

Práctica de Programación en Entornos de Datos 2021-2022

Instrucciones:

En esta práctica vamos a plantear un caso práctico que se tendrá que resolver con los contenidos vistos en el curso. El estudiante puede buscar y utilizar todas las funciones adicionales que crea necesarias, justificando su elección y explicando su uso.

La práctica se realizará sobre un notebook, que se entregará a través del apartado de entrega de tareas del curso virtual. Este notebook debe contener:

- El enunciado de cada apartado que se resuelva
- Justificación de las decisiones tomadas
- Código utilizado, mostrando el resultado de su ejecución, para poder responder a todas las preguntas que se realizan. Importante: el código debe funcionar en cualquier máquina sin necesidad de instalar nada. Para ello es importante usar rutas relativas y no absolutas.
- Cualesquiera comentarios sobre la realización de la práctica que el estudiante considere necesarios.

Enunciado:

Actualmente, la música se escribe principalmente utilizando siete notas. La primera es Do y le siguen, en este orden, Re, Mi, Fa, Sol, La y Si, que sería la séptima. La siguiente nota se llama nuevamente Do, pero es más agudo que el primero. Este Do está situado en una octava superior (el nombre de octava viene, precisamente, de que este DO sería la octava nota). Así, una octava es el conjunto de notas que van desde DO hasta SI.

Además de estas siete notas principales, existen cinco adicionales denominadas alteraciones. Por simplificar, se nombran utilizando una nota principal y añadiendo la palabra “sostenido”. Las cinco alteraciones serían: Do sostenido, Re sostenido, Fa sostenido, Sol sostenido y La sostenido. Es decir, cada octava tiene un total de 12 notas diferentes (siete principales y cinco alteraciones).

En un piano, las notas principales de una octava se corresponden con las teclas blancas, mientras que las alteraciones con las teclas negras. En la Figura 1 podemos ver las teclas de una octava con la nota que toca cada una de ellas.

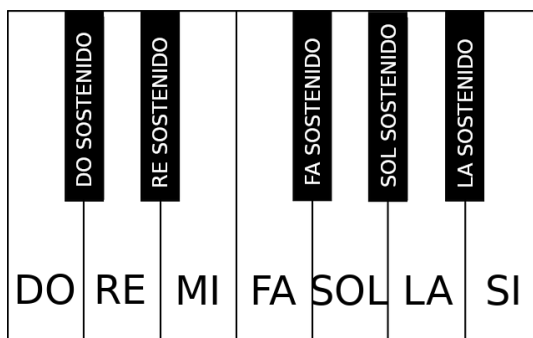


Figura 1: Notas de una octava del piano

El teclado completo de un piano se compone de 88 teclas entre blancas y negras y abarca siete octavas completas, numeradas de la 1 a la 7 y algunas notas extra: La, La sostenido y Si de lo que sería la octava 0 (por ser más grave que la octava 1) y el Do de lo que sería la octava 8 (por ser más aguda que la octava 7). En la Figura 2 podemos ver las 88 teclas del piano con las siete octavas completas resaltadas.

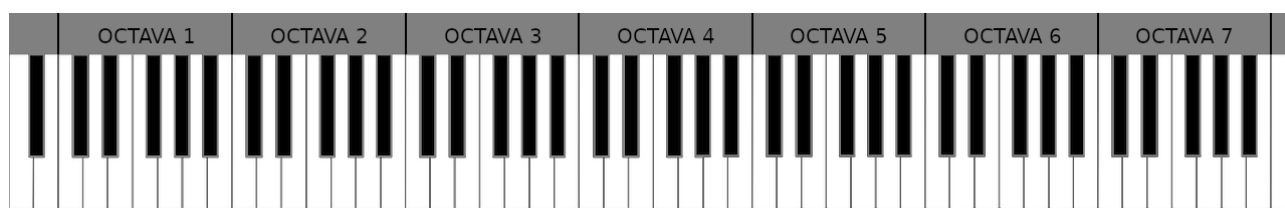


Figura 2: Las 88 teclas del piano y sus siete octavas completas

Disponemos de las partituras de diferentes piezas para piano escritas por diferentes compositores y nos piden realizar un estudio estadístico sobre ellas.

Cada partitura nos la dan ya preprocesada en un fichero de texto con la siguiente estructura:

- Primera línea: nombre de la partitura
- Segunda línea: compositor
- Resto de líneas: notas individuales que forman la partitura. Cada nota está representada en un string de dos o tres caracteres utilizando el sistema de notación anglosajón (en su variante inglesa/americana):
 - El primer carácter representa la nota C (Do), D (Re), E (Mi), F (Fa), G (Sol), A (La) y B (Si).
 - Si la nota es un sostenido, lleva a continuación el carácter #.
 - Por último, un carácter numérico del 0 al 8 representando la octava.

Así, las notas válidas van desde A0 (La de la octava 0) hasta C8 (Do de la octava 8), que se corresponden con las 88 teclas del piano. Por ejemplo, la nota D#3 representa el Re sostenido de la tercera octava, mientras que la nota G8 nunca aparecería, porque representaría el Sol de la octava ocho, que se sale del rango de teclas del piano.

Sin embargo, nos informan de que, debido a la tonalidad de algunas partituras, es posible que aparezcan lo que vamos a denominar *notas extrañas* como E# y B#. Aunque esas notas no se corresponden con ninguna de las teclas negras del piano (ya que serían Mi sostenido y Si sostenido respectivamente y, como podemos ver en la Figura 1, esas teclas no existen), musicalmente sí que tienen sentido, ya que un E# (Mi sostenido) es, en realidad, un F (Fa) de la misma octava, mientras que un B# (Si sostenido) es realmente un C (Do) de una octava superior.

Junto con el enunciado, se distribuye un fichero comprimido con los datos de varias partituras para utilizarlos como datos de entrada.

Se pide:

1. Construir un fichero CSV de 90 columnas en el que en cada línea tengamos la siguiente información:
 - La primera línea contiene los nombres de las columnas:
 - Primera columna: TITULO.
 - Segunda columna: AUTOR.
 - Sigüientes 88 columnas: los nombres de las notas (en la notación descrita) de las 88 teclas del piano, ordenadas desde la más grave (recordemos que es A0) hasta la más aguda (que sería C8).
 - Las sigüientes líneas contienen los datos de cada partitura
 - Primera columna: título de la partitura.
 - Segunda columna: autor de la partitura.

- Sigüientes 88 columnas: número de veces que la nota correspondiente aparece en la partitura.

Escribe el código que permita cargar el contenido de todos los ficheros situados en un directorio (cuyo nombre vendrá dado por parámetro) que contendrán las partituras y crear el CSV. Para ello, no olvides normalizar las *notas extrañas* E# y B# de acuerdo a lo comentado anteriormente.

2. Escribe el código que permita cargar un CSV como el creado en el ejercicio 1 en un DataFrame de Pandas. Trata de optimizar el espacio utilizado, incluye todas las explicaciones que consideres necesarias. El resto de ejercicios de la práctica realizarán su trabajo partiendo del DataFrame cargado en este ejercicio.
3. Queremos obtener respuesta a las siguientes preguntas:
 - a) Averiguar cuál es la tecla más popular de cada compositor individual y de todos los compositores de forma conjunta. Es decir: la tecla del piano que más veces aparece en todas las partituras de un compositor y en todas las partituras en global.
 - b) Averiguar cuál es la tecla más odiada de cada compositor individual y de todos los compositores de forma conjunta. Es decir: la tecla del piano que menos veces aparece en todas las partituras de un compositor y en todas las partituras en global.
 - c) Averiguar cuál es la octava favorita de cada compositor individual y de todos los compositores de forma conjunta. Es decir: la octava del piano (de 0 a 8) que más notas acumula entre todas las partituras de un compositor y entre todas las partituras en global.
 - d) Averiguar cuál es la octava más odiada de cada compositor individual y de todos los compositores de forma conjunta. Es decir: la octava del piano (de 0 a 8) que menos notas acumula entre todas las partituras de un compositor y entre todas las partituras en global.

Escribe el código python necesario para realizar los cálculos anteriores. Trata de escribir código reusable que pueda aprovecharse para más de un cálculo.

4. Sabiendo que, en términos musicales, la distancia entre dos teclas consecutivas del piano se denomina semitono (que es la mitad de un tono), queremos averiguar cual es la amplitud de cada partitura (que definiremos como la distancia entre la tecla más a la izquierda utilizada en la partitura y la tecla más a la derecha utilizada en la partitura) medida en tonos.

Por ejemplo, la distancia entre E4 (Mi de la octava 4) y A#4 (La sostenido de la octava 4) es de tres tonos (seis semitonos), ya que A#4 es la sexta tecla siguiente a E4 (las seis teclas siguientes a E4 son: F4, F#4, G4, G#4, A4 y, finalmente, A#4).

Para ello:

- a) Programa un método para obtener la nota más aguda (la tecla situada más a la derecha) que se utiliza en una partitura.
- b) Programa un método para obtener la nota más grave (la tecla situada más a la izquierda) que se utiliza en una partitura.
- c) Programa un método que calcule la amplitud de la partitura medida en tonos.
- d) Ahora deseamos averiguar cual es la partitura con mayor y menor amplitud en general y para cada compositor. Escribe el código necesario para realizar este cálculo.

5. Crea el código necesario para generar, de forma aleatoria, archivos con partituras en forma de fichero de texto como el utilizado en los ficheros de entrada. Ten en cuenta que en estos archivos pueden aparecer las *notas extrañas* (E# y B#), pero todas deberán estar en el rango de notas válido para un piano.