Experiment Title

Web address for GitHub repository

Your Name

Abstract

Experimental overview. This section should be no longer than 250 words.

Contents

List of Tables

List of Figures

<Note: set up autoreferencing for figures and tables in your document>

1 Research Question and Rationale

2 Dataset Information

3 Exploratory Data Analysis and Wrangling

```
World_Bank_Master <-read.csv("../Raw/WorldBank_Raw2_4.8.19.csv")

#Data Subset
World_Bank_Filter <- filter(World_Bank_Master, Indicator.Name == "Forest area (% of land

WorldBank_Gather <- gather(World_Bank_Filter, "Year", "Level", X1960:X2018)

WorldBank_Gather <- select(WorldBank_Gather, -Indicator.Code)

WorldBank_Spread <-  spread(WorldBank_Gather, Indicator.Name, Level)

#Format as character
WorldBank_Spread$Year <- as.character(WorldBank_Spread$Year)

#create string
WB_String <- substr(WorldBank_Spread$Year, 2, 5)

#Get rid of X in date
WorldBank_Spread$Year = WB_String

#Format as date
#WB_Fixed$Year <- as.Date(WB_Fixed$Year)
WorldBank_Spread$Year <- as.Date(WorldBank_Spread$Year, format = "%Y") #can I get it to

class(WorldBank_Spread$Year)
```

```
## [1] "Date"
```

```
#Change column names
names(WorldBank_Spread) <- c("Country", "Indicator.Code", "Year", "Electricity Access",

#Save processed file
#write.csv(WorldBank_Spread, row.names = FALSE, file = "../Processed/WorldBank_Process

Four_Countries <- filter(WorldBank_Spread, Country == "Brazil" | Country == "Spain" | Co

Five_Countries <- filter(WorldBank_Spread, Country == "Brazil" | Country == "Kenya" | Co

Six_Countries <- filter(WorldBank_Spread, Country == "Brazil" | Country == "Kenya" | Cou
```

```
colnames(WorldBank_Spread)
```

```
##  [1] "Country"             "Indicator.Code"      "Year"
##  [4] "Electricity Access"  "Agriculture"         "Ag.Methane"
```

```
##  [7] "Ag.NO2"               "Aquaculture"          "ArableLand"
## [10] "CO2Emissions"         "Forest"               "RenewableElectricity"
```

```
dim(WorldBank_Spread)
```

```
## [1] 15576    12
```

```
head(WorldBank_Spread)
```

```
##       Country Indicator.Code       Year Electricity Access Agriculture
## 1 Afghanistan            AFG 1960-04-10                 NA          NA
## 2 Afghanistan            AFG 1961-04-10                 NA    57.74592
## 3 Afghanistan            AFG 1962-04-10                 NA    57.83782
## 4 Afghanistan            AFG 1963-04-10                 NA    57.91441
## 5 Afghanistan            AFG 1964-04-10                 NA    58.01091
## 6 Afghanistan            AFG 1965-04-10                 NA    58.01397
##   Ag.Methane Ag.NO2 Aquaculture ArableLand CO2Emissions Forest
## 1         NA     NA          NA         NA      414.371     NA
## 2         NA     NA          NA   11.71767      491.378     NA
## 3         NA     NA          NA   11.79426      689.396     NA
## 4         NA     NA          NA   11.87085      707.731     NA
## 5         NA     NA          NA   11.94743      839.743     NA
## 6         NA     NA          NA   11.94743     1008.425     NA
##   RenewableElectricity
## 1                   NA
## 2                   NA
## 3                   NA
## 4                   NA
## 5                   NA
## 6                   NA
```

```
summary(WorldBank_Spread)
```

```
##            Country      Indicator.Code       Year
##  Afghanistan   :   59   ABW    :   59   Min.   :1960-04-10
##  Albania       :   59   AFG    :   59   1st Qu.:1974-04-10
##  Algeria       :   59   AGO    :   59   Median :1989-04-10
##  American Samoa:   59   ALB    :   59   Mean   :1989-04-09
##  Andorra       :   59   AND    :   59   3rd Qu.:2004-04-10
##  Angola        :   59   ARB    :   59   Max.   :2018-04-10
##  (Other)       :15222   (Other):15222
##  Electricity Access  Agriculture        Ag.Methane
##  Min.   :  0.00     Min.   : 0.2628   Min.   :      0
##  1st Qu.: 53.11     1st Qu.:20.5547   1st Qu.:    120
##  Median : 93.94     Median :37.3659   Median :   3300
##  Mean   : 75.04     Mean   :37.0790   Mean   : 117609
##  3rd Qu.:100.00     3rd Qu.:52.3930   3rd Qu.:  24198
```

```
##  Max.   :100.00     Max.   :93.4407   Max.   :3464398
##  NA's   :8618       NA's   :2521      NA's   :5056
##      Ag.NO2           Aquaculture          ArableLand
##  Min.   :      0.0   Min.   :        0   Min.   : 0.0012
##  1st Qu.:     86.9   1st Qu.:       68   1st Qu.: 3.5315
##  Median :   2302.9   Median :     3758   Median : 9.5558
##  Mean   :  63590.8   Mean   :  1601961   Mean   :13.1413
##  3rd Qu.:  15076.6   3rd Qu.:    95447   3rd Qu.:17.5690
##  Max.   :2242932.7   Max.   :106004184   Max.   :73.3886
##  NA's   :5056        NA's   :4696        NA's   :2658
##   CO2Emissions          Forest         RenewableElectricity
##  Min.   :     -81   Min.   :    0.00   Min.   :  0.000
##  1st Qu.:     964   1st Qu.:   12.50   1st Qu.:  0.465
##  Median :   11463   Median :   31.18   Median : 16.961
##  Mean   :  736069   Mean   :   42.70   Mean   : 28.211
##  3rd Qu.:  143107   3rd Qu.:   46.96   3rd Qu.: 49.255
##  Max.   :36138285   Max.   :16735.00   Max.   :100.000
##  NA's   :3321       NA's   :8717       NA's   :8738
```

```
summary(WorldBank_Spread$Agriculture)
```

```
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.    NA's
##  0.2628 20.5547 37.3659 37.0790 52.3930 93.4407    2521
```

```
summary(WorldBank_Spread$Forest)
```

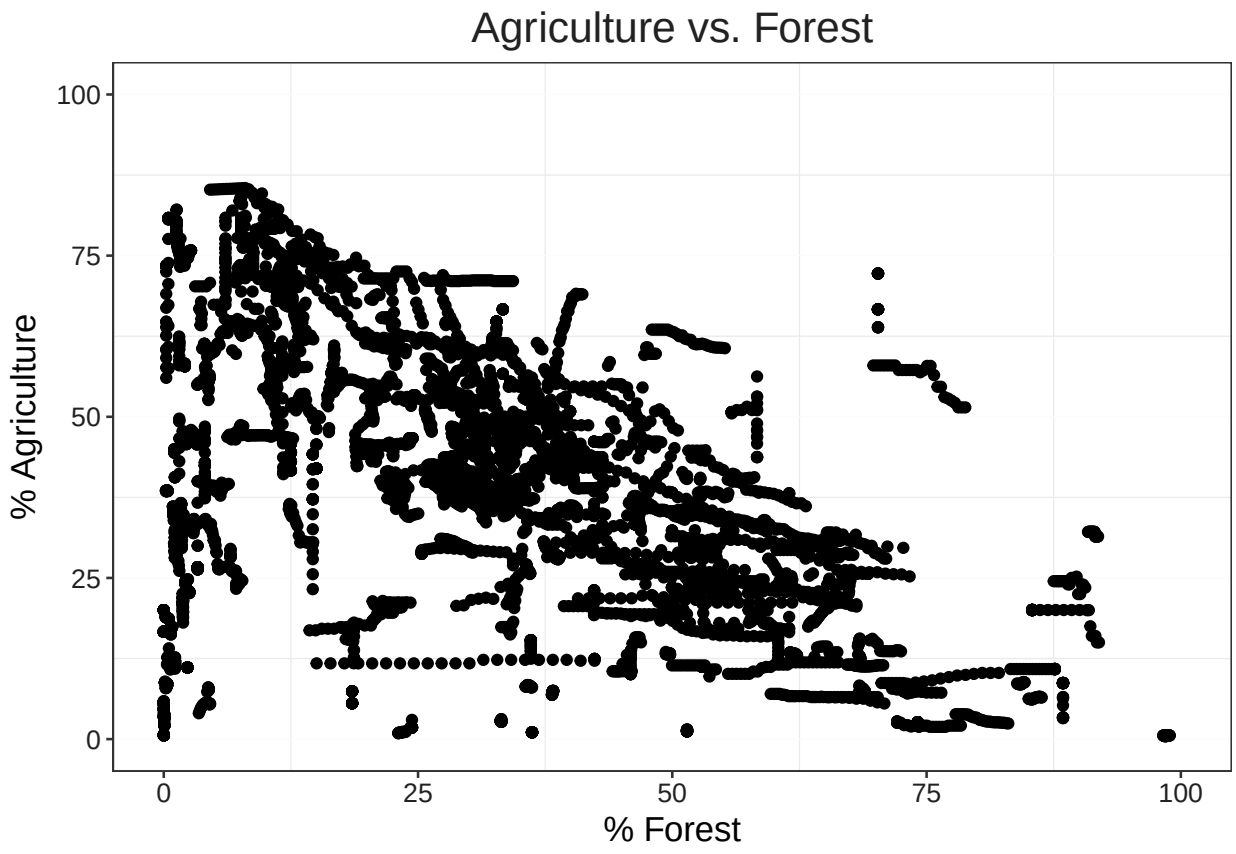
```
##     Min.  1st Qu.   Median     Mean  3rd Qu.     Max.     NA's
##     0.00    12.50    31.18    42.70    46.96 16735.00     8717
```

```
summary(WorldBank_Spread$`Renewable Electricity`)
```

```
## Length  Class   Mode
##      0   NULL   NULL
```

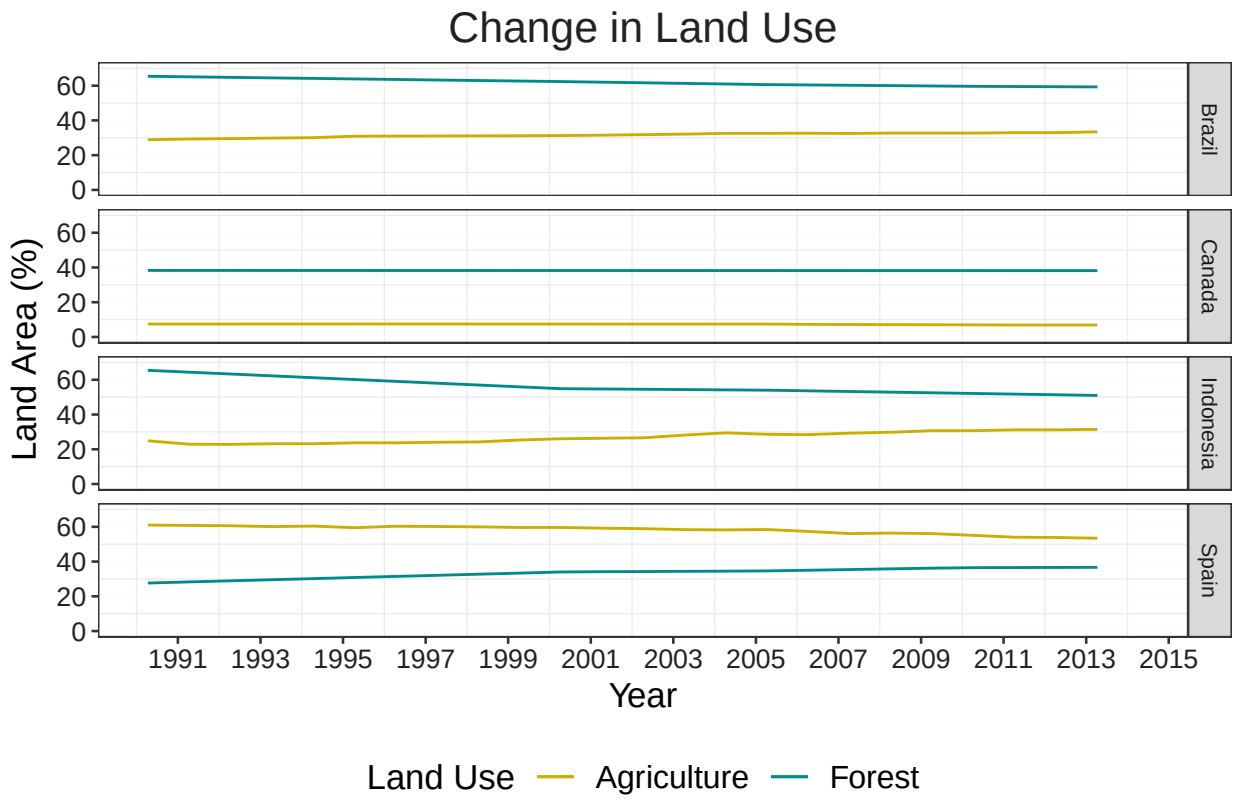
```
#Full Plot
AgVForest <-
  ggplot(WorldBank_Spread) +
  geom_point(aes(x = Forest, y = Agriculture)) +
  ggtitle("Agriculture vs. Forest") +
  ylab(expression("% Agriculture")) +
  xlab(expression("% Forest")) +
  scale_y_continuous(limits = c(0,100)) +
  scale_x_continuous(limits = c(0,100))
print(AgVForest)
```

```
## Warning: Removed 8897 rows containing missing values (geom_point).
```




```
LegendTitle1 <-"Land Use"

FourCountries.Facet <-
  ggplot(Four_Countries) +
  geom_line(aes(x = Year, y = Agriculture, color = "Agriculture")) +
  geom_line(aes(x = Year, y = Forest, color = "Forest")) +
  facet_grid(rows = vars(Country)) +
  ggtitle("Change in Land Use") +
  ylab(expression("Land Area (%)")) +
  scale_x_date(limits = as.Date(c("1990-04-09", "2014-04-09")),
  date_breaks = "24 months", date_labels = "%Y") +
  scale_y_continuous(limits = c(0,70)) +
  scale_color_manual(LegendTitle1,values = c("gold3", "darkcyan")) +
  labs(caption = "Data Source: World Bank")
print(FourCountries.Facet)
```

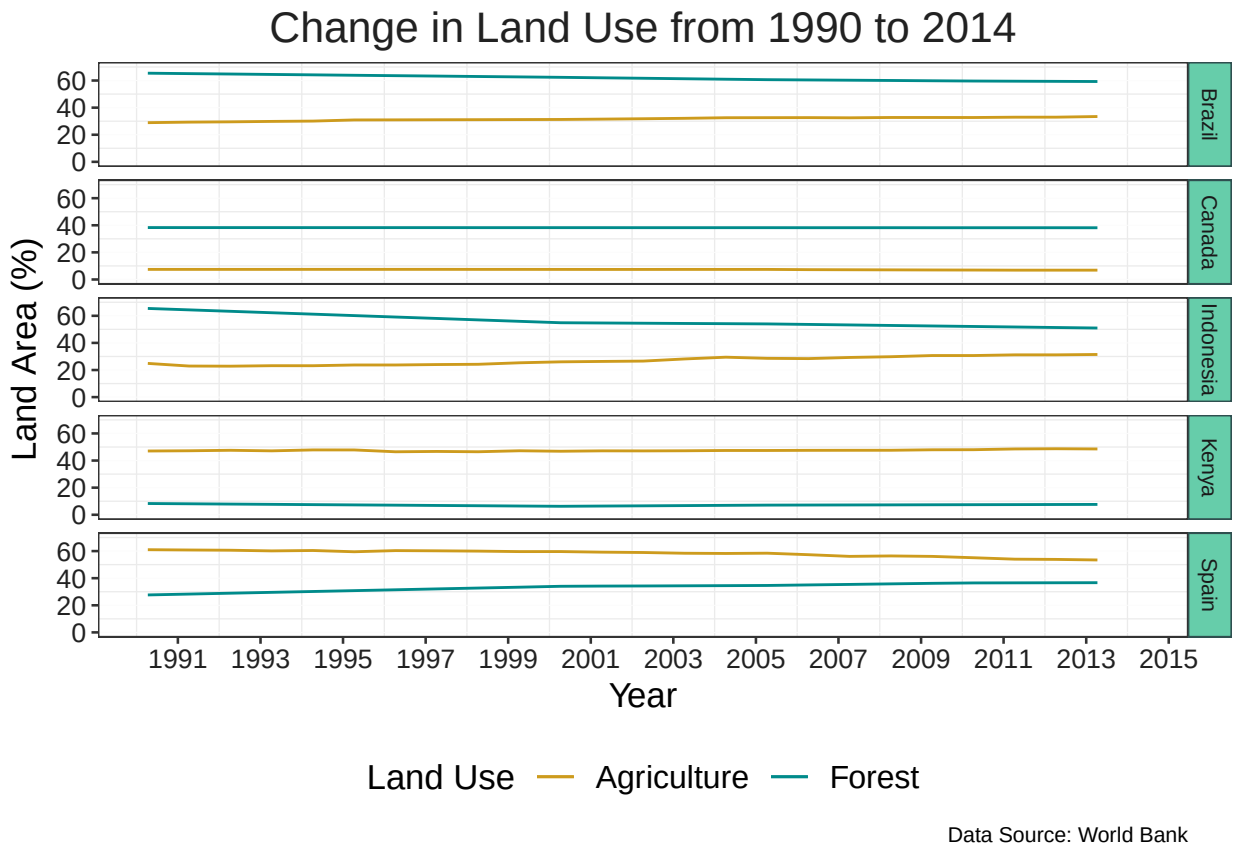
```
## Warning: Removed 35 rows containing missing values (geom_path).

## Warning: Removed 35 rows containing missing values (geom_path).
```




```
FiveCountries.Facet <-
  ggplot(Five_Countries) +
  geom_line(aes(x = Year, y = Agriculture, color = "Agriculture")) +
  geom_line(aes(x = Year, y = Forest, color = "Forest")) +
  facet_grid(rows = vars(Country)) +
  ggtitle("Change in Land Use from 1990 to 2014") +
  ylab(expression("Land Area (%)")) +
  scale_x_date(limits = as.Date(c("1990-04-09", "2014-04-09")),
  date_breaks = "24 months", date_labels = "%Y") +
  scale_y_continuous(limits = c(0,70)) +
  scale_color_manual(LegendTitle1,values = c("goldenrod3", "darkcyan")) +
  labs(caption = "Data Source: World Bank") +
  theme(strip.background = element_rect(fill= "aquamarine3", "darkslategray"))
print(FiveCountries.Facet)
```

```
## Warning: Removed 35 rows containing missing values (geom_path).
```

```
## Warning: Removed 35 rows containing missing values (geom_path).
```

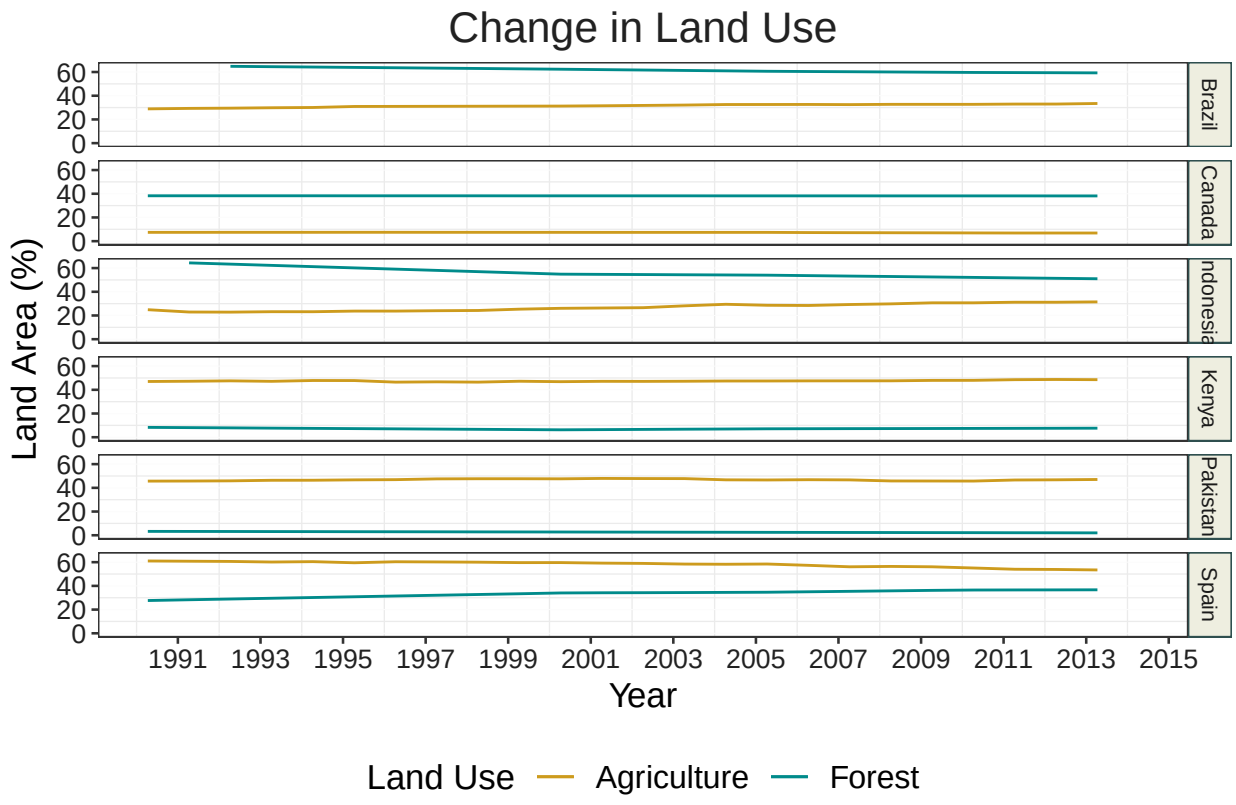



```
#seashell3 turquoise palegreen3 darkseagreen3 darkcyan aquamarine3
```

```
SixCountries.Facet <-
  ggplot(Six_Countries) +
  geom_line(aes(x = Year, y = Agriculture, color = "Agriculture")) +
  geom_line(aes(x = Year, y = Forest, color = "Forest")) +
  facet_grid(rows = vars(Country)) +
  ggtitle("Change in Land Use") +
  ylab(expression("Land Area (%)")) +
  scale_x_date(limits = as.Date(c("1990-04-09", "2014-04-09")),
  date_breaks = "24 months", date_labels = "%Y") +
  scale_y_continuous(limits = c(0,65)) +
  scale_color_manual(LegendTitle1,values = c("goldenrod3", "darkcyan")) +
  labs(caption = "Data Source: World Bank") +
  theme(strip.background = element_rect(fill= "ivory2", "darkslategray"))
print(SixCountries.Facet)
```

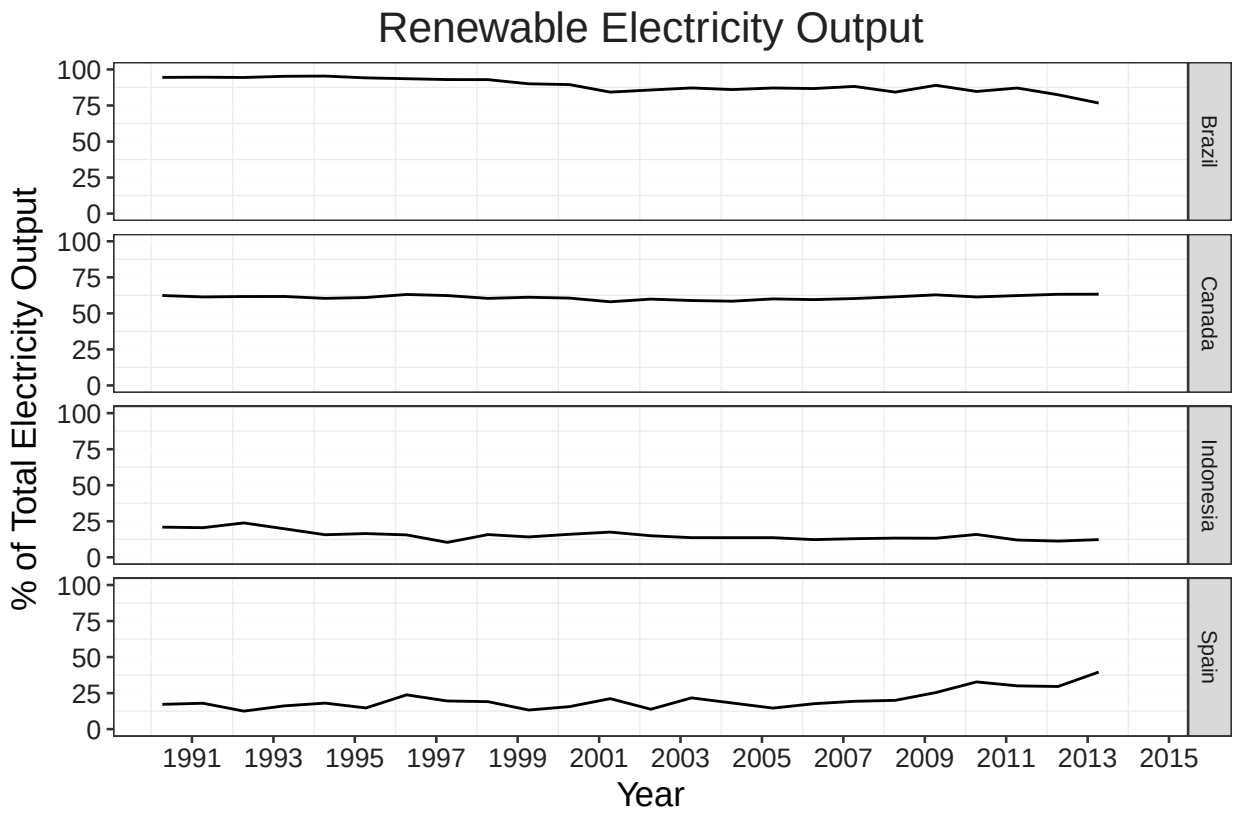
```
## Warning: Removed 35 rows containing missing values (geom_path).
```

```
## Warning: Removed 37 rows containing missing values (geom_path).
```



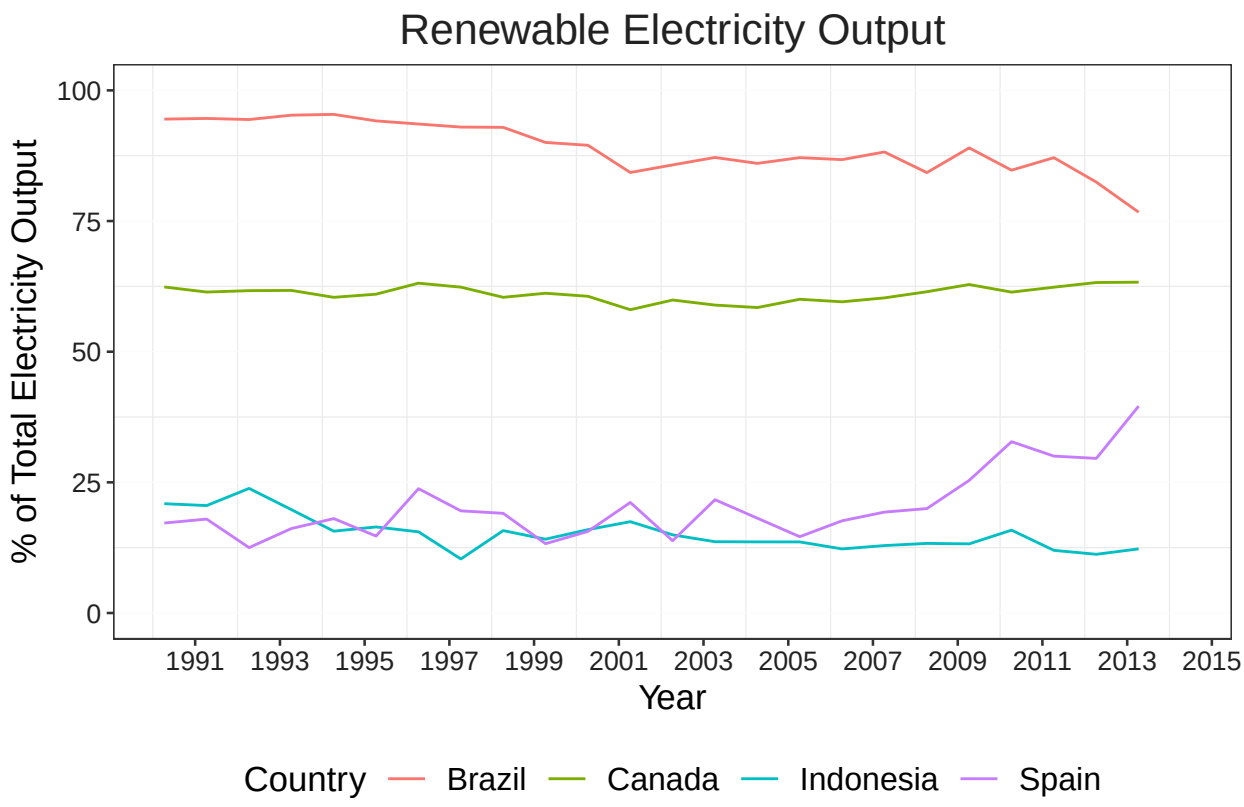

```
#Plot1
RenewableElectricity <-
  ggplot(Four_Countries) +
  geom_line(aes(x = Year, y = RenewableElectricity)) +
  facet_grid(rows = vars(Country)) +
  ggtitle("Renewable Electricity Output") +
  ylab(expression("% of Total Electricity Output")) +
  scale_x_date(limits = as.Date(c("1990-04-09", "2014-04-09")),
  date_breaks = "24 months", date_labels = "%Y") +
  scale_y_continuous(limits = c(0,100)) +
  labs(caption = "Data Source: World Bank")
print(RenewableElectricity)
```

```
## Warning: Removed 35 rows containing missing values (geom_path).
```




```
RE <-
  ggplot(data = Four_Countries, aes(x = Year, y = RenewableElectricity, color = Country))+
  geom_line()+
  ggtitle("Renewable Electricity Output") +
  ylab(expression("% of Total Electricity Output")) +
  scale_x_date(limits = as.Date(c("1990-04-09", "2014-04-09")),
  date_breaks = "24 months", date_labels = "%Y") +
  scale_y_continuous(limits = c(0,100)) +
  labs(caption = "Data Source: World Bank")
print(RE)
```

```
## Warning: Removed 140 rows containing missing values (geom_path).
```


Renewable Electricity Output
% of Total Electricity Output
100
75
50
25
0
1991
1993
1995
1997
1999
2001
2003
2005
2007
2009
2011
2013
2015
Year
Country
Brazil
Canada
Indonesia
Spain
Data Source: World Bank

```
# Plots
```

4 Analysis

5 Summary and Conclusions