Exploring South Asia : Exploratory Data Analysis of South Asian Countries

Ashwed Patil

Contents

Introduction

This report focuses on exploring development indicators of South Asian countries, particularly - **Population, Economic Growth and Environmental Development**. We investigate how South Asian countries have developed with respect to these indicators from 1960-2014 and how each individual nation has performed as compared to its counterparts.

Data

The data has been imported from the World Development Indicators dataset at Kaggle.

```
library(readr) #Read csv files
library(dplyr) #Data Manipulation
library(stringr) #String Manipulation
library(ggplot2) #Data Visualization
library(knitr)
opts_chunk$set(tidy.opts=list(width.cutoff = 60),tidy=TRUE)
#Reading the CSV file
indicators <- read.csv("~/world-development-indicators/Indicators.csv")

#Subset of South Asian countries
south_asia <- c("AFG","BGD", "BTN", "IND", "MDV", "NPL", "PAK", "LKA")
south_asia_indicators <- subset(indicators, indicators$CountryCode %in% south_asia)
```

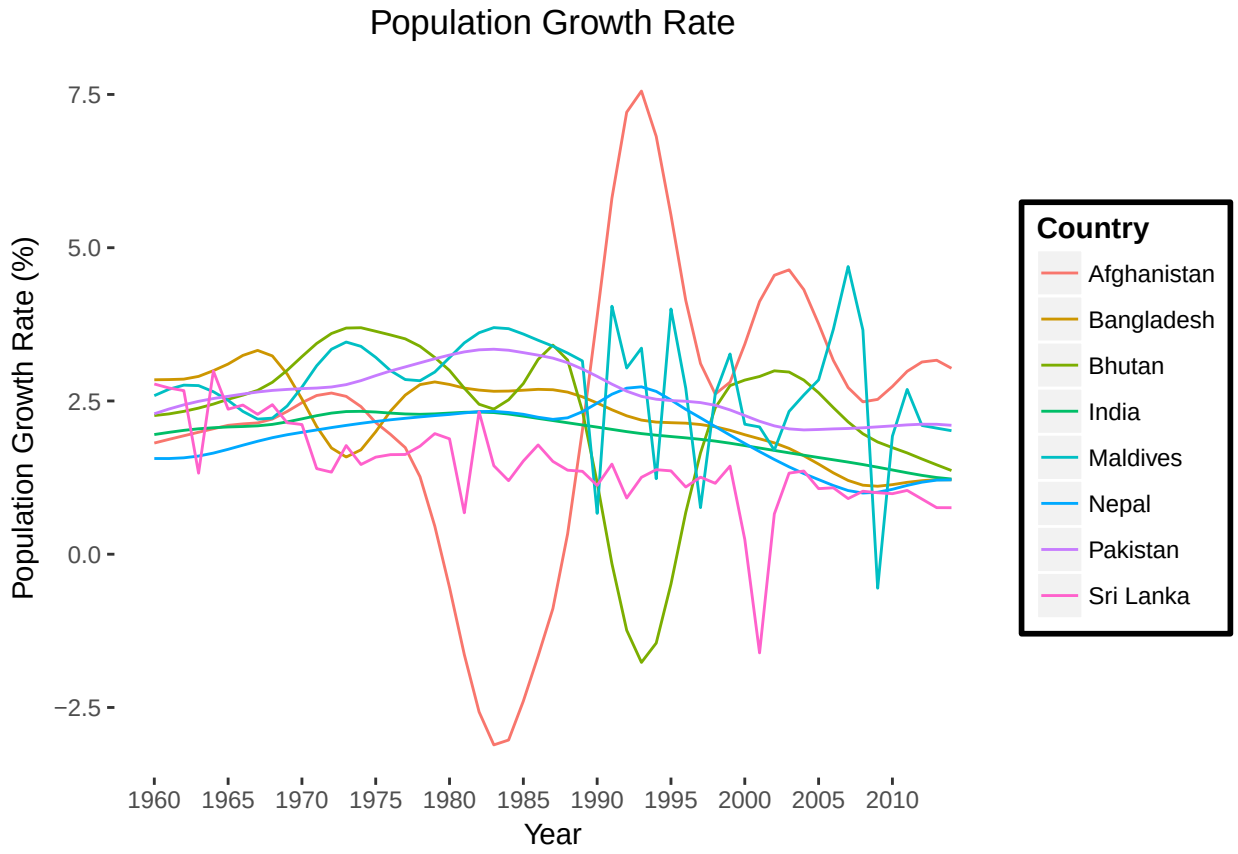
Population

Population Growth Rate

```
# Creating the Population subset
population <- subset(south_asia_indicators, south_asia_indicators$IndicatorCode ==
    "SP.POP.GROW")

# Creating a reusable theme function for visualization
my_theme <- theme(plot.title = element_text(hjust = 0.5), panel.grid.major = element_blank(),
    panel.grid.minor = element_blank(), panel.background = element_blank(),
    legend.title = element_text(face = "bold"), legend.background = element_rect(color = "black",
        size = 1))

# Visualizing Population Growth Rate
ggplot(data = population, aes(Year, Value)) + geom_line(aes(col = CountryName)) +
    scale_x_continuous(breaks = seq(1960, 2014, 5)) + ggtitle("Population Growth Rate") +
    ylab("Population Growth Rate (%)") + scale_color_discrete(name = "Country") +
    my_theme
```



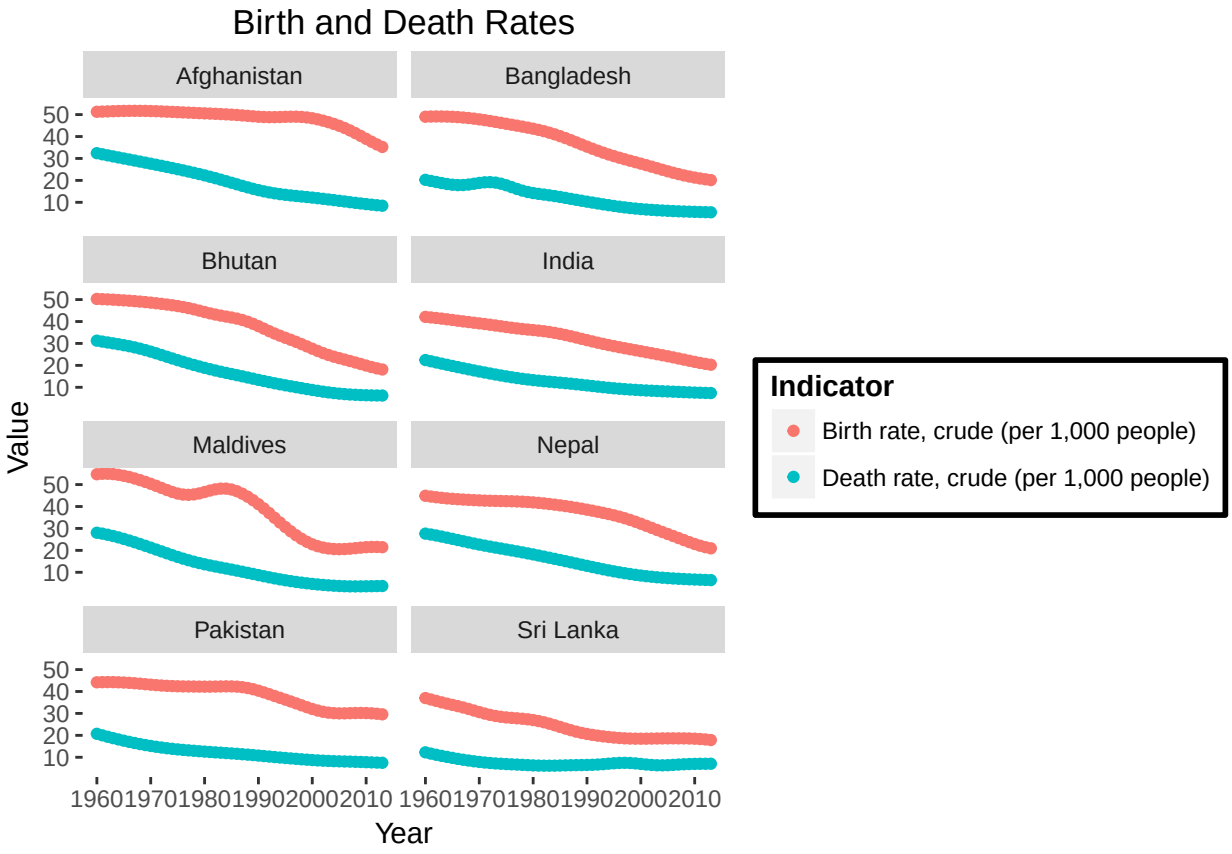
From the plot, it seems like most of the South Asian countries had an average population growth rate of 2% from 1960 to 2014. Afghanistan had a lot of variance in its population growth rate with a decline in the 1980-1985 period and a very high growth rate between 1990-1995. The decline can be partly attributed to the Soviet invasion of Afghanistan during this period. Bhutan saw a sharp decline in its population growth from the 1990 to 1995 due to the interethnic conflict that occured in the country during this period. Sri Lanka too

experienced a decline in its population growth rate in the 2000-2005 period due to the Sri Lankan civil war.

Birth and Death Rates

```
# Creating the Birth and Death Rates Indicators Subset
birth_death <- subset(south_asia_indicators, south_asia_indicators$IndicatorCode %in%
    c("SP.DYN.CBRT.IN", "SP.DYN.CDRT.IN"))

# Visualizing Birth and Death Rates
ggplot(data = birth_death, aes(Year, Value)) + geom_point(aes(col = IndicatorName)) +
    scale_x_continuous(breaks = seq(1960, 2014, 10)) + facet_wrap(~CountryName,
    ncol = 2) + scale_color_discrete(name = "Indicator") + my_theme +
    ggtitle("Birth and Death Rates")
```



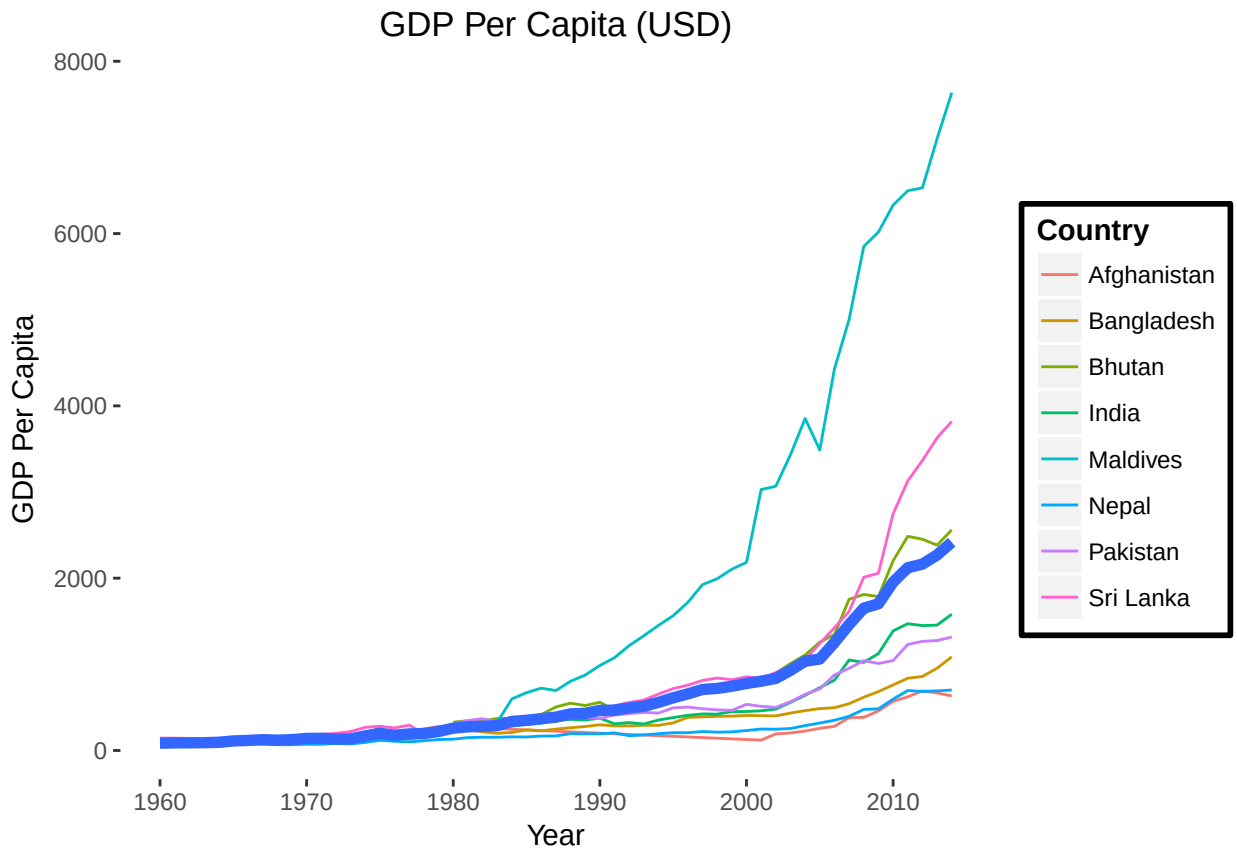
All South Asian countries seem to have experienced a decline in their birth rate during 1960 - 2015 except Maldives which saw a spike in its birth rate from 1980-1990. The death rate indicator follows a similar trend except for Sri Lanka which saw an increase in its death rate from 1995-2000 due to the civil war.

Economy

GDP Per Capita ($)

```
# GDP Per Capita Subset
gdp_per_capita <- subset(south_asia_indicators, south_asia_indicators$IndicatorCode ==
    "NY.GDP.PCAP.CD")

# Visualizing GDP Per Capita
ggplot(data = gdp_per_capita, aes(Year, Value)) + geom_line(aes(col = CountryName)) +
    geom_smooth(stat = "summary", fun.y = mean, linetype = 1,
        size = 2) + scale_x_continuous(breaks = seq(1960, 2014,
    10)) + ggtitle("GDP Per Capita (USD)") + ylab("GDP Per Capita") +
    scale_color_discrete(name = "Country") + my_theme
```

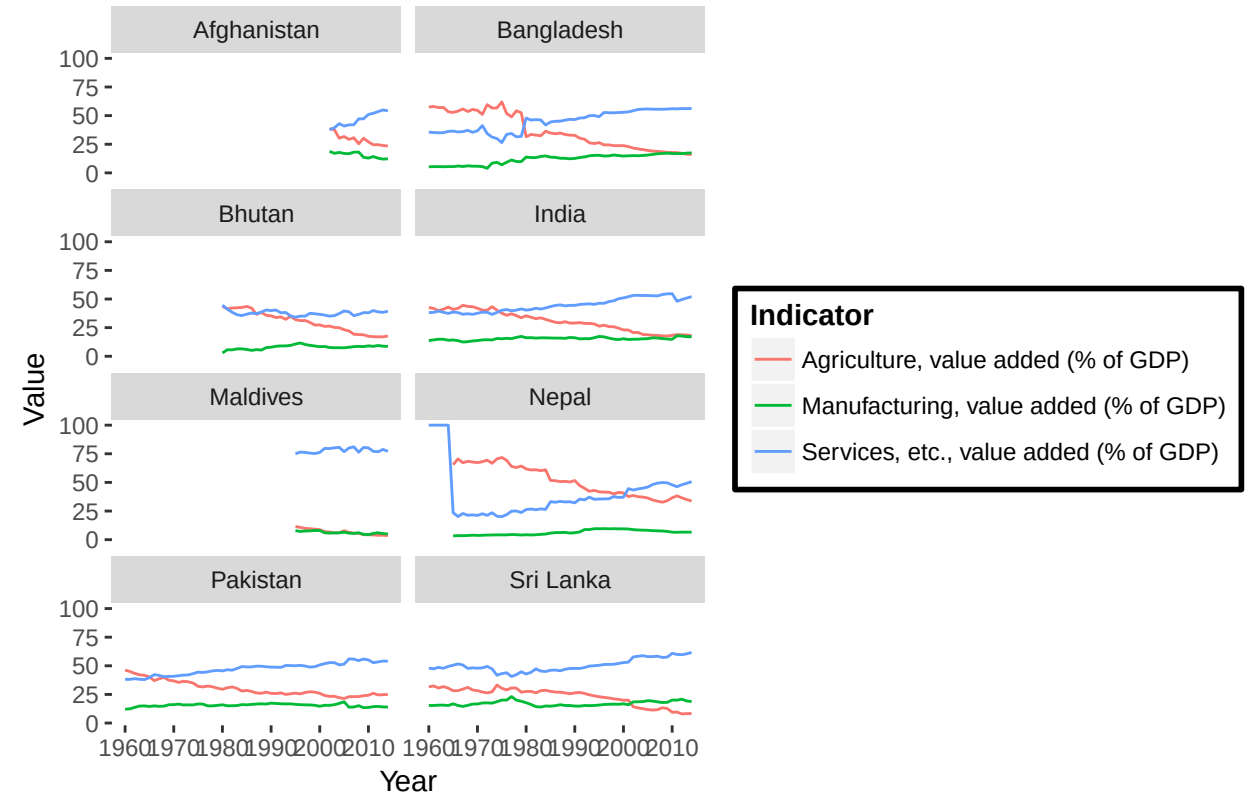


As we can infer from the plot, Maldives has the highest GDP Per Capita amongst all the South Asian countries at approximately $8000. Other nations seem to he struggling to improve their GDP per capita with most of them falling below the $3000 mark.

Contribution of Agriculture, Manufacturing and Service Industries

```
# Creating the sectors subset
sectors <- subset(south_asia_indicators, south_asia_indicators$IndicatorCode %in%
    c("NV.AGR.TOTL.ZS", "NV.IND.MANF.ZS", "NV.SRV.TETC.ZS"))
```

```
# Visualizing Sectors
ggplot(data = sectors, aes(Year, Value)) + geom_line(aes(col = IndicatorName)) +
    scale_x_continuous(breaks = seq(1960, 2014, 10)) + facet_wrap(~CountryName,
    ncol = 2) + scale_color_discrete(name = "Indicator") + my_theme +
    ggtitle(" Primary, Secondary and Tertiary Sector Contributions")
```

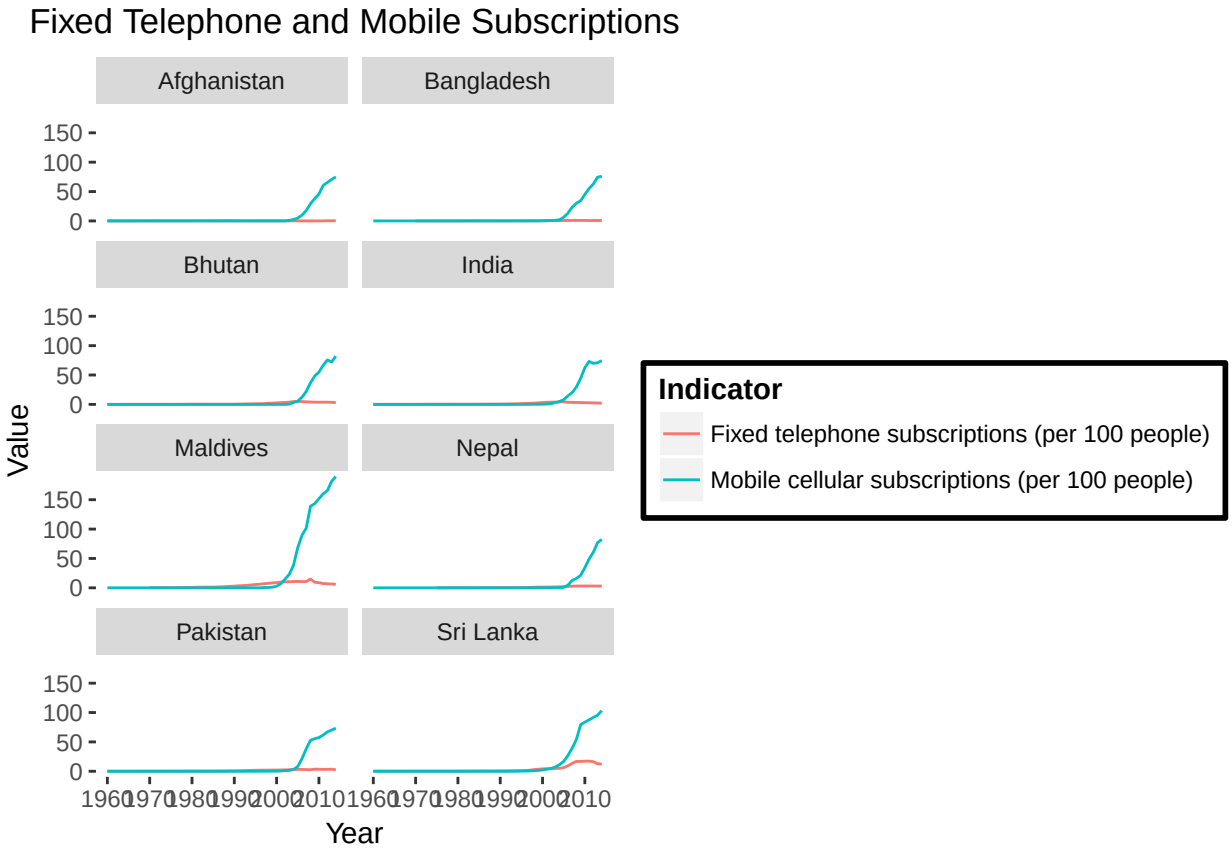


South Asia has been predominantly agricultural, but the plot indicates that service sector's contribution to the GDP of the respective countries has been increasing at a rapid rate while that of agriculture and manufacturing sectors continues to decline.

Telephone and Mobile Subscriptions

```
# Creating the subcription subset
tel_mobile <- subset(south_asia_indicators, south_asia_indicators$IndicatorCode %in%
    c("IT.MLT.MAIN.P2", "IT.CEL.SETS.P2"))

# Visualizing data
ggplot(data = tel_mobile, aes(Year, Value)) + geom_line(aes(col = IndicatorName)) +
    scale_x_continuous(breaks = seq(1960, 2014, 10)) + facet_wrap(~CountryName,
    ncol = 2) + scale_color_discrete(name = "Indicator") + my_theme +
    ggtitle("Fixed Telephone and Mobile Subscriptions")
```

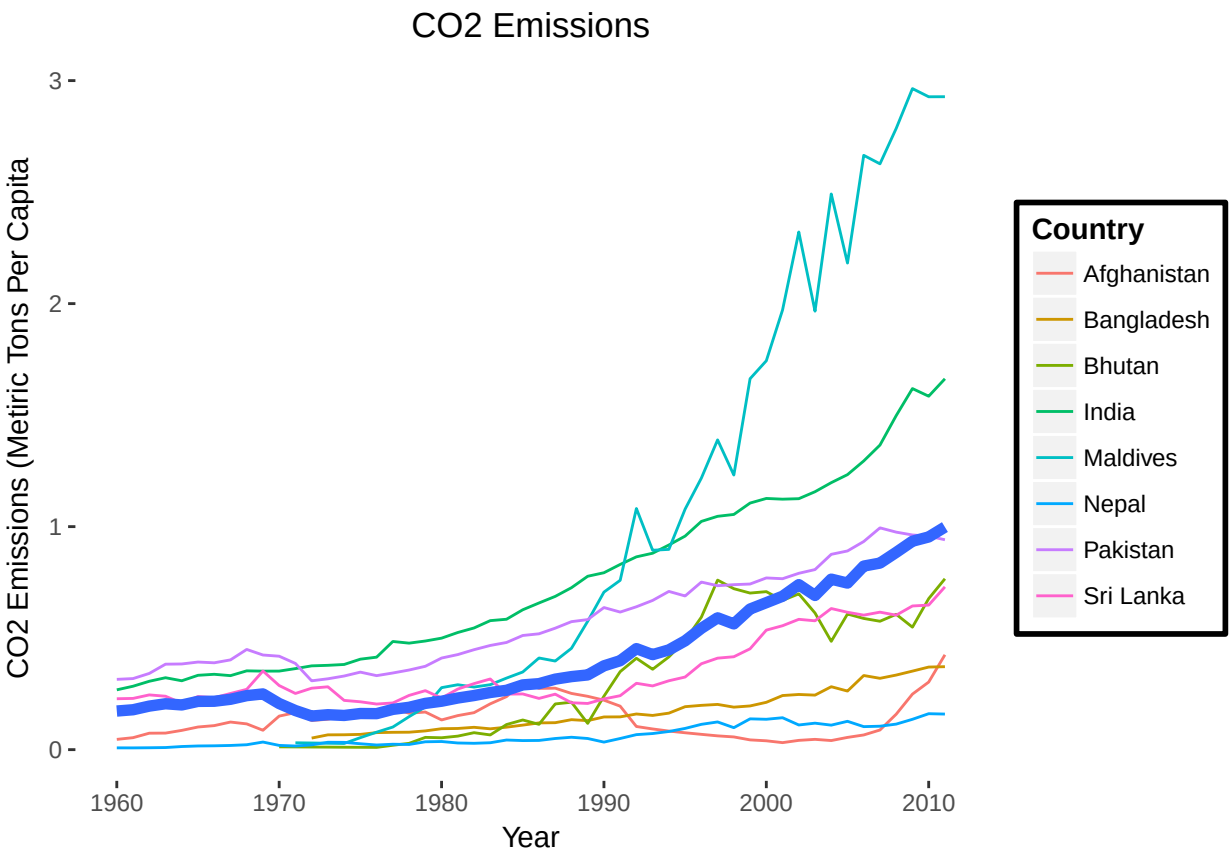


Mobile cellular subscriptions seem to have witnessed a massive growth in every South Asian country while fixed telephone subscription hasn't increased much. On the contrary, It has seen a decrease in some countries like Maldives and Bhutan.

Environment - CO2 Emissions

```
# CO2 Emissions Subset
co2 <- subset(south_asia_indicators, south_asia_indicators$IndicatorCode ==
    "EN.ATM.CO2E.PC")

# Visualizing CO2 Emissions
ggplot(data = co2, aes(Year, Value)) + geom_line(aes(col = CountryName)) +
    geom_smooth(stat = "summary", fun.y = mean, linetype = 1,
        size = 2) + scale_x_continuous(breaks = seq(1960, 2014,
    10)) + ggtitle("CO2 Emissions") + ylab("CO2 Emissions (Metiric Tons Per Capita") +
    scale_color_discrete(name = "Country") + my_theme
```



Maldives seems to be struggling in decreasing its CO2 Emissions, no wonder the country has been routinely ranked amongst the most affected nations by climate change. Bhutan has seen a relative decline in its CO2 emissions, thanks to the environmental friendly policies of its government.

Conclusion

South Asia seems to have progressed with mixed results, the population has grown rapidly but South Asian countries are struggling with improving their GDP per capita and environmental indicators. The telecommunications industry, though, has made a good breakthrough in the region.