Advanced Big Data Analytics	
Course Code: UE20CS581	Professor: Dr. K.V Subramaniam
Submitted By: Mehvish Hafiz	SRN: PES1PG20CS021
B. Shruthi Laya	SRN: PES1PG20CS011

Assignment: Creating an Inverted Index for Mahabharata text file.

Github Link: https://github.com/mehvish24/InvertedIndexMR

Understanding:

1. Installation of Hadoop & how to start & stop it.
2. Cleaning a text file i.e. removed stop words, punctuation, numbers etc.
3. Splitting the cleaned file into many sections.
4. Creating Directories in Hadoop file system & uploading data onto the file system.
5. Running Inverted Index on the data.
6. Understood how map reduce works.

Output of Inverted Index: