ESM 201 Assignment 1

Brad Anderson

February 9, 2018

library(tidyverse)

## Loading tidyverse: ggplot2  
## Loading tidyverse: tibble  
## Loading tidyverse: tidyr  
## Loading tidyverse: readr  
## Loading tidyverse: purrr  
## Loading tidyverse: dplyr

## Conflicts with tidy packages ----------------------------------------------

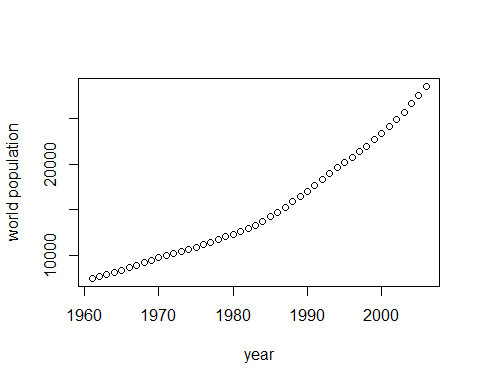
## filter(): dplyr, stats  
## lag(): dplyr, stats

countries <- read\_csv("C:/Users/Brad/github/eco1/data/countries.csv")

## Parsed with column specification:  
## cols(  
## country = col\_character(),  
## status = col\_character(),  
## year = col\_integer(),  
## food\_pc = col\_double(),  
## education = col\_double(),  
## hiv = col\_double(),  
## population = col\_double()  
## )

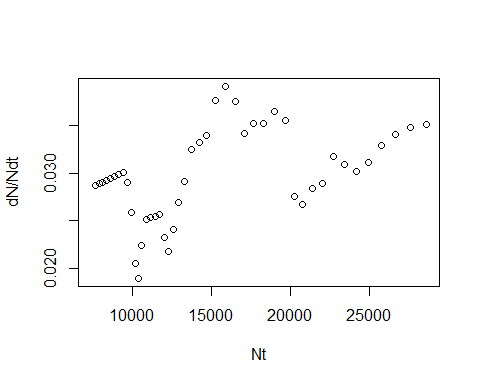
ugandadf <- countries %>%   
 filter( country == "Uganda")

plot(ugandadf$year, ugandadf$population, xlab="year", ylab="world population")

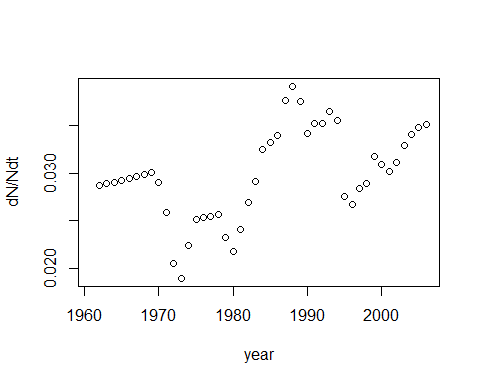


ugandadf$yearminus1<-(ugandadf$year-1) # add a column for the previous year   
  
# Create N\_(t-1) column   
ugandadf$popminus1<-ugandadf$population[match(ugandadf$yearminus1, ugandadf$year)]

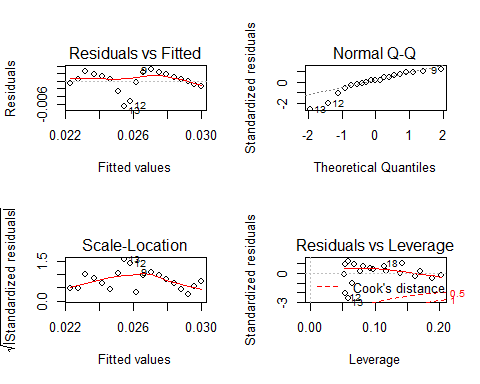
ugandadf$dNNdt<-log(ugandadf$population/ugandadf$popminus1)\*1/(ugandadf$year-ugandadf$yearminus1)  
  
#Plot results  
plot(ugandadf$population, ugandadf$dNNdt, xlab="Nt", ylab="dN/Ndt") # N OR N-1 HERE?



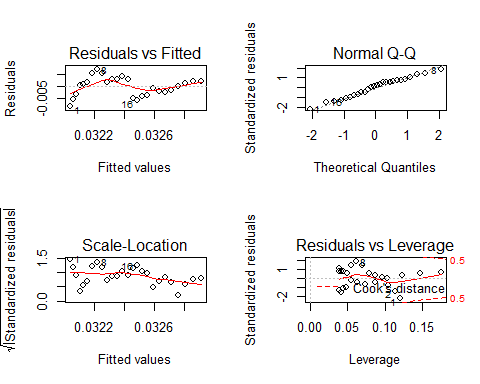
plot(ugandadf$year, ugandadf$dNNdt, xlab="year", ylab="dN/Ndt")



ugandalm\_1<-lm(dNNdt~population, data=ugandadf[ugandadf$year<=1980,])  
ugandalm\_2<-lm(dNNdt~population, data=ugandadf[ugandadf$year>1980,])  
  
par(mfrow=c(2,2))  
plot(ugandalm\_1)

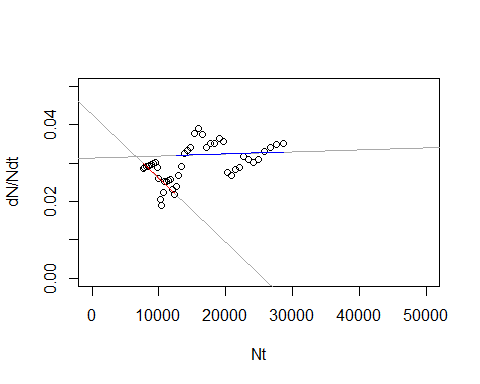


par(mfrow=c(2,2)) #Set parameters for plotting a 2x2 window  
plot(ugandalm\_2) # need to run this line at the same time as the one above to get 4-panel plot



Diagnostic plots seem OK.....

# make base plot   
plot(dNNdt~population, data=ugandadf, xlab="Nt", ylab="dN/Ndt", xlim=c(0, 50000), ylim=c(0,0.05))  
  
# add lines of best fit for both linear models   
abline(ugandalm\_1, col="darkgrey")  
abline(ugandalm\_2, col="darkgrey")  
  
# add within-sample predictions from both linear models   
dNNdt\_pred1<-predict(ugandalm\_1)  
dNNdt\_pred2<-predict(ugandalm\_2)   
  
lines(x=ugandadf$population[ugandadf$year<=1980 & ugandadf$year>1961], y=dNNdt\_pred1, col="red") # remove first year, for which there is no prediction (since there was no N-1)   
lines(x=ugandadf$population[ugandadf$year>1980], y=dNNdt\_pred2, col="blue")



uganda\_mlr <- lm(dNNdt ~ food\_pc + education + hiv + population, data = ugandadf)  
  
summary(uganda\_mlr)

##   
## Call:  
## lm(formula = dNNdt ~ food\_pc + education + hiv + population,   
## data = ugandadf)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -0.0075418 -0.0026207 0.0006135 0.0025456 0.0067158   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 3.021e-01 8.371e-02 3.609 0.000904 \*\*\*  
## food\_pc -2.506e-02 7.640e-03 -3.280 0.002265 \*\*   
## education 1.304e-02 4.465e-03 2.920 0.005928 \*\*   
## hiv 3.427e-04 2.482e-04 1.381 0.175606   
## population -2.518e-06 9.679e-07 -2.602 0.013262 \*   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 0.003864 on 37 degrees of freedom  
## (4 observations deleted due to missingness)  
## Multiple R-squared: 0.4092, Adjusted R-squared: 0.3453   
## F-statistic: 6.406 on 4 and 37 DF, p-value: 0.0005067

par(mfrow=c(2,2))  
plot(uganda\_mlr)

