



# Million Song Dataset

Home

## Field list

Submitted by millionsong on Fri, 01/14/2011 - 14:37

Below are a list of all fields available in the files of the dataset. The same list with data from a specific song is available [here](#). Another reference is the code: [display\\_song.py](#): if a field is displayed, the field exists and there should be a getter for it (if we forgot some in matlab or java, please let us know).  
For the analysis fields, we suggest you first read The Echo Nest [analyze documentation](#). The main audio features are 'segments\_pitches' and 'segments\_timbre'.

Field name	Type	Description	Link
<del>analysis_sample_rate</del>	float	sample rate of the audio used	<a href="#">url</a>
<del>artist_7digitalid</del>	int	ID from 7digital.com or -1	<a href="#">url</a>
artist familiarity	float	algorithmic estimation	<a href="#">url</a>
artist hottnesss	float	algorithmic estimation	<a href="#">url</a>
<del>artist_id</del>	string	Echo Nest ID	<a href="#">url</a>
artist latitude	float	latitude	
artist location	string	location name	
artist longitude	float	longitude	
<del>artist_mbid</del>	string	ID from musicbrainz.org	<a href="#">url</a>
artist mbtags	array string	tags from musicbrainz.org	<a href="#">url</a>
artist mbtags count	array int	tag counts for musicbrainz tags	<a href="#">url</a>
artist name	string	artist name	<a href="#">url</a>
<del>artist_playmeid</del>	int	ID from playme.com, or -1	<a href="#">url</a>
artist terms	array string	Echo Nest tags	<a href="#">url</a>
artist terms freq	array float	Echo Nest tags freqs	<a href="#">url</a>
artist terms weight	array float	Echo Nest tags weight	<a href="#">url</a>
<del>audio_md5</del>	string	audio hash code	
<del>bars_confidence</del>	array float	confidence measure	<a href="#">url</a>
bars start	array float	beginning of bars, usually on a beat	<a href="#">url</a>
<del>beats_confidence</del>	array float	confidence measure	<a href="#">url</a>
beats start	array float	result of beat tracking	<a href="#">url</a>
danceability	float	algorithmic estimation	
duration	float	in seconds	
<del>end_of_fade_in</del>	float	seconds at the beginning of the song	<a href="#">url</a>
energy	float	energy from listener point of view	
key	int	key the song is in	<a href="#">url</a>
key confidence	float	confidence measure	<a href="#">url</a>
loudness	float	overall loudness in dB	<a href="#">url</a>
mode	int	major or minor	<a href="#">url</a>
mode confidence	float	confidence measure	<a href="#">url</a>
release	string	album name	
<del>release_7digitalid</del>	int	ID from 7digital.com or -1	<a href="#">url</a>
<del>sections_confidence</del>	array float	confidence measure	<a href="#">url</a>
sections start	array float	largest grouping in a song, e.g. verse	<a href="#">url</a>
<del>segments_confidence</del>	array float	confidence measure	<a href="#">url</a>
segments loudness max	array float	max dB value	<a href="#">url</a>
segments loudness max time	array float	time of max dB value, i.e. end of attack	<a href="#">url</a>
segments loudness max start	array float	dB value at onset	<a href="#">url</a>
segments pitches	2D array float	chroma feature, one value per note	<a href="#">url</a>
segments start	array float	musical events, ~ note onsets	<a href="#">url</a>
segments timbre	2D array float	texture features (MFCC+PCA-like)	<a href="#">url</a>
similar artists	array string	Echo Nest artist IDs (sim. algo. unpublished)	<a href="#">url</a>
song hottnesss	float	algorithmic estimation	
<del>song_id</del>	string	Echo Nest song ID	
<del>start_of_fade_out</del>	float	time in sec	<a href="#">url</a>
tatums confidence	array float	confidence measure	<a href="#">url</a>
tatums start	array float	smallest rythmic element	<a href="#">url</a>
tempo	float	estimated tempo in BPM	<a href="#">url</a>
time signature	int	estimate of number of beats per bar, e.g. 4	<a href="#">url</a>
time signature confidence	float	confidence measure	<a href="#">url</a>
title	string	song title	
<del>track_id</del>	string	Echo Nest track ID	
<del>track_7digitalid</del>	int	ID from 7digital.com or -1	<a href="#">url</a>
year	int	song release year from MusicBrainz or 0	<a href="#">url</a>

### News

- April 25, 2012**  
The [MSD Challenge](#) has launched!
- October 20, 2011**  
We release the [Last.fm](#) dataset of tags and similarity!
- April 12, 2011**  
We release the [musiXmatch](#) dataset of lyrics!
- March 15, 2011**  
We release the [SecondHandSongs](#) dataset of cover songs!
- February 8, 2011**  
We release the dataset!  
(and get Dan to [blog](#))

### Quick links

- [LabROSA](#)
- [The Echo Nest](#)
- [Musicbrainz](#)
- [Infochimps](#)
- [7digital](#)
- [Last.fm](#)
- [musiXmatch](#)
- [SecondHandSongs](#)

### Main contact

- [MSD mailing list](#)