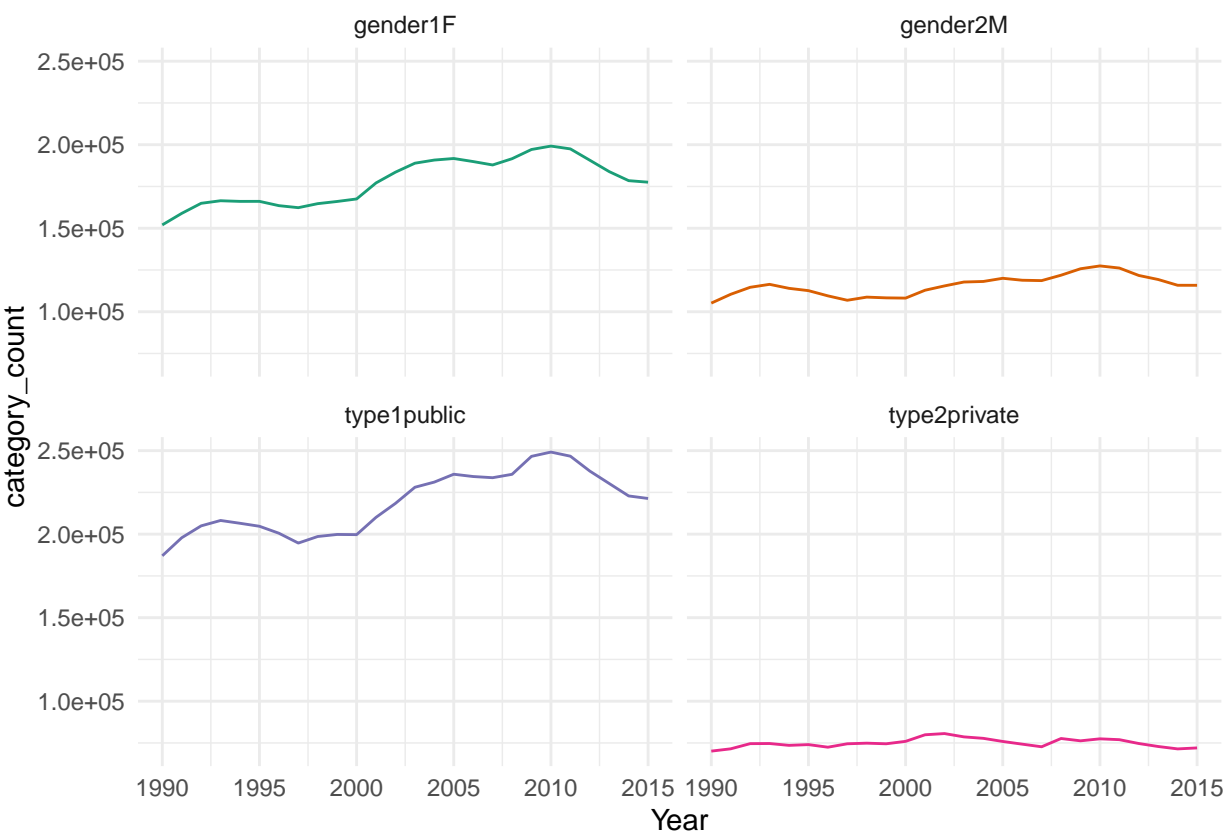
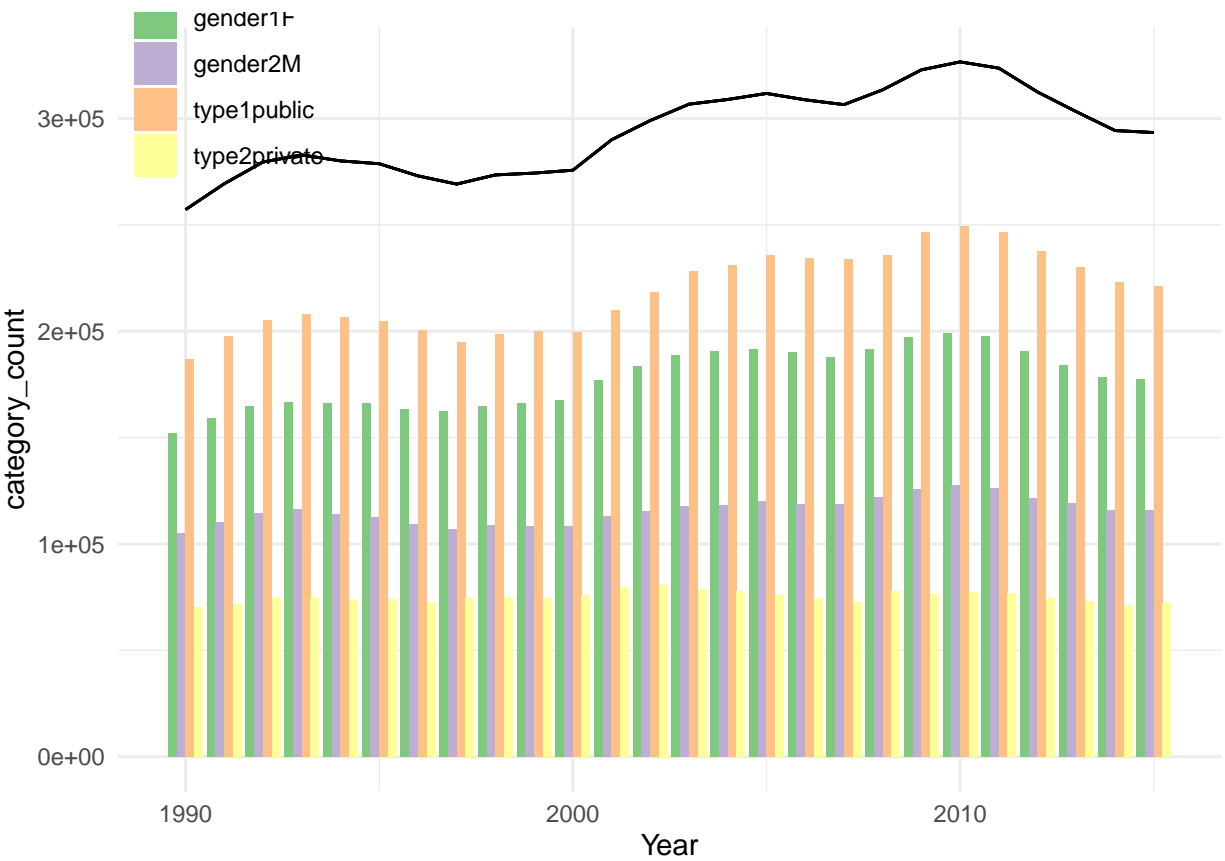


tidytuesday - HBCU Enrollment

Yuzhou Pan

2021-02-04

Setup and plots



What to improve

- Try mosaic plots
- Check `pivot_longer`, instead of conventional `gather`

Codes

```
hbcu_all <- readr::read_csv('https://raw.githubusercontent.com/rfordatascience/tidytuesday/master/data/2020/2020-01-13/hbcu.csv')

library(tidyverse)
library(gridExtra)
hbcu_all %>%
  filter(Year >= 1990) %>%
  select(c("Year", "Total enrollment", "Females", "Males",
           "Total - Public", "Total - Private")) %>%
  rename(gender1F = Females, gender2M = Males,
         type1public = `Total - Public`, type2private = `Total - Private`) %>%
  gather("category", "category_count", gender1F, gender2M, type1public, type2private) -> hbcu_gathered

hbcu_gathered %>%
  ggplot(aes(x = Year, y = category_count, fill = category)) +
  geom_bar(stat = "identity", position = "dodge") +
  geom_line(aes(x = Year, y = `Total enrollment`)) +
  theme_minimal() +
  theme(legend.position = c(0.1, 0.95)) +
  scale_fill_brewer(palette = "Accent")

hbcu_gathered %>%
  ggplot(aes(x = Year, y = category_count, col = category)) +
  geom_line(stat = "identity") +
  theme_minimal() +
  theme(legend.position = "none") +
  scale_color_brewer(palette = "Dark2") +
  scale_y_continuous(labels = function(x) format(x, scientific = TRUE)) +
  facet_wrap(~ category)
```