Date	10 October 2022
Team ID	PNT2022TMID40045
Project Name	Exploratory Analysis of Rainfall Data in India for Agriculture.
Maximum Marks	4 Marks

Solution Architecture:

Fragmentation

Clustering

mapper

Kmeans mapper class

reducer

Kmeans reducer class
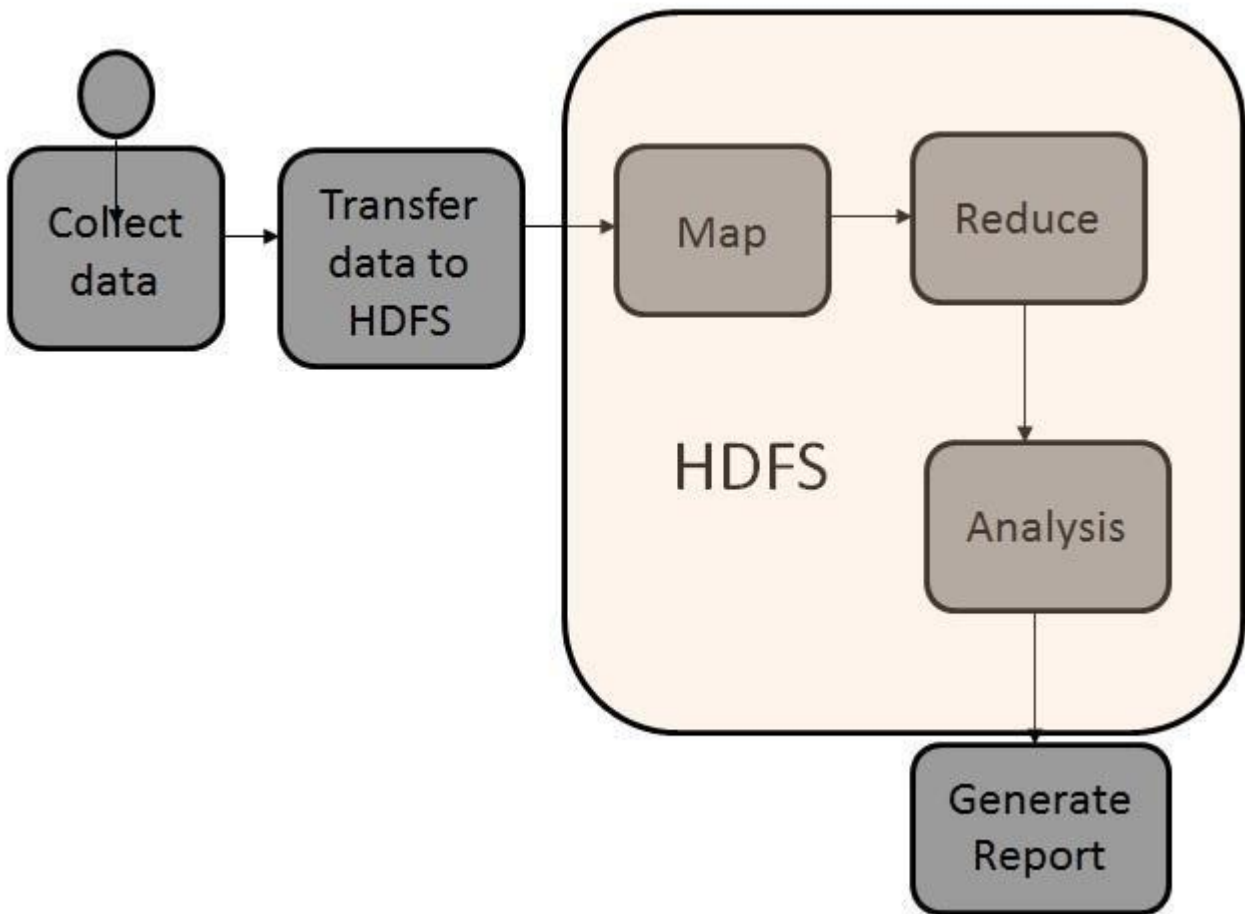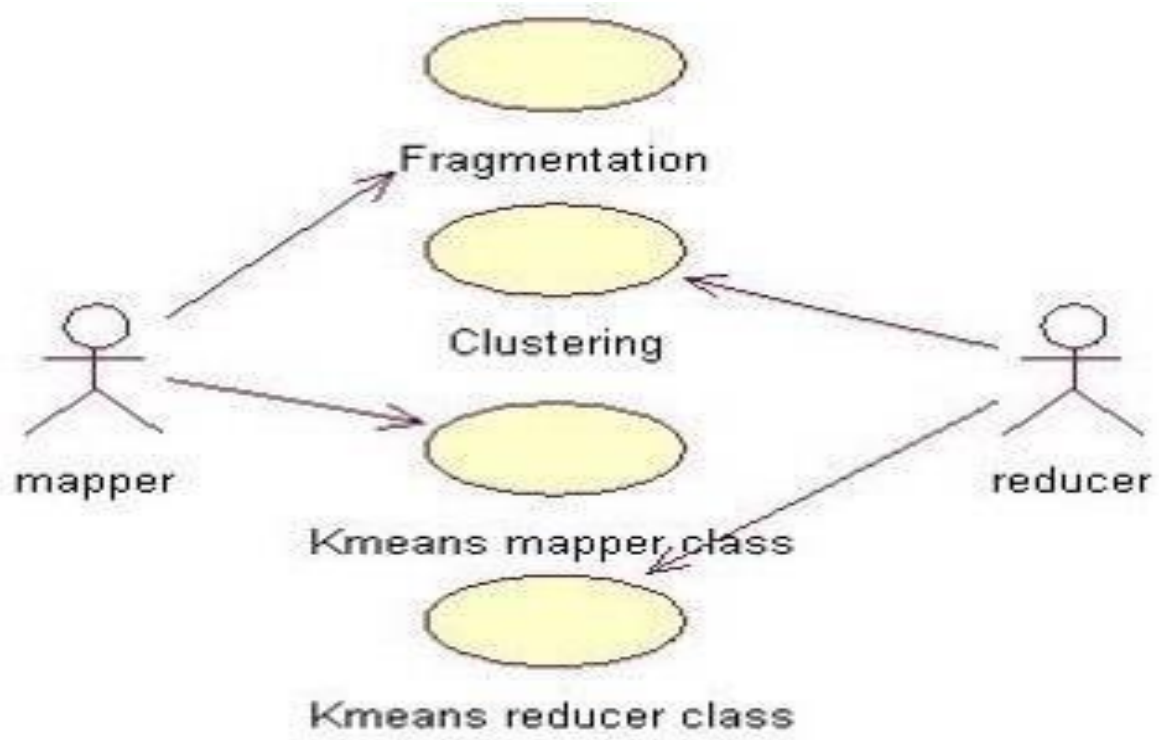
ANNUAL RAINFALL OF INDIA IN mm

S.NO	YEAR	ANN
1	1901	1030.8
2	1902	1038.4
3	1903	1195.9
4	1904	1025.1
5	1905	977.5
6	1906	1149.2
7	1907	1034.8
8	1908	1077.4
9	1909	1128.5
10	1910	1183.9
11	1911	1028.9
12	1912	1070.4
13	1913	1061.8
14	1914	1185.9
15	1915	1124.4
16	1916	1324.8
17	1917	1463.9
18	1918	1020.2
19	1919	1287.9
20	1920	1039.1
21	1921	1225
22	1922	1204.2
23	1923	1148.6
24	1924	1245.9
25	1925	1189.5
26	1926	1226.2
27	1927	1244.6
28	1928	1200.2
29	1929	1193.2
30	1930	1198.5
31	1931	1292.8
32	1932	1202.9
33	1933	1372
34	1934	1217.5
35	1935	1127.9

Indian agriculture mainly depend on the rainfall rate, major amount of cultivation is done based on the rain. If the rain fall rate is low then automatically the productivity in agriculture is also low. We are going to predict the future rain fall rate based on the past years. If we predict the future rain fall rate then automatically it will reflect on agriculture then it increases the income in agriculture. Agriculture productivity in India play important role in economy of country.

Indian agriculture is mainly based on the old traditional way of cultivation. We can replace the traditional way of cultivation with new way of using past data and statistics. To help the farmers in India our prediction and analysis. To predict the rain fall rate, Data Mining technique, clustering (modified k-Means) used. But now days the data size is increased dramatically and data is available in variety forms such as structured and un-structured for such reason we have gone through Hadoop Single Node Cluster to estimate the rainfall rate.