

a6-codeReview-jashangeet-shreysa

Futai-Shikha

October 30, 2017

Code Review:

- We were able to run the code and generate the output.
- In general, they implemented a good system to perform the tasks asked in the assignment. The code is good, easy to understand and very little modifications to be made.
- The Report has all the results for the assignment but it would be better if the code had all the results in the output folder.
- The ignored set of values could have been edited rather than taking the same example from the class notes:

```
val ignored = Set(
    "THAT", "WITH", "THE", "AND", "TO", "OF",
    "A", "IT", "SHE", "HE", "YOU", "IN", "I",
    "HER", "AT", "AS", "ON", "THIS", "FOR",
    "BUT", "NOT", "OR", "ME", "MY", "LP",
    "VERSION", "IS", "ALBUM", "AN", "THE")
```

Could have added some top prepositions, articles and conjunctions to get more accurate results.

- The code is highly organized and implements relevant functions to analyze data sets.
- When they retrieve the loudest, longest, fastest and hottest songs, they made composite keys, and use zip function to combine composite keys and values, which is good to understand and very flexible.

```
val top5loud = songtuple.zip(loud).map { x => (x._1, x._2) }.sortBy(_._2, false).take(5).toList
```

- When they retrieve the top 5 prolific, familiar and hottest artists, they shouldn't use composite keys (artistId, artistName), because one artist id indicates one artist and if we combine artist id and artist name we will get different artists name with the same artist id, but they are the same person!

```
var artistuple = records.map{ l => (l.getArtId(), l.getArtistName()) }
```

So, for the following tasks the results are not unique.

```
Top 5 most familiar artists:
Top 5 Hottest artists:
```

The results are something like:

```
Top 5 familiar artists:
```

```
Artist Name Artist Id
Familiarity Score Akon ARCGJ6U1187FB4D01F 1.0
Akon_ San Quinn_ JT the Bigga Figga ARCGJ6U1187FB4D01F 1.0
Akon / Eminem ARCGJ6U1187FB4D01F 1.0
Akon / Styles P ARCGJ6U1187FB4D01F 1.0
Akon / Wyclef Jean ARCGJ6U1187FB4D01F 1.0
```

```
Top 5 hottest Artists:
```

```
Artist Name Artist Id Hotness Score
Daft Punk ARF8HTQ1187B9AE693 0.997066533839045
Daft Punk ARF8HTQ1187B9AE693 0.997004803235357
Black Eyed Peas ARTDQRC1187FB4EFD4 0.982623202516712
Kanye West ARRH63Y1187FB47783 0.972399563931911
Kanye West / Jamie Foxx ARRH63Y1187FB47783 0.97239956393191
```

The artist Id's are not unique for these above results. We compared the results with our results and we believe the correct results are as follows:

artist_id,artist_hottnesss

ARRH63Y1187FB47783,1.0616411086142852 ARF8HTQ1187B9AE693,1.0197351620486825 ARTDQRC1187FB4EFD4,1.00486506
34179903 ARKU3Z61187FB51DCA,0.9208408851893838 ARJ7KF01187B98D717,0.9127957061307619

artist_id,artist_familiarity

ARCGJ6U1187FB4D01F,0.9902253716581705
ARY7WK51187B9B1941,0.976696123205106
ARUDYKB11F4C83C269,0.960481184780221
AREMCGW121318C5844,0.947327506634964
AR4OU721187FB4549D,0.944634995670516