PROBLEM STATEMENT: Design a distributed application using MapReduce which processes Music dataset. List out the number of unique listeners and no of times the track was shared with others. Use music dataset and process it using a pseudo distribution mode on Hadoop platform.

Step 1: Create a Java Project in Eclipse

1. **Create a New Java Project**:
    - Open Eclipse and go to **File** → **New** → **Java Project**.
    - Name the project (e.g., `Music`).
    - Click **Finish**.
2. **Create a Package**:
    - In the **Project Explorer**, right-click on `src` → **New** → **Package**.
    - Name the package (e.g., `Music`).
3. **Create the Classes**:
    - Right-click on the package you created → **New** → **Class**.
    - Create the following class:
        - **Music**.java

Step 2: Add the Code Eg. //PS_3_Music.java /// PS4_music.java

Step 3: Build the Project

1. **Build the Project**:
    - Click on **Project** → **Build Project** in Eclipse.
    - Make sure the project compiles without any errors.
2. **Check Build Path**:
    - Go to **Project Explorer** → Right-click on your project → **Build Path** → **Configure Build Path**.
    - Ensure that all the Hadoop JAR files you added are present in the **Libraries** section.
        i. hadoop-common (e.g., `hadoop-common.jar`)

ii. Hadoop-mapreduce-client-core (e.g., `hadoop-mapreduce-client-core-2.x.x.jar`)

Step 4: Create the JAR File

1. **Export to JAR**:
    - Go to **File** → **Export**.
    - Choose **Java** → **JAR file**.
    - Choose **Launch configuration** as `Music`.
    - Select the destination path and name your JAR file (e.g., `Musical.jar`).
    - Click **Finish**.

Step 5: Prepare the Input Files

1. **Import Input File (`Music_dataset.csv`)**:
    - The `Music_dataset.csv` should be placed in the HDFS directory.

    Upload the Input File to HDFS:

- Use the Hadoop shell to upload the input file to HDFS:

hdfs dfs -put /path/to/MusicFile.txt /user/cloudera/MusicFile.txt

Step 6: Configure the Run Configuration

1. **Set up Run Configuration**:
    - In Eclipse, go to **Run** → **Run Configurations**.
    - Select **Java Application** and click **New**.
    - In the **Main** tab, select the **Project** (your Hadoop project) and the **Main Class** (`W`).

Step 8: Run the Job

1. **Run the Hadoop Job**:
   [cloudera@quickstart ~]$ had oop jar /home/cloudera/Musical.jar Music.Music /user/cloudera/MusicFile.txt /user/cloudera/dir51

2. hdfs dfs -ls /user/cloudera/dir51

```
                Bytes Written=30
[cloudera@quickstart ~]$ hdfs dfs -ls /user/cloudera/dir51

Found 2 items
-rw-r--r--   1 cloudera cloudera          0 2025-04-26 23:53 /user/cloudera/dir51/_SUCCESS
-rw-r--r--   1 cloudera cloudera         30 2025-04-26 23:53 /user/cloudera/dir51/part-r-00000
[cloudera@quickstart ~]$
```

3. Hadoop fs -cat /user/cloudera/dir51/part-r-00000