

## bp\_artist.csv

### About this file

PRIMARY KEY: **artist\_id**

ADDITIONAL INFO:

- **artist\_id** referenced by tables **bp\_artist\_media**, **bp\_artist\_release**, **bp\_artist\_track** & **bp\_label\_artist**
  - **artist\_slug** is a portion of url query parameter to search the artist on beatport
  - **updated\_on** column is in PST
- 

## bp\_artist\_media.csv

### About this file

PRIMARY KEY: (**artist\_id**, **artist\_img\_id**)

FOREIGN KEY: **artist\_id** references **bp\_artist**

ADDITIONAL INFO:

- an artist may have more than 1 associated image. Although Beatport does not display more than 1 at any given time, historical images are stored in this table. This is done by design. To use the most up to date image, use most current **updated\_on** column when grouping by **artist\_id**
  - composite primary key
  - one-to-many relationship between artist and artist artwork
  - **artist\_img\_id** = 5539565 represents beatport's stock image for when artists have not provided an image of themselves.
  - The corresponding **artist\_img** is <https://geo-media.beatport.com/image/size/590x404/0dc61986-bccf-49d4-8fad-6b147ea8f327.jpg>
  - **updated\_on** column is in PST
- 

## bp\_artist\_release.csv

### About this file

PRIMARY KEY: (**release\_id**, **artist\_id**)

FOREIGN KEY:

1. **artist\_id** references **bp\_artist**
2. **release\_id** references **bp\_release**

**ADDITIONAL INFO:**

- **updated\_on** column is in PST
  - composite primary key
  - many-to-many relationship between artist and release
  - the significant decrease in count of **release\_id** between 400,000 - 800,000 should not be concerning. The majority of these releases have either been permanently removed or hidden from view on Beatport
- 

## bp\_artist\_track.csv

**About this file**

**PRIMARY KEY:** (**track\_id**, **artist\_id**, **is\_remixer**)

**FOREIGN KEY:**

1. **track\_id** references **bp\_track**
2. **release\_id** references **bp\_release**

**ADDITIONAL INFO:**

- an artist can be both an original artist and remixer of a single track
  - many-to-many relationship between tracks and artists.
  - the significant decrease in count of **track\_id** between 2 - 2.8 M should not be concerning. The majority of those tracks have either been permanently removed or hidden from view on Beatport.
- 

## bp\_genre.csv

**About this file**

**PRIMARY KEY:** **genre\_id**

**ADDITIONAL INFO:**

- **updated\_on** column is in PST
- **genre\_id** referenced by tables **bp\_track** & **bp\_subgenre**
- **song\_count** reflects the current song count per genre as of **updated\_on** as indicated from Beatport.com . It is not exactly representative of the number of songs in the database. If you were to view the change in **song\_count** over time, you might notice a net decrease when in reality there would be a net increase. This is due in part to songs being removed from

Beatport, **release\_date** getting changed, songs being hidden from view, or the genre being changed.

---

## bp\_key.csv

### About this file

PRIMARY KEY: **key\_id**

ADDITIONAL INFO:

- **updated\_on** column is in PST
  - **key\_id** referenced by table **bp\_track**
- 

## bp\_label.csv

### About this file

PRIMARY KEY: **label\_id**

ADDITIONAL INFO:

- **updated\_on** column is in PST
  - **label\_id** referenced by tables **bp\_label\_media**, **bp\_label\_artist**, **bp\_release** & **bp\_track**
- 

## bp\_label\_artist.csv

### About this file

PRIMARY KEY: (**label\_id**, **artist\_id**)

FOREIGN KEY:

1. **label\_id** references **bp\_label**
2. **artist\_id** references **bp\_artist**

ADDITIONAL INFO:

- **updated\_on** column is in PST
  - many-to-many relationship between labels and artists
  - composite primary key
-

## bp\_label\_media.csv

### About this file

PRIMARY KEY: (label\_id, label\_img\_id)

FOREIGN KEY: label\_id references bp\_label

#### ADDITIONAL INFO:

- standard image size is 500x500
  - a label may have more than 1 associated image. Although Beatport does not display more than 1 at any given time, historical images are stored in this table. This is done by design. To use the most up to date image, use most current updated\_on column when grouping by label\_id
  - composite primary key
  - label\_img\_id = 5539566 represents beatport's stock image for when labels have not provided an image to represent themselves.
  - The corresponding label\_img is [https://geo-media.beatport.com/image\\_size/500x500/cda9862c-cf92-4d13-ac65-7e9277181f51.jpg](https://geo-media.beatport.com/image_size/500x500/cda9862c-cf92-4d13-ac65-7e9277181f51.jpg)
  - updated\_on column is in PST
- 

## bp\_release.csv

### About this file

PRIMARY KEY: release\_id

FOREIGN KEY: label\_id references bp\_label

#### ADDITIONAL INFO:

- updated\_on column is in PST
  - release\_id referenced by tables bp\_artist\_release, bp\_release\_media, bp\_track\_media & bp\_track
  - the significant decrease in count of release\_id between 400,000 - 800,000 should not be concerning. The majority of these releases have either been permanently removed or hidden from view on Beatport
  - catalog doesn't appear to be used outside of Beatport.
  - might add columns upc (international release/album identifier), bpm\_range, release\_desc (description), track\_count
-

## bp\_release\_media.csv

### About this file

PRIMARY KEY: (release\_id, release\_img\_id)

FOREIGN KEY: release\_id references bp\_release

#### ADDITIONAL INFO:

- standard image size is 1400x1400
  - a release may have more than 1 associated image. Although Beatport does not display more than 1 at any given time, historical images are stored in this table. This is done by design. To use the most up to date image, use most current updated\_on column when grouping by release\_id
  - composite primary key
  - release\_img\_id = 5245821 represents beatport's stock image for when a release does not have any album artwork.
  - The corresponding release\_img is [https://geo-media.beatport.com/image\\_size/500x500/ab2d1d04-233d-4b08-8234-9782b34dcab8.jpg](https://geo-media.beatport.com/image_size/500x500/ab2d1d04-233d-4b08-8234-9782b34dcab8.jpg)
  - updated\_on column is in PST
- 

## bp\_subgenre.csv

### About this file

PRIMARY KEY: (genre\_id, subgenre\_id)

FOREIGN KEY: genre\_id references bp\_genre

#### ADDITIONAL INFO:

- composite primary key
  - there are a couple instances where the same subgenre belongs to 2 genres.
  - updated\_on column is in PST
  - composite (genre\_id, subgenre\_id) referenced by table bp\_track
- 

## bp\_track.csv

PRIMARY KEY: track\_id

FOREIGN KEY:

1. genre\_id references bp\_genre

2. **(genre\_id, subgenre\_id)** references **bp\_subgenre**
3. **key\_id** references **bp\_key**
4. **release\_id** references **bp\_release**
5. **label\_id** references **bp\_label**

#### ADDITIONAL INFO:

- **isrc** is potentially the single most important column in the entire database. The **isrc** (International Standard Recording Code) is used to find exact matches through Spotify API and ultimately link this Beatport Database to the MusicBrainz and Discogs databases.
- the significant decrease in count of **track\_id** between 2 - 4 M should not be concerning. The majority of those tracks have either been permanently removed or hidden from view on Beatport.
- bool value in **is\_remixed** does not necessarily indicate whether or not a track is remixed as there are many instances of a track being remixed with no remixers listed on Beatport. It more so references that a remixer was listed on the track i.e. the column **is\_remixer** in table **bp\_artist\_track** will be set to True.
- **release\_date** can vary between tracks with the same **release\_id**. This indicates an album/release where tracks came out in different phases, i.e. track being released as a single prior to the final album release date.
- **bpm** column must be taken with a grain of salt. **bpm** can be difficult to gauge for a song where it varies wildly. Another consideration is that some genres are more prone to having their **bpm** listed in either double-time or halftime which will either appear as twice or half of the standard **bpm** respectively.
- some instances where the **duration** column represented as a string in HH:MM:SS format is invalid, i.e. minutes or seconds will exceed '59'. These values have been left unmodified.
- **catalog** seems to be some sort of internal cataloging system between record labels and Beatport. I have found no clear indication that the same catalog numbers are used elsewhere.
- **label\_track\_id** is largely irrelevant considering it is used only 60% of the time. Might consider dropping in future updates. It appears to be something that is optionally set up between Beatport and a label for the label to be able to id their own tracks on Beatport.
- **is\_matched\_spot** column is a Boolean flag that indicates whether or not I have found a positive id between a Spotify and Beatport. Updates to this column will be coming soon as well as an entire database of Spotify values including audio features.
- **updated\_on** column is in PST

---

## bp\_track\_media.csv

### About this file

PRIMARY KEY: **track\_id**

FOREIGN KEY: **track\_id** references **bp\_track**

#### ADDITIONAL INFO:

- **wave\_img\_id** can be null if no waveform image exists
  - the UUID within a single track's columns **wave\_img** & **sample\_url** can be the same for some tracks but different for others. (Same example: [https://geo-media.beatport.com/image\\_size/1500x250/a98235a0-e6d3-4707-8abe-89ec4e70e828.png](https://geo-media.beatport.com/image_size/1500x250/a98235a0-e6d3-4707-8abe-89ec4e70e828.png) vs. <https://geo-samples.beatport.com/track/a98235a0-e6d3-4707-8abe-89ec4e70e828.LOFI.mp3>)
  - all sample related columns can be null if the mp3 sample does not exist on Beatport
  - **sample\_start** of 0 indicates the song is less than a minute and 30 seconds in length
  - standard image size is 1500x250
  - a track may NOT have more than 1 associated waveform image at any given time.
  - **updated\_on** column is in PST
-