

Statistics

Comp1 practice test

1. Given the data list below find the following: (calculate all by hand, show your work)

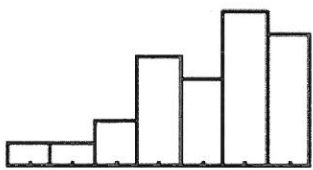
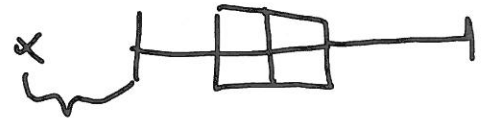
,1,2,3,3,5,7,5

- a) Mean
- b) Median
- c) Standard Deviation
- d) Q1
- e) Q3
- f) IQR
- g) Range
- h) Variance

2. Last weekend the police ticketed 18 men whose mean speed was 72 miles per hour, and 30 women whose mean was 64 mile per hour. Overall what was the mean speed of all the people ticketed?

3. Which is true of the data shown in the histogram?

- I. The distribution is skewed to the right. **F**
- II. The mean is probably smaller than the median. **T**
- III. We should use median and IQR to summarize these data. **T**



- a) I
- b) II
- c) III
- d) II and III
- e) I, II, and III

4. If we want to discuss any gaps and clusters in a data set, which of the following should not be chosen to display the data set?

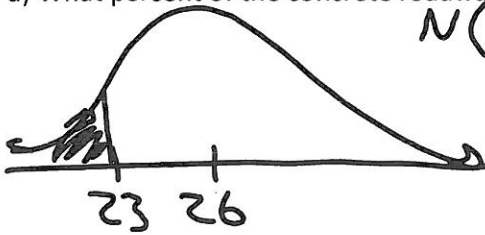
- a) histogram
- b) stem-and-leaf plot
- c) boxplot
- d) dotplot

5. Suppose that a normal model describes fuel economy (miles per gallon) for automobiles and that a Saturn has a standardize score (z-score) of 2.2. This means that saturns.....

- a) get 2.2 miles per gallon
- b) get 2.2 times the gas mileage of the average car
- c) get 2.2 mpg more than the average car
- d) have a standard deviation of 2.2 mpg
- e) achieve a fuel economy that is 2.2 standard deviations better than the average car.

9. A roadway construction process uses a machine that pours concrete onto the roadway and measures the thickness of the concrete so the roadway will up to the required depth in inches. The concrete thickness needs to be consistent across the road, but the machine isn't perfect and it is costly to operate. Since there's a safety hazard if the roadway is thinner than the minimum 23 inch thickness, the company sets the machine to average 26 inches for the batches of concrete. They believe the thickness level of the machine's concrete output can be described by a normal model with a standard deviation of 1.75 inches.

a) What percent of the concrete roadway is under the minimum depth?

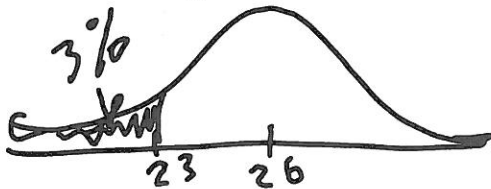


$N(26, 1.75)$

norm cdf
 lower: -10000
 upper: 23
 $\mu = 26$
 $\sigma = 1.75$
 $\approx 4.32\%$
 0.0432

b) The company's lawyers insist that no more than 3% of the output can be under the limit. Because of the ~~expense~~ expenses of operating the machine, they cannot afford to reset the mean to a higher value. Instead they will try to reduce the standard deviation to achieve the "only 3% under" goal. What standard deviation must they attain?

c) Explain what achieving a smaller standard deviation means in this context?



invNorm (area)
 area: 0.03
 $\mu = 0$
 $\sigma = 1$
 $= -1.88$

$$z = \frac{x - \bar{x}}{\sigma}$$

$$\frac{-1.88}{\sigma} = \frac{23 - 26}{\sigma}$$

$$\frac{-1.88\sigma}{\sigma} = \frac{-3}{\sigma}$$

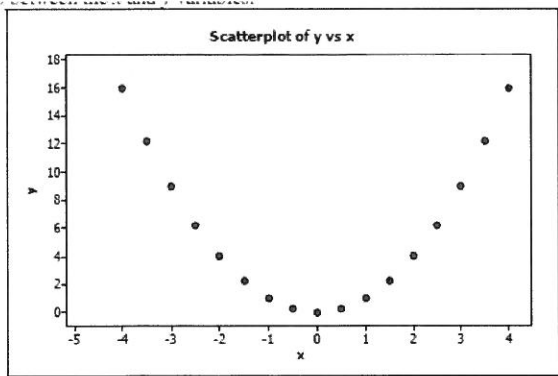
$$\frac{-1.88}{-1.88} = \frac{-3}{-1.88}$$

$$\sigma = 1.60$$

j) Make a histogram and boxplot of the following data by hand.

20,21,23,24,25,22,21,20,18,14,13,12,11,4,23,30,31,32,35,40,35

11. Is there an association between the x and y variables in the graph below?



a) How would you describe the form?

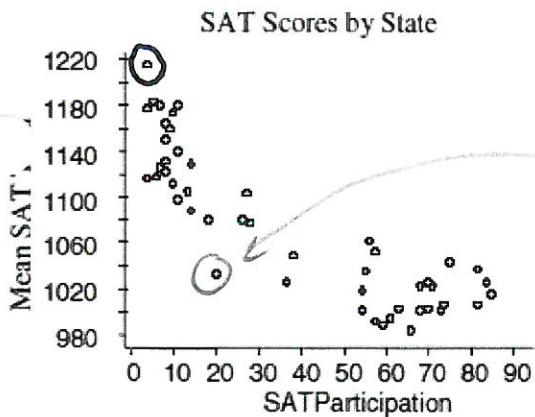
Quadratic, Parabolic

b) Would the correlation coefficient support a relationship in this scenario? If not why?

no because the association is not linear

c) What function would you use to straighten the data

square root because it is the inverse function of a quadratic.



a) Write a few sentences to describe the association.

linear (or exponential), negative, strong, 1 possible outlier

b) Estimate the correlation coefficient (r)

$r \approx -0.78$

c) If the point in the top left corner (4,1215) were removed, would the correlation become stronger, weaker, or remain about the same? Explain briefly.

Weaker, that point contributes a large negative value to $\sum \frac{z_x z_y}{n}$. If it is removed it makes $\sum \frac{z_x z_y}{n}$ smaller.

d) If the point in the very middle (38,1049) were removed, would the correlation become stronger, weaker, or remain out the same? Explain briefly.

Remain the same. $\sum \frac{z_x z_y}{n}$ is relatively unaffected by this point because $z_x z_y$ is small for this point. It is close to (\bar{x}, \bar{y}) .

15. A company manufactures a polypropylene rope in six different sizes. To assess the strength of the ropes they test a sample to see how much force the rope will hold without breaking. The table shows the results of the tests.

Diameter (mm)	Strength (kN)
4	60
7	157
10	254
12	334
15	551
20	938

a) Use your calculator to plot the data and describe the association.

b) Calculate the correlation coefficient.

$r \approx 0.98$