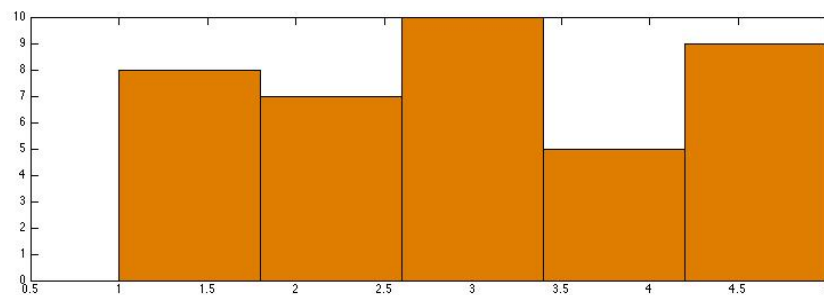


2 — Visualizing Data

2.1 Frequency

Definition 2.1 — Frequency. The frequency of a data set is the number of times a certain outcome occurs.

■ Example 2.1



This histogram shows the scores on students tests from 0-5. We can see no students scored 0, 8 students scored 1. These counts are what we call the frequency of the students scores. ■

2.1.1 Proportion

Definition 2.2 — Proportion. A proportion is the fraction of counts over the total sample. A proportion can be turned into a percentage by multiplying the proportion by 100.

■ **Example 2.2** Using our histogram from above we can see the proportion of students who scored a 1 on the test is equal to $\frac{8}{39} \approx 0.2051$ or 20.51% ■

2.2 Histograms

Definition 2.3 — Histogram. is a graphical representation of the distribution of data, discrete intervals (bins) are decided upon to form widths for our boxes.

R Adjusting the bin size of a histogram will compact (or spread out) the distribution.

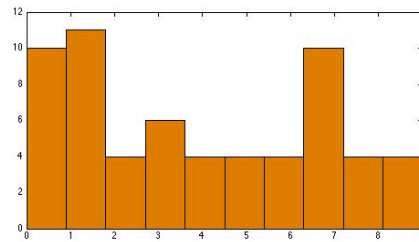


Figure 2.1: histogram of data set with bin size 1

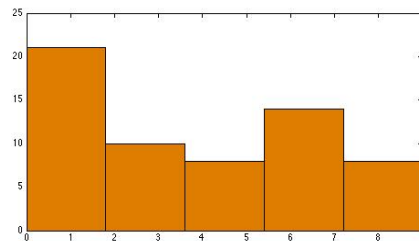


Figure 2.2: histogram of data set with bin size 2

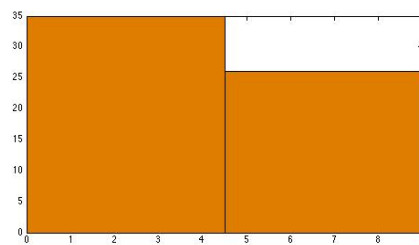


Figure 2.3: histogram of data set with bin size 5

2.2.1 Skewed Distribution

Definition 2.4 — Positive Skew. A positive skew is when outliers are present along the right most end of the distribution

Definition 2.5 — Negative Skew. A negative skew is when outliers are present along the left most end of the distribution

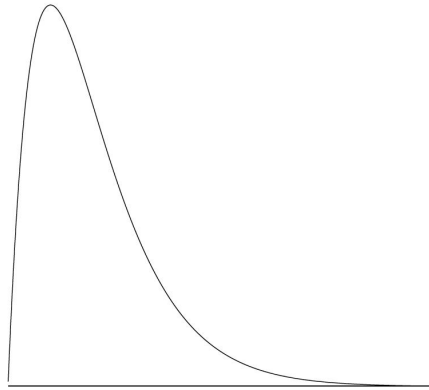


Figure 2.4: positive skew

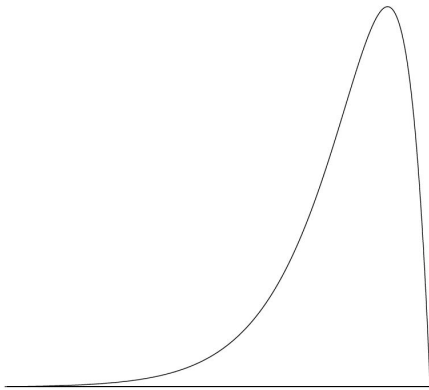


Figure 2.5: negative skew

2.3 Practice Problems

Problem 2.1 Kathleen counts the number of petals on all the flowers in her garden, create a histogram and describe the distribution of flower petals on Kathleen's flowers. Use a bin size of 2.

15	16	17
16	21	22
15	16	15
17	16	22
14	13	14
14	15	15
14	15	16
10	19	15
15	22	24
25	15	16

Table 2.1: Kathleen's petal counts

Problem 2.2 What number of petals seems most prominent in Kathleen's garden? What happens if we change the bin size to 5?

Problem 2.3 What does the skew in Kathleen's flower petal distribution seem to indicate?