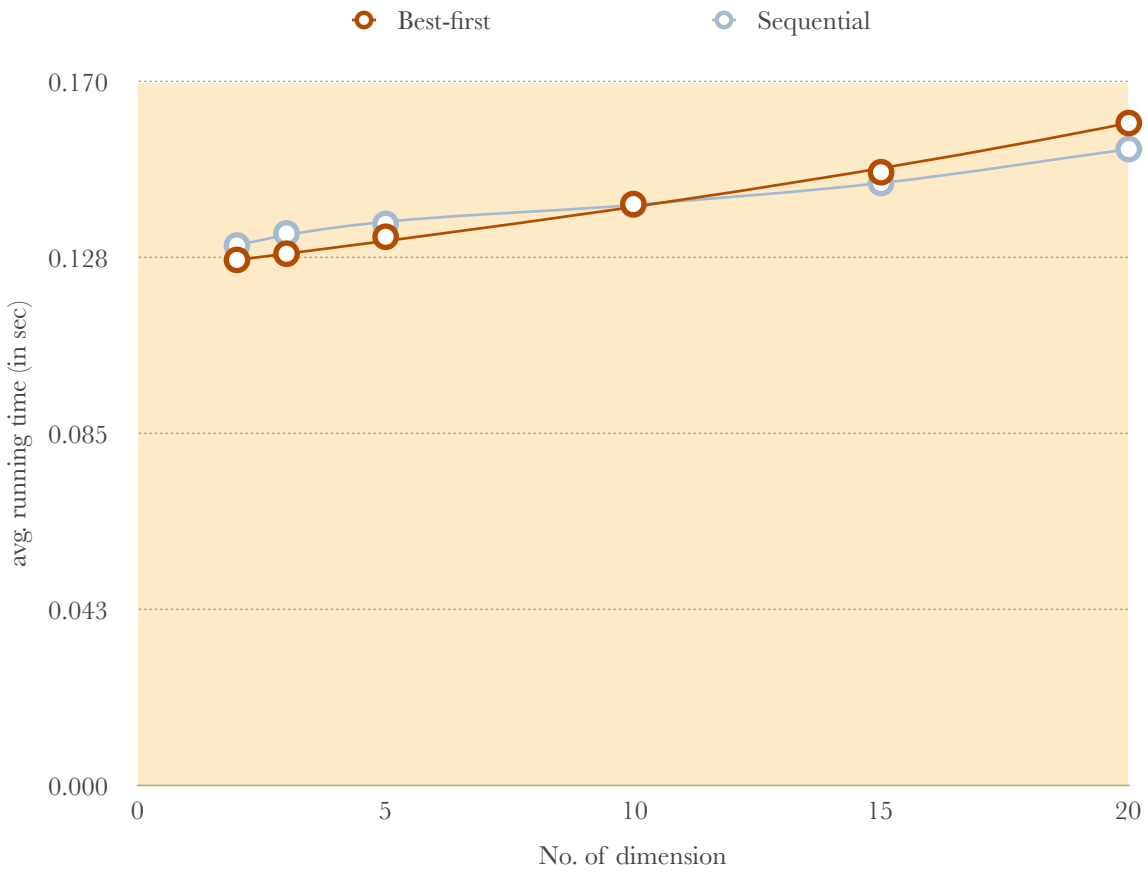
Ashley, Kartar
2014CS50280, 2014CS50286
April 15, 2018

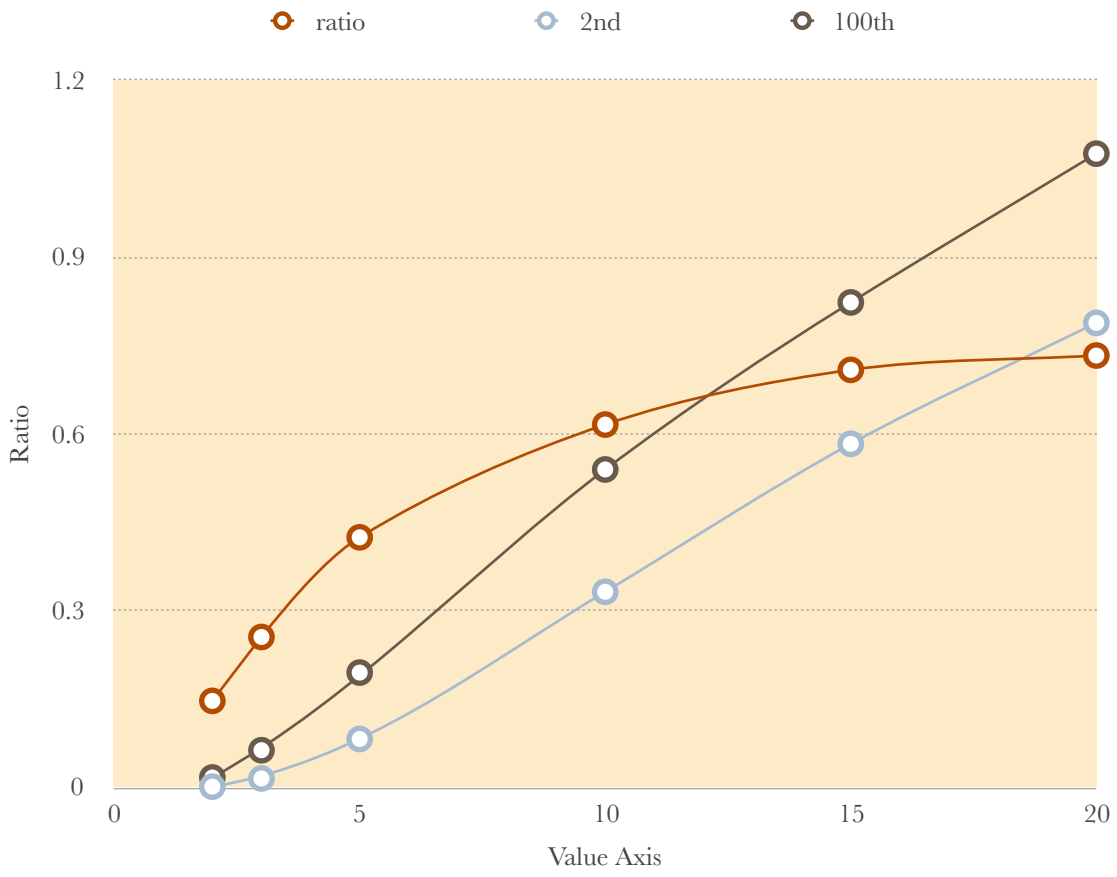
DBMS: Assignment 3

1. (a)



	Best-first	Sequential
X	Y	Y
2	0.1271	0.1306
3	0.1286	0.1335
5	0.1327	0.1358
10	0.1406	0.1404
15	0.1483	0.1457
20	0.1602	0.1539

1. (b)



	ratio	2nd	100th
X	Y	Y	Y
2	0.1479	0.0026	0.0181
3	0.2562	0.0164	0.0643
5	0.4251	0.0833	0.1959
10	0.6162	0.3328	0.5401
15	0.7089	0.5836	0.8232
20	0.7332	0.7885	1.0753

1. (c) Sequential scan complexity is O(n logk) . From the observation of **first plot**, the running time increases logarithmically with increasing in no. of dimension

(k) and n is same for all k. While for best first algorithm, the running time increases linearly with no. of dimensions -> O(kn*log n) . And, best first algorithm running time is higher than that of sequential scan for higher dimension (k>=10). This is due to scan the whole dimension of each point for computing MBR and distance between query point & MBR.

From the observation of **second plot**, the ratio decreases with decreasing in no. of dimension. It is due to the density is higher(or lower the volume) when the no. of dimension is lower. So, we can find k points near query point very quickly when the dimension is lower.

Further, if the density around query point is low, the volume of shape needed to encompass k nearest neighbors is large & consequently, the ratio is low.