Common Question Clarifications

Storage, Indexing, and Query Evaluation
Alternatives for data entries $k^*$

- Three alternatives
  1. Data record with key value $k$
  2. $<k, \text{rid of data record with search key } k>$
  3. $<k, \text{list of rids pf data record with search key } k>$

- Choice of alternatives for data entries is orthogonal to the indexing technique
  - Example of indexing techniques: B+-tree, hashing-based structures
  - Typically, index contains auxiliary information that directs searches to the desired data entries
Cost of query evaluation using index

- We have been focusing on #page IOs
  - Expensive, compared to operations in memory
- Two sources that lead to page IOs:
  1. Locating the data entries (if index are not fully in memory)
     - This is heavily determined by the indexing technique
  2. Cost of retrieving data from disk (if page is not in buffer pool)
     - Clustered vs unclustered can lead to different costs