Task No: Use Case 2 Date:22/10/25	Data analytics using MapReduce program on weather dataset	CO2 K3

AIM: To Construct MapReduce program to perform data analysis on weather dataset.

PROCEDURE:

Step 1: Set up Hadoop Cluster

Before you start, ensure you have a Hadoop cluster set up and running. You will need Hadoop installed, configured, and a Hadoop Distributed File System (HDFS) where your dataset is stored.

Step 2: Write Mapper and Reducer Classes

Create two Java classes, one for the Mapper and one for the Reducer.

Step 3: Configure and Run Hadoop Job.

Create a Hadoop job configuration and submit the job

Step 4: Compile and Package

Compile your Java code, create a JAR file, and include all dependencies (e.g., Hadoop libraries) in the JAR.

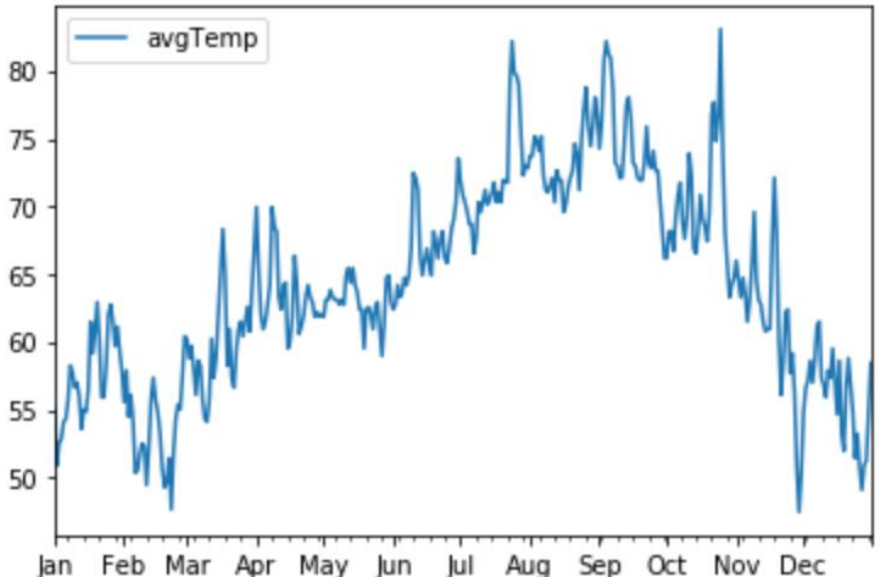
Step 5: Run the MapReduce Job

IMPLEMENTATION:

```
class WeatherAnalysis(MRJob):
    def mapper(self, _, line):
        try:
            fields = line.strip().split(',')
            if len(fields) == 8:
                year = fields[0].split('-')[0]
                temperature = float(fields[1])
                yield year, temperature
```

```
        except ValueError:
            pass
    def reducer(self, year, temperatures):
        temperatures = list(temperatures)
        average_temperature = sum(temperatures) / len(temperatures)
        yield year, round(average_temperature, 2)
if __name__ == '__main__':
    WeatherAnalysis.run()
```

OUTPUT:



RESULT:

Thus, Implemented Mapreducer Program to perform data analysis on weather dataset successfully.