Time Series Analysis in R

Nilotpal Sanyal

(nilotpal.sanyal@gmail.com)

Bayesian and Interdisciplinary Research Unit

(currently Interdisciplinary Statistical Research Unit)

Indian Statistical Institute

Exploratory time series commands

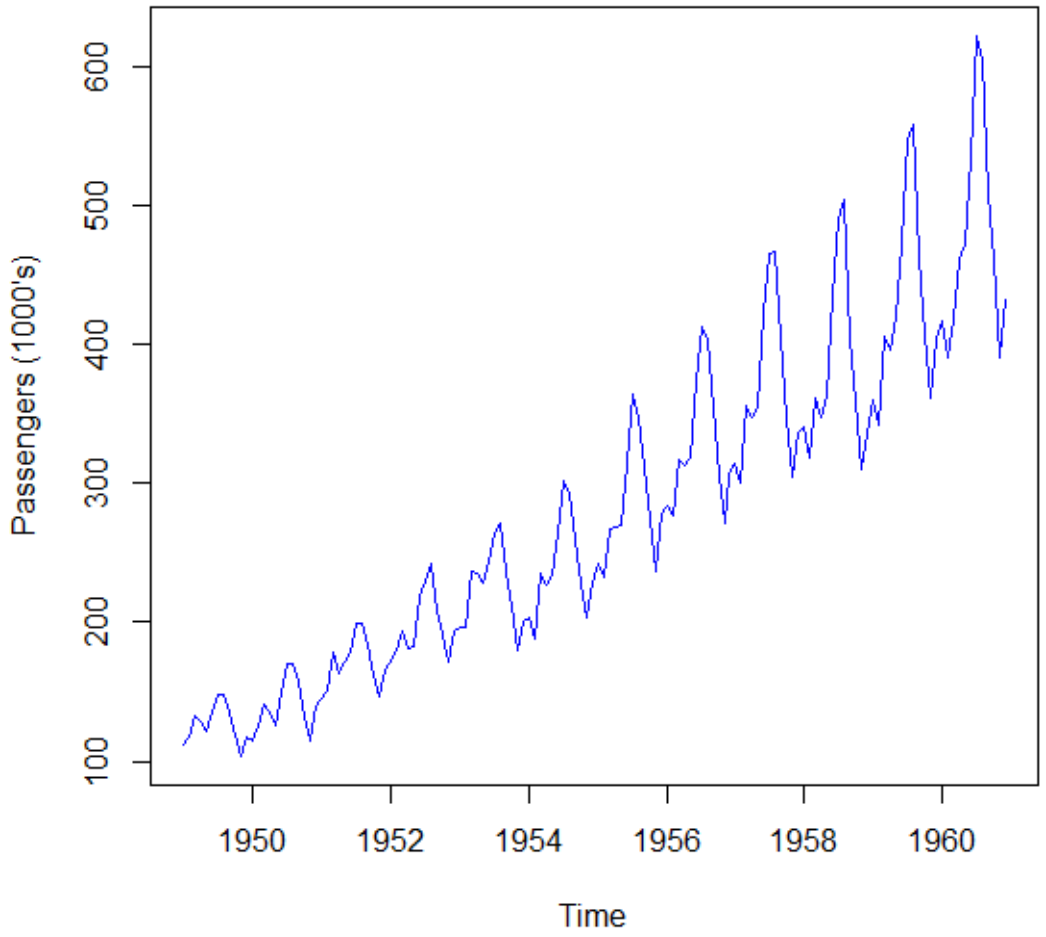
```
data(AirPassengers)          # loads the dataset 'AirPassengers'
AP <- AirPassengers
AP
     Jan  Feb Mar Apr May Jun Jul Aug Sep Oct Nov Dec
1949 112 118 132 129 121 135 148 148 136 119 104 118
1950 115 126 141 135 125 149 170 170 158 133 114 140
1951 145 150 178 163 172 178 199 199 184 162 146 166
1952 171 180 193 181 183 218 230 242 209 191 172 194
1953 196 196 236 235 229 243 264 272 237 211 180 201
1954 204 188 235 227 234 264 302 293 259 229 203 229
1955 242 233 267 269 270 315 364 347 312 274 237 278
1956 284 277 317 313 318 374 413 405 355 306 271 306
1957 315 301 356 348 355 422 465 467 404 347 305 336
1958 340 318 362 348 363 435 491 505 404 359 310 337
1959 360 342 406 396 420 472 548 559 463 407 362 405
1960 417 391 419 461 472 535 622 606 508 461 390 432
```

```
class(AP)          # shows the class of the object 'AP'
[1] "ts"           # time series object

start(AP)          # shows time series starting time-point
[1] 1949 1

end(AP)            # shows time series ending time-point
[1] 1960 12

frequency(AP)      # shows the number of observations per unit of time
[1] 12

deltat(AP)         # shows the time interval between observations
[1] 0.08333333

time(AP)           # gives the vector of times at which a time series was
sampled
```

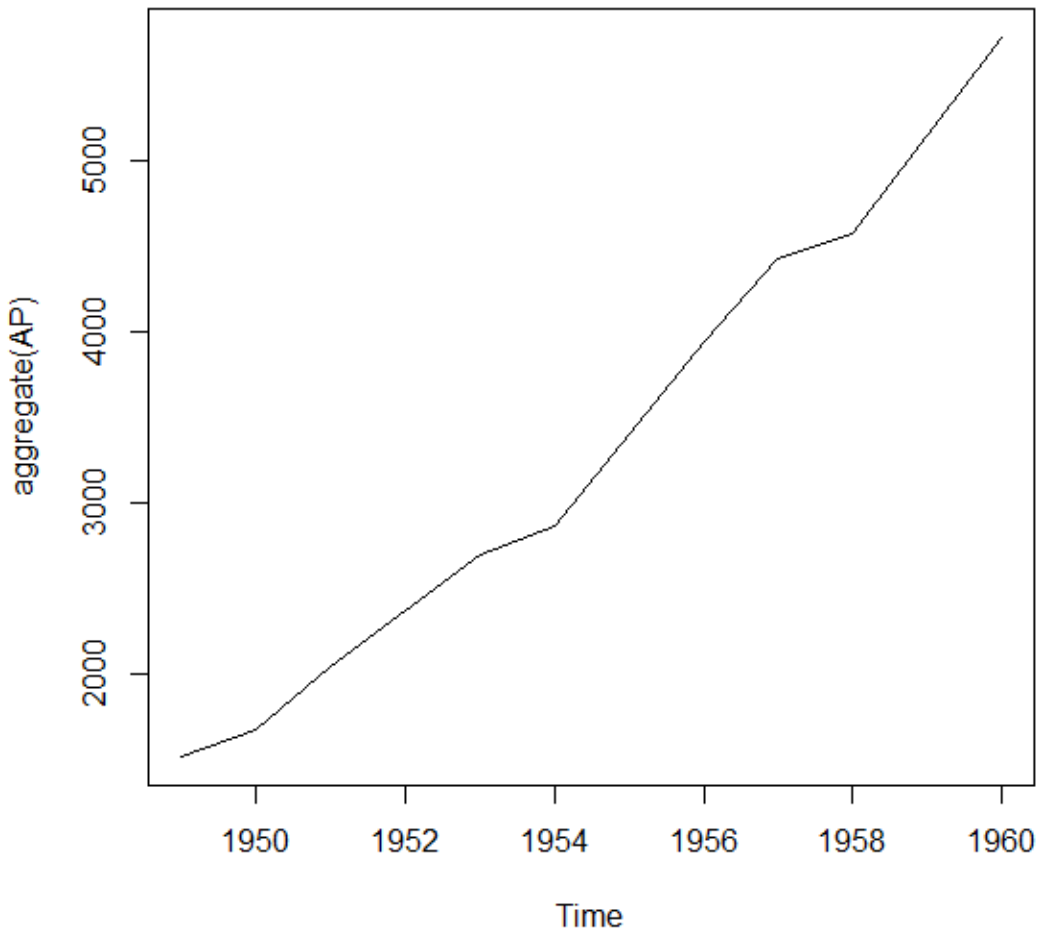
cycle(AP) # shows the positions in the cycle of each observation

	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
1949	1	2	3	4	5	6	7	8	9	10	11	12
1950	1	2	3	4	5	6	7	8	9	10	11	12
1951	1	2	3	4	5	6	7	8	9	10	11	12
1952	1	2	3	4	5	6	7	8	9	10	11	12
1953	1	2	3	4	5	6	7	8	9	10	11	12
1954	1	2	3	4	5	6	7	8	9	10	11	12
1955	1	2	3	4	5	6	7	8	9	10	11	12
1956	1	2	3	4	5	6	7	8	9	10	11	12
1957	1	2	3	4	5	6	7	8	9	10	11	12
1958	1	2	3	4	5	6	7	8	9	10	11	12
1959	1	2	3	4	5	6	7	8	9	10	11	12
1960	1	2	3	4	5	6	7	8	9	10	11	12

```
plot(AP, ylab = "Passengers (1000's)", col="blue")      # produces time series plot
```



```
plot(aggregate(AP))          # produces plots of aggregates
```



Extract specific parts from the time series

```
window(AP, start=c(1949,1), end=c(1950,10))
      Jan Feb Mar Apr May Jun  Jul  Aug Sep Oct Nov Dec
1949  112  118  132  129  121  135  148  148  136  119  104  118
1950  115  126  141  135  125  149  170  170  158  133

window(AP, start=c(1949,1), end=c(1951,2), frequency=6)
Time Series:
Start = c(1949, 1)
End = c(1951, 1)
Frequency = 6
 [1]  112  132  121  148  136  104  115  141  125  170  158  114  145

window(AP, start=c(1949,1), end=c(1951,2), frequency=4)
Qtr1 Qtr2 Qtr3 Qtr4
1949  112  129  148  119
1950  115  135  170  133
```

1951 145

How to compute trend and seasonal component by moving average in R?

```
decomAP <- decompose(AP)

names(decomAP)
[1] "x"        "seasonal" "trend"    "random"   "figure"   "type"

round(decomAP$trend, 1)        # rounded value of trend by m.a.
     Jan   Feb   Mar   Apr   May   Jun   Jul   Aug   Sep   Oct   Nov   Dec
1949    NA    NA    NA    NA    NA    NA 126.8 127.2 128.0 128.6 129.0 129.8
1950 131.2 133.1 134.9 136.4 137.4 138.8 140.9 143.2 145.7 148.4 151.5 154.7
1951 157.1 159.5 161.8 164.1 166.7 169.1 171.2 173.6 175.5 176.8 178.0 180.2
1952 183.1 186.2 189.0 191.3 193.6 195.8 198.0 199.8 202.2 206.2 210.4 213.4
1953 215.8 218.5 220.9 222.9 224.1 224.7 225.3 225.3 225.0 224.6 224.5 225.5
1954 228.0 230.5 232.2 233.9 235.6 237.8 240.5 244.0 247.2 250.2 253.5 257.1
1955 261.8 266.7 271.1 275.2 278.5 282.0 285.8 289.3 293.2 297.2 301.0 305.5
1956 310.0 314.4 318.6 321.8 324.5 327.1 329.5 331.8 334.5 337.5 340.5 344.1
1957 348.2 353.0 357.6 361.4 364.5 367.2 369.5 371.2 372.2 372.4 372.8 373.6
1958 375.3 377.9 379.5 380.0 380.7 381.0 381.8 383.7 386.5 390.3 394.7 398.6
```

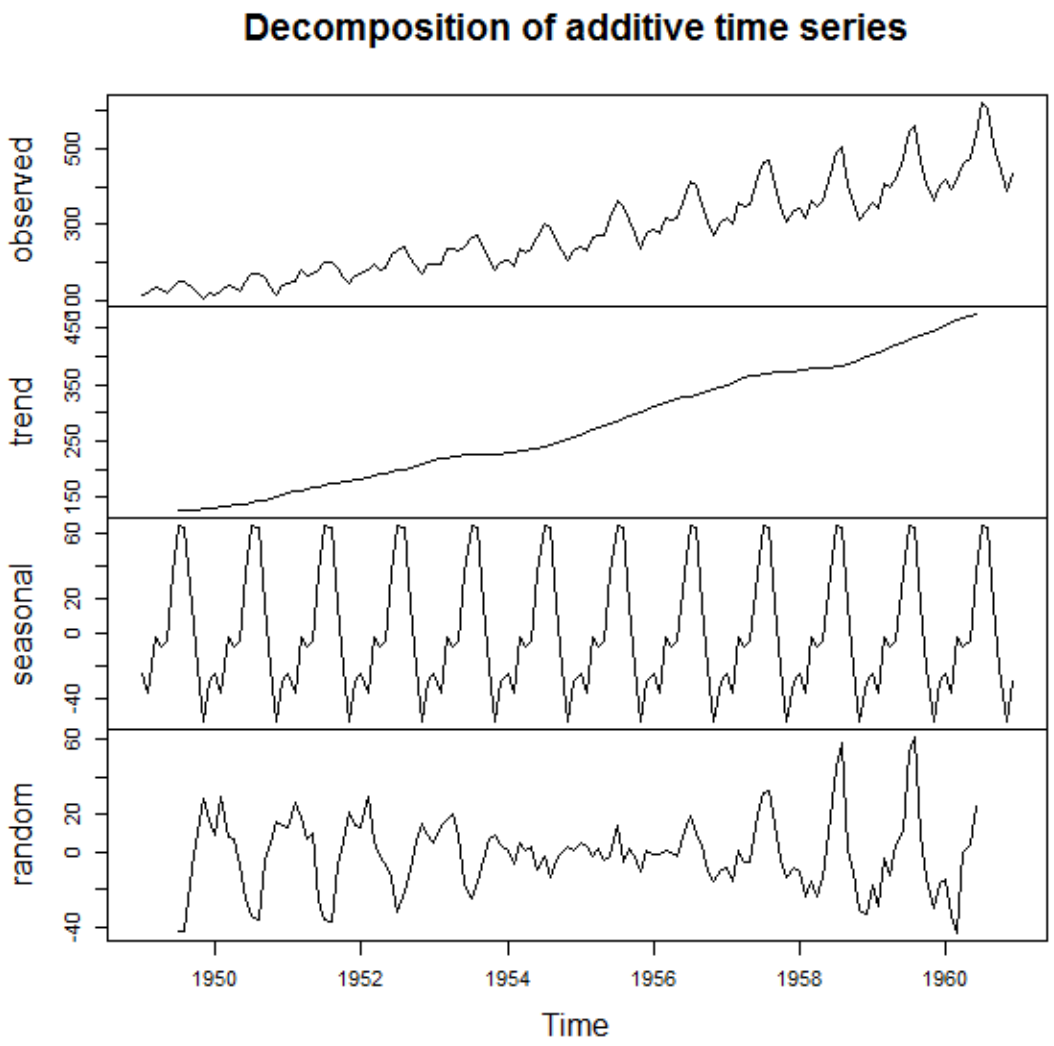
1959 402.5 407.2 411.9 416.3 420.5 425.5 430.7 435.1 437.7 441.0 445.8 450.6
1960 456.3 461.4 465.2 469.3 472.8 475.0 NA NA NA NA NA NA

round(decomAP$seasonal, 1) # rounded value of seasonal comp. by m.a.

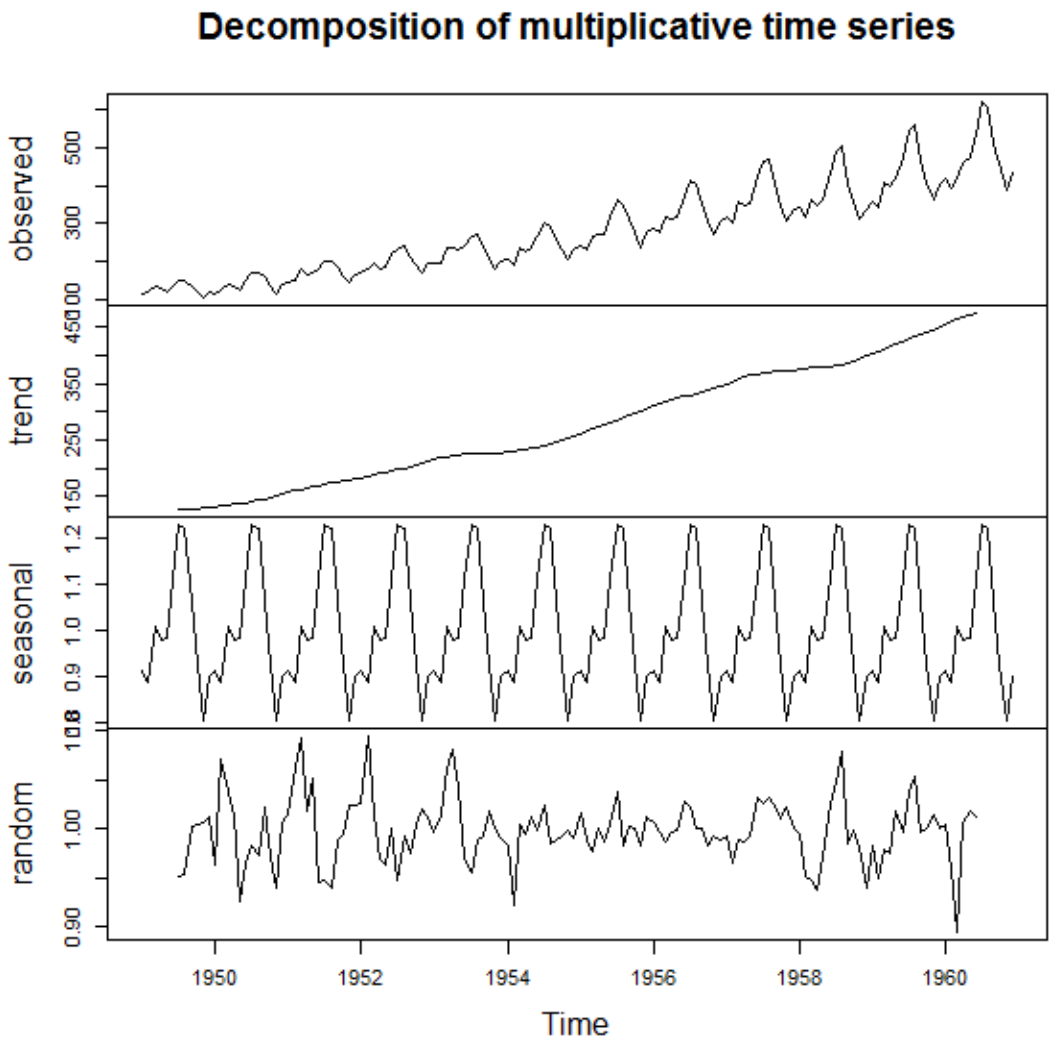
	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
1949	-24.7	-36.2	-2.2	-8.0	-4.5	35.4	63.8	62.8	16.5	-20.6	-53.6	-28.6
1950	-24.7	-36.2	-2.2	-8.0	-4.5	35.4	63.8	62.8	16.5	-20.6	-53.6	-28.6
1951	-24.7	-36.2	-2.2	-8.0	-4.5	35.4	63.8	62.8	16.5	-20.6	-53.6	-28.6
1952	-24.7	-36.2	-2.2	-8.0	-4.5	35.4	63.8	62.8	16.5	-20.6	-53.6	-28.6
1953	-24.7	-36.2	-2.2	-8.0	-4.5	35.4	63.8	62.8	16.5	-20.6	-53.6	-28.6
1954	-24.7	-36.2	-2.2	-8.0	-4.5	35.4	63.8	62.8	16.5	-20.6	-53.6	-28.6
1955	-24.7	-36.2	-2.2	-8.0	-4.5	35.4	63.8	62.8	16.5	-20.6	-53.6	-28.6
1956	-24.7	-36.2	-2.2	-8.0	-4.5	35.4	63.8	62.8	16.5	-20.6	-53.6	-28.6
1957	-24.7	-36.2	-2.2	-8.0	-4.5	35.4	63.8	62.8	16.5	-20.6	-53.6	-28.6
1958	-24.7	-36.2	-2.2	-8.0	-4.5	35.4	63.8	62.8	16.5	-20.6	-53.6	-28.6
1959	-24.7	-36.2	-2.2	-8.0	-4.5	35.4	63.8	62.8	16.5	-20.6	-53.6	-28.6
1960	-24.7	-36.2	-2.2	-8.0	-4.5	35.4	63.8	62.8	16.5	-20.6	-53.6	-28.6

```
round(decomAP$random, 1)  # rounded value of irregular component by m.a.
      Jan   Feb   Mar   Apr   May   Jun   Jul   Aug   Sep   Oct   Nov   Dec
1949   NA    NA    NA    NA    NA    NA -42.6 -42.1  -8.5  11.1  28.6  16.9
1950  8.5  29.1   8.3   6.6  -7.9 -25.2 -34.7 -36.0  -4.2   5.2  16.1  13.9
1951 12.6  26.6  18.4   6.9   9.8 -26.5 -36.1 -37.4  -8.0   5.8  21.6  14.5
1952 12.6  30.0   6.2  -2.3  -6.1 -13.2 -31.9 -20.6  -9.7   5.4  15.2   9.2
1953  4.9  13.7  17.3  20.1   9.4 -17.1 -25.2 -16.2  -4.5   7.1   9.1   4.1
1954  0.7  -6.3   5.0   1.1   2.9  -9.2  -2.3 -13.8  -4.7  -0.6   3.1   0.5
1955  4.9   2.5  -1.9   1.8  -4.0  -2.4  14.4  -5.2   2.2  -2.5 -10.4   1.2
1956 -1.2  -1.2   0.6  -0.7  -2.0  11.5  19.6  10.3   4.0 -10.9 -15.9  -9.5
1957 -8.5 -15.8   0.6  -5.3  -5.0  19.4  31.7  33.0  15.3  -4.8 -14.2  -9.0
1958 -10.5 -23.7 -15.3 -24.0 -13.2  18.6  45.3  58.5   1.0 -10.7 -31.1 -33.0
1959 -17.8 -29.0  -3.6 -12.3   4.0  11.1  53.5  61.1   8.8 -13.3 -30.2 -17.0
1960 -14.6 -34.2 -44.0  -0.3   3.8  24.6    NA    NA    NA    NA    NA    NA

decomAP$figure              # Seasonal figures only
[1]  -24.748737  -36.188131   -2.241162   -8.036616   -4.506313   35.402778
63.830808  62.823232  16.520202
[10] -20.642677 -53.593434 -28.619949
```

plot(decomAP) # Plots of different components



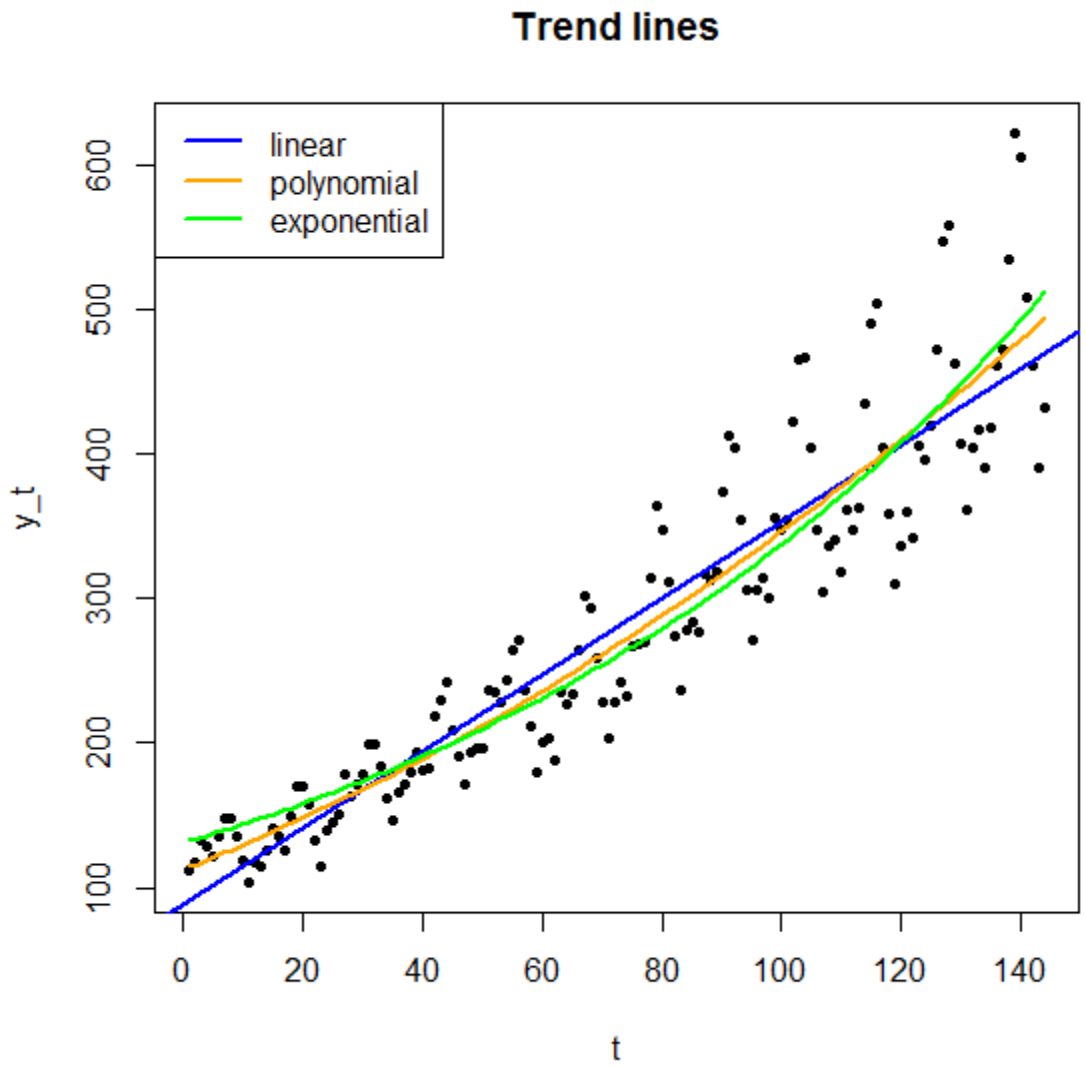
```
plot(decompose(AP,type="multiplicative"))
```



How to fit trend equation in R?

```
# time plot
y <- AP
x <- 1:length(AP)
plot(x, y, pch=20, ylab="y_t", xlab="t", main="Trend lines")

# linear trend line fit
fit <- glm(y~x)
co <- coef(fit)
abline(fit, col="blue", lwd=2)

# polynomial trend line fit
f <- function(x,a,b,d) {(a*x^2) + (b*x) + d}
fit <- nls(y ~ f(x,a,b,d), start = c(a=1, b=1, d=1))
co <- coef(fit)
curve(f(x, a=co[1], b=co[2], d=co[3]), add = TRUE, col="orange", lwd=2)
```

```
# exponential trend line fit
f <- function(x,a,b) {a * b^x}
fit <- nls(y ~ f(x,a,b), start = c(a=1, b=1))
co <- coef(fit)
curve(f(x, a=co[1], b=co[2]), add = TRUE, col="green", lwd=2)

#Add a descriptive legend to plot
legend("topleft", legend=c("linear","polynomial","exponential"),
col=c("blue","orange","green"), lwd=2)
```


Trend lines
linear
polynomial
exponential
y_t
100
200
300
400
500
600
0
20
40
60
80
100
120
140
t