Question: Following table shows the current (11th of November) statistics of Corona effected countries. Apply K-means algorithm to find the most effected countries, least effected countries and countries where pandemic situation is spreading but still under control. Use minimum three numerical data 3 columns.

#	Country, Other	Total Cases	Total Deaths	Total Recovered	Active Cases	Population
	World	52,283,356	1,286,979	36,603,966	14,392,411	
1	USA	10,633,715	246,732	6,621,711	3,765,272	331,706,290
2	India	8,684,039	128,164	8,064,548	491,327	1,384,900,866
3	Brazil	5,730,361	163,078	5,064,344	502,939	213,106,404
4	Poland	618,813	8,805	242,875	367,133	37,831,471
5	Chile	524,804	14,633	501,426	8,745	19,175,505
6	Iraq	508,508	11,482	436,657	60,369	40,543,343
7	Saudi Arabia	351,849	5,590	338,702	7,557	35,008,225
8	Pakistan	348,184	7,021	320,065	21,098	222,422,728
9	Romania	324,094	8,389	218,086	97,619	19,190,546
10	Israel	321,235	2,699	310,061	8,475	9,197,590
11	Qatar	134,887	233	131,926	2,728	2,807,805

X = Total Cases, Y = Total Deaths, Z = Active

Cases Cluster = (X, Y, Z) => (Total cases, Total deaths, Total Recovered)

1. K = 3 (Number of clusters)

2. Initial centroid: -

i. C1 => USA = (10633715, 246732, 6621711)
ii. C2 => Chile = (524804, 14633, 501426)
iii. C3 => Romania = (324094, 8389, 218086)

3. Calculate distance: Using Euclidean Distance $\sqrt{(x2-x1)2+(y2-y1)2+(z2-z1)2}$

Country's Distance	C1	C2	C3
USA	0.000	11819553.292	12138842.848
India	2428389.072	11125952.519	11466036.472
Brazil	5145412.560	6923877.546	7262084.285
Poland	11876212.361	275173.177	295759.964
Chile	11819553.292	0.000	347282.374
Iraq	11867183.842	66861.886	285991.909
Saudi Arabia	12051988.711	237593.086	123775.728
Pakistan	12062747.956	293466.105	186838.784
Romania	12138842.848	347282.374	0.000
Israel	12093127.039	279648.568	92195.177
Qatar	12345179.645	537375.918	208060.940

C1 (Most effected)	C2 (Least affected)	C3 (Spreading but under control)
USA	Poland	Saudi Arabia
India	Chile	Pakistan
Brazil	Iraq	Romania
		Israel
		Qatar

4. Compute the average pointes:

i. C1 => Average (USA, India, Brazil) = (8349371.667, 179324.666, 6583534.333)

ii. C2 => Average (Poland**,** Chile**,** Iraq) (550708.333, 11640.0, 393652.667)

iii. C3=> Average (Saudi Arabia**,** Pakistan, Romania, Israel, Qatar) (296049.8, 36284.0, 263768.0)

Country's Distance	C1	C2	C3
USA	2285656.507	11853733.686	12138164.991
India	1519217.272	11180665.490	11455085.044
Brazil	3027774.776	6976180.384	7252126.920
Poland	9999728.748	165469.598	324603.907
Chile	9911761.917	110883.199	330573.101
Iraq	9964507.773	60251.686	275034.966
Saudi Arabia	10148299.770	206363.624	97902.902
Pakistan	10161216.272	263371.825	148202.717
Romania	10244668.521	286684.983	60427.256
Israel	10190124.760	244388.050	62492.307
Qatar	10446678.049	491465.511	211318.328

C1 (Most effected)	C2 (Least affected)	C3 (Spreading but under control)
USA	Poland	Saudi Arabia
India	Chile	Pakistan
Brazil	Iraq	Romania
		Israel
		Qatar

Since the clusters in the new grouping remain the same, then K-means has found the optimal solution and terminates at this stage.