Scene info documentation

The challenge: Write code to pull out key features from images with the end goal of geolocating the image.

Approaches:

1. OLLAMA to read the images, outputting a general description as JSON.
2. Manually classifying 100 images
   1. Is there a child in the image?
   2. Is the image indoors or outdoors?

Data:

* <https://www.kaggle.com/datasets/itsahmad/indoor-scenes-cvpr-2019>
  + 15,000 random images from the MIT campus, with annotation for 2700 of them.
* <https://www.kaggle.com/datasets/fxmikf/interior-exterior-scene-classification/data>
  + 1000 interior and 1000 exterior images, taken as frames from various videos

**Directory and file explaination**

* **Code**
  + **manipulate\_files.py**
    - python functions to move files from point X to Y. The purpose of this was to get the MIT images all in one place, not in different folders.
    - Extract\_xml\_names is a function that gets the MIT annotations and puts them in their own file so that you know what each image contains.
* **Data**
  + MITImagesNotes.csv
    - For 2700 images with annotations, this has the filename and the list of annotations
* **Hand\_labelled\_csvs**
  + Descriptions.csv is 100 imags from Kaggle, hand-classified.
    - How many people are in this image?
    - Is this image indoors?
    - Is this image outdoors?
    - Is there a child in this image?
    - Does this image have readable text?
  + Child\_detect.csv
  + Indoor\_detect.csv
  + Indoor\_output\_input.csv
* **Output**
  + Output.csv
    - The result of calling OLLAMA on some images