```
### monthplot ###
monthplot(AP)
plot(AP)

# Création des variables muettes
vm = matrix(0,nr = length(AP), nc = 12)

for(i in 1:12){ v = rep(0,12)
v[i] = 1
vm[,i] = rep(v,12) }

mod = lm(log(AP)~temps+t+vm[,1]+vm[,2]+vm[,3]+vm[,4]+vm[,5]+vm[,6]
+vm[,7]+vm[,8]+vm[,9]+vm[,10]+vm[,11]+vm[,12])
summary(mod)

mod = lm(log(AP)~temps+t+vm[,1]+vm[,2]+vm[,3]+vm[,4]+vm[,5]+vm[,6]
+vm[,7]+vm[,8]+vm[,9]+vm[,10]+vm[,11]+vm[,12]-1)
summary(mod)

mod = lm(log(AP)~temps+t+vm[,1]+vm[,2]+vm[,3]+vm[,4]+vm[,5]+vm[,6]
+vm[,7]+vm[,8]+vm[,9]+vm[,10]+vm[,11])
summary(mod)

# Vérification
t1 = matrix(temps,nrow=144,ncol=1);
t2 = matrix(t,nrow=144,ncol=1)
X = cbind(rep(1,144),t1,t2,vm);X
Y = matrix(log(AP),nrow=144,ncol=1);Y
solve(t(X)%*%X)%*%(t(X)%*%Y)
# Suppression de la dernière variable
X = cbind(rep(1,144),t1,t2,vm[,1:11]);
res1=solve(t(X)%*%X)%*%(t(X)%*%Y)
# suprimer le vecteur 1
X = cbind(t1,t2,vm);
res2=solve(t(X)%*%X)%*%(t(X)%*%Y)

cbind(res1, res2)

### TD4 ###
age=c(35,45,55,65,75)
tension=c(114,124,143,158,166)
lm(tension~age)

X=cbind(rep(1,5),age);X
Y=matrix(tension,nrow=5,ncol=1);Y
solve(t(X)%*%X)%*%(t(X)%*%Y) # (X'X)^(-1)X'Y

AP =  AirPassengers
plot(AP); plot(AP, type = "p")
temps = time(AP);

mod1 = lm(AP~temps); summary(mod1);
shapiro.test(mod1$residuals)
# R2 = 0.8536, p-value = 3.419e-05
par(mfrow = c(3,2))
plot(mod1, which = 1); plot(mod1, which = 3)
mod2 = lm(log(AP)~temps); summary(mod2);
shapiro.test(mod2$residuals)
# R2 = 0.9015, p-value = 0.09457
plot(mod2, which = 1); plot(mod2, which = 3)
# mod2 : graphique 1  a une forme de bande -->
# var(e) est constante. on observe une courbure
# qui indique la relation parabolique

t = (temps-mean(temps))^2
mod3 = lm(log(AP)~temps+t);
summary(mod3);
shapiro.test(mod3$residuals)
# R2 = 0.9074, p-value = 0.04593
```

```
plot(mod3, which = 1); plot(mod3, which = 3)

par(mfrow = c(2,1))
plot(mod3, which = 1);
plot(mod3$residuals, type = "l")
plot(rnorm(144, m=0, s=0.1353), type = "l")

### TD3 exo2 ###
oeufs = read.table("U:/Documents/eco2324/oeufs.txt", header =T)

reg.oeufs=lm(POIDS~HAUTEUR,oeufs)
summary(reg.oeufs)
# P_chapeau = 28.526 * H-132.917 (modele estime)
# R2 = 0.9637
par(mfrow=c(2,2))
plot(reg.oeufs)
# graphique 1 : courbure importante-->E(e) n'est pas constante
# les residus ne sont pas indep.
shapiro.test(reg.oeufs$residuals)
# p-value = 0.03274 --> rejeter H0, e ne suit pas N(0, sigma2)
anova(reg.oeufs)
# sce = 7498.7, scr = 282.2
plot(oeufs$HAUTEUR,oeufs$POIDS)
abline(reg.oeufs)

# log - lineaire
lnh=log(oeufs$HAUTEUR)
lnp=log(oeufs$POIDS)
reg.lnOeufs=lm(lnp~lnh)

summary(reg.lnOeufs)
# ln(P)_chapeau = -1.67232 + 3.00357*ln(H)
# P_chapeau = exp(-1.67232)*H^(3.00357)
# R2 = 0.9898 > 0.9637 On ne peut pas conclure car
# on ne peut pas comparer ces deux R2
par(mfrow=c(2,2))
plot(reg.lnOeufs)

plot(reg.oeufs, which = 1)
plot(reg.oeufs, which = 3)
plot(reg.lnOeufs, which = 1)
plot(reg.lnOeufs, which = 3)
# les resultats sont ameliores selons les graphiques
shapiro.test(reg.lnOeufs$residuals)
# p-value = 0.378 > 0.05 : accepter H0
# les residus suivent une loi normale

anova(reg.lnOeufs)
# sce = 1.7389, scr = 0.0180
# nous ne pouvons pas comparer les R2 de ces modeles
plot(lnh,lnp)
abline(reg.lnOeufs)
# pseudo R2
lny=reg.lnOeufs$fitted.values
y=exp(lny) # P_chapeau calcule du 2eme modele
e=oeufs$POIDS-y
scr=sum(e^2)
pR2=1-scr/(7498.7+282.2)
pR2 # pseudo R2 = 0.987103 > 0.9637
# conclusion : mod2 est meilleur
```