PPHA 37040 Problem Set 1: Structural Transformation Training; Analysis of the Philippines

Nat Larsen, nclarsen

2025-01-23

Contents


```r
# Install Necessary Packages

# Install dplyr
if (!requireNamespace("dplyr", quietly = TRUE)) {
  install.packages("dplyr")
}

# Install ggplot2
if (!requireNamespace("ggplot2", quietly = TRUE)) {
  install.packages("ggplot2")
}

# Install dendextend
if (!requireNamespace("dendextend", quietly = TRUE)) {
  install.packages("dendextend")
}

# Install strucchange
if (!requireNamespace("strucchange", quietly = TRUE)) {
```

```r
  install.packages("strucchange")
}

# Install haven package to read the .dta file
if (!requireNamespace("haven", quietly = TRUE)) {
  install.packages("haven")
}

# Install tidyr
if (!requireNamespace("tidyr", quietly = TRUE)) {
  install.packages("tidyr")
}

# Install scales
if (!requireNamespace("scales", quietly = TRUE)) {
  install.packages("scales")
}

# Install ggrepel
if (!requireNamespace("ggrepel", quietly = TRUE)) {
  install.packages("ggrepel")
}
```

Data Preparation

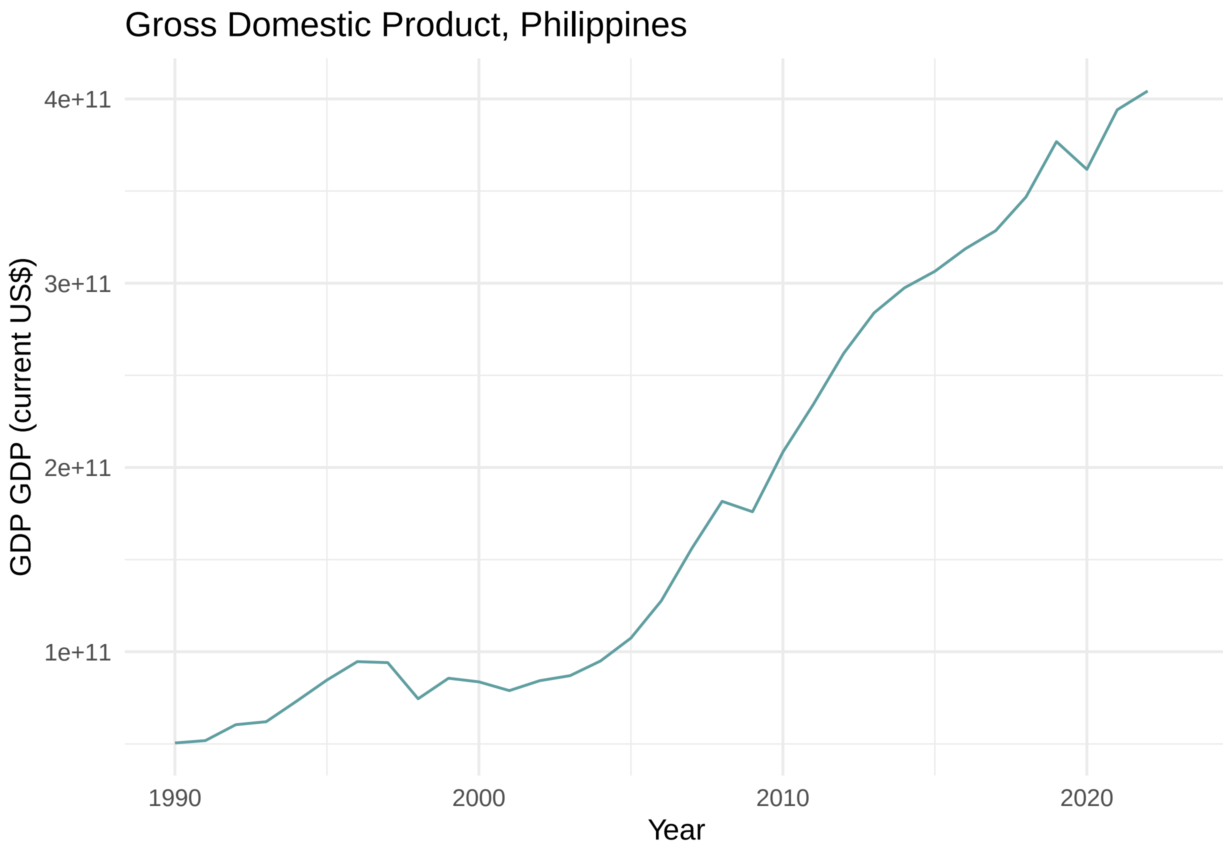
Load and Prepare Data

```r
# Read the .dta file
# .dta file must be in the same working directory as this .Rmd
# data <- read_dta("glmacro_master_alldata.dta")
data <- read_dta("/Users/natlarsen/Downloads/Master Data/glmacro_master_alldata.dta")
```

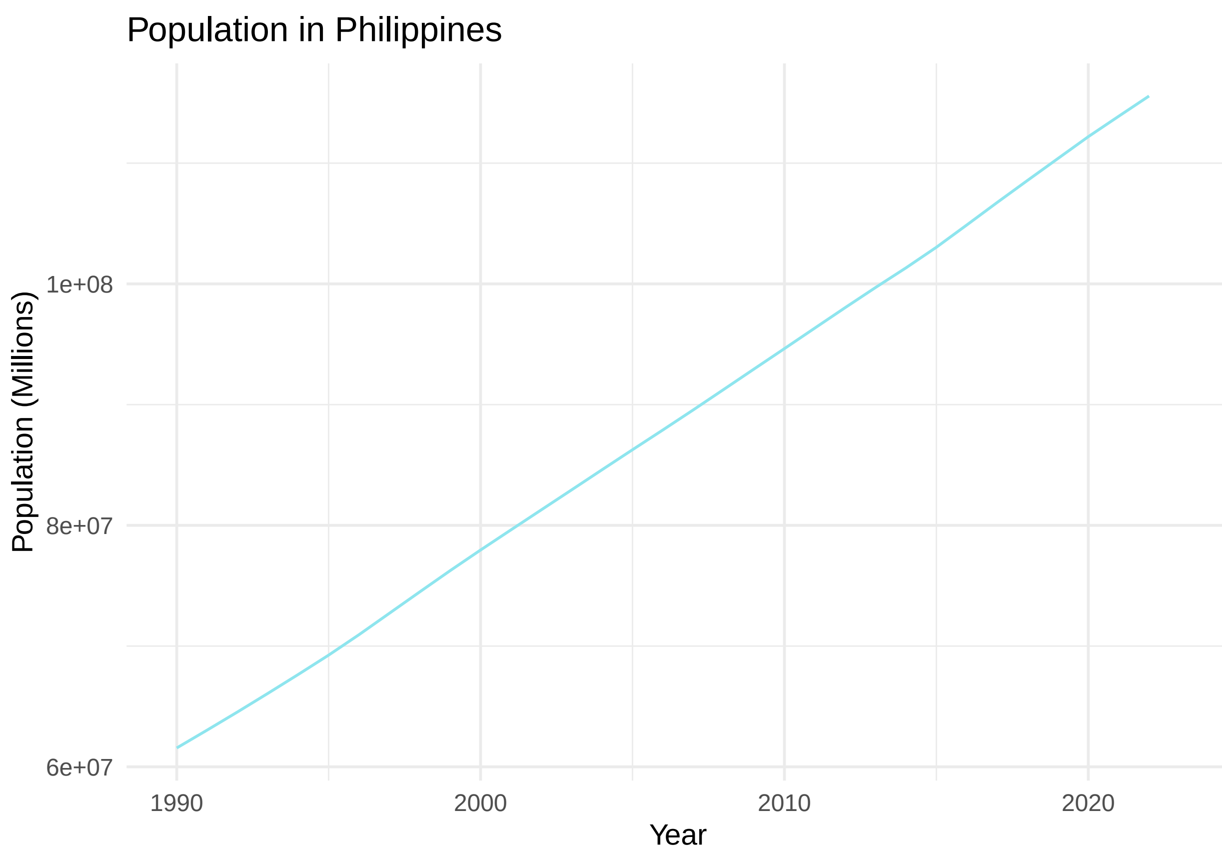
```r
# Select the Philippines as the country
country <- "Philippines"

# Add a dummy column for a chosen country
data <- data %>%
  mutate(dummy_chosen_country = ifelse(wb_countryname == country, 1, 0))
```

Question 1: History of Population and GDP (10 points)

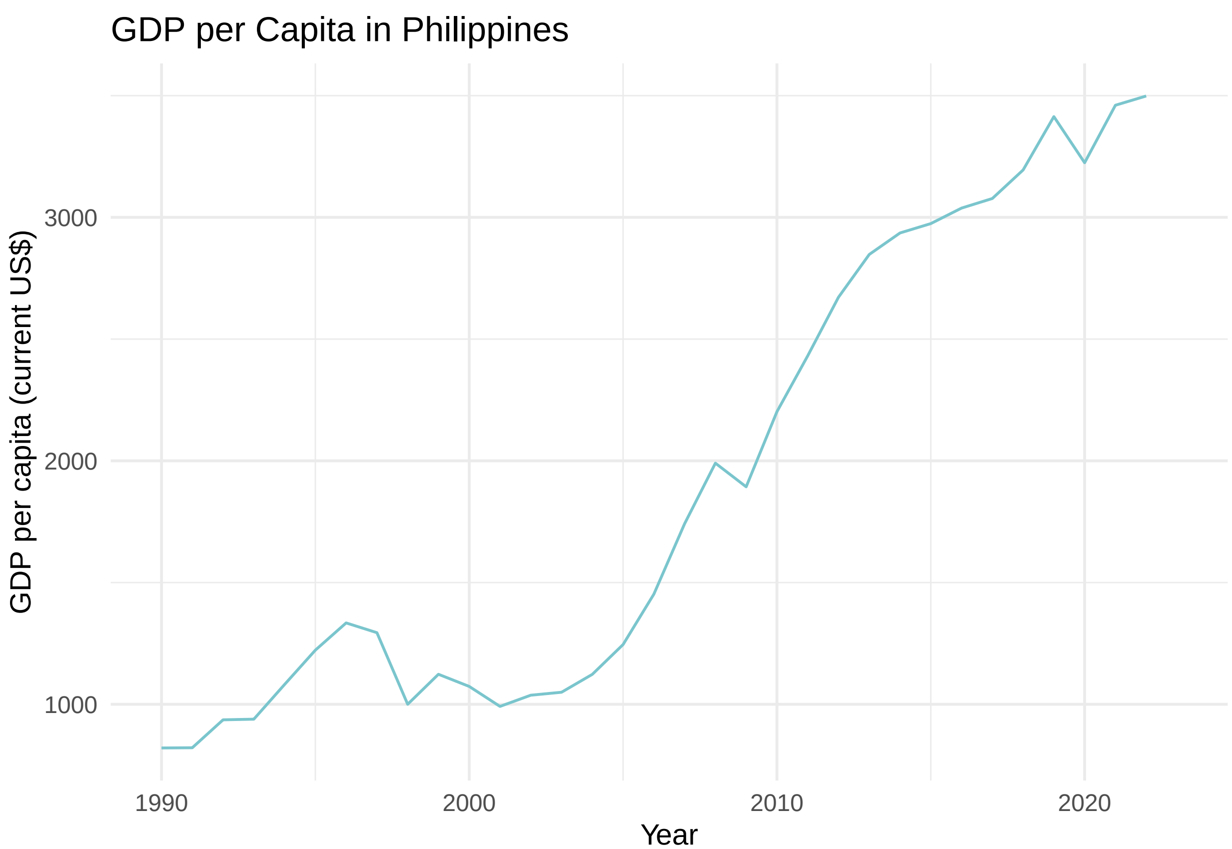
```r
# Filter the dataset and plot the GDP
ggplot(data %>% filter(wb_countryname == country & year >= 1990 & year <= 2023),
       aes(x = year, y = wdi_ny_gdp_mktp_cd)) +
  geom_line(color = "cadetblue") +
  labs(title = paste("Gross Domestic Product,", country),
       y = "GDP GDP (current US$)",
       x = "Year") +
  theme_minimal()
```



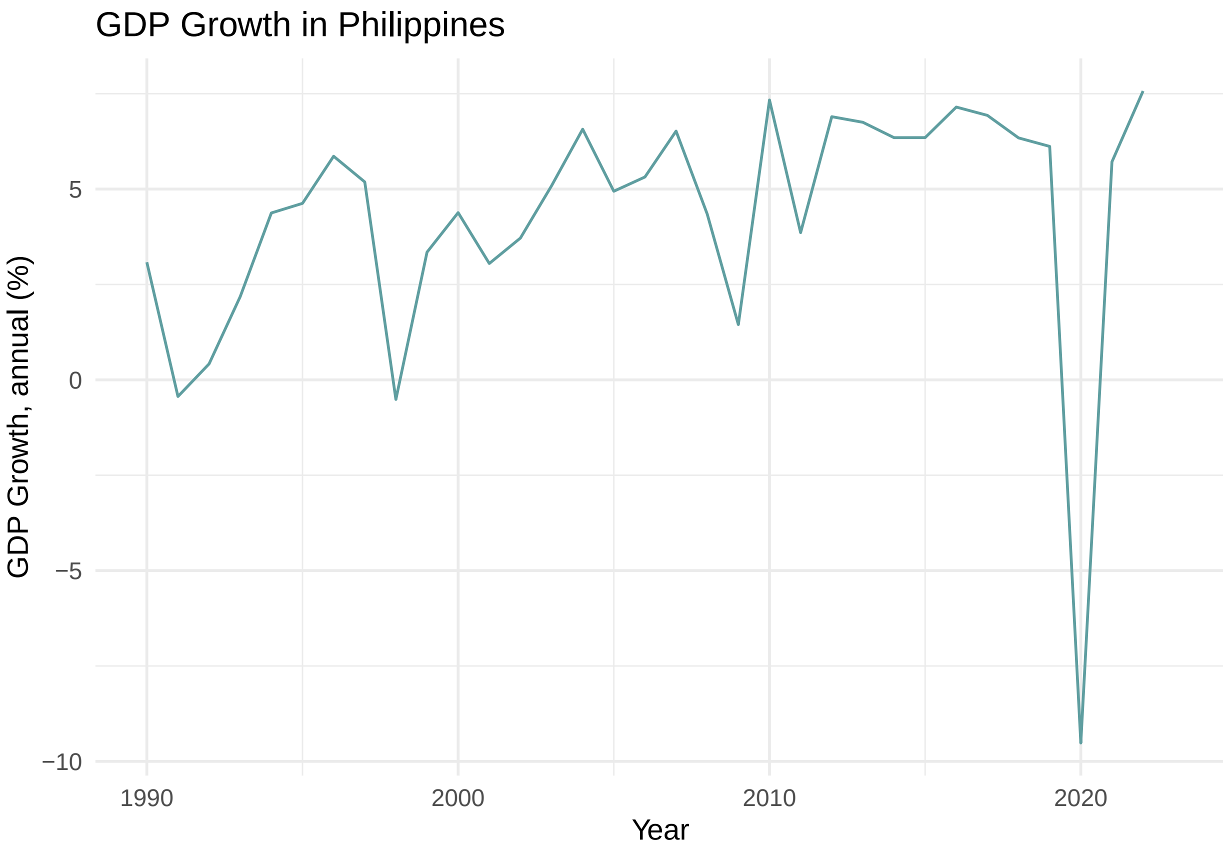

```r
# Plot Population Time Series
ggplot(data %>% filter(wb_countryname == country & year >= 1990 & year <= 2023),
       aes(x = year, y = wdi_sp_pop_totl)) +
  geom_line(color = "cadetblue2") +
  labs(title = paste("Population in", country),
       y = "Population (Millions)",
       x = "Year") +
  theme_minimal()
```




```
# Plot GDP per capita Time Series
ggplot(data %>% filter(wb_countryname == country & year >= 1990 & year <= 2023),
       aes(x = year, y = wdi_ny_gdp_pcap_cd)) +
  geom_line(color = "cadetblue3") +
  labs(title = paste("GDP per Capita in", country),
       y = "GDP per capita (current US$)",
       x = "Year") +
  theme_minimal()
```

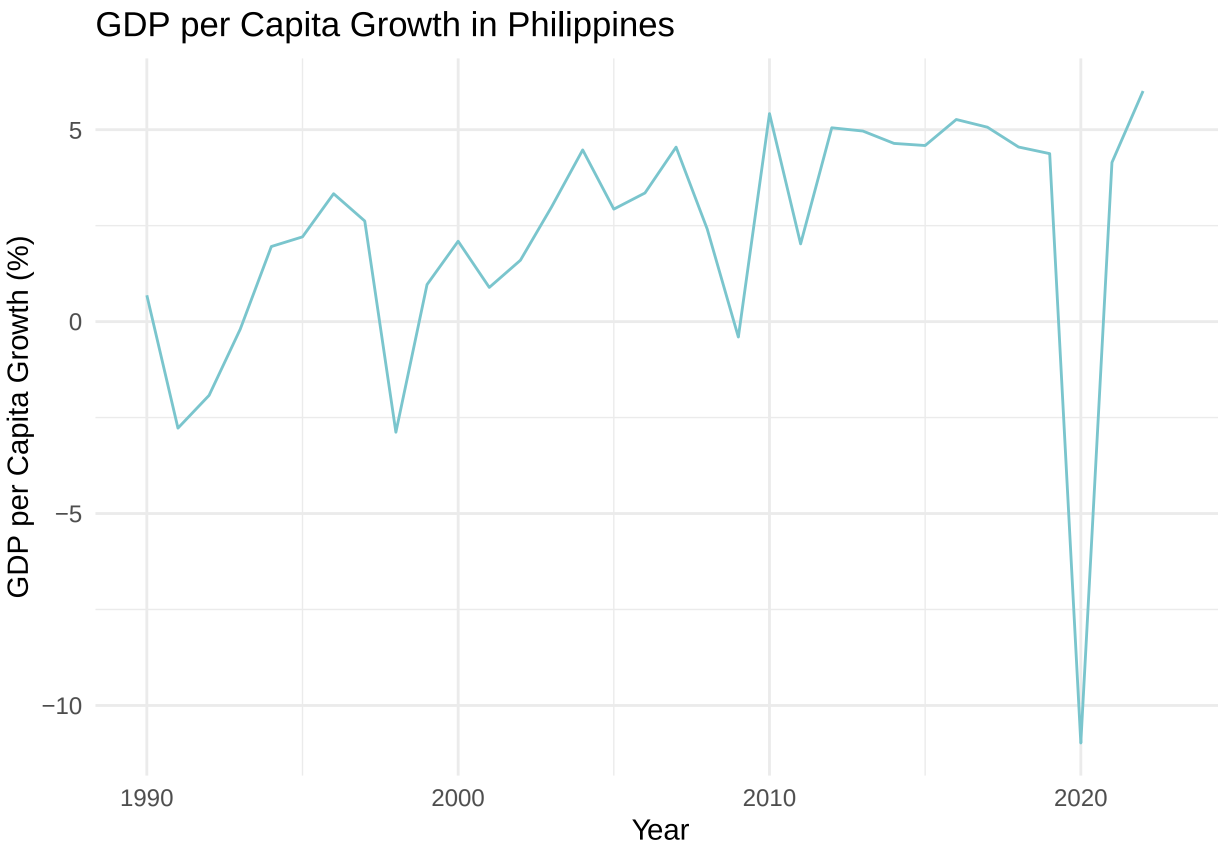



```r
# Plot Real GDP Growth
ggplot(data %>% filter(wb_countryname == country & year >= 1990 & year <= 2023),
       aes(x = year, y = wdi_ny_gdp_mktp_kd_zg)) +
  geom_line(color = "cadetblue") +
  labs(title = paste("GDP Growth in", country),
       y = "GDP Growth, annual (%)",
       x = "Year") +
  theme_minimal()
```




```r
# Plot Population Growth
ggplot(data %>% filter(wb_countryname == country & year >= 1990 & year <= 2023),
       aes(x = year, y = wdi_sp_pop_grow)) +
  geom_line(color = "cadetblue2") +
  labs(title = paste("Population Growth in", country),
       y = "Population Growth, annual (%)",
       x = "Year") +
  theme_minimal()
```


```r
# Plot GDP Per Capita Growth
ggplot(data %>% filter(wb_countryname == country & year >= 1990 & year <= 2023),
       aes(x = year, y = wdi_ny_gdp_pcap_kd_zg)) +
  geom_line(color = "cadetblue3") +
  labs(title = paste("GDP per Capita Growth in", country),
       y = "GDP per Capita Growth (%)",
       x = "Year") +
  theme_minimal()
```



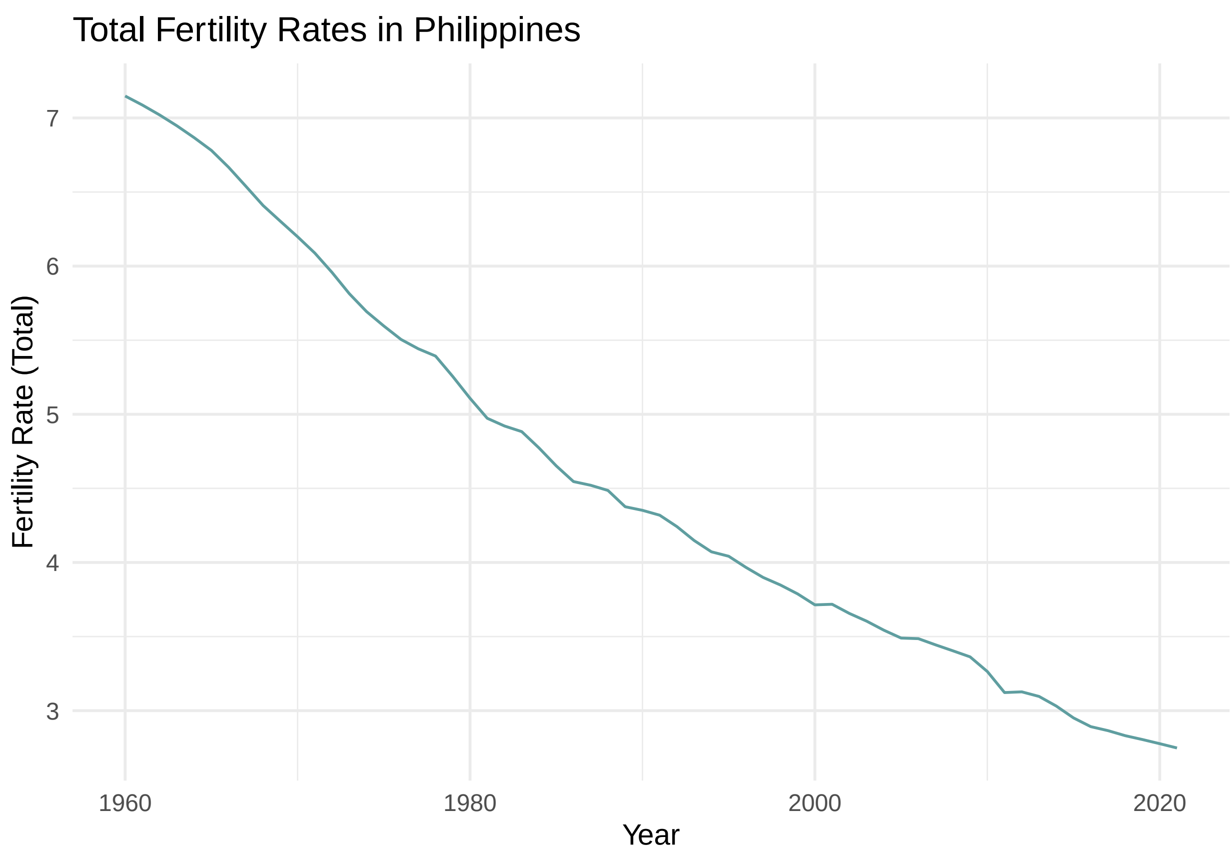
Between 1990 and 2023, the Philippines' **Real GDP** increased from 1 to 4 billion USD; **Population** consistently increased in population from 6 million to over 100 million; and **GDP per capita** increased, although less dramatically than real GDP, from under 1000 to around 3500 USD.

Real GDP Growth and **Real GDP per capita growth** were volatile, rising and then falling dramatically around 2000, 2010, and 2020, coinciding with the 1990s Asian Financial Crisis, the 2008 Global Financial Crisis, and the 2020 Pandemic. While the population is growing, the **Rate of Population Growth** is declining—after peaking around 1997, population growth shrunk from nearly 2.5% to current levels under 1.5% annually. While both real GDP and population are growing, the Philippines experienced more rapid rates of economic growth than population growth.

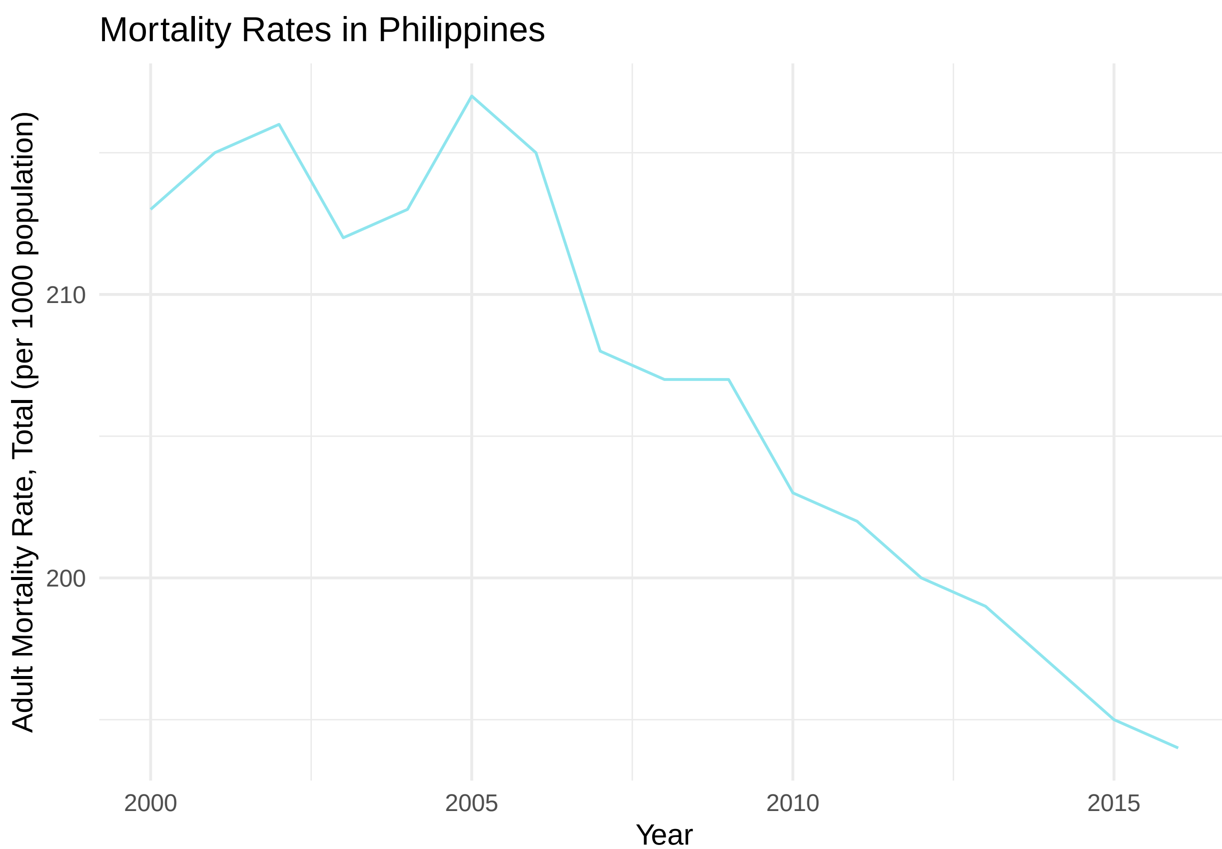
The Philippines' **growth trajectory** is then characterized by consistent GDP and population growth, yet vulnerability to external shocks. The Philippines' GDP growth is positive for improving human quality of life and development, and as we discussed in Lecture 1, compounds over time to sustain future growth and generations.

Question 2: Demographic Transition (10 points)

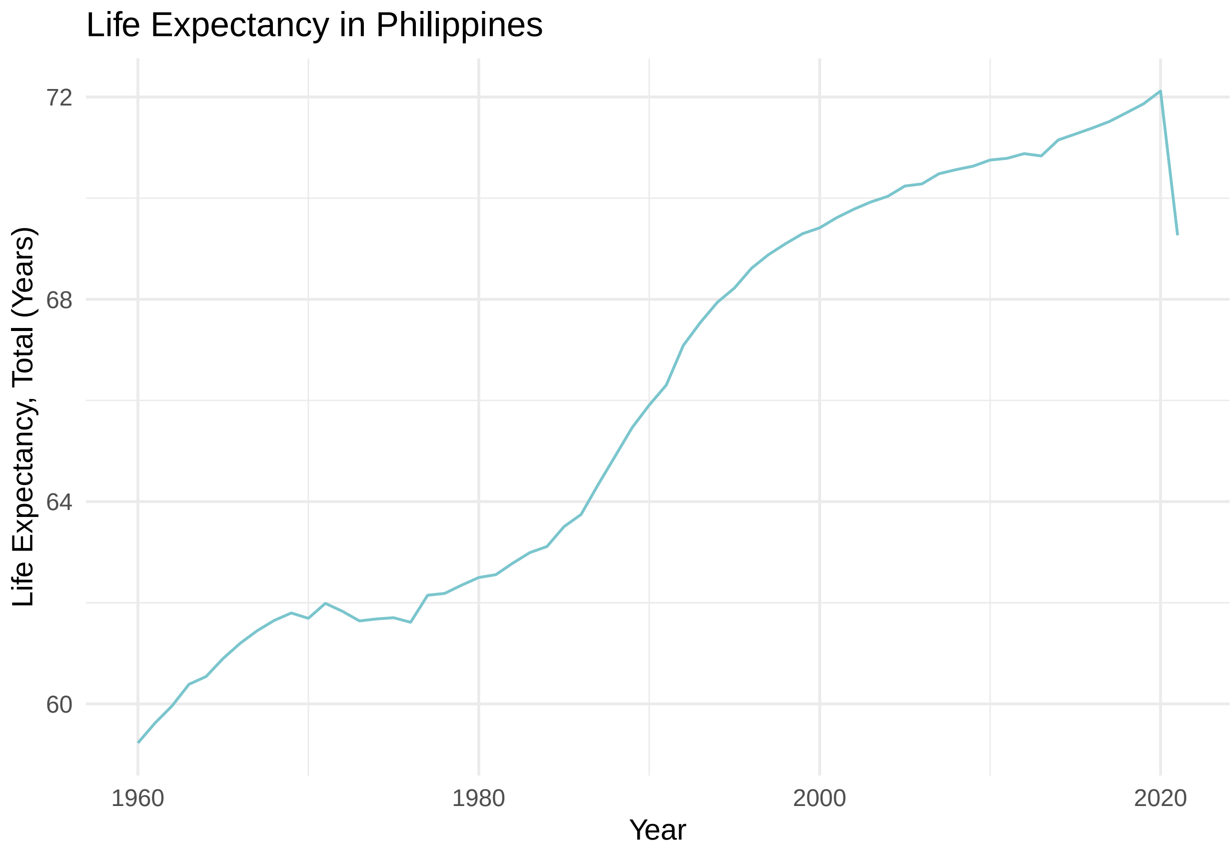
```r
# Fertility Rates Time Series
ggplot(data %>% filter(wb_countryname == country & !is.na(wdi_sp_dyn_tfrt_in)),
       aes(x = year, y = wdi_sp_dyn_tfrt_in)) +
  geom_line(color = "cadetblue") +
  labs(title = paste("Total Fertility Rates in", country),
       y = "Fertility Rate (Total)",
       x = "Year") +
  theme_minimal()
```



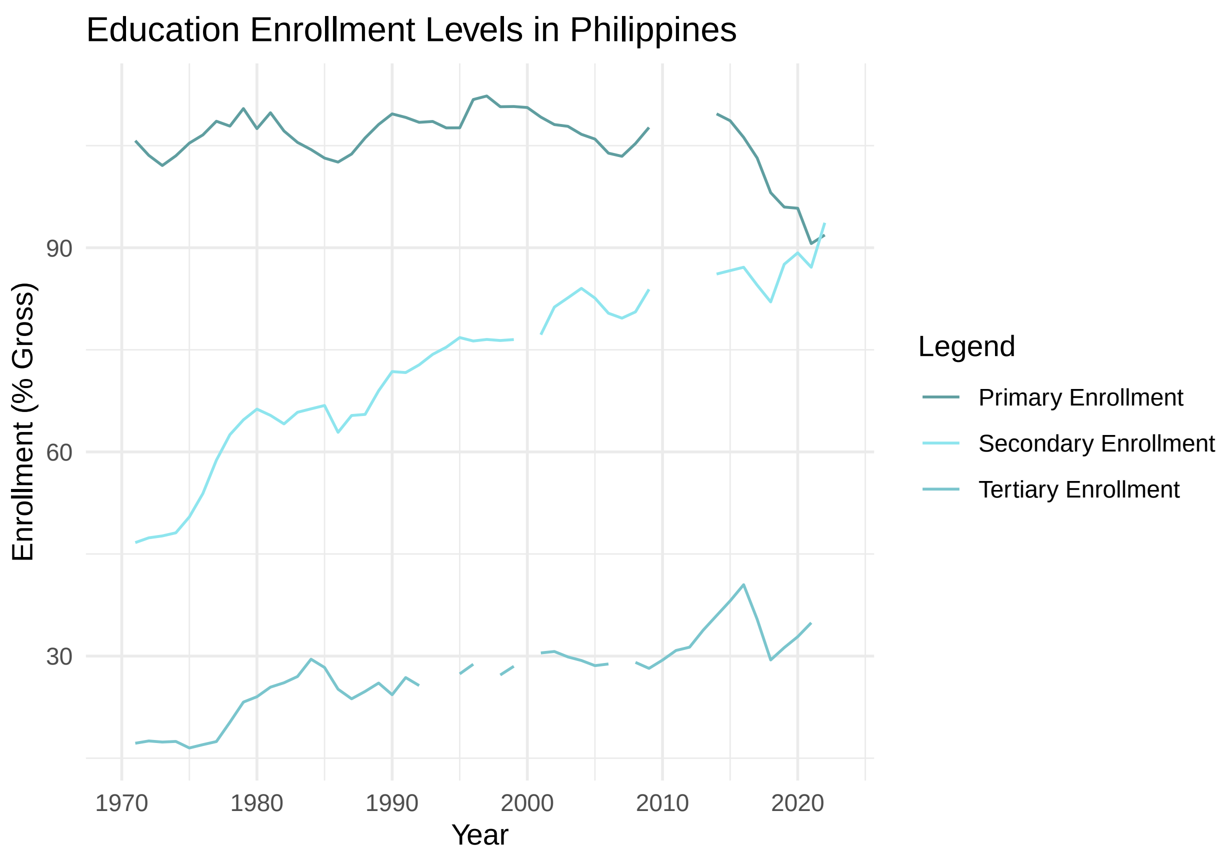

```r
# Mortality Rates Time Series
ggplot(data %>% filter(wb_countryname == country & !is.na(qog_who_mrt)),
       aes(x = year, y = qog_who_mrt)) +
  geom_line(color = "cadetblue2") +
  labs(title = paste("Mortality Rates in", country),
       y = "Adult Mortality Rate, Total (per 1000 population)",
       x = "Year") +
  theme_minimal()
```



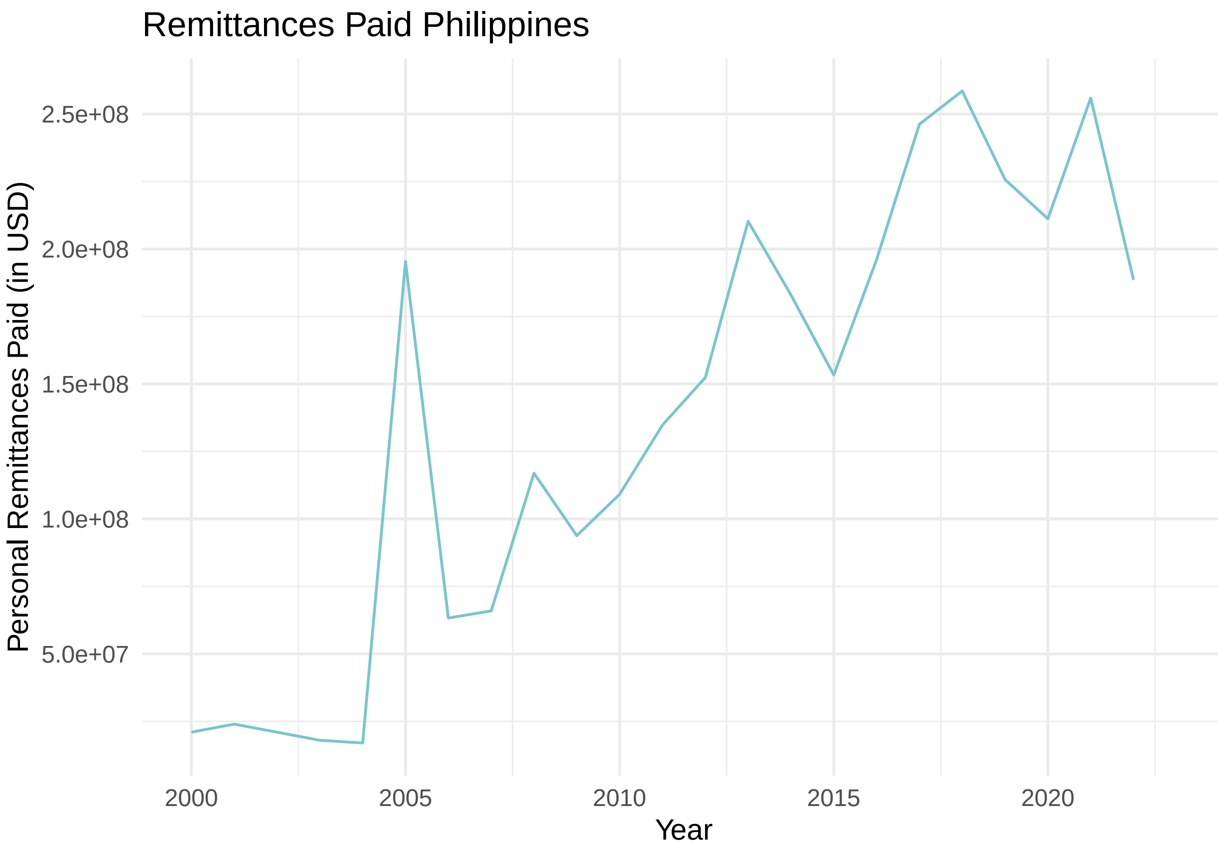

```r
# Life Expectancy Time Series
ggplot(data %>% filter(wb_countryname == country & !is.na(wdi_sp_dyn_le00_in)),
       aes(x = year, y = wdi_sp_dyn_le00_in)) +
  geom_line(color = "cadetblue3") +
  labs(title = paste("Life Expectancy in", country),
       y = "Life Expectancy, Total (Years)",
       x = "Year") +
  theme_minimal()
```




```r
# Plot Education Levels (Primary, Secondary, Tertiary)
ggplot(data %>% filter(wb_countryname == country & year >= 1970 & year <= 2023),
       aes(x = year)) +
  geom_line(aes(y = wdi_se_prm_enrr, color = "Primary Enrollment")) +
  geom_line(aes(y = wdi_se_sec_enrr, color = "Secondary Enrollment")) +
  geom_line(aes(y = wdi_se_ter_enrr, color = "Tertiary Enrollment")) +
  scale_color_manual(values = c("Primary Enrollment" = "cadetblue",
                                "Secondary Enrollment" = "cadetblue2",
                                "Tertiary Enrollment" = "cadetblue3")) +
  labs(title = paste("Education Enrollment Levels in", country),
       y = "Enrollment (% Gross)",
       x = "Year",
       color = "Legend") +
  theme_minimal()
```




```
# Additional Plot: Remittances
ggplot(data %>% filter(wb_countryname == country & year >= 2000 & year <= 2023),
       aes(x = year)) +
  geom_line(aes(y = wdi_bm_trf_pwkr_cd_dt), color = "cadetblue3") +
  labs(title = paste("Remittances Paid", country),
       y = "Personal Remittances Paid (in USD)",
       x = "Year") +
  theme_minimal()
```



The Philippines' **Total Fertility Rates** and **Mortality Rates** are declining, while **Life Expectancy** is increasing. This negative relationship between GDP and fertility aligns with the modern growth stage of development studied in Lecture 2, rather than the positive relationship observed in the Malthusian or post-Malthusian stages.

A demographic window opens when the decrease in death rates precedes that in birth rates—the people born outnumber those dying. Given its consistently increasing population, the Philippines is *presently experiencing a demographic window of opportunity*, as an expanding working population offers opportunities to develop production and growth. This demographic window of opportunity may close if fertility rates decline to the point of a demographic dependency, where an older population's needs exceed a working population's production.

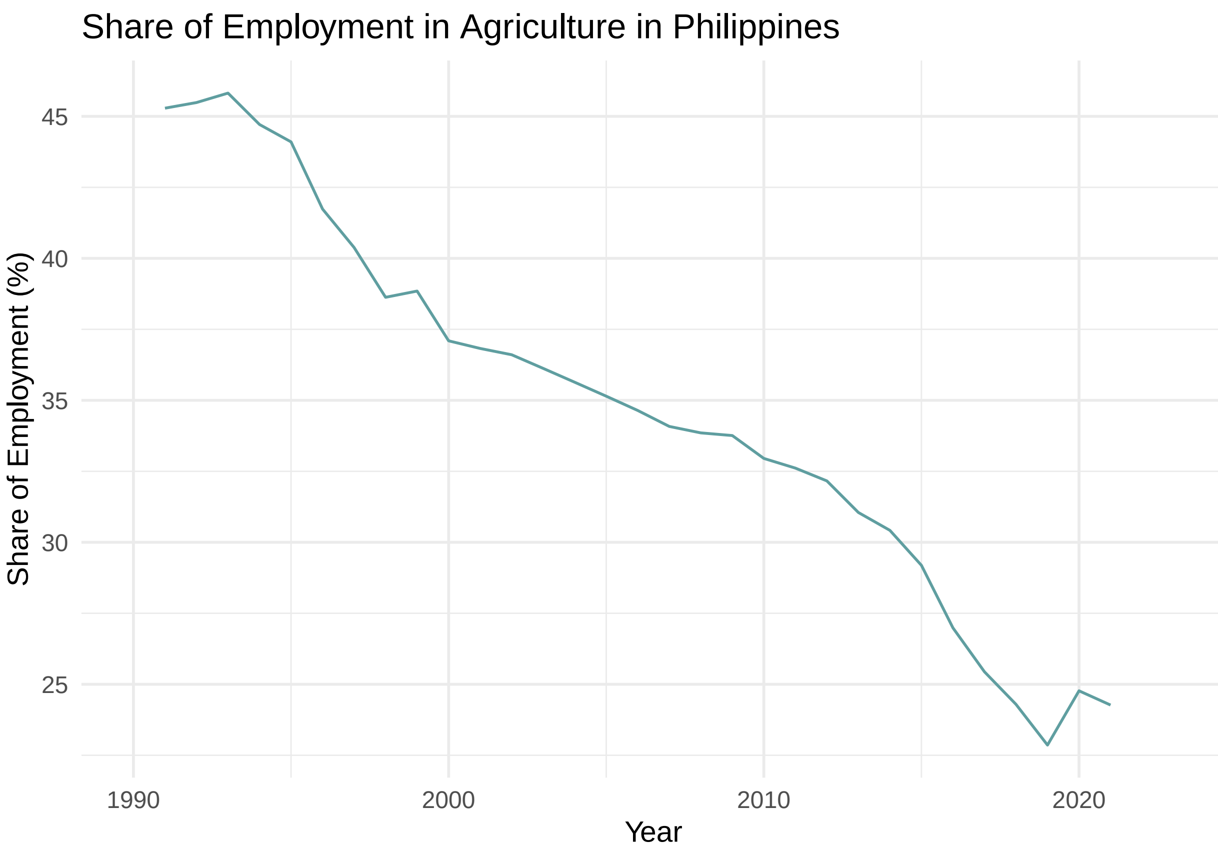
Since the 1970s, **Education Enrollment** increased, per the predictions of the demographic bonus for greater incentives for investment in a growing population; however, the lagging tertiary enrollment reveals further opportunities to expand education.

Having researched its importance for Filipino economic and human development, I additionally examined **Remittances Paid**, which increased dramatically since 2000. This rising inflow of funds from overseas boosts household income and national output, suggesting a role of international labor mobility and diffusion of know-how to fuel this observed growth.

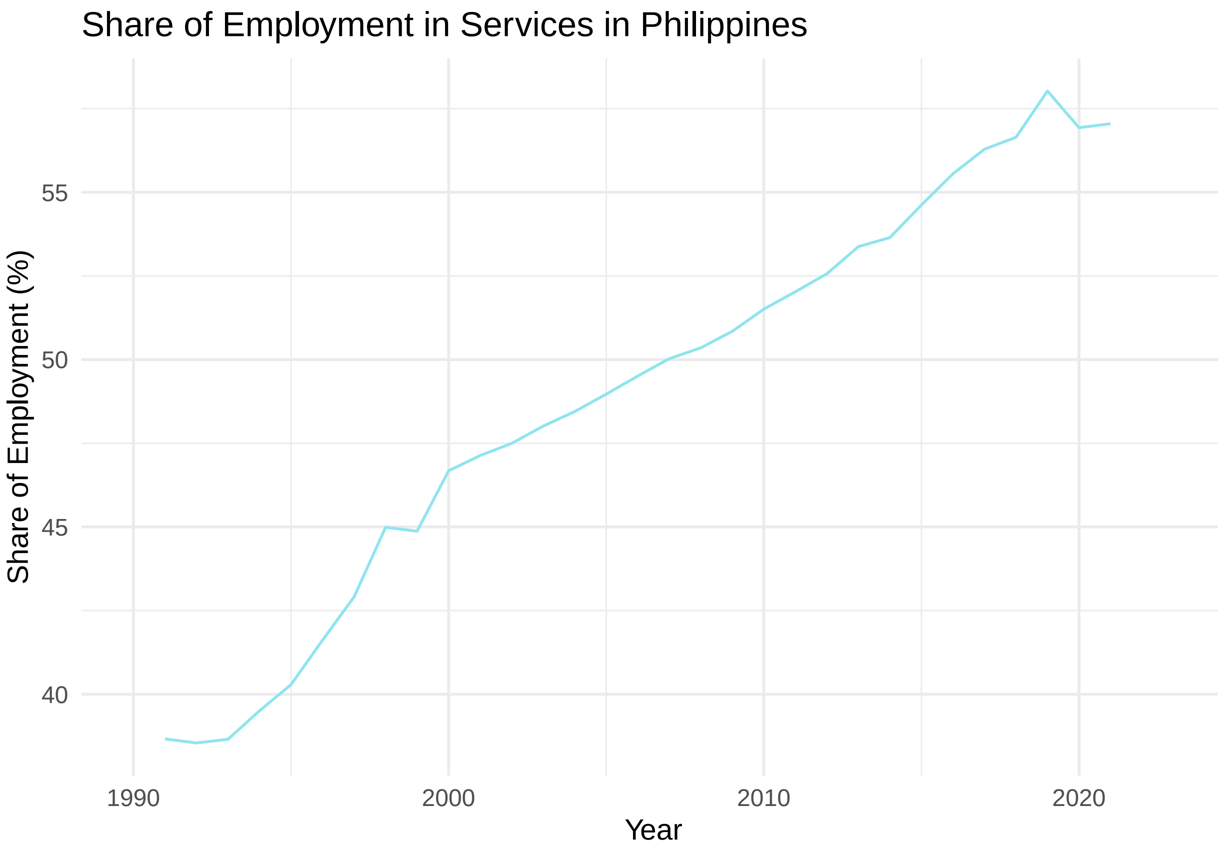
Question 3: History of Structural Transformation (20 points)

```r
# Employment Share in Agriculture
ggplot(data %>% filter(wb_countryname == country & year >= 1990 & year <= 2023),
       aes(x = year, y = wdi_sl_agr_empl_zs)) +
  geom_line(color = "cadetblue") +
  labs(title = paste("Share of Employment in Agriculture in", country),
```

```
       y = "Share of Employment (%)",
       x = "Year") +
  theme_minimal()
```



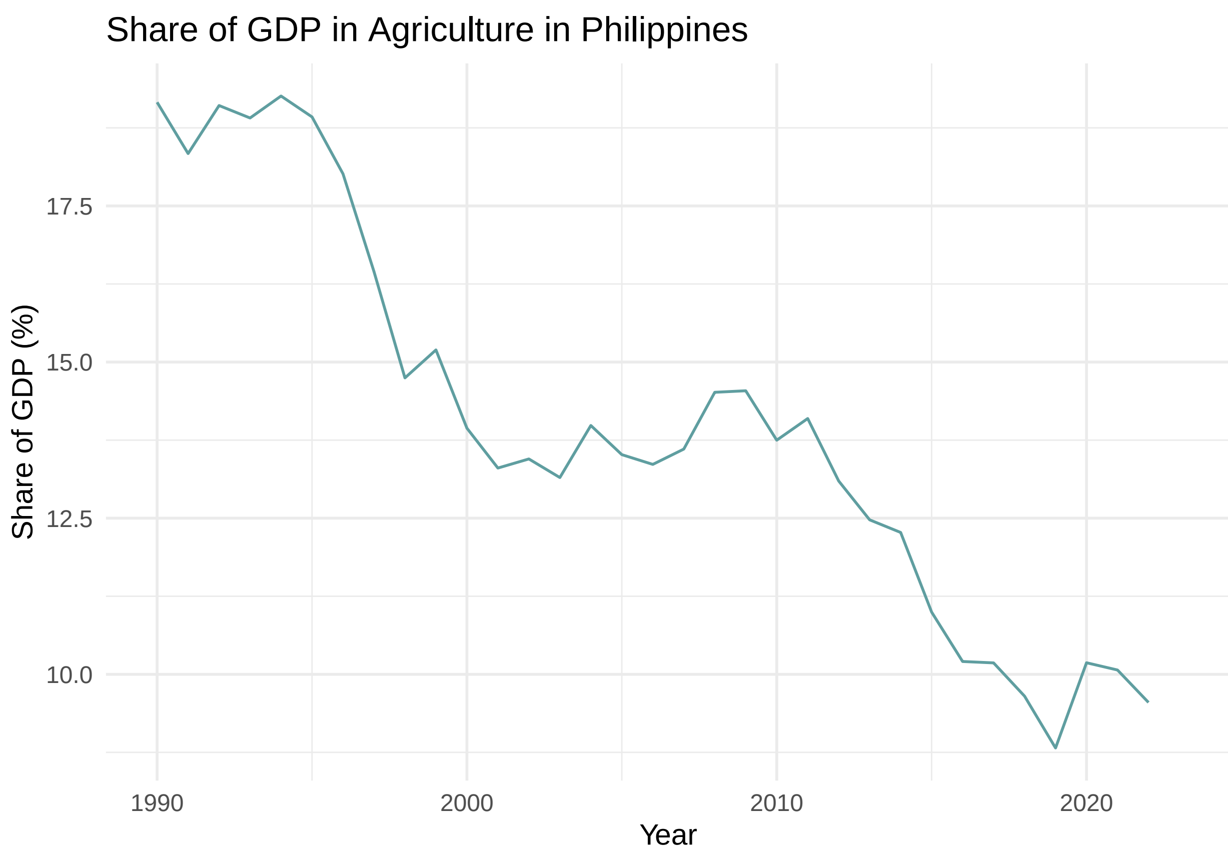

```
# Plot Employment Share in Services
ggplot(data %>% filter(wb_countryname == country & year >= 1990 & year <= 2023),
       aes(x = year, y = wdi_sl_srv_empl_zs)) +
  geom_line(color = "cadetblue2") +
  labs(title = paste("Share of Employment in Services in", country),
       y = "Share of Employment (%)",
       x = "Year") +
  theme_minimal()
```




```
# Plot Employment Share in Manufacturing
ggplot(data %>% filter(wb_countryname == country & year >= 1990 & year <= 2023),
       aes(x = year, y = wdi_sl_ind_empl_zs)) +
  geom_line(color = "cadetblue3") +
  labs(title = paste("Employment in Manufacturing in", country),
       y = "Share of Employment (%)",
       x = "Year") +
  theme_minimal()
```

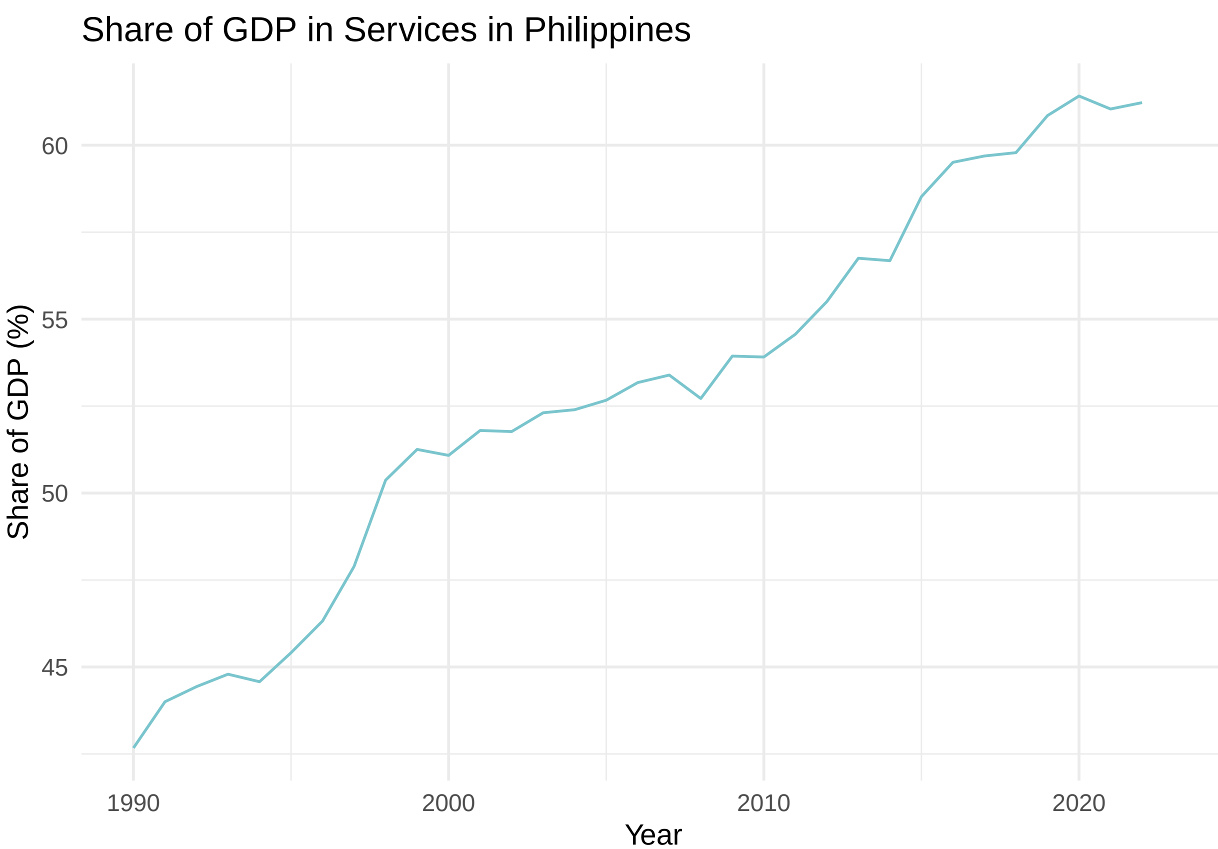

```r
# Convert total GDP from billions to US$
data <- data %>%
  mutate(total_gdp_usd = weo_ngdpd * 1e9)
```

```r
# Plot Share of GDP in Agriculture
ggplot(data %>% filter(wb_countryname == country & year >= 1990 & year <= 2023),
       aes(x = year, y = wdi_nv_agr_totl_zs)) +
  geom_line(color = "cadetblue") +
  labs(title = paste("Share of GDP in Agriculture in", country),
       y = "Share of GDP (%)",
       x = "Year") +
  theme_minimal()
```




```r
# Plot Share of GDP in Industry
ggplot(data %>% filter(wb_countryname == country & year >= 1990 & year <= 2023),
       aes(x = year, y = wdi_nv_ind_totl_zs)) +
  geom_line(color = "cadetblue2") +
  labs(title = paste("Share of GDP in Industry in", country),
       y = "Share of GDP (%)",
       x = "Year") +
  theme_minimal()
```


```r
# Plot Share of GDP in Services
ggplot(data %>% filter(wb_countryname == country & year >= 1990 & year <= 2023),
       aes(x = year, y = wdi_nv_srv_totl_zs)) +
  geom_line(color = "cadetblue3") +
  labs(title = paste("Share of GDP in Services in", country),
       y = "Share of GDP (%)",
       x = "Year") +
  theme_minimal()
```




```r
# Reshape GDP shares data of agriculture, industry, and services sector

data_long <- data %>%
  filter(wb_countryname == country & year >= 1990 & year <= 2023) %>%
  select(year, wdi_nv_agr_totl_zs, wdi_nv_ind_totl_zs, wdi_nv_srv_totl_zs) %>%
   pivot_longer(
    cols = c(wdi_nv_agr_totl_zs, wdi_nv_ind_totl_zs, wdi_nv_srv_totl_zs),
    names_to = "Sector",
    values_to = "Share"
  ) %>%
  mutate(Sector = recode(Sector,
                         wdi_nv_agr_totl_zs = "Agriculture",
                         wdi_nv_ind_totl_zs = "Industry",
                         wdi_nv_srv_totl_zs = "Services"))
```

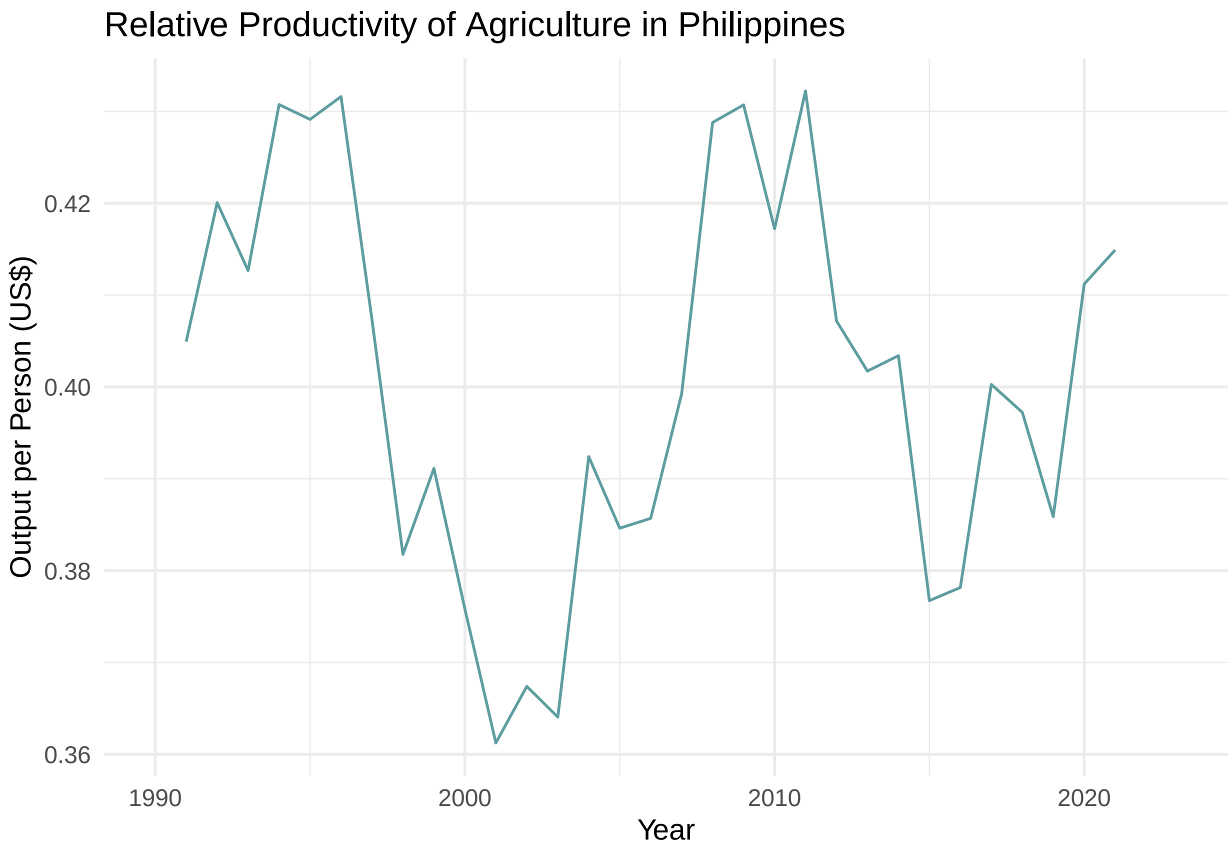
```r
# Plot Share of GDP by sector in stacked bar
data_long_filtered <- data_long %>%
  filter(year %% 5 == 0)  # Display data every five years

ggplot(data_long_filtered, aes(x = factor(year), y = Share, fill = Sector)) +
  geom_bar(stat = "identity", position = "stack") +
  scale_fill_manual(values = c("Agriculture" = "cadetblue",
                               "Industry" = "cadetblue2",
                               "Services" = "cadetblue3")) +
  labs(
    title = paste("Share of GDP by Sector in", country),
    x = "Year",
```

```r
    y = "% of GDP",
    fill = "Sector"
  ) +
  theme_minimal()
```

Plot Relative Productivity

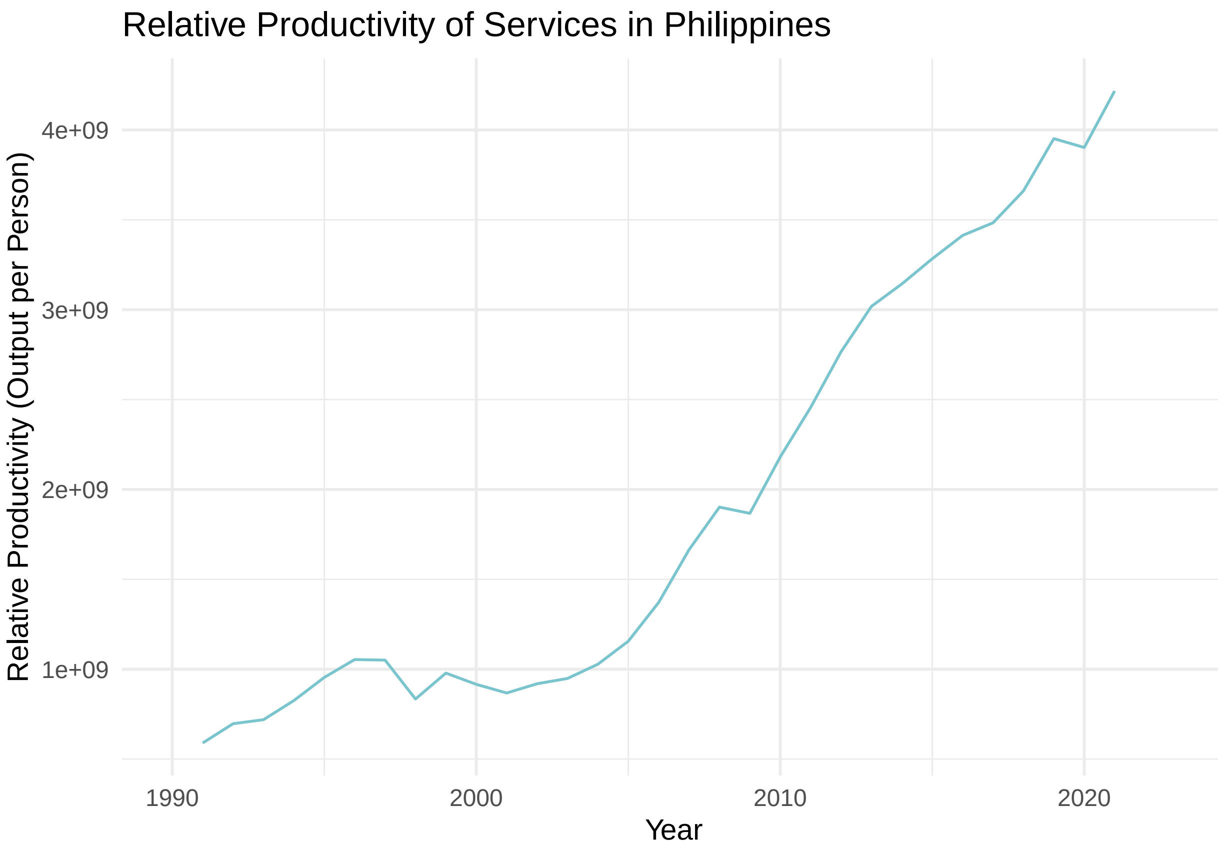
```r
# Calculate relative productivity of agriculture
data <- data %>%
  mutate(
    rel_prod_agriculture = wdi_nv_agr_totl_zs / wdi_sl_agr_empl_zs)
```

```r
# Plot relative productivity in agriculture

ggplot(data %>% filter(wb_countryname == country & year >= 1990 & year <= 2023),
       aes(x = year, y = rel_prod_agriculture)) +
  geom_line(color = "cadetblue") +
  labs(title = paste("Relative Productivity of Agriculture in", country),
       y = "Output per Person (US$)",
       x = "Year") +
  theme_minimal()
```




```r
# Calculate Relative Productivity for Manufacturing (Industry)
data <- data %>%
  mutate(
    manf_rp = wdi_nv_ind_manf_zs / wdi_sl_ind_empl_zs)

# Plot Relative Productivity for Manufacturing
ggplot(data %>% filter(wb_countryname == country & year >= 1990 & year <= 2023),
       aes(x = year, y = manf_rp)) +
  geom_line(color = "cadetblue2") +
  labs(
    title = paste("Relative Productivity of Manufacturing in", country),
    y = "Relative Productivity (Output per Person)",
    x = "Year"
  ) +
  theme_minimal()
```


```r
# Calculate Relative Productivity for Services
data <- data %>%
  mutate(
    manf_rp = wdi_nv_srv_totl_cd / wdi_sl_srv_empl_zs)

# Plot Relative Productivity for Services
ggplot(data %>% filter(wb_countryname == country & year >= 1990 & year <= 2023),
       aes(x = year, y = manf_rp)) +
  geom_line(color = "cadetblue3") +
  labs(
    title = paste("Relative Productivity of Services in", country),
    y = "Relative Productivity (Output per Person)",
    x = "Year"
  ) +
  theme_minimal()
```



Between 1990-2020, the Philippines' **Share of Employment in Agriculture** shrank to 20% while the **Share of Employment in Services** correspondingly grew to 57%. **Employment in Manufacturing** remained more stable, increasing slightly between 2010-2020 from 15% to 19%. Mirroring these employment trends, the **Share of GDP in Agriculture** decreased to under 10% while the **Share of GDP in Services** skyrocketed to over 60%. The **Share of GDP in Industry** stayed relatively the same, shrinking slightly. Meanwhile, **Relative Productivity** grew across **Agriculture**, **Services**, and **Industry**, revealing efficiency and output improvements per worker across sectors.

Comparing 1990 and 2020, the Philippines underwent a structural economic shift away from agriculture toward services, revealing an *acceleration* in services' employment and GDP share and a *deceleration* in agriculture's employment and GDP share; while relative productivity improved across sectors. These patterns remain consistent across time, rather than showing any significant volatility. This economic trajectory aligns with the Malthusian predictions that countries shift from land-intensive to capital-intensive economies with the growth of knowledge over time. Per the Solow model, productivity increases with technological advancement, as evident in the productivity improvements across sectors in the Philippines. As both models emphasize, the Philippines should continue investing in *technology* to sustain and further their development.

Question 4: Structural Breaks (10 points)

Filter Data for Chosen Country

```r
# Filter data for the chosen country
data_filtered <- data %>%
  filter(wb_countryname == country) %>%
  filter(!is.na(wdi_ny_gdp_pcap_kd_zg)) %>%
  rename(gdp_pc_g = wdi_ny_gdp_pcap_kd_zg)
```

Model: All Years

```r
model1 <- lm(gdp_pc_g ~ year, data = data_filtered)
summary(model1)
```

```
##
## Call:
## lm(formula = gdp_pc_g ~ year, data = data_filtered)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -13.5189  -0.6438   0.7174   1.8595   4.5351
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept) -39.11446   47.08122  -0.831    0.409
## year          0.02062    0.02364   0.872    0.387
##
## Residual standard error: 3.331 on 60 degrees of freedom
## Multiple R-squared:  0.01252,    Adjusted R-squared:  -0.003935
## F-statistic: 0.7609 on 1 and 60 DF,  p-value: 0.3865
```

Identify Structural Breaks

```r
break_test1 <- breakpoints(gdp_pc_g ~ year, data = data_filtered)
summary(break_test1)
```

```
##
##   Optimal (m+1)-segment partition:
##
## Call:
## breakpoints.formula(formula = gdp_pc_g ~ year, data = data_filtered)
##
## Breakpoints at observation number:
##
## m = 1       22
## m = 2    15 25
## m = 3    15 25       51
## m = 4    15 25    42 51
## m = 5    15 25 34 43 52
##
## Corresponding to breakdates:
##
## m = 1                     0.354838709677419
## m = 2    0.241935483870968 0.403225806451613
## m = 3    0.241935483870968 0.403225806451613
## m = 4    0.241935483870968 0.403225806451613                   0.67741935483871
## m = 5    0.241935483870968 0.403225806451613 0.548387096774194 0.693548387096774
##
## m = 1
## m = 2
## m = 3    0.82258064516129
## m = 4    0.82258064516129
## m = 5    0.838709677419355
```

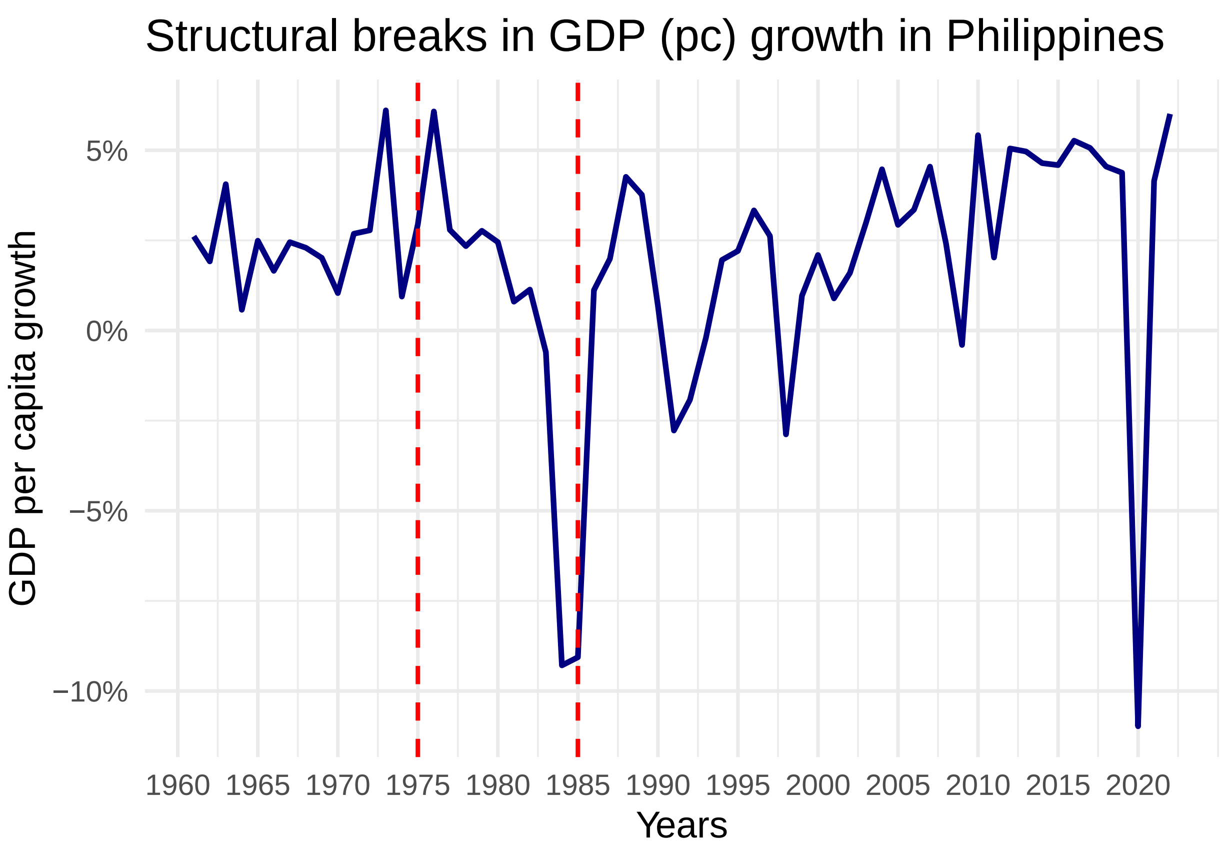
```
##
## Fit:
##
## m   0     1     2     3     4     5
## RSS 665.8 551.5 425.4 397.7 384.3 382.8
## BIC 335.5 336.2 332.5 340.7 351.0 363.1
```

```r
# Extract break years
breaks <- break_test1$breakpoints
break_years <- data_filtered$year[breaks]
break_years
```

```
## [1] 1975 1985
```

Structural Breaks Plot

```r
plot_breaks <- ggplot(data_filtered, aes(x = year, y = gdp_pc_g)) +
  geom_line(color = "navy", linewidth = 1) +
  geom_vline(xintercept = break_years, linetype = "dashed", color = "red",
             linewidth = 0.8) +
  scale_y_continuous(labels = scales::percent_format(scale = 1),
                     breaks = scales::pretty_breaks(n = 6)) +
  scale_x_continuous(breaks = seq(1960, 2021, by = 5)) +
  labs(y = "GDP per capita growth", x = "Years",
       title = paste0("Structural breaks in GDP (pc) growth in ", country)) +
  theme_minimal(base_size = 14)
plot_breaks
```



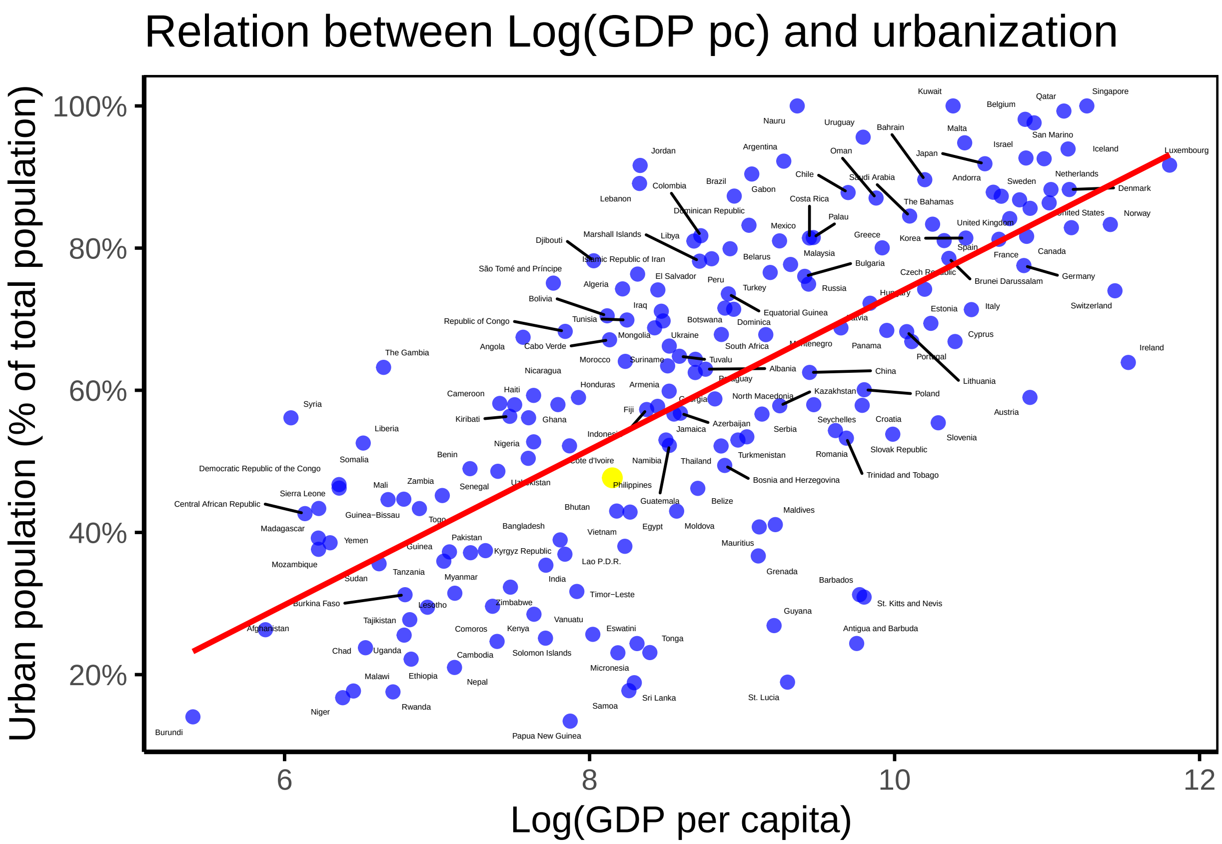
The algorithm identified two major structural breaks in *1975* and *1985*, which aligns with the drops in the GDP time series for the Philippines. Building on my observations from the prior questions, the 1975 break might relate to the Philippines' transition from an agrarian to an industrial economy as well as external economic factors such as the oil price shocks and the global recession of the 1980s. The 1985 break may also relate to internal turmoil from the 1986 People Power Revolution that destabilized the Philippines' political economy.

Although Questions 1 & 2 revealed overall patterns of quality of life improvements from increasing GDP per capita, life expectancy, and education in the past decades, these structural breaks suggest the importance of domestic and external factors in suddenly halting, or advancing, a country's growth trajectory. Viewing technology and the spread of ideas central to development, Solow, Kaldor, and Romer's theories may suggest that global recessions and crises that halt these processes could account for structural breaks in a country's economy. Further analysis could investigate the importance of external and internal factors in smaller countries. To what extent is the Philippines integrated into the global economy? What benefits or heightened susceptibility to global shocks does this bring?

Question 5: Structural transformation comparison with other countries (10 points)

```
# Add log of GDP per Capita
data$log_gdp_pc <- log(data$wdi_ny_gdp_pcap_cd)

# Add dummy column for the chosen country
data <- data %>%
  mutate(dummy_chosen_country = ifelse(wb_countryname == country, 1, 0))

# I am using 1960 as it is the first year for Philippines
data_filtered_1960 <- data %>%
  filter(year == 1960) %>%
  filter(!is.na(qog_wdi_popurb) & !is.na(wdi_ny_gdp_pcap_cd))

data_filtered_2021 <- data %>%
  filter(year == 2021) %>%
  filter(!is.na(qog_wdi_popurb) & !is.na(wdi_ny_gdp_pcap_cd))

#Plot Log(GDP) and Urban Population
ggplot(data_filtered_2021, aes(x = log_gdp_pc, y = qog_wdi_popurb)) +
  geom_point(color = "blue", size = 2, alpha = 0.7) +  # Scatter plot points
  geom_point(
    data = data_filtered_2021[data_filtered_2021$dummy_chosen_country == 1, ],
    aes(x = log_gdp_pc, y = qog_wdi_popurb),
    color = "yellow",
    size = 3
  ) +
  geom_text_repel(aes(label = wb_countryname), size = 1) +  # Repelled country labels
  geom_smooth(method = "lm", color = "red", se = FALSE, linetype = "solid") +
  scale_y_continuous(labels = scales::percent_format(scale = 1),
                     breaks = scales::pretty_breaks(n = 6)) +
  labs(
    title = "Relation between Log(GDP pc) and urbanization",
    y = "Urban population (% of total population)",
    x = "Log(GDP per capita)"
  ) +
```

```
theme_minimal(base_size = 14) +  # Minimal theme for a clean look
theme(
  panel.border = element_rect(color = "black", fill = NA, linewidth = 0.8),
  axis.line = element_line(color = "black", linewidth = 0.8),
  axis.ticks = element_line(color = "black", linewidth = 0.6),
  panel.grid.major = element_blank(),  # Remove major grid lines
  panel.grid.minor = element_blank()   # Remove minor grid lines
)
```



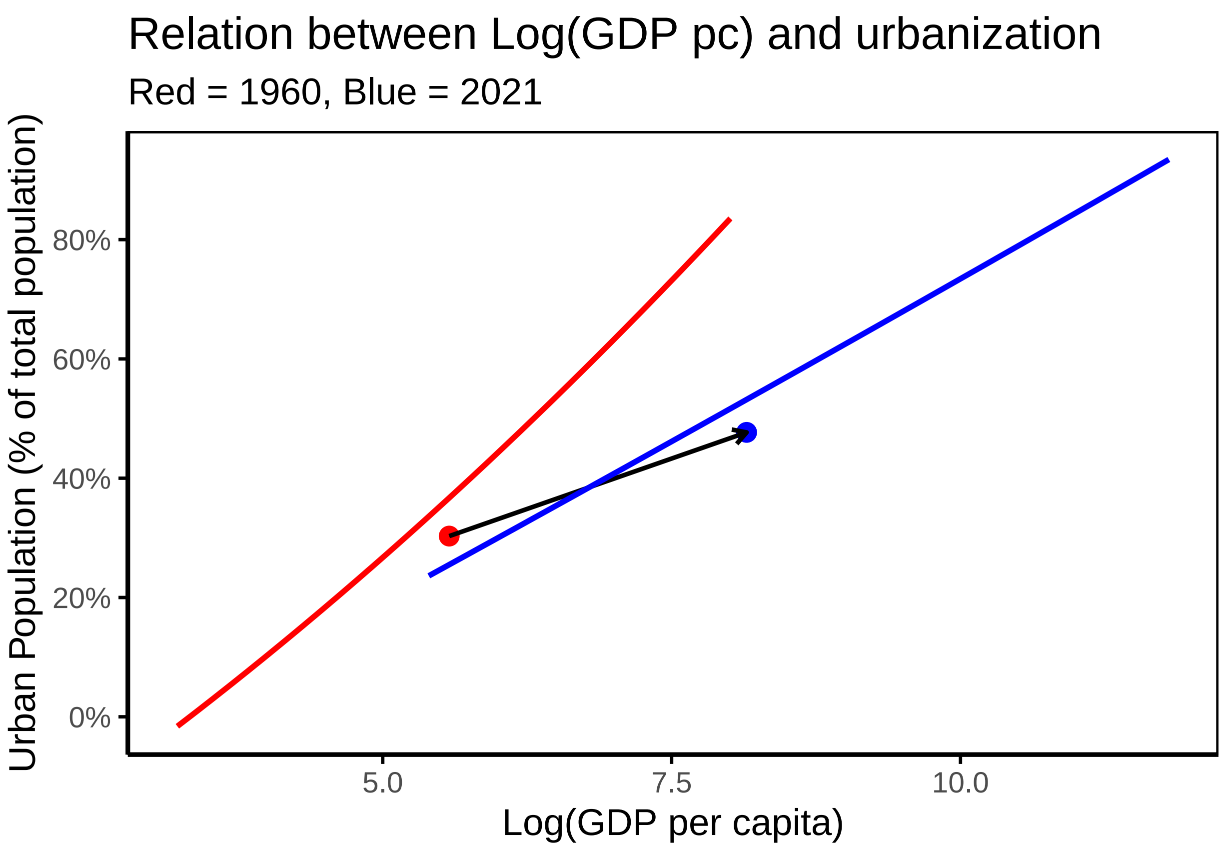
The scatter plot reveals that higher GDP per capita is associated with a higher urban population, which aligns with Romer's idea model expectation of a positive urban population and GDP relationship. However, the Philippines (yellow) falls under the line of best fit, suggesting that it is performing **worse than expected**—its urbanization rate is lower than its GDP per capita predicted. The surrounding clusters of dots suggest the Philippines is not the only country in this position—I added labels to my graph to identify that Bhutan, Belize, and Egypt are also nearby. On either side of the line, Burundi has the greatest extreme of low GDP per capita and urban population while Singapore maintains the highest GDP per capita and urban population.

Building on Romer's idea model, institutional differences may explain such disparities between countries. Especially given its strengthening income, demographics, and industry, the Philippines' lower urbanization level may be due to institutions failing to provide the necessary incentives to continue expanding. Following the discussion during the TA session about omitted variables, the urban population alone is not raising GDP per capita and vice versa—further analysis must interrogate the institutional factors growing GDP per capita but hindering the urban population at the same rate.

Question 6: Comparative Trends in Structural Transformation (10 points)

```r
ggplot() +
  # Highlight chosen country in 1960 (red point)
  geom_point(
    data = data_filtered_1960[data_filtered_1960$dummy_chosen_country == 1, ],
    aes(x = log_gdp_pc, y = qog_wdi_popurb),
    color = "red",
    size = 3
  ) +
  # Highlight chosen country in 2021 (blue point)
  geom_point(
    data = data_filtered_2021[data_filtered_2021$dummy_chosen_country == 1, ],
    aes(x = log_gdp_pc, y = qog_wdi_popurb),
    color = "blue",
    size = 3
  ) +
  # Add arrow connecting 1960 and 2021 points
  geom_segment(
    data = merge(
      data_filtered_1960[data_filtered_1960$dummy_chosen_country == 1, ],
      data_filtered_2021[data_filtered_2021$dummy_chosen_country == 1, ],
      by = "dummy_chosen_country"
    ),
    aes(
      x = log_gdp_pc.x, y = qog_wdi_popurb.x,
      xend = log_gdp_pc.y, yend = qog_wdi_popurb.y
    ),
    arrow = arrow(length = unit(0.2, "cm")),
    color = "black",
    linewidth = 0.8
  ) +
  # Polynomial regression for 1988 (red curve)
  geom_smooth(
    data = data_filtered_1960,
    aes(x = log_gdp_pc, y = qog_wdi_popurb),
    method = "lm",
    formula = y ~ poly(x, 2),
    color = "red",
    se = FALSE,
    linetype = "solid"
  ) +
  # Polynomial regression for 2021 (blue curve)
  geom_smooth(
    data = data_filtered_2021,
    aes(x = log_gdp_pc, y = qog_wdi_popurb),
    method = "lm",
    formula = y ~ poly(x, 2),
    color = "blue",
    se = FALSE,
    linetype = "solid"
  ) +
```

```r
scale_y_continuous(labels = scales::percent_format(scale = 1),
                   breaks = scales::pretty_breaks(n = 6)) +
# Axis labels
labs(
  title = "Relation between Log(GDP pc) and urbanization",
  subtitle = "Red = 1960, Blue = 2021",
  y = "Urban Population (% of total population)",
  x = "Log(GDP per capita)"
) +
# Minimal theme
theme_minimal(base_size = 14) +
theme(
  panel.border = element_rect(color = "black", fill = NA, linewidth = 0.8),
  axis.line = element_line(color = "black", linewidth = 0.8),
  axis.ticks = element_line(color = "black", linewidth = 0.6),
  panel.grid.major = element_blank(),
  panel.grid.minor = element_blank()
)
```



In 1960 (the first year GDP per capita data was available for the Philippines), the Philippines was below the urban population-GDP per capita curve, suggesting its urbanization levels were *lower* than expected. Its position in the middle of the red line further suggests its average performance globally.

In 2021, the Philippines was similarly *below* the curve, suggesting its urban population levels were still *lower* than expected for its GDP per capita. However, its position relative to other countries also deteriorated, as the Philippines falls on the lower half of the blue line, part of the bottom half of countries for urban population and GDP per capita.

Compared to the global trends, these curves reveal the Philippines' urbanization rate lagged behind other countries, its gap between actual and predicted urban population widening between 1960 and 2021. While its GDP per capita improved, the Philippines underperformed in urban population growth. The Malthusian model expects that greater population growth increases agricultural productivity—so perhaps the agricultural sectors' persistence in the Philippines continues to encourage rural development. The Kaldor and Romer theories explain lower urbanization levels due to a weak industrial sector and infrastructural development—while the services sector relatively expanded within the Philippines, the country may still be lagging on the global stage, revealing an opportunity to continue growing through investment in these areas.

Question 7: Conclusion (30 points)

Through the 20th century, the Philippines showed a growth trajectory characterized by growing GDP and population, a structural shift from agriculture towards services, and improving productivity. However, I also noted the country's susceptibility to external shocks that disrupted development and its urban population levels lower than expected by GDP per capita.

To analyze these patterns, economic theory from Malthus to Romer identifies development opportunities in the Philippines. The Philippines's decreasing death rates precede declining fertility to suggest a currently open demographic window. With an expanding working population, the Philippines can increase output, which in turn increases development spurring growth into the future. Additionally, the increase in education enrollment reveals greater investment in the population and production. Lastly, the shift from agriculture to services suggests an exchange of know-how and ideas, per the Romer and Hausmann lectures, capable of improving the Philippines' position in the product space and global exchange. Further evidence from the [Harvard Growth Lab] (https://atlas.hks.harvard.edu/explore/treemap?exporter=country-608) indicates that the Philippines' product complexity is improving rapidly and expects the country to continue diversifying.

While these observations greatly enhanced my understanding of its development, I still have questions about the Philippines' (1) institutions; (2) local-level disparities; and (3) global interconnectedness. In supplementary research, Villanueva noted the importance of recent private property, contract rights, and rule of law as well as investment in infrastructure and education to increase growth in the Philippines in recent decades ([Villanueva, 14] (https://research.ebsco.com/linkprocessor/plink?id=d24e4ee1-2a14-34e4-be53-7d55a5aa01f5)). Coupled with our understanding of Kaldor and Romer in class, to what extent have these institutions promoted urban and GDP per capita growth—and why might urban growth still fall behind? What policy interventions are necessary to continue promoting technological and human capital investment?

Additionally, examining the Philippines on the domestic scale (comparing current and past data within the Philippines) and on a global scale (comparing the Philippines' urban population and GDP per capita levels with other countries) revealed different perspectives on its growth—the Philippines changed a lot compared to itself decades ago, but changed less than other countries globally. This point interests me in further examining the local level in the Philippines—what inequalities persist between regions? Especially given the lower urban population levels, how do economic conditions differ between rural and urban regions? What is the significance of remittances inflow and foreign investment?

Additionally, what differs between the Philippines and other countries around the world in achieving urban population levels or GDP per capita growth? What accounts for this global-level inequality where the Philippines appears to lag? Comparing the Philippines to other countries on the global level suggests there is something unusual about the Philippines' comparatively low GDP per capita and lower-than-expected urban population levels. How does the rate of the Philippines' seemingly rapid GDP and population growth compare to other countries? What advantages does its open demographic opportunity present compared to other countries (such as Germany experiencing a demographic dependency crisis)?

This problem set exposed me to key data, patterns, and opportunities for development in the Philippines and I hope to continue investigating these trends through the sequential problem sets and the capstone project, positing the following growth question: **Given its consistently growing GDP and population, what factors hinder further educational attainment, urbanization, and structural economic transition in the Philippines? What institutional interventions may enable the Philippines to leverage its**

demographic window and growth opportunities to develop further—and increase its resilience to external shocks?