In [2]:
```
import numpy as np
import pandas as pd
import matplotlib.pyplot as plt
import seaborn as sns
```

In [3]:
```
df_tracks=pd.read_csv('tracks.csv')
df_tracks.head()
```

Out[3]:

	id	name	popularity	duration_ms	explicit	artists	id_artists	release_date	dancea
0	35iwgR4jXetI318WEWsa1Q	Carve	6	126903	0	['Uli']	['45tIt06XoI0Iio4LBEVpls']	1922-02-22	
1	021ht4sdgPcrDgSk7JTbKY	Capítulo 2.16 - Banquero Anarquista	0	98200	0	['Fernando Pessoa']	['14jtPCOoNZwquk5wd9DxrY']	1922-06-01	
2	07A5yehtSnoedViJAZkNnc	Vivo para Quererte - Remasterizado	0	181640	0	['Ignacio Corsini']	['5LiOoJbxVSAMkBS2fUm3X2']	1922-03-21	
3	08FmqUhxtyLTn6pAh6bk45	El Prisionero - Remasterizado	0	176907	0	['Ignacio Corsini']	['5LiOoJbxVSAMkBS2fUm3X2']	1922-03-21	
4	08y9GfoqCWfOGsKdwojr5e	Lady of the Evening	0	163080	0	['Dick Haymes']	['3BiJGZsyX9sJchTqcSA7Su']	1922	

In [4]:
```
#null values

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

Out[4]:
```
id                  0
name               71
popularity          0
duration_ms         0
explicit            0
artists             0
id_artists          0
release_date        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
dtype: int64
```

In [5]:
```
1 df_tracks.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 586672 entries, 0 to 586671
Data columns (total 20 columns):
 #   Column            Non-Null Count   Dtype
---  ------            --------------   -----
 0   id                586672 non-null  object
 1   name              586601 non-null  object
 2   popularity        586672 non-null  int64
 3   duration_ms       586672 non-null  int64
 4   explicit          586672 non-null  int64
 5   artists           586672 non-null  object
 6   id_artists        586672 non-null  object
 7   release_date      586672 non-null  object
 8   danceability      586672 non-null  float64
 9   energy            586672 non-null  float64
 10  key               586672 non-null  int64
 11  loudness          586672 non-null  float64
 12  mode              586672 non-null  int64
 13  speechiness       586672 non-null  float64
 14  acousticness      586672 non-null  float64
 15  instrumentalness  586672 non-null  float64
 16  liveness          586672 non-null  float64
 17  valence           586672 non-null  float64
 18  tempo             586672 non-null  float64
 19  time_signature    586672 non-null  int64
dtypes: float64(9), int64(6), object(5)
memory usage: 89.5+ MB
```

In [6]:
```
1 sorted_df=df_tracks.sort_values('popularity',ascending=True).head(10)
2 sorted_df
```

Out[6]:

	id	name	popularity	duration_ms	explicit	artists	id_artists	release_date	da
546130	181rTRhCcggZPwP2TUcVqm	Newspaper Reports On Abner, 20 February 1935	0	896575	0	['Norris Goff', 'Chester Lauck', 'Carlton Bric...	['3WCwCPDMpGzrt0Qz6quumy', '7vk8UqABg0Sga78Gl3...	1935-02-20	
546222	0yOCz3V5KMm8l1T8EFc60i	恋は水の上で	0	188440	0	['Hibari Misora']	['1m5pMY5blqJwdxJ7vqQtuN']	1949	
546221	0y48Hhwe52099UqYjegRCO	私の誕生日	0	173467	0	['Hibari Misora']	['1m5pMY5blqJwdxJ7vqQtuN']	1949	
546220	0xCmgtf9ka07hkZg3D6PaV	エル・チョクロ (EL CHOCLO)	0	205280	0	['Hibari Misora']	['1m5pMY5blqJwdxJ7vqQtuN']	1949	
546219	0tBXS3VuCPX7KWUFH2nros	恋は不思議なもの	0	185733	0	['Hibari Misora']	['1m5pMY5blqJwdxJ7vqQtuN']	1949	
546218	0qrKnQtYDVJhKFAXTHYVS9	ゆうべはどうしたの (WHATSA MALLA U)	0	183427	0	['Hibari Misora']	['1m5pMY5blqJwdxJ7vqQtuN']	1949	
546217	0nqsDxOeKSwEzp3AUQAAqS	Screen Director's Playhouse, Music For Million...	0	1767071	0	['Wilms Herbert', 'June Allyson', 'Joseph Kear...	['2rbm8QWvmnVwxFo84EVM1h', '4yW5adMgyIfHFzaL9i...	1949-04-10	
546216	0kGEdsxVLYjCdfxM9tbezd	ブルーマンボ	0	162147	0	['Hibari Misora']	['1m5pMY5blqJwdxJ7vqQtuN']	1949	
546215	0bc3PUZurUUXrY7yqoOxjq	Screen Director's Playhouse, Trade Winds direc...	0	1776652	0	['Wally Maher', 'Tay Garnett', 'Lurene Tuttle'...	['7hkhJTTI3VnUGVWUt8SJXT', '3kYeeIpRCgJz4fQYDv...	1949-05-19	
546214	0Wwm0ruSjYMIiWG0nyAI1F	Screen Director's Playhouse, It's A Wonderful ...	0	1767576	0	['Joseph Granby', 'Jimmy Stewart', 'Irene Tedr...	['6GK59BC4LJzqR0OpHAX2S3', '58BzBaExrnrx898sby...	1949-05-08	

In [7]:
```
df_tracks.describe().transpose()
```

Out[7]:

	count	mean	std	min	25%	50%	75%	max
popularity	586672.0	27.570053	18.370642	0.0	13.0000	27.000000	41.00000	100.000
duration_ms	586672.0	230051.167286	126526.087418	3344.0	175093.0000	214893.000000	263867.00000	5621218.000
explicit	586672.0	0.044086	0.205286	0.0	0.0000	0.000000	0.00000	1.000
danceability	586672.0	0.563594	0.166103	0.0	0.4530	0.577000	0.68600	0.991
energy	586672.0	0.542036	0.251923	0.0	0.3430	0.549000	0.74800	1.000
key	586672.0	5.221603	3.519423	0.0	2.0000	5.000000	8.00000	11.000
loudness	586672.0	-10.206067	5.089328	-60.0	-12.8910	-9.243000	-6.48200	5.376
mode	586672.0	0.658797	0.474114	0.0	0.0000	1.000000	1.00000	1.000
speechiness	586672.0	0.104864	0.179893	0.0	0.0340	0.044300	0.07630	0.971
acousticness	586672.0	0.449863	0.348837	0.0	0.0969	0.422000	0.78500	0.996
instrumentalness	586672.0	0.113451	0.266868	0.0	0.0000	0.000024	0.00955	1.000
liveness	586672.0	0.213935	0.184326	0.0	0.0983	0.139000	0.27800	1.000
valence	586672.0	0.552292	0.257671	0.0	0.3460	0.564000	0.76900	1.000
tempo	586672.0	118.464857	29.764108	0.0	95.6000	117.384000	136.32100	246.381
time_signature	586672.0	3.873382	0.473162	0.0	4.0000	4.000000	4.00000	5.000

In [8]:
```
most_popular=df_tracks.query('popularity>90', inplace=False).sort_values('popularity', ascending=False)
most_popular[:10]
```

Out[8]:

	id	name	popularity	duration_ms	explicit	artists	id_artists	release_date	dan
93802	4iJyoBOLtHqaGxP12qzhQI	Peaches (feat. Daniel Caesar & Giveon)	100	198082	1	['Justin Bieber', 'Daniel Caesar', 'Giveon']	['1uNFoZAHBGtllmzznpCI3s', '20wkVLutqVOYrc0kxF...	2021-03-19	
93803	7lPN2DXiMsVn7XUKtOW1CS	drivers license	99	242014	1	['Olivia Rodrigo']	['1McMsnEElThX1knmY4oliG']	2021-01-08	
93804	3Ofmpyhv5UAQ70mENzB277	Astronaut In The Ocean	98	132780	0	['Masked Wolf']	['1uU7g3DNSbsu0QjSEqZtEd']	2021-01-06	
92810	5QO79kh1waicV47BqGRL3g	Save Your Tears	97	215627	1	['The Weeknd']	['1Xyo4u8uXC1ZmMpatF05PJ']	2020-03-20	
92811	6tDDoYIxWvMLTdKpjFkc1B	telepatía	97	160191	0	['Kali Uchis']	['1U1el3k54VvEUzo3ybLPlM']	2020-12-04	
92813	0VjIjW4GlUZAMYd2vXMi3b	Blinding Lights	96	200040	0	['The Weeknd']	['1Xyo4u8uXC1ZmMpatF05PJ']	2020-03-20	
93805	7MAibcTli4IisCtbHKrGMh	Leave The Door Open	96	242096	0	['Bruno Mars', 'Anderson .Paak', 'Silk Sonic']	['0du5cEVh5yTK9QJze8zA0C', '3jK9MiCrA42lLAdMGU...	2021-03-05	
92814	6f3Slt0GbA2bPZlz0aIFXN	The Business	95	164000	0	['Tiësto']	['2o5jDhtHVPhrJdv3cEQ99Z']	2020-09-16	
91866	60ynsPSSKe6O3sfwRnIBRf	Streets	94	226987	1	['Doja Cat']	['5cj0lLjcoR7YOSnhnX0Po5']	2019-11-07	
92816	3FAJ6O0NOHQV8Mc5Ri6ENp	Heartbreak Anniversary	94	198371	0	['Giveon']	['4fxd5Ee7UefO4CUXgwJ7IP']	2020-03-27	

In [9]:
```
df_tracks.set_index("release_date",inplace=True)
df_tracks.index=pd.to_datetime(df_tracks.index)
df_tracks.head()
```

Out[9]:

release_date	id	name	popularity	duration_ms	explicit	artists	id_artists	dancea…
1922-02-22	35iwgR4jXetI318WEWsa1Q	Carve	6	126903	0	['Uli']	['45tIt06XoI0Iio4LBEVpls']	0.6
1922-06-01	021ht4sdgPcrDgSk7JTbKY	Capítulo 2.16 - Banquero Anarquista	0	98200	0	['Fernando Pessoa']	['14jtPCOoNZwquk5wd9DxrY']	0.6
1922-03-21	07A5yehtSnoedViJAZkNnc	Vivo para Quererte - Remasterizado	0	181640	0	['Ignacio Corsini']	['5LiOoJbxVSAMkBS2fUm3X2']	0.4
1922-03-21	08FmqUhxtyLTn6pAh6bk45	El Prisionero - Remasterizado	0	176907	0	['Ignacio Corsini']	['5LiOoJbxVSAMkBS2fUm3X2']	0.3
1922-01-01	08y9GfoqCWfOGsKdwojr5e	Lady of the Evening	0	163080	0	['Dick Haymes']	['3BiJGZsyX9sJchTqcSA7Su']	0.4

In [10]:
```
df_tracks[["artists"]].iloc[18]
```

Out[10]:
```
artists    ['Victor Boucher']
Name: 1922-01-01 00:00:00, dtype: object
```

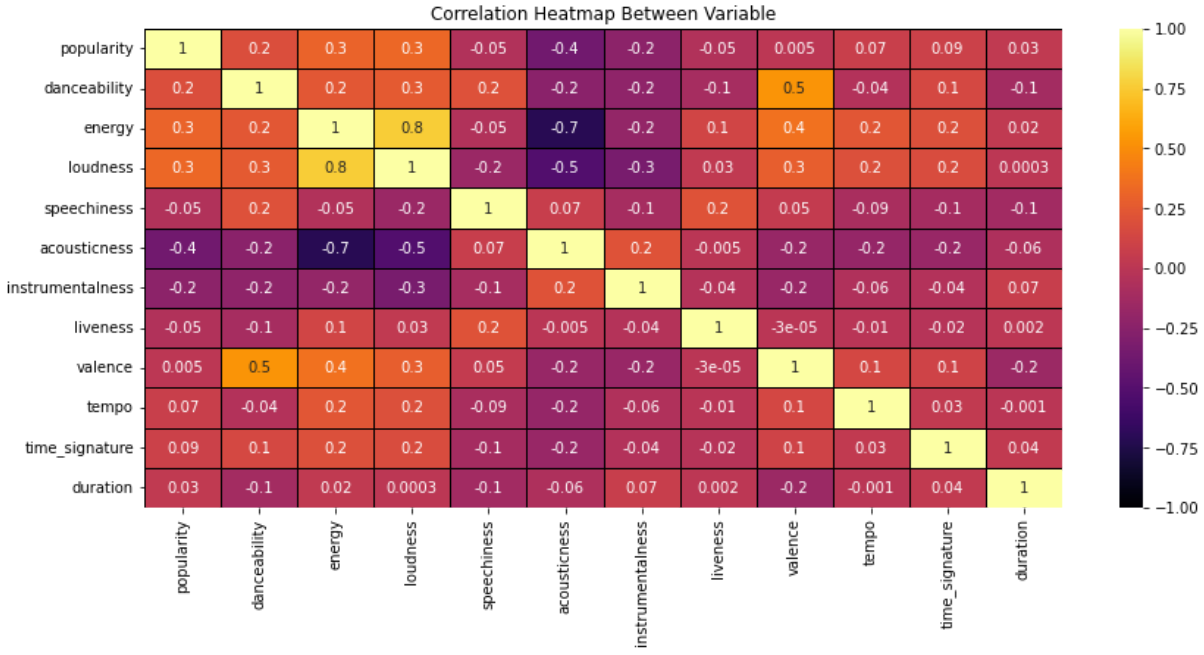
In [11]:
```
df_tracks["duration"]=df_tracks["duration_ms"].apply(lambda x: round(x/1000))
df_tracks.drop("duration_ms",inplace=True, axis=1)
```

In [12]:
```
df_tracks.duration.head()
```

Out[12]:
```
release_date
1922-02-22    127
1922-06-01     98
1922-03-21    182
1922-03-21    177
1922-01-01    163
Name: duration, dtype: int64
```

In [13]:
```
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=1, li
heatmap.set_title("Correlation Heatmap Between Variable")
heatmap.set_xticklabels(heatmap.get_xticklabels(),rotation=90)
```

Out[13]:
```
[Text(0.5, 0, 'popularity'),
 Text(1.5, 0, 'danceability'),
 Text(2.5, 0, 'energy'),
 Text(3.5, 0, 'loudness'),
 Text(4.5, 0, 'speechiness'),
 Text(5.5, 0, 'acousticness'),
 Text(6.5, 0, 'instrumentalness'),
 Text(7.5, 0, 'liveness'),
 Text(8.5, 0, 'valence'),
 Text(9.5, 0, 'tempo'),
 Text(10.5, 0, 'time_signature'),
 Text(11.5, 0, 'duration')]
```



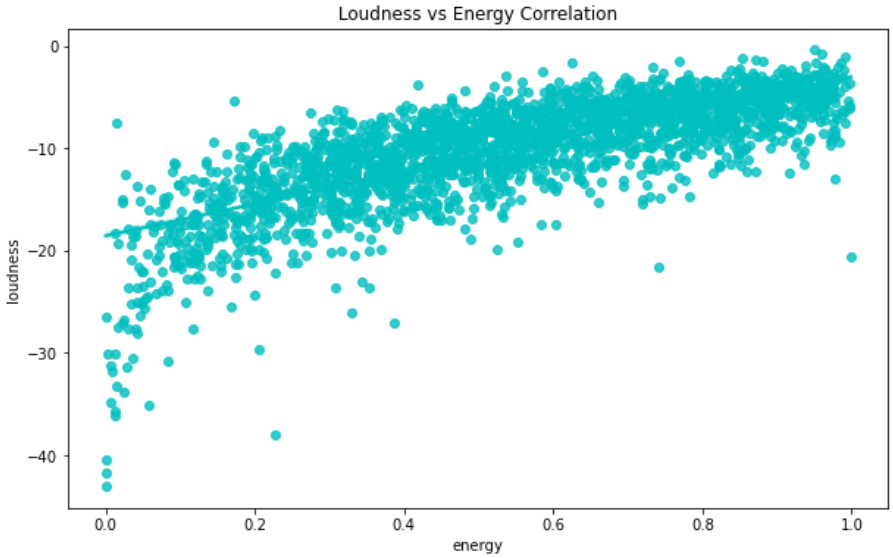
In [14]:
```
sample_df=df_tracks.sample(int(0.004*len(df_tracks)))
```

In [15]:
```
print(len(sample_df))
```

```
2346
```

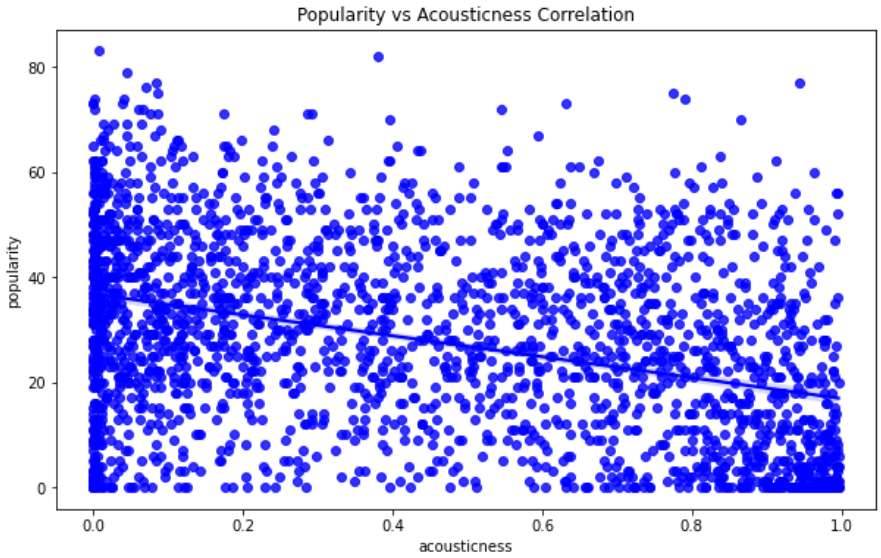
In [16]:
```
plt.figure(figsize=(10,6))
sns.regplot(data=sample_df, y="loudness",x="energy",color="c").set(title="Loudness vs Energy Correlation")
```

Out[16]:
```
[Text(0.5, 1.0, 'Loudness vs Energy Correlation')]
```



In [17]:
```
1 plt.figure(figsize=(10,6))
2 sns.regplot(data=sample_df, y="popularity",x="acousticness",color="b").set(title="Popularity vs Acousticnes
```

Out[17]: [Text(0.5, 1.0, 'Popularity vs Acousticness Correlation')]



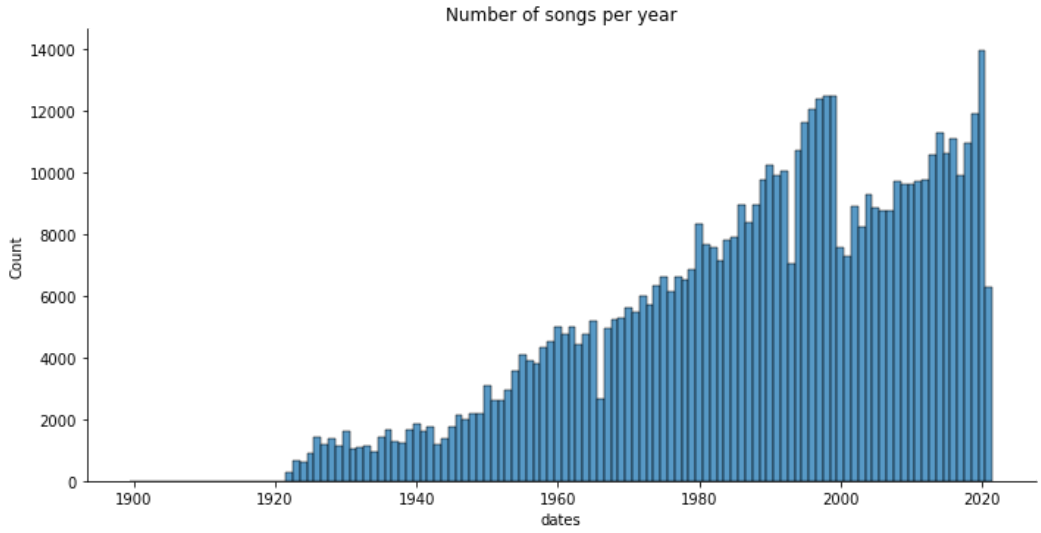
In [18]:
```
1 df_tracks['dates']=df_tracks.index.get_level_values('release_date')
2 df_tracks.dates=pd.to_datetime(df_tracks.dates)
3 years=df_tracks.dates.dt.year
```

In [24]:
```
1 #pip install --user seaborn==0.11.0
```

```
Collecting seaborn==0.11.0
  Downloading seaborn-0.11.0-py3-none-any.whl (283 kB)
Requirement already satisfied: scipy>=1.0 in c:\users\prashant pandey\anaconda3\lib\site-packages (from seabor
n==0.11.0) (1.7.1)
Requirement already satisfied: numpy>=1.15 in c:\users\prashant pandey\anaconda3\lib\site-packages (from seabo
rn==0.11.0) (1.22.4)
Requirement already satisfied: pandas>=0.23 in c:\users\prashant pandey\anaconda3\lib\site-packages (from seab
orn==0.11.0) (1.3.4)
Requirement already satisfied: matplotlib>=2.2 in c:\users\prashant pandey\anaconda3\lib\site-packages (from s
eaborn==0.11.0) (3.4.3)
Requirement already satisfied: cycler>=0.10 in c:\users\prashant pandey\anaconda3\lib\site-packages (from matp
lotlib>=2.2->seaborn==0.11.0) (0.10.0)
Requirement already satisfied: pillow>=6.2.0 in c:\users\prashant pandey\anaconda3\lib\site-packages (from mat
plotlib>=2.2->seaborn==0.11.0) (8.4.0)
Requirement already satisfied: kiwisolver>=1.0.1 in c:\users\prashant pandey\anaconda3\lib\site-packages (from
matplotlib>=2.2->seaborn==0.11.0) (1.3.1)
Requirement already satisfied: python-dateutil>=2.7 in c:\users\prashant pandey\anaconda3\lib\site-packages (f
rom matplotlib>=2.2->seaborn==0.11.0) (2.8.2)
Requirement already satisfied: pyparsing>=2.2.1 in c:\users\prashant pandey\anaconda3\lib\site-packages (from
matplotlib>=2.2->seaborn==0.11.0) (3.0.4)
Requirement already satisfied: six in c:\users\prashant pandey\anaconda3\lib\site-packages (from cycler>=0.10-
>matplotlib>=2.2->seaborn==0.11.0) (1.16.0)
Requirement already satisfied: pytz>=2017.3 in c:\users\prashant pandey\anaconda3\lib\site-packages (from pand
as>=0.23->seaborn==0.11.0) (2021.3)
Installing collected packages: seaborn
Successfully installed seaborn-0.11.0
Note: you may need to restart the kernel to use updated packages.
```

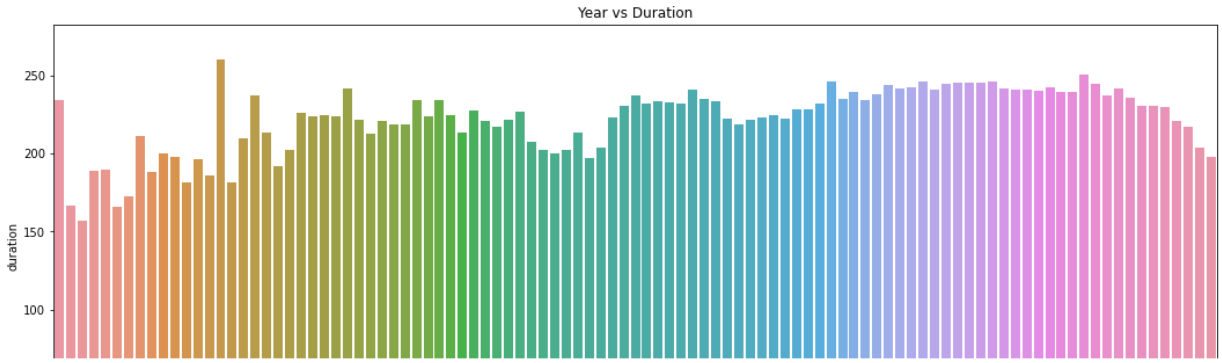
In [20]:
```
sns.displot(years,discrete=True, aspect=2,height=5,kind="hist").set(title="Number of songs per year")
```

Out[20]: <seaborn.axisgrid.FacetGrid at 0x20005828250>



In [21]:
```
total_dr=df_tracks.duration
fig_dims=(18,7)
fig,ax =plt.subplots(figsize=fig_dims)
fig=sns.barplot(x=years,y=total_dr, ax=ax, errwidth=False).set(title="Year vs Duration")
plt.xticks(rotation=90)
```

```
Text(96, 0, '2017'),
Text(97, 0, '2018'),
Text(98, 0, '2019'),
Text(99, 0, '2020'),
Text(100, 0, '2021')])
```



In [27]:
```
df_genre=pd.read_csv("spotifyFeatures.csv")
```

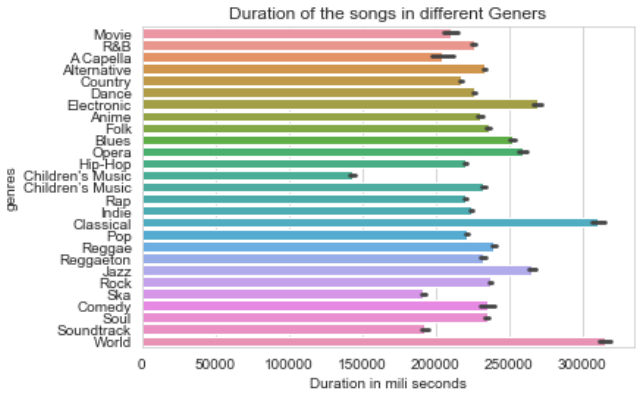
In [28]:
```
df_genre.head()
```

Out[28]:

	genre	artist_name	track_name	track_id	popularity	acousticness	danceability	duration_ms	energy	instrumentalnes
0	Movie	Henri Salvador	C'est beau de faire un Show	0BRjO6ga9RKCKjfDqeFgWV	0	0.611	0.389	99373	0.910	0.0(
1	Movie	Martin & les fées	Perdu d'avance (par Gad Elmaleh)	0BjC1NfoEOOusryehmNudP	1	0.246	0.590	137373	0.737	0.0(
2	Movie	Joseph Williams	Don't Let Me Be Lonely Tonight	0CoSDzoNIKCRs124s9uTVy	3	0.952	0.663	170267	0.131	0.0(
3	Movie	Henri Salvador	Dis-moi Monsieur Gordon Cooper	0Gc6TVm52BwZD07Ki6tIvf	0	0.703	0.240	152427	0.326	0.0(
4	Movie	Fabien Nataf	Ouverture	0IuslXpMROHdEPvSl1fTQK	4	0.950	0.331	82625	0.225	0.1(

In [31]:
```
plt.title("Duration of the songs in different Geners")
sns.color_palette("rocket", as_cmap=True)
sns.barplot(y='genre',x='duration_ms', data=df_genre)
plt.xlabel("Duration in mili seconds")
plt.ylabel("genres")
```

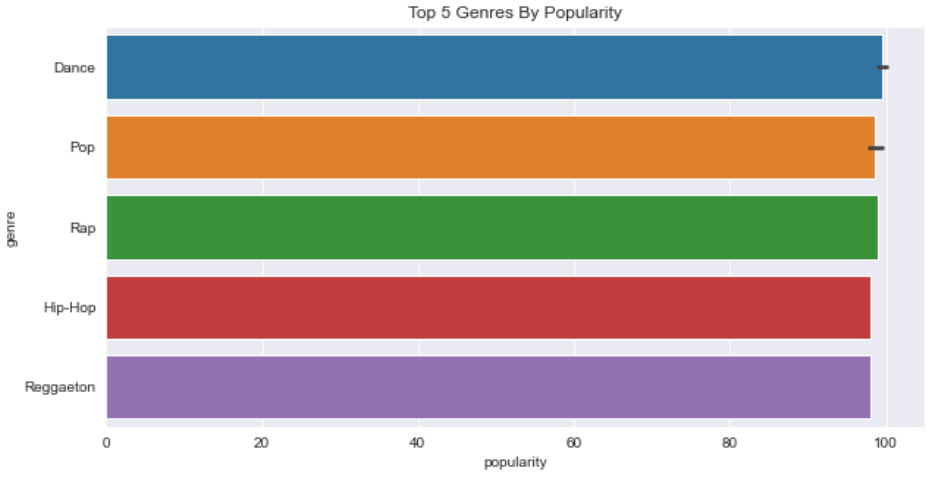
Out[31]: Text(0, 0.5, 'genres')



In [34]:
```
sns.set_style(style="darkgrid")
plt.figure(figsize=(10,5))
famous=df_genre.sort_values("popularity", ascending=False).head(10)
sns.barplot(y='genre',x='popularity', data=famous).set(title="Top 5 Genres By Popularity")

```

Out[34]: [Text(0.5, 1.0, 'Top 5 Genres By Popularity')]



In [ ]:
```
1
```