

1. De los datos proporcionados realiza los siguientes puntos:

- (a) Realiza el etiquetado de las variables de acuerdo a su tipo de variable
- (b) Eliminar duplicados
- (c) Completitud de las variables
- (d) Elimina las variables que tengan el 20 % o más de valores ausentes
- (e) ¿Cuántos de los registros en la variable "zip Code" son valores no válidos? , es decir, contienen letras
- (f) Elimina los registros que no tengas un "zip Code" válido, es decir contenga letras en los valores
- (g) ¿Cuántos de los registros en la variable "genero" son valores no validos? , es decir, contienen letras
- (h) Elimina los registros que no tengas un "genero" válido, es decir contenga letras en los valores
- (i) Haz una limpieza de la variable "name" , elimina caracteres especiales y todo debe estar en minúsculas
- (j) De la variable "artist" selecciona solo al primer artista que aparezca en la lista además de eliminar caracteres especiales
- (k) Normaliza la variable "genero" de tal forma que obtengas solo 8 categorías
- (l) Añade las siguientes columnas a tu tabla base: zip, lat, lng, city, state_name utilizando la tabla zips_practica
- (m) Convierte las variables "lat" y "lng" en tipo flotante y valida que la información sea consistente.
- (n) De la variable "city" y "state" elimina los dígitos que se encuentran dentro de las cadenas de texto
- (o) Crea una nueva variable llamada "state" que este conformada por "city", "state_name"
- (p) Los valores de la nueva variable "state" modificalos de cierta forma que todas sean minúsculas y sin acentos
- (q) Convierte los valores en la variable "release_date" a tipo datetime , además haz el conteo de las que no tienen la estructura necesaria para ser convertida en datetime y elimina esos registros

2. De la tabla tratada obten los siguientes datos:

- (a) Obten un DataFrame donde se muestre el conteo de registros por género
- (b) Genera una nueva variable que se llame "duration_minutos" que sea el valor de la variable "duration_ms" en minutos
- (c) ¿Cuáles son las 10 canciones más populares?
- (d) ¿Cuál es el promedio de duración en minutos y milisegundos?
- (e) ¿Cuál es el promedio y conteo de "energy" por "genero"?
- (f) ¿Qué canción tiene el "loudness" más bajo y cuál es la canción que tiene el "loudness" más alto?
- (g) Obtenga los 10 percentiles de las variables continuas