Question 1

Allotted time: 25 minutes (plus 5 minutes to submit)

An online company is hoping to sell a new product by placing a pop-up advertisement on their website. There are two different designs for the advertisement, and the company would like to determine which one is more effective, as measured by clicks and purchases. For the first 200 visitors to the updated website, half were randomly assigned to receive advertisement 1 and half to receive advertisement 2. The two-way table summarizes the results:

Customer Behavior

	Did not click	Clicked, no purchase	Clicked, made purchase	Total
Advertisement 1	70	8	22	100
Advertisement 2	64	20	16	100
Total	134	28	38	200

(a) Is this study an observational study or an experiment? Explain your answer.

(b) i. Provide one piece of evidence for why advertisement 1 is more effective.

ii. Provide one piece of evidence for why advertisement 2 is more effective.

(c) One customer from the study will be selected at random. Are the events "advertisement 1" and "made purchase" independent? Justify your answer based on probabilities calculated from the table above.

The online company conducted a test of the hypotheses

$$H_0 : p_1 - p_2 = 0$$
$$H_a : p_1 - p_2 \neq 0,$$

where p_1 is the proportion of customers similar to those in the study given advertisement 1 that would click and make a purchase and p_2 is the proportion of customers similar to those in the study given advertisement 2 that would click and make a purchase. The conditions for inference have been met.

(d) One of the conditions for inference that was met is that $n\hat{p}_c \geq 10$ and $n(1-\hat{p}_c) \geq 10$ for each group, where $\hat{p}_c$ is the combined (or pooled) proportion. Explain why it is necessary to satisfy this condition.

(e) The test resulted in a *p*-value of 0.2795. Interpret what this *p*-value measures in the context of this study.

(f) Based on the *p*-value, what conclusion should the online company make?

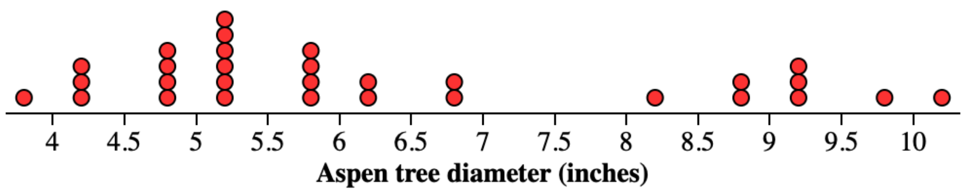
Question 2

Allotted time: 15 minutes (plus 5 minutes to submit)

A park ranger at a large national park wants to estimate the mean diameter of all the aspen trees in the park. The park ranger believes that due to environmental changes, the aspen trees are not growing as large as they were in 1975.

(a) Data collected in 1975 indicate that the distribution of diameter for aspen trees in this park was approximately normal with a mean of 8 inches and a standard deviation of 2.5 inches. Find the approximate probability that a randomly selected aspen tree in this park in 1975 would have a diameter less than 5.5 inches.

The park ranger selects a random sample of 30 aspen trees from the park in 2020 and measures their diameters. A dotplot of the diameters is shown below.



(b) Describe the distribution of aspen tree diameters from the sample.

(c) Assume the conditions for inference have been met. The park ranger uses the sample data to construct a 95% confidence interval for the mean diameter of all aspens in the park in 2020 as 5.6 to 7.0 inches. Calculate the point estimate and the margin of error.

(d) Based on the confidence interval, does the park ranger have convincing evidence that the mean diameter for all aspen trees in the park is different than 8 inches, as it was in 1975?

(e) Aspen trees tend to be smaller in the highlands of the park because they are subject to strong winds. While the number of aspen trees in the highlands of the park is about the same as the number of aspen trees in the lowlands of the park, the park ranger is concerned that 24 out of the 30 trees in the random sample came from the highlands. Propose a sampling method to address this concern and describe the benefit of using that sampling method.