Module 6: Unsupervised Learning

Assignment

edureka!

Module 6: Assignment

Analyze the information given in the following 'Insurance Policy dataset' to create clusters of persons falling in the same type.

The dataset can be downloaded from the link:

https://www.edureka.co/medias/4kpv5jqe37/download?media_file_id=189490981

Premiums Paid	Age	Days to Renew	Claims made	Income
2800	26	233	3890.076336	28000
2950	27	130	2294.444444	29500
3100	28	144	2564.545455	31000
3250	30	65	1978.26087	32500
3400	32	56	2009.090909	34000

The description of the attributes in the dataset are as follows:

- Premium Paid – Amount paid by the person
- Age – Indicates the age of the person
- Days to renew – Days remaining to renew the policy
- Claims made – Indicates the claims already made by the person
- Income – Net income of the person

Perform the following tasks on the dataset

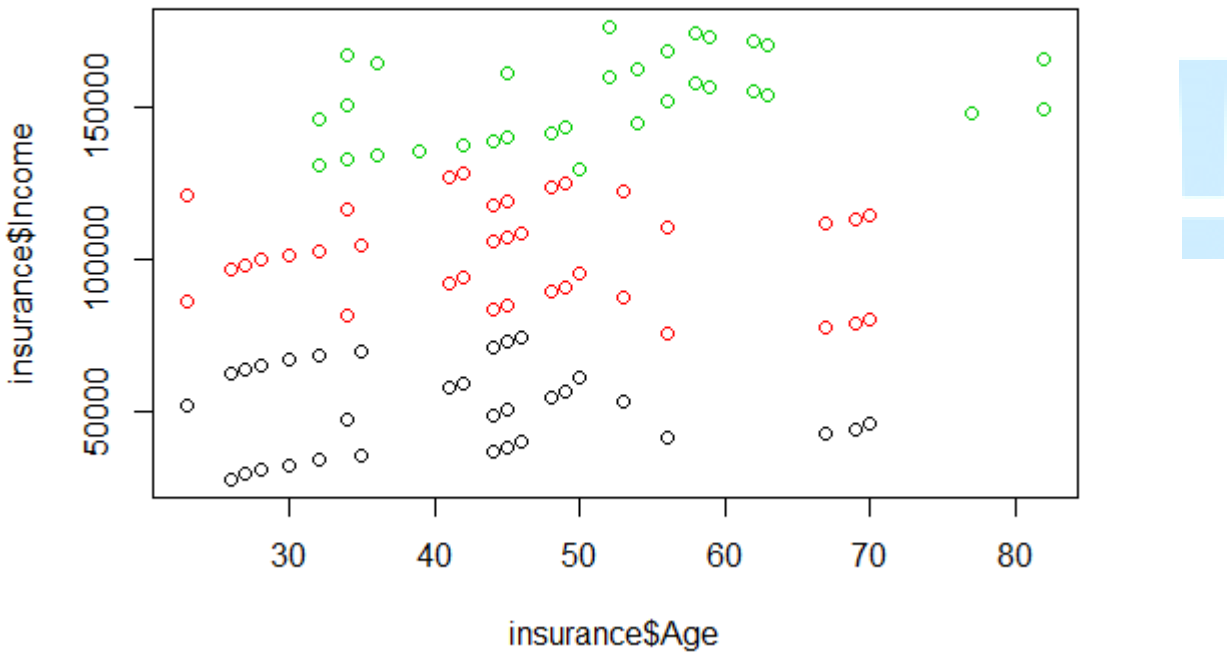
Task 1: Import the dataset into R

→ Attach the library xlsx into R to load the data .

Task 2: Perform K-means clustering on the data

→ Find the optimal number of clusters for our data

→ Plot age vs income while giving them cluster colors



Task 3: Perform C-means clustering on the data and save the membership dataframe in csv format.