Homework 2

Simon Zheng
260744353

March 12th, 2018

1 Indexing

1.1 sid index

1.1.1 (a)

Using the information we have:

$$\frac{40 courses}{1 dept} \times 100 depts \times 4 years \times \frac{2 terms}{1 year} \times \frac{1 courseoffer}{1 section} \times \frac{140 students}{1 courseoffer} \times \frac{1 university}{100000 students} \times 0.9$$
$$= 40.32$$
$$\approx 41 \text{ rid per data entry}$$

The size of the data entry:

$$8 bytes + (10 bytes \times 40.32) = 411.2 bytes$$

Since we index on the *sid* as key which is an 8 bytes (64 bits) int. A data entry also contains up to 41 rids as calculated previously and each is 10 bytes.
For the total number of data entries:

$$40.32 \times 100000 \ students = 4032000 \ data \ entries$$

Which is the maximum total amount of data entries we could get with the values we have.

1.1.2 (b) Number of leaves

$$\frac{412 \ bytes}{1 \ data \ entry} \times \frac{1 page}{4000 \ bytes} \times \frac{100}{60} = \frac{103 \ pages}{600 \ data \ entries}$$
$$\frac{103 \ pages}{600 \ data \ entries} \times 100000 \ data \ entries = 17166.\bar{6} \approx 17167 \ pages$$

1.1.3 (c) Number of intermediate nodes

We can fit all of the data in a tree of two intermediate nodes, both for min and max.