**What's a data frame?**

You may remember from the chapter about matrices that all the elements that you put in a matrix should be of the same type. Back then, your data set on Star Wars only contained numeric elements.

When doing a market research survey, however, you often have questions such as:

* 'Are your married?' or 'yes/no' questions (logical)
* 'How old are you?' (numeric)
* 'What is your opinion on this product?' or other 'open-ended' questions (character)
* ...

The output, namely the respondents' answers to the questions formulated above, is a data set of different data types. You will often find yourself working with data sets that contain different data types instead of only one.

A data frame has the variables of a data set as columns and the observations as rows. This will be a familiar concept for those coming from different statistical software packages such as SAS or SPSS.

# Quick, have a look at your data set

Wow, that is a lot of cars!

Working with large data sets is not uncommon in data analysis. When you work with (extremely) large data sets and data frames, your first task as a data analyst is to develop a clear understanding of its structure and main elements. Therefore, it is often useful to show only a small part of the entire data set.

So how to do this in R? Well, the function [head()](http://www.rdocumentation.org/packages/utils/functions/head) enables you to show the first observations of a data frame. Similarly, the function [tail()](http://www.rdocumentation.org/packages/utils/functions/head)prints out the last observations in your data set.

Both [head()](http://www.rdocumentation.org/packages/utils/functions/head) and [tail()](http://www.rdocumentation.org/packages/utils/functions/head) print a top line called the 'header', which contains the names of the different variables in your data set.

**Have a look at the structure**

Another method that is often used to get a rapid overview of your data is the function [str()](http://www.rdocumentation.org/packages/utils/functions/str" \t "_blank). The function [str()](http://www.rdocumentation.org/packages/utils/functions/str" \t "_blank) shows you the structure of your data set. For a data frame it tells you:

* The total number of observations (e.g. 32 car types)
* The total number of variables (e.g. 11 car features)
* A full list of the variables names (e.g. mpg, cyl ... )
* The data type of each variable (e.g. num)
* The first observations

Applying the [str()](http://www.rdocumentation.org/packages/utils/functions/str" \t "_blank) function will often be the first thing that you do when receiving a new data set or data frame. It is a great way to get more insight in your data set before diving into the real analysis.