



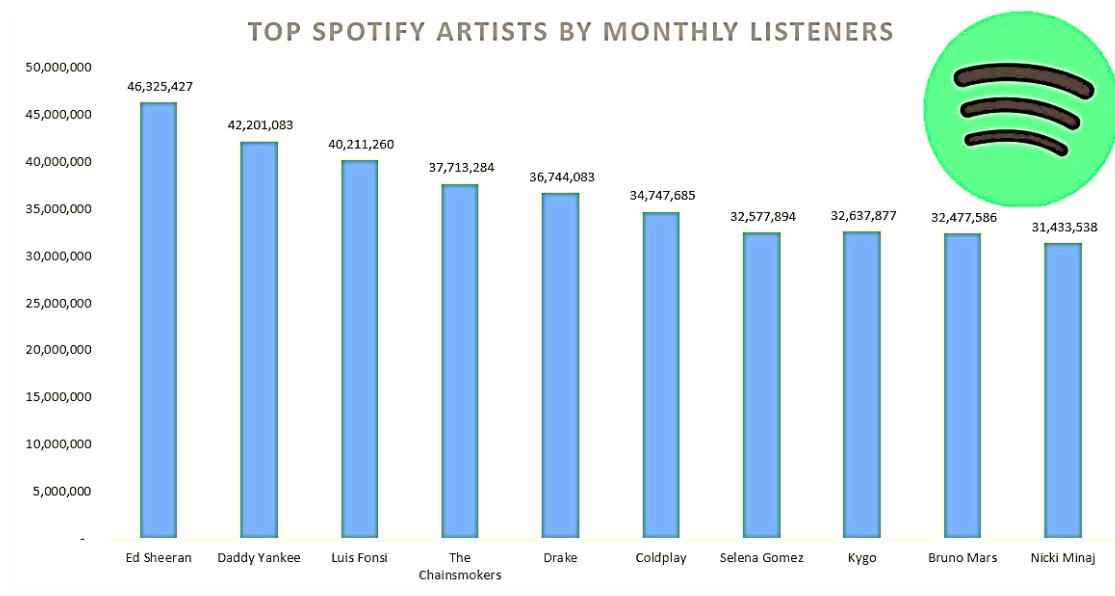
WHAT CAN SONGS TELL US ABOUT PEOPLE?

FÉLIX CAMACHO CRIADO

FEBRUARY 2019

ÍNDICE

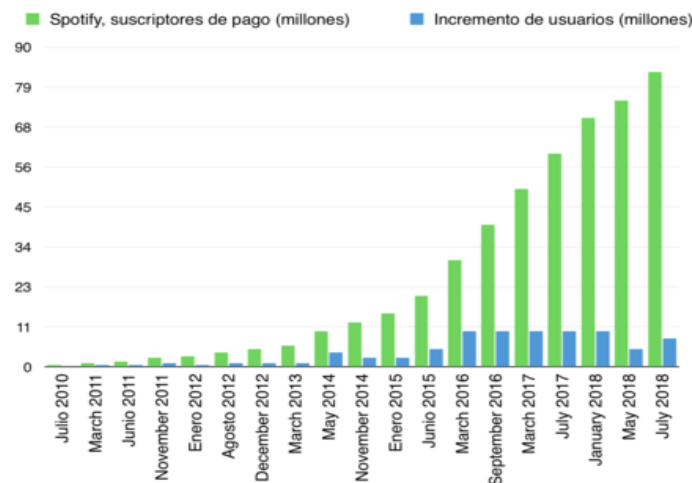
- ▶ INTRODUCCIÓN
- ▶ OBTENCIÓN DE INFORMACIÓN
- ▶ DESCRIPCIÓN DE LOS DATOS
- ▶ ALCANCE
- ▶ CONCLUSIONES



INTRODUCCIÓN

De igual manera que otras redes sociales, los servicios de música en streaming obtienen grandes cantidades de información de sus usuarios tales como sus hábitos musicales, desde donde están escuchando música o desde que dispositivo utilizan el servicio.

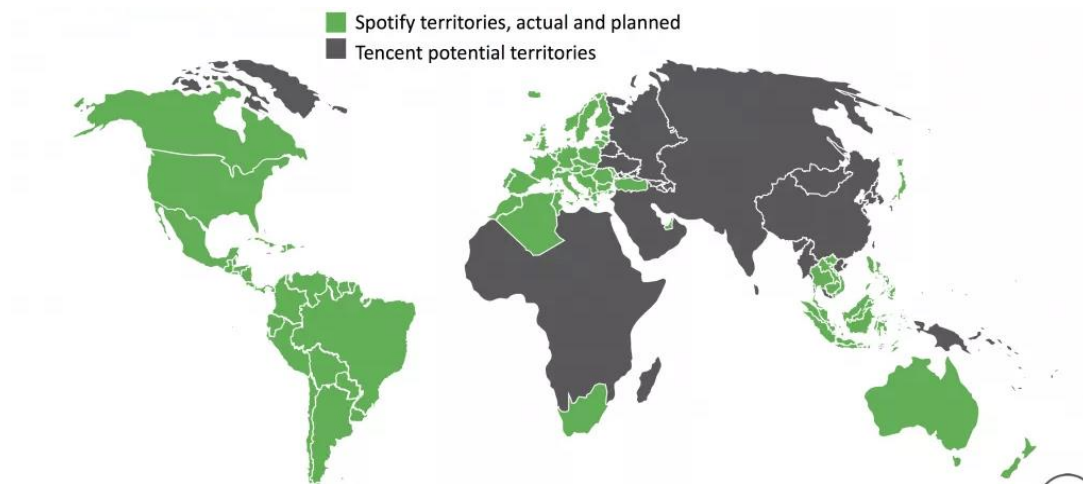
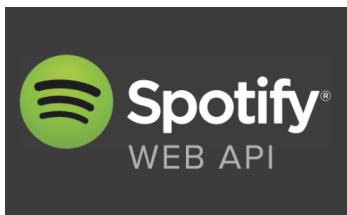
- ▶ Spotify es una aplicación multiplataforma empleada para la reproducción de música vía streaming.
- ▶ 180 millones de usuarios activos mensuales. 83 millones de usuarios de pago. 2Q 2018.
- ▶ Crecimiento positivo en el número de usuarios de pago durante los últimos años.
- ▶ Mayor clúster de Hadoop en Europa. 7000 empleados
- ▶ Objetivo: Obtener información relevante en función de las playlist más escuchadas en diferentes países.
- ▶ En lo personal, tengo especial interés en la música, utilizo diariamente la aplicación Spotify de escritorio y de móvil.



OBTENCIÓN DE INFORMACIÓN

Spotify ha desarrollado diversas funcionalidades para permitir que terceras personas tengan acceso a la biblioteca musical desde las aplicaciones que desarrollen. Para el desarrollo del proyecto se considerarán las siguientes recursos:

- ▶ Spotify for Developers. Acceso a API de Spotify. Funcionamiento a través de token.
- ▶ Spotify Charts. Ficheros CSV diarios con información del TOP 200 y TOP 50 VIRAL.
- ▶ Indicadores socio-demográficos, meteorológicos por país.
- ▶ Librerías personalizadas.
 - ▶ Spotifyr(R).
 - ▶ Spotipy (Python).



- ▶ Ambas librerías disponen de funciones definidas para obtener información de playlist públicas, albums así como características de las canciones.

DESCRIPCIÓN DE LOS DATOS

Spotify ha definido un sistema según el cual clasifica sus canciones en 9 atributos a los que asigna un valor entre 0.0 y 1.0 según sus características:

- ▶ **Acousticness (Acústica):** Una medida de confianza entre 0.0 y 1.0 de cuan acústica es una canción.
- ▶ **Danceability (Bailabilidad):** Una medida entre 0.0 y 1.0 de cuan apta es una canción para bailar.
- ▶ **Energy (Energía):** Representa la intensidad y actividad. Las variables que intervienen en este atributo incluyen el rango dinámico, volumen percibido, timbre y la entropía general de la canción.
- ▶ **Instrumentalness (Instrumentalidad):** Una medida de confianza entre 0.0 y 1.0 de si una canción incluye voz
- ▶ **Key (Clave):** Establece la clave en la que está la canción asignándole un número entero que representa su altura según la notación musical. Ej: 0=C, 1=C#/Db etc.
- ▶ **Liveness (En vivo):** Una medida entre 0.0 y 1.0 de si la canción fue grabada en vivo detectando la presencia de público.
- ▶ **Loudness (Sonoridad):** Determina el volumen medio de la canción en decibelios (dB).
- ▶ **Mode (Modo):** Indica el modo (mayor o menor) de la canción, es decir, el tipo de escala de la que deriva su contenido melódico. Un 1 representa una escala mayor y un 0 una menor.
- ▶ **Speechiness (Habla):** Establece intervalos entre 0.0 y 1.0 de proporción de la presencia de palabras habladas y música.
- ▶ **Tempo:** Indica el tempo medio de la canción en golpes por minuto (BPM).
- ▶ **Valence (Valencia):** Una medida entre 0.0 y 1.0 que describe cuan positiva es una canción. Una canción con una valencia alta será, en la mayoría de los casos, alegre animada o eufórica. En cambio, una canción con una valencia baja es más proclive a transmitir tristeza, depresión o ira.

ALCANCE

El alcance del proyecto aún no está definido, no obstante quisiera compartir algunas de las ideas que he pensado y las cuales en función de la viabilidad y de la información relevante que obtenga desarrollaré en menor o mayor medida

- ▶ Estudiar la correlación entre los atributos que componen las canciones
- ▶ Elaborar ranking de países en función de los atributos de sus canciones
- ▶ Basarse en el histórico de TOP playlist europeas para establecer que características tiene que tener una canción para entrar en el TOP de Europa.
- ▶ Investigar la relación del índice de felicidad de un país con los atributos de las canciones más escuchadas.
- ▶ Examinar el tiempo meteorológico de un país y su relación con los atributos de las canciones más escuchadas.
- ▶ Utilizar técnicas de clustering en playlist TOP 200 por país.
- ▶ Desarrollar mapas continentales con Shiny/Geoplot u otra librería de visualización de R o Python en función de los atributos de las canciones

CONCLUSIONES

A partir de los datos obtenidos de las playlist públicas compartidas por los usuarios de cada país y su combinación con datos externos pretendo dar respuestas aproximadas a preguntas como por ejemplo.

- ▶ ¿En los países con más horas de sol se escucha música considerada más positiva?
- ▶ ¿En los países con más precipitaciones anuales se escucha música considerada más deprimente?
- ▶ ¿Existe alguna relación entre los índices de felicidad de un país y que las canciones más escuchadas sean aquellas con el atributo Valence más alto?





GRACIAS POR SU
ATENCIÓN