# Garonzi_Homework_block3

2022-10-18

Income of International Organizations

The 'io_income_rs.csv' dataset contains funding data from some international organizations in Geneva. Please load the dataset, we will be working with it today. *Tip: Adding titles, legends, and themes to your plots helps us help you!*

Question 1

Create a bar plot displaying the top 10 overall donors and their total donations to all international organizations. Please color the donors by their type (i.e. public or private).

```r
io_income_rs%>%
  drop_na()%>%
  group_by(donor,type_donor)%>%
  summarise(total_donations= sum(amount_nominal,na.rm=TRUE))%>%
  arrange(-total_donations) %>%
  ungroup()%>%
slice_head(n=10)%>%
  ggplot(aes(x = reorder (donor, total_donations),
             y = total_donations,
             fill = type_donor)) +
  geom_col() +
   labs(title = "Top donors",
        x = "Donors",
        y = "Total donations",
        fill = "Type of donor") +
  theme_classic()+
  coord_flip()
```

Top donors
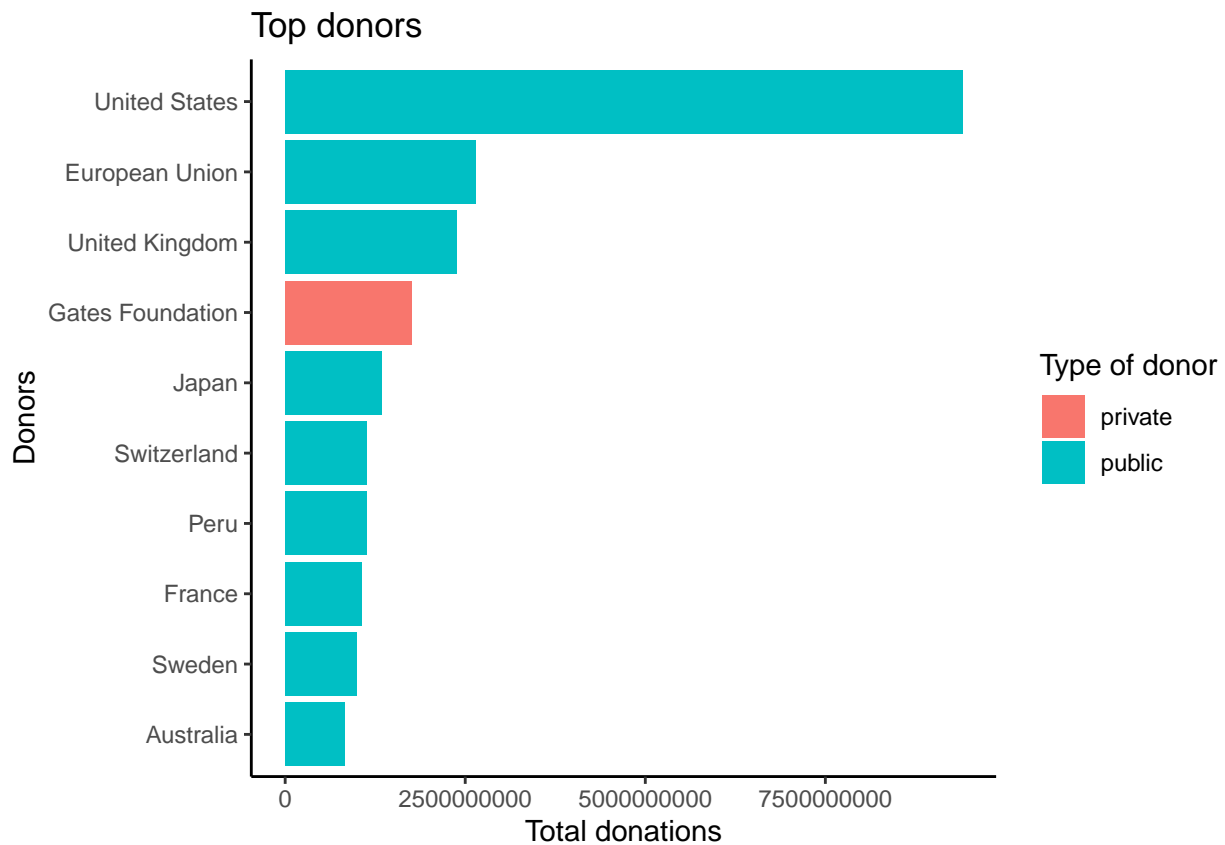
Question 2

Has the amount the US donates to all issue areas increased over time? Please illustrate this relationship in a line plot, colored by the respective issue areas.

```
io_income_rs%>%
  drop_na()%>%
  filter(donor=="United States")%>%
  group_by(issue_area,year)%>%
  summarise(amount_nominal=sum(amount_nominal, na.rm=TRUE))%>%
  ggplot(aes(x = year, y =amount_nominal)) +
  geom_line(aes(color = issue_area)) +
   labs(title = "US donations over time by issue area",
       x = "Year",
       y = "Donations amount") +
  theme_classic()+
  facet_wrap(~issue_area)
```

US donations over time by issue area
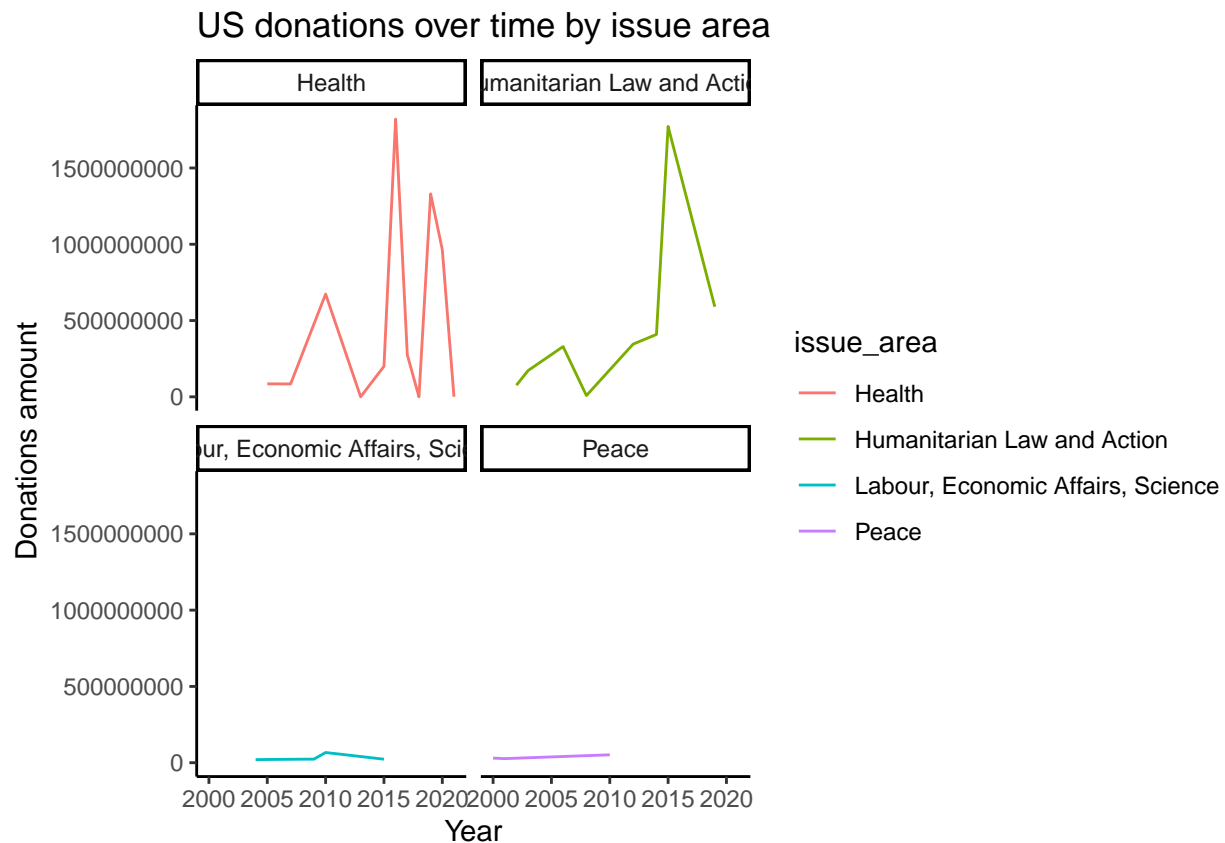
Question 3

Plot the distribution of all donations in the year 2000 and in the year 2020 comparing public and private donors. Are there outliers for either of these types of donors in 2000? What about in 2020? (Tip: box plots are great for distributions! Please treat year as factor and scale the nominal amount using "scale_y_log10(labels = scales::dollar)").

```
io_income_rs%>%
  drop_na(type_donor)%>%
  filter(year == 2000 | year == 2020)%>%
  group_by(type_donor,amount_nominal)%>%
  ggplot(aes(x = type_donor, y =amount_nominal, fill = type_donor)) +
  geom_boxplot()+
  facet_wrap(~year)+
  scale_y_log10(labels = scales::dollar)+
   labs(title = "Donations by type of donor in 2000 and 2020",
       x = "",
       y = "Donations amount",
       fill = "Type of donor") +
  theme_classic()
```

Donations by type of donor in 2000 and 2020
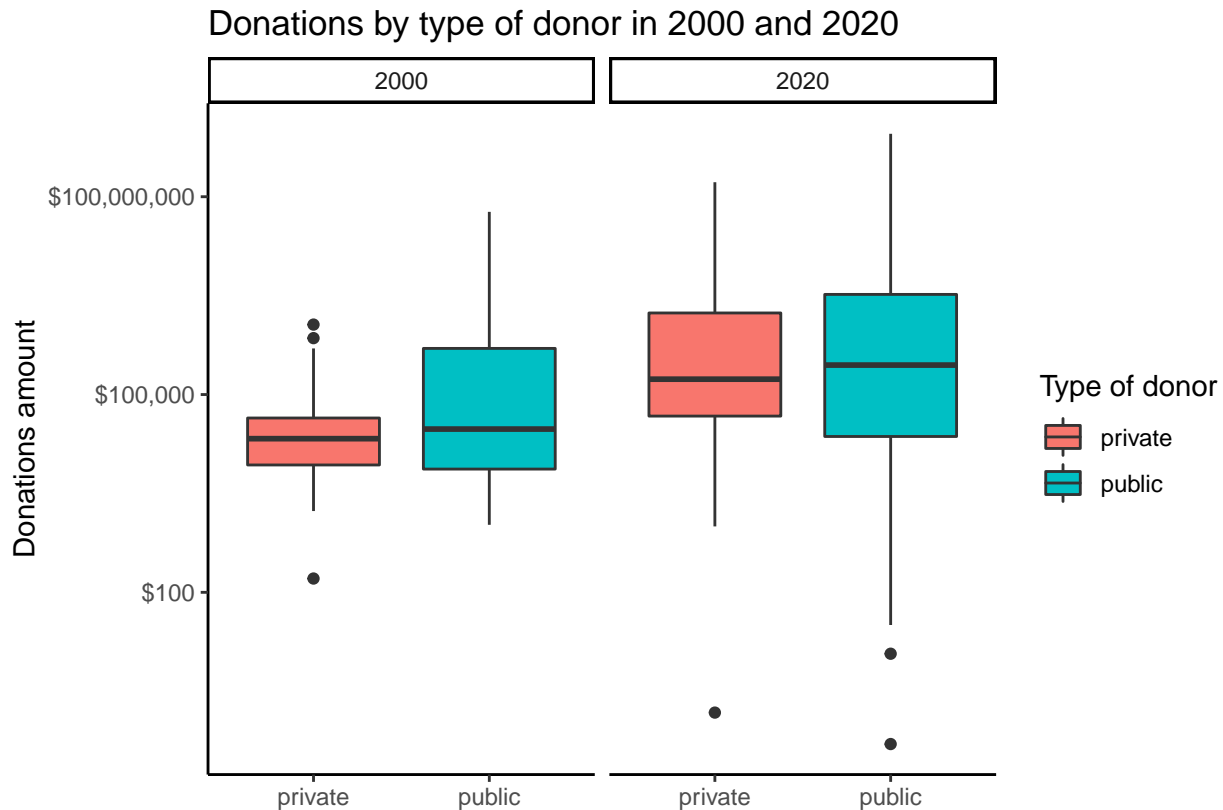
The plot shows that in the year 2000 there were some outliners within private donors, both with very high or very low values, while it does not show outliners for public donors. For the year 2020, the plot shows only outliners with very low values, this time for both public and private donors.

Question 4

Create four scatter plots, one for each issue area, containing all donation by year per donor type (Tip: use "facet_wrap(~issue_area)"). The shape of the points should reflect the donor type (Tip: use "geom_point(aes(shape = donor_type"))"). In each facet, please add a smoothed line to show the direction of the relationship (TIP: use"geom_smooth()"). The smoothed lines should be colored"red".

```
io_income_rs%>%
  drop_na()%>%
  ggplot(mapping = aes(x=year,y=amount_nominal))+
  geom_point(aes(shape=type_donor), alpha=.5)+
  facet_wrap(~issue_area, scales = "free_y")+ #
  geom_smooth(method="lm", color="red")+
  scale_y_log10()+
  labs(title = "",
      x = "",
      y = "Donations amount",
      shape = "Type of donor") +
  theme_classic()
```