Stat 184 Final project Report

Michael Yun

Research Question

How does the Regime type affect economic growth?

```
#load packages
library(tidyverse)
library(stargazer)
```

```
#load initial data
worldbankdata<-read.csv("https://raw.githubusercontent.com/Migueldesanta/STAT184/main/worldbank_data.cs
polity<-read.csv("https://raw.githubusercontent.com/Migueldesanta/STAT184/main/p5v2018.csv")
```

Data Description

World Development Indicators

Source: https://databank.worldbank.org/source/world-development-indicators#

GDP Growth: It measures the increase in a country's economic output over a specific period, which is 2010. And it's measured as a percentage change in the Gross Domestic Product.

Population: This represents the total number of people living in a country. Population is measured as an absolute number, counting all residents regardless of their legal status or citizenship.

Arable Land: Arable land refers to the land suitable for agriculture per person, which is key for growing crops. It's measured in hectares or acres.

Polity 5 Project

Source: https://view.officeapps.live.com/op/view.aspx?src=https%3A%2F%2Fwww.systemicpeace.org%2Finscr%2Fp5v2018.xls&wdOrigin=BROWSELINK

Polity 2 Score: The Polity2 score combines the "Autocracy" and "Democracy" scores from the Polity IV dataset, which evaluates various aspects of a political system's institutional characteristics. The scale is from -10 to 10, -10 represents most autocratic, 10 represents most democratic.

```
#Inspect Initial data of worldbankdata
head(worldbankdata)
```

```
##   Country.Name Country.Code                       Series.Name       Series.Code
## 1  Afghanistan          AFG             GDP growth (annual %) NY.GDP.MKTP.KD.ZG
## 2  Afghanistan          AFG      Population ages 15-64, total    SP.POP.1564.TO
## 3  Afghanistan          AFG Arable land (hectares per person) AG.LND.ARBL.HA.PC
## 4      Albania          ALB             GDP growth (annual %) NY.GDP.MKTP.KD.ZG
```

```
## 5      Albania          ALB      Population ages 15-64, total    SP.POP.1564.TO
## 6      Albania          ALB Arable land (hectares per person) AG.LND.ARBL.HA.PC
##   X2010..YR2010.
## 1    14.36244147
## 2       13722048
## 3    0.276448765
## 4    3.706938153
## 5        1963578
## 6    0.214897181
```

```
str(worldbankdata)
```

```
## 'data.frame':    798 obs. of  5 variables:
##  $ Country.Name  : chr  "Afghanistan" "Afghanistan" "Afghanistan" "Albania" ...
##  $ Country.Code  : chr  "AFG" "AFG" "AFG" "ALB" ...
##  $ Series.Name   : chr  "GDP growth (annual %)" "Population ages 15-64, total" "Arable land (hectares
##  $ Series.Code   : chr  "NY.GDP.MKTP.KD.ZG" "SP.POP.1564.TO" "AG.LND.ARBL.HA.PC" "NY.GDP.MKTP.KD.ZG"
##  $ X2010..YR2010.: chr  "14.36244147" "13722048" "0.276448765" "3.706938153" ...
```

```
names(worldbankdata)
```

```
## [1] "Country.Name"   "Country.Code"   "Series.Name"    "Series.Code"
## [5] "X2010..YR2010."
```

Comments: According to the inspection of worldbankdata, we have total 798 observations of 5 variables in the whole data set, which are "Country.Name","Country.Code" ,"Series.Name", "Series.Code" and "X2010..YR2010.". In this data set, each row(observation) represts the value of one indicator of a specific country in 2010. Some of the cases(data) are displayed above.

```
#Inspect Initial data of polity
head(polity)
```

```
##   p5   cyear ccode scode     country year flag fragment democ autoc polity
## 1  0 7001800   700   AFG Afghanistan 1800    0       NA     1     7     -6
## 2  0 7001801   700   AFG Afghanistan 1801    0       NA     1     7     -6
## 3  0 7001802   700   AFG Afghanistan 1802    0       NA     1     7     -6
## 4  0 7001803   700   AFG Afghanistan 1803    0       NA     1     7     -6
## 5  0 7001804   700   AFG Afghanistan 1804    0       NA     1     7     -6
## 6  0 7001805   700   AFG Afghanistan 1805    0       NA     1     7     -6
##   polity2 durable xrreg xrcomp xropen xconst parreg parcomp exrec exconst
## 1      -6      NA     3      1      1      1      3       3     1       1
## 2      -6      NA     3      1      1      1      3       3     1       1
## 3      -6      NA     3      1      1      1      3       3     1       1
## 4      -6      NA     3      1      1      1      3       3     1       1
## 5      -6      NA     3      1      1      1      3       3     1       1
## 6      -6      NA     3      1      1      1      3       3     1       1
##   polcomp prior emonth eday eyear eprec interim bmonth bday byear bprec post
## 1       6    NA     NA   NA    NA    NA      NA      1    1  1800     1   -6
## 2       6    NA     NA   NA    NA    NA      NA     NA   NA    NA    NA   NA
## 3       6    NA     NA   NA    NA    NA      NA     NA   NA    NA    NA   NA
## 4       6    NA     NA   NA    NA    NA      NA     NA   NA    NA    NA   NA
## 5       6    NA     NA   NA    NA    NA      NA     NA   NA    NA    NA   NA
```

```
## 6       6    NA     NA   NA    NA    NA      NA     NA   NA    NA    NA   NA
##   change d5 sf regtrans
## 1     88  1 NA       NA
## 2     NA NA NA       NA
## 3     NA NA NA       NA
## 4     NA NA NA       NA
## 5     NA NA NA       NA
## 6     NA NA NA       NA
```

```
str(polity)
```

```
## 'data.frame':    17574 obs. of  37 variables:
##  $ p5      : int  0 0 0 0 0 0 0 0 0 0 ...
##  $ cyear   : int  7001800 7001801 7001802 7001803 7001804 7001805 7001806 7001807 7001808 7001809 ..
##  $ ccode   : int  700 700 700 700 700 700 700 700 700 700 ...
##  $ scode   : chr  "AFG" "AFG" "AFG" "AFG" ...
##  $ country : chr  "Afghanistan" "Afghanistan" "Afghanistan" "Afghanistan" ...
##  $ year    : int  1800 1801 1802 1803 1804 1805 1806 1807 1808 1809 ...
##  $ flag    : int  0 0 0 0 0 0 0 0 0 0 ...
##  $ fragment: int  NA NA NA NA NA NA NA NA NA NA ...
##  $ democ   : int  1 1 1 1 1 1 1 1 1 1 ...
##  $ autoc   : int  7 7 7 7 7 7 7 7 7 7 ...
##  $ polity  : int  -6 -6 -6 -6 -6 -6 -6 -6 -6 -6 ...
##  $ polity2 : int  -6 -6 -6 -6 -6 -6 -6 -6 -6 -6 ...
##  $ durable : int  NA NA NA NA NA NA NA NA NA NA ...
##  $ xrreg   : int  3 3 3 3 3 3 3 3 3 3 ...
##  $ xrcomp  : int  1 1 1 1 1 1 1 1 1 1 ...
##  $ xropen  : int  1 1 1 1 1 1 1 1 1 1 ...
##  $ xconst  : int  1 1 1 1 1 1 1 1 1 1 ...
##  $ parreg  : int  3 3 3 3 3 3 3 3 3 3 ...
##  $ parcomp : int  3 3 3 3 3 3 3 3 3 3 ...
##  $ exrec   : int  1 1 1 1 1 1 1 1 1 1 ...
##  $ exconst : int  1 1 1 1 1 1 1 1 1 1 ...
##  $ polcomp : int  6 6 6 6 6 6 6 6 6 6 ...
##  $ prior   : int  NA NA NA NA NA NA NA NA NA NA ...
##  $ emonth  : int  NA NA NA NA NA NA NA NA NA NA ...
##  $ eday    : int  NA NA NA NA NA NA NA NA NA NA ...
##  $ eyear   : int  NA NA NA NA NA NA NA NA NA NA ...
##  $ eprec   : int  NA NA NA NA NA NA NA NA NA NA ...
##  $ interim : int  NA NA NA NA NA NA NA NA NA NA ...
##  $ bmonth  : int  1 NA NA NA NA NA NA NA NA NA ...
##  $ bday    : int  1 NA NA NA NA NA NA NA NA NA ...
##  $ byear   : int  1800 NA NA NA NA NA NA NA NA NA ...
##  $ bprec   : int  1 NA NA NA NA NA NA NA NA NA ...
##  $ post    : int  -6 NA NA NA NA NA NA NA NA NA ...
##  $ change  : int  88 NA NA NA NA NA NA NA NA NA ...
##  $ d5      : int  1 NA NA NA NA NA NA NA NA NA ...
##  $ sf      : int  NA NA NA NA NA NA NA NA NA NA ...
##  $ regtrans: int  NA NA NA NA NA NA NA NA NA NA ...
```

```
names(polity)
```

```
##  [1] "p5"       "cyear"    "ccode"    "scode"    "country"  "year"
```

```
##  [7] "flag"     "fragment" "democ"    "autoc"    "polity"   "polity2"
## [13] "durable"  "xrreg"    "xrcomp"   "xropen"   "xconst"   "parreg"
## [19] "parcomp"  "exrec"    "exconst"  "polcomp"  "prior"    "emonth"
## [25] "eday"     "eyear"    "eprec"    "interim"  "bmonth"   "bday"
## [31] "byear"    "bprec"    "post"     "change"   "d5"       "sf"
## [37] "regtrans"
```

Comments: According to the inspection of polity, we have total 17574 observations of 37 variables in the whole data set. In this data set, each row(observation) represts the value of some political indicators of a specific country in a specific year. Some of the cases(data) are displayed above.

Data Wrangling

Comments: Since here we only plan to use Population and Arable Land as controlled variables, Polity 2 Score as independent variable and GDP Growth as dependent variables in the year of 2010, We need do some data wrangling for the final data set.

```
# Pivot the data from long to wide format, ensuring each row corresponds to a single country
worldbankdata_wide <- worldbankdata %>%
  pivot_wider(
    names_from = `Series.Code`,
    values_from = `X2010..YR2010.`,
    id_cols =  `Country.Name`
  )
# Rename the columns as specified to 'gdpgrowth', 'population', 'arableland'
worldbankdata_wide_renamed <- worldbankdata_wide %>%
  rename(
    gdpgrowth = `NY.GDP.MKTP.KD.ZG`,
    population = `SP.POP.1564.TO`,
    arableland = `AG.LND.ARBL.HA.PC`,
    country = `Country.Name`)
# Select the necessary columns
polity_data_selected <- polity %>%
  select(country , polity2,year) %>%
  filter(year == 2010)
# Merge the two dataframes using an inner join on 'scode'
merged_data <- inner_join(polity_data_selected, worldbankdata_wide_renamed, by = "country")
merged_data<- merged_data %>%
  mutate(
    gdpgrowth = na_if(gdpgrowth, ".."),
    population = na_if(population, ".."),
    arableland = na_if(arableland, "..")
  ) %>%
  mutate(
    gdpgrowth = as.numeric(gdpgrowth),
    population = as.numeric(population),
    arableland = as.numeric(arableland)
  )
# Remove rows with any NA values
merged_data <- drop_na(merged_data)
# Create a new 'democracy' variable
merged_data$category <- cut(merged_data$polity2, breaks = c(-Inf, -6, 5, Inf),
```

```
labels = c("Strongly Autocracy", "Anocracy", "Strongly Democracy"),right = FALSE)
# Create three separate datasets
data_autocratic <- filter(merged_data, category == "Strongly Autocracy")
data_anocracy <- filter(merged_data, category == "Anocracy")
data_democratic <- filter(merged_data, category == "Strongly Democracy")
```

Comments: Here we merge the two data sets with wanted variables, and create a new variable called category based on the polity 2 score, -10 to -6 is strongly autocracy, -5 to 5 is anocracy, and -6 to 10 is democracy.

```
head(merged_data)
```

```
##     country polity2 year gdpgrowth population arableland           category
## 1   Albania       9 2010  3.706938    1963578  0.2148972 Strongly Democracy
## 2   Algeria       2 2010  3.600000   24238179  0.2092238           Anocracy
## 3    Angola      -2 2010  4.398376   12069114  0.1797623           Anocracy
## 4 Argentina       8 2010 10.125398   26042007  0.9317122 Strongly Democracy
## 5   Armenia       5 2010  2.200000    2073112  0.1522252 Strongly Democracy
## 6 Australia      10 2010  2.226538   14844429  1.1615850 Strongly Democracy
```

```
str(merged_data)
```

```
## 'data.frame':    137 obs. of  7 variables:
##  $ country   : chr  "Albania" "Algeria" "Angola" "Argentina" ...
##  $ polity2   : int  9 2 -2 8 5 10 10 -7 -5 8 ...
##  $ year      : int  2010 2010 2010 2010 2010 2010 2010 2010 2010 2010 ...
##  $ gdpgrowth : num  3.71 3.6 4.4 10.13 2.2 ...
##  $ population: num  1963578 24238179 12069114 26042007 2073112 ...
##  $ arableland: num  0.215 0.209 0.18 0.932 0.152 ...
##  $ category  : Factor w/ 3 levels "Strongly Autocracy",..: 3 2 2 3 3 3 3 1 2 3 ...
```

```
names(merged_data)
```

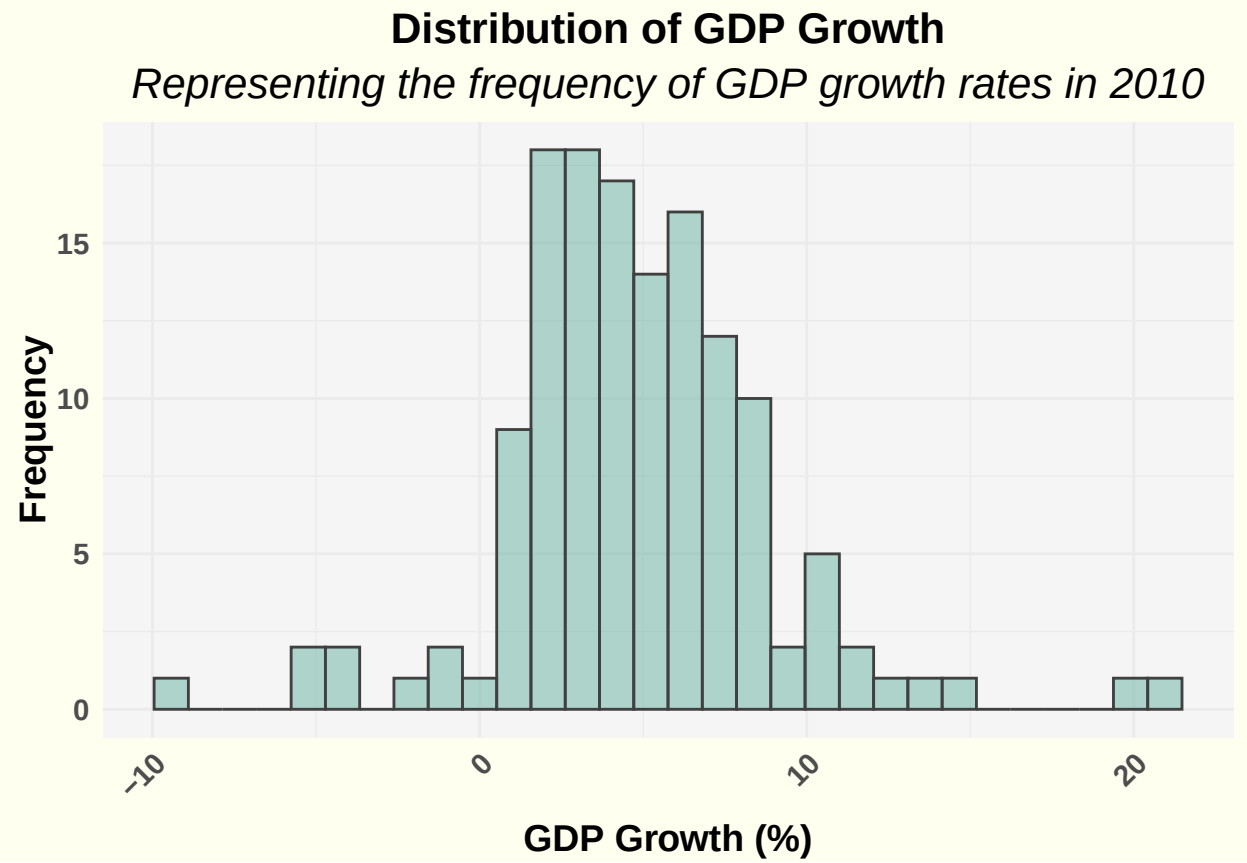
```
## [1] "country"    "polity2"    "year"       "gdpgrowth"  "population"
## [6] "arableland" "category"
```

Comments: According to the inspection of merged data, we have total 137 observations of 7 variables in the whole data set. In this data set, each row(observation) represts the values of independent, dependent and controlled varabales of a specific country in 2010. Some of the cases(data) are displayed above.

Exploratory data analysis

```
#create a histogram to display the distribution of GDP Growth
ggplot(merged_data, aes(x = gdpgrowth)) +
  geom_histogram(position = "identity", alpha = 0.5, bins = 30, fill = "#69b3a2", color = "#404040") +
  labs(title = "Distribution of GDP Growth",
       subtitle = "Representing the frequency of GDP growth rates in 2010",
       x = "GDP Growth (%)",
```

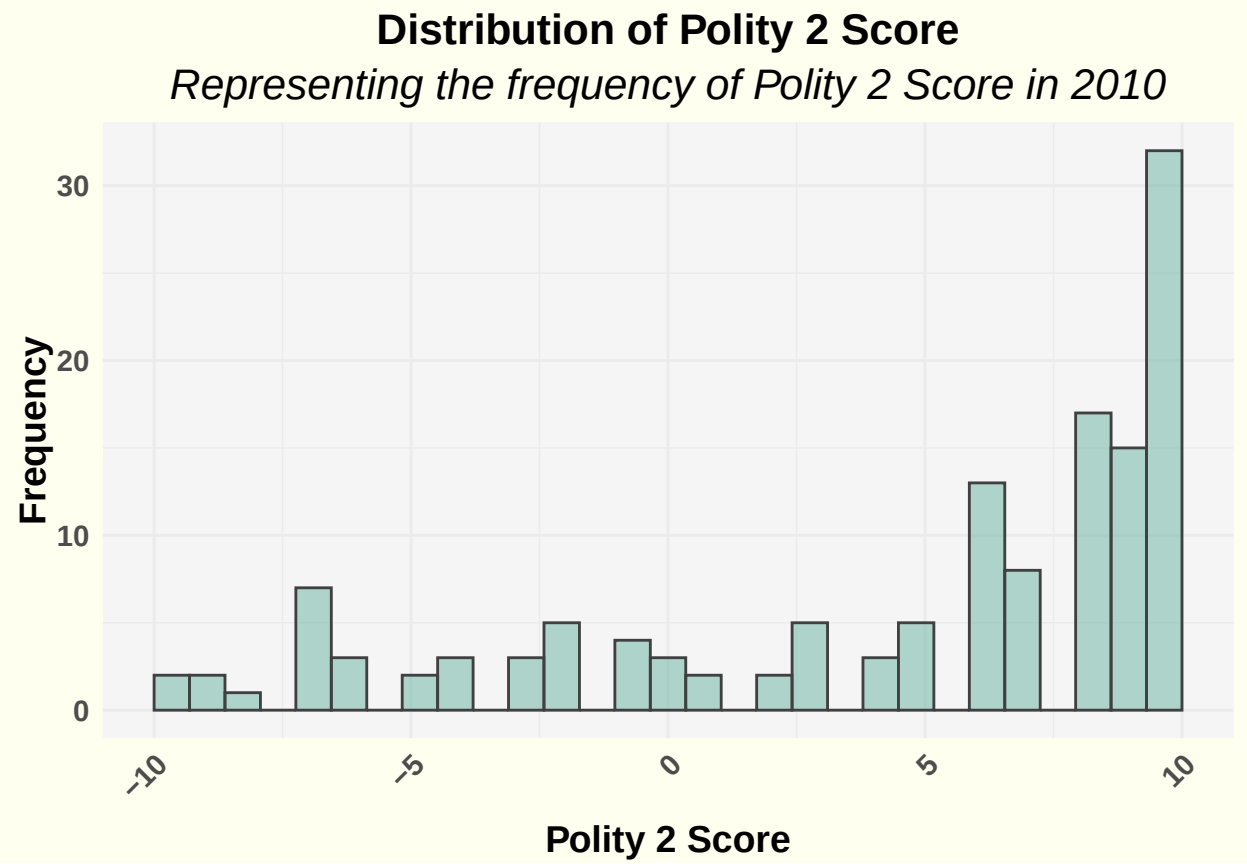
```
      y = "Frequency") +
  theme_minimal() +
  theme(
    text = element_text(size = 14, face = "bold"),
    plot.title = element_text(size = 16, face = "bold", hjust = 0.5),
    plot.subtitle = element_text(size = 16, face = "italic", hjust = 0.5),
    axis.title.x = element_text(vjust = -0.5),
    axis.title.y = element_text(vjust = 0.5),
    axis.text.x = element_text(angle = 45, hjust = 1),
    plot.background = element_rect(fill = "ivory", color = NA),
    panel.background = element_rect(fill = "whitesmoke", color = NA),
    legend.position = "none"
  )
```



Comments: The distribution of GDP growth rates is seen in the histogram. Most of the data is focused in the range of -10% to around 20%, with 0% to 10% being the most frequent occurrences. There is a noticeable peak, indicating that the dataset's average GDP growth rates fall between 2% and 3%. The distribution is right-skewed, which indicates that a few of the nations with the highest growth rates are tailing off to the right. If there are bars to the left of 0%, it means that certain nations' GDP growth has been negative. In general, the graph indicates a typical distribution of growth rates, with moderate growth being the most prevalent, strong growth being less often, and occasional contractions noted.

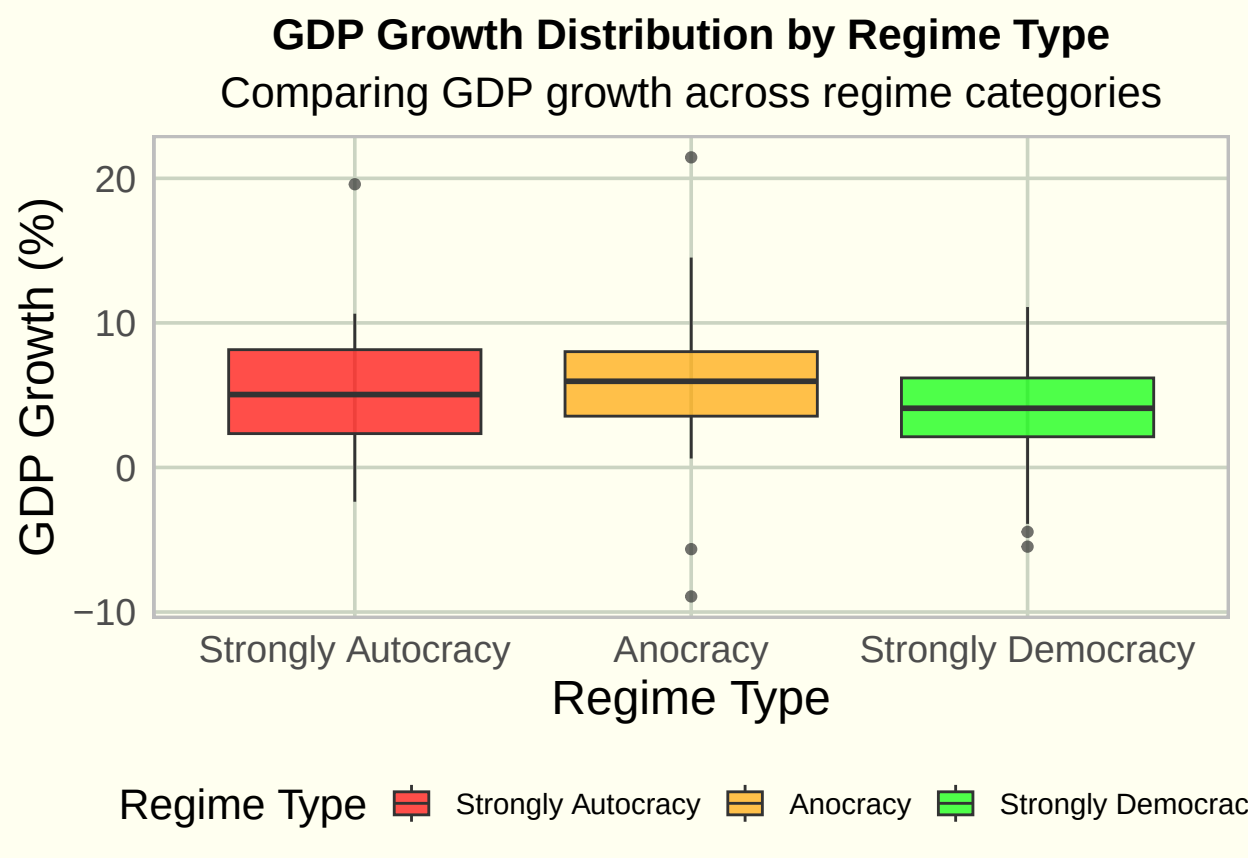
```
#create a histogram to display the distribution of GDP Growth
ggplot(merged_data, aes(x = polity2)) +
  geom_histogram(position = "identity", alpha = 0.5, bins = 30, fill = "#69b3a2", color = "#404040") +
  labs(title = "Distribution of Polity 2 Score",
```

```
      subtitle = "Representing the frequency of Polity 2 Score in 2010",
      x = "Polity 2 Score",
      y = "Frequency") +
theme_minimal() +
theme(
  text = element_text(size = 14, face = "bold"),
  plot.title = element_text(size = 16, face = "bold", hjust = 0.5),
  plot.subtitle = element_text(size = 16, face = "italic", hjust = 0.5),
  axis.title.x = element_text(vjust = -0.5),
  axis.title.y = element_text(vjust = 0.5),
  axis.text.x = element_text(angle = 45, hjust = 1),
  plot.background = element_rect(fill = "ivory", color = NA),
  panel.background = element_rect(fill = "whitesmoke", color = NA),
  legend.position = "none"
)
```



Comments: The frequency of Polity 2 Scores for a group of nations in 2010 is depicted in this histogram. A nation's degree of democracy and autocracy may be gauged using the Polity 2 Score, which has a range of -10 (very autocratic) to +10 (extremely democratic). The distribution is fairly bimodal, with a larger peak at the high end suggesting a concentration of highly democratic countries and one peak in the negative range indicating a cluster of more authoritarian states. The medium range has a heterogeneous distribution, indicating a mix of nations with traits of moderate government. Notably, the +10 score has the highest frequency, suggesting that a sizable number of nations have strong democratic characteristics in this specific year.

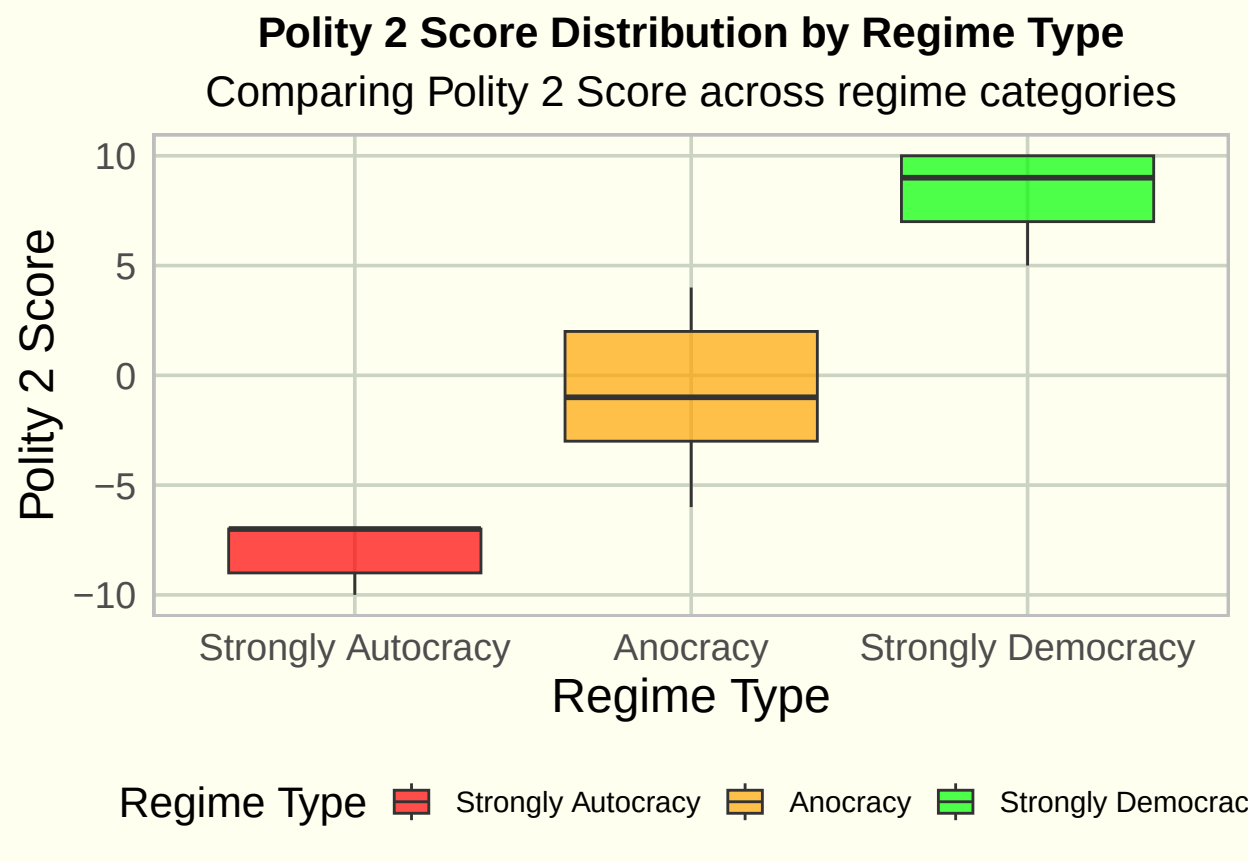
```r
#create a boxplot to display the distribution of GDP Growth by regime type
ggplot(merged_data, aes(x = category, y = gdpgrowth, fill = category)) +
  geom_boxplot(alpha = 0.7) +
  scale_fill_manual(values = c("red", "orange", "green")) +
  labs(title = "GDP Growth Distribution by Regime Type",
       subtitle = "Comparing GDP growth across regime categories",
       x = "Regime Type",
       y = "GDP Growth (%)",
       fill = "Regime Type") +
  theme_minimal(base_size = 14) +
  theme(
    plot.title = element_text(size = 16, face = "bold", hjust = 0.5),
    plot.subtitle = element_text(size = 16, hjust = 0.5),
    axis.title = element_text(size = 18),
    axis.text = element_text(size = 14),
    legend.title = element_text(size = 16),
    legend.text = element_text(size = 11),
    legend.position = "bottom",
    plot.background = element_rect(fill = "ivory", color = NA),
    panel.grid.major = element_line(color = "#CBD4C2"),
    panel.grid.minor = element_blank(),
    panel.border = element_rect(fill = NA, color = "gray", size = 1)
  )
```



Comments: The GDP growth percentages for the three different regime types are compared in the boxplot. Every box displays the interquartile range (IQR) of GDP growth for each regime; a line across the box denotes

the median figure. Strongly Autocracy regimes are represented by a red box with a smaller IQR than the other categories, suggesting less variation in GDP growth within this group. The orange-colored Anocracy group has a broader IQR, indicating higher variability. Strongly democratic regimes are represented by the green box, whose median value is larger than that of the other two groups. This suggests that GDP growth under strongly democratic regimes may be higher. Outliers are displayed in all categories as single points that go beyond the whiskers and reach the lowest and greatest values inside a range that has been computed.

```
#create a boxplot to display the distribution of Polity 2 Score by regime type
ggplot(merged_data, aes(x = category, y = polity2, fill = category)) +
  geom_boxplot(alpha = 0.7) +
  scale_fill_manual(values = c("red", "orange", "green")) +
  labs(title = "Polity 2 Score Distribution by Regime Type",
       subtitle = "Comparing Polity 2 Score across regime categories",
       x = "Regime Type",
       y = "Polity 2 Score",
       fill = "Regime Type") +
  theme_minimal(base_size = 14) +
  theme(
    plot.title = element_text(size = 16, face = "bold", hjust = 0.5),
    plot.subtitle = element_text(size = 16, hjust = 0.5),
    axis.title = element_text(size = 18),
    axis.text = element_text(size = 14),
    legend.title = element_text(size = 16),
    legend.text = element_text(size = 11),
    legend.position = "bottom",
    plot.background = element_rect(fill = "ivory", color = NA),
    panel.grid.major = element_line(color = "#CBD4C2"),
    panel.grid.minor = element_blank(),
    panel.border = element_rect(fill = NA, color = "gray", size = 1)
  )
```

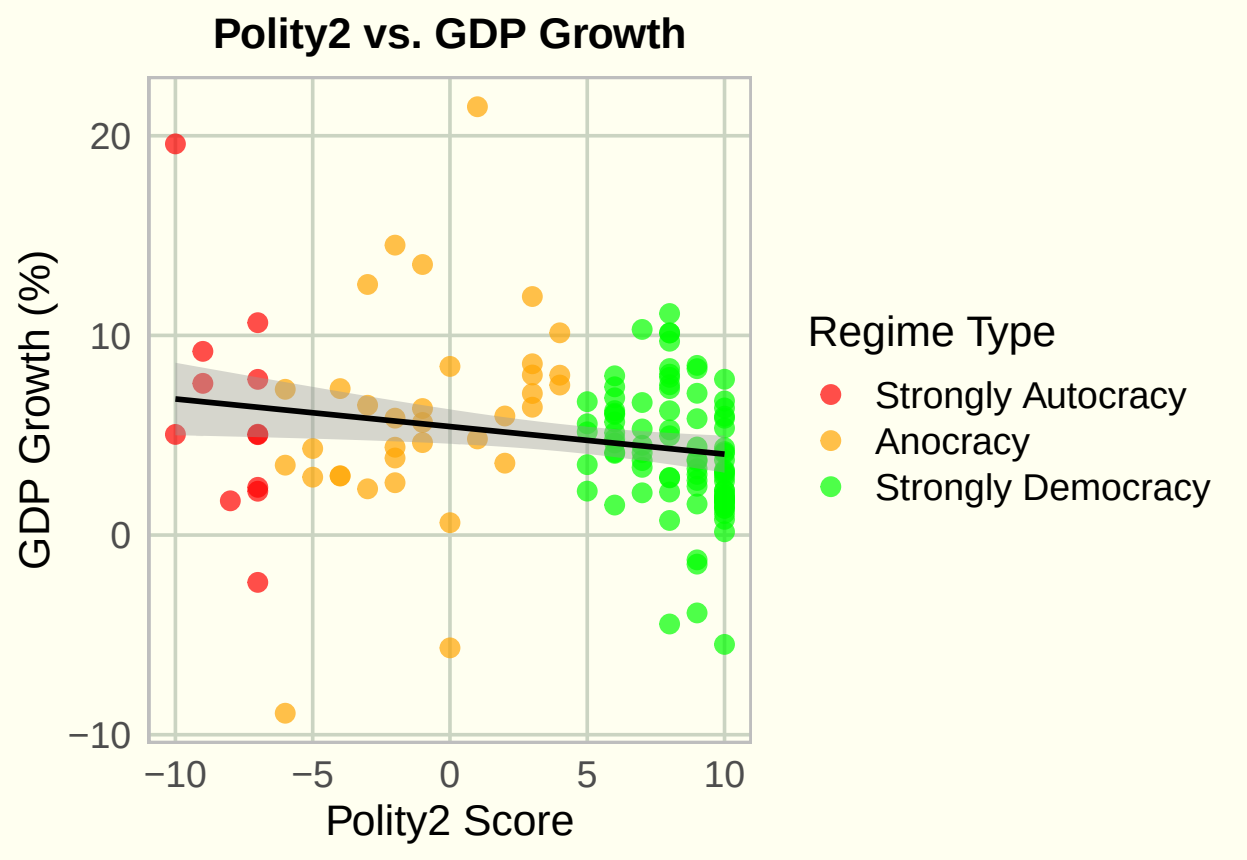


Comments: The distribution of Polity 2 scores among the three distinct regime groups is shown in the boxplot.A complete autocrat receives a score of -10, whereas a genuine democracy receives a score of +10. Strongly autocratic regimes are represented by the red box on the scale, which has limited variability and scores that are closely packed at the bottom. With a broader interquartile range spanning from negative to positive scores, the Anocracy group, shown in orange, has higher variability and a mixture of traits. With scores at the upper end of the spectrum, the green box for strongly democratic regimes indicates a concentration of democratic traits with less variability than an anocracy. Each regime type's appropriate color is used in the plot to visually assist in differentiating between the categories.

Data Visualization

```
# Create a scatter plot with 'category' used to color the points by regime type
ggplot(merged_data, aes(x = polity2, y = gdpgrowth, color = category)) +
  geom_point(alpha = 0.7, size = 3) +
  geom_smooth(method = "lm", se = TRUE, color = "black") +
  scale_color_manual(values = c("Strongly Autocracy" = "red",
                                "Anocracy" = "orange",
                                "Strongly Democracy" = "green")) +
  labs(title = "Polity2 vs. GDP Growth",
       x = "Polity2 Score",
       y = "GDP Growth (%)",
       color = "Regime Type") +
  theme_minimal(base_size = 14) +
  theme(
```

```
    plot.title = element_text(size = 16, face = "bold", hjust = 0.5),
    legend.title = element_text(size = 16),
    legend.text = element_text(size = 14),
    legend.position = "right",
    axis.title = element_text(size = 16),
    axis.text = element_text(size = 14),
    plot.background = element_rect(fill = "ivory", color = NA),
    panel.grid.major = element_line(color = "#CBD4C2"),
    panel.grid.minor = element_blank(),
    panel.border = element_rect(color = "gray", size = 1, fill = NA)
  )
```

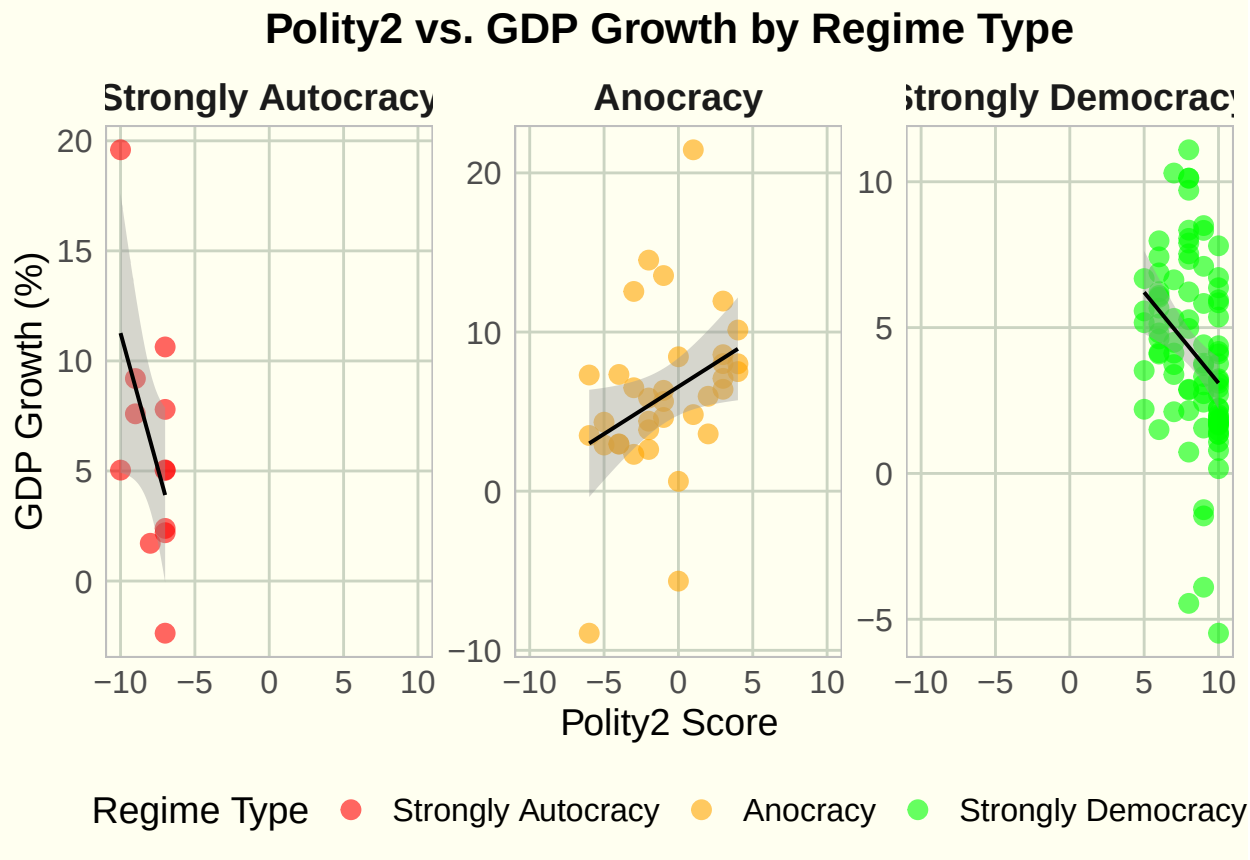


Comments: From the plot, we can see a concentration of autocratic regimes (red points) with lower Polity2 scores and a mix of GDP growth outcomes. Anocratic regimes (orange points) are spread across the middle of the Polity2 scale, also with varied GDP growth. Democracies (green points) are clustered towards higher Polity2 scores, with several achieving high GDP growth rates. A linear regression line with a gray confidence interval band suggests a general trend where higher Polity2 scores (more democratic regimes) might be associated with lower GDP growth rates, but the spread of points indicates variability and the presence of outliers.

```
# Create a scatter plot with faceted regime types
ggplot(merged_data, aes(x = polity2, y = gdpgrowth)) +
  geom_point(aes(color = category), alpha = 0.6, size = 3) +
  geom_smooth(method = "lm", se = TRUE, color = "black",size = 0.7) +
  scale_color_manual(values = c("Strongly Autocracy" = "red",
                                "Anocracy" = "orange",
```

```
                                "Strongly Democracy" = "green")) +
  labs(title = "Polity2 vs. GDP Growth by Regime Type",
       x = "Polity2 Score",
       y = "GDP Growth (%)") +
  facet_wrap(~ category, scales = "free_y") +
  theme_minimal() +
  theme(
    plot.title = element_text(size = 16, face = "bold", hjust = 0.5),
    strip.text = element_text(size = 14, face = "bold"),
    legend.position = "bottom",
    legend.title = element_text(size = 14),
    legend.text = element_text(size = 12),
    axis.title = element_text(size = 14),
    axis.text = element_text(size = 12),
    plot.background = element_rect(fill = "ivory", color = NA),
    panel.grid.major = element_line(color = "#CBD4C2"),
    panel.grid.minor = element_blank(),
    panel.border = element_rect(color = "gray", fill = NA, size = 0.5)
  ) +
  guides(color = guide_legend(title = "Regime Type"))
```

```
## Warning: Using ‘size‘ aesthetic for lines was deprecated in ggplot2 3.4.0.
## i Please use ‘linewidth‘ instead.
## This warning is displayed once every 8 hours.
## Call ‘lifecycle::last_lifecycle_warnings()‘ to see where this warning was
## generated.
```



Comments: In the "Strongly Autocracy" regime, the relationship between Polity2 scores and GDP growth appears negative trend, indicating that within "Strongly Autocracy" regime, higher polity 2 scores will result in lower GDP growth rates. In the "Anocracy" regime, the relationship between Polity2 scores and GDP growth appears positive trend, indicating that within "Anocracy" regime, higher polity 2 scores will result in higher GDP growth rates. In the "Strongly Autocracy" regime, the relationship between Polity2 scores and GDP growth appears positive trend, indicating that within "Strongly Autocracy" regime, higher polity 2 scores will result in lower GDP growth rates. This illustrate different association trend of dependent and independent variables among different regime types.

Model Fitting

Comments: To better understand the relationship between polity score and GDP Growth rate, we use population and arable land as controlled variables.

```
#model fitting
overall<-lm(gdpgrowth~polity2+arableland+population, data=merged_data)
autocracy<-lm(gdpgrowth~polity2+arableland+population, data = data_autocratic)
anocracy<-lm(gdpgrowth~polity2+arableland+population, data = data_anocracy)
democracy<-lm(gdpgrowth~polity2+arableland+population, data = data_democratic)
#regression table
stargazer(overall, autocracy, anocracy, democracy, type= "text",
title= "Regression Results",
dep.var.labels="GDP Growth(%)")
```

```
##
```

```
## Regression Results
## ==========================================================================================================
##                                                Dependent variable:
##                     ---------------------------------------------------------------------------------
##                                                    GDP Growth(%)
##                              (1)                  (2)                 (3)                  (4)
## ----------------------------------------------------------------------------------------------------------
## polity2                   -0.134**             -3.408**            0.620**              -0.639***
##                            (0.058)              (1.186)            (0.297)               (0.194)
##
## arableland                 0.670                10.900              2.882                0.142
##                           (1.384)              (8.516)            (2.836)               (1.372)
##
## population                0.000*               0.000*              0.000                0.000*
##                           (0.000)              (0.000)           (0.00000)              (0.000)
##
## Constant                  5.099***             -23.773*            5.731***             9.261***
##                           (0.560)              (10.344)           (1.354)               (1.629)
##
## ----------------------------------------------------------------------------------------------------------
## Observations                137                   12                  35                   90
## R2                         0.060                0.541               0.153                0.135
## Adjusted R2                0.039                0.369               0.071                0.105
## Residual Std. Error   4.016 (df = 133)     4.421 (df = 8)     5.115 (df = 31)      2.954 (df = 86)
## F Statistic         2.847** (df = 3; 133) 3.144* (df = 3; 8) 1.865 (df = 3; 31) 4.470*** (df = 3; 86)
## ==========================================================================================================
## Note:                                                                          *p<0.1; **p<0.05; ***p<0.0
```

Comments: Regarding the controlled variables, in the four models, the results show that non of these are statistically significant, so we can ignore these controlled variables. The polity 2 score coefficients(overall: -0.134,autocracy: -3.408,anocracy: 0.620,democracy:-0.639) are all statistically significant and are consistent with the scatter plot.

Conclusion

To sum up, for "Strongly Autocracy" regimes, a negative trend is observed, hinting that more democratic features could correlate with lower GDP growth within this category. Conversely, "Anocracy" regimes exhibit a positive trend, suggesting that as these regimes become less autocratic (higher Polity2 scores), they might experience higher GDP growth rates. And in "Strongly Democracy" regimes, suggests a negative correlation.

Regardless which regime type, the trend is negative hinting that higher polity score could correlate with lower GDP growth.