

1. (41 pts) Diet and blood pressure
 - (a) 5 pts. B
Individuals were randomly sampled. Treatments were not.
 - (b) 5 pts. haters
The sd for the haters (20.23) is larger than that for eaters (11.24)
 - (c) 5 pts. eaters
The se for the eaters $(11.24/\sqrt{16} = 2.81)$ is smaller than that for haters $(20.23/\sqrt{49} = 2.89)$.
Note: looking just at the sample size isn't sufficient. larger N \Rightarrow more precise only when the sd's are the same (or similar).
 - (d) 5 pts. 135.4 (or 135.3)
Notes: $se/3$ for eaters $= 2.81/3 = 0.94$, so report to the 10'ths place
 - (e) 5 pts. Yes, equal variances is appropriate
The ratio of sd's, $20.23/11.24$ is less than 2
 - (f) 3 pts. 63
Note: Calculated as $n_1 + n_2 - 2 = 49 + 16 - 2 = 63$
 - (g) 5 pts. > 0.05
The 95% confidence interval includes 25
 - (h) 3 pts. 99% CI is wider
Various explanations possible: 99% CI includes more values, so it's wider. 99% uses bigger T quantiles
 - (i) 5 pts. No
This is a causal claim. Treatments were not randomly assigned

2. (34 pts.) Apple blotches
 - (a) 3 pts. apple
 - (b) 3 pts. apple
 - (c) 3 pts. Yes
A randomization test only assumes independence. Treatments were randomly assigned
 - (d) 3 pts. 1.13
Note: Computed as $3.21 - 2.08$
 - (e) 5 pts. 0.026
17 randomized values < -1.13 . 8 values > 1.13 . Using a random sample of permutations so $p = (R+1)/(N+1) = 26/1001 = 0.026$
 - (f) 3 pts. Yes
The theoretical T distribution closely approximates the randomization distribution
Note: Removed from exam after grading.
 - (g) 5 pts. 0.59
Statement asks for median bleach / median tap. Log-scale estimate for bleach - tap is $0.472 - 0.996 = -0.524$. The multiplicative effect is $\exp(-0.524) = 0.592$.
Note: If you answered 1.69, you calculated the log-scale difference as tap - bleach $= 0.524$.

- (h) 3 pts. (0.40, 0.88)
Calculated as: $(\exp(-0.92), \exp(-0.13))$
- (i) 3 pts. results are the same for both analyses
Note: Because log transformation doesn't change the ranks of the observations
- (j) 3 pts. Doesn't produce confidence intervals (or doesn't easily)

3. (25 pts.)

- (a) 0 pts. Not graded
- (b) 5 pts. Paired T-test, $p < 0.0001$ (or the very small number from the printout)
Note: Data are paired. No sign of non-normality in the differences
- (c) 5 pts. Very strong evidence of a difference in means (or medians if T-test)
Note: If you said "strong evidence", you lost 1 point. Your conclusion should distinguish between some evidence and stronger evidence. See page 2 of week 1 notes or display 2.12 in the text. "Significant" lost more points because that just means $p < 0.05$.
- (d) 5 pts. -111.9, (-128.2, -95.6)
- (e) 10 pts. Answer depends on the choice of test - for paired T-test:

Assumption	Diagnostic	Decision	Explanation
Independence	study design	OK	Ponds randomly sampled
Normality	QQ plot of differences	OK	Dots fall along line

Notes: Here's why you might have lost points here:

Including equal variances as an assumption: It isn't relevant for a paired data analysis

Omitting independence: Saying paired is not sufficient, because paired data may or may not be independent. Imagine sampling each pond twice, once in the north and once in the south. Repeated in 2017. Data still paired, but there are two differences in each pond.

Bad explanation for independence: need to look at how the data were collected. Ponds were randomly sampled, so observations are independent. Sampling years can't be randomized. If there is an "e.u." it is the observation (2012 or 2017).

Bad explanation for normality: This has to be normality of the differences. Should be judged by a QQ plot of the differences and noting that the observations fall close to the line.

If you have questions about the grading, please come to office hours or ask during lab.