Machine Learning Model Proposed for E-Commerce

Name : Azis Muslim

Role : Data Scientist

Slides Points

Problem Statement

Dataset

Data Visualization

Linear Regression Model

Model Evaluation

Conclusions

Problem Statement

There are some merchants on e-commerce that issues for inventory management because they need to sell more fast moving product to decrease inventory cost. This mini research gonna be building machine learning model to predict ratings based on views and product price category. We hope this model gonna be beneficial for the merchant owner to decide which product should they sell to decrease the inventory cost.

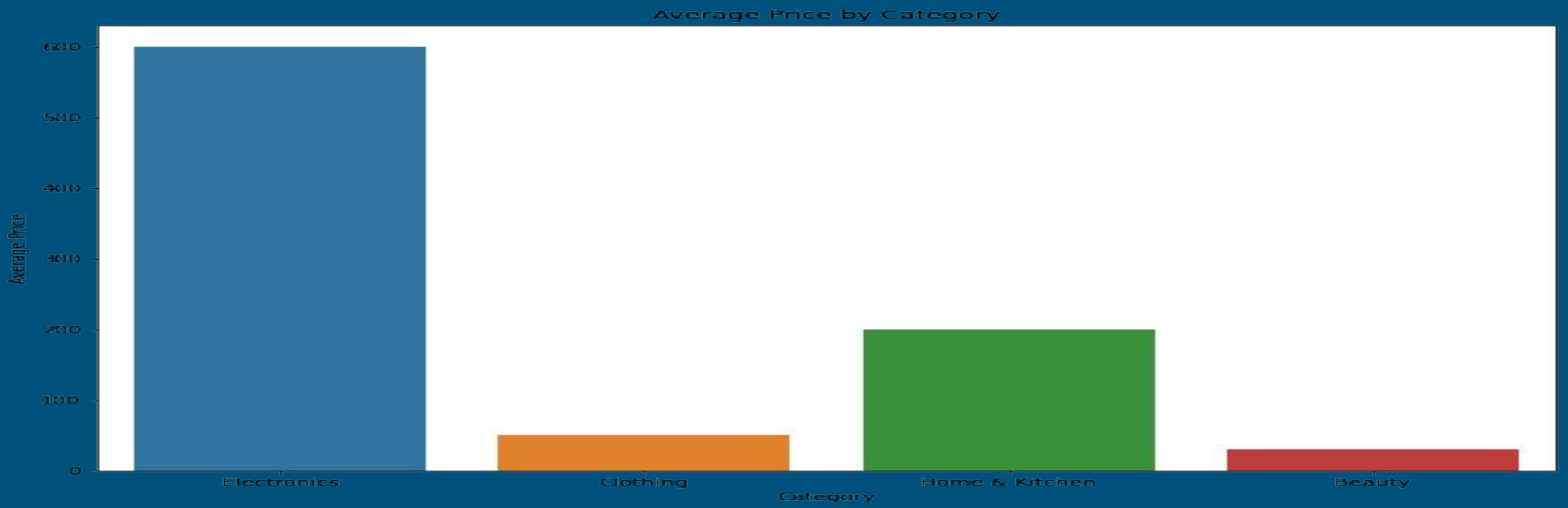
Dataset

01 Product Details

02 Purchase History

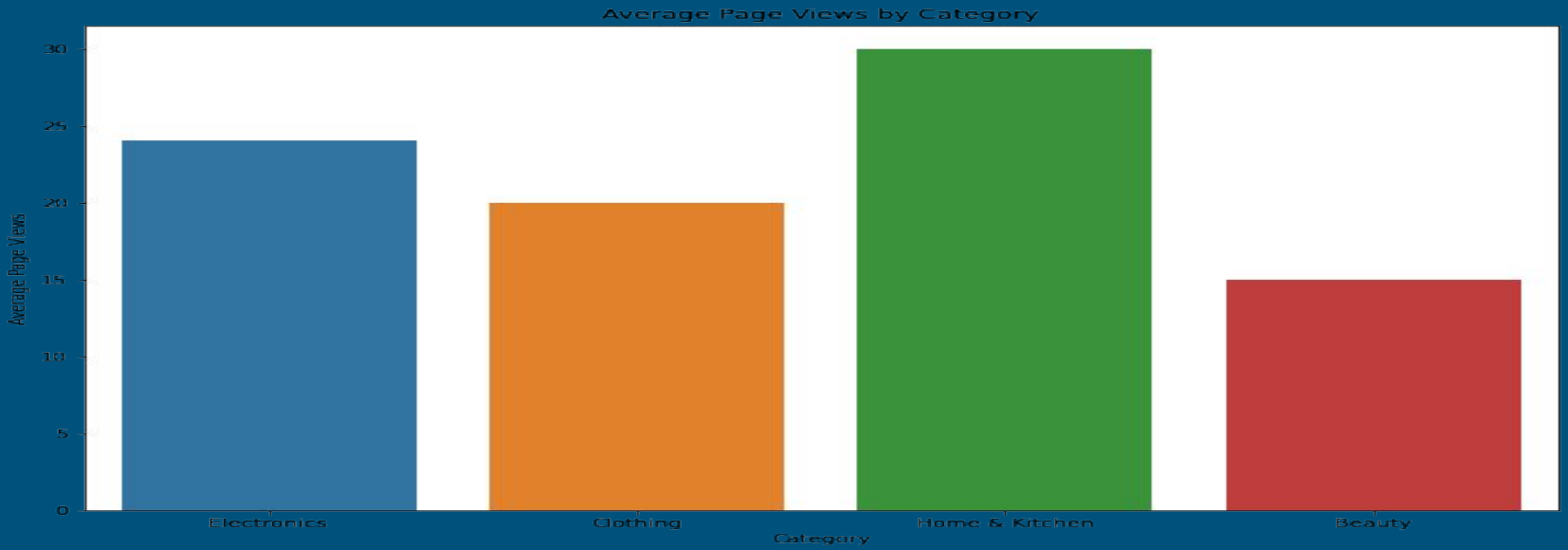
03 Customer Interactions

Data Visualization



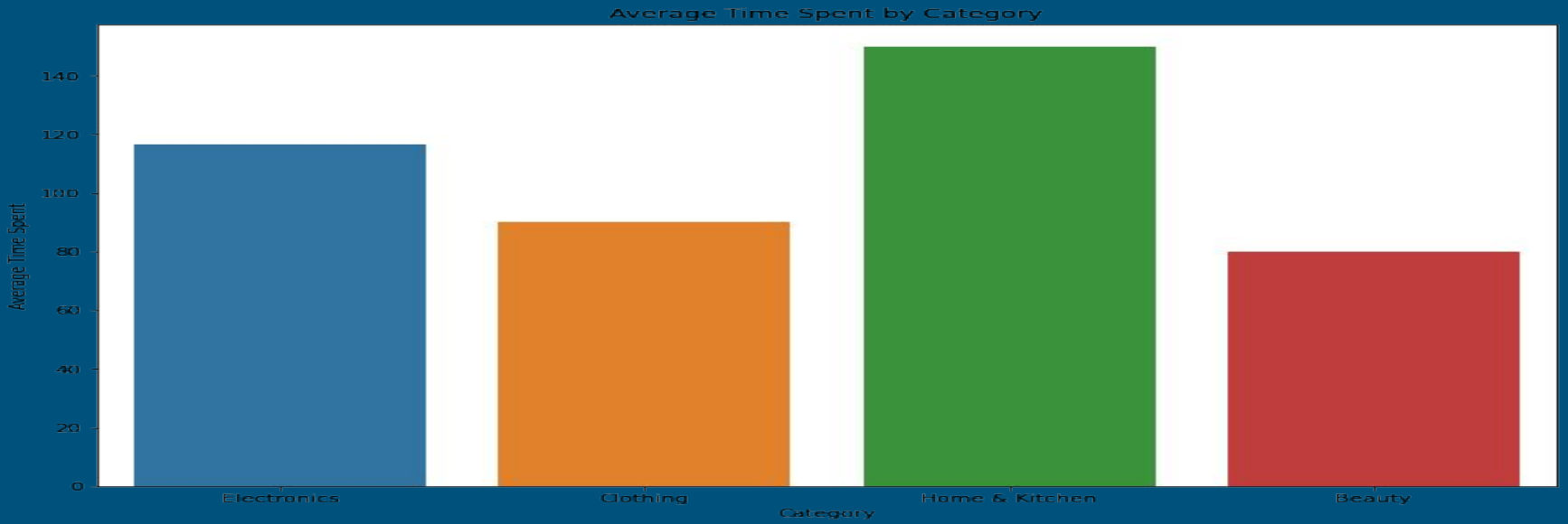
Based on the barplot above we can see that electronics have the higher price level followed by home & kithen category, clothing and beauty

Data Visualization



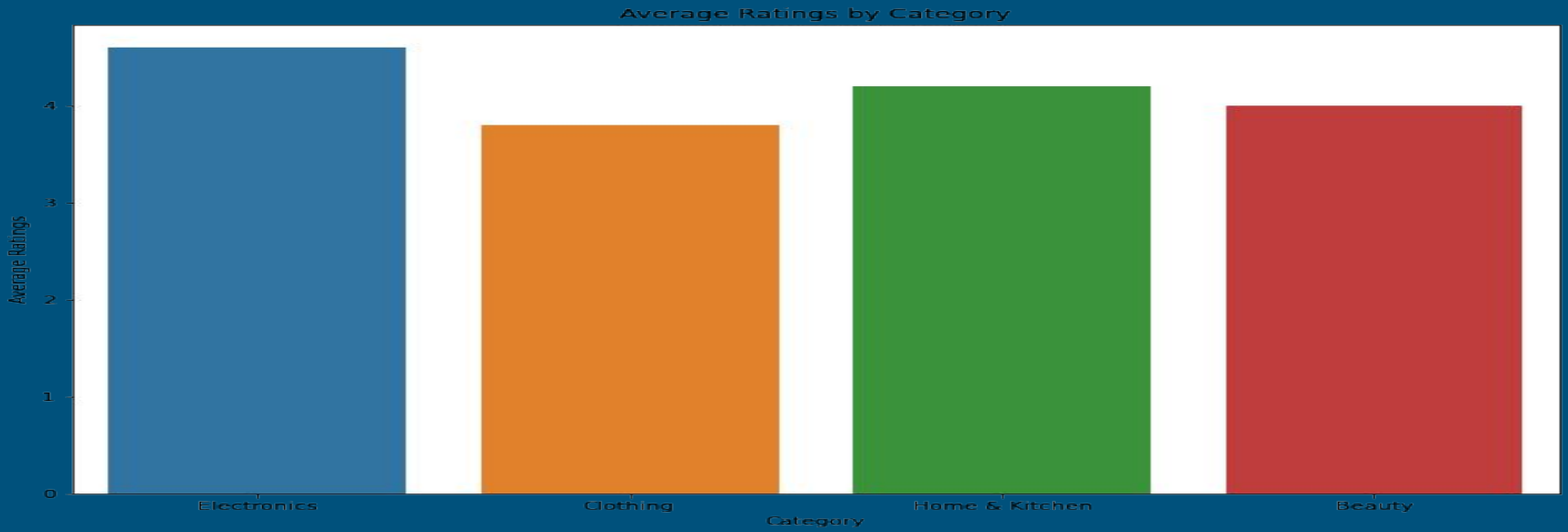
Based on the barplot above we can see that the most seen merchants category were home & kitchen followed by electronics, clothing, and beauty

Data Visualization



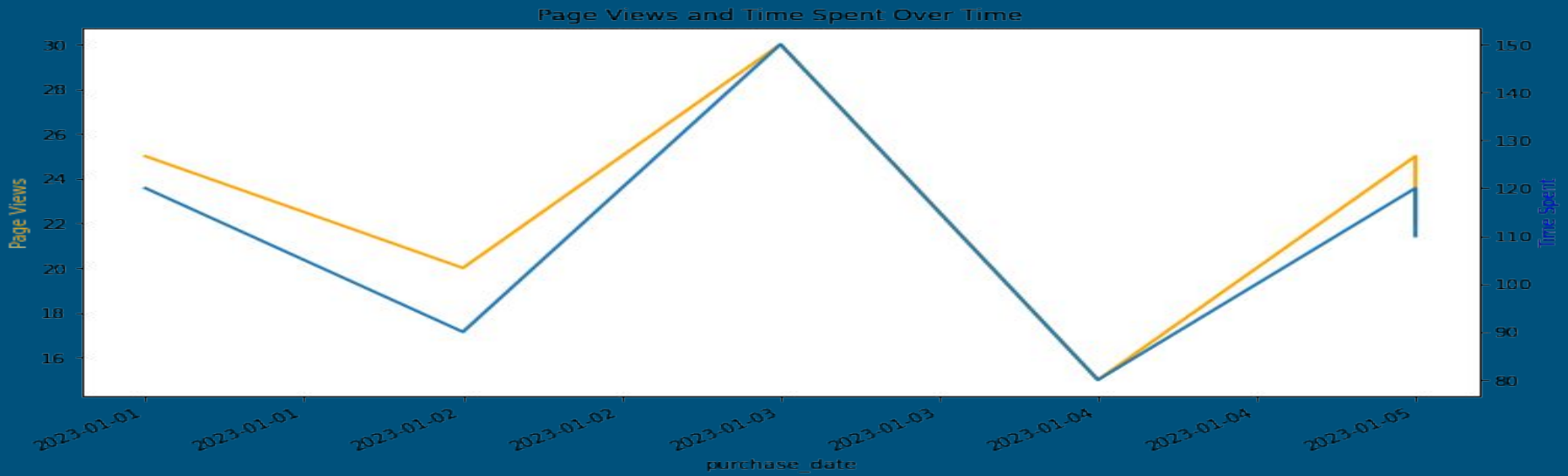
The barplot above show something similar for what we have already seen on views ranks by category where electronics have the higher price level followed by home & kitchen category, clothing and beauty.

Data Visualization



Based on the barplot above we can take the information that electronics got the highest ratings followed by home & kitchen, beauty, and clothing. There an insight here where in two previous barplot visualisation of time views and time spent based on category the ranks' order were home & kitchen, electronics, clothing, and beauty. When its plotted for average price based on category, beauty still in the last position meanwhile for the rating result they don't get least ratings score.

Data Visualization



Based on the chart above it can be seen that page views and time spent variable have collinearity.

Linear Regression Model

Linear regression is a statistical method used for modeling the relationship between a dependent variable and one or more independent variables. It assumes that this relationship can be expressed as a linear equation. In simpler terms, linear regression helps us understand how changes in one variable are associated with changes in another.

Model Evaluation

```
# Evaluate the model
mse = mean_squared_error(y_test, predictions)
print(f'Mean Squared Error: {mse}')
```

✓ 0.0s

```
Mean Squared Error: 0.0051852514049506625
```

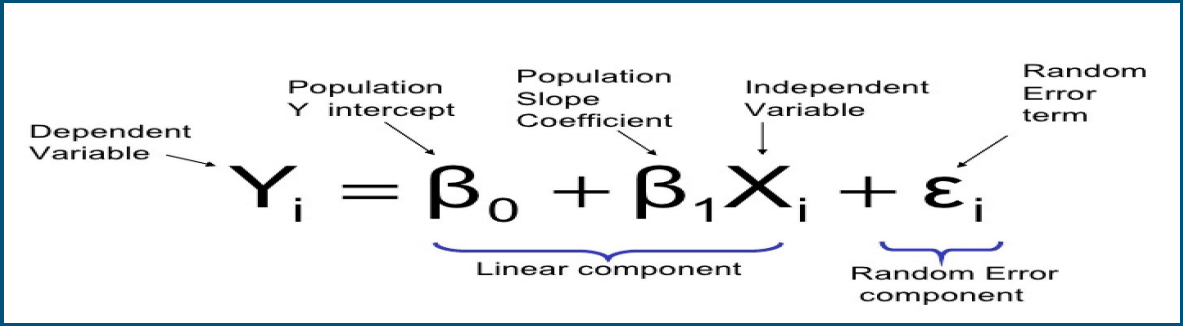
Mean Squared Error (MSE) is a commonly used metric in statistics and machine learning to measure the average squared difference between the predicted values and the actual values. It is a way to quantify the accuracy of a predictive model.

Conclusions

- There an insight here where in two previous barplot visualisation of time views and time spent based on category the ranks' order were home & kitchen, electronics, clothing, and beauty. When its plotted for average price based on category, beauty still in the last position meanwhile for the rating result they don't get least ratings score.

- The input data from the model do not provide ratings that are below 3.8, it would affect the model to give the ratings for 3.8 for the lowest ratings. More input data for next research would make the model predict better ratings score.

Appendix


Dependent Variable
Population Y intercept
Population Slope Coefficient
Independent Variable
Random Error term
$Y_i = \beta_0 + \beta_1 X_i + \varepsilon_i$
Linear component
Random Error component

Model Demo


free
tote
bag
mockups-design.com
Rating Prediction App
Enter Page Views:
21
Enter Price:
500.00
Predict Ratings
Predicted Ratings: 4.51