Analysis of ACS 2022

Lexun Yu   Colin Sihan Yang   Siddharth Gowda
Tanmay Shinde

October 10, 2024

Use the Spotify API to analyze data about ...

Table of contents

```
# install.packages("spotifyr")
# install.packages("tidyverse")
# install.packages("here")
# install.packages("tinytex")
# install.packages("dplyr")
# install.packages("tibble")
# install.packages("opendatatoronto")
# install.packages("janitor")
# install.packages("ggplot2")
# install.packages("knitr")
# install.packages("tesseract")
library(babynames)
```

```
Warning: package 'babynames' was built under R version 4.3.3
```

```
library(gh)
```

```
Warning: package 'gh' was built under R version 4.3.3
```

```r
library(here)
```

```
Warning: package 'here' was built under R version 4.3.3
```

```
here() starts at D:/Colin_Sihan_Yang/UT/STA304/spotifyanalysis
```

```r
library(httr)
```

```
Warning: package 'httr' was built under R version 4.3.3
```

```r
library(janitor)
```

```
Warning: package 'janitor' was built under R version 4.3.3
```

```
Attaching package: 'janitor'
```

```
The following objects are masked from 'package:stats':

    chisq.test, fisher.test
```

```r
library(jsonlite)
```

```
Warning: package 'jsonlite' was built under R version 4.3.3
```

```r
library(knitr)
```

```
Warning: package 'knitr' was built under R version 4.3.3
```

```r
library(lubridate)
```

```
Warning: package 'lubridate' was built under R version 4.3.3
```

```
Attaching package: 'lubridate'
```

```
The following objects are masked from 'package:base':

    date, intersect, setdiff, union
```

```r
library(pdftools)
```

```
Warning: package 'pdftools' was built under R version 4.3.3
```

```
Using poppler version 23.08.0
```

```r
library(purrr)
```

```
Warning: package 'purrr' was built under R version 4.3.3
```

```
Attaching package: 'purrr'
```

```
The following object is masked from 'package:jsonlite':

    flatten
```

```r
library(rvest)
```

```
Warning: package 'rvest' was built under R version 4.3.3
```

```r
library(spotifyr)
library(tesseract)
```

```
Warning: package 'tesseract' was built under R version 4.3.3
```

```r
library(tidyverse)
```

```
Warning: package 'tidyverse' was built under R version 4.3.3
```

```
Warning: package 'ggplot2' was built under R version 4.3.3
```

```
Warning: package 'tibble' was built under R version 4.3.3
```

```
Warning: package 'tidyr' was built under R version 4.3.3


Warning: package 'readr' was built under R version 4.3.3


Warning: package 'dplyr' was built under R version 4.3.3


Warning: package 'stringr' was built under R version 4.3.3


Warning: package 'forcats' was built under R version 4.3.3


-- Attaching core tidyverse packages ---------------------- tidyverse 2.0.0 --
v dplyr     1.1.4      v stringr 1.5.1
v forcats   1.0.0      v tibble  3.2.1
v ggplot2   3.5.1      v tidyr   1.3.1
v readr     2.1.5


-- Conflicts ------------------------------------------ tidyverse_conflicts() --
x dplyr::filter()         masks stats::filter()
x purrr::flatten()        masks jsonlite::flatten()
x readr::guess_encoding() masks rvest::guess_encoding()
x dplyr::lag()            masks stats::lag()
i Use the conflicted package (<http://conflicted.r-lib.org/>) to force all conflicts to beco
```

```r
library(usethis)
```

```
Warning: package 'usethis' was built under R version 4.3.3
```

```r
library(xml2)
```

```
Warning: package 'xml2' was built under R version 4.3.3
```

1 Get Spotify Access

```r
Sys.setenv(SPOTIFY_CLIENT_ID = 'fa302d336c68442a9c10250efda1f07a')
Sys.setenv(SPOTIFY_CLIENT_SECRET = 'd17f5eb4d5494842b4f8158df0654b4f')

access_token <- get_spotify_access_token()
```
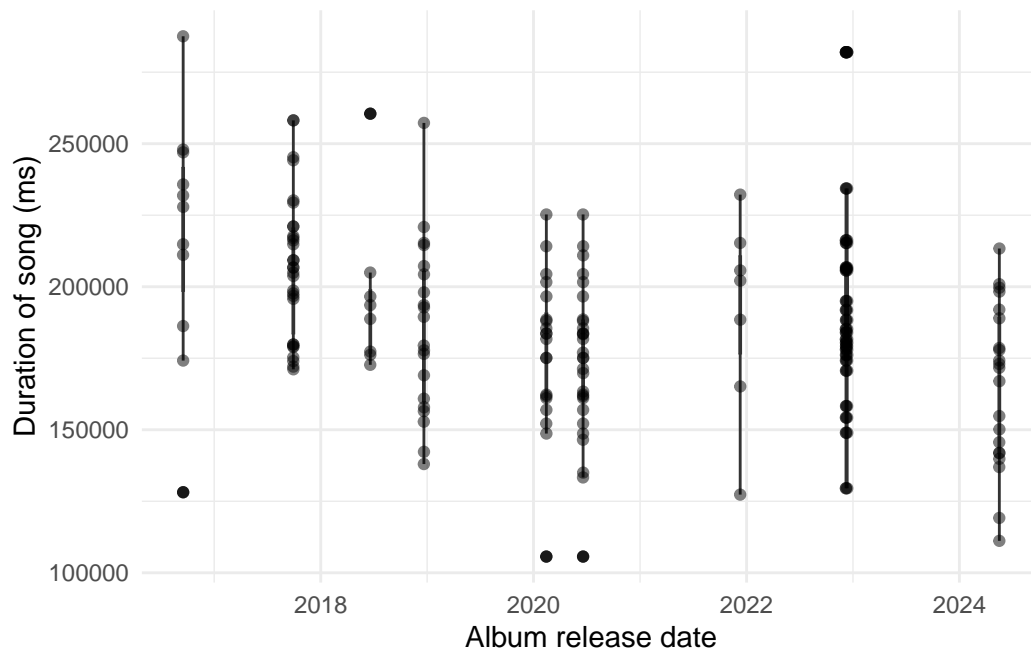
```r
radiohead <- get_artist_audio_features("radiohead")
saveRDS(radiohead, "radiohead.rds")
```

```r
boogie <- get_artist_audio_features("A Boogie Wit Da Hoodie")
saveRDS(boogie, "boogie.rds")
boogie = readRDS("boogie.rds")
```

```r
PTP = get_artist_audio_features("Pay money To my Pain")
saveRDS(PTP, "PTP.rds")
PTP = readRDS("PTP.rds")
```
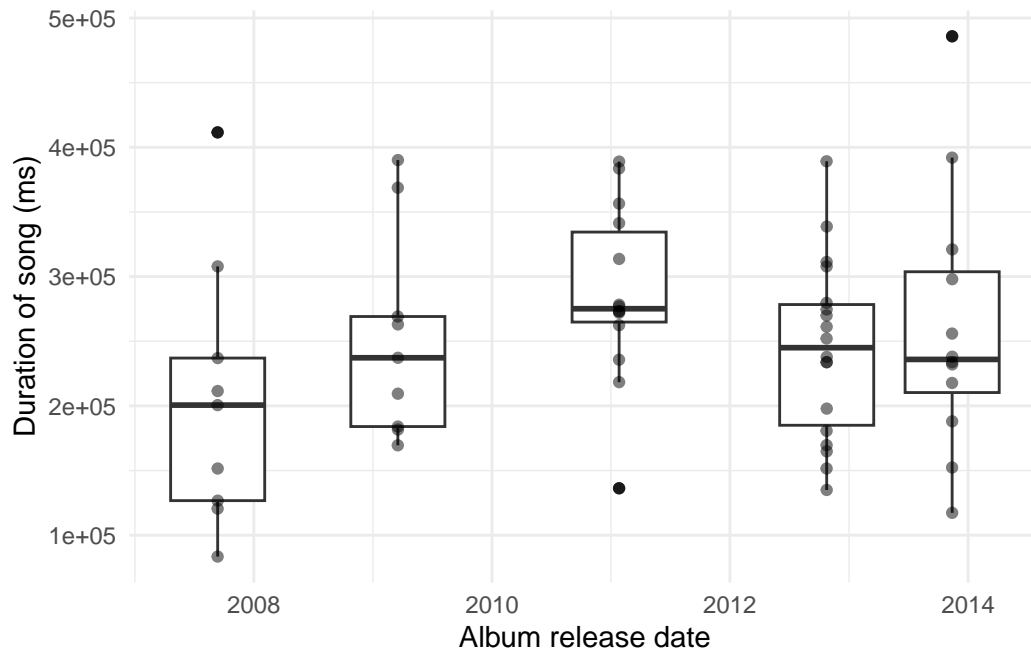
```r
boogie <- as_tibble(boogie)

boogie |>
  mutate(album_release_date = ymd(album_release_date)) |>
  ggplot(aes(
    x = album_release_date,
    y = duration_ms,
    group = album_release_date
  )) +
  geom_boxplot() +
  geom_jitter(alpha = 0.5, width = 0.3, height = 0) +
  theme_minimal() +
  labs(
    x = "Album release date",
    y = "Duration of song (ms)"
  )
```

```
PTP <- as_tibble(PTP)

PTP |>
  mutate(album_release_date = ymd(album_release_date)) |>
  ggplot(aes(
    x = album_release_date,
    y = duration_ms,
    group = album_release_date
  )) +
  geom_boxplot() +
  geom_jitter(alpha = 0.5, width = 0.3, height = 0) +
  theme_minimal() +
  labs(
    x = "Album release date",
    y = "Duration of song (ms)"
  )
```
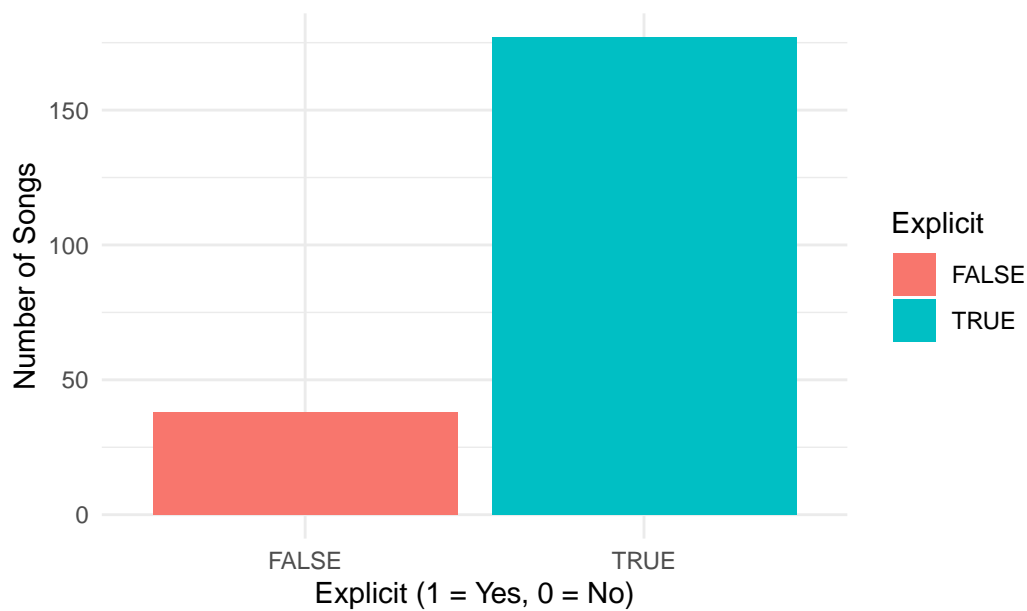
```
explicit_summary <- boogie %>%
  group_by(explicit) %>%
  summarise(count = n())

ggplot(explicit_summary, aes(x = as.factor(explicit), y = count, fill = as.factor(explicit)))
  geom_bar(stat = "identity") +
  labs(x = "Explicit (1 = Yes, 0 = No)", y = "Number of Songs",
       title = "Number of Explicit vs Non-Explicit Songs for boogie",
       fill = "Explicit") +
  theme_minimal()
```

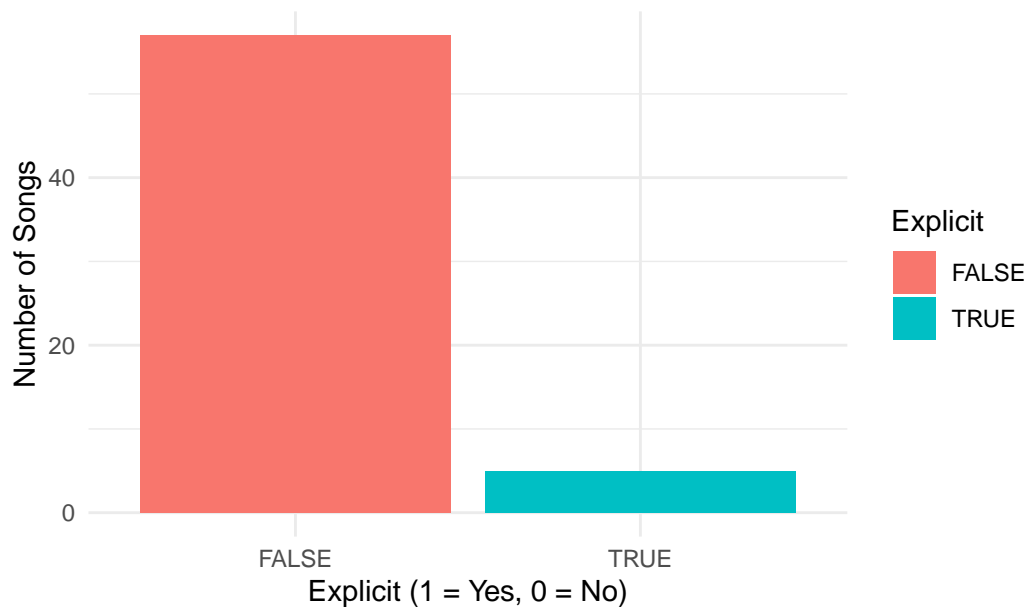## Number of Explicit vs Non−Explicit Songs for boogie


```r
explicit_summary <- PTP %>%
  group_by(explicit) %>%
  summarise(count = n())

ggplot(explicit_summary, aes(x = as.factor(explicit), y = count, fill = as.factor(explicit))
  geom_bar(stat = "identity") +
  labs(x = "Explicit (1 = Yes, 0 = No)", y = "Number of Songs",
       title = "Number of Explicit vs Non-Explicit Songs for pay my money to pain",
       fill = "Explicit") +
  theme_minimal()
```

Number of Explicit vs Non–Explicit Songs for pay my money to


```r
rbind(boogie, PTP, radiohead) |>
  select(artist_name, album_release_date, loudness) |>
  mutate(album_release_date = ymd(album_release_date)) |>
  ggplot(aes( x = album_release_date, y = loudness, color = artist_name)) +
  geom_point(alpha = 0.3) +
  geom_smooth() +
  theme_minimal() +
  facet_wrap(facets = vars(artist_name), dir = "v") +
  labs(
    x = "Album release date",
    y = "loudness",
    color = "Artist"
  ) +
  scale_color_brewer(palette = "Set1") +
  theme(legend.position = "bottom")
```

```
Warning: There was 1 warning in `mutate()`.
i In argument: `album_release_date = ymd(album_release_date)`.
Caused by warning:
!  8 failed to parse.
```

```
`geom_smooth()` using method = 'loess' and formula = 'y ~ x'
```

```
Warning: Removed 8 rows containing non-finite outside the scale range
(`stat_smooth()`).
```

```
Warning: Removed 8 rows containing missing values or values outside the scale range
(`geom_point()`).
```