

IDS Group Work Code Sahil Ali

Sahil Ali

2025-12-09

```
#This R markdown file will show the code for GDP growth vs NEET rates
```

```
countries <- read.csv("data sets/continents-according-to-our-world-in-data.csv",
                      stringsAsFactors = F)

youth_unemployment <- read.csv("data sets/youth-not-in-education-employment-training.csv",
                                 stringsAsFactors = F)
GDP_per_capita <- read.csv("data sets/gdp-per-capita-worldbank.csv",
                           stringsAsFactors = F)
```

Calculating the mean GDP per Capita growth rate by Continent Then joining with GDP per capita data

```
library(dplyr)
```

```
##
## Attaching package: 'dplyr'

## The following objects are masked from 'package:stats':
##     filter, lag

## The following objects are masked from 'package:base':
##     intersect, setdiff, setequal, union

all_gdp <- countries %>%
  left_join(GDP_per_capita, by = "Entity")
```

Removing and renaming certain columns, filtering out countries with no data

```
all_gdp <- all_gdp %>%
  select(-Year.x, -Code.x) %>%
  rename(Country = Entity) %>%
  rename(Year = Year.y) %>%
  rename(Code = Code.y) %>%
  rename(GDP_PC = GDP.per.capita..PPP..constant.2017.international...) %>%
  filter(!is.na(GDP_PC))
```

Calculating the GDP per capita growth rate for every year

```

all_gdp <- all_gdp %>%
  arrange(Country, Year) %>% # Make sure years are in order per country
  group_by(Country) %>% # Calculate within each country
  mutate(growth_rate = (GDP_PC - lag(GDP_PC)) / lag(GDP_PC) * 100) %>%
  ungroup()

continent_year_growth <- all_gdp %>%
  group_by(Continent, Year) %>%
  summarise(mean_growth_rate = mean(growth_rate, na.rm = TRUE),
            .groups = "drop") %>%
  filter(!is.na(mean_growth_rate))

```

Join with youth unemployment data

```

youth_unemployment2 <- countries %>%
  left_join(youth_unemployment, by = "Entity")

```

Removing and renaming certain columns, filtering out countries with no data

```

youth_unemployment2 <- youth_unemployment2 %>%
  select(-Year.x, -Code.x) %>%
  rename(Country = Entity) %>%
  rename(Year = Year.y) %>%
  rename(Code = Code.y) %>%
  rename(y_unemployment = Share.of.youth.not.in.education..employment.or.training..total....of.youth.pop)
  filter(!is.na(y_unemployment))

```

Calculating the mean youth unemployment rate by Continent

```

youth_unemployment2 <- youth_unemployment2 %>%
  arrange(Country, Year) %>% # Make sure years are in order per country
  group_by(Country) %>% # Calculate within each country
  mutate(growth_rate = (y_unemployment - lag(y_unemployment)) / lag(y_unemployment) * 100) %>%
  ungroup()

continent_year_unemployment <- youth_unemployment2 %>%
  group_by(Continent, Year) %>%
  summarise(mean_growth_rate = mean(growth_rate, na.rm = TRUE),
            .groups = "drop") %>%
  filter(!is.na(mean_growth_rate))

```

Creating one final data set

```

final_data <- continent_year_growth %>%
  rename(GDP_growth = mean_growth_rate) %>%
  inner_join(continent_year_unemployment %>%
    rename(Youth_unemployment_change = mean_growth_rate),
             by = c("Continent", "Year")) %>%
  filter(Year >= 2005, GDP_growth >= -7.5, Youth_unemployment_change <= 50)

```

Now plotting the graph

```

library(ggplot2)

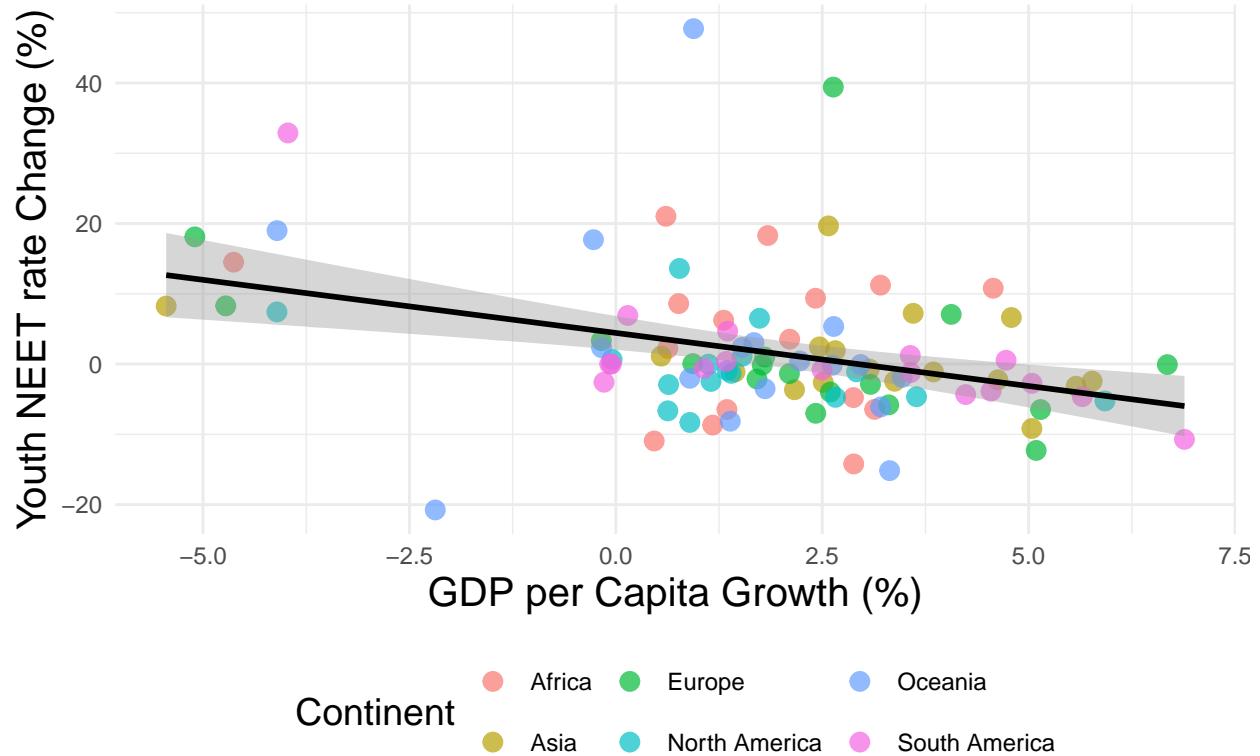
## Warning: package 'ggplot2' was built under R version 4.5.2

ggplot(final_data, aes(x = GDP_growth, y = Youth_unemployment_change)) +
  geom_point(aes(color = Continent), size = 3, alpha = 0.7) +
  geom_smooth(method = "lm", se = TRUE, color = "black") +
  labs(
    title = "Relationship between GDP per Capita Growth and Youth NEET rates",
    x = "GDP per Capita Growth (%)",
    y = "Youth NEET rate Change (%)") +
  theme_minimal() +
  theme(
    plot.title = element_text(hjust = 0.5, size = 18, face = "bold"),
    axis.title = element_text(size = 15),
    legend.title = element_text(size = 14),
    legend.position = "bottom")

```

`geom_smooth()` using formula = 'y ~ x'

Relationship between GDP per Capita Growth and Youth NEET rates



```
ggsave("GDP growth vs youth NEET Plot.png", width = 10, height = 6, dpi = 300)
```

`geom_smooth()` using formula = 'y ~ x'

```
#This R markdown file will show the code for mean growth by continent
```

```
countries <- read.csv("data sets/continents-according-to-our-world-in-data.csv",
                      stringsAsFactors = F)

GDP_per_capita <- read.csv("data sets/gdp-per-capita-worldbank.csv",
                           stringsAsFactors = F)
```

Join with GDP per capita data

```
library(dplyr)

all_gdp <- countries %>%
  left_join(GDP_per_capita, by = "Entity")
```

Removing and renaming certain columns, filtering out countries with no data

```
all_gdp <- all_gdp %>%
  select(-Year.x, -Code.x) %>%
  rename(Country = Entity) %>%
  rename(Year = Year.y) %>%
  rename(Code = Code.y) %>%
  rename(GDP_PC = GDP.per.capita..PPP..constant.2017.international...) %>%
  filter(!is.na(GDP_PC)) %>%
  filter(Year >= 2014)
```

Calculating the GDP per capita growth rate for every year

```
all_gdp <- all_gdp %>%
  arrange(Country, Year) %>% # Make sure years are in order per country
  group_by(Country) %>% # Calculate within each country
  mutate(growth_rate = (GDP_PC - lag(GDP_PC)) / lag(GDP_PC) * 100) %>%
  ungroup()

continent_year_growth <- all_gdp %>%
  group_by(Continent, Year) %>%
  summarise(mean_growth_rate = mean(growth_rate, na.rm = TRUE),
            .groups = "drop") %>%
  filter(!is.na(mean_growth_rate))
```

Now to plot the graphs

```
library(ggplot2)

ggplot(continent_year_growth, aes(x = Year, y = mean_growth_rate,
                                   color = Continent)) +
  geom_line(size = 0.7) +
  geom_point(size = 1.7) +
  labs(
    title = "Mean Growth in GDP per Capita for Each Continent",
    x = "Year",
    y = "Growth Rate (%)",
```

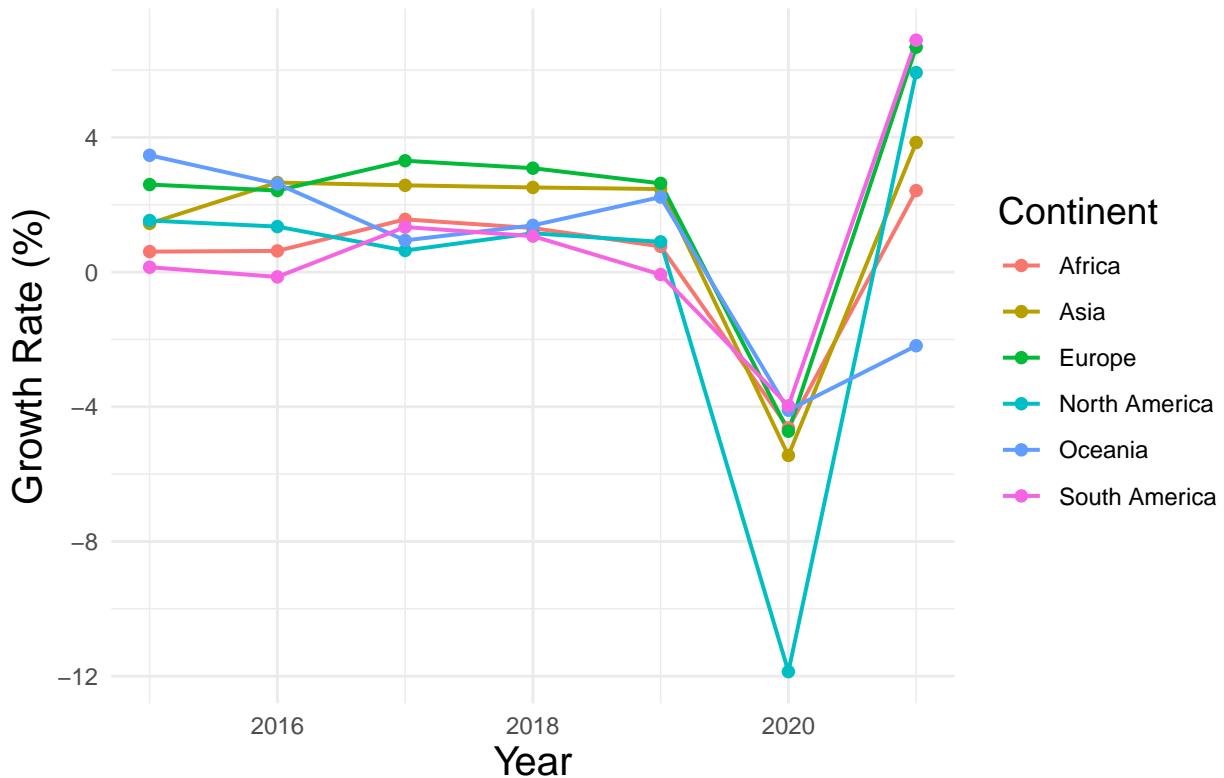
```

color = "Continent") +
theme_minimal() +
theme(
  plot.title = element_text(hjust = 0.5, size = 18, face = "bold"),
  axis.title = element_text(size = 15),
  legend.title = element_text(size = 14))

## 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.

```

Mean Growth in GDP per Capita for Each Continent



```
ggsave("Continent Mean Growth Plot.png", width = 10, height = 6, dpi = 300)
```

#This R markdown file will show the code for median growth by continent

```

countries <- read.csv("data sets/continents-according-to-our-world-in-data.csv",
                      stringsAsFactors = F)

GDP_per_capita <- read.csv("data sets/gdp-per-capita-worldbank.csv",
                           stringsAsFactors = F)

```

Joining the data sets

```

library(dplyr)

all_gdp <- countries %>%
  left_join(GDP_per_capita, by = "Entity")

```

Removing and renaming certain columns, filtering out countries with no data

```

all_gdp <- all_gdp %>%
  select(-Year.x, -Code.x) %>%
  rename(Country = Entity) %>%
  rename(Year = Year.y) %>%
  rename(Code = Code.y) %>%
  rename(GDP_PC = GDP.per.capita..PPP..constant.2017.international...) %>%
  filter(!is.na(GDP_PC)) %>%
  filter(Year >= 2014)

```

Calculating the GDP per capita growth rate for every year

```

all_gdp <- all_gdp %>%
  arrange(Country, Year) %>% # Make sure years are in order per country
  group_by(Country) %>% # Calculate within each country
  mutate(growth_rate = (GDP_PC - lag(GDP_PC)) / lag(GDP_PC) * 100) %>%
  ungroup()

continent_year_growth <- all_gdp %>%
  group_by(Continent, Year) %>%
  summarise(median_growth_rate = median(growth_rate, na.rm = TRUE),
            .groups = "drop") %>%
  filter(!is.na(median_growth_rate))

```

Now to plot the graphs

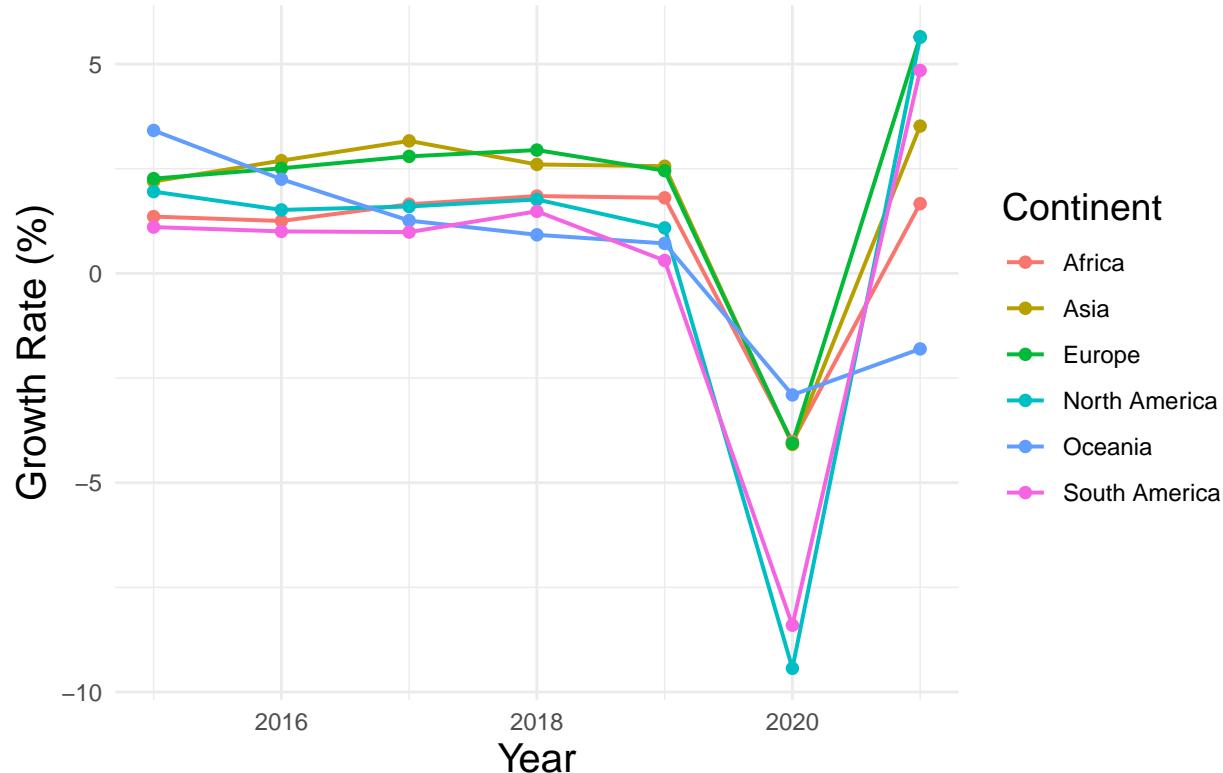
```

library(ggplot2)

ggplot(continent_year_growth, aes(x = Year, y = median_growth_rate,
                                    color = Continent)) +
  geom_line(size = 0.7) +
  geom_point(size = 1.7) +
  labs(
    title = "Median Growth in GDP per Capita for Each Continent",
    x = "Year",
    y = "Growth Rate (%)",
    color = "Continent") +
  theme_minimal() +
  theme(
    plot.title = element_text(hjust = 0.5, size = 18, face = "bold"),
    axis.title = element_text(size = 15),
    legend.title = element_text(size = 14))

```

Median Growth in GDP per Capita for Each Continent



```
ggsave("Continent Median Growth Plot.png", width = 10, height = 6, dpi = 300)
```

#This R markdown file will show the code for Youth NEET rates by continent

```
countries <- read.csv("data sets/continents-according-to-our-world-in-data.csv",
                      stringsAsFactors = F)
youth_unemployment <- read.csv("data sets/youth-not-in-education-employment-training.csv",
                               stringsAsFactors = F)
```

Calculating the mean youth unemployment rate by Continent Join with youth unemployment data

```
library(dplyr)

youth_unemployment2 <- countries %>%
  left_join(youth_unemployment, by = "Entity")
```

Removing and renaming certain columns, filtering out countries with no data

```
youth_unemployment2 <- youth_unemployment2 %>%
  select(-Year.x, -Code.x) %>%
  rename(Country = Entity) %>%
  rename(Year = Year.y) %>%
  rename(Code = Code.y) %>%
  rename(y_unemployment = Share.of.youth.not.in.education..employment.or.training..total....of.youth.pop)
  filter(!is.na(y_unemployment))
```

Calculate the change in youth unemployment every year

```
youth_unemployment_change <- youth_unemployment2 %>%
  filter(Year >= 2015) %>%
  arrange(Country, Year) %>% # Make sure order is correct
  group_by(Country) %>% # Calculate within each country
  mutate(pct_change = ((y_unemployment - lag(y_unemployment)) / lag(y_unemployment)) * 100) %>%
  ungroup()
```

Calculate the average youth unemployment change each year per continent

```
continent_avg <- youth_unemployment_change %>%
  group_by(Continent, Year) %>%
  summarise(avg_pct_change = mean(pct_change, na.rm = TRUE), .groups = "drop")
```

Now to plot the graph

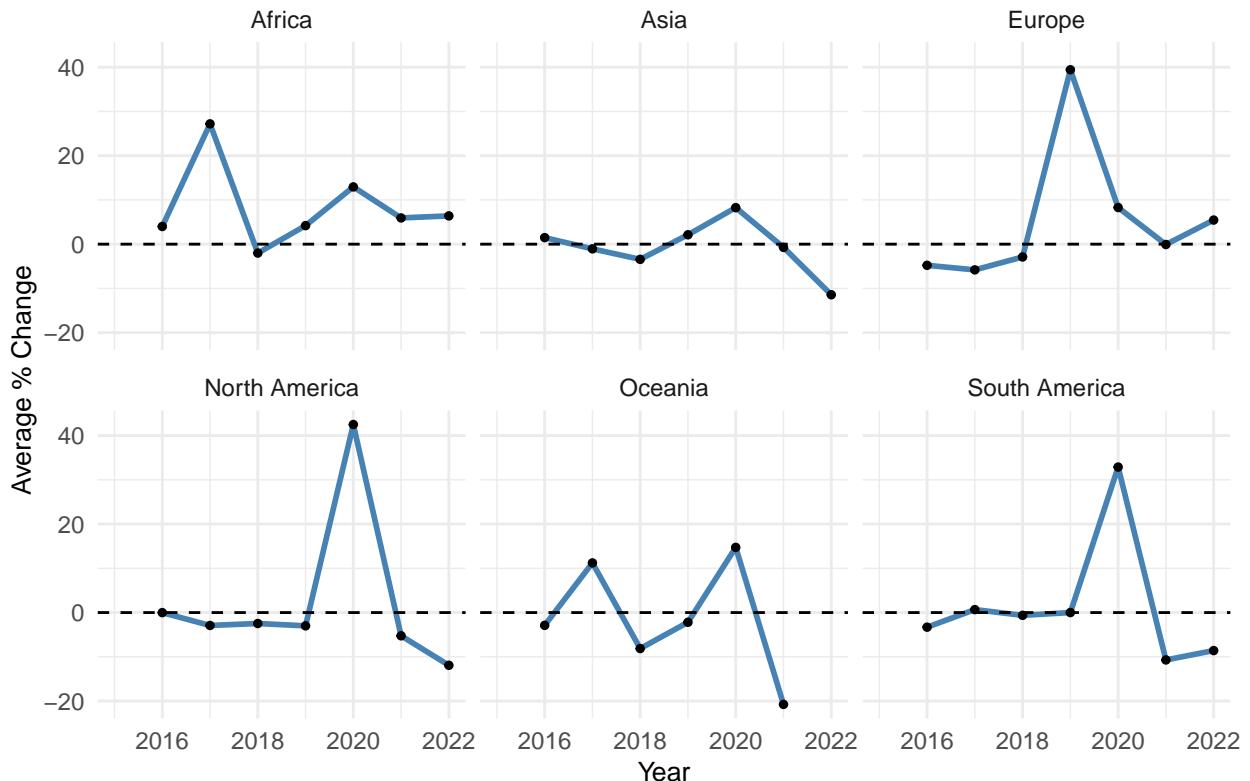
```
library(ggplot2)

ggplot(continent_avg, aes(x = Year, y = avg_pct_change)) +
  geom_line(size = 1, colour = "steelblue") +
  geom_point(size = 1) +
  geom_hline(yintercept = 0, linetype = "dashed") +
  facet_wrap(~Continent) +
  labs(
    title = "Average % Change in Youth NEET rates by Continent (from 2015)",
    x = "Year",
    y = "Average % Change") +
  theme_minimal() +
  theme(
    plot.title = element_text(hjust = 0.5, size = 12, face = "bold"),
    axis.title = element_text(size = 10),
    legend.title = element_text(size = 14))

## Warning: Removed 1 row 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()').
```

Average % Change in Youth NEET rates by Continent (from 2015)



```
ggsave("Youth NEET Rates Change Plot.png", width = 10, height = 6, dpi = 300)
```

```
## Warning: Removed 1 row 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()').
```

#This R markdown file will show the code for Youth NEET rates pre & post covid

```
countries <- read.csv("data sets/continents-according-to-our-world-in-data.csv",
                      stringsAsFactors = F)
youth_unemployment <- read.csv("data sets/youth-not-in-education-employment-training.csv",
                               stringsAsFactors = F)
```

Joining data sets

```
library(dplyr)

youth_unemployment2 <- countries %>%
  left_join(youth_unemployment, by = "Entity")
```

Removing and renaming certain columns, filtering out countries with no data

```

youth_unemployment2 <- youth_unemployment2 %>%
  select(-Year.x, -Code.x) %>%
  rename(Country = Entity) %>%
  rename(Year = Year.y) %>%
  rename(Code = Code.y) %>%
  rename(y_unemployment = Share.of.youth.not.in.education..employment.or.training..total....of.youth.pop)
  filter(!is.na(y_unemployment))

```

Calculating the average youth unemployment rate before and after 2020

```

youth_summary <- youth_unemployment2 %>%
  mutate(period = factor(
    ifelse(Year < 2020, "Before 2020", "After 2020"),
    levels = c("Before 2020", "After 2020"))) %>%
  group_by(Continent, period) %>%
  summarise(avg_youth_unemployment = mean(y_unemployment, na.rm = TRUE),
    .groups = "drop")

```

Coding the graphs

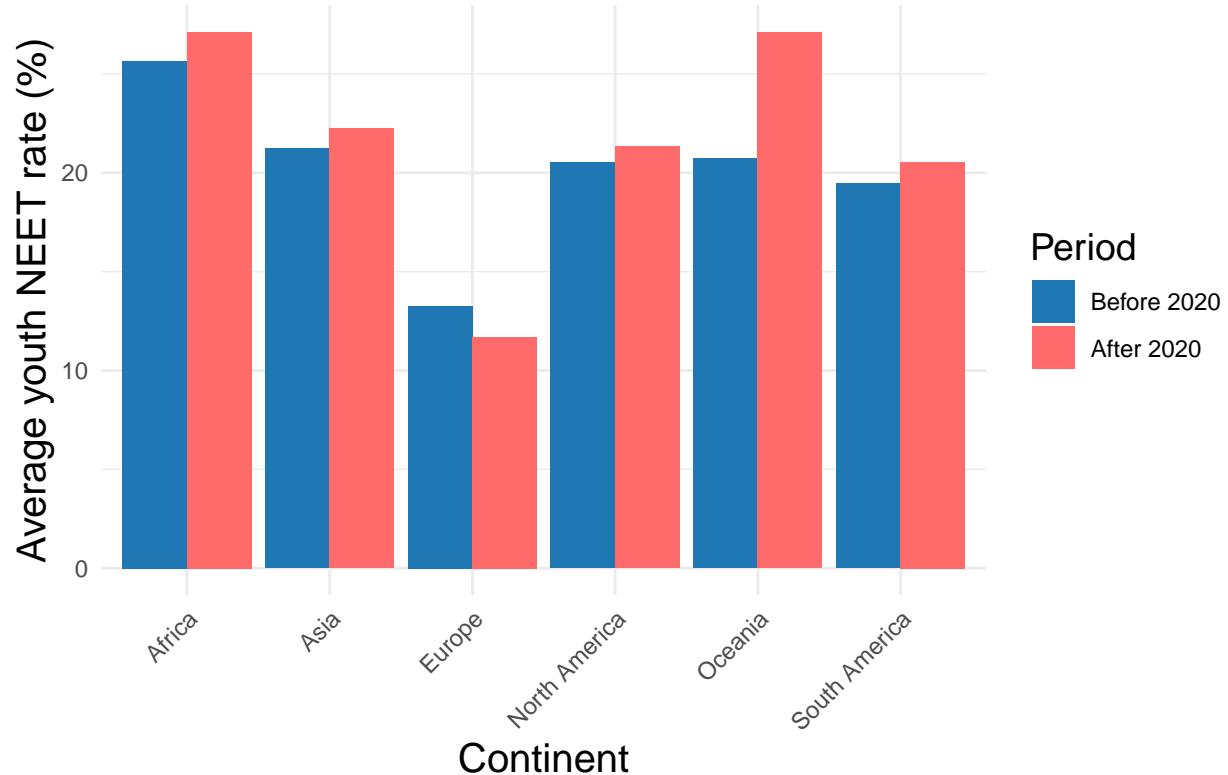
```

library(ggplot2)

ggplot(youth_summary, aes(x = Continent, y = avg_youth_unemployment,
  fill = period)) +
  geom_col(position = "dodge") +
  scale_fill_manual(values = c("Before 2020" = "#1f77b4",
    "After 2020" = "#FF6A6A")) +
  labs(
    title = "Youth NEET Rate Pre Covid Vs Post Covid",
    x = "Continent",
    y = "Average youth NEET rate (%)",
    fill = "Period") +
  theme_minimal() +
  theme(axis.text.x = element_text(angle = 45, hjust = 1),
    plot.title = element_text(hjust = 0.5, size = 18, face = "bold"),
    axis.title = element_text(size = 15),
    legend.title = element_text(size = 14))

```

Youth NEET Rate Pre Covid Vs Post Covid



```
ggsave("Youth NEET Bar Chart.png", width = 10, height = 6, dpi = 300)
```