**TECHNICAL REPORT:**

**PRODUCT DATASET CLEANING, SHORT TITLE OPTIMIZATION & DATA INSIGHTS**

1. **Introduction**  
   This project aimed to clean and optimize a product dataset, ensuring usability for SEO and analytical purposes. The process included data cleaning, generating concise SEO-friendly product titles, and exploratory data analysis (EDA). Key insights are presented with visualizations to emphasize data trends.
2. **Data Loading and Overview**

The dataset was successfully imported, and an initial inspection revealed issues such as missing values, duplicates, and inconsistent column naming conventions.

1. **Data Cleaning**
2. Missing values were replaced with defaults:
   1. Titles: "Unknown Product"
   2. Bullet Points: "No details available"
   3. Description: "No description available"
   4. Product Length: Median value of the column.
3. Duplicates were removed, and column names were standardized to lowercase and *snake\_case* format.
4. **Key Result:** Dataset now has consistent structure and no duplicates.
5. **SEO Optimization**

A concise version of product titles was generated by limiting them to the first five words. This optimization ensures better indexing for SEO while retaining relevance.  
**Example:**

* Original Title: "Best Ergonomic Chair with Adjustable Height"
* Short Title: "Best Ergonomic Chair with Adjustable"

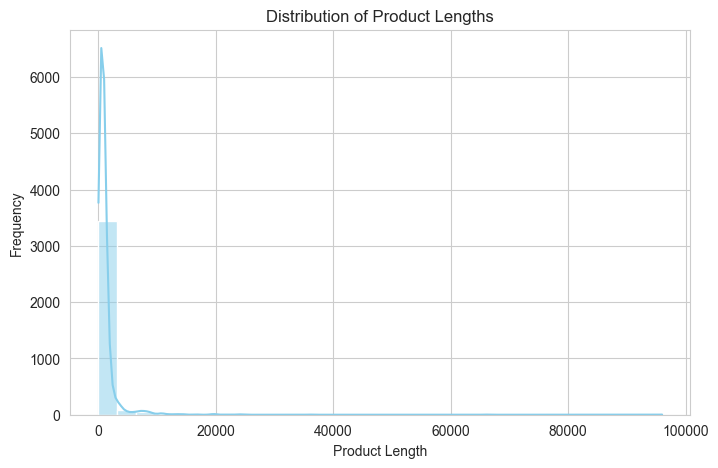
1. **Dataset Export**

The cleaned and optimized dataset was saved as a CSV file for future analysis or operational use.

* **File Path:** cleaned\_product\_\_data.csv

1. **Exploratory Data Analysis (EDA)**
2. **Distribution of Product Lengths**

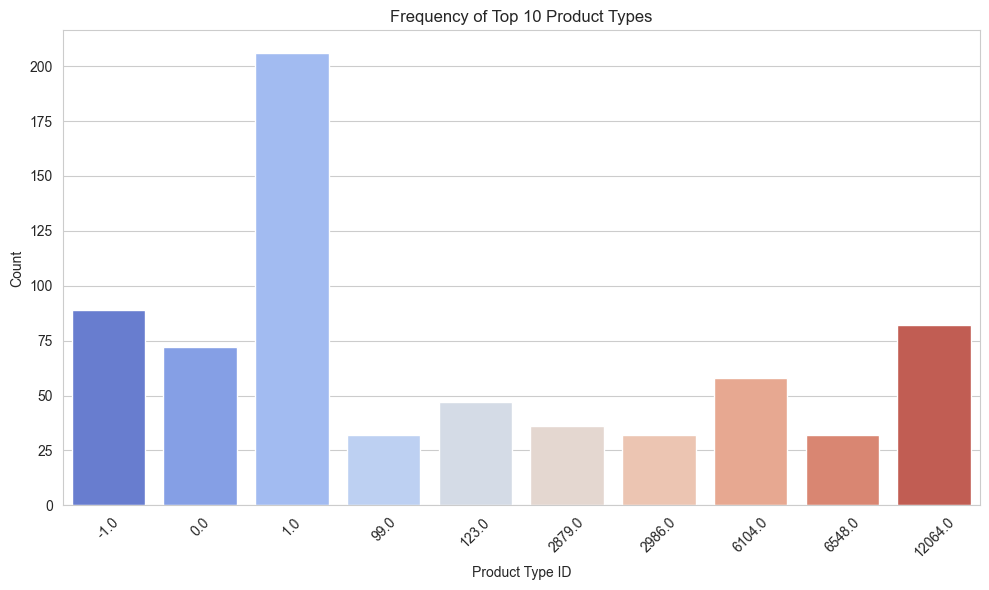
A histogram was generated to visualize the spread of product lengths. This revealed the most frequent ranges and highlighted potential outliers.



**Figure 1:** Distribution of Product Lengths

1. **Frequency of Top 10 Product Types**

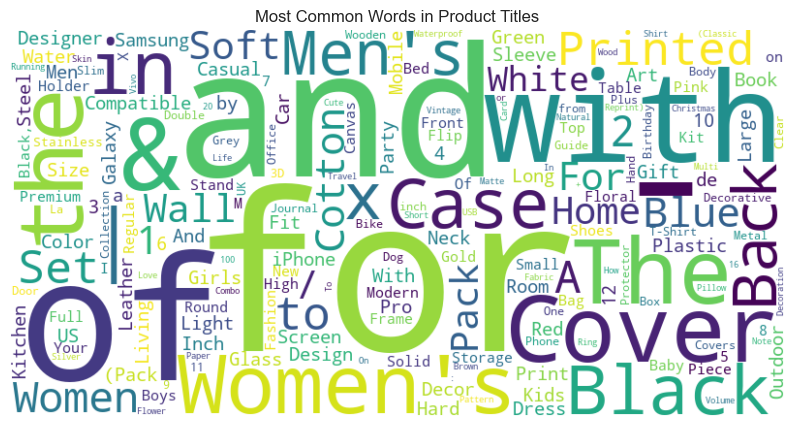
A bar plot was created to highlight the 10 most common product types, providing insight into dominant categories within the dataset.



**Figure 2:** Frequency of Top 10 Product Types

1. **Most Common Words in Product Titles**

Using a Word Cloud, the most frequently used terms in product titles were visualized, uncovering trends in naming conventions.

  
**Figure 3:** Most Common Words in Product Titles

1. **Conclusion**

The project successfully cleaned and optimized the dataset while uncovering key insights through visual exploration. The cleaned data and generated short titles are now ready for operational use or further analysis. Next steps could involve predictive modeling or category-specific analysis.