Q5)

a) $\lceil \log_{100} 10{,}000{,}000 \rceil = \lceil 3.5 \rceil = 4$ max depth

$\lceil \log_{200} 10{,}000{,}000 \rceil = \lceil 3.04 \rceil = 4$ min depth

b. & c) In the worst case, we assume all nodes are half full.

records	reads	writes	total (b)	total (c)
1-100	1	1	2*100	0
101-10,000	2	2	4*(10,000-100)	0
10,001 - 1,000,000	3	3	6*(1million-10K)	2*(1 million-10K)
1,000,001-10,000,000	4	4	8*(9 million)	4*9 million
	Total		77,979,800	37,980,000

In Part (C), we also need to write the first two layers of the index on disk at the end. That will additional 101 I/Os.

d) Assuming the data is sorted, the index can be constructed in a way that we fill the nodes fully from left to write, keeping the index frontier in memory, and when a node is written, we wont touch it again. The # of IOs will be the size of the index.

Index size = $\lceil 10{,}000{,}000/200 \rceil + \lceil 50{,}000/200 \rceil + \lceil 250/200 \rceil + 1$