

1. (5 points) Which of the following would be the most appropriate numerical summary of a skewed distribution?

- A. the median and quartiles
- B. the five number summary
- C. \bar{x} and s
- D. \bar{x} and the median
- E. \bar{x}

2. (5 points) What values can a correlation r can possibly take?

- A. $-1 \leq r \leq 1$
- B. $0 \leq r \leq 1$
- C. $r \geq 0$
- D. any real number

3. (5 points) Which of the following quantities is resistant to extreme observations?

- A. the mean
- B. standard deviation
- C. correlation
- D. the median

4. (5 points) If a distribution is skewed to the left,

- A. the mean and median are equal.
- B. the mean is greater than the median.
- C. the mean is less than the median.

5. (20 points) The most popular colors for cars and light trucks vary with region and over time. In North America white remains the top color choice, with black the top choice in Europe and silver the top choice in South America. Here is the distribution of the top colors for vehicles sold globally in 2010:

Color	Popularity
<i>Silver</i>	26%
<i>Black</i>	24%
<i>White</i>	16%
<i>Gray</i>	16%
<i>Red</i>	6%
<i>Blue</i>	5%
<i>Beige, brown</i>	3%
<i>Other</i>	

(a) Fill in the percent of vehicles that are in other colors. Make a graph to display the distribution of color popularity.

(b) Would it also be correct to make a pie chart?

6. (20 points) Keeping water clean. Keeping water supplies clean requires regular measurement of levels of pollutants. The measurements are indirect; a typical analysis involves forming a dye by a chemical reaction with the dissolved pollutant, then passing light through the solution and measuring its absorbance. To calibrate such measurements, the laboratory measures known standard solutions and uses regression to relate absorbance and pollutant concentration. This is usually done every day. Here is one series of data on the absorbance for different levels of nitrates. Nitrates are measured in milligrams per liter of water.

Nitrates	50	50	100	200	400	800	1200	1600	2000	2000
Absorbance	7.0	7.5	12.8	24.0	47.0	93.0	138.0	183.0	230.0	226.0

(a) Chemical theory says that these data should lie on a straight line. If the correlation is not at least 0.997, something went wrong and the calibration procedure is repeated. Plot the data and find the correlation. Must the calibration be done again?

(b) The calibration process sets nitrate level and measures absorbance. The linear relationship that results is used to estimate the nitrate level in water from a measurement of absorbance. What is the equation of the line used to estimate nitrate level?

(c) What does the slope of this line say about the relationship between nitrate level and absorbance? What is the estimated nitrate level in a water specimen with absorbance 40?

(d) Do you expect estimates of nitrate level from absorbance to be quite accurate? Why?

7. (20 points) Automated manufacturing operations are quite precise but still vary, often with distributions that are close to Normal. The width in inches of slots cut by a milling machine follows approximately the $N(0.8750, 0.0012)$ distribution. The specifications allow slot widths between 0.8720 and 0.8780 inch. What proportion of slots meet these specifications?

8. (20 points) According to the Census Bureau's 2010 Current Population Survey, the mean and median 2009 income of people at least 25 years old who had a bachelor's degree but no higher degree were \$46,931 and \$58,762.

Which of these numbers is the mean and which is the median? Explain your reasoning.