Session 5: Getting Started with MapReduce

Assignment 3

Session 5: Assignment 3

Table of Contents

1. Introduction

This assignment will help you to consolidate the concepts learnt in the session.

2. Problem Statement

We have a dataset of movie data analysis

Column Name:

Column 1: User ID

Column 2: Movie ID

Column 3: Rating

Column 4: Time stamp

The input file can be downloaded from the below link

[Movie data analysis](#)

Problem statement 1:

Find the number of ratings given by each user.

Problem statement 2:

Find the number of ratings received by each movie.

3. Output

N/A