*Big Data Hadoop ‘Assignment 5’*

DATASET

111115|222|0|1|0

111113|225|1|0|0

111117|223|0|1|1

111115|225|1|0|0

1st Column - UserId

2nd Column - TrackId

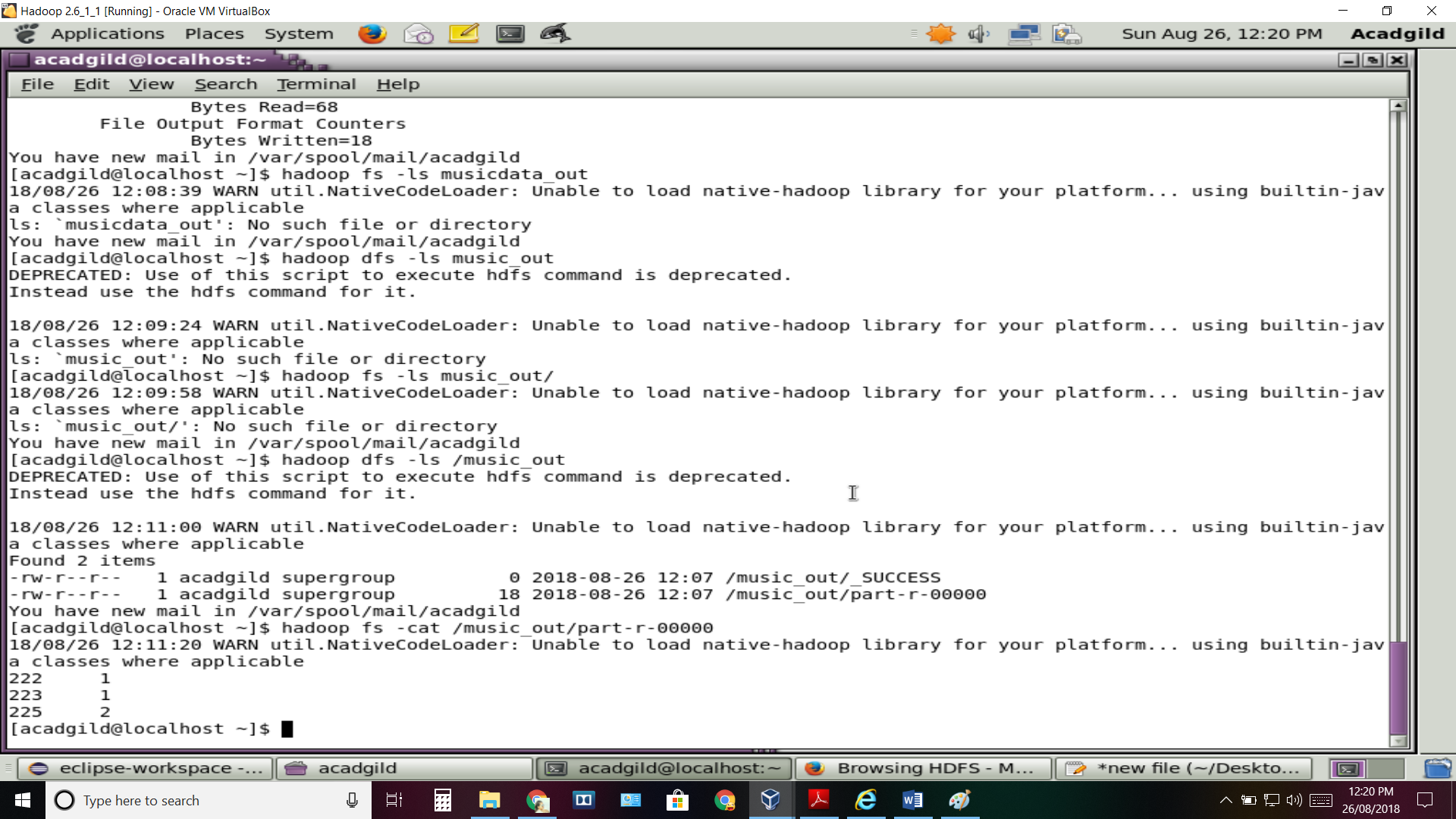
3rd Column - Songs Share status (1 for shared, 0 for not shared)

4th Column - Listening Platform (Radio or Web - 0 for radio, 1 for web)

5th Column - Song Listening Status (0 for skipped, 1 for fully heard)

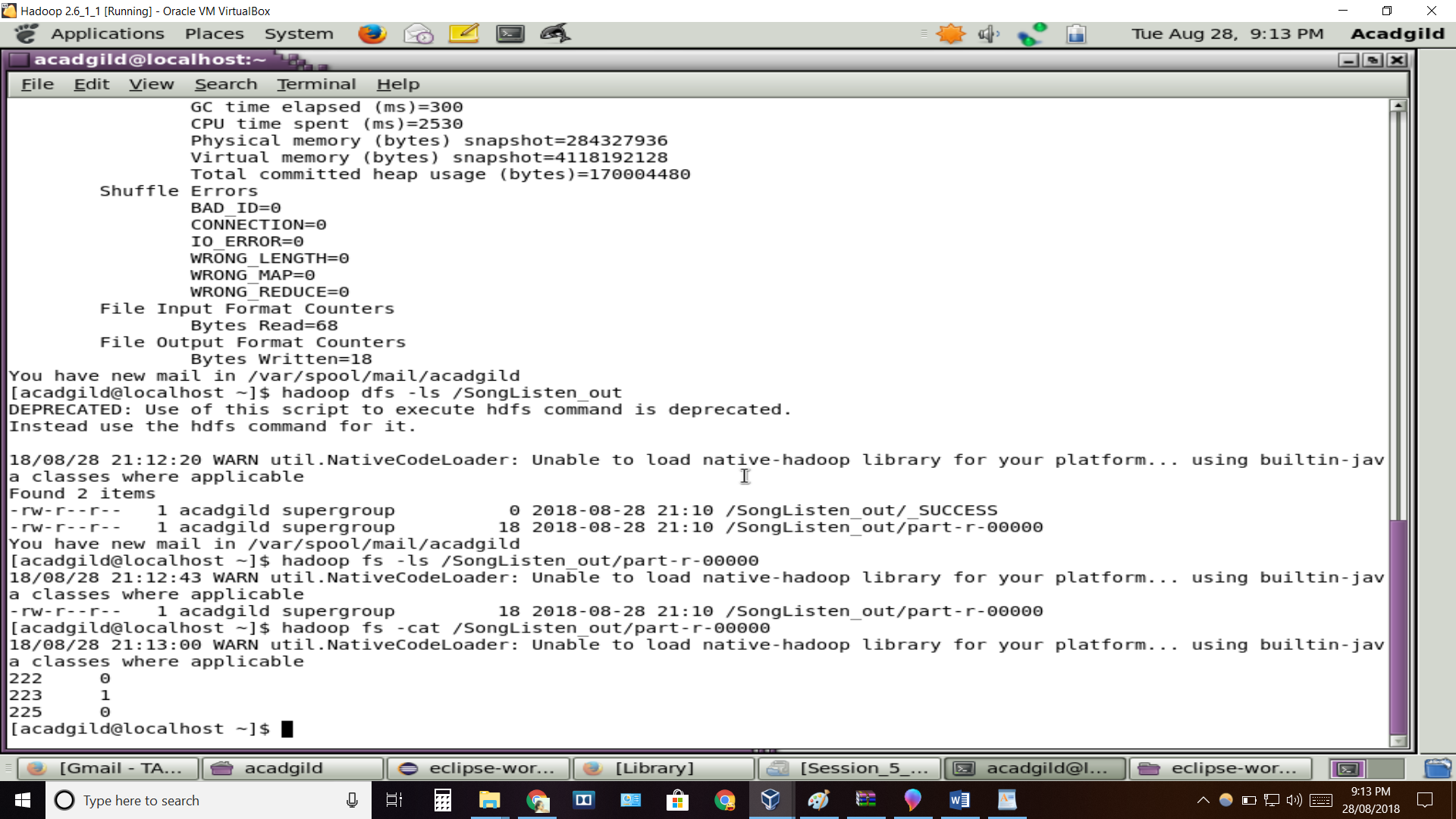
TASK 1: Find the number of unique listeners in the data set.

Output: Listeners are calculated unique per track record



Task 2: What are the number of times a song was heard fully.

Output: No. of times song fully heard per track.



Task 3: What are the number of times a song was shared.

Output: No. of time song is being shared by user.

