GOB: 1) Calculate Sum(RandInt(0,1,100)) and add your answer to the chart on the white board.

2) Sketch the binomial distribution for the number of tails you expect when flipping a fair coin 100 times.
Then calculate the mean and standard deviation.

3 a) You play roulette 100 times, betting $10 on red each time. What is the probability you come out ahead (win more than half of the time)? Show that this Binomial probability model can be approximated by a Normal model. Then use the applicable Normal model to solve the problem.

3 b) You play roulette 900 times, betting $10 on red each time. What is the probability you come out ahead? Show that this Binomial probability model can be approximated by a Normal model. Then use the applicable Normal model to solve the problem.

3 c) Explain what happened to your probability of winning as the number of bets increased. Why?

Part V Ch.18 Sampling distribution Models (SDMs)
If we were to summarize Statistics with just one word it would be variation.

To analyze categorical data, we often use counts or percents.
Now we will see another word used but if you get confused just replace it with percent.

Proportion	Ratio of: $\dfrac{\text{number of successes}}{\text{total}}$ [think percent]

Example 1: On October 26th, 2008 just a few days before the presidential election, two major polls showed different results. A Pew Research poll of 1198 likely voters showed that 53% supported Obama while 38% supported McCain. By contrast a poll conducted by IBD/TIPP at the same time showed that of 824 likely voters 47% supported Obama while 44% supported McCain, a spread of only 3% as compared to 15% in the other poll.
a) Was one poll wrong? Explain what is going on.

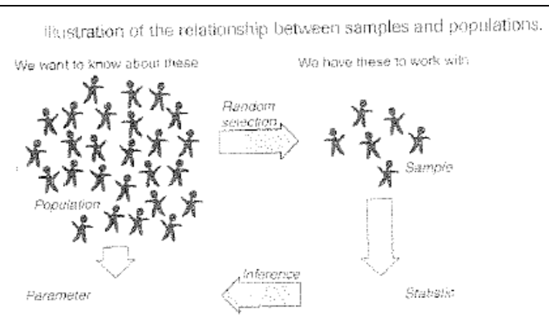
b) Should we be surprised to find that we could get outcomes this different from properly selected random samples drawn from the same population?

c) How can surveys conducted at the same time by organizations asking the same questions get different results?

It's because each survey is based on a different sample of about 1000 people. The proportions vary from sample to sample because the samples are composed of different people.

In order for a sample to fully and exactly describe the population it would have to be a census . Every other sample will vary from sample to sample because sampling variability will always be present. You might think that this would prevent us from learning anything reliable about a population by looking at a sample, but that's just not so. The fortunate fact is that sampling variability is not just unavoidable – it's predictable !

The link between probability and data is formed by the sampling distributions of statistics. A sampling distribution shows how a statistic would vary in repeated data production. When we can understand and predict the variability of our estimates, we've taken the essential step toward seeing past that variability, so we can understand the world.

We want to know the true population proportion, ___, but are often forced to work/estimate with a sample proportion, ___ .	p $\hat{p}$ 

Sampling variability (sampling error)	No sample fully and exactly describes the population; sample proportions and means will vary from sample to sample. It is not just unavoidable – it's predictable! (with SDMs)
Sampling Distribution Model (SDM)	Shows how a statistic (sample proportion or mean) would vary in repeated (think infinite) samples of size n. We used to focus on the data, and derive the statistics from it. Now we focus on the statistic itself. The sample proportion (or mean) becomes our datum, and in our imaginations we compare that statistic to all other values we might have obtained from all the other samples of size n we might have taken.

By thinking about what might happen if we were to draw many, many samples from the same population, we can learn a lot about how one particular sample will behave.

Sampling distribution models tame the variation in statistics enough to allow us to measure how close our computed statistic values (sample proportion or mean) are likely to be to the underlying parameters (population proportion or mean) they estimate.

The sample proportion, $\hat{p}$, does not have a binomial distribution because it is not the __________. But the SDM for a proportion appears to be _____ and _____. When certain conditions are met, the ________ is a good SDM for a proportion.	number of successes unimodal roughly symmetric Normal model
Assumptions / Conditions for using a normal model as the SDM for a proportion:	Assumptions: 1. Independent - sampled values must be independent of each other. Conditions: a) Randomization – SRS or at least representative and not biased. b) 10% Condition – If sampling w/o replacement Then $n \le 10\%$ of the population. 2. Sample Size - *n*, must be large enough. Conditions: a) Success/Failure - $np \ge 10$ and $nq \ge 10$.
Since the number of successes in the sample, X, is __________, we can obtain the mean and SD of the sample proportion by multiplying the mean and SD of the Binomial by the constant 1/n to get:	a Binomial random variable (n trials, probability p) $\mu(\hat{p}) = p \qquad \sigma(\hat{p}) = SD(\hat{p}) = \sqrt{\frac{pq}{n}}$
When we can understand and predict the variability of our estimates with SDMs, _______	we've taken the essential step toward seeing past that variability, so we can understand the world.

A4: pg. 432-438 / 6,8,15,16,26.