Python modules used to create the mushroom data in the folders of the parent directory (also available
in data, but be careful: If secondary_data_generation.py is run, some are overwritten withnew randomized instances).

The main modules are now shortly described in the order of their usual usage:

- data_set_categories.py

Container module, providing the file paths to the \data folder used by most modules
also provides hard-coded dicts, list and strings mainly used by primary_data_generation.py

- primary_data_generation.py

WARNING: Cannot be run since the used source book is not freely available. To run this module,
a EPUB copy of the book has to be acquired and the unpacked HTML files put into data/mushrooms_and_toadstools/.
The generated data set primary_data_generated.csv is available as well as manually edited an enriched
version primary_data_edited.csv, which read in used by the other modules.
Module used to read out a HTML version of the book Mushrooms & Toadstools by Patrick Hardin.
The results are written to a CSV to create a primary mushroom data based on different species,
usable for simulating hypothetical mushroom data.

- secondary_data_generation.py

Module used to simulate randomized hypothetical mushrooms based on the primary data created with
primary_data_generation.py. Results in secondary_data_generated.csv used in the following modules.

- mushroom_classifier.py

Module used to explorate, clean, encode and binary classify data with Naive Bayes, logistic regression
and LDA, including the evaluation of results with confusion matrix and scores.

- statistical_graphics.ipynb

Jupyter Notebook module used for visualization of bar plots, correlation heatmaps and ROC curves.

The remaining modules are utility modules imported and used by the main modules.