ReadMe

Unzip the delivered assignment submission folder

For NoCombiner

1. Create a maven project
2. Create a package with name mr.assignment2nc
3. Copy paste the provided class files
4. Update the pom.xml with the one provided
5. Add the provided the Makefile in the same location as pom.xml
6. Create an directory by name input
7. Place the 1991.csv file

To run locally:

Open the terminal from the containing Makefile

Run command – make alone

A new Output directory will be created with an output file and a success file

To run on AWS

Open the terminal from the containing Makefile

Create folder input in the AWS bucket specified in Makefile

Upload input files using the command - make upload-input-aws

Run Command – make cloud

Once the processing completes, output files can be downloaded using the command - make download-output-aws

For Combiner/In Mapper Combiner

Repeat the same steps as above except that the package name has to be mr.assignment2c, mr.assignment2IMC respectively

For Secondary Sort

Repeat the same steps as NoCombiner except for the following steps:

2. Create a package by name - mr.assignment2SS

7. Upload all 10 input files - (1880.csv to 1889.csv)