Background and Objectives:

This is a data scraping project that involves identification of insurance listings on products sold on Amazon. The aim is to scrape all products listed under the following categories on Amazon UK, France, Spain and Italy:

- Home All Products
- Furniture All Products
- Luggage All Products
- Electronic Toys All Products
- Sporting goods -
 - Odometers, Weight Benches, Home Gyms, Cycling Exclusives, Outdoor Navigation, Sports Technology Exclusives, Swimcomputers, Dumbbells & Weights, Training Equipments, Sports radars and Sailing Electronics and Navigation

Aim of the Exercise:

The goal is to scrape all products under the above mentioned categories and check the availability of Insurance products under these categories. The exercise aims to identify the following:

- 1. Number of products in each category
- 2. % of products displaying insurance
- 3. Various data points on insurance price, insurer, etc

Data Scraping Links:

The data scraping will take place on the following URLs:

- Amazon UK https://www.amazon.co.uk/
- Amazon Italy https://www.amazon.it/
- Amazon Spain https://www.amazon.es/
- Amazon France https://www.amazon.fr/

The data points to be collected from scraping along with sample data can be viewed in the excel file attached with this project brief document.

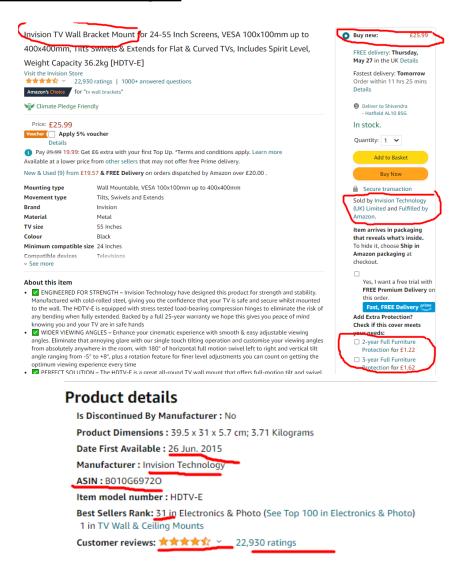
Example of Product Category Journey:

The below URL explains the data journey from Amazon Home Page to final product listing page:

- 1. https://www.amazon.co.uk/ref=nav_logo Amazon Home Page
- 2. https://www.amazon.co.uk/b?node=11052681 Home & Kitchen Category Page
- 3. https://www.amazon.co.uk/b/?node=10745681&ref = Oct s9 apbd odnav hd bw b kNJ3 7&pf rd r=9TGV03GVH792JXC56J9Z&pf rd p=f4b31bb4-dc17-5d60-824c-b 6888a5549e8&pf rd s=merchandised-search-10&pf rd t=BROWSE&pf rd i=1105 2681 Furniture Sub Category Listing Page
- 4. https://www.amazon.co.uk/b/?node=2850888031&ref = Oct s9 apbd odnav hd bw __bj5RR_8&pf rd r=R6VY85YVRJS1022E7811&pf rd p=b0cd6d2a-e7c0-588f-a484-b83eda3d976f&pf rd s=merchandised-search-10&pf rd t=BROWSE&pf rd i=1074_5681 Home Entertainment Furniture Sub-Category Page
- 5. https://www.amazon.co.uk/b/?node=310196011&ref =Oct s9 apbd odnav hd bw b36w2D1_0&pf rd r=5VHFSPQ383RFEG56XVB7&pf rd p=ee381a49-210c-5469-a0d5-eb89e89697b5&pf rd s=merchandised-search-10&pf rd t=BROWSE&pf rd i=2850888031 TV Mounts, Stands & Turntables products listing page
- 6. <a href="https://www.amazon.co.uk/Invision%C2%AE-Ultra-Swivel-Bracket-Mount/dp/B010G69720/ref=sr_1_3?_encoding=UTF8&c=ts&dchild=1&keywords=TV+Mounts%2C+St_ands+%26+Turntables&qid=1621851879&s=home-theater&sr=1-3&ts_id=310196011

 1 Actual Product listing Page

Screenshot of Data to be scraped:



Important Note:

The final data to be scraped from link 6 above, All details required are mentioned in the sample excel file attached with the brief document. All the data points mentioned in the Excel Document have to be collated for all 4 country links mentioned above.

In both these cases, we intend to capture all the details mentioned in the data points table above. The key identifier here would be the ASIN number, which is a unique number

pertaining to that particular product. Therefore it is vital to get all the information about the product as mentioned above.

Support:

- Point of contact
- Any technical or project related assistance in need be

Deliverables:

A CSV file for each of the following:

- 1. Amazon.co.uk Home
- 2. Amazon.co.uk Furniture
- 3. Amazon.co.uk Luggage
- 4. Amazon.co.uk Electronic Toys
- 5. Amazon.co.uk Sporting Goods