Lab Report 3

1710704

21/10/2019

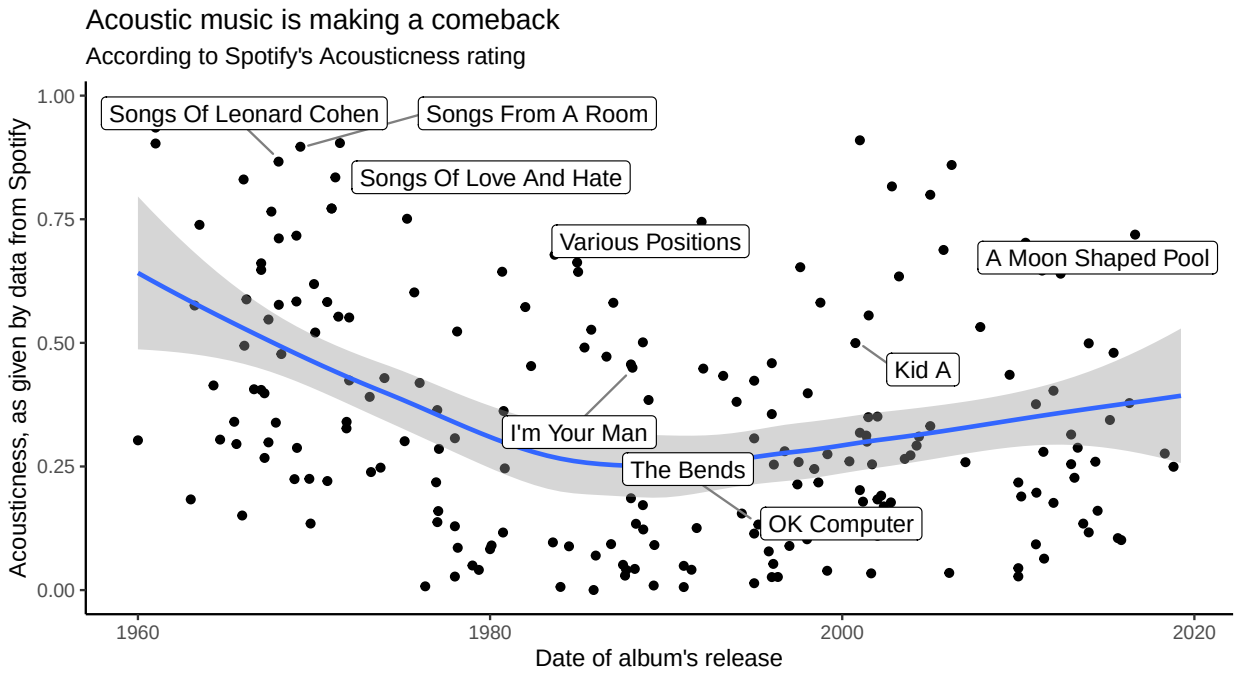
```
Spotify <- read_excel("edited_spotify.xlsx")
```

```
MyData = {Spotify %>%
    group_by(Artist,AlbumName,AlbumReleaseDate) %>%
    summarise(AlbumEnergy = mean(TrackEnergy),
              AlbumValence = mean(TrackValence),
              AlbumAcousticness = mean(TrackAcousticness) )}

MyData$AlbumReleaseDate = parse_date_time(MyData$AlbumReleaseDate,
                                          orders=c("y","ym","ymd"))

MyData <- mutate(MyData, Decade = (year(AlbumReleaseDate) - year(AlbumReleaseDate)%%10))
```

```
scatterPlot <- {MyData %>% ggplot(aes(MyData$AlbumReleaseDate,
                                      MyData$AlbumAcousticness))} +
                                    geom_point() + geom_smooth() +
                                  xlab("Date of album's release") +
              ylab("Acousticness, as given by data from Spotify") +
                     ggtitle("Acoustic music is making a comeback",
                             subtitle = "According to Spotify's Acousticness rating") +
  geom_label_repel(aes(label=ifelse(((Artist=="Leonard Cohen")| (Artist=="Radiohead")),
                                    as.character(AlbumName),"")),
                           box.padding   = 0.5, point.padding = 0.5,
                           segment.color = 'grey50') + theme_classic()
print(scatterPlot)
```

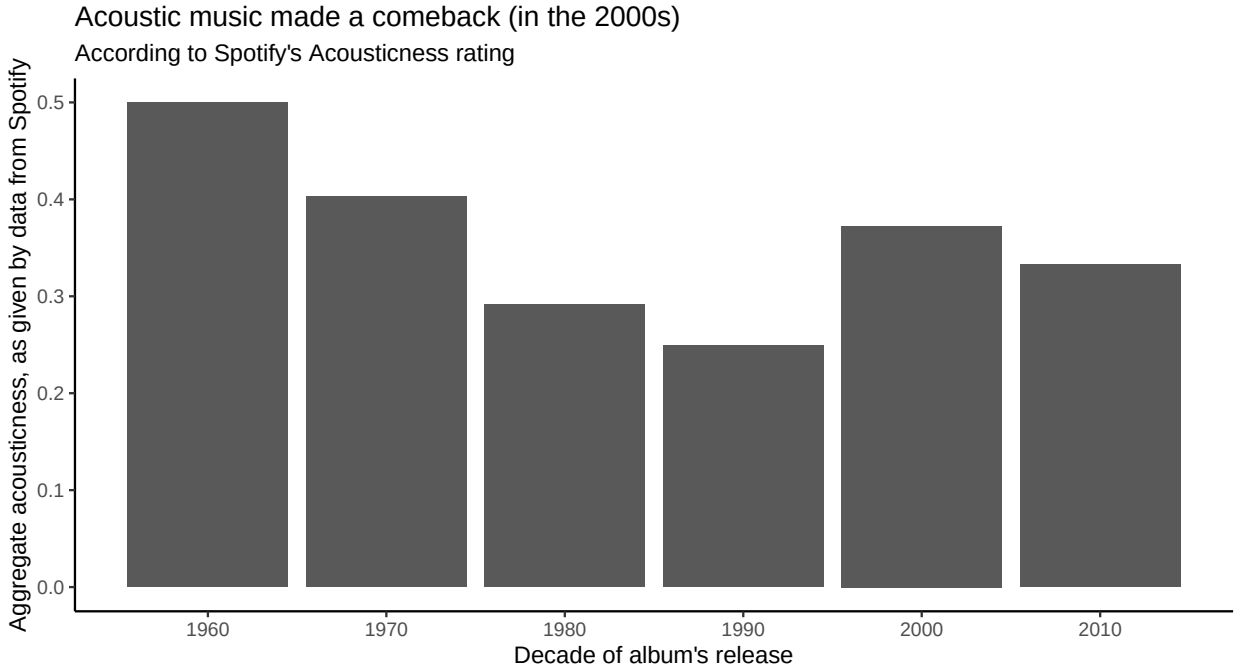


This first graph shows the relationship between an album's release date and its *acousticness*, defined by the Spotify Web API Guidance as " A confidence measure from 0.0 to 1.0 of whether the track is acoustic"; essentially this shows how the way albums are recorded has changed since 1960 (note that Wikipedia defines acoustic music as"music that solely or primarily uses instruments that produce sound through acoustic means, as opposed to electric or electronic means"). *Acousticness* is not so much a continuous scale of how much production a song has been through, but rather a rating of how confident Spotify is that a track is acoustic (with a rating close to 1.0 meaning there is a high likelihood that the given track is acoustic). For example, Leonard Cohen's *Songs From A Room* (1969) has an acousticness rating of 0.90, meaning Spotify is confident that it is acoustic, whereas his album *I'm Your Man* (1988) has an acousticness rating of 0.45 meaning Spotify has little confidence that the album is acoustic.

The graph shows a significant drop in the average acousticness of albums in our sample from 1960 to ~1990, and then a moderate rise from then until the most recent release (Billie Eilish, *WHEN WE ALL FALL ASLEEP, WHERE DO WE GO?* (2019)). No album has an acousticness rating of below 0.1 until 1976 (Ramones, *Ramones* (1976), with a rating of 0.008) and the album with the highest rating is Vashti Bunyan's *Just Another Diamond Day* (1970), with a rating 0.98. Between 1980 and 2001, no album has an acousticness rating higher than 0.72 (the highest being *Promenade* (1994) by The Divine Comedy, with a rating of 0.72).

It is possible that this dip in acousticness can be related to the rise of electronic music as a popular genre; the MIDI was invented in 1980, and digital synthesisers were largely popularised in the early 1980s. Electronic Dance Music became a popular genre in the 80s and 90s and had a large influence on pop music at the time, however (as we can see from the line on the plot) this influence has diminished since then.

```
aggregateData <- aggregate(MyData[,6], list(MyData$Decade), mean)
barPlot <- {aggregateData %>% ggplot(aes(aggregateData$Group.1,
                              aggregateData$AlbumAcousticness))} + geom_col() +
                                          xlab("Decade of album's release") +
              ylab("Aggregate acousticness, as given by data from Spotify") +
                     ggtitle("Acoustic music made a comeback (in the 2000s)",
                   subtitle = "According to Spotify's Acousticness rating") +
                  scale_x_continuous(breaks = scales::pretty_breaks(n = 6)) +
                     scale_fill_brewer(palette = "Blues") + theme_classic()
print(barPlot)
```



The plot above breaks down the average acousticness of albums in the sample by decade and shows that the pattern in the scatterplot may be misleading; in fact the average acousticness of albums from our sample released this decade is lower than that of albums released in the previous decade.

```
print(aggregateData)
```

```
##   Group.1 AlbumAcousticness
## 1    1960         0.4997527
## 2    1970         0.4032682
## 3    1980         0.2912585
## 4    1990         0.2493342
## 5    2000         0.3725584
## 6    2010         0.3324272
```

It is possible that acoustic music may be on the decline once again, or that grouping of the albums into decades may produce spurious patterns, or simply that this decade is experiencing a minor blip. Since there are only 211 albums analysed in this sample, each decade has an average of $\frac{211}{6} \approx 35$ albums to aggregate; therefore there is every possibility that the patterns shown in this second graph are are not representative of the entire population of albums. However, the pattern for the first 5 decades matches that of scatterplot, so it seems likely that it is simply the arbitrary grouping of albums by decade that produces this disparity.