

Case Project

# Product Analyst

---

Dataset : Womens Clothing E-Commerce

Link : <https://github.com/rahmat14/Case-Project---Product-Analyst->

# Dataset Review

---

Pada case project ini menggunakan dataset “Women Clothing E-Commerce” yang berisi ulasan penilaian kualitas beberapa produk. Case project ini akan menganalisis kualitas suatu produk berdasarkan pada informasi apakah produk diberikan rekomendasi atau tidak, maksud dari diberikan rekomendasi adalah produk tersebut mendapatkan penilaian baik untuk dijual.



# Informasi Dataset

---

Data yang digunakan merupakan data cross section yang terdapat 23.486 baris data dan terdiri atas 10 informasi sebagai berikut :

- Clothing ID : Kode unik pada suatu produk (Kategori Nominal)
- Age : Umur penilai (Kategori Ordinal)
- Title : Judul atas penilaian produk oleh penilai (String)
- Review Text : Ulasan pelanggan atas penilaian produk oleh penilai (String)
- Rating : Peringkat atas penilaian produk oleh penilai (Kategori Ordinal)
- Recommended IND : Rekomendasi penilaian produk oleh penilai (Kategori Ordinal)
- Positive Feedback : Jumlah penilai lain yang memberikan penilaian positif pada produk (Numerik Rasio)
- Division Name : Nama divisi dari suatu produk (Kategori Nominal)
- Department Name : Nama departemen dari suatu produk (Kategori Nominal)
- Class Name : Nama kelas dari suatu produk (Kategori Nominal)

# Tahapan Analisis

## Fokus pada Informasi Analisis

- Membuat tujuan analisis
- Membuat hipotesis analisis
- Menentukan software / tools analisis



## Eksplorasi Data Analisis

- Melakukan proses statistika deskriptif
- Melakukan proses analisis statistika infeerensia
- Membuat kesimpulan analisis



## Penyajian Data Analisis

- Menyusun hasil analisis secara sistematis
- Menyajikan hasil analisis dalam bentuk grafik
- Membuat dokumentasi hasil analisis

# Software Analisis



Microsoft Excel / Google  
Spreadsheet  
(Eksplorasi Data)



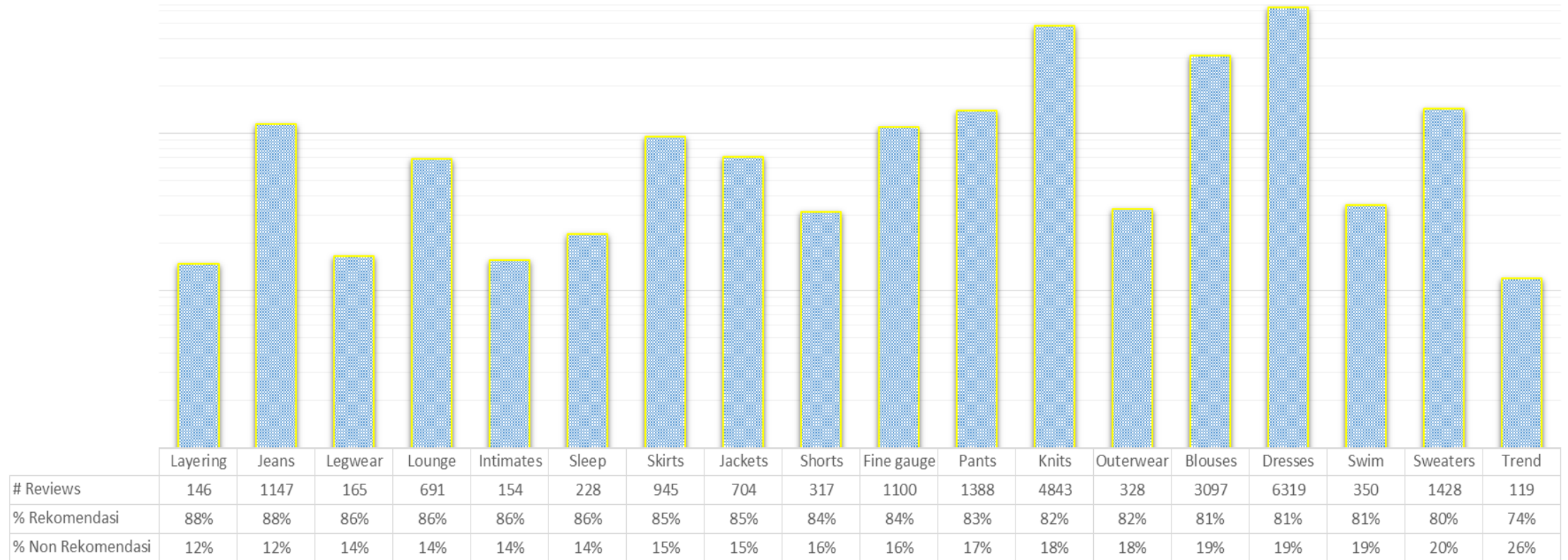
XL STAT  
(Data Modelling)



Microsoft Powerpoint  
(Penyajian Hasil Data)



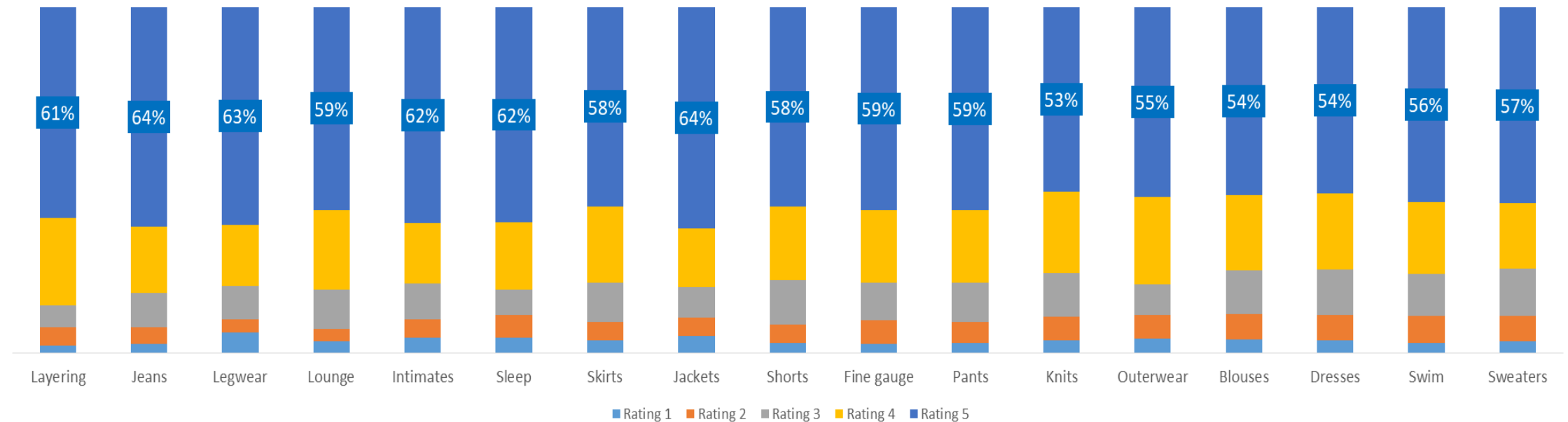
## Analisis v1



Penilaian pelanggan tidak menunjukkan perbedaan secara keseluruhan untuk merekomendasikan produk berdasarkan Class Product yang mencapai **80%** kecuali pada Class Product **Trend** yang mencapai **74%**. Penjelasan infomasi ini menyakinkan bahwa penilai memiliki karakteristik penilaian secara homogen

## Analisis v2

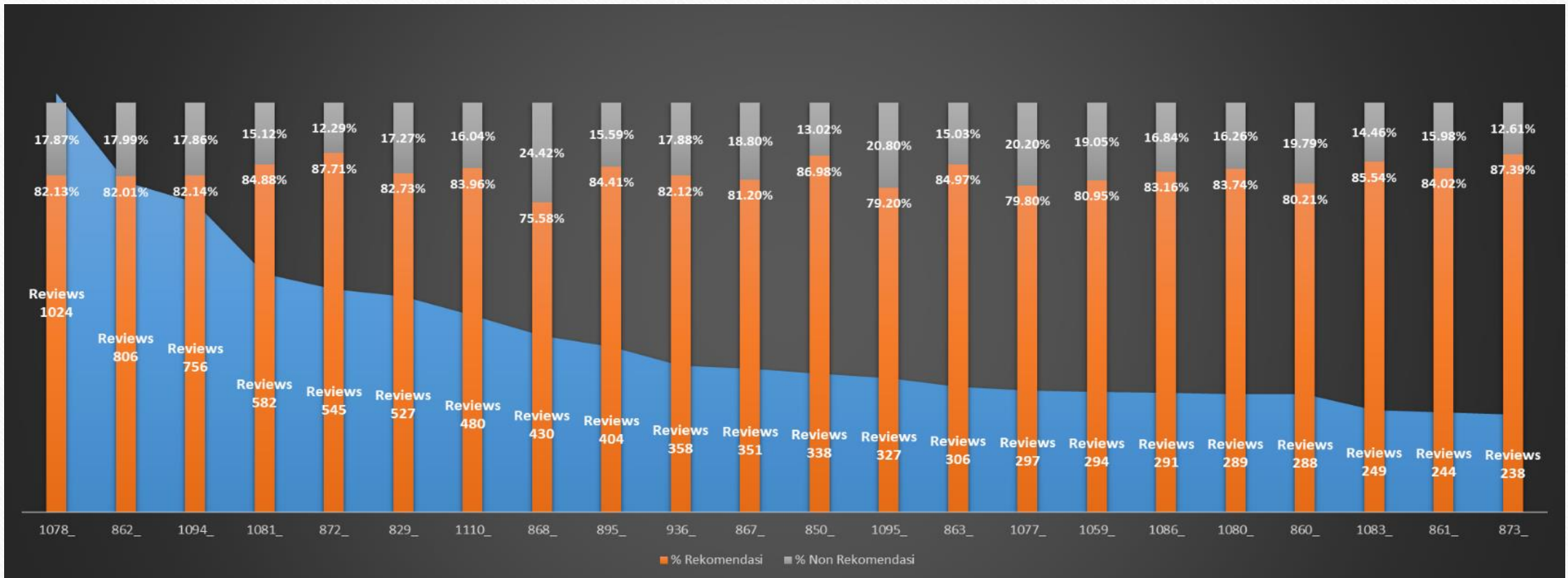
Proportion Rating of Product



Penilaian yang homogen dari keseluruhan penilai terpusat pada pemberian nilai **Rating 5** pada setiap produk berdasarkan Class Product melebihi **50%**. Informasi ini dapat kita jadikan sebagai standar dasar batas bawah dalam mengevaluasi produk



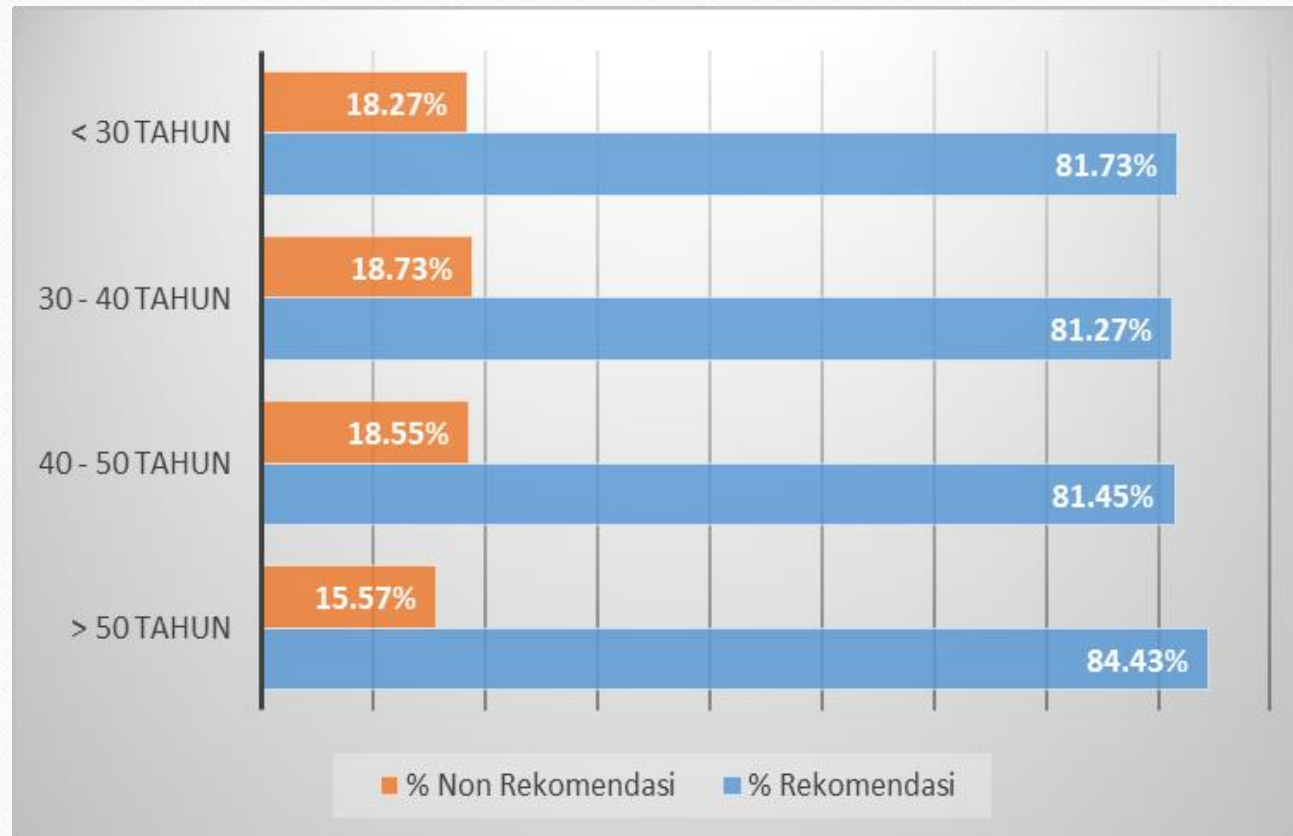
## Analisis v3



Secara subjektif diambil **22 produk** untuk dianalisis yang minimal memenuhi **1% dari total penilaian**. Informasi yang dapat disimpulkan bahwa tidak ada respon negatif / buruk dari penilaian yang dilakukan, hal ini dapat menjadi **masalah serius** bila penilai dari penilai yang memiliki **karakteristik homogen** ini berbanding terbalik terhadap penilaian sampel yang lebih luas lagi, maka diperlukan beberapa informasi tambahan untuk mendapatkan **informasi yang semakin detail** lagi

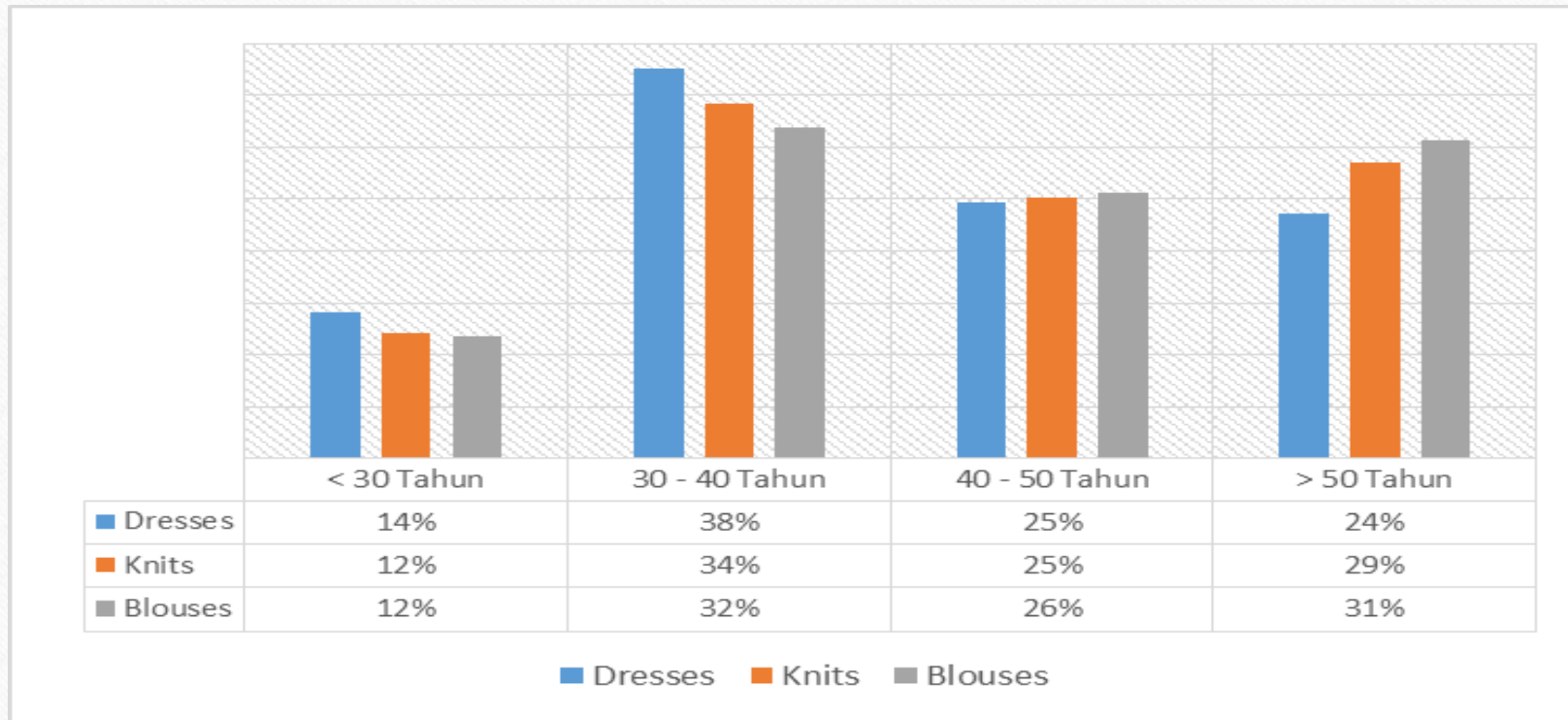


## Analisis v4



Kita mendapatkan informasi unik dari penilai berdasarkan kategori usia bahwa usia  $> 50$  tahun lebih **cenderung mudah** memberikan rekomendasi produk dan berbanding terbalik dengan usia  $< 30$  tahun yang lebih **selektif** dalam memberikan rekomendasi, sedangkan dalam 2 kategori usia 30 – 40 tahun dan usia 40 – 50 tahun memiliki karakteristik penilaian yang lebih **stabil**, maka dapat disimpulkan 2 kategori usia 30 – 40 tahun dan usia 40 – 50 tahun dapat dijadikan **standar validitas penilaian** dari segi usia

## Analisis v5



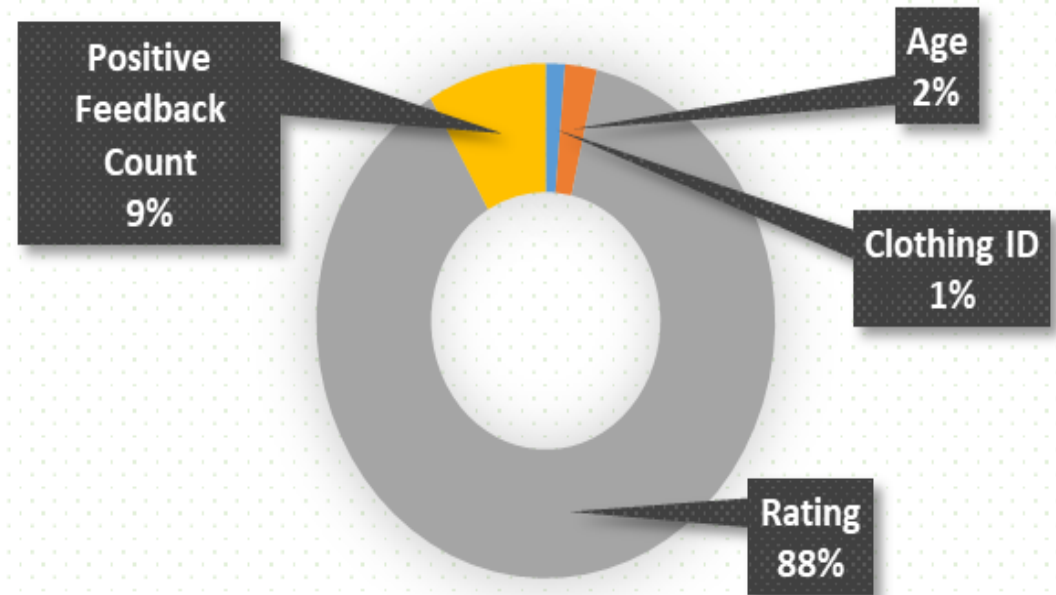
**3 top class product** dengan jumlah penilaian terbanyak yaitu Dresses, Knits, dan Blouses. Kategori usia < 30 Tahun dan usia 30 - 40 Tahun **memiliki karakteristik yang sama** yaitu lebih tertarik pada class product **Dresses** begitu juga pada kategori usia 40 – 50 Tahun dan usia > 50 Tahun yang lebih tertarik pada class product **Blouses**



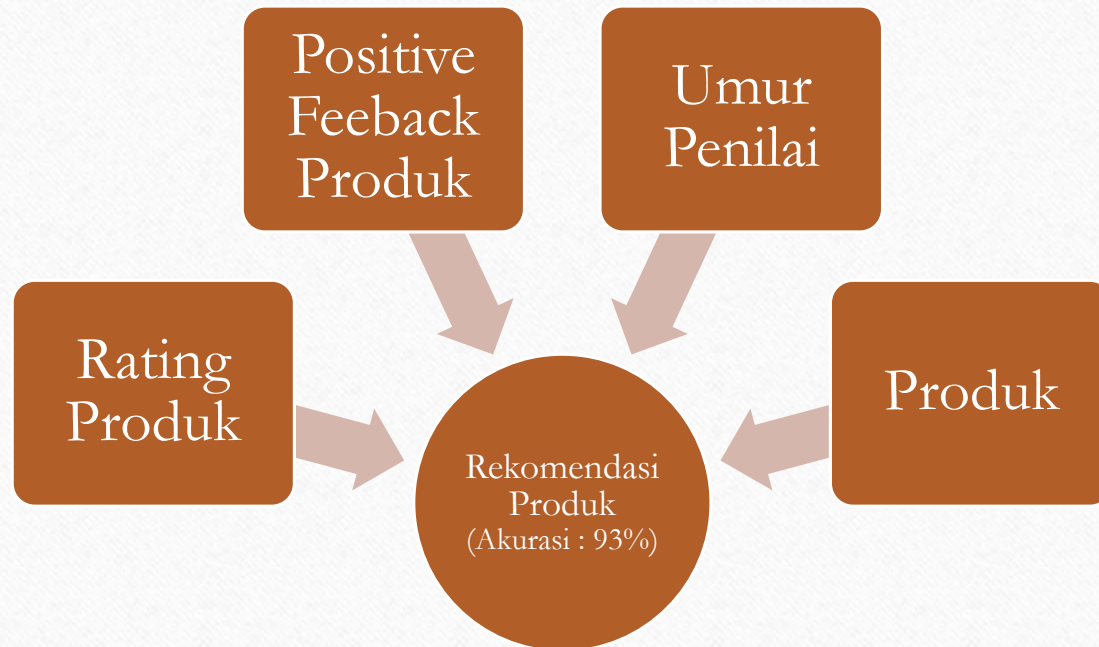
## Analisis v6

Penggunaan **Uji Korelasi** memberikan **informasi hubungan Pemberian Rekomendasi terhadap masing masing informasi terkait**. Dari data yang dimiliki menggambarkan bahwa pemberian Rekomendasi sangat dipengaruhi oleh informasi **Pemberian Rating**, kita dapat menyimpulkan bahwa **semakin tinggi rating** yang diberikan oleh penilai maka akan cenderung penilai akan memberikan rekomendasi atas produk tersebut

### Correlation Recommended Product



# Statistika Modelling



*Pred(Recommended IND*

$$\begin{aligned} &= 1 / (1 + \exp (-( -11.7672809304593 - 2.85425502940891 * \textit{Clothing ID} \\ &+ 1.03549575931649 * \textit{Age}_p + 23.8248168025808 * \textit{Rating} \\ &+ 0.516306577302527 * \textit{Positive Feedback Coun} \\ &- 0.132601845441967 * \textit{Class Name}))) \end{aligned}$$

Model prediksi ini dapat digunakan sebagai panduan kita dalam **mengkoreksi penilai dalam merekomendasikan produk**. Model ini dapat membantu kita mendapatkan informasi dari kombinasi total informasi terkait pemberian Rekomendasi produk



Terima kasih

---

Sangat Mengharapkan  
Komentar dan Saran