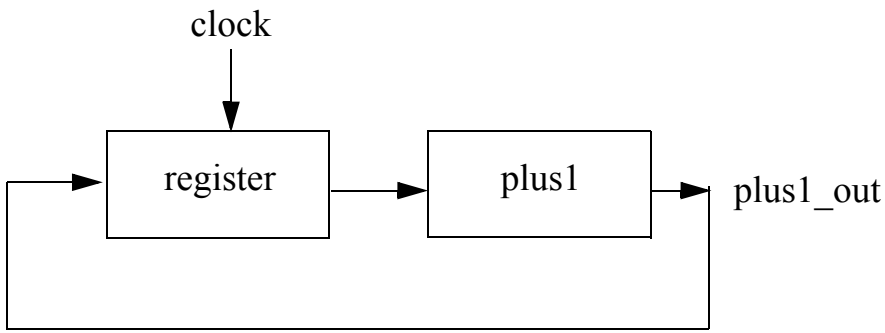
# Statistics

Why haven't we needed statistics so far in this class?

# Maximum clock frequency

Is there a limit to how fast you can set the clock?



## Speed binning of CPU chips

CPU chips vary in the clock rate at which they can run, due to physical variations. E.g., some processor chips in a batch may be able to run at 3.4 GHz, others only at 3 GHz

Chip manufacturers measure how fast each chip can run at and put them in bins

Let's say we're measuring the maximum clock rate for a chip

- try different clock rates, see if the chip works reliably on a set of programs
- use a variable-speed clock generator to try different clock rates
- clock generator has a speed dial, with marks every .05 GHz (2.40, 2.45, 2.50, 2.55, ... , 3.40)
- e.g. measure the max clock rate of a particular chip as 3.15 GHz

## Two attributes of a measurement

How accurate is this measurement?

- how close is your measurement to the true number?
- e.g. if true maximum clock rate for this chip is 3.16 GHz, you're accurate to within .01 GHz

How precise is this measurement?

- how exactly did you measure it, i.e. to what level of detail?
- what is the limit to precision with the above clock generator?

- is your precision guaranteed to be no worse than this?

Possible to be precise but not accurate

- if you measure many times and get 3.15 GHz each time, you're precise to within .05 GHz
- but what if the clock you're using is miscalibrated by .5 GHz, so the true max clock rate for this chip is 3.65 GHz?

Possible to be accurate but not precise?

Experimental error: difference between the measured value and the true value

# Communicating possible error

How you write a number implies a certain precision and accuracy

What if I say the max clock rate of this chip is 3.1415 GHz?

- what does that imply about how accurately and precisely I know the true max clock rate?

- how should I write the number?

Are zeroes significant?

- e.g. 3,150,000,000 Hz. Are the zeroes significant?

## Summarizing multiple data values

Let's say you measure a set of chips to have the following maximum clock rates:

```
3.00 GHz
3.10 GHz
3.10 GHz
3.20 GHz
3.40 GHz
3.50 GHz
3.80 GHz
4.00 GHz
4.40 GHz
```

Your boss asks you "How fast were the chips in the last batch?". What would you say?

## Averages

Mean

- arithmetic average
- add the values, divide by the number of values
- often abbreviated $\mu$

Median

- the 50th percentile
- half the values are higher; half are lower

Mode

- the most common value

# Standard deviation

Mean, median, mode all describe some kind of "average"

How to capture the amount of variation between values?

- e.g. large amount of variation may indicate a problem in the manufacturing process

Deviation for above set of data (from the mean of 3.50)

| value | deviation | deviation$^2$ |
|---|---|---|
| 3.00 | | |
| 3.10 | | |
| 3.10 | | |
| 3.20 | | |
| 3.40 | | |
| 3.50 | | |
| 3.80 | | |
| 4.00 | | |
| 4.40 | | |

Standard deviation is the square root of the average of the squared deviations

$$\sqrt{\frac{1}{n}\sum_{i=1}^{n}(x_i - \mu)^2}$$

Standard deviation is often abbreviated $\sigma$