

Sayeed Ahmed – SEC01 (NUID 002191535)

# Big Data System Engineering with Scala

## Spring 2023

### Assignment No. 2 - Spark



## 1) Getting tracks from Spotify:

```
//Getting tracks from Spotify
val token = "BQDPYXrbRYs6TwlvmlJAxkTAJIinZwasUdLmjM06Cq1XDEzKv8EP13N7mr6P1E4nGgOHYWDS4r5B-QGRcyxhp6M6nz60FLYK6FTP-DNXZB74L2PiSaYazVBhe5mWFDx6NDLJfQjZvnXvLN1cdAG0tN7EtS96e1pyvJZ4SY40LGN441PQndQWILjfkWT6KJNG"
val headers = Map("Authorization" -> s"Bearer $token")
val limit = 100
val offsets = List.range(0, 500, limit)
val urlList = offsets.map(offset => s"https://api.spotify.com/v1/playlists/$playlistId/tracks?offset=$offset&limit=$limit")
val responses = urlList.map(url => requests.get(url, headers = headers))
val tracks = responses.flatMap(response => ujson.read(response.text)("items").arr)
```

## 2) Part 1: Fetching top 10 longest songs

```
// Part 1: Fetching top 10 longest songs
val top10Tracks = tracks.sortBy(_("track")("duration_ms").num.toLong).reverse.take(10)
```

Output:

```
1 // Displaying the top 10 longest songs
2 println("Top 10 longest songs:")
3 for (track <- top10Tracks) {
4     val name = track("track")("name").str
5     val duration = track("track")("duration_ms").num.toLong
6     println(s"$name, $duration")
7 }
```

```
Top 10 longest songs:
The End - New Stereo Mix Advanced Resolution, 699533
Desolation Row, 681400
Free Bird, 550066
Purple Rain, 520786
Kashmir - 1990 Remaster, 508200
Iron Man - Live, 500693
Good Times, 495400
Stairway to Heaven - 1990 Remaster, 478173
L.A. Woman, 471160
A Whiter Shade Of Pale - Live, 466506
```

Command took 1.71 seconds -- by ahmed.say@northeastern.edu at 3/14/2023, 2:09:51 AM on test1

## 3) Part 2: Fetching artist details and sorting by follower count

```
//Part 2: Fetching artist details and sorting by follower count
val artistDetails = top10Tracks.flatMap(_("track")("artists").arr)
    .map(artist => {
        val artistId = artist("id").str
        val artistUrl = s"https://api.spotify.com/v1/artists/$artistId"
        val artistResponse = requests.get(artistUrl, headers = headers)
        val followerCount = ujson.read(artistResponse.text)("followers")("total").num
        (artist("name").str, followerCount)
    })
    .distinct
    .sortBy(_._2)
    .reverse
```

Output:

```
1 // Displaying artist details sorted by follower count
2 println("\nArtist details sorted by follower count:")
3 for ((artist, followers) <- artistDetails) {
4     println(s"$artist : $followers followers")
5 }
```

Artist details sorted by follower count:

Led Zeppelin : 1.3151926E7 followers

The Doors : 7106674.0 followers

Black Sabbath : 6808824.0 followers

Prince : 6646834.0 followers

Bob Dylan : 6020952.0 followers

Lynyrd Skynyrd : 4519016.0 followers

CHIC : 646066.0 followers

Procol Harum : 400525.0 followers

The Danish National Concert Orchestra and Choir : 75.0 followers

Command took 1.38 seconds -- by ahmed.say@northeastern.edu at 3/14/2023, 2:09:51 AM on test1