Multivariate Frequencies

The multivariate frequency values can be computed independently for each attribute. We can represent the frequency values for each attribute by a matrix, in which the number of rows is the number of values assumed by the attribute and the columns are frequency values

following frequency measures can be taken:

- absolute frequency
- relative frequency
- absolute cumulative frequency
- relative cumulative frequency.

Multivariate Data Visualization

for univariate and bivariate analysis, it is easier to under- stand data and experimental results when they are illustrated using visualization techniques.

new visualization approaches and techniques are continuously being created to deal with new types of data, new approaches to results interpretation and new data analysis tasks.

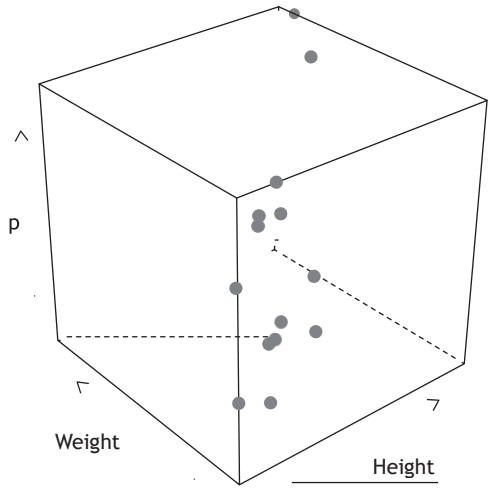


Figure 3.3 Plot for three attributes from the contacts data set.

One may ask how can we represent the relationships between more than three attributes. A straightforward approach would be to modify the three-dimensional graph

Multivariate Statistics

Statistical measures described for univariate and bivariate analysis, such as the mean and standard deviation, can easily be extended to multivariate analysis.

Location Multivariate Statistics

To measure the location statistics when there are several attributes we just measure the location of each attribute. Thus, the multivariate location statis- tical values can be computed independently for each attribute.

Dispersion Multivariate Statistics

dispersion statistics, such as the amplitude, interquartile range, mean absolute deviation and standard deviation

Example

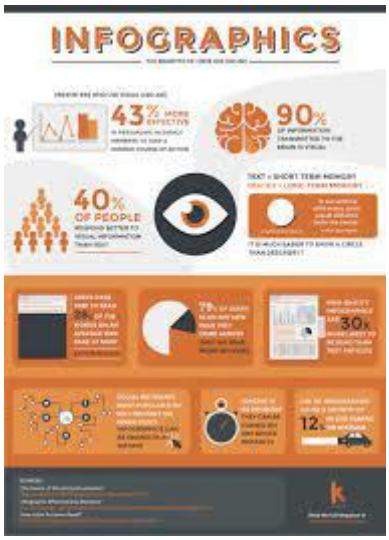
multivariate dispersion statistics for the attributes “maxtemp”, “height”, “weight” and “years” from the data set

Dispersion univariate statistics for quantitative attributes.

Dispersion statistics	**Maxtemp**	**Weight**	**Height**	**Years**
Amplitude	23.00	60.00	37.00	16.00
Interquartile range	11.75	17.50	14.75	9.50
MAD	7.41	14.09	11.12	6.67
s	7.45	17.38	11.25	5.66

Infographics and Word Clouds

it is common to highlight important facts by using infographics. It is important to understand the difference between data visualization and infographics.



both techniques transform data into an image, the info- graphic approach is subjective, is produced manually and is customized for a particular data set.

Word Clouds

A visualization tool frequently used in text mining to illustrate text data is the word cloud, which represents how often each word appears in a given text.

. The higher the frequency of a word in the text, the larger its size in a word cloud.

the words whose stem appear more often in the previous text are represented in a larger font size: this is the case for the words “text”, “word” and “cloud”.