



Carlos Cuevas Covarrubias
Centro de Investigación en Estadística y Matemáticas Aplicadas
ccuevas@anahuac.mx



Enrique Lemus Rodríguez
Centro de Investigación en Estadística y Matemáticas Aplicadas
elemus@anahuac.mx



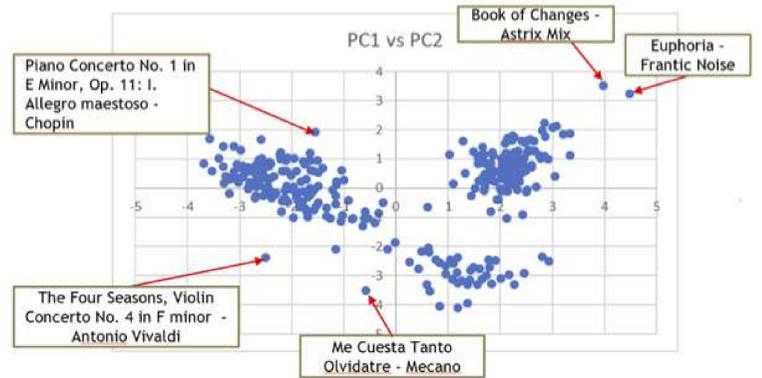
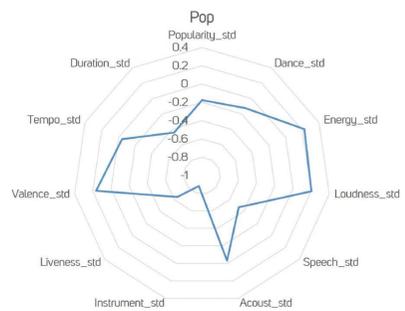
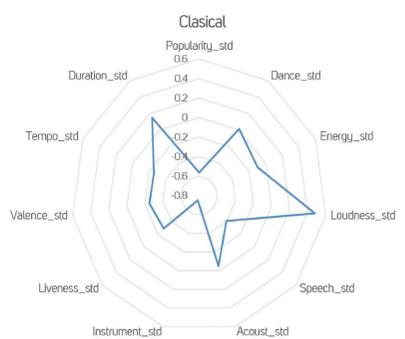
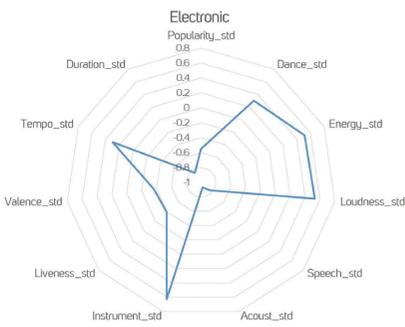
Lucía Straface
Centro de Investigación en Estadística y Matemáticas Aplicadas

INTRODUCCIÓN

We have a data base with Audio Features for 315 tracks, that belong to different musical genres:

- 122 Electronic Dance Music (Astrix, Frantic Noise, Bio Babas...)
- 148 classical music (Beethoven, Mozart, Chopin...)
- 45 pop music (Mecano, La Oreja de Van Gogh...)

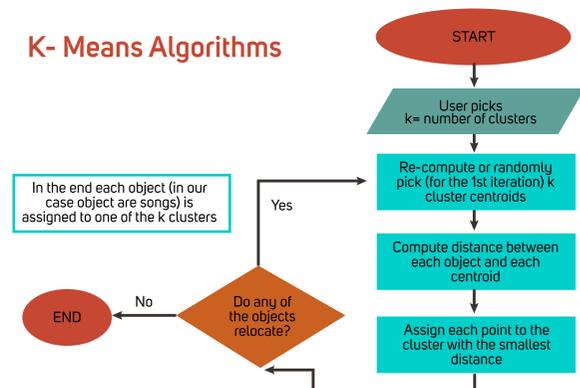
For each track, we have 11 variables: Popularity, Duration, Energy, Speechiness, Instrumentalness, Liveness, Acousticness Loudness, Tempo, Valence and Danceability.



There seem to be 3 groups, which happen to coincide with the 3 musical genres. Cluster Analysis refers to the task of grouping objects in such way that objects in the same group are similar to each other in terms of certain features, and somehow different to objects in other groups.

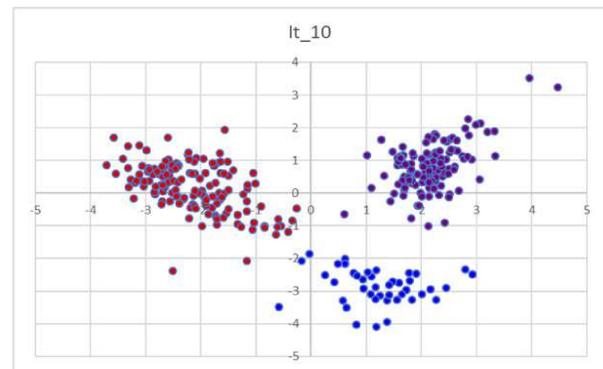
K-means Clustering aims to partition a data set into K clusters, with each observation belonging to the cluster with the nearest centroid.

K- Means Algorithms



RESULTADOS

The algorithm gives us the following classification:



DISCUSIÓN

Almost all the tracks were classified according to their genre. Only 3 tracks out of 315 were classified differently: Otro muerto, Eugenio Salvador Dalí, and Aire. All 3 tracks are form the same artist: Mecano.

We were able to form 3 playlists with 315 an it only took the algorithm a few seconds to do it, it would have taken us days to listen to all the tracks ourselves!

REFERENCIAS

1. Everitt BS, et al. Cluster Analysis. USA: Wiley; 2011.
2. Xu R, et al. Clustering. USA: Wiley; 2009.
3. De Dios S.J. Discovering similarities across my Spotify music using data, clustering and visualization. [online] Towards Data Science. 2017. Disponible en: <https://towardsdatascience.com/discovering-similarities-across-my-spotify-music-using-data-clustering-and-visualization-52b58e6f547b>