```
# -*- coding: utf-8 -*-
"""Kmeans.ipynb

Automatically generated by Colaboratory.

Original file is located at
    https://colab.research.google.com/drive/132w_By3xEifKsFpHwF8vtHYWTbguWNUb
"""

import numpy as np
import matplotlib.pyplot as plt
import pandas as pd
from sklearn.cluster import KMeans

data = pd.read_csv('driverdata.csv')

data

x = data.iloc[:,[2,3]].values

data.shape

data.info()

data.isnull().sum()

from sklearn.preprocessing import StandardScaler

sc = StandardScaler()

x = sc.fit_transform(x)

kmeans = KMeans(n_clusters = 3)
kmeans.fit(x)

y_kmeans = kmeans.fit_predict(x)

plt.scatter(x[y_kmeans == 0,0], x[y_kmeans == 0,1] , s = 20 , c= "red" , label = 'Cluster 1')
plt.scatter(x[y_kmeans == 1,0], x[y_kmeans == 1,1] , s = 20 , c = "blue" , label = 'Cluster 2')
plt.scatter(x[y_kmeans == 2,0], x[y_kmeans == 2,1] , s = 20 , c = "green" , label = 'Cluster 3')
plt.scatter(kmeans.cluster_centers_[:, 0], kmeans.cluster_centers_[:, 1] , s = 25,marker = 'x' , c = "black", label = 'Centroid')
plt.title("Cluster of Driver")
plt.xlabel("Score1")
plt.ylabel("Score2")
plt.legend()
plt.show()

data.hist()

plt.scatter(data['Score1'],data['Score2'],color = 'red')
```