Section 2.1

Definition: A **frequency table** partitions data into classes or intervals of equal width and shows how many data values are in each class. The class or intervals are constructed so that each data value falls into exactly one class. Here are the free throw percentages for a sample of 40 NBA players:

93	93	86	66	88	59	78	75
79	51	54	83	59	71	59	54
75	62	64	67	95	75	60	51
79	99	87	65	87	62	86	65
65	62	85	83	73	58	71	51

Exercise 1. What is the smallest free throw percentage above and the largest free throw percentage?

Let's construct a frequency table for the NBA players:

Free Throw Percentage	Frequency
51-60	
61-70	
71-80	
81-90	
91-100	

Here are some standard terms used in discussing and constructing frequency distributions.

Definition: The **lower class limit** is the lowest data value that can fit in a class.

Example: In the above example, the lower class limits are:

Definition: The **upper class limit** is the highest data value that can fit in a class.

Example: In the above example, the upper class limits are:

Definition: **Class boundaries** are the numbers used to separate the classes, but without the gaps created by class limits. To find **upper class boundaries**, add 0.5 unit to the upper class limits. To find **lower class boundaries**, subtract 0.5 unit from the lower class limits.

Example: In the above example, the class boundaries are:

Definition: The **class width** is the difference between the *lower* class limit of one class and the *lower* class limit of the next class.

Example: In the above example, the class width is:

When constructing a frequency table, there should be 5-10 classes.

Exercise 2. Construct two more frequency tables for the the previous data:

Free Throw Percentage	Frequency

Free Throw Percentage	Frequency

Class Exercise 1. Here are the test scores for 40 introductory statistics students:

87	70	92	66	63	80	77	81
94	81	69	75	88	97	93	84
88	92	87	62	94	67	60	88
68	91	100	82	83	68	64	83
84	97	88	80	61	72	63	98

Construct a frequency distribution for the above data.

Class Exercise 2. Here are the number of home runs for 28 major league hitters:

37	24	24	26	22	37	33
36	28	22	20	36	22	24
26	23	27	30	33	30	40
35	37	33	30	36	28	22

Construct a frequency distribution for the above data.

Relative Frequency Distribution

Definition: In a **relative frequency distribution**, the frequency of a class is replaced with a relative frequency (a proportion) or a percentage frequency (a percent). Percent frequencies are calculated as follows:

$$\text{percentage frequency} = (\text{class frequency})/(\text{sum of all frequencies}) \cdot 100\%.$$

Exercise 3. Construct a relative frequency distribution for the NBA data. Again here is the frequency distribution:

Free Throw Percentage	Frequency
51-60	
61-70	
71-80	
81-90	
91-100	

For each free throw percentage, calculate the relative frequency:

51-60

61-70

71-80

81-90

91- 100

Class Exercise 3. Here are the test scores for 40 introductory statistics students:

87	70	92	66	63	80	77	81
94	81	69	75	88	97	93	84
88	92	87	62	94	67	60	88
68	91	100	82	83	68	64	83
84	97	88	80	61	72	63	98

Construct a relative frequency distribution for the above data.

Class Exercise 4. Here are the number of home runs for 28 major league hitters:

37	24	24	26	22	37	33
36	28	22	20	36	22	24
26	23	27	30	33	30	40
35	37	33	30	36	28	22

Construct a relative frequency distribution for the above data.

We now a learn a method of graphically representing a frequency distribution.

Definition: A **histogram** is a graph consisting of bars of equal width drawn adjacent to each other (without gaps). The horizontal scale represents classes of quantitative data values and the vertical scale represents frequencies. The heights of the bars correspond to the frequency values.

A histogram is the most common way to graphically display frequency data.

Exercise 4. Construct a histogram for the NBA player data. Again, here is the frequency table:

Free Throw Percentage	Frequency
51-60	
61-70	
71-80	
81-90	
91-100	

.

Class Exercise 5. Construct a histogram for the home run data.

Definition: A **mound-shaped symmetric histogram** is a histogram in which both sides are (more or less) the same when the graph is folded vertically down the middle.

Definition: A **skewed left or a skewed right histogram** is a histogram in which one tail is

stretched out longer than the other. The direction of skewness is on the side of the *longer* tail. So, if the longer tail is on the left, we say the histogram is skewed to the left.

Exercise 5. Would shape would you expect for the distribution of annual incomes of adults?

Class Exercise 6. What shape would you expect for the distribution of life span for humans?

Class Exercise 7. What shape would you expect for the distribution of IQ scores?

Class Exercise 8. A university teacher saved every e-mail received from students in a large Introductory Statistics class during an entire term. He then counted, for each student who had sent him at least one-email, how many e-mails each student had sent. What shape would you expect for the distribution of e-mails?

Definition: A **uniform or rectangular histogram** is a histogram in which every class has equal frequency. From one point of view, a uniform distribution is symmetric with the added property that the bars are of the same height.

Definition: A **bimodal histogram** is a histogram in which the two classes with the largest frequencies are separated by at least one class.

Definition: **Outliers** in a data set are the data values that are very different from other measurements in the data set.

Definition: The **cumulative frequency** for a class is the sum of the frequencies for that class and all previous classes.

Exercise 6. Construct a cumulative frequency distribution for the NBA data. Again here is the frequency distribution:

Free Throw Percentage	Frequency
51-60	
61-70	
71-80	
81-90	
91-100	

Class Exercise 9. Construct a cumulative frequency distribution for the home run data.

Student Feedback

My teaching methods are (I hope) continually subject to improvement. If you have any comments, suggestions, or ideas, please email them to me at Sithparran.Vanniasegaram@evc.edu.

Homework

C problems

Section 2.1: 5, 15(a)-(e), 17(a)-(e), 19 (a)-(e)

B problems

Section 2.1: 1, 3, 7, 11, 13

A problems

Section 2.1: 9, 21, 23