What went well

* Creating a parser to go through the Excel files did not give issues when developing it since Python handles reading Excel files very well
* Developing an image scrapper also went well through some research into how to apply a URL search and extract images from search engines
* Making a function to resize and format the images after extraction went smoothly and will be applied after some more debugging the main program.
* Applying a URL check before extracting images from the web was also not difficult to apply. This feature was a great recommendation from the cyber team.

What didn’t go well

* Confirming accuracy when scrapping the photos off of the search engines has been an issue the team has faced. When gathering a sample of images for a specific product, the images vary which leads to potentially inaccurate product images
* Currently, the basis of the program is hard coded. When trying to change the number of images scrapped by the images caused some issues when trying to debug the program.
* Until we have an accurate way of authenticating the images, the amount of images scrapped will be hardcoded until a number is decided on.

What could be improved

* Communication between the team and the sponsor can be improved in the future. Consistent weekly meetings with the sponsor have been helpful but certain members of the sponsor’s team could be present for more of the meetings
* Despite the current program being functional for parsing and image scrapping, the runtime can get lengthy when large numbers of searches are requested through the Excel list
* Our program’s accuracy can be improved by the addition of the product description in the search. This was a suggestion advised to the team by the sponsor.

Challenges

* Most of the team was unfamiliar with programming in Python, so research into learning the necessary libraries and syntax was a priority.
* There were some issues when combining the separate Python programs in the GitHub repository. This was due to the team individually handling separate functions of the program locally on their devices before adding them to the repository.
* During early brainstorming for the product scrapping, an effective way of scrapping images was an issue for the team. Since some of the manufacturers did not have the product images on their websites, finding the images had to be done through multiple web searches rather than through their websites.