



RAD1509 –Estatística Aplicada à Administração II – Lista 06

Uma concessionária está interessada no estudo do perfil de seus clientes e resolveu desenvolver uma análise de agrupamento. Para esta análise a concessionária contém os seguintes dados dos clientes:

- Renda: Renda Mensal (em milhares de reais);
- NDep: Número de Dependentes;
- Vinc: Vínculo Empregatício:
 - Se o cliente possui vínculo formal considere Vinc = 1;
 - Se o cliente não possui vínculo empregatício formal considere Vinc = 0.

A Tabela 1 apresenta uma amostra com dados de 6 clientes da concessionária. Com base nestes dados faça uma análise de agrupamento, proceda como solicitado nos itens (1) a (3), listados a seguir.

Tabela 1. Dados de clientes de uma concessionária.

Cliente	Renda	NDep	Vinc	ID_2grupos	ID_3grupos
A	2,5	2	0		
B	0,3	7	1		
C	2,3	5	1		
D	2,3	3	0		
E	1,7	7	0		
F	2	1	1		

(1) Obtenha a matriz de dissimilaridade entre os clientes da Tabela 1 dada por uma matriz de distâncias. Considere a “Distância Euclidiana ao Quadrado” obtida a partir dos dados da Tabela 1, sem a necessidade de padronizar os dados. Informe cada valor da matriz com duas casas decimais.

(2) Considere o método hierárquico “Simples”, ou seja, o método da “menor distância”, e faça uma análise de agrupamentos:

(2.1) Apresente todas as etapas do procedimento de aglomeração, ou seja, as matrizes de distâncias a cada passo do desenvolvimento de aglomeração.

(2.2) Apresente o “Dendrograma” resultante do processo de aglomeração.

(3) A partir do desenvolvimento realizado em (2), estude os possíveis agrupamentos:

(3.1) Considere a formação de dois grupos de clientes: um grupo identificado como “1” e outro grupo identificado como “2”. Complete a coluna “ID_2grupos” da Tabela 1 com a identificação do grupo ao qual o cliente pertence (grupo 1 ou grupo 2).

(3.2) Informe a Distância entre os dois grupos descritos em (3.1).

(3.3) Considere a formação de três grupos de clientes: um grupo identificado como “1”, um grupo identificado como “2” e outro grupo identificado como “3”. Complete a coluna “ID_3grupos” da Tabela 1 com a identificação do grupo ao qual o cliente pertence (grupo 1, grupo 2 ou grupo 3).

(4) Considere o método K-means, faça uma análise de agrupamentos:

(4.1) Considere dois grupos e como configuração inicial os clientes A, B e C pertencentes ao mesmo Grupo. Informe os centroides dos grupos.

(4.2) Aplique o método K-means e informe o número de interações necessárias para convergir e qual a formação final.

(4.3) Apresente um gráfico de dispersão para as variáveis “Renda” no eixo-x e “NDep” no eixo-Y, com a configuração final e com os centroides dos dois grupos.

(4.4) Comente o gráfico apresentado e a análise realizada.