

Tables

Caleb Brooks
5/10/2018

```
library(tidyverse)
```

```
## -- Attaching packages ----- tidyverse 1.2.1 --
## √ ggplot2 2.2.1      √ purrr  0.2.4
## √ tibble  1.4.1      √ dplyr  0.7.4
## √ tidyr   0.7.2      √ stringr 1.2.0
## √ readr   1.1.1      √ forcats 0.2.0

## -- Conflicts ----- tidyverse_conflicts() --
## x dplyr::filter() masks stats::filter()
## x dplyr::lag()    masks stats::lag()
```

```
library(devtools)
library(sm)
```

```
## Warning in fun(libname, pkgname): couldn't connect to display ":0"
## Package 'sm', version 2.2-5.5: type help(sm) for summary information
```

```
library(rgl)
```

```
## Warning in rgl.init(initValue, onlyNULL): RGL: unable to open X11 display
## Warning: 'rgl_init' failed, running with rgl.useNULL = TRUE
```

```
library(httr)
library(xtable)
install_github("calebmbrooks/Rspotify")
```

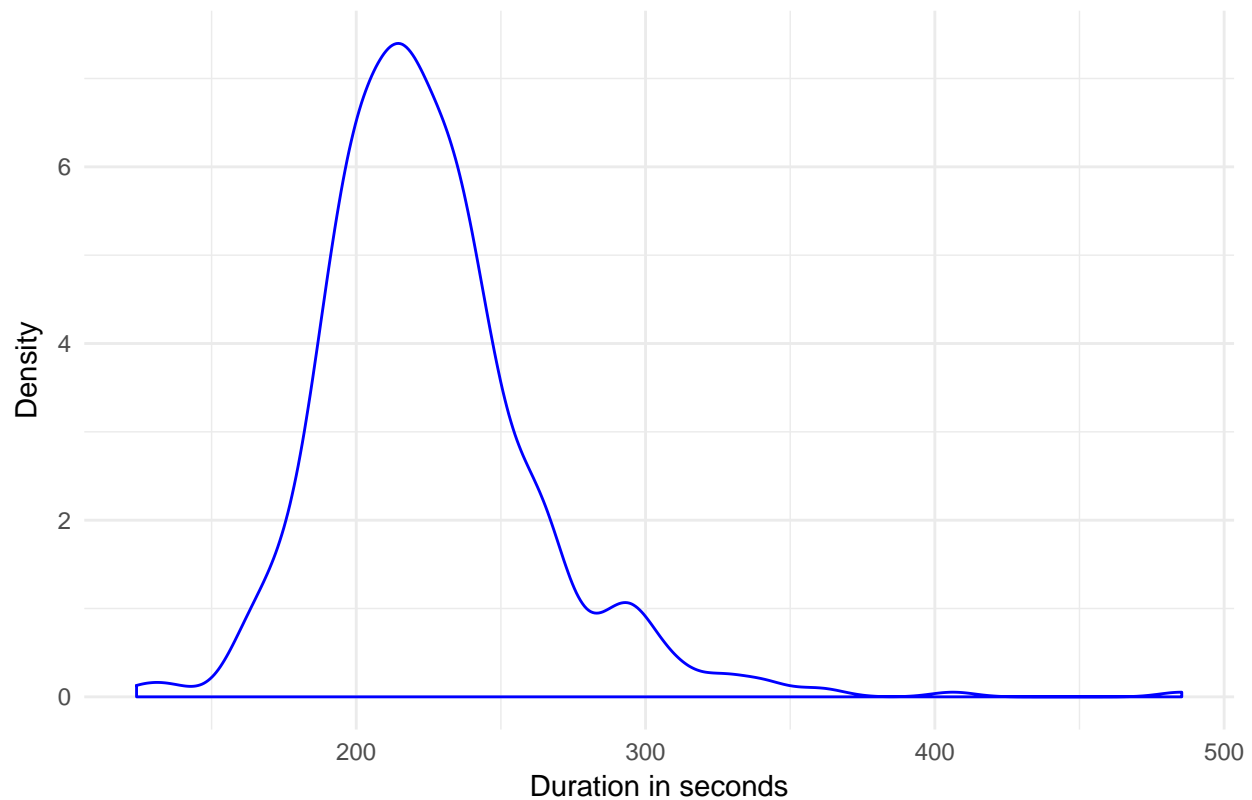
```
## Skipping install of 'Rspotify' from a github remote, the SHA1 (f23e1e0a) has not changed since last instal
## Use `force = TRUE` to force installation
```

```
library(Rspotify)
```

```
topAllTime <- readRDS("data/topAllTime.rds")
probs <- c(0.1, 0.25, 0.5, 0.75, 0.9)
quantiles <- quantile(topAllTime$duration_ms, prob=probs)
```

```
topAllTime %>%
  mutate(duration_s = duration_ms/1000) %>%
  ggplot(aes(x = duration_s))+
  geom_density(color = "Blue", aes(y = ..count..))+
  labs(title = "Distribution of song length", x = "Duration in seconds", y = "Density")+
  theme_minimal()
```

Distribution of song length



```
t1 <- topAllTime %>%
  group_by(artist) %>%
  summarise(numberOfSongs = n(),
            meanRank = mean(rank),
            sumReverseRank = sum(rank_desc),
            sumReverseRankSq = sum(rank_desc_sq)) %>%
  arrange(desc(sumReverseRankSq)) %>%
  head(n = 20)

topAllTime %>%
  group_by(artist) %>%
  summarise(avg_pos = (mean(rank)/(n()^2))) %>%
  arrange(avg_pos)
```

```
## # A tibble: 256 x 2
##   artist      avg_pos
##   <chr>      <dbl>
## 1 Calvin Harris 1.25
## 2 Ed Sheeran    1.29
## 3 Maroon 5      1.86
## 4 Drake         2.03
## 5 Coldplay      2.12
## 6 The Weeknd    2.53
## 7 The Chainsmokers 2.61
## 8 Ariana Grande 2.86
## 9 David Guetta  3.41
## 10 Luis Fonsi   3.62
```

```
## # ... with 246 more rows
```

```
t2 <- topAllTime %>%
  group_by(factor(time_signature)) %>%
  summarise(numTracks = n())
```

```
topAllTime %>%
  filter(time_signature == 5)
```

```
##           tracks           id popularity      artist
## 1   Chandelier 2s1sdSqGcKxpPr5lC17jAV      71      Sia
## 2   The Hills 25khomWgBVamSdKw7hzm3l      76  The Weeknd
## 3   Locked Away 7x5xYW5W420GPAdHUyyguy      71    R. City
## 4      Gold 6ZS07kPn8IMJFymyticbJO      74    Kiiara
## 5 Safe and Sound 1fzJyTCKeZuTSLByCsLRHl      13 Capital Cities
##           artist_full      artistId
## 1              Sia 5WU1DfRSoLAfcVSX1WnrxN
## 2      The Weeknd 1Xyo4u8uXC1ZmMpatF05PJ
## 3 R. City feat. Adam Levine 4TH4BHy0LdBi3dpBW4P2UX
## 4              Kiiara 4u5smJBskI6Adzv08PuiUP
## 5      Capital Cities 4gwpMTbLWtBUl0ijbVpuu
##           album      albumId danceability
## 1 1000 Forms Of Fear (Deluxe Version) 6FdNvo05sF4EKwCX9je1MH      0.399
## 2      Beauty Behind The Madness 28ZKQMoNBBOetKXZ97G2SN      0.356
## 3      What Dreams Are Made Of 4AUTMNSP56xicSznzKgzb      0.509
## 4      low kii savage 2EvTmiJyF3H4BoxNCL4HLh      0.600
## 5      In A Tidal Wave Of Mystery 0oGhveAbsimFjFVF5aIkjW      0.655
## energy key loudness mode speechiness acousticness instrumentalness
## 1 0.787 1 -2.880 1 0.0499 0.019700 6.07e-05
## 2 0.567 0 -7.011 0 0.0798 0.086100 0.00e+00
## 3 0.671 1 -5.709 1 0.0678 0.304000 0.00e+00
## 4 0.412 8 -9.343 1 0.3440 0.615000 2.50e-03
## 5 0.819 0 -4.852 1 0.0316 0.000176 3.74e-03
## liveness valence tempo duration_ms time_signature
## 1 0.0685 0.572 117.089 216120 5
## 2 0.1370 0.109 135.553 242253 5
## 3 0.0452 0.550 118.413 227480 5
## 4 0.1340 0.408 113.049 225882 5
## 5 0.1040 0.766 117.956 192790 5
##           uri
## 1 spotify:track:2s1sdSqGcKxpPr5lC17jAV
## 2 spotify:track:25khomWgBVamSdKw7hzm3l
## 3 spotify:track:7x5xYW5W420GPAdHUyyguy
## 4 spotify:track:6ZS07kPn8IMJFymyticbJO
## 5 spotify:track:1fzJyTCKeZuTSLByCsLRHl
##           analysis_url rank
## 1 https://api.spotify.com/v1/audio-analysis/2s1sdSqGcKxpPr5lC17jAV 53
## 2 https://api.spotify.com/v1/audio-analysis/25khomWgBVamSdKw7hzm3l 54
## 3 https://api.spotify.com/v1/audio-analysis/7x5xYW5W420GPAdHUyyguy 134
## 4 https://api.spotify.com/v1/audio-analysis/6ZS07kPn8IMJFymyticbJO 179
## 5 https://api.spotify.com/v1/audio-analysis/1fzJyTCKeZuTSLByCsLRHl 195
## rank_desc rank_desc_sq
## 1 503 253009
## 2 502 252004
## 3 422 178084
```

```
## 4      377      142129
## 5      361      130321
```

```
topAllTime %>%
  filter(time_signature == 3)
```

```
##                                     tracks
## 1                                     Perfect
## 2 Earned It (Fifty Shades Of Grey) - From The "Fifty Shades Of Grey" Soundtrack
## 3                                     Dangerous Woman
## 4                                     Like I'm Gonna Lose You
## 5                                     In The Night
## 6                                     Say Something
## 7                                     Dive
##
##           id popularity           artist
## 1 0tgVpDi06FyKpA1z0VMD4v          92      Ed Sheeran
## 2 2PIvq1pGrUjY007X5y1UpM          75      The Weeknd
## 3 7l94dyN2hX9c6wWcZQu0GJ          72      Ariana Grande
## 4 2YlZnw2ikdb837oKMKjBkW          81      Meghan Trainor
## 5 25KybV9BOUlvcnv7nN3Pyo          69      The Weeknd
## 6 6Vc5wAMmXdKIAM7WUoEb7N          78 A Great Big World
## 7 5lChrwmUPDJvedPQnIU8Ls          82      Ed Sheeran
##
##           artist_full           artistId
## 1      Ed Sheeran 6eUKZXaKkcviH0Ku9w2n3V
## 2      The Weeknd 1Xyo4u8uXC1ZmMpatF05PJ
## 3      Ariana Grande 66CXWjxzNUsdJxJ2JdwvnR
## 4      Meghan Trainor feat. John Legend 6JL8zeS1Nmi0ftqZTRgdTz
## 5      The Weeknd 1Xyo4u8uXC1ZmMpatF05PJ
## 6 A Great Big World feat. Christina Aguilera 5xKp3UyavIBUsGy3DQdXeF
## 7      Ed Sheeran 6eUKZXaKkcviH0Ku9w2n3V
##
##                                     album
## 1                                     ÷ (Deluxe)
## 2 Earned It (Fifty Shades Of Grey) [From The "Fifty Shades Of Grey" Soundtrack]
## 3                                     Dangerous Woman
## 4                                     Title (Deluxe)
## 5                                     Beauty Behind The Madness
## 6                                     Is There Anybody Out There?
## 7                                     ÷ (Deluxe)
##
##           albumId danceability energy key loudness mode speechiness
## 1 3T4tUhGYeRNVUGevb0wThu          0.599 0.448 8 -6.312 1 0.0232
## 2 6lBa3txRZWfiX6ZTEZlFCV          0.664 0.371 2 -5.959 0 0.0306
## 3 3OZgEyW4krCZ814pTJWr7          0.672 0.605 7 -5.353 1 0.0378
## 4 5W98Ab4VvQEufEE4Tie5fE          0.630 0.530 0 -7.259 1 0.0434
## 5 0P3oVJBFOv3TDXlYRhGL7s          0.317 0.683 7 -4.989 1 0.1340
## 6 1yOcLa4euMk9sV7rRJ89Dl          0.407 0.147 2 -8.822 1 0.0355
## 7 3T4tUhGYeRNVUGevb0wThu          0.761 0.386 4 -6.158 1 0.0399
##
##           acousticness instrumentalness liveness valence tempo duration_ms
## 1      0.1630          0.00e+00 0.1060 0.1680 95.050 263400
## 2      0.4040          0.00e+00 0.0971 0.4290 119.950 252227
## 3      0.0499          0.00e+00 0.3570 0.2970 134.015 235947
## 4      0.4000          0.00e+00 0.1770 0.4170 108.038 225053
## 5      0.0794          0.00e+00 0.0465 0.5380 167.615 235653
## 6      0.8570          2.89e-06 0.0913 0.0765 141.284 229400
## 7      0.3550          0.00e+00 0.0953 0.5260 134.943 238440
##           time_signature           uri
```

```

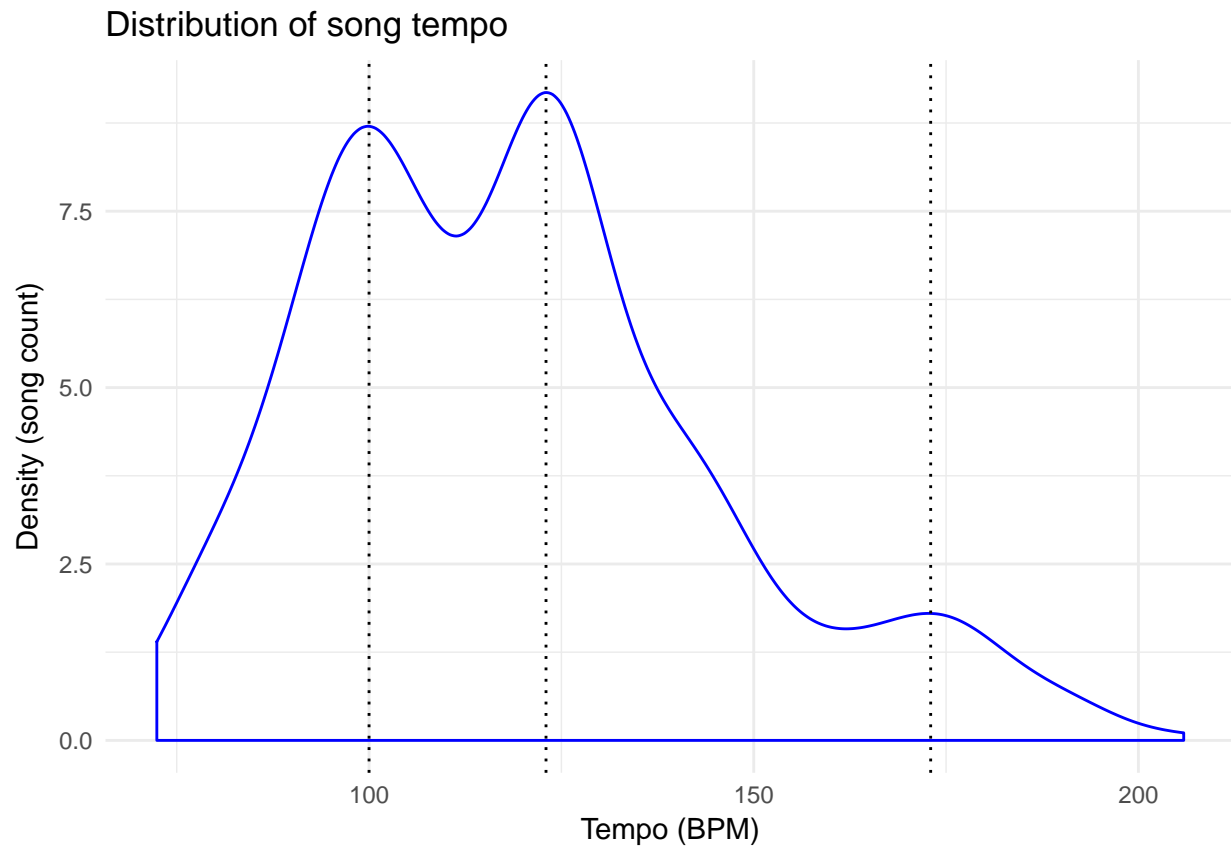
## 1          3 spotify:track:0tgVpDi06FyKpA1z0VMD4v
## 2          3 spotify:track:2PIvq1pGrUjY007X5y1UpM
## 3          3 spotify:track:7l94dyN2hX9c6wWcZQu0GJ
## 4          3 spotify:track:2YlZnw2ikdb837oKMKjBkW
## 5          3 spotify:track:25KybV9BOUlvcnv7nN3Pyo
## 6          3 spotify:track:6Vc5wAMmXdKIAM7WUoEb7N
## 7          3 spotify:track:5lChrwmUPDJvedPQnIU8Ls
##
##                                     analysis_url rank
## 1 https://api.spotify.com/v1/audio-analysis/0tgVpDi06FyKpA1z0VMD4v 85
## 2 https://api.spotify.com/v1/audio-analysis/2PIvq1pGrUjY007X5y1UpM 98
## 3 https://api.spotify.com/v1/audio-analysis/7l94dyN2hX9c6wWcZQu0GJ 214
## 4 https://api.spotify.com/v1/audio-analysis/2YlZnw2ikdb837oKMKjBkW 255
## 5 https://api.spotify.com/v1/audio-analysis/25KybV9BOUlvcnv7nN3Pyo 323
## 6 https://api.spotify.com/v1/audio-analysis/6Vc5wAMmXdKIAM7WUoEb7N 357
## 7 https://api.spotify.com/v1/audio-analysis/5lChrwmUPDJvedPQnIU8Ls 514
## rank_desc rank_desc_sq
## 1      471      221841
## 2      458      209764
## 3      342      116964
## 4      301       90601
## 5      233       54289
## 6      199       39601
## 7       42       1764

```

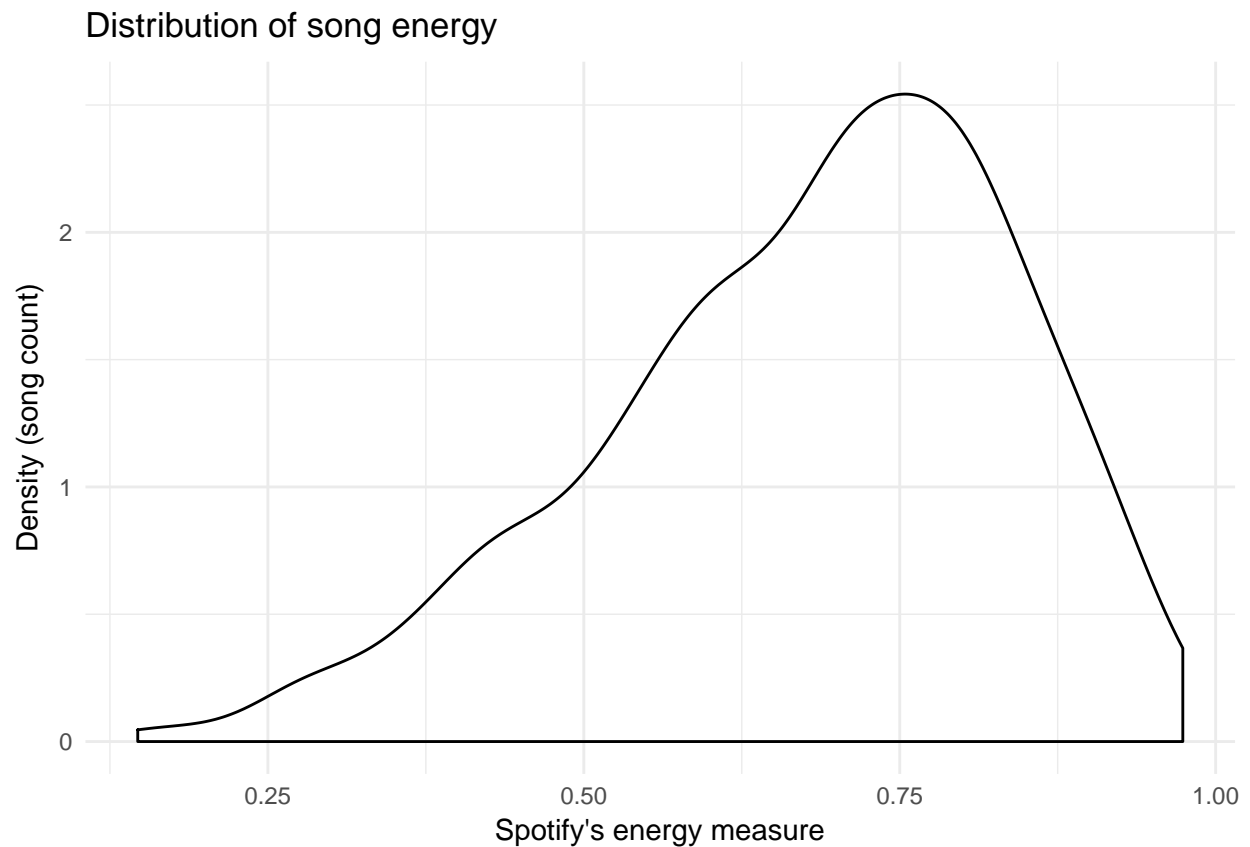
	artist	numberOfSongs	meanRank	sumReverseRank	sumReverseRankSq
1	Ed Sheeran	12	185.33	4448.00	1989768.00
2	Calvin Harris	14	245.21	4351.00	1629855.00
3	Drake	10	203.30	3527.00	1516005.00
4	The Weeknd	9	204.89	3160.00	1410352.00
5	The Chainsmokers	6	93.83	2773.00	1368143.00
6	Coldplay	10	212.50	3435.00	1220307.00
7	Maroon 5	12	268.00	3456.00	1209464.00
8	Ariana Grande	8	182.88	2985.00	1185313.00
9	Imagine Dragons	6	132.83	2539.00	1128117.00
10	Justin Bieber	6	193.67	2174.00	1010570.00
11	David Guetta	8	218.50	2700.00	963558.00
12	Bruno Mars	7	209.57	2425.00	949505.00
13	Rihanna	5	126.40	2148.00	940604.00
14	Major Lazer	4	111.00	1780.00	892882.00
15	Kygo	6	226.67	1976.00	861468.00
16	Twenty One Pilots	3	40.67	1546.00	797718.00
17	Jason Derulo	5	166.20	1949.00	775843.00
18	Sia	4	140.75	1661.00	755311.00
19	Avicii	5	179.60	1882.00	736886.00
20	DJ Snake	4	139.00	1668.00	728138.00

	factor(time_signature)	numTracks
1	1	1
2	3	7
3	4	542
4	5	5

```
topAllTime %>%
  ggplot(aes(x = tempo))+
  geom_density(color = "Blue", aes(y = ..count..))+
  labs(title = "Distribution of song tempo", x = "Tempo (BPM)", y = "Density (song count)")+
  theme_minimal()+
  geom_vline(xintercept = c(100, 123, 173), linetype = "dotted")
```



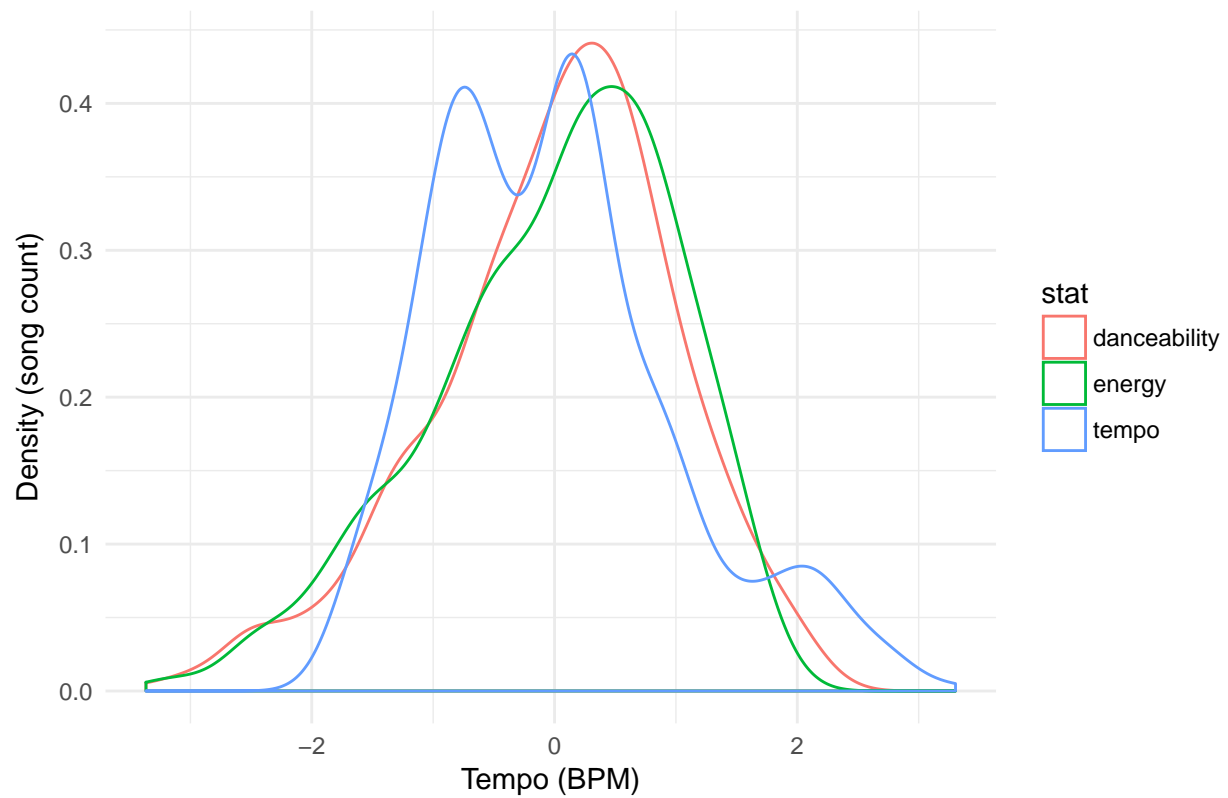
```
topAllTime %>%
  ggplot(aes(x = energy))+
  geom_density()+
  labs(title = "Distribution of song energy", x = "Spotify's energy measure", y = "Density (song count)")+
  theme_minimal()
```



```
topAllTime %>%  
  select(tracks, danceability, energy, tempo) %>%  
  mutate(danceability = scale(danceability), energy = scale(energy), tempo = scale(tempo)) %>%  
  gather(-tracks, key = "stat", value = "value") %>%  
  ggplot(aes(x = value, color = stat))+  
  geom_density()+  
  labs(title = "Distribution of song tempo", x = "Tempo (BPM)", y = "Density (song count)")+  
  theme_minimal()
```

```
## Warning: attributes are not identical across measure variables;  
## they will be dropped
```

Distribution of song tempo



```
topAllTime %>%  
  ggplot(aes(x = tempo, y = danceability))+  
  geom_point(aes(color= rank))+  
  geom_smooth(method = "loess")
```