4. A farmer wants to know whether there is a difference in the mean number of oranges that grow on trees fertilized with the current fertilizer, Brand C, or a new fertilizer, Brand N. The farmer randomly assigns 58 trees to be treated with Brand C and 58 trees to be treated with Brand N. The number of oranges from each tree is recorded. The summary statistics by fertilizer brand are shown in the table.

Summary Statistics for the Number of Oranges by Fertilizer Brand

	n	Mean	Standard Deviation
Brand C	58	141	15
Brand N	58	148	19

At the $\alpha = 0.05$ significance level, do the data provide convincing statistical evidence that the mean number of oranges on trees fertilized with Brand C is different from the mean number of oranges on trees fertilized with Brand N for all trees similar to those in the study?

AP Statistics: **2025**

6. Stefan, a psychologist, conducted a study to investigate the effect of time of day on reading comprehension in children. One hundred children volunteered, with their parents' consent, to participate in the study. Fifty of the children were randomly assigned to read a story at 9 a.m. and then answer 25 questions about it. The remaining 50 children were assigned to read the same story at 3 p.m. and answer the same 25 questions. The reading comprehension for each child was measured by a reading score, which was determined by the number of questions that were answered correctly about the story. Stefan is interested in comparing the mean reading scores for the two times of day. Table 1 shows the results of Stefan's study.

Table 1: Summary Statistics of Reading Scores

	n	Mean	Standard Deviation
9 a.m.	50	15.2	4.12
3 p.m.	50	17.9	4.43

Stefan found the conditions for inference were met and conducted a two-sample *t*-test for the difference in two population means. Let μ_{AM} represent the mean reading score for all children, similar to those in the study, who would read the story at 9 a.m. Let μ_{PM} represent the mean reading score for all children, similar to those in the study, who would read the story at 3 p.m. Stefan's hypotheses are as shown.

$$\text{H}_0\text{: } \mu_{\text{AM}} = \mu_{\text{PM}}$$
$$\text{H}_a\text{: } \mu_{\text{AM}} \neq \mu_{\text{PM}}$$

A. The *p*-value for Stefan's hypothesis test was 0.002. State an appropriate conclusion, at the 5 percent significance level, for Stefan's test in the context of the investigation. Justify your answer.

B. Explain why it was appropriate for Stefan to conduct a two-sample *t*-test for the difference in two population means instead of a paired *t*-test for the population mean difference.

C. Researchers are usually interested in the practical importance of their results as well as the statistical significance of the hypothesis test. The practical importance of the results indicates whether the observed results are meaningful in real life. For example, in an investigation of the heights of two groups of students, a difference in the two group means of 3.8 inches is much more meaningful, or has more practical importance, than a difference in the two group means of only 0.2 inches.

One indicator of practical importance is effect size. A common method for measuring effect size for the difference in two group means is Cohen's *d* coefficient. Cohen's *d* coefficient compares the absolute value of the difference in the means of the two groups to the pooled variability of the observed data values from the two groups.

Cohen's *d* coefficient can be calculated using $d = \dfrac{|\bar{x}_1 - \bar{x}_2|}{s_p}$, where s_p represents the pooled standard deviation, $\bar{x}_1$ represents the sample mean for the first group, and $\bar{x}_2$ represents the sample mean for the second group. When the sizes of the groups are equal, s_p is calculated as $s_p = \sqrt{\dfrac{s_1^{\,2} + s_2^{\,2}}{2}}$, where s_1 represents the sample standard deviation for the first group and s_2 represents the sample standard deviation for the second group.

Consider the summary statistics from Stefan's study in Table 1.

i. Calculate Cohen's *d* coefficient for Stefan's study. Show your work.

ii. Higher values of Cohen's *d* indicate greater practical importance and lower values of Cohen's *d* indicate less practical importance. Typically, we use the intervals listed in Table 2 to help interpret practical importance.

Table 2: Guidelines for Interpreting Cohen's *d* Coefficient

Cohen's *d* Coefficient	Practical Importance
$0 \le d \le 0.20$	Not very meaningful in real life
$0.20 < d < 0.80$	Somewhat meaningful in real life
$d \ge 0.80$	Very meaningful in real life

Based on your answer to part C (i) and the information in Tables 1 and 2, describe the practical importance of Stefan's results, in context.

D. Suppose the results of Stefan's study, summarized in Table 1, instead had a standard deviation for the 9 a.m. reading scores, s_1, and a standard deviation for the 3 p.m. reading scores, s_2, that were both greater than 4.43. Assume the group sample sizes and the means are not changed.

i. Would the Cohen's *d* coefficient in this new situation be smaller than, larger than, or the same as the Cohen's *d* coefficient calculated in part C (i)? Explain your answer.

ii. Does the Cohen's *d* coefficient described in part D (i) indicate that Stefan's observed difference in the means in the new situation would have more practical importance than, less practical importance than, or the same practical importance as what was originally determined in part C (ii)? Explain your answer.

STOP

END OF EXAM

AP Statistics: **2018**

4. The anterior cruciate ligament (ACL) is one of the ligaments that help stabilize the knee. Surgery is often recommended if the ACL is completely torn, and recovery time from the surgery can be lengthy. A medical center developed a new surgical procedure designed to reduce the average recovery time from the surgery. To test the effectiveness of the new procedure, a study was conducted in which 210 patients needing surgery to repair a torn ACL were randomly assigned to receive either the standard procedure or the new procedure.

(a) Based on the design of the study, would a statistically significant result allow the medical center to conclude that the new procedure causes a reduction in recovery time compared to the standard procedure, for patients similar to those in the study? Explain your answer.

(b) Summary statistics on the recovery times from the surgery are shown in the table.

Type of Procedure	Sample Size	Mean Recovery Time (days)	Standard Deviation Recovery Time (days)
Standard	110	217	34
New	100	186	29

Do the data provide convincing statistical evidence that those who receive the new procedure will have less recovery time from the surgery, on average, than those who receive the standard procedure, for patients similar to those in the study?

AP Statistics: **2014**

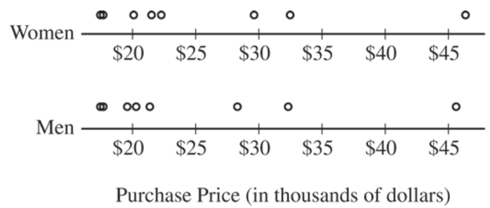
5. A researcher conducted a study to investigate whether local car dealers tend to charge women more than men for the same car model. Using information from the county tax collector's records, the researcher randomly selected one man and one woman from among everyone who had purchased the same model of an identically equipped car from the same dealer. The process was repeated for a total of 8 randomly selected car models.

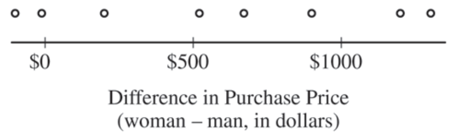
The purchase prices and the differences (woman – man) are shown in the table below. Summary statistics are also shown.

Car model	1	2	3	4	5	6	7	8
Women	\$20,100	\$17,400	\$22,300	\$32,500	\$17,710	\$21,500	\$29,600	\$46,300
Men	\$19,580	\$17,500	\$21,400	\$32,300	\$17,720	\$20,300	\$28,300	\$45,630
Difference	\$520	–\$100	\$900	\$200	–\$10	\$1,200	\$1,300	\$670

	Mean	Standard Deviation
Women	\$25,926.25	\$9,846.61
Men	\$25,341.25	\$9,728.60
Difference	\$585.00	\$530.71

Dotplots of the data and the differences are shown below.





Do the data provide convincing evidence that, on average, women pay more than men in the county for the same car model?