Chapter Project

The Breakfast of Champions

A nutritionist at the Food and Drug Administration is studying the effects of cereal marketing on family meal choices. In particular, she would like to understand how cereal manufacturers market their products in grocery stores. She became interested in doing this study after noticing how the cereal was being restocked one day in her local grocery store. The store personnel were restocking the cereal shelves based on a reference sheet that told them where everything was to be placed. The placement of each cereal brand seemed very deliberate.

Data

The data can be found at stat.hawkeslearning.com **Data Sets > Cereal Data.**

To gather data for her study, the nutritionist goes to the local grocery store and records data about cereal nutritional claims and shelf location for 77 cereals.

1. Identify the population and the sample in this scenario.
2. Consider the variables in the data set. Identify the variables that are qualitative and those that are quantitative.
3. Consider the variable *Shelf*. This variable is the shelf position of the cereal (bottom, middle, top) starting from the floor up. Based on your experience at the grocery store, do you think that the shelf position is related to the nutritional content of the cereal? Why or why not?

To see whether the shelf position is associated with one measure of nutritive value, the amount of sugar, look at the data for the variable *Sugars*. Compare the sugar content of cereals on each shelf by making a separate frequency table and histogram for the sugar content of the cereals on each shelf: a total of three frequency tables and three histograms. Use the sugar content values as they are – do not factor in the serving size. (The data for one of the cereals, Quaker Oatmeal, is missing. Just continue with what is available. That's the way it is in real life – values are missing, files are incomplete, etc.)

4. Make a frequency table for each shelf. (Hint: It might help to sort the variable Sugars by shelf location.)
   a. Use the same classes for all shelves.
   b. Use 6 classes.
5. Make a histogram for each shelf.
   a. Use graph paper and work neatly or use your calculator.
   b. Use the same scales for your histograms so you can compare the data easily.
   c. Title each histogram and label the axes.
6. Briefly describe the distribution in each histogram with respect to shape. Based on your histograms, which shelf position has cereals with the most sugar?
7. Consider your histograms for sugar content. Is the shelf position of a cereal related to its nutritive value as measured by sugar content? Explain your reasoning. What kinds of cereals are on each shelf?
8. What further data gathering would you recommend?
9. Does your analysis of the data generate any other questions?
10. Can you suggest how you might obtain data to answer these questions?

Source: https://www.kaggle.com/crawford/80-cereals/data