**Tugas Besar 2 Visualisasi Data**

**Nama: Timothy Kurnia**

**NIM: 41819010076**

**Soal:**

1. Dalam memuat dan menganalisa data alat atau library apa yang Anda gunakan. Berikan keuntungan dan kerugian menggunakan library tersebut.
2. Tuliskan syntax sederhana yang memanggil library memuat data dan syntax yang memuat data tersebut.
3. Jelaskan dengan menggunakan bahasa Anda sendiri apakah yang dimaksud dengan konsep kaidah Exploratory Data Analysis (EDA), dan kaidah filosofi dari Exploratory Data Analysis (EDA).
4. Jelaskan dengan menggunakan bahasa Anda sendiri apakah yang dimaksud dengan konsep
5. Univariate
6. Bivariate
7. Multivariate
8. Dalam konsep pemilihan warna pada grafik. Jelaskan konsep dasar pemilihan desain warna.

**Jawaban:**

1. Library yang digunakan untuk menganalisis data, salah satunya adalah library Numpy, yang dimana library ini merupakan singkatan dari Numerical Python yang merupakan library yang cukup banyak digunakan dalam proses analisis data. Numpy dapat digunakan untuk operasi vektor dan matriks, serta dapat mengelola array. Salah satu kelebihan dari library ini adalah dapat terintegrasi dengan beberapa bahasa pemrograman lainnya. Dan kekurangan dari library ini adalah, bentuk array yang harus homogen (memiliki tipe data yang sama), dan tidak memungkinkan adanya representasi, dan membutuhkan banyak waktu untuk melakukan komputasi.
2. Tuliskan syntax sederhana yang memanggil library memuat data dan syntax yang memuat data tersebut.

Jawaban:

Import numpy as np

theList=[100,200,300]

theArray=np.array[100,200,300])

thelabel=[‘Pertama’, ‘Kedua’, ‘Ketiga’]

theDict= {“Pertama”:100, “Kedua”:200, “Ketiga”:300}

theListSeries= pd.Series(data=theList, index=theLabel)

theArraySeries= pd.Series(theArray, theLabel)

theDictSeries= pd.Series(theDict)

print(theListSeries)

1. Exploratory Data Anlysis adalah bagian dari proses data sience. Yang dimana EDA menjadi sangat penting sebelum melakukan feature engineering dan modelling, karena dalam tahap ini kita harus memahami datanya terlebih dahulu. Exploratory Data Analysis memungkinkan analyst memahami isi data yang digunakan, mulai dari frekuensi, korelasi dan lainnya. Pada umumnya EDA dilakukan dengan beberapa cara:
2. Univariat Analysis, yaitu analisis deksriptif suatu variabel
3. Bivariat Analysis, yaitu analisis relasi dengan dua variabel yang biasanya dengan target variabel
4. Multivariat Analysis, yaitu analisis yang menggunakan lebih dari atau sama dengan tiga variabel
5. Jelaskan dengan menggunakan bahasa Anda sendiri apakah yang dimaksud dengan konsep

Jawaban:

1. Univariate

Unvariate adalah proses analisis data, yang dimana analisis deskriptif yang dilakukan adalah menggunakan satu variabel

1. Bivariate

Bivariate adalah proses analisis data, yang dimana analisis yang dilakukan analisis relasi dengan dua variabel yang biasanya dengan target variabel

1. Multivariate

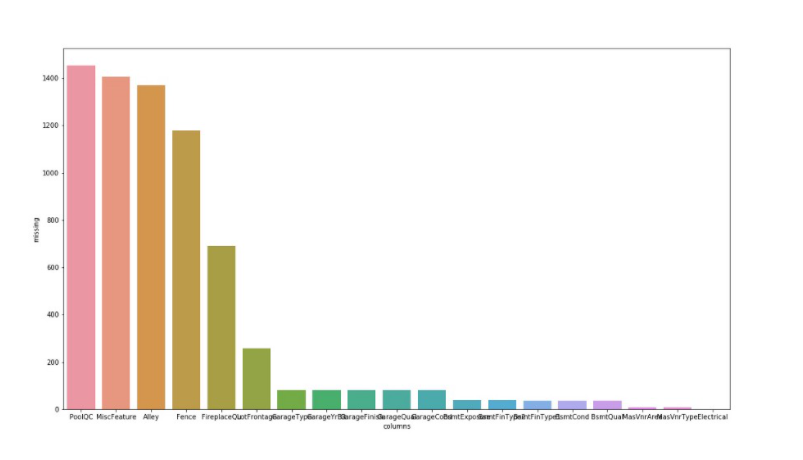
Multivariate adalah analisis data deskriptif, yang dimana pengolahan data dilakukan dengan mulativariabel atau menggunakan lebih dari atau sama dengan 3 variabel yang diolah.

1. Dalam konsep pemilihan warna pada grafik. Jelaskan konsep dasar pemilihan desain warna.

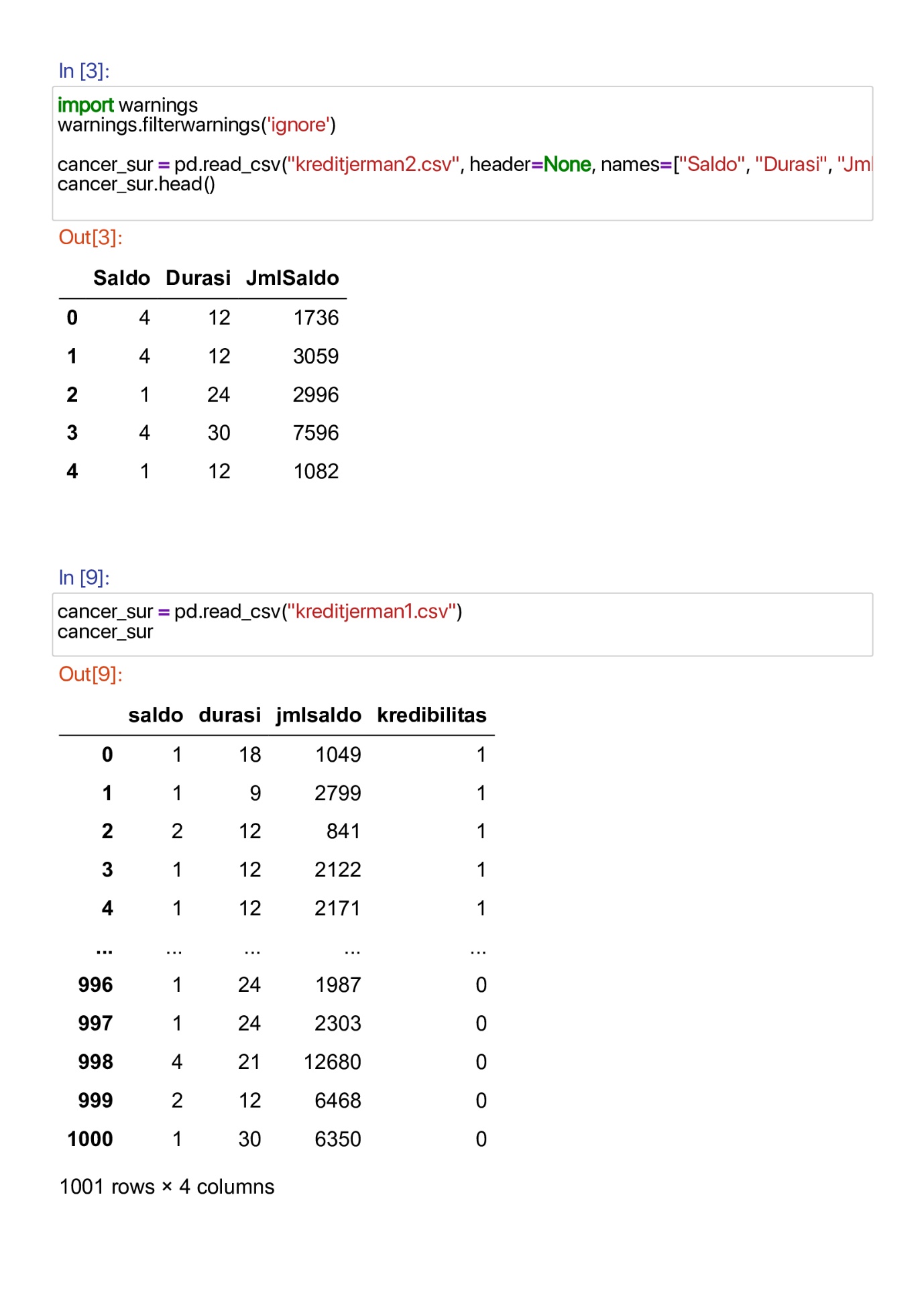
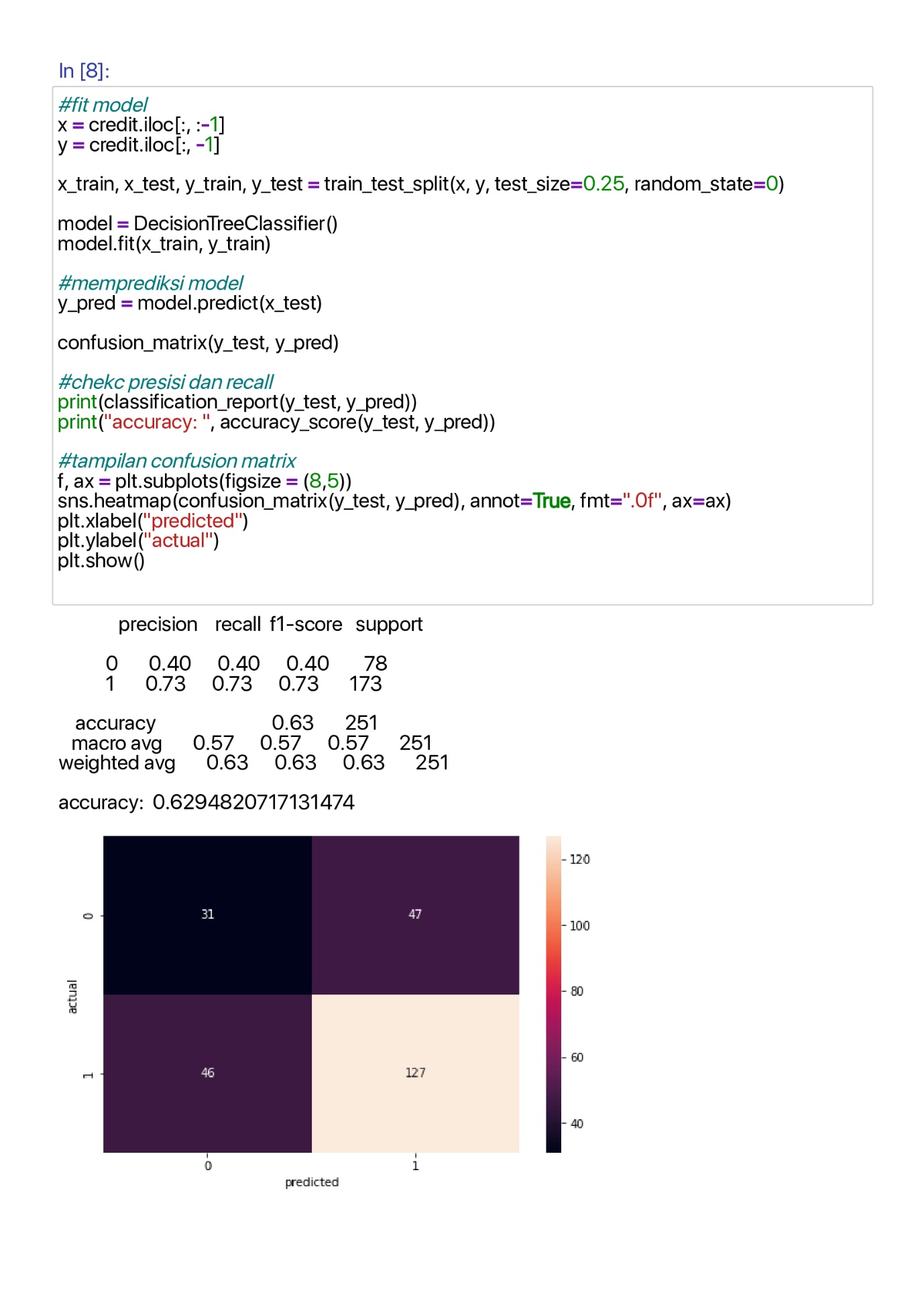
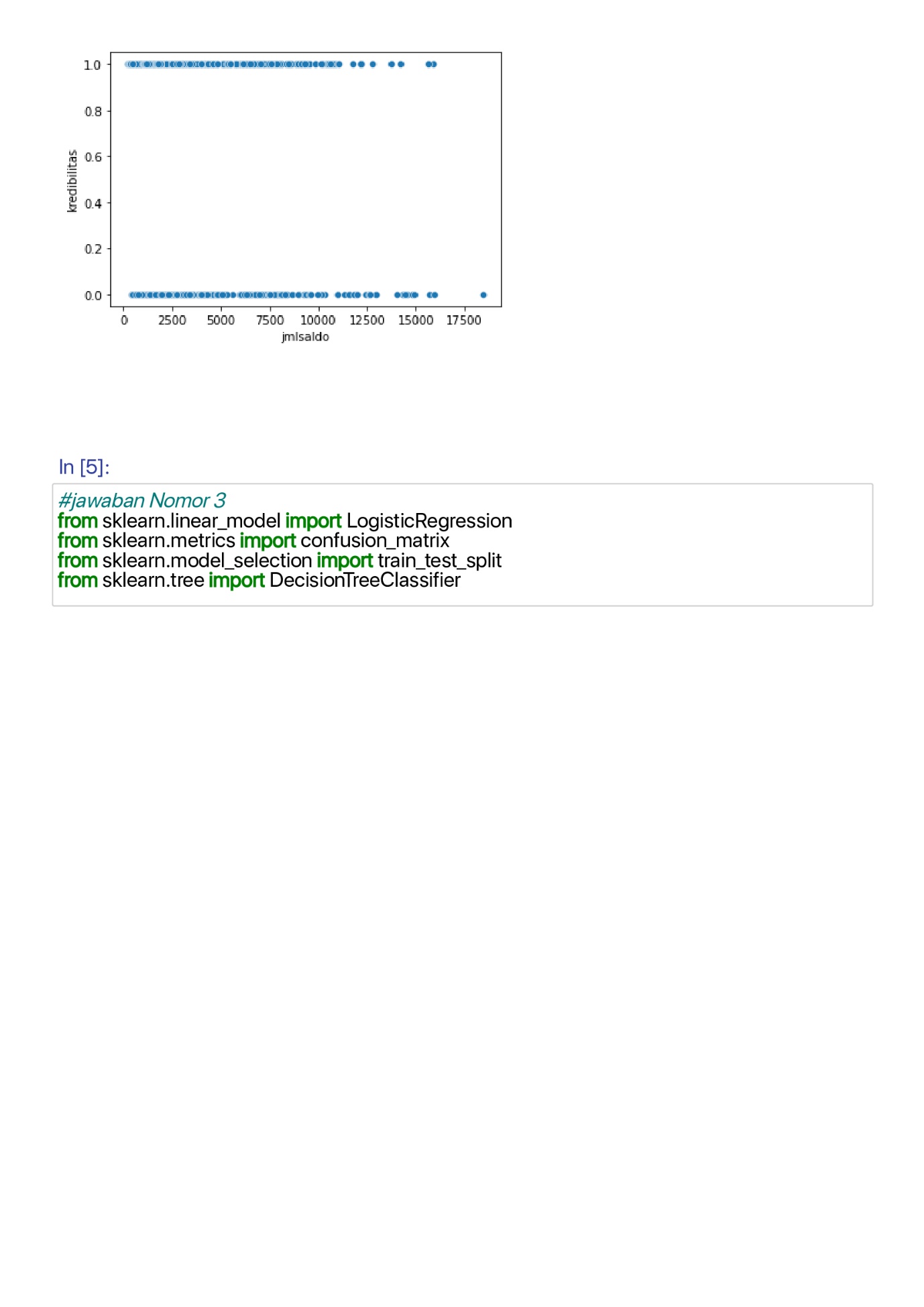
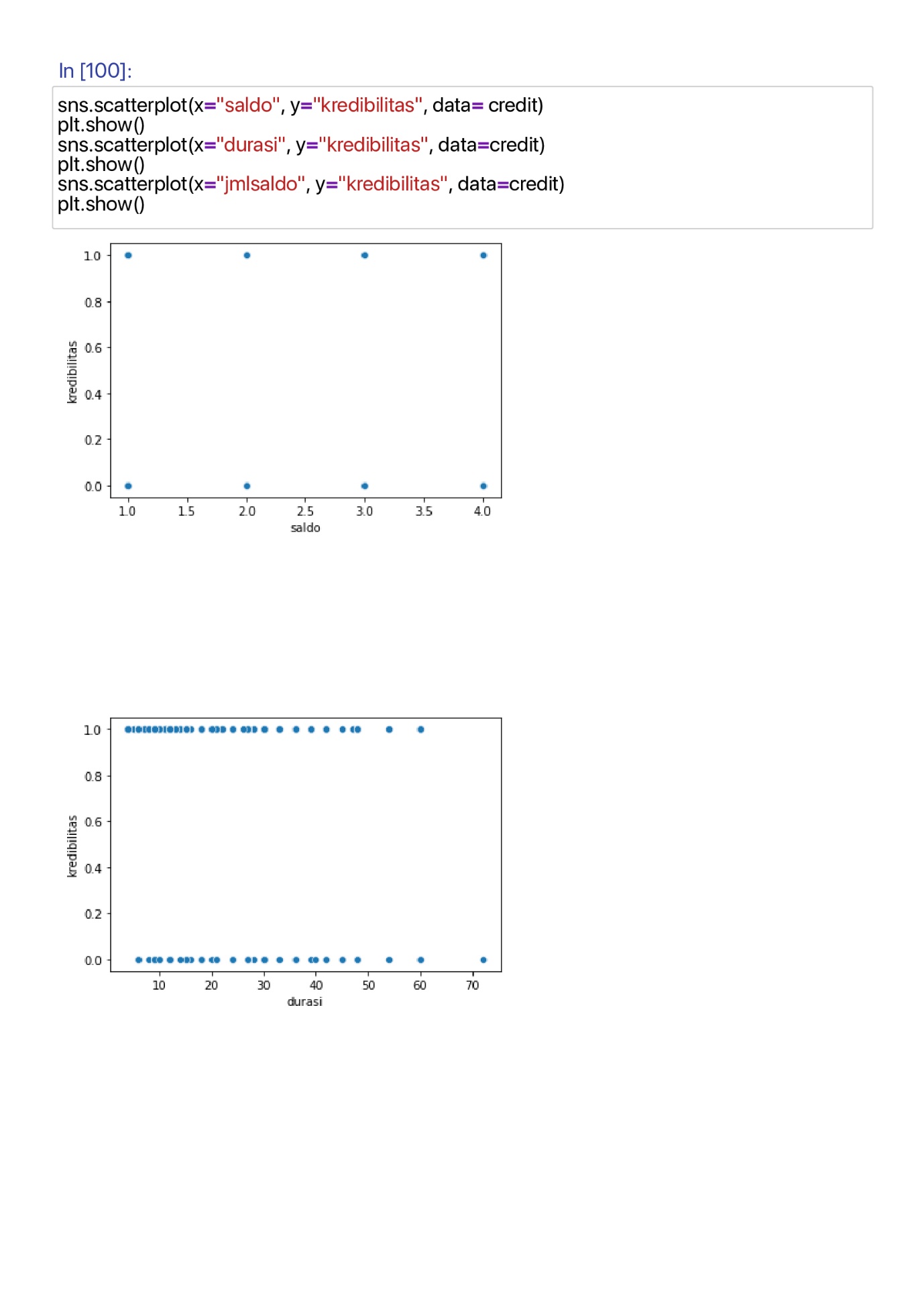
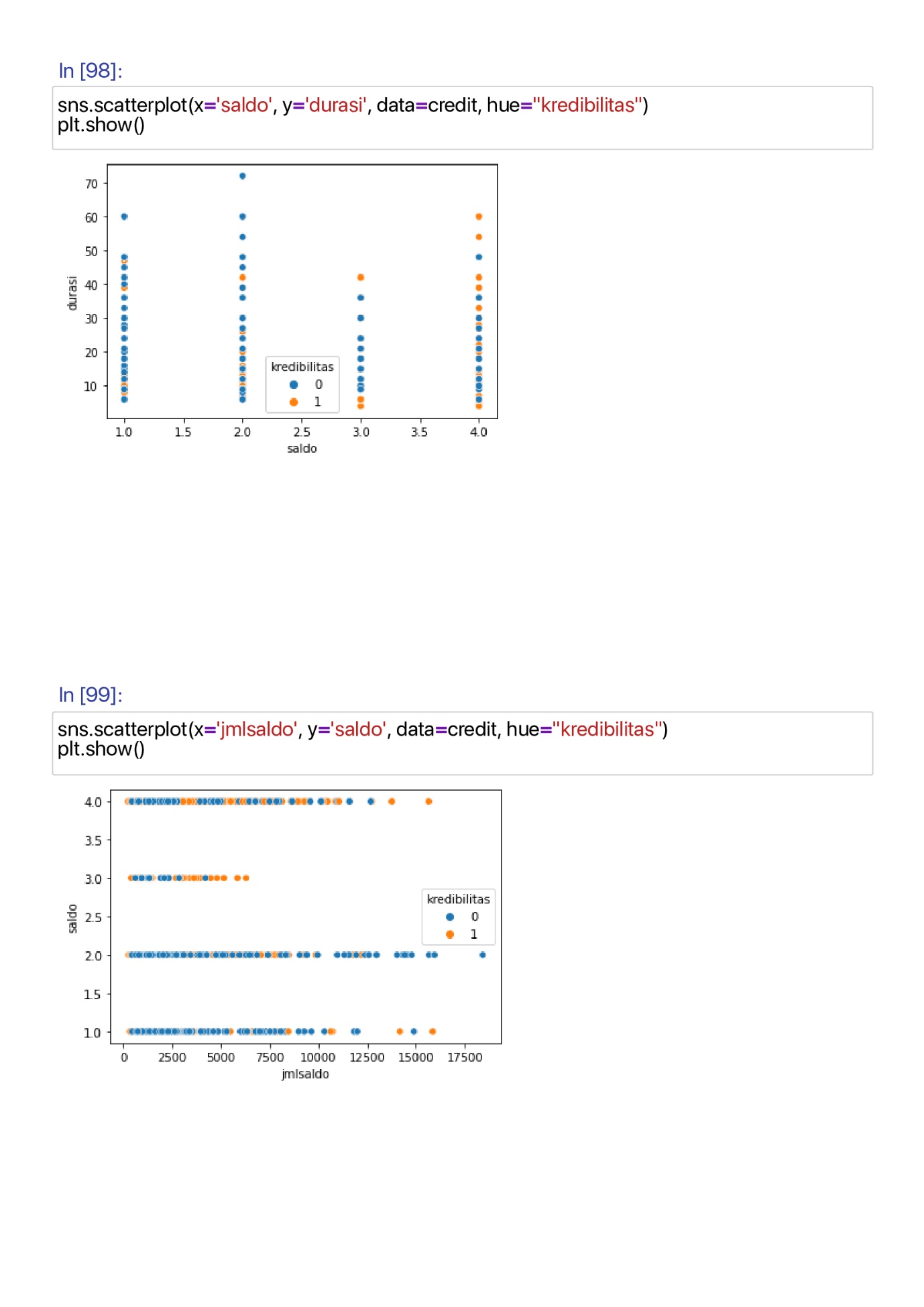
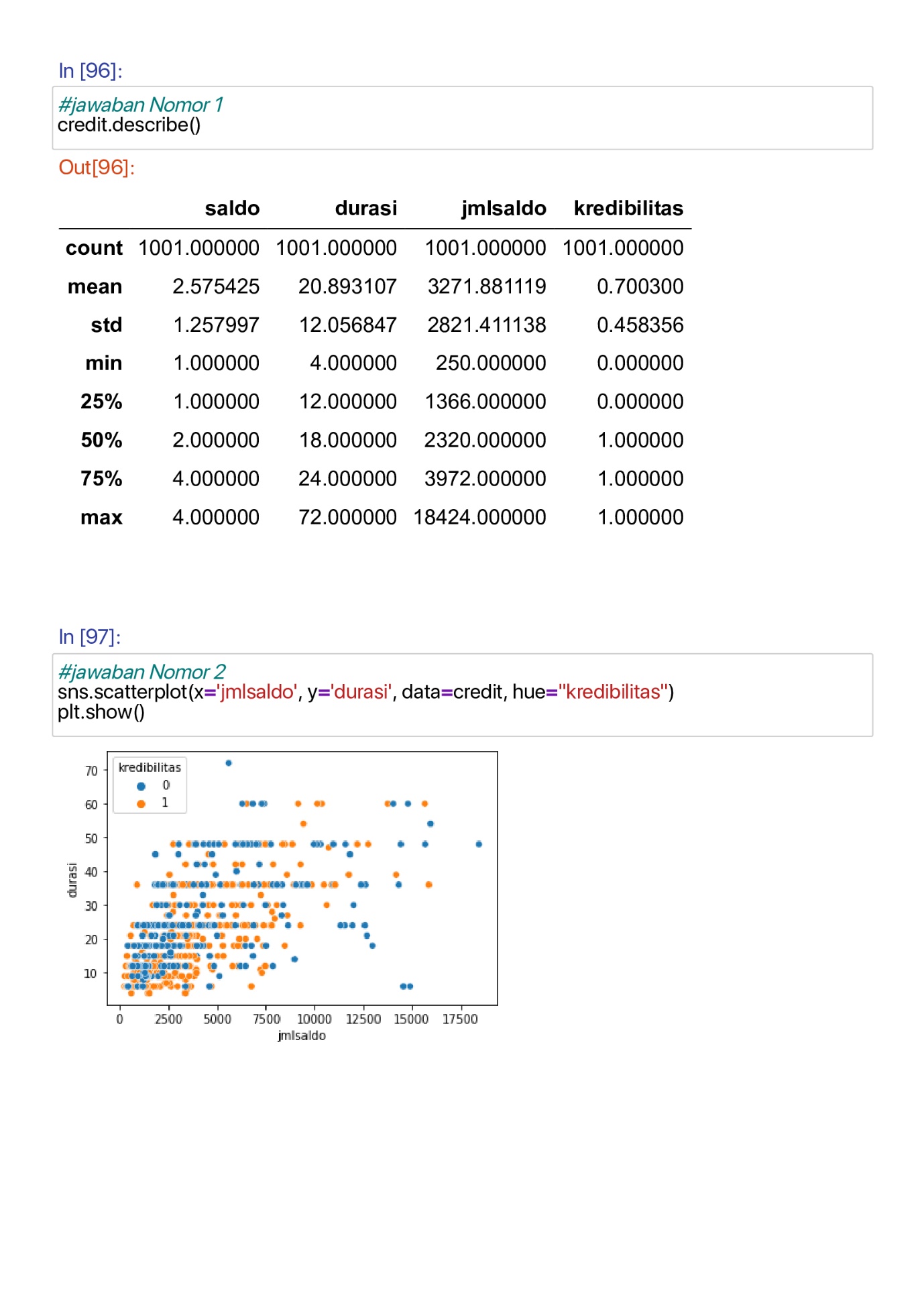
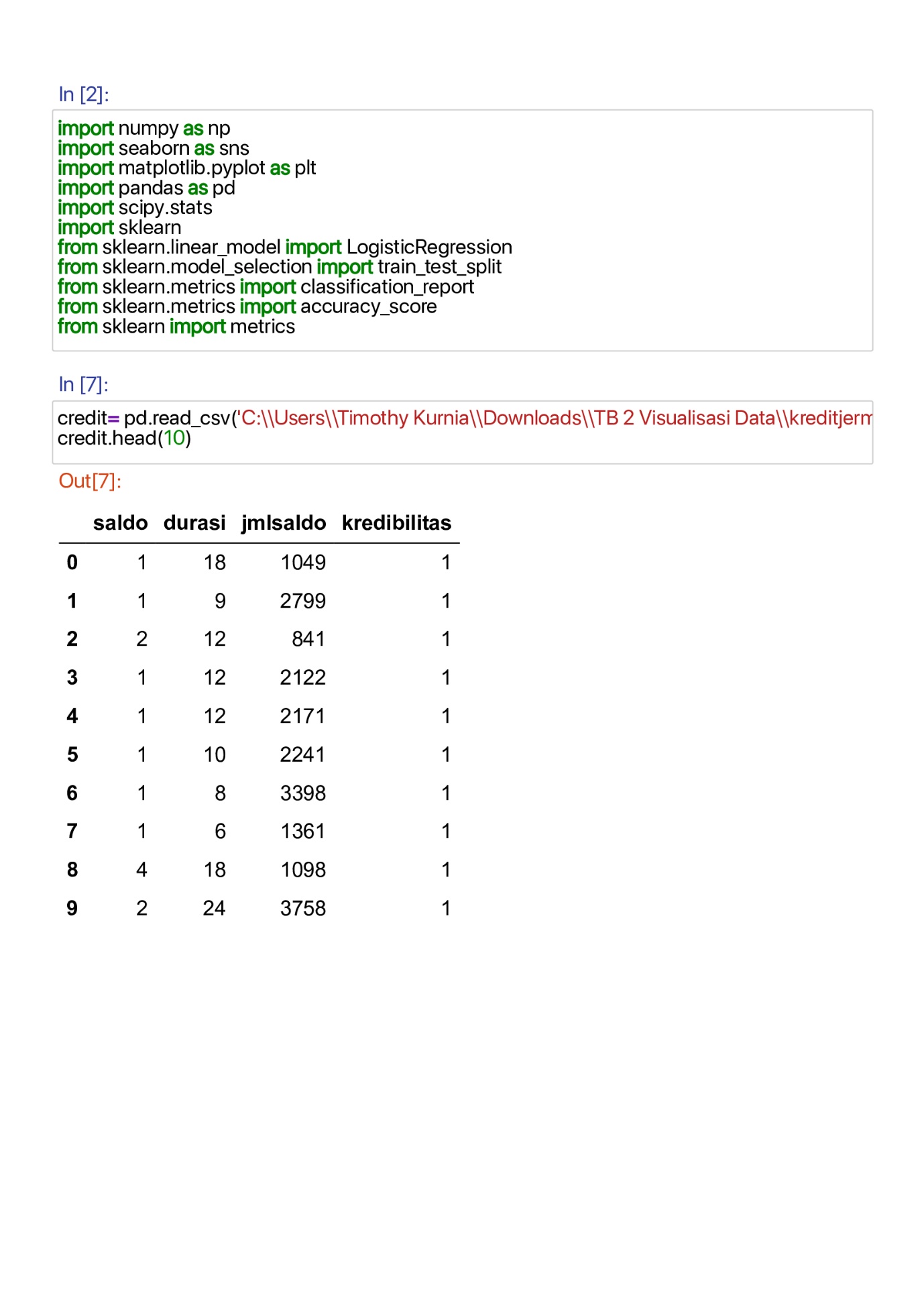
Jawaban:

Konsep pemilihan warna pada desain grafik yaitu, pemilihan warna yang dimana dapat menjelaskan perbedaan data yang jelas, contohnya pemilihan warna pada grafik yang dimana dengan data yang beragam, maka pemilihan warna pada grafik juga harus tepat, sehingga pembacaan grafik dapat menjadi lebih mudah, dengan adanya pemilihan warna yang tepat

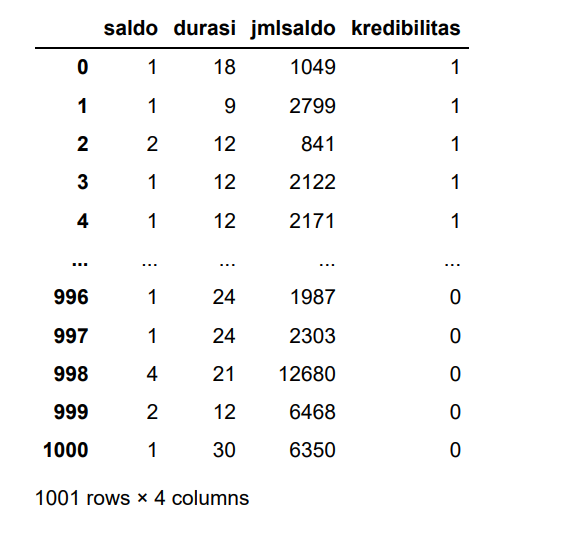


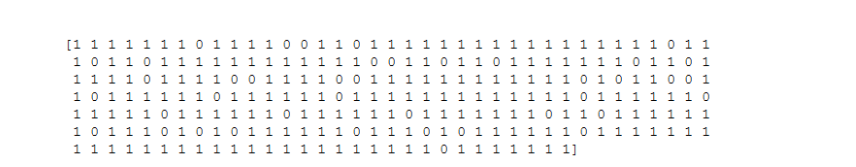


**Jawaban Soal Essai Menggunakan Python Menggunakan dataset Kreditjerman1 dan KreditJerman2 .csv**



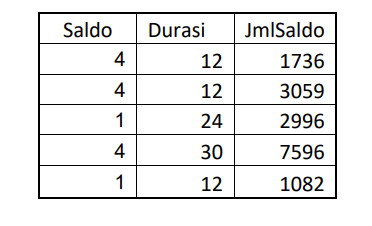
1. Buatlah plot data seperti tabel berikut: Anlisalah hasil di atas berikan penjelasan singkat dengan menggunakan Bahasa Anda sendiri.





Menurut saya, dari analisis data diatas, didapatkan hasil matriks biner seperti diatas, dan juga didapatkan deksripsi data dengan saldo, durasi dan jumlah kreditbilitas seperti diatas.

1. Dengan menggunakan model yang telah Anda bentuk pada soal e. Maka prediksilah apakah data kredit 5 nasabah baru berikut termasuk yang memiliki kreditbilitas.



Menurut saya, prediksi data kredit 5 nasabah baru seperti diatas, ada data yang mendapatkan kreditbilitas, dan ada data juga yang tidak mendapatkan kreditbilitas, dikarenakan dalam pengolahan data pada bagian e, ada data yang mendapatkan kredibilitas, dengan saldo dan durasi yang sama dan jumlah saldo yang berbeda, dan terdapat kredibilitas, namun juga ada data yang memiliki saldo dan durasi yang sama, namun tidak memiliki kreditbilitas, maka dapat disimpulkan, jika ada Sebagian data yang akan mendapatkan kreditbilitas, dan ada juga data yang tidak mendapatkan kreditbilitas, pada data 5 nasabah baru diatas.