**TUGAS   
PROYEK SAINS DATA**

**Analisis Dataset Stroke menggunakan Power Bi**

A picture containing text, sign

Description automatically generated

**Dosen pengampu :**

***EKA MALA SARI ROCHMAN, S.Kom.M.Kom***

**Disusun oleh :**

Agus Muhammad Al Hasib (210411100227)

**PROGRAM STUDI TEKNIK INFORMATIKA**

**FAKULTAS TEKNIK**

**UNIVERSITAS TRUNOJOYO MADURA**

**2023**

**Chirosis Prediction**

Data yang saya peroleh dari kaggle yaitu data sekunder yang berjudul “Cirrhosis Prediction Dataset” terdapat sejumlah data 418 data dan variable atribut sebanyak 20, target dari variable sendiri yaitu untuk memprediksi gejala pasien sirosis.

Setelah mendapatkan data kemudian dilakukan Langkah diantaranya yaitu

1. Cleaning data

Yaitu cara dilakukan pengecekan data yang hilang maupun data yang NA(Tidak ada).

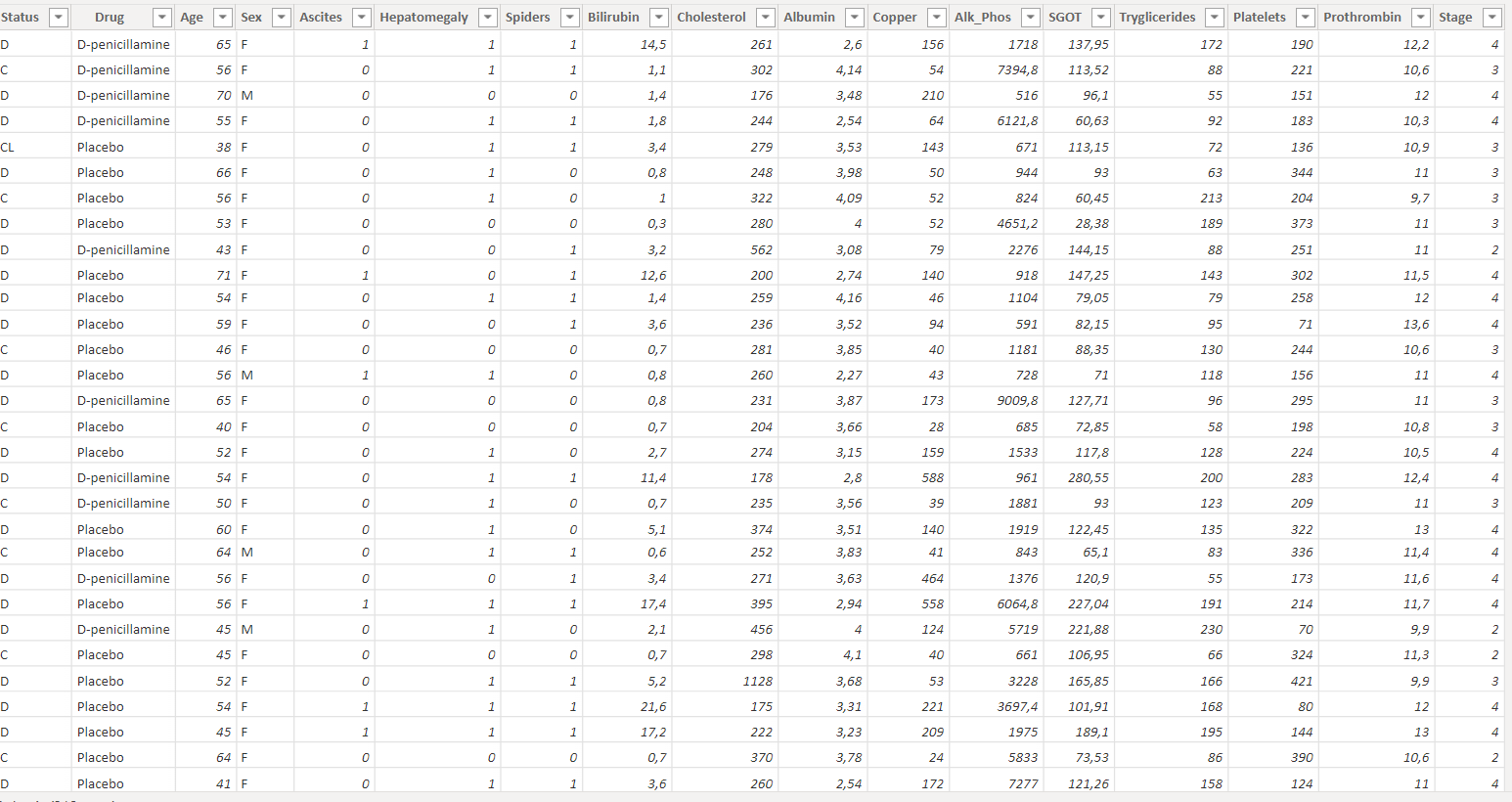
1. Transformasi data

Yaitu cara bagaimana data akan dilakukan pemberian label pada data yang nilainya bertipe atau berbentuk data numerik menjadi kategorik.

1. Balance data

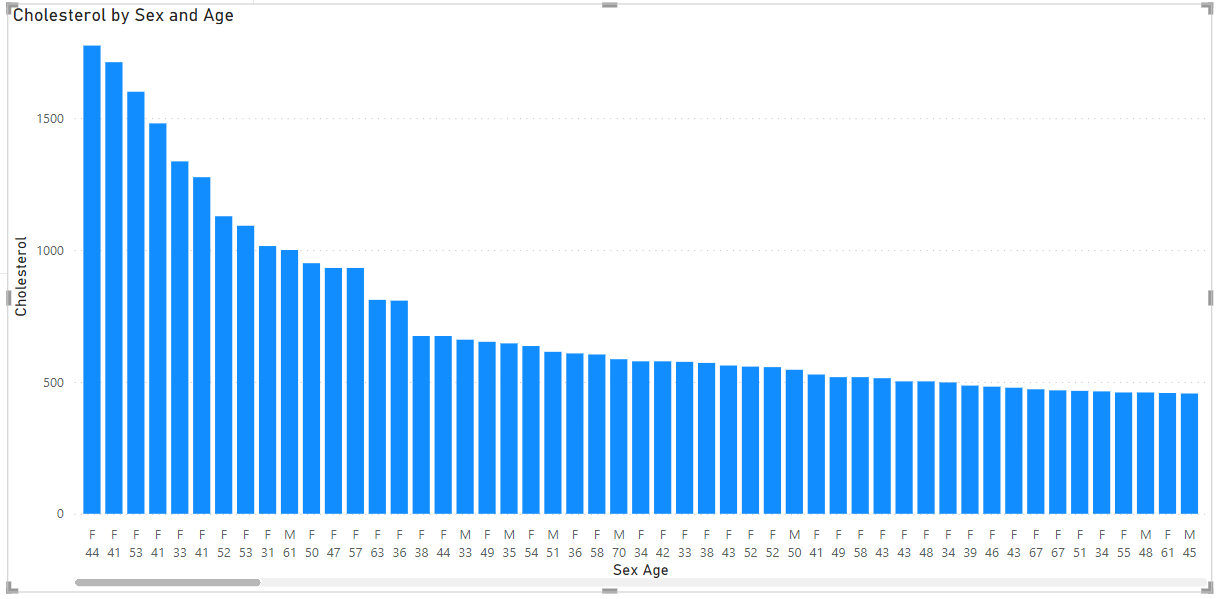
Yaitu cara untuk mengubah distribusi atau karakteristik data awal agar data tersbut bisa seimbang atau memiliki distribusi yang lebih mendekati normal

Dibawah ini adalah data yang telah melalui proses pre-processing data:



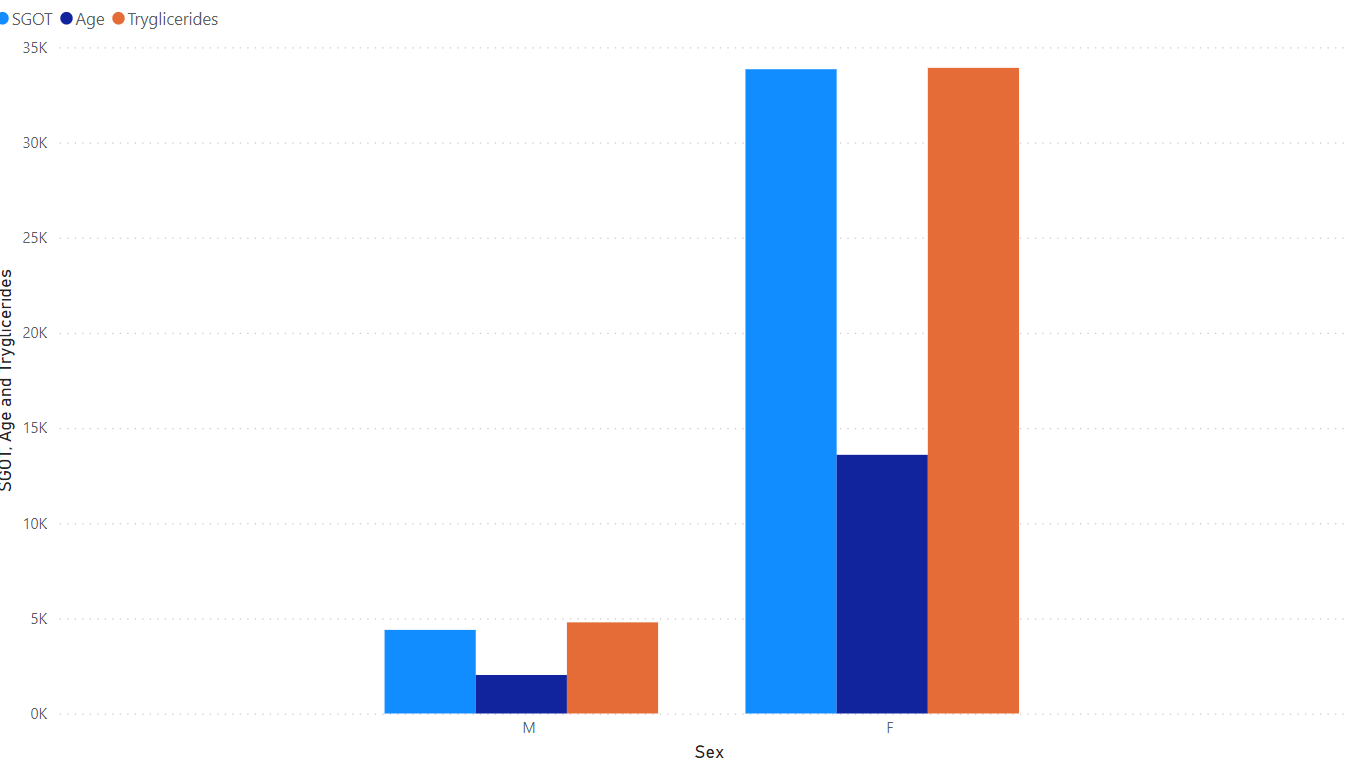
**Analisis pada aplikasi power BI**

1. Analisis berdasarkan umur, jenis kelamin dan colestrol



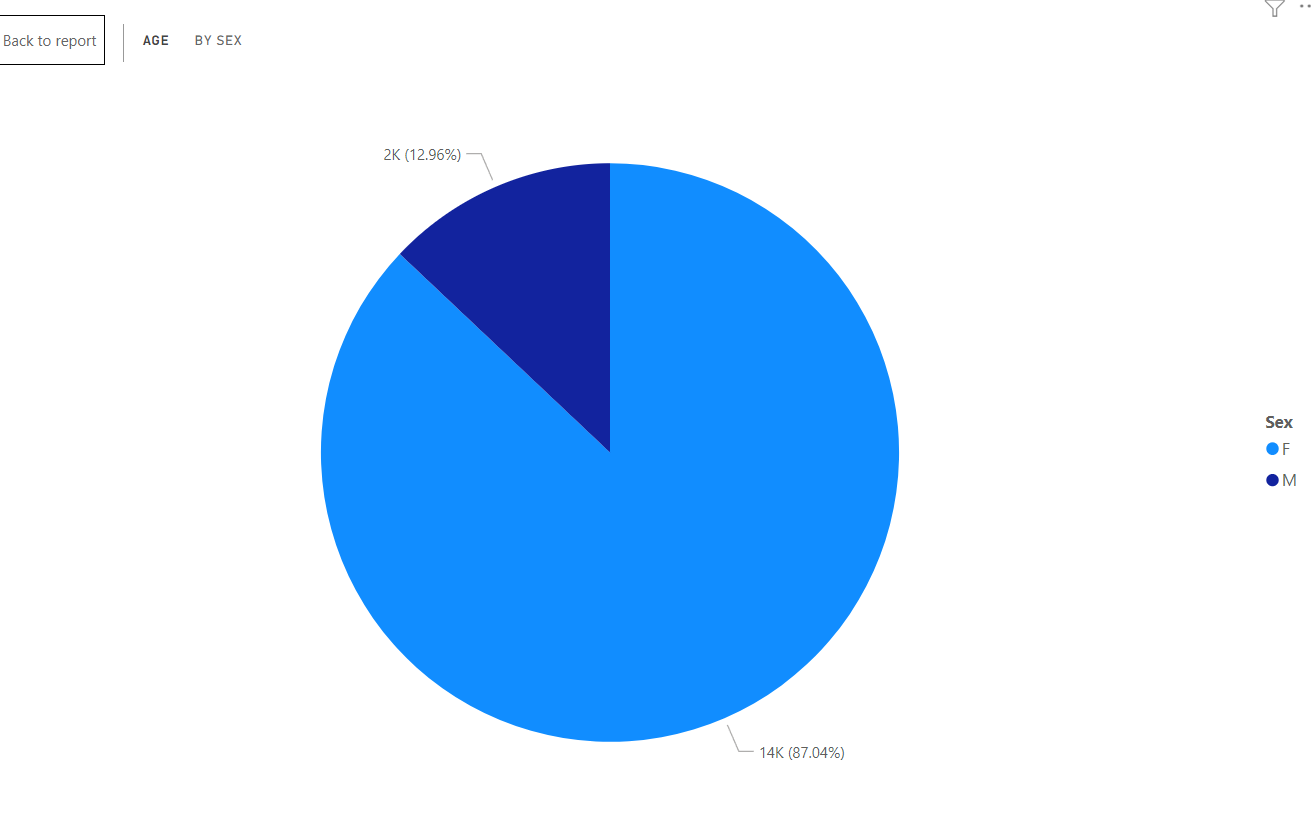
Dari data diatas menjelaskan bahwa rata-rata tertinggi yaitu pada usia 44 tahun dengan memiliki nilai colestrol paling tinggi yaitu 1500 dan yang paling berstatus yaitu Perempuan.

1. Analisis berdasarkan umur, jenis kelamin,SGOT, Tryglicerides



Dari diagram diatas menjelaskan bahwa Perempuan tinggi dari gejala SGOT, Tryglicerides.

1. Analisa berdasarkan umur dan jenis kelamin.



2. Cara Mengatasi Missing Value

missing value disebabkan karena informasi tentang objek tidak diberikan, sulit dicari, atau memang informasi tersebut tidak ada, hal ini akan menyebabkan menurunnya keakuratan dan kualitas data pada saat data diolah. Berikut beberapa cara untuk mengatasi missing value :

1. Mengisi dengan nilai 1 atau 0
2. Mengisi dengan nilai mean atau median
3. Menggunakan machine learning
4. Balancing data