

Proyecto DataViz - EMAS 2018

18 de junio de 2018

Proyecto de visualización de datos con ggplot2

Durante estos días estás aprendiendo aspectos teóricos y prácticos de la visualización de datos. En el proyecto se trata de que, a partir de una base de datos, plantees una pregunta de investigación. Esa pregunta deberá ser respondida utilizando gráficos. En el ámbito de las Ciencias Sociales estamos acostumbrados a que los datos provengan de encuestas, experimentos o indicadores socioeconómicos, sin embargo, cada vez existen más datos disponibles en la red, que aunque no han sido generados con el objeto de servir en la investigación, pueden ser utilizados para ese fin. En este ejercicio te propongo que utilices datos de Spotify.

Spotify, como otras redes sociales (*p. ej.* Twitter o Facebook), cuenta con una API que permite obtener datos de la aplicación. Una API corresponde con las siglas de *Application Programming Interface*, y se trata de una ventana a partir de la cual se pueden extraer datos de la plataforma. Por ejemplo, a través de la API de Spotify se puede descargar el listado completo de canciones de un artista con una serie de indicadores. Estos indicadores van desde el nombre de la canción o el álbum en el que se ubica, hasta una serie de medidas generadas por Spotify como *valence*, que se refiere al nivel de positividad que transmite la canción.

En estas instrucciones encontrarás como acceder y descargar los datos, posteriormente tendrás que plantear una sencilla pregunta de investigación y, a partir de los datos, dar respuesta a la misma. Recuerda que el miércoles podrás hacer una presentación corta (3 minutos) de tu gráfico al final de clase.

1. Ábrete una cuenta de desarrollador de Spotify

- Para ello necesitarás tener una cuenta en Spotify, que puedes abrirla en: <https://www.spotify.com/es/> haciendo click en **Registrarse**.
- Si ya tiene una cuenta en Spotify entra en la web <https://developer.spotify.com/dashboard/> y haz click en **LOG IN** e inicia sesión como si fuera tu cuenta habitual de Spotify.

2. Crear una aplicación

- El siguiente paso consiste en crear una aplicación, la cual generará las claves que te permitan descargar datos de Spotify desde R.
- Al crear una aplicación te pedirá datos como el nombre, la descripción y el tipo de aplicación. En nuestro caso, en nombre puede poner algo como “EMAS DataViz 2018”, en descripción puedes insertar un breve comentario sobre que vas a hacer “Análisis de datos de las canciones de tal o cual artista”. Por último, en qué estás creando, marca la casilla de **Desktop App** y dale a **NEXT**.
- En la siguiente pantalla te preguntas si vas a crear una aplicación comercial de algún tipo, que no es nuestro caso, por lo que tienes que marcar el **NO**.
- En la siguiente pantalla, marca todas las casillas y haz click en **SUBMIT**.
- Por último, en la pantalla que ha aparecido podrás ver **SHOW CLIENT SECRET**, así que haz click sobre esa palabra. A continuación tendrás dos números en pantalla, el **Client ID** y el **Client Secret**. No cierres la ventana, porque vas a necesitar esos dos números, que son los que te dan acceso a los datos de Spotify.

3. Descargar datos de R a través de Spotify

- Para las operaciones de descargar los datos vamos a utilizar un paquete que se llama *spotifyr*. Para utilizarlo necesitarás instalarlo primero, para después cargarlo:

```
install.packages("spotifyr")
```

```
library(spotifyr)
```

- En segundo lugar vas a crear dos variables de sistema, una para el **Client ID**, y otra para el **Client Secret**. Estas variables se almacenarán de forma permanente, y R las usará cada vez que le pidas acceder a los datos de spotify. Para crear las variables del sistema (recuerda que las claves tienen que ir entre comillas):

```
Sys.setenv(SPOTIFY_CLIENT_ID = "Copia aquí tu Client ID")
```

```
Sys.setenv(SPOTIFY_CLIENT_SECRET = "Copia aquí tu Client Secret")
```

- Ahora vamos a crear un objeto que nos de acceso a los datos de Spotify, para ello escribe y ejecuta la siguiente línea de código:

```
access_token <- get_spotify_access_token()
```

- Ya podemos descargar datos de Spotify, como por ejemplo los datos de las canciones de Camarón De La Isla. En este caso, la descarga tiene forma de data frame, y lo almaceno como objeto *camaron*. Una vez que hayas ejecutado el código con el artista de tu elección, te pedirá que vuelvas a teclear el nombre del artista en la consola (no en el script), y pulses **enter**.

```
camaron <- get_artist_audio_features('Cameron De La Isla')
```

- Una vez descargados los datos, verás que cada canción del artista tiene una serie de indicadores. Puedes encontrar una lista completa, con su descripción en: <https://developer.spotify.com/documentation/web-api/reference/tracks/get-audio-features/>
- Recuerda, que aparte de esta función, que sirve para descargar las todas las canciones del artista, el paquete *spotifyr* tiene otras como descargar las playlists de un usuario (*get_user_playlists*). Para ver todas las posibilidades del paquete puedes entrar en la documentación del mismo: <https://cran.r-project.org/web/packages/spotifyr/spotifyr.pdf>

4. Instrucciones del proyecto

Ahora puedes descargar unos datos que te parezcan interesantes, y generar una pregunta de investigación sencilla. Posteriormente, genera uno o varios gráficos para responder a la misma.