



ORGANIZING DATA WORKSHEET #1

1. An environmental engineering group has collected core samples from the earth to determine if they are contaminated. The samples were weighed, as follows:

15 40 69 65 10 60 34 20 43 40 10 25 44
65 32 45 30 36 45 50 35 39 50 46 55 31

- a. Is the data categorical or quantitative? If quantitative, specify discrete or continuous.
- b. Determine if the mean and standard deviation is an appropriate way to summarize this data. If so, find the mean and standard deviation with your calculator. If not, use the median and IQR and explain why you made your choice.
2. A farmer wonders if his crops grow better in sun or in shade. He measures the amount of fruit gathered from a sample of 50 trees growing in full sun and from a sample of 50 trees growing in mostly shade. The five-number summaries, (minimum, Q1, median, Q3, maximum), follow:
- Amount of fruit gathered from sunny trees: (10.8, 13.3, 22.1, 58.1, 100) bushels
Amount of fruit gathered from shady trees: (10.9, 18.5, 29.3, 61.5, 127) bushels
- a. On the same set of axes, create a boxplot for each type of tree. Compare the center, shape, spread, and outliers of amount of fruit from sunny trees to the amount from shady trees.
- b. The farmer wants to summarize the amount of fruit from each type of orchard with a mean and standard deviation. Is that appropriate? Explain.

3. Megan is an industrial engineer for Bowler Cola Company. She takes a random sample of cola cans from the production line each day to determine if the product meets various specifications. One of the measurements she records is the mass (in grams) of the filled cans. The following sorted data are from of a sample of 30 regular and diet cola cans.

Note: The data is sorted so it is easy to work with without a statistical calculator.

Regular

361	362	363	365	366	366	367	367	367	368
368	368	369	369	369	369	370	370	370	371
370	371	371	371	373	375	375	376	376	380

Diet

349	349	350	351	353	353	353	354	354	354
354	355	355	355	356	357	358	361	361	361
361	361	361	362	362	363	364	365	366	366

- Find the five number summary (min, third quartile, median, first quartile, max) for each soda.
- Make a boxplot for each type of soda on the same scale.
- Describe the center, shape, spread and any outliers of each boxplot.
- Compare the two samples.
- Each can is marked as containing 12 fluid ounces. Twelve ounces is about 341 grams. What would your analysis be for the Bowler Cola Company?

4. The city of Waynesboro is trying to decide whether to initiate a composting project where each residence would be provided with a dumpster for garden and yard waste. The city manager needs some measure of assurance that the citizens will participate before launching the project, so he chooses a random sample of 25 homes and provides them with the new dumpster for yard and garden waste. After one week the contents of each dumpster is weighed (in pounds) before processing. The sorted data is shown below:

0	0	0	0	1.7	2.6	2.9	4.2	4.4
5.1	5.6	6.4	8.0	8.9	9.7	10.1	11.2	13.5
15.1	16.3	17.7	21.4	22.0	22.2	36.5		

- a. Create a histogram for the data. Use an interval of 0 to 42 pounds on the x -axis and a class interval of 6 pounds.
- b. Describe the center, shape, spread and outliers.
- c. What is a better measure of center for this distribution the mean or median and why?
- d. What is a better measure of spread the standard deviation or IQR and why?
- e. The city can sell the compost, and engineers estimate the program will be profitable if each home averages at least 9 pounds of material. The city manager sees the mean is nearly 10 pounds and is ready to order dumpsters for every residence. What advice would you give him?