

Assignment

We appreciate your interest in joining **Circuit House** team as an intern.

Your enthusiasm and passion for **disrupting the Indian technology sector** have certainly caught our attention.

To test whether you have the potential to become a key contributor in our path to glory, we are handing out this assignment.

Tools/Skills required

1. Python
2. SQL
3. Power-point or Word to present your analysis.
4. Data analytics
5. Web scraping

Hiring process

Phase 1 Initial interview: Focused on getting to know your strengths, skill sets and past professional or academic experience.

Phase 2: Assignment

Phase 3: Present your findings in a Powerpoint or Word format

Phase 4: Interview with another core member/s

Assignment:

1. Data scraping/cleaning

Circuit house is evaluating a potential entry in the **Smart lock** market. Since this is a fairly new category there are very few estimates present in terms of market size and key players. One way to estimate the key players is by looking at major brands on online portals. The task assigned to you is to scrape online portals (Flipkart or Amazon, you can choose any one portal) and create a database of key products. Below are the fields required in the database.

Brand name	String
Price	Integer
Rating	Float
Rating count	Integer
Review count	Integer
Ranking (where does the product appear in the search)	Integer
URL	String

Note: Limit scraping to 20 pages of the search results

2. Analysis

Based on your collected data analyze the below

- Number of brands in the segment
- Count of SKUs per brand
- Relative ranking: Create a formula to rank brands based on the rank of the SKUs in the search results. E.g.: Brand A has 4 SKUs in the search results and Brand B has 6 SKUs in the search results, below are their ranking in search results

Brand A	Rank	Brand B	Rank
SKU 1	2	SKU 1	10
SKU 2	12	SKU 2	15
SKU 3	25	SKU 3	30
SKU 4	40	SKU 4	45
		SKU 5	70
		SKU 6	110

As per above table, Brand A should have better relative ranking than Brand B

- Relative rating: Use same logic as above to calculate the relative rating (not count) of a brand.
- Price distribution of SKUs

Price Band	SKU count
<INR 4999	
INR 3000-4999	
INR 5000-9999	
INR 10000-14999	
INR 15000-19999	
Greater than 20000	

3. Presentation

Present the above data with commentary and graphical representation in a ppt or word format.

Not more than 10 slides/4 pages

Evaluation (Out of 100)

1. Data scraping/cleaning	30
2. Analysis	40
3. Presentation	30

Reach out to Abhishek (+91-9886186654) if you have any questions.