SPORTIFY MAP

By Victor Chang

```
In [5]:  import pandas as pd

         general_songs = pd.read_excel("C:/Users/Acer/Documents/Portafolio/playlist_songs
         general_songs
```

Out[5]:

	Song Name	Artists	Genres	Duration (sec)	Country_2
0	CARNIVAL	$¥$, $KanyeWest$, $TyDolla$ ign	NaN	264.33	US
1	What's Up?	4 Non Blondes	NaN	295.53	US
2	Lejos	6 Voltios	rock en español	372.32	PE
3	I Ran (So Far Away)	A Flock Of Seagulls	new wave, synthpop	308.93	GB
4	Dancin (feat. Luvli) - Krono Remix	Aaron Smith, Krono, Luvli	NaN	198.05	US
...	...	...	...	...	...
1997	Azul	Zoé	rock en español, mexican rock, mexican indie, ...	194.23	MX
1998	Arrullo De Estrellas	Zoé	rock en español, mexican rock, mexican indie, ...	252.69	MX
1999	Labios Rotos - En Vivo Desde México / 2010	Zoé	rock en español, mexican rock, mexican indie, ...	243.04	MX
2000	Soñé - En Vivo	Zoé	rock en español, mexican rock, mexican indie, ...	226.63	MX
2001	LOVELY BASTARDS	ZWE1HVNDXR, yatashigang	phonk, brazilian phonk	116.10	BY

2002 rows × 5 columns

In [6]:
```python
paises = pd.read_excel("C:/Users/Acer/Documents/Portafolio/playlist_songs_with_c
paises
```

Out[6]:

	Country_name	Country_name_large	Code_2	Code_3
0	Afghanistan	the Islamic Republic of Afghanistan	AF	AFG
1	Åland Islands	Åland	AX	ALA
2	Albania	the Republic of Albania	AL	ALB
3	Algeria	the People's Democratic Republic of Algeria	DZ	DZA
4	American Samoa	American Samoa	AS	ASM
...	...	...	...	...
266	Wallis and Futuna	the Territory of the Wallis and Futuna Islands	WF	WLF
267	Western Sahara	the Sahrawi Arab Democratic Republic	EH	ESH
268	Yemen	the Republic of Yemen	YE	YEM
269	Zambia	the Republic of Zambia	ZM	ZMB
270	Zimbabwe	the Republic of Zimbabwe	ZW	ZWE

271 rows × 4 columns

In [7]:
```python
# Unir general_songs con paises usando Country_2 y Code_2
general_songs = general_songs.merge(
    paises[['Code_2', 'Country_name']],  # Solo las columnas necesarias
    left_on='Country_2',
    right_on='Code_2',
    how='left'
)

# Si no quieres mantener Code_2 en el resultado:
general_songs.drop(columns='Code_2', inplace=True)

# Verificar
general_songs.head()
```

Out[7]:

	Song Name	Artists	Genres	Duration (sec)	Country_2	Country_name
0	CARNIVAL	$¥$, $KanyeWest$, $TyDolla$ ign	NaN	264.33	US	United States of America (the)
1	What's Up?	4 Non Blondes	NaN	295.53	US	United States of America (the)
2	Lejos	6 Voltios	rock en español	372.32	PE	Peru
3	I Ran (So Far Away)	A Flock Of Seagulls	new wave, synthpop	308.93	GB	United Kingdom of Great Britain and Northern I...
4	Dancin (feat. Luvli) - Krono Remix	Aaron Smith, Krono, Luvli	NaN	198.05	US	United States of America (the)

In [8]:
```python
general_songs.to_excel("C:/Users/Acer/Documents/Portafolio/playlist_songs_with_c
```

In [13]:
```python
general_songs = pd.read_excel("C:/Users/Acer/Documents/Portafolio/playlist_songs
general_songs
```

Out[13]:

	Song Name	Artists	Genres	Duration (sec)	Country_2	Country_name
0	CARNIVAL	¥$, KanyeWest, TyDolla$ ign	NaN	264.33	US	United States of America
1	What's Up?	4 Non Blondes	NaN	295.53	US	United States of America
2	Lejos	6 Voltios	rock en español	372.32	PE	Peru
3	I Ran (So Far Away)	A Flock Of Seagulls	new wave, synthpop	308.93	GB	United Kingdom of Great Britain and Northern I..
4	Dancin (feat. Luvli) - Krono Remix	Aaron Smith, Krono, Luvli	NaN	198.05	US	United States of America
...	...	...	...	...	...	..
1997	Azul	Zoé	rock en español, mexican rock, mexican indie, ...	194.23	MX	Mexico
1998	Arrullo De Estrellas	Zoé	rock en español, mexican rock, mexican indie, ...	252.69	MX	Mexico
1999	Labios Rotos - En Vivo Desde México / 2010	Zoé	rock en español, mexican rock, mexican indie, ...	243.04	MX	Mexico
2000	Soñé - En Vivo	Zoé	rock en español, mexican rock, mexican indie, ...	226.63	MX	Mexico
2001	LOVELY BASTARDS	ZWE1HVNDXR, yatashigang	phonk, brazilian phonk	116.10	BY	Belarus

2002 rows × 6 columns

◄ ▬▬▬▬▬▬▬▬▬▬▬▬▬▬▬▬▬▬▬▬▬▬▬ ►

In [14]:
```python
# 1) Contar canciones por país (ISO-2)
counts = (general_songs
          .dropna(subset=['Country_2'])
          .groupby('Country_2', as_index=False)
          .size()
          .rename(columns={'size': 'n_songs'}))

# 2) Pasar de ISO-2 a ISO-3 y nombre del país
map_df = counts.merge(
    paises[['Code_2', 'Code_3', 'Country_name']],
    left_on='Country_2', right_on='Code_2', how='left'
)

# (opcional) Revisar códigos sin mapeo
sin_mapeo = map_df[map_df['Code_3'].isna()]['Country_2'].unique()

# 3) Mapa con Plotly
import plotly.express as px
fig = px.choropleth(
    map_df,
    locations='Code_3',              # requiere ISO-3
    color='n_songs',
    hover_name='Country_name',
    title='Número de canciones por país',
    color_continuous_scale='Viridis'  # opcional
)
fig.update_layout(coloraxis_colorbar_title='Canciones')
fig.show()

# 4) Guardar a HTML (interactivo)
fig.write_html('mapa_canciones_por_pais.html')
```

In [15]:

```python
import numpy as np
import pandas as pd
import plotly.express as px
import plotly.graph_objects as go
# Métricas para el hover
total = map_df['n_songs'].sum()
map_df['share'] = map_df['n_songs'] / total
map_df['rank']  = map_df['n_songs'].rank(method='dense', ascending=False).astype

# 3) Figura base
fig = px.choropleth(
    map_df,
    locations='Code_3',                 # ISO-3
    color='n_songs',
    hover_name='Country_name',
    color_continuous_scale='YlOrRd',    # paleta más "choropleth"
    title='Número de canciones por país',
)

# Hover más informativo
fig.update_traces(
    customdata=np.stack([map_df['Country_name'], map_df['rank'], map_df['share']
    hovertemplate=(
        "<b>%{customdata[0]}</b><br>" +
        "Canciones: %{z:,}<br>" +
        "Participación: %{customdata[2]:.1%}<br>" +
        "Ranking: %{customdata[1]}<extra></extra>"
    )
```

```python
)

# Geografía y estilo
q95 = map_df['n_songs'].quantile(0.95)
fig.update_geos(
    projection_type='natural earth',
    showcountries=True, countrycolor="#eaeaea",
    showcoastlines=True, coastlinecolor="#cfcfcf",
)
fig.update_layout(
    coloraxis_colorbar_title='Canciones',
    coloraxis=dict(cmin=0, cmax=q95),   # atenúa outliers muy altos
    margin=dict(l=10, r=10, t=60, b=10)
)

# 4) Dropdown para cambiar paleta al vuelo
paletas = ['YlOrRd','Viridis','Plasma','Turbo','Blues','Greens']
fig.update_layout(
    updatemenus=[dict(
        buttons=[dict(label=p, method='restyle',
                      args=[{'colorscale':[p]}]) for p in paletas],
        direction='down', x=0.02, y=1.05, showactive=True
    )]
)

fig.show()
fig.write_html('mapa_canciones_por_pais_interactivo.html')
```

In [17]:
```python
promedio = general_songs['Duration (sec)'].mean()
print(promedio)
```

231.7786763236763

In [19]:
```python
import plotly.express as px

# Calcular media y desviación por país
stats = (general_songs
         .groupby('Country_name', as_index=False)
         .agg(
             mean_duration=('Duration (sec)', 'mean'),
             std_duration=('Duration (sec)', 'std')
         ))

# Gráfico de barras con barras de error
fig = px.bar(
    stats,
    x='Country_name',
    y='mean_duration',
    error_y='std_duration',  # Desviación estándar como barra de error
    title='Promedio de duración de canciones por país (±1 desviación estándar)',
    labels={'mean_duration': 'Duración promedio (segundos)', 'Country_name': 'Pa
)

fig.update_layout(
    xaxis_tickangle=45,
    height=600
)

fig.show()
```

In [ ]: