DEPARTMENT OF COMPUTER SCIENCE & ENGINEERING
Symbiosis Institute of Technology, Pune
Symbiosis International (Deemed University)

A DSBI REPORT

ON

“Advanced Seller Insights: Amazon Performance Evaluation”

A DSBI report submitted in partial fulfillment of the requirements for the degree of

BACHELOR OF TECHNOLOGY in COMPUTER SCIENCE & ENGINEERING

Submitted By:

Sr. No.	Name	PRN	Division
1	Abhishek Rajput	22070122007	CSE A
2	Armaan Jeswani	22070122029	CSE A
3	Arnav Jain	22070122030	CSE A
4	Atul Goyal	22070122039	CSE A

UNDER THE GUIDANCE OF:
Dr. Deepak Dharrao

1. Abstract

The research creates multifactored mathematical scoring metrics to evaluate Amazon sellers whose performance incorporates various evaluation components. Traditional methods of assessing buyers primarily use ratings and general customer feedback although these methods fail to identify key differences between seller performance. The study's research benefit originates from its method which combines rating analysis and brand variance metrics with negative feedback identification patterns and award recognition for exceptional products. Rating systems become smarter through new built-in functions that overcome these systems' shortcomings to evaluate trustworthiness and excellence of vendors and products.

The dynamic data presentation system functions together with the Power BI performance score visualization to let users evaluate seller performance by examining product types coupled with feedback trends and brand selections. The active system delivers advanced evolving information beyond what static evaluation systems can produce. The research makes available extensive datasets containing vendor information and customer feedback evaluations. The data processing system uses Python to generate a specialized score that results from all data handling procedures including engineering and cleaning operations. The quality control system enhancement with selected seller criteria creates standardized evaluation methods for e-commerce platforms to achieve scalable efficient seller assessments that deliver better customer satisfaction.

2. Introduction

The project achieves seller performance evaluation by uniting business intelligence methods with machine learning algorithms. Traditional star rating methods applied by Amazon to evaluate sellers present a major weakness because they overlook essential product-related and service quality elements.

Business analytic operations rely on a Power BI tool as the main component in this project design. Power BI received selection over Tableau because the team preferred its superior Microsoft integration features and more powerful data functionalities. Users that access BI tools through these platforms can generate insights using filtering features although they lack advanced technical experience.

Businesses achieve improved operational performance through the combination of Power BI and ML because the system automatically conducts analytics and generates straightforward visuals that depict complicated data relationships. The researchers utilized their forecast results to show ML score predictions together with Power BI in this work.

Examples of How ML and Power BI Enhance Business:

1. Power BI together with ML provides retail businesses with demand forecasts which help them achieve maximum inventory levels.
2. Power BI displays consumer buying data through segmentation algorithms which ML uses to cluster customer groups.
3. Real-time monitoring of ML techniques through Power BI dashboards allows SaaS companies to detect financial fraud.
4. The displayed Power BI dashboards contain fraudulent transaction identification results through finance analysis techniques.
5. Power BI utilizes ML models alongside it to predict sales which creates visual goal tracking dashboards that show sales information.

3. Problem Statement

E-commerce websites normally base their seller evaluation approach on markers like customer reviews and delivery speed measures alongside star ratings. The establishment of these minimal seller evaluation systems proves completely ineffective at reflecting vendor performance output. The current system allows unperceivable suppliers who give exemplary support to escape notice even if they possess few ratings whereas sellers with numerous postings but inadequate service get dismissed.

These systems demonstrate total deficiencies in both transparency and adjustable performance metrics. There is limited focus on the analysis of important issues such as brand diversity and category-wise consistency and long-term feedback trends. Lack of vendor regulations creates a situation where below-average sellers stay unmonitored thereby threatening platform security as well as customer trust. The expanding market of Amazon requires an automated multi-attribute scoring system that can assess vendor performance based on their volume.

3.1 Objectives:

The principal focus of this project requires the development of a scalable explainable understandable method which evaluates Amazon retail outlets based on actual data records. A new analytical system should replace traditional star ratings by measuring multiple seller metrics which include how popular their products are and their brand variety and customer satisfaction rates.

1. The program utilizes Python along with data manipulation libraries pandas and numpy to organize and prepare the unprocessed Amazon seller information.

2. Engineer three features that track seller activities by measuring rating scores alongside long-term feedback changes and counting brand products.

3. Creating a weighted scoring system requires uniting at least three reasonable evaluation variables into one definite reporting method.

4. Engineering an interactive Power BI dashboard that enables visualizing seller scores with interface options for time-based or category selection and displays abnormal seller patterns and trends for better analytics.

5. The platform will present recommendations based on measures of sales promotion together with demotion functions and removal functions and focused support resources.

6. A Python-based process will preprocess and clean the raw Amazon seller dataset through the utilization of pandas and numpy data manipulation libraries.

3.2 Stakeholders and Use Cases

Different parties in the e-commerce system find practical value in the implemented framework. The scoring model and dashboard serve various strategic and operational needs.

Key Stakeholders:

• **Category Managers:** The scoring model enables sellers assessment for promotional activities including special listings and marketing campaigns.

• **Quality Assurance Teams:** Determine underperforming sellers through monitoring seller behaviors then put them forward for review or removal processes.

• **Business Intelligence Analysts:** These business analysts should evaluate longitudinal data about sellers and their categories to generate vital reports for organizational leadership.

• **Sellers Themselves**: Sellers achieve better visibility of their performance metrics through the system which helps them compare their results to competition.

• **Investors & Operations Teams**: Check the reliability of merchant partners and their future operations when making choices regarding partnerships or acquisitions.

4. Importance of Power BI and Machine Learning in Business

Power BI and Machine Learning have redefined current corporate actors particularly in industries which generate extensive data such as e-commerce. The combination of structured and unstructured data that companies produce becomes efficient practical knowledge through the use of Power BI and ML tools.

This project known as Amazon Seller Evaluation required the essential use of Power BI and ML to tackle authentic business problems.

4.1 Why Power BI is Important in This Project:

The software provides interactive dashboards and visual analytics which require no programming skills to allow stakeholders a fast understanding of seller performance data.

Visible from the graphic presentation were seller performance results alongside top performers and underperforming sellers while seasonal complaint patterns and score drop trends were displayed.

4.2 Why Machine Learning is Important in This Project:

Regression-based scoring models form part of the wide range of ML techniques that enable the analysis of historical data to predict future seller outcomes. Feature engineering alongside ML constructs allows multiple elements such as hero product ratings and brand count and negative feedback to be merged into a single performance scoring system.

The addition of ML models in the future will allow detection of high-risk sellers and performance drops so that platform efficacy can become more proactive.

5. Proposed architecture

5.1 Schematic Representation of the Project:

Data Collection: The obtained data from Amazon seller listings contains negative comments and brand differences together with product ratings that serves to generate seller performance statistics for graphical project visualization.

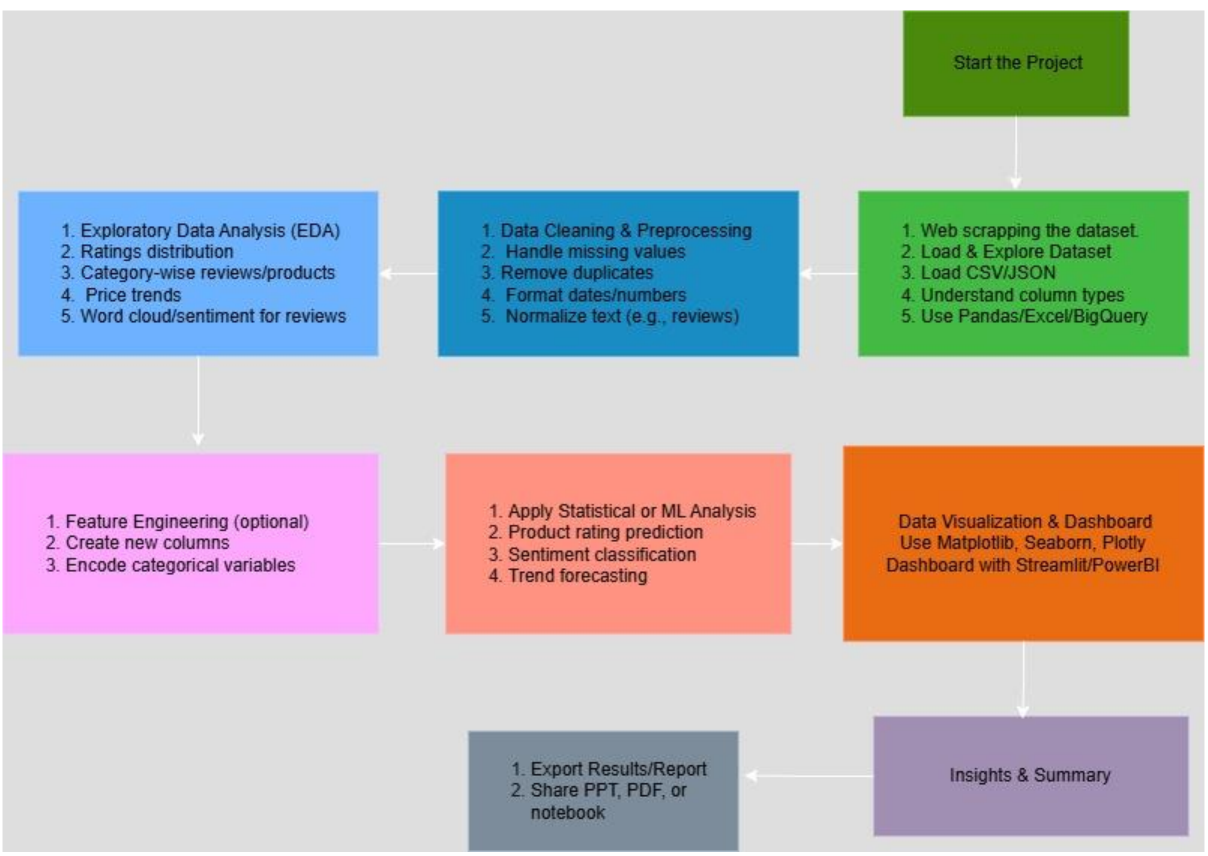
Data Processing: Python libraries like pandas, numpy were used in data processing to prepare raw data.

Feature Engineering: The feedback patterns and brand influence received customization to support the scoring model in its operation.

Performance Score Calculation: The weighted score system sorts sellers based on negative comments together with hero product ratings and brand variety.

Visualization: The data collection occurs through Power BI whereas visualization is displayed through interactive dashboards and visualizations.

Diagram/Flowchart:



5.2 Proposed Model for End-to-End Solution

The proposed model uses a comprehensive end-to-end method for operation.

Data Collection: Public data from Scrape Amazon sellers allows one to collect insights about product volumes and ratings together with negative feedback proportions and hero product performances and brand choices.

Data Processing: Python's data processing tools allow one to conduct standardization and missing value management and pre-processing while cleaning and omitting outliers from the data.

Scoring Model: The scoring system looks at seller performance through a weight-based mechanism determined by established parameters.

Front-End (Power BI): The results display performance scores and feedback distribution through Power BI which represents them alongside time-based and category-based trends.

6. Innovation in the project

The project stands unique because it built a multi-attribute performance scoring system which looks beyond ratings to include multiple important features.

- **Hero product ratings**: Commitment ratings on hero products serve as vital signs that demonstrate both customer trust and the achievement of products in market success.
- **Negative feedback percentage**: Negative feedback percentage measures reliability development through time-based percentages of negative feedback.
- **Brand diversity**: A company's brand diversity measures its growth level and size making it align with enduring brand diversification success.
- **Product Ratings Volume**: The total number of ratings for all products sold by a seller determines their presence and engagement with customers on the platform.
- **Sales Volume**: Sales Volume column demonstrates the seller’ ability to produce revenue.

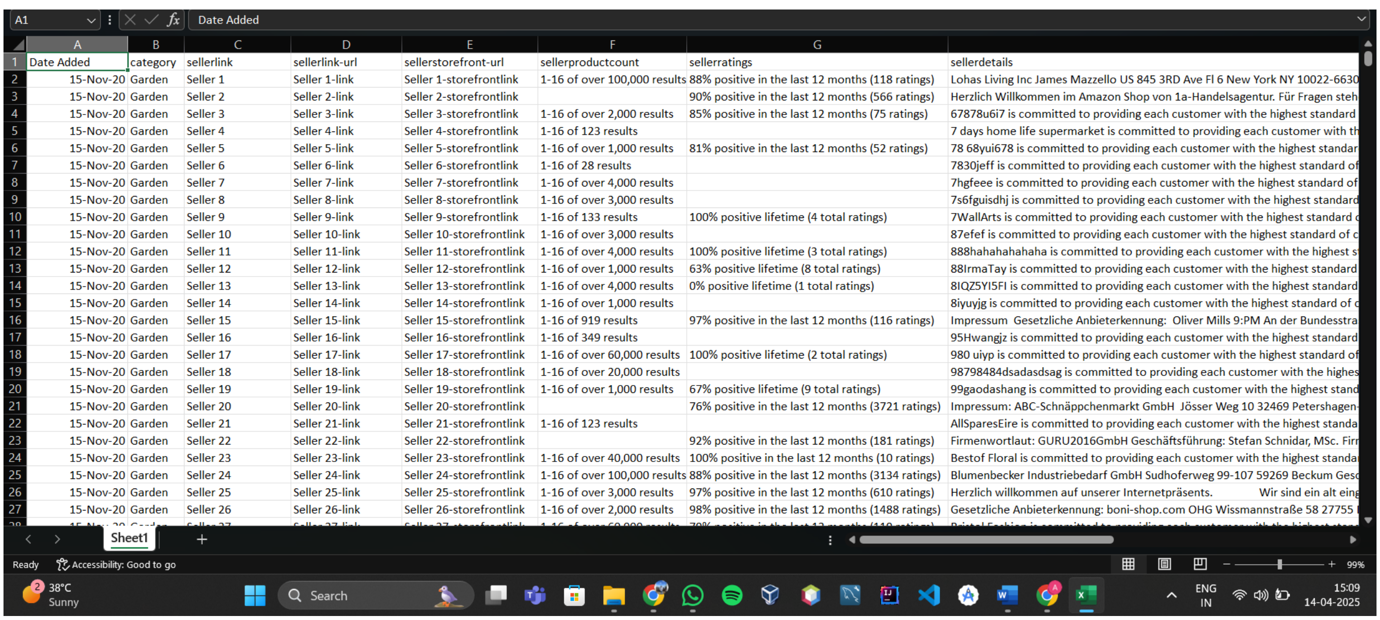
Traditional rating systems get surpassed by this approach which uses long-term trends and seller behavior measurements to explore various aspects regarding seller success so addressing many concerns. The system advances traditional rating solutions by examining various seller behaviors to present an improved deep understanding of seller quality and durability over time.

7. Workflow/Implementation

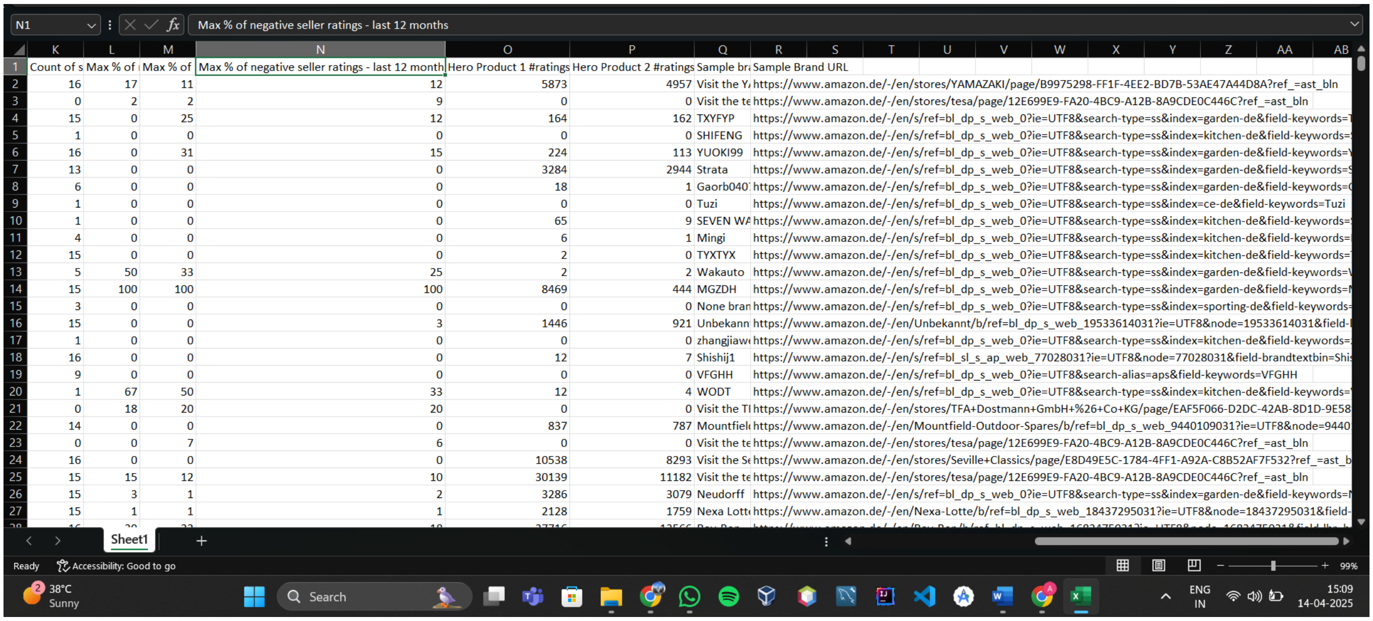
The dataset used in this project contains detailed information on Amazon merchants and a set of criteria needed for multi-dimensional performance analysis.

7.1 Dataset Overview (Tabular Form):

Attribute	Description
Seller Link	Unique URL of the seller storefront
Seller Product Count	Total number of products listed by the seller
Seller Ratings	Overall rating score and negative feedback percentages
Hero Product 1 Ratings	Number of ratings on the top-selling product
Number of Brands	Number of distinct brands handled by the seller
Negative Feedback %	Percentage of negative feedback
Negative Feedback % (365 days)	Percentage of negative feedback over the last year



	A	B	C	D	E	F	G	
1	Date Added	category	sellerlink	sellerlink-url	sellerstorefront-url	sellerproductcount	sellerratings	sellerdetails
2	15-Nov-20	Garden	Seller 1	Seller 1-link	Seller 1-storefrontlink	1-16 of over 100,000 results	88% positive in the last 12 months (118 ratings)	Lohas Living Inc James Mazzello US 845 3RD Ave Fl 6 New York NY 10022-6630
3	15-Nov-20	Garden	Seller 2	Seller 2-link	Seller 2-storefrontlink		90% positive in the last 12 months (566 ratings)	Herzlich Willkommen im Amazon Shop von 1a-Handelsagentur. Für Fragen steh
4	15-Nov-20	Garden	Seller 3	Seller 3-link	Seller 3-storefrontlink	1-16 of over 2,000 results	85% positive in the last 12 months (75 ratings)	67878u6i7 is committed to providing each customer with the highest standard
5	15-Nov-20	Garden	Seller 4	Seller 4-link	Seller 4-storefrontlink	1-16 of 123 results		7 days home life supermarket is committed to providing each customer with th
6	15-Nov-20	Garden	Seller 5	Seller 5-link	Seller 5-storefrontlink	1-16 of over 1,000 results	81% positive in the last 12 months (52 ratings)	78 68yui678 is committed to providing each customer with the highest standar
7	15-Nov-20	Garden	Seller 6	Seller 6-link	Seller 6-storefrontlink	1-16 of 28 results		7830jeff is committed to providing each customer with the highest standard of
8	15-Nov-20	Garden	Seller 7	Seller 7-link	Seller 7-storefrontlink	1-16 of over 4,000 results		7hgfeee is committed to providing each customer with the highest standard of
9	15-Nov-20	Garden	Seller 8	Seller 8-link	Seller 8-storefrontlink	1-16 of over 3,000 results		7s6fguisdhj is committed to providing each customer with the highest standard
10	15-Nov-20	Garden	Seller 9	Seller 9-link	Seller 9-storefrontlink	1-16 of 133 results	100% positive lifetime (4 total ratings)	7WallArts is committed to providing each customer with the highest standard o
11	15-Nov-20	Garden	Seller 10	Seller 10-link	Seller 10-storefrontlink	1-16 of over 3,000 results		87efef is committed to providing each customer with the highest standard of c
12	15-Nov-20	Garden	Seller 11	Seller 11-link	Seller 11-storefrontlink	1-16 of over 4,000 results	100% positive lifetime (3 total ratings)	888hahahahahaha is committed to providing each customer with the highest s
13	15-Nov-20	Garden	Seller 12	Seller 12-link	Seller 12-storefrontlink	1-16 of over 1,000 results	63% positive lifetime (8 total ratings)	88IrmaTay is committed to providing each customer with the highest standard
14	15-Nov-20	Garden	Seller 13	Seller 13-link	Seller 13-storefrontlink	1-16 of over 4,000 results	0% positive lifetime (1 total ratings)	8IQZ5YI5FI is committed to providing each customer with the highest standard
15	15-Nov-20	Garden	Seller 14	Seller 14-link	Seller 14-storefrontlink	1-16 of over 1,000 results		8iyuyjg is committed to providing each customer with the highest standard of c
16	15-Nov-20	Garden	Seller 15	Seller 15-link	Seller 15-storefrontlink	1-16 of 919 results	97% positive in the last 12 months (116 ratings)	Impressum Gesetzliche Anbieterkennung: Oliver Mills 9:PM An der Bundesstra
17	15-Nov-20	Garden	Seller 16	Seller 16-link	Seller 16-storefrontlink	1-16 of 349 results		95Hwangjz is committed to providing each customer with the highest standard
18	15-Nov-20	Garden	Seller 17	Seller 17-link	Seller 17-storefrontlink	1-16 of over 60,000 results	100% positive lifetime (2 total ratings)	980 uiyp is committed to providing each customer with the highest standard of
19	15-Nov-20	Garden	Seller 18	Seller 18-link	Seller 18-storefrontlink	1-16 of over 20,000 results		98798484dsadasdsag is committed to providing each customer with the highes
20	15-Nov-20	Garden	Seller 19	Seller 19-link	Seller 19-storefrontlink	1-16 of over 1,000 results	67% positive lifetime (9 total ratings)	99gaodashang is committed to providing each customer with the highest stand
21	15-Nov-20	Garden	Seller 20	Seller 20-link	Seller 20-storefrontlink		76% positive in the last 12 months (3721 ratings)	Impressum: ABC-Schnäppchenmarkt GmbH Jösser Weg 10 32469 Petershagen-
22	15-Nov-20	Garden	Seller 21	Seller 21-link	Seller 21-storefrontlink	1-16 of 123 results		AllSparesEire is committed to providing each customer with the highest standa
23	15-Nov-20	Garden	Seller 22	Seller 22-link	Seller 22-storefrontlink		92% positive in the last 12 months (181 ratings)	Firmenwortlaut: GURU2016GmbH Geschäftsführung: Stefan Schnidar, MSc. Firr
24	15-Nov-20	Garden	Seller 23	Seller 23-link	Seller 23-storefrontlink	1-16 of over 40,000 results	100% positive in the last 12 months (10 ratings)	Bestof Floral is committed to providing each customer with the highest standa
25	15-Nov-20	Garden	Seller 24	Seller 24-link	Seller 24-storefrontlink	1-16 of over 100,000 results	88% positive in the last 12 months (3134 ratings)	Blumenbecker Industriebedarf GmbH Sudhoferweg 99-107 59269 Beckum Gesc
26	15-Nov-20	Garden	Seller 25	Seller 25-link	Seller 25-storefrontlink	1-16 of over 3,000 results	97% positive in the last 12 months (610 ratings)	Herzlich willkommen auf unserer Internetpräsents. Wir sind ein alt eing
27	15-Nov-20	Garden	Seller 26	Seller 26-link	Seller 26-storefrontlink	1-16 of over 2,000 results	98% positive in the last 12 months (1488 ratings)	Gesetzliche Anbieterkennung: boni-shop.com OHG Wissmannstraße 58 27755



N1 Max % of negative seller ratings - last 12 months

	K	L	M	N	O	P	Q	R
1	Count of s	Max % of	Max % of	Max % of negative seller ratings - last 12 month	Hero Product 1 #ratings	Hero Product 2 #ratings	Sample bra	Sample Brand URL
2	16	17	11	12	5873	4957	Visit the Y	https://www.amazon.de/-/en/stores/YAMAZAKI/page/B9975298-FF1F-4EE2-BD7B-53AE47A44D8A?ref_=ast_bln
3	0	2	2	9	0	0	Visit the te	https://www.amazon.de/-/en/stores/tesa/page/12E699E9-FA20-4BC9-A12B-8A9CDE0C446C?ref_=ast_bln
4	15	0	25	12	164	162	TXYFYP	https://www.amazon.de/-/en/s/ref=bl_dp_s_web_0?ie=UTF8&search-type=ss&index=garden-de&field-keywords=
5	1	0	0	0	0	0	SHIFENG	https://www.amazon.de/-/en/s/ref=bl_dp_s_web_0?ie=UTF8&search-type=ss&index=kitchen-de&field-keywords=
6	16	0	31	15	224	113	YUOKI99	https://www.amazon.de/-/en/s/ref=bl_dp_s_web_0?ie=UTF8&search-type=ss&index=garden-de&field-keywords=
7	13	0	0	0	3284	2944	Strata	https://www.amazon.de/-/en/s/ref=bl_dp_s_web_0?ie=UTF8&search-type=ss&index=garden-de&field-keywords=
8	6	0	0	0	18	1	Gaorb040	https://www.amazon.de/-/en/s/ref=bl_dp_s_web_0?ie=UTF8&search-type=ss&index=garden-de&field-keywords=
9	1	0	0	0	0	0	Tuzi	https://www.amazon.de/-/en/s/ref=bl_dp_s_web_0?ie=UTF8&search-type=ss&index=ce-de&field-keywords=Tuzi
10	1	0	0	0	65	9	SEVEN WA	https://www.amazon.de/-/en/s/ref=bl_dp_s_web_0?ie=UTF8&search-type=ss&index=kitchen-de&field-keywords=
11	4	0	0	0	6	1	Mingi	https://www.amazon.de/-/en/s/ref=bl_dp_s_web_0?ie=UTF8&search-type=ss&index=kitchen-de&field-keywords=
12	15	0	0	0	2	0	TYXTYX	https://www.amazon.de/-/en/s/ref=bl_dp_s_web_0?ie=UTF8&search-type=ss&index=kitchen-de&field-keywords=
13	5	50	33	25	2	2	Wakauto	https://www.amazon.de/-/en/s/ref=bl_dp_s_web_0?ie=UTF8&search-type=ss&index=garden-de&field-keywords=
14	15	100	100	100	8469	444	MGZDH	https://www.amazon.de/-/en/s/ref=bl_dp_s_web_0?ie=UTF8&search-type=ss&index=garden-de&field-keywords=
15	3	0	0	0	0	0	None bran	https://www.amazon.de/-/en/s/ref=bl_dp_s_web_0?ie=UTF8&search-type=ss&index=sporting-de&field-keywords=
16	15	0	0	3	1446	921	Unbekann	https://www.amazon.de/-/en/Unbekannt/b/ref=bl_dp_s_web_19533614031?ie=UTF8&node=19533614031&field-
17	1	0	0	0	0	0	zhangjiaw	https://www.amazon.de/-/en/s/ref=bl_dp_s_web_0?ie=UTF8&search-type=ss&index=kitchen-de&field-keywords=
18	16	0	0	0	12	7	Shishij1	https://www.amazon.de/-/en/s/ref=bl_sl_s_ap_web_77028031?ie=UTF8&node=77028031&field-brandtextbin=Shi
19	9	0	0	0	0	0	VFGHH	https://www.amazon.de/-/en/s/ref=bl_dp_s_web_0?ie=UTF8&search-alias=aps&field-keywords=VFGHH
20	1	67	50	33	12	4	WODT	https://www.amazon.de/-/en/s/ref=bl_dp_s_web_0?ie=UTF8&search-type=ss&index=kitchen-de&field-keywords=
21	0	18	20	20	0	0	Visit the T	https://www.amazon.de/-/en/stores/TFA+Dostmann+GmbH+%26+Co+KG/page/EAF5F066-D2DC-42AB-8D1D-9E58
22	14	0	0	0	837	787	Mountfiel	https://www.amazon.de/-/en/Mountfield-Outdoor-Spares/b/ref=bl_dp_s_web_9440109031?ie=UTF8&node=9440
23	0	0	7	6	0	0	Visit the te	https://www.amazon.de/-/en/stores/tesa/page/12E699E9-FA20-4BC9-A12B-8A9CDE0C446C?ref_=ast_bln
24	16	0	0	0	10538	8293	Visit the S	https://www.amazon.de/-/en/stores/Seville+Classics/page/E8D49E5C-1784-4FF1-A92A-C8B52AF7F532?ref_=ast_b
25	15	15	12	10	30139	11182	Visit the te	https://www.amazon.de/-/en/stores/tesa/page/12E699E9-FA20-4BC9-A12B-8A9CDE0C446C?ref_=ast_bln
26	15	3	1	2	3286	3079	Neudorff	https://www.amazon.de/-/en/s/ref=bl_dp_s_web_0?ie=UTF8&search-type=ss&index=garden-de&field-keywords=
27	15	1	1	1	2128	1759	Nexa Lott	https://www.amazon.de/-/en/Nexa-Lotte/b/ref=bl_dp_s_web_18437295031?ie=UTF8&node=18437295031&field-

```python
for i, col in enumerate(df.columns):
    print(f"{i+1}. {col}")
```

```
1. Date Added
2. category
3. sellerlink
4. sellerlink-url
5. sellerstorefront-url
6. sellerproductcount
7. sellerratings
8. sellerdetails
9. seller business name
10. businessaddress
11. Count of seller brands
12. Max % of negative seller ratings - last 30 days
13. Max % of negative seller ratings - last 90 days
14. Max % of negative seller ratings - last 12 months
15. Hero Product 1 #ratings
16. Hero Product 2 #ratings
17. Sample brand name
18. Sample Brand URL
```

7.2 Exploratory Data Analysis (Data Cleaning and Preparation):

The primary dataset was cleaned and pre-processed using Python libraries, pandas and numpy which ensured the precision, consistency, and dependability of the scoring model depended on this last phase.

- **Preprocessing**: In Preprocessing removing of duplicates, null values, and outliers helps to clear the data. Standardize columns depending on both statistically and functionally homogeneity.
- **Exploratory Data Analysis (EDA)**: By EDA, we viewed the ratings, product counts, and feedback % distribution graphically became simple. Search histograms and boxes for deviations and abnormalities.
- **ML Algorithms Used**:
  - **Scoring Formula**: Based on Scoring Formula, custom grading system, brand diversity, and hero product was evaluated.
  - **Why ML**: Based on past data trends, forecasted sales using either linear regression or ensemble technique.
- **Database**: Whether simple connecting with Power BI is possible will rely on the CSV file keeping the dataset.

```
df.isnull().sum()
```

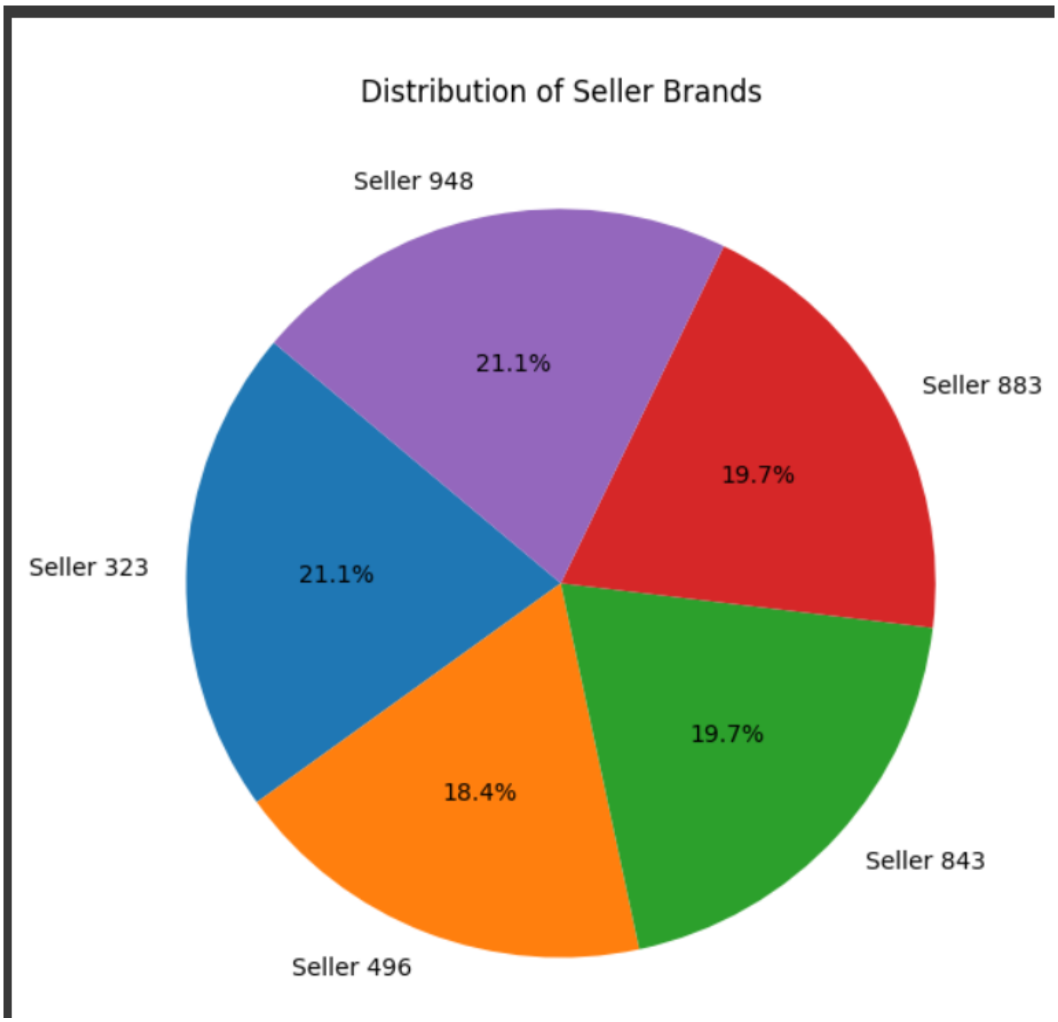
	0
Date Added	0
category	0
sellerlink	0
sellerlink-url	0
sellerstorefront-url	0
sellerproductcount	326
sellerratings	428
sellerdetails	27
seller business name	56
businessaddress	75
Count of seller brands	0
Max % of negative seller ratings - last 30 days	0
Max % of negative seller ratings - last 90 days	0
Max % of negative seller ratings - last 12 months	0

7.3 Methodology:

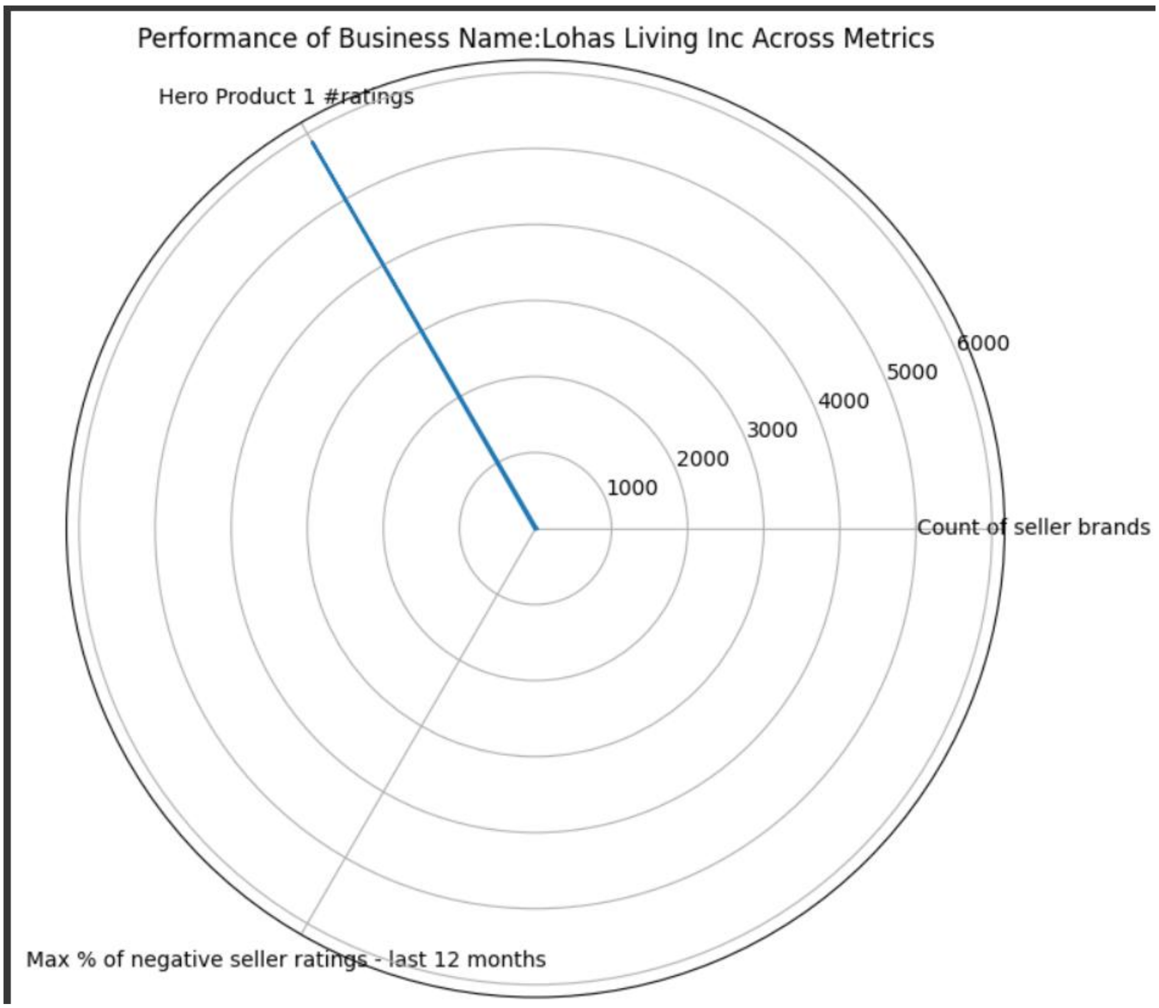
The project involves transformation of raw seller data into insightful information by using a precise, multi-phase strategy comprising analytical processing and visual reporting. Every stage of the treatment promised consistent, unambiguous, instructive results.

1. Data Collection:
   The dataset was scraped from the publicly available Amazon seller storefronts. Among the data were brand counts, ratings, product counts, feedback %, hero product statistics, and seller linkages.
2. Data Cleaning:
   First Discovered and then dealt with missing numbers, incorrect formatting, and outliers. Text fields were standardized; numerical fields were tuned to fit computation.
3. Feature Engineering:
   Among the ones added were normalized feedback ratings, brand influence weights, and a single performance score—new factors aimed to offer a full picture of vendor quality.
4. Scoring Model Application:
   seller got a different Performance based on hero product engagement, brand diversity, and negative feedback trends by using the Scoring Model.
5. Export for Visualization:
   The cleaned and improved data was exported into CSV form to demonstrate a correct Power BI interface.
6. Dashboard Development:
   To allow a dynamic analysis of seller performance, the structured data was visulaized using charts, slicers, score tables, and category filters thereby.

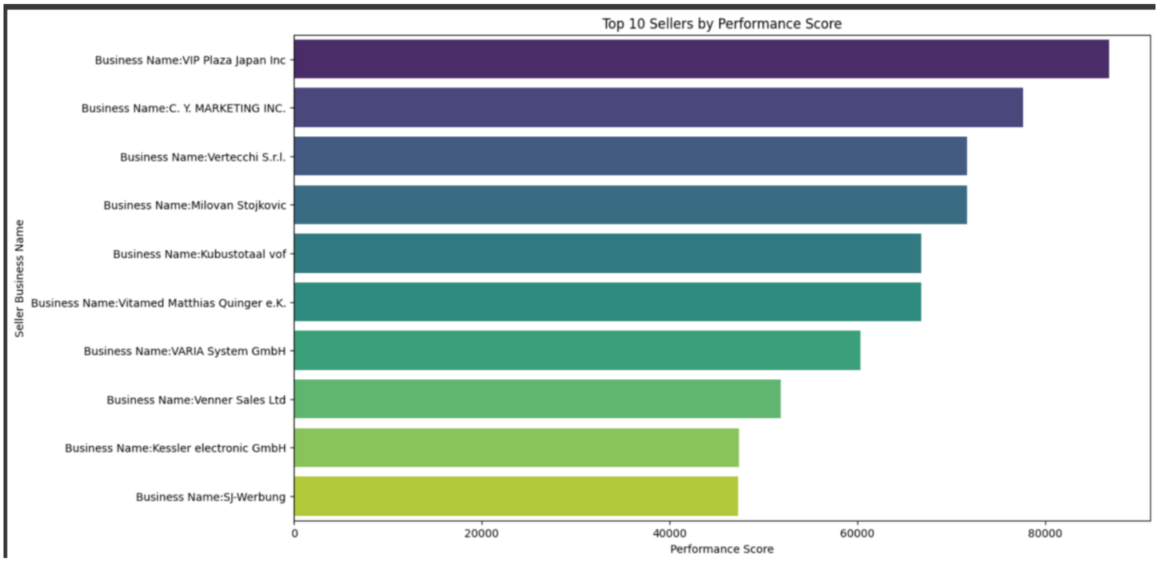
```
import matplotlib.pyplot as plt
import numpy as np

top_sellers = df.groupby('sellerlink').sum(numeric_only=True).sort_values(by=['Hero Product 1 #ratings'], ascending=False).head(5)
relevant_sellers = df[df['sellerlink'].isin(top_sellers.index)]

categories = ['Hero Product 1 Ratings', 'Hero Product 2 Ratings', 'Max % Negative Ratings (Last 12 Months)', 'Count of Seller Brands']
radar_data = []

for seller in relevant_sellers['sellerlink'].unique():
    seller_data = relevant_sellers[relevant_sellers['sellerlink'] == seller]
    values = [
        seller_data['Hero Product 1 #ratings'].values[0],
        seller_data['Hero Product 2 #ratings'].values[0],
        seller_data['Max % of negative seller ratings - last 12 months'].values[0],
        seller_data['Count of seller brands'].values[0]
    ]
    radar_data.append(values)

radar_data = np.array(radar_data)
radar_data = radar_data / radar_data.max(axis=0)


num_vars = len(categories)

angles = np.linspace(0, 2 * np.pi, num_vars, endpoint=False).tolist()
```



```python
import matplotlib.pyplot as plt

negative_ratings_data = df[['sellerlink', 'Max % of negative seller ratings - last 12 months']].dropna()


if not negative_ratings_data.empty:

    negative_ratings_data = negative_ratings_data.sort_values(by='Max % of negative seller ratings - last 12 months', ascending=False)


    plt.figure(figsize=(15, 7))
    plt.barh(negative_ratings_data['sellerlink'], negative_ratings_data['Max % of negative seller ratings - last 12 months'], color='red')
    plt.xlabel('Max % of Negative Ratings (Last 12 Months)')
    plt.ylabel('Seller')
    plt.title('Negative Ratings Percentage for Each Seller')
    plt.gca().invert_yaxis()
    plt.tight_layout()
    plt.show()
else:
    print("No data available to plot.")
```



```
# Best Seller Identification
best_sellers = df[['seller business name', 'Hero Product 1 #ratings', 'Max % of negative seller ratings - last 12 months', 'Count of seller brands']].copy()
best_sellers['Performance Score'] = best_sellers['Hero Product 1 #ratings'] - best_sellers['Max % of negative seller ratings - last 12 months'] * 10 + best_sellers
best_sellers = best_sellers.sort_values(by='Performance Score', ascending=False).head(10)

plt.figure(figsize=(14, 8))
sns.barplot(x='Performance Score', y='seller business name', data=best_sellers, palette='viridis', hue='seller business name', legend=False)
plt.title('Top 10 Sellers by Performance Score')
plt.xlabel('Performance Score')
plt.ylabel('Seller Business Name')
plt.show()
```



7.4 Performance Score Formula Used:

Performance Score = (Hero Product 1 Ratings) − (Max % Negative Seller Ratings over 12 Months × 10) + (Count of Seller Brands × 5)

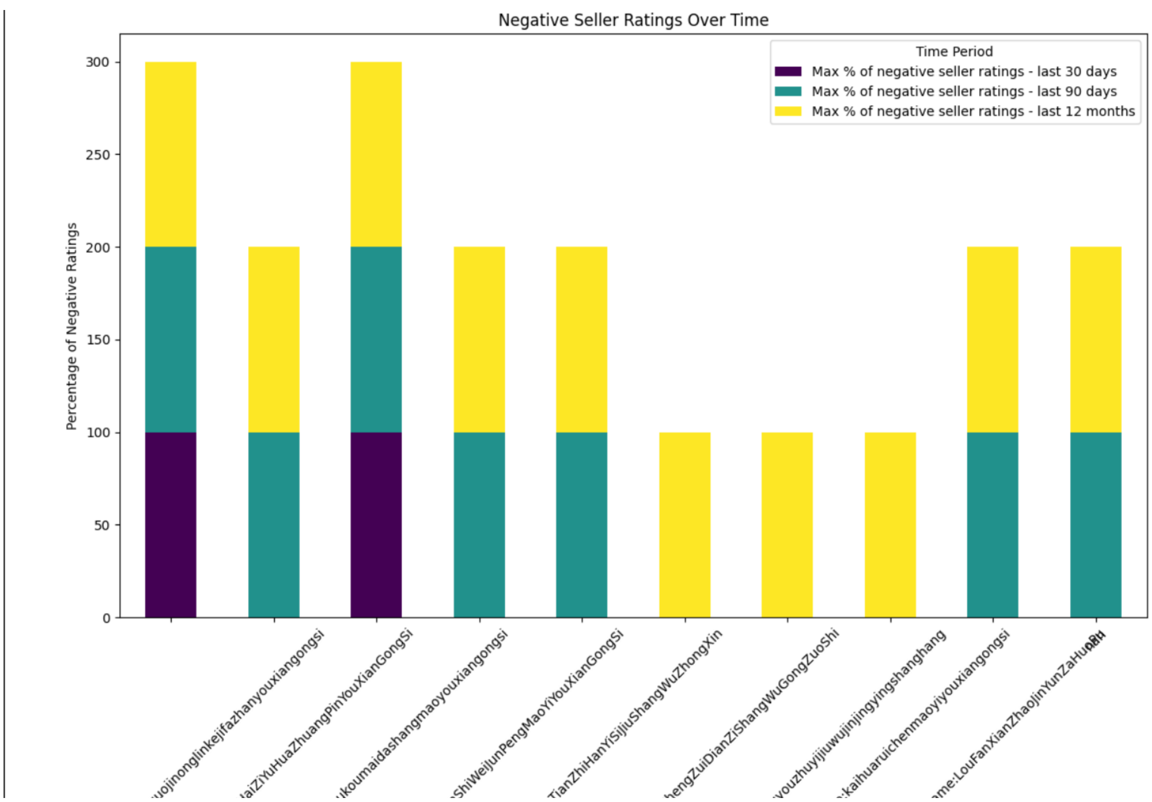
Explanation of Components:

- Hero Product 1 Ratings:
  Showcases consumer interaction with the most sold item by the seller.Usually, A higher rating reflects trust and popularity.
- Max % Negative Seller Ratings – Last 12 Months:
  This is the highest percentage of unfavorable evaluations gathered during the previous 12 months. A penalty is taken since a higher value indicates declining quality of services.
- Count of Seller Brands:
  The brand count of the vendor shows the number of brands under their control. Usually perceived favorably and indicating diversity and organizational maturity is a larger count.

Weight Selection Rationale:

- Negative reviews underwent a 10-penalty multiplier to underline the need of consistency and customer satisfaction.
- The brand count was multiplied by five to indicate its relevance without discounting the influence of product reviews or assessments.

```
# Distribution of Negative Ratings
plt.figure(figsize=(14, 8))
negative_ratings = df[['seller business name', 'Max % of negative seller ratings - last 30 days',
                       'Max % of negative seller ratings - last 90 days',
                       'Max % of negative seller ratings - last 12 months']].sort_values(by='Max % of negative seller ratings - last 12 months', ascending=False).head(10)

negative_ratings.set_index('seller business name').plot(kind='bar', stacked=True, figsize=(14, 8), cmap='viridis')
plt.title('Negative Seller Ratings Over Time')
plt.xlabel('Seller Business Name')
plt.ylabel('Percentage of Negative Ratings')
plt.xticks(rotation=45)
plt.legend(title='Time Period')
plt.show()
```



7.5. Predictive Modeling for Seller Performance

Using a predictive modeling technique, we estimated future selling performance relying on past data patterns, so improving the multi-attribute performance rating system. Using Random Forests, a ML model, complexity of many attributes like brand variation, negative feedback, product evaluations, and sales volume. Having been trained on historical performance data, the model generated learning patterns suggesting how the behavior of a seller will change with time.

The Random Forest model was used to forecast based on their historical data to check whether a seller would remain in the "at-risk," advance, or keep their performance level. The computer generated a risk score for every vendor, thereby implying the probability of future performance decrease. High risk score suppliers could be under observation or the object of proactive action.

Outcome and Impact:

Early warnings: The program informed vendors displaying indicators of impending performance drop which allowed participants to react pro-actively, supporting one another or altering their visibility.

Operational Efficiency: By means of automatic hazard detection, the model lessened the need for on-hand inspections, so saving time and money in the process of seller assessment.

At last, the predictive model improved our scoring system and provided a forward-looking view that helped us in tackling problems before they became more important and better manage salespeople.

8. Results

Power BI was used to create interactive dashboards that graphically show seller performance data following the establishment of performance ratings and export of a cleaned dataset. The dashboards were supposed to allow users quickly notice trends, investigate specific sites or eras, and evaluate vendors.

Key Dashboard Components:

- Seller Leaderboard:
  Performance ratings guide this bar graph showing the top and bottom performing sellers. Users could now more quickly spot which stores are succeeding and which would require more work.
- Score vs. Brand Count: This scatter or column visual showed the benefit of brand diversity by proving the link between performance scores and the number of brands each vendor manages.
- Feedback Distribution View:
  Feedback Distribution View is created by charts displaying the negative review distribution over 30, 90, and 365 days. This permits one to observe over time trends in service quality.
- Hero Product Comparison Table:
  This matrix table helps determine whether a strong hero product is linked with general seller success by means of a comparison between sellers' ratings of hero products and their overall performance score.
- Filters and Slicers:
  This let data be filtered with category, time period, brand count range, and score range which increases usability and simplifies focused research.
- Category-wise Summary:
  Seeing visuals showing aggregated figures by product category helps managers assess seller distribution and performance inside particular segments, category-wise.

```
df['Performance Score'] = (
    df['Hero Product 1 #ratings'] -
    (df['Max % of negative seller ratings - last 12 months'] * 10) +
    (df['Count of seller brands'] * 5)
)

best_seller_df = df[['sellerlink', 'Performance Score']].sort_values(by='Performance Score', ascending=False)


best_seller_name = best_seller_df.iloc[0]['sellerlink']
best_seller_score = best_seller_df.iloc[0]['Performance Score']

print(f"The best seller is: {best_seller_name} with a Performance Score of {best_seller_score}")
```

```
The best seller is: Seller 948 with a Performance Score of 86876
```

```
[78] df['Performance Score'].max()
```

```
86876
```

Seller Country Location
Australia
Austria
Mail IDs
-
Seller Business Name
Klaus Haberl Gmbh
Lackner & Urnitsch Ges.M....
performance score
-5
86935
Hero Product 1 #ratings
0
86856
Sum of performance score by Seller Business Name
Lackner & Urni...
Klaus Haberl G...
Wasser- Und ...
Vitapur Gmbh
Trendagentur ...
Walter Werkze...
0K
2K
Sum of performance score
Seller Business Name
Sum of Hero Product 1 #ratings by Seller Business Name
Lackner & Urn...
Klaus Haberl ...
Wasser- Und ...
Vitapur Gmbh
Trendagentur ...
Walter Werkz...
0K
2K
Sum of Hero Product 1 #ratings
Seller Business Name
Sum of Negative seller ratings - 12 months by Seller Business Name
(0%)
0
(47.06%)
(0%)
(0%)
Seller Business Na...
Wasser- Und G...
Walter Werkzeu...
Trendagentur K...
Vitapur Gmbh
Klaus Haberl G...
Lackner & Urnit...
Sum of Negative seller ratings - 12 months by Seller Business Name
(0%)
0
(47.06%)
(0%)
Seller Busines...
Wasser- Un...
Walter Wer...
Trendagent...
Vitapur Gm...
Klaus Habe...
Lackner & ...
Sum of mean of negative ratings by Seller Business Name
0.04
0.02
0.00
Sum of mean of negative ...
Walter Wer...
Trendagen...
Wasser- U...
Vitapur G...
Klaus Habe...
Lackner & ...
Seller Business Name
Sum of performance score and Sum of Count of Seller Brands by Hero Product 1 #...
2391
1323
88%
99%
1178
1760
100%
98%
9...
Sum of performance score by Seller Business Name
(0%)
(0%)
(0%)
1K
(18.42%)
(0%)
Seller Busine...
Lackner & ...
Klaus Habe...
Wasser- U...
Vitapur G...
Trendagen...
Walter Wer...

9. Output and Insights

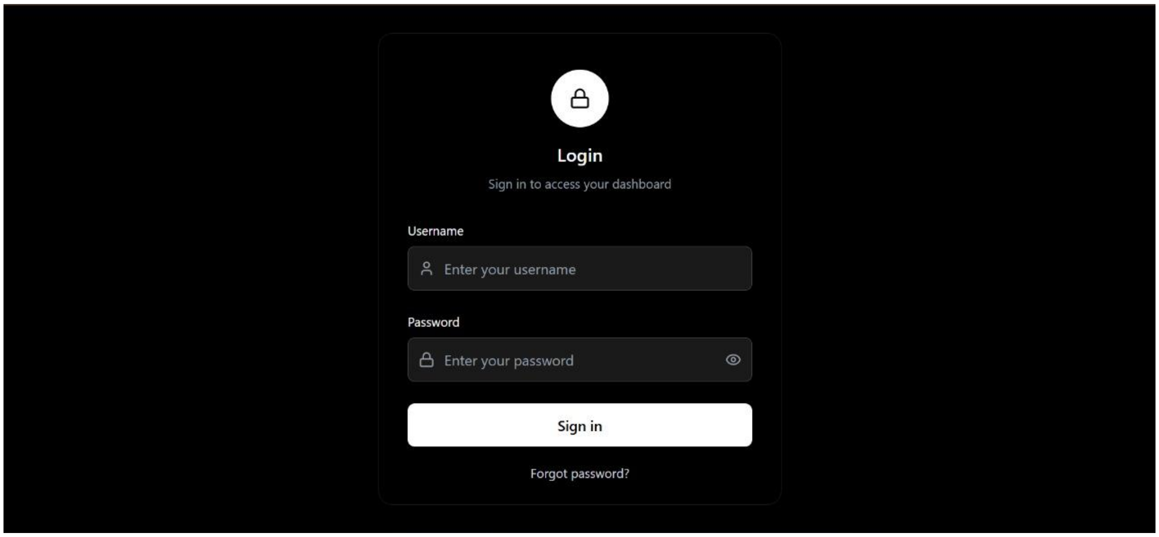
9.1 Key Results:

- Best sellers were hero products with great ratings and few negative reviews. This revealed that a good way to evaluate a seller's whole performance is by looking at a popular product and high caliber of services.
- A few merchants with a lot of inventory did not rank very high. High percentages of unfavorable evaluations over 90 or 365 days lowered their overall score despite their size, implying that scale is not always a fair gauge of good performance.
- Those who handled several brands typically scored better. This helps to justify the idea that a varied portfolio of brands helps sellers to act professionally and preserve standards of quality.
- Variables like feedback window, product number, or brand diversity on the Power BI dashboard help one find both high and low performers easily.
- For example, some new vendors showed quite good performance as their hero product engagement rate was high and there were no unfavorable remarks in the past months.

9.2 Key Insights:

- Attract and keep customers depend on hero product performance, hence it is a major score indicator.
- Since they reflect continuous service quality, long-term unfavorable reviews—such as the 365-day period—turned out to be more dependable than short-term assessments.
- Though brand count by itself does not ensure a good score, when combined with less negative reviews it usually corresponds with greater performance.
- Trends peculiar to several categories revealed: the average negative input was generally greater in various categories, implying that more tight quality control is needed in these sectors.

These results support the significance of dashboard-based data in impacting seller-related actions as well as the efficiency of the grading system.


Login
Sign in to access your dashboard
Username
Enter your username
Password
Enter your password
Sign in
Forgot password?


Welcome, User!
You are now viewing the OTT dashboard
AMAZON SELLER DATA ANALYSIS
This project evaluates Amazon sellers using a custom scoring model developed on real-world data, processed using Python and shown with Power BI. Developing a multi-attribute performance score based on brand variation and feedback patterns in addition to ratings is the novelty of this research. By means of Power BI, this project visualizes the scores and enables interactive analysis to provide actionable information about seller behavior—which is not possible with present instruments. The project generates an extensive seller evaluation system that might be implemented into e-commerce platforms to improve consumer experience and quality control. Scraped from publicly available seller data from Amazon, the dataset provides the foundation for this study guided by business intelligence and data science to support decisions.

10. Conclusion

This project effectively created a scalable, open, transparent method based on actual market data for assessing Amazon vendors. It presented a whole picture of seller performance by merging data science and business intelligence tools, much beyond conventional star ratings or feedback summaries.

Based on genuine elements including hero product ratings, negative feedback trends, and brand volatility, the unique performance score allows a fair and consistent evaluation of s ellers in all categories and times.

Among other stakeholders category managers, quality control teams, BI analysts, and actual sellers, the Power BI dashboard allows them to view these data. All things considered, this project has the ability to greatly enhance the way e-commerce systems track, assist, and maximize its seller environment. Apart from providing a path for intervention and development, it helps spot outstanding performers.

11. References

[1] Microsoft Corporation, "Power BI documentation," 2024.

[2] A. Ferrari and M. Russo, *The definitive guide to DAX: Business intelligence with Microsoft Excel, SQL Server Analysis Services, and Power BI*, 2nd ed. Redmond, WA: Microsoft Press, 2021.

[3] A. Mehta, *Mastering Microsoft Power BI: Expert techniques for effective data analytics and business intelligence*. Birmingham, U.K.: Packt Publishing, 2022.

[4] R. Puls, *Analyzing data with Power BI and Power Pivot for Excel*. Redmond, WA: Microsoft Press, 2020.

[5] P. Gorissen, "Power BI data analysis and visualization," 2020.

[6] U. Kumar, *Data analytics with Power BI: A practical guide to storytelling with data*. New York, NY: Apress, 2023.

[7] M. Helfert, "Business intelligence–past, present, and future," in *Lecture Notes in Business Information Processing*, vol. 96, Springer, 2011, pp. 1–14.

[8] S. Few, *Now you see it: Simple visualization techniques for quantitative analysis*. Oakland, CA: Analytics Press, 2009.

[9] W. McKinney, *Python for data analysis: Data wrangling with Pandas, NumPy, and IPython*, 2nd ed. Sebastopol, CA: O'Reilly Media, 2017.