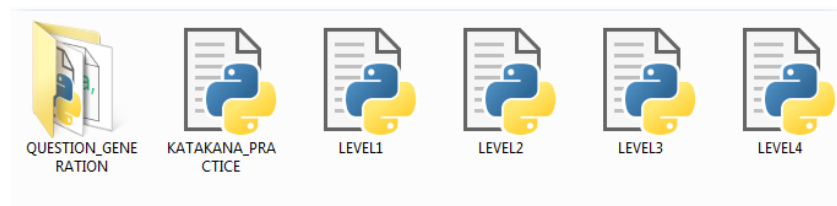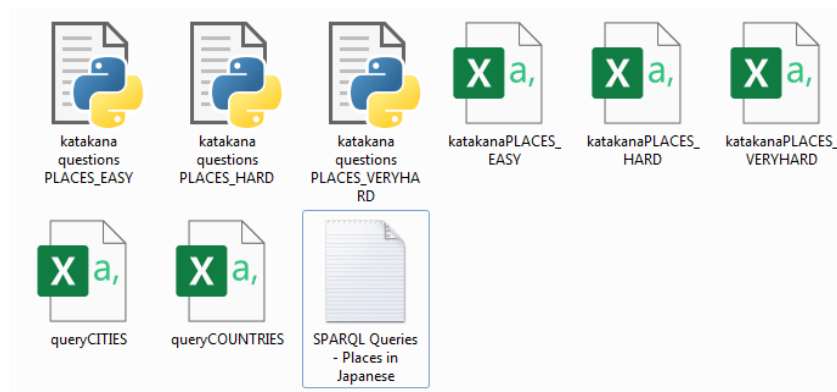# WHAT IS WHAT IN THE FOLDER

MAIN FOLDER:

The question generation folder has all the files used to create questions with hints and distractors.

The **KATAKANA_PRACTICE.py** file is the main file, the one to be run to actually use the practice tool. It is **recommended to run it with IDLE** or a similar Python editor instead of a terminal **to display all the characters** properly (at least running it on the Windows Command Processor did not display all the characters). This file imports a function from the LEVEL file corresponding to the level selected by the user after running the main file.

- **LEVEL1**: the user is given a **clue**; **distractors** are only of **similar length** to the correct option, but **morphologically varied**.
- **LEVEL2**: the user is given a **clue**; **distractors**, in their Japanese equivalent**, share the character** presented as a clue, but in random positions.
- **LEVEL3**: the user is given **no clue**; **distractors**, despite being of **similar length** to the correct option, are **morphologically varied**.
- **LEVEL4**: the user is given **no clue**; **distractors**, in their Japanese equivalent, are **morphologically similar**.

QUESTION GENERATION FOLDER:

This folder contains a **txt** file with the code of the **queries** that were run to obtain the data on cities and countries in Japanese and English from **Wikidata**.

It also contains the two **csv** files obtained **from the queries**.

Then, there are **three py files** that **extract questions** based on the information **from the query csv** files. The folder also includes the **resulting csv** files with the question data.

- **EASY**: distractors are only of **similar length** to the correct option, but **morphologically varied**.
- **HARD**: distractors, in their Japanese equivalent, **share the character** presented as a clue, but in random positions.
- **VERY_HARD**: distractors, in their Japanese equivalent, are **morphologically similar**.

DISCLAIMER: At least in Windows, the question generation programs create a csv that has extra line breaks. The same programs then correct this problem. The "bad" versions of the csv files have not been included (not needed).