NewmanHW3

Joshua Newman

## 2022-10-20

Joshua Newman Homework 3
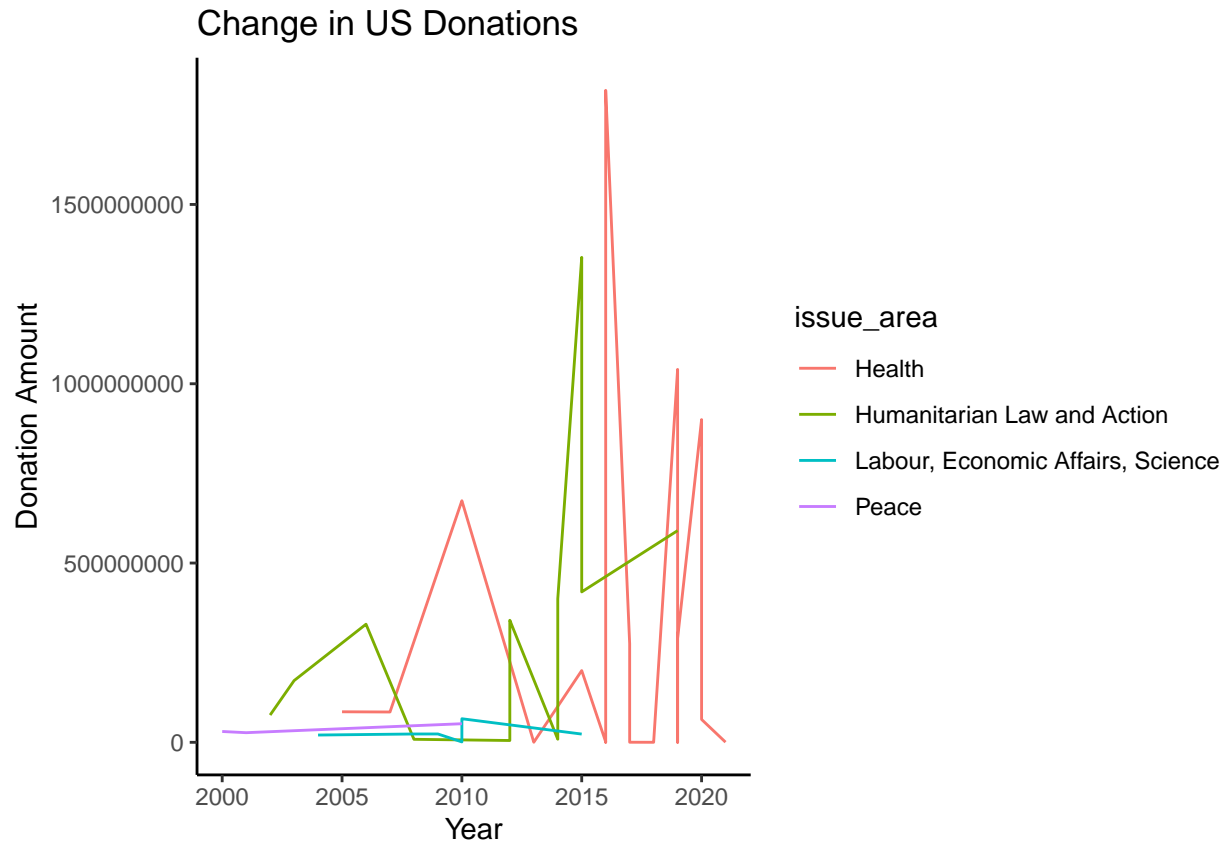
Question 1

```r
order_by_amount <- aggregate(io_income_rs$amount_nominal,
                             list(io_income_rs$donor), FUN=sum)
order_by_amount$year = NULL
order_by_amount <- rename(order_by_amount, "donor" = Group.1)
order_by_amount <- rename(order_by_amount, "amount_nominal" = x)
order_by_amount <- order_by_amount[order(-order_by_amount$amount_nominal),]
order_by_amount <- slice(order_by_amount, c(1, 2, 3, 4, 5, 6, 7, 8, 9, 10))
order_by_amount <- left_join(order_by_amount, io_income_rs)
order_by_amount$year = NULL
order_by_amount$issue_area = NULL
order_by_amount$type_donor <- ifelse(grepl("United States|European Union|
                                            |United Kingdom|Japan|Switzerland|
                                            |Peru|France|Sweden|Australia",
                              order_by_amount$donor), "Public",
                              order_by_amount$type_donor)
order_by_amount$type_donor <- ifelse(grepl("Gates Foundation",
                                            order_by_amount$donor), "Private",
                                 order_by_amount$type_donor)

order_by_amount %>%
  ggplot(aes(x = amount_nominal,
             y = donor,
             fill = as.factor(type_donor))) +
  geom_col(position= "dodge") +
  labs(title = "Top 10 Donors",
       x = "Amount Donated",
       y = "Donors",
       fill = "Type of Donor") +
  scale_fill_manual(values = c("lightblue", "lightgreen")) +
  theme_classic()
```

Top 10 Donors
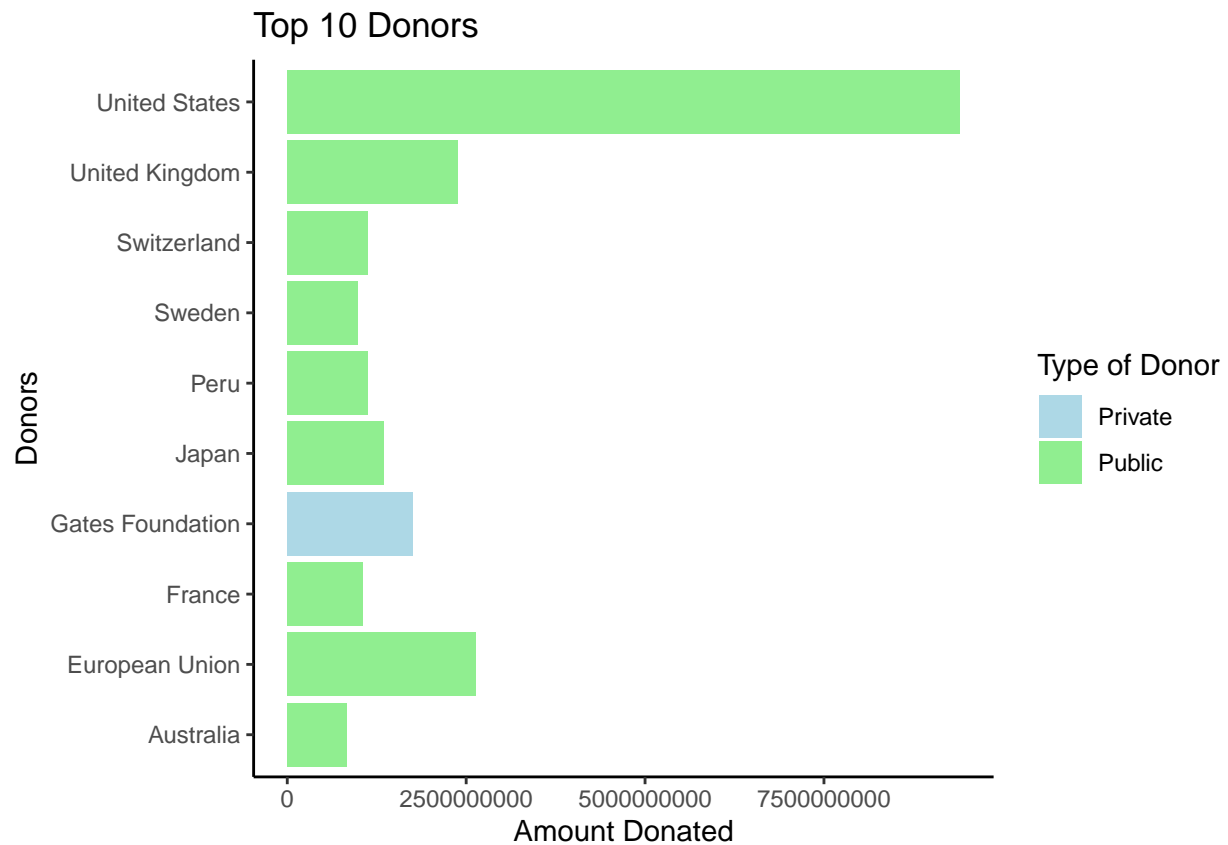
Question 2

```
io_income_rs %>%
  filter(donor == "United States") %>%
  group_by(issue_area) %>%
  ggplot(aes(x = year, y = amount_nominal)) +
  geom_line(aes(color = issue_area)) +
  labs(title = "Change in US Donations",
       x = "Year",
       y = "Donation Amount") +
  theme_classic()
```

Change in US Donations

Based on this line chart, US donations in each year vary wildly.

Question 3

```
io_income_rs %>%
  filter(year == 2000 | year == 2020) %>%
  ggplot(aes(x = type_donor, y = amount_nominal, fill = type_donor)) +
  scale_y_log10(labels = scales::dollar) +
  geom_boxplot(aes(color = type_donor)) +
  facet_wrap(~year) +
  labs(title = "Public and Private Donations in 2000 and 2020",
       x = "Year",
       y = "Donation Amount") +
  theme_classic()
```

Public and Private Donations in 2000 and 2020
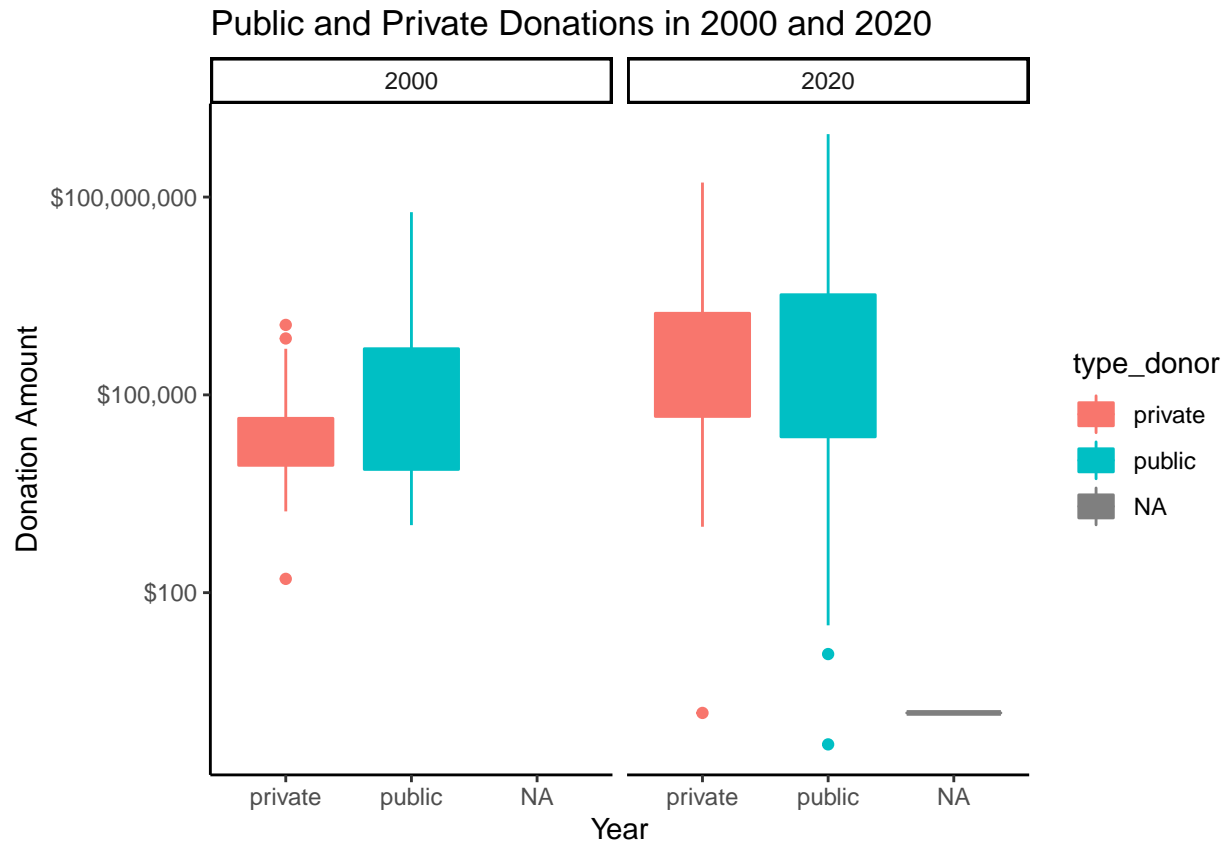
According to the graph, in the year 2000, private donors had two high outliers and one low outlier. Public donors had no outliers. In 2020, private only had one low outlier, while public had two low outliers.

Question 4

```
io_income_rs %>%
  ggplot(aes(x = year, y = amount_nominal)) +
  geom_point(aes(shape = type_donor)) +
  geom_smooth(color = "red") +
  facet_wrap(~issue_area) +
  labs(title = "Donations per Year per Issue Area",
       x = "Year", y = "Amount Donated") +
  scale_y_log10(labels = scales::dollar) +
  theme_classic()
```

Donations per Year per Issue Area