B+ tree Problem

Assume that you have 100,000 student records each of length 100 bytes, which fill to capacity data pages of 2048 bytes each.

To index records on their unique STID key (so there are 100,000 different key values) we build B+ tree. This is a dense index since bottom level nodes of the B+ tree hold a pointer to each record in the data page (plus a pointer to the next node in the B+ tree). Assume that keys use 20 bytes and pointers also use 20 bytes. The page size is still 4096 bytes.

A. How many pages are there in the data file.

B. Describe the number of nodes at each level of the B+ tree, for both the best case and the worst case situation (i.e., least pages and most pages used, respectively).

C. Now, say that we use the B+ tree to support a range search query on the index, and that this search returns 72 records. Estimate the number of pages needed to answer the query when
(i) the student records where placed in the file sorted STID, and
(ii) the student records were stored in the file sorted by their SS#

Assume that you have 100,000 student records each of length 100 bytes, which fill to capacity data pages of 2048 bytes each. To index records on their unique STID key (so there are 100,000 different key values) we build B+ tree. This is a dense index since bottom level nodes of the B+ tree hold a pointer to each record in the data page (plus a pointer to the next node in the B+ tree). Assume that keys use 20 bytes and pointers also use 20 bytes. The page size is still 2048 bytes.

A. How many pages are there in the data file:
Each page holds 20 records (Records are not split across pages). We have 5000 pages.

B. Describe the number of nodes at each level of the B+ tree, for both the best case and the worst case situation (i.e., least pages and most pages used, respectively).

- *One can fit 50 key-pointer pairs, and an additional pointer. Thus n=51.*
  Best case:
  100, 000/50 = 2000 leaf nodes, ⌈2000/51⌉ = 40 at the next level, and then the root.
- *Worst case:*
  100000/25 = 4000 leaf nodes,
- *At the next level: 4000/26 = 153.8. Do we take the ceiling: no we take the floor. 153/26 = 5.9 so we have 5 blocks. Then there is the root.*

C(i): 4+1+5= 10 C(ii): 4+1+72= 77