**TUGAS   
PROYEK SAINS DATA**

**Analisis Dataset Stroke menggunakan Power Bi**



**Dosen pengampu :**

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

**Disusun oleh :**

Adz Dzikry Pradana Putra (210411100093)

**PROGRAM STUDI TEKNIK INFORMATIKA**

**FAKULTAS TEKNIK**

**UNIVERSITAS TRUNOJOYO MADURA**

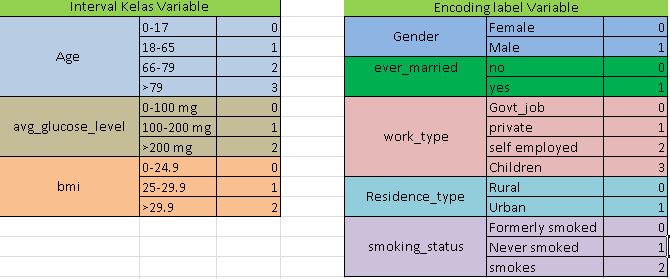
**2023**

**Stroke Prediction**

Data yang digunakan adalah data sekunder yang diperoleh dari situs kaggle yang berjudul “Stroke Prediction Dataset” dengan sejumlah data sebanyak kurang lebih 3000 data dan variable sebanyak 11, dengan 10 variable atribut dan 1 variable target dari 10 variable ini 7 bertipe data kategorik dan 3 bertipe data numerik .Variabel targetnya yaitu memprediksi pasien terkena stroke dan tidak terkena stroke.

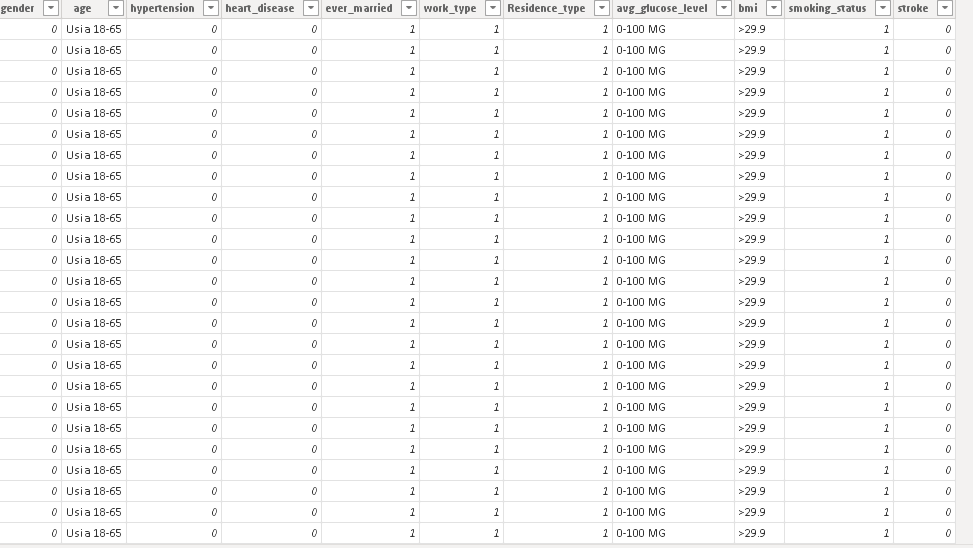
Setelah itu akan dilakukan beberapa langkah antara lain :

* Pembersihan data dimana dilakukan pengecekan data hilang dan data duplikat
* tranformasi data dimana dilakukan pemberian label pada data bertipe kategorik yang berbentuk string, dan juga mengubah bentuk data numerik menjadi kategorik dengan interval.



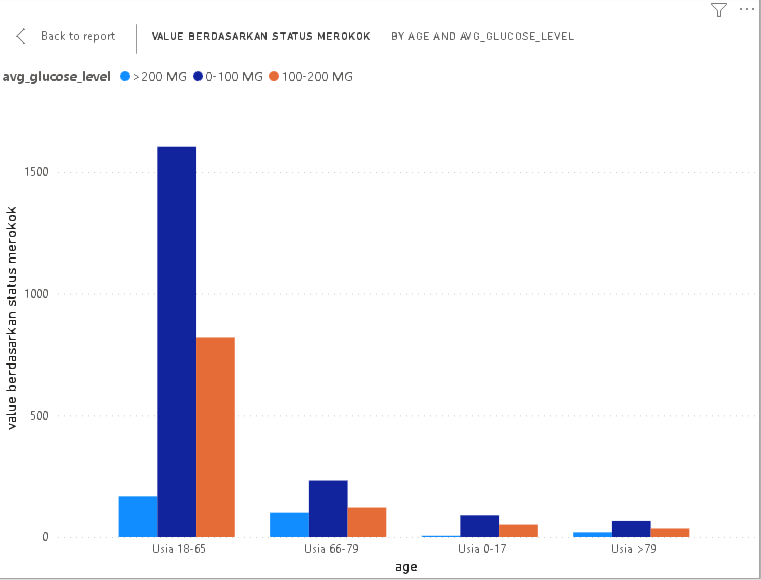
*Hasil setelah transformasi data berdasarkan jurnal dan sesuai ketentuan WHO(World Health Organization)*

Berikut adalah data setelah melalui proses pre-processing :



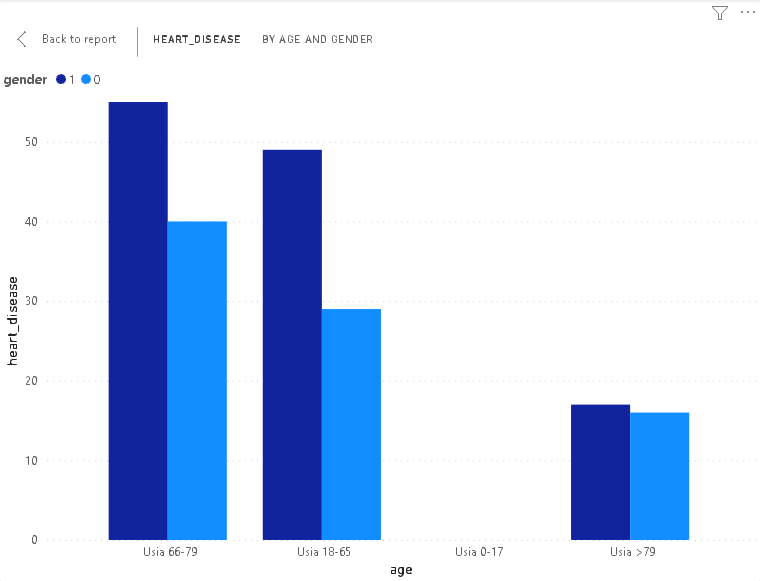
**Analisis Parameter Pada Aplikasi Power BI**

* Analisis berdasarkan umur,level glukosa dan status merokok sebagai value



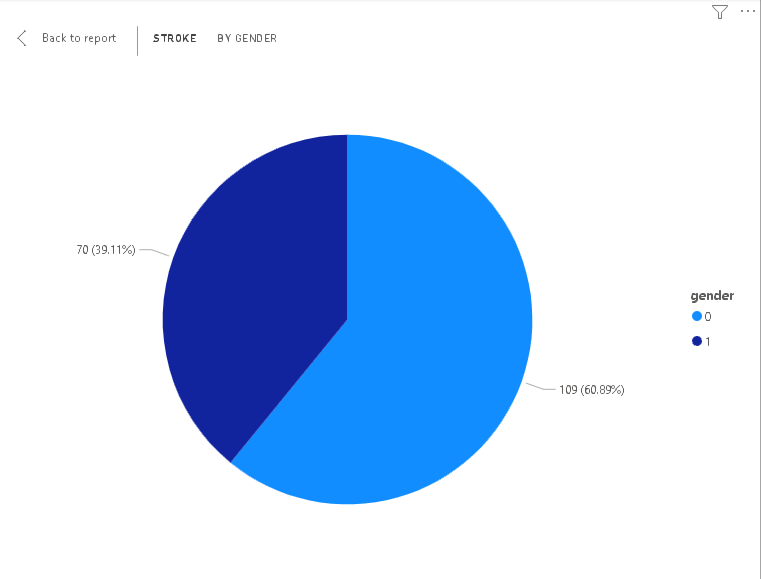
Dari diagram visual tersebut dapat menjelaskan bahwa rata-rata tertinggi pada usia 18-65 dengan memiliki level glukosa 0-100 mg terdapat 1604 data yang berstatus merokok .

* Analasis berdasarkan umur,gender dan penyakit jantung sebagai value



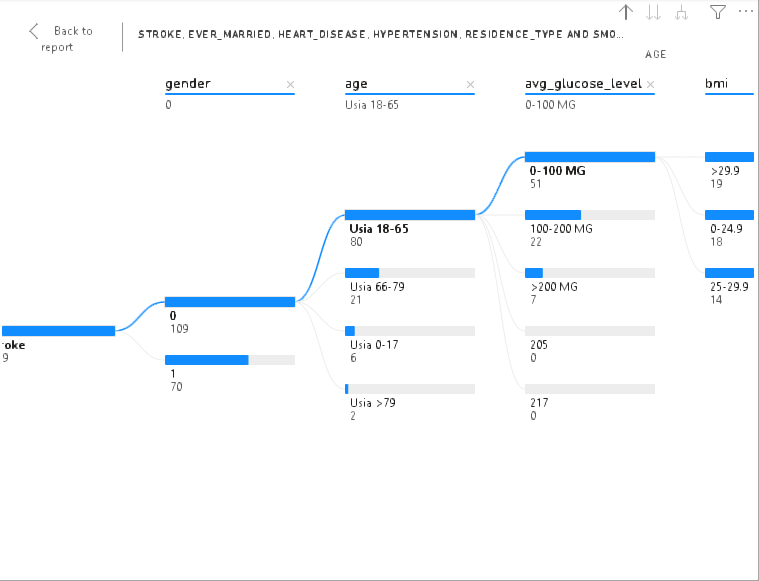
Dari diagram visual tersebut dapat menjelaskan bahwa pada rata-rata usia 66-79 dengan bergender perempuan terdapat 55 data yang mengalami penyakit jantung

* Analisis berdasarkan gender dan stroke sebagai value



Dari kedua parameter tersebut disimpulkan bahwa yang bergender perempuan menghasilkan 60.89% stroke dan bergender laki-laki menghasilkan 39.11%.

* Analisis berdasarkan decomposition tree berdasarkan semua paramater



Bahwa stroke paling banyak dialami oleh gender 0 (Female) yang berusia 18-65 dan memiliki level glukosa 0-100mg dan memiliki msi >29.9 berdasarkan dataset dan diagram diatas