

Project Datasets

You'll be working with two datasets that reside in S3. Here are the S3 links for each:

- Song data: `s3://udacity-dend/song_data`
- Log data: `s3://udacity-dend/log_data`

Log data json path: `s3://udacity-dend/log_json_path.json`

Song Dataset

The first dataset is a subset of real data from the [Million Song Dataset](#). Each file is in JSON format and contains metadata about a song and the artist of that song. The files are partitioned by the first three letters of each song's track ID. For example, here are filepaths to two files in this dataset.

```
song_data/A/B/C/TRABCE1128F424C983.json
song_data/A/A/B/TRABJL12903CDCF1A.json
```

And below is an example of what a single song file, TRAABJL12903CDCF1A.json, looks like.

```
{"num_songs": 1, "artist_id": "ARJIE2Y1187B994AB7", "artist_latitude": null, "artist_longitude":
```

Log Dataset

The second dataset consists of log files in JSON format generated by this [event simulator](#) based on the songs in the dataset above. These simulate app activity logs from an imaginary music streaming app based on configuration settings.

The log files in the dataset you'll be working with are partitioned by year and month. For example, here are filepaths to two files in this dataset.

```
log_data/2018/11/2018-11-12-events.json
log_data/2018/11/2018-11-13-events.json
```

And below is an example of what the data in a log file, 2018-11-12-events.json, looks like.

	artist	auth	firstName	gender	BirthDate	lastName	length	level	location	method	page	registration	sessionId	song	status	ts	userAgent	userId
0	None	Logged In	Celeste	F	0	Williams	None	free	Klamath Falls, OR	GET	Home	1.5410109e+12	438	None	200	1541990217795	"Mozilla/5.0 (Windows NT 6.1; WOW64) AppleWebKit...	53
1	Paramore	Logged In	Sylvia	F	0	Crut	99.16036	free	Washington-- Arlington, Alexandria, DC-MD-VA	PUT	NextSong	1.5410288e+12	345	Major/Two Landed/one	200	1541990258795	"Mozilla/5.0 (Macintosh; Intel Mac OS X 10_10_4) AppleWebKit...	10
2	Berry Tuckwell@University of St. Martin-in-the-Pa...	Logged In	Celeste	F	1	Williams	271715873	free	Klamath Falls, OR	PUT	NextSong	1.5410109e+12	438	Horn Concerto No. 4 in E-flat major, Op. 34, No. 1, Beethoven...	200	1541990254795	"Mozilla/5.0 (Windows NT 6.1; WOW64) AppleWebKit...	53
3	Gary Allen	Logged In	Celeste	F	2	Williams	27122567	free	Klamath Falls, OR	PUT	NextSong	1.5410109e+12	438	Nothing On But The Radio	200	1541990254795	"Mozilla/5.0 (Windows NT 6.1; WOW64) AppleWebKit...	53
4	None	Logged In	Jacqueline	F	0	Lynch	None	paid	Atlanta-Randy Berry/Roswell, GA	GET	Home	1.5402224e+12	389	None	200	1541990714795	"Mozilla/5.0 (Macintosh; Intel Mac OS X 10_10_4) AppleWebKit...	29