

## Frequency distributions, Histograms, and Related Topics

**Example 1:** page 43/2. *Medical: Glucose Testing*

The following data represent glucose blood levels (mg/100 ml) after a 12-hour fast for a random sample of 70 women.

1. Find the class width for six classes
2. Make a frequency table showing class limits, class boundaries, midpoints, frequencies, and relative frequencies.
3. Draw a histogram
4. Draw a relative-frequency histogram

45	66	83	71	76	64	59	59	76	82	80	81	85	77	82	90
87	72	79	69	83	71	87	69	81	76	96	83	67	94	101	94
89	94	73	99	93	85	83	80	78	80	85	83	84	74	81	70
65	89	70	80	84	77	65	46	80	70	75	45	101	71	109	73
73	80	72	81	63	74										

$$\frac{(\text{largest data value} - \text{smallest data value})}{(\text{desired number of classes})} = \frac{(109 - 45)}{6} = 10.67, \text{ thus the class width} = 11$$

**lower class limit** - is the lowest data value that can fit in a class

**upper class limit** - is the highest data value that can fit in a class

**class width** - is the difference between the lower class limit of one class and the lower class limit of the next class

Let's begin to build frequency table:

Class limits lower - upper	Class Boundaries lower - upper	Frequency	Relative Frequency	Class Midpoint
▶ 45 -				
56 -				
67 -				
78 -				
89 -				
100-				

smallest  
data value

$$45+11=56$$

$$56+11=67$$

....

$$55+11=66$$

$$66+11=77$$

....

Class limits lower - upper	Class Boundaries lower - upper	Frequency	Relative Frequency	Class Midpoint
▶ 45 - 55				
56 - 66				
67 - 77				
78 - 88				
89 - 99				
100- 110				

smallest  
data value

$$45+11=56$$

$$56+11=67$$

....

$$55+11=66$$

$$66+11=77$$

....

- fill in the *upper class limits* so that the classes span the entire range of data

Class limits lower - upper	Class Boundaries lower - upper	Frequency	Relative Frequency	Class Midpoint
45 - 55				$(45+55)/2= 50$
56 - 66				$(56+66)/2= 61$
67 - 77				$(67+77)/2= 72$
78 - 88				$(78+88)/2= 83$
89 - 99				$(89+99)/2= 94$
100- 110				$(100+110)/2=105$

$$\text{midpoint} = \frac{(\text{lower class limit} + \text{upper class limit})}{2}$$

Class limits lower - upper	Class Boundaries lower - upper	Frequency	Relative Frequency	Class Midpoint
45 - 55		3		$(45+55)/2= 50$
56 - 66		<del>    </del>    7		$(56+66)/2= 61$
67 - 77		<del>    </del> <del>    </del> <del>    </del> <del>    </del>    22		$(67+77)/2= 72$
78 - 88		<del>    </del> <del>    </del> <del>    </del> <del>    </del> <del>    </del>    27		$(78+88)/2= 83$
89 - 99		<del>    </del>     8		$(89+99)/2= 94$
100- 110		3		$(100+110)/2=105$

*frequency* (for each class) is the number of data values that fall into that class  
(*tally* data)

Class limits lower - upper	Class Boundaries lower - upper	Frequency	Relative Frequency	Class Midpoint
45 - 55	44.5 - 55.5	3		$(45+55)/2= 50$
56 - 66	55.5 - 66.5	7		$(56+66)/2= 61$
67 - 77	66.5 - 77.5	22		$(67+77)/2= 72$
78 - 88	77.5 - 88.5	27		$(78+88)/2= 83$
89 - 99	88.5 - 99.5	8		$(89+99)/2= 94$
100- 110	99.5 - 110.5	3		$(100+110)/2=105$

There is a space between the upper limit of one class and the lower limit of the next class ( 55 and 56 ).

The halfway points of these intervals are called class boundaries (  $(55+56/2) = 55.5$  )

*To find class boundaries (for integer data):*

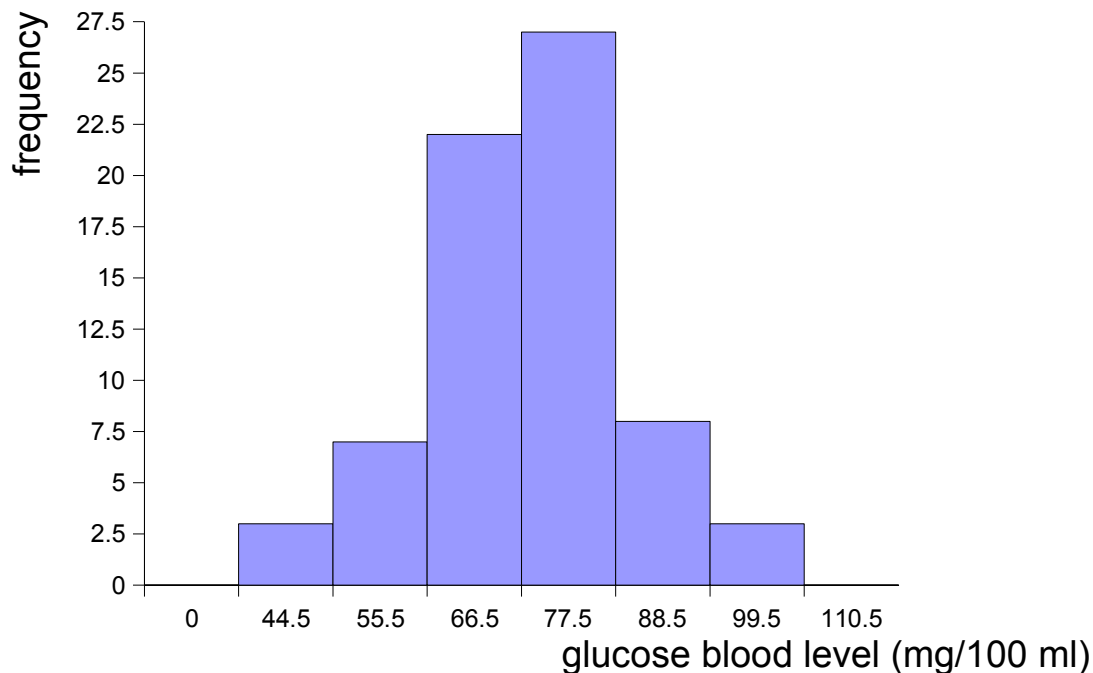
upper class boundaries : add 0.5 to the upper class limit

lower class boundaries : subtract 0.5 from the lower class limit

Class limits lower - upper	Class Boundaries lower - upper	Frequency	Relative Frequency	Class Midpoint
45 - 55	44.5 - 55.5	3		$(45+55)/2= 50$
56 - 66	55.5 - 66.5	7		$(56+66)/2= 61$
67 - 77	66.5 - 77.5	22		$(67+77)/2= 72$
78 - 88	77.5 - 88.5	27		$(78+88)/2= 83$
89 - 99	88.5 - 99.5	8		$(89+99)/2= 94$
100- 110	99.5 - 110.5	3		$(100+110)/2=105$

Right now we can build histogram for the given data:

Histogram for the glucose blood levels



Class limits lower - upper	Class Boundaries lower - upper	Frequency	Relative Frequency	Class Midpoint
45 - 55	44.5 - 55.5	3	$3/70 = 0.043$	$(45+55)/2= 50$
56 - 66	55.5 - 66.5	7	$7/70 = 0.1$	$(56+66)/2= 61$
67 - 77	66.5 - 77.5	22	$22/70= 0.314$	$(67+77)/2= 72$
78 - 88	77.5 - 88.5	27	$27/70= 0.386$	$(78+88)/2= 83$
89 - 99	88.5 - 99.5	8	$8/70 = 0.114$	$(89+99)/2= 94$
100- 110	99.5 - 110.5	3	$3/70 = 0.043$	$(100+110)/2=105$

Total frequencies:

70

$0.43+0.1+ \dots + 0.043= 1$

Let's build relative frequency histogram.

$$\text{relative frequency} = \frac{f}{n} = \frac{(\text{class frequency})}{(\text{total of all frequencies})}$$

Relative frequency histogram for the glucose blood levels

