Revenue Potential in Amazon UK

Exploratory Data Analysis

Name: Example 2

## **Problem Statement:**

Since the mid 1990, the popularity of Online Shopping has taken the world by storm. Due to the rise in popularity, local retailers are experiencing a decline in sales. You are hired to help an art supply retailer to identify possible opportunities to create an online shop through Amazon FBA program. You are tasked to investigate the pros and cons of using Amazon FBA as a distribution method for their store as well as revenue potential for items sold in similar product categories and whether there is a chance for them to generate revenues between £10,000 to £30,000+.

## Business Impact:

Exploring this data could inform retailers and opportunists of the potential risk and reward associated with Amazon FBA program. By examining past and existing Amazon listings, this study can provide an approximate earning potential as well as success and failure rates in a specific product category.

# **General Dataset Information:**

File Name: amz\_uk\_processed\_data.csv

Description: Amazon UK Products Dataset 2023

Dataset Details: 1,680,129 Rows & 10 Columns

Size: 635,804KB ( 620MB )

Source: Kaggle - [Dataset Link](https://www.kaggle.com/datasets/asaniczka/amazon-uk-products-dataset-2023)

## **Target Features:**

The problem statement outlines the essential metrics of finding revenue potential for using Amazon FBA program. The key features in question that can be found within the dataset:.

1. **Reviews** - Concrete evidence of items purchased by the user.
2. **Price** - Use to identify unit price sold for the listing.
3. **Stars** - Use to gauge popularity of an item.
4. **isBestSeller** - Use to gauge popularity of an item.
5. **boughtInLastMonth** - Possible gauge for potential revenues.
6. **CategoryName** - Essential for identifying item type to help retailers.
7. **RevenueByReview** - Feature Engineered column for revenue approximation.
8. **PotentialReview** - Feature Engineered column for potential future revenue.
9. **StarBins** - Feature Engineered column that groups star ratings together.

To answer the problem statement. These nine metrics in combination will be used to conduct the analysis.

# List of Analysis

## **Analysis #1 - Total Records & Average Unit Price**

To bring into perspective, the number of records was used to demonstrate the research effort in deriving the conclusion of revenue potential in the Art & Supply product category. After isolating the records by its Product Category, there are 46,324 records across 4 categories that the retailer can use to position its product. The average unit price is essential for the retailer to determine pricing strategy.

## **Analysis #2 - Review Distribution**

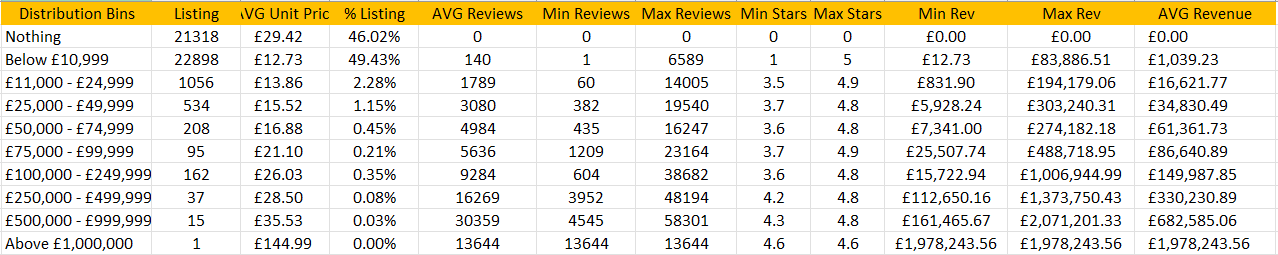
Using the records, a new table can be tabulated by grouping the “Reviews” attribute to determine how many listings are successful. The listings are grouped together by their appropriate review range. Using the bin identification, total listing count, as well as average revenues can be derived for analysis. The result demonstrates the number of reviews a listing is required to accumulate to acquire reasonable size revenue. As per the above diagram. It requires 900-999 reviews to generate revenue of £10,000+, which consists of only 0.47% of the listing being able to achieve these results. Since the original problem statement mentions a revenue range of £10,000 to £30,000. Further breakdown is needed.

## **Analysis #3 - Basic Statistics**

Using basic statistics, we can determine the spread of the listings to see whether it conforms to our previous analysis using the review bins. The central tendency shows a very high range of revenue (£1,967,242.57) with a very small mean (£50,227.61) and median (£24,935.58) range. The variance and standard deviation also suggest a wide range of dispersion. The combination of the data tells us that there is a disproportionate success to failure ratio. Further investigation is needed.

## **Analysis #4 - Revenue Distribution**

The review distribution might’ve revealed what a successful listing might look like, but it is not an accurate indicator of success. The statistics for listings with revenues of £11,000 contain a high level of dispersion. To further expand on the analysis. Revenue distribution table is created to show how many listings fall within a particular revenue range.



To provide a thorough analysis. All revenue ranges are considered starting from £0 and gradually increment towards £1,000,000. This distribution chart shows a clear picture that the majority of the listings makes less than £11,000, approximately 95.45%. The retailer as per our problem statement request wanted revenue potential of £10,000 to £30,000, which consist of 3.43% of the listings. As we approach a higher revenue bracket, success level drops exponentially.

To achieve the level of success requested by the retailer. An average of 1,789 to 3,080 reviews ( verify buyer ) is needed. This information can then be compared against their local retail traffic to determine the course of action needed to achieve the result.

# Conclusion

Out of the 46,324 records across 4 categories. There is a disproportionate success to failure ratio with 46.02% of listings making nothing, and 49.43% of listings making less than £11,000. To achieve reasonable success on Amazon FBA, collectively there is 4.55% of listings making above £11,000. All calculations performed in this analysis are an approximation. Note that there could be numerous unaccounted transactions that could increase the chances of success and revenue potential.