

Statistics for grouped data

grouped data refers to the arrangement of raw data with a wide range of values into groups. This process makes the data more manageable. Graphs and frequency diagrams can then be drawn showing the class intervals chosen instead of individual values. An estimate, \bar{x} , of the mean of the population from which the data are drawn can be calculated from the grouped data as:

$$\bar{x} = \frac{\sum fx}{\sum f}$$

In this formula, x refers to the mid-point of the class intervals, and f is the class frequency. Note that the result of this will be different from the sample mean of the ungrouped data.

Class limits	Class midpoint	frequency
\$240 - 259.99	\$250	7
\$260 - 279.99	\$270	20
\$280 - 299.99	\$290	33
\$300 - 319.99	\$310	25
\$320 - 339.99	\$330	11
\$340 - 359.99	\$350	4
		Total = 100

Cumulative frequency

The graph of a cumulative frequency distribution is called an ogive (pronounced “o-jive”). For the less-than type of cumulative distribution, this graph indicates the cumulative frequency below each exact class limit of the frequency distribution. When such a line graph is smoothed, it is called an ogive curve.

Relative frequency

A relative frequency distribution is one in which the number of observations associated with each class has been converted into a relative frequency by dividing by the total number of observations in the entire distribution. Each relative frequency is thus a proportion, and can be converted into a percentage by multiplying by 100.

One of the advantages associated with preparing a relative frequency distribution is that the cumulative distribution and the ogive for such a distribution indicate the cumulative proportion (or percentage) of observations up to the various possible values of the variable. A percentile value is the cumulative percentage of observations up to a designated value of a variable.