**DIGITAL SKILLED UP AFRICAN-INCUBATOR HUB**

**DATA ANALYSIS CLASS   
CAPSTONE PROJECTS**

**JULY 2025**

**CASE 1**

**CASE ONE IS ON EXCEL: AMAZON**

The following step were taken:

**1)Data cleaning Stage**

**\* Remove Duplicate in the Product ID column**

Click any cell within your dataset and apply Ctrl + A to highlight all the data

while still highlighted, click your Data Tab

Go to Remove Duplicate tool that is under "Data Tool" group.

Click Unselect All first, then only select product\_id

Once you do that click OK

**\*Category Column splitting**

Copy and Paste the Category Column to a "playing ground" environment. By adding new worksheet and adopt Text to Columns.

Copy onto your clipboard the common delimited (|) in the overall categories’ details. Once you copy it, it’s already glued to your clipboard. Then, highlight your column A where you now have your category details, and click Text to Columns. dialogue box would come up once., then click Next.

Click any of the details and go into the formula bar. Highlight it and copy it, Paste the copied delimiter into this Other space. Then click on Finish.

Then rename the category as wished. Click on the new columns copy and paste on the main sheet.

**\*Data Validation**

Use your filter tool to validate each column’s values

To locate filter, click data tab, your Filter in the sort and filter group tab.

Then identify and corrected the following:

If you proceed to column I, meant for actual price, you will notice that the data type is incorrect. The correct value should be having 139,900.

Column K has this symbol (\) on it and it was replaced with 0

**2) Calculated Columns**

**Total Potential Revenue**

Click on column M to insert a new column beside column L.

Then label it like this

How to derive those values according to the instruction is to multiply actual price by rating count Next is column for % Discount Buckets of <50% and 50% or More

=IF(J2>=50%,"Yes","No") for the 1st product and double-click the formula down for the rest.

**Next is Price Range Bucket**

You can Insert the column beside discounted Price

Before you do this, you need to apply your currency to conform with what is on the Case Study. By applying this India currency on all price related columns. You can get it this way👇

Highlight all your price related columns at once and apply this format on them

To change the currency So once you've done that, go back to the new column

Copy this currency ₹ as you will need it as a text for your formula.

Then you go =IF (H2<200,"<₹200”, IF (OR (H2=200, H2<=500),"₹200 - ₹500",">₹500")) If you cannot copy the Indian currency, just copy it from this formula

**3)Data Table**

Turn it to a table by applying Ctrl + T

Question 7 &12 used formula

**4) Analysis and Presentation**

\*Analysis was done using Pivot table and Formula

\*Presentation using Pivot Chart