Every 15 Minutes

L Lathrop

1/19/2017

## Introduction

In the summer of 2012, more than 100 maurauders on horseback crossed from Chad into Cameroon's Bouba Ndjidah National park with one mission: to kill as many elephants as possible without getting caught. Using rocket-propelled grenades and AK-47s, they mowed down hundreds of elephants, decimating entire herds, all for one reason: ivory.

The illegal killing of elephants in Africa for their ivory takes places sometimes on a small scale, with poisoned watermelons, and other times on a large scale, using belt-fed automatic weapons. But regardless of the method, when an elephant comes into contact with a poacher, more often than not, the poacher gets what he came for. This is why between 2010 and 2012 alone, some 100,000 elephants were slaughtered. There only "crime" that they hold the ivory that the world cannot seem to resist.

This project seeks to understand the socio-economic factors that may contribute to poaching. It is crucial for us to understand these contributing factors so that they may be remedied, thus allowing for the continuation of elephants in the wild. For it has been estimated that, at the current rate of killing, all elephants in the wild will be gone within 10 years.

Several researchers have studied the problem, and their findings will be taken into account in the model. For instance, Lemieux and Clark (2009) studied the effect of regulated and unregulated markets and showed that the presence of unregulated markets, either in-country or in a bordering country, had a significant impact on poaching activity.

The initial dataset included socio-economic and governance data on 34 countries with elephant populations in Central, Western, Eastern, and Southern Africa. These data came from the International Monetary Fund (IMF), the UN Development Programme (UNDP), the Unesco Institute for Statistics, The World Bank, and Transparency International (see the codebook for specifics). Information regarding regulated and unregulated ivory markets in Africa came from Lemieux and Clark (2009), and conflict data came from the Uppsala Conflict Data Programme (UCDP). Data concerning elephant populations and poaching came from the Monitoring the Illegal Killing of Elephants (MIKE) program of CITES.

Together, these data will allow a comprehensive assessment of one side of the poaching equation, namely, conditions in Africa tat may contribute to illegal elephant killing. What these data do not address is the demand side of the equation. That is, we do not have the benefit of trafficking and sales data whereby illegal ivory is transmitted and sold to consumers, largely in Asian and Southeast Asian countries. Unfortunately such data is not readily accessible and gathering such data would be beyond the scope of this project.

# LOAD IN THE NECESSARY PACKAGES  
library(ggplot2)  
library(ggthemes)  
library(GGally)  
library(grDevices)  
library(colorRamps)  
library(dplyr)  
library(RColorBrewer)  
library(reporttools)  
library(stargazer)  
library(fpp)  
library(xtable)  
library(plotly)  
library(Cairo)  
library(MASS)  
library(car)  
library(Amelia)  
library(corrplot)  
library(caret)  
options(xtable.floating = FALSE)  
options(xtable.timestamp = "")

## Data preprocessing and exploratory analysis

# READ IN DATASET  
elephants <- read.csv("elephant\_master\_reduced.csv", header = TRUE)

sapply(elephants, function(x) sum(is.na(x)))

## country year   
## 0 0   
## ISO2 ISO3   
## 14 0   
## region subregionid   
## 0 0   
## cap.lat cap.long   
## 0 0   
## NGDP\_RPCH NGDPD   
## 0 0   
## PCPI PCPIPCH   
## 0 0   
## GGX\_NGDP GGXWDG\_NGDP   
## 0 0   
## BCA HDI   
## 0 21   
## GNI Resource.Depletion   
## 21 191   
## Adult.literacy Primary.ed.enrollment   
## 13 42   
## Mean.Schooling Total.pop   
## 191 21   
## Pop.MultiDim.Povert Deprivation.Intensity   
## 38 38   
## Pop.Below.National.Poverty PPP.125.day   
## 38 38   
## International.Dev.Aid Corruption.Perception.Index   
## 21 21   
## Reg.Market Unreg.Market   
## 0 0   
## Reg.Market.Bordering Unreg.Market.Bordering   
## 0 0   
## Voice.Accountability Political.Stability   
## 21 21   
## Government.Effectiveness Rule.Law   
## 21 21   
## Corruption.Control Reg.Quality   
## 21 21   
## Armed.Conflict Non.State.Conflict   
## 0 0   
## Non.State.Conflict.Deaths PIKE.regional   
## 0 14   
## Definite.Possible Elephant.range   
## 241 241   
## Tot.Carcasses Illegal.Carcasses   
## 0 0   
## Percent.Illegal   
## 0

# CREATE VARIABLES TO ASSESS LEVEL OF ILLEGAL KILLING OF ELEPHANTS  
elephants$High.Illegal <- factor(with(elephants, ifelse((Percent.Illegal > 0.5), 1, 0)))  
table(elephants$High.Illegal)

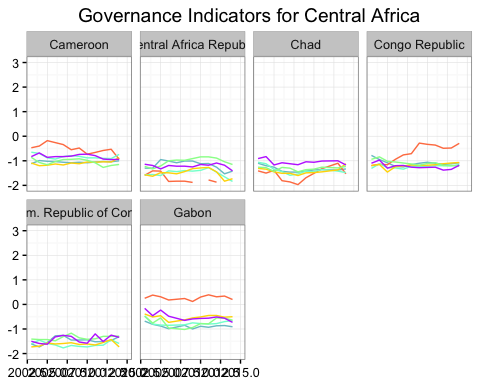
##   
## 0 1   
## 161 140

There are several missing values in the dataset, so imputation will be necessary.

Note that the mean number of elephant carcasses for a given year is 49.3189369 and the mean number of illegal carcasses is 24.5847176, indicating that approximately half of all elephant deaths recorded are due to illegal poaching.

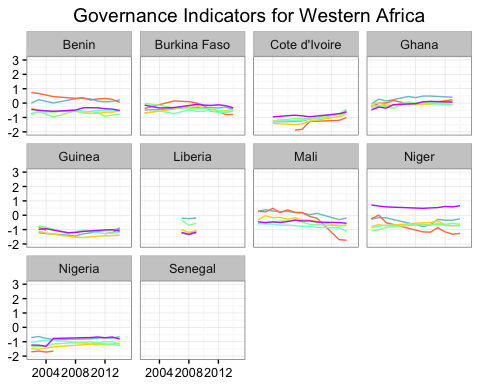
### Governance indicators by region

elephantsFC <- filter(elephants, subregionid=="FC")  
elephantsFW <- filter(elephants, subregionid=="FW")  
elephantsFE <- filter(elephants, subregionid=="FE")  
elephantsFS <- filter(elephants, subregionid=="FS")  
  
governanceFC <- ggplot(data=elephantsFC, aes(x=year))+  
 geom\_line(aes(y=Voice.Accountability), color="cadetblue3")+  
 geom\_line(aes(y=Political.Stability), color="coral")+  
 geom\_line(aes(y=Government.Effectiveness), color="aquamarine")+  
 geom\_line(aes(y=Rule.Law), color="gold")+   
 geom\_line(aes(y=Corruption.Control), color="palegreen")+  
 geom\_line(aes(y=Reg.Quality), color="darkorchid1")+  
 ylim(-2, 3) + xlab(NULL) + ylab(NULL) +  
 ggtitle("Governance Indicators for Central Africa") + theme\_bw()  
governanceFC + facet\_wrap(~country, ncol = 4)

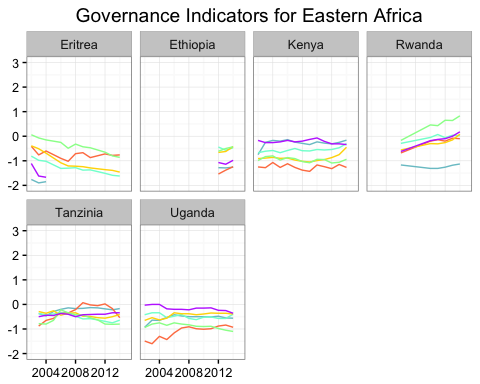
 

governanceFW <- ggplot(data=elephantsFW, aes(x=year))+  
 geom\_line(aes(y=Voice.Accountability), color="cadetblue3")+  
 geom\_line(aes(y=Political.Stability), color="coral")+  
 geom\_line(aes(y=Government.Effectiveness), color="aquamarine")+  
 geom\_line(aes(y=Rule.Law), color="gold")+  
 geom\_line(aes(y=Corruption.Control), color="palegreen")+  
 geom\_line(aes(y=Reg.Quality), color="darkorchid1")+  
 ylim(-2, 3) + xlab(NULL) + ylab(NULL) +  
 ggtitle("Governance Indicators for Western Africa")+ theme\_bw()  
governanceFW + facet\_wrap(~country, ncol = 4)

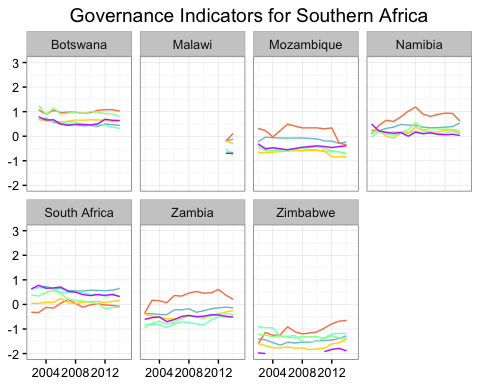
## geom\_path: Each group consists of only one observation. Do you need to  
## adjust the group aesthetic?  
## geom\_path: Each group consists of only one observation. Do you need to  
## adjust the group aesthetic?  
## geom\_path: Each group consists of only one observation. Do you need to  
## adjust the group aesthetic?  
## geom\_path: Each group consists of only one observation. Do you need to  
## adjust the group aesthetic?  
## geom\_path: Each group consists of only one observation. Do you need to  
## adjust the group aesthetic?  
## geom\_path: Each group consists of only one observation. Do you need to  
## adjust the group aesthetic?

governanceFE <- ggplot(data=elephantsFE, aes(x=year))+  
 geom\_line(aes(y=Voice.Accountability), color="cadetblue3")+  
 geom\_line(aes(y=Political.Stability), color="coral")+  
 geom\_line(aes(y=Government.Effectiveness), color="aquamarine")+  
 geom\_line(aes(y=Rule.Law), color="gold")+  
 geom\_line(aes(y=Corruption.Control), color="palegreen")+  
 geom\_line(aes(y=Reg.Quality), color="darkorchid1")+  
 ylim(-2, 3) + xlab(NULL) + ylab(NULL) +  
 ggtitle("Governance Indicators for Eastern Africa")+ theme\_bw()  
governanceFE + facet\_wrap(~country, ncol = 4)

governanceFS <- ggplot(data=elephantsFS, aes(x=year))+  
 geom\_line(aes(y=Voice.Accountability), color="cadetblue3")+  
 geom\_line(aes(y=Political.Stability), color="coral")+  
 geom\_line(aes(y=Government.Effectiveness), color="aquamarine")+  
 geom\_line(aes(y=Rule.Law), color="gold")+  
 geom\_line(aes(y=Corruption.Control), color="palegreen")+  
 geom\_line(aes(y=Reg.Quality), color="darkorchid1")+  
 ylim(-2, 3) + xlab(NULL) + ylab(NULL) +  
 ggtitle("Governance Indicators for Southern Africa")+ theme\_bw()  
governanceFS + facet\_wrap(~country, ncol = 4)

From these visualizations we can see that governance indicators show a great deal of volatility, especially with respect to the countries of central Africa.

# NORMALIZE AND STANDARDIZE THE DATA  
centered.elephants <- data.frame(scale(elephants[,c(9:28,33:38,41:47)]))  
centered.elephants <- cbind(elephants[,c(1:8, 29:32,39:40, 48)], centered.elephants)  
centered.elephants$High.Illegal <- as.integer(centered.elephants$High.Illegal)  
  
summary(centered.elephants)

## country year ISO2 ISO3   
## Kenya : 14 Min. :2002 KE : 14 KEN : 14   
## Namibia : 14 1st Qu.:2005 UG : 14 NAM : 14   
## South Africa: 14 Median :2009 ZA : 14 UGA : 14   
## Uganda : 14 Mean :2009 ZM : 14 ZAF : 14   
## Zambia : 14 3rd Qu.:2012 ZW : 14 ZMB : 14   
## Zimbabwe : 14 Max. :2015 (Other):217 ZWE : 14   
## (Other) :217 NA's : 14 (Other):217   
## region subregionid cap.lat cap.long   
## Central Africa :73 FC:73 Min. :-25.580 Min. :-17.29   
## Eastern Africa :62 FE:62 1st Qu.:-15.280 1st Qu.: 7.32   
## Southern Africa:83 FS:83 Median : 0.200 Median : 17.04   
## Western Africa :83 FW:83 Mean : -2.353 Mean : 17.45   
## 3rd Qu.: 9.020 3rd Qu.: 31.02   
## Max. : 15.190 Max. : 38.55   
##   
## Reg.Market Unreg.Market Reg.Market.Bordering  
## Min. :0.0000 Min. :0.000 Min. :0.0000   
## 1st Qu.:0.0000 1st Qu.:0.000 1st Qu.:0.0000   
## Median :0.0000 Median :0.000 Median :0.0000   
## Mean :0.1827 Mean :0.196 Mean :0.6213   
## 3rd Qu.:0.0000 3rd Qu.:0.000 3rd Qu.:0.0000   
## Max. :1.0000 Max. :1.000 Max. :3.0000   
##   
## Unreg.Market.Bordering Armed.Conflict Non.State.Conflict High.Illegal   
## Min. :0.000 Min. :0.000 Min. :0.0000 Min. :1.000   
## 1st Qu.:1.000 1st Qu.:0.000 1st Qu.:0.0000 1st Qu.:1.000   
## Median :2.000 Median :0.000 Median :0.0000 Median :1.000   
## Mean :1.625 Mean :0.186 Mean :0.1561 Mean :1.465   
## 3rd Qu.:2.000 3rd Qu.:0.000 3rd Qu.:0.0000 3rd Qu.:2.000   
## Max. :4.000 Max. :1.000 Max. :1.0000 Max. :2.000   
##   
## NGDP\_RPCH NGDPD PCPI PCPIPCH   
## Min. :-8.5293 Min. :-0.3618 Min. :-0.5352 Min. :-5.8036   
## 1st Qu.:-0.3300 1st Qu.:-0.3096 1st Qu.:-0.3403 1st Qu.:-0.3793   
## Median : 0.1156 Median :-0.2634 Median :-0.2831 Median :-0.1315   
## Mean : 0.0000 Mean : 0.0000 Mean : 0.0000 Mean : 0.0000   
## 3rd Qu.: 0.4595 3rd Qu.:-0.1620 3rd Qu.:-0.1089 3rd Qu.: 0.2270   
## Max. : 6.0386 Max. :10.4778 Max. : 6.2945 Max. :10.8290   
##   
## GGX\_NGDP GGXWDG\_NGDP BCA HDI   
## Min. :-2.3099 Min. :-0.9185 Min. :-4.54036 Min. :-1.9909   
## 1st Qu.:-0.6149 1st Qu.:-0.5280 1st Qu.:-0.11323 1st Qu.:-0.7003   
## Median :-0.1572 Median :-0.2934 Median : 0.06001 Median :-0.1235   
## Mean : 0.0000 Mean : 0.0000 Mean : 0.00000 Mean : 0.0000   
## 3rd Qu.: 0.4249 3rd Qu.: 0.1494 3rd Qu.: 0.16136 3rd Qu.: 0.4949   
## Max. : 5.1196 Max. : 9.5444 Max. : 8.32849 Max. : 2.3208   
## NA's :21   
## GNI Resource.Depletion Adult.literacy   
## Min. :-0.78873 Min. :-0.78585 Min. :-1.8109   
## 1st Qu.:-0.58528 1st Qu.:-0.60130 1st Qu.:-0.4468   
## Median :-0.40793 Median :-0.35802 Median : 0.1428   
## Mean : 0.00000 Mean : 0.00000 Mean : 0.0000   
## 3rd Qu.:-0.01034 3rd Qu.: 0.04883 3rd Qu.: 0.9773   
## Max. : 3.30277 Max. : 3.59869 Max. : 1.3620   
## NA's :21 NA's :191 NA's :13   
## Primary.ed.enrollment Mean.Schooling Total.pop   
## Min. :-2.2945 Min. :-1.6722 Min. :-0.7589   
## 1st Qu.:-0.6960 1st Qu.:-0.8276 1st Qu.:-0.6433   
## Median : 0.4682 Median : 0.1059 Median :-0.3322   
## Mean : 0.0000 Mean : 0.0000 Mean : 0.0000   
## 3rd Qu.: 0.7722 3rd Qu.: 0.5060 3rd Qu.: 0.1736   
## Max. : 1.2227 Max. : 2.1062 Max. : 5.1226   
## NA's :42 NA's :191 NA's :21   
## Pop.MultiDim.Povert Deprivation.Intensity Pop.Below.National.Poverty  
## Min. :-0.77386 Min. :-1.8583 Min. :-1.88540   
## 1st Qu.:-0.57802 1st Qu.:-0.8221 1st Qu.:-0.68390   
## Median :-0.36804 Median :-0.1059 Median : 0.05715   
## Mean : 0.00000 Mean : 0.0000 Mean : 0.00000   
## 3rd Qu.:-0.02326 3rd Qu.: 0.5798 3rd Qu.: 0.66509   
## Max. : 3.59793 Max. : 2.1646 Max. : 1.91336   
## NA's :38 NA's :38 NA's :38   
## PPP.125.day International.Dev.Aid Corruption.Perception.Index  
## Min. :-1.89838 Min. :-0.7880 Min. :-1.4858   
## 1st Qu.:-0.57562 1st Qu.:-0.6316 1st Qu.:-0.6962   
## Median :-0.04553 Median :-0.2990 Median :-0.3014   
## Mean : 0.00000 Mean : 0.0000 Mean : 0.0000   
## 3rd Qu.: 0.80659 3rd Qu.: 0.3669 3rd Qu.: 0.2908   
## Max. : 2.14421 Max. : 7.3977 Max. : 3.1531   
## NA's :38 NA's :21 NA's :21   
## Voice.Accountability Political.Stability Government.Effectiveness  
## Min. :-2.2123 Min. :-2.3099 Min. :-1.94399   
## 1st Qu.:-0.8154 1st Qu.:-0.7518 1st Qu.:-0.82597   
## Median : 0.1171 Median : 0.0425 Median :-0.01403   
## Mean : 0.0000 Mean : 0.0000 Mean : 0.00000   
## 3rd Qu.: 0.6751 3rd Qu.: 0.8174 3rd Qu.: 0.38557   
## Max. : 1.7948 Max. : 1.9631 Max. : 2.42605   
## NA's :21 NA's :21 NA's :21   
## Rule.Law Corruption.Control Reg.Quality   
## Min. :-1.8022 Min. :-1.4243 Min. :-2.3549   
## 1st Qu.:-0.8068 1st Qu.:-0.7361 1st Qu.:-0.7376   
## Median : 0.1324 Median :-0.2157 Median : 0.1326   
## Mean : 0.0000 Mean : 0.0000 Mean : 0.0000   
## 3rd Qu.: 0.6060 3rd Qu.: 0.4095 3rd Qu.: 0.6048   
## Max. : 2.2276 Max. : 3.1580 Max. : 1.9437   
## NA's :21 NA's :21 NA's :21   
## Non.State.Conflict.Deaths PIKE.regional Definite.Possible   
## Min. :-0.234 Min. :-1.6331 Min. :-0.53052   
## 1st Qu.:-0.234 1st Qu.:-0.7531 1st Qu.:-0.51666   
## Median :-0.234 Median :-0.2788 Median :-0.42619   
## Mean : 0.000 Mean : 0.0000 Mean : 0.00000   
## 3rd Qu.:-0.234 3rd Qu.: 1.0344 3rd Qu.:-0.01505   
## Max. : 9.719 Max. : 1.7745 Max. : 3.91046   
## NA's :14 NA's :241   
## Elephant.range Tot.Carcasses Illegal.Carcasses   
## Min. :-0.8947 Min. :-0.53091 Min. :-0.48580   
## 1st Qu.:-0.7287 1st Qu.:-0.48785 1st Qu.:-0.46604   
## Median :-0.5119 Median :-0.36944 Median :-0.38700   
## Mean : 0.0000 Mean : 0.00000 Mean : 0.00000   
## 3rd Qu.: 0.4491 3rd Qu.:-0.03573 3rd Qu.:-0.05108   
## Max. : 3.0532 Max. : 7.07978 Max. : 8.38660   
## NA's :241   
## Percent.Illegal   
## Min. :-1.33852   
## 1st Qu.:-1.00876   
## Median : 0.06352   
## Mean : 0.00000   
## 3rd Qu.: 0.90474   
## Max. : 1.46556   
##

Several of the variables still appear to be highly skewed, so further transformation will be required, as the Amelia II package imputes missing values assuming a normal distribution.

# EXAMINE DATA FOR HIGHLY CORRELATED VARIABLES  
clrs <- brewer.pal(10, "Spectral")  
cors <- cor(centered.elephants[,9:47], use = "pairwise")  
corrplot <- corrplot.mixed(cors, col = clrs, number.cex=0.3,  
 tl.pos="lt", tl.cex=0.5, tl.col="black",  
 tl.srt=45)



# REMOVE HIGHLY CORRELATED VARIABLES  
pd.elephants <- subset(centered.elephants,   
 subset=TRUE, select=c(1:22,24:28,30:34,36:47))

Linear regression was run on each of the highly correlated variables to determine which variables should be removed. Three variables were deleted from the dataset. We also see that of the observed values of total elephant carasses in relation to illegal carcasses, 252 observations meet the condition of more than 50% of the total were killed illegally.

# SEPARATE DATA INTO TRAINING AND TESTING SETS  
inVal <- createDataPartition(pd.elephants$High.Illegal, p = 0.7, list=FALSE)  
eleph.train <- pd.elephants[inVal,]  
eleph.test <- pd.elephants[-inVal,]  
  
# IMPUTE MISSING VALUES  
library(snow)  
set.seed(1357)  
elephants.out.train <- amelia(eleph.train, m=1, frontend = FALSE,   
 idvars = c("ISO2", "ISO3", "region",   
 "subregionid", "cap.lat", "cap.long"),  
 ts = "year", cs = "country",   
 polytime = 0, intercs = TRUE, p2s = 1,   
 parallel="snow", ncpus = 3,   
 empri = .01\*nrow(centered.elephants))  
  
# Amelia II: Multiple Imputation  
# (Version 1.7.4, built: 2015-12-05)  
# Copyright (C) 2005-2017 James Honaker, Gary King and Matthew Blackwell  
# Refer to http://gking.harvard.edu/amelia/ for more information

# SPOT CHECK VARIABLES WITH HIGH MISSINGNESS  
par(mfrow=c(3,2))  
tscsPlot(elephants.out.train, cs = "Kenya", main = "Kenya",  
 var = "Resource.Depletion", ylim = c(-2,0))  
tscsPlot(elephants.out.train, cs = "Botswana", main = "Botswana",  
 var = "Resource.Depletion", ylim = c(-2,0))  
tscsPlot(elephants.out.train, cs = "Cameroon", main = "Cameroon",  
 var = "Definite.Possible", ylim = c(-2,0))  
tscsPlot(elephants.out.train, cs = "Chad", main = "Chad",  
 var = "Definite.Possible", ylim = c(-2,0))  
tscsPlot(elephants.out.train, cs = "Zimbabwe", main = "zimbabwe",  
 var = "Primary.ed.enrollment", ylim = c(-2,2))  
tscsPlot(elephants.out.train, cs = "Congo Republic", main = "Congo Republic",  
 var = "Primary.ed.enrollment", ylim = c(-2,2))

