**Project Title:**

Ecommerce Website Comparison Database

**Team Members:** Jovan andShola

**Project Description/Outline:**

We decided to perform a website comparison of vast electronical items ranging from household electronics to personal electronic items and gadgets. We will source our datasets by web scraping popular ecommerce websites to extract data, transform data by cleaning datasets to create a clear structure and load data into a relational database (postgresql). From our datasets we will compare consumer experience by ratings of items and price differences between categories of the same items.

**Areas of Research**

* Televisions
* Mobile Phones
* Headphones
* Tablets and Kindles
* Washing Machines
* Tumble Driers
* Washer Dryers
* Fridge and Freezers
* Microwaves
* Vacuum Cleaners
* Laptops and PCs
* PC Monitors
* Printers
* Apple Watches
* Smart Watches

**Data Sources:**

Method: Beautiful Soup and Splinter

Product List: Webscraping Currys.co.uk

Currys(Prices/Reviews): Webscaping Currys.co.uk

Very(Prices/Reviews):Webscraping Very.co.uk

Currys is seen as the most popular electronic store in the UK, so with that being the case, our Product list will be created using their stock. With their items being most popular and most used, it makes a useful dataset of products. Using these products, the Very and Currys website will be webscraped to find prices, reviews and review count, which created two dataframes.

**Cleaning:**

Method: pandas

The dataframes will be cleaned by removal of any null values. The data types will also be checked for the correct types. The price will be a float, rating will be a float and the rating count will be an integer. The columns will also be adjusted for a more suitable label.

**Load:**

Method: sqlalchemy

The data will be loaded into a relational database, postgresql. Using the database it will the then be joined creating a table with both very and curry products ready for analysis.

**Research Questions to Answer:**

* Can we find an ecommerce website with significantly lower prices?
* Which ecommerce website has lower prices by category?
* Do reviews differ between both websites we compare?

**Rough Breakdown of Tasks:**

* Create code to web scrap for data
* Clean data sets to transform and create clear structure
* Load data into chosen database
* Perform queries on database to analyse product prices
* Collaboratively review findings
* Write description of findings based on analyse
* Create professional README.md document
* Generate repository and collaboratively upload to Github

**Findings**

The findings we expect to see are:

* Currys will be cheaper as it is more established with it is also have large stores
* Currys will have worse reviews, this is due it being larger and more trusted therefore customers have higher expectations and trust. Which allows customer’s opinions to sway towards more negative.
* There will also be a filtered analysis which will have a minimum of 50 reviews to give a better view of the product and service, which may create more similar ratings.