**Assignment for Data Analysts - Python only**

1. Scrape the product list page of the amazon (choose shoes range between Rs. 1000 - 2500)

What to scrape:

* Title of the product
* Ratings
* Number of reviews
* Price of the product
* Free delivery or not

1. Clean the urls (refer the given excel sheet)

Example:

Raw url:  <https://technology/gadgets/apple-iphone-12-review-an-iterative-update-over-iphone-11-11605525335131.html>

Cleaned url: [technology/gadgets/apple-iphone-12-review-an-iterative-update-over-iphone-11](https://technology/gadgets/apple-iphone-12-review-an-iterative-update-over-iphone-11-11605525335131.html)

Basically, we removed “https://” and “11605525335131.html”, similarly it should be done for all urls

1. Make a new dataframe which contains only those rows which have the number of returning\_users greater than new\_users (refer excel sheet)

**NOTE:**

**Submission of this assignment has to be in the form of a Google Colab. Any other submission (PY code / Excel / CSV) will not be entertained.**