# Time Series Linear Regression

## Code

* Just change the nValid to 36 and 50 to get the 36 and 50 months graphs
* The “magic” here is the formula used to create the “train.lm.trig”

|  |
| --- |
| library("forecast")  library("zoo")  # import data and create time series object  Amtrak.data <- read.csv("Amtrak data.csv")  ridership.ts <- ts(Amtrak.data$Ridership, start = c(1991, 1), end = c(2004, 3), freq = 12)  # separate training and validation data  nValid <- 36 # set to 36 or 50 months  nTrain <- length(ridership.ts) - nValid  train.ts <- window(ridership.ts, start = c(1991, 1), end = c(1991, nTrain))  valid.ts <- window(ridership.ts, start = c(1991, nTrain + 1), end = c(1991, nTrain + nValid))  ### Time Series Linear Regression  train.lm.trig <- tslm(train.ts ~ trend + I(trend^2) + I(sin(2\*pi\*trend/1)) + I(cos(2\*pi\*trend/1)))  #summary(train.lm.trig)  train.lm.trig.pred <- forecast(train.lm.trig, h = nValid, level = 95)  # trailing and center moving average  ma.trailing <- rollmean(ridership.ts, k = 12, align = "right")  ma.centered <- ma(ridership.ts, order = 12)  ### Time Series Linear Regression  plot(train.lm.trig.pred, ylim = c(1300, 2600), ylab = "Ridership", xlab = "Time", bty = "l", xaxt = "n", xlim = c(1991,2006.25), main = "", flty = 2)  axis(1, at = seq(1991, 2006, 1), labels = format(seq(1991, 2006, 1)))  lines(train.lm.trig.pred$fitted, lwd = 2, col = "blue")  lines(valid.ts)  lines(c(2004.25 - 3, 2004.25 - 3), c(0, 3500))  lines(c(2004.25, 2004.25), c(0, 3500))  text(1996.25, 2500, "Training")  text(2002.75, 2500, "Validation")  text(2005.25, 2500, "Future")  arrows(2004 - 3, 2450, 1991.25, 2450, code = 3, length = 0.1, lwd = 1,angle = 30)  arrows(2004.5 - 3, 2450, 2004, 2450, code = 3, length = 0.1, lwd = 1,angle = 30)  arrows(2004.5, 2450, 2006, 2450, code = 3, length = 0.1, lwd = 1, angle = 30)  lines(ma.trailing, lwd = 2, lty = 2)  lines(ma.centered, lwd = 2) |

