

```
In [ ]: import pandas as pd

data = pd.read_csv("csv/new_dataset.csv")
data.head()
```

```
Out[ ]:
```

	track_id	artists	track_name	popularity	duration_ms	explici
0	5SuOikwiRyPMVoIQDJUGSV	Gen Hoshino	Comedy	73	230666	Fals
1	4qPNDBW1i3p13qLCt0Ki3A	Ben Woodward	Ghost - Acoustic	55	149610	Fals
2	1iJBSr7s7jYXzM8EGcbK5b	Ingrid Michaelson;ZAYN	To Begin Again	57	210826	Fals
3	6lfxq3CG4xtTiEg7opyCyx	Kina Grannis	Can't Help Falling In Love	71	201933	Fals
4	5vjLSffimiIP26QG5WcN2K	Chord Overstreet	Hold On	82	198853	Fals

```
In [ ]: data.describe()
```

```
Out[ ]:
```

	popularity	duration_ms	danceability	energy	key	lou
count	114000.000000	1.140000e+05	114000.000000	114000.000000	114000.000000	114000.0
mean	33.238535	2.280292e+05	0.566800	0.641383	5.309140	-8.2
std	22.305078	1.072977e+05	0.173542	0.251529	3.559987	5.0
min	0.000000	0.000000e+00	0.000000	0.000000	0.000000	-49.5
25%	17.000000	1.740660e+05	0.456000	0.472000	2.000000	-10.0
50%	35.000000	2.129060e+05	0.580000	0.685000	5.000000	-7.0
75%	50.000000	2.615060e+05	0.695000	0.854000	8.000000	-5.0
max	100.000000	5.237295e+06	0.985000	1.000000	11.000000	4.5

Modify / Clean Dataset

```
In [ ]: #data = data.drop('track_name',axis=1)
data
```

Out []:

	track_id	artists	popularity	duration_ms	explicit	dan
0	5SuOikwiRyPMVoIQDJUgSV	Gen Hoshino	73	230666	False	
1	4qPNDBW1i3p13qLCt0Ki3A	Ben Woodward	55	149610	False	
2	1iJBSr7s7jYXzM8EGcbK5b	Ingrid Michaelson;ZAYN	57	210826	False	
3	6lfxq3CG4xtTiEg7opyCyx	Kina Grannis	71	201933	False	
4	5vjLSffimiIP26QG5WcN2K	Chord Overstreet	82	198853	False	
...
113995	2C3TZjDRiAzdyViavDJ217	Rainy Lullaby	21	384999	False	
113996	1hlz5L4IB9hN3WRYPOCGPw	Rainy Lullaby	22	385000	False	
113997	6x8ZfSoqDjuNa5SVP5QjvX	Cesária Evora	22	271466	False	
113998	2e6sXL2bYv4bSz6VTdnfLs	Michael W. Smith	41	283893	False	
113999	2hETkH7cOfqzmz3LqZDHzf5	Cesária Evora	22	241826	False	

114000 rows × 18 columns

In []: %pwd

In []: path = '/root/AAI-540/Module2/csv/new_dataset.csv'
data.to_csv(path,index=False)

In []: data2 = pd.read_csv("csv/new_dataset.csv")
data2.head()

Out []:

	track_id	artists	popularity	duration_ms	explicit	danceabili
0	5SuOikwiRyPMVoIQDJUgSV	Gen Hoshino	73	230666	False	0.6
1	4qPNDBW1i3p13qLCt0Ki3A	Ben Woodward	55	149610	False	0.4
2	1iJBSr7s7jYXzM8EGcbK5b	Ingrid Michaelson;ZAYN	57	210826	False	0.4
3	6lfxq3CG4xtTiEg7opyCyx	Kina Grannis	71	201933	False	0.2
4	5vjLSffimiIP26QG5WcN2K	Chord Overstreet	82	198853	False	0.6

In []: %%html

<p>Shutting down your kernel for this notebook to release resources.</p>
<button class="sm-command-button" data-commandlinker-command="kernelmenu:shutdown">

<script>

```
try {  
  els = document.getElementsByClassName("sm-command-button");  
  els[0].click();  
}  
catch(err) {  
  // NoOp  
}  
</script>
```

Shutting down your kernel for this notebook to release resources.

In []: