Closing event

EMOS Coding lab 2024

Statistics Explained through literate programming

28 June 2024

1.Welcome and opening

Albrecht Wirthmann

Eurostat, Head of Methodology, Innovation in official statistics

European Commission

EMOS goals

1. Promote cooperation between statistical offices and universities to offer content in official statistics in master's programmes

2. Build a European community in official statistics

Agenda

13:00 – 13:10	**Welcome and opening**	Albrecht Wirthmann *Eurostat, Head of Methodology, Innovation in official statistics*
13:10 – 13:20	**Coding lab 2024**	Tina Steenvoorden *EMOS secretariat* Andrea GALLELLI *Eurostat, Culture, and sport statistics*
13:20 – 14:20	**Presentation of the results by the Coding Lab participants**	
	Consumer prices of recreational and sporting goods and services	Ricardo Dias de Carbalho *University of Porto*
	Quality of life indicators - overall experience of life	Gabriella Manuti *University of Rome La Sapienza*
	Young people - digital world	Joseph Nyajuoga *Örebro University*
	Quality of life indicators - natural and living environment	Maripaz Venegas Gonzalez *Complutense University of Madrid*
	Consumer prices of cultural goods and services	Sona Yavrumyan *University of Bologna*
	Culture statistics - cultural employment	Nikola Quaresimin, *University of Bamberg*
14:20 – 14:35	**Coding Lab discussion**	Matyas Tamas MESZAROS *Eurostat, Statistical methodology and data integration*
14:35 – 14:45	**What EMOS has to offer in the next months**	Tina Steenvoorden *EMOS secretariat*

European Commission

2. EMOS Coding Lab 2024

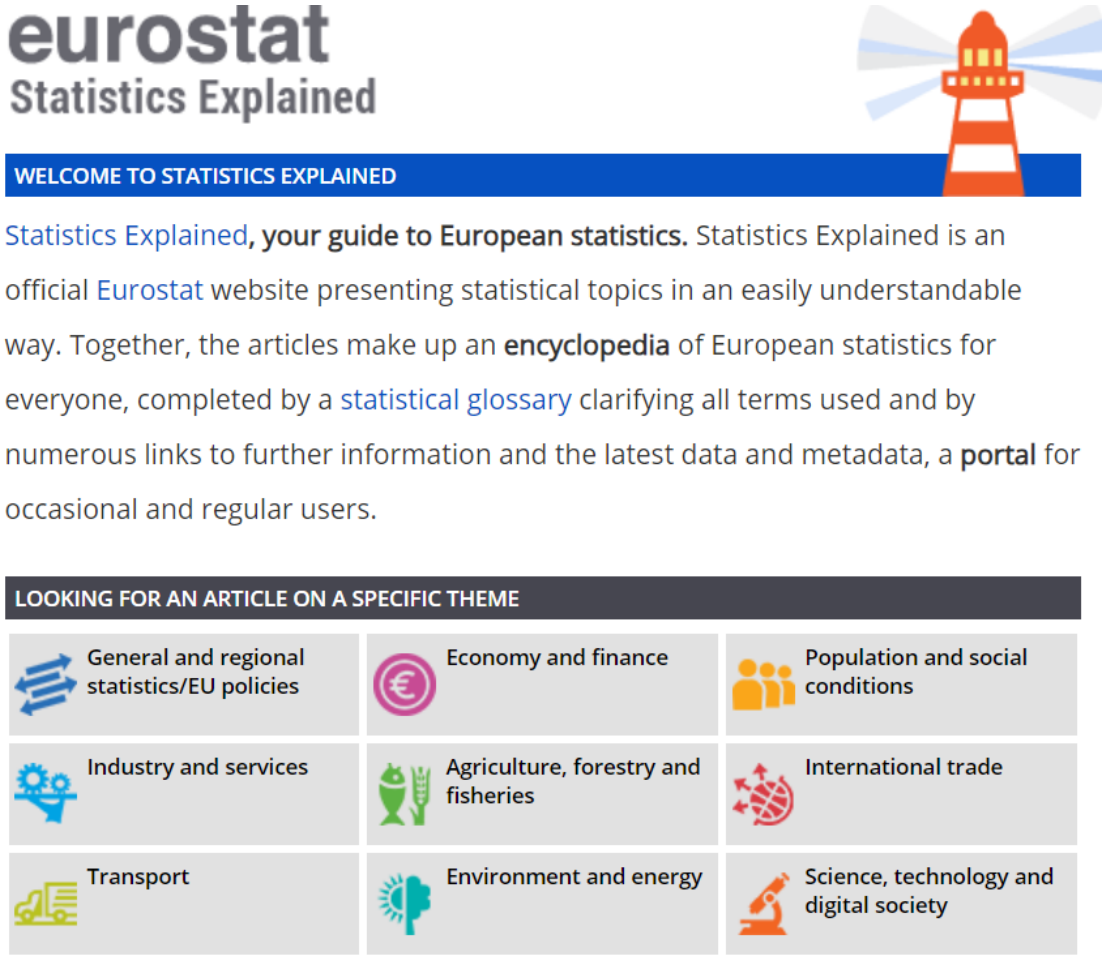
Tina Steenvoorden
EMOS secretariat

Andrea GALLELLI
Eurostat F1, culture and sport statistics

European Commission

Coding lab

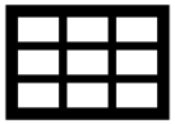
- 3rd edition of the EMOS Coding Lab entitled ‘Statistics Explained through literate programming’
- Call launched in February
- 13 students, 9 universities
- March – June 2024



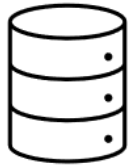
Organisation of work

- Kick off meeting end of March
- Division of groups – spread of knowledge
- MS teams as a tool for communication
- Regular meetings with mentors

Producing Statistics Explained Articles

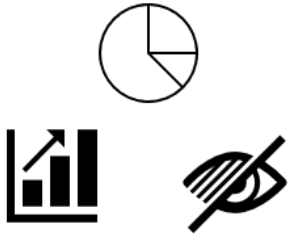


1. Data production



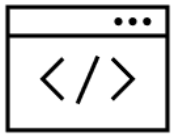
2. Upload on ‘Eurobase’

3. Download the data



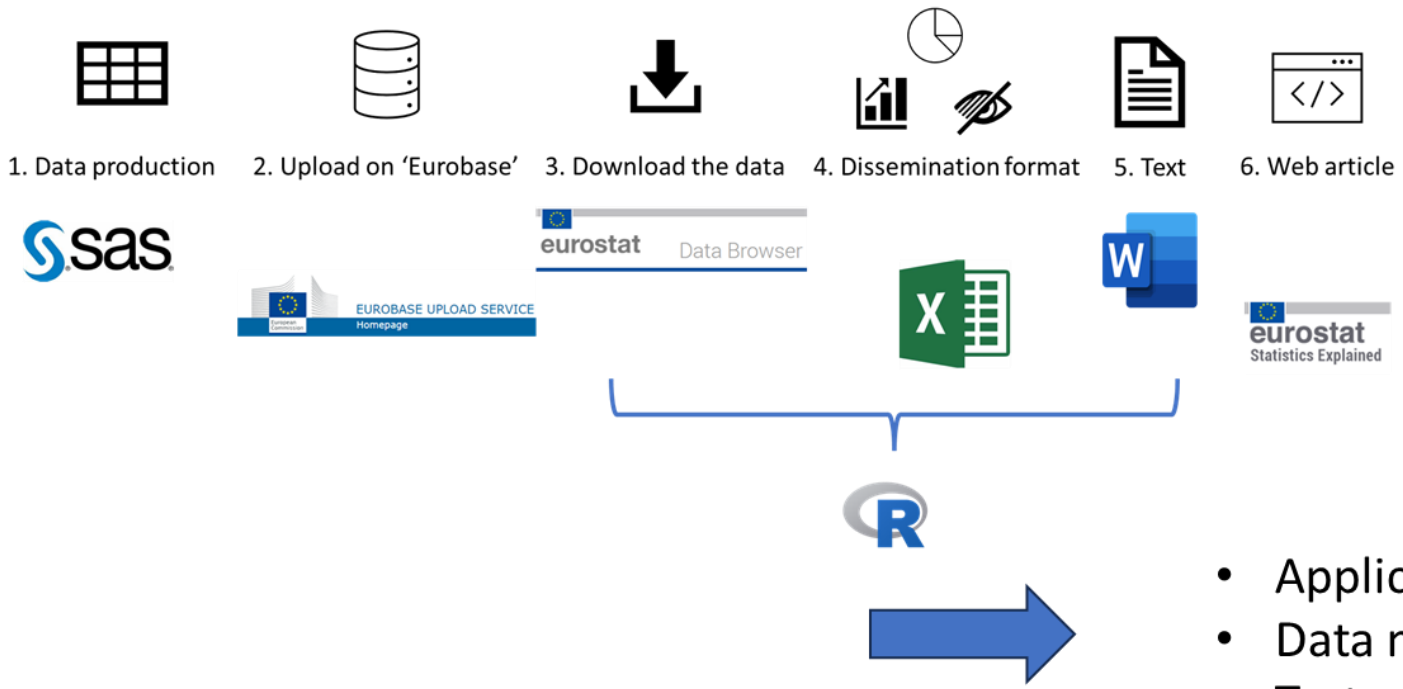
4. Dissemination format

5. Text



6. Web article

Producing Statistics Explained Articles

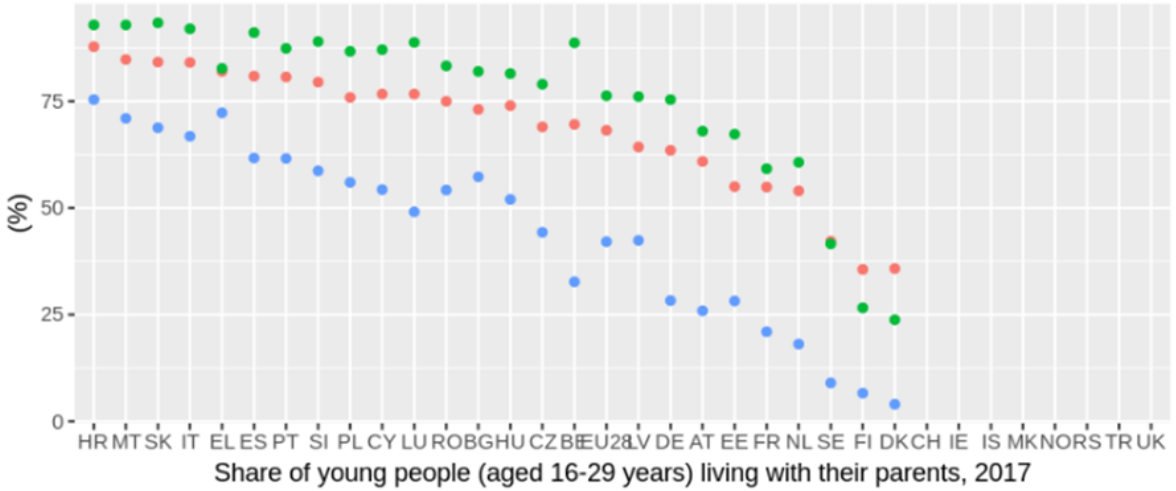


- Application programming interface (API)
- Data mining and visualization
- Text editor (Markdown)

Layout guidelines and accessibility

```
            aes(x=reorder(geo, -values), y=values, fill=sex, col=age)) +
 geom_point() +
 xlab("Share of young people (aged 16-29 years) living with their parents, 2017") +
 ylab("(%)")
print(p)
```

```
Warning message:
"Removed 24 rows containing missing values (geom_point)."
```

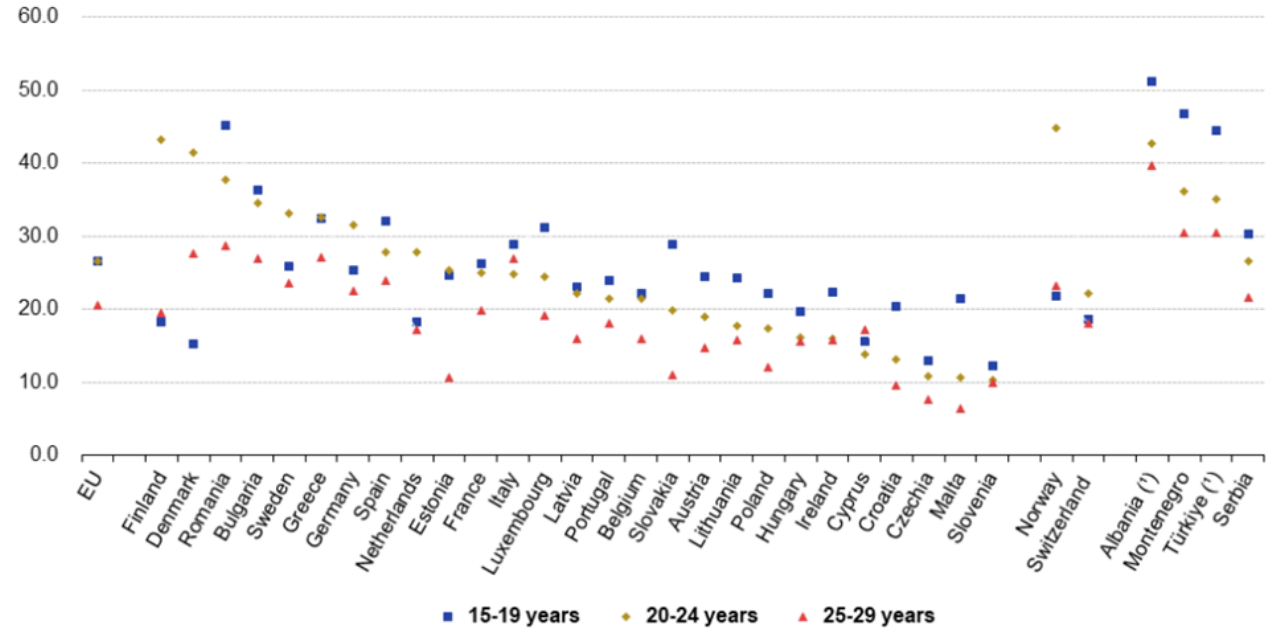



```
id = "ilc_peps01"
# id <-search_eurostat("People at risk of poverty or social exclusion by age and sex")$code[1]
dat <-get_eurostat(id, time_format = "num", filters=list(unit="PC", geo="EU28", age="Y16-29"))
head(dat)
```

unit	age	sex	geo	time	values
PC	Y16-29	F	EU28	2007	NA
PC	Y16-29	F	EU28	2008	NA
PC	Y16-29	F	EU28	2009	NA

Share of persons at risk of poverty or social exclusion for population aged 15-29 years, by detailed age groups, 2022

(%)



(1) Data from 2021

Source: Eurostat (online data code: ilc_peps01n)

eurostat

Figure 1: Share of people at risk of poverty or social exclusion for population aged 15-29 years, by detailed age group, 2022

European Commission

3. Presentation of the results by the Coding Lab participants

Consumer prices of recreational and sporting goods and services	Ricardo Dias de Carbalho *University of Porto*
Quality of life indicators - overall experience of life	Gabriella Manuti *University of Rome La Sapienza*
Young people - digital world	Joseph Nyajuoga *Örebro University*
Quality of life indicators - natural and living environment	Maripaz Venegas Gonzalez *Complutense University of Madrid*
Consumer prices of cultural goods and services	Sona Yavrumyan *University of Bologna*
Culture statistics - cultural employment	Nikola Quaresimin, *University of Bamberg*

Consumer prices of recreational and sports

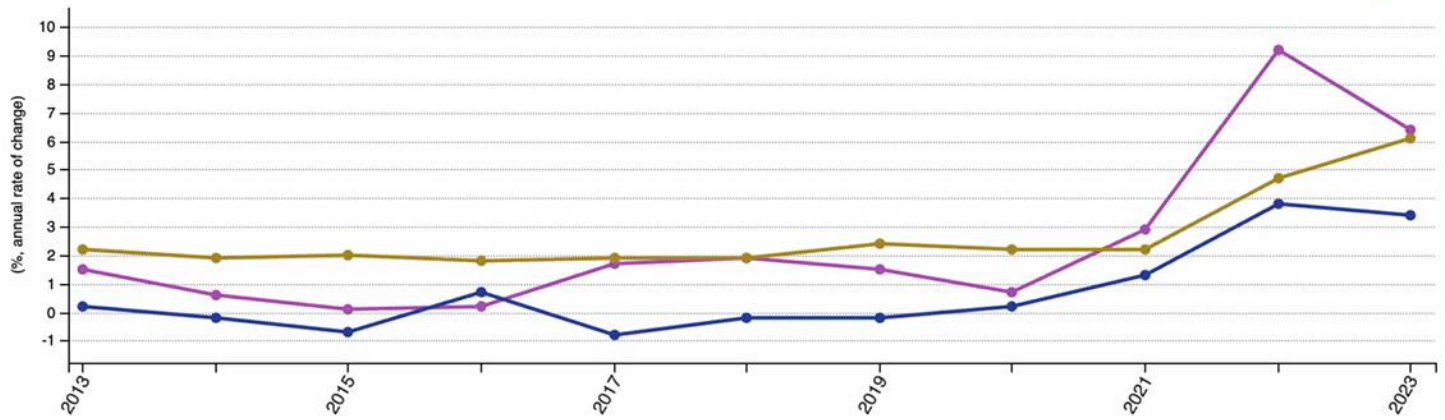
Ricardo Carvalho

Faculty of Economics - University of Porto

European Commission

Original article

The all-items HICP and the HICPs for sporting goods and services, EU, 2013-2023



Source: Eurostat (online data code: prc_hicp_aind)

eurostat

This article gives an overview of how prices of recreational and sporting goods and services have changed based on the harmonised index of consumer prices (HICP).

Full article −

HICP for sporting goods and services – trends at EU level

HICP for sporting goods and services – focus on countries

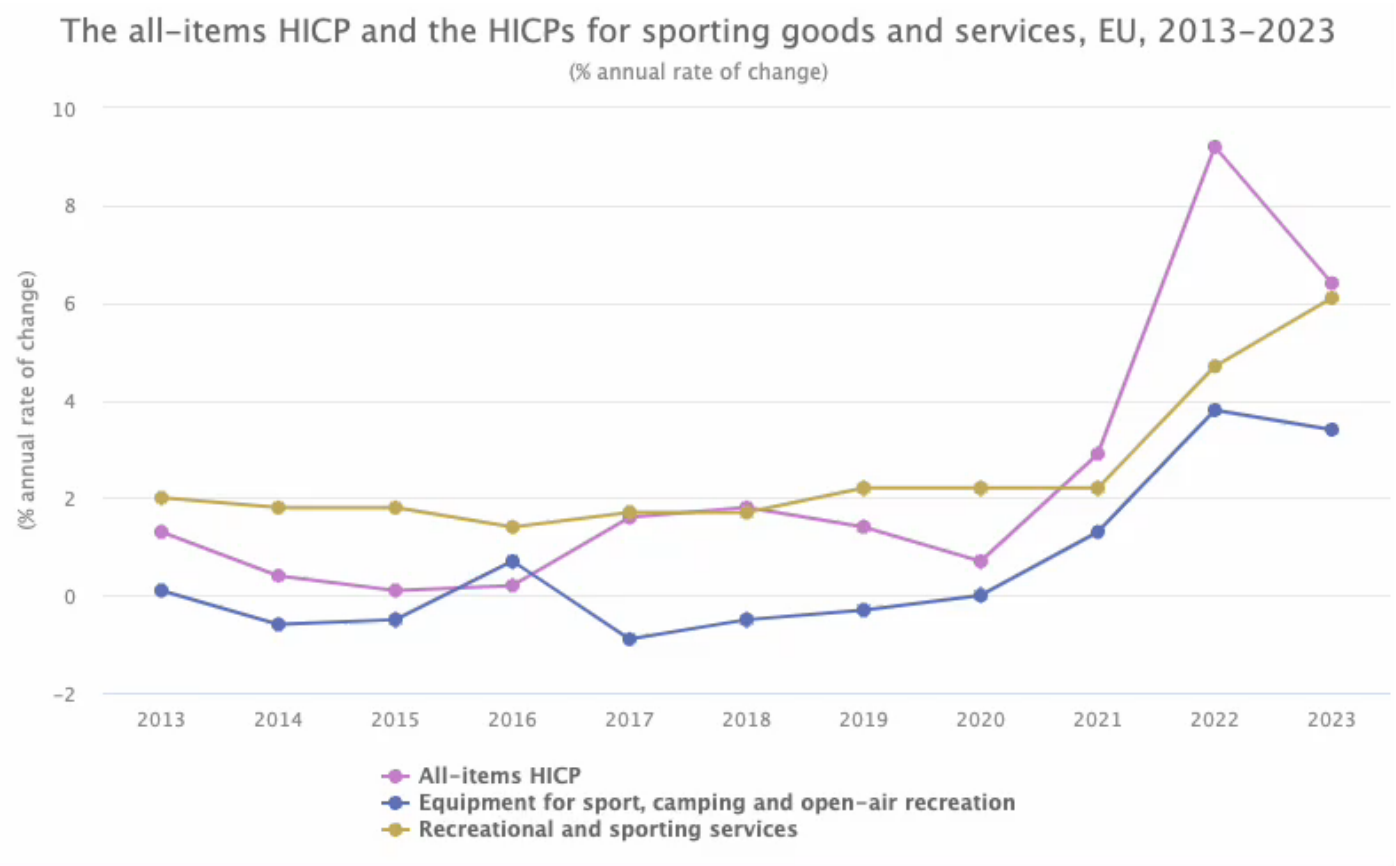
European Commission

Structure

1. Each plot
    a. Display plot
    b. Code for choosing parameters, filters
    c. Code for the plot/table

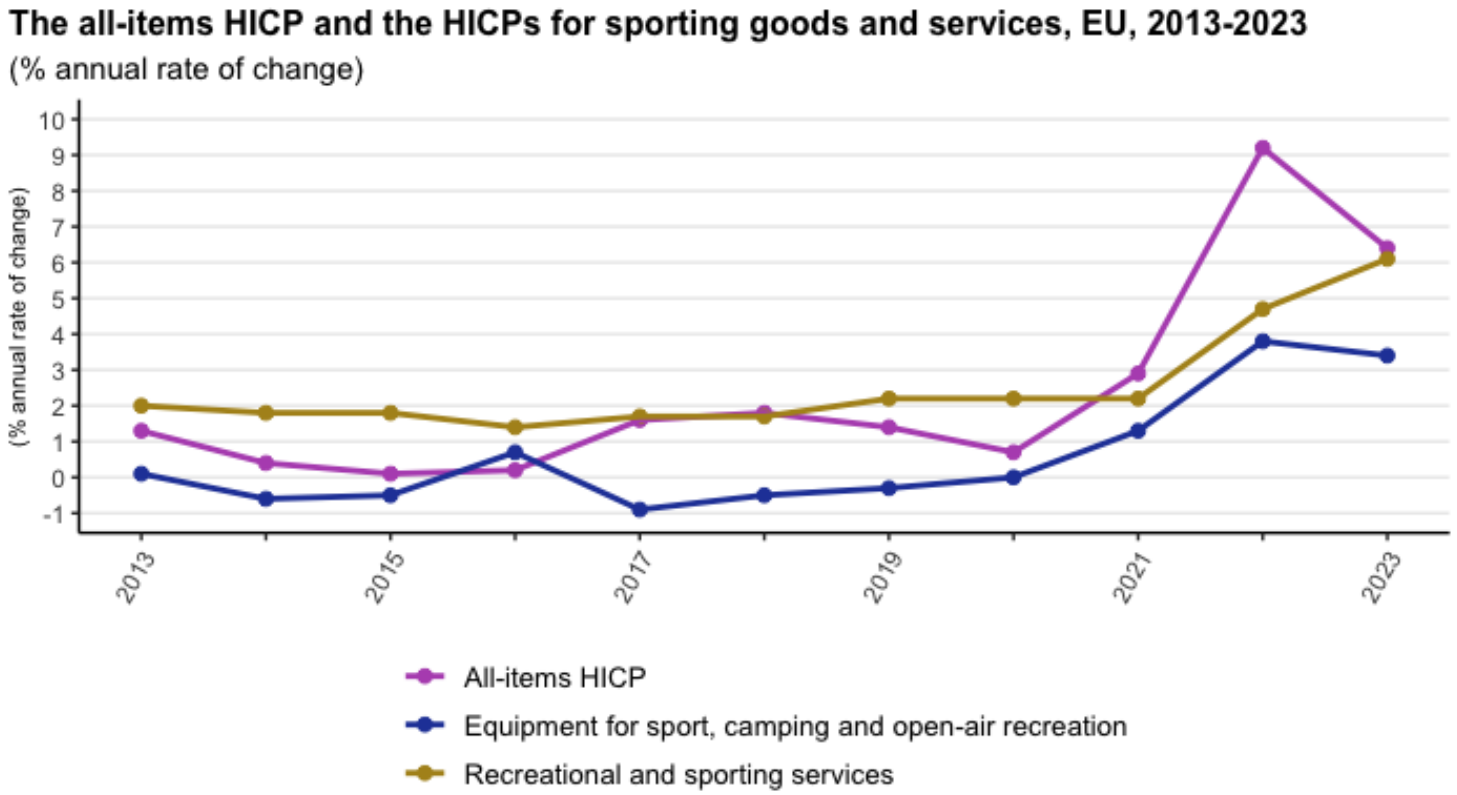
2. Code for exporting into Excel

The all-items HICP and the HICPs for sporting goods and services, EU, 2013-2023



European Commission

The all-items HICP and the HICPs for sporting goods and services, EU, 2013-2023



Note: the data refer to the official EU aggregate. Its country coverage changes in line with the addition of new EU Member States and integrates them using a chain-linked index formula.
Source: Eurostat (online data code: prc_hicp_aind)

European Commission

The all-items HICP and the HICPs for sporting goods and services, EU, 2013-2023

```
#Choose dataset and filters here
dataset <- "prc_hicp_aind"

categories <- c("All-items HICP",
                "Equipment for sport, camping and open-air recreation",
                "Recreational and sporting services")

start_year <- 2013
end_year <- 2023

geo <- "EU27_2020"
unit <- "RCH_A_AVG"

#Plot title, subtitle, caption here
title <- "The all-items HICP and the HICPs for sporting goods and services, EU, 2013-2023"
subtitle <- "(% annual rate of change)"
caption = "Note: the data refer to the official EU aggregate. Its country coverage changes in line with the addition of new EU
Member States and integrates them using a chain-linked index formula.
Source: Eurostat (online data code: prc_hicp_aind)"

#Choose colors for plot here. Eurostat pallete B
euro_palette <- c("#B655BD", "#2644A7", "#B09120", "#672DC4", "#388AE2", "#AF155C")

palette <- c("All-items HICP" = euro_palette[1],
             "Equipment for sport, camping and open-air recreation" = euro_palette[2],
             "Recreational and sporting services" = euro_palette[3])
```

The all-items HICP and the HICPs for sporting goods and services, EU, 2013-2023

```
interactive_plot <- hchart(prc_hicp_aind_eu, "line", hcaes(x = time, y = values, group = coicop)) %>%
  hc_title(text = title) %>%
  hc_subtitle(text = subtitle) %>%
  hc_xAxis(title = list(text = ""), categories = unique(prc_hicp_aind_eu$time), labels = list(format = "{value}")) %>%
  hc_yAxis(title = list(text = subtitle), min = floor(min(prc_hicp_aind_eu$values)), max = ceiling(max(prc_hicp_aind_eu$values))) %>%
  hc_tooltip(shared = TRUE, crosshairs = TRUE, pointFormat = '<b>{series.name}</b>: {point.y}<br/>') %>%
  hc_legend(align = "center", verticalAlign = "bottom", layout = "vertical") %>%
  hc_colors(euro_palette) %>%
  hc_plotOptions(line = list(marker = list(symbol = "circle", enabled = TRUE)))
```

The all-items HICP and the HICPs for sporting goods and services, EU, 2013-2023

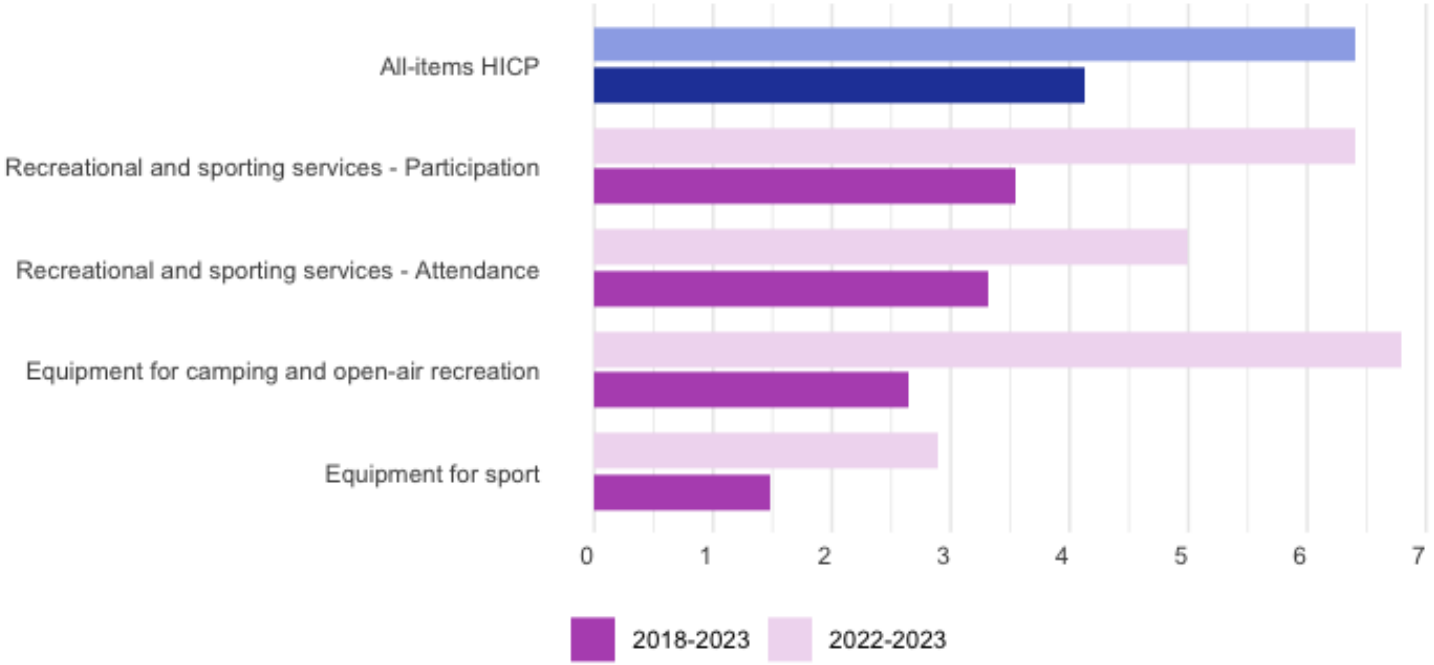
```
plot1 <- ggplot(data = prc_hicp_aind_eu, aes(x = time, y = values, group = coicop, color = coicop)) +
  geom_point(size = 2, shape = 21, aes(fill = coicop)) +
  geom_line(linewidth = 1) +

  labs(title = title,
       subtitle = subtitle,
       x = "",
       y = subtitle,
       caption = caption) +
  theme_bw() +
  theme(plot.title = element_text(size = 12, face = "bold", hjust = 0),
        axis.title.y = element_text(size = 8),
        axis.title.x = element_blank(),
        legend.text = element_text(size = 10),
        legend.position = "bottom",
        legend.title = element_blank(),
        legend.direction = "vertical",
        panel.grid.major.x = element_blank(),
        panel.grid.minor.x = element_blank(),
        panel.grid.minor.y = element_blank(),
        panel.border = element_blank(),
        axis.line = element_line(color = "black"),
        axis.text.x = element_text(angle = 60, hjust = 1),
        plot.title.position = "plot",
        plot.caption = element_text(size = 7, hjust = 0),
        plot.caption.position = "plot") +
  scale_y_continuous(limits = c(floor(min(prc_hicp_aind_eu$values)), ceiling(max(prc_hicp_aind_eu$values))), breaks = seq(floor(min(prc_hicp_aind_eu$values)), ceiling(max(prc
  scale_x_continuous(breaks = start_year:end_year, labels = function(x) ifelse(x %% 2 == 1, x, "")) +
  scale_color_manual(values = palette) +
  scale_fill_manual(values = palette)
```

European Commission

Harmonised indices of consumer prices for selected sporting goods and services

Harmonised indices of consumer prices for selected sporting goods and services, EU, annual average rates of change 2018-2023 and 2022-2023

(%)



Note: the data refer to the official EU aggregate. Its country coverage changes in line with the addition of new EU Member States and integrates them using a chain-linked index formula.
Source: Eurostat (online data code: prc_hicp_aind)

European Commission

Harmonised indices of consumer prices for selected sporting goods and services

```
#Choose dataset and filters here
dataset <- "prc_hicp_aind"

group <- "All-items HICP"
categories <- c(group,
                "Recreational and sporting services - Participation",
                "Recreational and sporting services - Attendance",
                "Equipment for camping and open-air recreation", "Equipment for sport","")

start_year_1 <- 2018
start_year_2 <- 2022
end_year <- 2023

geo <- "EU27_2020"
unit <- "RCH_A_AVG"

#Define plot title, subtitle and colors here
title <- "Harmonised indices of consumer prices for selected sporting goods and services, EU,\nannual average rates of chang
subtitle <- "(%)"

#color of year group 1 in each category
color_y1 <- "#B656BD"
#color of year group 2 in each category
color_y2 <- "#f0dcf1"
#color of year group 1 in all
color_y1_all <- "#2644A7"
#color of year group 2 in all
color_y2_all <- "#9cade8"
```

European Commission

Harmonised indices of consumer prices for selected sporting goods and services

```
prc_hicp_aind_rch_eu <- get_eurostat_data(dataset, filters = list(geo = geo, coicop = categories, unit = unit), date_filter=seq(start_year_1,end_year,1), label = T)
prc_hicp_aind_rch_eu$time <- as.numeric(as.character(prc_hicp_aind_rch_eu$time))

year_group_1 <- paste(start_year_1, end_year, sep = "-")
year_group_2 <- paste(start_year_2, end_year, sep = "-")

prc_hicp_aind_g1 <- prc_hicp_aind_rch_eu[time >= start_year_1+1 & time <= end_year, .(mean(values)), by = coicop]
names(prc_hicp_aind_g1)[2] <- year_group_1
prc_hicp_aind_g2 <- prc_hicp_aind_rch_eu[time >= start_year_2+1 & time <= end_year, .(mean(values)), by = coicop]
names(prc_hicp_aind_g2)[2] <- year_group_2

prc_hicp_aind_merged <- merge(prc_hicp_aind_g1, prc_hicp_aind_g2, on = "coicop")
prc_hicp_aind_merged <- reshape2::melt(prc_hicp_aind_merged, id.vars = "coicop")

group_year_group_1 <- paste(group, year_group_1)
group_year_group_2 <- paste(group, year_group_2)
categories_year_group_1 <- paste("Others", year_group_1)
categories_year_group_2 <- paste("Others", year_group_2)

prc_hicp_aind_merged$group <- ifelse(prc_hicp_aind_merged$coicop == group&prc_hicp_aind_merged$variable == year_group_1, group_year_group_1,
                                     ifelse(prc_hicp_aind_merged$coicop == group&prc_hicp_aind_merged$variable == year_group_2, group_year_group_2,
                                            ifelse(prc_hicp_aind_merged$variable == year_group_1, categories_year_group_1, categories_year_group_2)))


unique_coicop <- unique(as.character(prc_hicp_aind_merged$coicop))
unique_coicop <- unique_coicop[unique_coicop != group]
ordered_levels <- c(unique_coicop, group)
prc_hicp_aind_merged$coicop <- factor(prc_hicp_aind_merged$coicop, levels = ordered_levels)
```

Harmonised indices of consumer prices for selected sporting goods and services

```
plot2 <- ggplot(prc_hicp_aind_merged, aes(x = coicop, y = value, fill = group)) +
  geom_bar(stat = "identity", position = position_dodge(width = 0.8), width = 0.7) +
  labs(title = title,
       subtitle = subtitle,
       fill = "coicop",
       caption = caption) +
  scale_fill_manual(values = colors,
                    breaks = c(categories_year_group_1,categories_year_group_2),
                    labels = c(year_group_1,year_group_2)) +
  theme_minimal() +
  coord_flip() +
  theme(legend.position = "bottom",
        legend.direction = "horizontal",
        legend.title = element_blank(),
        axis.text.x = element_text(hjust = 1),
        axis.title.x = element_blank(),
        axis.title.y = element_blank(),
        plot.title = element_text(size = 12, face = "bold"),
        plot.caption = element_text(size = 7, hjust = 0),
        plot.caption.position = "plot",
        panel.grid.major.y = element_blank(),
        plot.title.position = "plot",
        legend.justification = c(0,0)
        ) +
  scale_y_continuous(breaks = seq(0,ceiling(max(prc_hicp_aind_merged$value)),1))

print(plot2)
dev.off()
```

HICP for sporting goods and services – focus on countries

	All-items HICP		Equipment for sport		Equipment for camping and open-air recreation		Recreational and sporting services - Attendance		Recreational and sporting services - Participation	
Countries	2018-2023	2022-2023	2018-2023	2022-2023	2018-2023	2022-2023	2018-2023	2022-2023	2018-2023	2022-2023
European Union - 27 countries (from 2020)	**4.1**	**6.4**	**1.5**	**2.9**	**2.6**	**6.8**	**3.3**	**5.0**	**3.5**	**6.4**
Belgium	3.5	2.3	1.7	4.6	2.5	7.0	4.2	6.8	3.0	5.8
Bulgaria	5.6	8.6	2.2	6.5	NaN	NaN	NaN	NaN	8.8	15.8
Czechia	7.2	12.0	5.0	7.5	0.2	7.6	9.5	22.6	8.3	14.2
Denmark	3.0	3.4	-1.7	-1.7	-3.0	-7.9	4.3	3.6	3.1	4.7
Germany	3.9	6.0	2.1	0.9	2.6	5.6	3.3	4.5	2.6	5.0
Estonia	6.9	9.1	4.1	5.9	NaN	NaN	6.9	12.8	8.6	10.6
Ireland	3.2	5.2	-2.8	1.5	0.0	4.5	3.2	4.2	3.3	6.0
Greece	2.7	4.2	-0.5	1.8	-0.1	1.4	0.7	2.8	0.8	3.2
Spain	3.0	3.4	0.0	-0.7	0.0	NaN	5.4	9.0	1.0	2.8
France	3.1	5.7	1.9	3.5	3.4	8.1	1.7	1.9	1.9	4.3
Croatia	4.5	8.4	0.5	2.5	-1.2	4.6	7.9	13.6	5.6	14.1

HICP for sporting goods and services – focus on countries

```
#Choose dataset and filters here
dataset <- "prc_hicp_aind"

categories <- c("All-items HICP",
                "Equipment for sport",
                "Equipment for camping and open-air recreation",
                "Recreational and sporting services - Participation",
                "Recreational and sporting services - Attendance"
                )

start_year_1 <-2018
start_year_2 <- 2022
end_year <- 2023

unit <- "RCH_A_AVG"

#Define header colors here
color_a <- "#f0dcf1"
color_b <- "#e1bae4"
```

European Commission

HICP for sporting goods and services – focus on countries

```
prc_hicp_aind <- get_eurostat_data(dataset, filters = list(unit = unit, coicop = categories), date_filter=seq(start_year_1,end_year,1), label = T)

prc_hicp_aind$time <- as.numeric(as.character(prc_hicp_aind$time))
pivot_table<- dcast(prc_hicp_aind, geo + coicop ~ time, value.var = "values", fun.aggregate = mean)

start_1_col <- which(names(pivot_table) == start_year_1) + 1
start_2_col <- which(names(pivot_table) == start_year_2) + 1
end_col <- which(names(pivot_table) == end_year)

pivot_table[, start_1_col:end_col] <- sapply(pivot_table[, start_1_col:end_col], as.numeric)

year_group_1 <- paste(start_year_1, end_year, sep = "-")
year_group_2 <- paste(start_year_2, end_year, sep = "-")

#Computing mean of year groups
if (ncol(pivot_table[, start_1_col:end_col, drop = FALSE]) == 1) {
  pivot_table[[year_group_1]] <- pivot_table[, end_col]
} else {
  pivot_table[[year_group_1]] <- round(rowMeans(pivot_table[, start_1_col:end_col], na.rm = TRUE),1)
}

if (ncol(pivot_table[, start_2_col:end_col, drop = FALSE]) == 1) {
  pivot_table[[year_group_2]] <- pivot_table[, end_col]
} else {
  pivot_table[[year_group_2]] <- round(rowMeans(pivot_table[, start_2_col:end_col], na.rm = TRUE),1)
}
```

HICP for sporting goods and services – focus on countries

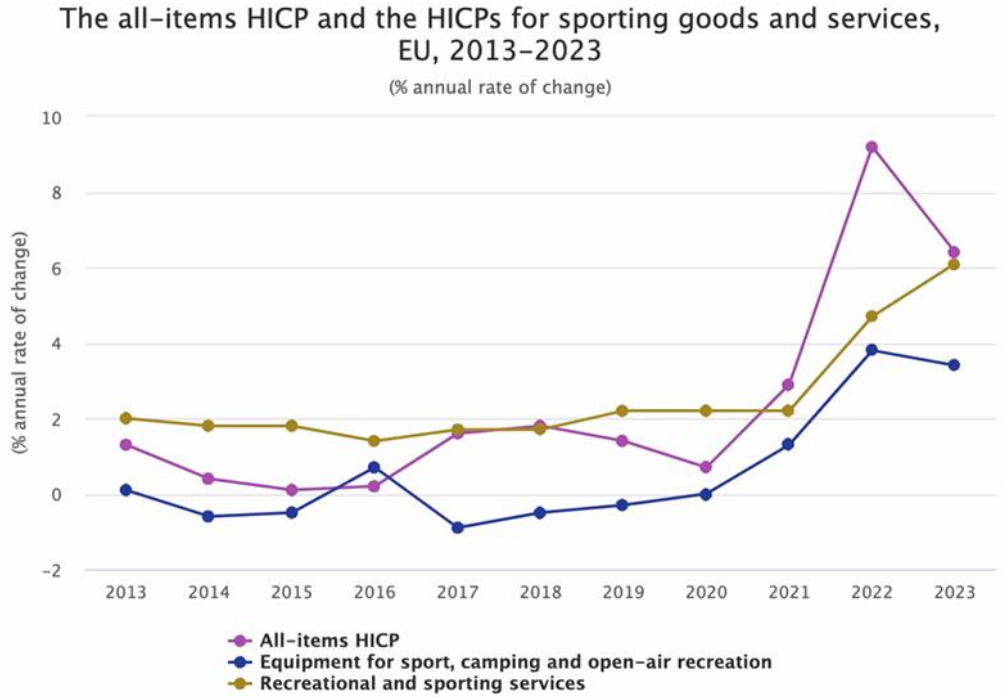
```
#Creating table
kable_styling <- kable(final_df, format = "html", escape = F, digits = 1,
                       row.names = FALSE) %>%
  add_header_above(header, background = color_a) %>%
  row_spec(row = 0, background = color_a, extra_css = "white-space: nowrap; border: none;") %>%
  row_spec(row = 1, background = color_b, bold = TRUE, extra_css = "border: none;") %>%
  column_spec(1, bold=TRUE)

groups <- unique(groups_info)
group_indices <- split(seq_len(nrow(final_df)), groups_info)

for (group in groups) {
  kable_styling <- kable_styling %>%
    group_rows(" ", min(group_indices[[group]]), max(group_indices[[group]]),
               label_row_css = "border-bottom: 1px solid black; margin: 0; padding: 0;")
}

kable_styling <- kable_styling %>%
  kable_styling(bootstrap_options = c("striped", "hover"),
                full_width = F, font_size = 12, fixed_thead = TRUE)

kable_styling
```

Export to Excel/png

```
# EXCEL ------------------------------------------------------------------------
wb <- createWorkbook()
sheet_1 <- createSheet(wb, sheetName = "HICP by year")
addPicture("plot1.png", sheet_1, scale = 1, startRow = 1, startColumn = 1)
sheet_2 <- createSheet(wb, sheetName = "Harmonised indices")
addPicture("plot2.png", sheet_2, scale = 1, startRow = 1, startColumn = 1)
sheet_3 <- createSheet(wb, sheetName = "Countries")
addDataFrame(final_df, sheet_3, row.names = FALSE, startRow = 1, startColumn = 1)

saveWorkbook(wb, "Consumer prices of recreational and sporting goods and services.xlsx")


#Export the table as html and png
library(webshot)
library(htmltools)
webshot::install_phantomjs(force = TRUE)

save_kable(kable_styling, file = "table1.html")
webshot("table1.html", "table1.png", selector = "table")
```

Article

Consumer prices of recreational and sporting goods and services



Source: Eurostat (online data code: prc_hicp_aind)

This article gives an overview of how prices of recreational and sporting goods and services have changed based on the harmonised index of consumer prices (HICP).

Thank you!

Ricardo Carvalho

Faculty of Economics - University of Porto

European Commission

Consumer prices of recreational and sporting goods and services	Ricardo Dias de Carbalho *University of Porto*
Quality of life indicators - overall experience of life	Gabriella Manuti *University of Rome La Sapienza*
Young people - digital world	Joseph Nyajuoga *Örebro University*
Quality of life indicators - natural and living environment	Maripaz Venegas Gonzalez *Complutense University of Madrid*
Consumer prices of cultural goods and services	Sona Yavrumyan *University of Bologna*
Culture statistics - cultural employment	Nikola Quaresimin, *University of Bamberg*

Quality of life indicators - Overall experience of life

Gabriella Manuti

Sapienza University of Rome | ENSAI

European Commission

Overall life satisfaction, 2022

```
plot0 <- ggplot(countries,aes(x = reorder(geo, values[geo != "European Union - 27 countries (from 2020)"],
decreasing = TRUE), y = values)) +
  geom_col(width = 0.6, fill = "#2644A7") +
  geom_hline(yintercept = EU_value, color = "#B09120",linetype = "solid",linewidth =1) +
  scale_y_continuous(expand = c(0, 0), limits = c(0, 10),breaks = seq(0, 10, by = 1)) +
  scale_x_discrete(labels = countries$geo) +
  coord_cartesian(ylim = c(0, 10)) +
  theme(panel.grid.major.y = element_line(color = "gray", linetype = "dashed"),
        panel.grid.minor = element_blank(),
        panel.grid.major.x = element_blank(),
        axis.text.x = element_text(angle = 60, vjust = 1, hjust = 1, color="black"),
        axis.text.y = element_text(vjust = 0.5, hjust = 0, color="black"),
        panel.background = element_rect(fill = "white"),
        axis.line.x = element_line(colour = "black", linetype = "solid"),
        axis.ticks = element_line(),
        axis.ticks.length.y  = unit(0, "cm"),
        plot.title = element_text(hjust = 0, vjust = 0.5, face = "bold"),
        plot.title.position = "plot",
        plot.margin = unit(c(1, 1, 3, 1), "lines"),
        plot.caption.position = "plot",
        plot.caption = element_text(size = 7, hjust = 0),
        legend.box.margin = margin(t = 10, unit = "pt"),
        legend.position = "bottom",
        legend.direction = "horizontal",
        legend.title = element_blank(),)+
  labs(x = NULL, y = NULL, title="Overall life satisfaction, 2022", subtitle= "(mean scale 0-10)", caption=
"Figure 1: Overall life satisfaction, 2022\n(mean scale 0-10)\nSource: Eurostat: ilc_pw01")+
  scale_fill_manual(name = NULL,
                    values = c("Overall life satisfaction" = "#2644A7", "EU average" = "#B09120"),
                   labels = c("Overall life satisfaction", "EU average"))
```

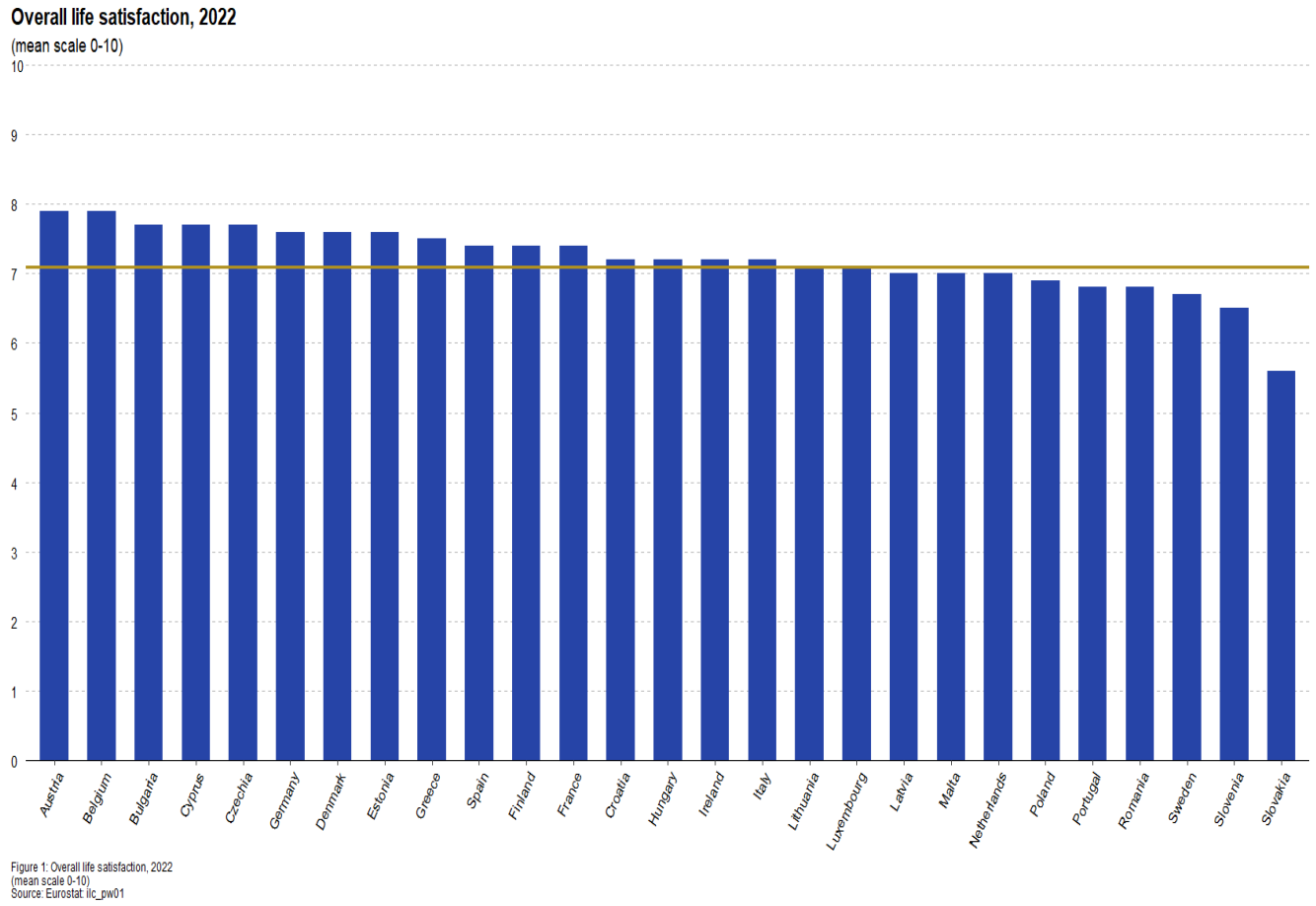


Figure 1: Overall life satisfaction, 2022
(mean scale 0-10)
Source: Eurostat: ilc_pw01

European Commission

Overall life satisfaction by age group, 2022

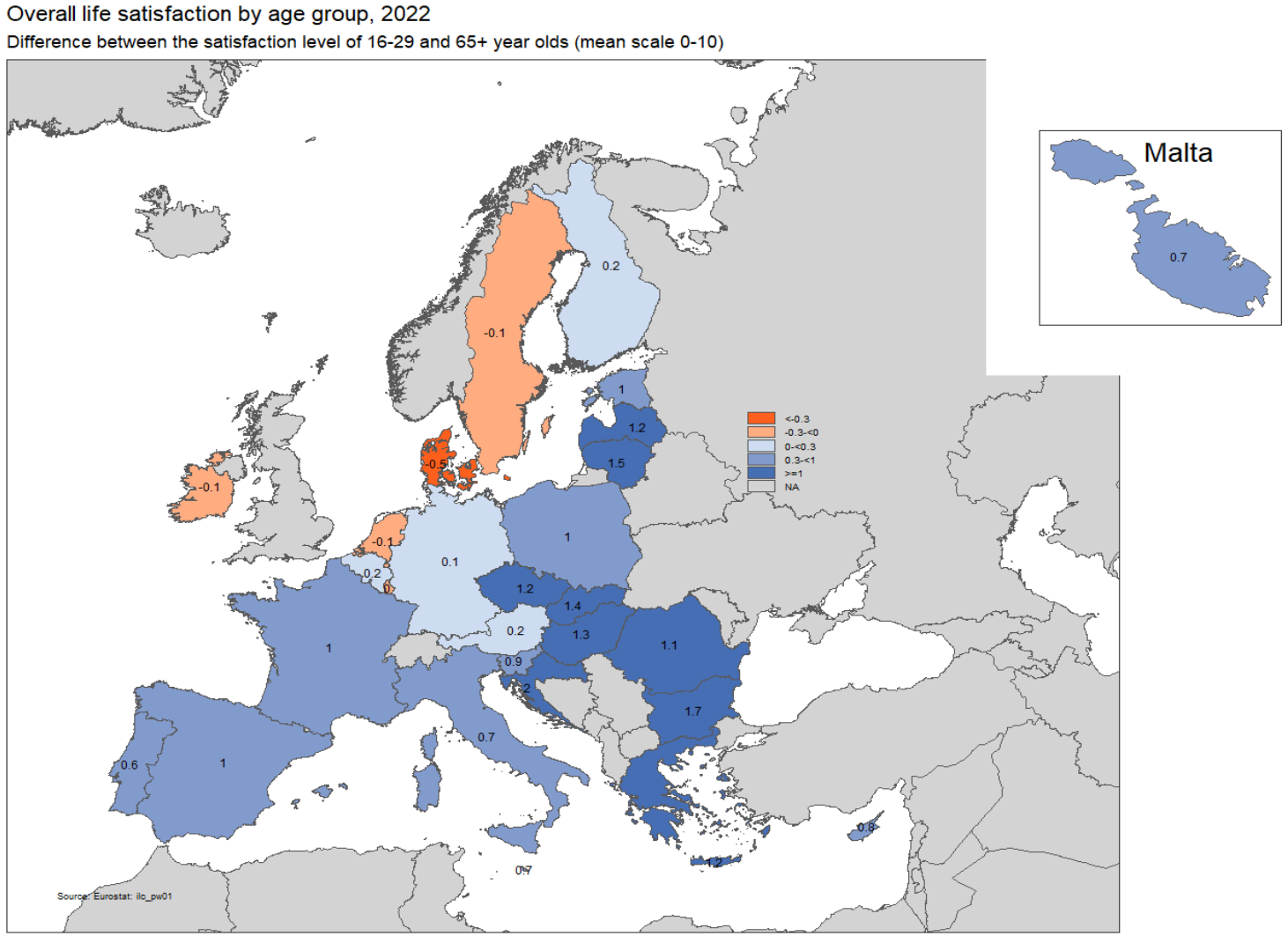
```
countries_diff <- countries_map %>%
  pivot_wider(names_from = age, values_from = values) %>%
  mutate(difference = `Y16-19` - `Y_GE65`)

merged_data$bins <- cut(
  merged_data$difference,
  breaks = c(-Inf, -0.3, 0, 0.3, 1, Inf),
  labels = c("<-0.3", "-0.3-<0", "0-<0.3", "0.3-<1", ">=1"),
  include.lowest = TRUE
)

malta_map1 <- ggplot(malta_data) +
  geom_sf(aes(fill = bins)) +
  geom_sf_text(aes(label = round(difference, 1)), size = 2.5, color = "black") +
  scale_fill_manual(values = c(
    "<-0.3" = "#FF5F1D",
    "-0.3-<0" = "#FFAC81",
    "0-<0.3" = "#CDDCF0",
    "0.3-<1" = "#829BCD",
    ">=1" = "#466EB4"),
    na.value = "lightgrey") +
  labs(fill = "") +
  guides(fill = guide_colorbar(reverse = TRUE)) +
  theme_bw() +
  theme(axis.text.x = element_blank(),
        axis.text.y = element_blank(),
        axis.title.x = element_blank(),
        axis.title.y = element_blank(),
        panel.grid = element_blank(),
        axis.ticks = element_blank(),
        legend.position = "none",
        plot.margin = unit(c(3, 1, 1, 1), "cm") ) +
  annotate("text", x = -Inf, y = Inf, label = "Malta", vjust = 1.5, hjust = -1.5, size = 6, fontface = "plain")
```

Overall life satisfaction by age group, 2022

```
europe_map <- ggplot(merged_data) +
  geom_sf(aes(fill = bins)) +
  geom_sf_text(aes(label = round(difference, 1)), size = 2.5, color = "black") +
  scale_fill_manual(values = c(
    "<-0.3" = "#FF5F1D",
    "-0.3-<0" = "#FFAC81",
    "0-<0.3" = "#CDDCF0",
    "0.3-<1" = "#829BCD",
    ">=1" = "#466EB4"),
  na.value = "lightgrey"
  ) +
  theme_bw() +
  coord_sf(crs = 3035, xlim = c(2377294, 7453440), ylim = c(1313597, 5628510)) +
  labs(title = "Overall life satisfaction by age group, 2022",
       subtitle= "Difference between the satisfaction level of 16-29 and 65+ year olds (mean scale 0-10)") +
  guides(fill = guide_legend(keywidth = unit(0.6, "cm"), keyheight = unit(0.3, "cm"), drop = TRUE)) +
  annotate("text", x = 2380000, y = 1300000,
           label = "Source: Eurostat: ilc_pw01",
           size = 2, color = "black", hjust = 0) +
  theme(legend.position = c(0.7, 0.55),
        legend.text = element_text(size = 6),
        legend.title = element_blank(),
        legend.key.size = unit(0.3, "cm"),
        legend.background = element_blank(),
        axis.text.x = element_blank(),
        axis.text.y = element_blank(),
        axis.title.x = element_blank(),
        axis.title.y = element_blank(),
        panel.grid = element_blank(),
        axis.ticks = element_blank(),
        panel.background = element_blank()
  )
map1<- ggdraw(europe_map) +
  draw_plot(malta_map1, 0.6, 0.6, .4, .4)
```

European Commission

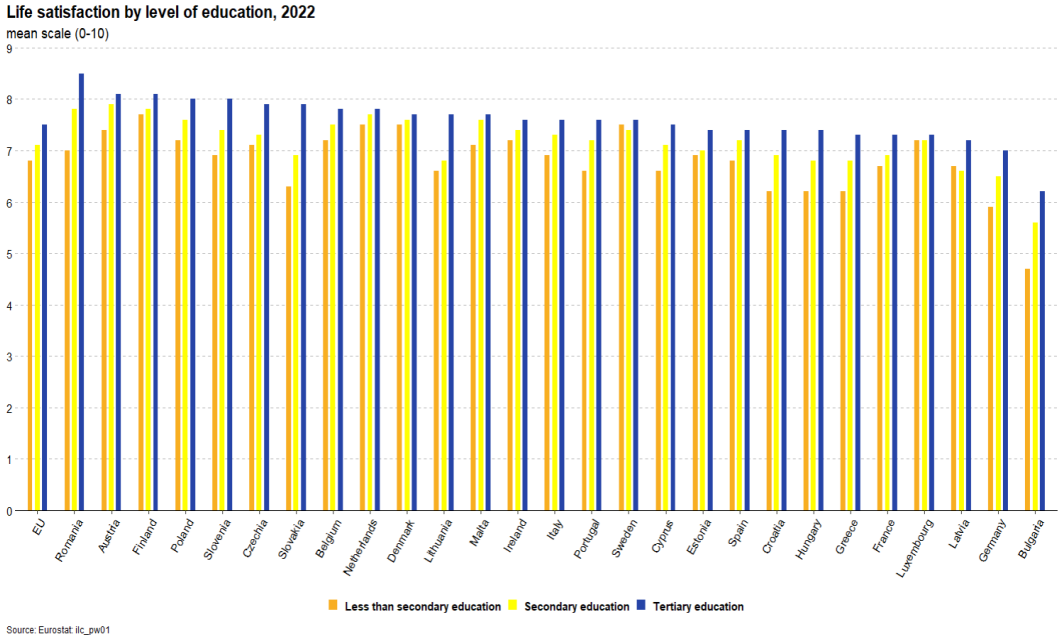
Overall life satisfaction by agre group, 2022



European Commission

Overall life satisfaction by education level

```r
#countries subset to select the EU_value and the other countries for the plot
eu_data <- subset(education_countries, geo == "European Union - 27 countries (from 2020)")
country_data <- subset(education_countries, geo != "European Union - 27 countries (from 2020)")

#order the countries by the tertiary education level
country_data <- country_data %>%
  arrange(desc(ifelse(isced11 == "Tertiary education (levels 5-8)", values, NA)))

#re-arrange the dataset
ordered_data <- bind_rows(eu_data, country_data)
```

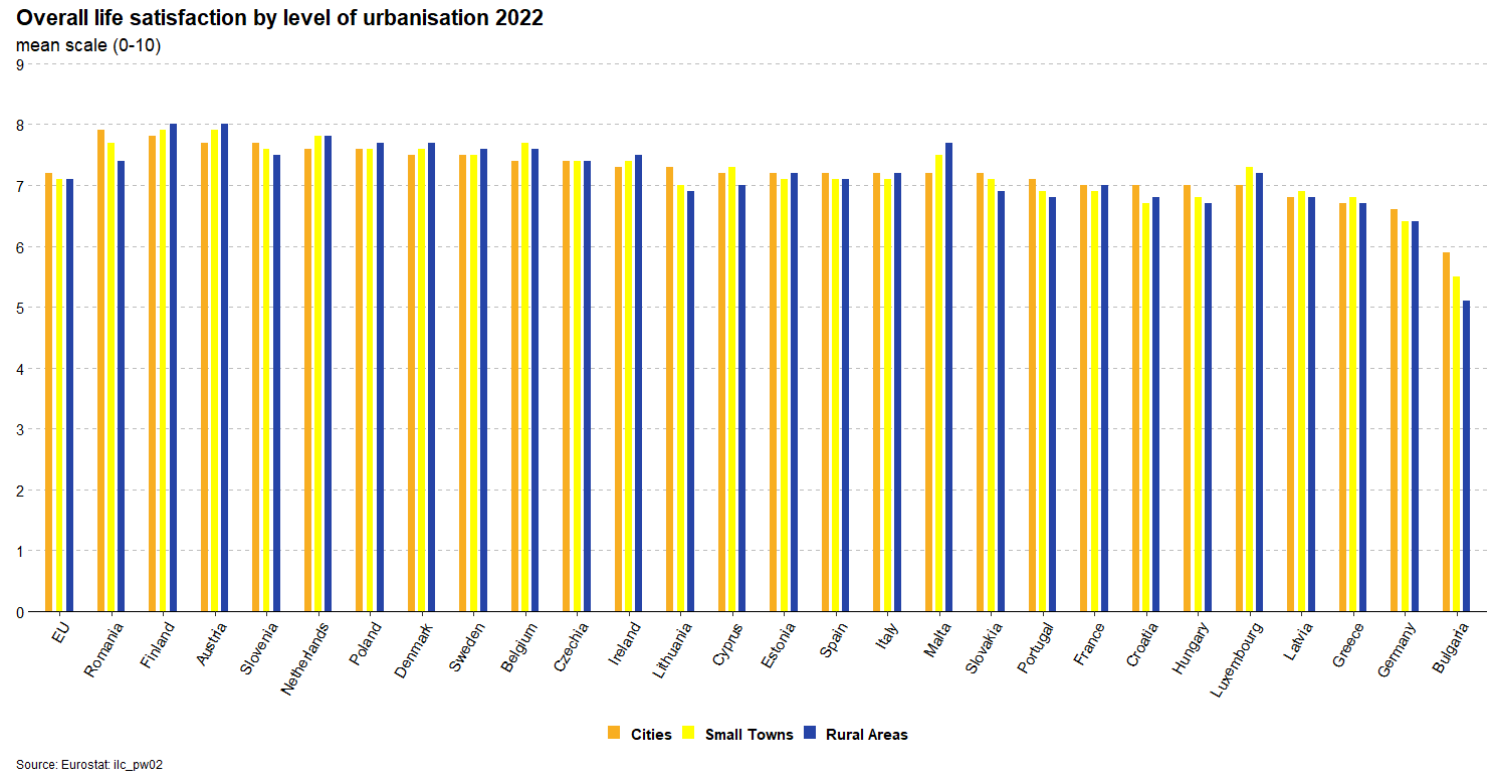



```r
plot2 <- ggplot(ordered_data, aes(x = factor(geo, levels = country_names), y = values, fill = isced11)) +
  scale_fill_manual(values = custom_colors, labels = c("Less than secondary education", "Secondary education",
"Tertiary education")) +
  geom_bar(stat = "identity", position = position_dodge(width = 0.6), width = 0.4) +
  labs(title = "Life satisfaction by level of education, 2022",subtitle ="mean scale (0-10)", x = NULL, y = NULL,
caption="Source: Eurostat: ilc_pw01") +
  scale_y_continuous(expand = c(0, 0), limits = c(0, 9), breaks = seq(0, 9, by = 1)) +
  coord_cartesian(ylim = c(0, 9)) +
  theme(plot.title = element_text(hjust = 0, vjust = 0.5, face = "bold"),
        plot.title.position = "plot",
        plot.margin = unit(c(1, 1, 3, 1), "lines"),
        plot.caption.position = "plot",
        plot.caption = element_text(size = 7, hjust = 0),
        panel.grid.major.y = element_line(color = "gray", linetype = "dashed"),
        panel.background = element_rect(fill = "white"),
        panel.grid.major.x = element_blank(),
        panel.grid.minor = element_blank(),
        axis.text.x = element_text(angle = 60, vjust = 1, hjust = 1, color = "black"),
        axis.text.y = element_text(vjust = 0.5, hjust = 0, color = "black"),
        axis.ticks = element_line(),
        axis.ticks.length.y = unit(0, "cm"),
        axis.line.x = element_line(colour = "black", linetype = "solid"),
        legend.position = "bottom",
        legend.title = element_blank(),
        legend.text= element_text(face="bold")) +
  guides(fill = guide_legend(keywidth = unit(0.3, "cm"), keyheight = unit(0.2, "cm")))+
  scale_shape_manual(values = 16)+
  scale_x_discrete(labels = function(x) ifelse(x == 'European Union - 27 countries (from 2020)', 'EU', x))
```

European Commission

Overall life satisfaction by urbanisation level

```r
#countries subset to select the EU_value and the other countries for the plot
eu_data_urb <- subset(urbanisation_countries, geo == "European Union - 27 countries (from 2020)")
country_data_urb <- subset(urbanisation_countries, geo != "European Union - 27 countries (from 2020)")

#order the countries by the cities degree of urbanisation
country_data_urb <- country_data_urb %>%
  arrange(desc(ifelse(deg_urb == "Cities", values, NA)))

#re-arrange the data
ordered_data_urb <- bind_rows(eu_data_urb, country_data_urb)
country_names_urb <- unique(ordered_data_urb$geo)
```



Source: Eurostat: ilc_pw02

European Commission

Overall life satisfaction by demographic characteristics at EU level, 2018 and 2022

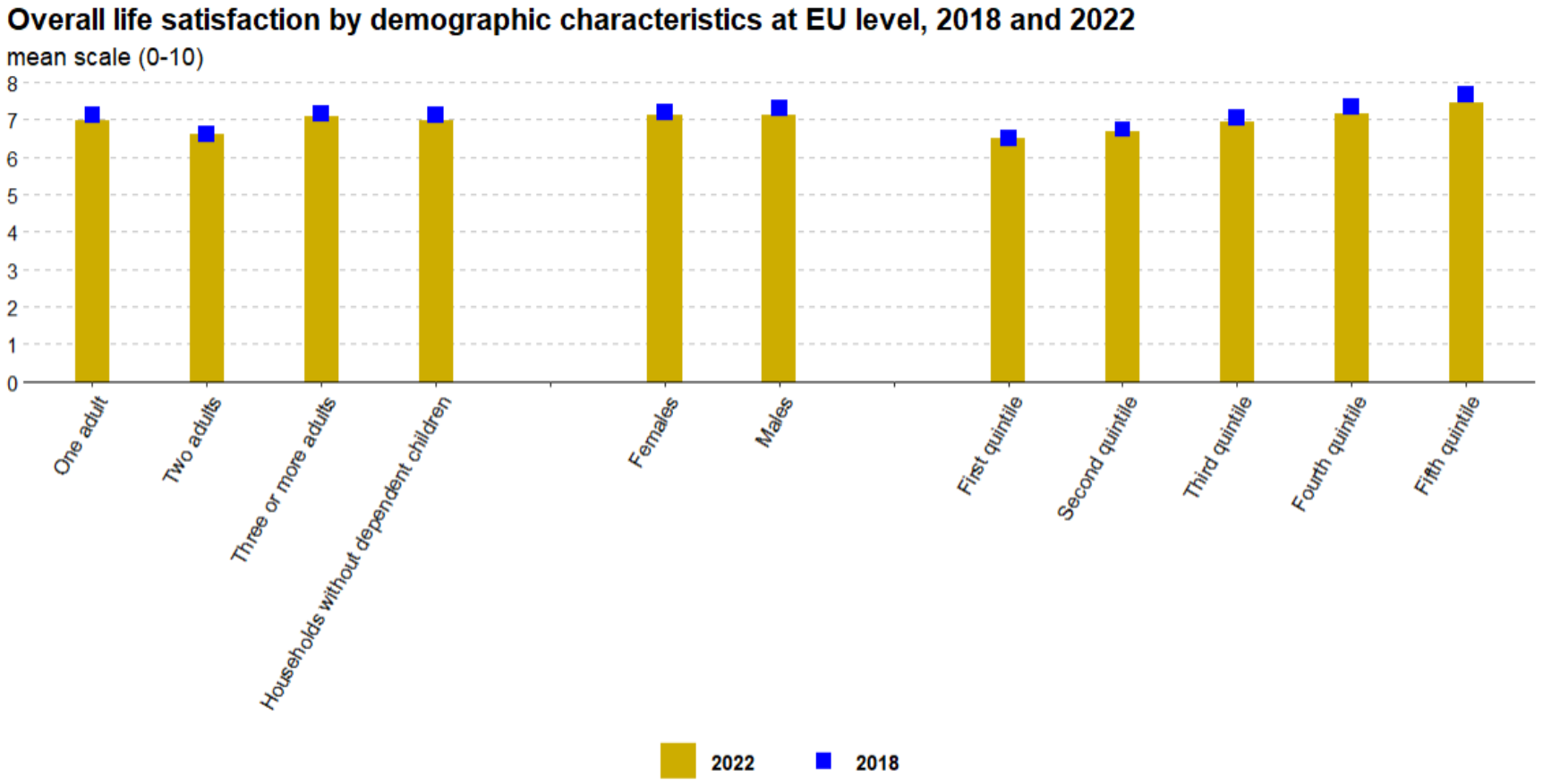
```
#data on the household composition
hhcomp_data <- household_countries %>%
  filter(hhcomp != "Total") %>%
  mutate(category = hhcomp)

#data on the income
quant_inc_data <- household_countries %>%
  filter(quant_inc != "Total") %>%
  mutate(category = quant_inc)

#data on the gender (ilc_pw01)
gender_data <- subset(ilc_pw01, sex %in% c("Males", "Females") & age == "16 years or over" & time %in% c("2018",
"2022") & geo == "European Union - 27 countries (from 2020)") %>%
  mutate(category = sex)

combined_data <- bind_rows(hhcomp_data,gender_data, quant_inc_data)

#only one row per category and year by aggregating the data
aggregated_data <- combined_data %>%
  group_by(category, time) %>%
  summarise(values = mean(values, na.rm = TRUE)) %>%
  ungroup()

#split the aggregated data in 2018 and 2022
data_2018 <- filter(aggregated_data, time == 2018)
data_2022 <- filter(aggregated_data, time == 2022)

#add spacing rows for a better visualization in the plot
cat_1_space <- length(unique(hhcomp_data$category))
cat_2_space <- cat_1_space + length(unique(gender_data$category))
```

Overall life satisfaction by demographic characteristics at EU level, 2018 and 2022

```
data_2022 <- data_2022 %>%
  mutate(category = as.character(category)) %>%
  add_row(category = "", time = factor(2022, levels = levels(data_2022$time)), values = NA, .before = (cat_1_space+1)) %>%
  add_row(category = " ", time = factor(2022, levels = levels(data_2022$time)), values = NA, .before = (cat_2_space+2))

data_2018 <- data_2018 %>%
  mutate(category = as.character(category)) %>%
  add_row(category = "", time = factor(2018, levels = levels(data_2018$time)), values = NA, .before = (cat_1_space+1)) %>%
  add_row(category = " ", time = factor(2018, levels = levels(data_2018$time)), values = NA, .before = (cat_2_space+2))

data_2022$category <- factor(data_2022$category, levels = unique(data_2022$category))
data_2018$category <- factor(data_2018$category, levels = unique(data_2018$category))

plot4 <- ggplot() +
  geom_bar(data = data_2022, aes(x = category, y = values, fill = time),
           stat = "identity", width = 0.3) +
  geom_point(data = data_2018, aes(x = category, y = values, color = "2018"),
             shape = 15, size = 3) +
  labs(title = "Overall life satisfaction by demographic characteristics at EU level, 2018 and 2022",
       subtitle = "mean scale (0-10)", x = NULL, y = NULL,caption="All data are estimates\nSource: Eurostat: ilc_pw01,
ilc_pw02") +
  scale_y_continuous(expand = c(0, 0), limits = c(0, 8), breaks = seq(0, 8, by = 1)) +
  coord_cartesian(ylim = c(0, 8)) +
  theme(plot.title = element_text(hjust = 0, vjust = 0.5, face = "bold"),
        plot.title.position = "plot",
        plot.margin = unit(c(1, 1, 3, 1), "lines"),
        plot.caption.position = "plot",
        plot.caption = element_text(size = 7, hjust = 0),
        panel.grid.major.y = element_line(color = "gray", linetype = "dashed"),
        panel.background = element_rect(fill = "white"),
        panel.grid.major.x = element_blank(),
        panel.grid.minor = element_blank(),
        axis.text.x = element_text(angle = 60, vjust = 1, hjust = 1, color = "black"),
        axis.text.y = element_text(vjust = 0.5, hjust = 0, color = "black"),
        axis.ticks = element_line(),
        axis.ticks.length.y = unit(0, "cm"),
        axis.line.x = element_line(colour = "black", linetype = "solid"),
        legend.position = "bottom",
        legend.title = element_blank(),
        legend.text = element_text(face = "bold")) +
  scale_fill_manual(values = c("2022" = "gold3")) +
  scale_color_manual(values = c("2018" = "blue")) +
  guides(fill = guide_legend(title = "Year"))
```

European Commission

Overall life satisfaction by demographic characteristics at EU level, 2018 and 2022



All data are estimates
Source: Eurostat: ilc_pw01, ilc_pw02

European Commission

Consumer prices of recreational and sporting goods and services	Ricardo Dias de Carbalho *University of Porto*
Quality of life indicators - overall experience of life	Gabriella Manuti *University of Rome La Sapienza*
Young people - digital world	Joseph Nyajuoga *Örebro University*
Quality of life indicators - natural and living environment	Maripaz Venegas Gonzalez *Complutense University of Madrid*
Consumer prices of cultural goods and services	Sona Yavrumyan *University of Bologna*
Culture statistics - cultural employment	Nikola Quaresimin, *University of Bamberg*

Quality of life indicators - natural and living environment

Joseph Nyajuoga

Örebro University, Sweden

Maripaz González

Complutense University of Madrid, Spain

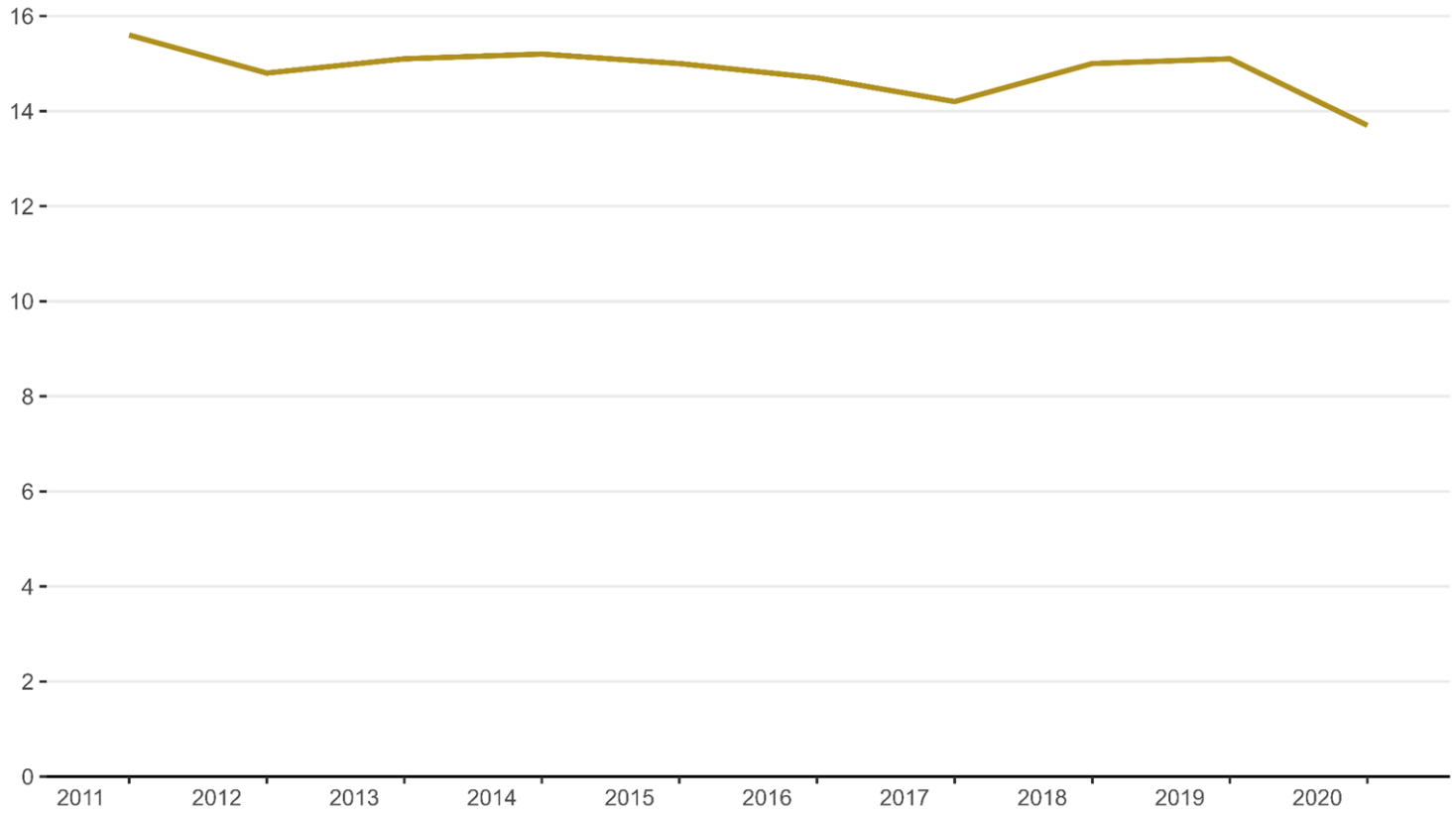
Recommended packages

- restatapi
- giscoR
- data.table
- chron
- reshape2
- kableExtra
- tidyverse
- tmap
- plotly
- highcharter

Line graph

Population reporting exposure to pollution, grime or other environmental problems, EU (¹) 2011-2020

(%)



(¹) Estimate.
Source: Eurostat (online data code: ilc_mddw02)

European Commission

Retrieving data

Parameters

```
id1<-"ilc_mddw02"
date1 = 2011:2020
```

API connection

```
data1 = get_eurostat_data(
  id1,
  filters = list(incgrp = "TOTAL",
                 hhtyp = "TOTAL",
                 geo = "EU27_2020"),
  date_filter = date1,
  label = TRUE,
  ignore.case = TRUE,
  exact_match = FALSE,
  perl = TRUE,
  stringsAsFactors = FALSE,
  force_local_filter = F,
  keep_flags = T
  )
```

Data wrangling

```
data1 = data1 |>
        group_by(time) |>
        summarise(exp = mean(values)) |>
        ungroup()
```

Plotting

Figure's text

```
title <- paste0("Population reporting exposure to
pollution, grime or other\nenvironmental
problems, EU (\U00B9) ", date1[1], "-",
date1[length(date1)])
subtitle <- "(%)"
caption <- paste0("\n(\U00B9) Estimate.",
                  "\nSource: Eurostat (online
data code: ", id1, ")")
```

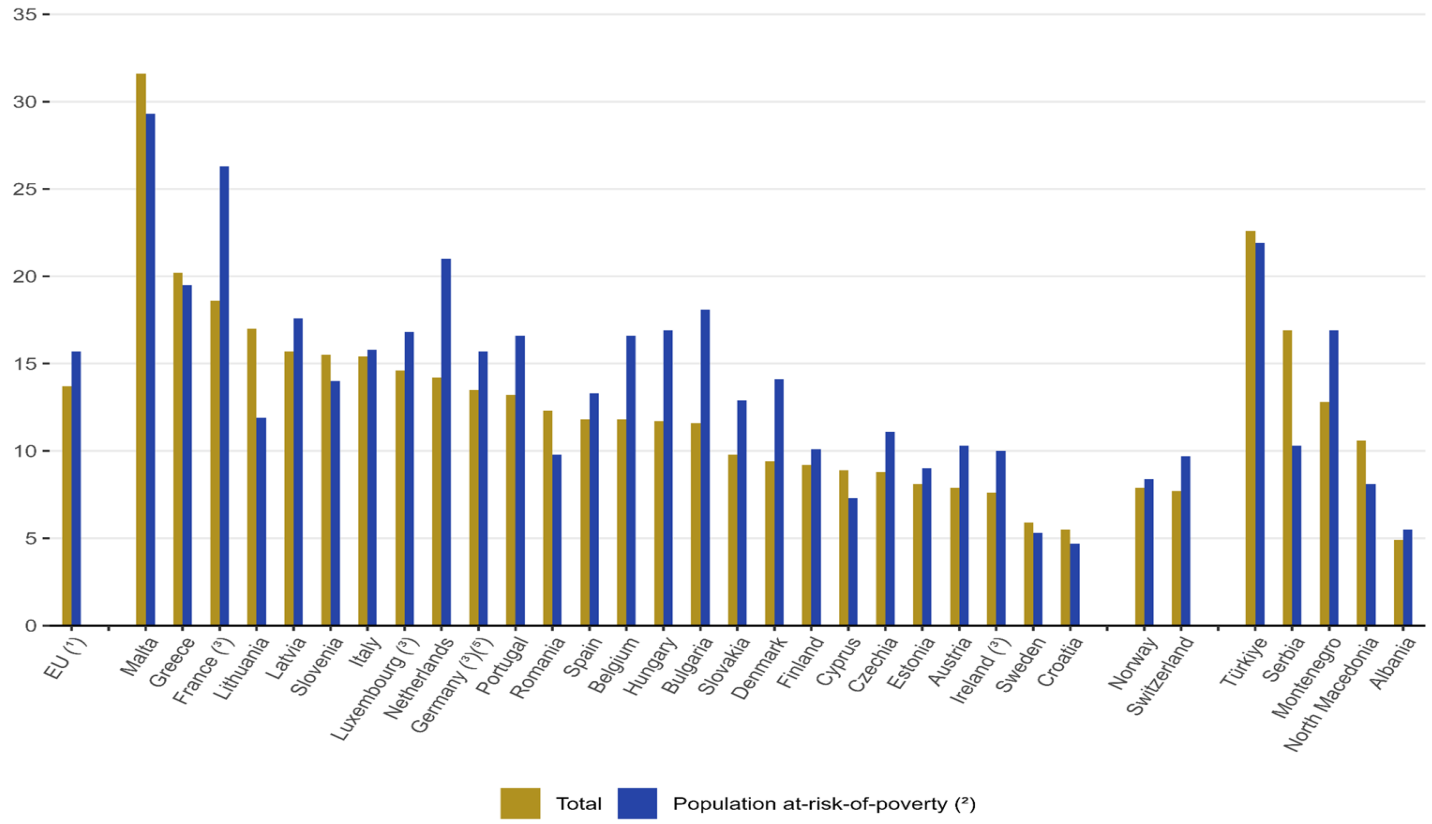
Creating the plot

```
p1 <- ggplot(data = data1, aes(x = time, y = exp,
group = 1)
  ) +
  geom_line(color = line_color,linewidth = 1) +
  scale_y_continuous(
    expand = c(0, 0),
    limits = c(0, max(data1$exp) + 2),
    breaks = seq(0, max(data1$exp) + 2, 2)
  )
  + AESTHETICS
  +  labs(
    title = title,
    subtitle = subtitle,
    caption = caption)
```

Bar graph

Population reporting exposure to pollution, grime or other environmental problems, by income situation, 2020 (% share)

(%)



(1) Estimate.
(2) People living below the national poverty threshold (60% of median equivalised income).
(3) Break in series.
(4) 2019 instead of 2020.
(5) Low reliability.
(6) 2018 instead of 2020.
Source: Eurostat (online data code: ilc_mddw02)

Retrieving data

Parameters

```
id2 <- "ilc_mddw02"
date2 <- 2020
```

API connection (geocodes and data)

```
# Get country codes for EU and EFTA
eu_cc <- get("cc", envir =
.restatapi_env)$EU27_2020
efta_cc <- c("CH", "NO", "IS", "LI")

# Fetch the data from Eurostat
data2 <- get_eurostat_data(
  id2,
  filters = list(
    incgrp = c("TOTAL", "B_MD60"),
    freq = "Annual",
    hhtyp = "TOTAL",
    unit = "Percentage"
  ),
  date_filter = date2,
  label = TRUE,
  ignore.case = TRUE,
  exact_match = FALSE,
  perl = TRUE,
  stringsAsFactors = FALSE,
  force_local_filter = F,
  keep_flags = TRUE
)
```

Modificating dataset for plotting

Geocodes

```
# Get the data structure definition (DSD) for
geographic codes
dsd <- get_eurostat_dsd(id2) |>
  as.data.table() |>
  filter(concept == "geo") |>
  select(code, name)

# Merge the main data with DSD to get full names
for geos
datafig2 <- merge(data2, dsd, by.x = "geo", by.y
= "name", all.x = TRUE)
```

Specific order of countries

```
# Order the names for EU, EFTA, and other
countries
name_ord_eu <- datafig2 |>
  filter(code %in% eu_cc, incgrp == "Total") |>
  arrange(values) |>
  pull(geo)

(...)

# Combine the ordered names
name_ord <- c(name_ord_othr, "   ", name_ord_efta,
"   ", name_ord_eu, "    ", "EU")
```

Modificating dataset for plotting

Final details for labelling data and categories' order

```
# Prepare the final dataset for plotting
datafig2 <- datafig2 |>
  select(geo, incgrp, values) |>
  mutate(
    incgrp = recode(incgrp,
                    "Below 60% of median equivalised income"
= "Population at-risk-of-poverty (\U00B2)")) |>
  add_row(
    geo = c("  ", "   ", "    "),
    incgrp = "Total",
    values = NA
  )

# Factorize the geo and ind_type columns
datafig2$geo <- factor(datafig2$geo, levels = rev(name_ord))
datafig2$incgrp <- factor(datafig2$incgrp, levels =
c("Total",
                                        "Population
at-risk-of-poverty (\U00B2)"))

# Recode specific geo values for the final plot
datafig2 <- datafig2 |>
  mutate(
    geo = recode(geo,
                 "France" = "France (\U00B3)",
                 "Luxembourg" = "Luxembourg (\U00B3)",
                 "Germany" = "Germany (\U00B3)(\U2075)",
                 "Ireland" = "Ireland (\U00B3)",
                 "EU" = "EU (\U00B9)")
  )
```

Plotting

Figure's text

```
title <- paste0("Population reporting exposure to pollution,
grime or other \nenvironmental problems, by income
situation, ", date2, " (% share)")
subtitle <- "(%)"
caption <- paste0("\n(\U00B9) Estimate.",
                  "\n(\U00B2) People living below the
national poverty threshold (60% of median equivalised
income).",
                  "\n(\U00B3) Break in series.",
                  "\n(\U2074) 2019 instead of 2020.",
                  "\n(\U2075) Low reliability.",
                  "\n(\U2076) 2018 instead of 2020.",
                  "\nSource: Eurostat (online data code: ",
id2, ")")
```

Creating the plot

```
p2 <- ggplot(data = datafig2, aes(x = geo, y = values,
fill = incgrp)) +
  geom_bar(stat = "identity", position = "dodge", width =
0.5) +
  AESTHETICS+
    labs(
    title = title,
    subtitle = subtitle,
    caption = caption)
```

Young people - digital world

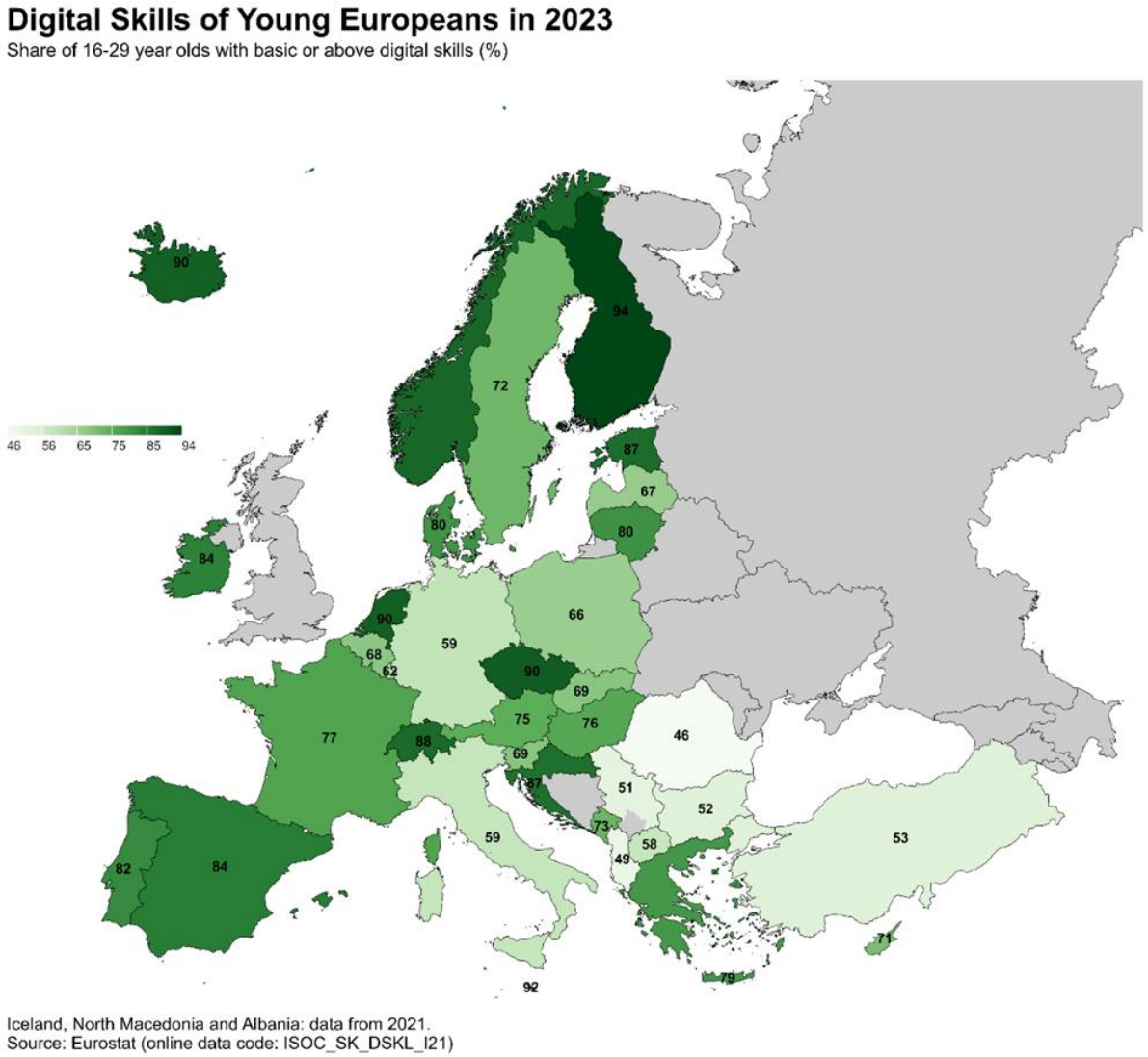
Joseph Nyajuoga

Örebro University, Sweden

Maripaz González

Complutense University of Madrid, Spain

Map



Iceland, North Macedonia and Albania: data from 2021.
Source: Eurostat (online data code: ISOC_SK_DSKL_I21)

European Commission

Nuts of European countries

Selecting and classifying countries

```r
nuts3_sf <- giscoR::gisco_get_nuts(
  nuts_level = "0",
  resolution = "3",
  year = "2021"
)

countries_sf <- giscoR::gisco_get_countries(
  resolution = "3",
  region = c("Asia", "Europe")
)

# FILTER COUNTRIES
non_eu_list <- c(
  "AM", "AZ", "BA",
  "BY", "GE", "MD",
  "RU", "UA", "XK"
)

eu_list <- unique(nuts3_sf$CNTR_CODE)

eu_sf <- countries_sf %>%
  dplyr::filter(CNTR_ID %in% eu_list)

non_eu_sf <- countries_sf %>%
  dplyr::filter(CNTR_ID %in% non_eu_list)
```

Indicator data

API connection

```
# DIGITAL SKILLS DATA
indicator_df <- get_eurostat_data(
  map_id,
  filters = list(
    ind_type = c("Y16_29"),
    indic_is = "I_DSK2_BAB",
    freq = "Annual",
    unit = c("PC_IND"),
    date_filter = c(2021:2023)
  ),
  label = FALSE,
  ignore.case = TRUE,
  exact_match = TRUE,
  perl = TRUE,
  stringsAsFactors = FALSE,
  force_local_filter = FALSE,
  keep_flags = TRUE
)
```

Combining nuts and indicator data

Adding geographical information to the indicator values

```
indicator_filtered_df <- indicator_df %>%
  dplyr::select(geo, time, values) %>%
  dplyr::rename("NUTS_ID" = "geo")

# Data frame TO WIDE FORMAT
indicator_wide_df <- indicator_filtered_df %>%
  tidyr::pivot_wider(names_from = time, values_from = values)

# REPLACE MISSING VALUES
indicator_df_final <- indicator_wide_df %>%
  dplyr::mutate(values = dplyr::if_else(is.na(`2023`), `2021`,
`2023`)) %>%
  dplyr::select(NUTS_ID, values)

# MERGE NUTS3 Spatial Forms AND Data Frame
mapping_sf <- nuts3_sf %>%
  dplyr::left_join(indicator_df_final, by = "NUTS_ID")
```

Indicator categories

Defining Breaks & Colors

```
ni <- classInt::classIntervals(mapping_sf$values, n = 5, style
= "equal")$brks
brk <- ni %>%
  append(max(mapping_sf$values)) %>%
  head(-1)
breaks <- c(min(mapping_sf$values), brk) %>%
  tail(-1)

cols <- hcl.colors(n = 6, palette = "Greens", rev = TRUE)
```

Geographical information

Defining the bbox range

```
crs_lambert <- "+proj=laea +lat_0=52 +lon_0=10 +x_0=4321000
+y_0=3210000 +datum=WGS84 +units=m +no_defs"

xmin <- -10.66
ymin <- 34.5
xmax <- 45
ymax <- 71

bbox <- sf::st_sfc(sf::st_polygon(list(cbind(
  c(xmin, xmax, xmax, xmin, xmin),
  c(ymin, ymin, ymax, ymax, ymin)
))), crs = 4326)

lambert_bbox <- sf::st_transform(bbox, crs = crs_lambert)
bb <- sf::st_bbox(lambert_bbox)
```

Plotting

Layering the sfs using the geom_sf

```
p <- ggplot(data = mapping_sf) +
  geom_sf(data = subset(eu_sf, CNTR_ID ==
"RS"),
          fill = "grey80", color = "black",
size = .15) +
  geom_sf(mapping = aes(fill = values),
          color = NA, size = 0) +
  geom_sf(data = eu_sf, color = "black",
          size = .15, fill = "transparent") +
  geom_sf(data = non_eu_sf, color = "black",
          size = .15, fill = "grey80") +
```

Breaks, bbox, colors and annotations

```
  geom_sf_text(data = mapping_sf,
               aes(label = round(values, 0)),

               size = 4,
               color = "black",
               check_overlap = TRUE,
               nudge_y = 0.8,,
               fontface = "bold",
               path = T) +
  coord_sf(crs = crs_lambert, xlim =
c(bb["xmin"], bb["xmax"]), ylim =
c(bb["ymin"], bb["ymax"])) +
  scale_fill_gradientn(
    colors = cols,
    breaks = breaks,
    labels = round(breaks, 0),
    limits = c(min(mapping_sf$values),
max(mapping_sf$values)),
    na.value = "grey80"
```

Plotting

Legend

```
guides(
  fill = guide_colorbar(
    direction = "horizontal",
    barheight = unit(2.5, units = "mm"),
    barwidth = unit(50, units = "mm"),
    label.position = "bottom",
    label.hjust = 0,
    nrow = 1,
    byrow = TRUE
  )
) +
```

Theme and text

```
  theme_void() +
  theme OPTIONS
  ) +
  labs(
    caption = paste0("Iceland, North
Macedonia and Albania: data from 2021.",
                     "\nSource: Eurostat (online
data code: ", map_id, ")\n"),
    title = "Digital Skills of Young
Europeans in 2023",
    subtitle = "Share of 16-29 year olds with
basic or above digital skills (%)\n"
  )
```

Consumer prices of recreational and sporting goods and services	Ricardo Dias de Carbalho *University of Porto*
Quality of life indicators - overall experience of life	Gabriella Manuti *University of Rome La Sapienza*
Young people - digital world	Joseph Nyajuoga *Örebro University*
Quality of life indicators - natural and living environment	Maripaz Venegas Gonzalez *Complutense University of Madrid*
Consumer prices of cultural goods and services	Sona Yavrumyan *University of Bologna*
Culture statistics - cultural employment	Nikola Quaresimin, *University of Bamberg*

Consumer Prices of Cultural Goods and Services

Sona Yavrumyan

University of Bologna

European Commission

Article Outline

- The article reviews the changes in consumer prices for cultural goods and services in the EU from 2013 to 2023, using the Harmonised Index of Consumer Prices (HICP). Over this period, newspapers and periodicals had a significant price increases across all EU countries, surpassing overall inflation. In contrast, the prices of books and cultural services rose less rapidly and slower. Recording media prices generally decreased, with occasional increases in specific years. Similarly, prices for information processing equipment and equipment for the reception, recording, and reproduction of sound and picture mostly saw a downward trend, with minor increases in some years. Country-specific variations were also noted.

Plot 1

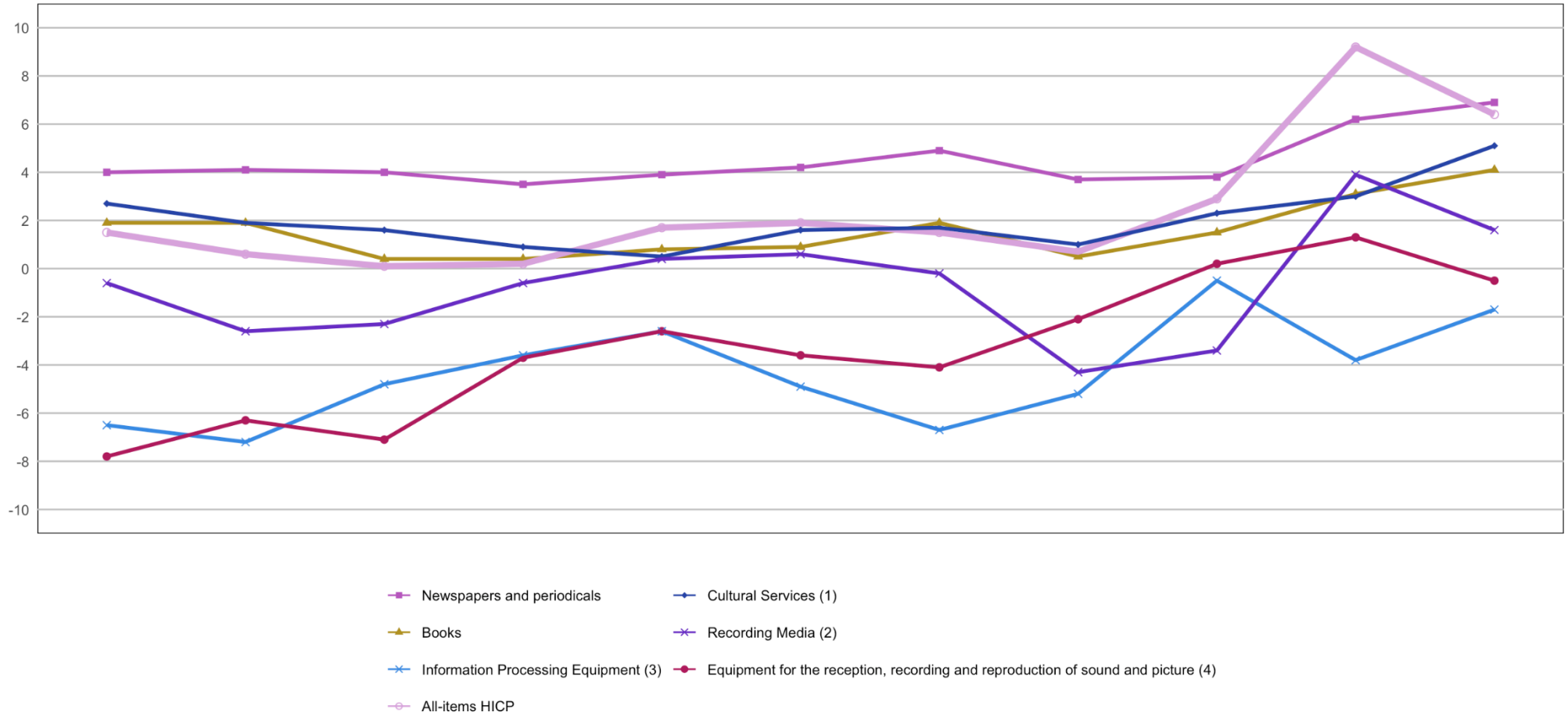
```r
# Get the id and read the data through dsd
id <- "prc_hicp_aind"

# Specify the years of interest
time_EU <- seq(from = 2013, to = 2023, by = 1)

# Specify the 7 categories of interest
cult_var_names <- c("CP0952", "CP0951", "CP0913", "CP00", "CP0914", "CP0911", "CP0942")
data_api <- get_eurostat_data(id,
                              filters = list(geo = "EU", coicop = cult_var_names,
                                             unit = "RCH_A_AVG"),
                              date_filter = time_EU, label = TRUE)

# Create the dataset with left join by category (coicop)
# Now define all the variables seperately

                      ##### PLOT1 ####

time_graph1 <- unique(data_api$time)

all_hicp <- data_api %>%
  filter(coicop == "All-items HICP") %>%
  select(values)

equip <- data_api %>%
  filter(coicop ==
"Equipment for the reception, recording and reproduction of sound and picture") %>%
  select(values)

info_equip <- data_api %>%
  filter(coicop == "Information processing equipment") %>%
  select(values)

rec_media <- data_api %>%
  filter(coicop == "Recording media") %>%
  select(values)

cult_serv <- data_api %>%
  filter(coicop == "Cultural services") %>%
  select(values)

books <- data_api %>%
  filter(coicop == "Books") %>%
  select(values)

newspapers <- data_api %>%
  filter(coicop == "Newspapers and periodicals") %>%
  select(values)

# Bind cols and create a new column for data frame names
combined_cols <- bind_cols(as_tibble(time_graph1), all_hicp, equip, info_equip,
                           rec_media, cult_serv, books, newspapers)
```

```r
# Define codes for each variable with palette 2 from Eurostat apart
# from all_hicp which didn't have a colour from the palettes
color_palette <- c(
  all_hicp = "#d7a2db",
  equip = "#af155c",
  info_equip = "#388ae2",
  rec_media = "#672dc4",
  cult_serv = "#2644a7",
  books = "#b09120",
  newspapers = "#b656bd"
)
# Create the plot
plot1 <- ggplot(combined_data_long, aes(x = time, y = value,
                                        linewidth = variable, size = variable)) +
  geom_point() +
  geom_line() +
  guides(color=guide_legend(ncol=2), shape=guide_legend(ncol=2), linewidth=F, size=F) +
  scale_shape_manual(values=c(15,17,4,21,18,4,19), labels=labels_original) +
  scale_linewidth_manual(values = c(1.1,1.1,1.1,1.8,1.1,1.1,1.1)) +
labs(title=
"Harmonised indices of consumer prices for selected cultural goods and services, EU, 2013-
subtitle="(%, annual rate of change)", x = "", y = "", color = "Variable",
shape = "Variable",  cols='markers') +
  scale_x_continuous(breaks = seq(2013, 2023, by = 1)) +
  scale_y_continuous(breaks = seq(-10, 12, by = 2), limits = c(-10,10)) +
  scale_color_manual(values = color_palette, labels=labels_original) +
  scale_size_manual(values=c(1.8,1.8,1.8,2,1.8,1.8,1.8)) +
  theme_minimal() +
  theme(panel.background = element_rect(fill = "white"), # Set background to white
        panel.grid.major.y = element_line(color = "gray"), # Set major grid lines to gray
        panel.grid.minor = element_blank(),  # Remove minor grid lines
        panel.grid.major.x = element_blank(),  # Remove vertical grid lines
        panel.grid.minor.x = element_blank(),
        plot.title = element_text(face = "bold"),
        plot.caption = element_text(size = 8, hjust = 0, vjust = 5),
        legend.position = "bottom", legend.title = element_blank())

logo_file <- "logo_RGB-POS.png"  # Eurostat logo
logo <- readPNG(logo_file)
logo_grob <- rasterGrob(logo, interpolate=T)

# Create a text grob for the footnote
footnote_grob <- textGrob(cap, x = unit(0, "npc"), y = unit(1, "npc"),
                          just = c("left", "top"), gp = gpar(fontsize = 8))

# Combine the footnote and logo in a horizontal layout
footnote_with_logo <- arrangeGrob(footnote_grob, logo_grob, ncol = 2, widths = c(4, 1))

# Combine the plot and the footnote with logo in a vertical layout
combined_grob <- arrangeGrob(plot1, footnote_with_logo, ncol = 1, heights = c(4, 1))

# Display the combined plot with footnote and logo
grid.newpage()
grid.draw(combined_grob)
```

European Commission

Harmonised indices of consumer prices for selected cultural goods and services, EU, 2013-2023

(%, annual rate of change)



Note: average indices reflect changes in EU membership (when a Member State joins the EU, its HICPs are chained with the aggregate index at the time of the accession).
(1) Includes cinemas, theatres, concerts; museums, libraries, zoological gardens; television and radio fees, hire of equipment and accessories for culture; and other cultural services.
(2) Includes records, CDs, DVDs, tapes, cassettes, etc.
(3) Includes personal computers, visual display units, printers and miscellaneous accessories accompanying them; also computer software packages such as operating systems, applications, languages, etc.
(4) Includes TV sets, CD players, stereo systems, radios, etc.
Source: Eurostat (online data code: prc_hicp_aind)

eurostat

European Commission

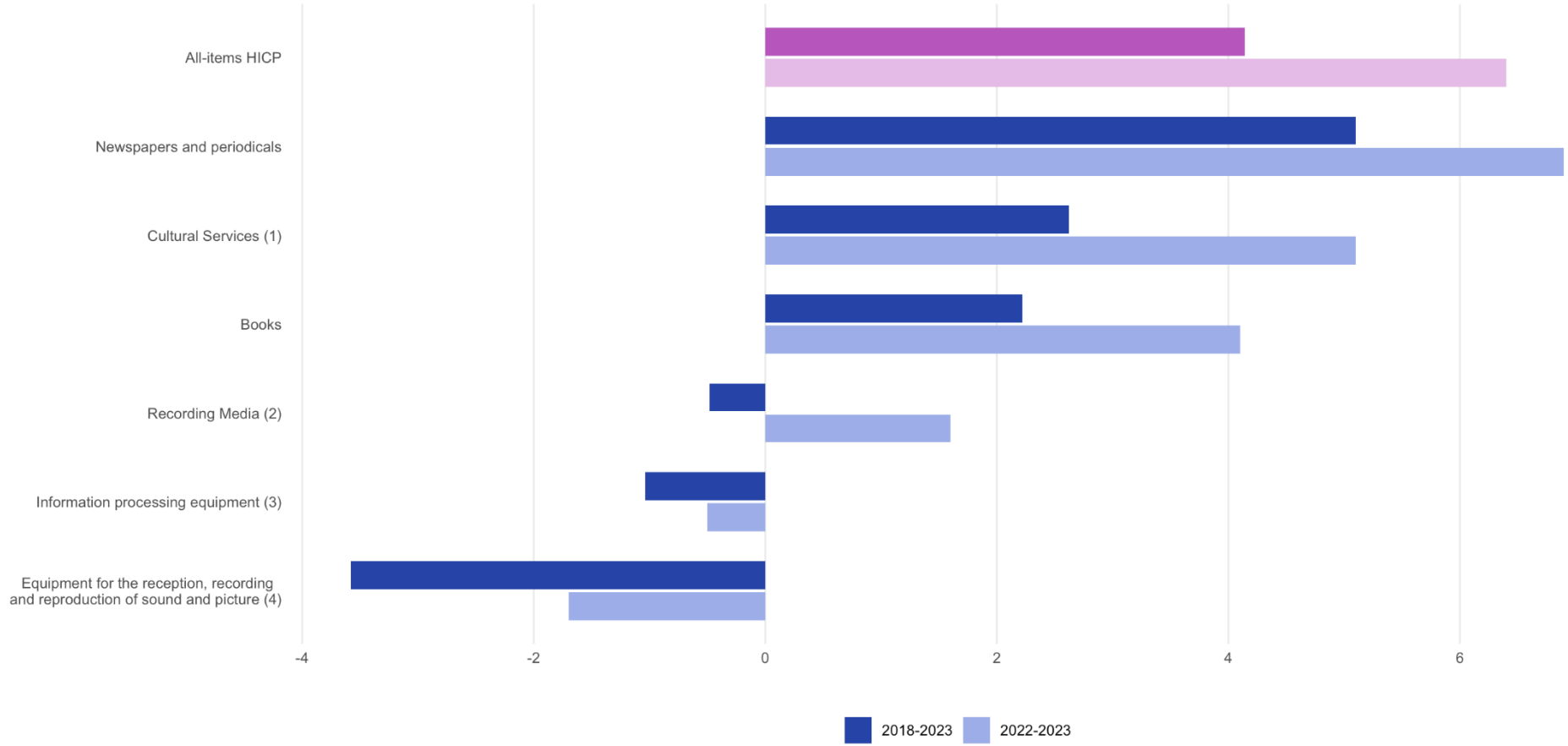
Plot 2

```r
# Filter 2018-2023 and 2022-2023 since that's what's required
values_2022_2023 <- data_api2 %>%
  filter(time == 2023) %>%                          # Filter for the years 2022 to 2023
  group_by(coicop) %>%                              # Group by the category column
  summarize(value = mean(values, na.rm = TRUE)) %>% # Calculate the mean for each category
  select(coicop, value) %>%
  mutate(date = rep("2022-2023", 7)) %>%
  arrange(coicop)

values_2018_2023 <- data_api2 %>%
  filter(time > 2018 & time <= 2023) %>%            # Filter for the years 2018 to 2023
  group_by(coicop) %>%                              # Group by the category column
  summarize(value = mean(values, na.rm = TRUE)) %>% # Calculate the mean for each category
  select(coicop, value) %>%
  mutate(date = rep("2018-2023", 7)) %>%
  arrange(coicop)

# Join
data_long_2 <- bind_rows(values_2018_2023, values_2022_2023)
data_long2 <- data_long_2 %>%
  mutate(Color_Group = ifelse(coicop == "All-items HICP", paste(coicop, date),
paste("Other", date))) %>% mutate_if(is.character,as.factor)

# Redefine the names of the categories as needed
var_categories <- unique(data_api$coicop) %>% as.character()
var_categories[2] <- "Equipment for the reception, recording
and reproduction of sound and picture (4)"
var_categories[3] <- "Information processing equipment (3)"
var_categories[4] <- "Recording Media (2)"
var_categories[5] <- "Cultural Services (1)"

var_cat <- bind_rows(as_tibble(var_categories), as_tibble(var_categories))
colnames(var_cat) <- "category"
data_long2 <- bind_cols(data_long2, var_cat) %>% select(-coicop)

data_long2 <- data_long2 %>%
  arrange(date) %>%
  arrange(value) %>%
  arrange(desc(Color_Group)) %>%
  mutate_if(is.character, as.factor)

data_long2$category <- factor(data_long2$category, levels=unique(data_long2$category))
```

```r
# Plot the second graph

plot2 <- ggplot(data_long2, aes(x = category, y = value, fill = Color_Group)) +
  geom_bar(stat = "identity", position = position_dodge2(width = 0.8, reverse = TRUE),
width = 0.7) +
  coord_flip() +  # This flips the axes so categories are on the y-axis
  scale_fill_manual(values = c(
    "All-items HICP 2022-2023" = "#e2bbe5",
    "All-items HICP 2018-2023" = "#b656bd",
    "Other 2022-2023" = "#9cade8",
    "Other 2018-2023" = "#2644a7"
  ),
  breaks = c("Other 2018-2023","Other 2022-2023"),
  labels = c("2018-2023","2022-2023")) +  # Customize the colors for each specific group
  labs(title =
"Harmonised Indices of Consumer Prices for Selected Cultural Goods and Services",
       subtitle = "EU, Annual Average Rates of Change 2018-2023 and 2022-2023 (%)",
       x = "",
       y = ""
      ) +
  theme_minimal() +
  theme(legend.position = "bottom", legend.title = element_blank(),
        plot.title = element_text(size = 14, face = "bold"),
        plot.subtitle = element_text(size = 12),
        plot.caption = element_text(size = 7, hjust = 0),
        panel.grid.major.y = element_blank()) +
  scale_y_continuous(breaks = c(seq(-4, 6, by = 2)), minor_breaks= c(seq(-4, 6, by = 2)))

# Display the plot with footnote and logo
footnote_grob2 <- textGrob(cap2, x = unit(0, "npc"), y = unit(1, "npc"),
                           just = c("left", "top"), gp = gpar(fontsize = 8))

# Combine the footnote and logo in a horizontal layout
footnote_with_logo2 <- arrangeGrob(footnote_grob2, logo_grob, ncol = 2, widths = c(4, 1))
# Combine the plot and the footnote with logo in a vertical layout
combined_grob2 <- arrangeGrob(plot2, footnote_with_logo2, ncol = 1, heights = c(4, 1))

# Display the combined plot
grid.newpage()
grid.draw(combined_grob2)
```

Harmonised Indices of Consumer Prices for Selected Cultural Goods and Services

EU, Annual Average Rates of Change 2018-2023 and 2022-2023 (%)



Note: average indices reflect changes in EU membership (when a Member State joins the EU, its HICPs are chained with the aggregate index at the time of the accession).
(1) Includes cinemas, theatres, concerts; museums, libraries, zoological gardens; television and radio fees, hire of equipment and accessories for culture; and other cultural services.
(2) Includes records, CDs, DVDs, tapes, cassettes, etc.
(3) Includes personal computers, visual display units, printers and miscellaneous accessories accompanying them; also computer software packages such as operating systems, applications, languages, etc.
(4) Includes TV sets, CD players, stereo systems, radios, etc.
Source: Eurostat (online data code: prc_hicp_aind)

eurostat

European Commission

Table

```r
# Define the dataset for the table and filter to include required countries and regions
data_api_all <- get_eurostat_data(id,
                                  filters = list(coicop = cult_var_names,
                                                 unit = "RCH_A_AVG"),
                                  date_filter = time_EU, label = TRUE)

# data_api_all <- left_join(data_api_all, labels, by="coicop")
# view(data_api_all)
unique(data_api_all$geo)

data_api_all_2 <- data_api_all %>%
  filter(
    !geo %in% c("United Kingdom", "United States") &
      !str_detect(geo, "^(Euro area|European Economic|European Union -)"))

data_api_all_2 <- data_api_all_2 %>% mutate(time=as.numeric(as.character(time)))

# Sort the times again 2018-2023 and 2022-2023
values_2022_2023_all <- data_api_all_2 %>%
  filter(time == 2023) %>%
  group_by(coicop, geo) %>%
  summarize(value = mean(values, na.rm = TRUE)) %>% # Calculate the mean for each category
  select(coicop, value, geo) %>%
  mutate(date = rep("2022-2023")) %>%
  arrange(coicop)

values_2018_2023_all <- data_api_all_2 %>%
  filter(time > 2018 & time <= 2023) %>%            # Filter for the years 2018 to 2023
  group_by(coicop, geo) %>%                            # Group by the category column
  summarize(value = mean(values, na.rm = TRUE)) %>% # Calculate the mean for each category
  select(coicop, value, geo) %>%
  mutate(date = rep("2018-2023")) %>%
  arrange(coicop)

# Bind and redefine the categories as needed
combo <- bind_rows(values_2018_2023_all, values_2022_2023_all) %>%
  mutate(coicop = str_replace_all(coicop, c(
    "Equipment for the reception, recording and reproduction of sound and picture" =
"Equipment for the reception, recording  \nand reproduction of sound and picture (4)",
    "Information processing equipment" = "Information processing equipment (3)",
    "Recording media" = "Recording Media (2)",
    "Cultural services" = "Cultural Services (1)"
  )))
```

```r
#Get the elements to 13 from 7 and save in reverse order and arrange
labels_graph2 <- rev(data_long2$category[7:13])

combo$coicop <- factor(combo$coicop, levels=labels_graph2)
combo2 <- combo %>%
  arrange(date) %>%
  arrange(coicop)

levels(combo2$geo)[17] <- "EU"

country_names <- c("Albania (5)" = "Albania", "Montenegro (5)" = "Montenegro",
"North Macedonia (5)" = "North Macedonia",
"Kosovo (5)(6)" = "Kosovo*", "Serbia (5)" = "Serbia",
"Turkiye (5)" = "Turkiye", "Czech Republic" = "Czechia")

# Convert the 'geo' column to a factor and recode it
combo2$geo <- fct_recode(factor(combo2$geo), !!!country_names)

combo3 <- combo2 %>%
  pivot_wider(names_from = c(coicop, date), values_from = value)

# Specified order of country names
country_order <- c
("EU", "Belgium", "Bulgaria", "Czech Republic", "Denmark", "Germany",
"Estonia", "Ireland", "Greece", "Spain", "France", "Croatia", "Italy", "Cyprus", "Latvia",
"Lithuania", "Luxembourg", "Hungary", "Malta", "Netherlands", "Austria", "Poland",
"Portugal", "Romania", "Slovenia", "Slovakia", "Finland", "Sweden", "Iceland", "Norway",
"Switzerland", "Montenegro (5)",
"North Macedonia (5)", "Albania (5)", "Serbia (5)",
"Turkiye (5)", "Kosovo (5)(6)")

# Convert the 'geo' column to a factor with the specified levels
combo3$geo <- factor(combo3$geo, levels = country_order)

# Sort the dataset by the 'geo' column
combo3 <- combo3[order(combo3$geo), ]
colnames(combo3) <- c(" ",rep(c("2018-2023","2022-2023"), times=7))


# Define the header and the footnote
header <- c(1, rep(2,7))
names(header) <- c(" ", as.character(labels_graph2))

names(header)[3] <- "Newspapers
and periodicals"
names(header)[4] <- "Cultural
Services (1)"
names(header)[6] <- "Recording
Media (2)"
```

Table

```
footnote_text <- paste0(c(
"(1) Includes cinemas, theatres, concerts; museums, libraries, zoological gardens;",
"television and radio fees, hire of equipment and accessories for culture; and other",
" cultural services. (2) Includes records, CDs, DVDs, tapes, cassettes, etc.",
"(3) Includes personal computers, visual display units, printers and",
"miscellaneous accessories",
"accompanying them; also computer software packages such as operating",
"systems, applications, languages, etc.",
"(4) Includes TV sets, CD players, stereo systems, radios, etc.",
"(5) Definition differs, see metadata", "href='https://ec.europa.eu/eurostat/cache/",
"metadata/en/prc_hicp_esms.htm'",
"(6) This designation is without prejudice to positions on status, and is in line with",
"UNSCR 1244/1999 andthe ICJ Opinion on the Kosovo declaration of independence.")

# Create the table
final_table <- kable(combo3, format = "html", escape = F, digits = 1,
caption = "The all-items HICP and HICPs for cultural goods and services") %>%
  group_rows(" ", 1, 1) %>%
  group_rows(" ", 2, 28) %>%
  group_rows(" ", 29, 31) %>%
  group_rows(" ", 32, 36) %>%
  group_rows(" ", 37, 37) %>%
  add_header_above(header, background = "#f0dcf1", include_empty = F) %>%
  row_spec(row = 0, background = "#f0dcf1") %>%
  row_spec(row = 1, background = "#e1bae4") %>%
  kable_styling(bootstrap_options = c("striped", "hover", "condensed"),
                full_width = F, font_size = 12, fixed_thead = TRUE) %>%
  footnote(general = footnote_text, escape = F)

# Output the table
final_table
write_xlsx(combo3, path = "Table_HICP.xlsx")
```

The all-items HICP and HICPs for cultural goods and services, annual average rates of change 2018-2023 and 2022-2023 (%)

	All-items HICP		Newspapers and periodicals		Cultural Services (1)		Books		Recording Media (2)		Information processing equipment (3)		Equipment for the reception, recording and reproduction of sound and picture (4)	
	2018-2023	2022-2023	2018-2023	2022-2023	2018-2023	2022-2023	2018-2023	2022-2023	2018-2023	2022-2023	2018-2023	2022-2023	2018-2023	2022-2023
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

eurostat

Saving datasets in Excel for Further Use

Consumer prices of cultural goods and services

Harmonised indices of consumer prices for selected cultural goods and services, EU, annual average rates of change 2018-2023 and 2022-2023

(%)

Category	Year										
	2013	2014	2015	2016	2017	2018	2019	2020	2021	2022	2023
All-items HICP	1.5	0.6	0.1	0.2	1.7	1.9	1.5	0.7	2.9	9.2	6.4
Equipment for the reception, recording and reproduction of sound and picture	-6.5	-7.2	-4.8	-3.6	-2.6	-4.9	-6.7	-5.2	-0.5	-3.8	-1.7
Information processing equipment	-7.8	-6.3	-7.1	-3.7	-2.6	-3.6	-4.1	-2.1	0.2	1.3	-0.5
Recording media	-0.6	-2.6	-2.3	-0.6	0.4	0.6	-0.2	-4.3	-3.4	3.9	1.6
Cultural services	2.7	1.9	1.6	0.9	0.5	1.6	1.7	1	2.3	3	5.1
Books	1.9	1.9	0.4	0.4	0.8	0.9	1.9	0.5	1.5	3.1	4.1
Newspapers and periodicals	4	4.1	4	3.5	3.9	4.2	4.9	3.7	3.8	6.2	6.9

Category	Year	
	2018-2023	2022-2023
Books	2.22	4.1
Cultural Services (1)	2.62	5.1
Equipment for the reception, recording and reproduction of sound and picture (4)	=AVERAGE(I13:M13)	-1.7
Newspapers and periodicals	5.1	6.9
Information processing equipment (3)	-1.04	-0.5
Recording Media (2)	-0.48	1.6
All-items HICP	4.14	6.4

European Commission

Consumer prices of recreational and sporting goods and services	Ricardo Dias de Carbalho *University of Porto*
Quality of life indicators - overall experience of life	Gabriella Manuti *University of Rome La Sapienza*
Young people - digital world	Joseph Nyajuoga *Örebro University*
Quality of life indicators - natural and living environment	Maripaz Venegas Gonzalez *Complutense University of Madrid*
Consumer prices of cultural goods and services	Sona Yavrumyan *University of Bologna*
Culture statistics - cultural employment	Nikola Quaresimin, *University of Bamberg*

Culture statistics – cultural employment

Nikola Quaresimin

Otto-Friedrich Universität Bamberg

European Commission

Global Settings

```
# install and load packages
pacman::p_load(restatapi,     # for Eurostat data
               giscoR,        # for geographical Data
               ggplot2,       # for graphics
               tidyverse,     # for data cleaning
               cowplot,       # for layering with "ggdraw"
               writexl,       # for exporting data as excel files
               kableExtra,    # for table formatting
               webshot,       # for saving tables as png
               knitr,         # for tables
               magick         # to read in logo
)

webshot::install_phantomjs(force = T) # for saving tables as png
```

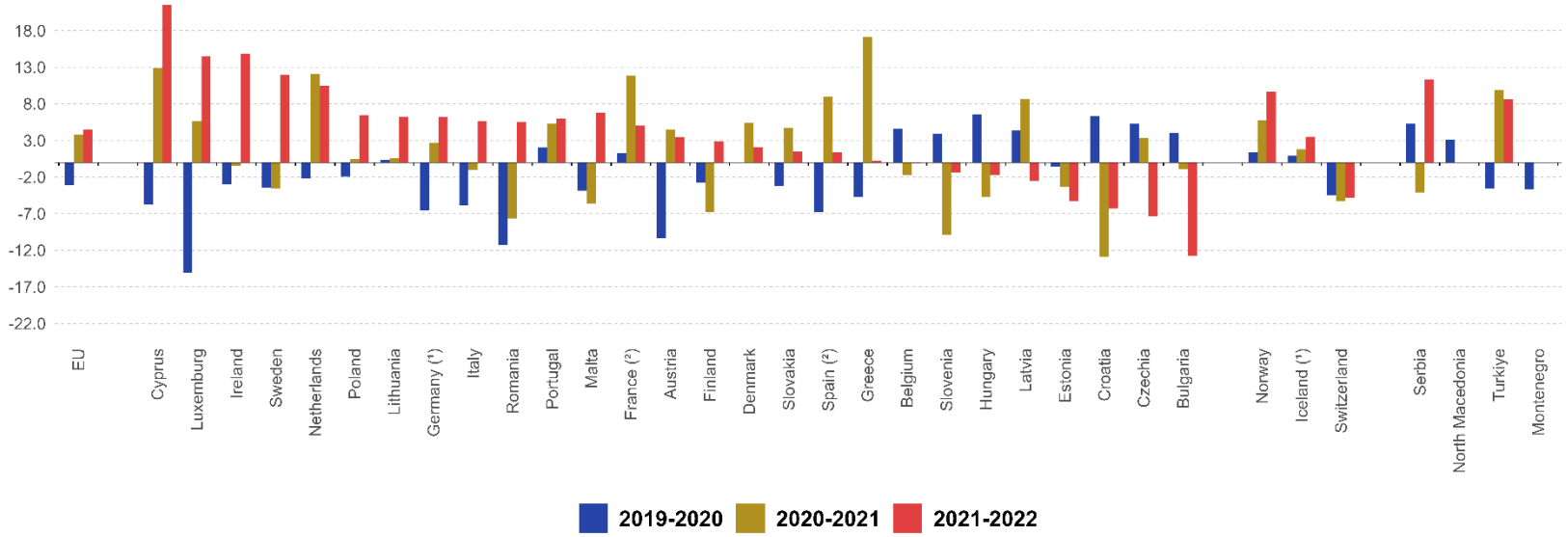
```
# color palette
clrs <- read.csv("estat_colour_themes.txt",sep="\t",stringsAsFactors = F)
image(c(1:6),c(1:nrow(clrs)),matrix(1:(6*nrow(clrs)),6,nrow(clrs)),col=as.vector(t(apply(as.matrix(clrs[,c(3:8)]),
2,rev))), xlab = "", ylab = "", xaxt = "n", yaxt = "n", bty = "n",
      main = "Eurostat colour palettes")
text(c(0.5), c(1:nrow(clrs)), adj=c(0,1),rev(clrs$theme_desc), col = clrs$text_light)
text(c(1.5), c(1:nrow(clrs)), adj=c(0,1),rev(clrs$theme_desc), col = clrs$text_dark)
```

```
# logo
logo <- image_read("png/Eurostat_Newlogo.png")


# Typography
header_size <- 18
text_size <- 9
# legend_size <-    # legend text sizes are different throughout the different graphics
```

European Commission

Figure 1

Cultural employment – annual rates of change, 2020-2022
(%)



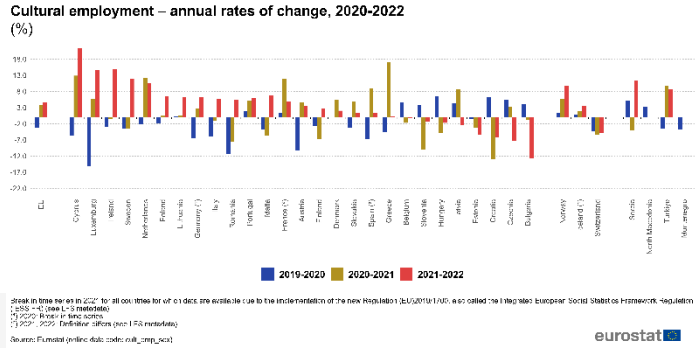
Break in time series in 2021 for all countries for which data are available due to the implementation of the new Regulation (EU)2019/1700, also called the Integrated European Social Statistics Framework Regulation (IESS FR) (see LFS metadata)
(¹) 2020: Break in time series.
(²) 2021, 2022: Definition differs (see LFS metadata).

Source: Eurostat (online data code: cult_emp_sex)

eurostat

European Commission

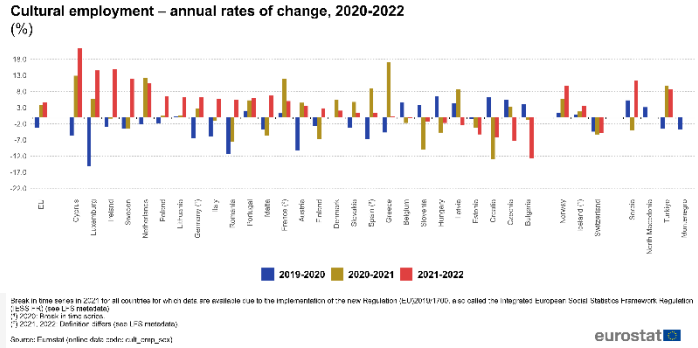
Figure 1




```
# settings
id1 <- "cult_emp_sex"
date2 <- 2019:2022
col_fig1 <- as.vector(clrs[1, 3:5])



title_fig1 <-   paste0("Cultural employment – annual rates of change, ",
                       date2[2],"-",date2[length(date2)])
subtitle_fig1 <-  "(%)"
caption_fig1 <-  paste0("\n", "Break in time series in 2021 for all countries for which data are available due to
the implementation of the new Regulation (EU)2019/1700, also called the Integrated European Social Statistics Fram
ework Regulation",
                        "\n(IESS FR) (see LFS metadata)",
                        "\n(\U00B9) 2020: Break in time series.",
                        "\n(\U00B2) 2021, 2022: Definition differs (see LFS metadata).",
                        "\n",
                        "\nSource: Eurostat (online data code: ",id1,")")



# read in data
cult_emp_fig1 <- get_eurostat_data(id1, filters = list(time = date2, sex = "T",
                                                       unit = "THS_PER", label = T))
```

European Commission

Figure 1

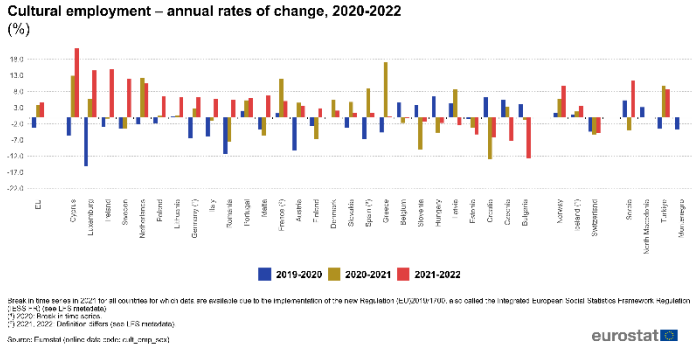



```
# calculate the rate of change: (value_currentyear-value_previousyear)/value_previousyear*100
cult_emp_fig1 <-  cult_emp_fig1 %>%
                          group_by(geo) %>%
                          mutate(rate_of_change = (values - lag(values)) / lag(values) * 100) %>%
                          filter(!is.na(rate_of_change))


# add NA lines for North Macedonia and Montenegro
NA_fig1 <- data.frame(
  geo = c("MK", "MK", "ME", "ME"),
  unit = "THS_PERS",
  sex = "T",
  time = c(2021, 2022, 2021, 2022),
  values = NA,
  rate_of_change = NA)

NA_fig1[, 1:4] <- lapply(NA_fig1[, 1:4], factor)


dt_sep<-data.table::data.table(geo=c(" "," "," "),unit = "THS_PERS", sex="T",time = c(2020, 2021, 2022),values=
rep(NA,3),rate_of_change=rep(NA,3))

dt_sep$time<-as.factor(dt_sep$time)

cult_emp_fig1 <- bind_rows(NA_fig1, cult_emp_fig1,dt_sep)
```

European Commission

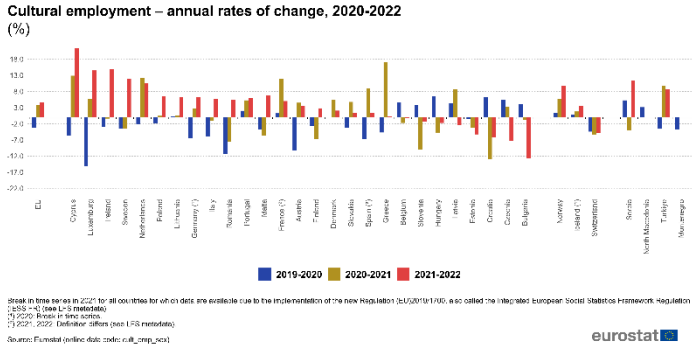
Figure 1

```
# Specify the desired order of countries
desired_order <- c("EU27_2020"," ","CY", "LU", "IE", "SE", "NL", "PL", "LT",
                   "DE", "IT", "RO", "PT", "MT", "FR", "AT", "FI", "DK",
                   "SK", "ES", "EL",
                   "BE", "SI", "HU", "LV", "EE", "HR",
                   "CZ", "BG","  ", "NO", "IS", "CH", "   ", "RS", "MK", "TR", "ME")


# Reorder the 'geo' variable with 'desired order'
cult_emp_fig1$geo <- factor(cult_emp_fig1$geo, levels = desired_order)


# add country labels
geo_labels <- c("EU27_2020" = "EU"," "=" " , "CY" = "Cyprus", "LU" = "Luxemburg",
                "IE" = "Ireland", "SE" = "Sweden", "NL" = "Netherlands",
                "PL" = "Poland", "LT" = "Lithuania",
                "DE" = "Germany", "IT" = "Italy", "RO" = "Romania",
                "PT" = "Portugal", "MT" = "Malta", "FR" = "France",
                "AT" = "Austria", "FI" = "Finland", "DK" = "Denmark",
                "SK" = "Slovakia", "ES" = "Spain", "EL" = "Greece",
                "BE" = "Belgium", "SI" = "Slovenia", "HU" = "Hungary",
                "LV" = "Latvia", "EE" = "Estonia", "HR" = "Croatia",
                "CZ" = "Czechia", "BG" = "Bulgaria", "  "="  ","NO" = "Norway",
                "IS" = "Iceland", "CH" = "Switzerland", "   "="   ", "RS" = "Serbia",
                "MK" = "North Macedonia", "TR" = "Turkiye", "ME" = "Montenegro")
```



European Commission

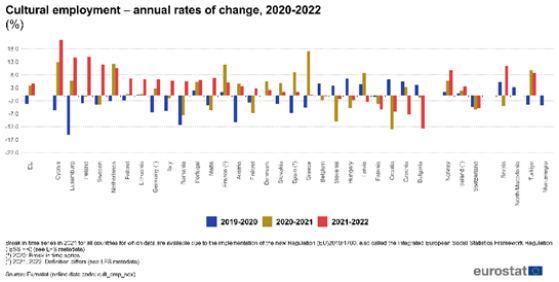
Figure 1




```r
# Rename specific countries for footnotes
geo_labels["DE"] <- "Germany (\U00B9)"
geo_labels["FR"] <- "France (\U00B2)"
geo_labels["ES"] <- "Spain (\U00B2)"
geo_labels["IS"] <- "Iceland (\U00B9)"



# Define the tick positions and lengths
tick_positions <- seq(0.5, length(unique(cult_emp_fig1$geo)) + 0.5, by = 1)
tick_length <- 0.5
```

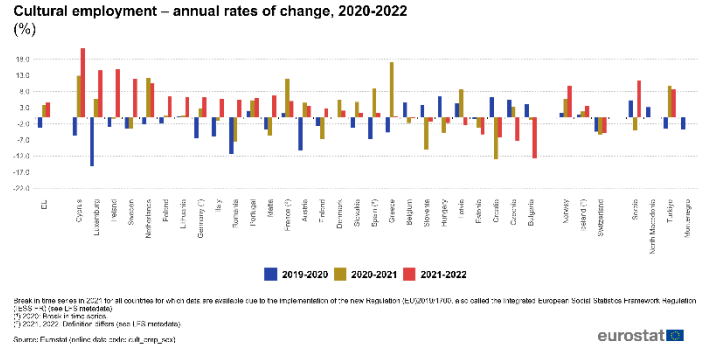
European Commission

Figure 1




```
# Barplot
fig1 <- ggplot(cult_emp_fig1 , aes(x = geo, y = rate_of_change,
                fill = factor(time, levels = c("2020", "2021", "2022")))) +
  geom_hline(yintercept = seq(-22, 18, by = 5), color = "grey", linetype = "dashed",linewidth = 0.2)+
  geom_bar(stat = "identity", position = position_dodge2(width = 0.7, preserve = "single", padding = 0),
           width = 0.7) +
  geom_hline(yintercept = 0, color = "black", linewidth = 0.2) +
  geom_segment(data = data.frame(x = tick_positions), aes(x = x, xend = x, y = 0, yend = -tick_length),
               color = "black", linetype = "solid", linewidth = 0.3, inherit.aes = FALSE) +
  labs(title = title_fig1,
       subtitle = subtitle_fig1,
       caption = caption_fig1,
        fill = "Year") +
  scale_y_continuous(breaks = seq(-22, 18, by = 5),  labels = function(x) sprintf("%.1f", x)) +
  scale_x_continuous(breaks = seq(1, length(unique(cult_emp_fig1$geo))), labels = geo_labels) +
  scale_x_discrete(labels = geo_labels) +
  scale_fill_manual(values = c("2020" = col_fig1[[1]], "2021" = col_fig1[[2]], "2022" = col_fig1[[3]]),
                    labels = c("2020" = "2019-2020", "2021" = "2020-2021", "2022" = "2021-2022")) +
  theme_bw() +
  theme(legend.background = element_blank(),
        legend.title = element_blank(),
        legend.text = element_text(family = "Arial", face = "bold", size = 12),
        legend.position = "bottom",
        plot.title = element_text(face = "bold", size = header_size, hjust = 0),
        plot.title.position = "plot",
        plot.subtitle = element_text(hjust = 0, size = header_size),
        plot.caption.position = "plot",
        axis.title.x = element_blank(),
        axis.title.y = element_blank(),
        axis.text.x = element_text(angle = 90, vjust = 0.5, hjust = 1, size = text_size),
        axis.text.y = element_text(size = text_size),
        axis.ticks.x.bottom = element_blank(),
        axis.ticks.y =  element_blank(),
        panel.grid = element_blank(),
        panel.background = element_blank(),
        plot.caption=element_text(hjust = 0, size = text_size),
        panel.border = element_blank()
      )
```

Figure 1




```r
# Combine the plot and the logo
plot_gtable <- ggplotGrob(fig1)
```

```r
fig1 <- ggdraw() +
  draw_grob(plot_gtable) +
  draw_image(logo, x = 0.99, y = -0.04, width = 0.15, height = 0.15, hjust = 1, vjust = 0)
```

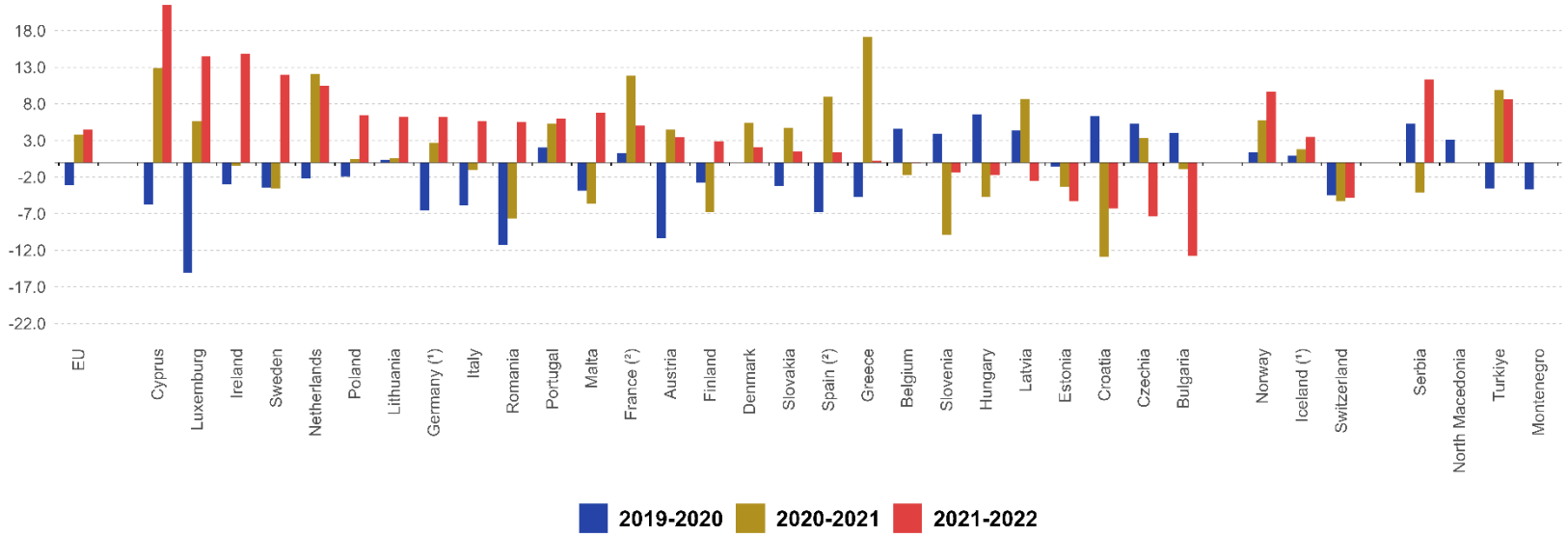
```r
# export png
ggsave("png/fig1.png", plot = fig1, width = 12, height = 6,
       dpi = 1000, create.dir = T)

# include png in Markdown
knitr::include_graphics("png/fig1.png")
```

```r
# export data in excel
write_xlsx(cult_emp_fig1, path = "excel/data_cult_emp_fig1.xlsx")
```

European Commission

Figure 1

Cultural employment – annual rates of change, 2020-2022
(%)



Break in time series in 2021 for all countries for which data are available due to the implementation of the new Regulation (EU)2019/1700, also called the Integrated European Social Statistics Framework Regulation (IESS FR) (see LFS metadata)
(¹) 2020: Break in time series.
(²) 2021, 2022: Definition differs (see LFS metadata).

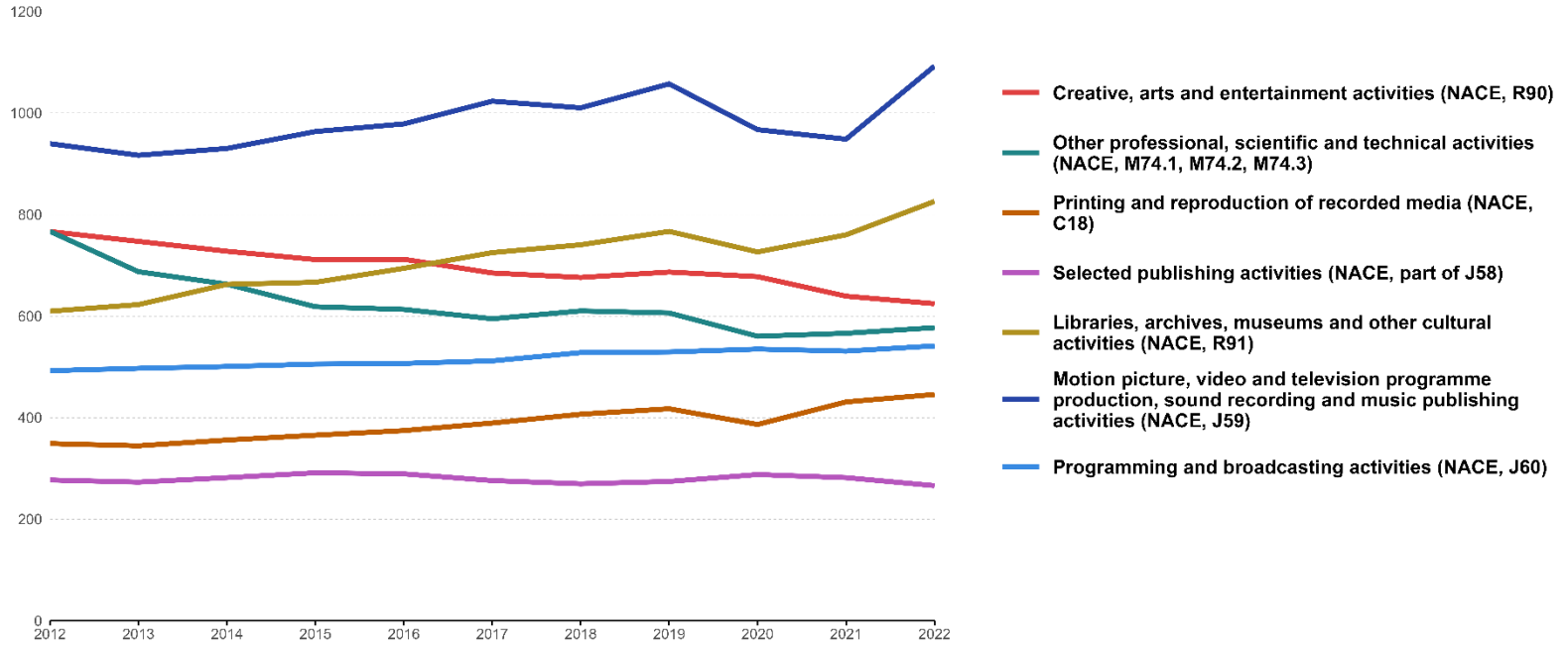
Source: Eurostat (online data code: cult_emp_sex)

eurostat

European Commission

Figure 2

Evolution of cultural employment by selected NACE Rev. 2 activities, EU, 2012-2022
(thousands)



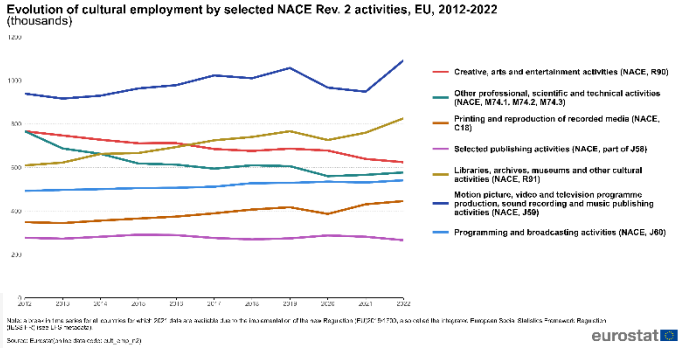
Note: a break in time series for all countries for which 2021 data are available due to the implementation of the new Regulation (EU)2019/1700, also called the integrated European Social Statistics Framework Regulation (IESS FR) (see LFS metadata).

Source: Eurostat(online data code: cult_emp_n2)

eurostat

European Commission

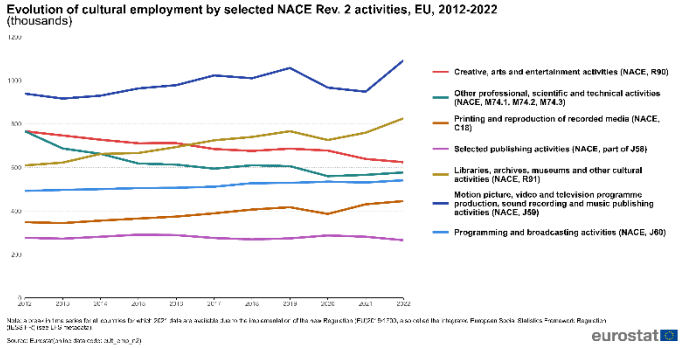
Figure 2




```
# settings
id2 <- "cult_emp_n2"
date4 <- 2012:2022

# title, subtitle, caption, labels
title_fig2 <-  "Evolution of cultural employment by selected NACE Rev. 2 activities, EU, 2012-2022"
subtitle_fig2 <-  "(thousands)"
caption_fig2 <-  paste0("\n",
                        "Note: a break in time series for all countries for which 2021 data are available due to t
he implementation of the new Regulation (EU)2019/1700, also called the integrated European Social Statistics Frame
work Regulation",
                        "\n(IESS FR) (see LFS metadata).
                        ","\n",
                        "\nSource: Eurostat(online data code: ", id2, ")")
labels_fig2 <- c("Creative, arts and entertainment activities (NACE, R90)",
                 "Other professional, scientific and technical activities \n(NACE, M74.1, M74.2, M74.3)",
                 "Printing and reproduction of recorded media (NACE,\nC18)",
                 "Selected publishing activities (NACE, part of J58)",
                 "Libraries, archives, museums and other cultural\nactivities (NACE, R91)",
                 "Motion picture, video and television programme\nproduction, sound recording and music publishing
\nactivities (NACE, J59)",
                 "Programming and broadcasting activities (NACE, J60)")
```

European Commission

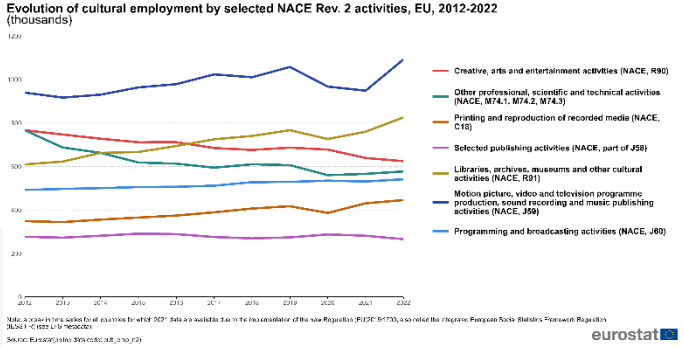
Figure 2




```
# Define colors
col_fig2 <- c('R90' = clrs[1, 3],
              'M74' = clrs[1, 4],
              'C18' = clrs[1, 5],
              'J58' = clrs[1, 6],
              'R91' = clrs[1, 7],
              'J59' = clrs[1, 8],
              'J60' = clrs[2, 3])


# read in data
cult_emp_n2 <- get_eurostat_data(id2, filters = list(time = date4,nace_r2 = c('R90', 'M74', 'C18', 'J58', 'R91',
'J59', 'J60'), geo = "EU27_2020"))
```

European Commission

Figure 2



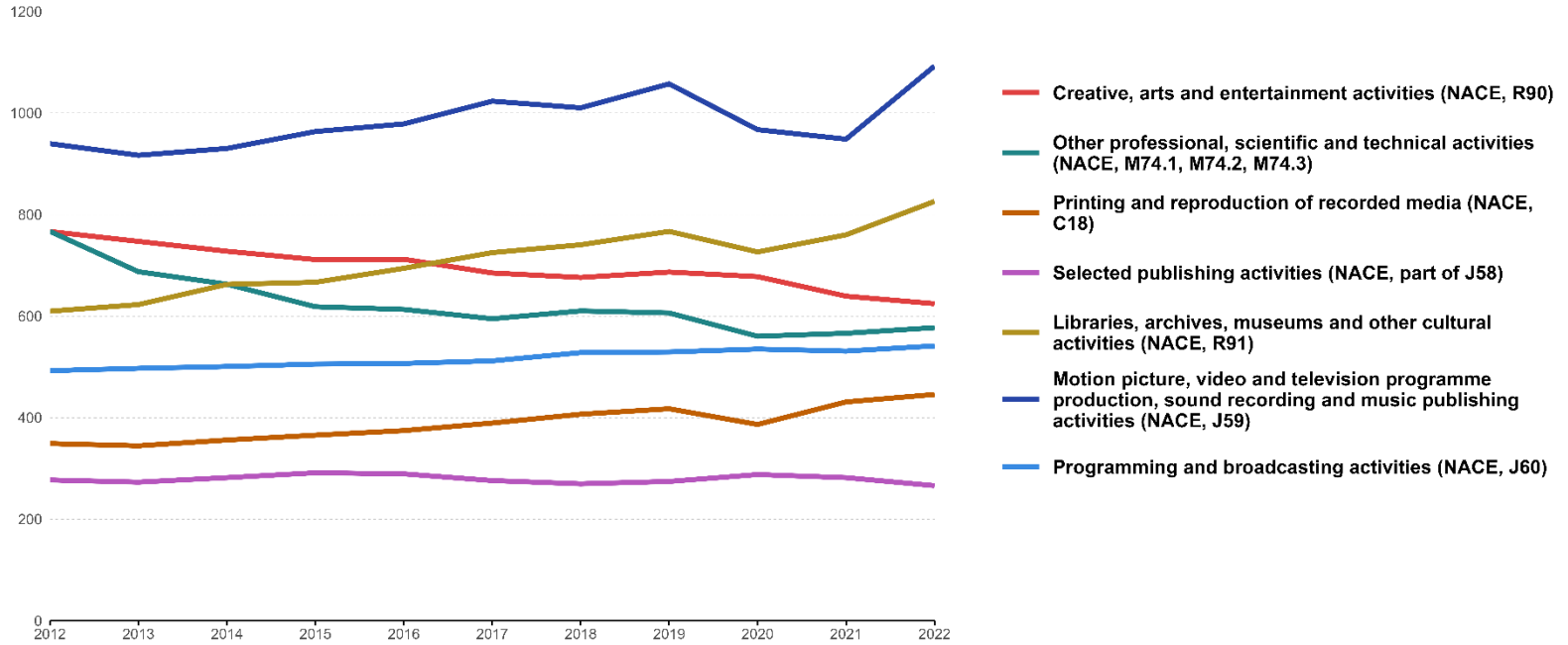

```
# Plot
fig2 <- cult_emp_n2 %>%
  ggplot(aes(x = time, y = values, group = nace_r2, color = nace_r2)) +
  geom_hline(yintercept = seq(0, 1200, by = 200), color = "grey", linetype = "dashed", linewidth = 0.2) +
  geom_line(size = 1.5) +
  labs(title = title_fig2,
       subtitle = subtitle_fig2,
       caption = caption_fig2) +
  scale_y_continuous(breaks = seq(0, 1200, by = 200), expand = c(0, 0)) +
  scale_x_discrete(expand = c(0, 0)) +
  scale_color_manual(values = col_fig2, labels = labels_fig2) +
  theme(legend.background = element_blank(),
        legend.title = element_blank(),
        legend.text = element_text(face = "bold", size = 12),
        legend.key.spacing.y = unit(0.3, "cm"),
        legend.key.size = unit(2, "lines"),
        legend.margin = margin(-1.6, 1, 0, 1, "cm"),
        plot.title = element_text(face = "bold", size = header_size),
        plot.subtitle = element_text(margin = margin(t = -0.2, b = 0.5, unit = "cm"), size = header_size),
        plot.title.position = "plot",
        axis.title.x = element_blank(),
        axis.title.y = element_blank(),
        axis.ticks.x.bottom = element_line(),
        axis.line.x = element_line(),
        axis.text = element_text(size = text_size),
        panel.grid = element_line(),
        axis.ticks = element_blank(),
        panel.background = element_blank(),
        plot.caption = element_text(hjust = 0, size = text_size),
        plot.caption.position = "plot")
```

```
# Combine the plot and the logo
plot_gtable <- ggplotGrob(fig2)

fig2 <- ggdraw() +
  draw_grob(plot_gtable) +
  draw_image(logo, x = 0.99, y = -0.04, width = 0.15, height = 0.15, hjust = 1, vjust = 0)


# export png
ggsave("png/fig2.png", plot = fig2, width = 14, height = 7,
       dpi = 1000, create.dir = T)

# export data in excel
write_xlsx(cult_emp_n2, path = "excel/data_cult_emp_fig2.xlsx")

# include png in markdown
knitr::include_graphics("png/fig2.png")
```

European Commission

Figure 2

Evolution of cultural employment by selected NACE Rev. 2 activities, EU, 2012-2022
(thousands)



Note: a break in time series for all countries for which 2021 data are available due to the implementation of the new Regulation (EU)2019/1700, also called the integrated European Social Statistics Framework Regulation (IESS FR) (see LFS metadata).

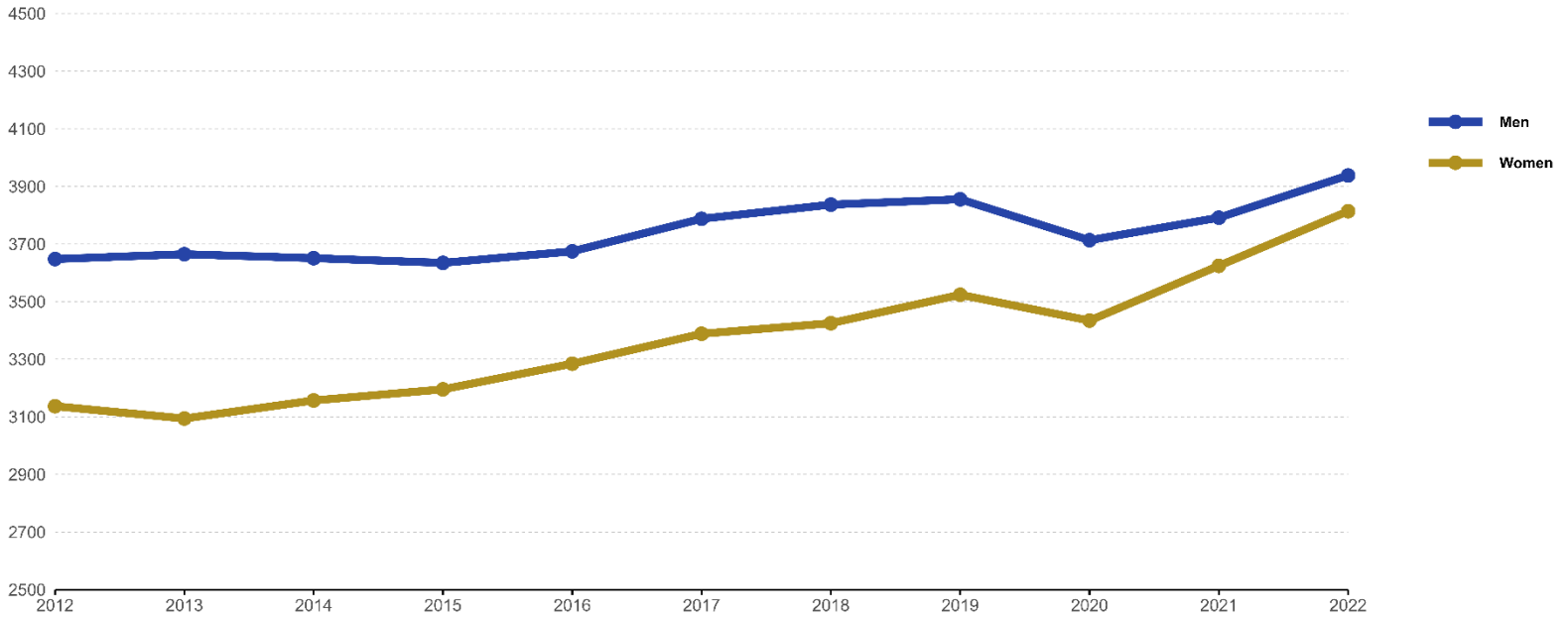
Source: Eurostat(online data code: cult_emp_n2)

eurostat

European Commission

Figure 4

Evolution of cultural employment in the EU by sex, 2012 - 2022
(thousands)



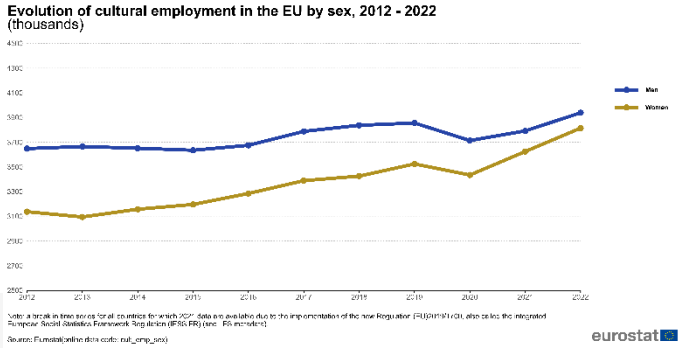
Note: a break in time series for all countries for which 2021 data are available due to the implementation of the new Regulation (EU)2019/1700, also called the integrated European Social Statistics Framework Regulation (IESS FR) (see LFS metadata).

Source: Eurostat(online data code: cult_emp_sex)

eurostat

European Commission

Figure 4

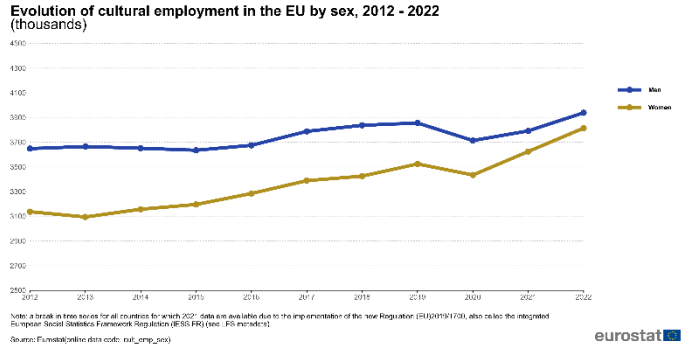



```
#settings
#id1 <- cult_emp_sex    # has been defined for previous graphic
date6 <- 2012:2022

# title, subtitle, caption
title_fig4 <-  paste("Evolution of cultural employment in the EU by sex,", date6[1], "-", date6[length(date6)])
subtitle_fig4 <-  "(thousands)"
caption_fig4 <-  paste0("\n",
                  "Note: a break in time series for all countries for which 2021 data are available due to the im
plementation of the new Regulation (EU)2019/1700, also called the integrated \nEuropean Social Statistics Framewor
k Regulation (IESS FR) (see LFS metadata).",
                  "\n",
                 "\nSource: Eurostat(online data code: ",id1,")")


#colour
col_fig4 <- c("Men" = clrs[1, 3], "Women" = clrs[1, 4])

# read in data
cult_emp_fig4 <- get_eurostat_data(id1, filters = list(time = date6, geo = "EU27_2020",
                                                       unit = "THS_PER", sex = c("M", "F")))
# create factor for sex with the right order and labels
cult_emp_fig4$sex <- factor(cult_emp_fig4$sex, levels = c("M", "F"),
       labels = c("Men", "Women"))
```

European Commission

Figure 4

```r
# the line plot
fig4 <- cult_emp_fig4 %>%
  ggplot( aes(x=time, y= values, group= sex, color = sex)) +
  geom_hline(yintercept = seq(2500, 4500, by = 200), color = "grey", linetype = "dashed", linewidth = 0.2) +
  geom_line(size = 2) +
  geom_point(size = 3) +
  labs(title = title_fig4,
       subtitle = subtitle_fig4,
       caption = caption_fig4) +
  scale_y_continuous(breaks = seq(2500, 4500, by = 200), limits = c(2500, 4500), expand = c(0, 0)) +
  scale_x_discrete(expand = c(0, 0)) +
  scale_color_manual(values = col_fig4) +
  theme(legend.text = element_text(face = "bold", size = 8),
        legend.key.size = unit(2.5, "lines"),
        plot.title = element_text(face = "bold", size = header_size),
        plot.subtitle = element_text(margin = margin(t = -0.2, b = 0.5, unit = "cm"), size = header_size),
        plot.title.position = "plot",
        legend.background = element_blank(),
        legend.margin = margin(-6, 1, 0, 1, "cm"),
        legend.key.spacing.y = unit(-0.5, "cm"),
        legend.title = element_blank(),
        axis.title.x = element_blank(),
        axis.title.y = element_blank(),
        axis.ticks.x.bottom = element_line(),
        axis.line.x = element_line(),
        axis.text = element_text(size = text_size),
        panel.grid = element_line(),
        axis.ticks = element_blank(),
        panel.background = element_blank(),
        plot.caption=element_text(hjust = 0, size = text_size),
        plot.caption.position = "plot") +
    coord_cartesian(clip = 'off')
```

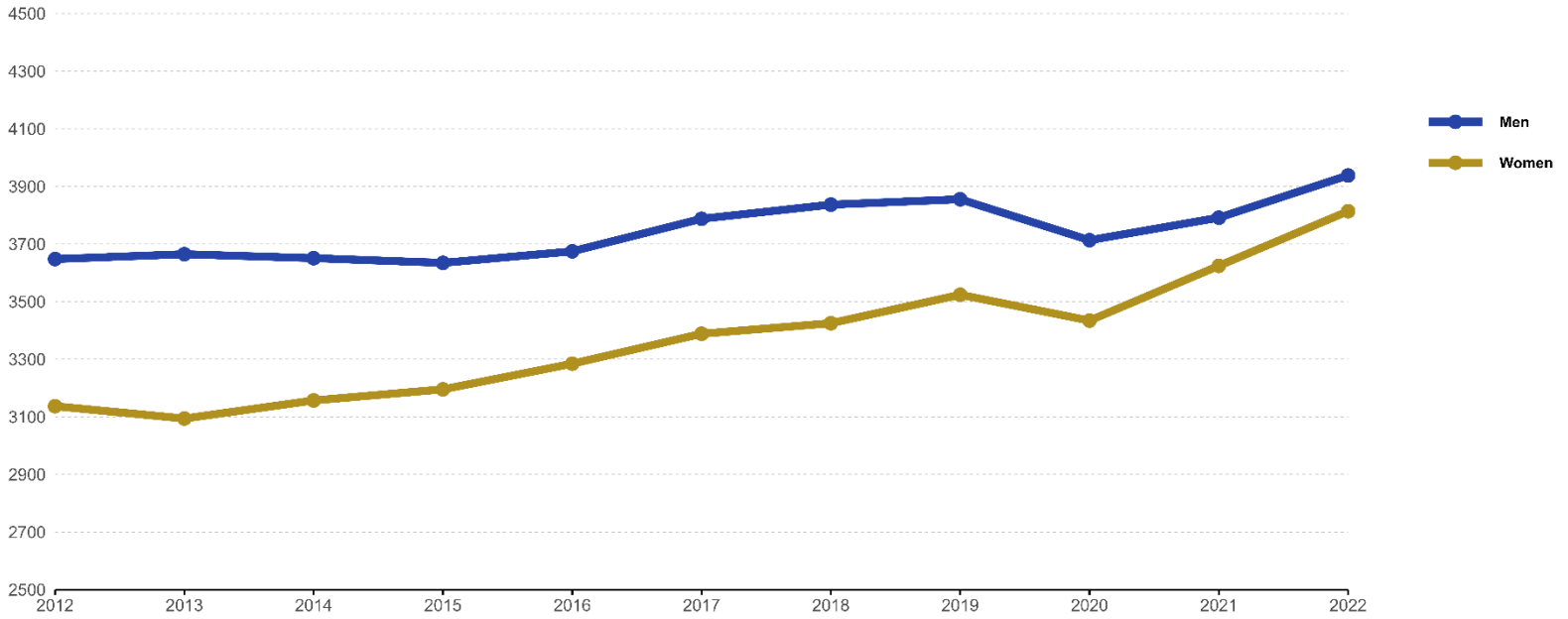



```r
# Combine the plot and the logo
plot_gtable <- ggplotGrob(fig4)

fig4 <- ggdraw() +
  draw_grob(plot_gtable) +
  draw_image(logo, x = 0.99, y = -0.04, width = 0.15, height = 0.15, hjust = 1, vjust = 0)



# export png
ggsave("png/fig4.png", plot = fig4, width = 12, height = 6,
       dpi = 1000, create.dir = T)

# export data in excel
write_xlsx(cult_emp_fig4, path = "excel/data_cult_emp_fig4.xlsx")


#include png in markdown
knitr::include_graphics("png/fig4.png")
```

European Commission

Figure 4

Evolution of cultural employment in the EU by sex, 2012 - 2022
(thousands)



Note: a break in time series for all countries for which 2021 data are available due to the implementation of the new Regulation (EU)2019/1700, also called the integrated European Social Statistics Framework Regulation (IESS FR) (see LFS metadata).

Source: Eurostat(online data code: cult_emp_sex)

eurostat

European Commission

4. Coding Lab discussion

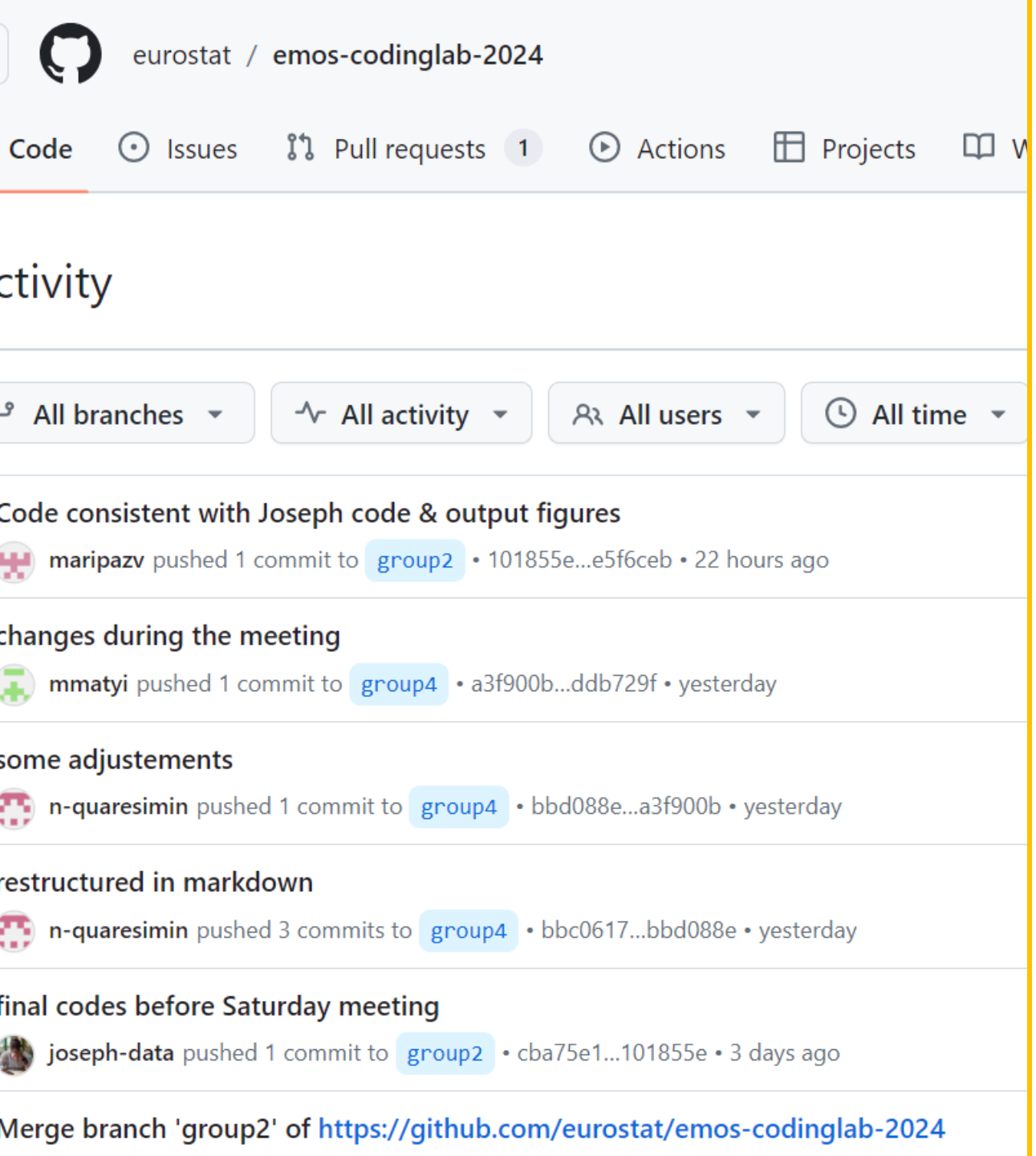
Matyas Tamas MESZAROS
Eurostat, Statistical methodology and data integration

European Commission

The outcome of the 2024 coding lab

- Better understanding of the data and metadata
- Improved programming skills in R
- Experience to work on code in groups
- Appealing visualizations following publication requirements

- Reusable public GitHub repository: eurostat/emos-codinglab-2024 (github.com)



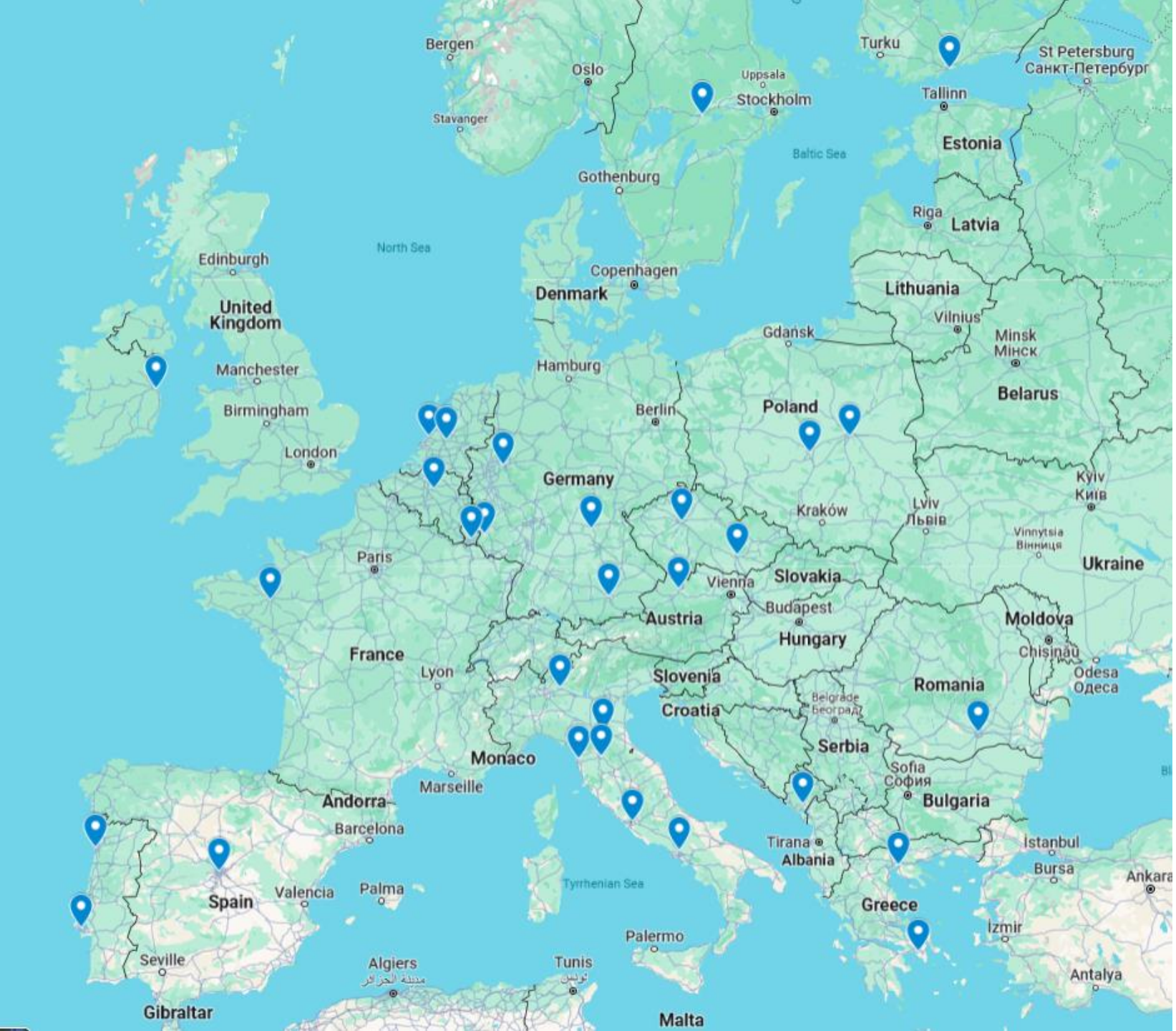
Q&A

European Commission

5. What EMOS has to offer in the next months

Tina Steenvoorden
EMOS secretariat

European Commission



EUROPEAN MASTER IN OFFICIAL STATISTICS
EMOS

EMOS today

31 EMOS-labelled master's programmes

17 European countries

European Commission

Master’s thesis competition

MTC in general

- highlights official statistics as a research topic and puts forward young talents with innovative contributions
- **award ceremony** organised at the NTTS conference taking place every two years

EMOS Master’s thesis competition 2024

- **opening** of the call for applications: September 2024
- **award ceremony**: March 2025
- **eligible candidates**: students graduating in 2023 and 2024

2023 EMOS Master thesis competition winners' video testimonials

European Big Data Hackathon

EBDH in general

- teams compete with data products combining official statistics and new source data to support policymakers in a pressing challenge
- **award ceremony** organised at the NTTS conference

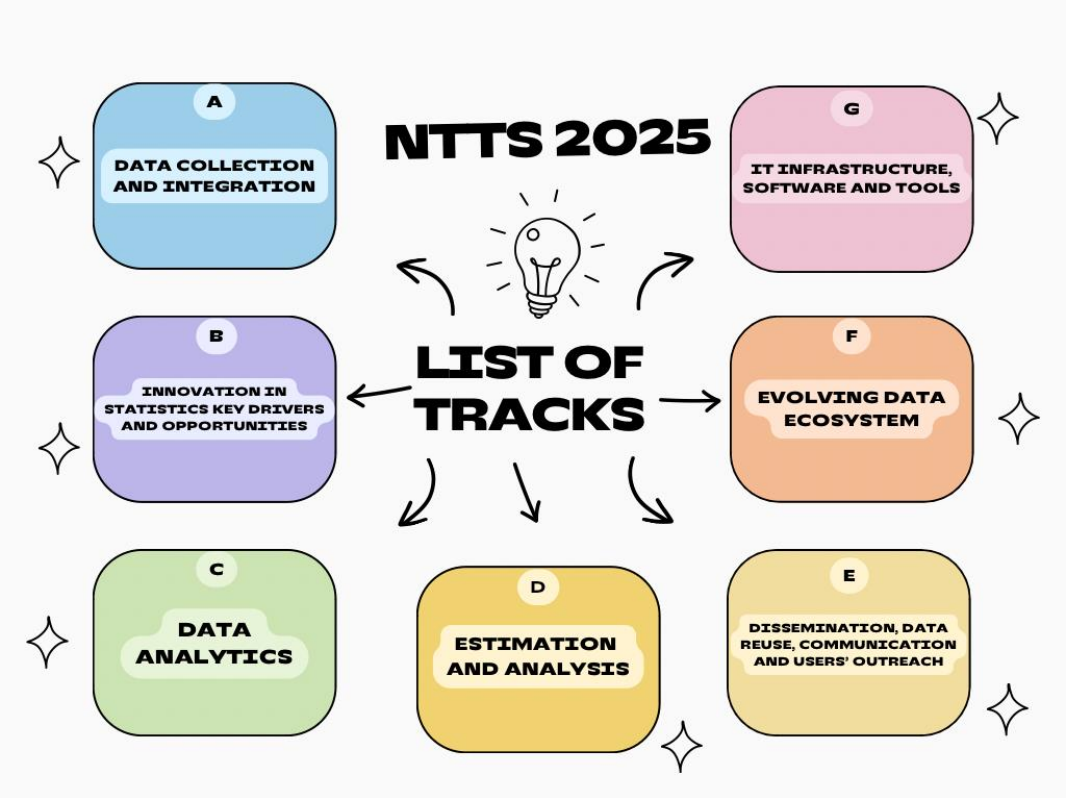
EBDH 2025

- **opening** of the team registrations: October 2024
- EBDH takes place in March 2025
- Volunteers for evaluation panel and participation as a team

Conference on New Technologies and Techniques for Statistics 2025

KEY DATES

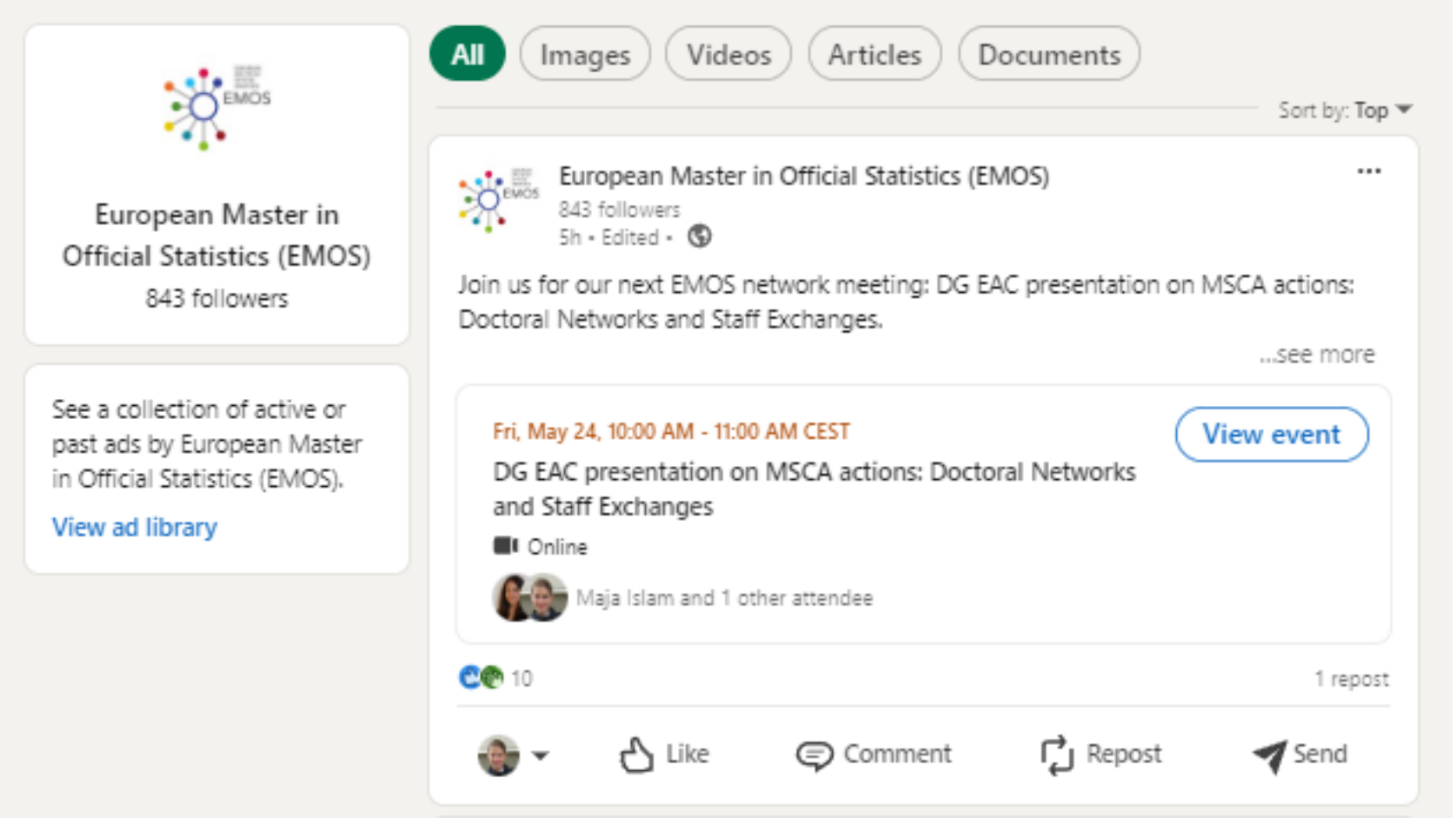
- Opening call for abstracts: 1 June 2024
- Deadline for submission: 30 Sept 2024
- Notification of acceptance: 16 Nov 2024
- Registration deadline: 15 February 2025
- Accepted speakers are expected to participate in-person in Brussels
- NTTS takes place 11-13 March 2025
- NTTS 2025 is free of charge for participants



Full list of topics: https://cros.ec.europa.eu/dashboard/ntts-2025

European Commission

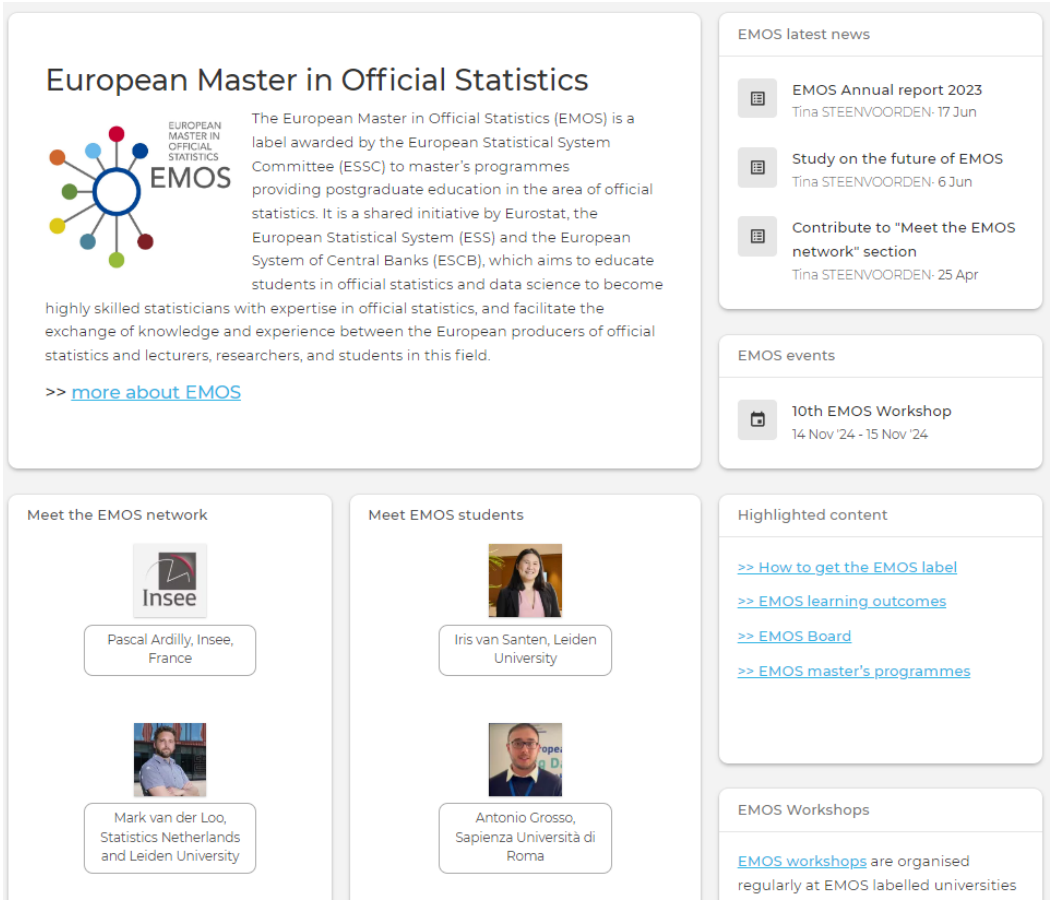
EMOS on LinkedIn



Meet EMOS students

- New EMOS website launched early this year
- Section on the EMOS website
- Students invited to contribute
- To be featured, send us your contribution:

  https://ec.europa.eu/eusurvey/runner/Meet_EMOS_students



European Commission

Thank you!

Keep in touch:

@EU_EMOS

@europeanmasterinofficialst492

www.linkedin.com/company/theemos

ESTAT-EMOS@ec.europa.eu