# ETL-Project

Top Spotify Tracks of 2018 – CSV

<https://www.kaggle.com/nadintamer/top-spotify-tracks-of-2018>

Billboard Top 100 1958-2018

<https://data.world/kcmillersean/billboard-hot-100-1958-2017/workspace/file?filename=Hot+Stuff.csv>

Top Spotify Table

|  |  |  |  |
| --- | --- | --- | --- |
| Song | Artist | Position | Time Duration |
|  |  |  |  |

Billboard Top 100 Table

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Song | Artist | # of Weeks | Avg Position | Peek Position |
|  |  |  |  |  |

Necessary Transformations

* Copy only the columns needed into a new DataFrame.
* Use AVG, Count, Max functions for Billboard data
* Reformat Spotify song names to remove feature artist(s)
* Join on song name
* Rename columns to fit the tables created in the database.

Join Table

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Spotify Position | Song | Artist | # of Billboard weeks | Avg Billboard position | Top Billboard position | Song Duration |
|  |  |  |  |  |  |  |