

```
1 # CREATE DATA FOR JOYPLOT
2
3 ## Load Data ####
4 genre_df <- read_csv('../data/cleaned_skatedata.csv')
5
6 ## Create Dataframe of genre percentages by year
7 genre_pct_by_year <- genre_df %>%
8   select(genre, year) %>%
9   filter(year > 1988) %>%
10  group_by(year) %>%
11  mutate(year_cnt = n()) %>%
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 %>%
26   group_by(activity) %>%
27   filter(!is.na(activity)) %>%
28   arrange(activity, time) %>%
29   mutate(p_peak = p/max(p), # Normalize as percentage of peak popularity
30          p_smooth = (lag(p_peak) + p_peak + lead(p_peak)) / 3, # Moving average
31          p_smooth = coalesce(p_smooth, p_peak)) %>% # When there's no lag or lead, we
32   get NA. Use the pointwise data
33   ungroup() %>%
34   mutate(activity = reorder(activity, p, FUN=which.max))
35
36 ## Save data to tsv ####
37 write_tsv(joyplot_data, "data.tsv")
```