**Step 1.** Read a flat file( which has Amazon url and unique id. Read the file scrap amazon.com content. File will have 1000 links

**Step 2.** **In First run:** Extract following items

**Fields to be extracted:** ID(Sequence number) | Run date(sysdate) | Product title | Number of Rating | Highest Number of Rating(Calculated) | Ratin | Highest Rating(Calculated) | BSR | Price | Sellers | Lowsest Price Among all seller(Calculated) | ASIN | URL **(Refer attached excel)**

**In Consequence run:**

Update all the above items along with that update highest rating and highest number of rating.

\* Price need to consider as Product price + Shipping price

**Below example will explain what is highest rating and highest number of ratings.**

**First Run Example:**

Product title| price| rating| highest rating| number of rating| highest number of rating | sellers | Lowest Price

Ball | 2 | 4.5 | 4.5 | 19 | 19 | [{“Seller1 Name”, “1”}, {“Seller2 Name”, “2”}] | 1

Pen | 22 | 4.5 | 4.5 | 19 | 19 | [{“SellerXX1 Name”, “21”}, {“Seller2 Name”, “22”}] | 21

**Next Run Example:**

Product title| price| rating| highest rating| number of rating| highest number of rating | sellers | Lowest Price

Ball | 2 | 4.4 | 4.5 | 20 | 20 | [{“Seller1 Name”, “1”}, {“Seller2 Name”, “2”}] | 1

Pen | 22 | 3.5 | 4.5 | 10 | 19 | [{“SellerXX1 Name”, “21”}, {“Seller2 Name”, “22”}] | 21

**Step 3**. Use free proxy service to prevention amazon restriction on running server.

**Step 4**. Deploy it in digital ocean for manual run.