

Trabalho 01

Arthur Linhares Madureira
Victor Henrique Silva Ribeiro

29 de setembro de 2023

1 Tarefa 1

Na primeira tarefa, inicialmente foram extraídas informações acerca da duração das músicas no dataset (duração mínima, média e máxima). No entanto, essas medidas são muito sensíveis à presença de outliers. Para contornar esse empecilho, as músicas de duração muito fora do padrão foram removidas utilizando a metodologia IQRR. Com isso, o valor da média foi consideravelmente reduzido, enquanto o valor mínimo e valor máximo tiveram uma brusca alteração.

Amostra	Mínimo	Média	Máximo
Todas as Músicas	0	234,408 seg	10.435,467 seg
Músicas Não Outliers	107,582 seg	226,899 seg	349,583 seg

Tabela 1: Estatísticas sobre duração das músicas.

2 Tarefa 2

Na segunda tarefa, buscou-se os cinco artistas com maiores aparições em playlists. Esses artistas foram: Drake, Rihanna, Kanye West, The Weeknd, Kendrick Lamar. O artista mais relevante, utilizando essa métrica, foi Drake, que aparece em cerca de 20000 playlists. É interessante observar que, para todos os artistas analisados, o número de playlists aumentou de forma muito rápida no período que compreende 2011-2017.

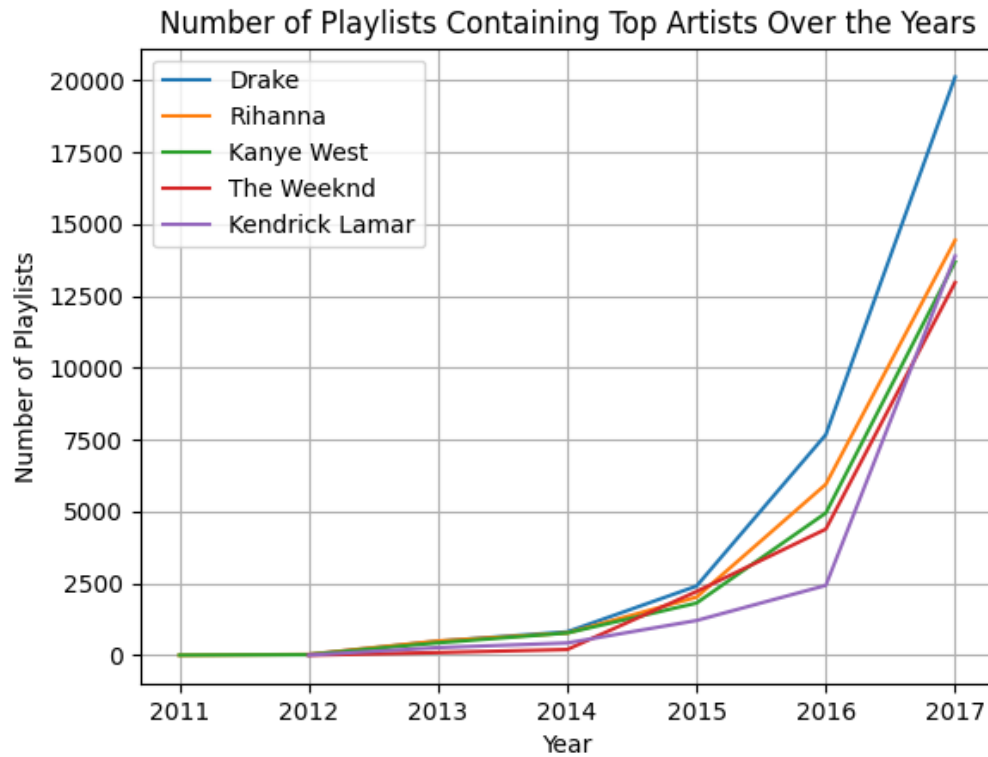


Figura 1: Artistas com maiores aparições em playlists

3 Tarefa 3

O ponto no gráfico de Função de Distribuição Cumulativa (CDF) correspondente, aproximadamente, a $x = 0,2$ possui $y = 0,8$. Ou seja, cerca de 80% das playlists demonstram uma prevalência de artistas igual ou inferior a 0,2. Esse padrão sugere que a maioria das playlists exibe uma baixa prevalência de artistas, indicando que múltiplos artistas contribuem com uma fração modesta do total de faixas em cada playlist. Essa observação revela uma notável diversidade na seleção de artistas, onde cada artista contribui com uma parcela diminuta das faixas em uma playlist típica. Este cenário sugere uma apreciação por uma ampla variedade de artistas e pode indicar uma preferência por uma experiência musical mais diversificada entre os usuários.

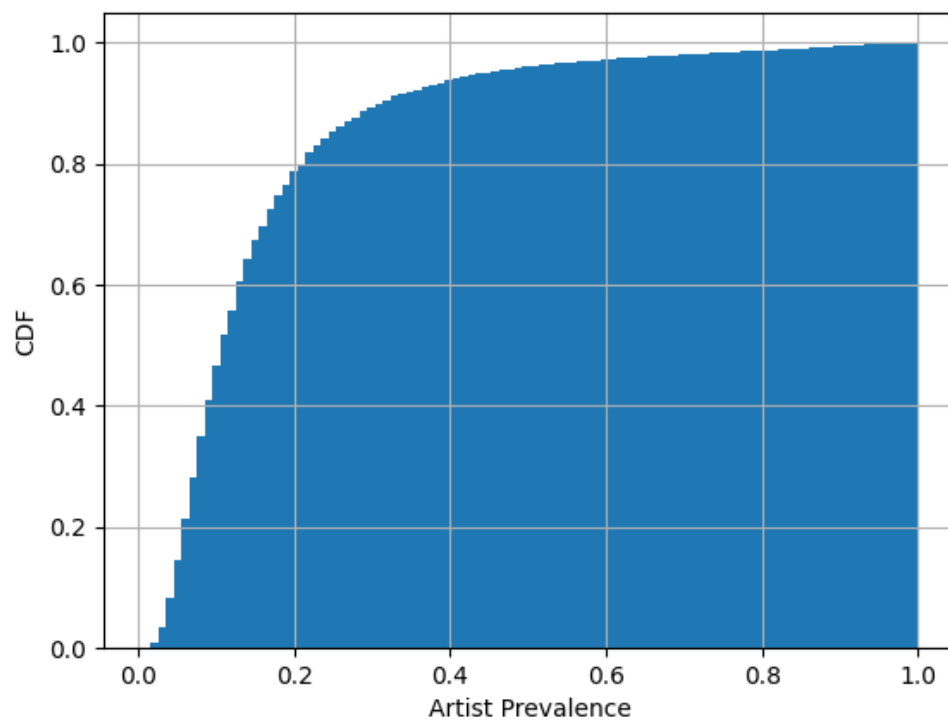


Figura 2: Função de Distribuição Cumulativa (CDF).