



Nome: \_\_\_\_\_

Nº USP \_\_\_\_\_

**1 A Tabela 1** Apresenta as “distâncias” entre 5 consumidores. Estas distâncias foram determinadas com base em indicadores sócio-econômicos. A matriz de distâncias é simétrica. Com o objetivo de verificar semelhanças entre os consumidores realize uma análise de agrupamentos de forma a obter dois grupos.

**1.1** Apresente processo de aglomeração por meio do método de “Ligação Completa” (ou vizinho mais distante). Apresente as distâncias em cada etapa.

**1.2** Apresente o resultado para dois grupos na forma de um dendrograma e informe a distância entre os dois grupos obtidos.

**Tabela 1.** Distâncias

|    | C1 | C2 | C3 | C4 | C5 |
|----|----|----|----|----|----|
| C1 | 0  |    |    |    |    |
| C2 | 12 | 0  |    |    |    |
| C3 | 19 | 13 | 0  |    |    |
| C4 | 22 | 20 | 33 | 0  |    |
| C5 | 24 | 8  | 30 | 32 | 0  |

**2** Considere 5 empresas com as características descritas na Tabela 2. Agrupe as 5 empresas em 2 grupos através da técnica de análise de agrupamentos K-means.

**2.1** Inicialmente considere as empresas A e C como sendo do grupo 1 e as outras empresas (B, D e E) como sendo do grupo 2. Obtenha os centros dos grupos 1 e 2. Apresente um gráfico de dispersão que contenha as empresas e os centros dos grupos nas dimensões PL e AT.

**2.2** Aplique o método K-means e verifique a solução estável para dois grupos. Ou seja, indique os integrantes de cada grupo.

**2.3** Apresente o gráfico de dispersão com a solução final, com o centro dos grupos obtidos e as empresas.

**Tabela 2.** Empresas

| Id | PL (R\$ K) | AT (R\$ K) |
|----|------------|------------|
| A  | 1164       | 1620       |
| B  | 8993       | 26470      |
| C  | 3220       | 7594       |
| D  | 3327       | 9726       |
| E  | 5142       | 12384      |

**3** Considerando-se as variáveis rentabilidade (rent), liquidez\_geral (liq) e endividamento (end), uma empresa classificou seus clientes e 5 categorias: Grupos 1, 2, 3, 4 e 5. Como resultado de uma análise discriminante a empresa obteve uma tabela com os coeficientes de cinco funções de classificação, uma para cada categoria de classificação de clientes (tabela ao lado)

**Classification Function Coefficients**

|            | grupos  |         |         |         |         |
|------------|---------|---------|---------|---------|---------|
|            | 1       | 2       | 3       | 4       | 5       |
| end        | ,733    | ,560    | ,963    | ,612    | ,582    |
| liq        | 34,733  | 17,195  | 23,418  | 33,137  | 60,393  |
| rent       | ,428    | ,230    | ,475    | ,069    | ,449    |
| (Constant) | -42,561 | -18,886 | -50,459 | -35,026 | -71,100 |

**3.1** Classifique, as três empresas abaixo, em uma das categorias (um dos grupos).

| Empresa | end | liq | rent |
|---------|-----|-----|------|
| X       | 6,6 | 2,2 | 1,5  |
| Y       | 4,1 | 1,2 | 4,2  |
| Z       | 2,1 | 1,0 | 7,9  |

4 A partir do resultado de entrevistas com consumidores, foi realizada uma regressão logística de forma a obter a probabilidade do consumidor ser Leal à marca Nike. A variável "Leal" foi obtida a partir da Lealdade, por:

$$\begin{aligned} \text{Leal} &= 1 \text{ se Lealdade} > 4 \text{ e} \\ \text{Leal} &= 0 \text{ se Lealdade} \leq 4. \end{aligned}$$

| Variável  | Beta  |
|-----------|-------|
| Constante | -2,00 |
| Idade     | 0,02  |
| Intenção  | 0,25  |

As variáveis explicativas utilizadas foram Idade e Intenção de compra. Os betas obtidos no modelo logístico são apresentados no quadro ao lado.

4.1 Considerando os entrevistados listados na Tabela 3, aplique a função logística e obtenha:

4.1.1 A função linear  $g(x)$  para cada entrevistado

4.1.2 A probabilidade de que entrevistado seja Leal à marca Nike:  $P(\text{Leal})$ .

4.1.3 A Previsão de que cada entrevistado seja Leal à marca Nike. Escolha o critério com base em  $P(\text{Leal})$ .

4.2 Como você interpreta os valores de Betas neste modelo, ou seja, qual a influência da idade e da intenção da compra para a probabilidade do consumidor ser leal à marca Nike?

4.3 Forneça o percentual de acerto do modelo com base nos resultados obtidos no item 4.1 e nos valores originais da Tabela 3.

| Entrevistado | Idade | Intenção | Lealdade | Leal | $g(x)$ | $P(\text{Leal})$ | Prev (Leal) |
|--------------|-------|----------|----------|------|--------|------------------|-------------|
| A            | 43    | 5        | 6        |      |        |                  |             |
| B            | 34    | 4        | 3        |      |        |                  |             |
| C            | 25    | 4        | 5        |      |        |                  |             |
| D            | 37    | 3        | 2        |      |        |                  |             |
| E            | 30    | 5        | 6        |      |        |                  |             |
| F            | 27    | 4        | 2        |      |        |                  |             |
| G            | 42    | 7        | 5        |      |        |                  |             |
| H            | 24    | 1        | 1        |      |        |                  |             |
| I            | 24    | 1        | 1        |      |        |                  |             |
| J            | 40    | 5        | 7        |      |        |                  |             |
| K            | 30    | 1        | 3        |      |        |                  |             |
| L            | 24    | 2        | 4        |      |        |                  |             |
| M            | 39    | 3        | 4        |      |        |                  |             |
| N            | 20    | 6        | 7        |      |        |                  |             |