§ 2.1 Describing Qualitative Data

Definition: A *class* is one of the catagories into which qualitative data can be classified.

Definition: The *class frequency* is the number of observations in the data set that fall into a particular class.

Definition: The *class relative frequency* is the frequency divided by the total number of observations in the data set; that is,

$$\text{Class relative frequency} = \frac{\text{class frequency}}{n}$$

where n is the total number of observations.

Example A catagorical survey is taken on how much trust people place in information they read on the internet. Forty people are asked to rate their trust based on the scale:

A = trust in everything they read.
M = trust in most of what they read.
H = Trust in about half of what they read.
S = trust in a small portion of what they read.

Construct a frequency and relative frequency chart for the following data

M	M	M	A	H	M	S	M	H	M
S	M	M	M	M	A	M	M	A	M
M	M	H	M	M	M	H	M	H	M
A	M	M	M	H	M	M	M	M	M

Class	Frequency (f)	Relative Frequency (f/n)
A	4	$4/40 = 0.1 = 10\%$
M	28	$28/40 = 0.7 = 70\%$
H	6	$6/40 = 0.15 = 15\%$
S	2	$2/40 = 0.05 = 5\%$
total:	40	1.00

Definition: the class percentage is the class relative frequency multiplied by 100; that is "written as a percentage".

Example Students in a biology class participate in an exam and the results are classified based on the following chart

Grade (x)	Frequency (f)	Relative Frequency (f/n)
A: 90-100		.08
B: 80-89	36	
C: 70-79	90	
D: 60-69	30	.15
F: 0-59	28	
Total		

Calculate the missing entries.

We don't know, are not given, how many students earned an A or the total number that took the exam. However we do know that 30 students earned a D grade and that was 15% of the total number of students.

Thus, we can use "Rel. Freq. = $\frac{\text{Freq.}}{n}$" to calculate n, the total number of students that took the exam:

$$.15 = \frac{30}{n} \Rightarrow n = \frac{30}{.15} = 200$$

From this, there are two ways to calculate the number of students that earned an A grade:

(i) frequency of A grad = total minus other class frequencies $= 200 - (36+90+30+28) = 16$

(ii) Rel. Freq. A $= \frac{\text{Freq. A}}{n} \Leftrightarrow$

$$.08 = \frac{f}{200} \Leftrightarrow$$

$$f = (.08)(200) = 16$$

Note that both ways generate the same value. ☺

The remaining entries in the chart are then

Rel. Freq. B $= \frac{36}{200} = .18$

Rel. Freq. C $= \frac{90}{200} = .45$

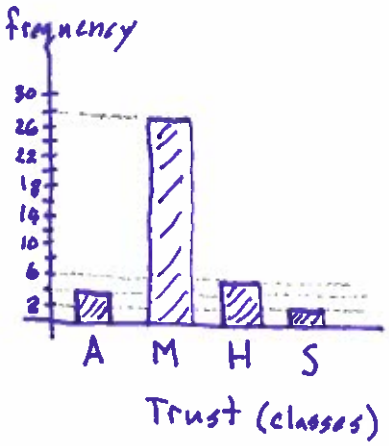
Rel. Freq. F $= \frac{28}{200} = .14$

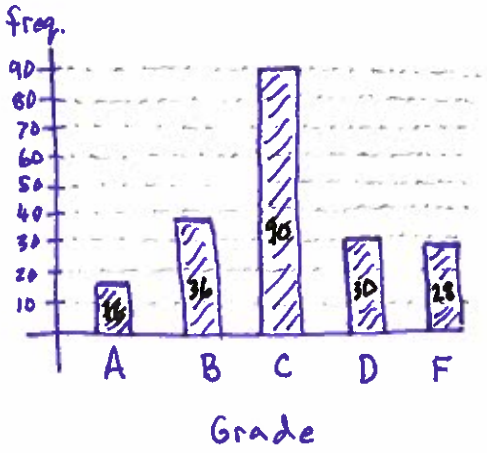
Check to make sure the total of the Relative Frequencies is 1.00:

$$.08 + .18 + .45 + .15 + .14 = 1.00$$

Often it is "easier" to see if the data is put into a graphical form. Three types of graphs we look at in this section are Bar Graphs, Pie Charts, and Pareto diagrams.

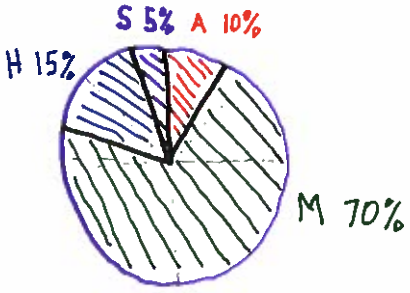
Bar Graphs for the above examples



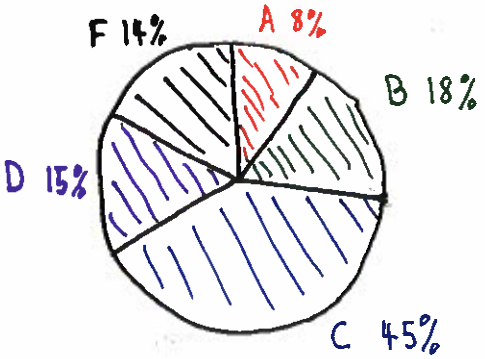


Pie Charts for the above examples
(Requires figuring the angle or parts before hand)

A: 10% of 360° = 36°
M: 70% of 360° = 252°
H: 15% of 360° = 54°
S: 5% of 360° = 18°



A: .08 × 360° = 28.8°
B: .18 × 360° = 64.8°
C: .45 × 360° = 162°
D: .15 × 360° = 54°
F: .14 × 360° = 50.4°



Pareto Diagrams for our examples

(Like a Bar Graph only Largest to smallest freq., left to right.)