

# Primer Concepto -Análisis del estado de las bandas musicales en la escena colombiana

Rodrigo Castillo, Alejandra Campo, Gran Maestro Chamán.

October 21, 2021

## Contents

<b>1</b>	<b>Qué base de datos escogieron.</b>	<b>1</b>
1.1	Metodología de la recolección de datos sobre bandas en la escena colombiana . . . . .	2
1.1.1	SIMUS . . . . .	2
1.1.2	BOMM (Bogotá Music Market) . . . . .	3
1.1.3	Bandas o cantantes famosos . . . . .	4
1.1.4	Sin embargo, lo que se pretende con estos datos es extraer solamente los nombres de las bandas activas en la escena colombiana, pues, posteriormente, a partir de estas bandas se obtendrá la información de spotify, por lo que unificaré los datos. . . . .	5
1.2	Merge con Spotify: . . . . .	6

## 1 Qué base de datos escogieron.

La idea de este proyecto es hacer un análisis del estado de las bandas musicales en la escena musical colombiana, sin embargo, aunque existen muchas bases de datos y formas diversas de recolectar datos sobre música, no existe un dataset sólido sobre la música en Colombia que cuente con suficientes features para aplicar análisis estadístico de datos en ellos. La mayoría de las plataformas (como spotify) que cuentan con una API no dan información sobre la región en donde se encuentra un artista.

Sin embargo, no tiene gracia hacer análisis de datos sobre grandes datasets que no tienen influencia sobre los miembros del grupo y sobre los cuales ya existen análisis muy complejos hechos por gente mas experimentada, por lo

que el grupo se dispuso a hacer su propio dataframe de la música en Colombia.

## 1.1 Metodología de la recolección de datos sobre bandas en la escena colombiana

### 1.1.1 SIMUS

Existen diversos datos sobre bandas colombianas , por ejemplo existe una entidad que se llama el SIMUS en el cuál las bandas musicales que tienen contrataciones tienen que registrarse legalmente, esta base de datos cuenta con 882 bandas que se han registrado legalmente.

```
import pandas as pd
datos_simus = pd.read_excel("bandas_simus.xls")
print(datos_simus.head(10))
print(datos_simus.shape)
```

	Nombre	Tipo Agrupación
0	Ensamble musical casa de la cultura San Alberto	Conjuntos
1	Charlie Rueda & The Jazzmorgans	Agrupaciones de Música Popular Urbana
2	Gente pescaito	Conjuntos
3	los Andes	Conjuntos
4	CORO MUNICIPAL DE MALAMBO	Coros
5	CORO ATRES	Coros
6	A TRES TRIO	Trios
7	CUERPO Y ALMA LATIN ROCK BAND	Agrupaciones de Música Popular Urbana
8	banda escuela 16 de julio	Bandas
9	gaita casa de la cultura de san estanislao	Agrupaciones de Música Tradicional

(882, 6)

En esta base de datos cuenta con las columnas:

- nombre de la banda
- tipo de agrupación
- país(todas son colombianas)
- Departamento del país
- Municipio
- Estado de la banda

### 1.1.2 BOMM (Bogotá Music Market)

Existe una revista llamada BOMM es una plataforma de promoción y circulación organizada por la Cámara de Comercio de Bogotá, como parte de su programa de apoyo a las Industrias Creativas y Culturales, en esta página existe un catálogo de las bandas que han participado en eventos en el año 2021 (<https://www.bogotamusicmarket.com/Artistas?year=2021>) (no existen mas años y este input es vulnerable a RCE injection por si alguien quiere decirle a los de BOMM). En BOMM no existe propiamente un Dataset, sin embargo, están todas los artistas inscritos, por lo que se puede hacer un simple programa que los scrapee:

[illegible]

```

except:
    driver.close()

bomm_data = pd.DataFrame(bomm_list)
print(bomm_data)
print(bomm_data.shape())

done
      band_name
0    8bm - 8bits Memory
1      Ácido Pantera
2      Afro Legends
3  Agrupación Guarura
4      Aka Rezzo
..          ...
252      Yooko
253      Zafat
254    Zalama Crew
255    Zatélithe
256      Zultan

[257 rows x 1 columns]

```

### 1.1.3 Bandas o cantantes famosos

Con el fin del proyecto, una de las cosas que se busca analizar es que factores hacen que una banda sea famosa y que otra no , por lo que también añadiré una lista de cantantes muy famosos colombianos al dataset también scrapeado de <https://www.ranker.com/list/bands-from-colombia/reference>

**nota:**el script que scrapeo estos datos está en el repositorio pero no se puede llamar desde el notebook pues es un poco mañoso

```

import json
import pandas as pd
with open("famous_data.json") as file:
    data = json.load(file)
famous_dataframe = pd.DataFrame(data)
print(famous_dataframe)

```

```

      name

```

```

0           Shakira
1           Juanes
2       Mateo Camargo
3   Alfredo Gutierrez
4       Alex González
..           ...
100   Gabriel Torregrosa
101       Rafael Rodríguez
102       Fredys Arrieta
103       Alejandro Palacio
104   David Escobar Gallego

```

```
[105 rows x 1 columns]
```

1.1.4 Sin embargo, lo que se pretende con estos datos es extraer solamente los nombres de las bandas activas en la escena colombiana, pues, posteriormente, a partir de estas bandas se obtendrá la información de spotify, por lo que unificaré los datos.

```

import pandas as pd

colombia_scene_list = []
for i in famous_dataframe["name"]:
    data = {"name":i,
            "font":"famous"}
    colombia_scene_list.append(data)

for i in bomm_data["band_name"]:
    data = {"name":i,
            "font":"bomm"}
    colombia_scene_list.append(data)

for i in datos_simus["Nombre"]:
    data = {"name":i,
            "font":"simus"}
    colombia_scene_list.append(data)

colombia_state_names = pd.DataFrame(colombia_scene_list)
print(colombia_state_names)

```

	name	font
0	Shakira	famous
1	Juanes	famous
2	Mateo Camargo	famous
3	Alfredo Gutierrez	famous
4	Alex González	famous
...	...	...
1239	Semillero escuela Municipal de música de Piendamó	simus
1240	Chirimia del Pacifico Colombiano	simus
1241	ESCUELA DE FORMACION MUSICAL PAEZ VIVE	simus
1242	Pal' Lereo Pabla	simus
1243	Mar a fuera	simus

[1244 rows x 2 columns]

## 1.2 Merge con Spotify: