Kayla Bracall DAT 129 – Spring 2020

**Final Project – Web Scraping**

**Project Overview**

This project aims to find which type of yarn is most expensive, and which weight of yarn is most commonly sold. This is done by scraping information from yarn.com. The following criteria was examined:

|  |  |
| --- | --- |
| **Yarn Types** | **Yarn Weights** |
| **Alpaca** | **Lace** |
| **Bamboo** | **Fingering** |
| **Cashmere** | **Sock** |
| **Cotton** | **Sport** |
| **Luxury** | **DK** |
| **Polyester** | **Worsted** |
| **Silk** | **Bulky** |
| **Wool** | **Super Bulky** |

**Hypothesis**

My hypothesis is that cashmere would be the most expensive yarn material. Based on my own buying habits, I hypothesized that worsted weight yarn would be the most common weight sold. After scraping the site and finding the average price of each type of material, cashmere yarn was found to be the most expensive, on average. However, DK weight yarn was found to be the most common weight sold.

**Process**

1. Determine criteria on webpage to be scraped
2. Build code that scrapes for price
3. Find average price per material
4. Build code that scrapes for weight and add to dictionary
5. Add all yarn materials to dictionary
6. Build code that “sorts” dictionaries
7. Display first value in each dictionary
8. Write yarn dictionary to CSV file

**Output**

