

```
In [1]: import numpy as np
import pandas as pd
import psycopg2
from sklearn.cluster import KMeans
import matplotlib.pyplot as plt
import sqlalchemy
from sqlalchemy.ext.automap import automap_base
from sqlalchemy.orm import Session
from sqlalchemy import create_engine, func, inspect, desc
from config import username, password
```

```
In [2]: # create engine
engine = create_engine(f'postgresql+psycopg2://{username}:{password}@aws-gt-daviz-finalpg-001.cloqvwuqbywl.us-east-1.rds.amazonaws.com:5432/spotify_db')
```

```
In [3]: # reflect an existing database into a new model
Base = automap_base()

# reflect the tables
Base.prepare(engine, reflect=True)
```

```
In [4]: # We can view all of the classes that automap found
Base.classes.keys()
```

```
Out[4]: ['artistgenre', 'kmodel', 'artists', 'genres', 'tracks', 'years']
```

```
In [5]: # Save references to each table
Artistgenre = Base.classes.artistgenre
Artists = Base.classes.artists
Genres = Base.classes.genres
Tracks = Base.classes.tracks
Years = Base.classes.years
```

```
In [6]: # Look at columns in tables
inspector = inspect(engine)
columns = inspector.get_columns('tracks')
for column in columns:
    print(column["name"], column["type"])
```

```
acousticness DOUBLE PRECISION
artists TEXT
danceability DOUBLE PRECISION
duration_ms BIGINT
energy DOUBLE PRECISION
explicit BIGINT
id TEXT
instrumentalness DOUBLE PRECISION
key BIGINT
liveness DOUBLE PRECISION
loudness DOUBLE PRECISION
mode BIGINT
name TEXT
popularity BIGINT
release_date TEXT
speechiness DOUBLE PRECISION
tempo DOUBLE PRECISION
valence DOUBLE PRECISION
year BIGINT
```

```
In [7]: # Create our session (link) from Python to the DB
session = Session(engine)
```

```
In [8]: tracks_q = session.query(Tracks.id, Tracks.name, Tracks.artists, Tracks.acousticness,
                                Tracks.danceability, Tracks.energy, Tracks.instrumentalness,
                                Tracks.valence, Tracks.popularity, Tracks.year,
                                Tracks.key, Tracks.liveness, Tracks.loudness, Tracks.tempo).\
                                filter(Tracks.popularity > 20).all()

# store results in dataframe
tracks_kDF = pd.DataFrame(tracks_q, columns=['id', 'name', 'artists', 'acousticness',
                                             'danceability', 'energy', 'instrumentalness',
                                             'valence', 'popularity', 'year',
                                             'key', 'liveness', 'loudness', 'tempo'])

# Key: 0 is C natural, 1 is C#, 2 is D♭ and so on up to 11, which is B♭
# Loudness: Values typical range between -60 and 0 db.
```

```
In [9]: tracks_kDF['popularity'] = tracks_kDF['popularity']/1000
tracks_kDF['year'] = tracks_kDF['year']/2021/10
tracks_kDF['key'] = tracks_kDF['key']/11
tracks_kDF['loudness'] = tracks_kDF['loudness']/60
tracks_kDF['tempo'] = tracks_kDF['tempo']/244
```

```
In [10]: tracks_kDF.describe()
```

Out[10]:

	acousticness	danceability	energy	instrumentalness	valence	popularity
count	98468.000000	98468.000000	98468.000000	98468.000000	98468.000000	98468.000000
mean	0.372530	0.548568	0.555895	0.114267	0.544981	0.042022
std	0.334264	0.173603	0.253578	0.262269	0.261449	0.014184
min	0.000000	0.000000	0.000000	0.000000	0.000000	0.021000
25%	0.049800	0.431000	0.365000	0.000000	0.336000	0.031000
50%	0.281000	0.557000	0.568000	0.000090	0.555000	0.040000
75%	0.679000	0.674000	0.763000	0.018500	0.766000	0.052000
max	0.996000	0.988000	1.000000	1.000000	1.000000	0.100000

```
In [11]: session.close()
```

```
In [13]: len(tracks_kDF)
```

Out[13]: 98468

In [14]: *# Kmeans*

```
X = tracks_kDF.select_dtypes(np.number)
number_cols = list(X.columns)
tracks_fitted = KMeans(n_clusters=2500, init='k-means++', verbose=5).fit(X)
```

```
Initialization complete
Iteration 0, inertia 3700.2725651819264
Iteration 1, inertia 3281.2610617202263
Iteration 2, inertia 3179.330418229141
Iteration 3, inertia 3130.7326142655097
Iteration 4, inertia 3101.293585184495
Iteration 5, inertia 3082.3693310523377
Iteration 6, inertia 3068.807838449143
Iteration 7, inertia 3058.494896408551
Iteration 8, inertia 3050.783196877869
Iteration 9, inertia 3045.089265043015
Iteration 10, inertia 3040.6564367119463
Iteration 11, inertia 3037.0718612169494
Iteration 12, inertia 3034.3439200157886
Iteration 13, inertia 3032.0373310889368
Iteration 14, inertia 3030.1068622051994
Iteration 15, inertia 3028.484544714439
Iteration 16, inertia 3027.123711807117
Iteration 17, inertia 3025.900456385234
Iteration 18, inertia 3024.844413739896
Iteration 19, inertia 3023.9423437843884
Iteration 20, inertia 3023.1839103793027
Iteration 21, inertia 3022.4799262295924
Iteration 22, inertia 3021.870464186583
Iteration 23, inertia 3021.390859861625
Iteration 24, inertia 3021.0294096139273
Iteration 25, inertia 3020.727051783075
Iteration 26, inertia 3020.5079174621646
Iteration 27, inertia 3020.3067689079103
Iteration 28, inertia 3020.135058275098
Iteration 29, inertia 3020.0101889211533
Iteration 30, inertia 3019.9019285302757
Iteration 31, inertia 3019.8233477463405
Iteration 32, inertia 3019.7437377452056
Iteration 33, inertia 3019.706086759648
Iteration 34, inertia 3019.6692854026073
Iteration 35, inertia 3019.6393778791
Iteration 36, inertia 3019.613223898844
Iteration 37, inertia 3019.58943762318
Iteration 38, inertia 3019.5467528224917
Iteration 39, inertia 3019.5088076217303
Iteration 40, inertia 3019.4879684596253
Iteration 41, inertia 3019.4767961909665
Iteration 42, inertia 3019.4635280896546
Iteration 43, inertia 3019.4568704268886
Iteration 44, inertia 3019.456009015041
Converged at iteration 44: center shift 2.337063414064299e-29 within toleranc
e 4.5491510301177605e-06
Initialization complete
Iteration 0, inertia 3706.063270219865
Iteration 1, inertia 3280.273312469025
Iteration 2, inertia 3180.492004892706
Iteration 3, inertia 3131.961740663411
Iteration 4, inertia 3103.704473273612
Iteration 5, inertia 3085.4091189534597
Iteration 6, inertia 3072.093079294105
Iteration 7, inertia 3062.5326934881946
```

```
Iteration 8, inertia 3055.1773101573212
Iteration 9, inertia 3048.9550449254584
Iteration 10, inertia 3043.8521397176637
Iteration 11, inertia 3039.483068335052
Iteration 12, inertia 3035.904384711288
Iteration 13, inertia 3032.9356377728536
Iteration 14, inertia 3030.417432691227
Iteration 15, inertia 3028.585468579036
Iteration 16, inertia 3027.0351596166315
Iteration 17, inertia 3025.685496413792
Iteration 18, inertia 3024.583100148839
Iteration 19, inertia 3023.6578424122354
Iteration 20, inertia 3022.9041874600034
Iteration 21, inertia 3022.3014842586135
Iteration 22, inertia 3021.7702012924087
Iteration 23, inertia 3021.311831203422
Iteration 24, inertia 3020.8862605933446
Iteration 25, inertia 3020.4179814706335
Iteration 26, inertia 3020.0032844070197
Iteration 27, inertia 3019.6824369438305
Iteration 28, inertia 3019.43774740221
Iteration 29, inertia 3019.2213772707987
Iteration 30, inertia 3019.043397604579
Iteration 31, inertia 3018.8741432979446
Iteration 32, inertia 3018.6780352765822
Iteration 33, inertia 3018.533067783294
Iteration 34, inertia 3018.4006278974343
Iteration 35, inertia 3018.2683792392063
Iteration 36, inertia 3018.173109345543
Iteration 37, inertia 3018.073826999846
Iteration 38, inertia 3017.9877503689954
Iteration 39, inertia 3017.8887771866234
Iteration 40, inertia 3017.8228752210134
Iteration 41, inertia 3017.7778112999013
Iteration 42, inertia 3017.745719806867
Iteration 43, inertia 3017.714685551795
Iteration 44, inertia 3017.68572712043
Iteration 45, inertia 3017.6702715650704
Iteration 46, inertia 3017.654588704049
Iteration 47, inertia 3017.641732245016
Iteration 48, inertia 3017.6032101990077
Iteration 49, inertia 3017.5840895695387
Iteration 50, inertia 3017.572719088114
Iteration 51, inertia 3017.5380762524424
Iteration 52, inertia 3017.499335119023
Iteration 53, inertia 3017.444227625345
Iteration 54, inertia 3017.411950191208
Iteration 55, inertia 3017.3932464829522
Iteration 56, inertia 3017.3855015087966
Iteration 57, inertia 3017.380294283656
Iteration 58, inertia 3017.3789798673706
Iteration 59, inertia 3017.378112736955
Converged at iteration 59: center shift 2.535039389761676e-29 within toleranc
e 4.5491510301177605e-06
Initialization complete
Iteration 0, inertia 3700.6660060948516
Iteration 1, inertia 3278.7486333746915
```

Iteration 2, inertia 3178.9964210739017
Iteration 3, inertia 3130.596132154662
Iteration 4, inertia 3101.893788568396
Iteration 5, inertia 3083.6157387831922
Iteration 6, inertia 3070.937846679243
Iteration 7, inertia 3061.4411549916103
Iteration 8, inertia 3054.12337677557
Iteration 9, inertia 3048.2392166701807
Iteration 10, inertia 3043.803059163777
Iteration 11, inertia 3040.281319255449
Iteration 12, inertia 3037.351944546411
Iteration 13, inertia 3035.0123307890326
Iteration 14, inertia 3033.1992412110167
Iteration 15, inertia 3031.5681113493915
Iteration 16, inertia 3030.147173133418
Iteration 17, inertia 3029.023586662679
Iteration 18, inertia 3028.137292488244
Iteration 19, inertia 3027.4012562719886
Iteration 20, inertia 3026.7362354806282
Iteration 21, inertia 3026.1564025978787
Iteration 22, inertia 3025.6358534062815
Iteration 23, inertia 3025.2410700435926
Iteration 24, inertia 3024.9112553675527
Iteration 25, inertia 3024.645609725499
Iteration 26, inertia 3024.3942232343647
Iteration 27, inertia 3024.218173290921
Iteration 28, inertia 3024.049231409456
Iteration 29, inertia 3023.864895968386
Iteration 30, inertia 3023.733093685217
Iteration 31, inertia 3023.63322174365
Iteration 32, inertia 3023.5310963786646
Iteration 33, inertia 3023.4045660273314
Iteration 34, inertia 3023.2762750868956
Iteration 35, inertia 3023.1675865872194
Iteration 36, inertia 3023.0576092673914
Iteration 37, inertia 3022.980808911364
Iteration 38, inertia 3022.89973367682
Iteration 39, inertia 3022.8348554340137
Iteration 40, inertia 3022.790201709355
Iteration 41, inertia 3022.75012132382
Iteration 42, inertia 3022.7167777189748
Iteration 43, inertia 3022.695145173618
Iteration 44, inertia 3022.673420021223
Iteration 45, inertia 3022.6407123622766
Iteration 46, inertia 3022.613189622849
Iteration 47, inertia 3022.592775931592
Iteration 48, inertia 3022.5762147134087
Iteration 49, inertia 3022.5408755648573
Iteration 50, inertia 3022.504108121116
Iteration 51, inertia 3022.4841355713456
Iteration 52, inertia 3022.4628323066468
Iteration 53, inertia 3022.4518474788556
Iteration 54, inertia 3022.4391344223172
Iteration 55, inertia 3022.3952912651116
Iteration 56, inertia 3022.3682914746714
Iteration 57, inertia 3022.3347687235187
Iteration 58, inertia 3022.2895376466745

Iteration 59, inertia 3022.2660318216476
Iteration 60, inertia 3022.2555460713515
Iteration 61, inertia 3022.2513802870585
Iteration 62, inertia 3022.237284807157
Iteration 63, inertia 3022.232800147582
Iteration 64, inertia 3022.22865131249
Iteration 65, inertia 3022.2268027792834
Converged at iteration 65: center shift 2.413269807607492e-29 within tolerance 4.5491510301177605e-06
Initialization complete
Iteration 0, inertia 3701.7102495904187
Iteration 1, inertia 3277.5379668652727
Iteration 2, inertia 3180.3328880606737
Iteration 3, inertia 3133.7220343884724
Iteration 4, inertia 3104.8788552008596
Iteration 5, inertia 3085.011976756169
Iteration 6, inertia 3070.989852414189
Iteration 7, inertia 3060.434048336815
Iteration 8, inertia 3052.3489891652735
Iteration 9, inertia 3046.1821763714365
Iteration 10, inertia 3041.3880159865516
Iteration 11, inertia 3037.628500281483
Iteration 12, inertia 3034.2944328949143
Iteration 13, inertia 3031.339141026413
Iteration 14, inertia 3028.5722900886108
Iteration 15, inertia 3026.2809250542964
Iteration 16, inertia 3024.3657354957168
Iteration 17, inertia 3022.7060871755343
Iteration 18, inertia 3021.2662345481726
Iteration 19, inertia 3020.0268298467545
Iteration 20, inertia 3018.9394251011668
Iteration 21, inertia 3018.0398830037725
Iteration 22, inertia 3017.2544514309916
Iteration 23, inertia 3016.6147505962886
Iteration 24, inertia 3015.9393532986965
Iteration 25, inertia 3015.367956956756
Iteration 26, inertia 3014.8845022971186
Iteration 27, inertia 3014.562521446805
Iteration 28, inertia 3014.2826960733723
Iteration 29, inertia 3014.1248642817213
Iteration 30, inertia 3013.96488093471
Iteration 31, inertia 3013.8214793313623
Iteration 32, inertia 3013.7147693122274
Iteration 33, inertia 3013.628622840812
Iteration 34, inertia 3013.5676910999605
Iteration 35, inertia 3013.531982021356
Iteration 36, inertia 3013.478024200039
Iteration 37, inertia 3013.4364327122935
Iteration 38, inertia 3013.4131945236045
Iteration 39, inertia 3013.394480874506
Iteration 40, inertia 3013.371817957489
Iteration 41, inertia 3013.3419653825513
Iteration 42, inertia 3013.3150956304726
Iteration 43, inertia 3013.2884572345793
Iteration 44, inertia 3013.2638923281565
Iteration 45, inertia 3013.248291896254
Iteration 46, inertia 3013.240246710745


```
Iteration 47, inertia 3013.236520270301
Iteration 48, inertia 3013.2274196351827
Iteration 49, inertia 3013.2164359284816
Iteration 50, inertia 3013.207854985346
Iteration 51, inertia 3013.1954019575087
Iteration 52, inertia 3013.183254412056
Iteration 53, inertia 3013.1758759384156
Iteration 54, inertia 3013.166060938448
Iteration 55, inertia 3013.159007520165
Iteration 56, inertia 3013.154051480938
Iteration 57, inertia 3013.1436612032903
Iteration 58, inertia 3013.138438980181
Iteration 59, inertia 3013.1332408195835
Iteration 60, inertia 3013.127492715314
Iteration 61, inertia 3013.1244068408987
Converged at iteration 61: center shift 2.585749888809916e-29 within toleranc
e 4.5491510301177605e-06
Initialization complete
Iteration 0, inertia 3702.7554160636946
Iteration 1, inertia 3281.1717662938577
Iteration 2, inertia 3180.409066158064
Iteration 3, inertia 3132.89773504229
Iteration 4, inertia 3105.2125034054634
Iteration 5, inertia 3086.591321240971
Iteration 6, inertia 3073.047791115025
Iteration 7, inertia 3063.473792492169
Iteration 8, inertia 3056.051740262659
Iteration 9, inertia 3049.9960368598145
Iteration 10, inertia 3044.9033991358124
Iteration 11, inertia 3040.777592902029
Iteration 12, inertia 3037.5443918642313
Iteration 13, inertia 3034.7338790207937
Iteration 14, inertia 3032.4754552646623
Iteration 15, inertia 3030.8601953669486
Iteration 16, inertia 3029.4553821854374
Iteration 17, inertia 3028.232464748257
Iteration 18, inertia 3027.1850367366733
Iteration 19, inertia 3026.3423425302854
Iteration 20, inertia 3025.662624285713
Iteration 21, inertia 3025.0593214340674
Iteration 22, inertia 3024.4659315037093
Iteration 23, inertia 3024.032914703325
Iteration 24, inertia 3023.678492218338
Iteration 25, inertia 3023.3720087850156
Iteration 26, inertia 3023.0759166846938
Iteration 27, inertia 3022.826531916569
Iteration 28, inertia 3022.662809513436
Iteration 29, inertia 3022.4856748584075
Iteration 30, inertia 3022.278279931777
Iteration 31, inertia 3022.0613956087936
Iteration 32, inertia 3021.731606580992
Iteration 33, inertia 3021.4809342210474
Iteration 34, inertia 3021.2791186411614
Iteration 35, inertia 3021.117051856547
Iteration 36, inertia 3020.9623215146034
Iteration 37, inertia 3020.78908183209
Iteration 38, inertia 3020.653146592016
```

```
Iteration 39, inertia 3020.5493259832974
Iteration 40, inertia 3020.463695825878
Iteration 41, inertia 3020.367960351228
Iteration 42, inertia 3020.3012643174347
Iteration 43, inertia 3020.2630996073494
Iteration 44, inertia 3020.231742597747
Iteration 45, inertia 3020.196143812109
Iteration 46, inertia 3020.1607489012727
Iteration 47, inertia 3020.1109519716965
Iteration 48, inertia 3020.0590557811274
Iteration 49, inertia 3019.9942260802027
Iteration 50, inertia 3019.9595898494917
Iteration 51, inertia 3019.9059750372867
Iteration 52, inertia 3019.8433714393886
Iteration 53, inertia 3019.7708329470424
Iteration 54, inertia 3019.7045561054765
Iteration 55, inertia 3019.6474812496776
Iteration 56, inertia 3019.61797316842
Iteration 57, inertia 3019.604904261766
Iteration 58, inertia 3019.5993911236164
Iteration 59, inertia 3019.5924573962293
Iteration 60, inertia 3019.5909217012136
Iteration 61, inertia 3019.5900289991832
Converged at iteration 61: center shift 2.3722187166472842e-29 within toleran
ce 4.5491510301177605e-06
Initialization complete
Iteration 0, inertia 3699.0799465320006
Iteration 1, inertia 3278.3819424522217
Iteration 2, inertia 3178.466166782394
Iteration 3, inertia 3129.6272356535587
Iteration 4, inertia 3100.8375833311547
Iteration 5, inertia 3082.3882795058603
Iteration 6, inertia 3069.263836898137
Iteration 7, inertia 3059.31923102042
Iteration 8, inertia 3051.5709634470422
Iteration 9, inertia 3045.623124774408
Iteration 10, inertia 3040.965012458599
Iteration 11, inertia 3037.1577083851007
Iteration 12, inertia 3034.0128512470624
Iteration 13, inertia 3031.3706039347057
Iteration 14, inertia 3029.046903547245
Iteration 15, inertia 3027.080241023299
Iteration 16, inertia 3025.5589960896277
Iteration 17, inertia 3024.3047586352372
Iteration 18, inertia 3023.143878743841
Iteration 19, inertia 3022.222477198175
Iteration 20, inertia 3021.391352380539
Iteration 21, inertia 3020.565090148276
Iteration 22, inertia 3019.9261669508796
Iteration 23, inertia 3019.3453072678653
Iteration 24, inertia 3018.8664444772817
Iteration 25, inertia 3018.4151940004895
Iteration 26, inertia 3017.9378986655433
Iteration 27, inertia 3017.5131039511634
Iteration 28, inertia 3017.0900774574943
Iteration 29, inertia 3016.6524656287347
Iteration 30, inertia 3016.3046254079495
```

```
Iteration 31, inertia 3016.100365484463
Iteration 32, inertia 3015.9545517049887
Iteration 33, inertia 3015.838784900845
Iteration 34, inertia 3015.711912195053
Iteration 35, inertia 3015.603740844246
Iteration 36, inertia 3015.4779552773803
Iteration 37, inertia 3015.3411725772235
Iteration 38, inertia 3015.225778653554
Iteration 39, inertia 3015.1570504403744
Iteration 40, inertia 3015.1129793795008
Iteration 41, inertia 3015.064745388846
Iteration 42, inertia 3015.040238230417
Iteration 43, inertia 3015.0194506453354
Iteration 44, inertia 3015.002271526754
Iteration 45, inertia 3014.9845970511255
Iteration 46, inertia 3014.979737418412
Iteration 47, inertia 3014.9709169962066
Iteration 48, inertia 3014.9585283166575
Iteration 49, inertia 3014.9574896868435
Converged at iteration 49: center shift 2.4449427462513026e-29 within toleran
ce 4.5491510301177605e-06
Initialization complete
Iteration 0, inertia 3707.436579708424
Iteration 1, inertia 3284.8830041108263
Iteration 2, inertia 3183.1549694729088
Iteration 3, inertia 3134.7273841784013
Iteration 4, inertia 3106.8217559317372
Iteration 5, inertia 3087.5171710058526
Iteration 6, inertia 3073.9740879256215
Iteration 7, inertia 3063.8988560710486
Iteration 8, inertia 3056.0779742416526
Iteration 9, inertia 3050.1756035772128
Iteration 10, inertia 3045.7620502075024
Iteration 11, inertia 3042.1358694223727
Iteration 12, inertia 3039.169594757885
Iteration 13, inertia 3036.7570052774176
Iteration 14, inertia 3034.860607694204
Iteration 15, inertia 3033.166465616247
Iteration 16, inertia 3031.6030573854964
Iteration 17, inertia 3030.3709897766653
Iteration 18, inertia 3029.3769996867
Iteration 19, inertia 3028.48281945093
Iteration 20, inertia 3027.69077121967
Iteration 21, inertia 3027.005019394694
Iteration 22, inertia 3026.3806387755344
Iteration 23, inertia 3025.8478277175705
Iteration 24, inertia 3025.4025698149926
Iteration 25, inertia 3025.101337169128
Iteration 26, inertia 3024.82653926297
Iteration 27, inertia 3024.5752171871886
Iteration 28, inertia 3024.3420592601674
Iteration 29, inertia 3024.149799875266
Iteration 30, inertia 3024.0026329315415
Iteration 31, inertia 3023.8584516007663
Iteration 32, inertia 3023.7631838006064
Iteration 33, inertia 3023.658306211181
Iteration 34, inertia 3023.5854685090058
```

```
Iteration 35, inertia 3023.4964033599854
Iteration 36, inertia 3023.414283024605
Iteration 37, inertia 3023.3336388207567
Iteration 38, inertia 3023.275872156043
Iteration 39, inertia 3023.215621318051
Iteration 40, inertia 3023.1604790829097
Iteration 41, inertia 3023.0956921200027
Iteration 42, inertia 3023.05885758174
Iteration 43, inertia 3023.037753371713
Iteration 44, inertia 3023.027606800173
Iteration 45, inertia 3023.022422776976
Iteration 46, inertia 3023.020663530417
Iteration 47, inertia 3023.017211806942
Iteration 48, inertia 3023.0122435389285
Iteration 49, inertia 3023.006598132991
Converged at iteration 49: center shift 1.9920641150300885e-29 within tolerance 4.5491510301177605e-06
Initialization complete
Iteration 0, inertia 3710.4931319096654
Iteration 1, inertia 3281.739337813283
Iteration 2, inertia 3179.4569536760214
Iteration 3, inertia 3129.857613978698
Iteration 4, inertia 3100.178782709438
Iteration 5, inertia 3081.2352533108124
Iteration 6, inertia 3067.688613229443
Iteration 7, inertia 3057.6875998793003
Iteration 8, inertia 3050.5936445479147
Iteration 9, inertia 3045.024151842602
Iteration 10, inertia 3040.4531785589793
Iteration 11, inertia 3036.7083937620905
Iteration 12, inertia 3033.6171527322776
Iteration 13, inertia 3031.0648970685775
Iteration 14, inertia 3029.014483678286
Iteration 15, inertia 3027.2200826935323
Iteration 16, inertia 3025.72292262125
Iteration 17, inertia 3024.4544009966176
Iteration 18, inertia 3023.414606141631
Iteration 19, inertia 3022.4872419494395
Iteration 20, inertia 3021.657362442796
Iteration 21, inertia 3020.9157834361035
Iteration 22, inertia 3020.3271242337223
Iteration 23, inertia 3019.7738504362724
Iteration 24, inertia 3019.314268837811
Iteration 25, inertia 3018.952876581257
Iteration 26, inertia 3018.611630066699
Iteration 27, inertia 3018.322088467087
Iteration 28, inertia 3018.0644884573735
Iteration 29, inertia 3017.8141701003597
Iteration 30, inertia 3017.610740554922
Iteration 31, inertia 3017.4427287732888
Iteration 32, inertia 3017.2752365446286
Iteration 33, inertia 3017.1083115219417
Iteration 34, inertia 3016.97275081587
Iteration 35, inertia 3016.860073820989
Iteration 36, inertia 3016.721396599132
Iteration 37, inertia 3016.572645792222
Iteration 38, inertia 3016.4138264726653
```

Iteration 39, inertia 3016.294514293035
Iteration 40, inertia 3016.2015198609765
Iteration 41, inertia 3016.103724467367
Iteration 42, inertia 3016.050756758687
Iteration 43, inertia 3015.987819759852
Iteration 44, inertia 3015.9259792493563
Iteration 45, inertia 3015.8710293433437
Iteration 46, inertia 3015.797801962287
Iteration 47, inertia 3015.7374273717787
Iteration 48, inertia 3015.6676445977923
Iteration 49, inertia 3015.5792209442607
Iteration 50, inertia 3015.494524204539
Iteration 51, inertia 3015.4319108639525
Iteration 52, inertia 3015.371624638187
Iteration 53, inertia 3015.317631252385
Iteration 54, inertia 3015.24805750084
Iteration 55, inertia 3015.152100199898
Iteration 56, inertia 3015.0933534219716
Iteration 57, inertia 3015.03287034604
Iteration 58, inertia 3015.0048954677704
Iteration 59, inertia 3014.9926481564503
Iteration 60, inertia 3014.9871143123587
Iteration 61, inertia 3014.982599747526
Converged at iteration 61: center shift 2.2216213951645167e-29 within tolerance 4.5491510301177605e-06
Initialization complete
Iteration 0, inertia 3698.7054345023985
Iteration 1, inertia 3276.5118362211174
Iteration 2, inertia 3175.8275930053037
Iteration 3, inertia 3127.5874326824155
Iteration 4, inertia 3098.74261253699
Iteration 5, inertia 3078.926331145376
Iteration 6, inertia 3065.3220850871257
Iteration 7, inertia 3055.1224705569794
Iteration 8, inertia 3047.8331661636576
Iteration 9, inertia 3042.0709942408375
Iteration 10, inertia 3037.8940063877753
Iteration 11, inertia 3034.608912879601
Iteration 12, inertia 3031.7928192446434
Iteration 13, inertia 3029.475923099306
Iteration 14, inertia 3027.693169794705
Iteration 15, inertia 3026.173209702298
Iteration 16, inertia 3024.8924030207368
Iteration 17, inertia 3023.8195977766063
Iteration 18, inertia 3022.8647846315966
Iteration 19, inertia 3021.9709216058377
Iteration 20, inertia 3021.23339124334
Iteration 21, inertia 3020.7012724511196
Iteration 22, inertia 3020.2935744829533
Iteration 23, inertia 3019.9107237561047
Iteration 24, inertia 3019.5784717804736
Iteration 25, inertia 3019.2837713327262
Iteration 26, inertia 3019.0106750889345
Iteration 27, inertia 3018.7410831662
Iteration 28, inertia 3018.491863846509
Iteration 29, inertia 3018.273509856567
Iteration 30, inertia 3018.122823977469

```
Iteration 31, inertia 3017.9908117481923
Iteration 32, inertia 3017.881856711994
Iteration 33, inertia 3017.752516460321
Iteration 34, inertia 3017.6497309952833
Iteration 35, inertia 3017.5704475081884
Iteration 36, inertia 3017.512699205523
Iteration 37, inertia 3017.455169283846
Iteration 38, inertia 3017.398065033887
Iteration 39, inertia 3017.354933697791
Iteration 40, inertia 3017.3366053910745
Iteration 41, inertia 3017.3219053306725
Iteration 42, inertia 3017.3106921334256
Iteration 43, inertia 3017.3021604559967
Iteration 44, inertia 3017.2953344085267
Iteration 45, inertia 3017.281679736709
Iteration 46, inertia 3017.270356000214
Iteration 47, inertia 3017.2663444440036
Iteration 48, inertia 3017.264888197617
Converged at iteration 48: center shift 2.3296428685352238e-29 within toleran
ce 4.5491510301177605e-06
Initialization complete
Iteration 0, inertia 3697.9202656482285
Iteration 1, inertia 3277.243082138155
Iteration 2, inertia 3178.0337063015936
Iteration 3, inertia 3130.9500744409593
Iteration 4, inertia 3103.2568557898044
Iteration 5, inertia 3084.723851715987
Iteration 6, inertia 3071.4509411836148
Iteration 7, inertia 3061.7221780747063
Iteration 8, inertia 3054.2685765669567
Iteration 9, inertia 3048.2411657966227
Iteration 10, inertia 3043.5976672177803
Iteration 11, inertia 3039.8669918727787
Iteration 12, inertia 3036.9206434326143
Iteration 13, inertia 3034.659235963653
Iteration 14, inertia 3032.6170738112137
Iteration 15, inertia 3030.9392732781725
Iteration 16, inertia 3029.3465646861914
Iteration 17, inertia 3028.016159517495
Iteration 18, inertia 3026.849367276596
Iteration 19, inertia 3025.8231245387624
Iteration 20, inertia 3025.1557486089373
Iteration 21, inertia 3024.4221423194117
Iteration 22, inertia 3023.8225470120647
Iteration 23, inertia 3023.3679425528762
Iteration 24, inertia 3023.008834094402
Iteration 25, inertia 3022.6871727803787
Iteration 26, inertia 3022.462197590018
Iteration 27, inertia 3022.2349371656733
Iteration 28, inertia 3021.9681676479527
Iteration 29, inertia 3021.7334434266154
Iteration 30, inertia 3021.5366485209515
Iteration 31, inertia 3021.3869064192545
Iteration 32, inertia 3021.2777200727114
Iteration 33, inertia 3021.1331627181858
Iteration 34, inertia 3020.9254108775817
Iteration 35, inertia 3020.7597725222786
```

```

Iteration 36, inertia 3020.5879591427697
Iteration 37, inertia 3020.459672237637
Iteration 38, inertia 3020.3537417755656
Iteration 39, inertia 3020.2727676786703
Iteration 40, inertia 3020.2055690959555
Iteration 41, inertia 3020.1313501486893
Iteration 42, inertia 3020.07585370395
Iteration 43, inertia 3020.0594078900185
Iteration 44, inertia 3020.03816668188
Iteration 45, inertia 3020.0003897806832
Iteration 46, inertia 3019.973839618177
Iteration 47, inertia 3019.958658229459
Iteration 48, inertia 3019.9446564265
Iteration 49, inertia 3019.9363753050566
Iteration 50, inertia 3019.9140754383075
Iteration 51, inertia 3019.896577037439
Iteration 52, inertia 3019.881910500799
Iteration 53, inertia 3019.8762042049616
Iteration 54, inertia 3019.8721230832116
Iteration 55, inertia 3019.8692584009837
Converged at iteration 55: center shift 2.6523550157486445e-29 within tolerance 4.5491510301177605e-06

```

```
In [15]: song_cluster_labels = tracks_fitted.predict(X)
```

```
In [16]: tracks_kDF['cluster_label'] = song_cluster_labels
```

```
In [17]: # save model to csv
model_k = tracks_kDF[['id', 'name', 'artists', 'acousticness',
                      'danceability', 'energy', 'instrumentalness',
                      'valence', 'popularity', 'year',
                      'key', 'liveness', 'loudness', 'tempo',
                      'cluster_label']]
model_k.to_csv('chosenKmodel.csv', index=False)
```

Explore Model

```
In [18]: tracks_kDF.loc[tracks_kDF['name'] == "Heartbreak Hotel"]
```

Out[18]:

		id	name	artists	acousticness	danceability	energy	instrumentalness
1686	6xNwKKNYZcvgV3XTlwsgNio	Heartbreak Hotel	Heartbreak Hotel	['Elvis Presley']	0.840	0.696	0.210	
21935	2lfE8FhnPPEBSS4rL1qhBb	Heartbreak Hotel	Heartbreak Hotel	['Elvis Presley']	0.843	0.694	0.295	

◀ ▶

```
In [20]: tracks_kDF['search'] = tracks_kDF["artists"].str.find('Uriah Heep')

# All Time Low, Taylor Swift, Fleetwood Mac, Luke Bryan, Dan + Shay
```

```
In [21]: tracks_kDF.loc[tracks_kDF['search'] != -1].sort_values(by = 'popularity', ascending=False)
```


Out[21]:

		id	name	artists	acousticness	danceability	energy	instru
19547	3WAzdrpkaA9kGbl2SG88n4		Easy Livin'	['Uriah Heep']	0.002330	0.216	0.8760	
46800	0U3go7hV3vJQHsDS5Pi2O8		Lady In Black - Alternate Version	['Uriah Heep']	0.008430	0.503	0.5340	
16290	2AEW4oatiphhXuGRteMLqx		Stealin'	['Uriah Heep']	0.000896	0.328	0.4860	
26917	1pUgfxRPGR9HHc2ozHOjMa		The Wizard - 2017 Remastered	['Uriah Heep']	0.003350	0.416	0.6760	
75316	0Z80GSrYzoXlf0mCpBlcml		Free Me	['Uriah Heep']	0.043600	0.759	0.5130	
74316	6A4YSJ2IUysfT6Y63ZNQBR		Rain - 2017 Remastered	['Uriah Heep']	0.965000	0.296	0.0279	
90909	4Lm7bFL3LgRSVVQeh317d2		Come Away Melinda - 2016 - Remaster	['Uriah Heep']	0.602000	0.299	0.2270	
82893	2IR3UXzacEdTN38rTQpn1G		Sunrise - 2017 Remastered	['Uriah Heep']	0.014200	0.376	0.6160	
74274	7HEA2JdVohL9m9DIXBwyXS		Circle of Hands - 2017 Remastered	['Uriah Heep']	0.008190	0.191	0.5090	
73985	65jjvicwOPcgfKw6x7irMm		Gypsy - 2016 - Remaster	['Uriah Heep']	0.000889	0.220	0.4980	
65318	10jaZ7UjuwFDpfgNXmDpEB		Rainbow Demon - 2017 Remastered	['Uriah Heep']	0.019200	0.276	0.4990	
74270	5IR4Tz9BNOLaWns2Ja8i7a		The Magician's Birthday - 2017 Remastered	['Uriah Heep']	0.003830	0.315	0.8760	
65377	0PyqmFI4A9w53n8l9VEeO3		Traveller in Time - 2017 Remastered	['Uriah Heep']	0.039400	0.425	0.6750	
91042	31D3Z706PJ8spuY2zBDjO1		Bird of Prey - 2016 - Remaster	['Uriah Heep']	0.002720	0.401	0.6640	
82985	3Q8PIY9eipV8Y8O4bkLxVd		Paradise - 2017 Remastered	['Uriah Heep']	0.089200	0.594	0.3750	
82917	3qwwwxepJUZ8ypHo7ZnRO8l		Easy Livin' - 2017 Remastered	['Uriah Heep']	0.000270	0.218	0.8230	

	id	name	artists	acousticness	danceability	energy	instru
91169	5aOIVLaFBhp8b7bCyHAka	Sweet Lorraine - 2017 Remastered	['Uriah Heep']	0.016700	0.405	0.6000	

```
In [22]: tracks_kDF.loc[tracks_kDF['search'] != -1]["cluster_label"].unique()
```

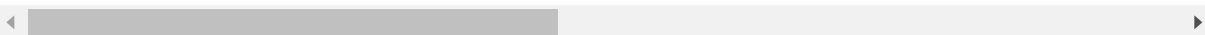
Out[22]: array([1424, 1652, 740, 1590, 1001, 323, 917, 758, 2308, 2323, 531, 863, 1665, 365, 642, 749, 2345])

```
In [23]: recommended = tracks_kDF.loc[tracks_kDF['cluster_label']== 1652].sort_values(b  
y = 'popularity', ascending=False)  
recommended
```

Out[23]:

		id	name	artists	acousticness	danceability	ener
5500	753KutoAy00apPspIMRetG		Let There Be Rock	['AC/DC']	0.000209	0.332	0.7
19547	3WAzdrpkaA9kGbl2SG88n4		Easy Livin'	['Uriah Heep']	0.002330	0.216	0.8
19731	4c6yZLnA730uJgQf9FTW2M		Heresy	['Pantera']	0.000034	0.304	0.9
20123	3vjHuVFZGC0Bm2Ud0wNpgs		Sweat	['TOOL']	0.000302	0.338	0.8
15913	1luUe3A7EfJmvZH3EljtwR		Hocus Pocus - Original Single Version	['Focus']	0.051400	0.266	0.8
52122	4OGcBC7ekhVszg4YdQug7O		You Can't Quit Me Baby	['Queens of the Stone Age']	0.023000	0.306	0.7
30505	6HqnBoikcyjy711kbrfhlr		Poison Was The Cure - Remastered 2004	['Megadeth']	0.000147	0.278	0.8
42439	6qRRnwgJADWFbWUbVtaKNI		Bad Dog No Biscuits	['SEATBELTS']	0.005320	0.400	0.9
51117	6WBdwdLLif4kuG3s6ot8uB		Wolverine Blues	['Entombed']	0.000457	0.348	0.9
96018	3B46jXophS1p38YQzu5A1K		Force Your Way	['Nobuo Uematsu']	0.000009	0.405	0.7
78560	5aMsCHy4nyk0LKJnk63aCr		He's My Thing - 2007 Remaster	['Babes In Toyland']	0.000008	0.232	0.9
59317	7a2Y8ocNI3hFk0LTnfhN7M		Bad Penny	['Big Black']	0.000006	0.436	0.9
36544	0ng5q14OCwukb4W09AuuQv		Ibiza Bar	['Pink Floyd']	0.000643	0.110	0.7
59179	4m4Z6Rkz0XJclrxz9vtMBQ		Starpower	['Sonic Youth']	0.008020	0.310	0.8
76866	5dfoQUgmrP6KhRkuTLFb1Y		The Living End	['The Jesus and Mary Chain']	0.000001	0.206	0.7
76953	6ej8EFrgYq86kPfFti4VfX		Never Understand	['The Jesus and Mary Chain']	0.000034	0.204	0.9
77864	5ilKaDGB9NcbgepirpKQFK		Burning Fight	['Inside Out']	0.000046	0.238	0.7
94012	7xXGnYqfPeGCu4SSggDjtR		Back to Shalla-Bal	['Joe Satriani']	0.001140	0.308	0.9
76599	1RyEhDluS4ulmGan8BuoRh		Procreation (Of the Wicked)	['Celtic Frost']	0.000003	0.269	0.7
85066	3isS9hdnmyMePPrKm6qEQU		The Third of the Storms (Evoked Damnation)	['Hellhammer']	0.000602	0.173	0.8
94196	4oqZhaz3506QfeMXc2sXaL		Kaleidoscope - 2001 Remaster	['Ride']	0.000023	0.148	0.8

	id	name	artists	acousticness	danceability	energy
68123	5IBxOsRIIkWEJh4vFNR5gk	Peace Sells - Randy Burns Mix	['Megadeth']	0.014800	0.288	0.9
77050	4xa7uy2h4QX6IIUxMCi0Bi	The Conjuring - Randy Burns Mix	['Megadeth']	0.000596	0.227	0.9
93010	0kLm1VTFWFSzmMmcSy7zOz	Morbid Tales	['Celtic Frost']	0.000001	0.179	0.9
64653	6ghf1333jb7cJ9sQCgMRWI	Not Right - 2019 Remaster	['The Stooges']	0.002480	0.303	0.7
92192	0Z6TxhJh5U7zzyrf2w7udy	It's Cold Outside	['Stiv Bators']	0.005370	0.414	0.9



In [24]: `tracks_kDF.describe()`

Out[24]:

	acousticness	danceability	energy	instrumentalness	valence	popularity
count	98468.000000	98468.000000	98468.000000	98468.000000	98468.000000	98468.000000
mean	0.372530	0.548568	0.555895	0.114267	0.544981	0.042022
std	0.334264	0.173603	0.253578	0.262269	0.261449	0.014184
min	0.000000	0.000000	0.000000	0.000000	0.000000	0.021000
25%	0.049800	0.431000	0.365000	0.000000	0.336000	0.031000
50%	0.281000	0.557000	0.568000	0.000090	0.555000	0.040000
75%	0.679000	0.674000	0.763000	0.018500	0.766000	0.052000
max	0.996000	0.988000	1.000000	1.000000	1.000000	0.100000



In [26]: `len(recommended)`

Out[26]: 26

```
In [27]: for i, row in recommended.iterrows():  
         print(f"{row['name']} - {row['artists']}")
```

```
Let There Be Rock - ['AC/DC']  
Easy Livin' - ['Uriah Heep']  
Heresy - ['Pantera']  
Sweat - ['TOOL']  
Hocus Pocus - Original Single Version - ['Focus']  
You Can't Quit Me Baby - ['Queens of the Stone Age']  
Poison Was The Cure - Remastered 2004 - ['Megadeth']  
Bad Dog No Biscuits - ['SEATBELTS']  
Wolverine Blues - ['Entombed']  
Force Your Way - ['Nobuo Uematsu']  
He's My Thing - 2007 Remaster - ['Babes In Toyland']  
Bad Penny - ['Big Black']  
Ibiza Bar - ['Pink Floyd']  
Starpower - ['Sonic Youth']  
The Living End - ['The Jesus and Mary Chain']  
Never Understand - ['The Jesus and Mary Chain']  
Burning Fight - ['Inside Out']  
Back to Shalla-Bal - ['Joe Satriani']  
Procreation (Of the Wicked) - ['Celtic Frost']  
The Third of the Storms (Evoked Damnation) - ['Hellhammer']  
Kaleidoscope - 2001 Remaster - ['Ride']  
Peace Sells - Randy Burns Mix - ['Megadeth']  
The Conjuring - Randy Burns Mix - ['Megadeth']  
Morbid Tales - ['Celtic Frost']  
Not Right - 2019 Remaster - ['The Stooges']  
It's Cold Outside - ['Stiv Bators']
```

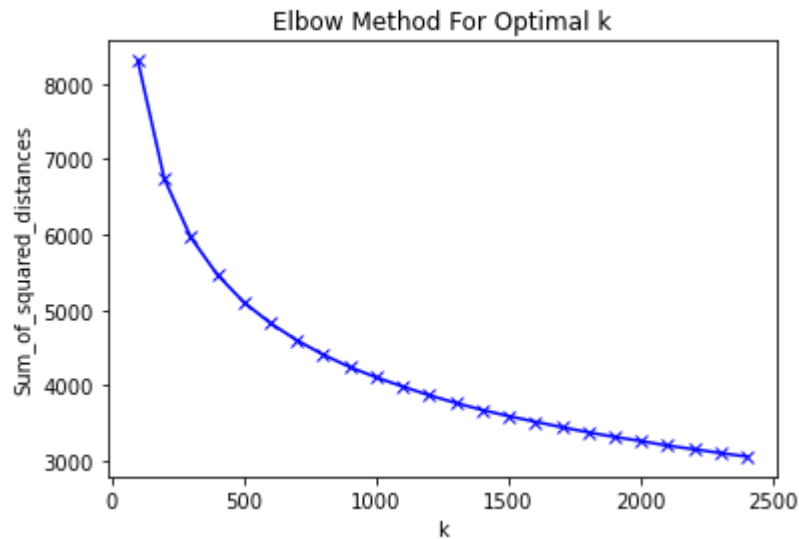
```
In [ ]:
```

```
In [28]: Sum_of_squared_distances = []  
K = range(100,2500,100)  
for k in K:  
    km = KMeans(n_clusters=k)  
    km = km.fit(X)  
    Sum_of_squared_distances.append(km.inertia_)  
print(k)
```

100
200
300
400
500
600
700
800
900
1000
1100
1200
1300
1400
1500
1600
1700
1800
1900
2000
2100
2200
2300
2400

```
In [29]: # sample elbow after added columns
plt.plot(K, Sum_of_squared_distances, 'bx-')
plt.xlabel('k')
plt.ylabel('Sum_of_squared_distances')
plt.title('Elbow Method For Optimal k')

plt.savefig('ElbowMethod-addedCol-max91.png')
plt.show()
```



```
In [30]: for i in range(1,10):
          print(f'{K[i]} Clusters - {Sum_of_squared_distances[i]} inertia')
```

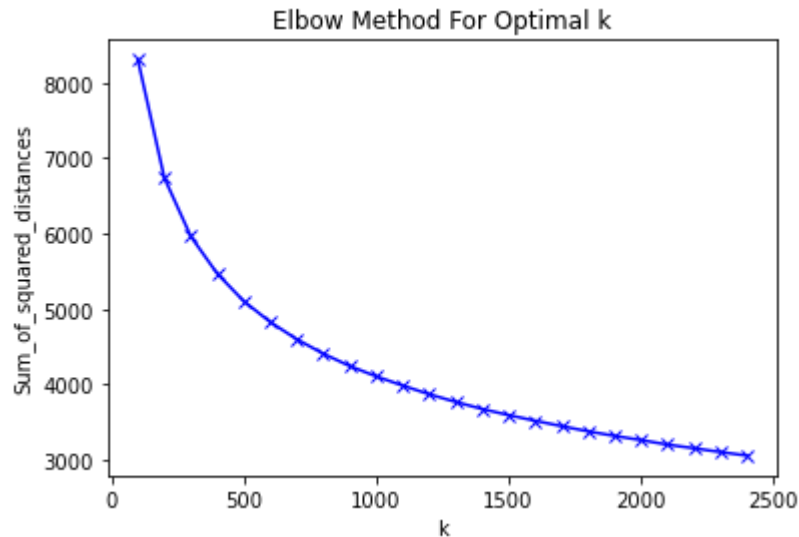
```
200 Clusters - 6740.0254928079585 inertia
300 Clusters - 5963.570418735913 inertia
400 Clusters - 5468.723442265337 inertia
500 Clusters - 5105.18997897541 inertia
600 Clusters - 4828.43150479651 inertia
700 Clusters - 4600.7228094078555 inertia
800 Clusters - 4409.217425200976 inertia
900 Clusters - 4245.692699446985 inertia
1000 Clusters - 4108.799642848685 inertia
```



```
In [33]: #don't mess with this one
import matplotlib.pyplot as plt

plt.plot(K, Sum_of_squared_distances, 'bx-')
plt.xlabel('k')
plt.ylabel('Sum_of_squared_distances')
plt.title('Elbow Method For Optimal k')

plt.savefig('ElbowMethodK.png')
plt.show()
```



```
In [32]: for i in range(1,10):
          print(f'{K[i]} Clusters - {Sum_of_squared_distances[i]} inertia')

200 Clusters - 6740.0254928079585 inertia
300 Clusters - 5963.570418735913 inertia
400 Clusters - 5468.723442265337 inertia
500 Clusters - 5105.18997897541 inertia
600 Clusters - 4828.43150479651 inertia
700 Clusters - 4600.7228094078555 inertia
800 Clusters - 4409.217425200976 inertia
900 Clusters - 4245.692699446985 inertia
1000 Clusters - 4108.799642848685 inertia
```

In []:

In []:

In []:

```
In [ ]: # Visualizing the Clusters with PCA
import plotly.express as px
from sklearn.decomposition import PCA

song_embedding = PCA(n_components=2).fit_transform(X)
projection = pd.DataFrame(columns=['x', 'y'], data=song_embedding)
projection['title'] = tracks_kDF['name']
projection['cluster'] = tracks_kDF['cluster_label']

fig = px.scatter(projection, x='x', y='y', color='cluster', hover_data=['x',
'y', 'title'])
fig.show()
```

```
In [ ]:
```