Tidy-week12

Xingnuo Zhang

3/20/2021

```r
#install.packages("tidytuesdayR")
library("tidytuesdayR")
tuesdata <- tidytuesdayR::tt_load('2021-03-16')
```

```
## --- Compiling #TidyTuesday Information for 2021-03-16 ----

## --- There is 1 file available ---

## --- Starting Download ---

##
##  Downloading file 1 of 1: `games.csv`

## --- Download complete ---
```

```r
tuesdata <- tidytuesdayR::tt_load(2021, week = 12)
```

```
## --- Compiling #TidyTuesday Information for 2021-03-16 ----

## --- There is 1 file available ---

## --- Starting Download ---

##
##  Downloading file 1 of 1: `games.csv`

## --- Download complete ---
```

```r
games <- tuesdata$games
```

```r
#install.packages("dplyr")
library(dplyr)
```

```
##
## Attaching package: 'dplyr'

## The following objects are masked from 'package:stats':
##
##     filter, lag

## The following objects are masked from 'package:base':
##
##     intersect, setdiff, setequal, union
```

```r
library(tidyverse)
```

```
## -- Attaching packages --------------------------------------- tidyverse 1.3.0 --

## v ggplot2 3.3.3      v purrr   0.3.4
## v tibble  3.1.0      v stringr 1.4.0
```

```
## v tidyr    1.1.3     v forcats 0.5.1
## v readr    1.4.0

## -- Conflicts ----------------------------------------- tidyverse_conflicts() --
## x dplyr::filter() masks stats::filter()
## x dplyr::lag()    masks stats::lag()
```
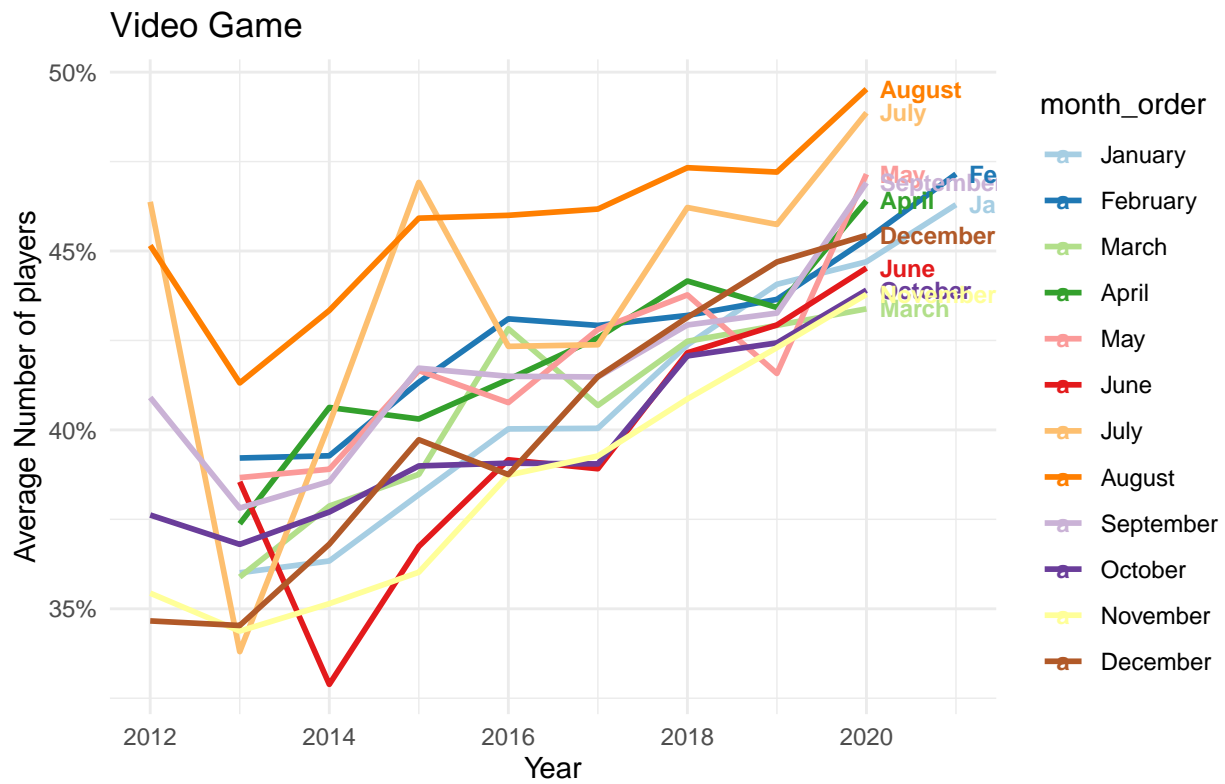
```r
library(ggplot2)
library(extrafont)
```

```
## Registering fonts with R
```

```r
games %>%
  left_join(data.frame(month = month.name, month_id = 1:12), by = "month") %>%
  mutate(ap_perc = as.numeric(str_remove(avg_peak_perc, '%')),
    month_order = factor(month_id, ordered = TRUE, levels = 1:12, labels = month.name)) %>%
  group_by(year, month_order) %>%
  summarise(avg_peak_per = mean(ap_perc, na.rm = TRUE),
            sd = sd(ap_perc, na.rm = TRUE)) %>%
  arrange(year, month_order) %>%
  group_by(month_order) %>%
  mutate(y_txt = ifelse(row_number() == n(), avg_peak_per, NA),
         x_txt = ifelse(row_number() == n(), year, NA)) %>%
  ggplot(aes(x = year, y = avg_peak_per)) +
  geom_line(aes(group = month_order, col = month_order), size = 1) +
  geom_text(aes(x_txt, label = paste0("  ", month_order), col = month_order),
    size = 3,
    hjust = 0,
    fontface = "bold")+
  scale_y_continuous(labels = scales::percent_format(scale = 1, accuracy = 1)) +
  scale_color_brewer(palette = "Paired") +
  theme_minimal() +
  labs(x = "Year",
    y = "Average Number of players",
    title = "Video Game",
    caption = str_c("Created by: Xingnuo Zhang, U of T\n",
                    "Data Source: StreamCharts"))
```

```
## `summarise()` has grouped output by 'year'. You can override using the `.groups` argument.
```

```
## Warning: Removed 92 rows containing missing values (geom_text).
```

Video Game

Created by: Xingnuo Zhang, U of T
Data Source: StreamCharts