OTIMIZAÇÃO DAS VENDAS DE PRODUTOS DA PANIFICAÇÃO

Joaquim Eduardo de Moura Nicácio[1]

RESUMO: O emprego da estatística multivariada da análise de componentes principais, também conhecida como transformada de Hotelling ou transformada de Karhunen-Loève, na redução da dimensão dos dados observados sem perda substancial de informação, permite uma análise do comportamento das vendas de produtos da indústria da panificação, através de grupos específicos de produtos correlacionados. Como consequência, melhora o planejamento da produção e aprimora a análise de custos de produção desses produtos. A validação dessa redução de dados é dada com a análise de agrupamento também chamada de análise de cluster, que também é uma estatística multivariada.
PALAVRAS CHAVE: Estatística Multivariada, Análise de Componentes Principais, Comportamento das vendas, Análise de Custos de Produção, Análise de Cluster.

SUMMARY: The use of multivariate statistical analysis of principal components, also known as Hotelling transform or Karhunen-Loève transform, reducing the size of the observed data without substantial loss of information, allows an analysis of the behavior of sales of industry products bread, through specific product groups correlated. As a result, improved production planning and cost analysis enhances production of these products. The validation of this data reduction is given to the cluster analysis also called cluster analysis, which is also a multivariate statistics.
KEYWORDS: Multivariate Statistics, Principal Components Analysis, Behavior Sales, Production Cost Analysis, Cluster Analysis.

1 – INTRODUÇÃO

O processo de otimização através da maximização de uma *função objetivo* sujeita a alguma restrição, conhecido como Análise de Componentes Principais (ACP) ou transformada de Hotelling ou ainda Karhunen-Loève, permite a redução de dados sem perda significativa de informação.

Essa ACP foi desenvolvida por Karl Pearson e Harold Hotelling em 1933 a partir de um trabalho de um psicólogo, Spearman, em 1904, onde este tentou mostrar a partir de muitos testes diferentes que havia um fator geral da inteligência humana (fator G) e um conjunto de fatores específicos. Os fatores são entidades (construtos) hipotéticas, não diretamente observáveis que se infere ou se constrói a partir do agrupamento das variáveis. Estas são chamadas *variáveis estatísticas.*

Em 1947 Thurstone provou a inconsistência de tais conclusões criando o método de componentes principais. Foi produzido, em princípio, um modelo de análise delas sem a necessidade de hipóteses iniciais. O modelo geral das componentes principais é:

$$\begin{aligned} y_1 &= a_{11}x_1 + a_{12}x_2 + \cdots + a_{1p}x_p \\ y_2 &= a_{21}x_1 + a_{22}x_2 + \cdots + a_{2p}x_p \\ &\vdots \\ y_p &= a_{p1}x_1 + a_{p2}x_2 + \cdots + a_{pp}x_p \end{aligned} \qquad (1)$$

[1] Prof. Ms. no Departamento de Ciências Contábeis da UFMT.

onde $x_1, x_2, \ldots, x_p$ são as variáveis medidas e *correlacionadas entre si* , $y_1, y_2, \ldots, y_p$ são as variáveis estatísticas não correlacionadas que designam as componentes principais; a_{ij} onde $i = 1, \ldots, p \; e \; j = 1, \ldots, p$ são os pesos (ponderação) de cada componente na equação geral, que definem as novas variáveis y_p e são escolhidos de forma a explicar a máxima variação dos dados. Esses **pesos** são *autovetores* - $\boldsymbol{V}$ associados aos *autovalores* - $\boldsymbol{\Lambda}$ da matriz de correlação dos dados estandardizados, isto é: centrados e reduzidos onde $C = V\sqrt{\Lambda}$. Também são chamados *coordenadas das variáveis no novo sistema de eixos.*

Ela nada mais é que o cálculo dos autovalores e respectivos autovetores associados, de uma **matriz de variância-covariância** ou de uma **matriz de coeficientes de correlação** entre variáveis que seja *simétrica e definida positiva*.

Também é considerada uma técnica de transformações de variáveis, isto é: se cada variável medida (*score*) pode ser considerada como um eixo de variabilidade, estando correlacionada com outras variáveis, esta análise transforma os dados de tal modo que mantenha a mesma variabilidade existente, com o mesmo número de eixos originais, porém *não mais correlacionados entre si*.

A necessidade de se saber quais produtos têm comportamentos de vendas semelhantes, é determinante para a melhora do planejamento da produção com consequente redução de custos, pois, otimiza a utilização de recursos sem necessidade de formar estoque uma vez que o produto é produzido e praticamente vendido, reduzindo em muito a possibilidade de ficar aguardando a venda.

Nessas condições temos um problema bem definido: é possível reduzir a dimensão dos dados de forma a permitir conhecer quais produtos são consumidos e o comportamento deles, ou seja, com que frequência os produtos são consumidos?

A hipótese é que com a aplicação da análise multivariada de componentes principais (ACP), ocorrerá à redução de dados transparecendo um *mix* de produtos que são mais demandados. A validação deste *mix* se dará após análises de agrupamento.

A importância dessa pesquisa reside no fato da indústria da panificação passar a ter uma projeção de custos de produção e, consequentemente, de vendas de produtos associada ao hábito dos consumidores no dia a dia das suas atividades.

O objetivo geral é dotar a administração da produção de meios científicos de planejamento de custos de produtos através do comportamento das suas vendas. O objetivo específico é permitir que a atividade análise de custos de produção passe a ter um critério bem definido de comparação de custos projetados com os realizados.

2 – REVISÃO DA LITERATURA

A necessidade de obtenção de custos de produção é exigida pelo planejamento da atividade *produção* em uma indústria de panificação assim como em qualquer outra indústria. Considerando os aspectos inflacionários existentes em qualquer economia e a grande concorrência existente no mercado, a projeção de custos futuros permite uma otimização de recursos pela indústria da panificação principalmente quando se conhecem os hábitos do consumidor.

Sob essa ótica, a estimação de custos de matérias primas de pães, bolos, salgados e doces, favorece uma administração calcada em princípios científicos com a utilização de estatística multivariada no diagnóstico de produção e vendas.

Para que se proceda a uma análise de comportamento de vendas de produtos, é condição necessária e suficiente que exista alta correlação entre eles. Estas correlações darão origem a uma *matriz simétrica definida positiva* ponto de partida para se efetuar a análise de componentes principais.

Para obter esta matriz, é necessária a organização dos dados amostrados em uma matriz **X** (*n* x *m*) onde as linhas representam as amostras (dias) ou unidades estatísticas e, as colunas, as variáveis-produtos, medidas em unidades (quantidade de unidades de produto vendido).

$$X = \begin{bmatrix} x_{11} & x_{12} \dots & x_{1m} \\ x_{21} & x_{22} \cdots & x_{2m} \\ \vdots & \vdots & \vdots \\ x_{n1} & x_{n2} \cdots & x_{nm} \end{bmatrix} \qquad (2)$$

Apesar das variáveis–produtos representarem suas medidas em unidades inteiras, como consequência, poderem ser comparadas entre si, em se tratando de proporcionalidade, é conveniente centrar e reduzir cada variável, ou seja, subtrair de cada variável a sua média e, após, dividir pelo desvio padrão da variável da amostra, um processo designado como estandardização ou transformação Z. "Essa transformação geométrica que permite passar de X a X_R é uma homotetia de centro em G (centro de gravidade da nuvem de pontos) e de proporção 1 ÷ s_x", Crivisqui (1999, p. 26).

$$X_R = \frac{(x_i - \bar{x})}{s_X};\ \forall i \qquad \textbf{(3)}$$

A matriz de correlação **R** quando os dados foram centrados e reduzidos, isto é: estandardizados, torna-se igual à matriz **Σ** de variância-covariância porque a variância neste caso é igual a 1. Essa matriz permite reduzir o número de variáveis sob as seguintes condições:

> Se não houver nenhuma correlação entre uma variável e as demais, então ela pode ser eliminada; Se houver uma correlação pequena (por convenção: R < 0,2 indica uma associação muito baixa; entre 0,2 e 0,3 baixa; entre 0,4 e 0,69 moderada; entre 0,7 e 0,89 alta; e por fim entre 0,9 e 1 uma associação muito alta) então pode reduzir o número de variáveis. Pestana e Gageiro (2000, p. 146).

Uma medida do grau de associação entre variáveis é o determinante da matriz de correlação. Quanto mais próximo de zero maior o grau de linearidade entre as variáveis. Este é um pressuposto para a aplicação da Análise das Componentes Principais.

A derivação das componentes principais deve obedecer a um critério de otimização que é a maximização de uma função de várias variáveis sujeita a uma ou mais restrições, sendo empregado o método dos multiplicadores de Lagrange.

> Na estatística multivariada e em outras áreas aplicadas, é muitas vezes necessária a maximização de uma forma quadrática. Devido à forma quadrática Q = **x'**R**x** poder ser feita arbitrariamente grande tomando-se os valores dos elementos de **x** grandes, é necessário maximizar Q condicionada a alguma restrição no comprimento de **x**. Uma conveniente alternativa é tomar uma solução normalizada de **x**, ou seja, uma solução tal que **x** tenha comprimento unitário. Então a maximização da forma quadrática Q pode ser transformada na maximização da razão: $\lambda = \frac{\mathbf{x}'\mathrm{R}\mathbf{x}}{\mathbf{x}'\mathbf{x}}$ para toda matriz **R** simétrica real. Desde que $\lambda = \frac{\mathbf{x}'\mathrm{R}\mathbf{x}}{\mathbf{x}'\mathbf{x}}$, então para um ponto estacionário qualquer **i**, $\left(\mathrm{R} - \lambda_i \mathbf{I}\right)\mathbf{x_i} = \mathbf{0}$. Ferreira (1996, p. 55).

Devemos lembrar que λ é a matriz diagonal dos autovalores de *R* e **x'x** é a matriz identidade.

"A aplicação deste método resulta na definição da equação característica de autovetores da matriz de variância-covariância **Σ** que definem os eixos fatoriais (componentes principais) em direção e comprimento pela resolução da equação:

$$\left[\Sigma - \lambda_i \boldsymbol{I}\right]\boldsymbol{a}_i = \boldsymbol{0} \quad \textbf{(4)}$$

Onde **I, $\mathbf{a_i}$** e **0** são vetores." Reis (2001, p. 264). O número de eixos fatoriais é determinado por ***i*** que varia de 1 até *n*.

Para que esta equação tenha uma solução para **a** que não seja a solução nula, é necessário que a matriz $\left[\Sigma - \lambda\mathbf{I}\right]$ seja uma matriz singular, isto é, seu determinante $\left|\Sigma - \lambda\mathbf{I}\right| = 0$ e neste caso, existirá uma solução não nula para **a**, se e só se λ for um valor próprio de **Σ**.

Como se trata de um problema de maximização, as raízes características da equação devem ser positivas. Para que isso ocorra, a matriz **R** ou **Σ** deve ser definida positiva, isto é: "uma matriz **Σ** é chamada uma matriz definida positiva se todos determinantes dos menores principais de **Σ** forem positivos" Bolch e Huang (1974, p. 21).

$$R_{11} > 0, \begin{vmatrix} R_{11} & R_{12} \\ R_{21} & R_{22} \end{vmatrix} > 0, \begin{vmatrix} R_{11} & R_{12} & R_{13} \\ R_{21} & R_{22} & R_{23} \\ R_{31} & R_{32} & R_{33} \end{vmatrix} > 0, ..., |R| > 0 \quad \textbf{(5)}$$

Para exemplificar vamos considerar uma matriz hipotética **Σ** e uma matriz **I** tais que:

$$\Sigma = \begin{bmatrix} 2 & 1 \\ 1 & 3 \end{bmatrix} \text{ e a matriz } I = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

A equação característica dos autovetores $\left[\Sigma - \lambda I\right]\boldsymbol{a} = \boldsymbol{0}$ pode ser escrita como $\phi = a'\Sigma a - \lambda\left(a'a - 1\right)$ onde $a = \begin{bmatrix} a_1 \\ a_2 \end{bmatrix}$ e $a'a = 1$ que é a restrição e os autovetores estejam normatizados, ϕ é a função objetivo a ser maximizada e λ são os multiplicadores indeterminados de Lagrange e $\Sigma = \begin{bmatrix} 2 & 1 \\ 1 & 3 \end{bmatrix}$.

Para determinarmos o máximo da função ϕ devemos achar suas derivadas parciais e igualá-las a zero. Uma vez que o vetor **a** = $\begin{bmatrix} a_1 \\ a_2 \end{bmatrix}$ temos:

$$Z = a'\Sigma a = \begin{bmatrix} a_1 & a_2 \end{bmatrix} \begin{bmatrix} 2 & 1 \\ 1 & 3 \end{bmatrix} \begin{bmatrix} a_1 \\ a_2 \end{bmatrix} = \quad \textbf{(6)}$$

$Z = 2a_1^2 + 2a_1a_2 + 3a_2^2$. Determinando as derivadas parciais de Z tem-se:

$$\frac{\partial Z}{\partial a} = \begin{bmatrix} \frac{\partial Z}{\partial a_1} \\ \frac{\partial Z}{\partial a_2} \end{bmatrix} = \begin{bmatrix} 4a_1 + 2a_2 \\ 2a_1 + 6a_2 \end{bmatrix} = 2\begin{bmatrix} 2 & 1 \\ 1 & 3 \end{bmatrix} \begin{bmatrix} a_1 \\ a_2 \end{bmatrix} \quad \textbf{(7)}$$

Aplicando o multiplicador de Lagrange na função objetivo, tem-se:

$$\phi = 2a_1^2 + 2a_1a_2 + 3a_2^2 - \lambda\left(a_1^2 + a_2^2 - 1\right) \quad \textbf{(8)}$$

Efetuando as derivadas parciais e igualando a zero, obtêm-se:

$$\frac{\partial \phi}{\partial a_1} = 4a_1 + 2a_2 - 2\lambda a_1 = 0$$

(9)

$$\frac{\partial \phi}{\partial a_2} = 2a_1 + 6a_2 - 2\lambda a_2 = 0$$

(10)

A equação característica dos autovalores $\left[\Sigma - \lambda \boldsymbol{I}\right] = 0$ torna-se:

$\left[\begin{bmatrix} 2 & 1 \\ 1 & 3 \end{bmatrix} - \lambda \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}\right] = \begin{bmatrix} (2-\lambda) & 1 \\ 1 & (3-\lambda) \end{bmatrix} = 0$. Achando o determinante de $\left[\Sigma - \lambda \boldsymbol{I}\right] = 0$

Obtemos: $\lambda^2 - 5\lambda + 5 = 0$ que tem como raízes (multiplicadores, agora determinados, de Lagrange) λ_1 **= 3,618034** e λ_2 = **1,381966**. Aplicando λ_1, para maximizar, em **(9)** e em **(10)** obtêm-se:

$2a_2 - 3{,}236068a_1 = 0$ fazendo a_1= -1, então: $a_1 = 0{,}618034$ **(11)**

$2a_1 - 1{,}236068a_2 = 0$ fazendo a_2 = 1 então: $a2 = -1{,}618034$ **(12)**

A equação característica dos autovetores é $[\Sigma - \lambda I]\boldsymbol{a} = 0$ e para λ_1 = 3,618034 suas equações simultâneas são:

$$\begin{bmatrix} (2-\lambda_1) & 1 \\ 1 & (3-\lambda_1) \end{bmatrix} \begin{bmatrix} a_{11} \\ a_{21} \end{bmatrix} = 0 \Rightarrow \begin{Bmatrix} -1.618a_{11} + a_{21} = 0 \\ a_{11} - 0.618034a_{21} = 0 \end{Bmatrix} \quad \textbf{(13)}$$

Para sua solução pode existir inúmeros valores que resolvam esse sistema indeterminado.

> "Um dos procedimentos normalmente adotados é dar o valor **um** para o elemento a_{11} na primeira equação e determinar o valor de a_{21} ignorando a segunda equação. Repetir o procedimento para a outra raiz própria, porém considerando na segunda equação $\mathbf{a_{22}}$ **= 1**". Valentin (2000, p. 81).

$$\begin{bmatrix} 2-\lambda_2 & 1 \\ 1 & 3-1.381966 \end{bmatrix} \begin{bmatrix} a_{21} \\ a_{22} \end{bmatrix} = 0 \Rightarrow \begin{Bmatrix} 0.618034a_{21} + a_{22} = 0 \\ a_{21} + 1.618034a_{22} = 0 \end{Bmatrix}$$

(14)

A matriz dos autovetores torna-se: $\boldsymbol{a} = \begin{bmatrix} 1 & -1{,}618034 \\ 1{,}618034 & 1 \end{bmatrix}$.

Para que esses dois vetores possam ser comparados, é preciso estar numa mesma escala. Para isso, deve-se normaliza-los, dividindo cada elemento pela

raiz quadrada da soma dos quadrados dos elementos. A matriz dos autovetores normalizados torna-se igual a:

$$\boldsymbol{a} = \begin{bmatrix} \frac{1}{\sqrt{3,618034}} & \frac{-1,618034}{\sqrt{3,618034}} \\ \frac{1,618034}{\sqrt{3,618034}} & \frac{1}{\sqrt{3,61834}} \end{bmatrix} = \begin{bmatrix} 0,525731 & -0,850651 \\ 0,850651 & 0,525731 \end{bmatrix}$$

(15)

Após isso pode ser observado que:

1. Esses dois vetores são ortogonais, pois o produto escalar de suas coordenadas é igual a zero, isto é:

$$\begin{bmatrix} 0,525731 & 0,850651 \end{bmatrix} \begin{bmatrix} -0,850651 \\ 0,525731 \end{bmatrix} = 0$$

(16)

2. Os elementos desses dois vetores de norma 1 são os **cossenos diretores** dos ângulos que eles fazem com o sistema de origem;
3. Esses vetores correspondem a uma rotação dos eixos originais de um ângulo $\alpha = 58°19'$, lembrando que a matriz de rotação anti-horária **T** é:

$$\boldsymbol{T} = \begin{bmatrix} \cos\alpha & -sen\alpha \\ sen\alpha & \cos\alpha \end{bmatrix} = \begin{bmatrix} 0,525731 & -0,850651 \\ 0,850651 & 0,525731 \end{bmatrix}$$

(17)

Com respeito às raízes características ou raízes próprias é necessário conhecer duas propriedades recomendadas por Bolch e Huang (1974, p. 30):

1. O produto das raízes características de uma matriz simétrica (nada mais é do que a matriz diagonal) é igual ao determinante desta matriz:

$$|D| = |\mathbf{a}'R\mathbf{a}| = |\mathbf{a}'\mathbf{a}|\,|R| = |R| \qquad \text{desde que} \quad |\mathbf{a}'\mathbf{a}| = \mathbf{I}$$

(18)

$$\begin{vmatrix} 2 & 1 \\ 1 & 3 \end{vmatrix} = \begin{vmatrix} 3{,}618024782 & \\ & 1{,}381962495 \end{vmatrix} = 5$$

2. A soma das raízes características de uma matriz simétrica é igual ao traço da matriz:

$$\text{tr}\,\mathbf{D} = \text{tr}(\mathbf{a}'R\mathbf{a}) = \text{tr}(\mathbf{a}'\mathbf{a}R) = \text{tr}R$$

(19)

$$2 + 3 = 3{,}618024782 + 1{,}381962495 = 5$$

A interpretação dos fatores deve ser iniciada pela matriz fatorial inicial não rotacionada, pois contém cargas fatoriais para cada variável em cada fator o que permite ao pesquisador verificar a melhor combinação linear de variáveis.

"As *cargas fatoriais* são a correlação de cada variável com o fator. Elas indicam o grau de correspondência entre a variável e o fator, sendo que cargas maiores fazem a variável representativa do fator". Hair et al. (2005, p. 103).

As cargas fatoriais são calculadas, segundo Valentin (2000, p. 83), de acordo com uma matriz de covariância Σ e a matriz diagonal das raízes quadradas dos autovalores D. O produto $\boldsymbol{\Sigma D} = f = \begin{bmatrix} EixoI & EixoII \\ f1 & f2 \\ f1.2 & f2.2 \end{bmatrix}$.

Ainda de acordo com esse autor, a soma dos *quadrados das coordenadas* de uma variável sobre os dois eixos , isto é: o vetor linha $\begin{bmatrix} f_1^2 & f_2^2 \end{bmatrix}$ é *igual a variância desta variável*, e que a *soma dos quadrados*

das duas variáveis sobre um determinado eixo *é igual a variância deste eixo*, isto é: seu autovalor $\begin{bmatrix} f_1^2 \\ f_{1.2}^2 \\ = \lambda \end{bmatrix}$.

A comunalidade de uma variável é a proporção da sua variância que pode ser explicada pelo método das componentes principais. Na análise de componentes principais como não se supõe a existência de nenhum fator comum, a comunalidade toma como valor inicial 1 (variância total igual a 1 devido a redução e centralização dos dados na matriz de correlação). "Em termos estatísticos, a variância total pode ser decomposta numa parcela explicada pela comunalidade e uma parcela não explicada, denominada de especificidade". Hair et al. (2005, p.101).

Um outro pressuposto da Análise de Componentes Principais é a estatística KMO:

> A estatística de Kaiser-Meyer-Olkin (KMO) é uma medida de adequação amostral, que é um índice que compara a magnitude dos coeficientes de correlação observados com a magnitude dos coeficientes de correlação parcial. Ela é definida conforme Merino; Díaz (2002, p. 413) da seguinte maneira:

$$KMO = \frac{\sum_{i \neq j} r_{ij}^2}{\sum_{i \neq j} r_{ij}^2 + \sum_{i \neq j} r_{ij,\boldsymbol{m}}^2}$$

(20)

Onde $\sum_{i \neq j} r_{ij}^2$ é a soma dos quadrados de cada elemento fora da diagonal principal da matriz de correlação dos dados brutos originais e $\sum_{i \neq j} r_{ij,\boldsymbol{m}}^2$ é a soma dos quadrados de cada elemento fora da diagonal principal da matriz de correlação parcial dos dados brutos originais.

> A medida de Kaiser-Meyer-Olkin (KMO) é uma medida de diagnóstico da homogeneidade das variáveis. Para valores superiores a 0,9 nesta medida, a análise fatorial é muito boa, para valores entre 0,8 e 0,9 a análise fatorial é boa. Isto porque o KMO perto de 1 indica coeficientes de correlação parciais pequenos. Leal (2005).

O teste de esfericidade de Bartlett, um outro pressuposto da ACP, comprova que a matriz de correlação se ajusta a matriz identidade **I**, ou seja, ausência de correlação significativa entre as variáveis. Isto significa que a nuvem de pontos se ajusta a uma esfera perfeita, expressando assim a hipótese nula por:

H_0: **R = I**

H_0: $|\mathbf{R}| = 1$

A fórmula correspondente assume a seguinte expressão:

$$\chi^2 = -\left[(n-1) - \frac{1}{6} * (2 * v + 5)\right] * \ln |R|$$

(21)

Onde, n = tamanho amostral, v = número de variáveis, ln = logaritmo neperiano, R = matriz de correlação. Se a hipótese nula é aceita (p > 0,05) significa que as variáveis não estão inter-relacionadas e, portanto, não faz sentido levar a cabo uma análise fatorial.

> Quando a análise de componentes principais é aplicada a partir de uma matriz de correlações, uma forma de avaliar

> o grau de adequação dos dados à aplicação deste tipo de análise consiste em testar se, na população, essa matriz é uma matriz identidade, ou seja, não existem correlações significativas entre as variáveis. A aplicação da análise de componentes principais pressupõe que se rejeite a hipótese nula: H_0: R = I ou H_0: $\lambda_1 = \lambda_2 = \ldots = \lambda_p$. Reis (2001, p. 278).

As correlações entre variáveis também podem ser analisadas computando as correlações parciais entre elas, ou seja, as correlações entre variáveis quando os efeitos de outras variáveis são levados em consideração.

Quanto à correlação parcial (SPIEGEL, 1972, p. 451) assim se pronuncia: "é importante, frequentemente, medir a correlação entre uma variável dependente e uma independente particular, quando todas as outras implicadas se conservam constantes, isto é, quando se removem os efeitos de todas as outras variáveis".

Em se tratando de análise multivariada de componentes principais, não existe a condição de variável dependente ou independente o que não invalida a forma de calcular a correlação parcial, cuja fórmula para a correlação parcial entre X_1 e X_2, conservando-se X_3 constante é:

$$r_{12,3} = \frac{r_{12} - r_{13}r_{23}}{\sqrt{\left(1 - r_{13}^2\right)\left(1 - r_{23}^2\right)}}$$

(22)

> Se existem fatores "verdadeiros" nos dados, a correlação parcial deverá ser pequena, pois a variável pode ser explicada pelos fatores (variáveis estatísticas com cargas para cada variável). Se as correlações parciais são altas, então não há fatores latentes "verdadeiros" e a análise fatorial é inadequada Hair et al. (1998, p. 98).

A medida de adequação MAS, que é um outro pressuposto da ACP, é similar ao pressuposto KMO. Sua fórmula para cada j-iésima variável de MSA vem dada pela seguinte fórmula:

$$\mathrm{MSA}(\mathrm{j}) = \frac{\sum_{i \neq j} \boldsymbol{r}_{ij}^{2}}{\sum_{i \neq j} \boldsymbol{r}_{ij}^{2} + \sum_{i \neq j} \boldsymbol{r}_{ij,\boldsymbol{m}}^{2}}$$

(23)

Onde $MSA_{(j)}$ será determinado pela soma dos quadrados dos elementos formadores da coluna (*j*), exceto o valor da diagonal, que é 1, da matriz de correlação dos dados brutos originais, ou seja: $\sum_{i \neq j} \boldsymbol{r}_{ij}^{2}$, e quanto à $\sum_{i \neq j} \boldsymbol{r}_{ij,\boldsymbol{m}}^{2}$ é a soma dos quadrados dos elementos formadores da coluna (*j*), exceto o valor da diagonal, que é 1, da matriz de correlação parcial. Após isso, esse valor passa a substituir o valor unitário da coluna (*j*).

De acordo com os autores: Merino; Díaz (2002, p. 413), se o valor da medida de adequação amostral (MSA) for reduzido (os valores abaixo de 0,6 são considerados medíocres) e pode não ser pertinente utilizar a análise fatorial com estes dados.Essa estatística permite uma análise da fatorabilidade dos dados.

Um dos critérios de extração do número de fatores é o de Kaiser, denominado critério K1, isto é, reter uma quantidade de fatores que possuem autovalores iguais ou acima de 1. Também foi utilizado o gráfico *Scree plot* ou prova de sedimentação de Cattell. A confirmação do número de fatores a ser retidos se dá com uma análise de agrupamento ou análise de cluster.

A interpretação de cada componente é feita com base nos pesos ou correlações das variáveis e estaria simplificada se cada variável tivesse um peso relativamente mais elevado para *apenas uma* das componentes, e pesos pequenos ou próximos de zero para todas as restantes componentes principais.

3 – MATERIAIS E MÉTODOS

A empresa na qual os produtos foram observados, tem sua sede em Cuiabá e é conhecida como Pão e Delícias. Foram sorteados, dentre 130 produtos dela, doze destes contemplando as produções de pães, bolos, biscoitos, salgados, massa de pizza, e foram codificados como P1 a P12, de

acordo com o apêndice, para resguardar o segredo de empresa. Em seguida tiveram suas quantidades de venda anotadas durante um mês que foi novembro de 2011, devido às condições operacionais que não permitiriam obter uma amostragem bem mais ampla.

Após inserção desses dados, chamados variáveis, no *software* Matlab foi extraído uma matriz de correlações com a finalidade de detectar variáveis com baixos índices, o que permitiu efetuar uma redução deles.

Esses foram estandardizados, isto é: centrados e reduzidos (média zero e variância 1) para que a matriz de Correlação-**R** fosse idêntica à matriz de Variância-Covariância-**Σ**. Em seguida, foi determinado o valor do determinante da matriz de correlação cujo valor foi 0.0061. Quando o seu valor está próximo de zero, indica que as variáveis utilizadas estão linearmente relacionadas e isto permite concluir que a análise fatorial é uma técnica pertinente para analisar essas variáveis.

Essa condição favorecida pela capacidade de processamento do computador permitiu verificar se a matriz de correlação teve determinante diferente de zero. Também foi verificado se ela era uma **matriz definida positiva**, que tem seus autovalores (**λ**) maiores que zero.

Em seguida foi executado o teste de Kaiser-Meyer-Olkin cuja matriz de correlações parciais foi obtida através da aplicação da seguinte fórmula, Valentin (2000, p.47)

$$r_p = \left[D_{R^{-1}}^{-0.5} * R^{-1} * D_{R^{-1}}^{-0.5} \right]$$

(24)

"Onde R^{-1} é a matriz inversa da matriz de coeficientes de correlação **R** entre todas as variáveis, e $D_{R^{-1}}^{-0.5}$ é uma matriz diagonal, onde a diagonal é formada pelo inverso da raiz quadrada dos termos da diagonal de R^{-1}". A interpretação do índice KMO deve ser de acordo com o quadro 1, (REIS, 2001):

KMO	Análise das Componentes Principais

1,00 – 0,90	Muito Boa
0,80 – 0,90	Boa
0,70 – 0,80	Média
0,60 – 0,70	Razoável
0,50 – 0,60	Má
< 0,50	Inaceitável

Quadro 1 – Intervalos do teste de KMO

Considerado *médio* o teste de KMO, foi executado ainda o teste MSA que se tornou a diagonal principal da matriz de correlação parcial e, ainda foi executado o teste de esfericidade de Bartlett que fornece a probabilidade estatística de que a matriz de correlação tenha correlações significantes entre pelo menos algumas variáveis.

Diante disso viabilizou-se a aplicação do método estatístico das componentes principais. O cálculo dos autovalores ou raízes características e dos autovetores característicos associados foi elaborado com a utilização do programa matemático Matlab 6.5 que já fornece os autovetores normalizados, isto é: $\mathbf{v_i'v_i} = 1$ e $\mathbf{v_i'v_j} = 0$ para $j \neq i$.

Foi adotado o critério de Kaiser, para retenção dos autovalores iguais ou maiores que 1. Também foi elaborado o gráfico *scree plot* de Cattell.

Os pesos ou fatores de ponderação quando apresentados em trabalhos científicos de componentes principais, normalmente são os vetores originais transformados de acordo com a seguinte lei de transformação:

$$\mathbf{v_i^*} = \lambda^{\frac{1}{2}}.\mathbf{v_i} \qquad \textbf{(25)}$$

Estes vetores são tais que a soma dos quadrados dos seus elementos são iguais ao correspondente valor próprio $\boldsymbol{\lambda_i}$ e não igual a **1,** uma vez que

$$\left(\boldsymbol{v_i^*}\right)'\boldsymbol{v_i^*} = \left(\sqrt{\lambda}_i.\boldsymbol{v_i}\right)'.\left(\sqrt{\lambda}_i.\boldsymbol{v_i}\right) = \lambda_i \boldsymbol{v_i'v_i} = \lambda_i \qquad \textbf{(26)}$$

O que significa no entender de (REIS, 2001, p. 270), que os elementos das componentes mais importantes, são *ponderados* de modo mais significativo. Ainda de acordo com a autora:

Se a matriz dos pesos $\mathbf{f} = \left\lfloor \mathbf{v}_1^* \quad \mathbf{v}_2^* \quad \ldots \quad \mathbf{v}_i^* \right\rfloor$ então $\mathbf{f = V.\Lambda^{½}}$ e $\mathbf{R = f.f'}$ **(27)**

Os vetores transformados podem ser interpretados de duas maneiras:

- São pesos ou ponderações das variáveis iniciais nos fatores (componentes) respectivos;
- E simultaneamente, os seus elementos medem as correlações entre as componentes e as variáveis originais estandardizadas.

Dessa forma:

$$\mathrm{Cov}\left(Y_j, X_i\right) = \mathrm{Cov}\left(Y_j, \sum_{j=1}^{p} v_{ij}\, Y_j\right) = v_{ij}\, \mathrm{Var}(Y_j) = v_{ij}\lambda_j$$

(28)

Correlação (

$$Y_j, X_i) = \frac{\mathrm{Cov}\left(Y_j, X_i\right)}{\sqrt{\mathrm{Var}\left(Y_j\right).\mathrm{Var}\left(X_i\right)}} = \frac{\lambda_j v_{ij}}{\sqrt{\lambda_j}} = v_{ij}\sqrt{\lambda_j} \qquad \textbf{(29)}$$

A confirmação do número de fatores retidos se dará pelo método de agrupamento ou análise de cluster que tiver o maior coeficiente cofenético que é o coeficiente *r* de Pearson, calculado entre os índices de similaridade ou distância da matriz original e os índices reconstituídos com base no dendograma.

Para melhor simplificação dos resultados, no sentido de ter agrupamentos mais homogêneos em cada componente, foi executada a *rotação rígida dos eixos ortogonais*, mediante o critério Varimax. Neste critério as distâncias e ângulos entre as variáveis permanecem inalteradas. Não afeta os valores das comunalidades ou correlações.

4 – RESULTADOS

As regras comumente utilizadas pelos pesquisadores indicam a necessidade de sete a dez observações por variável. Neste caso, seria

necessário ter 120 observações. Como a amostra de variáveis foi composta de 372 observações, esse indicador foi plenamente atendido Freitas e Andrade (2004).

Antes de realizar a análise de componentes principais, com o objetivo de verificar a existência de fatores subjacentes as 12 variáveis, foram observados alguns índices que forneceram informações sobre a indicação ou não da solução fatorial.

A matriz de correlação indicou as variáveis P5, P9, P10, P12 com coeficientes menores ou iguais a 0,423. Essas variáveis foram retiradas da análise, significando que 66,66% foram significativas com correlações superiores a 0,423 conforme tabela 1.

O determinante da matriz de correlação foi de 0.0061, valor este que indica alta correlação entre as variáveis e que existirá mais de um fator.

Tabela 1

r - Matriz de Correlação Linear de Pearson

P1	P2	P3	P4	P6	P7	P8	P11
1.000	0.6757	0.5565	0.3368	0.4283	0.2104	0.0171	0.6048
0.6757	**1.000**	0.4799	0.2932	0.2988	0.3274	0.0705	0.5889
0.5565	0.4799	**1.000**	0.7888	0.5524	0.2802	0.3988	0.8746
0.3368	0.2932	0.7888	**1.000**	0.6393	0.2808	0.5726	0.6390
0.4283	0.2988	0.5524	0.6393	**1.000**	0.4330	0.2302	0.5106
0.2104	0.3274	0.2802	0.2808	0.4330	**1.000**	0.0730	0.2681
0.0171	0.0705	0.3988	0.5726	0.2302	0.0730	**1.000**	0.2356
0.6048	0.5889	0.8746	0.6390	0.5106	0.2681	0.2356	**1.000**

Fonte: Construção do autor.

A fórmula utilizada no teste de esfericidade de Bartlett foi a (21). Executando com os valores obtidos foi encontrado com 28 graus de liberdade:

$$\chi^2 = [\ 30 - 1 - 0{,}167 \times (2 \times 8 + 5)] \times (-4{,}017) = \mathbf{102{,}434}.$$

A matriz de correlações parciais ou matriz anti-imagem (**r_m**), tabela 2, apresenta em sua diagonal principal valores, medida de adequação amostral (MAS), considerados médios.

Tabela 2

r_m - Matriz de Correlações Parciais ou Matriz Anti-imagem

P1	P2	P3	P4	P6	P7	P8	P11
0.7638	-0.5202	-0.2213	0.1425	-0.2666	0.1556	0.1422	-0.0157
-0.5202	**0.7305**	0.1261	0.0140	0.1477	-0.2778	-0.0800	-0.3088
-0.2213	0.1261	**0.7482**	-0.4860	0.1024	-0.0666	-0.1147	-0.7085
0.1425	0.0140	-0.4860	**0.7606**	-0.4441	0.0159	-0.4317	0.0568
-0.2666	0.1477	0.1024	-01.444	**0.7720**	-0.3413	0.1170	-0.0921
0.1556	-0.2778	-0.0666	0.0159	-0.3413	**0.7131**	0.0546	0.0717
0.1422	-0.0800	-0.1147	-0.4317	0.1170	0.0546	**0.6926**	0.1617
-0.0157	-0.3088	-0.7085	0.0568	-0.0921	0.0717	0.1617	**0.7802**

Fonte: Construção do autor.

Os vetores principais da matriz de correlação foram calculados pelo programa matemático Matlab 6.5 cujos resultados estão na tabela 3 em ordem contrária do pacote matemático. O produto da matriz **V** pela matriz simétrica **V**' gerou uma matriz diagonal unitária, demonstrando a ortogonalidade dos vetores.

Tabela 3

V - Vetores próprios da matriz de correlação

V1	**V2**	**V3**	**V4**	**V6**	**V7**	**V8**	**V11**
0.1090	-0.1191	0.6396	-0.4507	-0.0203	0.2208	0.4396	0.3454

-0.1190	-0.1100	-0.6115	-0.1337	0.5019	0.1261	0.4548	0.3275
-0.7520	-0.0000	0.1607	0.3840	-0.1244	0.1860	-0.1044	0.4475
0.3113	-0.7442	-0.1207	0.0380	-0.1374	0.0035	-0.3894	0.4034
-0.1050	0.3136	-0.2987	-0.5068	-0.5131	-0.3793	-0.0898	0.3597
0.0401	-0.0189	0.2571	0.2506	0.3351	-0.8309	0.1223	0.2270
0.0226	0.3127	0.1362	-0.3074	0.5673	0.1346	-0.6374	0.2059
0.5462	0.4727	-0.0528	0.4647	-0.1262	0.2199	0.0897	0.4326

Fonte: Construção do autor.

Os autovalores próprios da matriz de correlação, também foram calculados pelo Matlab 6.5, cujos resultados estão na tabela 4. A soma deles, que é o traço da matriz, é igual a 8.

Tabela 4
Λ - Autovalores da matriz de correlação

λ_1	λ_2	λ_3	λ_4	λ_5	λ_6	λ_7	λ_8
0,0837	0,1885	0,2684	0,4507	0,6181	0,9454	**1.3383**	**4.1069**

Fonte: Construção do autor

A matriz dos pesos ou coordenadas das variáveis $f_{m,m} = V_{m,m}\sqrt{\Lambda_{m,m}}$ (matriz da estrutura fatorial) onde **V** é a matriz dos vetores característicos e **Λ** é a matriz dos autovalores está representada abaixo.

$$f = \begin{bmatrix} 0.0315 & -0.0517 & 0.3314 & -0.3026 & -0.0160 & 0.2147 & 0.5085 & 0.7000 \\ -0.0344 & -0.0477 & -0.3168 & -0.0898 & 0.3946 & 0.1226 & 0.5261 & 0.6637 \\ -0.2176 & -0.0000 & 0.0832 & 0.2578 & -0.0978 & 0.1808 & -0.1208 & 0.9069 \\ 0.0901 & -0.3231 & -0.0625 & 0.0255 & -0.1081 & 0.0034 & -0.4505 & 0.8175 \\ -0.0304 & 0.1362 & -0.1547 & -0.3402 & -0.4034 & -0.3688 & -0.1039 & 0.7289 \\ 0.0116 & -0.0082 & 0.1332 & 0.1682 & 0.2635 & -0.8079 & 0.1415 & 0.4600 \\ 0.0065 & 0.1358 & 0.0706 & -0.2064 & 0.4460 & 0.1309 & -0.7374 & 0.4173 \\ 0.1580 & 0.2052 & -0.0274 & 0.3120 & -0.0992 & 0.2138 & 0.1038 & 0.8766 \end{bmatrix}$$

O quadro 3 resumo dos resultados diz o seguinte: a primeira componente principal explica 51,134% da *variância total inicial*. A segunda componente principal explica 16,73%.

	Componente 1	Componente 2	...	**Comunalidade:** h_i^2
P1	$0{,}7000^2$	$0{,}5085^2$	...	$0{,}0315^2 + (-0{,}0517)^2 + ... = 1$
P2	$0{,}6637^2$	$0{,}5261^2$	...	$-0{,}0344^2 + (-0{,}0477)^2 + ... = 1$
P3	$0{,}9069^2$	$-0{,}1208^2$		$-0{,}2176^2 + (-0{,}0000)^2 + ... = 1$
P4	$0{,}8175^2$	$-0{,}4505^2$		$0{,}0901^2 + (-0{,}3231)^2 + ... = 1$
P6	$0{,}7289^2$	$-0{,}1039^2$		$-0{,}0304^2 + 0{,}1362^2 + ... = 1$
P7	$0{,}4600^2$	$0{,}1415^2$		$0{,}0116^2 + (-0{,}0082)^2 + ... = 1$
P8	$0{,}4173^2$	$-0{,}7374^2$		$0{,}0065^2 + 0{,}1358^2 + ... = 1$
P11	$0{,}8766^2$	$0{,}1038^2$		$0{,}1580^2 + 0{,}2052^2 + ... = 1$
λ_j^*	**4,1067**	**1,3382**	...	* As correlações elevadas ao

% de variância explicada	51,34%	16,73%		Quadrado e somadas (em coluna) reproduzem os autovalores.

Quadro 3 – Resumo dos resultados pelo Matlab 6.5

O critério de Kaiser reteve duas componentes, redundando em 63,845% da variância total. O gráfico *scree plot* ou prova de sedimentação de Cattell, figura 1, mostra uma representação gráfica da magnitude dos autovalores (λ_i) em seu eixo vertical, de forma decrescente. No seu eixo horizontal apresenta o número de autovalores calculados. A sua ordem também é contrária à da tabela 4.

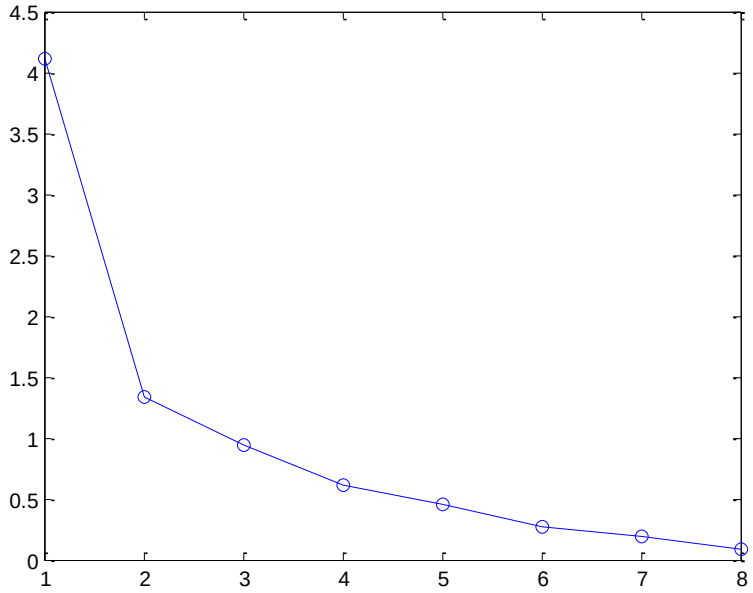

Figura 1 – Gráfico *Scree Plot* de Cattell

O dendograma conforme figura 2, pelo método de *weighted*, cujo coeficiente cofenético foi 0,6693, considerado baixo, confirma o número de componentes principais, pois, apresenta dois grupos.

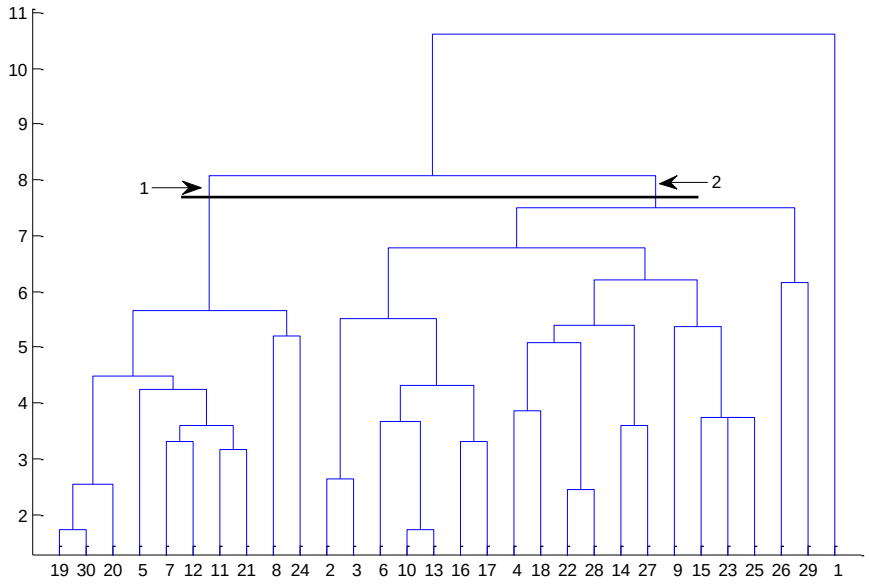

Figura 2 – Dendrograma

5 – DISCUSSÃO

A utilização de *software* facilita em muito o trabalho de pesquisa, no entanto apresenta determinadas particularidades que podem trazer alguma confusão para o seu usuário, como critérios de cálculos diferentes dos tradicionais. Por isso foi utilizada as fórmulas matemáticas e estatísticas apresentadas no texto.

A produção de produtos partindo de uma concepção própria do proprietário nem sempre é a melhor solução. Um pouco mais de metodologia, aprimora o conhecimento adquirido do empresário junto a seus clientes.

A utilização de metodologia estatística melhora o emprego da tecnologia da informação, pois, transforma *a informação* em conhecimento e quem tem conhecimento tem vantagem competitiva.

Saber o que vai vender e quando vender otimiza o custo de produção, uma vez que os recursos são utilizados na hora e quantidade certa não ocorrendo perdas e nem paralisação de capital na forma de estoques fazendo surgir o custo de oportunidade.

Manter a sazonalidade da formação de custos permite uma análise mais apurada deles, pois podem ser comparados na unidade de tempo e isto é possível quando se conhece o que vai ser produzido e quanto vai ser produzido.

6 - CONCLUSÃO

A formação de grupos específicos de produtos, na análise de vendas, permite que o empresário planeje melhor sua produção, evitando perdas que não são contabilizadas oficialmente, mas existentes pela aplicação de recursos na formação de estoques que giram menos cujo recurso financeiro se fosse aplicado no mercado renderiam juros imediatamente, pois são custos de oportunidade.

O conhecimento desses grupos homogêneos de venda permite que se analise qual é o perfil do consumidor deles e, consequentemente qual é o perfil dos demais consumidores. Admitir uma perda de rentabilidade na manutenção do consumidor é um aspecto que necessita ser conhecido para que ocorra uma administração mais científica.

Essa metodologia pode ser aplicada de forma mais generalizada para as atividades de venda no varejo, pois favorece uma análise exploratória com muita segurança revelando hábitos do consumidor, existência de produtos complementares e suplementares, otimização dos recursos com um todo, no planejamento de compras, estoques e produção.

A confirmação pelo método de cluster apresentou a formação de dois grupos, refletindo a confirmação dos dois eixos principais, apesar do coeficiente cofenético ser o maior entre os métodos:
Average, S*ingle, Complete, Ward* presentes no *software* Matlab, mas, considerado baixo pela maioria dos pesquisadores.

7 – REFERÊNCIAS

BOLCH, Ben W. et al. **Multivariate Statistical Methods for Business and Economics.** New Jersey: Prentice-Hall, 1974.

CRIVISQUI, Eduardo. **II Seminário, por videoconferência, de Métodos Exploratórios de Estatística Multivariada.** Florianópolis, Universidade Federal de Santa Catarina, 1999.

FERREIRA, Daniel Furtado. **Análise Multivariada.** Lavras: Universidade Federal de Lavras, 1996.

FREITAS, Isa Aparecida et al. **Construção e validação de Escala de Crenças sobre o Sistema de Treinamento.** Estudos de Psicologia, v.9, n.3, p.479-488. 2004

HAIR, Joseph F.; TATHAM, Ronald L. et al. **Análise Multivariada de Dados.** 5. ed. Porto Alegre: Bookman, 2005.

LEAL, Susana Cristina Henriques. **MG-05. A BANCA EM PORTUGAL – UMA APLICAÇÃO DA ANÁLISE FACTORIAL.** Disponível em: <http://docentes.esgs.pt/aep/Investigacao_ficheiros/MG-05.pdf > **Acesso em:** 12 de setembro de 2005.

MERINO, Antonio Pardo et al. **SPSS 11 Guía para el análisis de datos.** México: McGraw Hill, 2002.

PESTANA, Maria Helena et al. **Análise de Dados para Ciências Sociais: A Complementaridade do SPSS.** 2.ed. Lisboa: Edições Silabo, 2000.

REIS, Elizabeth. **Estatística Multivariada Aplicada.** 2 ed. Lisboa: Edições Silabo, 2001.

SPIEGEL, Murray R. **Estatística.** Traduzido por Pedro Cosentino. Rio de Janeiro: McGraw-Hill, Coleção Schaum.

VALENTIN, Jean Louis. **Ecologia Numérica: Uma introdução à análise multivariada de dados ecológicos.** Rio de Janeiro: Interciência, 2000.

Professor Zavalla, essa é a tabela que falei.

APÊNDICE 1

Tabela 5
Variáveis vendidas no mês de novembro de 2005

dias	P1	P2	P3	P4	P5	P6	P7	P8	P9	P10	P11	P12
01	**350**	**140**	**480**	**30**	6	**12**	**6**	**50**	40	48	**70**	71
02	**356**	**142**	**476**	**25**	6	**9**	**4**	**46**	39	48	**68**	71
03	**357**	**143**	**477**	**26**	4	**9**	**3**	**45**	38	47	**69**	70
04	**360**	**142**	**479**	**27**	5	**9**	**4**	**45**	38	50	**70**	70
05	**355**	**139**	**473**	**25**	5	**8**	**4**	**44**	40	48	**67**	71
06	**355**	**140**	**475**	**26**	8	**10**	**7**	**46**	40	49	**69**	75
07	**354**	**139**	**475**	**25**	7	**8**	**3**	**47**	39	49	**68**	73
08	**355**	**141**	**472**	**25**	5	**9**	**5**	**48**	38	49	**65**	75
09	**357**	**143**	**477**	**27**	3	**10**	**7**	**50**	38	49	**69**	70
10	**356**	**140**	**475**	**26**	4	**9**	**7**	**46**	41	48	**66**	72
11	**353**	**139**	**474**	**25**	5	**9**	**2**	**45**	39	49	**65**	73

12	**354**	**139**	**475**	**25**	8	**8**	**0**	**46**	38	47	**67**	74
13	**356**	**140**	**475**	**25**	8	**8**	**7**	**45**	39	47	**66**	74
14	**360**	**142**	**480**	**29**	8	**13**	**6**	**50**	38	49	**70**	76
15	**356**	**139**	**478**	**28**	4	**12**	**5**	**50**	39	49	**68**	74
16	**356**	**139**	**477**	**26**	6	**12**	**6**	**46**	39	49	**68**	73
17	**355**	**139**	**475**	**26**	7	**12**	**5**	**47**	37	49	**66**	71
18	**359**	**141**	**479**	**27**	7	**11**	**4**	**48**	41	50	**70**	71
19	**352**	**139**	**474**	**25**	3	**8**	**4**	**46**	37	48	**66**	71
20	**351**	**139**	**473**	**25**	8	**9**	**5**	**45**	37	48	**65**	73
21	**354**	**140**	**473**	**26**	8	**8**	**2**	**47**	39	48	**66**	74
22	**357**	**141**	**478**	**28**	4	**12**	**6**	**45**	40	49	**69**	72
23	**355**	**139**	**477**	**29**	5	**11**	**4**	**47**	39	50	**68**	75
24	**352**	**139**	**474**	**27**	8	**8**	**4**	**50**	38	49	**65**	75
25	**355**	**140**	**477**	**28**	3	**9**	**6**	**49**	39	49	**68**	75
26	**355**	**140**	**479**	**27**	5	**6**	**4**	**49**	40	49	**68**	74
27	**359**	**142**	**480**	**30**	7	**12**	**7**	**47**	40	47	**70**	74
28	**358**	**142**	**477**	**28**	7	**13**	**6**	**46**	40	48	**68**	74
29	**356**	**139**	**476**	**28**	5	**9**	**0**	**49**	39	47	**67**	71

30	**35 2**	**13 9**	**47 4**	**25**	5	**7**	**5**	**45**	38	48	**66**	71

Fonte: Construção do autor