Sri Indu
College of Engineering & Technology
UGC Autonomous Institution
Recognized under 2(f) & 12(B) of UGC Act 1956,
NAAC, Approved by AICTE &
Permanently Affiliated to JNTUH
Estd.2001

Ministry of Education

BRAIN O VISION

Sri Indu College of Engineering and Technology

HACKATHON - 2023

Team : 60

PALLE PRANAY

ALLI MAHESH

MEKALA SAI GNANESHWAR

YANAM MADHUKAR

VOGULAPU VAISHNAVI

S BHARGAVI

Industry: Education

Department: Inventory and Procurement

(Procurement and inventory management is the process of deciding what to buy and when)

PROBLEM STATEMENT

To acquiring and maintaining a collection of books in the market that meets to diverse needs and preferences.

SOLUTION

Based on the historical borrowing data and feedback of the customers. We will increase the number of helpful books in market.

Advantages:

- Paper waste is reduced.
- Storage waste will be reduced.
- Demand will be increases for the books
- Cost efficiency

ABSTRACT

The inventory procurement process in marketing books is a critical component of the publishing and distribution industry. This abstract provides a concise overview of the key aspects involved in managing book inventory for marketing purposes. It encompasses activities such as forecasting demand, supplier selection, order management, storage, and distribution logistics. Efficient inventory procurement in book marketing is essential for meeting customer demands, reducing costs, and optimizing sales. This abstract highlights the significance of streamlined inventory management strategies and their impact on the success of marketing books in a competitive marketplace.

INTRODUCTION

The education industry is a vast and dynamic sector that plays a pivotal role in shaping the future of societies and individuals. Within this industry, the Department of Inventory Procurement plays a crucial role in ensuring the efficient and effective management of educational resources. This introduction provides an overview of the significance of inventory procurement within the education industry and outlines its multifaceted responsibilities, which encompass acquiring, managing, and distributing a wide range of educational materials, equipment, and supplies.

Features:

- Author
- Language
- Price
- Rating
- Best Selling

DATASET

For Building our prediction model, We had used a data set from the Kaggle website. There are five attributes in this dataset , which has been divided into training data and testing data. The remaining data is used for testing , with the remaining 20% used for model training. The training dataset is used to create a prediction model for divers needed and preference books and the test set is used to assess the regression model. The data sets discription is displayed in the table below

Author	Writer of the Book
Language	Which language does the Book was published
Price	The cost of the Book
Rating	Whether the book has high rating or low rating
Best Selling	Whether the book has the best selling

FEATURE ENGINEERING:

According to table, that data set consist of five variables each of these features can help us estimate

The customers Interests

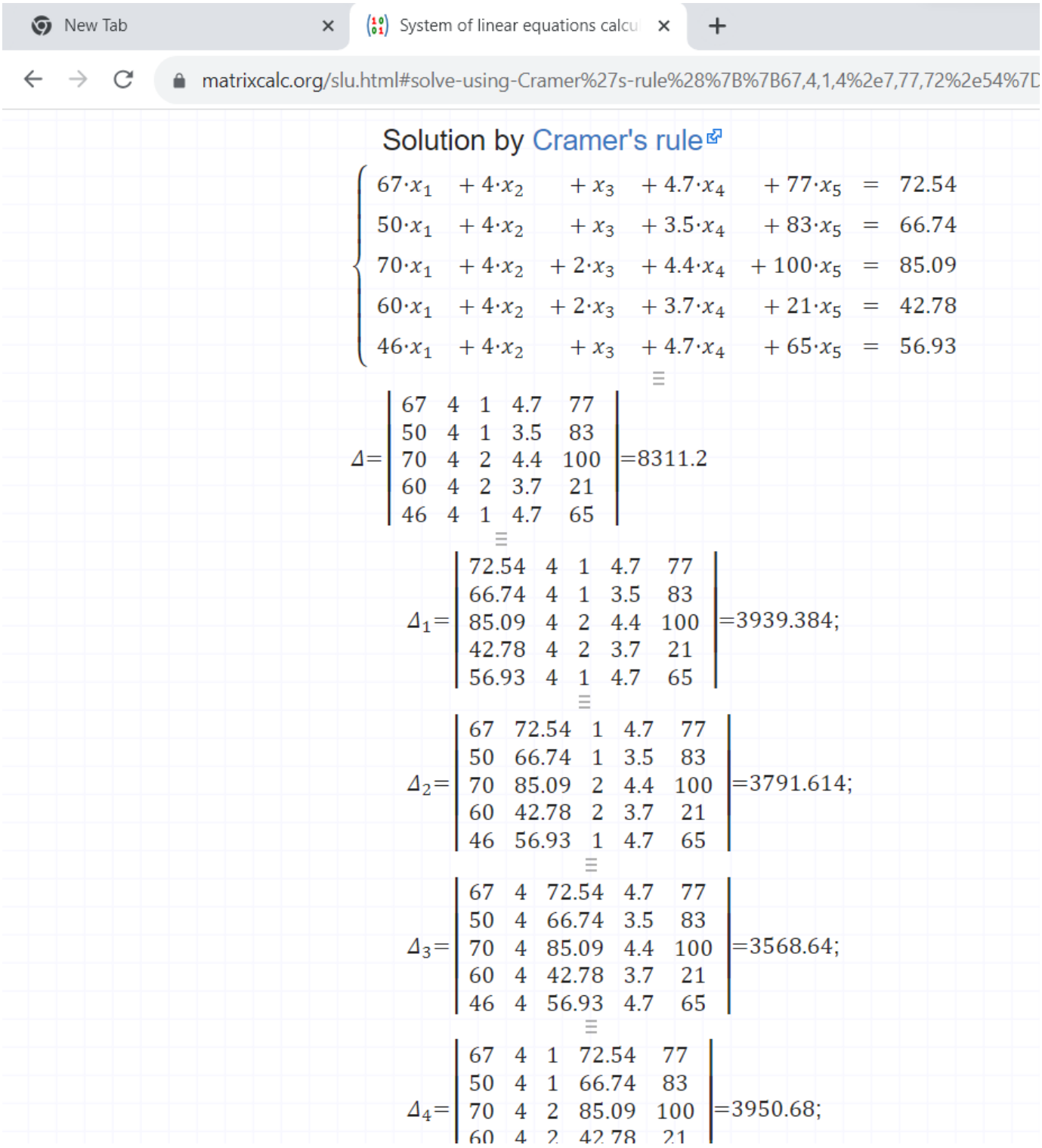
	A	B	C	D	E	F
1	Author's[Popularity per 100]	Language[4.Eng,3.Tel,2.Hin,1.Oth]	Cost[1.Low,2.Medium,3.Expensive]	Rating[5]	Best Selling[M]	Customer Interest
2	67	4	1	4.7	77	72.5
3	50	4	1	3.5	83	66.74528302
4	70	4	2	4.4	100	85.09433962
5	60	4	2	3.7	21	42.78301887
6	80	4	2	2.9	40	60.80188679
7	46	4	1	4.7	65	56.93396226
8						

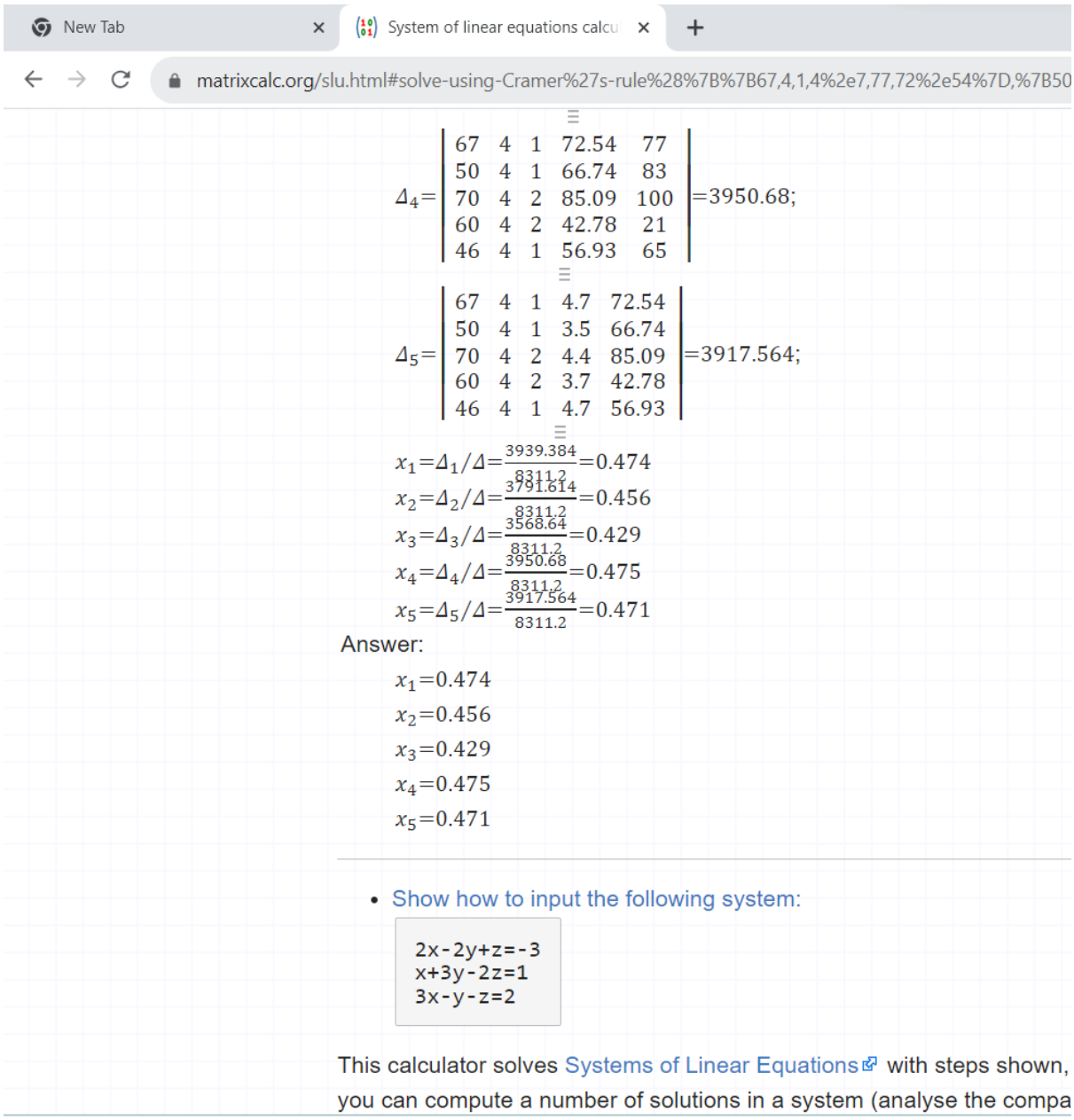
Note: We get the data from the Kaggle

Source Link...

Document of the Link....

Matrix Calculation:





Note:

By Cramer's Rule Feature engineering matrix will be solved

[Source Link....]

PREPROCESSING OF DATA

According to Table, the data sets consists of five variables

Each of these features can help us to estimate the customer interest ,which is dependent variable to some extent. To effectively apply to the data to the ML algorithms ,it is examined and updated in these stage.

In order to represent into numerical the features are transformed into numerical values.

The author of the book is categorical as the popularity. The popularity will be taken as per 100 . The categories of the features are taken in the columns such as language, price, rating, best selling. We encoded the customers interest in the books. According to the given dataset.

We now analyse the other independent variable along with the dependent variable

```
In [74]: plt.scatter(df['Languages'],df['Buyer Rating'] ,color='b')
Out[74]: <matplotlib.collections.PathCollection at 0x286af41e230>
```

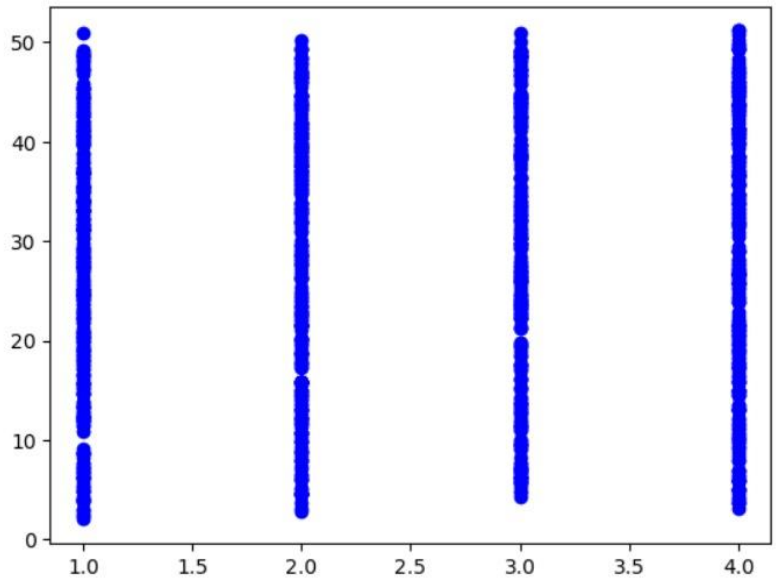


Fig .1 Language to customer interest plot

```
In [73]: plt.scatter(df['Price'],df['Buyer Rating'] ,color='b')
Out[73]: <matplotlib.collections.PathCollection at 0x286af357910>
```

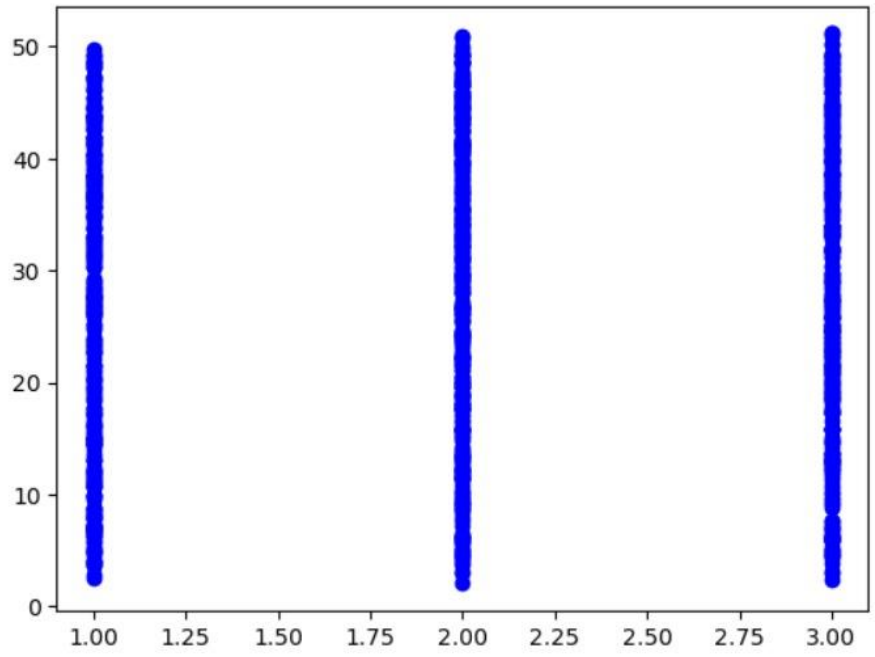


Fig .2 Price to customer interest plot

```
In [70]: plt.scatter(df['Rating'],df['Buyer Rating'] ,color='b')
Out[70]: <matplotlib.collections.PathCollection at 0x286af369cf0>
```

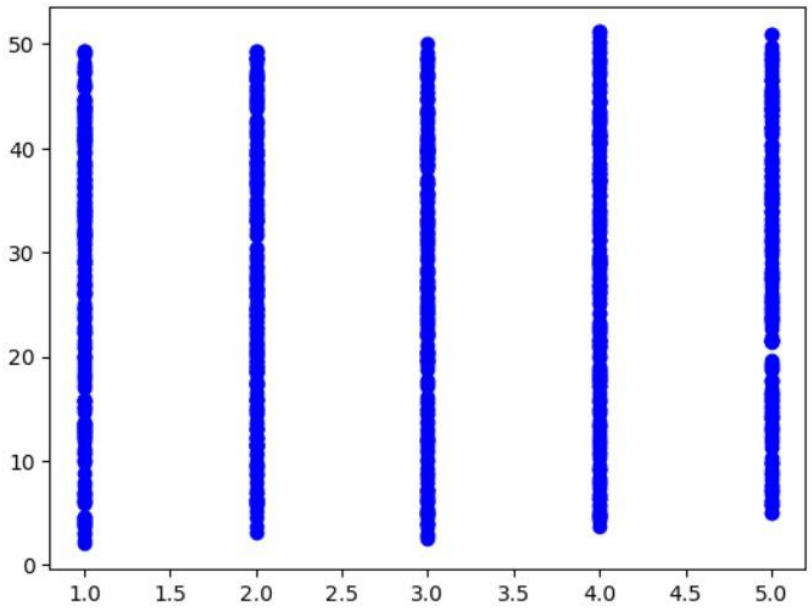


Fig .3 Rating per customer interest plot

```
In [75]: plt.scatter(df['Best Selling'],df['Buyer Rating'] ,color='b')
Out[75]: <matplotlib.collections.PathCollection at 0x286af3cfa90>
```

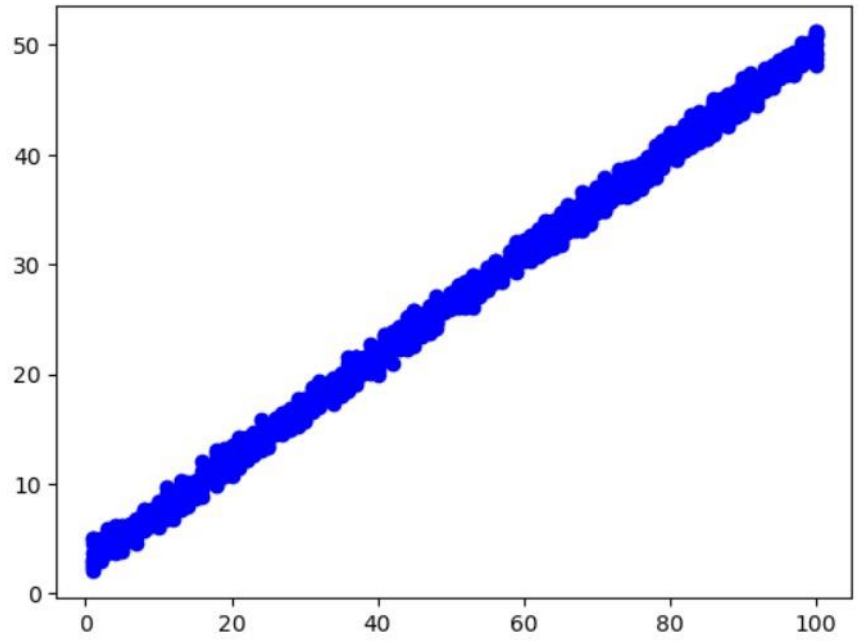


Fig .3 Best Selling per customer interest plot

PREDICTIVE MODELLING APPROACH

A machine learning approach is a good option for predicting best book because it can manage the dataset's complicated linkages and patterns. Regression analysis algorithm is example of machine learning approaches that can be used. The task requirements and dataset properties determine the technique to use.

REGRESSION ANALYSIS

Since regression analysis is a widely-used method for forecasting continuous variables, it is pertinent for estimating customer interest on buying books. It simulates the link between the dependent variable and the independent variables. Regression models can be as basic as linear regression or as complex as multiple regression, which enables the simultaneous inclusion of many factors. Interpretable coefficients from regression analysis show the amount and direction of the association between variables and medical costs. It can assist in determining the main causes of decreasing demand of book.

LINEAR REGRESSION

In [60]:
```
from sklearn.linear_model import LinearRegression
X_train, X_test, Y_train, Y_test = train_test_split(df[['Best Selling', 'Buyer Rating']], df['Buyer Rating'], test_size=0.25)
mlr=LinearRegression()

mlr.fit(X_train,Y_train)

pred_mlr = mlr.predict(X_test)
```

In [61]:
```
from sklearn.metrics import r2_score

r2_score = r2_score(Y_test, pred_mlr)

print("R2 score:", r2_score)
```

```
R2 score: 1.0
```

MODEL TRAINING AND EVALUATION

Model training and evaluation are critical steps in the process of customer interest prediction. These steps involve training the predictive models using the dataset and assessing their performance to determine their effectiveness. Here's an overview of model training and evaluation

1.Splitting the Dataset:

The first step is to split the dataset into training and testing sets. The training set is used to train the models, while the testing set is used to evaluate their performance. Typically, a random or stratified split

is performed, ensuring that both sets represent the overall characteristics of the data

```
In [60]: from sklearn.linear_model import LinearRegression
         X_train, X_test, Y_train, Y_test = train_test_split(df[['Best Selling', 'Buyer Rating']], df['Buyer Rating'], test_size=0.25)
         mlr=LinearRegression()

         mlr.fit(X_train,Y_train)

         pred_mlr = mlr.predict(X_test)
```

2. Model Training:

Train the predictive models using the training set. The specific approach depends on the chosen algorithms,17 such as regression analysis, decision tree algorithms, or machine learning techniques. Regression Analysis: Fit the regression models using the training set. Adjust the model parameters to optimize the fit, such as through ordinary least squares or iterative optimization algorithms.

3. Model Evaluation:

Evaluate the trained models using the testing set to assess their performance and generalization capabilities. Various evaluation metrics can be

utilized,18 depending on the nature of the problem and the specific goals:

- Mean Absolute Error (MAE): Measure the average absolute difference between the predicted customers interest and the actual expenses in the testing set. It provides an indication of the average prediction error.

- Root Mean Square Error (RMSE): Calculate the square root of the average squared difference between predicted and actual health expenses. RMSE is more sensitive to large errors compared to MAE.

- R-squared (R^2): Determine the proportion of customer interest that is explained by the models. A higher R-squared value indicates a better fit.

- Accuracy Measures: In classification scenarios, accuracy, precision, recall, or F1-score can be used to evaluate the performance of the models

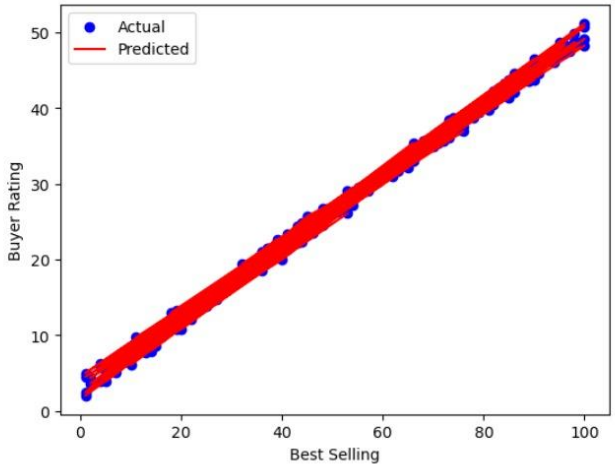
RESULT

The Regression analysis model was take for problem to be resolved and the model was highly accurate and offered insightful information about the variables affecting resolution

OUTPUT:

In [72]:
```
import matplotlib.pyplot as plt

plt.scatter(X_test["Best Selling"], Y_test, color='b', label='Actual')
plt.plot(X_test["Best Selling"], pred_mlr, color='r', label='Predicted')
plt.xlabel("Best Selling")
plt.ylabel("Buyer Rating")
plt.legend()
plt.show()
```



REFERENCE

The real time data we collected : pdf

Source Code

Kaggle

Good reach

D12 Blog