First argument: JSON folder (input)

Second argument: Excel file (output)

Iterate over every LabelMe JSON in the input folder

Iterate over shape objects in “shapes” in every JSON

Only consider shapes that are narrower than 500 pixels

For every shape that remains only consider one final output value:

* human\_output.human\_corrected\_text IF EXISTS
* If does not exist, AND shape is “text\_cell” or “column\_header”, then openai\_output.response
* If does not exist, AND shape is “numerical\_cell”, then tesseract\_output.ocr\_score

Write out the contents into the excel, as follows: For every super\_row calculate the max number of rows of the final output, that will be the vertical size of the text block for EACH SHAPE in that super\_row. So every new block in the excel that is in the same super\_row will start at the same line in the excel. The horizontal size of the text block is 1, as each “super\_column” corresponds to an excel column. Write out every shape’s content accordingly.

Once you are ready with the first JSON, go on with the next one, and append to the same excel sheet, right where you left off with the previous one