

Lesson 3: Representing a Data Collection (distribution)

Definitions:

Distribution → a set/group of data values (#'s that measures things)

e.x. a set of grades = x

dist → $\{88\%, 80\%, 79\%, 75\%\}$

Variable → a letter/symbol that represents the element/data values that vary in a distribution.

3 different ways to describe/qualify a variable:

Qualitative → a variable/distribution measured in words.
e.x. Teacher ratings
e.x. $\{\text{good, bad, excellent}\}$

Discrete quantitative → a variable/distribution measured in numbers that are integers. (whole number)
e.x. number of fb friends
 $\{300, 800, 120\}$

Continuous quantitative → a variable/dist. measured in numbers that are real numbers
 \mathbb{R} (whole numbers and decimals)

e.x. people's height
 $\{5'7, 6'1, 163.5\text{cm}\}$

Question 1:

Anne is an elementary school teacher and loves to pick berries in the summer. The other day she asked her 26 students to guess what her favourite berry was. Here are the results obtained.

Strawberry, raspberry, strawberry, blueberry, cherry, currant, strawberry, strawberry, raspberry, blueberry, strawberry, strawberry, raspberry, raspberry, blueberry, cherry, currant, raspberry, cherry, strawberry, raspberry, raspberry, strawberry, cherry, cherry, blueberry.

→ fruit

a distribution of raw data

a) What does the variable in this distribution represent?

→ fruit / berry let x represent the favourite fruit.

b) Is the variable discrete quantitative, continuous quantitative, or qualitative?

→ qualitative because the variable is measured in words.

c) Represent this distribution of results using the below frequency table.

ANNE'S FAVOURITE BERRIES

nota bene:

BERRY = x	FREQUENCY →
Strawberry	8
Raspberry	7
Blueberry	4
Cherry	5
Current	2

→ the number of times a data values appears in a distribution.

d) How many people's favourite berry is raspberry?

7 people

Do question 2/3 from Part A, handout 1.

PART B: REPRESENTING A GIVEN DISTRIBUTION WITH A RELATIVE FREQUENCY TABLE

Question 1:

A German teacher asks her new students how many languages they speak fluently. Here are the results obtained.

4, 2, 3, 2, 1, 4, 3, 2, 2, 3, 5, 1, 2, 3, 3, 4, 2, 3, 3, 1.

a) What does the variable in this distribution represent?

let x represent the number of languages

b) Is the variable discrete quantitative, continuous quantitative, or qualitative?

discrete quant.

c) Represent this distribution of results using the below relative frequency table.

dist { 4, 2, 3, 2, 1, 4, 3, 2, 2, 3, 5, 1, 2, 3, 3, 4, 2, 3, 3, 1 }

relation \rightarrow connection

TABLE 4.8 - CREATING A RELATIVE FREQUENCY TABLE ACCORDING TO THE NUMBER OF LANGUAGES SPOKEN

LANGUAGE	FREQUENCY	RELATIVE FREQUENCY	PERCENTAGE (%)
1	3	$\frac{3}{20} = 0.15$	15%
2	6	$\frac{6}{20} = 0.3$	30%
3	7	$\frac{7}{20} = 0.35$	35%
4	3	$\frac{3}{20} = 0.15$	15%
5	1	$\frac{1}{20} = 0.05$	5%
Total	20	1	100%

should equal / should equal

(Percentage)

Relative frequency of a data value: how often/frequent the value appears in relation to the whole (the distribution)

$$RF = \frac{\text{frequency/part}}{\text{total/whole}}$$

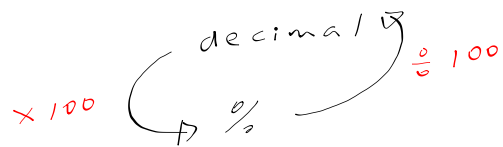
$$\% = \frac{\text{part}}{\text{whole}} \times 100\%$$

Tip: Construct a total row.

d) How many people speak three languages or more?

$7 + 3 + 1 = 11$ people

Do question 2/3 from Part B of 1st handout.



Determining Measures of Central Tendency
of a listed Distribution

trend/most frequent data value in center.

. in order
. from smallest to largest

Question 1:

Chantal is enrolled in adult education and is very proud of having finished her Secondary IV math. Here are the results of her eight exams.

{ 82%, 88%, 86%, 97%, 85%, 75%, 95%, 85% } 95%

Determine the following with respect to the above distribution:

- a) The minimum: 75%
- b) The maximum: 97%
- c) The range: = max - min = 97 - 75 = 22%
- d) The mean: _____
- e) The median: _____
- f) The mode: 85%

Definition: Mode: most frequent data value

mean/average: one data values that summarizes trend of distribution

$$\bar{x} = \frac{x_1 + x_2 + x_3}{3}$$

$$\bar{x} = \frac{x_1 + x_2 + x_3 + x_4}{4}$$

$$\bar{x} = 95\%$$

$$\bar{x} = \frac{\text{sum of values}}{\text{number of values}}$$

82%, 88%, 86%, 97%, 85%, 75%, 95%, 85%

$$\bar{x} = \frac{(82 + 88 + 86 + 97 + 85 + 75 + 95 + 85)}{8}$$

$$\bar{x} = 86.6\%$$

Try questions/leave median for later.