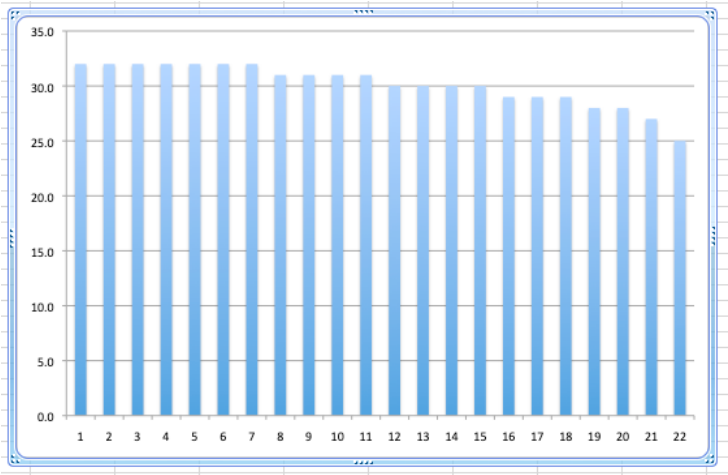
Problem set 1.

This is a rough guide to how you might answer the questions on this problem set. Graded out of 32.



General observations:

Many people gave answers that are way too precise. There is no precise value for miles of roads, so anything computed from that can't have very many figures. The price of gas varies by pennies every day and is different everywhere in the country, so anything more accurate than say $2 or $2.50 or $3 is excessively precise. And so on.

You can do precise arithmetic then throw away most of the digits at the end, but it's just as good to throw away most of the digits at the beginning; that way you can do the computation in your head or on paper, without a calculator.

Some confusion of bits, bytes, and what can be stored in them. If a byte is only large enough to hold one alphabetic character, it's not big enough to hold a picture! JPEG images are typically going to be several MB.

Be careful of units: lots of confusion of MB, GB and TB, usually by missing a factor of 1000 somewhere. Arithmetic errors crop up here too.

Area is not a linear dimension: if the length or width of something doubles, the area goes up by 4.

Problem 1: [12 pts, 3 each]

(a) Bytes to store your academic record?

An example record:

```
Brian Kernighan 1969 COS109 Computers in our World C- (4x)
```

which is about 15 + 5 + 4*35, so roughly 150-200 characters.

Answer should be somewhere 100-300, maybe less if you don't include the course name.

(b) Bytes for all current undergrads?

200 * 5000 = 1,000,000, so something in the range 200K to 1M?

(c) An idea for compression

collect / replace course name by just number
collect all names for each course
collect all names for each year
etc.

Any scheme that recognizes repeated data.
Schemes that compress COS into CS or the like are less effective than those that only use the course name once, etc.

Think of putting all the data in a spreadsheet, then sorting by some column(s).
Any column where the sorting yields a bunch of identical entries is a good target, since you only need to store the unique values.

(d) Accurate, conservative, optimistic, and why.

If it takes 500KB for all students for one semester, that's 1MB for a year. 250 years * 1 MB is 250 MB.
It's very conservative.

Problem 2: [12 pts, 3 each]

(a) Miles of road in the US? (Km is ok too)

If the USA is 3000 miles wide x 1500 miles high, there are 1500 E-W roads that are 3000 miles long and 3000 N-S roads that are 1500 miles long. That's 2 * 1500 * 3000 = 9 million. Probably too high, but anything in the 1M to 10M range is ok.

A fair number of people didn't explain how they got from area to length of roads. One way is to say that each one-mile square has four miles of road around the edge, but then you have to divide the total by two to avoid double-counting.

(b) Cost of gas?

Gas is $2-3/gallon, a car gets 20-25 mpg, so 10-15 cents/mile, so $100K to $1M.

(c) Disk storage needed, in TB?

1M miles = 5B feet so 5 * 10^9 images * 5*10^6 bytes/image
= 25 * 10^15 bytes = 25,000 TB.

(d) Cost of disk storage?

Disk is about $20/TB, so $20 * 25 * 1000 so about $500 K.

Problem 3: [8 pts, 2 each]

(a) How many GB of secondary storage does your computer have?

128, 256, 512 GB are typical.

(b) How many GB of RAM?

4, 8, 16 GB are typical

(c) Pixels on the screen?

1440 * 900 = 1.3 Mp, or whatever

(d) Pixels on your cellphone screen?

Anything reasonable. Cell phone screen pixels are much denser.