**Assignment 21.2**

|  |
| --- |
|  |
|  |

|  |
| --- |
| **Mapper Class -** |
|  |

|  |
| --- |
|  |
|  |

|  |
| --- |
|  |
|  |

|  |
| --- |
| Mapper class has got four methods |
|  |

|  |
| --- |
| 1. Setup |
|  |

|  |
| --- |
| 1. Map |
|  |

|  |
| --- |
| 1. Run |
|  |

|  |
| --- |
| 1. Clean up |
|  |

|  |
| --- |
|  |
|  |

|  |
| --- |
|  |
|  |

|  |
| --- |
|  |
|  |

|  |
| --- |
| **Setup Method**   * As name suggests it is the first method executed before any of the other method is executed, this method is used for the setting up initial setups if needed before the actual mapping starts. e.g finding the cache file and using the that file can be done in the setup phase of the mapping .   **Map Method**   * In this method the input is taken from the file line by line and converts in the key value format and sent to the reducer. Map method is called for each input split spawned by the input format from the job. |
|  |

|  |
| --- |
|  |
|  |

|  |
| --- |
| **Run Method**   * The default run() method simply takes each key / value pair supplied by the context and calls the map() method. This method is overridden very rarely. |
|  |

|  |
| --- |
|  |
|  |

|  |
| --- |
| **Clean up method-** |
|  |
|  |

|  |
| --- |
| * During cleanup() is that you clean up any resources you may have allocated. There are other uses too, which is to flush out any accumulation of aggregate results. |
|  |

|  |
| --- |
|  |
|  |

|  |
| --- |
| **Flow** |
|  |

|  |
| --- |
| Setup ---> Map -----> Clean up |
|  |
|  |

|  |
| --- |
|  |
|  |

|  |
| --- |
|  |
|  |

|  |
| --- |
|  |
|  |

|  |
| --- |
|  |
|  |

|  |
| --- |
|  |
|  |

|  |
| --- |
|  |
|  |

|  |
| --- |
|  |
|  |

|  |
| --- |
| **Reducer Class -** |
|  |

|  |
| --- |
|  |
|  |

|  |
| --- |
| As we know the reducer code reads the outputs generated by the different mappers as <Key,Value> pairs. The Reducer interface expects four generics, which define the types of the input and output key value pairs. The first two parameters define the intermediate key and value types, the second two define the final output key and value types. |
|  |

|  |
| --- |
|  |
|  |

|  |
| --- |
| Reducer has got same **Setup , run** and **clean up** method as in that of the mapper class with same functionality |
|  |

|  |
| --- |
|  |
|  |

|  |
| --- |
| **Reduce method** –  The main task of the reducer class is to perform user operation on all the mapper key value pairs sort and shuffle results and to combine these results into one output. |
|  |