Project Report – Bill, Christy, and Michal – Website Keyword Analysis

**## Project Report**

At the end of the week, your team will submit a Final Report that describes the following:

\* **\*\*E\*\***xtract: your original data sources and how the data was formatted (CSV, JSON, pgAdmin 4, etc).

\* **\*\*T\*\***ransform: what data cleaning or transformation was required.

\* **\*\*L\*\***oad: the final database, tables/collections, and why this was chosen.

Extract: We took an exported CSV of keywords from Wordstream as well as lists of keywords scraped from websites to compare.

Transform: We compared the string lists using Python functions and then stored the resultant ratings (amounts of hits between both lists) in a variable for storage.

Load: Having combined the results together into dictionaries, we used Pandas dataframes to assemble a dictionary of dictionaries to be loaded into MongoDB using PyMongo. This was chosen because the data was ultimately all text information and MongoDB is an efficient and useful system for list-like objects.