David Atunwa IDS Assignment

```
library(tidyverse)
```

```
## -- Attaching core tidyverse packages ------------------------ tidyverse 2.0.0 --
## v dplyr     1.1.4     v readr     2.1.5
## v forcats   1.0.1     v stringr   1.6.0
## v ggplot2   4.0.0     v tibble    3.3.0
## v lubridate 1.9.4     v tidyr     1.3.1
## v purrr     1.2.0
## -- Conflicts ------------------------------------------ tidyverse_conflicts() --
## x dplyr::filter() masks stats::filter()
## x dplyr::lag()    masks stats::lag()
## i Use the conflicted package (<http://conflicted.r-lib.org/>) to force all conflicts to become errors
```

```
library(dplyr)
library(purrr)
library(broom)
library(reshape2)
```

```
##
## Attaching package: 'reshape2'
##
## The following object is masked from 'package:tidyr':
##
##     smiths
```

```
library(patchwork)
library(zoo)
```

```
##
## Attaching package: 'zoo'
##
## The following objects are masked from 'package:base':
##
##     as.Date, as.Date.numeric
```

```
df_class <- read_csv("Classification_by_Income.csv")
```

```
## Rows: 269 Columns: 4
## -- Column specification --------------------------------------------------------
## Delimiter: ","
## chr (4): Economy, Code, Region, Income group
##
## i Use `spec()` to retrieve the full column specification for this data.
## i Specify the column types or set `show_col_types = FALSE` to quiet this message.
```

```r
df_gdp <- read_csv("GDP_Per_Capita.csv")
```

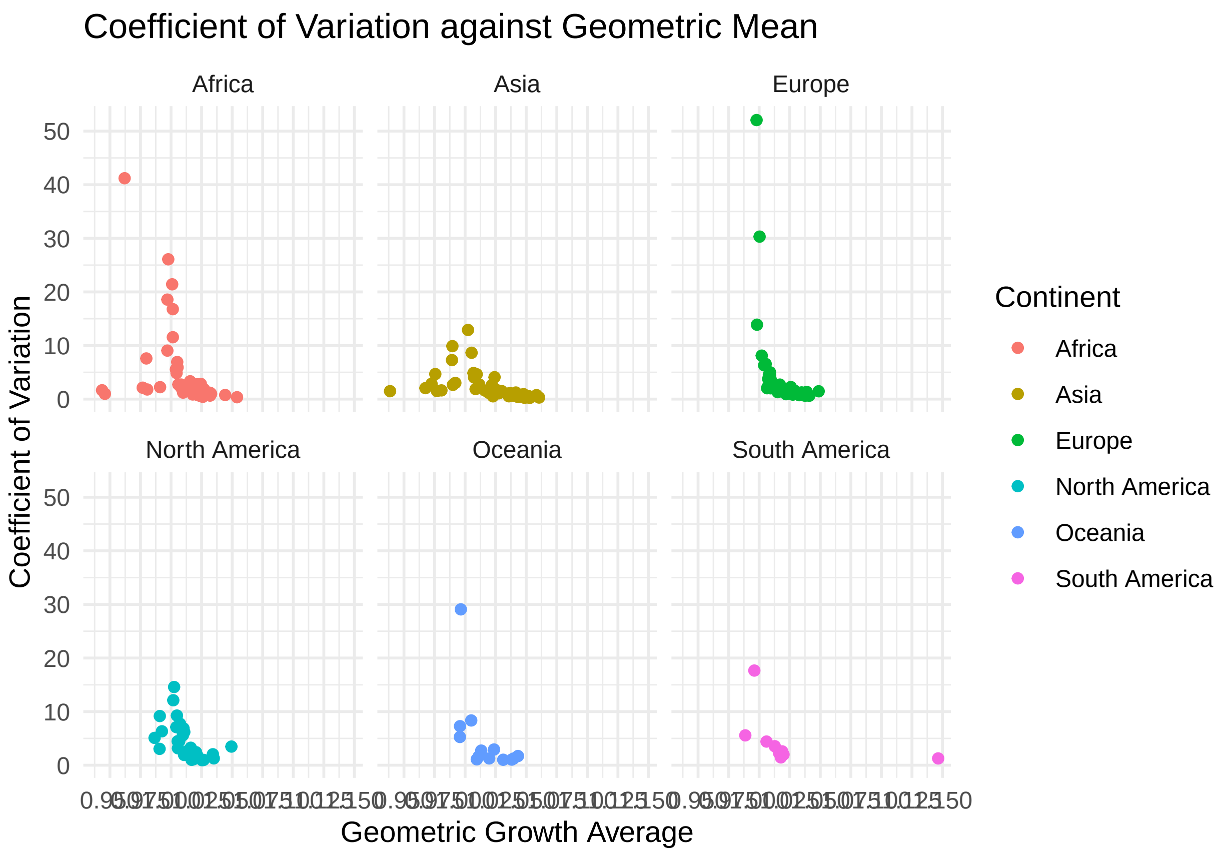
```
## New names:
## Rows: 268 Columns: 68
## -- Column specification
## -------------------------------------------------------- Delimiter: "," chr
## (3): Data Source, World Development Indicators, ...3 dbl (65): ...4, ...5,
## ...6, ...7, ...8, ...9, ...10, ...11, ...12, ...13, .....
## i Use `spec()` to retrieve the full column specification for this data. i
## Specify the column types or set `show_col_types = FALSE` to quiet this message.
## * `` -> `...3`
## * `` -> `...4`
## * `` -> `...5`
## * `` -> `...6`
## * `` -> `...7`
## * `` -> `...8`
## * `` -> `...9`
## * `` -> `...10`
## * `` -> `...11`
## * `` -> `...12`
## * `` -> `...13`
## * `` -> `...14`
## * `` -> `...15`
## * `` -> `...16`
## * `` -> `...17`
## * `` -> `...18`
## * `` -> `...19`
## * `` -> `...20`
## * `` -> `...21`
## * `` -> `...22`
## * `` -> `...23`
## * `` -> `...24`
## * `` -> `...25`
## * `` -> `...26`
## * `` -> `...27`
## * `` -> `...28`
## * `` -> `...29`
## * `` -> `...30`
## * `` -> `...31`
## * `` -> `...32`
## * `` -> `...33`
## * `` -> `...34`
## * `` -> `...35`
## * `` -> `...36`
## * `` -> `...37`
## * `` -> `...38`
## * `` -> `...39`
## * `` -> `...40`
## * `` -> `...41`
## * `` -> `...42`
## * `` -> `...43`
## * `` -> `...44`
## * `` -> `...45`
## * `` -> `...46`
## * `` -> `...47`
## * `` -> `...48`
## * `` -> `...49`
## * `` -> `...50`
## * `` -> `...51`
## * `` -> `...52`
## * `` -> `...53`
## * `` -> `...54`
## * `` -> `...55`
## * `` -> `...56`
## * `` -> `...57`
## * `` -> `...58`
## * `` -> `...59`
## * `` -> `...60`
## * `` -> `...61`
## * `` -> `...62`
## * `` -> `...63`
## * `` -> `...64`
## * `` -> `...65`
## * `` -> `...66`
## * `` -> `...67`
## * `` -> `...68`
```

```r
df_cont <- read_csv("continents-according-to-our-world-in-data.csv")
```

```
## Rows: 285 Columns: 4
## -- Column specification --------------------------------------------------------
## Delimiter: ","
## chr (3): Entity, Code, Continent
## dbl (1): Year
##
## i Use `spec()` to retrieve the full column specification for this data.
## i Specify the column types or set `show_col_types = FALSE` to quiet this message.
```

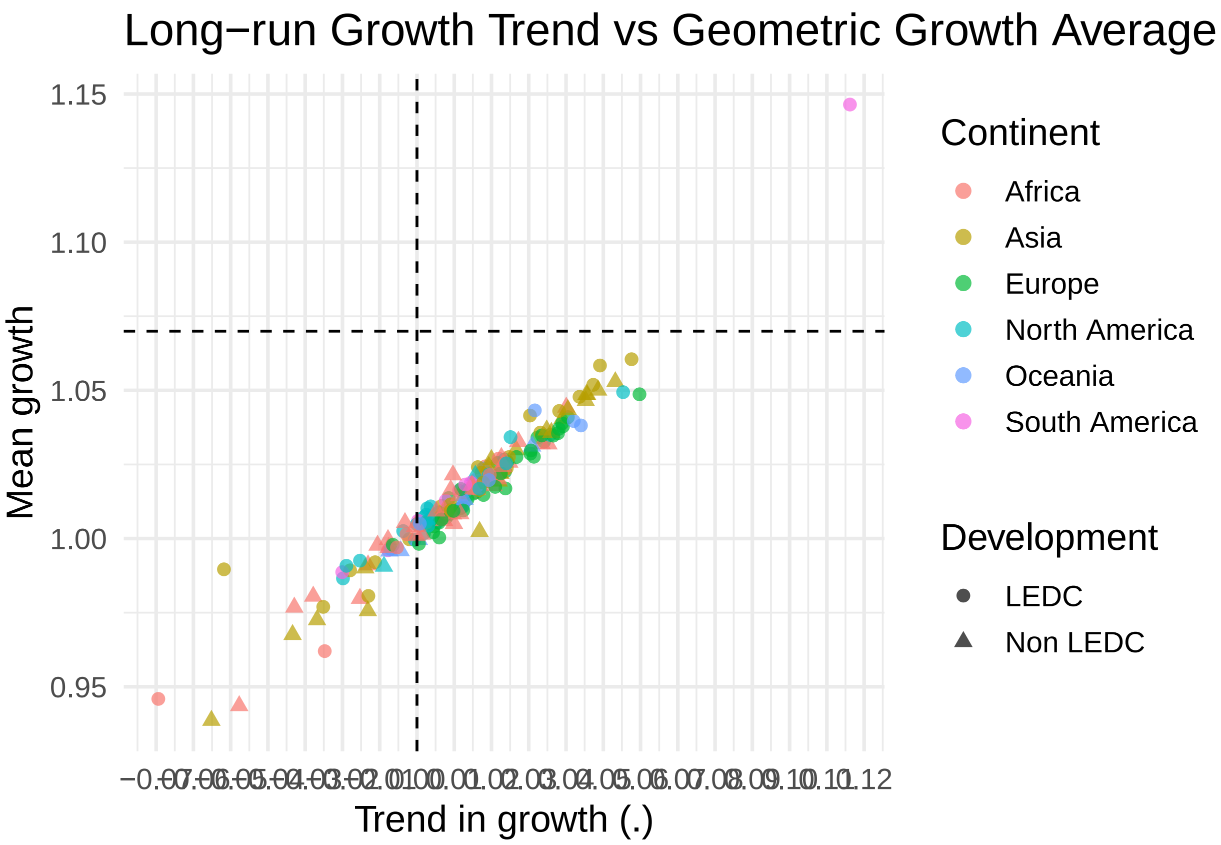
```r
#Cleaning GDP Per Capita File

new_columns <-  2010:2024
old_columns <- paste0("...", 54:68)
df_gdp <- df_gdp %>% rename_with(~as.character(new_columns), .cols = all_of(old_columns))
df_gdp <- df_gdp %>% select(c("Data Source", "World Development Indicators", as.character(2010:2024))) %
  filter(`World Development Indicators` %in% df_cont$Code)
df_gdp <- df_gdp[rowSums(is.na(df_gdp)) < 2,] %>% rename("Code" = "World Development Indicators")

df_gdp <- df_gdp %>% gather(, key="Year", value="GDP_Per_Capita", as.character(2010:2024)) %>%
  arrange(Code)
df_gdp <- df_gdp %>% group_by(Code) %>%
  mutate(Annual_Growth_Rate_PC = ((`GDP_Per_Capita` - lag(`GDP_Per_Capita`)) / lag(`GDP_Per_Capita`)) *
         Growth_Factor = ((`GDP_Per_Capita` / lag(`GDP_Per_Capita`)))) %>%
  ungroup()
df_gdp <- left_join(df_gdp, df_cont[,c(2,4)], by = "Code")
df_gdp <- left_join(df_gdp, df_class[, c(2,4)], by = "Code")
df_gdp$Year <- as.numeric(df_gdp$Year)
```

```r
#Regressing log per-capita GDP to estimate long-run growth trajectories

trend_results <- df_gdp %>%
  group_by(Code) %>%
  nest() %>%
  mutate(model = map(data, ~ lm(log(GDP_Per_Capita) ~ Year, data = .x)),
         results = map(model, tidy)) %>%
  unnest(results) %>%
  filter(term == "Year") %>%
  select(Code, beta = estimate, p_value = p.value)


#Computing metrics of volatility

df_main <- df_gdp %>%
  select(c(2,5,6)) %>%
  group_by(Code) %>%
  summarise(`Geometric Growth Average` = exp(mean(log(Growth_Factor), na.rm = TRUE)),
            `Standard Deviation` = sd(Annual_Growth_Rate_PC, na.rm = TRUE),
            `Coefficient of Variation` = sd(Annual_Growth_Rate_PC, na.rm = TRUE) / abs(mean(Annual_Grow
df_main <- left_join(df_main, df_cont[,-3], by = "Code")
df_main <- left_join(df_main, df_class[, c(2,4)], by = "Code")
df_main <- left_join(df_main, trend_results, by = "Code")


# (ONE) Coefficeint of Variation Against Growth
ggplot(data = df_main[df_main$`Coefficient of Variation` < 100, ]) +
  geom_point(mapping = aes(x = `Geometric Growth Average`,
                          y = `Coefficient of Variation`,
                          color = Continent)) +
  scale_x_continuous(n.breaks = 10) +
  facet_wrap(~ Continent) +
  labs(title = "Coefficient of Variation against Geometric Mean") +
  theme_minimal()
```



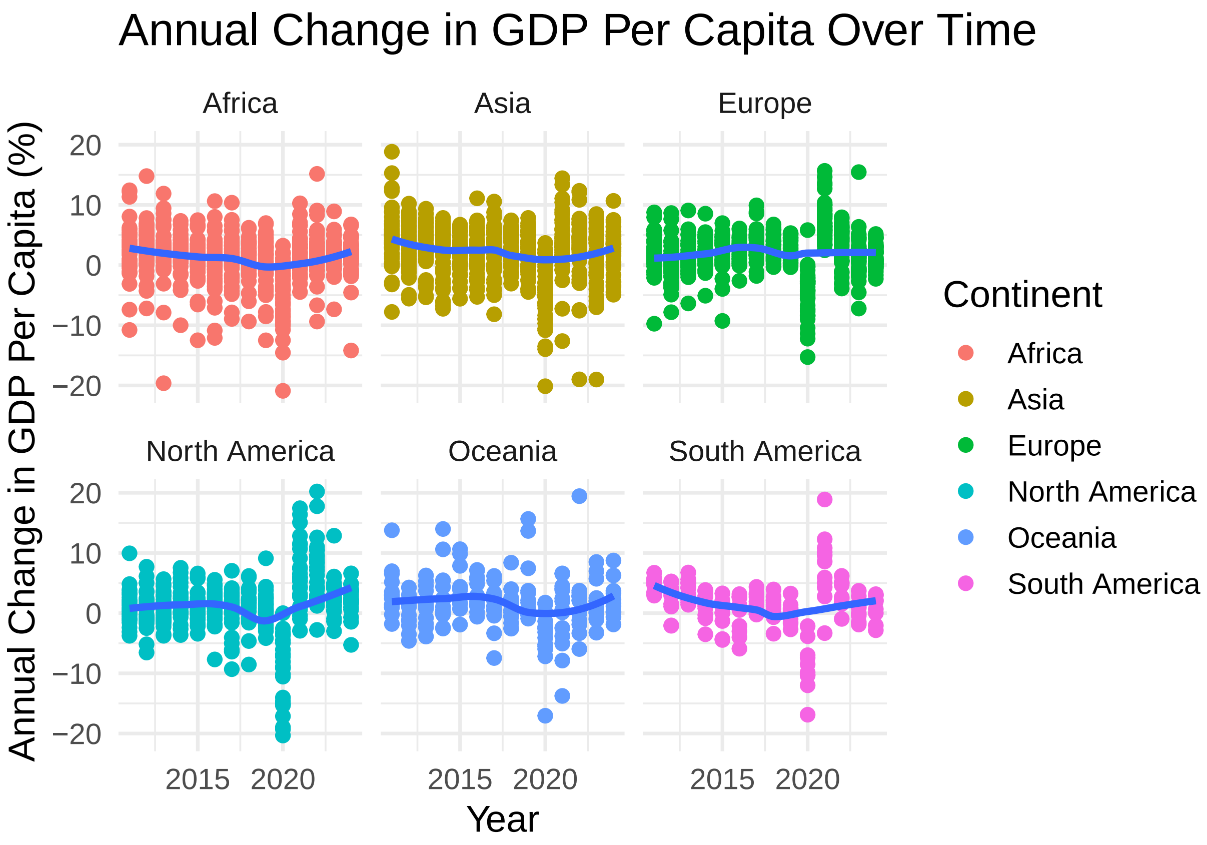
```
# (TWO) Long-run Growth Trend against Geometric Growth Rate

ggplot(df_main[df_main$beta < 2.0,], aes(x = beta, y = `Geometric Growth Average`,
                                         colour = Continent,
                                         shape = as.vector(
                                           df_main[df_main$beta < 2.0,]$`Income group`== "Low income" |
  geom_point(size = 2, alpha = 0.7, show.legend = TRUE) +
  geom_vline(xintercept = 0, linetype = "dashed", linewidth=0.5) +
  geom_hline(yintercept = 1.07, linetype = 2, linewidth=0.5) +
  labs(
    x = "Trend in growth ( )",
    y = "Mean growth",
    title = "Long-run Growth Trend vs Geometric Growth Average",
    shape = "Development") +
  scale_x_continuous(n.breaks = 15) +
  scale_shape_manual(values = c("TRUE" = 17,
                                "FALSE" = 16),
                     labels = c("LEDC", "Non LEDC", "")) +
  theme_minimal(base_size = 14)
```

```
## Warning: Removed 1 row containing missing values or values outside the scale range
## (`geom_point()`).
```

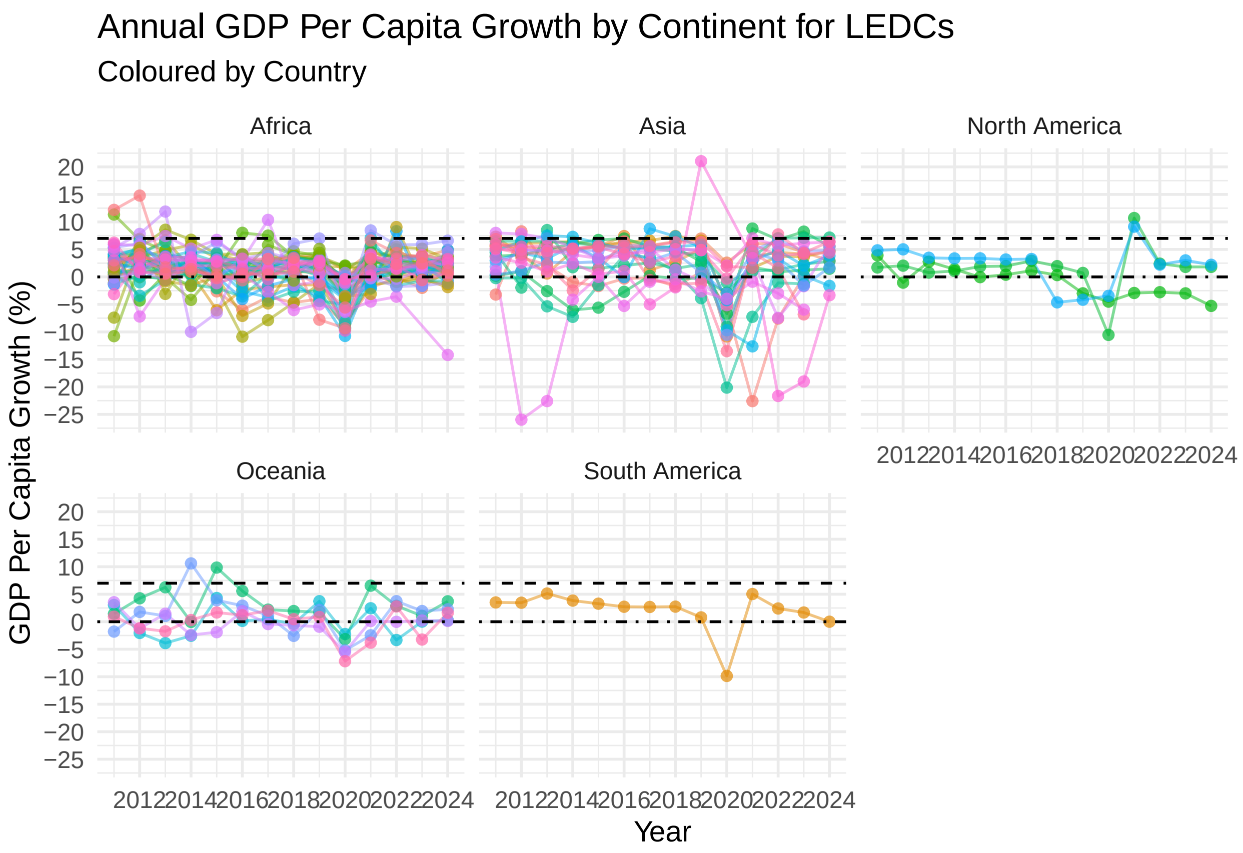


```
# (THREE) Annual Change in GDP Per Capita Over Time

ggplot(data = filter(df_gdp, abs(Annual_Growth_Rate_PC) < 21)) +
  geom_point(mapping = aes(x = Year, y = `Annual_Growth_Rate_PC`, color = Continent)) +
  facet_wrap( ~ Continent) +
  geom_smooth(mapping = aes(x = Year, y = Annual_Growth_Rate_PC), se = FALSE) +
  labs(title = "Annual Change in GDP Per Capita Over Time",
       y= "Annual Change in GDP Per Capita (%)") +
  theme_minimal(base_size = 14)
```

```
## `geom_smooth()` using method = 'loess' and formula = 'y ~ x'
```



```
# (FOUR) Annual GDP per Capita Growth by Continent for LEDCs

ggplot(data = filter(df_gdp, `Income group` == "Lower middle income" | `Income group` == "Low income", a
       aes(x = Year, y = Annual_Growth_Rate_PC, colour = `Data Source`)) +
  geom_line(alpha = 0.5, show.legend = FALSE) +      # line for each country
  geom_point(alpha = 0.7, size = 1.5, show.legend = FALSE) +  # points for each year
  geom_hline(yintercept = 7, linetype = "dashed", colour = "black", linewidth = 0.5) +
  geom_hline(yintercept = 0, linetype = "dotdash", colour = "black", linewidth = 0.5) +
  labs(title = "Annual GDP Per Capita Growth by Continent for LEDCs",
       subtitle = "Coloured by Country",
       y = "GDP Per Capita Growth (%)",
       x = "Year") +
  facet_wrap(~ Continent) +
  scale_y_continuous(n.breaks = 10) +
  scale_x_continuous(n.breaks = 10) +
  theme_minimal()
```



```r
# Computing Average Growth Rates for All Countries in Each Continent and All LEDCs
continent_avg <- df_gdp %>%
  group_by(Continent, Year) %>%
  summarise(
    Geometric_Growth_Average = exp(mean(log(Growth_Factor), na.rm = TRUE)),
    n_countries = n()
  ) %>%
  ungroup() %>% mutate(Geometric_Growth_Rate = (Geometric_Growth_Average -1) * 100)
```

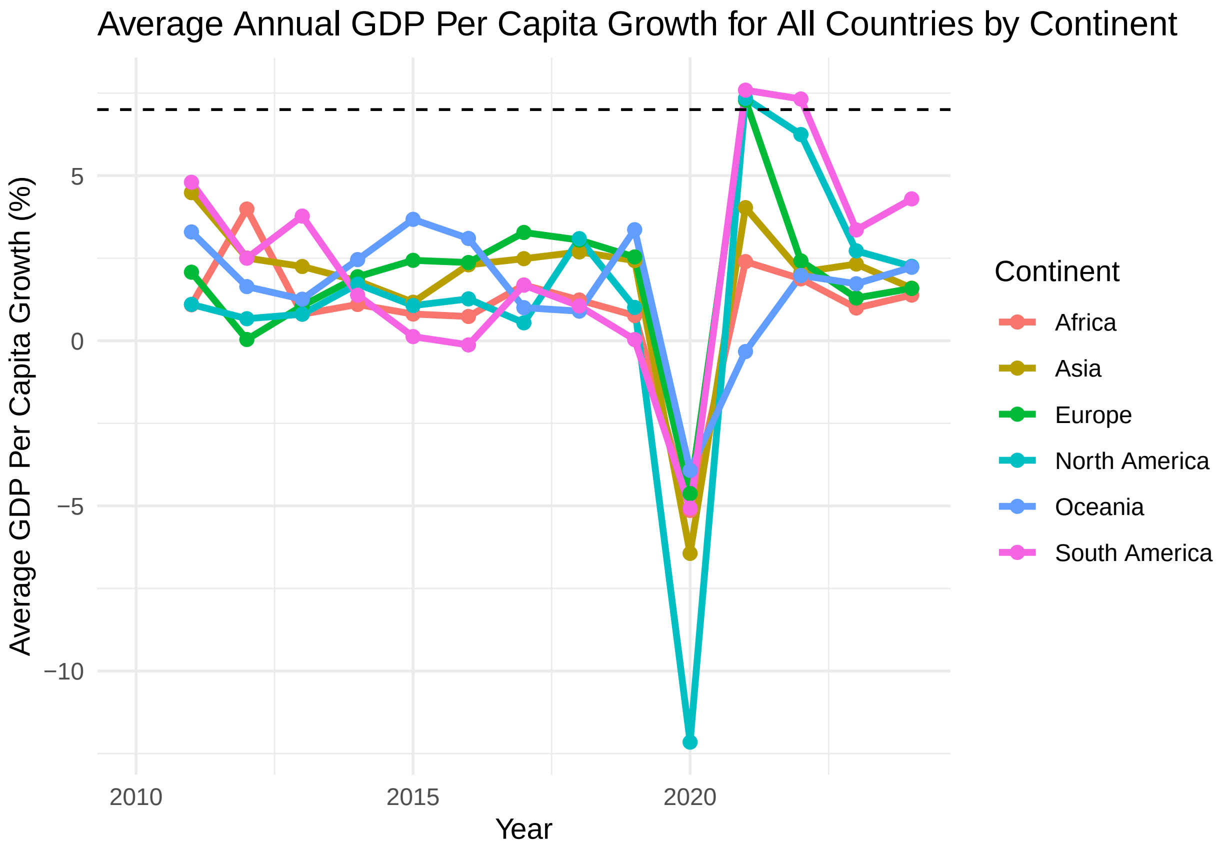
```
## `summarise()` has grouped output by 'Continent'. You can override using the
## `.groups` argument.
```

```r
continent_avg_for_ledcs <- df_gdp[df_gdp$`Income group` == "Low income" | df_gdp$`Income group` == "Low
  group_by(Continent, Year) %>%
  summarise(
    Geometric_Growth_Average = exp(mean(log(Growth_Factor), na.rm = TRUE)),
    n_countries = n()
  ) %>%
  ungroup() %>% mutate(Geometric_Growth_Rate = (Geometric_Growth_Average -1) * 100)
```

```
## `summarise()` has grouped output by 'Continent'. You can override using the
## `.groups` argument.
```

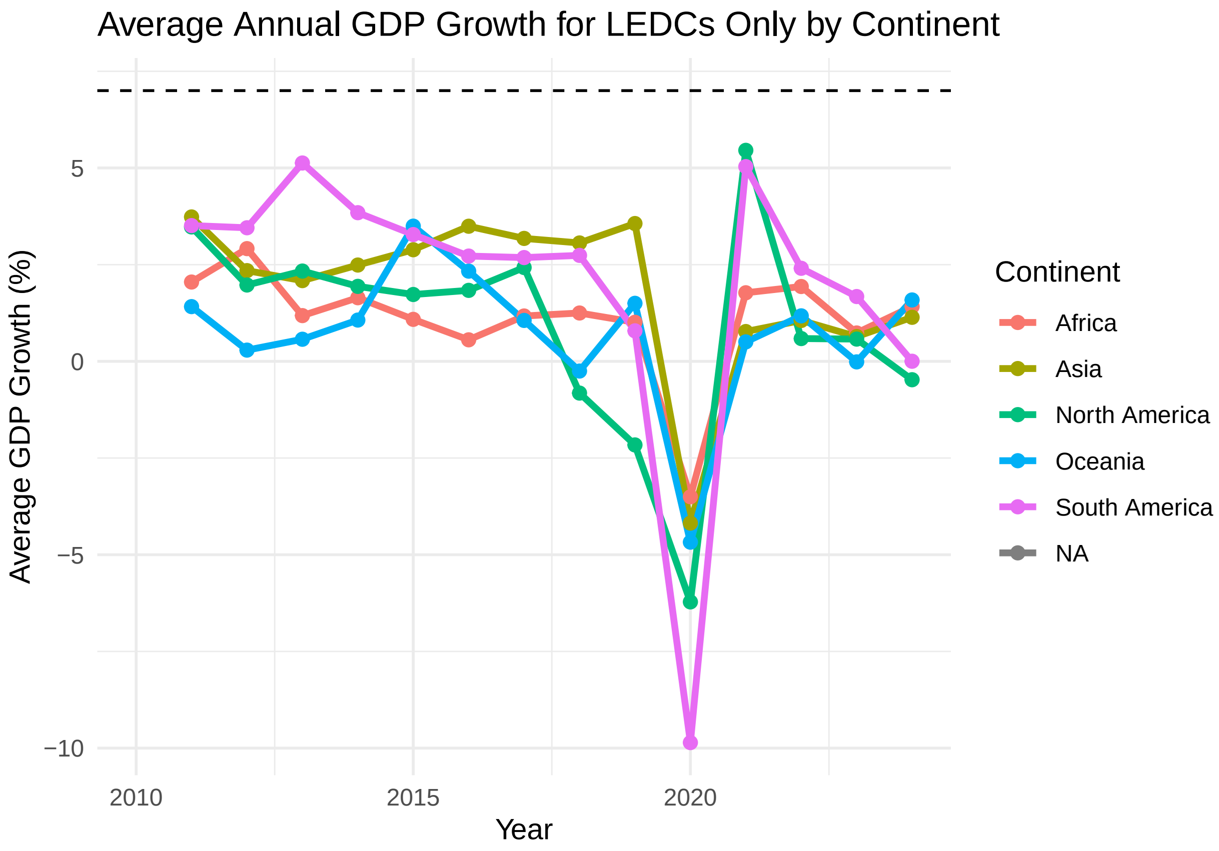
```r
# (FIVE) Average Annual GDP Per Capita Growth For All Countries by Continent

ggplot(continent_avg, aes(x = Year, y = Geometric_Growth_Rate, colour = Continent)) +
  geom_line(linewidth = 1.2) +
  geom_point(size = 2) +
  geom_hline(yintercept = 7, linetype = "dashed", colour = "black") +
  labs(
    title = "Average Annual GDP Per Capita Growth for All Countries by Continent",
    y = "Average GDP Per Capita Growth (%)",
    x = "Year"
  ) +
  theme_minimal()
```

```
## Warning: Removed 6 rows containing missing values or values outside the scale range
## (`geom_line()`).

## Warning: Removed 6 rows containing missing values or values outside the scale range
## (`geom_point()`).
```



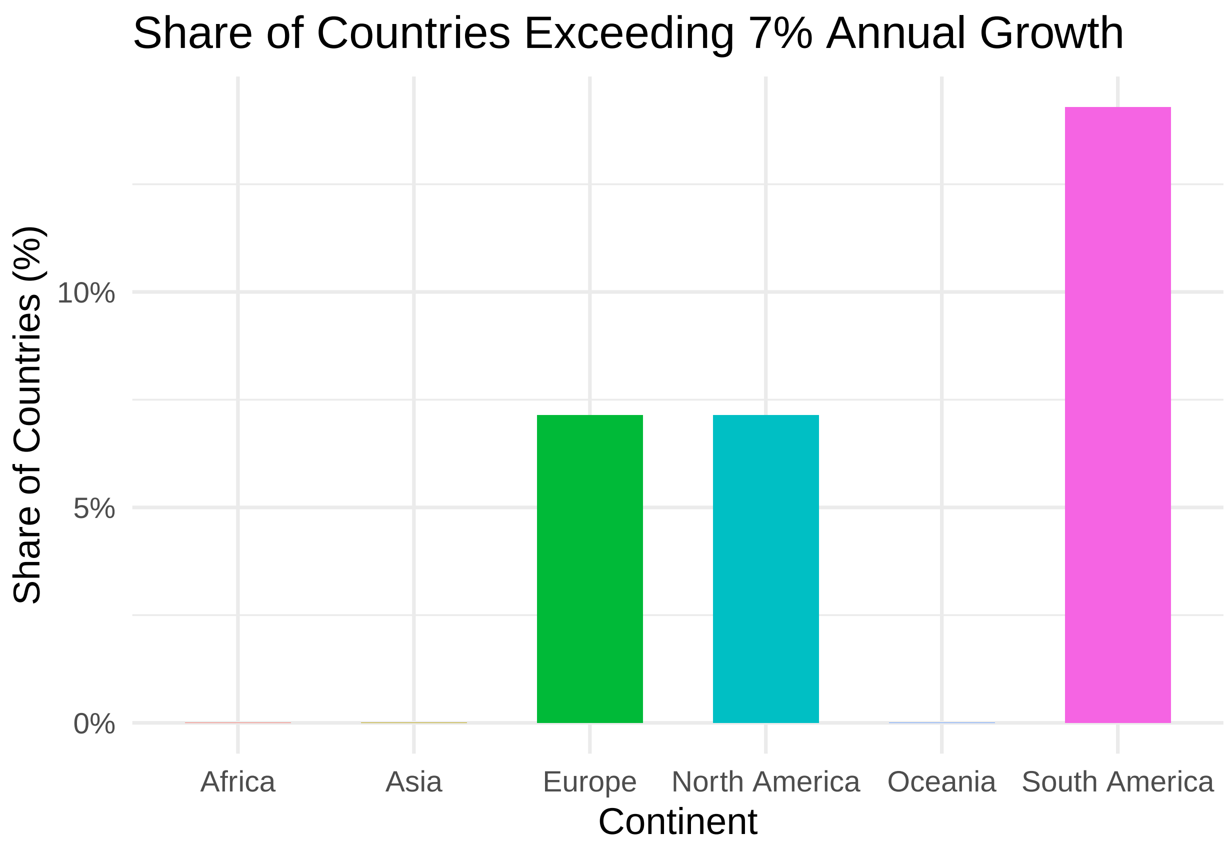
```r
# (SIX) AVergae Annual GDP Per Capita Growth For all Countries by Continent

ggplot(continent_avg_for_ledcs, aes(x = Year, y = Geometric_Growth_Rate, colour = Continent)) +
  geom_line(linewidth = 1.2) +
  geom_point(size = 2) +
  geom_hline(yintercept = 7, linetype = "dashed", colour = "black") +
  labs(
    title = "Average Annual GDP Growth for LEDCs Only by Continent",
    y = "Average GDP Growth (%)",
    x = "Year"
  ) +
  theme_minimal()
```

```
## Warning: Removed 6 rows containing missing values or values outside the scale range
## (`geom_line()`).
## Removed 6 rows containing missing values or values outside the scale range
## (`geom_point()`).
```

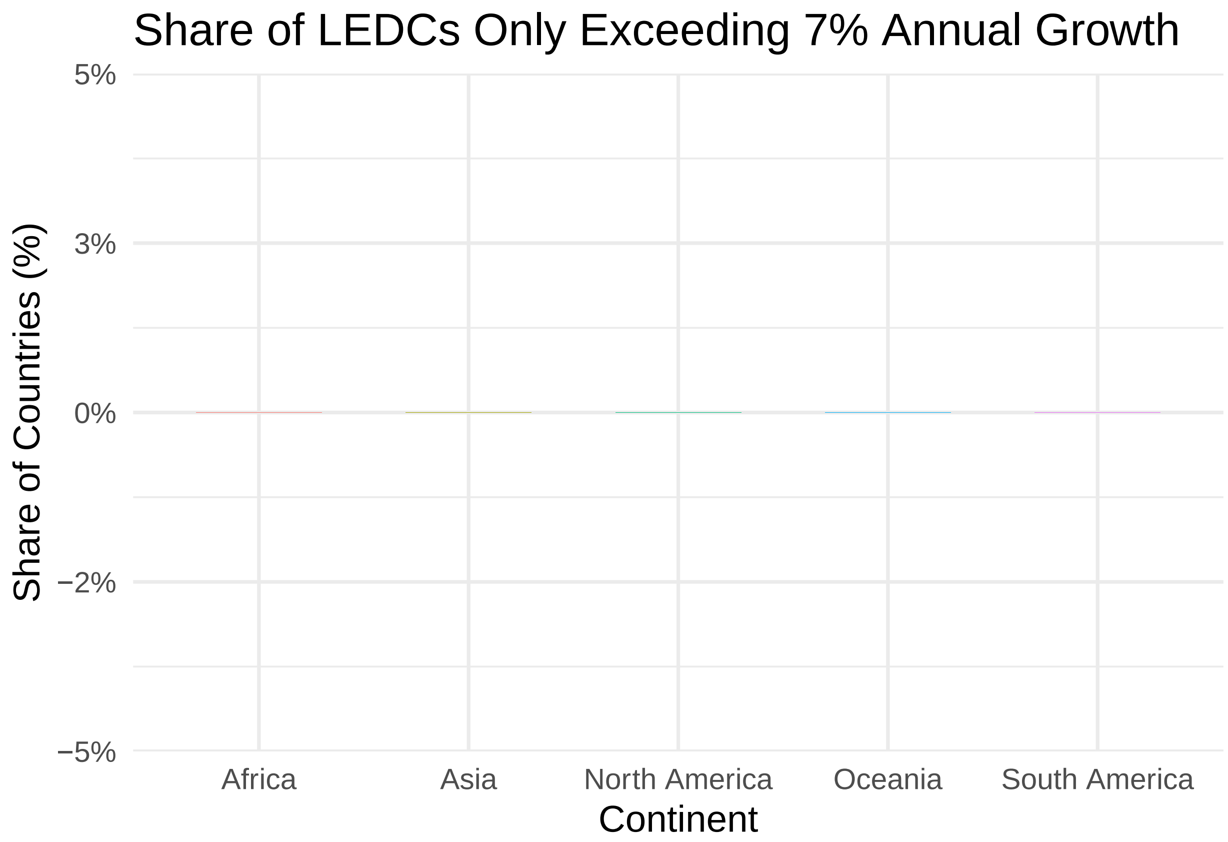


```
# Calculating Percentage of the Time Period for Which Average Growth is Above 7%
continent_share_above7 <- continent_avg %>%
  mutate(above7 = Geometric_Growth_Rate >= 7) %>%
  group_by(Continent) %>%
  summarise(share_above7 = mean(above7, na.rm = TRUE))

continent_share_above7_for_ledcs <- continent_avg_for_ledcs %>%
  mutate(above7 = Geometric_Growth_Rate >= 7) %>%
  group_by(Continent) %>%
  summarise(share_above7 = mean(above7, na.rm = TRUE))
```

```r
# (SEVEN) BAR CHART - Share of Countries Exceeding 7% Annual Growth

ggplot(continent_share_above7, aes(x = Continent, y = share_above7, fill = Continent)) +
  geom_col(width = 0.6) +
  scale_y_continuous(labels = scales::percent_format(accuracy = 1), n.breaks = 5) +
  labs(
    title = "Share of Countries Exceeding 7% Annual Growth",
    y = "Share of Countries (%)",
    x = "Continent"
  ) +
  theme_minimal(base_size = 14) +
  theme(legend.position = "none")
```



```r
# (EIGHT) BAR CHART - Share of LEDCS Exceeding 7% Annual Growth

ggplot(na.omit(continent_share_above7_for_ledcs), aes(x = Continent, y = share_above7, fill = Continent
  geom_col(width = 0.6) +
  scale_y_continuous(labels = scales::percent_format(accuracy = 1), n.breaks = 5) +
  labs(
    title = "Share of LEDCs Only Exceeding 7% Annual Growth",
    y = "Share of Countries (%)",
    x = "Continent"
  ) +
  theme_minimal(base_size = 14) +
  theme(legend.position = "none")
```

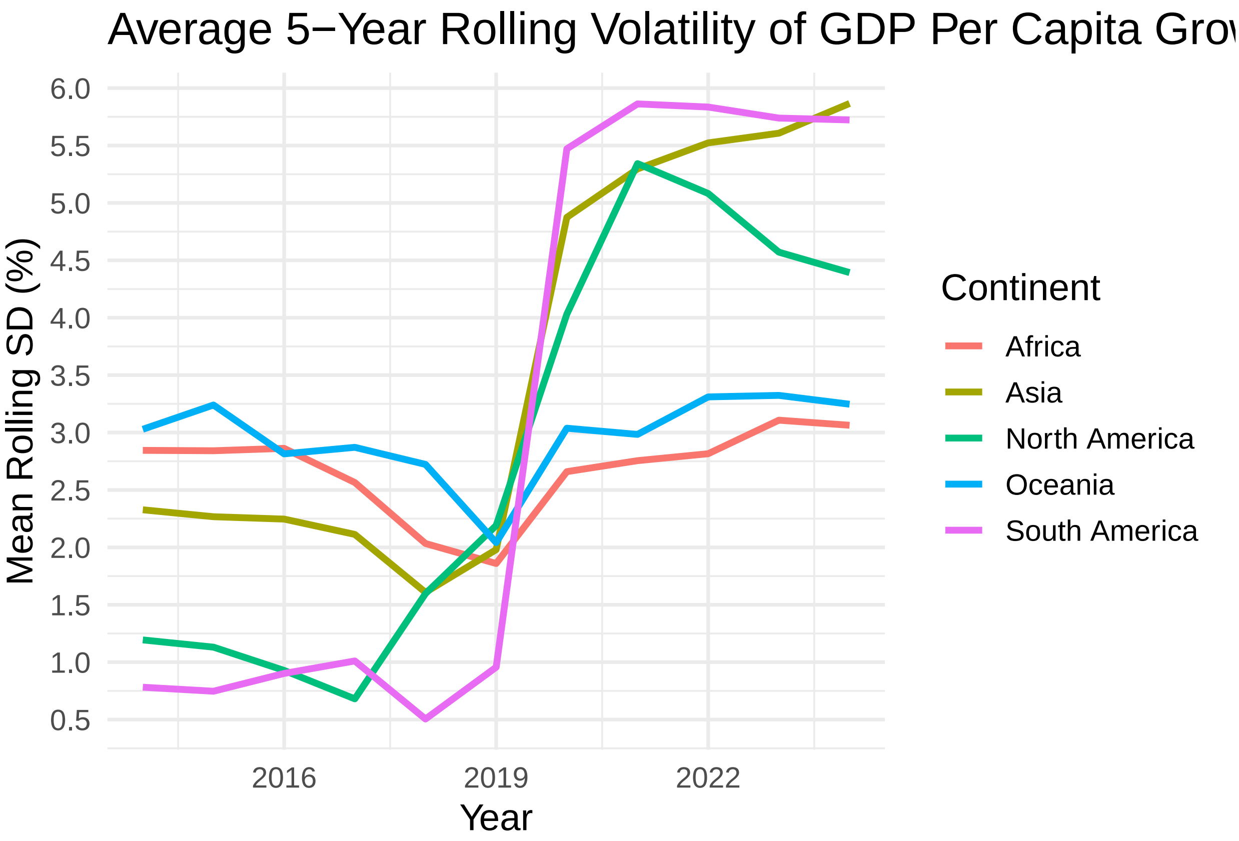


```
# Calculating Average Rolling Volatility for Each Continent
df_vol <- df_gdp %>%
  filter(`Income group` %in% c("Low income", "Lower middle income")) %>%
  group_by(Code) %>% arrange(Year) %>% mutate(
    rolling_sd = rollapply(
      Annual_Growth_Rate_PC,
      width = 5,
      FUN = function(x) sd(x, na.rm = TRUE),
      fill = NA,
      align = "right"
    )
  )


df_vol <- df_vol %>%
  group_by(Continent, Year) %>%
  summarise(mean_vol = mean(rolling_sd, na.rm = TRUE))
```

```
## `summarise()` has grouped output by 'Continent'. You can override using the
## `.groups` argument.
```

```
# (NINE) Average 5-Year Rolling Volatility of GDP Per Capita Growth in LEDCs

ggplot(df_vol[df_vol$Year > 2013,]) +
  geom_line(size = 1.2, aes(Year, mean_vol, color = Continent)) +
  labs(
    title = "Average 5-Year Rolling Volatility of GDP Per Capita Growth in LEDCs",
    y = "Mean Rolling SD (%)"
  ) +
  theme_minimal(base_size = 14) +
  scale_y_continuous(n.breaks = 10)
```

```
## 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.
```

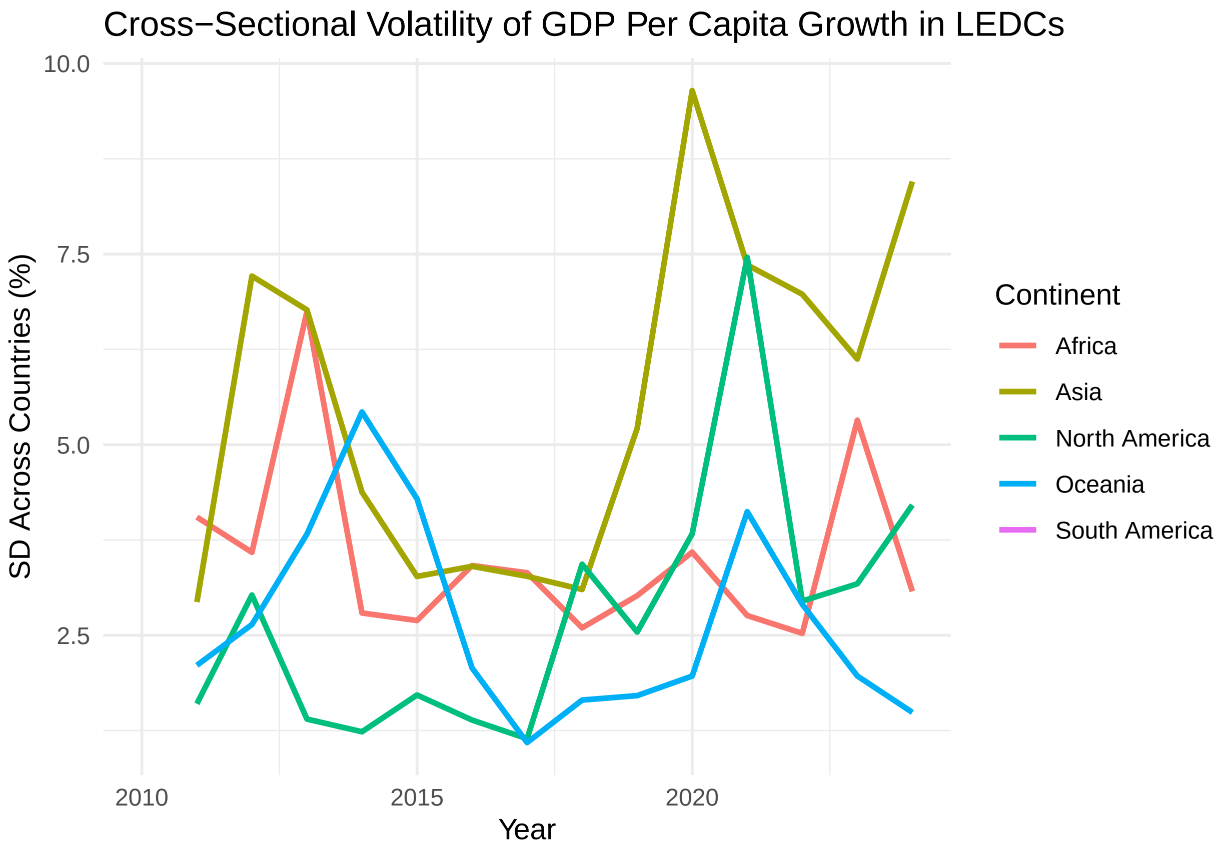


```
# (TEN) Cross Sectional Volatility Graph
df_cross <- df_gdp %>%
  filter(`Income group` %in% c("Low income", "Lower middle income")) %>%
  group_by(Continent, Year) %>%
  summarise(
    sd_growth = sd(Annual_Growth_Rate_PC, na.rm = TRUE)
  )
```

```
## `summarise()` has grouped output by 'Continent'. You can override using the
## `.groups` argument.
```

```r
ggplot(df_cross, aes(Year, sd_growth, color = Continent)) +
  geom_line(size = 1) +
  labs(title = "Cross-Sectional Volatility of GDP Per Capita Growth in LEDCs",
       y = "SD Across Countries (%)") +
  theme_minimal()
```

```
## Warning: Removed 19 rows containing missing values or values outside the scale range
## (`geom_line()`).
```


Cross−Sectional Volatility of GDP Per Capita Growth in LEDCs
SD Across Countries (%)
Year
Continent
Africa
Asia
North America
Oceania
South America