```
import numpy as np
import pandas as pd
import matplotlib.pyplot as plt
import seaborn as sns
```

```
df_tracks = pd.read_csv('/content/track.csv')
df_tracks.head()
```

	Unnamed: 0	track_id	artists	album_name	track_name	pop
0	0	5SuOikwiRyPMVoIQDJUgSV	Gen Hoshino	Comedy	Comedy	
1	1	4qPNDBW1i3p13qLCt0Ki3A	Ben Woodward	Ghost (Acoustic)	Ghost - Acoustic	
2	2	1iJBSr7s7jYXzM8EGcbK5b	Ingrid Michaelson;ZAYN	To Begin Again	To Begin Again	
3	3	6lfxq3CG4xtTiEg7opyCyx	Kina Grannis	Crazy Rich Asians (Original Motion Picture Sou...	Can't Help Falling In Love	
4	4	5vjLSffimiIP26QG5WcN2K	Chord Overstreet	Hold On	Hold On	

5 rows × 21 columns

```
pd.isnull(df_tracks).sum()
```

```
Unnamed: 0          0
track_id            0
artists             1
album_name          1
track_name          1
popularity          0
duration_ms         0
explicit            0
danceability        0
energy              0
key                 0
loudness            0
mode                0
speechiness         0
acousticness        0
instrumentalness    0
liveness            0
valence             0
tempo               0
time_signature      0
track_genre         0
dtype: int64
```

```
df_tracks.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 114000 entries, 0 to 113999
Data columns (total 21 columns):
 #   Column            Non-Null Count   Dtype  
---  ------            --------------   -----  
 0   Unnamed: 0        114000 non-null  int64  
 1   track_id          114000 non-null  object 
 2   artists           113999 non-null  object 
 3   album_name        113999 non-null  object 
 4   track_name        113999 non-null  object 
 5   popularity        114000 non-null  int64  
 6   duration_ms       114000 non-null  int64  
 7   explicit          114000 non-null  bool
 8   danceability      114000 non-null  float64
 9   energy            114000 non-null  float64
 10  key               114000 non-null  int64
 11  loudness          114000 non-null  float64
 12  mode              114000 non-null  int64
 13  speechiness       114000 non-null  float64
 14  acousticness      114000 non-null  float64
 15  instrumentalness  114000 non-null  float64
 16  liveness          114000 non-null  float64
 17  valence           114000 non-null  float64
 18  tempo             114000 non-null  float64
 19  time_signature    114000 non-null  int64
 20  track_genre       114000 non-null  object
dtypes: bool(1), float64(9), int64(6), object(5)
memory usage: 17.5+ MB
```

```
sorted_df = df_tracks.sort_values('popularity', ascending = True).head(10)
sorted_df
```

	Unnamed: 0	track_id	artists	album_name	track_r
56999	56999	4zJLkYcHGjWuf2Ui0eVAso	Håkan Hellström	LUGNA LÅTAR	Det kom aldri över för
68346	68346	3bvSoEdHqhDv5jmIP0qflo	Brray	Homecoming Latin Party	Bich Con C
68347	68347	25WU3f3gv6ATdPtSG5cybu	Don Omar;Juan Magán	Perreo Tenebroso Vol. 4	No S Modas Ella S Mo
19647	19647	0EBP3qfDpCZE2skX3PfbTf	Tracy Lawrence	Chillin' It - Mellow Day Country	Just and
68361	68361	5Ct6xnkAQJ7lZXp7vHM8uF	Chris Jedi;Ozuna;Brytiago	Perreo Tenebroso Vol. 4	Bip
19645	19645	5wW7fkOaNCKHM8NvEDf8Kh	Big & Rich;Bon Jovi	Chillin' It - Mellow Day Country	Born A
19644	19644	3yz60wsJlg630mjAfU4qlv	Steve Earle	Good Times Country	Go Ama
19643	19643	4BviPcJC1obuSn4sKkyRqG	Tracy Lawrence	Country Car Hits	Excit
19642	19642	1up0F8kK6mgObzqWi9Myom	Steve Earle	Finest Country	Jerusa
19641	19641	4FQDsvQMu5ny7mY8ehv4vP	Sugarland	Christmas Country Songs 2022	W Wonder

10 rows × 21 columns

```
df_tracks.describe().transpose()
```

	count	mean	std	min	25%	
Unnamed: 0	114000.0	56999.500000	32909.109681	0.000	28499.75000	5699
popularity	114000.0	33.238535	22.305078	0.000	17.00000	3
duration_ms	114000.0	228029.153114	107297.712645	0.000	174066.00000	2129
danceability	114000.0	0.566800	0.173542	0.000	0.45600	
energy	114000.0	0.641383	0.251529	0.000	0.47200	
key	114000.0	5.309140	3.559987	0.000	2.00000	
loudness	114000.0	-8.258960	5.029337	-49.531	-10.01300	
mode	114000.0	0.637553	0.480709	0.000	0.00000	
speechiness	114000.0	0.084652	0.105732	0.000	0.03590	
acousticness	114000.0	0.314910	0.332523	0.000	0.01690	

```
most_popular = df_tracks.query('popularity > 90', inplace = False).sort_values('popularity',ascending = False)
most_popular[:10]
```

	Unnamed: 0	track_id	artists	album_name	track_nam
20001	20001	3nqQXoyQOWXiESFLlDF1hG	Sam Smith;Kim Petras	Unholy (feat. Kim Petras)	Unholy (feat. Kim Petras
81051	81051	3nqQXoyQOWXiESFLlDF1hG	Sam Smith;Kim Petras	Unholy (feat. Kim Petras)	Unholy (feat. Kim Petras
51664	51664	2tTmW7RDtMQtBk7m2rYeSw	Bizarrap;Quevedo	Quevedo: Bzrp Music Sessions, Vol. 52	Quevedo Bzrp Music Sessions Vol. 5
81210	81210	4uUG5RXrOk84mYEfFvj3cK	David Guetta;Bebe Rexha	I'm Good (Blue)	I'm Good (Blue
89411	89411	5ww2BF9slyYgNOk37BlC4u	Manuel Turizo	La Bachata	La Bachata
20008	20008	4uUG5RXrOk84mYEfFvj3cK	David Guetta;Bebe Rexha	I'm Good (Blue)	I'm Good (Blue
88410	88410	5ww2BF9slyYgNOk37BlC4u	Manuel Turizo	La Bachata	La Bachata
30003	30003	4uUG5RXrOk84mYEfFvj3cK	David Guetta;Bebe Rexha	I'm Good (Blue)	I'm Good (Blue
67356	67356	5ww2BF9slyYgNOk37BlC4u	Manuel Turizo	La Bachata	La Bachata
68303	68303	5ww2BF9slyYgNOk37BlC4u	Manuel Turizo	La Bachata	La Bachata

10 rows × 21 columns

```
df_tracks [["artists"]].iloc[18]
```

```
artists    Jason Mraz;Colbie Caillat
Name: 18, dtype: object
```

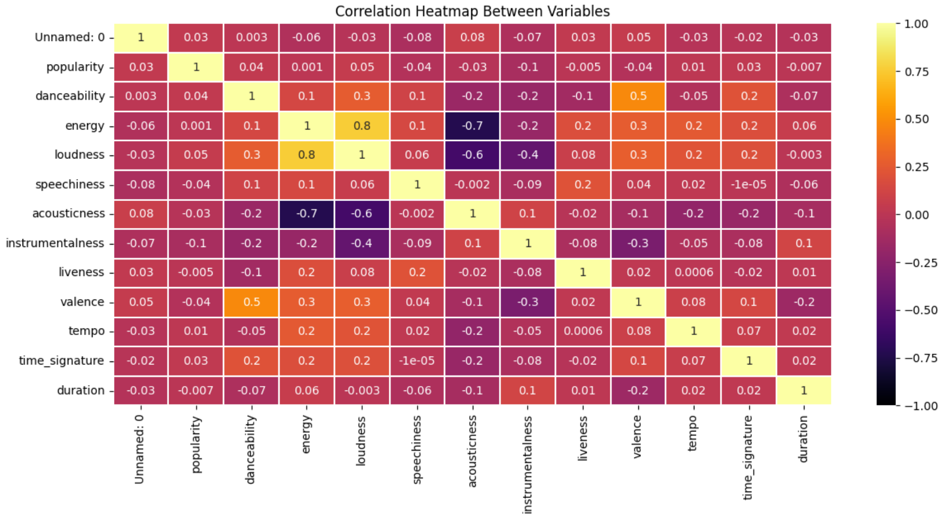
```
df_tracks["duration"] = df_tracks["duration_ms"].apply(lambda x : round(x/1000))
df_tracks.drop("duration_ms", inplace = True, axis =1)
df_tracks.duration.head()
```

```
0    231
1    150
2    211
```

```
3    202
4    199
Name: duration, dtype: int64
```

```
corr_df = df_tracks.drop(["key", "mode", "explicit"], axis=1).corr(method="pearson")

plt.figure(figsize=(14, 6))

heatmap = sns.heatmap(corr_df, annot=True, fmt=".1g", vmin=-1, vmax=1, center=0, cmap="inferno", linewidths=0.1, linecol

heatmap.set_title("Correlation Heatmap Between Variables")

heatmap.set_xticklabels(heatmap.get_xticklabels(), rotation=90)

plt.show()
```

```
<ipython-input-10-8135ad29342c>:1: FutureWarning: The default value of numeric_onl
  corr_df = df_tracks.drop(["key", "mode", "explicit"], axis=1).corr(method="pears
```



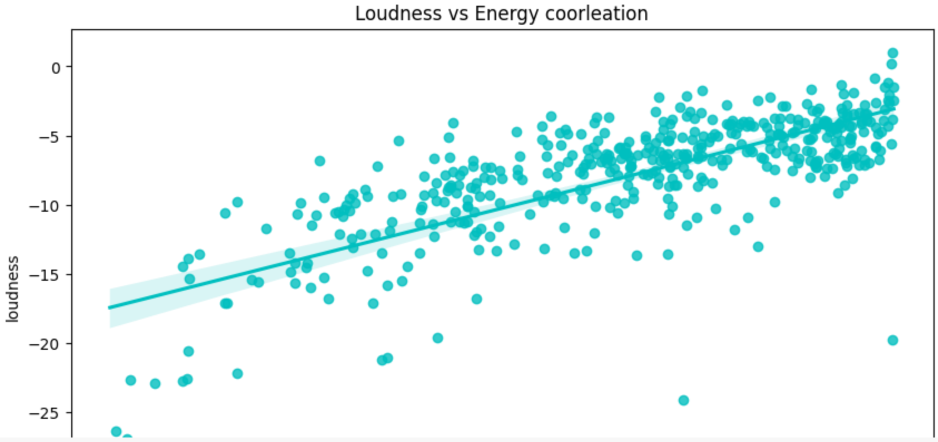
	Unnamed: 0	popularity	danceability	energy	loudness	speechiness	acousticness	instrumentalness	liveness	valence	tempo	time_signature	duration
Unnamed: 0	1	0.03	0.003	-0.06	-0.03	-0.08	0.08	-0.07	0.03	0.05	-0.03	-0.02	-0.03
popularity	0.03	1	0.04	0.001	0.05	-0.04	-0.03	-0.1	-0.005	-0.04	0.01	0.03	-0.007
danceability	0.003	0.04	1	0.1	0.3	0.1	-0.2	-0.2	-0.1	0.5	-0.05	0.2	-0.07
energy	-0.06	0.001	0.1	1	0.8	0.1	-0.7	-0.2	0.2	0.3	0.2	0.2	0.06
loudness	-0.03	0.05	0.3	0.8	1	0.06	-0.6	-0.4	0.08	0.3	0.2	0.2	-0.003
speechiness	-0.08	-0.04	0.1	0.1	0.06	1	-0.002	-0.09	0.2	0.04	0.02	-1e-05	-0.06
acousticness	0.08	-0.03	-0.2	-0.7	-0.6	-0.002	1	0.1	-0.02	-0.1	-0.2	-0.2	-0.1
instrumentalness	-0.07	-0.1	-0.2	-0.2	-0.4	-0.09	0.1	1	-0.08	-0.3	-0.05	-0.08	0.1
liveness	0.03	-0.005	-0.1	0.2	0.08	0.2	-0.02	-0.08	1	0.02	0.0006	-0.02	0.01
valence	0.05	-0.04	0.5	0.3	0.3	0.04	-0.1	-0.3	0.02	1	0.08	0.1	-0.2
tempo	-0.03	0.01	-0.05	0.2	0.2	0.02	-0.2	-0.05	0.0006	0.08	1	0.07	0.02
time_signature	-0.02	0.03	0.2	0.2	0.2	-1e-05	-0.2	-0.08	-0.02	0.1	0.07	1	0.02
duration	-0.03	-0.007	-0.07	0.06	-0.003	-0.06	-0.1	0.1	0.01	-0.2	0.02	0.02	1

```
sample_df = df_tracks.sample(int(0.004*len(df_tracks)))
print(len(sample_df))
```

```
456
```

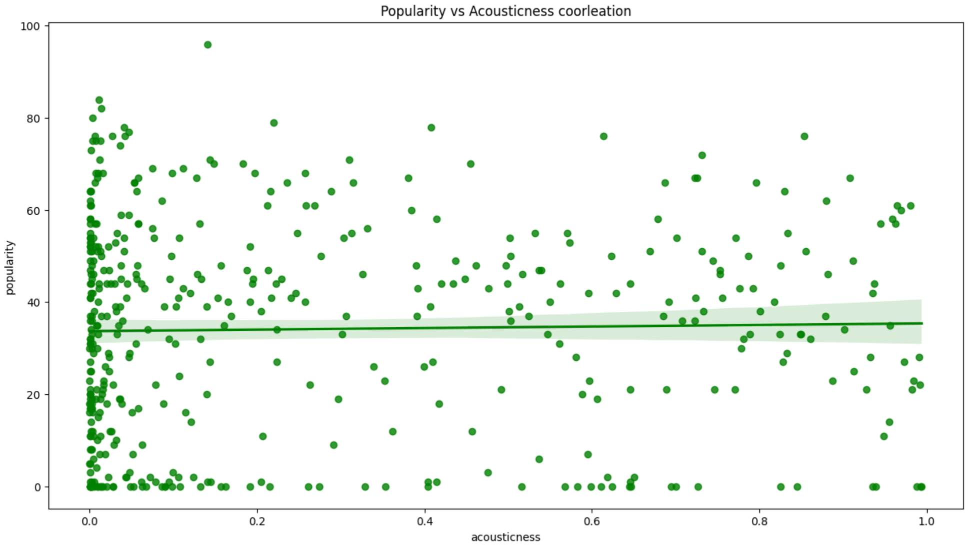
```
plt.figure(figsize = (10,6))
sns.regplot(data = sample_df, y = "loudness", x = "energy", color = "c").set(title = "Loudness vs Energy coorleation")
```

```
[Text(0.5, 1.0, 'Loudness vs Energy coorleation')]
```



```
plt.figure(figsize = (15,8))
sns.regplot(data = sample_df, y = "popularity", x = "acousticness", color = "g").set(title = "Popularity vs Acousticness
```

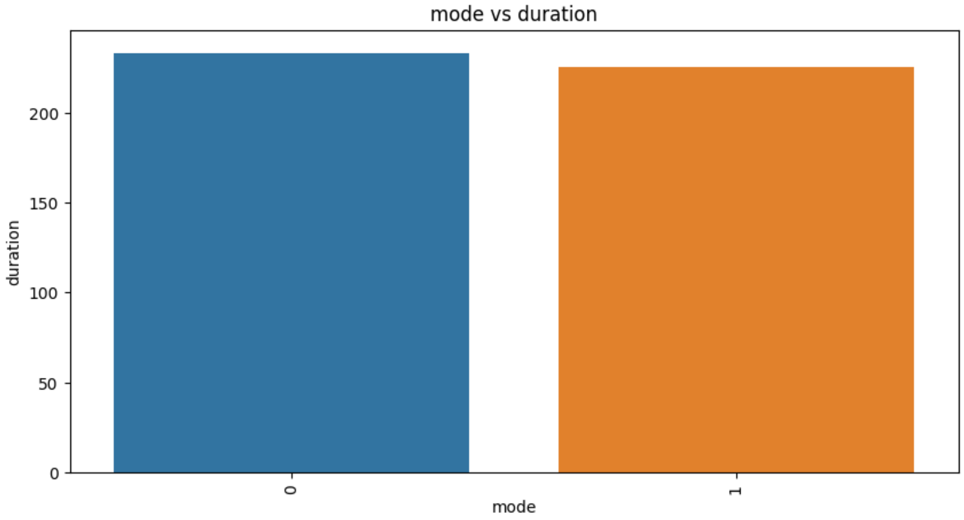
```
[Text(0.5, 1.0, 'Popularity vs Acousticness coorleation')]
```



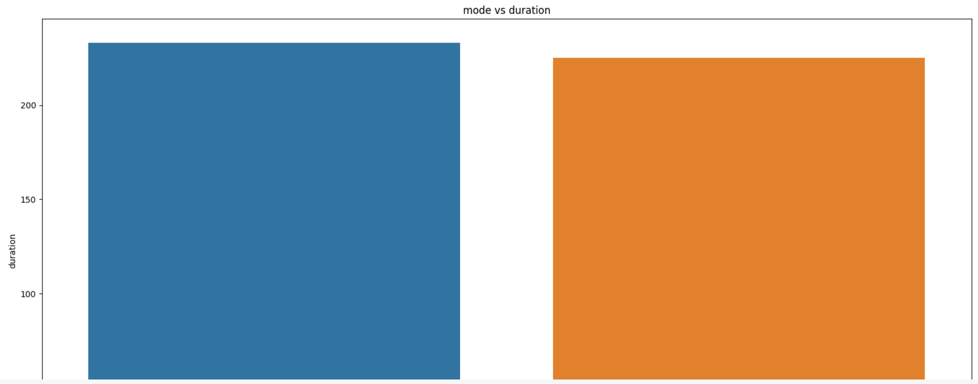
```
!pip install seaborn==0.11.0
```

```
Collecting seaborn==0.11.0
  Downloading seaborn-0.11.0-py3-none-any.whl (283 kB)
     ━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━ 283.1/283.1 kB 6.2 MB/s eta 0:00:00
Requirement already satisfied: numpy>=1.15 in /usr/local/lib/python3.10/dist-package
Requirement already satisfied: scipy>=1.0 in /usr/local/lib/python3.10/dist-packages
Requirement already satisfied: pandas>=0.23 in /usr/local/lib/python3.10/dist-packag
Requirement already satisfied: matplotlib>=2.2 in /usr/local/lib/python3.10/dist-pac
Requirement already satisfied: contourpy>=1.0.1 in /usr/local/lib/python3.10/dist-pa
Requirement already satisfied: cycler>=0.10 in /usr/local/lib/python3.10/dist-packag
```

```
fig_dims = (10, 5)
fig, ax = plt.subplots(figsize=fig_dims)
sns.barplot(x='mode', y='duration', data=df_tracks, ax=ax, errwidth=False)
plt.xticks(rotation=90)
plt.title("mode vs duration")
plt.show()
```



```
fig_dims = (20, 10)
fig, ax = plt.subplots(figsize=fig_dims)
sns.barplot(x='mode', y='duration', data=df_tracks, ax=ax, errwidth=False)
plt.xticks(rotation=60)
plt.title("mode vs duration")
plt.show()
```



```
df_genre = pd.read_csv('/content/SpotifyAudioFeaturesNov2018.csv')
df_genre.head()
```

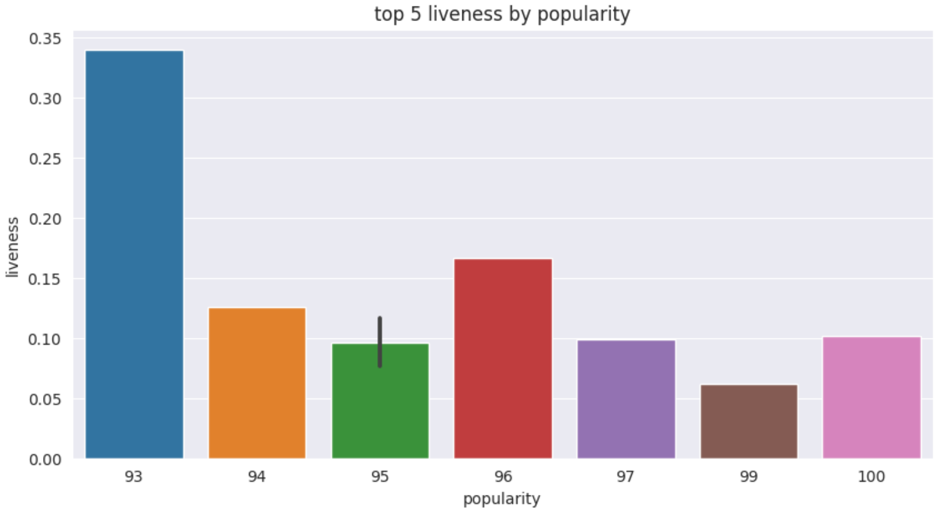
	artist_name	track_id	track_name	acousticness	danceability
0	YG	2RM4jf1Xa9zPgMGRDiht8O	Big Bank feat. 2 Chainz, Big Sean, Nicki Minaj	0.005820	0.743
1	YG	1tHDG53xJNGsItRA3vfVgs	BAND DRUM (feat. A$AP Rocky)	0.024400	0.846
2	R3HAB	6Wosx2euFPMT14UXiWudMy	Radio Silence	0.025000	0.603
3	Chris Cooq	3J2Jpw61sO7l6Hc7qdYV91	Lactose	0.029400	0.800
4	Chris Cooq	2jbYvQCyPgX3CdmAzeVeuS	Same - Original mix	0.000035	0.783

```
df_genre.isnull().sum()
```

```
artist_name         0
track_id            0
track_name          0
acousticness        0
danceability        0
duration_ms         0
energy              0
instrumentalness    0
key                 0
liveness            0
loudness            0
mode                0
speechiness         0
tempo               0
time_signature      0
valence             0
popularity          0
dtype: int64
```

```
sns.set_style(style = "darkgrid")
plt.figure(figsize = (10,5))
famous = df_genre.sort_values("popularity", ascending = False).head(10)
sns.barplot(y = 'liveness', x = 'popularity', data = famous).set(title = "top 5 liveness by popularity")
```

```
[Text(0.5, 1.0, 'top 5 liveness by popularity')]
```



✓ 0s completed at 12:31 PM