**APPAREL VALUATOR**

**A MINI PROJECT REPORT**

*Submitted by*

**Lakshana G**

**CB.SC.I5DAS20018**

***In partial fulfilment of the requirements for the award of the degree of***

**Integrated Master of Science**

**in**

**Data Science**

******

**Amrita School of Physical Sciences**

**Department of Mathematics**

**Amrita Vishwa Vidyapeetham**

**Coimbatore – 641112**

**May 2024**

**DEPARTMENT OF MATHEMATICS**

**AMRITA SCHOOL OF PHYSICAL SCIENCES,**

Coimbatore - 641112

******

**BONAFIDE CERTIFICATE**

This is to certify that the project report entitled “**Apparel Valuator**” submitted by **Lakshana G (CB.SC.I5DAS20018)** in partial fulfillment of the requirements for the award of the degree of **Integrated Master of Science** in **Data Science** is a bonafide record of the work carried out under my guidance and supervision at Amrita School of Physical Sciences, Amrita Vishwa Vidyapeetham, Coimbatore

Project Coordinator

Designation

Chairperson

Department of Mathematics

Dr. K. Somasundaram

The project was evaluated by us on:

Internal Examiner Internal Examiner

**DEPARTMENT OF MATHEMATICS**

**AMRITA SCHOOL OF PHYSICAL SCIENCES,**

Coimbatore - 641112

******

# DECLARATION

I, **Lakshana G(CB.SC.I5DAS20018)**, hereby declare that the Mini Project entitled “**Apparel Valuator**”, is the record of the original work done by me. To the best of knowledge, this work has not formed the basis of the award of any degree/ diploma/ associateship/ fellowship/ or a similar award to any candidate in any university.

**Date:** 03-05-2024

**Place:** Coimbatore

**Signature of the Student**

COUNTERSIGNED

**Dr. J. Geetha**

Project Coordinator,

Department of Mathematics,

School of Physical Sciences,

Amrita Vishwa Vidyapeetham,

Coimbatore-641112, Tamil Nadu, India

**ACKNOWLEDGMENTS**

I would like to express my special thanks **to Dr. K. Somasundaram**, Chairman, Department of Mathematics, Amrita School of Physical Sciences, for giving me a golden opportunity to carry out this wonderful project with the project titled “**Apparel Valuator**”. This project has allowed me to learn, discover and experiment a lot of new things.

I would also like to thank my class advisor, **Dr J.** **Geetha**, Department of Mathematics, Amrita School of Physical Sciences, for supporting me with this project idea and guiding me throughout the completion of the project.

I would like to thank the teachers in the Department of Mathematics for encouraging me to do this project.

I sincerely thank all the staff for their invaluable teachings over the years and the constant inspiration for my project work.

I am especially indebted to my beloved parents and family for their blessings and inspiration.

**Place:** Coimbatore

**Date:** 03-05-2024 LAKSHANA G

**Contents**

1. **Certificate**………………………………………………………………. **ii**
2. **Declaration**………………………………………………………………**iii**
3. **Acknowledgement**……………………………………………………….**iv**
4. **Abstract**………………………………………………………………….1
5. **Introduction**……………………………………………………………..**2**
6. **Implementation** i. Preparing custom Dataset………….………………………………..…3

ii. Cleaning and Preprocessing…………………………………………. 4

iii. Exploratory Data analysis………………………………………........5

iv. Review Sentiment Analysis...……………………………………….21

v. Building and Evaluating ML Models…………………………….….23

vi. Fine-Tuning the Model……………………………………………...26

vii. Interactive User-friendly Website…………………………………..27

1. **Conclusion**……………………………………………………………...**31**
2. **Reference**……………………………………………………………….

**Abstract**

In the online marketplace, buying clothes from the web has become a major trend because it has several advantages for purchasers and allows vendors to use profitable prospects even if they are located in distant places. My project intends to assist sellers in pricing their men’s and women’s clothing effectively. Furthermore, I have designed a simple user-friendly website that can be used by sellers with different educational backgrounds and levels of proficiency in English language skills. This tool allows merchants to create accounts easily, change them whenever necessary as well as get useful standardized prices for their garments thus giving them power to do good marketing which will lead into higher revenues.

The model that has been implemented is trained on a custom dataset and involves fine-tuning various parameters of the regression models. Multiple models have been implemented and assessed, with careful consideration given to the profound impact these tools can have on the livelihoods of sellers. Furthermore, my project incorporates diverse visualizations aimed at empowering both customers and sellers to discern market trends, understand customer behaviours, and identify preferences within the dynamic landscape of online shopping.

After close examination, I found out that this model not only meets but exceeds performance standards hence it can give correct estimates about how much people should pay for clothes accurately. Moreover, Python Flask framework was used during the development of our website where SQL Alchemy also played an important role since most parts were made using HTML alongside CSS so as to make it look better while at the same time working better too.

In addition, the website was meticulously designed with the Python Flask framework, SQL Alchemy, making extensive use of HTML and CSS to improve its functionality and aesthetic appeal. These technologies significantly improved the entire user experience and guaranteed the efficacy of the product by enabling the creation of a smooth and intuitive interface.

**5. Introduction**

There are many online platforms and businesses that cater to various fashion interests in the men’s and women’s market which is why buyers often find it difficult to deal with price disparities on the internet nowadays. Finding the ideal balance between value perception and affordability becomes crucial for vendors who want to satisfy customers and maximise revenues. In order to take on this challenge head-on, I have developed an advanced pricing strategy designed especially for the ever-changing apparel market. ‘’

The "Price Valuator" project aims to revolutionize the process of apparel pricing by leveraging machine learning algorithms, specifically regression models, and advanced natural language processing techniques. By harnessing data on various attributes such as brand, features, specifications, type, material, and available sizes, the project predicts optimal prices for clothing items.

Using advanced regression models and language models together, accurate price estimates are generated by this system hence giving sellers the power to make informed choices. For this purpose, a user-friendly website has been created so as to enhance accessibility and usability where sellers can input clothing details and get personalized price recommendations designed around what they have for sale.

Different models have been tried out and evaluated, being aware that these instruments can strongly affect the sellers’ lives. Also, in my project there are various visual representations which are meant to give power to both buyers and sellers so that they may notice market shifts, interpret client actions as well as recognize changing tastes across e-commerce.

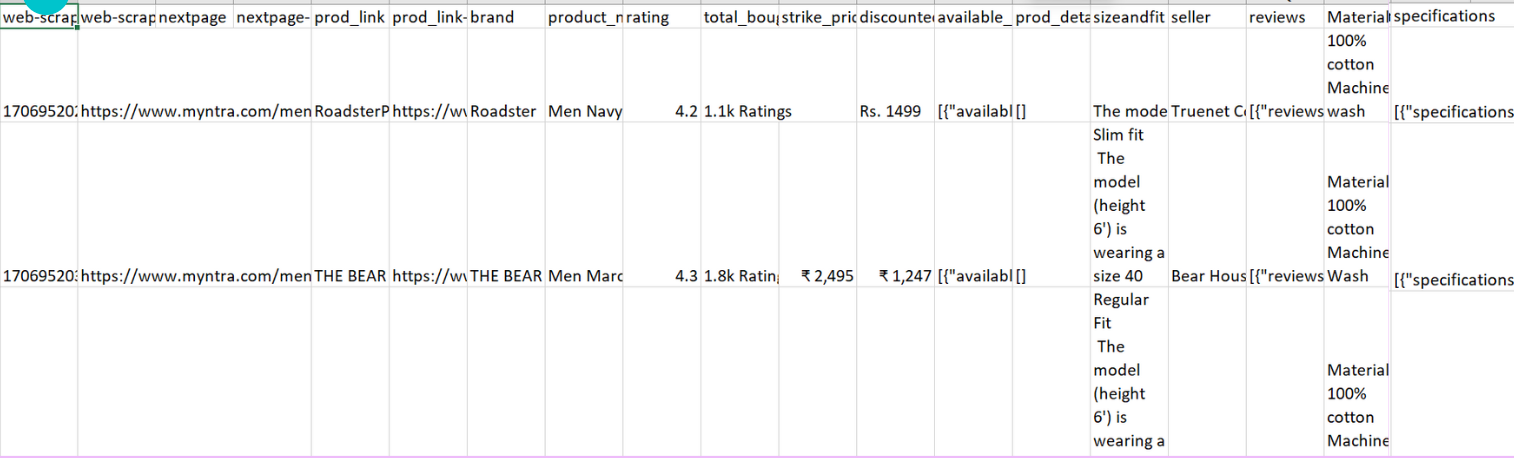
**6. Implementation**

1. **Preparing custom Dataset**

To acquire my dataset, I used Web Scraper extension, a free Chrome extension that is used to scrape data in JSON format. When correctly changed, it scrapes every piece of a website, given the elements that need to be scraped, i.e the items that we are interested in from the website. Myntra is one of India's most prominent clothes e-commerce platforms, and I scraped data from it.

For menswear, I scraped information from pages that featured t-shirts, jackets, jeans, blazers, casual shirts, formal shirts, sweatshirts, sweaters, kurtas, pyjamas, sherwanis, suits, formal suits, track pants and casual trousers among others. For ladies’ wear on the other hand, my focus was on dresses, blazers, lehenga, cholis, jumpsuits, jeans, sarees, jackets, skirts, sweatshirts, ethnic wear, leggings, shrugs, ethnic tops, tunics and more.  
  
On the whole my dataset had 13,869 rows of data where each row represented a single unique item of clothing. The set has two subsets: men’s clothing which accounted for 6,978 records and women’s clothing comprising 6,889 records. In these records there are seventeen columns in total that were extracted with some attributes being product link, brand name, rating, total bought, strike price, discounted price, available sizes, product details, size & fit, seller reviews, material care, specifications, type etcetera.

The reviews column is one of our dataset's unique features. Despite the abundance of datasets on the internet, none of them have a reviews column. An extra element that provides several significant details concerning the clothing goods is the review section.



1. **Data Cleaning and Preprocessing**

In order to make the dataset ready for analysis, several preprocessing steps were performed on the raw data. Such steps include:

**Extraction and Combination**: A lot of entries in this dataset were presented in unstructured formats such as lists or dictionaries which contained clothing details. These values were extracted from their original locations and merged into well-formatted columns that could be easily analysed.

**Dealing with links**: Links are usually found in some cells mostly representing features of a product like product discount, product details. Therefore, this part deals with how these URLs were handled since they would not add any value to data analysis; unless further processed to extract useful information such as discount percentage for instance while preparing the dataset.

**Character Removal**: This step involves removing special characters commonly found within textual input, including currency symbols like the Rupee sign and emoticons present in review columns. The main purpose behind this action is to enable machines interpret numbers accurately when fed with numerical text or any other form of alphanumeric strings.

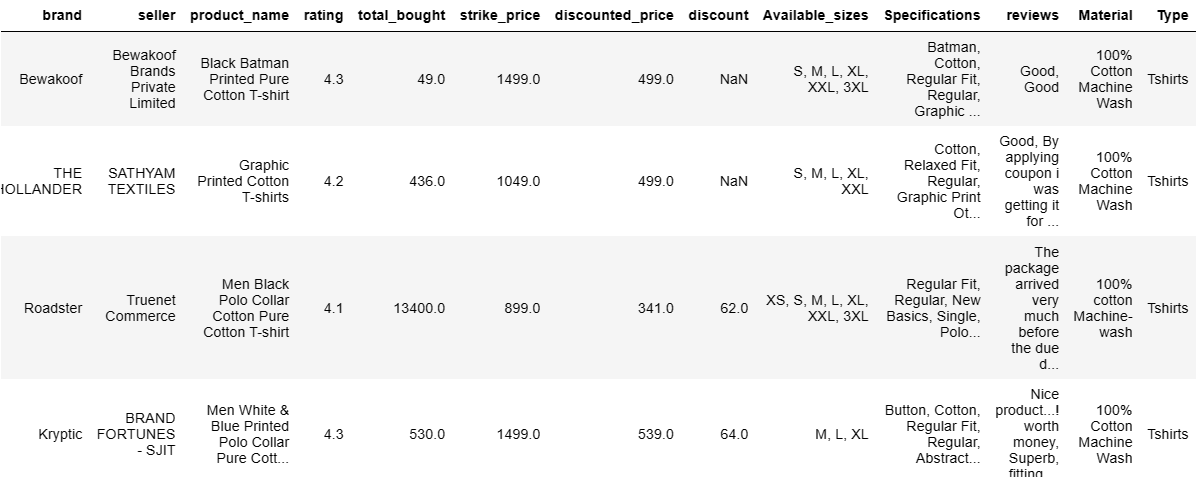
**Handling Null Values**: There were missing records identified under strike price column; ratings column; discount column and lastly reviews column. To address this issue, we imputed those empty fields by substituting them with appropriate placeholders numbers or words based on whether it was categorical variable or numeric variable respectively. For numerical variables mean was used as an alternative while categorical mode served better imputation method.

**Column Merging**: There were overlapping details in columns like product information and specifications for the same piece of clothing. In order to avoid duplication and make it simpler, these columns were combined which allowed for compact storage and analysis of important information.

**TF-IDF** or term frequency-inverse document frequency is a statistical measure used to evaluate the relevance of a word in a document compared to an entire corpus. TF-IDF establishes this by examining word frequency in a document (TF) and scaling it according to the frequency of that same word across all documents within said corpus (IDF). In machine learning, we will employ a TF-IDF vectorizer to turn words into numbers for use as inputs into machine-learning algorithms at some later stage. At some point after this, we’re going to have our project take those product details and run them through the Machine Learning Algorithm but first things first let’s just get them vectorized using TF-IDF Vectorizer in order for them not only be able pass through but also make sense once they do so.

The dataset was transformed into a structured and standardized form that is suitable for deep analysis and modelling aimed at unearthing valuable pricing patterns as well as consumer tastes within the fashion industry. This preparation process involved many steps.

Dataset after cleaning:



1. **Exploratory Data Analysis**

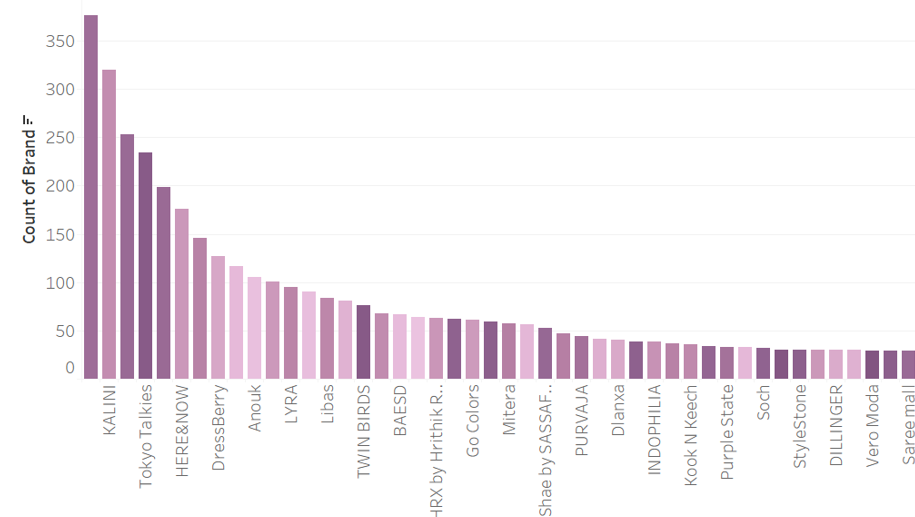
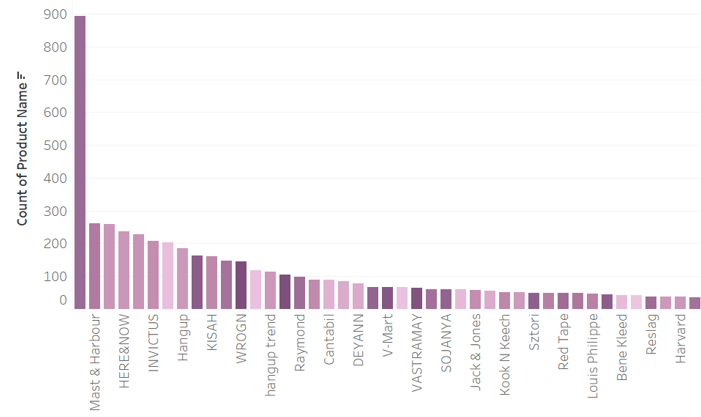
Exploratory data analysis (EDA) is an important first step in the data analysis process, involving the processes of summarizing, visualizing, and understanding the underlying patterns and trends in the data. By analysing data from multiple perspectives and using descriptive statistics, data visualization, and statistical techniques, EDA helps to make informed decisions.

Different graphic techniques have been used to compare important features and trends in men’s and women’s clothing. Bar plots, bubble plots, 3D visualizations, and word clouds are used to visualize the different features in the dataset. These images provide valuable insight into the distributions, relationships and patterns in the data, and help to understand the unique characteristics of men’s and women’s clothing

Tableau was selected as the primary tool for data visualization due to certain limitations encountered with Python when handling large datasets for visualization purposes. Tableau offers intuitive data visualization, facilitating easy exploration and analysis. With a user-friendly interface, it empowers users to create interactive dashboards, uncover insights, and make informed decisions, promoting data-driven decision-making.

1. **Discover the Brand with the Most Items**

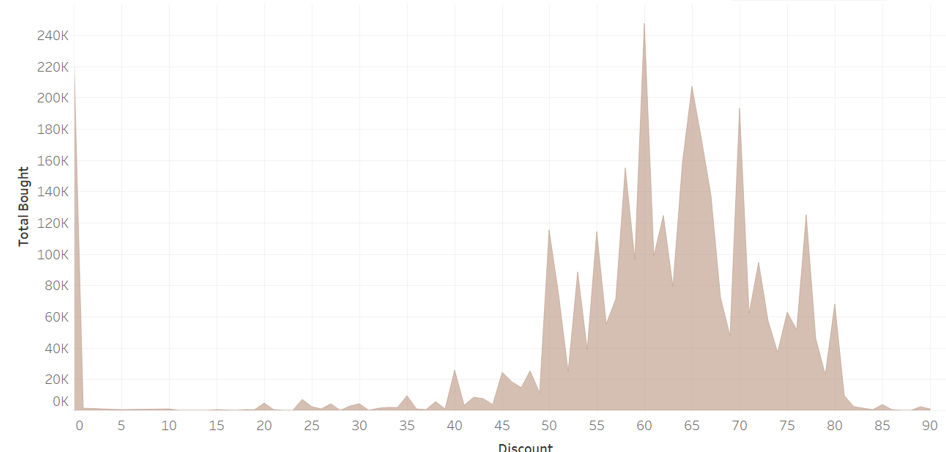
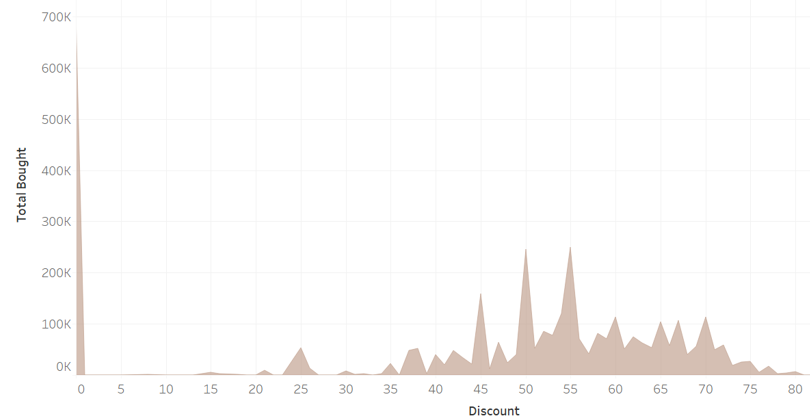
**Women Men**

**** ****

Kalini and Roadster are the bands with the most products featured in the women's section according to the samples taken from the Myntra website. The brands with the greatest number of apparel items mentioned in the men's section are Roadster, Mast and Harbour.

1. **The most well-liked and well-received discount among customers**

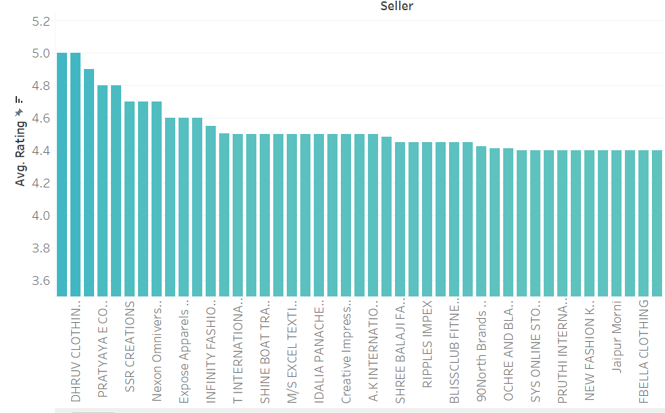
**Women Men**

** **

247287 female buyers purchased apparel at a 60% discount. and 206,000 female customers purchased during the 65% discount. This indicates that women are more concerned with the dress requirements than the associated savings. The well-received reductions for guys were 50% and 55%.

1. **Seller’s average ratings**

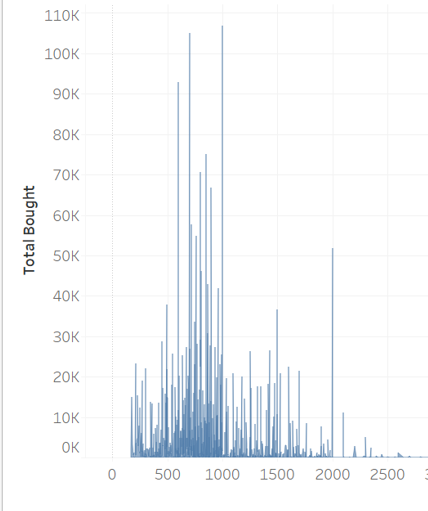
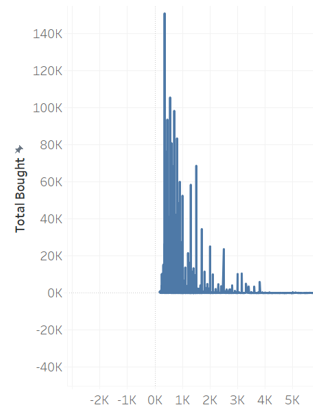
**Women Men**

** **

In the statistics for both men and women, a greater proportion of brands have an average rating higher than 3.5. Pyxis Brand Technologies Private limited has received the highest possible rating of five from men and women for Swasthik wear, Marks and Spenser.

1. **Most purchased price range**

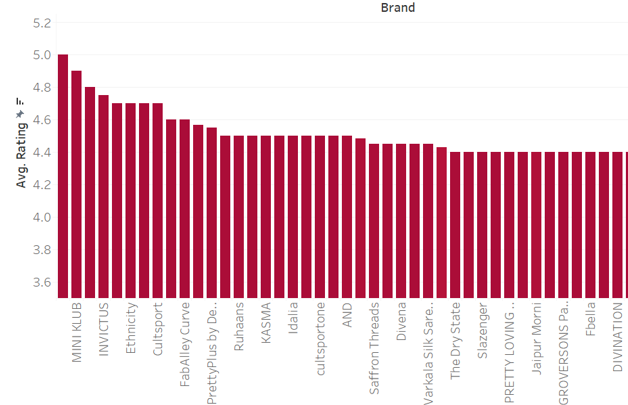
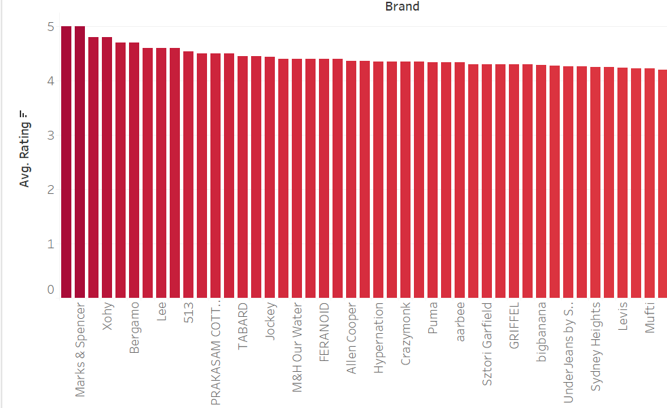
**Women Men**

****  ****

The data makes it quite evident that women spend more money on clothes than men do. According to the plan, the most popular prices for women are for Rs. 999, Rs. 699, and Rs. 500, while the most popular prices for males are worth Rs. 349, Rs. 549.

1. **Brand’s Average Rating**

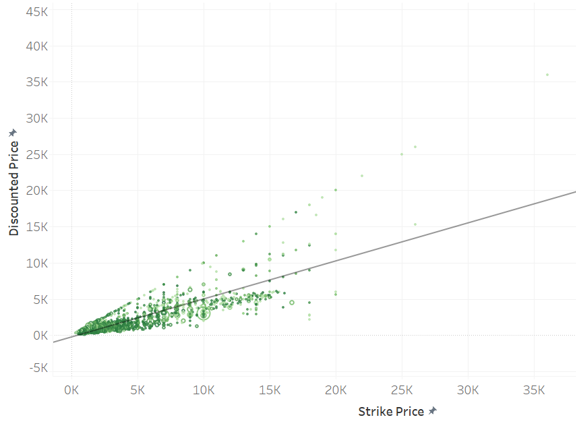
**Women Men**

****  ****

The two brands with the highest ratings, Karagini and Mini Klub, are 4.9 and 5.0, respectively. These Indian companies offer women's ethnic clothing. For men, the brands with five stars are Spenser, Marks, and Never Loose. Marks and Spencer is a well-liked brand and a men’s top pick.

1. **Relationship between discount price and strike price**

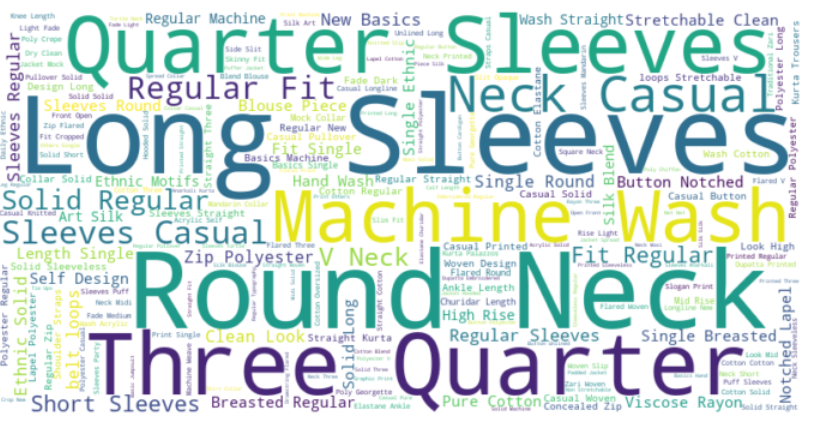
**Women Men**

** **

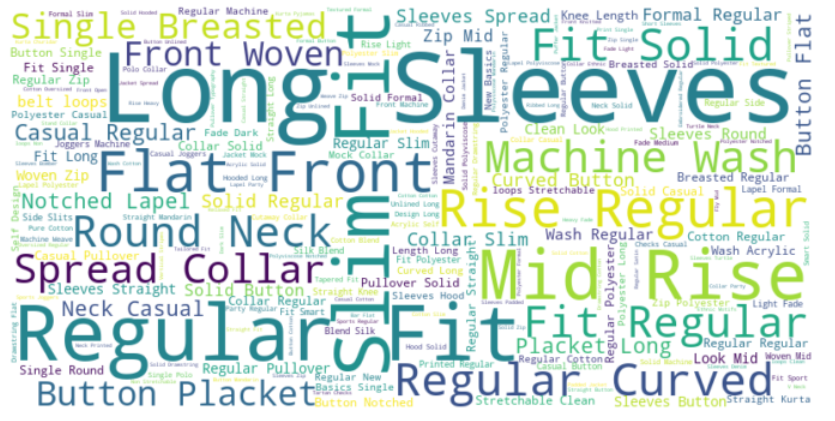
The two graphs demonstrate a linear relationship between the discounted price and the strike price. However, when we delve more into the plot, we can observe that the data points in the male and female datasets are narrower and more divided, respectively. Women's apparel is discounted even at a higher striking price, whereas men's clothing products are reduced at a price far less than the product's strike price.

1. **Word cloud of women and men specification**

**Women**



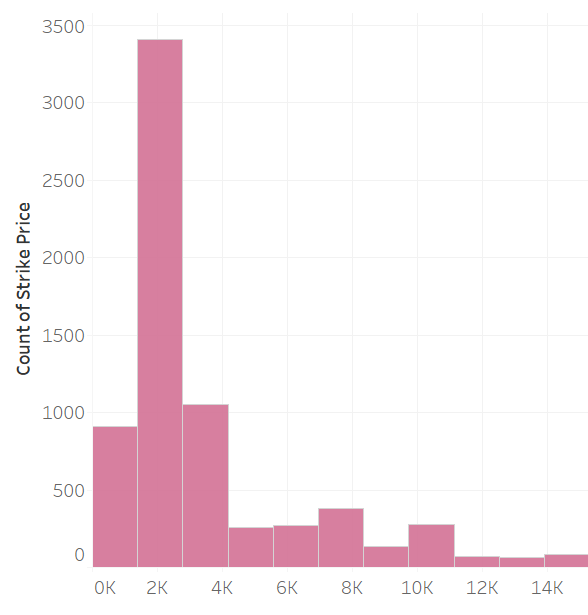
**Men**



More frequently occurring words appear larger in a word cloud, which graphically displays the frequency of terms in a text. It's helpful for rapidly determining which terms appear most frequently in a dataset and for revealing important themes or subjects.  
The most common criteria for women's clothing are long sleeves, round necks, and three-quarter sleeves. While the most common clothing requirements for males are Regular, Slim Fit, Long Sleeves, and Woven.

1. **Distribution of Strike price**

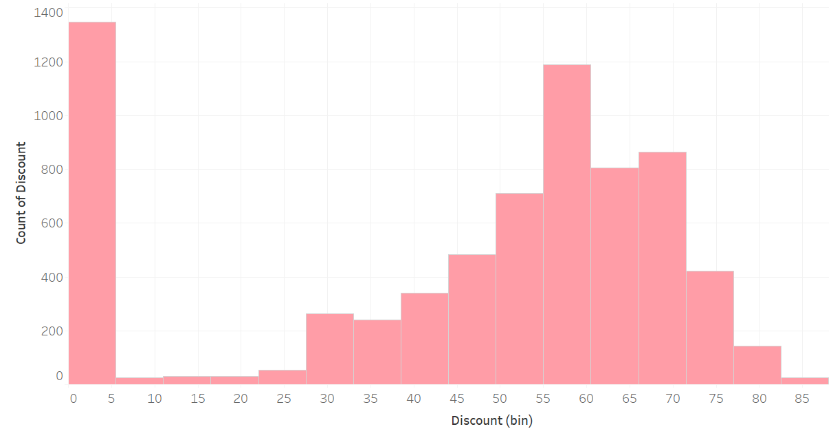
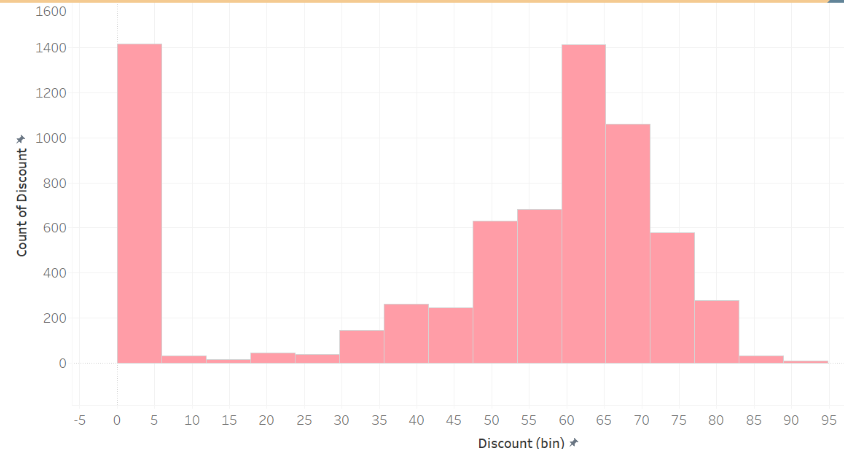
**Women Men**

** **

Clothes for women are more expensive than clothes for men. The majority of women's clothes is priced between 1500 and 4000 Rupees. On the other hand, the majority of men's clothes costs between 2000 and 2500 Rupees.

1. **Popular discounts**

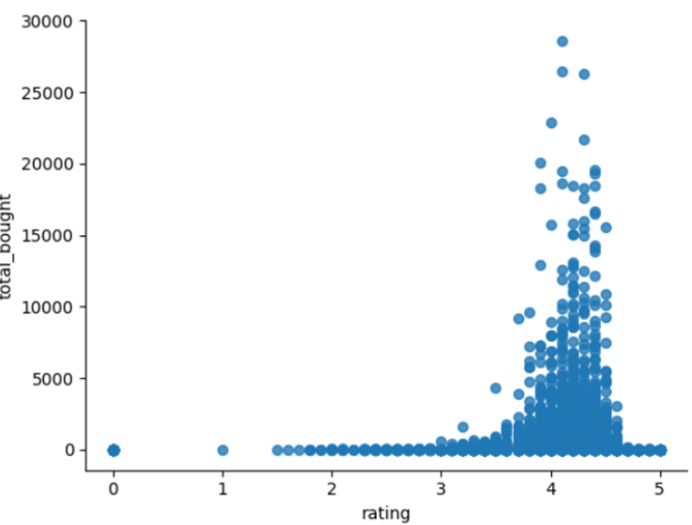
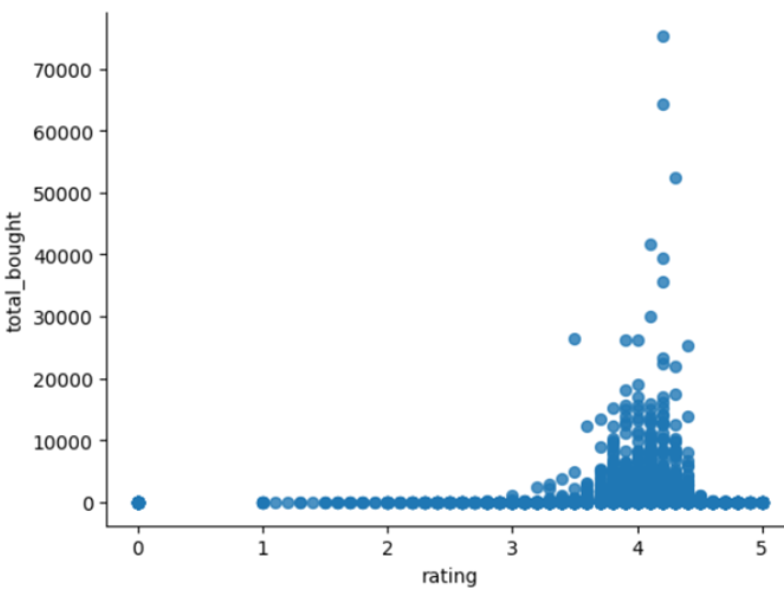
**Women Men**

****

As can be seen in the area chat above, the most common discounts offered are 60% and 65% for women's clothing and 60% and 65% for men's clothing. However, as the bar on the left makes evident, the majority of the apparel in our sample is not discounted.

1. **Rating vs Total bought**

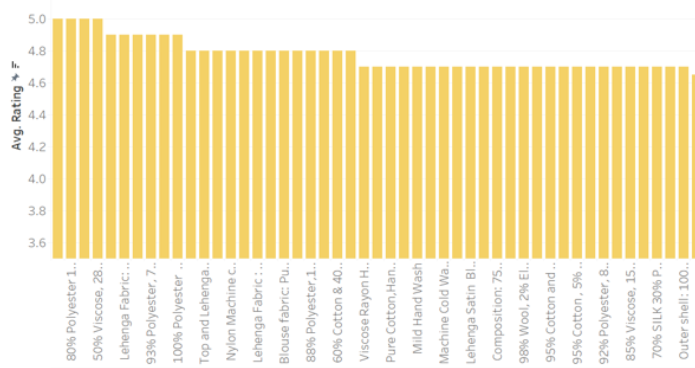
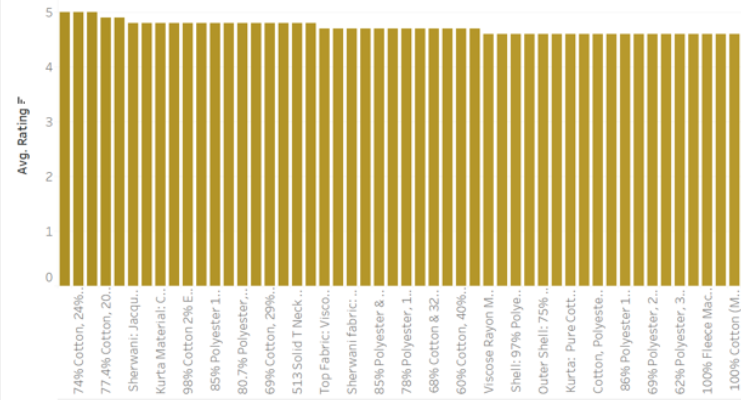
**Women Men**

** **

Five stars indicate that the product is outstanding in terms of both quality and appearance, while a rating of zero indicates that the product is not worthy enough to be purchased. Therefore, the number of people who buy certain things increases along with the positive reviews for those items. Positive product reviews encourage more people to purchase the item. Most women purchase clothing when there are more than 3 stars for a particular item. Men only buy apparel items that have ratings close to or above 4 stars

1. **Customers’ preferred type of material**

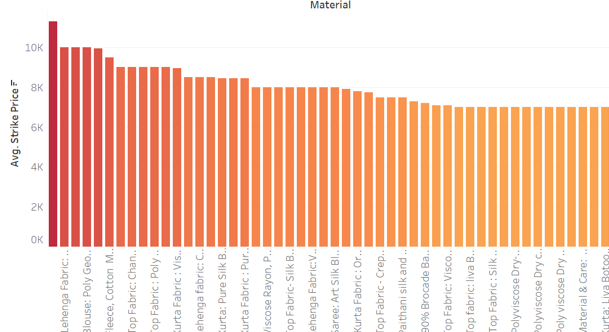
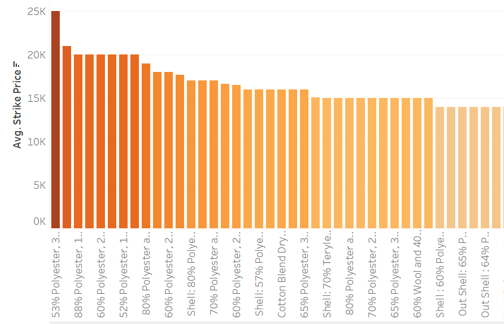
**Women Men**

** **

An apparel’s quality, longevity, and wearer's comfort are all largely determined by its material. It also affects things like maintenance, requirements and suitability for various seasons or weather conditions. Women prefer a blend of polyester and cotton because of its durability. Men prefer cotton and nylon the most.

1. **Which material is more expensive**

**Women Men**

**** ****

For an extended duration, the cost of clothing remains stable. The more costly types of materials in the women's division are silk, poly-georgette, and lehenga textiles. While spandex, wool, and polyester are more costly materials for men's apparel

1. **Tree Map**

**Women**

****

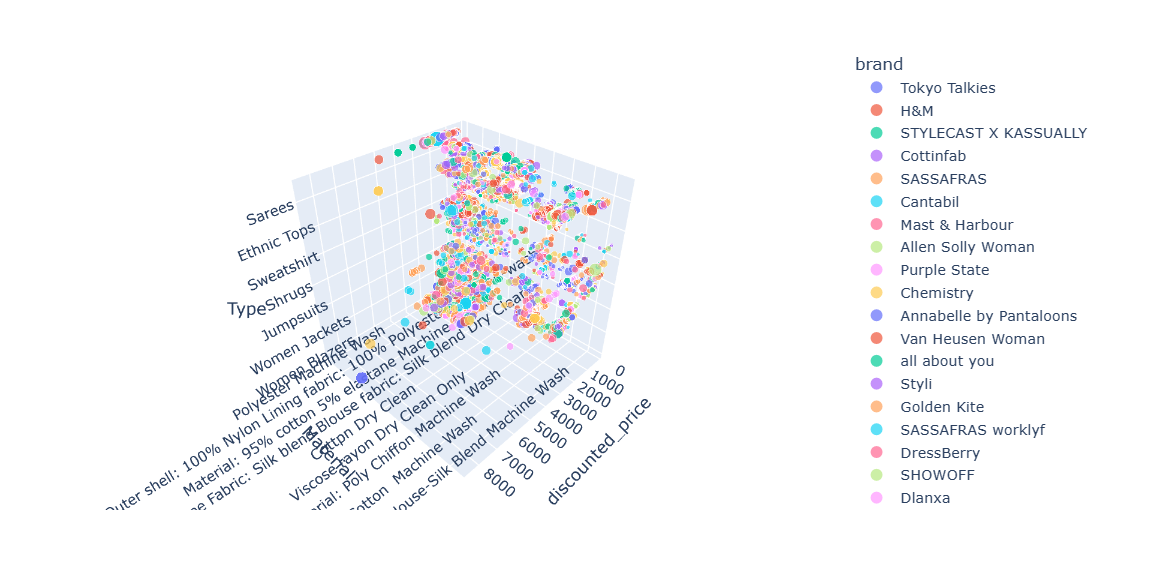
**Men**

****

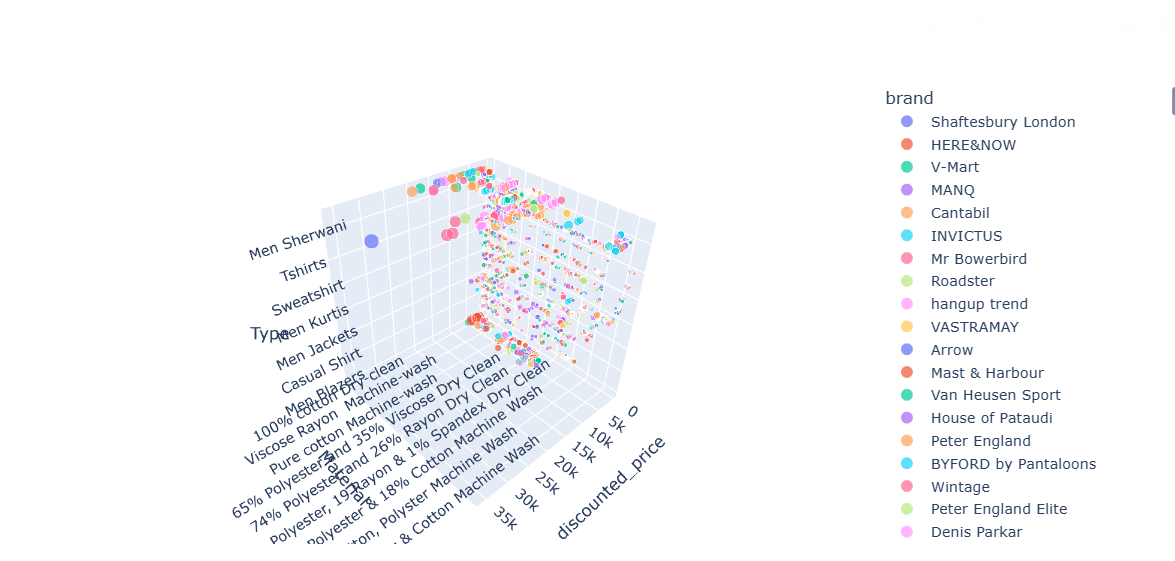
Using a series of rectangles, each rectangle's size is proportionate to a certain feature, like value or frequency, a tree map visualises hierarchical data. It is helpful in showing the data's hierarchical structure and making comparisons between various categories or tiers within the hierarchy easier. The retailers and the outfits they offer for both men and women are represented by this tree map.

1. **3D plots**

**Women**



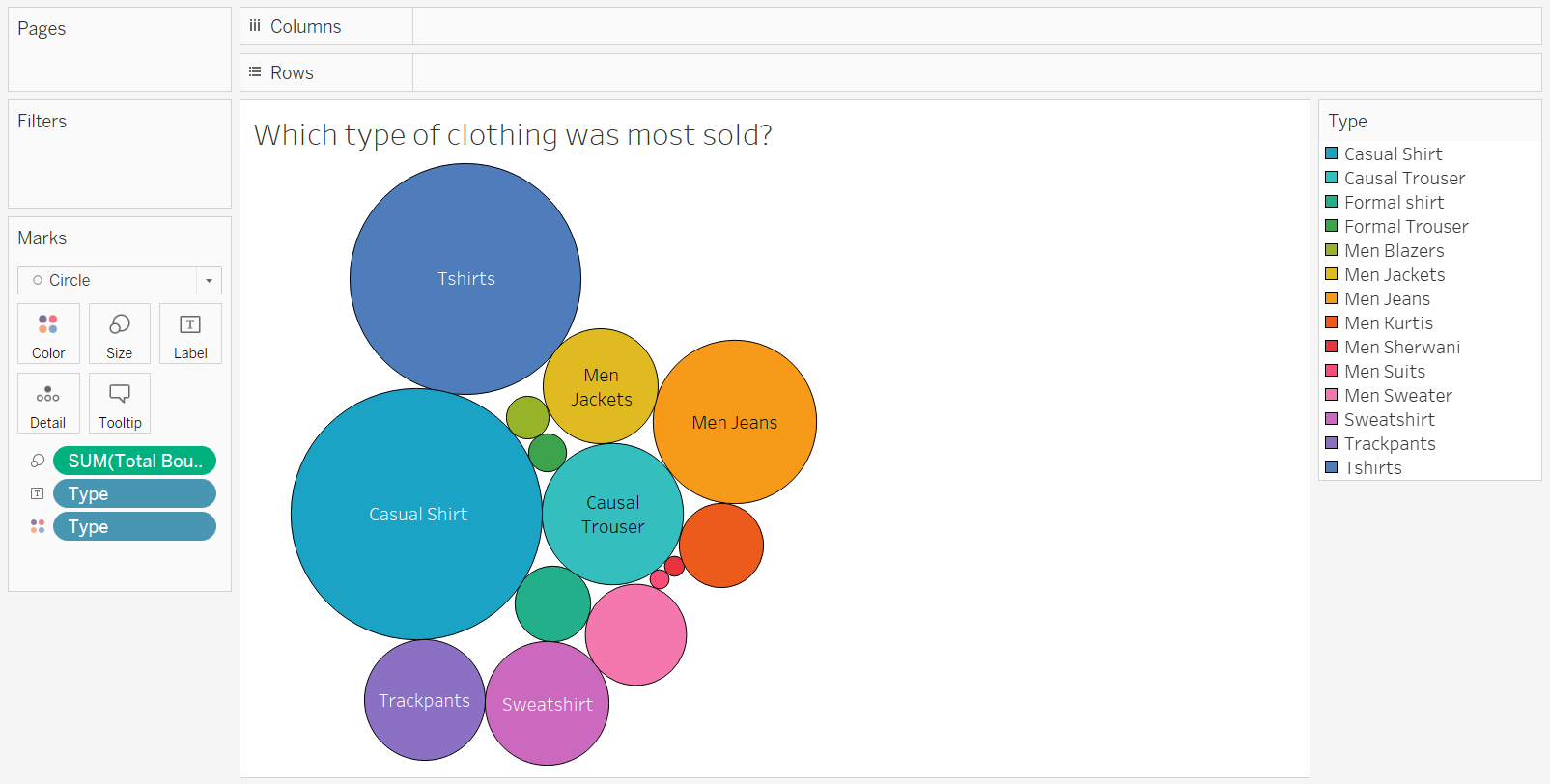
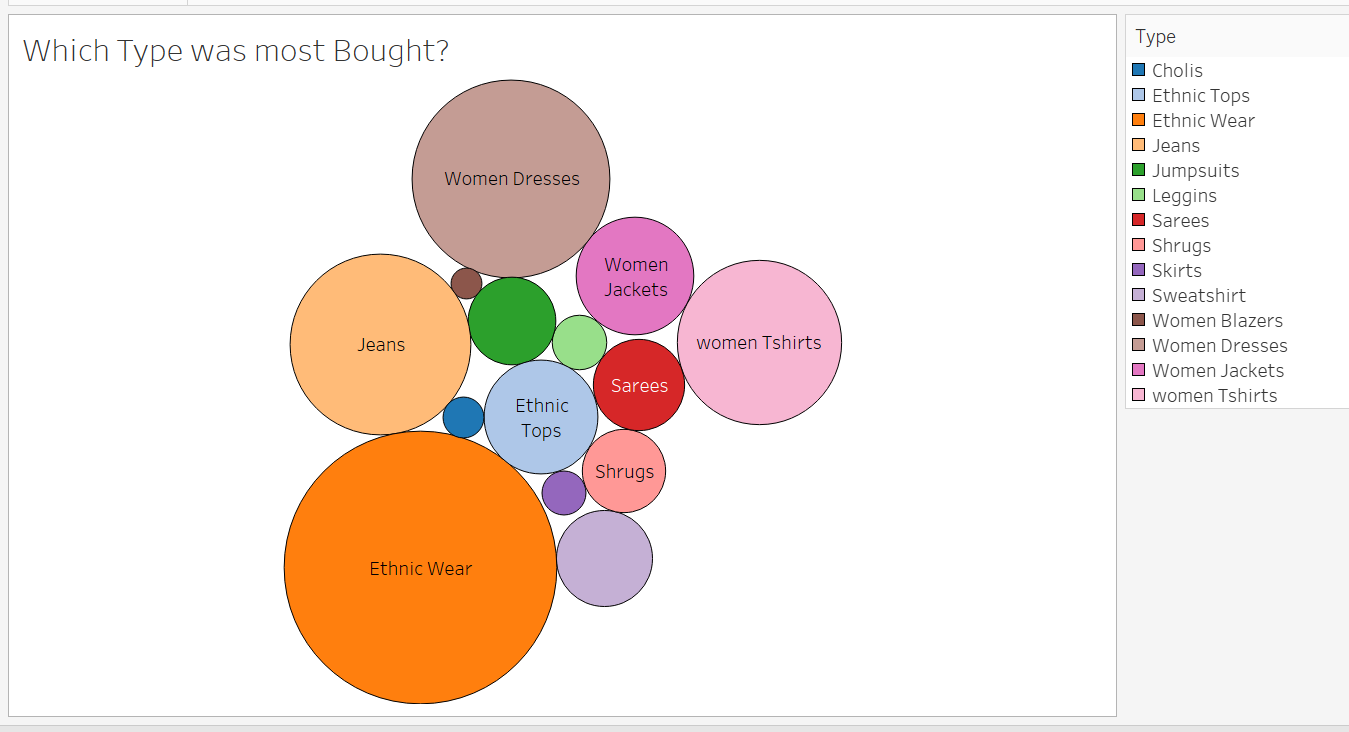
**Men**



3D plots are excellent for identifying clusters. This a three-dimensional plot that compares brand, outfit type, and discounted price. However, as we can see, the data points are more dispersed and do not support any conclusions

1. **Which type was most bought?**

**Women Men**



By summing up all of the consumers who have purchased, the bubble plot indicates the most popular form of clothing. Women purchase the majority of clothing, including dresses, jeans, and ethnic apparel. Men purchase more t-shirts, casual shirts, and jeans than women.

1. **Review Sentiment Analysis**

The review column is a significant feature of the dataset. Reviews that are favourable draw more buyers, and there's a great probability the product will be sold in the future. A greater number of neutral or negative comments reduces the likelihood that the client would purchase that item. Additionally, the seller's and the brand's reputation suffers, which hinders their growth and profitability.

After extraction, the review column has been cleaned. The value "No review" has been used in lieu of all the null values. There are several reviews for each entry in the dataset. Emojis have also been eliminated by employing the emoji library. Regular expressions have been used to remove all special characters.

**Usage of LLMs for sentiment analysis**

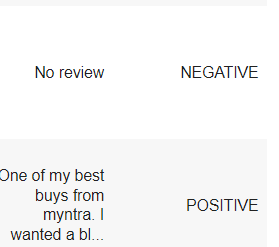
A Large Language Model (LLM) is a type of deep learning model, usually based on transformer architecture, that is trained on large amounts of text data to understand human-like speech and enable this model to perform various natural language processing applications including text generation, sensitivity analysis and language translation.

I have used the "**juliensimon/reviews-sentiment-analysis**" model for analysing the sentiment of review column. "juliensimon/reviews-sentiment-analysis" is a typical example of an LLM optimized for sentiment analysis, focused primarily on analysing the sentiment of repetitive products, since it is trained on amazon product review dataset. Pre-trained transformer-based language models are best -tuned to a large dataset of labelled product reviews and learn how to categorize emotions into positive, negative or neutral groups

To conduct a sentiment analysis, the model takes a text, such as a sentence, as input, and assigns a numerical score that reflects the sentiment expressed in the text. The score represents the confidence of the model in classifying emotions as positive, negative or neutral. The model achieves this by using known patterns and relationships in the textual data, enabling it to recognize and classify emotions based on the language used in the input.

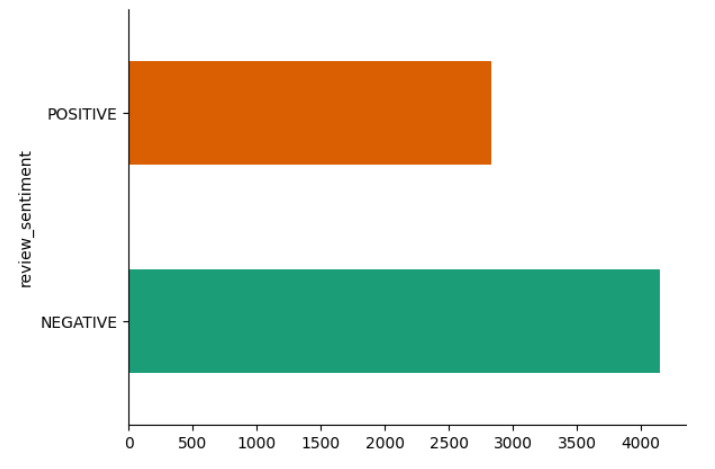
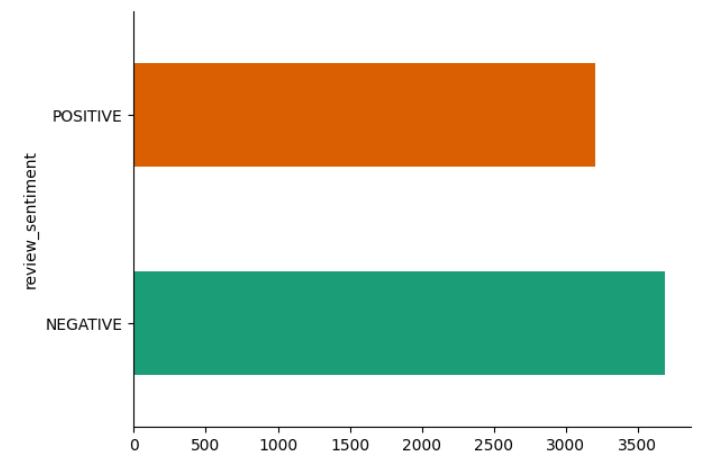
Since each record has several reviews, the model receives each review and determines the sentiment of the record based on it. The corresponding sentiment for each review sentence in a record will be labelled as positive, negative and neutral. The class that receives the majority of votes will be designated as the sentiment for the product.



****

**Sentiment Analysis**

**Women Men**



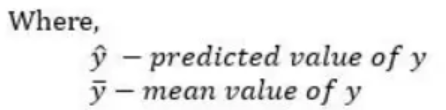
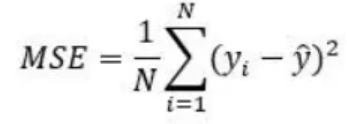
There are more positive reviews for women's apparel compared to that of men’s. Myntra undoubtedly has the best women's collections! The negative reviews are more in both the graph since null values were replaced by "no reviews" and it is labeled as negative sentiment by the model.

1. **Building and evaluating ML Models**

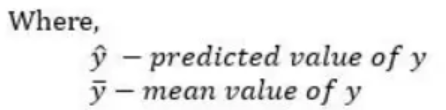
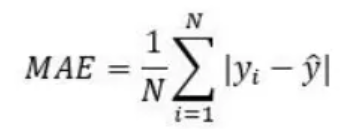
To predict the price of the apparel, columns such as brand, seller, product name, rating, total bought, strike price, discounted price, Available sizes, specifications, Material, Type were used to train the model. All these columns were combined into a single feature for effective analysis. The TF-IDF vectorizer function was used to convert the combined columns that influence the apparel item's pricing into vectors. The dataset was then subjected to a number of different regression models in order to forecast the pricing of clothing items.

Some of the evaluation metrics that are widely used to evaluate the performance of regression models are Mean Squared Error (MSE), Mean Absolute Error (MAE), and R-squared (R^2).

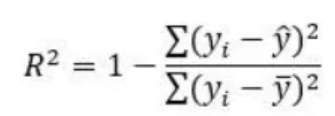
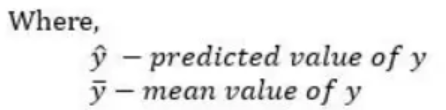
**Mean Squared Error (MSE):** MSE measures the mean of the squared difference between the actual values in the dataset and predicted values predicted by the modifier. It gives an idea of the average mean squared deviation of the predicted values from the actual values. Lower MSE values indicate better model performance.



**Mean Absolute Error:** This helps to calculate the average absolute difference between actual and estimated values of the dataset. It gives the average absolute deviation of the estimated values from the actual values. Similar to MSE, the lower the MAE value, the better the model performance will be.



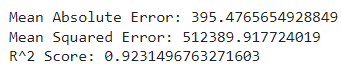
**R-Squared Error:** R-squared is a statistical measure of how much of the variation in the dependent variable (target) can be explained by a model's independent variables (features). It ranges from 0 to 1, and 1 represents a perfectly well-predicting model of the target variable through the features; 0 means that the model does not explain any of the variability in the target variable beyond that which is expected by random fluctuation. Higher R-squared values indicate better model performance.

**Fitting Machine Learning Models**

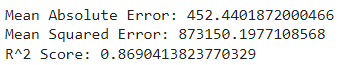
**Random Forest Regressor:**

A part of ensemble learning methods is the Random Forest regressor. It combines several decision trees to make a prediction. It receives a prediction from every tree, which it then averages to obtain the ultimate or the final prediction. When it comes to making decisions, they are strong and straightforward. This technique works well with both large and small datasets because it is resistant to overfitting. It is capable of handling both numerical and categorical variables, and without requiring the one-hot encoding of categorical columns, it can transfer these features straight to the machine learning model. The following are the outcomes of applying the random forest regressor.



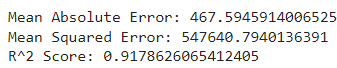
**Decision Tree Regressor:**

For complex regression challenges, this supervised learning approach is employed. It creates a decision-making framework equivalent to a single tree. Every leaf node represents a prediction value, and every interior node represents a decision. Overfitting may result from its ability to identify complex patterns in the data, although regularisation strategies can help prevent this from happening. It is robust to outliers in the data. The following are the outcomes of applying the decision tree regressor.



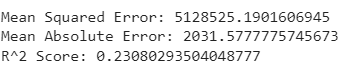
**XG Boost Regressor:**

Many people are aware of XGBoost Regressor for its exceptional performance and scalability. It can handle large datasets with high-dimensional features. It analyses the significance of every factor and focuses decisions on the most important features. Each new tree is trained to minimize loss of the previous decision trees. The following are the outcomes of applying the XGB regressor.



**Ada Boost Regressor:**

Adaptive boosting regressor is also an ensemble learning method. It combines the weak learners in sequential manner. Weights are adjusted based on the errors made by the previous weak learners. The prediction of each weak learner is taken based on its performance. The following are the outcomes of applying the Ada Boost regressor.



1. **Fine-Tuning the model.**

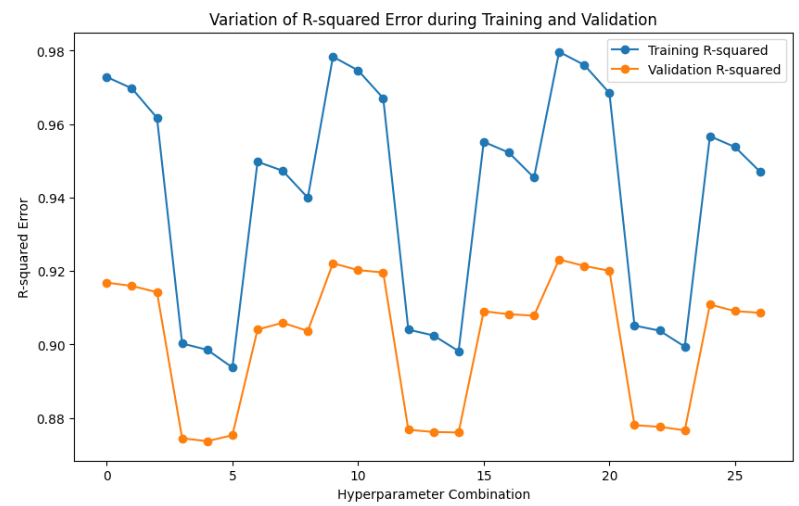
The Random Forest regressor yields good results when compared with the decision tree regressor, Ada boost regressor, and XG boost regressor, according to the results. When compared to other models, R-squared error is the highest and MSE and MAE are the lowest. Therefore, adjusting a few of the model's parameters could improve the outcome.Here are a few parameters that can be fine-tuned:

**Learning\_rate**: a tool for process of optimization.

**max\_depth**: This specifies the maximum depth for every tree.

**n\_estimators**: is the number of trees that the model will construct.

The most widely used method for determining the ideal parameter combination given a list of parameter values is **GridSearchCV**. It determines the best combination to lower MAE and MSE and to increase R square value. GridSearchCV estimates the model's performance by cross-validating the training data for every set of hyperparameters, assisting in preventing overfitting

.   


Here we can see that, R-Squared value is highest when n\_estimators is 100, and min\_sample\_split is 10.

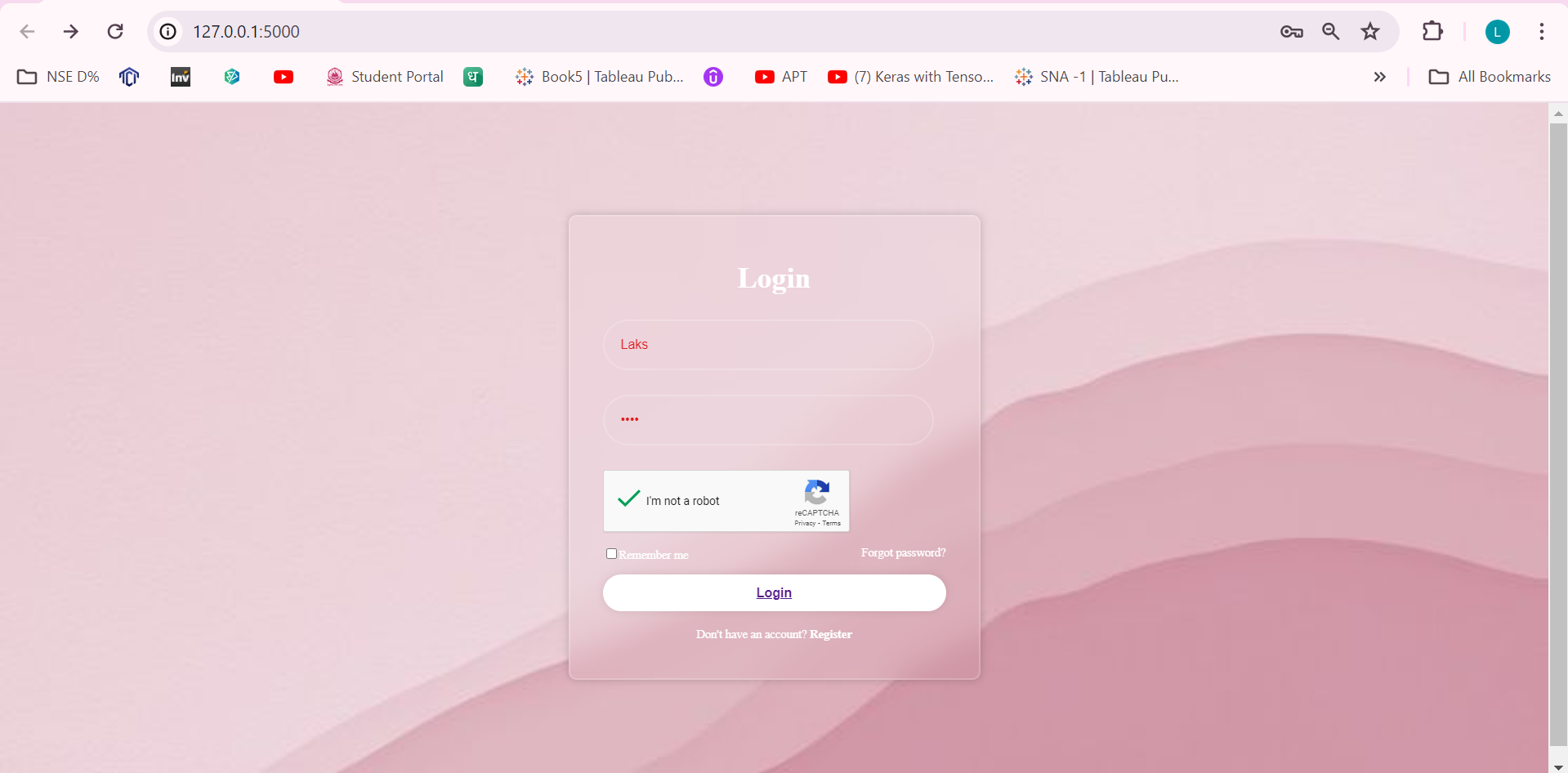
1. **Interactive User-friendly website**

Python Flask was used to develop a user-friendly website, while SQL Alchemy was used for the backend. There is a table called seller in the database. It consists of passwords and user IDs. The Flask application can be used to add users and update passwords. The website contains minimal information to be input, making it easier for even less educated people to use it efficiently to get best retail prices for their clothing items.

**Login Page:**

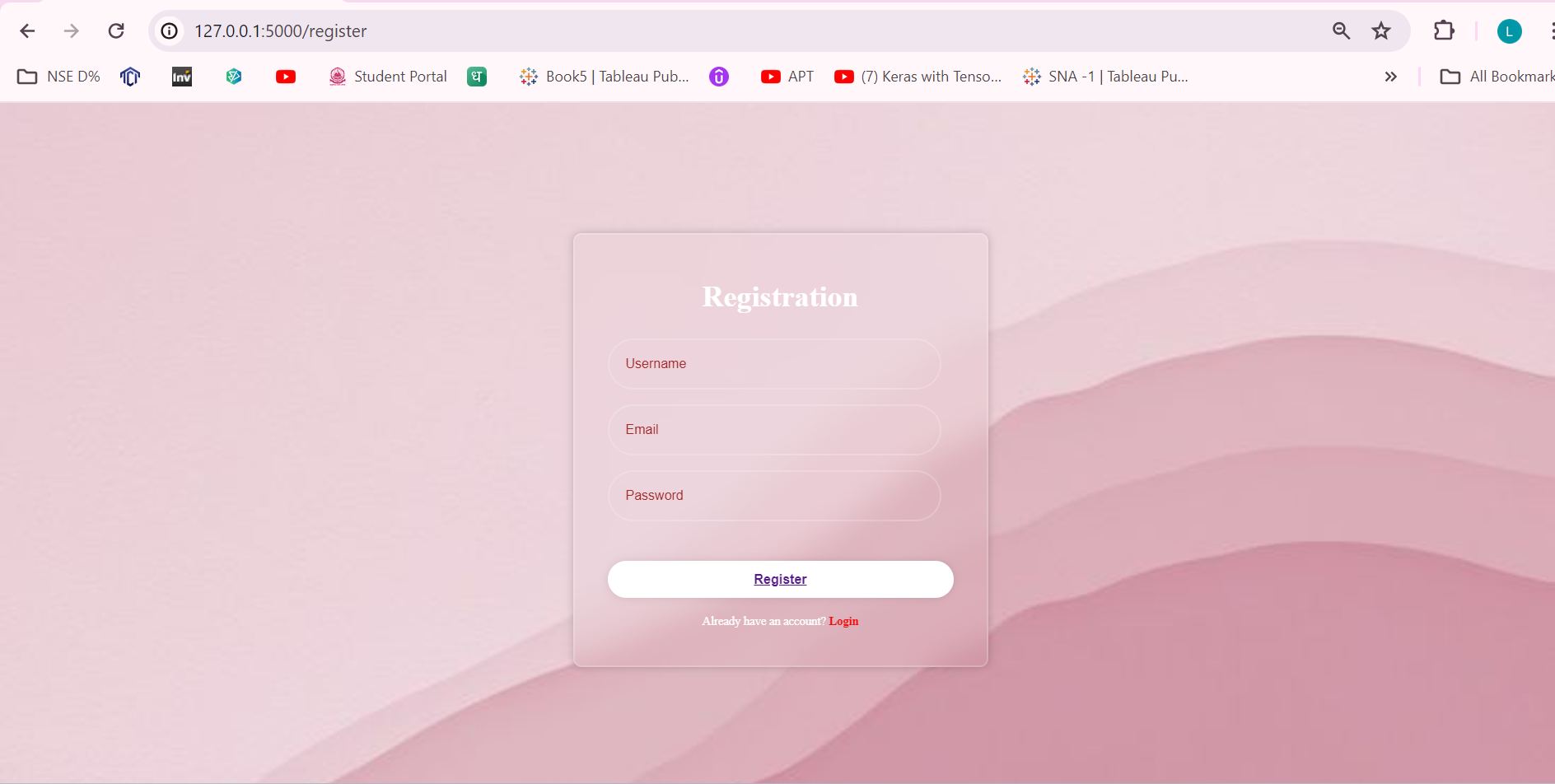
The login page receives user input and verifies it in the database. This uses Google ReCAPTCHA version 2. Google reCAPTCHA v2 is a security tool that confirms a user is human and not a bot in order to shield websites from spam and abuse. Integrating Google ReCaptcha V2: I first registered my website on the Google ReCAPTCHA official website to obtain the API keys, which include site key and secret keys, in order to obtain ReCAPTCHA v2authentication for my website.

Users are prompted with a straightforward checkbox to verify their identity, The following page is navigated if the user ID and password entered match a record in the database and the captcha is verified correctly.



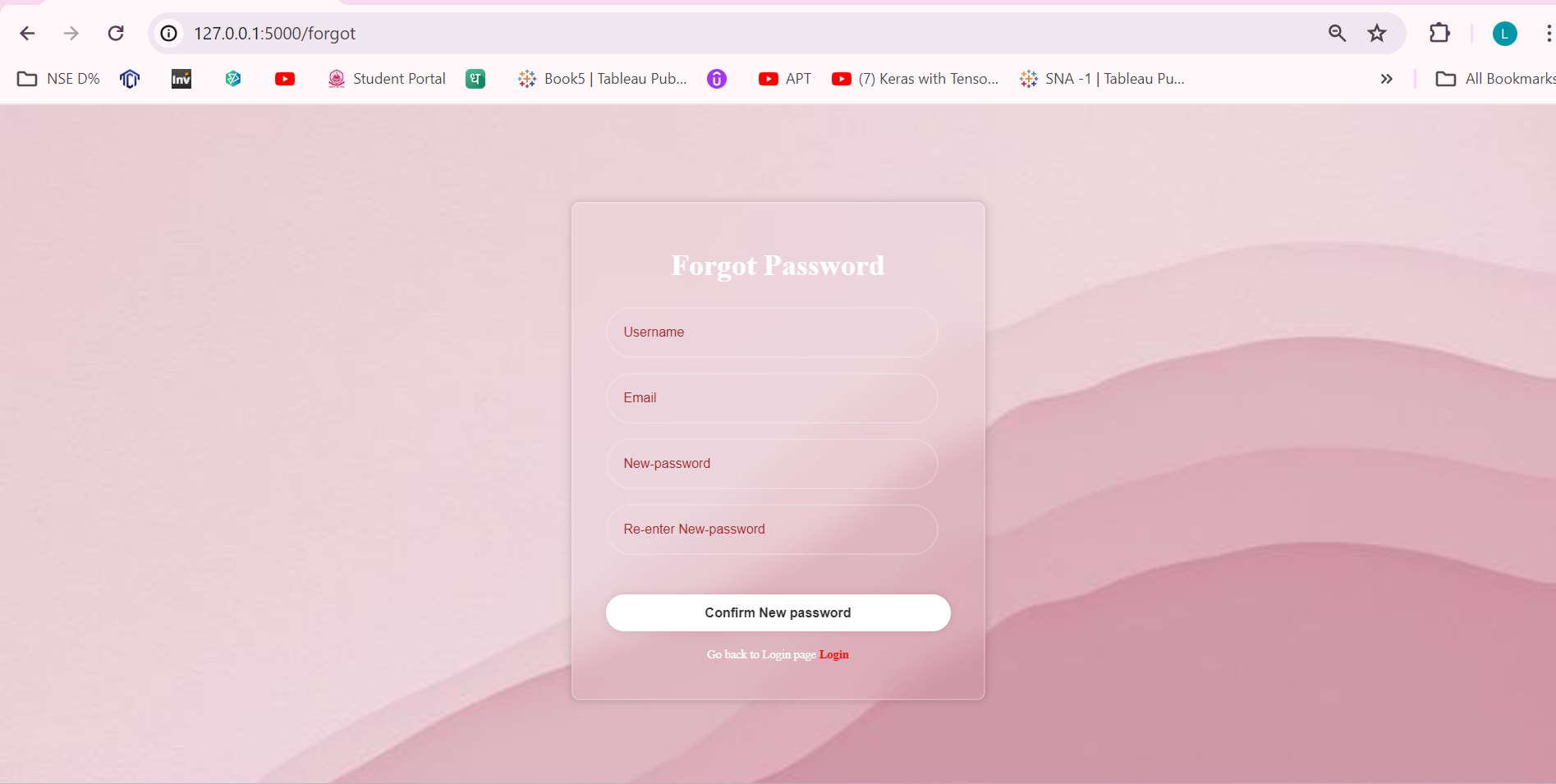
**Registration Page**

A new user can register by clicking the link provided on the login page if they are interested in using our product. By selecting the "Register" button, he can register his information, which will be stored in the backend. He can click the Login link to return to the login page after registering.



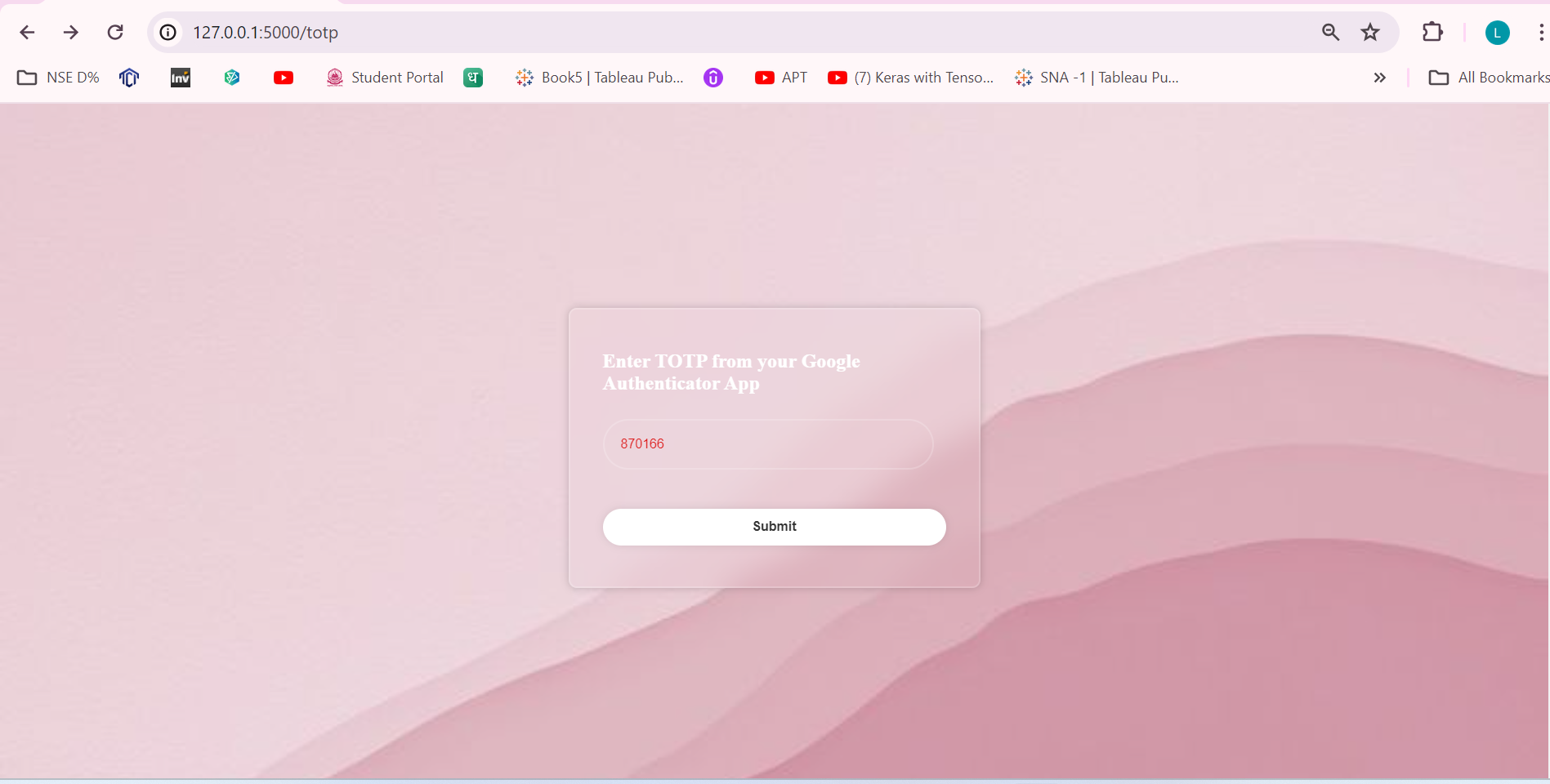
**Forgot Password Page**

An infrequent user can click the forgot password option on the login page to reset their password if they happen to forget it. A SQL query is used in the database that will update the password of that particular user. Then, after changing his password, he can log in to our website effortlessly.



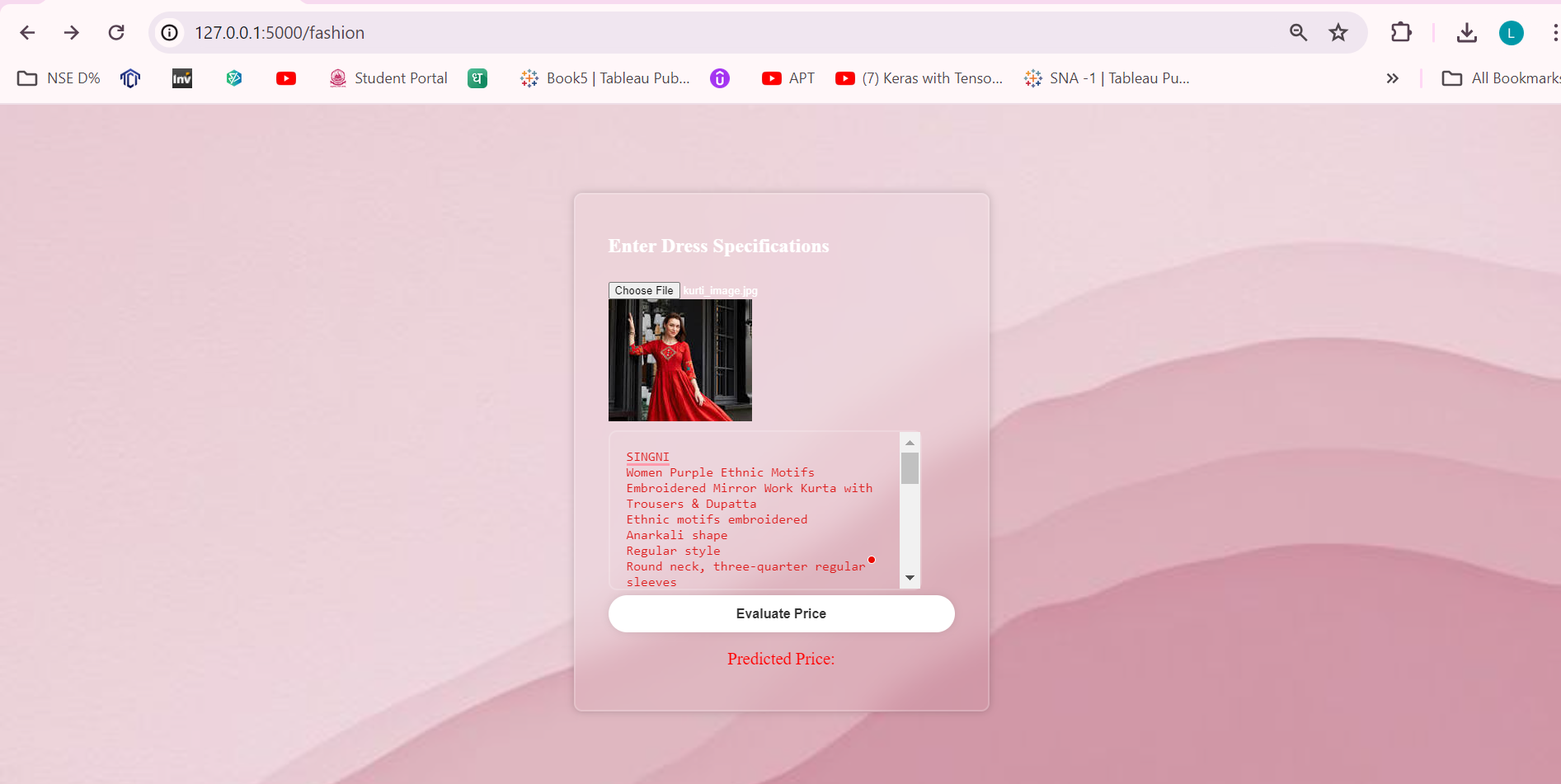
**Time-based-one-time-password**

To ensure that the correct individual is using the product, a time-based one-time password has been set up. The Microsoft Authenticator is linked to this security feature where a 6-digit OTP is updated every 30 seconds. So, it only logs you in to our prediction page if the password you entered matches.



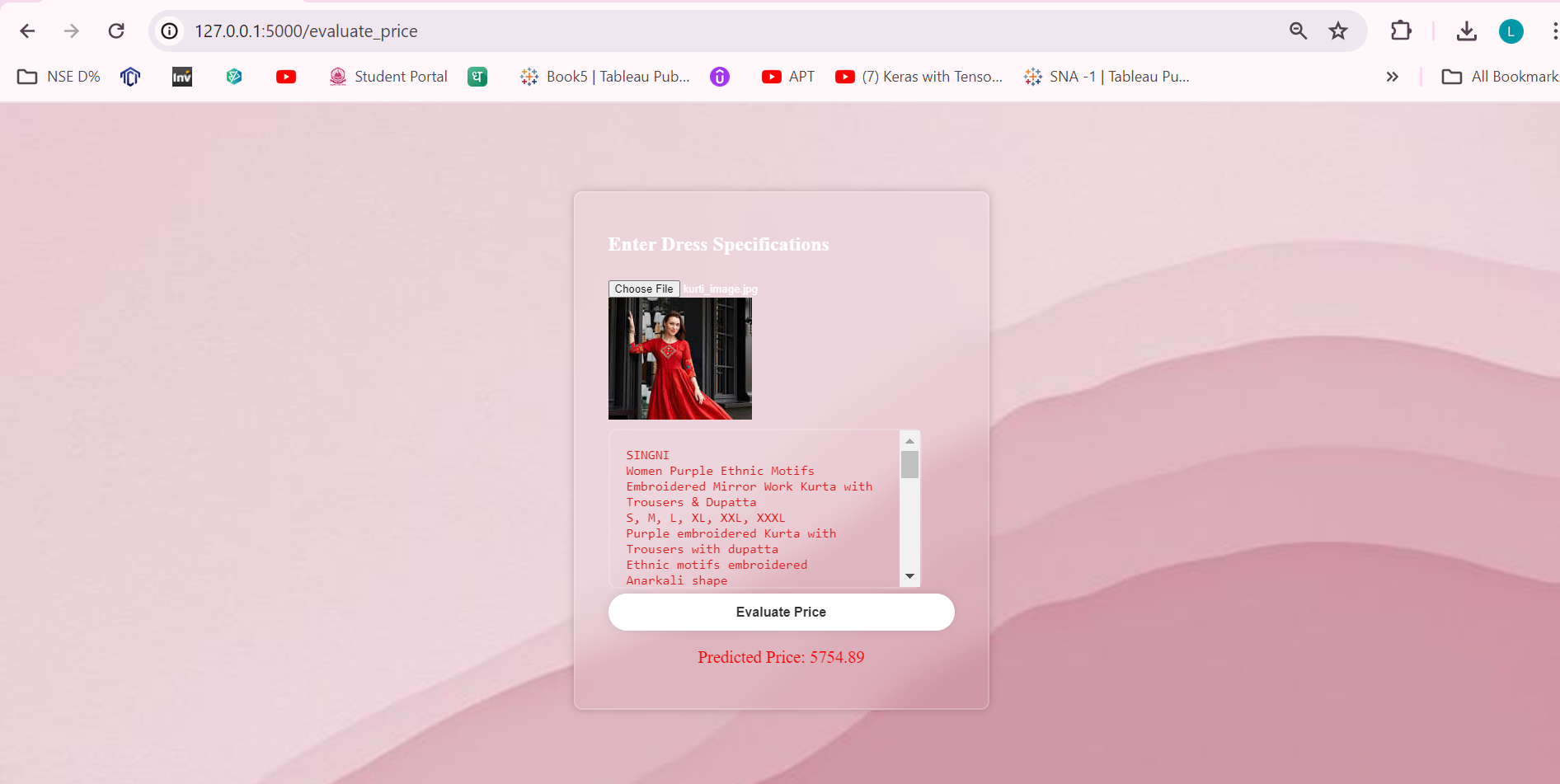
**Input Page**

The main model for the prediction on this straightforward yet powerful prediction page is the Random Forest regressor. To estimate the price of a product, the user can enter all the product specifications he is aware of. In order to confirm that he provides the right specs, he can also upload a photograph. Having numerous input text boxes on this page is an additional implementation method, but the user might not comprehend or might even think that filling up each text box takes a long time.



**Output Page**

By selecting the "evaluate price" button, the customer will find the best deal on his clothing.



1. **Conclusion**

Modern online sellers’ price their apparel items using a combination of budget and luxury pricing strategies. The cost of each article of apparel that is sold online affects consumers and shapes their initial decision to purchase or not. Value-based pricing is the approach I have applied to every article of clothing that retailers have put online. There are a lot of established traditional techniques, such as the backward pricing method, the keystone markup method, and the absorption pricing method, but they take a lot of time to compute each time, especially for large brands and the volume of textile materials produced nowadays. It takes less time and is a simple tool to price things effectively using machine learning models. My project clothing evaluator is a state-of-the-art tool that sellers from remote locations and online fashion retailers may use to price their products profitably and persuade clients of their worth.

Additionally, the clothes merchants can get vital insights from the visualisations that were displayed as part of the exploratory data analysis, as they indicate important client buying patterns and trends. By being aware of the kinds of discounts that encourage clients to make larger purchases, the kinds of fabrics they favour, the most popular clothing attires, and so forth. These visualisation approaches can also be used to forecast future clothing demand. Another important thing to keep in mind is how much customers are influenced by the sentiment expressed in clothing reviews. Positive reviews help build a positive reputation for the product and brand, which boosts sales.

1. **Reference**

[1] N. S. Sutiawan and I. G. B. B. Nugraha, "Online price prediction system of consumption commodities," 2017 International Conference on Information Technology Systems and Innovation (ICITSI), Bandung, Indonesia, 2017, pp. 145-150, doi: 10.1109/ICITSI.2017.8267933.

[2] T. K. Shivaprasad and J. Shetty, "Sentiment analysis of product reviews: A review," 2017 International Conference on Inventive Communication and Computational Technologies (ICICCT), Coimbatore, 2017, pp. 298-301, doi: 10.1109/ICICCT.2017.7975207.

[3] Data-Driven Retail Excellence: Machine Learning for Demand Forecasting and Price Optimization by Vinit Taparia, Piyush Mishra, Nitik Gupta, Hitesh Chandiramani [2024: Vol 12 Iss 1](https://journal.riverpublishers.com/index.php/JGEU/issue/view/43) pp 68–98

[4] Y. Zhang and S. Liou, "Wardrobe Rentals: A Study of the Fashion Industry under Sharing Economy," 2021 9th International Conference on Orange Technology (ICOT), Tainan, Taiwan, 2021, pp. 1-4, doi: 10.1109/ICOT54518.2021.9680630.

[5] S. K, B. S, K. R, V. P. V and M. A, "Customization of User Experience in Fashion Technology," 2023 3rd International Conference on Innovative Practices in Technology and Management (ICIPTM), Uttar Pradesh, India, 2023, pp. 1-6, doi: 10.1109/ICIPTM57143.2023.10117970.

[6] Z. Wang and X. Liu, "Statistical Analysis and Big Data Based Intelligent Fashion Prediction Model," 2021 IEEE Conference on Telecommunications, Optics and Computer Science (TOCS), Shenyang, China, 2021, pp. 878-882, doi: 10.1109/TOCS53301.2021.9688919.