Case Study

Identification of promising e-commerce acquisition targets.

TASK

Your task is to sanitize and analyze a data set to profile the sellers present, and develop selection criteria to identify the best or most promising Amazon merchants in this dataset, that the Acquisitions team should reach out to, and acquire.

Approach

Analyze the dataset provided on the next slide using Tableau and present your insights (all data sanity should be done in Tableau or Tableau prep- feel free to use custom SQL as needed). Please do not do any manual data cleaning (only formula-based sanitization) , and do not use excel for this exercise, only use Tableau and SQL.

Results

Please present your insights in the form of a Tableau Dashboard or story. We will assess you on your data parsing process / SQL queries, as well as the insights and selection criteria you develop to analyze the dataset and present the best sellers. We will ask you to present your results in a joint call via screen share.

Garden Category Data from Amazon.

This is a sample dataset of publicly available info on e-commerce sellers in the Garden category in the Amazon marketplace.

Download the data set

Sample Longlist

Download the Tableau free trial

Click here to download a 14-day trial

version of Tableau to complete the

assignment

Preparing the data set for your analysis

Before you develop your selection criteria, you will need sanitize the data, and parse out useable data from the unstructured text. Here are some tips:

- The column **sellerproductcount** gives you the count of products in the form '1-16 of over 100,000 results' , and you can parse out the product count 100,000 .
- **sellerratings** - this columns gives you the % and count of positive ratings (e.g. 88% positive in the last 12 months (118 ratings)) if parsed correctly
- **sellerdetails** - you can use this text to parse out phone numbers, and email IDs of merchants, where available, so our team can reach out to them.
- **businessaddress** - this will give you the business locations of the sellers. You can parse them to identify if a seller is registered in the US , Germany (DE), or China (CN).
- **Hero Product 1 #ratings and Hero Product 2 #ratings** - these 2 columns give you the number of ratings of the 2 'hero products' or bestselling products of this seller.