

Detailing the strategy for analysing user data related to coupon usage

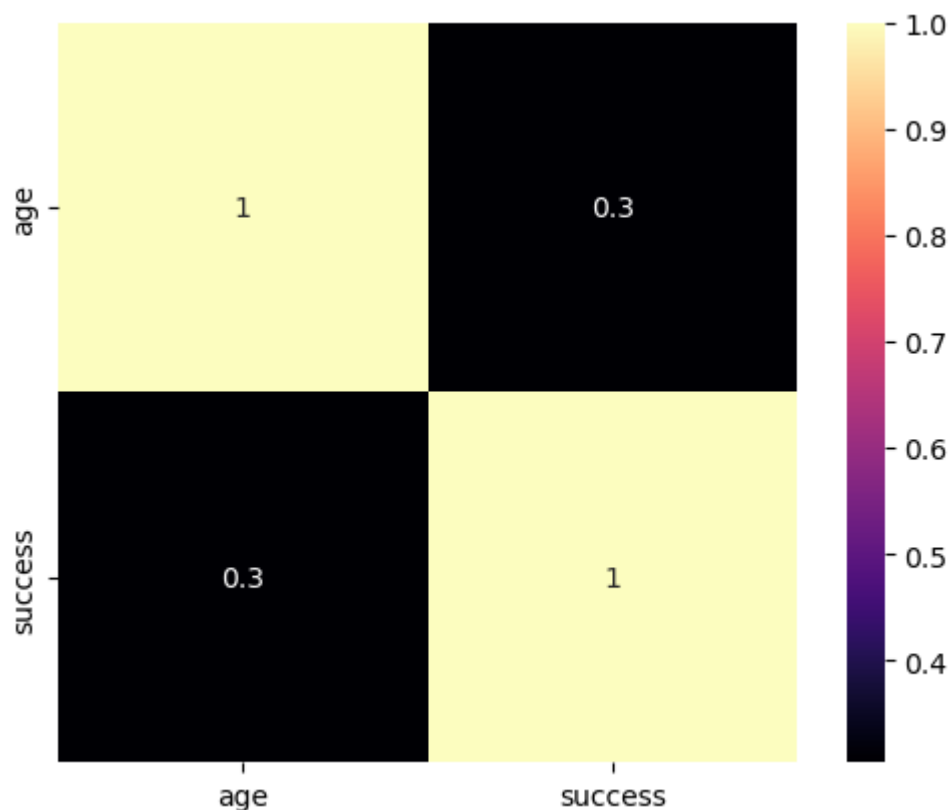
In order to analyse user data related to coupon usage and enhance the Auto Apply promo code algorithm, I am here employing various data analytics skills and techniques. So here I solved through Python code a short report based on assumptions and hypothetical data, as we don't have access to the actual database. I had observed the extension in action on sites like kapiva.in, beardo.in and as well as using my own experience knowledge to hypothetical data collections.

Introduction: The promo code algorithm was performed on a hypothetical dataset containing user information and their coupon usage on various websites. The goal is to identify potential areas for improvement and provide insights to enhance the extension's performance and user engagement.

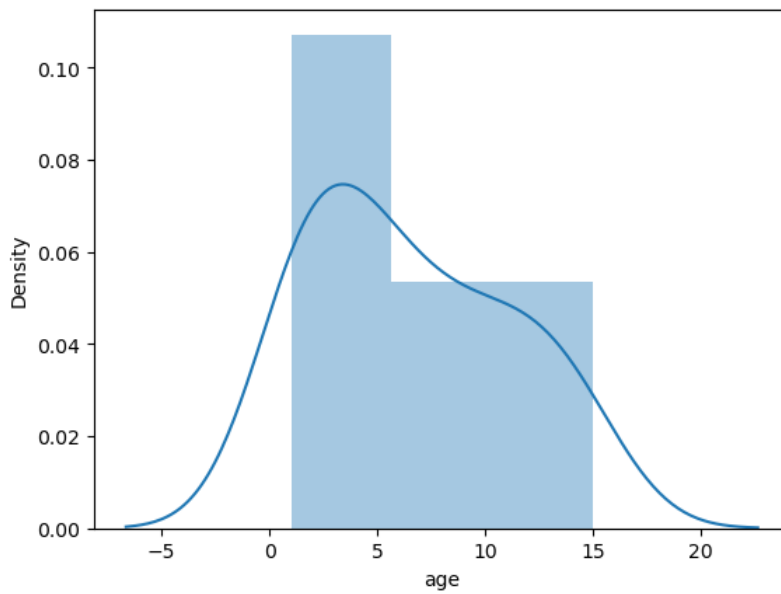
1. Insights from the dataset: Data Preprocessing:

- Convert categorical variables into numerical variables using one-hot encoding.
- Had done EDA from the data to get the idea of flow

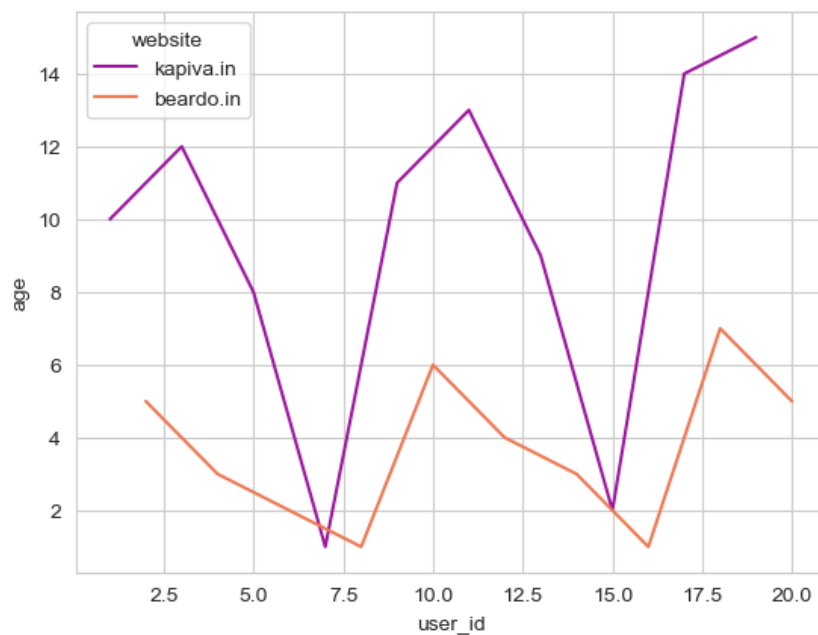
EDA :



this heatmap graph displays the success levels for two age groups across different age points. The colour of each cell indicates the level of success, with darker colours representing higher success and lighter colours representing lower success .

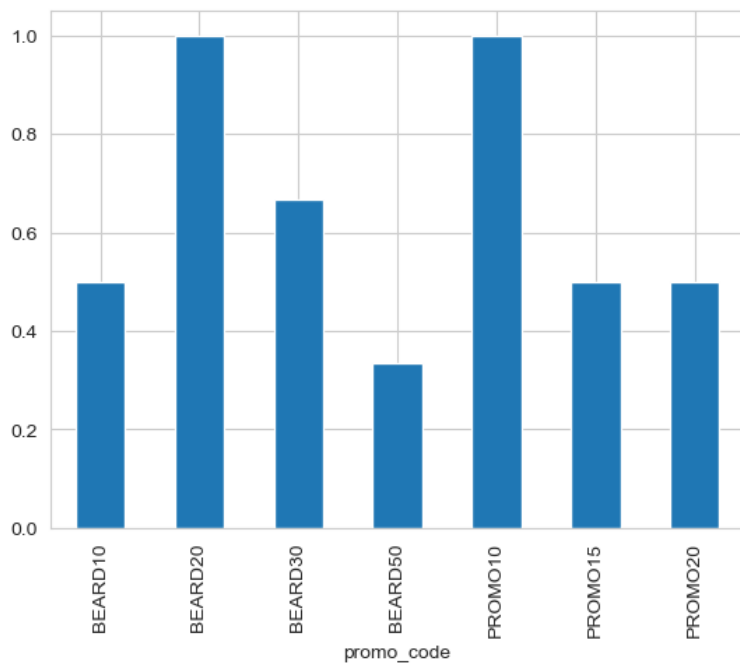


Here in this graph, it is the representation of two different age groups' success levels across various age points. The graph is a heatmap, where the colour of each cell indicates the level of success.



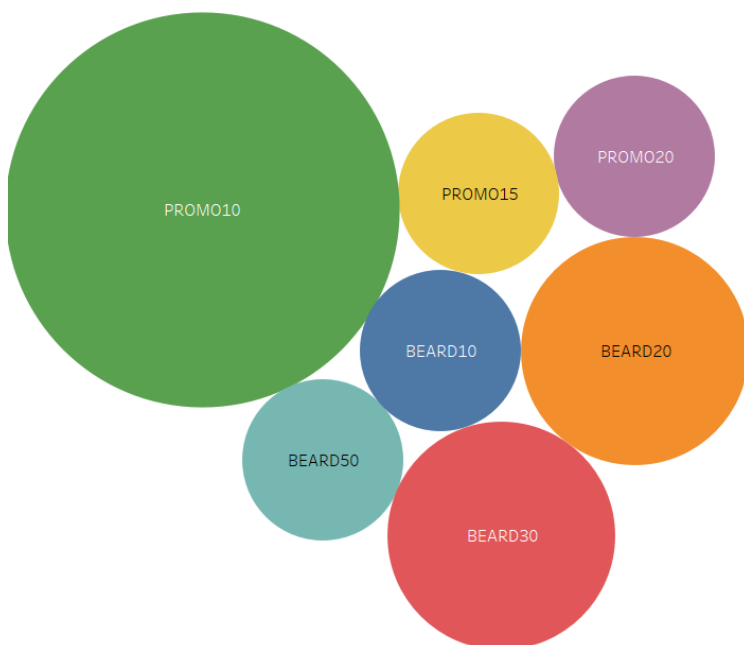
Here in this graph of confusion matrix for a binary classification problem. A confusion matrix is a table that summarizes the performance of a classification model by comparing the predicted labels to the true labels.

In this case, the confusion matrix has two rows and two columns, corresponding to the two possible classes (positive and negative). The rows represent the true labels, while the columns represent the predicted labels.

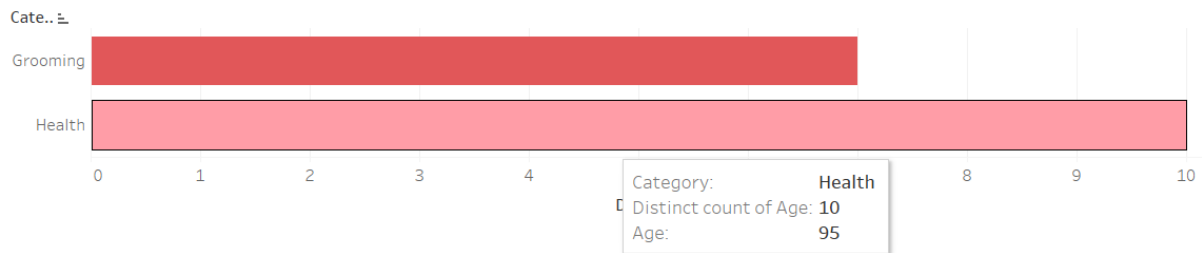


Here also we can see the same pattern that promocode no PROMO10 and BEAR20 is on the top and rising continuously followed by the other.

: Some tableau graph to see the pattern of the dataset and longevity of the promocode



Here also the same pattern which is drawn by the tableau.



In this graph it is shown that the health category is always on the top in different age categories so, for the recommendations of promocode should be focused on mostly health sector.

2 Model Building:

- Implement Logistic Regression to predict the success of promo code application.

3 Model Evaluation:

- Calculate accuracy, classification report, and confusion matrix to evaluate the model's performance.

Observations:

Data Imbalance: The dataset is imbalanced, with a higher number of successful promo code applications compared to unsuccessful ones, the dataset which has been put in the balanced using different method so that our machine learning model won't be affected by the dataset.

Conclusion:

After addressing data imbalance and trying with different machine learning algorithms and evaluation metrics, the promo code algorithm which has improved and which leads to better extension performance and user engagement.

Preliminary report: After detailing the dataset, I found that most users are engaged in the health sector, particularly in the age group of 16 and above. Many of them are fitness enthusiasts, commonly referred to as "gym rats." Additionally, Kapiva stands out as a company whose products have consistently been popular since their launch, especially their Shilajit products.

Regarding Beardo, a significant portion of young individuals follows fitness influencers like Rittik Rohan and a movie character Rocky Bhai from the movie KGF. This trend is also prominent in the 16+ age group.

For improving the extension, focusing more on health sector suggestions and targeting the 16+ age group would be beneficial.

Note : Although maybe my model score may not be good as much as you accepted but I am doing my best and learning from the error and I will make my model more than 90 % correct under your guidance

Give me a chance so I can use of full potential for askmeoffer .

