

NAME: Answer Key

Directions: Be sure to show all of your work. An answer alone will not receive any credit. You must show a formula or how you arrived at your answer. Partial credit will be given on all problems.

²⁴ 1. True/False questions.

- T F ³a. The Median and Trimmed Average are both robust measures of location.
- T F ³b. A population is only a set of numeric measurements.
- T F ³c. To be useful about making conclusions about the population, a sample should be representative of the population.
- T F ³d. A symmetric distribution would never have outliers.
- T F ³e. The range and IQR are both robust measures of spread.
- T F ³f. Normal distributions are always symmetric.
- T F ³g. Symmetric distributions are always Normal.
- T F ³h. Sample spaces are a list of possible outcomes from an experiment.

24
24

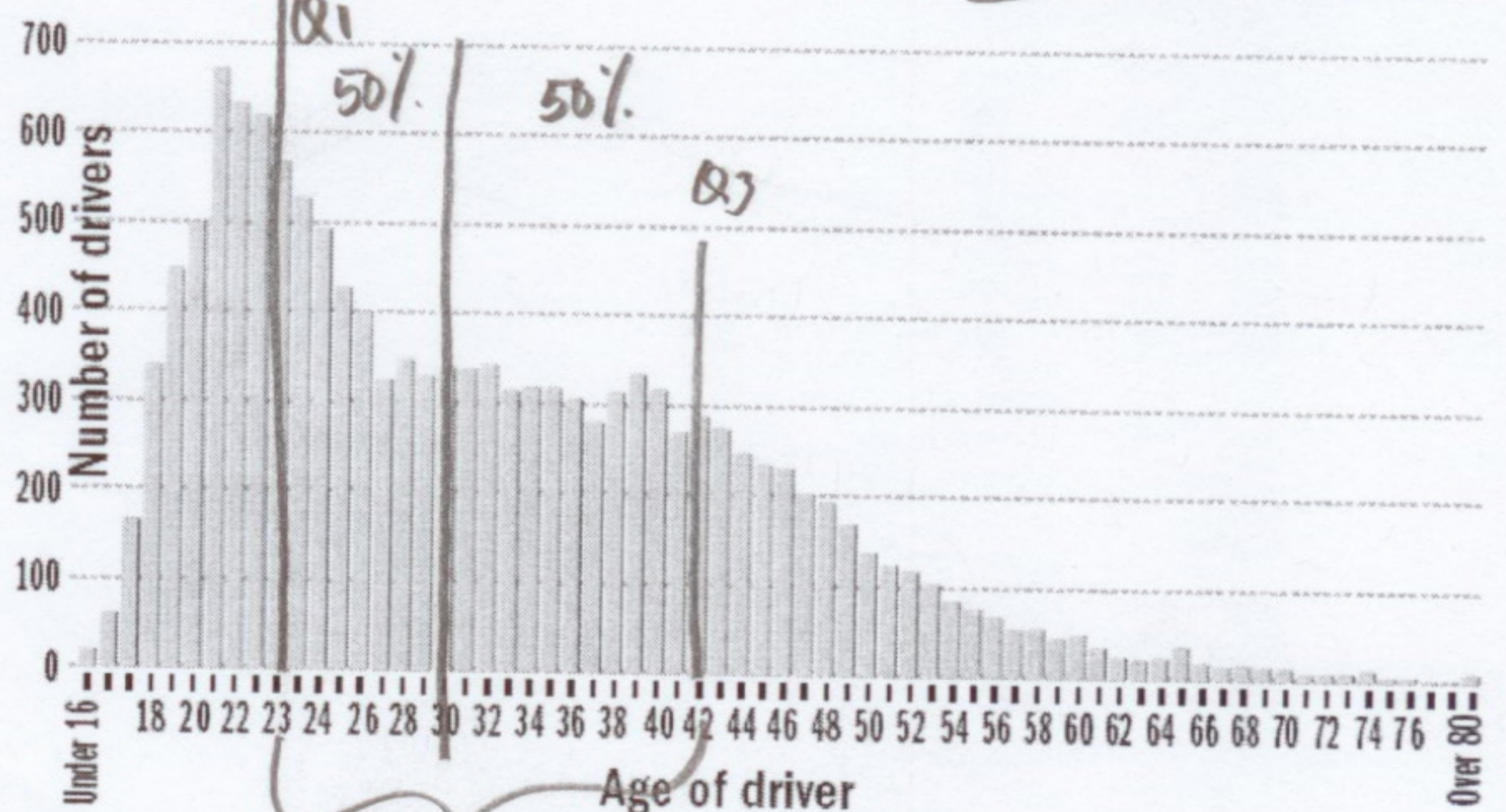
¹⁵ 2. Use the figure below about drunk-driving crashes from the Cincinnati *Enquirer* to answer the True/False questions below.

- T F ³a. This distribution is skewed left.
- T F ³b. The variance would be larger than the IQR.
- T F ³c. The average would be greater than the median.
- T F ³d. The IQR is approximately 50.
- T F ³e. Because of the outliers, the IQR would be a better measure of spread than the SD.

15
15

Young and reckless

Median age of a drunken driver involved in a crash was 29. The curve peaked at 21.



$IQR = 42 - 23 = 19$

⁵ 3. For the weekly rainfall data in SW Ohio seen in a later problem, the variance of weekly total rainfalls is approximately 1.2030. For an observation to be declared an outlier, how far away from the average would it need to be?

using z-score $sd = \sqrt{var} = \sqrt{1.2030} = 1.0968$

$|z| \geq 3 \text{ or } 4 \Rightarrow \frac{|obs - avg|}{sd} \geq 3 \text{ or } 4 \Rightarrow |obs - avg| \geq 3 \text{ or } 4 \times sd$

or $3 \text{ or } 4 \times \sqrt{1.2030}$ or $3 \text{ or } 4 \times 1.0968$

so $|distance| > 3.2904 \text{ or } 4.3875$

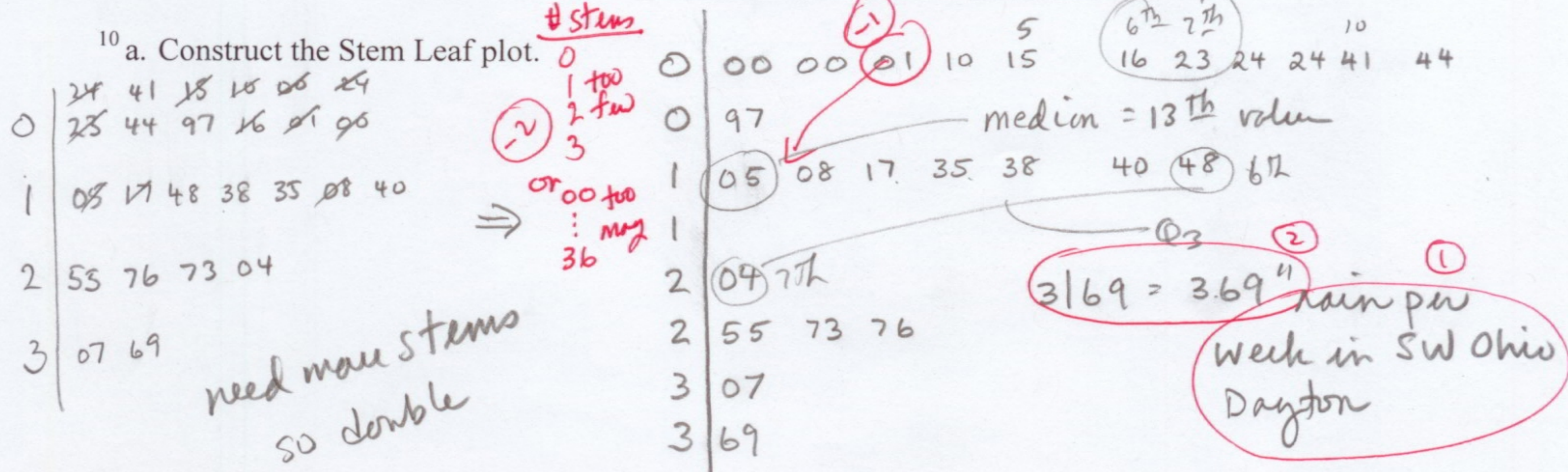
(3) if obs > 3sd

5
5

~~or if obs~~

35 4. Weekly rainfalls in SW Ohio have varied quite a bit this year. Below are the weekly (Sunday thru Saturday) rainfall totals for the Dayton, Ohio NOAA site; use this information to answer the questions that follow.

March	2.55	0.23	2.76	1.05	
April	0.44	0.97	0.16	1.17	
May	1.48	3.07	0.24	0.41	
June	3.69	1.38	0.15	1.35	
July	1.08	2.73	0.10	2.04	0.00
August	1.40	0.01	0.00	0.24	



10 b. Obtain the Five Number Summary, **BY HAND AND SHOW YOUR WORK!!!!** max = 3.69 min = 0.00

- Med = 13th value since $n=25$ odd so median = 1.05
- Q_1 = median of 12 values below Median so avg of 6th + 7th = $.16 + .23 / 2 = 0.1950$
- Q_3 = " of " above median so avg of $\frac{1.48 + 2.04}{2} = 1.76$ so {0.00, 0.1950, 1.05, 1.76, 3.69}

5 c. What **TWO** facts about your 5# summary clearly illustrate this distribution is not normal? Since normals are symmetric any indication of skew \Rightarrow Not normal. Since skew \Rightarrow non symmetry of distances btwn 5# summary. Relative to what = -2

So distance btwn max \rightarrow med \neq min \rightarrow med ; max $\rightarrow Q_3 \neq$ min $\rightarrow Q_1$; $Q_3 \rightarrow$ med $\neq Q_1 \rightarrow$ med

① 2.64 \neq 1.05 ② 1.93 \neq 0.195 ③ 0.71 \neq 0.8550

5 d. To summarize the center of this distribution of rainfall amounts, would you use the average or median? Why? BRIEFLY!!!!

B/c not symmetric distrib skewed, median better measure of center

avg = -2 no outlier

5 e. Using the upper fence, are there any outliers on the upper end of the distribution?

Upper fence = $Q_3 + 1.5(IQR) = Q_3 + 1.5(Q_3 - Q_1) = 1.76 + 1.5(1.76 - 0.195)$

$= 1.76 + 1.5(1.5650) = 1.76 + 2.3475 = 4.1075$

no outliers since max is within upper fence.