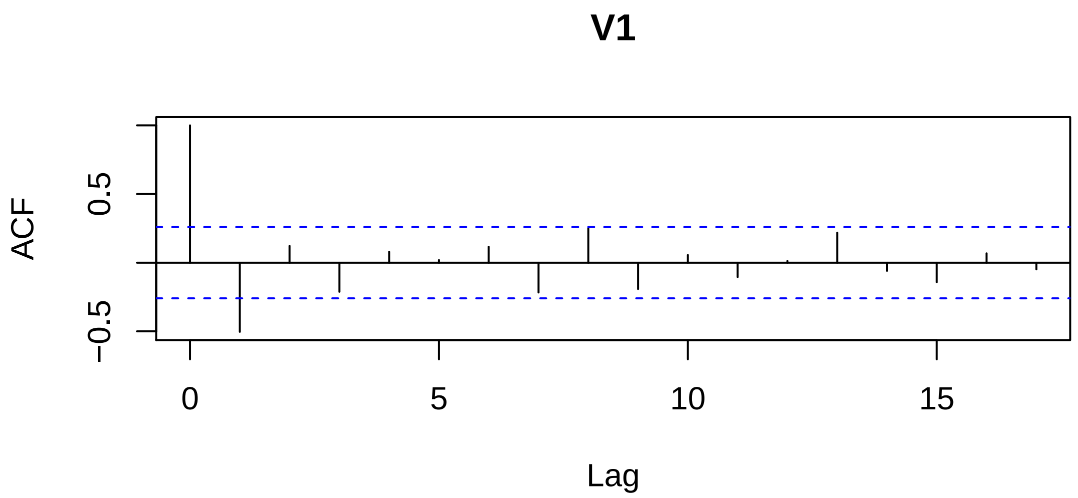
```r
# 2.1

# a)

a <- read.table("http://gamma.mini.pw.edu.pl/~szymanowskih/lab2/OSHORTS.txt")
head(a,2)
```

```
##    V1
## 1  78
## 2 -58
```

```r
acf(a)   # ma(1)
```




```r
v <- as.numeric(as.matrix(a))-mean(as.matrix(a))
Box.test(v,lag=20,type="Ljung")   # czyli to nie jest bialy szum
```

```
##
##  Box-Ljung test
##
## data:  v
## X-squared = 38.18, df = 20, p-value = 0.008424
```

```r
# b)

ma1 <- arima(a,c(0,0,1))
ma1$coef
```

```
##        ma1 intercept
##    -0.8473   -4.7798
```

```r
Box.test(ma1$res,lag=20,type="Ljung")
```

```
##
##  Box-Ljung test
##
## data:  ma1$res
## X-squared = 21.81, df = 20, p-value = 0.3509
```

```r
# c) -> uzyjemy statystyki walda do tego testu

ma1$var.coef
```

```
##                 ma1  intercept
## ma1        0.01453    0.02001
## intercept  0.02001    1.05404
```

```
stat <- as.numeric(ma1$coef[2]/sqrt(ma1$var.coef[2,2]))
p.val <- pnorm(stat)
p.val
```

```
## [1] 1.614e-06
```

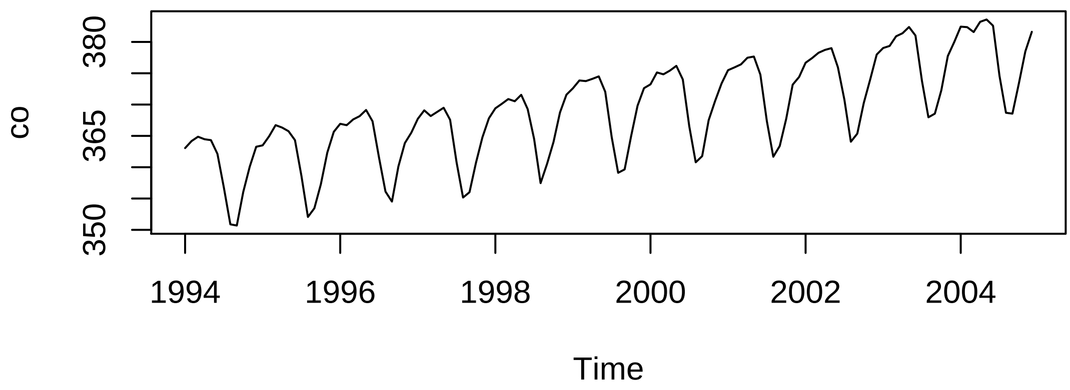
```
# odrzucamy hipoteze zerowa -> srednia jest ujemna

# d)
# przedzial ufnosci dla thety: theta to wspolczynnik w modelu ma(1)
# xt = epst + theta*eps(t-1) + mi

qu <- qnorm(0.975)
ma1$coef[1]+sqrt(ma1$var.coef[1,1])*c(-qu,qu)
```

```
## [1] -1.0835 -0.6111
```

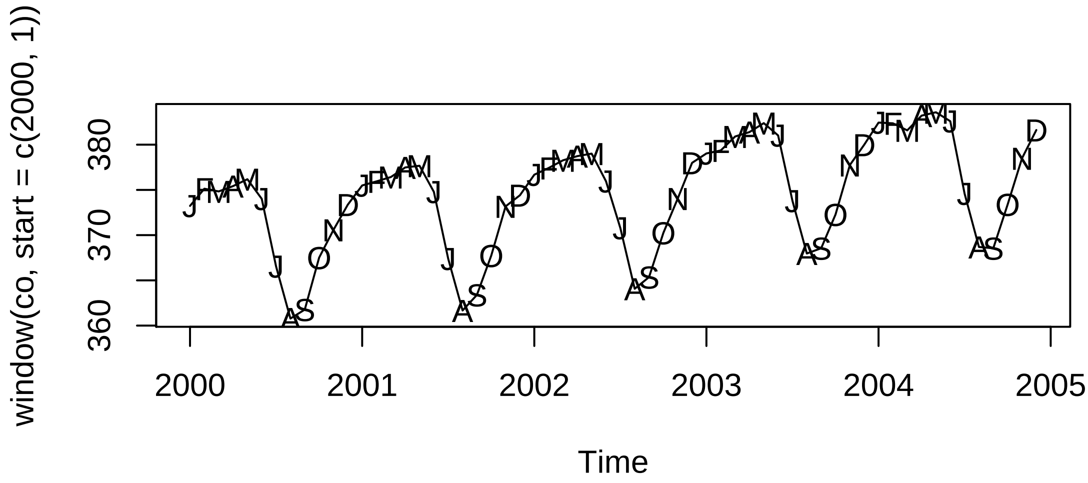
```
# 2.2

library("MASS")

# a)

co <- dget("http://gamma.mini.pw.edu.pl/~szymanowskih/lab2/co2.dput")
head(co)
```

```
## [1] 363.1 364.2 364.9 364.5 364.3 362.1
```

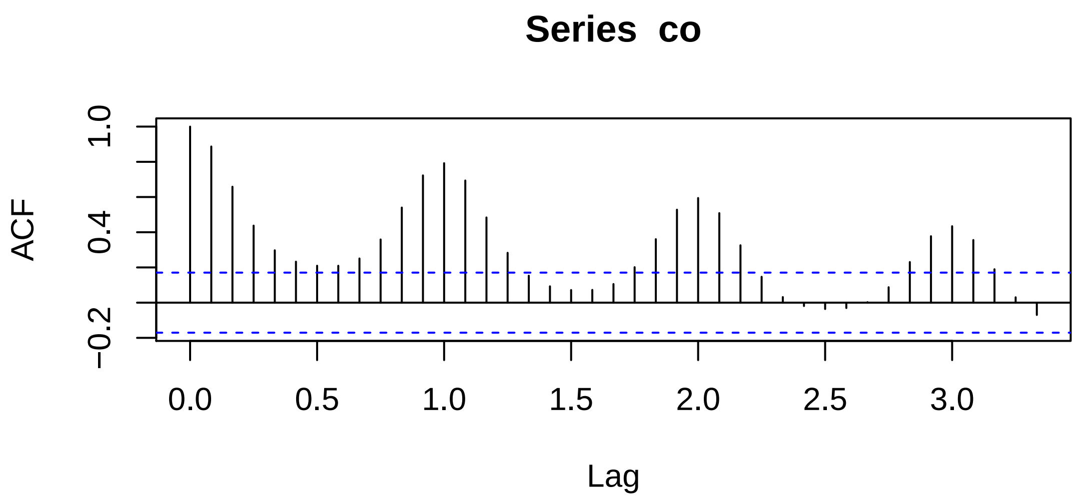
```
# b)

plot(co)  # co widac: sezonowosc i trend
```



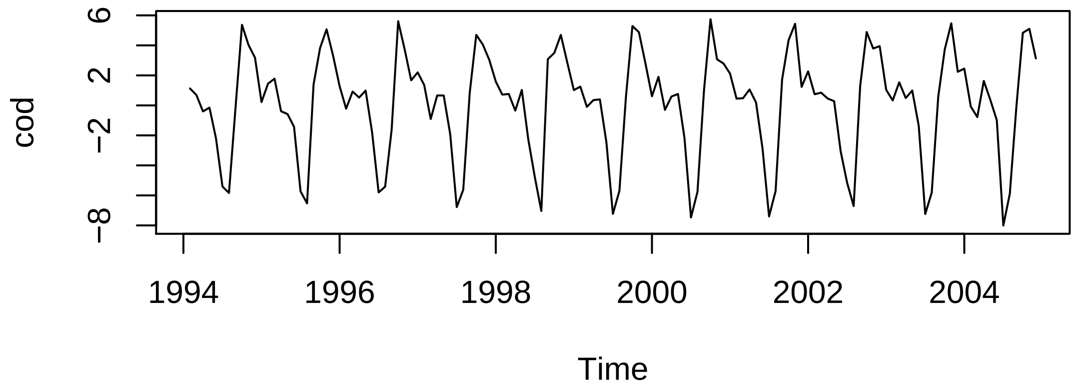

```
# c)

plot(window(co,start=c(2000,1)))
Month=c("J", "F","M","A","M","J","J","A","S","O","N","D" )
points(window(co, start=c(2000,1)), pch=Month)
```



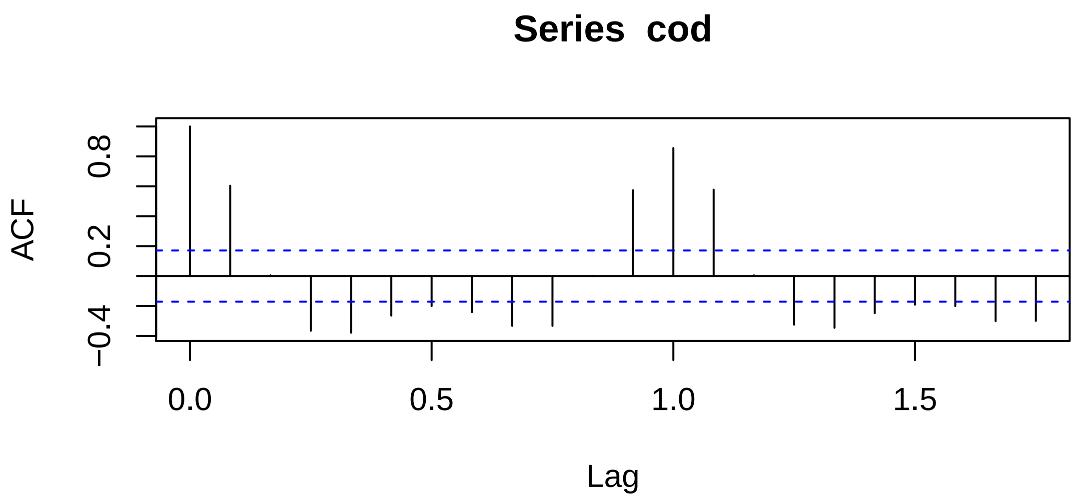

```
# d)

acf(co,lag.max=40)    # widac niestacjonarnosc i sezonowosc
```



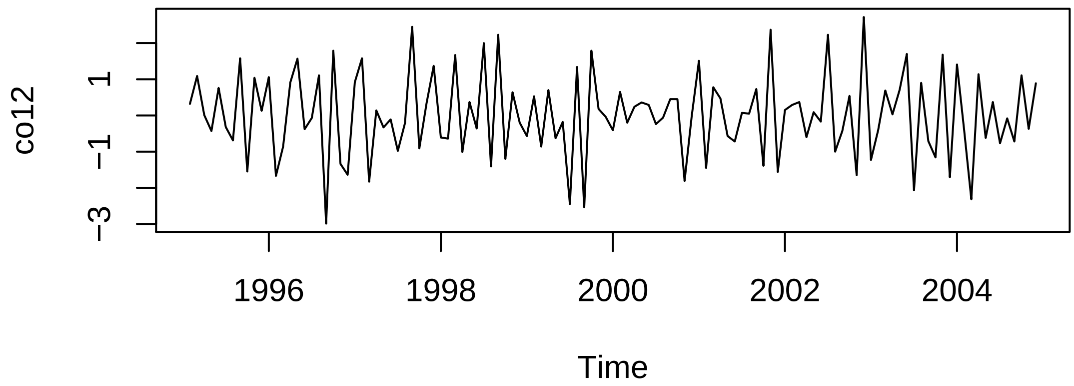

```
# zeby usunac trend: zroznicujmy

# e)

cod <- diff(co)
plot(cod)    # trend usuniety, ale mamy okresowosc
```



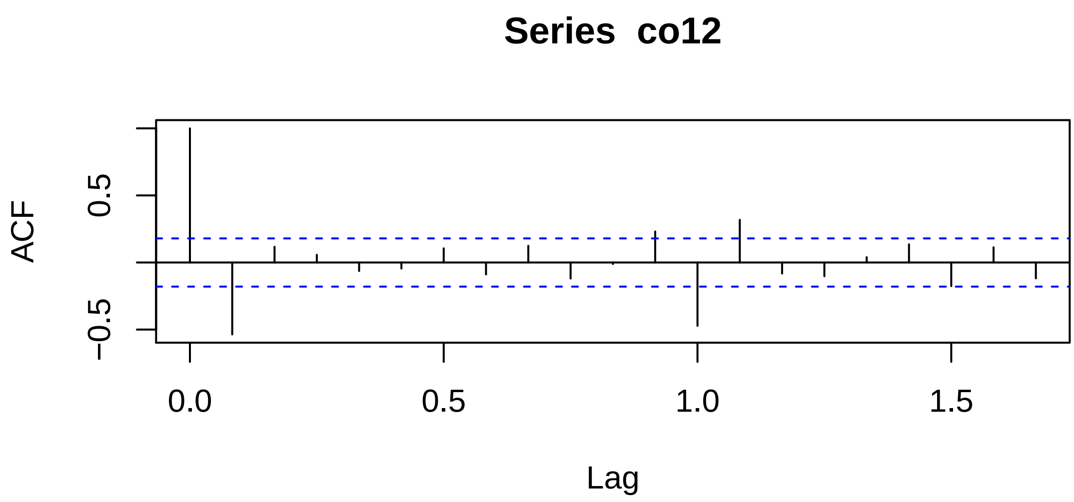

```
# f)

acf(cod)
```



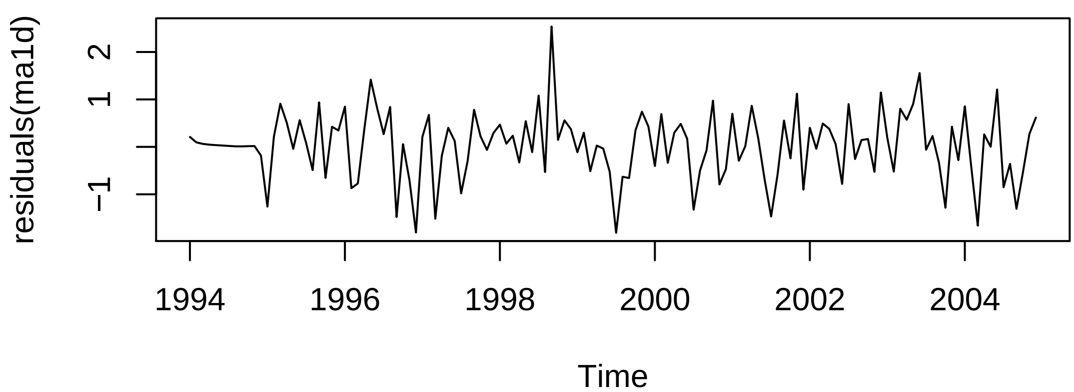

```
# g)

co12 <- diff(cod,lag=12)
plot(co12)
```



```
# lepiej pod tym wzgledem, ze nie widac regularnosci,
# czyli szereg wyglada na stacjonarny

# h)

acf(co12)
```

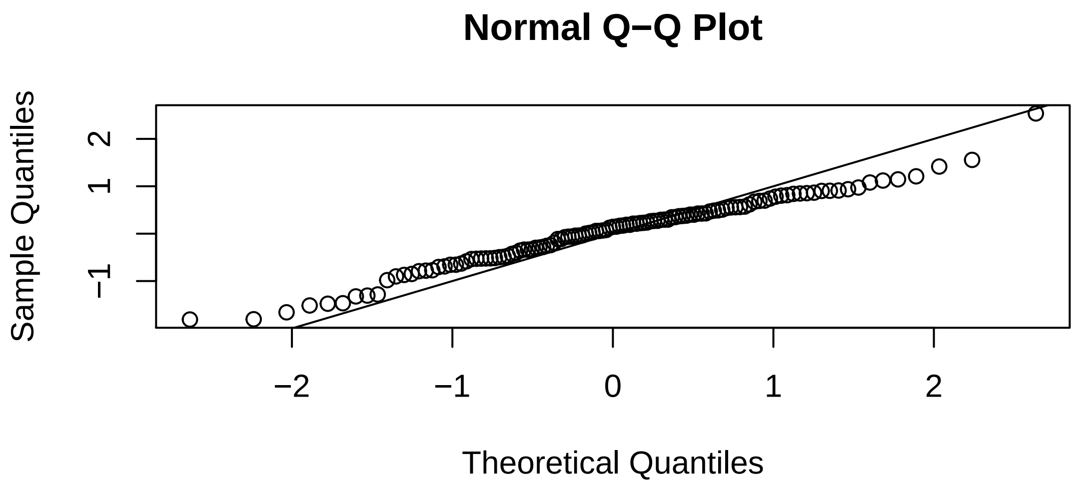



```
# ten rysunek nam mowi, ze pierwsza i dwunasta sa istotne i jeszcze dwie,
# ale je olewamy :D

# i)

# dopasujemy model z pierwszym i dwunastym - model sezonowy

ma1d <- arima(co,order=c(0,1,1),seasonal=list(order=c(0,1,1),12))

# j)

plot(residuals(ma1d))
```

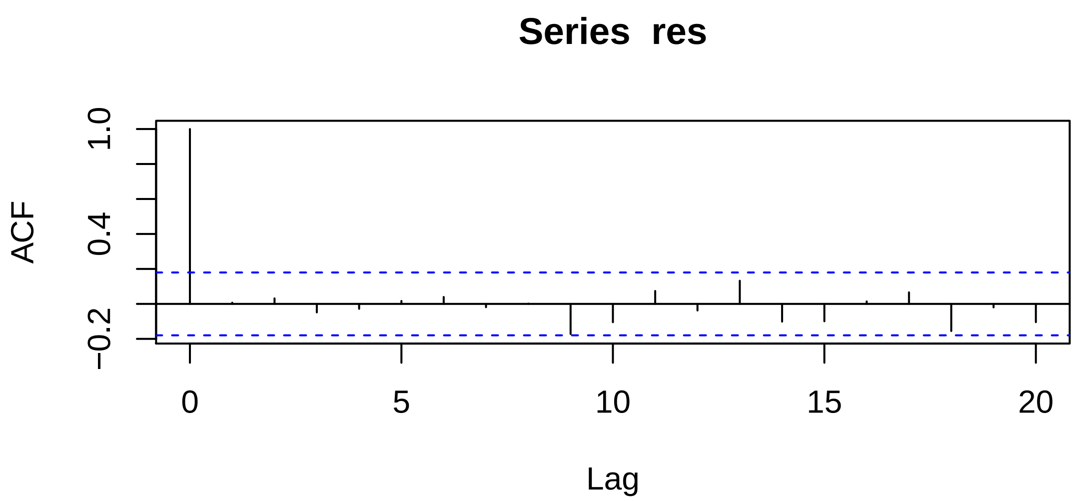



```
# dlaczego pierwsze trzynascie jest podejrzanych?
# bo dla pierwszych trzynastu brak nam danych - R je sobie przybliza,
# dlatego sie pojawiaja w ogole
```

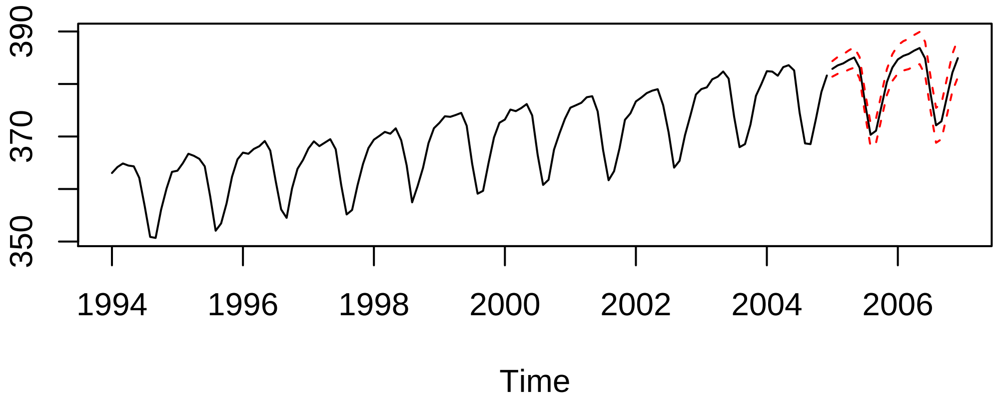
```r
# k)

res <- residuals(ma1d)[-c(1:13)]
qqnorm(res)
abline(0,1)
```




```r
Box.test(res,type="Ljung",lag=20)

##
##  Box-Ljung test
##
## data:  res
## X-squared = 17.94, df = 20, p-value = 0.5915

shapiro.test(res)    # uznajemy, ze rezidua maja rozklad normalny

##
##  Shapiro-Wilk normality test
##
## data:  res
## W = 0.982, p-value = 0.1134

# l)

acf(res)
```



```r
# m)

pred24 <- predict(ma1d,n.ahead=24)
ts.plot(pred24$pred,co,pred24$pred+2*pred24$se,
        pred24$pred-2*pred24$se,col=c("black","black","red","red"),
        lty=c(1,1,2,2))
```




```r
# n)

pred48 <- predict(ma1d,n.ahead=48)
ts.plot(pred48$pred,co,pred48$pred+2*pred48$se,
        pred48$pred-2*pred48$se,col=c("black","black","red","red"),
        lty=c(1,1,2,2) ,xlim=c(2004,2009))
```