

Your grade will depend on the depth and quality of diverse statistical perspectives in your responses, so take your time on the free responses. Be sure to follow all directions carefully. Informal writing is fine.

First circle the following statement: **A: I can do it!**

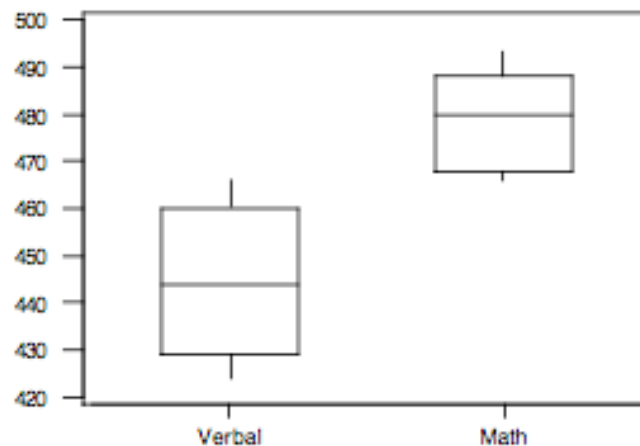
**(Analyzing Probability and Chance in HIV Testing)** In 2016, approximately 36,700,000 people worldwide were living with HIV out of 7,600,000,000 total (source <http://www.amfar.org/worldwide-aids-stats/>)

- a) What is the probability that someone is HIV+? **Show work, leave as a decimal, and round to 4 decimal places.**
- b) What is the probability that a someone is HIV-? **Show work, leave as a decimal, and round to 4 decimal places.**
- c) Suppose we have 100,000 representatives from around the world and the ELISA test correctly tests positive for 95% of those who have the disease, and has a false negative of 5%. It correctly tests negative for 99% of those who do not have the disease, and has a false positive of 1%. Compute the number of people in each cell. **Show work and round to whole numbers of people.**

	Test +	Test -
Person is HIV+		
Person is HIV-		
Total		

- d) Approximately what percentage of people who test positive is not actually HIV+? **Show work.**
- e) Say the cost was going to be a \$10 per person as a deductible. Would you support legislation for mandatory HIV testing worldwide using the ELISA HIV test? Explain and take the cost per life into consideration.

**(Investigating & Interpreting Key Features of Graphical Boxplots)** Say that **every** year from 1967 to 1981, average SAT scores declined or stayed the same. Yearly data of decreasing verbal and math scores from 1967 to 1981 were used to create:



- a) USE only the initial beginning and ending data to label the **hi** and **lo** of verbal scores and the year each occurred on the verbal boxplot:
- | Year | Verbal | Math |
|------|--------|------|
| 1967 | 466    | 492  |
| 1981 | 424    | 466  |
- b) USE a ruler or straight edge to very carefully draw ONE line across to the y-axis from the **median verbal score** on the verbal boxplot. **Note: this score does NOT end in a 5 (draw your line very precisely!).**
- c) Label the year that this median score occurred on the verbal boxplot after calculating the median (not the average!).
- d) Use only the work you did above in order to determine how many points **verbal** scores declined from the 1967 to the median year? **Show work.**
- e) Similarly, determine how many points **verbal** scores declined from the median year to 1981? **Show work.**
- f) Use **only parts e) and f)** to give a positive advertising spin on decreasing **verbal** SAT scores. Ie, the fact that they are decreasing is negative, but use parts d) and e) only to say something positive about the decline.
- g) Use only the math and verbal boxplots to give a positive advertising spin on decreasing **mathematics** SAT scores.
- h) What are the difficulties in studying whether there are gender or racial biases on the SAT? Relate your answer to our class discussions on stereotype vulnerability, and explain how we can help get rid of the threat/vulnerability by changing the way we test.

**(Analyzing Probability and Chance in Gallup)** On April 4, 2017, Gallup published poll results on its web site under the headline, "Affordable Care Act Gains Majority Approval for First Time." Of 1,023 adults surveyed, 55% of them responded "approve" to the question, "Do you generally approve or disapprove of the 2010 Affordable Care Act, signed into law by President Obama, that restructured the U.S. healthcare system?" The article also notes that the ACA had never before showed majority support in Gallup polling, but that 48% of the sample said "approved" the first time the current version of the question was asked in November 2012.

- a) If this was a simple random sample of the 1023 adults in 2017, what would the conservative 95% confidence interval margin of error be? **Show work.**
- b) Gallup gives a 95% confident margin of error of plus or minus 3% for the 2012 poll, which had 48% of the sample "approved." Give the lower and upper boundaries for 95% confidence intervals.
- c) Give the lower and upper boundaries for 95% confidence intervals for the "approve" results for the 2017 poll, which had 55% of the sample "approved" and a margin of error plus or minus 4%.
- d) Assume for this question that Gallup's sample is indeed representative of the population (e.g., no bias). If you take into account Gallup's margin of errors when you interpret the poll results, is it statistically valid for Gallup to make the statements it did in the headline "Affordable Care Act Gains Majority Approval for First Time? Critique and explain.
- e) Assume little to no bias and truly a random sample. If a polling company conducted 100 such polls with a 95% confidence interval, then about how many of them are likely to include the true population percentage?
- f) Is there any way to know which intervals from the 100 polls contain the true percentage and which ones don't?
- g) Explain what 95% confidence interval means.
- h) Explain what margin of error means.
- i) Explain why a one-time phone call for a survey has too much bias in it. How could we still use phones but create a somewhat better sample?