# About Marico

Marico is committed to grow its digital business and create digital brands across geographies & to create engagement on digital. Thus, it’s important for us to constantly keep an eye on market, understand the space, innovate and succeed. To do so, Digital transformation team is trying to create a quicker & agile way to automate insights from available data.



## Global Context of ecommerce -

* Revenue in the eCommerce market is projected to reach US $4.11tn in 2023.
* Revenue is expected to show an annual growth rate (CAGR 2023-2027) of 11.51%, resulting in a projected market volume of US$6.35tn by 2027.
* With a projected market volume of US$1,487.00bn in 2023, most revenue is generated in China.
* In the eCommerce market, the number of users is expected to amount to 5.29bn users by 2027.
* User penetration will be 57.2% in 2023 and is expected to hit 66.6% by 2027.
* With this growth rate spotted, Marico plans to expand its ecommerce with a growth rate to the tune of 4X

## Internet penetration is on rise in India.

* Active internet user base in India is 759 Mn as of Dec 2022, by 2025 the number is expected to grow to 900 million. This is for the first time that most Indians have become active internet users. 399 million are from rural India, while 360 million are from urban India.
* Over 50% of internet users research online before making a product purchase: Kantar study
* 9/10 offline buyers are using online touchpoints for information in their purchase journey: Google
* This rise in digitization is leading to rise in focus of building Digital brands within Marico ecosystem – organically or Inorganically.
* Success of digital brands is dependent heavily on data driven marketing & commerce

# The Problem -

Marico works extensively with 3rd party vendors to enable technological capabilities. An agency which crawls the Category wise Top Best Seller SKU data for Marico is present. It looks at SKU’s and extracts information from webpages, tags that data according to various parameters & creates Category Codes for us.

However, Marico wants to use newest AI models & Crawling capabilities to get to most accurate solutions to these real-life problems.

* **Team Size –** Upto 4 pax
* **Category –** Hair Oils
* **Rounds –** A single round contest with 2 levels of submission. First cut submission will be eliminatory basis accuracy of results.
* **Datasets Provided –** 
  1. Trained/Tagged Data – USA (Subset)
  2. Testing/Untagged Data – UUAE, India
* **Aim –** 
  1. Create an integrated Crawling/NLP/Image Processing solution which apes the Manual tagging of parameters for SKU at highest accuracy. (Explained “tagging” process in Appendix)
  2. Do the needful for all 2 Testing dataset with best possible accuracy.
  3. Output format –
     1. Raw Data file (Via Data Crawl) – Given in zip file (Same as Training Dataset)
     2. Category Codes (Via Data Wrangling)– Data intersection of certain parameters which will help jury decide accuracy. Can be done via python/excel. Given in zip file.
     3. All codes and other necessary files
* **Evaluation Criteria –** 
  1. Accuracy in tagging all 3 Testing dataset via your solution
     1. Category codes will help Jury understand the accuracy levels as we already have solutions to provided testing datasets.
     2. Team needs to focus on the most holistic method to collect and tag data
  2. Robustness of model to work for Other Categories (Apart from Hair Oils)
* **Note to Teams –** 
  1. Deeper Dives - Analysis of Sub Segments of category can help get to most accurate solution. Abnormal High/Low needs to be checked if the tagging is accurate.
  2. Use of entirety of webpage to extract data for tagging (text, Images, etc)
  3. Ownership of all submission (Codes, Data, etc) will be of Marico India Limited.
* **Timeline –** 
  1. First Cut Submission (Code, Raw Data, Category Codes, etc) – 16th Dec EOD
  2. Feedback by Jury on Results – 17 to 18th Dec
  3. Final Submission – 22nd Dec, 2023 EOD

# Rewards & Recognition

Marico believes in rewarding the talent which seeks to push boundaries and find solutions. Staying true to that philosophy, Marico will reward the Top performers in an all-rounded manner.

* Certificate of Appreciation to the Top 10 Teams across campuses.
* Cash Reward of Rs. 10,000 (Winner), Rs. 5,000 (Runners-Up)
* Gift Hamper from the Family of Marico Brands to the Top 5 Teams.
* A chance for the Top 10 teams to interact with Top Marico Leadership.

# Appendix

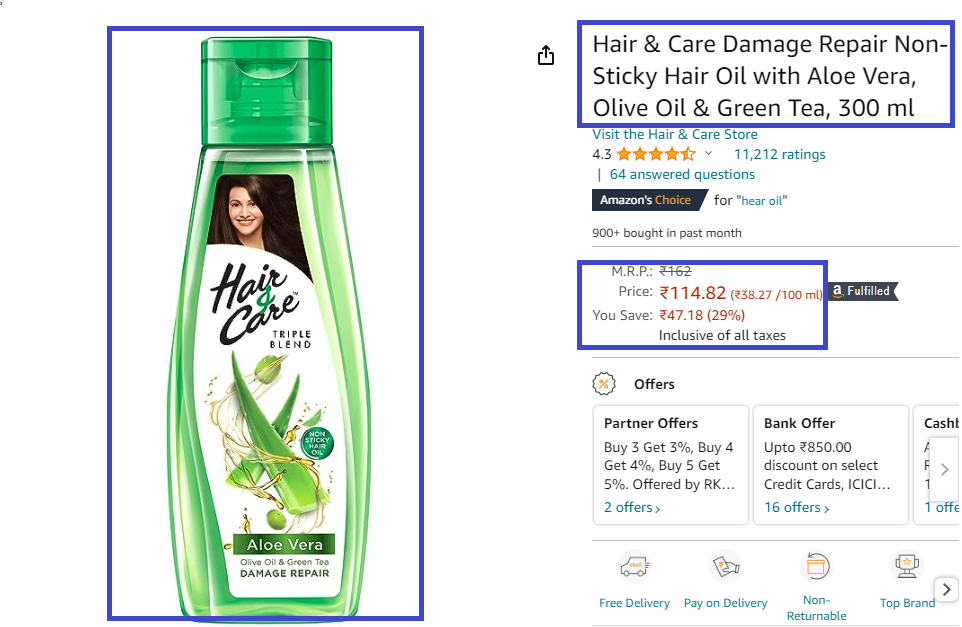
## Tagging Process -

1. Amazon Webpage for top seller SKU in Hair oil Category

A screenshot of a website

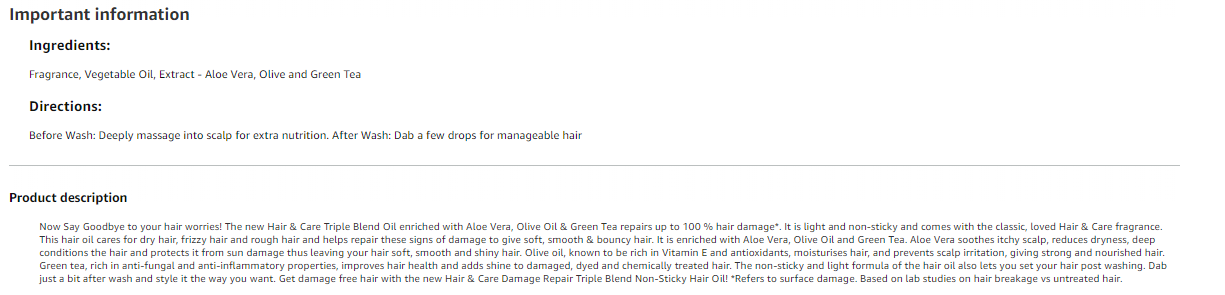
Description automatically generated

1. Use all possible placeholders of information to get data which will be tagged into various parameters. Some are shown below -



A screenshot of a product

Description automatically generated



A screenshot of a computer

Description automatically generated

## 