GDP Per Capita, Price Levels and Population For European Countries

```
knitr::opts_chunk$set(warning = FALSE, error = FALSE, message = FALSE,
                      fig.align = 'center', fig.height = 5, fig.width = 5, tidy=TRUE)
```

Introduction

For our project we were particularly interested in the relation between GDP per capita and price levels in different European countries. We got our data from the Eurostat website. In the code below we begin by importing our dataset from the Eurostat website and cleaning it by removing the columns we do not need.

```
##MERGE DATASETS
total <- merge(gdp_data,price_data,by=c("TIME_PERIOD","geo"))

#total <- merge(total,pop_data,by=c("TIME_PERIOD","geo"))

total <- na.omit(total)

#total <- subset(total, select = -c(`LAST UPDATE`))
#total <- subset(total, select = -c(`LAST UPDATE.x`))
#total <- subset(total, select = -c(`LAST UPDATE.y`))
#total <- subset(total, select = -label)

total <- total %>% filter(geo !="EU27_2020")
total <- total %>% filter(geo !="EA19")
total <- total %>% filter(geo !="EA20")

total <- right_join(total, eu_countries, by = join_by(geo == code))
```

Average Price Levels in Different European Countries Between 2011 and 2022

Below we create an initial plot showing the average price levels in different European countries between the years 2012 to 2022. As seen below it is evident that Denmark (DK) has the highest average price levels while Bulgaria has the lowest average price levels.

```
#CREATE AN INITIAL PRICE PLOT
ggplot(total, aes(x =
                    reorder(x=geo,
                            X=price_levels),
                            y=price_levels)) + xlab("Countries") + ylab("Price Levels") +
  geom_bar(stat = "summary", fun="mean") + coord_flip() + theme(legend.position = "none")
```
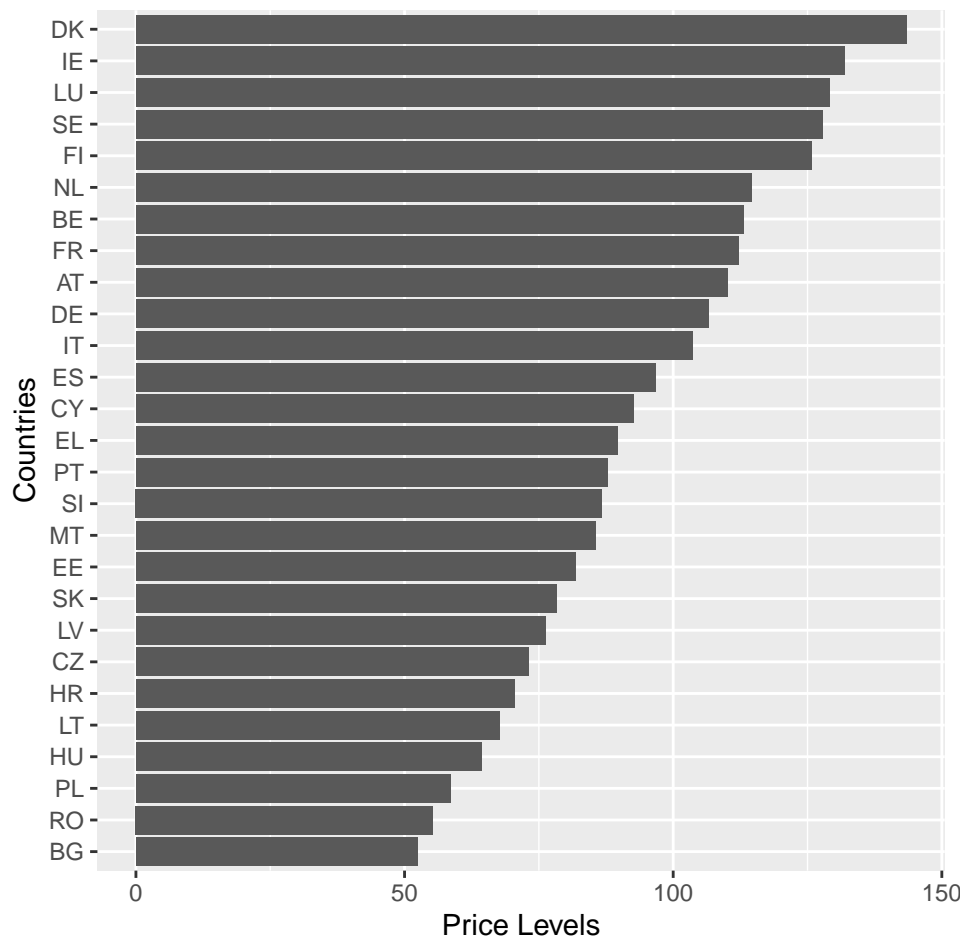
Price and GDP Levels in Different European Regions

We decided to get an approximation of price levels in different European regions to see if a certain region of the continent performs better in these metrics than others. As there is no defined regions, we used ChatGPT to help us with dividing the different parts of Europe. This is shown below.

```
##CREATE A REGION BASED COLUMN


western_europe <- c("BE", "DK", "DE", "IE", "FR", "IT", "LU", "NL", "PT", "ES")
northern_europe <- c("EE", "LV", "LT", "FI", "SE")
eastern_europe <- c("BG", "CZ", "HR", "HU", "PL", "RO", "SK","SI")
southern_europe <- c("EL", "CY", "MT", "AT")

total$regions <- ifelse(total$geo %in% western_europe,
                "Western Europe",
                ifelse(total$geo %in% eastern_europe, "Eastern Europe",
                ifelse(total$geo %in% southern_europe, "Southern Europe","Northern Europe")))
```

The plot below shows the average price levels in different European regions over the years 2011 to 2022. As we can see, Denmark performs the best on average and Bulgaria performs the worst.

```
##CREATE A REGION BASED PRICE PLOT


ggplot(total, aes(x =
                reorder(x=geo,
```
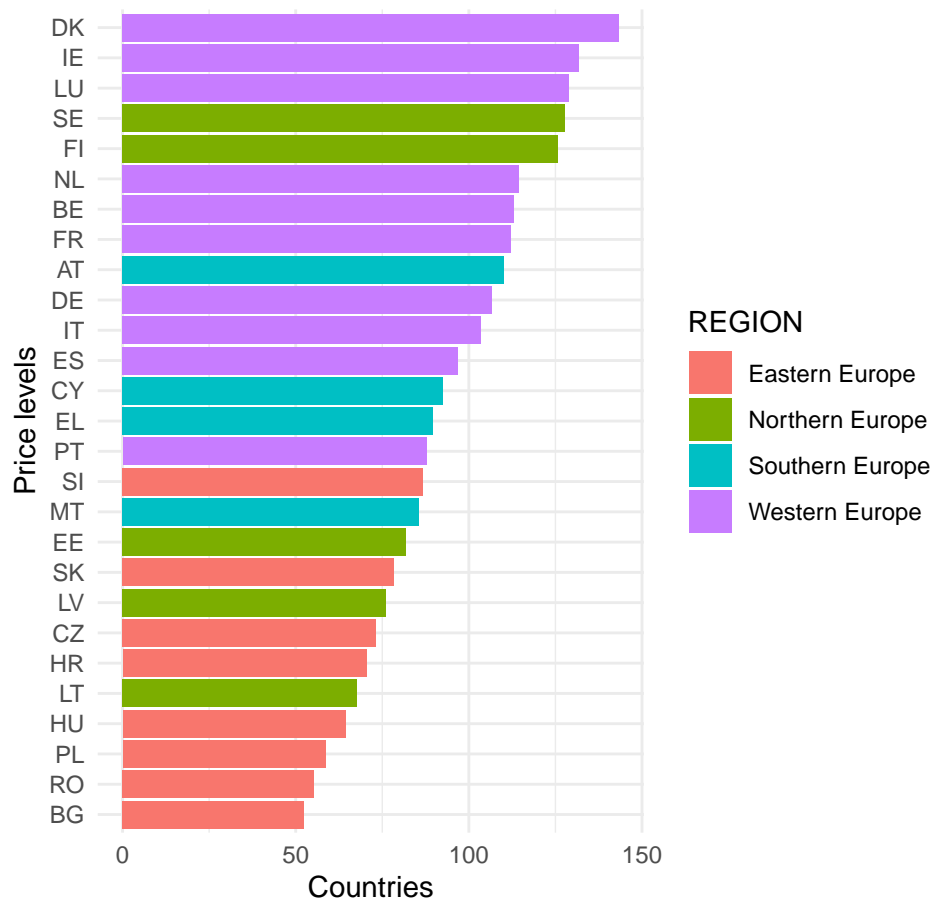
```
                                X=price_levels),
                            y=price_levels,
                            fill=regions)) + labs(x = "Price levels", y = "Countries", fill = "REGION"
    geom_bar(stat = "summary", fun="mean") + coord_flip() + theme_minimal()
```
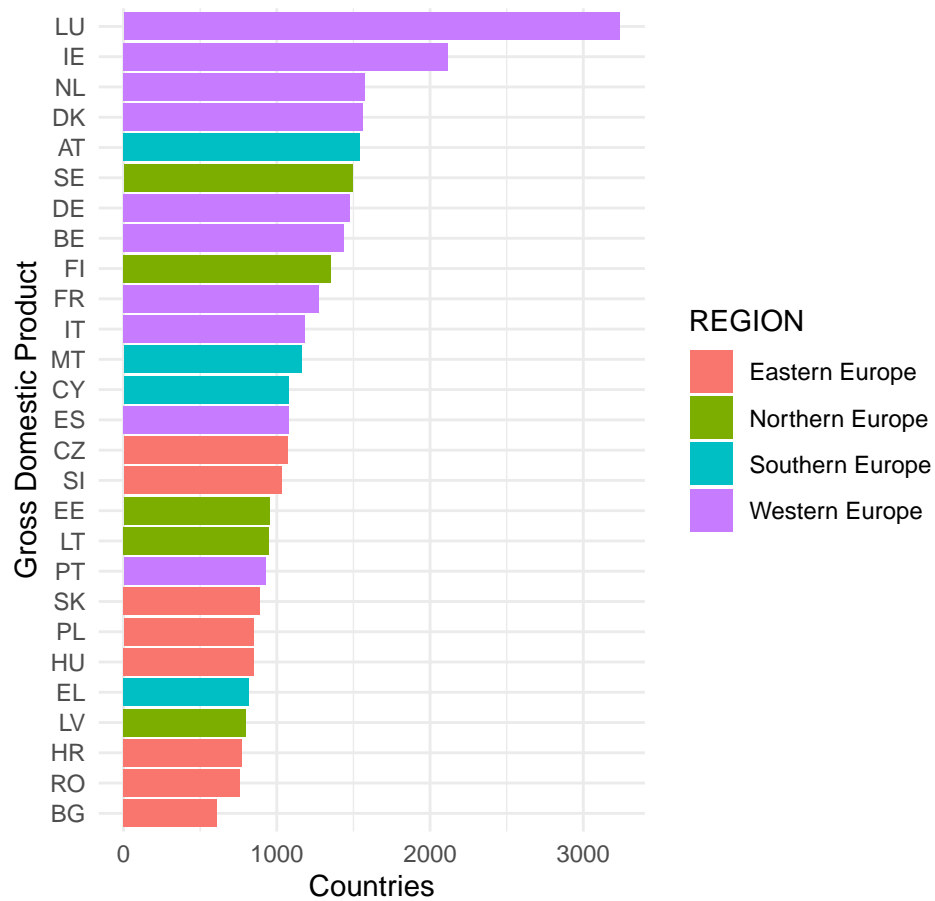

ice levels of various European countries separated by region

The plot below shows the average GDP per capita in different European regions over the years 2011 to 2022. As we can see, Luxembourg performs the best on average and Bulgaria performs the worst. We can understand this by the fact that GDP is calculated per capita and Luxembourg has one of the smallest populations.

```
#CREATE A REGION BASED GDP PLOT
ggplot(total, aes(x =
                    reorder(x=geo,
                            X=GDP),
                        y=GDP,
                        fill=regions)) + labs(x = "Gross Domestic Product", y = "Countries", fill =
```

the GDP of various European countries separated by region

Geospatial Plots

Here we visualise the data onto an actual map. As shown below, Luxembourg and the UK have the highest GDP.

```
#SHP_27
SHP_0 <- get_eurostat_geospatial(resolution = 10,
                                 nuts_level = 0,
                                 year = 2016)

EU28 <- eu_countries %>%
  select(geo = code, name)
SHP_27 <- SHP_0 %>%
  select(geo = NUTS_ID, geometry) %>%
  inner_join(EU28, by = "geo") %>%
  arrange(geo) %>%
  st_as_sf()

gdp_map <- total %>%
  filter(TIME_PERIOD == 2022) %>%
  select(geo, GDP) %>%
  inner_join(SHP_27, by = "geo") %>%
  st_as_sf()
```

```
gdp_map %>%
  ggplot(aes(fill = GDP)) +
  geom_sf() +
  scale_fill_continuous(type = "viridis") +
  scale_x_continuous(limits = c(-10, 35)) +
  scale_y_continuous(limits = c(35, 65)) +
  theme_void() +
  labs(caption="Varying GDP per European country in 2022")
```
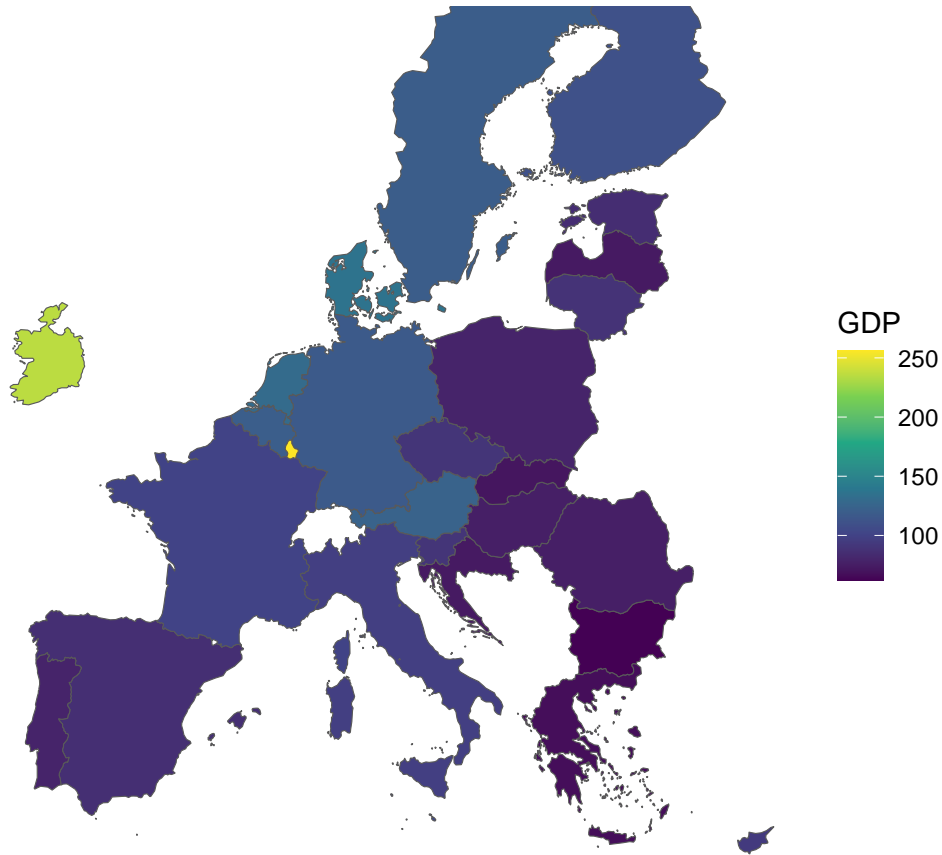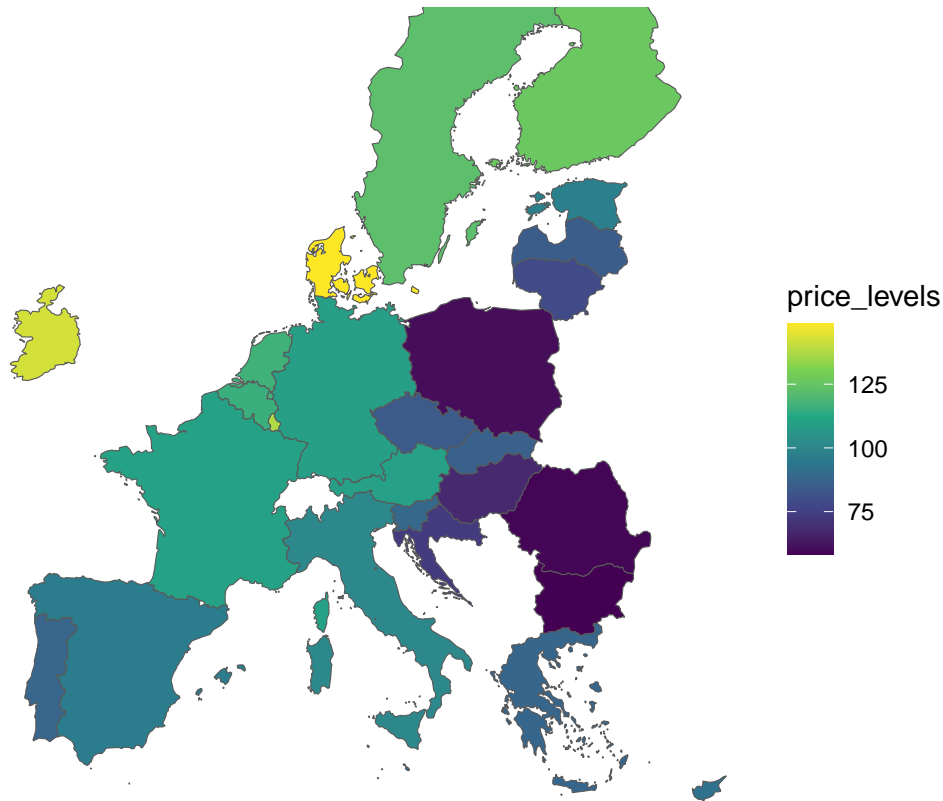

Varying GDP per European country in 2022

We can also see that Denmark, Luxembourg and the UK have the highest price levels on average.

```
price_map <- total %>%
  filter(TIME_PERIOD == 2022) %>%
  select(geo, price_levels) %>%
  inner_join(SHP_27, by = "geo") %>%
  st_as_sf()

price_map %>%
  ggplot(aes(fill = price_levels)) +
  geom_sf() +
  scale_fill_continuous(type = "viridis") +
  scale_x_continuous(limits = c(-10, 35)) +
  scale_y_continuous(limits = c(35, 65))  + theme_void() +
  labs(caption="Varying price levels per European country")
```

Varying price levels per European country

Spaghetti Plot

The spaghetti plot below illustrates the GDP per capita over time for the European countries. It is clear that the GDP for Ireland increases over time dramatically and the GDP for Luxembourg is the highest which makes sense since it is a rich small country and this is GDP per capita.

```
total$new_name = ifelse (total$name == "Ireland", "Ireland", ifelse(total$name == "Luxembourg", "Luxembo
# spagetti plot using total and gdp
ggplot(total, aes(x = TIME_PERIOD, y = GDP, group = name, color = new_name)) +
  geom_line() +
  labs(title = "Spaghetti Plot",
       x = "Year",
       y = "Value",
       caption = "Spaghetti plot showing the progression of GDP from 2012 to 2022") +
  theme_minimal() +
  scale_color_manual(name = "Country",
                     values = c("Ireland" = "green", "Luxembourg" = "lightblue", "Other" = "grey"))
```

Spaghetti Plot
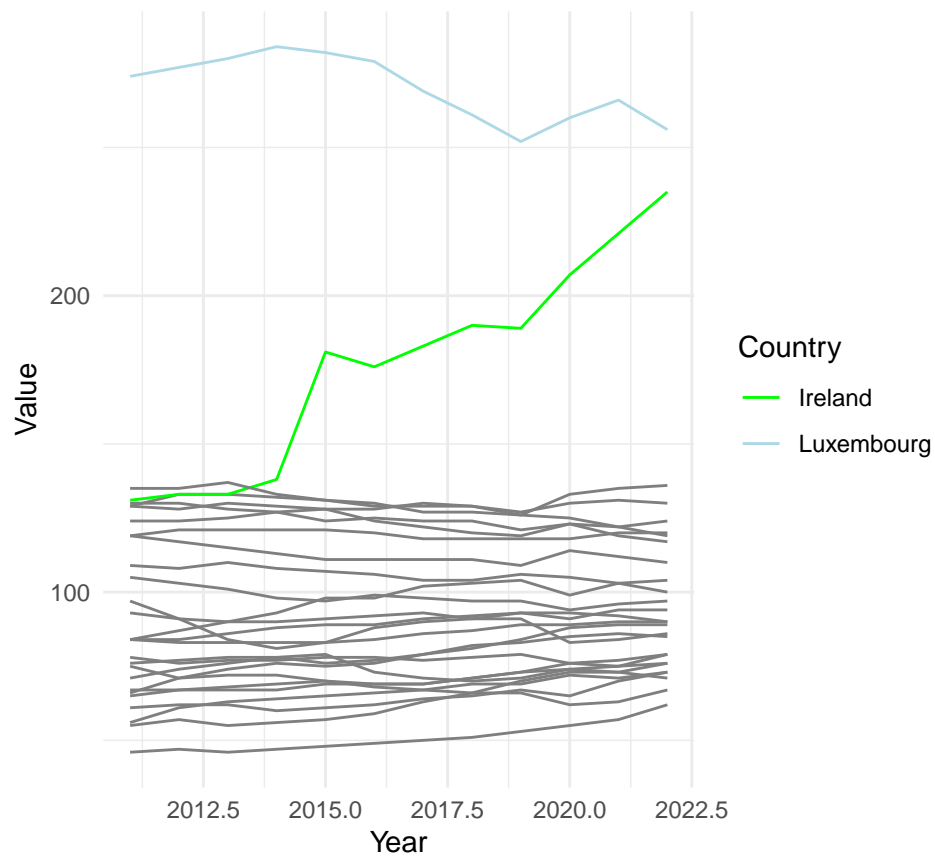
Spaghetti plot showing the progression of GDP from 2012 to 2022

Scatterplot

The figure below displays the scatter plot for the respective GDP value and price level for the European countries. The GDP value is on the x-axis and price level is on the y-axis. The colours represent the regions of Europe. Each dot on the scatter plot represents the price level and GDP for a specific country at a certain year. There is a positive trend between GDP and price level as one would expect.

```r
#new_total<- total$TIME_PERIOD=="2022"
new_total <- subset(total, TIME_PERIOD == 2022)

ggplot(new_total, aes(x=GDP, y=price_levels, color=regions)) + geom_point() + labs(caption = "Scatter pl
```

plot showing GDP vs. price levels for the year 2022 by region

The figure above displays the simplified scatter plot for the European countries for GDP and Average price levels. This time each dot represents the respective GDP and average price level for each country in a specific region averaged over all 10 the years. The highest GDP and average price level corresponds to Western Europe and the lowest GDP and average price level.