

Frequency Distribution

A frequency distribution lists data values along with their corresponding frequencies. The frequency distribution helps to give us a picture of the data.

What is frequency?

The number of times something occurs.

In this Example: Suppose we want to discuss the adults who moved to Phoenix in 2012 from outside of the state of Arizona to become residents. The ages in years of those adults are as follows.

25	20	32	31	45	69	70	76	27	45
46	47	37	38	22	31	33	51	55	46
63	41	80	35	36	36	65	25	25	34
37	59	59	45	44	34	29	27	29	32
72	64	63	36	48	37	31	21	21	22

We want to find the frequency distribution of the ages of the adults who moved to Phoenix in 2012 from outside the state of Arizona with the intent to become residents.

To develop a frequency distribution, we need to start with the number of class we want to use. Here we will look at 6 classes. The number of classes is used to determine the width of each class. Find the difference between the maximum data value and the minimum data value and divide that by the desired number of classes.

$$\text{class width} = \frac{\text{max} - \text{min}}{\# \text{ of classes}} = \frac{80 - 20}{6} = 10$$

Now we want to find the starting value of the classes. This is the minimum data value.

Starting value = 20

Now we find the lower class limit for each class. The lower class limit of the first class is the starting value which is the minimum data value. To find the lower limit of the other classes add the class width to the lower limit of the preceding class.

$$\text{Lower limit of 2}^{\text{nd}} \text{ class} = 20 + 10 = 30$$

$$\text{Lower limit of 3}^{\text{rd}} \text{ class} = 30 + 10 = 40$$

$$\text{Lower limit of 4}^{\text{th}} \text{ class} = 40 + 10 = 50$$

$$\text{Lower limit of 5}^{\text{th}} \text{ class} = 50 + 10 = 60$$

Lower limit of 6th class = $60 + 10 = 70$

Now find the upper class limits.

Upper class limit of 1st class = one unit less than the lower class limit of 2nd class

The units in this problem are years so we will subtract 1 year.

Upper class limit of 1st class = $30 - 1 = 29$

Upper class limit of 2nd class = $40 - 1 = 39$

Upper class limit of 3rd class = $50 - 1 = 49$

Upper class limit of 4th class = $60 - 1 = 59$

Upper class limit of 5th class = $70 - 1 = 69$

The upper class limit of the 6th class (which is the last class) is the lower class limit of the 6th class plus the class width.

Upper class limit of 6th class = $69 + 10 = 79$

Put the class limits into a table

Classes	Frequency
20 – 29	
30 – 39	
40 – 49	
50 – 59	
60 – 69	
70 – 79	

Now we find the frequency for each class. Count the number of data values that fall into each class. This is put in the table under the frequency column.

Classes	Frequency
20 – 29	12
30 – 39	16
40 – 49	9
50 – 59	4
60 – 69	5
70 – 79	4

To verify that your count was correct add up the frequency column. The total should be the same as the number of data points, which is 50 for this situation.

Classes	Frequency
20 – 29	12
30 – 39	16
40 – 49	9
50 – 59	4
60 – 69	5
70 – 79	4
Total	50

Relative Frequency is a percentage representation of the frequency.

$$\text{Relative Frequency} = \frac{\text{Frequency}}{\text{number of data values}}$$

Classes	Frequency	Relative Frequency
20 – 29	12	$\frac{12}{50}(100\%) = 24\%$
30 – 39	16	$\frac{16}{50}(100\%) = 32\%$
40 – 49	9	$\frac{9}{50}(100\%) = 18\%$
50 – 59	4	$\frac{4}{50}(100\%) = 8\%$
60 – 69	5	$\frac{5}{50}(100\%) = 10\%$
70 – 79	4	$\frac{4}{50}(100\%) = 8\%$
Total	50	100%

Cumulative Frequency is the addition of the frequencies of the class plus all the classes that came before it.

Classes	Frequency	Relative Frequency	Cumulative Frequency
20 – 29	12	$\frac{12}{50}(100\%) = 24\%$	12
30 – 39	16	$\frac{16}{50}(100\%) = 32\%$	12+16=28
40 – 49	9	$\frac{9}{50}(100\%) = 18\%$	9+28=37
50 – 59	4	$\frac{4}{50}(100\%) = 8\%$	4+37=41
60 – 69	5	$\frac{5}{50}(100\%) = 10\%$	5+41=46
70 – 79	4	$\frac{4}{50}(100\%) = 8\%$	4+46=50
Total	50	100%	

Once you have a Frequency Distribution, a Histogram can be made to get the picture of the data. A **histogram** is a bar graph in which the horizontal scale represents classes of data values and the vertical scale represents frequencies. The height of each bar is the frequency and the width of each bar is the class width.