

Melanin-Centered Skincare Recommendation System

Presented By:

THE GLOWCODE



THE TEAM



ESTHER CHERUIYOT

*Group Leader &
Data Processing Lead*

- Data Loading: Prepare and load datasets for project use.
- EDA
- Documentation (Reports): Create project reports, update project proposals, and compile results.
- ReadMe on GitHub



BRIAN GITHINJI

*Modeling & Recommendation
System Specialist*

- Modeling: Develop and fine-tune machine learning models (including collaborative and content-based filtering) for the recommendation system.
- Machine Learning Model to Complement Overall Recommendation System: Develop a secondary model to strengthen recommendation accuracy.



GRACE GITAU

*Feature Engineering &
Deployment Specialist*

- Feature Engineering: Clean and engineer features from the loaded data to prepare it for analysis and modeling
- Exploratory Data Analysis (EDA).
- Deployment: Manage the deployment process, ensuring that the system functions effectively in its deployed environment.



MAUREEN IMANENE

*Sentiment Analysis &
Presentation Specialist*

- Sentiment Analysis: Preprocess data & Conduct sentiment analysis on product reviews to help refine the recommendation model.
- Deep Learning Model for Sentiment Prediction: Build an LSTM model to predict sentiment.
- Non-Technical Presentation

Project Summary

- **Objective:** Address the gap in skincare recommendations tailored for women of color with unique needs like hyperpigmentation, dryness, and sensitivity.
- **Solution:** A recommendation system using AI techniques (content-based filtering, collaborative filtering, and sentiment analysis) to generate personalized skincare suggestions.
- **Impact:** Improves accessibility to effective skincare solutions and enhances satisfaction for women with melanin-rich skin.



Business Problem

- **Problem Statement:** Current recommendation systems often overlook specific needs of women of color, leading to less effective skincare experiences and results.
- **Goal:** Develop a targeted AI-driven solution that addresses unique skincare concerns by leveraging advanced techniques and integrating user-specific features.



Product Skincare

Project Objectives

- Develop a melanin-centered skincare recommendation system using deep learning, tailored for Black women's unique skin needs.
- Utilize content-based and collaborative filtering along with sentiment analysis to enhance recommendation accuracy.
- Deploy an accessible Streamlit interface for personalized, user-friendly skincare suggestions.

Skincare
Product

Stakeholders

Skincare
Product

- **Users:** Women of color seeking tailored skincare solutions.
- **Skincare Brands:** Interested in insights on products suited for melanin-rich skin.
- **Healthcare Professionals:** Dermatologists who may use the system to recommend products.

Data Understanding

Data sources:

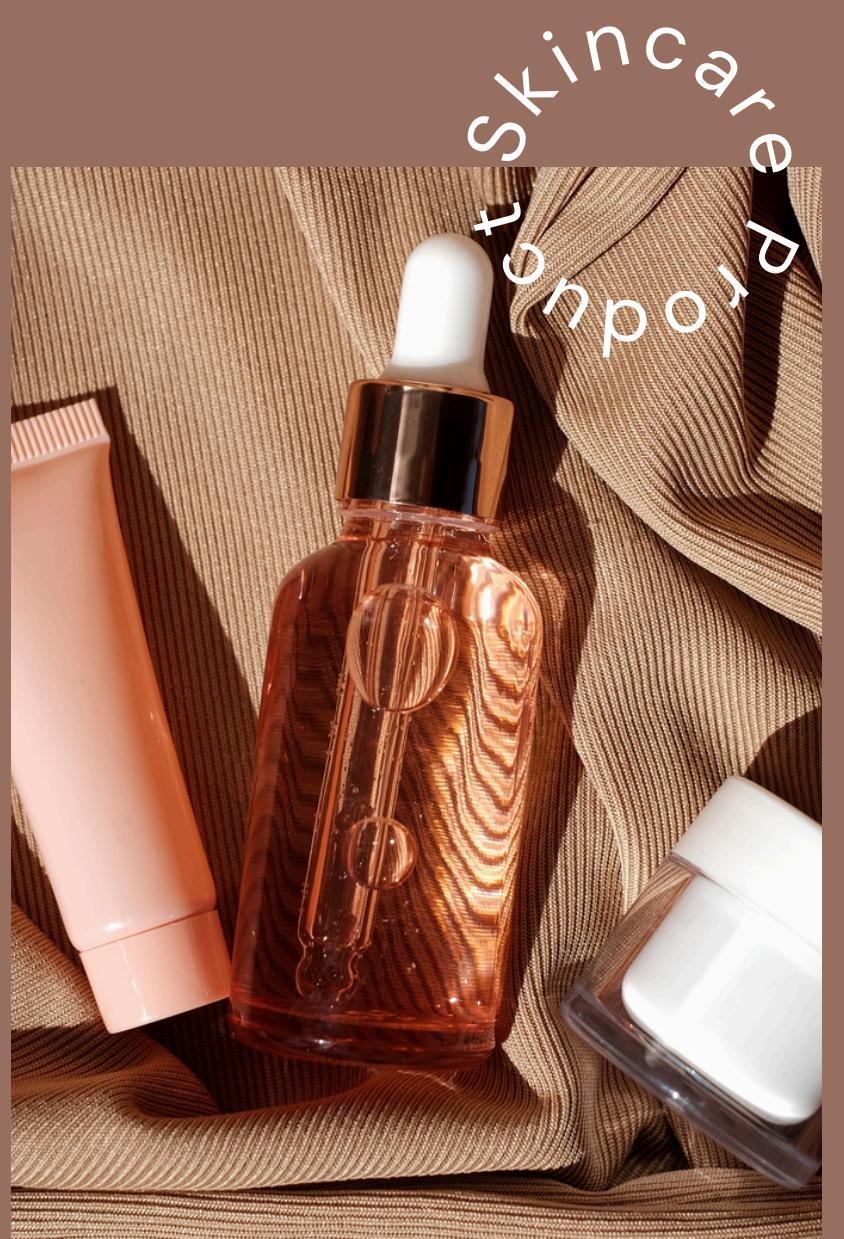
We scraped data on Sephora products and user reviews .

Dataset Overview:

We had 57,367 rows & 29 columns after cleaning.

Key features include:

- **Product Features:** product_id, product_name, brand_name, ingredients, rating, price_ksh, new, out_of_stock, highlights.
- **Review Features:** author_id, rating, review_text, skin_type, skin_tone, and helpfulness.



Metrics of Success

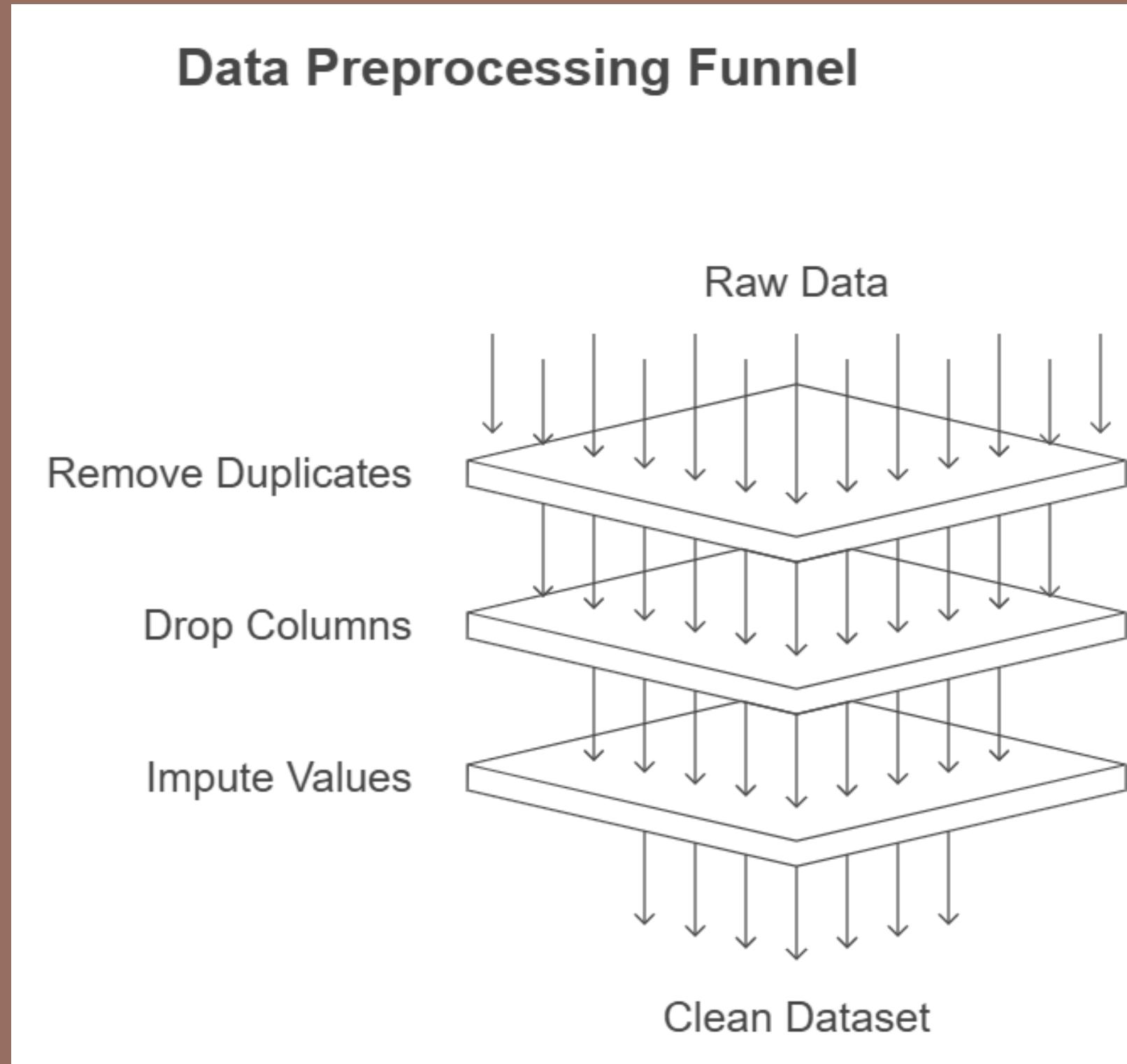
A. Accuracy Metrics:

- Precision@K: 80% relevance in top recommendations.
- Recall@K: 75% of relevant items captured in top recommendations.
- F1 Score: 0.77 (balance of precision and recall).

B. Coverage Metrics:

- Hit Rate @5: 85% of users find a relevant product in the top 5.
- NDCG: 90% of relevant items ranked early in recommendations.

Data Preprocessing Steps



1. Data Loading & Merging datasets.
2. Preprocessing Steps:
 - Checked for duplicates and handled missing values.
 - We dropped columns with over 50% missing values. We were left with 57,367 rows & 29 columns.
 - Imputed missing numerical values with median, and string columns with placeholders.

Feature Engineering

Skin Tone Classification: Created a 'skin_tone_category' to classify users as 'melanated' or 'non-melanated.'

Additional Features: Converted date formats, added 'price_ksh' for local currency analysis, and categorized products by price tier.



EDA

We performed;

**UNIVARIATE ANALYSIS, BIVARIATE ANALYSIS and
MULTIVARIATE ANALYSIS.**

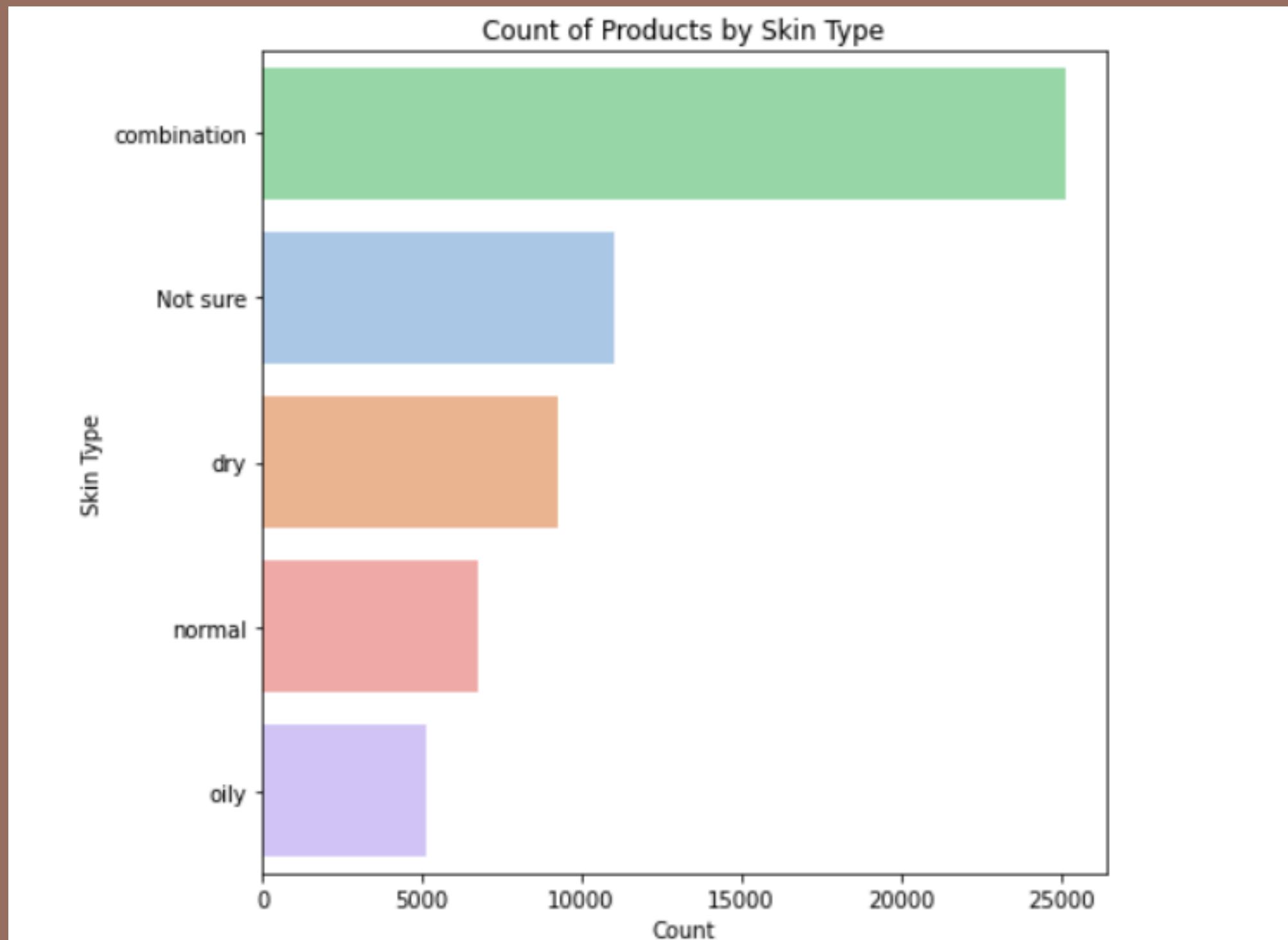
We found that;

- **Rating Distribution:** Most products have high ratings, indicating general customer satisfaction.
- **Price Distribution:** Majority of products are low-to-mid price, making them budget-friendly.
- **Skin Type Insights:** Combination skin products are the most common, followed by dry and normal skin products.



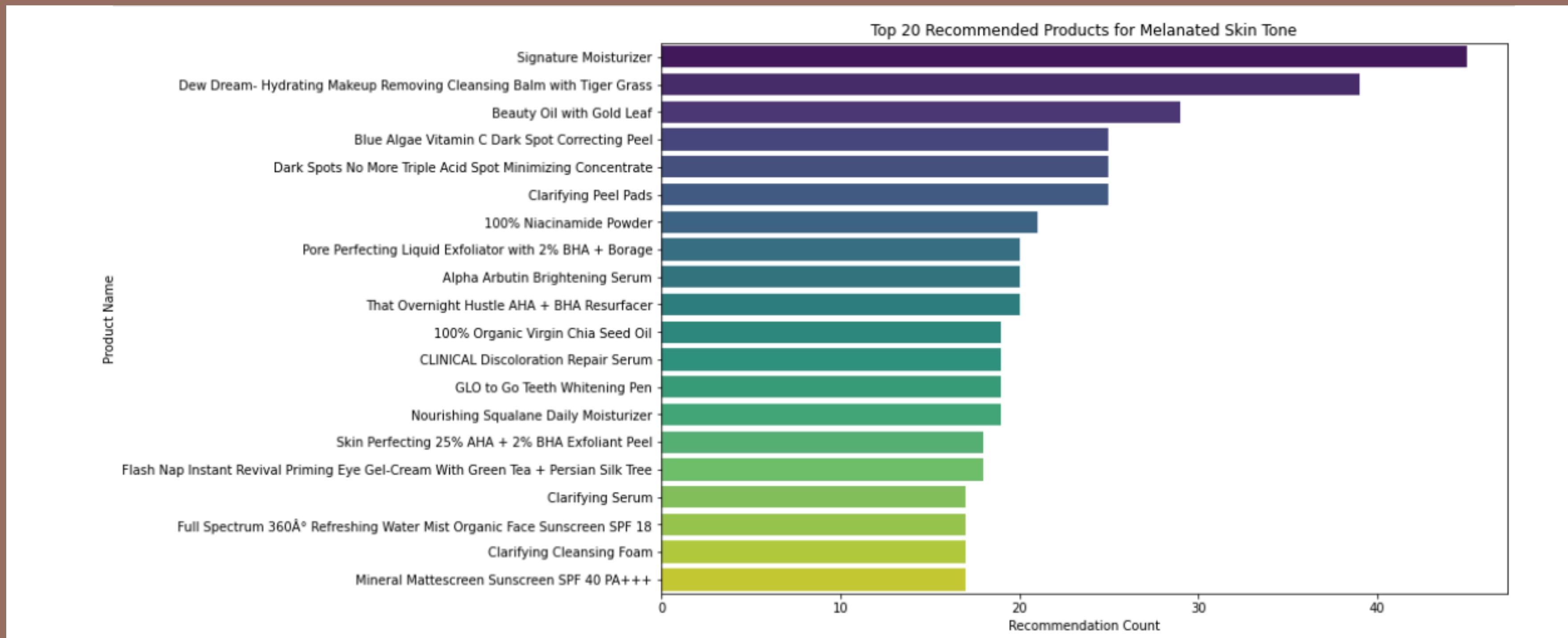
Product Skincare

Count by Skin-type

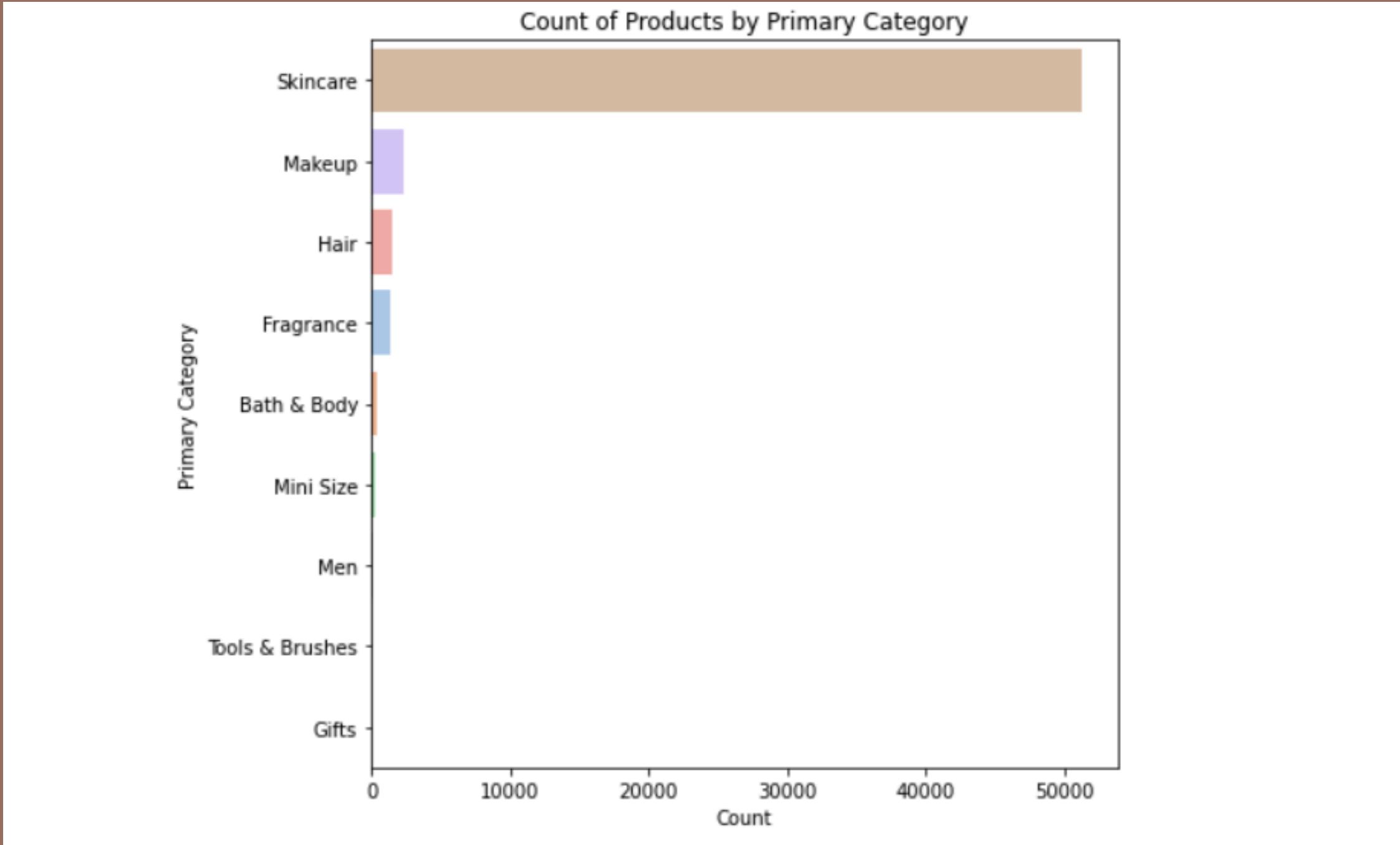


- Combination skin products are the most popular, followed by those for dry, normal, and oily skin types.
- This trend highlights the need to prioritize products for combination and dry skin, aligning with the primary needs of the Black women demographic.

Top recommended products

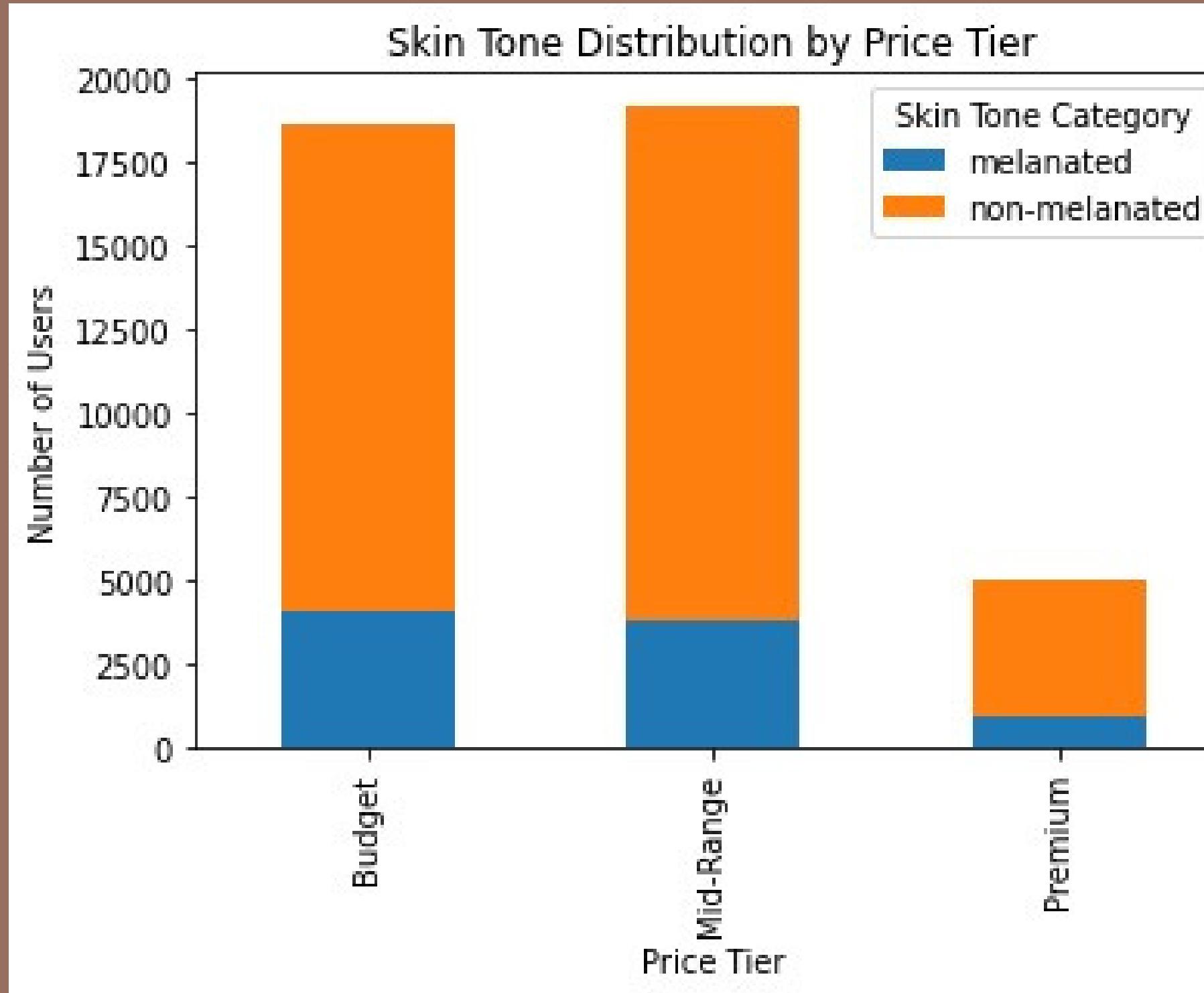


Distribution of products by primary category



- The dataset's focus is clearly on skincare, aligning with the project's goal of providing skincare recommendations.
- Limited data on non-skincare categories suggests fewer insights for those areas.
- This distribution underscores the dataset's strong suitability for a skincare-focused recommendation system.

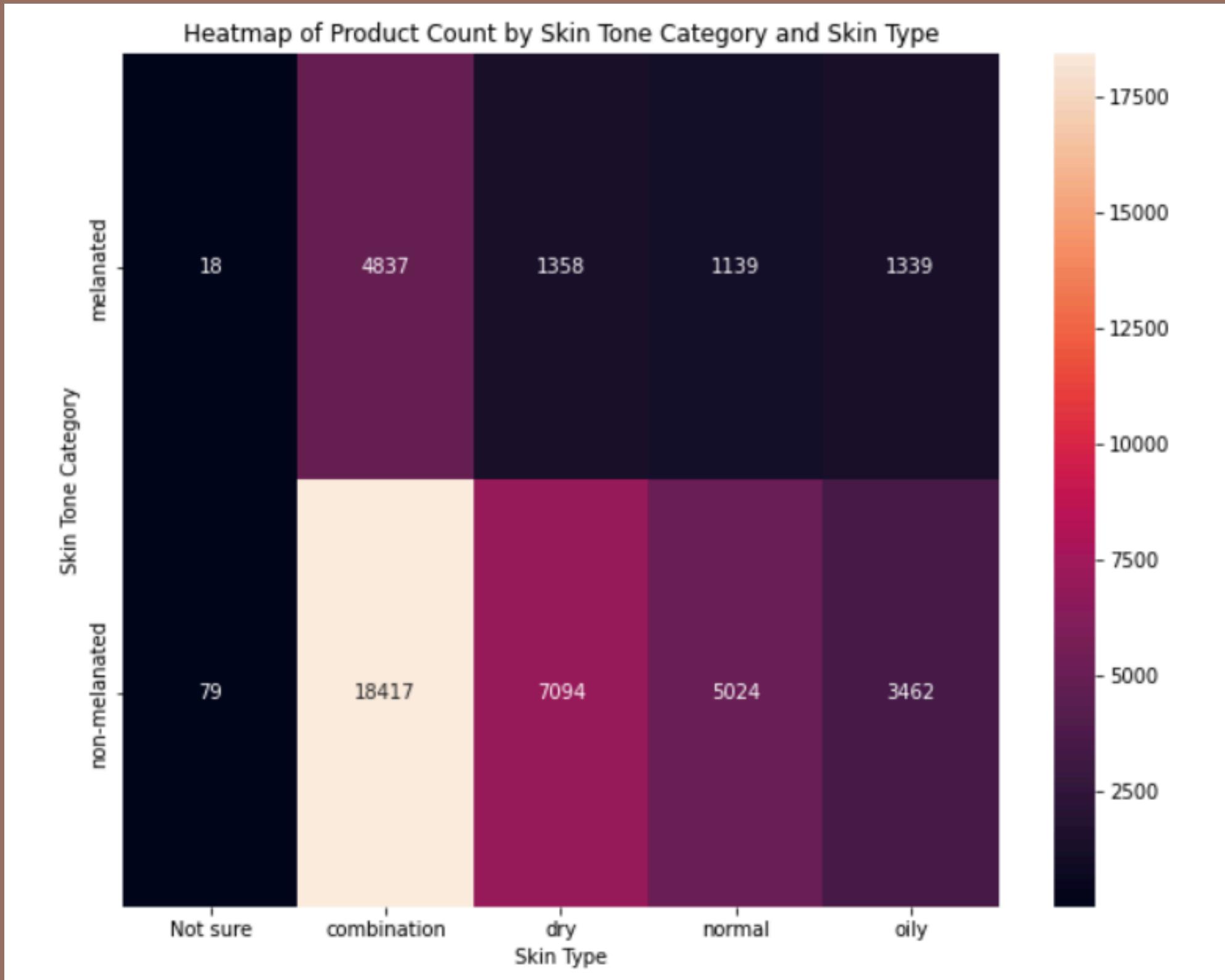
Skin Tone Distribution by price Tier



This chart shows that most users prefer Budget and Mid-Range products, with non-melanated users dominating all price tiers.

Premium options are less popular overall, especially among melanated users, highlighting a potential gap in accessible high-end products for this group.

Correlation HeatMap

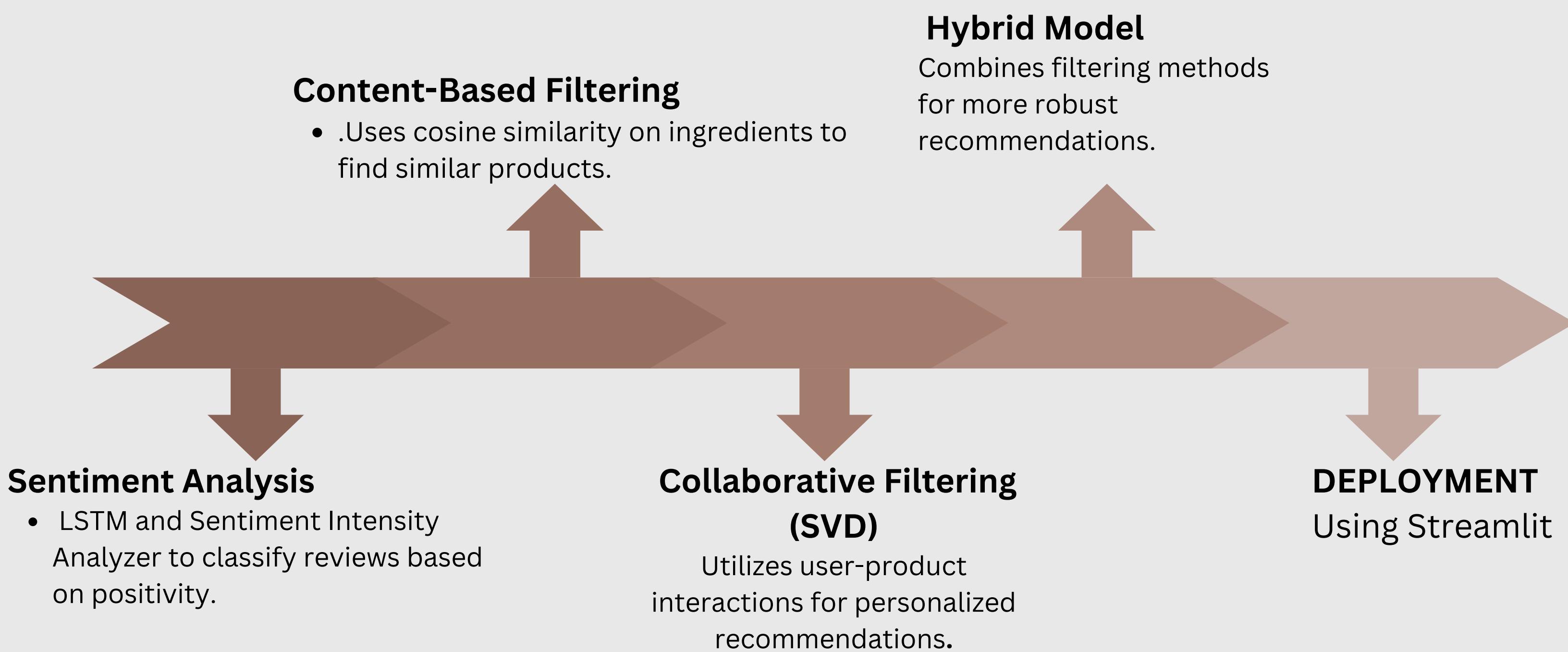


This heatmap shows product availability across skin tone categories and skin types. Key insights include:

- Combination Skin:** Products for combination skin are the most abundant, especially for non-melanated users (18,417 products), compared to 4,837 for melanated users.
- Dry, Normal, and Oily Skin:** There are significantly fewer products for these skin types across both skin tone categories, with non-melanated users having more options in each category.
- Gaps for Melanated Skin:** Melanated users have fewer product options across all skin types, indicating a potential market gap for products catering to their specific skin needs.

This analysis highlights an opportunity to diversify offerings for melanated skin, particularly for non-combination skin types.

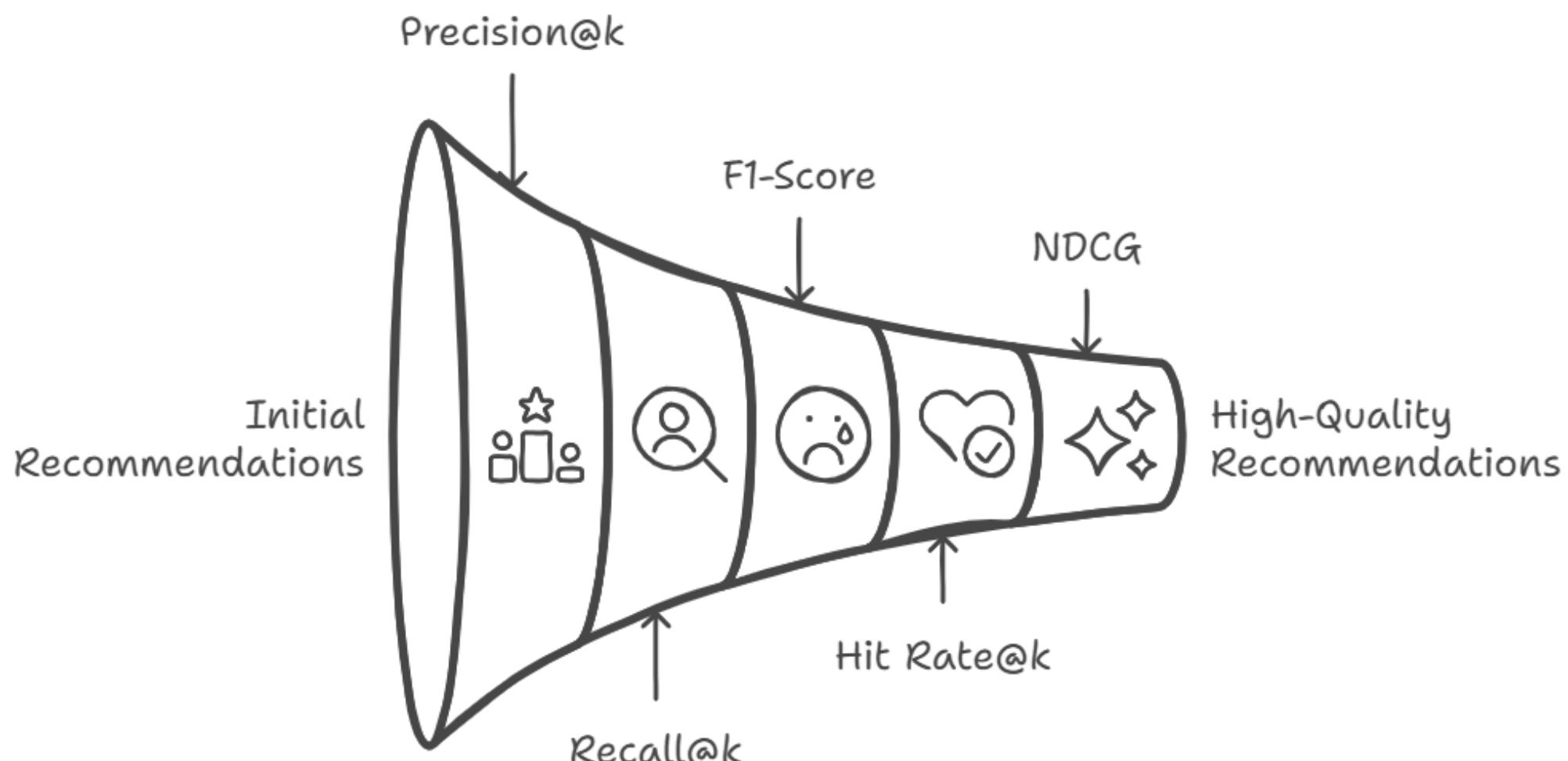
Modeling Techniques



Model Evaluation Metrics

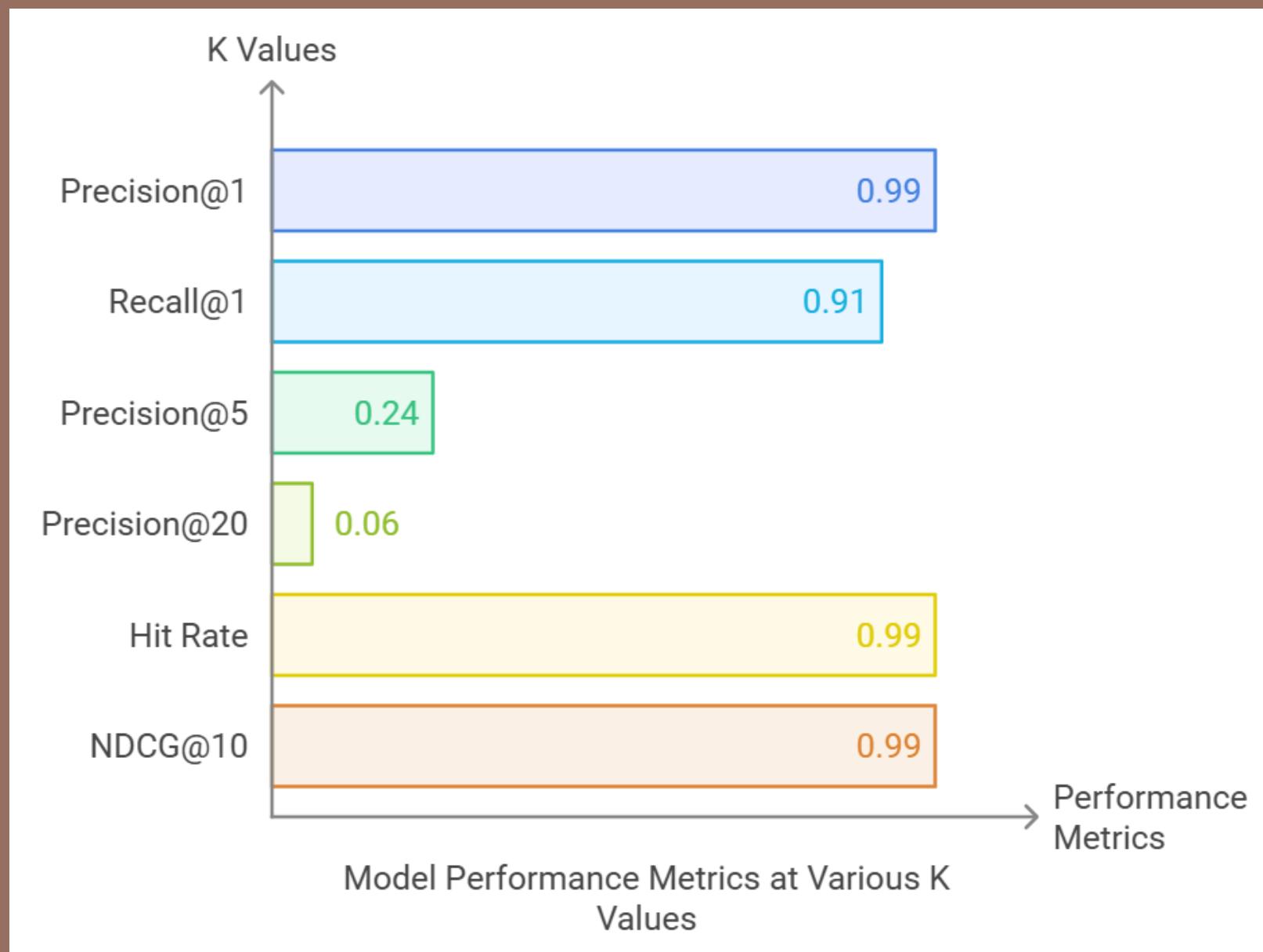
Product Skincare

Refining Recommendation Quality



- **Precision@k:** Measures the relevance of the top K-recommended products.
- **Recall@k:** Evaluates the system's effectiveness in covering user preferences in the top K recommendations.
- **F1-Score:** Ensures a balance between precision and recall.
- **Hit Rate@k:** Determines if users find a relevant product in the top 5.
- **Normalized Discounted Cumulative Gain (NDCG):** Considers the position of relevant items, rewarding relevant items appearing earlier in the recommendations list.

Model Performance



1. High Precision and Recall at K=1:

Precision@1 of 0.99 and Recall@1 of 0.91 indicate that the model is very accurate when recommending a single top item. Users are very likely to find a relevant item in the top 1 recommendation.

2. Performance Drops as K Increases:

Precision drops significantly as K increases (e.g., Precision@5 of 0.24 and Precision@20 of 0.06). This suggests that as more items are recommended, fewer of them are relevant.

3. Hit Rate@k == 0.99

With a Hit Rate of 0.99 across all values of K , the model is consistently recommending at least one relevant item within the top K recommendations for almost every user.

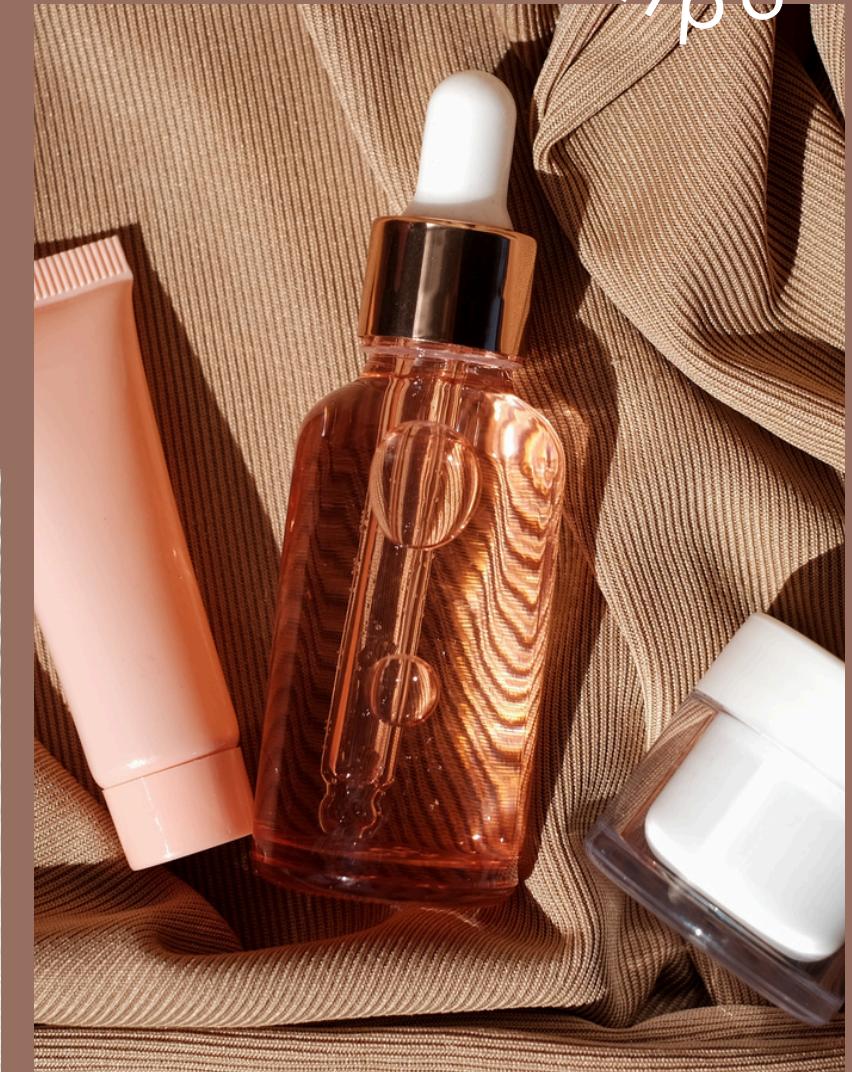
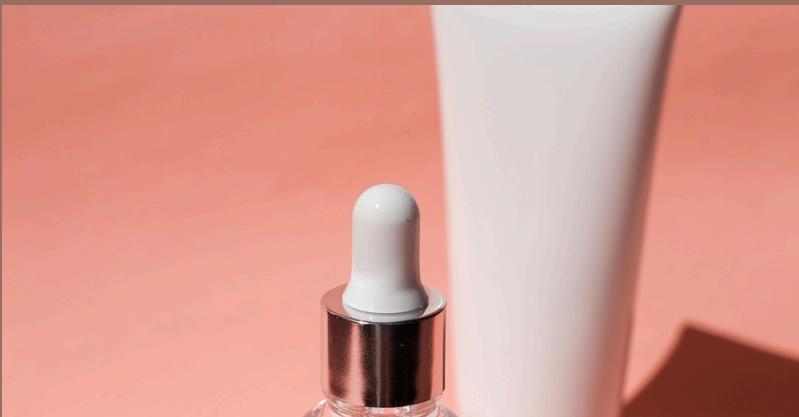
4. The NDCG@10: 0.99

Result is excellent, indicating that the ranking provided by your recommendation model closely matches the ideal order, where the most relevant items appear at the top.

Deployment

Why Streamlit?

- 1. Ease of Use:** Minimal code needed for rapid prototyping.
- 2. Real-Time Interaction:** Instant updates based on user input.
- 3. Seamless Integration:** Works well with Python libraries like `scikit-learn` and `pandas`.
- 4. No Frontend Required:** Handles both backend and frontend, saving time.
- 5. Interactive:** Built-in widgets for user input and tailored recommendations.
- 6. User-Friendly Display:** Simple ways to visualize results.
- 7. Quick Deployment:** One-click deployment for easy sharing.

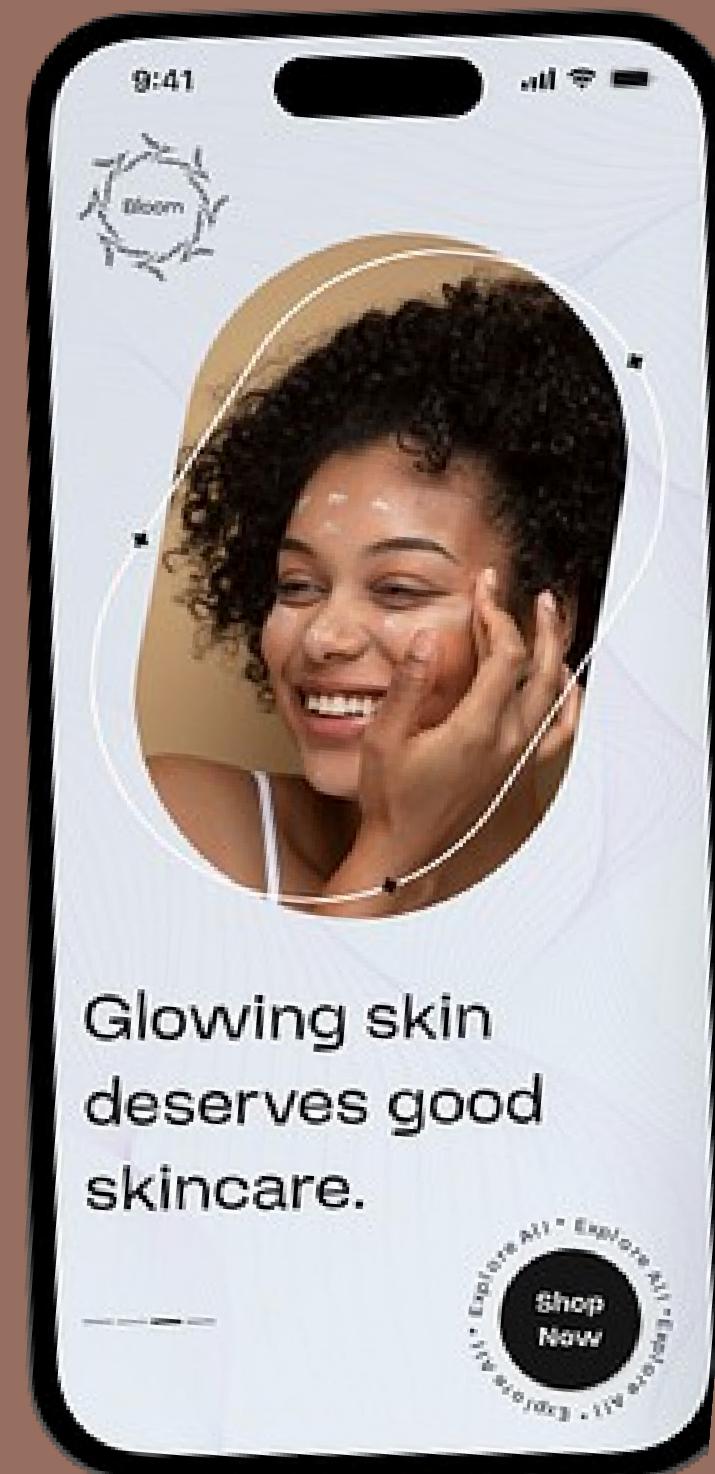


Skincare
Product

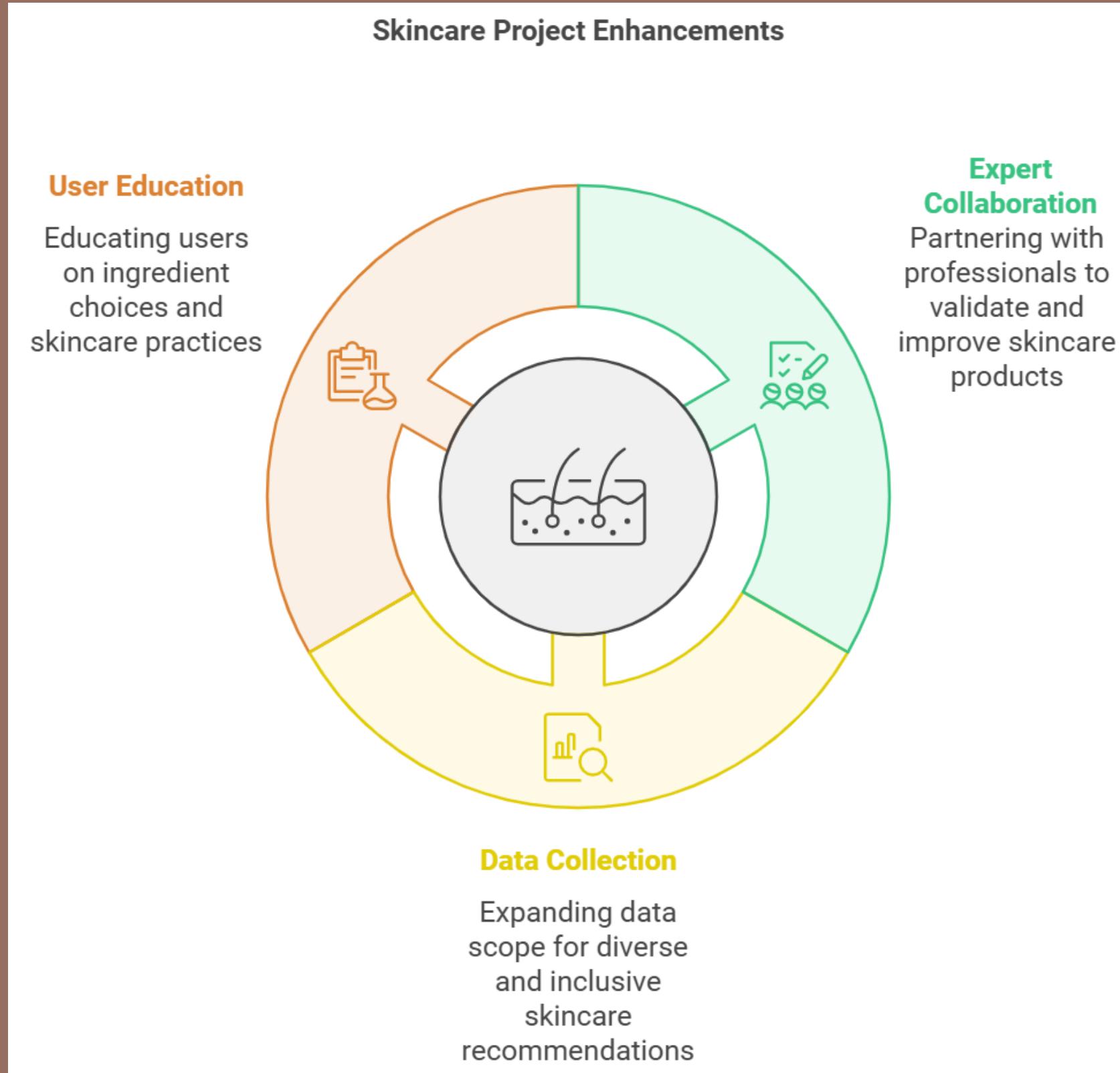
TECH DEMO

Here's the link to our user Interface.

[skincareRecommenderApp](#)



Next Steps



- **User Feedback Integration:** Allow users to rate recommendations, creating a feedback loop for model refinement.
- **Partnership with Experts:** Collaborate with skincare professionals for validation and improvement.
- **Enhanced Data Collection:** Broaden dataset scope for better diversity in recommendations.
- **User Education:** Provide educational content on ingredient choices and skincare practices.

Conclusion

- **Impact:** The AI-driven skincare recommendation system addresses a major gap for women of color, enhancing satisfaction and efficacy.
- **Future Potential:** Demonstrates how AI can make skincare more inclusive, setting a foundation for wider application and expansion.



Thank you.

