

## **DAT - Data Science**

## Experiencia

IMPORTANTE: Algunos comentaron que "les hubiera gustado ver más código en la de IA", y también que "sería bueno entender un poco más lo que se está haciendo en las actividades de código".

Hemos tomado nota, y por eso esta experiencia ha sido aprender cómo afrontar un tutorial de un tema nuevo paso a paso, buscando información sobre las piezas que nos faltan. Y de paso aprender conceptos cómo *clustering* o *k-means*.

Por favor, los que no pudisteis atender a la clase, mirad la grabación que está en la CARPETA. No os limitéis a hacer simplemente copy+paste sin entender (ese era precisamente el objetivo de la experiencia).

> Has quedado fascinado con Moodplay, el recomendador interactivo de música basado en los estados de ánimo, que usa Data Science y Machine Learning, y que desarrolló el profesor Denis Parra y su equipo.

Así que decides seguir un tutorial que hace una cosa parecida: detectar estados de ánimo en la música que escuchas, agrupar las canciones, y crear listas de Spotify para cada estado de ánimo.



## Competencias

#### **Especificas:**

 Conocimientos básicos de data science, tales como conceptos (clustering), algoritmos (k-means), y herramientas (scikit-learn)

#### Transversales:

 Seguir un tutorial de un tema desconocido de forma autónoma y buscar proactivamente la información necesaria.

#### **BKNs**:

• ¡Aprendí a hacer mi propia lista de **Spotify** basado en estados de ánimo!



### Material

El tutorial original que seguiremos en clase es Building Mood-Based Spotify Playlists using K-Means Clustering.

Sin embargo, para el entregable usaremos un Google Colab con una versión adaptada del tutorial que hemos preparado para que puedas hacerlo con nuestros datos, sin necesidad de tener Spotify. El Google Colab que está en CANVAS.

Deberás copiar el Google Colab para poder modificarlo! Para esto abre el Colab y ve a Archivo > Guardar una copia en Drive.



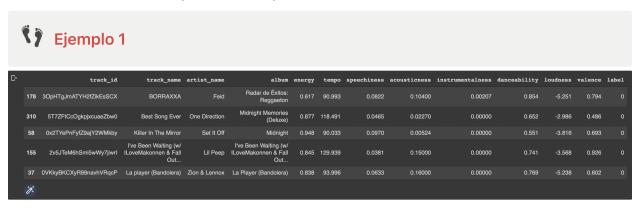
Problemas?: Si tienes dudas, pregunta en DISCORD. Igualmente, si no tienes computador o estás muy perdido, recuerda buscar en clase o usar DISCORD para encontrar otra persona que si pueda/sepa con la que trabajar juntos.





Para entregar: Link en CANVAS

Siguiendo el tutorial del Colab Adaptado, se te pide entregar una **captura de pantalla** con un ejemplo de canciones de alguna de las playlists que generarías (no hace falta los 4 clusters, con 1 basta).





<u>Aa</u> ID	# %	<b>■</b> Descripción
<u>R1</u>	25%	Se genera la tabla de de La captura muestra un canciones
<u>R2</u>	50%	Las canciones están o En la tabla solo hay car (mismo tipo)
R3	25%	Se muestran los atribu Se ven los atributos en accoustiness, entre otr

**SUM 100%** 



#### ■ Default view

<u>Aa</u> ID	# %	■ Descripción
<u>R1</u>	25%	Se genera la tabla de la captura muestra un canciones
<u>R2</u>	50%	Las canciones están o En la tabla solo hay car (mismo cluster)
<u>R3</u>	25%	Se muestran los atribu Se ven los atributos en accoustiness, entre otr

sum 100%







<u>Aa</u> ID	# %	<b>■</b> Descripción
<u>R1</u>	25%	La tabla muestra las ca playlist
R2	50%	Hay canciones del labe mezcladas.
R3	25%	Se ven los atributos en accoustiness, entre otr

sum 100%

# **No It Yourself (DIY)**



Para evaluar: Recuerda para que te evalúen los DIY y te den la nota bonus correspondiente, debes compartirlos en DISCORD siguiendo las instrucciones en CANVAS. ¡Así motivarás a los demás y podrás buscar sinergias!

### ■ Default view

<u>Aa</u> Num	<b>∷</b> Nivel	<b>■</b> DYI
1	Casual	Hacer el tutorial pero u
2	Pro	Haz el último paso del Spotify con librerías de
3	Casual	Investiga e interactúa s
4	Pro	Usa alguna de las libre proyectos (e.g., Plotly,
<u>5</u>	Pro	Avanza en algún curso Science Specialization
6	Pro	Realiza otros tutoriales Data Science), investig paso, como hicimos en
7	Casual	Piensa en datos de tu v (e.g., notas, mensajes películas IMDB,) y pi (predicción, clasificació interesante hacer.
8	Pro	Aprende a hacer web s páginas de internet que
9	Casual	Mira videos para apren concretos de Data Min o KNN.
10	Casual	Investiga sobre que co sub-grupo <u>HAIVis</u> (Hur es parte del IALAB.
···		y cualquier otra idea

# Exploración



### TIC2433 - Minería de Datos

Programa en la CARPETA

- ¿Cómo puedo sacar conclusiones si tengo muchísimos datos?
- ¿Qué técnicas de Clustering (K-means), Clasificación (e.g., KNN), y Reducción de la dimensionalidad (e.g., PCA)?
- ¿Como puedo obtener datos directamente de páginas web (e.g., web scrapping)?

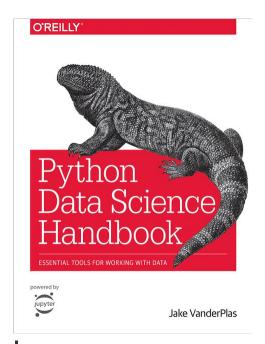


#### IIC3633 - Sistemas Recomendadores

Programa en la CARPETA

- ¿Cómo se pueden crear un sistema que recomiende?
- ¿Qué algoritmos de recomendación existen y cómo puedo diseñar mis propias técnicas?
- ¿Qué particulares hay para recomendar en diferentes dominios como películas, música, imágenes, o redes sociales?

Bibliografía



Python Data Science Handbook Jacob T. Vanderplas O'Reilly (2017) Link Bibliotecas UC