

## 1.6: Exercises

A product development manager at the campus bookstore wants to make sure that the backpacks being sold there are strong enough to carry the heavy books students carry around campus. The manager decides she will collect some data on how heavy are the bags/packs/suitcases students are carrying around at the moment, by stopping the next 100 people she meets at the center of campus and measuring.

What are the individuals in this study? What is the population? Is there a sample – what is it? What is the variable? What kind of variable is this?

During a blood drive on campus, 300 donated blood. Of these, 136 had blood of type  $O$ , 120 had blood of type  $A$ , 32 of type  $B$ , and the rest of type  $AB$ .

Answer the same questions as in the previous exercise for this new situation.

Now make at least two visual representations of these data.

Go to [the Wikipedia page](#) for “Heights of Presidents and Presidential Candidates of the United States” and look only at the heights of the presidents themselves, in centimeters ( $cm$ ).

Make a histogram with these data using bins of width 5. Explain how you are handling the edge cases in your histogram.

1cm

Suppose you go to the supermarket every week for a year and buy a bag of flour, packaged by a major national flour brand, which is labelled as weighing  $1kg$ . You take the bag home and weigh it on an extremely accurate scale that measures to the nearest  $1/100^{th}$  of a gram. After the 52 weeks of the year of flour buying, you make a histogram of the accurate weights of the bags. What do you think that histogram will look like? Will it be symmetric or skewed left or right (which one?), where will its center be, will it show a lot of variation/spread or only a little? Explain why you think each of the things you say.

What about if you buy a  $1kg$  loaf of bread from the local artisanal bakery – what would the histogram of the accurate weights of those loaves look like (same questions as for histogram of weights of the bags of flour)?

If you said that those histograms were symmetric, can you think of a measurement you would make in a grocery store or bakery which would be skewed; and if you said the histograms for flour and loaf weights were skewed, can you think of one which would be symmetric? (Explain why, always, of course.) [If you think one of the two above histograms was skewed and one was symmetric (with explanation), you don't need to come up with another one here.]

Twenty sacks of grain weigh a total of  $1003kg$ . What is the mean weight per sack?

Can you determine the median weight per sack from the given information? If so, explain how. If not, give two examples of datasets with the same total weight but different medians.

For the dataset  $\{6, -2, 6, 14, -3, 0, 1, 4, 3, 2, 5\}$  which we will call  $DS_1$ , find the mode(s), mean, and median.

Define  $DS_2$  by adding 3 to each number in  $DS_1$ . What are the mode(s), mean, and median of  $DS_2$ ?

Now define  $DS_3$  by subtracting 6 from each number in  $DS_1$ . What are the mode(s), mean, and median of  $DS_3$ ?

Next, define  $DS_4$  by multiplying every number in  $DS_1$  by 2. What are the mode(s), mean, and median of  $DS_4$ ?

Looking at your answers to the above calculations, how do you think the mode(s), mean, and median of datasets must change when you add, subtract, multiply or divide all the numbers by the same constant? Make a specific conjecture!

1cm

There is a very hard mathematics competition in which college students in the US and Canada can participate called the **William Lowell Putnam Mathematical Competition**. It consists of a six-hour long test with twelve problems, graded 0 to 10 on each problem, so the total score could be anything from 0 to 120.

The median score last year on the Putnam exam was 0 (as it often is, actually). What does this tell you about the scores of the students who took it? Be as precise as you can. Can you tell what fraction (percentage) of students had a certain score or scores? Can you figure out what the quartiles must be?

Find the range,  $IQR$ , and standard deviation of the following sample dataset:

$$DS_1 = \{0, 0, 0, 0, 0, .5, 1, 1, 1, 1, 1\} \quad . \quad (1.6.1)$$

Now find the range,  $IQR$ , and standard deviation of the following sample data:

$$DS_2 = \{0, .5, 1, 1, 1, 1, 1, 1, 1, 1, 1\} \quad . \quad (1.6.2)$$

Next find the range,  $IQR$ , and standard deviation of the following sample data:

$$DS_3 = \{0, 0, 0, 0, 0, 0, 0, 0, 0, .5, 1\} \quad . \quad (1.6.3)$$

Finally, find the range,  $IQR$ , and standard deviation of sample data  $DS_4$ , consisting of 98 0s, one .5, and one 1 (so like  $DS_3$  except with 0 occurring 98 times instead of 9 time).

What must be true about a dataset if its range is 0? Give the most interesting example of a dataset with range of 0 and the property you just described that you can think of.

What must be true about a dataset if its  $IQR$  is 0? Give the most interesting example of a dataset with  $IQR$  of 0 and the property you just described that you can think of.

What must be true about a dataset if its standard deviation is 0? Give the most interesting example of a dataset with standard deviation of 0 and the property you just described that you can think of.

Here are some boxplots of test scores, out of 100, on a standardized test given in five different classes – the same test, different classes. For each of these plots,  $A - E$ , describe qualitatively (in the sense of §3.4) but in as much detail as you can, what must have been the histogram for the data behind this boxplot. Also sketch a possible such histogram, for each case.

---

1.6: Exercises is shared under a [not declared](#) license and was authored, remixed, and/or curated by LibreTexts.