

```
1 public class KMeans {
2     static final int numClusters = ...;
3     static final double EPSILON = 0.01;
4     static List<Point> points;
5     static Map<Integer, Point> centers;
6
7     public static void main(String[] args) {
8         points = readFile("cluster.dat");
9         centers = Point.randomDistinctCenters(points);
10        MapReduce<List<Point>, Integer, List<Point>, Point> mapReduce
11            = new MapReduce<>();
12        mapReduce.setMapperSupplier(KMeans.Mapper::new);
13        mapReduce.setReducerSupplier(KMeans.Reducer::new);
14        mapReduce.setInput(splitInput(points, numWorkerThreads));
15        double convergence = 1.0;
16        while (convergence > EPSILON) {
17            Map<Integer, Point> newCenters = mapReduce.call();
18            convergence = distance(centers, newCenters);
19            centers = newCenters;
20        }
21        displayOutput(centers);
22    }
23    static class Mapper extends Mapper<List<Point>, Integer, List<Point>> {
24        public Map<Integer, List<Point>> compute() {
25            Map<Integer, List<Point>> map = new HashMap<>();
26            for (Point point : input) {
27                int myCenter = closestCenter(centers, point);
28                map.putIfAbsent(myCenter, new LinkedList<>());
29                map.get(myCenter).add(point);
30            }
31            return map;
32        }
33    }
34    static class Reducer extends Reducer<Integer, List<Point>, Point> {
35        public Point compute() {
36            List<Point> cluster = new LinkedList<>();
37            for (List<Point> list : valueList) {
38                cluster.addAll(list);
39            }
40            return Point.barycenter(cluster);
41        }
42    }
43 }
```

FIGURE 17.7 A MapReduce-based KMeans application.