**BFS Assignment 1: ReadMe File**

The MapReduce Program consists of the following classes:

1. **File Generator Class**

This class generates the input file for the MapReduce program. The input file is already generated and is located in the Input folder. Sample data for the input file is as follows:

<Node> <tab> <adjacency list>|<weight>|<node color>|<parent>

1. **Driver Class**

The driver class sets the configuration parameters (Input / Output file, Mapper class, Reducer class, Partitioner class etc.) for MapReduce jobs through the Job object. Hadoop uses these configurations to run the MapReduce program.

This Assignment has 2 jobs (1 – Graph summarization, 2 – Graph processing).

* Job 1 uses 1 Mapper class (Mapper2.class)
* Job 2 uses 1 Mapper (Mapper1.class), 1 Reducer (Reducer1.class) and 1 Partitioner class (Partitioner1.class)

The Input parameters for the Driver class is the Input and Output folder paths.

1. **Mappers and Reducers Classes**

Mapper2.class

This Mapper takes the randomly generated Input file as the input and generates an Output file which contains the nodes and corresponding adjacency list. It also outputs to the screen the Total number of Nodes and Edges in the Input graph.

Mapper1.class

This Mapper takes the randomly generated Input file as the input and processes it to generate key value pairs. In this Mapper, the key is the node ID and the value is the corresponding text after the tab in the Input file.

Reducer1.class

The Reducer1 class takes input from Mapper1.class and further processes

i.e., generates a single output line for multiple values of the same key (Node

ID) based on the following rule:

* The node with the darkest color and the minimum distance value.

The Mapper and Reducer classes are further chained in the Driver class until all the nodes (lines) in the input file are visited.

1. **Partitioner Class**

The Partitioner1.class creates 3 partitions based on the values of node ID. The Partitioner takes input from the Mapper1.class and generates 3 partitions for each of the 3 Reducer tasks defined in the Driver class.