Insights to find the best Amazon merchants

1. Data Cleaning and Extraction

- From **sellerdetails** in the dataset extracted **Email address** using SQL.
- From **sellerdetails** in the dataset extracted **Phone Numbers** using Power BI.
- From **sellerratings** in the dataset extracted **rating Percentage** using SQL.
- From **sellerproductcount** in the dataset extracted **Total Produc**t that seller has using SQL.
- From **sellerdetails** in the dataset extracted **Phone Numbers** Using DAX formula in Power BI.
- From **businessaddress** in the dataset extracted **Country Code** using Dax formula in Power BI.

2. Insights To Find Best Seller

- These are the columns to in this data which are crucial to find best sellers:
    - **Hero Product 1 #ratings , Hero Product 2 #ratings**
    - **SellerProductsCount**
    - **Max % of negative seller ratings - last 30 days, 90 days, 12 Months.**
- A Seller with maximum interactions and fewer negative ratings is considered as a good Seller.

A. Max Hero Rating

- Since we have data for two hero products, I considered the maximum hero product rating from the two columns, as it reflects the highest level of user interactions with the seller, whether the rating is good or bad.
- Now I have **MaxHeroRating,** I use Bucketing technique and divide the sellers into A, B, C, D.

Condition	BucketMaxHeroRating	Total No.of seller in this Bucket
MaxHeroRating >= 5000	**A**	**113 sellers**
MaxHeroRating >= 1000	**B**	**126 sellers**
MaxHeroRating >= 100	**C**	**258 sellers**
MaxHeroRating < 100	**D**	**1287 sellers**

B. Max % of negative seller ratings

- Sellers with less negative rating status are Good sellers and High negative sellers are Bad sellers.

Condition	CustomerRatingStatus	Total No.of seller in this Bucket
Max % of negative seller ratings - last 12 months, 30 days, 90 days <= 25%	Good	1401 sellers
Max % of negative seller ratings - last 12 months, 30 days, 90 days > 25%	Bad	383 sellers

C. Products count

- Since each seller's product count varies significantly, I used a bucketing technique to categorise the sellers into A, B, C, and D based on their product count.

Condition	BucketProductCount	Total No.of seller in this Bucket
ProductCount] >= 7000	A	253 Sellers
ProductCount] >= 3000	B	238 Sellers
ProductCount] >= 1000	C	277 Sellers
ProductCount] > 0	D	363 Sellers
ProductCount] = 0	E	270 Sellers

- Sellers with maximum interactions, fewer negative ratings, and a higher product count are considered the best merchants.
- I created a cross-tabulation between **BucketProductCount** and **CustomerRatingStatus,** these are the results.

BucketProductCount	Bad	Total	Bad Percentage
A	98	351	27%
B	67	305	21%
C	85	362	23%
D	77	440	17%

- From the above insights, it is evident that sellers with fewer products are being classified as the best sellers, which introduces bias. Therefore, I have decided to exclude **BucketProductCount** from the analysis**.**

- So, Best way to find Best Seller is maximum interactions(**HeroProductRating)** and fewer negative ratings(**Max % of negative seller ratings).**

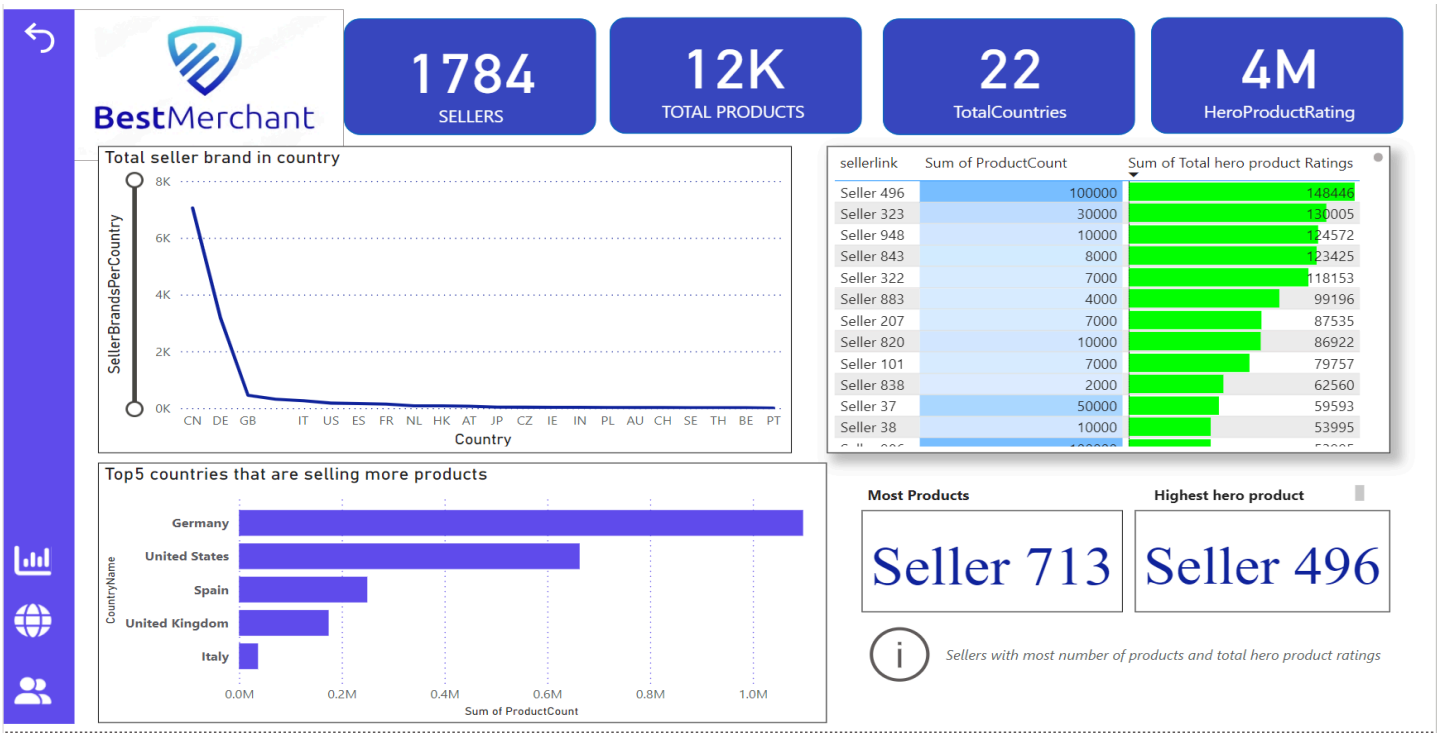
D. **Final segments**

- These are parameters I considered to find a best Seller.

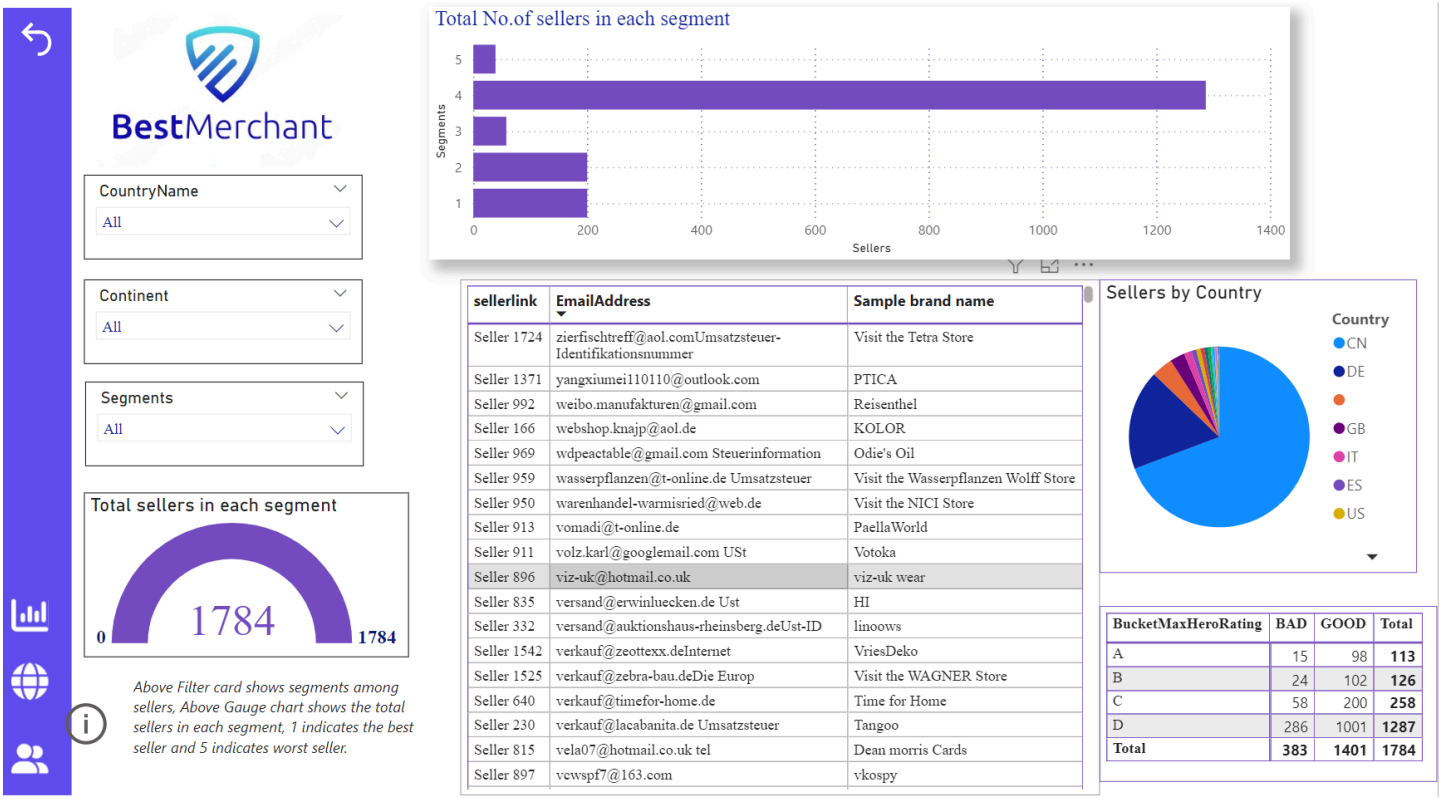
BucketMaxHeroRating	CustomerRatingStatus	Final Tile	Total No.of seller in this Bucket
A - B	Good	1	200 sellers
C	Good	2	200 sellers
C	Bad	3	58 sellers
D	Good/Bad	4	1287 sellers
A - B	Bad	5	39 sellers

3. Visualisation

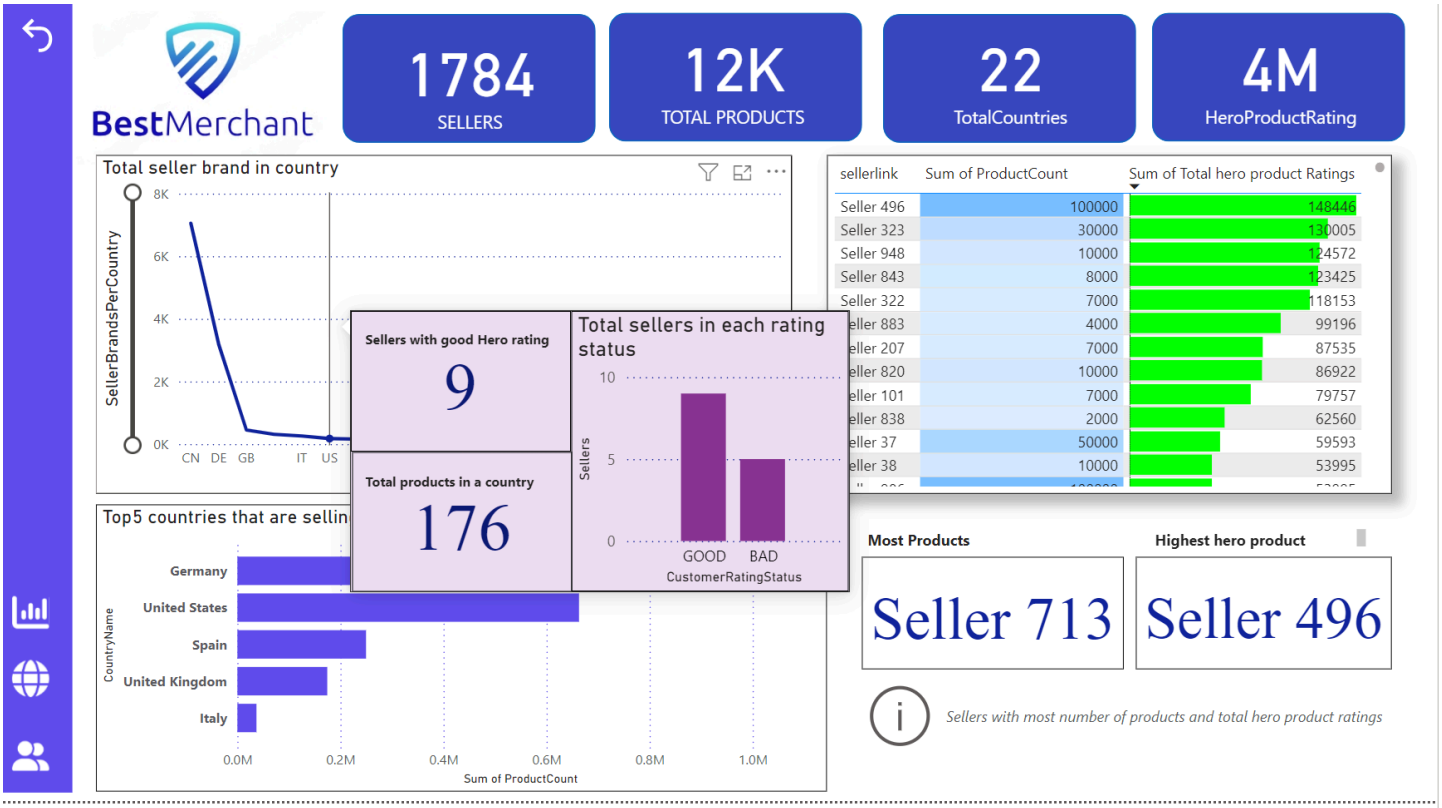
- I developed three dashboard pages for this data. The first page is the Executive Dashboard, providing overall insights from the data. I added several KPIs, bar charts, a line chart, and a table. To emphasise key points in the table, I applied conditional formatting.



- The second dashboard is the Seller Dashboard, where I filtered sellers based on the segments I considered. Additionally, I included filtering options based on Country Name, Continent, and segments.



- Additionally, I created a navigation pane to easily move between dashboards and added customised tooltips to provide additional information from the data.



- I am attaching sql file, power Bi file, Best sellers data for review.