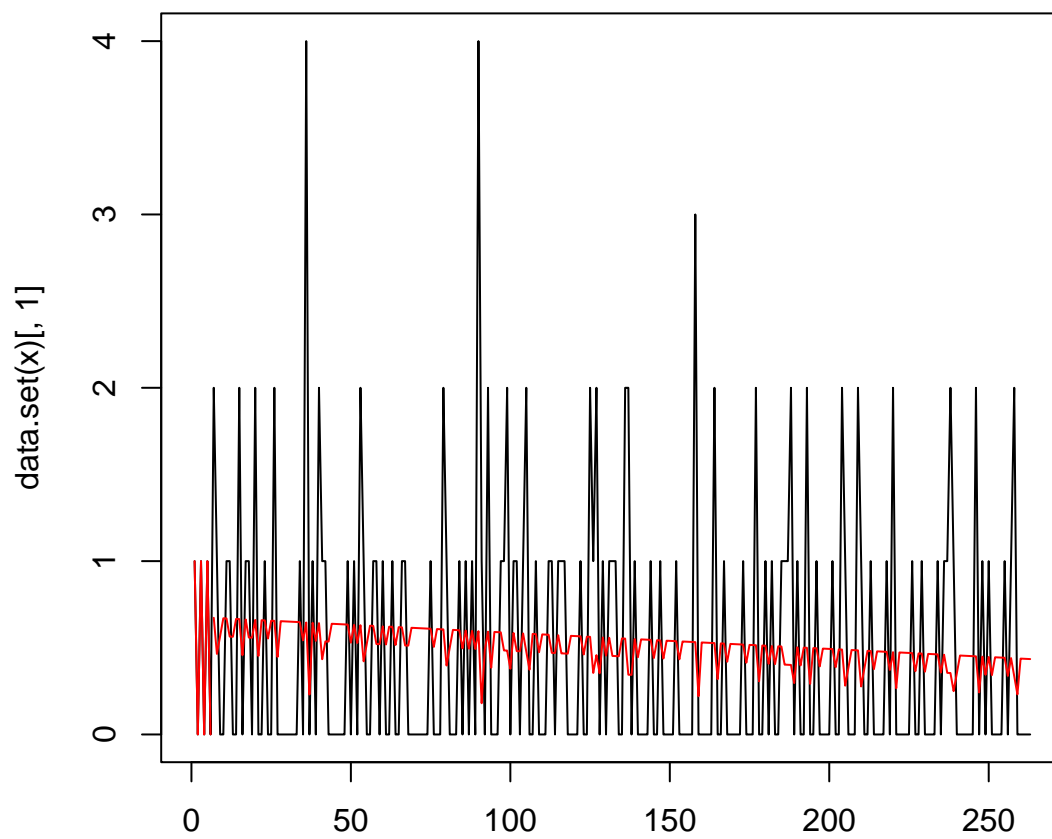
```
## Class 'cpt.reg' : Changepoint Regression Object
##      ~~~~ : S4 class containing 12 slots with names
##      cpttype date version data.set method test.stat pen.type pen.value minseglen cpts ncpts
##
## Created on   : Sun Jan 29 03:56:54 2023
##
## summary(.) :
## -----
## Created Using changepoint version 2.2.4
## Changepoint type      : Change in regression
## Method of analysis    : PELT
## Test Statistic       : Normal
## Type of penalty       : MBIC with value, 33.43292
## Maximum no. of cpts   : Inf
## Changepoint Locations : 6
```

```
cpts(cpt.death$trendar1cpt)
```

```
## [1] 6
```

```
### Plot model
```

```
plot(cpt.death$trendar1cpt)
```

```
### Test for model fitness
```

```
cpt.death.fit = envcpt(death,models=c("trendar1cpt"))$trendar1cpt
```

```
## Fitting 1 models
```

```
## |
```

```
resid.death = death - (rep(param.est(cpt.death.fit)$beta[,1],times=seg.len(cpt.death.fit))) - (rep(param
```

```
## Warning in death - (rep(param.est(cpt.death.fit)$beta[, 1], times =
```

```
## seg.len(cpt.death.fit)): longer object length is not a multiple of shorter
```

```
## object length
```

```
## Warning in rep(param.est(cpt.death.fit)$beta[, 2], times =
```

```
## seg.len(cpt.death.fit)) * : longer object length is not a multiple of shorter
```

```
## object length
```

```
corrected.death = resid.death + param.est(cpt.death.fit)$beta[2,2] + param.est(cpt.death.fit)$beta[2,1]
```

```
shapiro.test(corrected.admission)
```

```
##
```

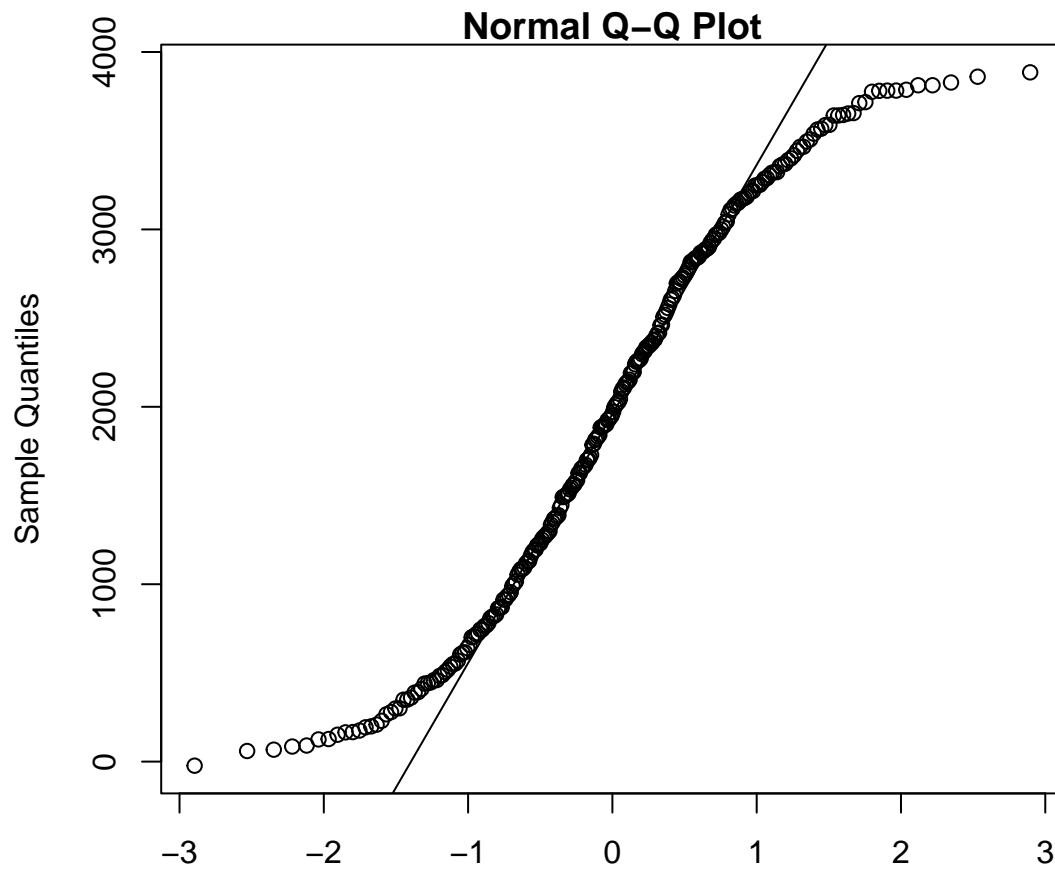
```
## Shapiro-Wilk normality test
```

```
##
## data:  corrected.admission
## W = 0.95534, p-value = 2.989e-07

ks.test(resid.admission, pnorm, mean = mean(corrected.admission), sd = sd(corrected.admission))

##
## Asymptotic one-sample Kolmogorov-Smirnov test
##
## data:  resid.admission
## D = 0.95244, p-value < 2.2e-16
## alternative hypothesis: two-sided

qqnorm(corrected.admission)
qqline(corrected.admission)
```



2.3.4 ED discharge diagnosis

```
my_theme <- function() {
  # Colors
```

```

color.background = "white"
color.text = "#22211d"

# Begin construction of chart
theme_bw(base_size=15) +

  # Format background colors
  theme(panel.background = element_rect(fill=color.background, color=color.background)) +
  theme(plot.background = element_rect(fill=color.background, color=color.background)) +
  theme(panel.border = element_rect(color=color.background)) +
  theme(strip.background = element_rect(fill=color.background, color=color.background)) +

  # Format the grid
  theme(panel.grid.major.y = element_blank()) +
  theme(panel.grid.minor.y = element_blank()) +
  theme(axis.ticks = element_blank()) +

  # Format the legend
  theme(legend.position = "none") +

  # Format title and axis labels
  theme(plot.title = element_text(color=color.text, size=20, face = "bold")) +
  theme(axis.title.x = element_text(size=14, color="black", face = "bold")) +
  theme(axis.title.y = element_text(size=14, color="black", face = "bold", vjust=1.25)) +
  theme(axis.text.x = element_text(size=10, vjust=0.5, hjust=0.5, color = color.text)) +
  theme(axis.text.y = element_text(size=10, color = color.text)) +
  theme(strip.text = element_text(face = "bold")) +

  # Plot margins
  theme(plot.margin = unit(c(0.35, 0.2, 0.3, 0.35), "cm"))
}

```

```

rank <- read.csv('dxrank.csv')

rank$Period2 <- as.factor(rank$Period2)

df.rank <- rank %>%
  group_by(Period) %>%
  arrange(Period2, Rank, Diagnosis) %>%
  mutate(ranking = row_number()) %>%
  as.data.frame()

head(df.rank)

```

2.3.4.1 Bump chart top 10 ED discharge diagnosis

##	Period	Period2	Diagnosis	Rank
## 1	PreMCO	1	Mental, behavioral and neurodevelopmental disorders	1
## 2	PreMCO	1	Pregnancy, childbirth, and the puerperium	2
## 3	PreMCO	1	Poisoning/ ingestion	3
## 4	PreMCO	1	Oral cavity and salivary glands	4
## 5	PreMCO	1	Abuse, neglect, and other maltreatment	5

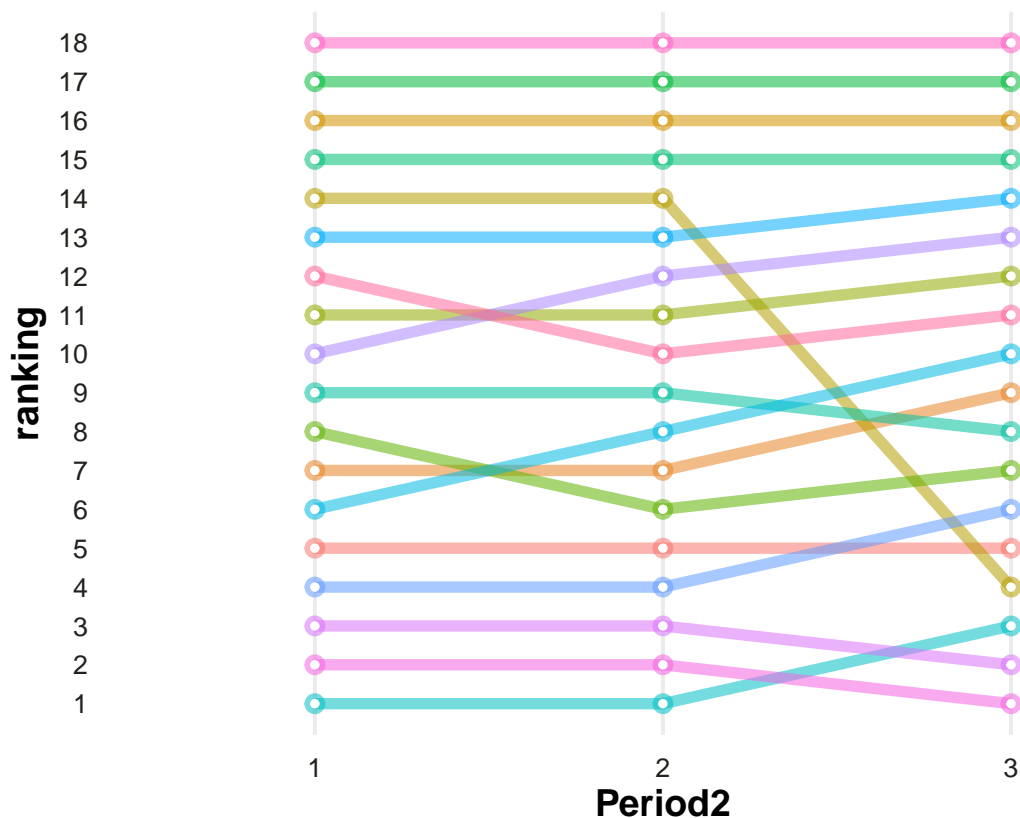
```
## 6 PreMCO      1      Musculoskeletal system and connective tissue      6
## Patients Percentage ranking
## 1      138 0.08332025      1
## 2      185 0.11169744      2
## 3      472 0.28497941      3
## 4      582 0.35139411      4
## 5      978 0.59048700      5
## 6     1182 0.71365607      6
```

```
library(ggbump)
```

```
show.top.n <- 10
```

```
ggplot(data = df.rank, aes(x = Period2, y = ranking, group = Diagnosis)) +
  geom_line(aes(color = Diagnosis, alpha = 1), size = 2) +
  geom_point(aes(color = Diagnosis, alpha = 1), size = 3) +
  geom_point(color = "#FFFFFF", size = 1) +
  scale_y_reverse(breaks = 1:18) +
  coord_cartesian(ylim = c(1,18)) +
  theme(legend.position = "none") +
  my_theme()
```

```
## Warning: Using 'size' aesthetic for lines was deprecated in ggplot2 3.4.0.
## i Please use 'linewidth' instead.
```



```
### Primary model
```

```
cpt.respi <- envcpt(respi)
```

2.3.4.2 Diagnosis: Respi

```
## Fitting 12 models
```

```
## |
```

```
|
```

```
BIC(cpt.respi)
```

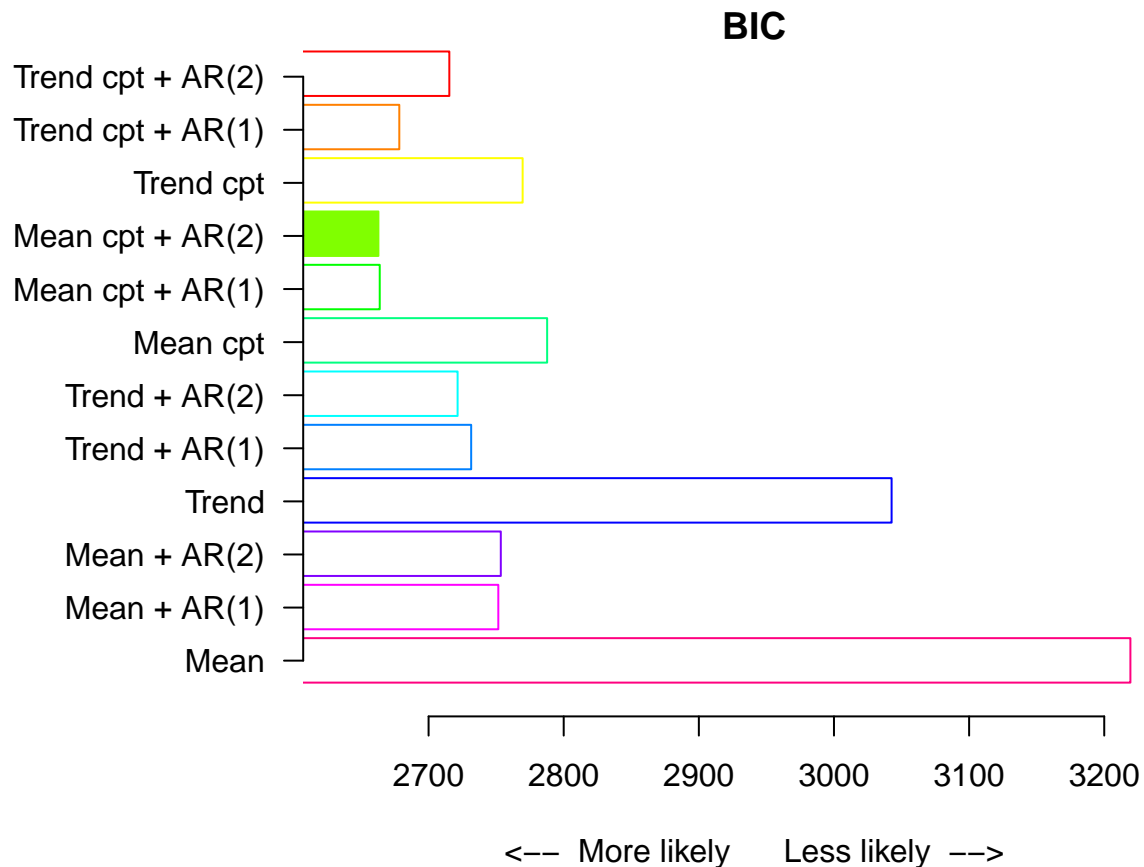
```
##      mean      meancpt      meanar1      meanar2      meanar1cpt      meanar2cpt
## 3219.356 2787.755 2751.611 2753.458 2663.856 2662.868
##      trend      trendcpt      trendar1      trendar2      trendar1cpt      trendar2cpt
## 3042.656 2769.607 2731.536 2721.489 2678.349 2715.359
```

```
which.min(BIC(cpt.respi))
```

```
## meanar2cpt
```

```
##      6
```

```
plot(cpt.respi,type="bic")
```



```
### Selection of model based on best BIC
```

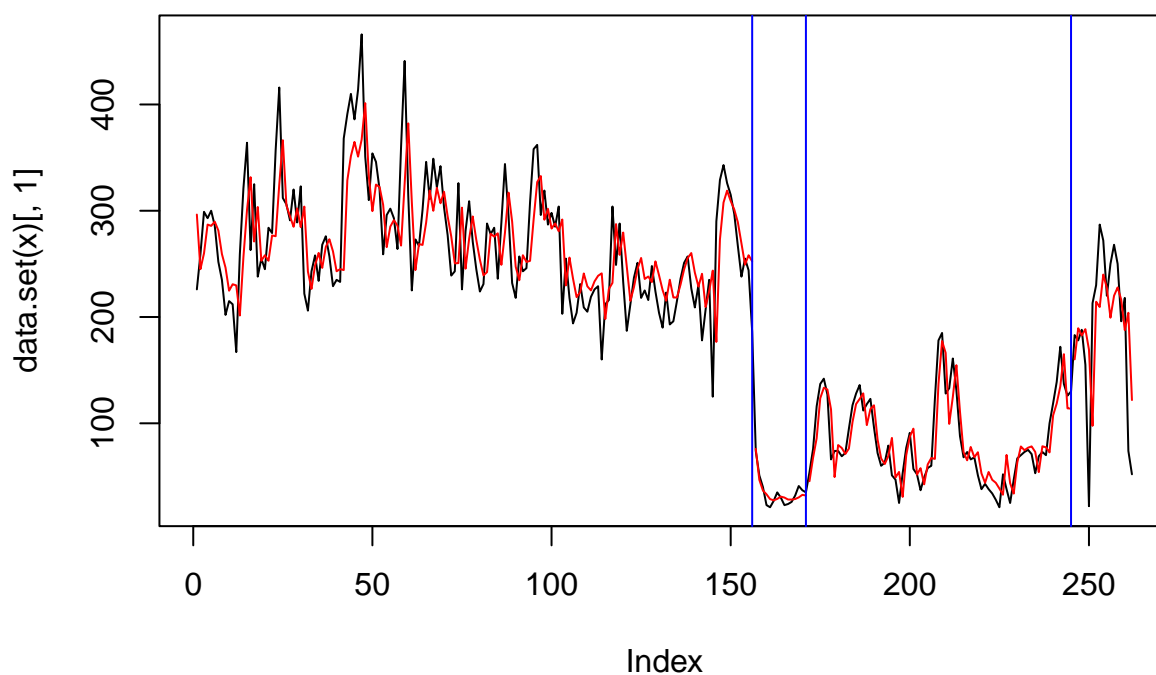
```
cpt.respi$meanar2cpt
```

```
## Class 'cpt.reg' : Changepoint Regression Object
##      ~~~ : S4 class containing 12 slots with names
##      cpttype date version data.set method test.stat pen.type pen.value minseglen cpts ncpts
##
## Created on   : Sun Jan 29 03:56:54 2023
##
## summary(.) :
## -----
## Created Using changepoint version 2.2.4
## Changepoint type      : Change in regression
## Method of analysis    : PELT
## Test Statistic       : Normal
## Type of penalty       : MBIC with value, 33.41007
## Maximum no. of cpts   : Inf
## Changepoint Locations : 156 171 245
```

```
cpts(cpt.respi$meanar2cpt)
```

```
## [1] 156 171 245
```

```
plot(cpt.respi[[which.min(BIC(cpt.respi))+1]]) + abline(v=cpts(cpt.respi$meanar2cpt), col='blue')
```



```
## integer(0)
```

```
### Primary model
```

```
cpt.cns <- envcpt(cns)
```

2.3.4.3 Diagnosis: CNS

```
## Fitting 12 models
```

```
##      |
```

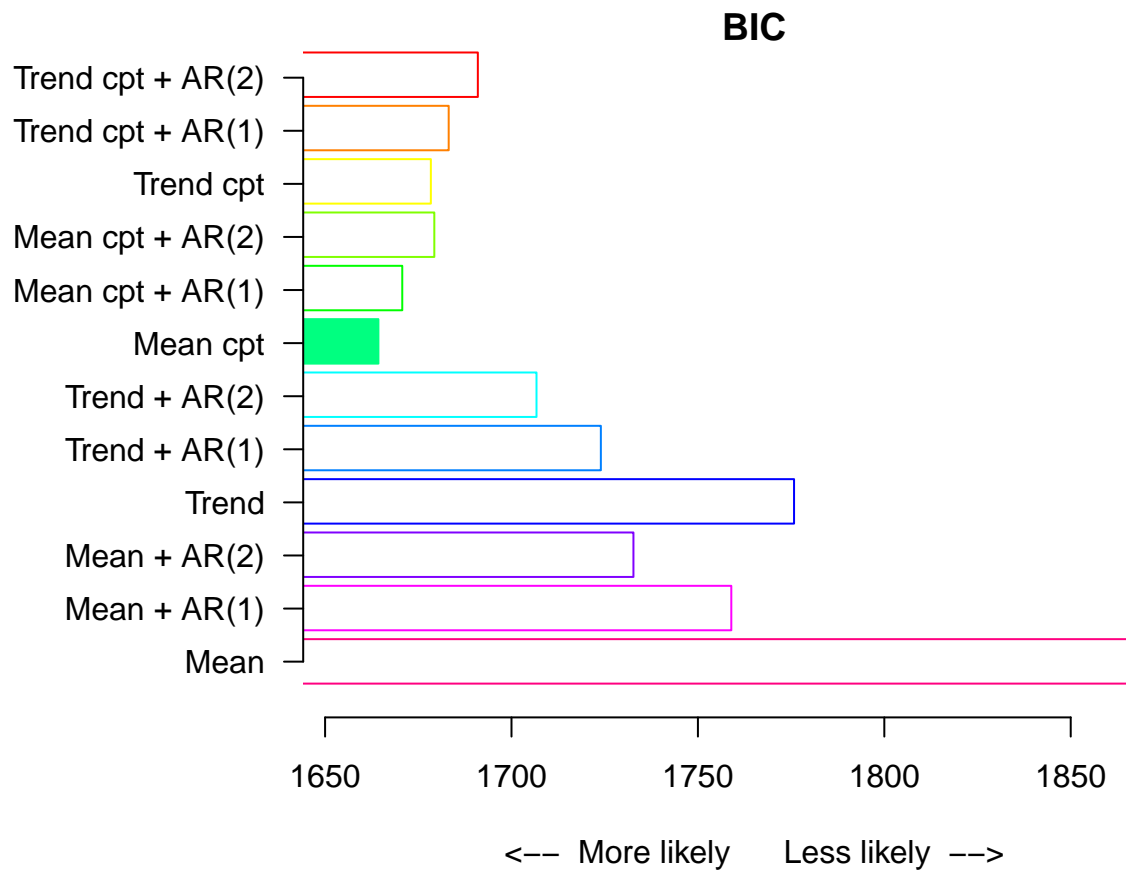
```
BIC(cpt.cns)
```

##	mean	meancpt	meanar1	meanar2	meanar1cpt	meanar2cpt
##	1866.011	1664.296	1758.957	1732.702	1670.694	1679.294
##	trend	trendcpt	trendar1	trendar2	trendar1cpt	trendar2cpt
##	1775.786	1678.365	1723.964	1706.705	1683.145	1690.965

```
which.min(BIC(cpt.cns))
```

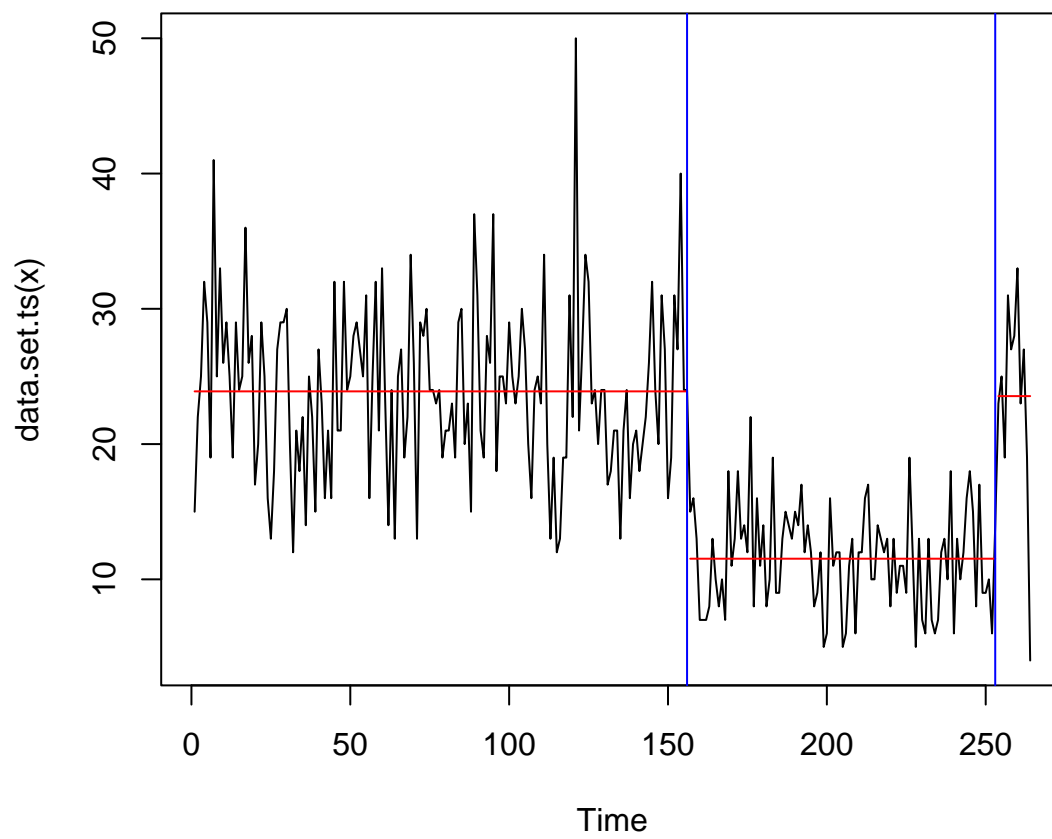
```
## meancpt  
##      2
```

```
plot(cpt.cns,type="bic")
```



```
### Selection of model based on best BIC
```

```
plot(cpt.cns[[which.min(BIC(cpt.cns))+1]]) + abline(v=cpts(cpt.cns$meancpt), col='blue')
```

```
## integer(0)
```

```
### Primary model
```

```
cpt.cd <- envcpt(cd)
```

2.3.4.4 Diagnosis: Fever and CD

```
## Fitting 12 models
```

```
## |
```

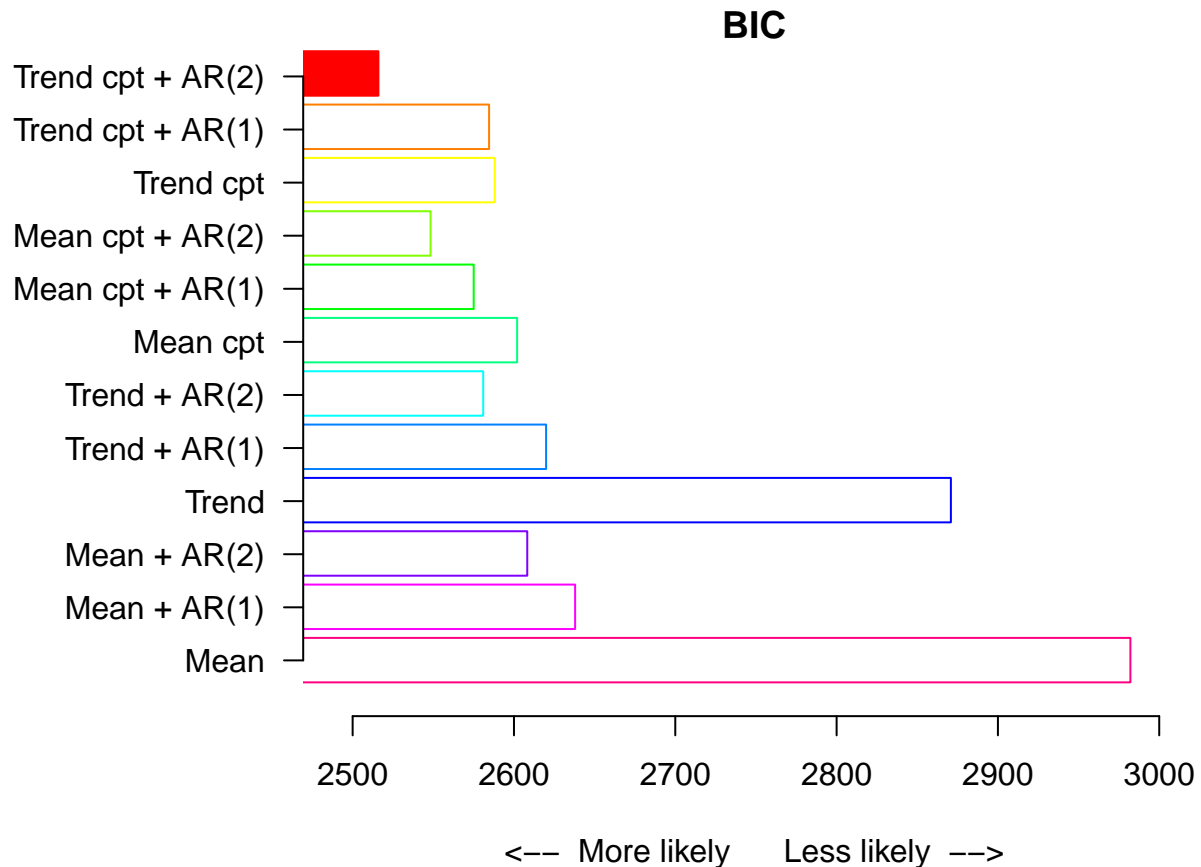
```
BIC(cpt.cd)
```

##	mean	meancpt	meanar1	meanar2	meanar1cpt	meanar2cpt
##	2982.126	2601.957	2637.910	2608.229	2575.019	2548.292
##	trend	trendcpt	trendar1	trendar2	trendar1cpt	trendar2cpt
##	2870.835	2588.102	2619.923	2580.919	2584.582	2515.961

```
which.min(BIC(cpt.cd))
```

```
## trendar2cpt
##          12
```

```
plot(cpt.cd,type="bic")
```



```
### Selection of model based on best BIC
```

```
cpt.cd$trendar2cpt
```

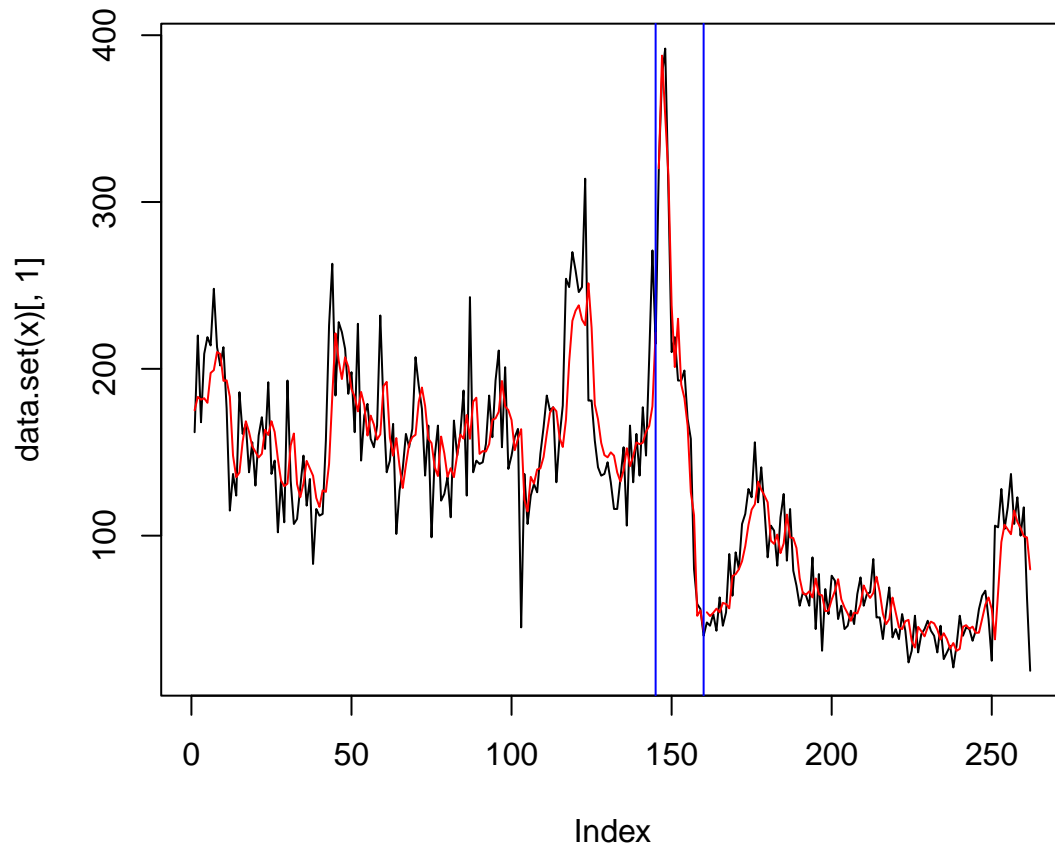
```
## Class 'cpt.reg' : Changepoint Regression Object
##      ~~      : S4 class containing 12 slots with names
##               cpttype date version data.set method test.stat pen.type pen.value minseglen cpts ncpts
##
## Created on   : Sun Jan 29 03:56:54 2023
##
## summary(.)   :
## -----
## Created Using changepoint version 2.2.4
## Changepoint type      : Change in regression
## Method of analysis     : PELT
## Test Statistic        : Normal
## Type of penalty        : MBIC with value, 38.97841
```

```
## Maximum no. of cpts : Inf
## Changepoint Locations : 145 160
```

```
cpts(cpt.cd$trendar2cpt)
```

```
## [1] 145 160
```

```
plot(cpt.cd[[which.min(BIC(cpt.cd))+1]]) + abline(v=cpts(cpt.cd$trendar2cpt), col='blue')
```



```
## integer(0)
```

```
### Primary model
```

```
cpt.trauma <- envcpt(trauma)
```

2.3.4.5 Diagnosis: Burn or trauma

```
## Fitting 12 models
```

```
## |
```

```
|
```

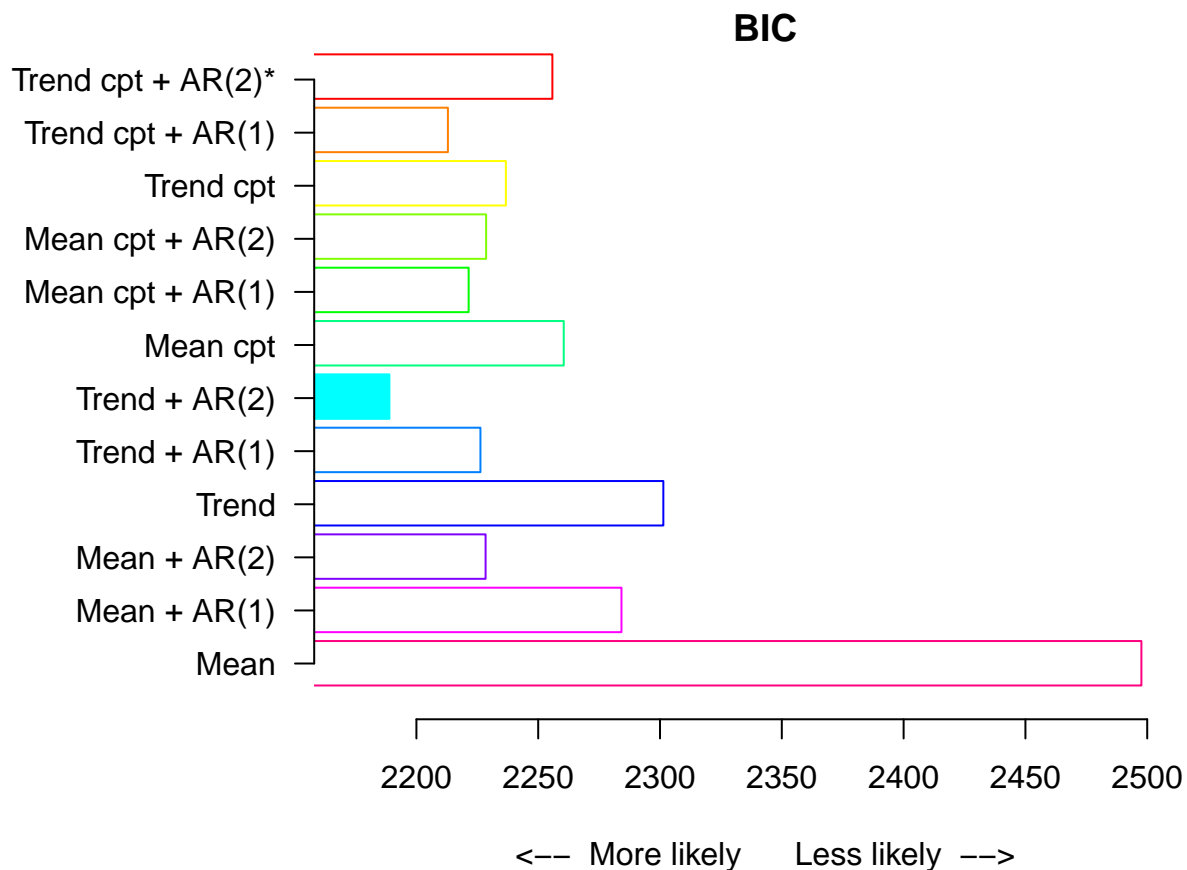
```
BIC(cpt.trauma)
```

```
##      mean      meancpt      meanar1      meanar2      meanar1cpt      meanar2cpt
##  2497.612    2260.486    2284.184    2228.405    2221.447    2228.618
##      trend      trendcpt      trendar1      trendar2      trendar1cpt      trendar2cpt
##  2301.375    2236.732    2226.285    2188.913    2212.933    2255.824
```

```
which.min(BIC(cpt.trauma))
```

```
## trendar2
##      10
```

```
plot(cpt.trauma,type="bic")
```



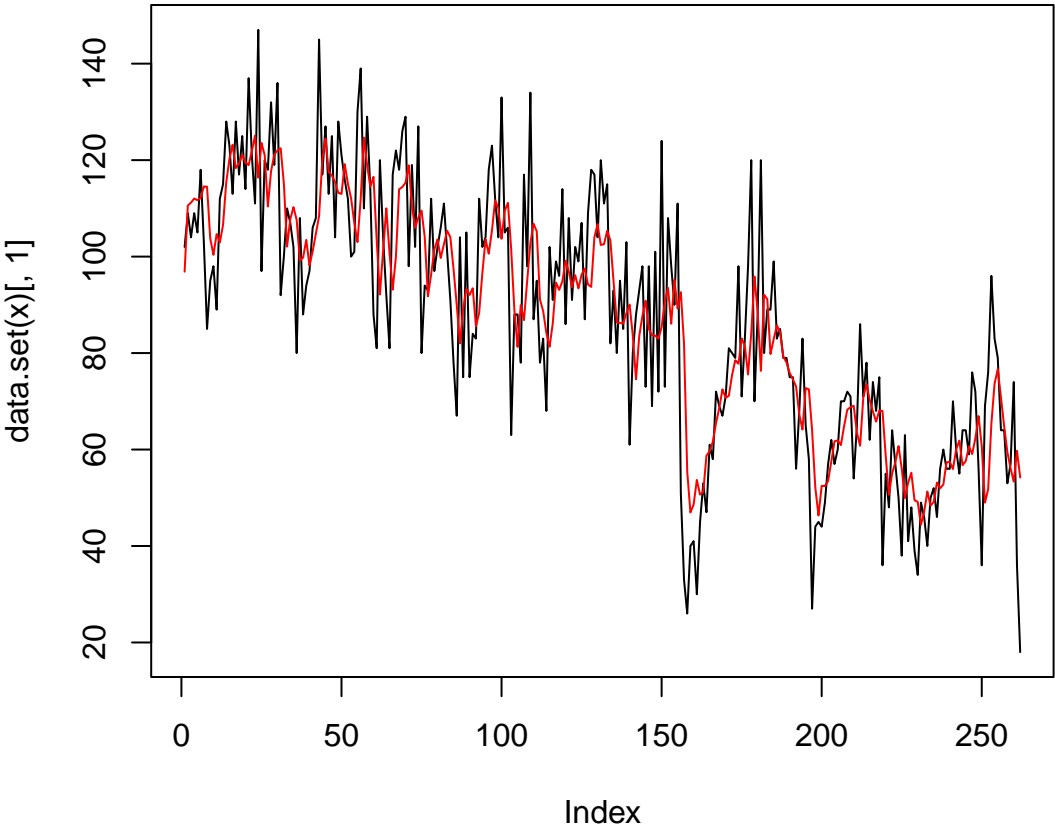
```
### Selection of model based on best BIC
```

```
cpt.trauma$trendar2
```

```
##
## Call:
## lm(formula = data[-c(1:2)] ~ c(1:(n - 2)) + data[2:(n - 1)] +
##      data[1:(n - 2)])
##
```

```
## Coefficients:
##      (Intercept)      c(1:(n - 2))  data[2:(n - 1)]  data[1:(n - 2)]
##      43.19028      -0.09785      0.29959      0.34939
```

```
plot(cpt.trauma$trendar2cpt)
```



```
### Primary model
cpt.gi <- envcpt(gi)
```

2.3.4.6 Diagnosis: GI

```
## Fitting 12 models
```

```
## |
```

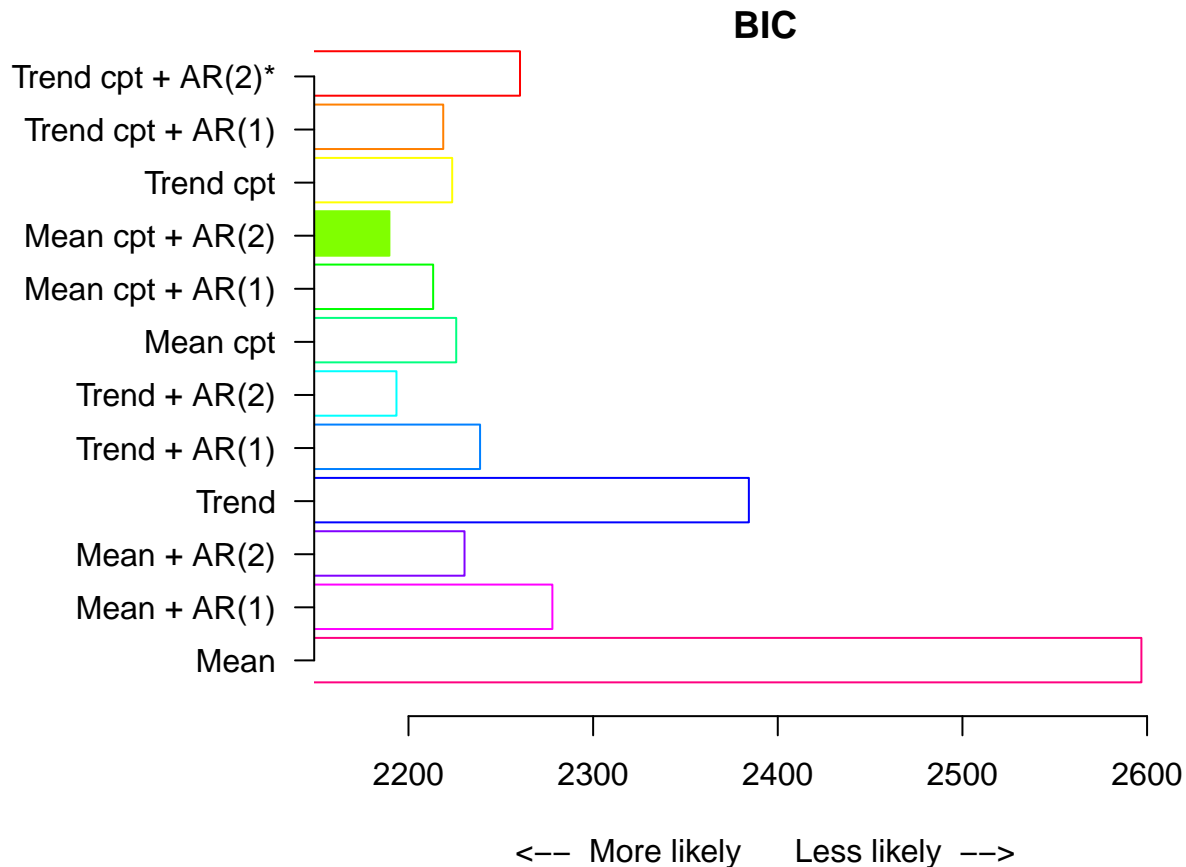
```
BIC(cpt.gi)
```

##	mean	meancpt	meanar1	meanar2	meanar1cpt	meanar2cpt
##	2596.921	2225.853	2277.936	2230.352	2213.366	2189.665
##	trend	trendcpt	trendar1	trendar2	trendar1cpt	trendar2cpt
##	2384.335	2223.670	2238.771	2193.487	2218.827	2260.398

```
which.min(BIC(cpt.gi))
```

```
## meanar2cpt
##          6
```

```
plot(cpt.gi,type="bic")
```



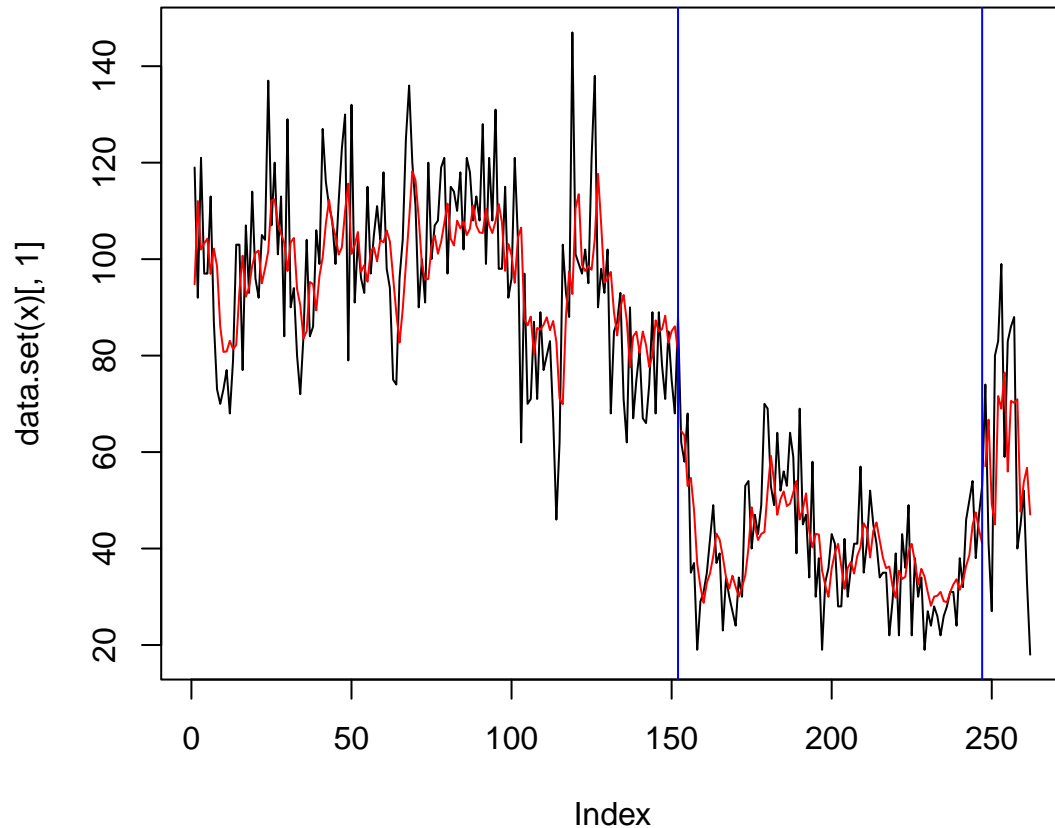
```
### Selection of model based on best BIC
```

```
cpt.gi$meanar2cpt
```

```
## Class 'cpt.reg' : Changepoint Regression Object
##      ~~      : S4 class containing 12 slots with names
##               cpttype date version data.set method test.stat pen.type pen.value minseglen cpts ncpts
##
## Created on   : Sun Jan 29 03:56:54 2023
##
## summary(.)   :
## -----
## Created Using changepoint version 2.2.4
## Changepoint type      : Change in regression
## Method of analysis    : PELT
## Test Statistic       : Normal
## Type of penalty       : MBIC with value, 33.41007
```

```
## Maximum no. of cpts : Inf
## Changepoint Locations : 152 247
```

```
plot(cpt.gi[[which.min(BIC(cpt.gi))+1]]) + abline(v=cpts(cpt.gi$meanar2cpt), col='blue')
```



```
## integer(0)
```

```
### Primary model
```

```
cpt.neo <- envcpt(neo)
```

2.3.4.7 Diagnosis: Neonatal

```
## Fitting 12 models
```

```
## |
```

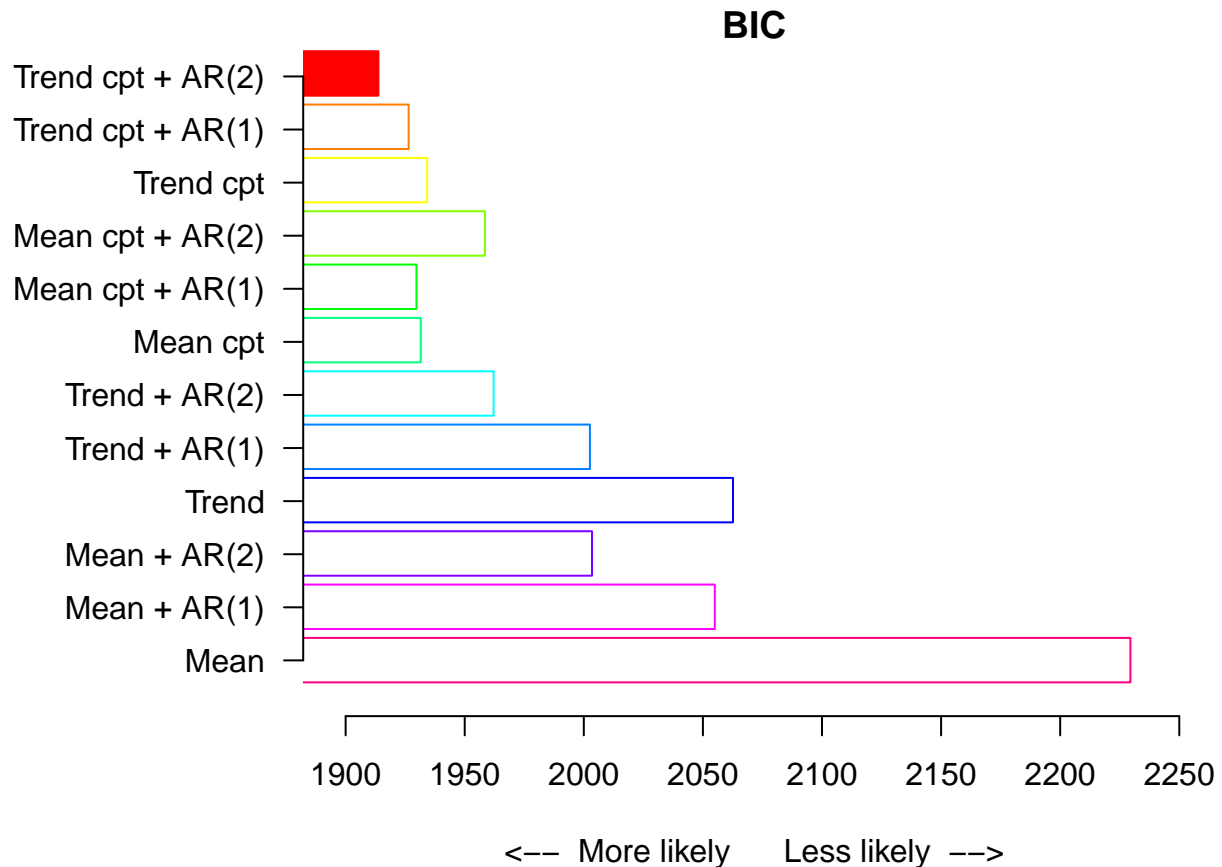
```
BIC(cpt.neo)
```

##	mean	meancpt	meanar1	meanar2	meanar1cpt	meanar2cpt
##	2229.480	1931.546	2054.998	2003.440	1929.740	1958.466
##	trend	trendcpt	trendar1	trendar2	trendar1cpt	trendar2cpt
##	2062.640	1934.209	2002.572	1962.132	1926.456	1913.762

```
which.min(BIC(cpt.neo))
```

```
## trendar2cpt
##          12
```

```
plot(cpt.neo,type="bic")
```



```
### Selection of model based on best BIC
```

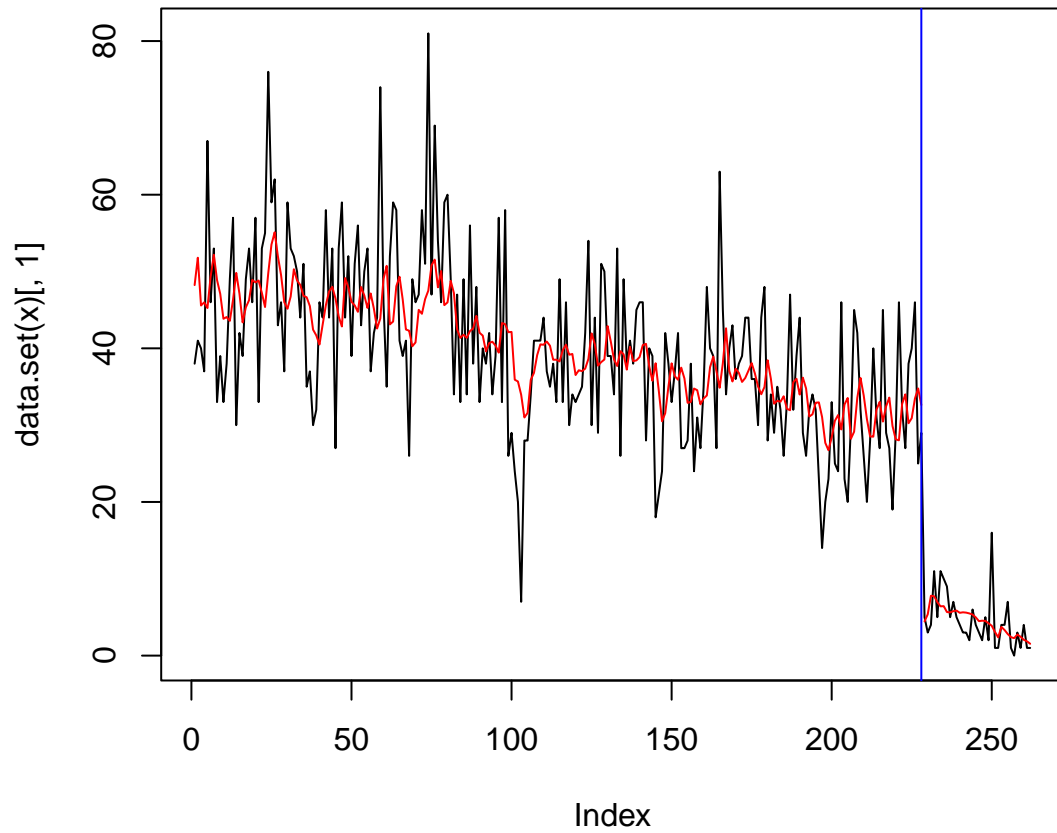
```
cpt.neo$trendar2cpt
```

```
## Class 'cpt.reg' : Changepoint Regression Object
##      ~~      : S4 class containing 12 slots with names
##              cpttype date version data.set method test.stat pen.type pen.value minseglen cpts ncpts
##
## Created on   : Sun Jan 29 03:56:54 2023
##
## summary(.)   :
## -----
## Created Using changepoint version 2.2.4
## Changepoint type      : Change in regression
## Method of analysis    : PELT
## Test Statistic       : Normal
## Type of penalty       : MBIC with value, 38.97841
```



```
## Maximum no. of cpts : Inf  
## Changepoint Locations : 228
```

```
plot(cpt.neo[[which.min(BIC(cpt.neo))+1]]) + abline(v=cpts(cpt.neo$trendar2cpt), col='blue')
```



```
## integer(0)
```