Q1 Does the content of the index node include the pointer addresses for the key-pointer pairs, or will printing the keys be sufficient?

Keys only will do.

Q2 Does the content of data blocks refer to all records in the data blocks and their content?

The "tconst" values of those records within the data block will do.

Q3: Are we allowed to use Python for the B+ Tree project?

Ans: Any programming language can be used as long as the size information could be collected and presented in a clear way. Note that B+ tree structure uses the concepts of pointers quite often. Projects based on different programming languages will be assessed equally.

Q4: When packing fields into records, is it okay to have a fixed format with fixed length? Will doing so cause marks to be penalised?

Ans: The way how the packing is done is flexible. As long as the design could be well justified, there should be no penalty.

Q5: What do you mean by size information exactly?

Ans: By size, it means the storage (in terms of Bytes).

Q6: In experiments, it is asked to report the size of database. Does the database include the index?

Ans: By the size of the database, it means the sum of the size of the relational data and that of the index.

Q7: Does 1 MB = 10^6 B or 1MB = 2^20 B?

Ans: 1 MB = 10^6 Bytes

Q8: Is it required to consider other data types than those in the dataset?

Ans: It is good enough to consider only the data types in the dataset.

Q9: Is it required to consider other datasets than the provided dataset?

Ans: It is good enough to consider only the provided dataset.

Q10: Should the index at the 1st level be a clustered index?

Ans: No, it is not required to build a clustered index at the 1st level.

Q11: According to the project description, the B+ tree shall be stored in the main memory with the node bounded by the block size. Does it mean that the tree shall be stored in the storage component designed in part 1?

Ans: No. It's just that the size of each node should be bounded by the block size.