Problem 1

1.1

Size = Header + name + student + birthday + state

= 4 + 1 + 1 + 8 + 1 = 15 B

1.2.

4 + 1 + 1 + 8 + 1 = 15 B

↓ word-alignment

(4 + 1 + 1 + 2) + (8) + (1 + 7) = 24 B

At page level:

slot count + free space pointer = 8 + 8 = 16B

For each record:

record pointer + record length = 8 + 8 = 16B

Totally we have:

$16 + (24 + 16)x = 16 + 40x \quad (x \in N)$

$\because 16 + 40x \leq 2048 \quad (x \in N)$

$\therefore x \leq 50.8 \quad (x \in N)$

$\therefore \max x = 50$

$\therefore$ maximum number = 50

1.3. reorder:

birthday + Record Header + name + student + state + padding

= (8) + (4 + 1 + (1 + 1 + 1) = 16B

$\therefore 16 + (16 + 16)x \leq 2048$

$x \leq 63.5$

$\therefore \max x = 63$

Problem 2.

2.1

DSM: for size: 100 Pages
for credits: 100 Pages
$\therefore$ total: 200 pages

NSM: read both size and credits: all tuples are needed
$\because$ worst cases
$\therefore$ 500 pages.

2.2.

DSM: min: 1 for id, 3 for $\begin{matrix}\text{name}\\ \text{instructor}\\ \text{size}\end{matrix}$ = 4 pages

max: 100 for id, 2x3 for $\begin{matrix}\text{name}\\ \text{instructor}\\ \text{size}\end{matrix}$ = 106 pages

NSM: min 1 for all = 1 pages

max: worst cases: 500 pages.

Problem 3: Bloom Filter:

3.1:

insert 7:

$h_1(7) = 7$ 0.7 mod 10=

$h_2(7) = 0$

set 0th and 7th to 1:

[1 0 0 0 0 0 1 0 0]

insert 27

$h_1(27) = 7$

$h_2(27) = 2$

set 7th and 2nd to 1:

[1 0 1 0 0 0 1 0 0]

3.2

Because: Hash functions can collide: different element may map to same bits
When number of elements increases, more bits are set to 1. causing bit overlaps.

3.3.

the false positive probability p for a Bloom Filter is approximately

$$p \approx (1 - e^{-\frac{kn}{m}})^k$$

where m is the size of the bit array
n is the number of elements inserted
k is the number of hash functions

we choose $k = \frac{m}{n} \ln 2$

$\Rightarrow p \approx (\frac{1}{2})^k < 0.01$

$\Rightarrow k > \log_2(100) \approx 6.64 \quad k \in N^+$

$\therefore$ set $k = 7$

$\therefore m = \frac{kn}{\ln 2} = \frac{7000}{\ln 2} \approx 10100$

$\therefore$ we can choose $m \approx 10000$ and $k = 7$