**Christmas Project**

#1 Using Link: <https://www.digitalocean.com/community/tutorials/how-to-work-with-web-data-using-requests-and-beautiful-soup-with-python-3>

Had to set up virtual python programming environment. Used:

Create the python virtual environment.

$ python3 -m venv venv

Activate the virtual environment.

$ source venv/bin/activate

(venv) $

Installed Requests and Beautiful Soup packages (term for library in python) using pip tool.

#2 Wrote up code for Web-scraper: <https://www.digitalocean.com/community/tutorials/how-to-scrape-web-pages-with-beautiful-soup-and-python-3>

Wrote Code in VS Code. Used terminal to run (entered python3 main.py in command line).

#3 Wrote up requirements after follow-up call:

1. Implement a web scraper that pulls content from a FAQ webpage and puts it into a text file
2. Implement a supervised key word extraction algorithm
   1. Which takes in a premade list of words
   2. Searches through the contents in the text file and returns the paragraphs where that word appears
   3. The returned paragraphs will be ranked based on number of times that word appeared and give special priority if the word appeared in the question

I have finished the web-scraper part for now (I used my school’s website as the test webpage) I may have to alter the code slightly for the actual FAQ webpages. I will now implement the algorithm. I will assume that in the text file, I have a way of identifying whether the content is a question or an answer (so that I can give special priority).