**Laporan Tugas Besar**

**Pembelajaran Mesin**

Oleh : Aldiyan Farhan N – 1301180344

Link Video Presentasi :

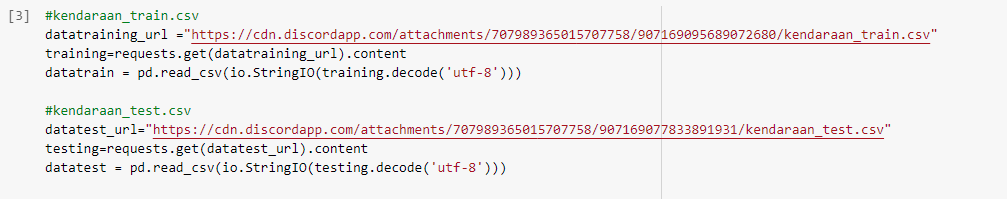
https://youtu.be/elJHlqG0c0s

1. **Permasalahan**

Diberikannya data pembelian kendaraan untuk mengelompokkan pelanggan berdasarkan data pelanggan di dealer tanpa memperhatikan label kelas apakah pelanggan tertarik untuk membeli kendaraan baru atau tidak. Permasalahan untuk tugas besar 1 ini adalah bagaimana data yang diberikan dapat di clusterisasi.

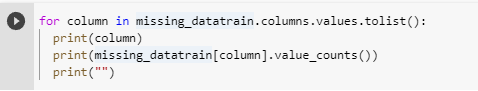
1. **Explorasi dan Persiapan Data**
   1. Memasukan File

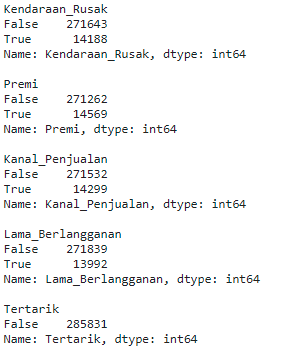
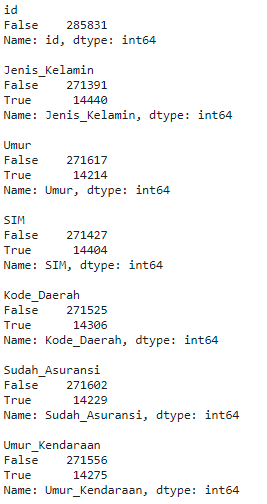
Memasukan data dengan melakukan pembanggilan URL yang memuat data set yang akan di explorasi :



* 1. Pengecekan Missing Value

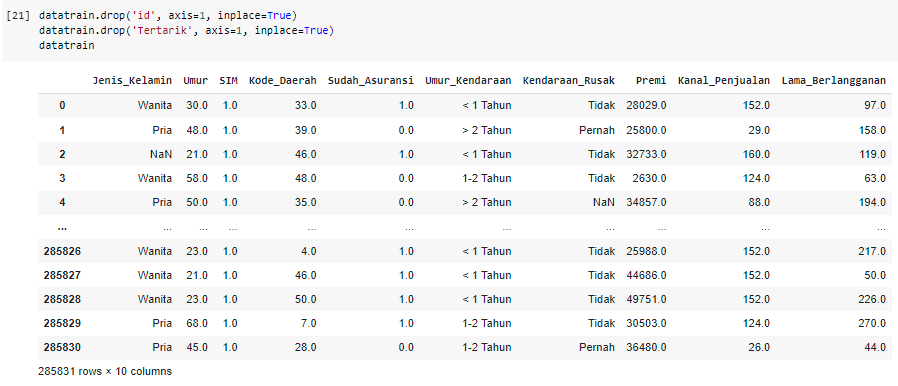
Melakukan pengecekan untuk adanya value yang bernilai kosong/null. Data yang bernilai kosong/null akan mempengaruki pemrosesan data yang menyebabkan pemrosesan tidak berjalan optimal





* 1. Drop Coloum

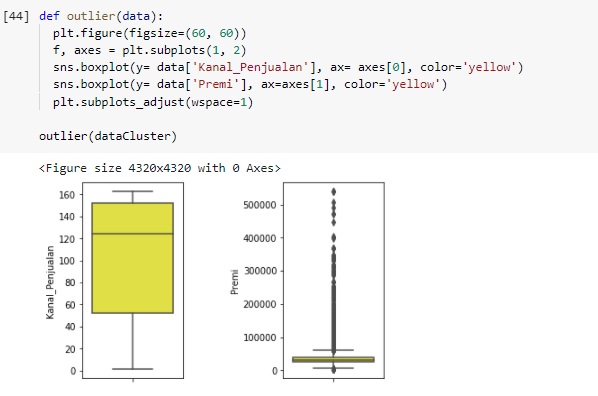
Dilakukan untuk menghapus kolom yang tidak digunakan untuk menghilangkan redundan





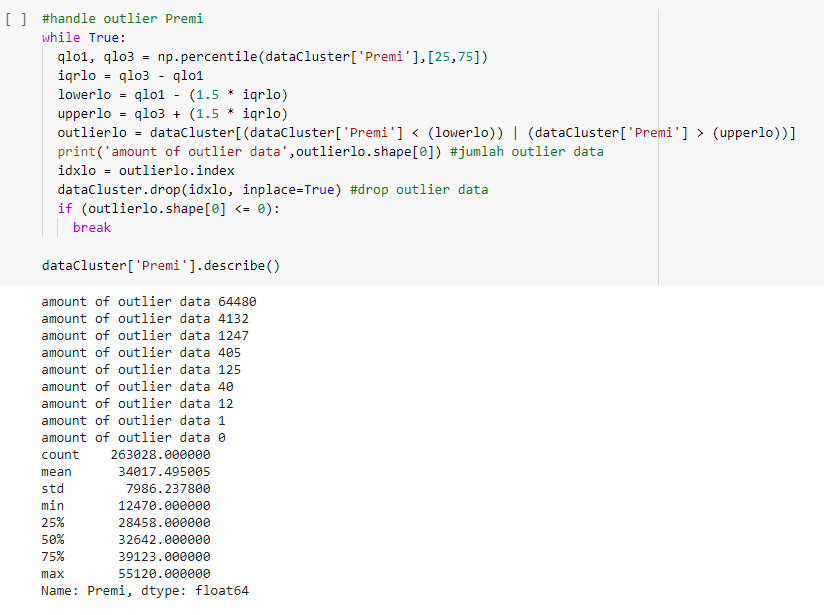
* 1. Pengecekan Outlier

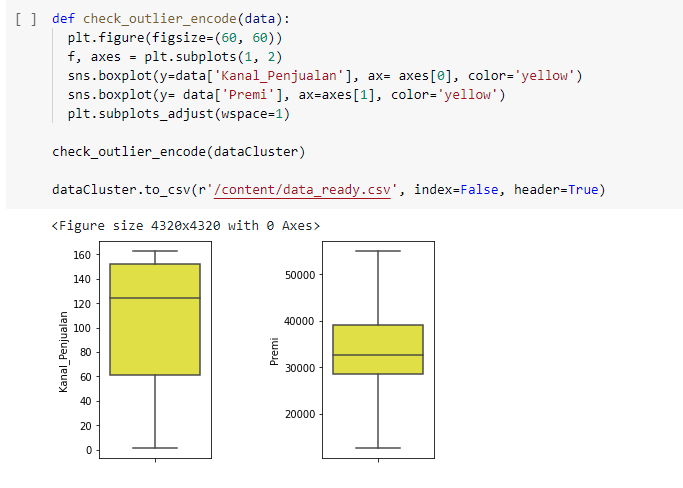
Setelah selesai dengan missing value, tahap selanjutnya adalah mencari kolom dimana terdapat data outlier yang kemudian di modelkan dengan menggunakan boxplot supaya lebih mudah melihat dimana outlier tersebut berada



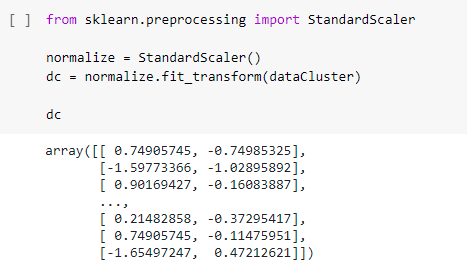
* 1. Data Cleansing

Melakukan data cleansing kemudian cek kembali adakah outlier di premi

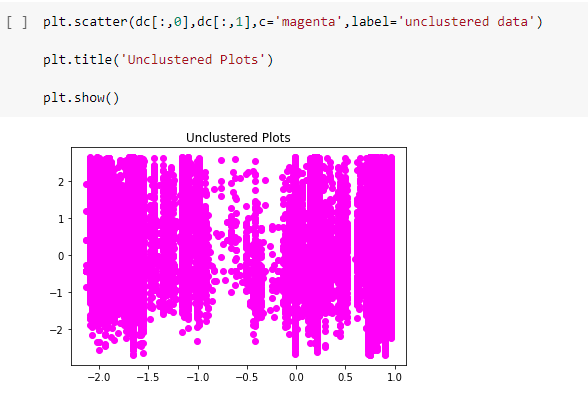




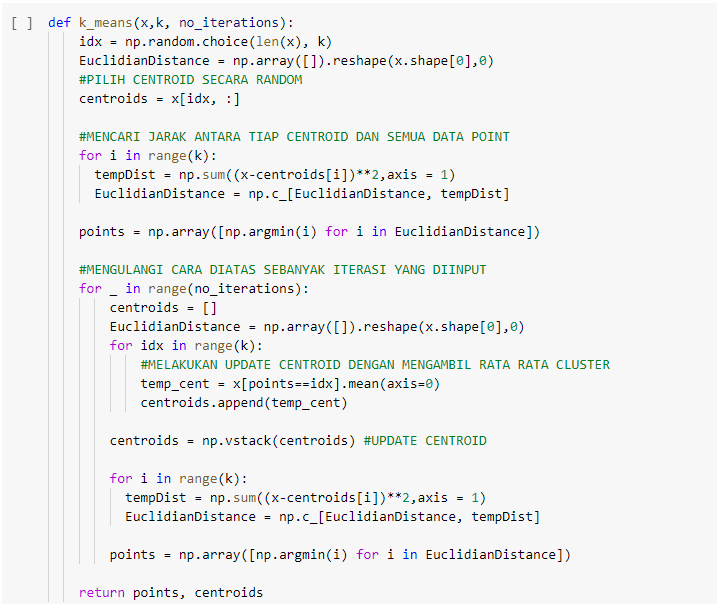
* 1. Normalisasi



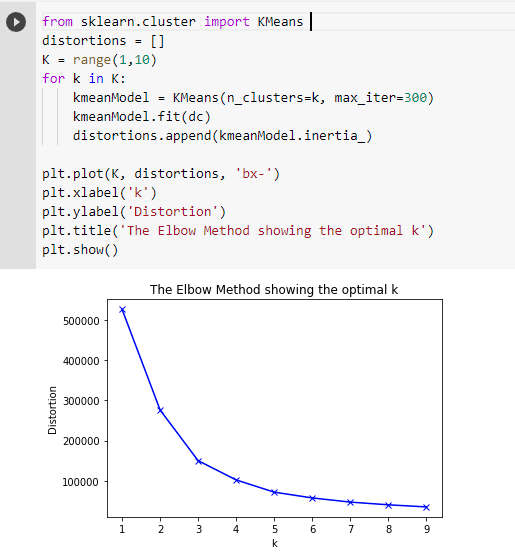
1. **Pemodelan**
   1. Unclustered Plot



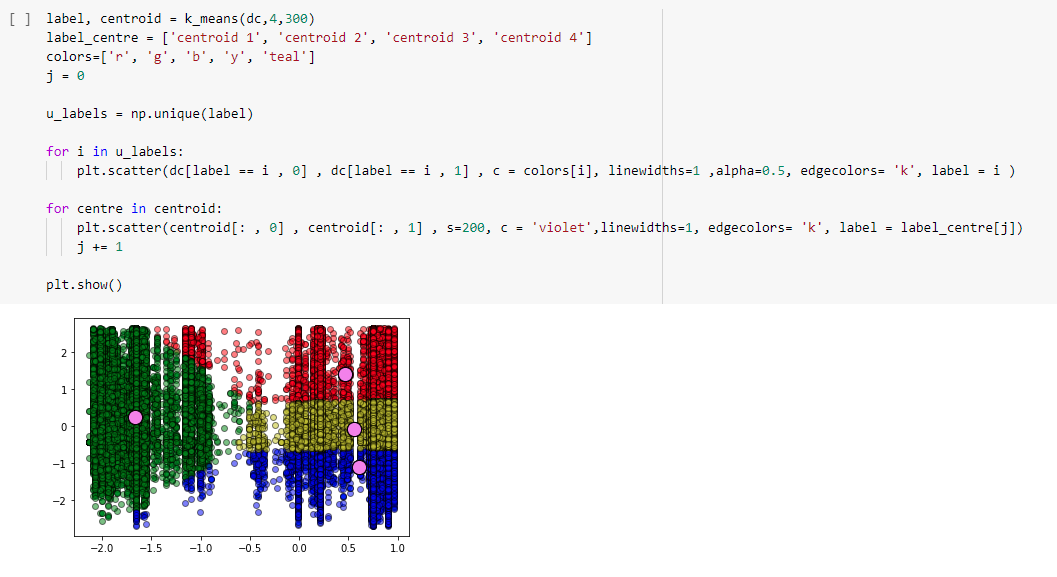
* 1. K-Means Functions

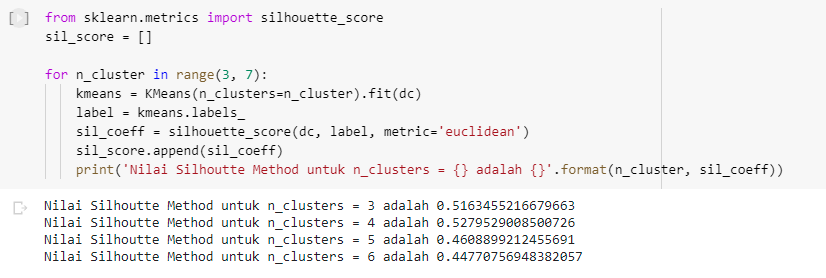


* 1. Elbow Method



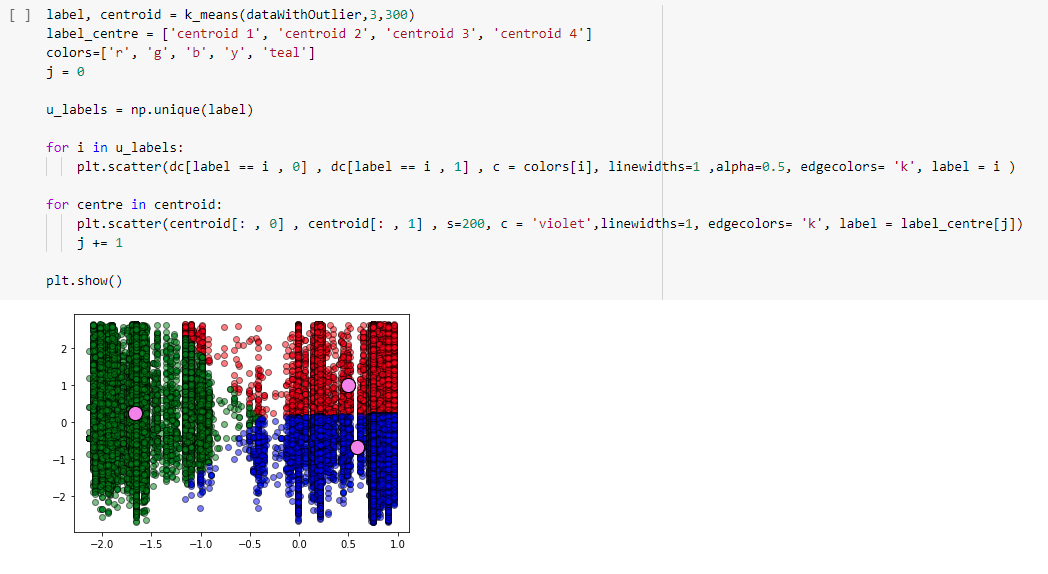
* 1. Plotting Hasil K-Means





1. **Observasi**
   1. Eksperimen Dengan Nilai K Yang Didapat

K = 3



K = 4

