The `MASS` library contains the `Boston` dataframe, which records `medv` (median house value) for 506 neighborhoods around Boston. It also includes several variables that can be used as predictors, such as `rm` (average number of rooms per house), `age` (average age of houses), and `lstat` (percent of households with low socioeconomic status). In this example, we will use `lstat` and fit several models with nonlinear terms to predict `medv`.

1. Fit a simple regression model with `lstat` as predictor
2. Verify regression assumptions.
3. Fit a quadratic regression model with `lstat` and as predictor
4. Use the `poly` function to fit higher order polynomial models
5. Compare the simple regression model with one with `log(rm)` as predictor

```
# mlrpoly.r

library(MASS)      # Boston
library(PASWR2)    # checking.plots()

d0 = Boston
m1=lm(medv~lstat,d0)
checking.plots(m1)
# some curvature in residuals vs fitted

# Non-linear Transformations of the Predictors

m2=lm(medv~lstat+I(lstat^2),d0)
checking.plots(m2)
# curvature is no longer there
par(mfrow=c(2,2))
plot(m2)

# compare m1 and m2
summary(m1)
summary(m2)
# m2 is better

m5=lm(medv~poly(lstat,5),d0)
checking.plots(m5)
# no curvature

# compare
summary(m5)
# residual std error always decreases when adding more terms
# adj-R2 increased

m7=lm(medv~poly(lstat,7),d0)
checking.plots(m7)
# adj-R2 decreased

# other possible transformations
m8 = lm(medv~rm,d0)
checking.plots(m8)

m9 = lm(medv~log(rm),d0)
checking.plots(m9)

m9b = lm(medv~I(rm^2),d0)
checking.plots(m9b)

# compare
summary(m8)
summary(m9)
summary(m9b)
```