Exploratory data analysis, EDA, is a philosophy about how one can approach the analysis of a data set. It is not a fixed collection of tools that one can use to analyze a data set, but it is a general approach which promotes looking at data in different ways without having any assumptions. It is usually used when one receives a new data set, and wishes to learn something about the underlying structure of the data set, or wishes to discover outliers or trends in the data. In general, EDA is about looking at the data that is presented to one, in many different ways. Both visual and non-visual methods are used in EDA. Examples of non-visual methods are simply to calculate the mean, median, variance, quartiles etc. of the data, while visual methods could be a boxplot, scatterplot, which quickly allows one to discover outliers and trends, or even a simple bar chart. The boxplot is a quick way to present some of the calculated properties mentioned above, such as the median and some of the quartiles. The boxplot was actually developed by John Tukey, who is widely regarded as one of the big promoters of EDA, and it was presented in his book "Exploratory Data Analysis", as a method one could use while doing EDA. Another approach one can use within EDA is principal component analysis, PCA, which is described in another section.

When doing EDA, one will often be able to use the acquired knowledge to formulate new hypotheses about the data, which can then be used to make more specific visualizations or calculate more specific properties of the data set, but it is not a guaranteed outcome. EDA will not always produce new knowledge, which is not a flaw in the approach, but rather a natural consequence of the approach.