8/3/2019 joyplot.R

```
# CREATE DATA FOR JOYPLOT
 1
 2
   ## Load Data ####
 3
   genre_df <- read_csv('../data/cleaned_skatedata.csv')</pre>
4
5
   ## Create Dataframe of genre percentages by year
 6
7
    genre_pct_by_year <- genre_df %>%
      select(genre, year) %>%
8
9
      filter(year > 1988) %>%
10
      group_by(year) %>%
      mutate(year cnt = n()) %>%
11
12
      ungroup() %>%
13
      group_by(genre, year, year_cnt) %>%
14
      summarise(gen cnt = n()) %>%
15
      ungroup() %>%
16
      group_by(genre, year) %>%
17
      summarise(genre_pct = gen_cnt / year_cnt) %>%
18
      ungroup() %>%
19
      arrange(year) %>%
20
      rename(activity = genre,
21
             time = year,
22
             p = genre_pct)
23
24
   ## Create smooth dataset of areas for joyplot ####
25
    joyplot data <- genre pct by year %>%
      group_by(activity) %>%
26
27
      filter(!is.na(activity)) %>%
28
      arrange(activity, time) %>%
      mutate(p_peak = p/max(p), # Normalize as percentage of peak popularity
29
             p smooth = (lag(p peak) + p peak + lead(p peak)) / 3, # Moving average
30
             p smooth = coalesce(p smooth, p peak)) %>% # When there's no lag or lead, we
31
    get NA. Use the pointwise data
32
      ungroup() %>%
33
      mutate(activity = reorder(activity, p, FUN=which.max))
34
   ## Save data to tsv ####
35
   write tsv(joyplot data, "data.tsv")
36
37
```