

1. Сделать таблицу artists в Hive и вставить туда значения, используя датасет

```
create table artists
(mbid string,
artist_mb string,
artist_lastfm string,
country_mb string,
country_lastfm string,
tags_mb array<string>,
tags_lastfm array<string>,
listeners_lastfm double,
scrobbles_lastfm double,
ambiguous_artist boolean)
row format
delimited fields terminated by ','
collection items terminated by ';'
tblproperties ("skip.header.line.count"="1");

load data inpath '/artists.csv' overwrite into table artists;
```

The screenshot shows the Hue web interface running on a local host. The main query editor displays the command: `SELECT * FROM artists LIMIT 10;`. The results pane shows 10 rows of data from the 'artists' table. The columns are: `artists.mbid`, `artists.artist_mb`, `artists.artist_lastfm`, and `artists.` (likely `country_mb`). The data includes entries for Coldplay, Radiohead, Red Hot Chili Peppers, Rihanna, Eminem, The Killers, Kanye West, Nirvana, Muse, and Queen.

	artists.mbid	artists.artist_mb	artists.artist_lastfm	artists.
1	cc197bad-dc9c-440d-a5b5-d52ba2e14234	Coldplay	Coldplay	United Kingdom
2	a74b1b7f-71a5-4011-9441-d0b5e4122711	Radiohead	Radiohead	United Kingdom
3	8bfac288-c0c5-448d-9573-c33ea2aa5c30	Red Hot Chili Peppers	Red Hot Chili Peppers	United States
4	73e5e69d-3554-40d8-8516-00cb38737a1c	Rihanna	Rihanna	United States
5	b95ce3ff-3d05-4e87-9e01-c97b66af13d4	Eminem	Eminem	United States
6	95e1ead9-4d31-4808-a7ac-32c3614c116b	The Killers	The Killers	United States
7	164f0d73-1234-4e2c-8743-d77bf2191051	Kanye West	Kanye West	United States
8	5b11f4ce-a62d-471e-81fc-a69a8278c7da	Nirvana	Nirvana	United States
9	9c9f1380-2516-4fc9-a3e6-f9f61941d090	Muse	Muse	United Kingdom
10	0383daff-2a4e-4d10-a46a-e9e041da8eb3	Queen	Queen	United Kingdom

2. HiveQL

a) Найти исполнителя с максимальным числом скробблов

```
SELECT artists.artist_lastfm
FROM artists
WHERE artists.scrobbles_lastfm in
      (SELECT Max(artists.scrobbles_lastfm) from artists);
```

	artist_lastfm
1	The Beatles

b) Найти самый популярный тэг на ластфм

```
SELECT tag, count(*) tags_count
FROM artists
      LATERAL VIEW explode(tags_lastfm) item_val AS tag
GROUP BY tag
ORDER BY tags_count DESC
LIMIT 1;
```

	tag	tags_count
1	seen live	81278

c) Найти самые популярные исполнители 10 самых популярных тегов ластфм

```
SELECT DISTINCT t1.artist_mb, t1.listeners_lastfm
FROM (SELECT artist_mb, listeners_lastfm, tagg
      FROM artists
            LATERAL VIEW explode(tags_lastfm) artists AS tagg
      WHERE artists.tagg in
            (SELECT tags_table.tag
              from (SELECT tag, count(*) tags_count
                    FROM artists
                          LATERAL VIEW explode(tags_lastfm) item_val AS tag
                    GROUP BY tag
                    ORDER BY tags_count DESC
                    LIMIT 10
              ) tags_table)) t1
ORDER BY t1.listeners_lastfm DESC
LIMIT 10;
```

	artist_mb	listeners_lastfm
1	Coldplay	5381567
2	Radiohead	4732528
3	Red Hot Chili Peppers	4620835
4	Rihanna	4558193
5	Eminem	4517997
6	The Killers	4428868
7	Kanye West	4390502
8	Nirvana	4272894
9	Muse	4089612
10	Queen	4023379

- d) Найти самые популярные теги ластфм в самой музыкальной стране (где больше всего исполнителей)

```
SELECT tag, country_lastfm, count(*) as count
FROM artists
      LATERAL VIEW explode(tags_lastfm) artists AS tag
WHERE artists.country_lastfm in
      (SELECT countries.country_lastfm
        FROM (SELECT country_lastfm, count(*) as count
              FROM artists
              WHERE length(country_lastfm) > 0
              GROUP BY country_lastfm
              ORDER BY count DESC
              LIMIT 1) countries)
GROUP BY tag, country_lastfm
ORDER BY count DESC
LIMIT 10;
```

< < 10 rows > > ↺ ■ ↻ ★			
	tag	country_lastfm	count
1	american	United States	17926
2	USA	United States	12347
3	seen live	United States	9603
4	rock	United States	8177
5	All	United States	7199
6	alternative	United States	5830
7	indie	United States	5219
8	pop	United States	4508
9	experimental	United States	4382
10	electronic	United States	4123