

1.1 GRAPHICAL DISPLAYS

Histograms A useful first step in the representation of observed data is to reduce it to a type of bar chart. Consider, for example, the data presented in Table 1.1.1. These numbers represent the live loads observed in a New York warehouse. To anticipate the typical, the extreme, and the long-term behavior of structural members and footings in such structures, the engineer must understand the nature of load distribution. Load variability will, for example, influence relative settlements of the column footings. The values vary from 0 to 229.5 pounds per square foot (psf). Let us divide this range into 20-psf intervals, 0 to 19.9, 20.0 to 39.9, etc., and tally the number of occurrences in each interval.

Plotting the frequency of occurrences in each interval as a bar

Table 1.1.1 Floor-load data*

Bay	Base- ment	1st	2d	3d	4th	5th	6th	7th	8th	9th
A	0	7.8	36.2	60.6	64.0	64.2	79.2	88.4	38.0	72.7
B	72.2	72.6	74.4	21.8	17.1	48.5	16.8	105.9	57.2	75.7
C	225.7	42.5	59.8	41.7	39.9	55.5	67.2	122.8	45.2	62.9
D	55.1	55.9	87.7	59.2	63.1	58.8	67.7	90.4	43.3	55.2
E	36.6	26.0	90.5	23.0	43.5	52.1	102.1	71.7	4.1	37.3
F	129.4	66.4	138.7	127.9	90.9	46.9	197.5	151.1	157.3	197.0
G	134.6	73.4	80.9	53.3	80.1	62.9	150.8	102.2	6.4	45.4
H	121.0	106.2	94.4	139.6	152.5	70.2	111.8	174.1	85.4	83.0
I	178.8	30.2	44.1	157.0	105.3	87.0	50.1	198.0	86.7	64.6
J	78.6	37.0	70.7	83.0	179.7	180.2	60.6	212.4	72.2	86.0
K	94.5	24.1	87.3	80.6	74.8	72.4	131.1	116.1	53.6	99.1
L	40.2	23.4	8.4	42.6	43.4	27.4	63.8	18.4	16.2	58.7
M	92.2	49.8	50.9	116.4	122.9	132.3	105.2	160.3	28.7	46.8
N	99.5	106.9	55.9	136.8	110.4	123.5	92.4	160.9	45.4	96.3
O	88.5	48.4	62.3	71.3	133.2	92.1	111.7	67.9	53.1	39.7
P	93.2	55.0	80.8	143.5	122.3	184.2	150.0	57.6	6.8	53.3
Q	96.1	54.8	63.0	228.3	139.3	59.1	112.1	50.9	158.6	139.1
R	213.7	65.7	90.3	198.4	97.3	155.1	165.4	155.3	229.5	73.0
S	137.6	62.5	156.5	154.1	134.3	81.6	194.4	155.1	89.3	73.4
T	79.8	68.7	58.6	141.6	100.7	106.0	131.1	157.4	80.2	65.0
U	78.5	118.2	126.4	33.8	124.6	78.9	146.0	100.3	97.8	75.3
V	24.8	55.6	135.6	56.3	66.9	72.2	105.4	98.9	101.7	58.2

* Observed live loads (in pounds per square foot); bay size: 400 ± 6 ft².

Source: J. W. Dunham, G. N. Brekke, and G. N. Thompson [1952], "Live Loads on Floors in Buildings," Building Materials and Structures Report 133, National Bureau of Standards, p. 22.

Data Reduction

A necessary first step in any engineering situation is an investigation of available data to assess the nature and the degree of the uncertainty. An unorganized list of numbers representing the outcomes of tests is not easily assimilated. There are several methods of organization, presentation, and reduction of observed data which facilitate its interpretation and evaluation.

It should be pointed out explicitly that the treatment of data described in this chapter is in no way dependent on the assumptions that there is randomness involved or that the data constitute a random sample of some mathematical probabilistic model. These are terms which we shall come to know and which some readers may have encountered previously. The methods here are simply convenient ways to reduce raw data to manageable forms.

In studying the following examples and in doing the suggested problems, it is important that the reader appreciate that the data are "real" and that the variability or scatter is representative of the magnitudes of variation to be expected in civil-engineering problems.

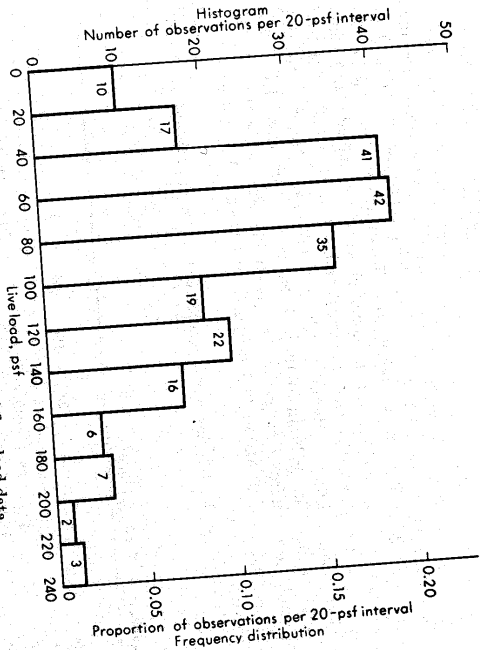


Fig. 1.11 Histogram and frequency distribution of floor-load data.

yields a *histogram*, as shown in Fig. 1.1.1. The height, and more usefully, the area, of each bar are proportional to the number of occurrences in that interval. The plot, unlike the array of numbers, gives the investor an immediate impression of the range of the data, its most frequently occurring values, and the degree to which it is scattered about the central or typical values. We shall learn in Chap. 2 how the engineer can predict analytically from this shape the corresponding curve for the total load on a column supporting, say, 20 such bays.

If the scale of the ordinate of the histogram is divided by the total number of data entries, an alternate form, called the *frequency distribution*, results. In Fig. 1.1.1, the numbers on the right-hand scale were obtained by dividing the left-hand scale values by 220, the total number of observations. One can say, for example, that the proportion of loads observed to lie between 120 and 139.9 psf was 0.10. If this scale were divided by the interval length (20 psf), a *frequency density distribution* would result, with ordinate units of "frequency per psf." The area under this histogram would be unity. This form is preferred when different sets of data, perhaps with different interval lengths, are to be compared with one another.

The *cumulative frequency distribution*, another useful graphical representation of data, is obtained from the frequency distribution by calculating the successive partial sums of frequencies up to each interval division point. These points are then plotted and connected by straight lines to form a nondecreasing (or monotonic) function from zero to unity.

In Fig. 1.1.2, the cumulative frequency distribution of the floor-load data, the values of the function at 20, 40, and 60 psf were found by forming the partial sums $0 + 0.0455 = 0.0455$, $0.0455 + 0.0775 = 0.1230$, and $0.1230 + 0.1880 = 0.3090$.† From this plot, one can read that the proportion of the loads observed to be equal to or less than 139.9 psf was 0.347. After a proper balancing of initial costs, consequences of poor performance, and these frequencies, the designer might conclude that a beam supporting one of these bays must be stiff enough to avoid deflections in excess of 1 in. in 99 percent of all bays. Thus the design should be checked for deflections under a load of 220 psf.

Some care should be taken in choosing the width of each interval † When constructing the cumulative frequency distribution, one can avoid the arbitrariness of the intervals by plotting one point per observation, that is, by plotting i/n versus $x^{(i)}$, where $x^{(i)}$ is the i th in ordered list of data. (see Fig. 1.2.1).

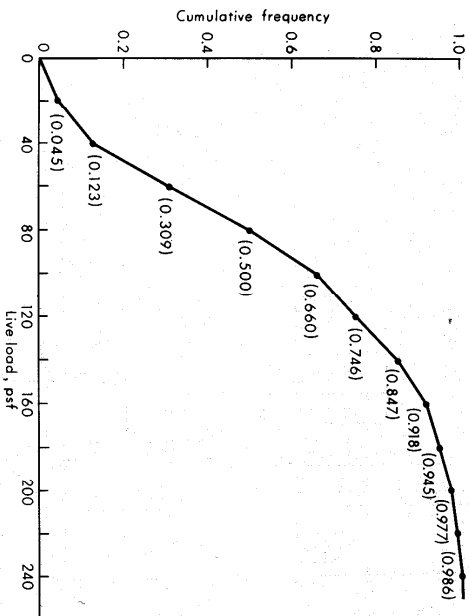


Fig. 1.12 Cumulative frequency distribution of floor-load data.