**Etapa 1 – Pré-processamento dos dados**

1. Carregar a base;
2. Verificar as informações;
3. Verificar valores ausentes;
4. Verificar duplicados;
5. Renomear colunas para Português: 'artista', 'musica', 'duraçao\_ms', 'explicito', 'ano', 'popularidade', 'dancabilidade', 'energia', 'tonalidade', 'volume', 'modo', 'discurso’, ‘acustica', 'instrumentalidade', 'ao\_vivo', 'valência', 'tempo’, ‘gênero'.
6. Contar a quantidade de gêneros musicais por registro;
7. Alterar dados com tipos incorretos;
8. Remover colunas desnecessárias;
9. Separar colunas numéricas e categóricas;
10. Contar o total de gêneros musicais;
11. Criar coluna para representar o principal gênero musical de cada canção.

**Perguntas**

1. Quantos registros estão presentes na base?
2. Quais são as colunas numéricas e as categóricas?

**Etapa 2 – EDA**

1. Estatísticas descritivas;
2. Em média, qual é o tempo (batidas por minuto) das músicas?
3. Em média, as músicas têm quantos gêneros musicais diferentes? Qual o maior observado para uma mesma música?
4. Quantos gêneros musicais são representados na base de dados? Gere a visualização para responder;
5. Quais os três principais gêneros musicais presentes na base de dados?
6. Cruze as variáveis numéricas com o gênero musical principal e destaque, considerando a média, qual gênero se destaca em cada variável. (Ex: em relação ao discurso, em média, as músicas de hip-hop são as que mais se destacam (0,15)).
7. Quais os 10 artistas mais populares no geral?
8. Correlacione as variáveis numéricas, e destaque as relações com valores elevados (próximos de 1 ou -1).
9. Há mais músicas em modo menor (0) ou maior (1)?
10. Há mais músicas taxadas como explícitas ou não explícitas?
11. Em média, a diferença entre músicas explícitas e não explícitas é muita alta?
12. Quantos artistas estão representados? E quais os dez mais comuns?