

Tarea Grande 1

Profesor Denis Parra

Anunciada: 26 de marzo de 2018

Indicaciones

- Fecha de Entrega: 9 de abril de 2018.
 - Debes entregar tu tarea en tu repositorio GitHub privado asignado para esta evaluación.
 - Cada hora o fracción de atraso descuenta 0,5 puntos de la nota que obtengas.
 - La tarea es *individual*. La copia será evaluada con nota 1 en el la tarea, además de las sanciones disciplinarias correspondientes.
-

Objetivo

El objetivo de esta tarea es que aprendas a:

- Utilizar Python para manipular datos.
- Crear una página web sencilla que contemple el uso de HTML, CSS y Javascript, para confeccionar un *dashboard*.

En particular, tendrás que manipular datos de los hábitos musicales de un usuario real de [LastFM](#), para utilizarlos en la confección de un *dashboard* que muestre información interesante sobre la persona estudiada.

Descripción de los datos

Junto a este enunciado, se te entregaron dos archivos más:

1. `dataset_final.csv`: contiene todas las canciones que fueron escuchadas por un usuario desde enero de 2017 hasta el 21 de marzo de 2018. Los campos que contiene el archivo son:

- a) Date: fecha en que se escuchó la canción, en formato *Unix Epoch Time*.
 - b) Track: título de la canción.
 - c) Artist: intérprete de la canción.
2. `artists_genres_images.csv`: archivo que contiene el género musical y la imagen de portada de todos los artistas presentes en el archivo anterior. Sus campos son:
- a) Artist: artista.
 - b) Genre: género musical.
 - c) Image_Link: enlace a la imagen de portada del artista.

Instrucciones

Parte 1: Manipulación de los datos y generación de gráficos con Python (4.5 puntos)

En la primera parte de tu tarea, trabajarás con los datos entregados para obtener lo necesario para la confección del *dashboard*. **El entregable de esta sección es un *jupyter notebook* (extensión `.ipynb`).** De no cumplir con el tipo de archivo, recibirás una penalización en tu nota. Además, se recomienda fuertemente que utilices `pandas` y `matplotlib` para realizar lo siguiente:

1. Convertir las fechas a un formato más legible (YYYY-mm-DD), como por ejemplo 2018-03-21 23:03:17 **(0.2 puntos)**.
2. Confeccionar un gráfico de barras donde en el eje x se vean los nombres de las 10 canciones más escuchadas y en el y la cantidad de veces que fue escuchada cada una **(0.5 puntos)**.
3. Obtener los 5 artistas más escuchados y disponer de esta información en un DataFrame cuyas columnas sean: nombre del artista, veces que fue escuchado y enlace a su imagen¹. Posteriormente, guardar esta información en un archivo `.csv`, generado con Python **(0.3 puntos)**.
4. Generar un gráfico de torta de los géneros musicales más escuchados por el usuario. Deben agregar los nombres y valores porcentuales a cada género en el gráfico. **(0.5 puntos)**.
5. Graficar la evolución de la cantidad de canciones escuchadas por mes **(1 punto)**. [Ejemplo](#).
6. Realizar un gráfico de torta donde se evidencie la cantidad de canciones escuchadas según día de la semana. Se debe poder identificar cuantas canciones escucha y qué día las escucha. **(1 punto)**.

¹Para este ítem, es conveniente que investigues sobre el método `merge` de `pandas`.

7. Generar un gráfico de barras en el que se vea cuántas canciones se escucharon según hora del día² (**1 punto**). [Ejemplo](#).

Para poder utilizar los gráficos en la Parte 2, estos deben ser exportados a un archivo de imagen.

Parte 2: *Dashboard* (1.5 puntos)

Con los datos y gráficos ya obtenidos, debes construir una página web que contenga lo siguiente:

1. Gráfico de las 10 canciones más escuchadas.
2. Tabla con los 5 artistas más escuchados y sus respectivas imágenes. Esta información debe ser extraída del archivo `.csv` generado en la parte anterior.
3. Gráfico de los géneros musicales más escuchados.
4. Gráfico de la evolución mensual de la cantidad de canciones escuchadas.
5. Gráfico de la cantidad de canciones escuchadas según día de la semana.
6. Gráfico de la cantidad de canciones escuchadas según hora del día.

Como la confección de los gráficos será evaluada en la parte 1, la distribución de puntaje de esta parte es:

1. La página es *responsive* y los elementos que la componen están correctamente alineados y no se sobreponen unos con otros (**0.75 puntos**).
2. Uso de JavaScript para cargar los datos `.csv` generado en la parte 1 y así construir la tabla (**0.75 puntos**).

Formato de entrega

Para evaluar con facilidad la parte 1 de esta tarea, deberás realizarla en un *jupyter notebook* y subir este archivo de extensión `.ipynb` a tu repositorio asignado. En cuanto a la parte 2, debes incluir **todos** los archivos que permitan visualizar tu página correctamente. En el repositorio base de la tarea, encontrarás una estructura propuesta que te ayudará a cumplir con el formato esperado.

²Para este ítem y los dos anteriores, se recomienda emplear los métodos `set_index` y `resample` de `pandas`.