AP Statistics Final Review – Vocabulary

- 1. Census Collection of data from the entire population or close to. (e.g. US Census)
- 2. Observational Study observing and recording what is happening. For example, one student took Pepsi, another student selected milk. Can't prove anything.
- 3. Survey Informal way to collect data. Very biased if it's controversial, such as asking how much you drink. Question wording can be biased too.
- 4. Experiment A treatment has to be imposed on the subjects. For example, some group takes a drug and another group takes a placebo or a different drug.
- 5. Bias systematically favoring one outcome over another.
- Simple Random Sample place the names in a hat, mix them, select a certain number of names. SRS consists of n individuals from the population chosen in such a way that every SET of n individuals has an equal chance to be the sample actually chosen.
- 7. Voluntary Response Sample Send out he survey and those who want to, will reply. Usually only those with a strong opinion for OR against, will reply. This leads to biased results.
- 8. Population versus Sample a sample is a portion of the entire population. The sample is used to draw conclusions about the entire population.
- Confounding Too many things happened too many variables makes it confusing to tell what caused what. For example, you were speeding and using your cell phone and changed the radio station. It's hard to know which action caused you to drive off the road.
- 10. Blind & Double Blind Blind means the doctor knows, but the patient does not. Double blind means neither knows and this eliminates bias; this is the better method.
- 11. Stratified Random Sample Divide the population into strata, groups of individuals that are similar in some way that might affect their responses. Then choose a separate SRS from each strata.
- 12. Multi-stage Sampling For example, randomly select schools, randomly select classes, then randomly select students.
- 13. Treatment what you are doing to the (tomato plants) or fish tanks, etc. Each treatment can have a number of factors.
- 14. Control verses Control Group Control Group are those who received the placebo. The control group provides a baseline for comparing the effects of the other treatments. The Control is a part of experimental design. A proper comparative design ensures that influences other than the experimental treatments operate equally on all groups.
- 15. a. Placebo sometimes a fake pill, but could be a thoroughly tested drug (for example ibuprofen, to compare to a new anti-inflamatory drug). The placebo is given to the control group.

b. Placebo Effect – the confounding which occurs when patients get better because they expect the treatment to work even though they did not receive the active treatment.

- 16. Blocking when you are designing an experiment and you know ahead of time that males and females react differently to a drug, then you should block on gender because this reduces the variation. There will be less variation within gender. Block on known variables that effect the subjects in the study.
- 17. Matched Pairs match students will similar grades and assign one of each pair to the new experimental biology class. Or, have the subject be their own matched pair, have them solve puzzles without caffeine, then a few days later solve puzzles with caffeine.
- 18. An experiment needs:
- 1) Randomization use impersonal chance to assign experimental unit to treatments. Why? This helps create roughly equivalent groups of experimental units by balancing the effects of lurking variables that are not controlled on the treatment groups.
- 2) Control for lurking variables that might affect the response.
- 3) Replication use enough experimental units in each group so that any differences in the effects of the treatments can be distinguished from chance differences between groups.
- 19. Lurking Variables an unknown variable that causes issues in the results of the study.
- 20. Factors different levels, such as strength of the drugs: placebo group, 100mg, 200mg, & 400mg.
- 21. Cluster Sample divide the population into groups or clusters. Randomly select SOME of the clusters. ALL individuals in the chose clusters are included in the sample.
- 22. Statistically significant an observed effect so large that it would rarely occur by chance.

Correlation does not prove causation