Homework 3: Writing Component
USING MAPREDUCE TO ANALYZE THE MILLION SONG DATASET
VERSION 1.0

DUE DATE: Friday April 19th, 2019 @ 5:00 pm

Q1. Assume that you have complete control to reorganize and distribute the dataset. What is the fewest number of MapReduce jobs that you would need to implement the complete set of tasks in HW3-PC? Note that your program should emit outputs for each task into a separate file.

[300-400 words]

Q2. You are designing a new streaming service that can scale to millions of users and stream songs that are personalized to a user. How would you extend assignments HW1-PC, HW2-PC, and HW3-PC to accomplish this goal?

[400-500 words]

Q3. What if a family has a shared account on your streaming service? How would your streaming service recommend songs so that each family member is satisfied? Please design your own measure for satisfaction.

[300-400 words]

Q4. You are starting a music production company and you are working with local artists in town. How would you perform micro-adjustments to a song so that it is a commercial success in different countries? How would you avoid concept drifts – what is in vogue today may not be in the near future?

[300-400 words]

Q5. The kind of music a person likes may change/evolve over time. Describe a potential scheme to design "interest trajectories" that allows you to recommend songs that a person may like in the future.

[300-400 words]

1   Grading

Homework 3 accounts for 20 points towards your final course grade. This written component accounts for 20% of the points set aside for HW3 i.e. this assignment accounts for 4% of your cumulative course grade. This assignment is graded on a 20-point scale with each question accounting for 4points.

2   What to Submit

You should submit a PDF document. Please use the following naming convention: HW3-WC-Firstname-Lastname.pdf.

The folder set aside for this assignment's submission using checkin is **HW3-WC**