Code Summary

Group C9

2025-12-10

Loading Required Libraries

```r
library(tidyverse)
library(dplyr)
library(ggplot2)
library(purrr)
library(patchwork)
```

Data Import and Preprocessing

Importing Data Sets

```r
Continent_Classification <- read.csv("data sets/continents-according-to-our-world-in-data.csv")
GDP_per_capita <- read.csv("data sets/gdp-per-capita-worldbank.csv")
LDC <- read.csv("data sets/LDC.csv")
World_Population <- read.csv("data sets/world population.csv")
NEET <- read.csv("data sets/youth-not-in-education-employment-training.csv")
```

Rename Columns

```r
colnames(GDP_per_capita) <- c("Country", "Code", "Year", "GDP_Per_Capita")
colnames(Continent_Classification) <- c("Country", "Code", "Year", "Continent")
colnames(World_Population) <- c("Country", "Code", "Year", "Population")
```

Convert Columns to Correct Types

```r
GDP_per_capita <- GDP_per_capita %>%
  mutate(
    Country = as.character(Country),
    Code = as.character(Code),
    Year = as.integer(Year),
    GDP_Per_Capita = map_dbl(GDP_Per_Capita, as.numeric)
  )
```

```
Continent_Classification <- Continent_Classification %>%
  mutate(
    Country = as.character(Country),
    Code = as.character(Code),
    Year = as.integer(Year),
    Continent = as.character(Continent)
  )

World_Population <- World_Population %>%
  mutate(
    Country = as.character(Country),
    Code = as.character(Code),
    Year = as.integer(Year),
    Population = as.numeric(Population)
  )
```

Data Combination and Processing

Combine Data Sets

```
GDP_Continent_Population_Combined <- GDP_per_capita %>%
  left_join(Continent_Classification %>% select(Code, Continent), by = "Code") %>%
  left_join(World_Population %>% select(Code, Year, Population), by = c("Code", "Year"))

GDP_Continent_Population_Combined <- GDP_Continent_Population_Combined %>%
  drop_na()
```

LDC Classification

```
GDP_Continent_Population_Combined <- GDP_Continent_Population_Combined %>%
  mutate(LDC_Status = Country %in% LDC$Country & Year %in% LDC$Year)
```

Add Growth Rate

```
GDP_Continent_Population_Combined <- GDP_Continent_Population_Combined %>%
  arrange(Country, Year) %>%
  group_by(Country) %>%
  mutate(GDP_growth_rate = (GDP_Per_Capita - lag(GDP_Per_Capita)) / lag(GDP_Per_Capita) * 100) %>%
  ungroup()
```

Determining Fair Growth Targets

Calculate Continental Means (2009-2021)

```r
# Europe
europemean <- GDP_Continent_Population_Combined %>%
  filter(Continent == "Europe") %>%
  filter(Year >= 2009 & Year <= 2021) %>%
  summarise(mean_gdp_pc = mean(GDP_Per_Capita, na.rm = TRUE))

# Asia
asiamean <- GDP_Continent_Population_Combined %>%
  filter(Continent == "Asia") %>%
  filter(Year >= 2009 & Year <= 2021) %>%
  summarise(mean_gdp_pc = mean(GDP_Per_Capita, na.rm = TRUE))

# Africa
africamean <- GDP_Continent_Population_Combined %>%
  filter(Continent == "Africa") %>%
  filter(Year >= 2009 & Year <= 2021) %>%
  summarise(mean_gdp_pc = mean(GDP_Per_Capita, na.rm = TRUE))

# North America
northamericamean <- GDP_Continent_Population_Combined %>%
  filter(Continent == "North America") %>%
  filter(Year >= 2009 & Year <= 2021) %>%
  summarise(mean_gdp_pc = mean(GDP_Per_Capita, na.rm = TRUE))

# South America
southamericamean <- GDP_Continent_Population_Combined %>%
  filter(Continent == "South America") %>%
  filter(Year >= 2009 & Year <= 2021) %>%
  summarise(mean_gdp_pc = mean(GDP_Per_Capita, na.rm = TRUE))

# Oceania
oceaniamean <- GDP_Continent_Population_Combined %>%
  filter(Continent == "Oceania") %>%
  filter(Year >= 2009 & Year <= 2021) %>%
  summarise(mean_gdp_pc = mean(GDP_Per_Capita, na.rm = TRUE))
```

Combine Continental Means

```r
means_df <- data.frame(
  Continent = c("Europe", "Asia", "Africa", "North America", "South America", "Oceania"),
  mean_gdp_pc = c(
    europemean$mean_gdp_pc,
    asiamean$mean_gdp_pc,
    africamean$mean_gdp_pc,
    northamericamean$mean_gdp_pc,
    southamericamean$mean_gdp_pc,
    oceaniamean$mean_gdp_pc
  )
)
```

Target Growth Analysis

```
TargetGrowth <- GDP_Continent_Population_Combined %>%
  filter(GDP_growth_rate > -10, GDP_growth_rate < 10, GDP_Per_Capita < 60000) %>%
  ggplot(aes(x = GDP_Per_Capita, y = GDP_growth_rate)) +
  geom_point(alpha = 0.3, size = 0.5) +
  geom_smooth(method = "lm", se = FALSE, color = "black") +
  geom_vline(data = means_df,
             aes(xintercept = mean_gdp_pc, color = Continent),
             linetype = "dashed", size = 0.7) +
  labs(title = "GDP per Capita vs GDP Growth Rate with Continent Means",
       x = "GDP per Capita", y = "GDP Growth Rate (%)",
       color = "Continent") +
  theme_minimal() +
  theme(plot.title = element_text(hjust = 0.5),
        legend.position = "bottom")

print(TargetGrowth)
```

Predicted Growth Rates

```
# Fit the linear model on filtered data
plot_df <- GDP_Continent_Population_Combined %>%
  filter(GDP_growth_rate > -10,
         GDP_growth_rate < 10,
         GDP_Per_Capita < 60000)

m <- lm(GDP_growth_rate ~ GDP_Per_Capita, data = plot_df)

# Predict growth at each mean GDP per capita
means_df$predicted_growth_at_mean <- predict(m, newdata = data.frame(GDP_Per_Capita = means_df$mean_gdp.

# Output results
result <- means_df %>%
  mutate(
    mean_gdp_pc = round(mean_gdp_pc, 2),
    predicted_growth_at_mean = round(predicted_growth_at_mean, 3)
  )

print(result)
```

Continental Growth Analysis (2009-2021)

Filter Years and Compute Continental Growth

```
# Filter the years from 2009-2021
GDP_Continent_Population_Combined <- GDP_Continent_Population_Combined %>%
  filter(Year >= 2009 & Year <= 2021)
```

```r
# Compute Continent Growth
continent_growth <- GDP_Continent_Population_Combined %>%
  group_by(Continent, Year) %>%
  summarise(
    weighted_gdp_pc = sum(GDP_Per_Capita * Population, na.rm = TRUE) / sum(Population, na.rm = TRUE),
    average_gdp_pc  = mean(GDP_Per_Capita, na.rm = TRUE),.groups = "drop"
  ) %>%
  arrange(Continent, Year) %>%
  group_by(Continent) %>%
  mutate(weighted_growth = (weighted_gdp_pc / lag(weighted_gdp_pc) - 1) * 100) %>%
  mutate(average_growth = (average_gdp_pc / lag(average_gdp_pc) - 1) * 100) %>%
  ungroup()
```

Check for Missing Data

```r
na_counts <- GDP_Continent_Population_Combined %>%
  group_by(Continent, Year) %>%
  summarise(
    na_GDP_growth_rate = sum(is.na(GDP_growth_rate)),.groups = "drop"
  ) %>%
  arrange(Continent, Year)

print("NA counts by continent and year:")
print(na_counts)
```

Continental Growth Visualization

Individual Continental Plots

```r
# Africa (threshold: 3%)
Africa_growth <- ggplot(subset(continent_growth, Continent == "Africa"), aes(x = Year)) +
  geom_line(aes(y = weighted_growth, color = "Weighted")) +
  geom_point(aes(y = weighted_growth, color = "Weighted"), size = 0.75) +
  geom_line(aes(y = average_growth, color = "Average")) +
  geom_point(aes(y = average_growth, color = "Average"), size = 0.75) +
  geom_hline(yintercept = 3, linetype = "dashed", color = "red", linewidth = 0.5) +
  labs(title = "Africa", y = "Growth Rate (%)", x = NULL) +
  theme_minimal()

# Asia (threshold: 2%)
Asia_growth <- ggplot(subset(continent_growth, Continent == "Asia"), aes(x = Year)) +
  geom_line(aes(y = weighted_growth, color = "Weighted")) +
  geom_point(aes(y = weighted_growth, color = "Weighted"), size = 0.75) +
  geom_line(aes(y = average_growth, color = "Average")) +
  geom_point(aes(y = average_growth, color = "Average"), size = 0.75) +
  geom_hline(yintercept = 2, linetype = "dashed", color = "red", linewidth = 0.5) +
  labs(title = "Asia", y = "Growth Rate (%)", x = NULL) +
  theme_minimal()
```

```r
# Europe (threshold: 1.6%)
Europe_growth <- ggplot(subset(continent_growth, Continent == "Europe"), aes(x = Year)) +
  geom_line(aes(y = weighted_growth, color = "Weighted")) +
  geom_point(aes(y = weighted_growth, color = "Weighted"), size = 0.75) +
  geom_line(aes(y = average_growth, color = "Average")) +
  geom_point(aes(y = average_growth, color = "Average"), size = 0.75) +
  geom_hline(yintercept = 1.6, linetype = "dashed", color = "red", linewidth = 0.5) +
  labs(title = "Europe", y = "Growth Rate (%)", x = NULL) +
  theme_minimal()

# North America (threshold: 1.8%)
NorthAmerica_growth <- ggplot(subset(continent_growth, Continent == "North America"), aes(x = Year)) +
  geom_line(aes(y = weighted_growth, color = "Weighted")) +
  geom_point(aes(y = weighted_growth, color = "Weighted"), size = 0.75) +
  geom_line(aes(y = average_growth, color = "Average")) +
  geom_point(aes(y = average_growth, color = "Average"), size = 0.75) +
  geom_hline(yintercept = 1.8, linetype = "dashed", color = "red", linewidth = 0.5) +
  labs(title = "North America", y = "Growth Rate (%)", x = NULL) +
  theme_minimal()

# South America (threshold: 2%)
SouthAmerica_growth <- ggplot(subset(continent_growth, Continent == "South America"), aes(x = Year)) +
  geom_line(aes(y = weighted_growth, color = "Weighted")) +
  geom_point(aes(y = weighted_growth, color = "Weighted"), size = 0.75) +
  geom_line(aes(y = average_growth, color = "Average")) +
  geom_point(aes(y = average_growth, color = "Average"), size = 0.75) +
  geom_hline(yintercept = 2, linetype = "dashed", color = "red", linewidth = 0.5) +
  labs(title = "South America", y = "Growth Rate (%)", x = NULL) +
  theme_minimal()

# Oceania (threshold: 2%)
Oceania_growth <- ggplot(subset(continent_growth, Continent == "Oceania"), aes(x = Year)) +
  geom_line(aes(y = weighted_growth, color = "Weighted")) +
  geom_point(aes(y = weighted_growth, color = "Weighted"), size = 0.75) +
  geom_line(aes(y = average_growth, color = "Average")) +
  geom_point(aes(y = average_growth, color = "Average"), size = 0.75) +
  geom_hline(yintercept = 2, linetype = "dashed", color = "red", linewidth = 0.5) +
  labs(title = "Oceania", y = "Growth Rate (%)", x = NULL) +
  theme_minimal()
```

Combined Continental Growth Plot

```r
# Combine all six with one shared legend
Growth_Combined_Plot <- (Africa_growth | Asia_growth | Europe_growth) /
  (NorthAmerica_growth | SouthAmerica_growth | Oceania_growth) +
  plot_annotation(title = "GDP per Capita Growth by Continent",
                  theme = theme(plot.title = element_text(hjust = 0.5))) +
  plot_layout(guides = "collect") &
  theme(legend.position = "bottom")

print(Growth_Combined_Plot)
```

LDC Analysis

Share of LDCs Meeting the 7% Target

```r
ldc_target_share <- GDP_Continent_Population_Combined %>%
  filter(LDC_Status) %>%
  group_by(Continent, Year) %>%
  summarise(
    n_ldc         = n_distinct(Country),
    n_meet        = n_distinct(Country[GDP_growth_rate >= 7]),
    share_meeting = if_else(n_ldc > 0, 100 * n_meet / n_ldc, NA_real_),
    .groups = "drop"
  )
```

```r
LDCShare <- ggplot(ldc_target_share, aes(x = Year, y = share_meeting, color = Continent)) +
  geom_line(size = 0.5) +
  geom_point(size = 0.7) +
  labs(
    title = "Share of LDCs Meeting the UN  7% GDP Growth Target by Continents",
    x = "Year",
    y = "% of LDCs meeting target",
    color = "Continent"
  ) +
  theme_minimal(base_size = 12) +
  theme(
    legend.position = "bottom",
    plot.title = element_text(hjust = 0.5)
  )

print(LDCShare)
```

LDC GDP Per Capita Growth Rates

```r
ldc_continent_growth <- GDP_Continent_Population_Combined %>%
  filter(LDC_Status) %>%
  group_by(Continent, Year) %>%
  summarise(
    weighted_gdp_pc_ldc = sum(GDP_Per_Capita * Population, na.rm = TRUE) /
      sum(Population, na.rm = TRUE),
    average_gdp_pc_ldc  = mean(GDP_Per_Capita, na.rm = TRUE),
    .groups = "drop"
  ) %>%
  arrange(Continent, Year) %>%
  group_by(Continent) %>%
  mutate(
    weighted_growth_ldc = (weighted_gdp_pc_ldc / lag(weighted_gdp_pc_ldc) - 1) * 100,
    average_growth_ldc  = (average_gdp_pc_ldc  / lag(average_gdp_pc_ldc)  - 1) * 100
  ) %>%
  ungroup()
```

Africa LDC Growth

```
Africa_LDC <- ggplot(subset(ldc_continent_growth, Continent == "Africa"), aes(x = Year)) +
  geom_line(aes(y = weighted_growth_ldc, color = "Weighted"), linewidth = 1) +
  geom_point(aes(y = weighted_growth_ldc, color = "Weighted"), size = 1.5) +
  geom_line(aes(y = average_growth_ldc,  color = "Average"), linewidth = 1, linetype = "dashed") +
  geom_point(aes(y = average_growth_ldc,  color = "Average"), size = 1.5) +
  scale_color_manual(values = c("Weighted" = "#1f77b4", "Average" = "#ff7f0e")) +
  labs(title = "Africa LDC GDP per Capita Growth", y = "Growth (%)", x = NULL, color = "Measure") +
  theme_minimal() +
  theme(
    legend.position = "bottom",
    panel.grid.minor = element_blank(),
    plot.title = element_text(hjust = 0.5)
  )

print(Africa_LDC)
```

Asia LDC Growth

```
ldc_continent_growth_Asia <- ldc_continent_growth %>%
  filter(Continent == "Asia", Year %in% 1996:2021)

Asia_LDC <- ggplot(ldc_continent_growth_Asia, aes(x = Year)) +
  geom_line(aes(y = weighted_growth_ldc, color = "Weighted"), linewidth = 1) +
  geom_point(aes(y = weighted_growth_ldc, color = "Weighted"), size = 1.5) +
  geom_line(aes(y = average_growth_ldc,  color = "Average"), linewidth = 1, linetype = "dashed") +
  geom_point(aes(y = average_growth_ldc,  color = "Average"), size = 1.5) +
  scale_color_manual(values = c("Weighted" = "#1f77b4", "Average" = "#ff7f0e")) +
  labs(title = "Asia LDC GDP per Capita Growth", y = "Growth (%)", x = NULL, color = "Measure") +
  theme_minimal() +
  theme(
    legend.position = "bottom",
    panel.grid.minor = element_blank(),
    plot.title = element_text(hjust = 0.5)
  )

print(Asia_LDC)
```

Oceania LDC Growth

```
Oceania_LDC <- ggplot(subset(ldc_continent_growth, Continent == "Oceania"), aes(x = Year)) +
  geom_line(aes(y = weighted_growth_ldc, color = "Weighted"), linewidth = 1) +
  geom_point(aes(y = weighted_growth_ldc, color = "Weighted"), size = 1.5) +
  geom_line(aes(y = average_growth_ldc,  color = "Average"), linewidth = 1, linetype = "dashed") +
  geom_point(aes(y = average_growth_ldc,  color = "Average"), size = 1.5) +
  scale_color_manual(values = c("Weighted" = "#1f77b4", "Average" = "#ff7f0e")) +
  labs(title = "Oceania LDC GDP per Capita Growth", y = "Growth (%)", x = NULL, color = "Measure") +
  theme_minimal() +
```

```r
  theme(
    legend.position = "bottom",
    panel.grid.minor = element_blank(),
    plot.title = element_text(hjust = 0.5)
  )

print(Oceania_LDC)
```

Distribution of LDC Growth Rates

```r
# Common y-limits across the three continents
ylim_range <- ldc_continent_growth %>%
  filter(Continent %in% c("Africa", "Asia", "Oceania")) %>%
  pull(weighted_growth_ldc) %>%
  range(na.rm = TRUE)

# Africa
Africa_box <- ggplot(subset(ldc_continent_growth, Continent == "Africa" & !is.na(weighted_growth_ldc)),
                     aes(x = "Africa", y = weighted_growth_ldc)) +
  geom_boxplot(fill = "#1f77b4") +
  coord_cartesian(ylim = ylim_range) +
  labs(title = "Africa", x = NULL, y = "Growth Rate (%)") +
  theme_minimal() +
  theme(plot.title = element_text(hjust = 0.5))

# Asia
Asia_box <- ggplot(subset(ldc_continent_growth, Continent == "Asia" & !is.na(weighted_growth_ldc)),
                   aes(x = "Asia", y = weighted_growth_ldc)) +
  geom_boxplot(fill = "#1f77b4") +
  coord_cartesian(ylim = ylim_range) +
  labs(title = "Asia", x = NULL, y = NULL) +
  theme_minimal() +
  theme(plot.title = element_text(hjust = 0.5))

# Oceania
Oceania_box <- ggplot(subset(ldc_continent_growth, Continent == "Oceania" & !is.na(weighted_growth_ldc)
                      aes(x = "Oceania", y = weighted_growth_ldc)) +
  geom_boxplot(fill = "#1f77b4") +
  coord_cartesian(ylim = ylim_range) +
  labs(title = "Oceania", x = NULL, y = NULL) +
  theme_minimal() +
  theme(plot.title = element_text(hjust = 0.5))

# Combine into one row with a shared title
LDC_Distribution_Combined <- (Africa_box | Asia_box | Oceania_box) +
  plot_annotation(
    title = "Distribution of Weighted LDC GDP per Capita Growth",
    theme = theme(plot.title = element_text(hjust = 0.5))
  )

print(LDC_Distribution_Combined)
```

Regional Comparisons

Caribbean vs Rest of North America

```r
GDP_Continent_Population_Combined_NorthAmerica <- GDP_Continent_Population_Combined %>%
  filter(Continent == "North America")

caribbean <- c("ATG","BHS","BRB","CUB","DMA","DOM","GRD","HTI","JAM","KNA","LCA","VCT","TTO")

CaribbeanNorthAmerica <- GDP_Continent_Population_Combined_NorthAmerica %>%
  mutate(Group = if_else(Code %in% caribbean, "Caribbean", "Rest of North America")) %>%
  group_by(Year, Group) %>%
  summarise(
    weighted_gdp_pc = sum(GDP_Per_Capita * Population, na.rm = TRUE) /
      sum(Population, na.rm = TRUE),
    .groups = "drop"
  ) %>%
  arrange(Group, Year) %>%
  group_by(Group) %>%
  mutate(
    weighted_growth = (weighted_gdp_pc / lag(weighted_gdp_pc) - 1) * 100
  ) %>%
  ungroup()
```

```r
CaribbeanNorthAmericaGraph <- ggplot(CaribbeanNorthAmerica, aes(x = Year, y = weighted_growth, color = (
  geom_line(size = 0.75) +
  geom_point(size = 1.5, alpha = 0.8) +
  scale_color_manual(values = c("Caribbean" = "#D62728",
                                "Rest of North America" = "#7F7F7F")) +
  labs(
    title = "Caribbean vs Rest of North America",
    x = "Year",
    y = "GDP per capita growth (%)",
    color = NULL
  ) +
  theme_minimal(base_size = 12) +
  theme(
    legend.position = "bottom",
    plot.title = element_text(hjust = 0.5),
    plot.subtitle = element_text(hjust = 0.5)
  )

print(CaribbeanNorthAmericaGraph)
```

Africa: Diversified vs Commodity-Dependent Economies

```r
diversified_codes <- c(
  "BWA","CPV","DJI","EGY","SWZ","GHA","KEN","MUS","MAR","NAM",
  "RWA","SEN","SYC","ZAF","TZA","TUN"
)
```

```r
commodity_codes <- c(
  "DZA","AGO","BEN","BFA","BDI","CMR","CAF","TCD","COM","COG","CIV","COD",
  "GNQ","ETH","GAB","GMB","GIN","GNB","LSO","LBR","LBY","MDG","MWI","MLI",
  "MRT","MOZ","NER","NGA","STP","SLE","SOM","SDN","TGO","UGA","ZMB","ZWE"
)

DiversifiedCommodity <- GDP_Continent_Population_Combined %>%
  mutate(Group = case_when(
    Code %in% diversified_codes ~ "Diversified",
    Code %in% commodity_codes   ~ "Commodity-dependent",
    TRUE                        ~ NA_character_
  )) %>%
  filter(!is.na(Group)) %>%
  group_by(Year, Group) %>%
  summarise(
    weighted_gdp_pc = sum(GDP_Per_Capita * Population, na.rm = TRUE) /
      sum(Population, na.rm = TRUE),
    .groups = "drop"
  ) %>%
  arrange(Group, Year) %>%
  group_by(Group) %>%
  mutate(
    weighted_growth = (weighted_gdp_pc / lag(weighted_gdp_pc) - 1) * 100
  ) %>%
  ungroup()
```

```r
DiversifiedCommodityGraph <- ggplot(DiversifiedCommodity,
                                    aes(x = Year, y = weighted_growth, color = Group)) +
  geom_line(size = 0.75) +
  geom_point(size = 1.5, alpha = 0.8) +
  scale_color_manual(values = c("Diversified" = "#1f77b4", "Commodity-dependent" = "#D62728")) +
  labs(
    title = "Diversified vs Commodity-dependent",
    x = "Year",
    y = "GDP per capita growth (%)",
    color = NULL
  ) +
  theme_minimal(base_size = 12) +
  theme(
    legend.position = "bottom",
    plot.title = element_text(hjust = 0.5)
  )

print(DiversifiedCommodityGraph)
```

Haiti vs North America

```r
haiti_growth <- GDP_Continent_Population_Combined %>%
  filter(Code == "HTI") %>%
  arrange(Year) %>%
  group_by(Country) %>%
```

```
  mutate(weighted_gdp_pc = GDP_Per_Capita,
         weighted_growth  = (weighted_gdp_pc / lag(weighted_gdp_pc) - 1) * 100) %>%
  ungroup() %>%
  select(Year, Country, weighted_growth) %>%
  mutate(Group = "Haiti")

na_growth <- GDP_Continent_Population_Combined %>%
  filter(Continent == "North America") %>%
  group_by(Year) %>%
  summarise(
    weighted_gdp_pc = sum(GDP_Per_Capita * Population, na.rm = TRUE) /
      sum(Population, na.rm = TRUE),
    .groups = "drop"
  ) %>%
  arrange(Year) %>%
  mutate(weighted_growth = (weighted_gdp_pc / lag(weighted_gdp_pc) - 1) * 100,
         Group = "North America (weighted avg)") %>%
  select(Year, Group, weighted_growth)

HaitiNorthAmerica <- haiti_growth %>%
  select(Year, Group, weighted_growth) %>%
  bind_rows(na_growth) %>%
  filter(!is.na(weighted_growth))
```

```
HaitiNorthAmericaPlot <- ggplot(HaitiNorthAmerica, aes(x = Year, y = weighted_growth, color = Group)) +
  geom_line(size = 0.75) +
  geom_point(size = 1.5, alpha = 0.9) +
  geom_hline(yintercept = 7, linetype = "dashed", color = "black", alpha = 0.6) +
  scale_color_manual(values = c("Haiti" = "#D62728",
                                "North America" = "#1f77b4")) +
  labs(title = "Haiti vs North America",
       x = "Year", y = "GDP per capita growth (%)", color = NULL) +
  theme_minimal(base_size = 12) +
  theme(legend.position = "bottom",
        plot.title = element_text(hjust = 0.5)) +
  annotate("text", x = min(HaitiNorthAmerica$Year, na.rm = TRUE), y = 7,
           label = "7% target", hjust = -0.1, vjust = -0.5, size = 3.2, color = "black")

print(HaitiNorthAmericaPlot)
```

Oceania: Australia & New Zealand vs Rest of Pacific Islands

```
GDP_Continent_Population_Combined_Oceania <- GDP_Continent_Population_Combined %>%
  filter(Continent == "Oceania")

anz_codes <- c("AUS", "NZL")

Oceania_ANZ_vs_Rest <- GDP_Continent_Population_Combined_Oceania %>%
  mutate(Group = if_else(Code %in% anz_codes, "Australia & New Zealand", "Rest of Pacific Islands")) %>%
  group_by(Year, Group) %>%
  summarise(
```

```
    weighted_gdp_pc = sum(GDP_Per_Capita * Population, na.rm = TRUE) /
      sum(Population, na.rm = TRUE),
    .groups = "drop"
  ) %>%
  arrange(Group, Year) %>%
  group_by(Group) %>%
  mutate(
    weighted_growth = (weighted_gdp_pc / lag(weighted_gdp_pc) - 1) * 100
  ) %>%
  ungroup()
```

```
Oceania_ANZ_vs_Rest_Graph <- ggplot(Oceania_ANZ_vs_Rest, aes(x = Year, y = weighted_growth, color = Grou
  geom_line(size = 0.75) +
  geom_point(size = 1.5, alpha = 0.85) +
  scale_color_manual(values = c("Australia & New Zealand" = "#1f77b4",
                                "Rest of Pacific Islands" = "#D62728")) +
  labs(
    title = "Australia & New Zealand vs Rest of Pacific Islands",
    x = "Year",
    y = "GDP per capita growth (%)",
    color = NULL
  ) +
  theme_minimal(base_size = 12) +
  theme(
    legend.position = "bottom",
    plot.title = element_text(hjust = 0.5)
  )

print(Oceania_ANZ_vs_Rest_Graph)
```

NEET and GDP Growth Analysis

Load NEET Data

```
GDP_NEET <- read.csv("data sets/gdp-per-capita-worldbank.csv")
NEET <- read.csv("data sets/youth-not-in-education-employment-training.csv")
Continents_NEET <- read.csv("data sets/continents-according-to-our-world-in-data.csv")
```

Africa: NEET vs GDP Growth

```
# Filter the data of NEET in Africa in 2010-2021
youth_NEET_Africa <- NEET %>%
  filter(Year >= 2010, Year <= 2020) %>%
  inner_join(Continents_NEET, Year, by = "Entity") %>%
  select(-Year.y, -Code.y) %>%
  filter(Continent == "Africa")

# Filter the data of GDP in Africa in 2010-2021
GDP_Africa <- GDP_NEET %>%
```

```r
  filter(Year >= 2009, Year <= 2020) %>%
  inner_join(Continents_NEET, Year, by = "Entity") %>%
  select(-Year.y, -Code.y) %>%
  filter(Continent == "Africa")

# Calculate GDP growth rate
GDP_Growth_Africa <- GDP_Africa %>%
  group_by(Year.x) %>%
  summarise(Total_GDP = sum(GDP.per.capita..PPP..constant.2017.international...)) %>%
  mutate(GDP_Growth = ((Total_GDP - lag(Total_GDP)) / lag(Total_GDP)) * 100)

# Calculate mean NEET
mean_youth_NEET_Africa <- youth_NEET_Africa %>%
  group_by(Year.x) %>%
  summarise(mean_NEET = mean(Share.of.youth.not.in.education..employment.or.training..total....of.youth

# Combine the two data set for plot
GDP_Growth_NEET_Africa <- inner_join(GDP_Growth_Africa, mean_youth_NEET_Africa, by = "Year.x")


# Plot NEET and GDP growth
Africa_NEET_GDP <- ggplot(GDP_Growth_NEET_Africa, aes(x = Year.x)) +
  geom_point(aes(y = mean_NEET, colour = "Mean NEET"), size = 1.8) +
  geom_line(aes(y = mean_NEET, colour = "Mean NEET"), size = 0.8) +
  geom_point(aes(y = GDP_Growth, colour = "GDP Growth"), size = 1.8) +
  geom_line(aes(y = GDP_Growth, colour = "GDP Growth"), size = 0.8) +
  scale_x_continuous(breaks = 2010:2020) +
  labs(title = "GDP Growth per Capita VS Mean Youth NEET in Africa (2010-2020)",
       x = "Year",
       y = "GDP Growth per Capita and Mean Youth NEET")

print(Africa_NEET_GDP)
```

Asia: NEET vs GDP Growth

```r
# Filter the data of NEET in Asia in 2010-2021
youth_NEET_Asia <- NEET %>%
  filter(Year >= 2010, Year <= 2020) %>%
  inner_join(Continents_NEET, Year, by = "Entity") %>%
  select(-Year.y, -Code.y) %>%
  filter(Continent == "Asia")

# Filter the data of GDP in Asia in 2010-2021
GDP_Asia <- GDP_NEET %>%
  filter(Year >= 2009, Year <= 2020) %>%
  inner_join(Continents_NEET, Year, by = "Entity") %>%
  select(-Year.y, -Code.y) %>%
  filter(Continent == "Asia")

# Calculate GDP growth rate
GDP_Growth_Asia <- GDP_Asia %>%
  group_by(Year.x) %>%
```

```r
  summarise(Total_GDP = sum(GDP.per.capita..PPP..constant.2017.international...)) %>%
  mutate(GDP_Growth = ((Total_GDP - lag(Total_GDP)) / lag(Total_GDP)) * 100)

# Calculate mean NEET
mean_youth_NEET_Asia <- youth_NEET_Asia %>%
  group_by(Year.x) %>%
  summarise(mean_NEET = mean(Share.of.youth.not.in.education..employment.or.training..total....of.youth

# Combine the two data set for plot
GDP_Growth_NEET_Asia <- inner_join(GDP_Growth_Asia, mean_youth_NEET_Asia, by = "Year.x")


# Plot NEET and GDP growth
Asia_NEET_GDP <- ggplot(GDP_Growth_NEET_Asia, aes(x = Year.x)) +
  geom_point(aes(y = mean_NEET, colour = "Mean NEET"), size = 1.8) +
  geom_line(aes(y = mean_NEET, colour = "Mean NEET"), size = 0.8) +
  geom_point(aes(y = GDP_Growth, colour = "GDP Growth"), size = 1.8) +
  geom_line(aes(y = GDP_Growth, colour = "GDP Growth"), size = 0.8) +
  scale_x_continuous(breaks = 2010:2020) +
  labs(title = "GDP Growth per Capita VS Mean Youth NEET in Asia (2010-2020)",
       x = "Year",
       y = "GDP Growth per Capita and Mean Youth NEET")

print(Asia_NEET_GDP)
```

Europe: NEET vs GDP Growth

```r
# Filter the data of NEET in Europe in 2010-2021
youth_NEET_Europe <- NEET %>%
  filter(Year >= 2010, Year <= 2020) %>%
  inner_join(Continents_NEET, Year, by = "Entity") %>%
  select(-Year.y, -Code.y) %>%
  filter(Continent == "Europe")

# Filter the data of GDP in Europe in 2010-2021
GDP_Europe <- GDP_NEET %>%
  filter(Year >= 2009, Year <= 2020) %>%
  inner_join(Continents_NEET, Year, by = "Entity") %>%
  select(-Year.y, -Code.y) %>%
  filter(Continent == "Europe")

# Calculate GDP growth rate
GDP_Growth_Europe <- GDP_Europe %>%
  group_by(Year.x) %>%
  summarise(Total_GDP = sum(GDP.per.capita..PPP..constant.2017.international...)) %>%
  mutate(GDP_Growth = ((Total_GDP - lag(Total_GDP)) / lag(Total_GDP)) * 100)

# Calculate mean NEET
mean_youth_NEET_Europe <- youth_NEET_Europe %>%
  group_by(Year.x) %>%
  summarise(mean_NEET = mean(Share.of.youth.not.in.education..employment.or.training..total....of.youth
```

```
# Combine the two data set for plot
GDP_Growth_NEET_Europe <- inner_join(GDP_Growth_Europe, mean_youth_NEET_Europe, by = "Year.x")


# Plot NEET and GDP growth
Europe_NEET_GDP <- ggplot(GDP_Growth_NEET_Europe, aes(x = Year.x)) +
  geom_point(aes(y = mean_NEET, colour = "Mean NEET"), size = 1.8) +
  geom_line(aes(y = mean_NEET, colour = "Mean NEET"), size = 0.8) +
  geom_point(aes(y = GDP_Growth, colour = "GDP Growth"), size = 1.8) +
  geom_line(aes(y = GDP_Growth, colour = "GDP Growth"), size = 0.8) +
  scale_x_continuous(breaks = 2010:2020) +
  labs(title = "GDP Growth per Capita VS Mean Youth NEET in Europe (2010-2020)",
       x = "Year",
       y = "GDP Growth per Capita and Mean Youth NEET")

print(Europe_NEET_GDP)
```

North America: NEET vs GDP Growth

```
# Filter the data of NEET in North America in 2010-2021
youth_NEET_NorthAmerica <- NEET %>%
  filter(Year >= 2010, Year <= 2020) %>%
  inner_join(Continents_NEET, Year, by = "Entity") %>%
  select(-Year.y, -Code.y) %>%
  filter(Continent == "North America")

# Filter the data of GDP in North America in 2010-2021
GDP_NorthAmerica <- GDP_NEET %>%
  filter(Year >= 2009, Year <= 2020) %>%
  inner_join(Continents_NEET, Year, by = "Entity") %>%
  select(-Year.y, -Code.y) %>%
  filter(Continent == "North America")

# Calculate GDP growth rate
GDP_Growth_NorthAmerica <- GDP_NorthAmerica %>%
  group_by(Year.x) %>%
  summarise(Total_GDP = sum(GDP.per.capita..PPP..constant.2017.international...)) %>%
  mutate(GDP_Growth = ((Total_GDP - lag(Total_GDP)) / lag(Total_GDP)) * 100)

# Calculate mean NEET
mean_youth_NEET_NorthAmerica <- youth_NEET_NorthAmerica %>%
  group_by(Year.x) %>%
  summarise(mean_NEET = mean(Share.of.youth.not.in.education..employment.or.training..total....of.youth

# Combine the two data set for plot
GDP_Growth_NEET_NorthAmerica <- inner_join(GDP_Growth_NorthAmerica, mean_youth_NEET_NorthAmerica, by =

# Plot NEET and GDP growth
NorthAmerica_NEET_GDP <- ggplot(GDP_Growth_NEET_NorthAmerica, aes(x = Year.x)) +
  geom_point(aes(y = mean_NEET, colour = "Mean NEET"), size = 1.8) +
  geom_line(aes(y = mean_NEET, colour = "Mean NEET"), size = 0.8) +
  geom_point(aes(y = GDP_Growth, colour = "GDP Growth"), size = 1.8) +
```

```r
  geom_line(aes(y = GDP_Growth, colour = "GDP Growth"), size = 0.8) +
  scale_x_continuous(breaks = 2010:2020) +
  labs(title = "GDP Growth per Capita VS Mean Youth NEET in North America (2010-2020)",
       x = "Year",
       y = "GDP Growth per Capita and Mean Youth NEET")

print(NorthAmerica_NEET_GDP)
```

Oceania: NEET vs GDP Growth

```r
# Filter the data of NEET in Oceania in 2010-2021
youth_NEET_Oceania <- NEET %>%
  filter(Year >= 2010, Year <= 2020) %>%
  inner_join(Continents_NEET, Year, by = "Entity") %>%
  select(-Year.y, -Code.y) %>%
  filter(Continent == "Oceania")

# Filter the data of GDP in Oceania in 2010-2021
GDP_Oceania <- GDP_NEET %>%
  filter(Year >= 2009, Year <= 2020) %>%
  inner_join(Continents_NEET, Year, by = "Entity") %>%
  select(-Year.y, -Code.y) %>%
  filter(Continent == "Oceania")

# Calculate GDP growth rate
GDP_Growth_Oceania <- GDP_Oceania %>%
  group_by(Year.x) %>%
  summarise(Total_GDP = sum(GDP.per.capita..PPP..constant.2017.international...)) %>%
  mutate(GDP_Growth = ((Total_GDP - lag(Total_GDP)) / lag(Total_GDP)) * 100)

# Calculate mean NEET
mean_youth_NEET_Oceania <- youth_NEET_Oceania %>%
  group_by(Year.x) %>%
  summarise(mean_NEET = mean(Share.of.youth.not.in.education..employment.or.training..total....of.youth

# Combine the two data set for plot
GDP_Growth_NEET_Oceania <- inner_join(GDP_Growth_Oceania, mean_youth_NEET_Oceania, by = "Year.x")

# Plot NEET and GDP growth
Oceania_NEET_GDP <- ggplot(GDP_Growth_NEET_Oceania, aes(x = Year.x)) +
  geom_point(aes(y = mean_NEET, colour = "Mean NEET"), size = 1.8) +
  geom_line(aes(y = mean_NEET, colour = "Mean NEET"), size = 0.8) +
  geom_point(aes(y = GDP_Growth, colour = "GDP Growth"), size = 1.8) +
  geom_line(aes(y = GDP_Growth, colour = "GDP Growth"), size = 0.8) +
  scale_x_continuous(breaks = 2010:2020) +
  labs(title = "GDP Growth per Capita VS Mean Youth NEET in Oceania (2010-2020)",
       x = "Year",
       y = "GDP Growth per Capita and Mean Youth NEET")

print(Oceania_NEET_GDP)
```

South America: NEET vs GDP Growth

```r
# Filter the data of NEET in South America in 2010-2021
youth_NEET_SouthAmerica <- NEET %>%
  filter(Year >= 2010, Year <= 2020) %>%
  inner_join(Continents_NEET, Year, by = "Entity") %>%
  select(-Year.y, -Code.y) %>%
  filter(Continent == "South America")

# Filter the data of GDP in South America in 2010-2021
GDP_SouthAmerica <- GDP_NEET %>%
  filter(Year >= 2009, Year <= 2020) %>%
  inner_join(Continents_NEET, Year, by = "Entity") %>%
  select(-Year.y, -Code.y) %>%
  filter(Continent == "South America")

# Calculate GDP growth rate
GDP_Growth_SouthAmerica <- GDP_SouthAmerica %>%
  group_by(Year.x) %>%
  summarise(Total_GDP = sum(GDP.per.capita..PPP..constant.2017.international...)) %>%
  mutate(GDP_Growth = ((Total_GDP - lag(Total_GDP)) / lag(Total_GDP)) * 100)

# Calculate mean NEET
mean_youth_NEET_SouthAmerica <- youth_NEET_SouthAmerica %>%
  group_by(Year.x) %>%
  summarise(mean_NEET = mean(Share.of.youth.not.in.education..employment.or.training..total....of.youth

# Combine the two data set for plot
GDP_Growth_NEET_SouthAmerica <- inner_join(GDP_Growth_SouthAmerica, mean_youth_NEET_SouthAmerica, by =
```

```r
# Plot NEET and GDP growth
SouthAmerica_NEET_GDP <- ggplot(GDP_Growth_NEET_SouthAmerica, aes(x = Year.x)) +
  geom_point(aes(y = mean_NEET, colour = "Mean NEET"), size = 1.8) +
  geom_line(aes(y = mean_NEET, colour = "Mean NEET"), size = 0.8) +
  geom_point(aes(y = GDP_Growth, colour = "GDP Growth"), size = 1.8) +
  geom_line(aes(y = GDP_Growth, colour = "GDP Growth"), size = 0.8) +
  scale_x_continuous(breaks = 2010:2020) +
  labs(title = "GDP Growth per Capita VS Mean Youth NEET in South America (2010-2020)",
       x = "Year",
       y = "GDP Growth per Capita and Mean Youth NEET")

print(SouthAmerica_NEET_GDP)
```

NEET Proportion Analysis

Change in Proportion of Youth NEET (2015-2020)

```r
# Load data
GDP_prop <- read.csv("data sets/gdp-per-capita-worldbank.csv")
NEET_prop <- read.csv("data sets/youth-not-in-education-employment-training.csv")
```

```r
Continents_prop <- read.csv("data sets/continents-according-to-our-world-in-data.csv")

# Filter the data of NEET from 2015-2020
NEET_filtered <- NEET_prop %>%
  filter(Year %in% c(2015, 2020)) %>%
  inner_join(Continents_prop, Year, by = "Entity") %>%
  select(-Year.y, -Code.y) %>%
  rename(NEET = Share.of.youth.not.in.education..employment.or.training..total....of.youth.population.)

# Find Continent Average
continent_neet <- NEET_filtered %>%
  group_by(Continent, Year.x) %>%
  summarise(mean_NEET = mean(NEET, na.rm = TRUE))

# Form data set of proportion in NEET and not in NEET
continent_proportion <- continent_neet %>%
  mutate(
    NEET = mean_NEET,
    In_EET = 100 - mean_NEET
  ) %>%
  select(Continent, Year.x, NEET, In_EET)

# Change to long format
continent_long <- continent_proportion %>%
  pivot_longer(cols = c("NEET", "In_EET"),
               names_to = "Category",
               values_to = "Proportion")
```

```r
# Plot the graph
NEET_Proportion_Plot <- ggplot(continent_long,
      aes(x = factor(Year.x),
          y = Proportion,
          fill = Category)) +
  geom_col(position = "stack") +
  facet_wrap(~ Continent, nrow = 1) +
  scale_y_continuous(limits = c(0, 100)) +
  labs(
    title = "Change in Proportion of Youth NEET in Each Continent (2015-2020)",
    x = "Continents",
    y = "Proportion (%)"
  ) +
  theme_minimal()

print(NEET_Proportion_Plot)
```

NEET Trends Across Continents

Mean Youth NEET Percentage (2010-2020)

```r
# Load data
youth_NEET_trends <- read.csv("data sets/youth-not-in-education-employment-training.csv")
```

```r
continents_trends <- read.csv("data sets/continents-according-to-our-world-in-data.csv")

# Tidy columns
youth_NEET_continent_1 <- inner_join(youth_NEET_trends, continents_trends, by = "Entity")
youth_NEET_continent <- youth_NEET_continent_1 %>%
  select(-Code.y, -Year.y)

# Rename columns
youth_NEET_continent <- youth_NEET_continent %>%
  rename(NEET = Share.of.youth.not.in.education..employment.or.training..total....of.youth.population.,
         Year = Year.x)

# Calculate mean
mean_NEET_continents <- youth_NEET_continent %>%
  group_by(Continent, Year) %>%
  summarise(mean_NEET = mean(NEET, na.rm = TRUE))


# Plotting
NEET_Trends_Plot <- mean_NEET_continents %>%
  filter(Year >= 2010, Year <= 2020) %>%
  ggplot(aes(x = Year, y = mean_NEET, colour = Continent)) +
  geom_point(size = 1.8) +
  geom_line(size = 0.8) +
  scale_x_continuous(breaks = 2010:2020) +
  labs(title = "Mean Youth NEET Percentage Across Continents from 2010 to 2020",
       y = "Mean NEET")

print(NEET_Trends_Plot)
```