

```
In [ ]: import pandas as pd
import plotly.express as px
import plotly.graph_objects as go
```

```
In [ ]: df = pd.read_excel('MusicList.xlsx')

df = df.loc[:, ~df.columns.str.contains('^Unnamed')]

df = df[df['top genre'].notna()]

# replace every genre description with a general description of what they are
# e.g. chicago rap -> rap
genres = ["hip hop", "pop", "rock", "country", "drill", "jazz", "turkish", "lo-fi", "cl

for i in range(len(genres)):
    df.loc[df['top genre'].str.contains(genres[i]), 'top genre'] = genres[i]
```

```
In [ ]: # we need to get rid of some genres that are weirdly defined
df.loc[df['top genre'].str.contains('hip-hop'), 'top genre'] = "hip hop"
df.loc[df['top genre'].str.contains('electro'), 'top genre'] = "electronic"
df.loc[df['top genre'].str.contains('oyun-havasi'), 'top genre'] = "turkish"
df.loc[df['top genre'].str.contains('arabesk'), 'top genre'] = "turkish"
df.loc[df['top genre'].str.contains('beatlesque'), 'top genre'] = "rock"
df.loc[df['top genre'].str.contains('bass'), 'top genre'] = "electronic"
df.loc[df['top genre'].str.contains('edm'), 'top genre'] = "electronic"
df.loc[df['top genre'].str.contains('baglama'), 'top genre'] = "turkish"
df.loc[df['top genre'].str.contains('dubstep'), 'top genre'] = "electronic"

#drop if it is not in the list of genres
df = df[df['top genre'].isin(genres)]

df
```

[illegible]

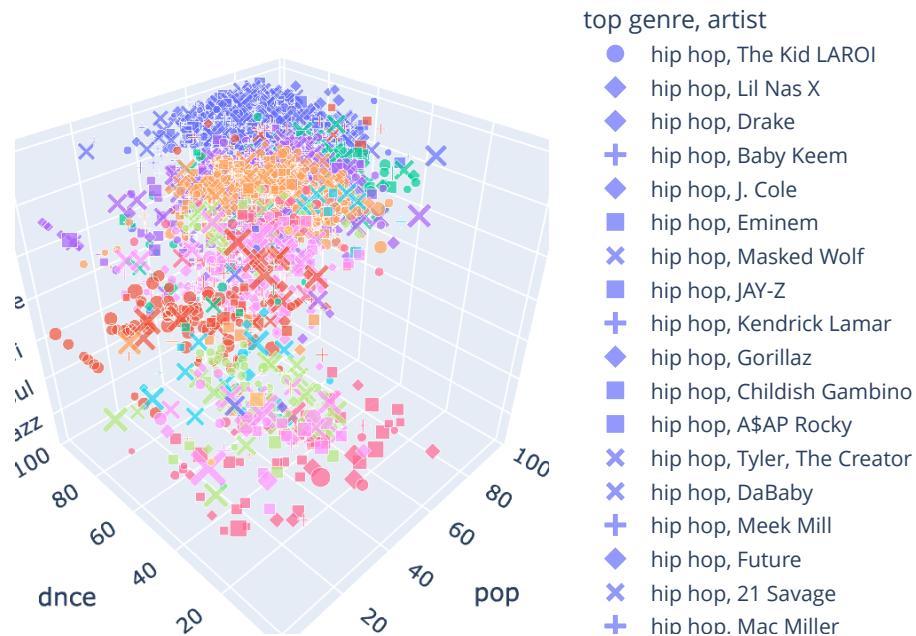
	title	artist	top genre	year	added	bpm	nrngy	dnce	dB	live	val	dur	acous
3196	Carmen Suite No.2/ Habanera (1.Akt) (Bizet)	Georges Bizet	classical	2009	2021-11-10	123	1	61	-26	18	49	132	94
3197	Valse sentimentale, Op.51, No.6	Pyotr Ilyich Tchaikovsky	classical	2016	2021-11-10	90	1	21	-35	7	6	180	98
3201	Trap Nilüfer	Orkundk	hip hop	2020	2020-10-29	145	31	46	-15	57	24	219	26
3202	The Court Of The Crimson King (Including The R...	King Crimson	rock	1969	2020-10-25	135	35	34	-12	33	7	563	7
3203	West End Girls - Ben Liebrand 9 Course Suite	Pet Shop Boys	pop	2020	2020-10-24	121	78	75	-6	37	53	518	5

2576 rows × 15 columns

```
In [ ]: fig = px.scatter_3d(df, x="pop", y="dnce", z="top genre", color="top genre", symbol="a
fig.update_scenes(xaxis_autorange="reversed")
fig.update_scenes(yaxis_autorange="reversed")
fig.update_scenes(zaxis_autorange="reversed")
fig.show()

fig.write_image("./fig1.svg")
```

![]



```
In [ ]: fig2 = px.pie(df, values="pop", names="top genre", title="Popularity by Genre", color_d
fig2.update_traces(textposition='inside', textinfo='percent+label')
fig2.show()

fig2.write_image("./fig2.svg")
```



Popularity by Genre

