

# Eksploratorna analiza Spotify skupa podataka

... Ivan Josip Kardum, 36557629 Damjan Crnković

2026-01-29

```
data <- read.csv("spotify_songs.csv", sep = ",", fileEncoding = "ISO-8859-1")

glimpse(data)

## Rows: 32,833
## Columns: 23
## $ track_id
## $ track_name
## $ track_artist
## $ track_popularity
## $ track_album_id
## $ track_album_name
## $ track_album_release_date
## $ playl
<chr> "6f807x0ima9a1j3VPbc7VN", "0r7CVbZTWZgbTCYdfa~  
<chr> "I Don't Care (with Justin Bieber) - Loud Lux~  
<chr> "Ed Sheeran", "Maroon 5", "Zara Larsson", "Th~  
<int> 66, 67, 70, 60, 69, 67, 62, 69, 68, 67, 58, 6~  
<chr> "2oCs0DGTsR098Gh5ZS12Cx", "63rPS0264uRjW1X5E6~  
<chr> "I Don't Care (with Justin Bieber) [Loud Luxu~  
<chr> "2019-06-14", "2019-12-13", "2019-07-05", "20~  
<chr> "Pop Remix", "Pop Remix", "Pop Remix", "Pop R~  
<chr> "37i9dQZF1DXcZDD7cfEKhW", "37i9dQZF1DXcZDD7cf~  
<chr> "pop", "pop", "pop", "pop", "pop", "po~  
<chr> "dance pop", "dance pop", "dance pop", "dance~  
<dbl> 0.748, 0.726, 0.675, 0.718, 0.650, 0.675, 0.4~  
<dbl> 0.916, 0.815, 0.931, 0.930, 0.833, 0.919, 0.8~  
<int> 6, 11, 1, 7, 1, 8, 5, 4, 8, 2, 6, 8, 1, 5, 5, ~  
<dbl> -2.634, -4.969, -3.432, -3.778, -4.672, -5.38~  
<int> 1, 1, 0, 1, 1, 0, 0, 1, 1, 1, 1, 0, 0, ~  
<dbl> 0.0583, 0.0373, 0.0742, 0.1020, 0.0359, 0.127~  
<dbl> 0.10200, 0.07240, 0.07940, 0.02870, 0.08030, ~  
<dbl> 0.00e+00, 4.21e-03, 2.33e-05, 9.43e-06, 0.00e~  
<dbl> 0.0653, 0.3570, 0.1100, 0.2040, 0.0833, 0.143~  
<dbl> 0.518, 0.693, 0.613, 0.277, 0.725, 0.585, 0.1~  
<dbl> 122.036, 99.972, 124.008, 121.956, 123.976, 1~  
<int> 194754, 162600, 176616, 169093, 189052, 16304~  
  
head(data, 10)  
  
##          track_id          track_name
## 1 6f807x0ima9a1j3VPbc7VN I Don't Care (with Justin Bieber) - Loud Luxury Remix
## 2 0r7CVbZTWZgbTCYdfa2P31           Memories - Dillon Francis Remix
## 3 1z1Hg7Vb0AhHDiEmnDE791      All the Time - Don Diablo Remix
## 4 75FpbthrwQmzH1BJLuGdC7    Call You Mine - Keanu Silva Remix
## 5 1e8PAfcKUYoKkxPhrHqw4x   Someone You Loved - Future Humans Remix
## 6 7fvUMiyapMsRRxr07cU8Ef Beautiful People (feat. Khalid) - Jack Wins Remix
## 7 20Ay1PUDDfwRGfe01YqlCQ        Never Really Over - R3HAB Remix
## 8 6b1RNvAcJJqH73eZ04BLAB Post Malone (feat. RANI) - GATTĀ\u009cSO Remix
```

```

## 9 7bF6tC03gFb8INrEDcjNT5          Tough Love - TiÃ«sto Remix / Radio Edit
## 10 1IXGILkPm0tOCNeq00kCPa        If I Can't Have You - Gryffin Remix
##   track_artist track_popularity      track_album_id
## 1     Ed Sheeran           66 2oCs0DGTsR098Gh5ZS12Cx
## 2     Maroon 5            67 63rPS0264uRjW1X5E6cWv6
## 3     Zara Larsson         70 1HoSmj2eLcsrR0vE9gThr4
## 4 The Chainsmokers       60 1nqYs0ef1yKKuGOVchbsk6
## 5     Lewis Capaldi        69 7m7vv9wlQ4i0LFuJiE2zsQ
## 6     Ed Sheeran           67 2yiy9cd2QktrNvWC2EUi0k
## 7     Katy Perry            62 7INHYSeusaFlyrHSNxm8qH
## 8     Sam Feldt             69 6703SRPsLkS4bPtMFFJes1
## 9     Avicii                68 7CvAfGvq4R1IwEbT9o8Iav
## 10    Shawn Mendes          67 4QxzbffSsVryEQwvPFEV5Iu
##   track_album_name
## 1 I Don't Care (with Justin Bieber) [Loud Luxury Remix]
## 2                               Memories (Dillon Francis Remix)
## 3                               All the Time (Don Diablo Remix)
## 4                               Call You Mine - The Remixes
## 5 Someone You Loved (Future Humans Remix)
## 6 Beautiful People (feat. Khalid) [Jack Wins Remix]
## 7                               Never Really Over (R3HAB Remix)
## 8 Post Malone (feat. RANI) [GATTÃ\u009cSO Remix]
## 9                               Tough Love (TiÃ«sto Remix)
## 10    If I Can't Have You (Gryffin Remix)
##   track_album_release_date playlist_name      playlist_id playlist_genre
## 1     2019-06-14    Pop Remix 37i9dQZF1DXcZDD7cfEKhW      pop
## 2     2019-12-13    Pop Remix 37i9dQZF1DXcZDD7cfEKhW      pop
## 3     2019-07-05    Pop Remix 37i9dQZF1DXcZDD7cfEKhW      pop
## 4     2019-07-19    Pop Remix 37i9dQZF1DXcZDD7cfEKhW      pop
## 5     2019-03-05    Pop Remix 37i9dQZF1DXcZDD7cfEKhW      pop
## 6     2019-07-11    Pop Remix 37i9dQZF1DXcZDD7cfEKhW      pop
## 7     2019-07-26    Pop Remix 37i9dQZF1DXcZDD7cfEKhW      pop
## 8     2019-08-29    Pop Remix 37i9dQZF1DXcZDD7cfEKhW      pop
## 9     2019-06-14    Pop Remix 37i9dQZF1DXcZDD7cfEKhW      pop
## 10    2019-06-20    Pop Remix 37i9dQZF1DXcZDD7cfEKhW      pop
##   playlist_subgenre danceability energy key loudness mode speechiness
## 1     dance pop      0.748  0.916   6   -2.634    1   0.0583
## 2     dance pop      0.726  0.815   11  -4.969    1   0.0373
## 3     dance pop      0.675  0.931   1   -3.432    0   0.0742
## 4     dance pop      0.718  0.930   7   -3.778    1   0.1020
## 5     dance pop      0.650  0.833   1   -4.672    1   0.0359
## 6     dance pop      0.675  0.919   8   -5.385    1   0.1270
## 7     dance pop      0.449  0.856   5   -4.788    0   0.0623
## 8     dance pop      0.542  0.903   4   -2.419    0   0.0434
## 9     dance pop      0.594  0.935   8   -3.562    1   0.0565
## 10    dance pop      0.642  0.818   2   -4.552    1   0.0320
##   acousticness instrumentalness liveness valence tempo duration_ms
## 1     0.1020      0.00e+00  0.0653  0.518 122.036   194754
## 2     0.0724      4.21e-03  0.3570  0.693 99.972   162600
## 3     0.0794      2.33e-05  0.1100  0.613 124.008   176616
## 4     0.0287      9.43e-06  0.2040  0.277 121.956   169093
## 5     0.0803      0.00e+00  0.0833  0.725 123.976   189052
## 6     0.0799      0.00e+00  0.1430  0.585 124.982   163049
## 7     0.1870      0.00e+00  0.1760  0.152 112.648   187675

```

```
## 8      0.0335    4.83e-06   0.1110   0.367 127.936    207619
## 9      0.0249    3.97e-06   0.6370   0.366 127.015    193187
## 10     0.0567    0.00e+00   0.0919   0.590 124.957    253040
```

## uredivanje podataka

```
data$track_album_release_date <- as.Date(data$track_album_release_date) # mozda posix
factcols <- c("playlist_genre", "playlist_subgenre", "playlist_name", "track_artist", "mode")
data[factcols] <- lapply(data[factcols], as.factor)
```

safsdffgg