Objective:

The primary goal of this test is to understand 3 aspects -

1. Is the shelf empty or not?

2. What time the picture was taken?

3. What is the item that is missing from the shelf?

Requirements:

Library Required:

- a. Tensorflow 1.13.1
- b. ImageAI
- c. Tesseract
- d. OpenCV

Note: All the libraries can be install using pip

To install Tesseract, you also have to download its binary file to install in windows

Link – https://pypi.org/project/pytesseract/

If you getting error code 535 while sending emails please refer to below url and activate the less-secure option:

Link - https://www.google.com/settings/security/lesssecureapps

How to Run:

Go to Main.ipynb , search for **Start form here** start executing all the code blocks but before that execute **only Library block** and **Misc Func block**

Methodology:

To achieve the mentioned objective, I've used ImageAI python framework for object detection and Tesseract python framework for OCR purpose.

Steps:

1. Firstly, I've resized the image and compressed to faster the training process
2. For each processed image annotation has been done to help identify the objects so that model can train on it.
3. In training a pretrained model is used to improve accuracy which is a trained-on YOLO dataset.

4. Now model is used to detect objects from the given image, and it is then sorted in the decreasing order of detection probability. From the sorted list of objects only the top-N objects will be used whose accuracy is equal to or more than 80%.
5. After detecting I've performed OCR using Tesseract python framework, which will detect the text from the objects and the text is now manipulated to answer the objectives.
6. A message is created which is sent to specified email.

Conclusion:

Since the dataset of original images given is very sparse and to build a higher accuracy model we need plenty of images of a particular class which was not the case, hence the accuracy to detect objects where poor, but it is able to detect Shelf Label, Location, Time, Empty Spaces in the image as there were enough images to train for that.