Carla Harrell DSC540 Final Project 02/28/2020

In my final project, we were required to produce at least 15 to 20 variables form an API or Webpage and perform 2 Data Transformation/Clean-Up Steps. I chose two different techniques to get the experience from both. I chose an API from spotify.com and extracting product information from Walmart.

**Spotify**

Data was retrieved from Spotify on an Artist named Coldplay. The objective was to locate the most well-known songs being played on Spotify by the artist. Spotify requires API permissions & credentials that must be requested through their developer's code. The credentials are in the “dashboard” under select “create client id. Python module — spotipy — was imported and pip install Spotify had to be installed. Data was extracted from Coldplay’s song and another artist who may collaborated with him. Fifteen variables pertaining to various songs and albums by this artist were created. I removed duplicates and perform cleaning. At the end I decided to place a graph just to see which song was played most often. The API is JSON file and it provides most of the variable names.

The data recovered from here furnished the music craftsman with data on the recurrence of their songs are being played, frequency, sound data and more. The artist will have the option to take the information to make new melodies dependent on their fans. In the event artist just released an album, week after week to month to month investigation ought to be performed utilizing this code to monitor what is playing. This can help craftsman with improving music on the off chance that they recognize what individuals are tuning in to.

**Webscraping**

Data was scraped from Walmart on female body spray using beautiful soup. It was very difficult to retrieve 15 – 20 variables and I had to split the data in order to create variables. For example, the product description gave the name of the fragrance, brand, and the ounces. I removed the brand and created a variable named brand. Walmart does have a developer code, but it is used for vendors only. Once tried to inspect the elements the data was hidden, and I had to right click on each item that I wanted to scrape to bring up the div tabs. Most of the items are disallowed, but you can scrape some data. I was successful in retrieving my 15 variables.

The information retrieved from Walmart can provide information on the types of body spray, cost, ratings, shipping cost, etc. It can give the store things to keep on the rack’s dependent on surveys, yet information would require further investigation such as revenues being generated in order to perform a thorough investigation of the products. Weekly and monthly analysis should be performed depending upon the stores policy.

References:

<https://www.walmart.com/>

<https://developer.spotify.com/>