Conceito A

02/12/83

Visto

UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL

DEPARTAMENTO DE ESTATÍSTICA

"ANÁLISE DE SÉRIES TEMPORAIS"

MÉTODO DE DECOMPOSIÇÃO - CENSUS II

SÉRGIO COSTA BARATOJO

TRABALHO DE CONCLUSÃO DO CURSO

DE BACHARELADO EM ESTATÍSTICA

PROF. ORIENTADOR: SÉRGIO FISCHER

Porto Alegre, Novembro de 1983

AGRADECIMENTOS

Aos amigos: Fernando Steigleder
Juarez Simioni e
Walter Edmilson Farioli.

Ao professor: Reinaldo Castro Souza.

Ã Carmen Beatriz Sandoval pelo trabalhos de datilografia.

ÍNDICE

APRESENTAÇÃO

Este trabalho surgiu da necessidade que a instituição financeira, na qual estagiei, possuia de obter detalhes sobre o comportamento de algumas séries econômicas, a fim de estimar valores futuros.

Primeiramente é apresentado uma visão geral do método clássico de decomposição por média móvel, que serve como uma introdução e um auxílio para o perfeito entendimento do processo desenvolvido na parte final deste trabalho.

Foi escolhido então um programa muito importante - CENSUS MARK II -, desenvolvido por Julius Shiskin na "Bureau of the Census of the U.S. Departament of Commerce", para se fazer a decomposição e análise das séries de interesse. Esta metodologia, com algumas adaptações a realidade local e uma aplicação são mostradas detalhadamente.

Finalmente, foi desenvolvido um programa de computador referente ao método acima descrito ("Census Mark II") na linguagem APL ("A planing language") cuja listagem correspondente é mostrada no apêndice.

1) MÉTODO DE DECOMPOSIÇÃO

1.1 - INTRODUÇÃO

Qualquer indústria, instituição, empresa ou Governo tem necessidade de estudar séries de valores, pertinentes a sua área, para fazer planos para o futuro . Esta série de valores, de uma variável, observada em sucessivos e iguais períodos de tempo é o que chamamos de série temporal. Para uma definição mais rigorosa de séries temporais, recomendamos ao leitor interessado: Box & Jenkins, (1970). 1

Matematicamente se Z_t é uma série temporal; então podemos denotá-la como abaixo: 2

$$Z_t = f(t)$$

onde f(t) é uma função conhecida do tempo e

t = 1, 2, ... é o índice de tempo.

São exemplos de série temporal:

a) série de médias mensais de depósitos à vista em uma Agência Bancária.
b) Série de índices diários da Bolsa de Valores de Porto Alegre.

Não podemos esperar que o estudo do histórico destas variáveis, nos dê previsões exatas, isto é, que os dados futuros sejam exatamente os que prevemos. Temos que nos conscientizar de que os modelos de análise são apenas instrumentos para auxiliar na tomada de decisões. 3

Vários são os métodos existentes na literatura para a análise e previsão de séries temporais, conforme pode ser visto no "survey" apresentado recentemente por SOUZA, (1981). 4
Neste trabalho estudamos um deles, conhecido pelo nome de

Método de Decomposição.

O Método de Decomposição pode ser considerado como o pioneiro dos métodos univariados (baseados numa única série) e parte do princípio de que uma série temporal é composta 5 de quatro componentes distintas: tendência (T_t), sazonalidade (S_t), variação cíclica (C_t) e uma componente aleatória (e_t).

Assim temos que:

$$Z_t = f(T_t, S_t, C_t, e_t)$$

De acordo com a natureza da função f(t), podemos 6 ter os seguintes tipos de modelos:

(i) MODELO ADITIVO

$$Z_t = T_t + S_t + C_t + e_t$$

(ii) MODELO MULTIPLICATIVO

$$Z_t = T_t \times S_t \times C_t \times e_t$$

(iii) MODELO MISTO

$$Z_t = T_t \times S_t \times C_t + e_t$$

1.2 - TENDÊNCIA

Por tendência entende-se uma influência de longo prazo, relativamente permanente, que indica um crescimento ou uma diminuição sistemática da magnitude do valor da variável.

Existem séries que chamamos de estacionárias (série 7 sem tendência), pois são invariantes no tempo, tendo uma média constante e os valores observados flutuando em torno desta. Podemos verificar se a componente tendencial está presente, através dos pontos plotados ou de vários testes de correlação, entre o tempo (crescente) e a variável (crescente, decrescente ou constante). (Exemplo: coeficiente de correlação ordinal de Kendall. Ver outros testes em Morretin & Toloi, (1981 a)).

Após a constatação da existência de tendência, temos que estimá-la. Os métodos mais utilizados são:

a) Suavizar os valores da série ao redor de um ponto, para estimar a tendência naquele ponto. (Veremos o procedimento na seção sobre decomposição por média móvel);

b) Ajustar uma função do tempo, como um polinômio, uma exponencial ou outra função suave de t.

Um procedimento para se identificar o modelo matemático que melhor se ajusta à série em estudo é apresentado a seguir:

Seja Δ operador diferença, i.é, $\Delta Z_t = Z_t - Z_{t-1}$

Usando a notação de Box & Jenkins, (1970) para o operador diferença em função do operador de atraso "B" definido por: $B^k Z_t = Z_{t-k}$, temos:

$$\Delta y_t = y_t - y_{t-1} = y_t - By_t = (1-B)\, y_t$$

Analogamente, definimos a diferença de ordem genérica d como abaixo:

$$\Delta^d y_t = (1-B)^d y_t$$

Se $Z = at + b$

Dando um acréscimo às variáveis t e Z teremos:

$$Z + \Delta Z = a(t + \Delta t) + b$$

$$\Delta Z = at + a\Delta t + b - (at + b)$$

$$\Delta Z = a\Delta t \cong \text{Constante}$$

a = constante

t = intervalos de tempo em que observamos as variáveis. Tb é constante

10

Logo concluimos que, quando a primeira diferença dos dados estudados forem aproximadamente constante, a função que melhor se ajusta a série é uma reta, isto é, a tendência será representada pela equação:

$$T_t = at + b$$

Sendo agora: $Z = at^2 + bt + c$, teremos, considerando os acréscimos a t e Z:

$$Z + \Delta Z = a(t + \Delta t)^2 + b(t + \Delta t) + c$$

$$Z + \Delta Z = a(t^2 + 2t\Delta t + (\Delta t)^2) + bt + b\Delta t + c$$

$$\Delta Z = 2at\Delta t + a(\Delta t)^2 + b\Delta t$$

Considerando agora em outro acréscimo teríamos:

$$\Delta Z + \Delta^2 Z = 2a(t + \Delta t)\Delta t + a(\Delta t)^2 + b\Delta t$$

$$\Delta^2 Z = 2at\Delta t + 2a(\Delta t)^2 + a(\Delta t)^2 + b\Delta t - \Delta Z$$

$$\Delta^2 Z = 2a(\Delta t)^2 \cong \text{constante}$$

a = constante

t = intervalo de tempo em que obs. as variáveis igualmente constante

Portanto, quando a segunda diferença dos dados estudados forem aproximadamente constante, a função que menor se ajusta a série é uma polinomial do 2º grau, isto é, a tendên cia será representada pela equação:

$$T_t = at^2 + bt + c$$

Segundo este raciocínio obtemos a tabela abaixo:

COLUNAS TESTE		FUNÇÃO DE AJUSTAMENTO
$\Delta Z \cong$ Const.	$\rightarrow$	$Z = at + b$
$\Delta^2 Z \cong$ Const.	$\rightarrow$	$Z = at^2 + bt$
$\Delta^n Z \cong$ Const.	$\rightarrow$	$Z = a_n t^n + a_{n-1} t^{n-1} + \ldots + a_0$
$\Delta\log Z \cong$ Const.	$\rightarrow$	$Z = ab^t$
$\Delta\log(\log Z) \cong$ Const.	$\rightarrow$	$Z = 10^{ab^t}$
$\frac{\Delta\log Z}{\Delta\log t} \cong$ Const.	$\rightarrow$	$Z = at^n$

Para verificamos qual das colunas teste é a mais (11)
constante calculamos a dispersão percentual, de cada uma, em relação a média e escolhemos a menor.

A dispersão percentual em relação a média é dada (12)
por:

$$DM = \left|\frac{M - VA}{M}\right| \times 100$$

Sendo:

M = Média

VA = Valor que mais se afasta da média.

Finalmente para acharmos os coeficientes, da equação escolhida, usamos a teoria de mínimos quadrados que não será discutida neste trabalho. Temos, assim, a tendência para quando t= 1, ..., N e podemos estimá-la para t= N+1, N+2,...

Para remover a componente tendencial dos dados originais dividimos ou subtraímos destes a tendência estimada, dependendo da natureza do modelo, isto é, aditivo e/ou multiplicativo como é mostrado a seguir:

MODELO ADITIVO

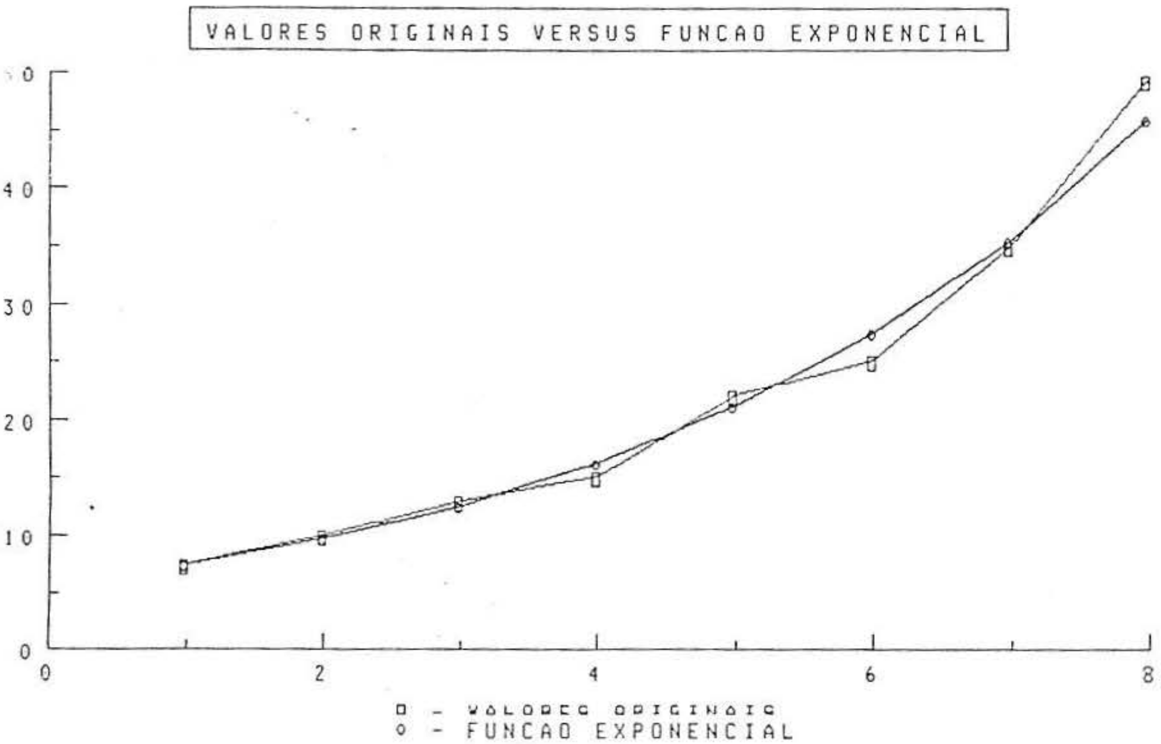
$$Z_t - T_t = T_t + S_t + C_t + e_t - T_t$$

$$= S_t + C_t + e_t$$

MODELO MULTIPLICATIVO

$$\frac{Z_t}{T_t} = \frac{T_t \times S_t \times C_t \times e_t}{T_t} = S_t \times C_t \times e_t$$

Na página seguinte é apresentado um exemplo da determinação da tendência por ajuste de funções. O relatório é a saída do programa "Ajuste", por nós desenvolvido.



```
LINE 246482   COLUNA 1                                              INPUT
      T←1 2 3 4 5 6 7 8
      DADOS←7.5 10 13 15 22 25 35 49.5

      T AJUSTE DADOS
DISPERCAO PERCENTUAL DA RETA, POLIN. 2. E 3. GRAU, EXPON. TIPO  1 E 2  E FUNCAO
       POTENCIA
141.6666667 500 1275 52.58078572 57.07948964 107.5333819
A FUNCAO QUE MELHOR SE AJUSTA E UMA EXPONENCIAL DO TIPO 1
COEFICIENTES DA EXPONENCIAL DO TIPO 1
5.759729087 1.296591602
PROJECAO DA TENDENCIA PARA OS PROXIMOS DOZE VALORES
   59.65280849   77.34533051   100.285306    130.0290855   168.5946203
       218.5983686   283.4328091   367.4966      476.4930052   617.8168289
       801.0561119   1038.642627
```

1.3 - SAZONALIDADE

Sazonalidade pode ser definida como sendo uma influência de curta duração, que todos os anos (ou em intervalos de tempo menores) se repetem nas mesmas épocas, e que sistematicamente pressiona o valor da variável sempre num mesmo sentido.

Como exemplo de fenômenos que ocorrem regularmente de ano para ano, isto é, de fenômenos sazonais, temos:

a) Aumento das vendas no comércio, na época do Natal;

b) Aumento das vendas de sorvetes no verão ...

São vários os fatores que originam as variações sazonais, entre eles as condições climáticas, os costumes sociais e as festividades religiosas.

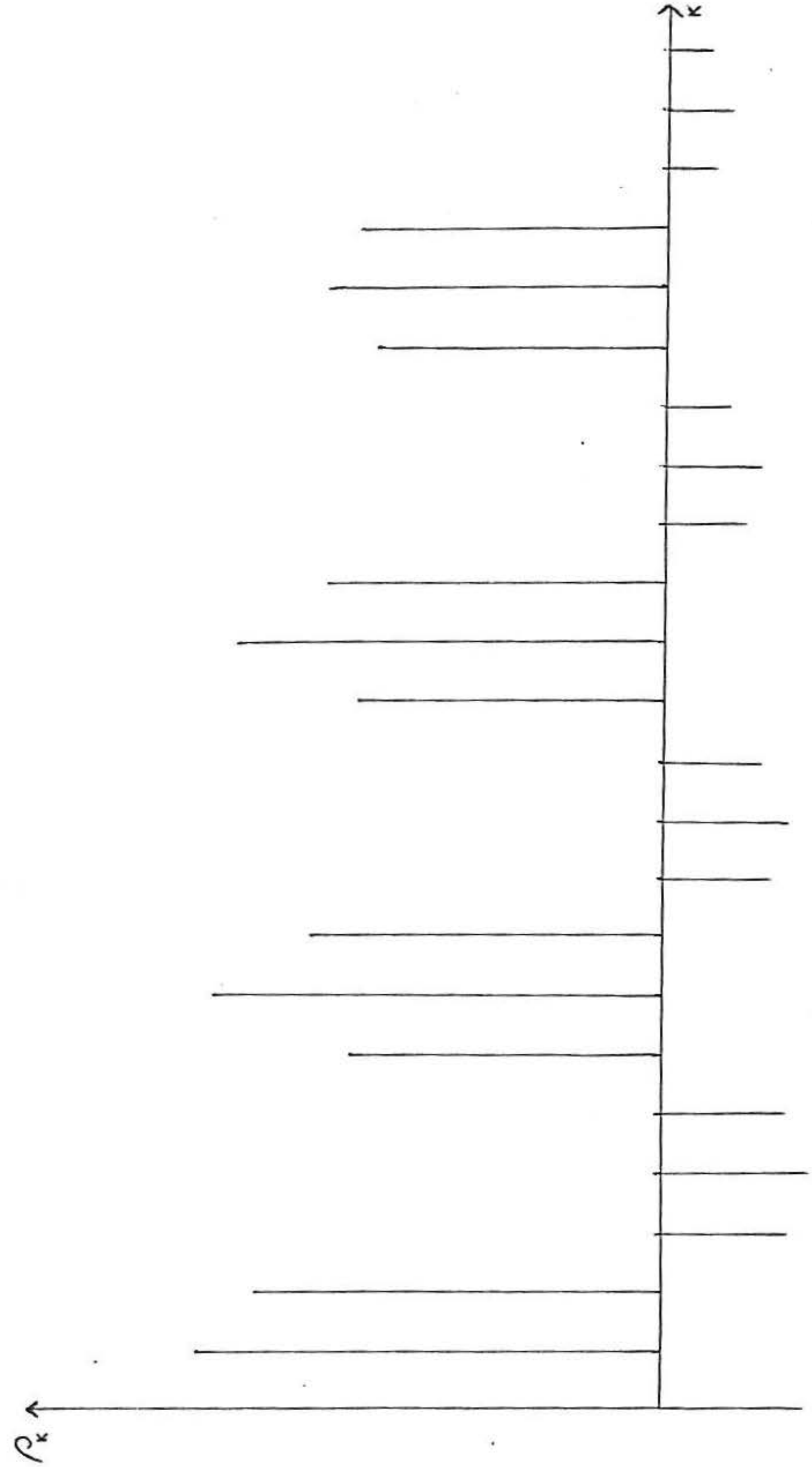
Muitas vezes, podemos verificar a existência de sazonalidade, em uma série, através de uma simples inspeção do seu gráfico. Quando isto não é possível e nosso problema é simplesmente captar sua ocorrência, devemos analisar seu correlograma. Correlograma é o gráfico da função de autocorrelação, isto é, da função que nos dá um grau de intensidade da correlação existente entre observações vizinhas da série.

A autocorrelação de defasagem "k" e definida por:

$$\rho_k = \frac{COV\ (X_t,\ X_{t+K})}{VAR\ (X_t)}$$; sua estimativa será

$$\hat{\rho}_k = \frac{\sum_{t-1}^{N-k} (X_t - \overline{X})(X_{t+k} - \overline{X})}{\sum_{t-1}^{N-K} (X_t - \overline{X})^2}$$

13



CORRELOGRAMA DE UMA SÉRIE COM SAZONALIDADE DE TAMANHO 6

Outra maneira de constatarmos a existência de sazonalidade, em uma série, seria através dos estimadores dos coeficientes sazonais, conforme mostramos a seguir:

Define-se um estimador do fator sazonal, representado por:

$$\hat{\rho}_i \ , \ i = 1,2, \ldots, s$$

onde "s" é o comprimento do período sazonal (s = 12 série com sazonalidade anual, s = 4 série com sazonalidade trimestral, etc.) como:

$$\hat{\rho}_i = \frac{\bar{X}_i}{\bar{\bar{X}}}$$

onde:

$\bar{\bar{X}}$ = média da série

$\bar{X}_i$ = média dos "X" na posição i de períodos diferentes

Exemplo: Vide página seguinte.

VENDAS (UNID.FÍSICAS) DO PRODUTO "A", PELA EMPRESA "B"

MESES	1967	1968	1969	Σ	$\overline{X}_i$	$\hat{\rho}_i = \overline{X}_i / \overline{\overline{X}}$
JAN	158	143	226	527	175,67	1,48
FEV	175	201	240	616	205,33	1,73
MAR	121	122	191	434	144,67	1,22
ABR	34	35	52	121	40,33	0,34
MAI	33	31	42	106	35,33	0,30
JUN	10	18	28	56	18,67	0,16
JUL	11	14	17	42	14,00	0,12
AGO	84	108	108	300	100,00	0,85
SET	164	145	175	484	161,33	1,36
OUT	126	126	214	466	155,33	1,31
NOV	124	143	219	466	162,00	1,37
DEZ	160	198	269	627	209,00	1,76
Σ	1.200	1.284	1.781	4.265		12

$$\overline{\overline{X}} = \frac{4.265}{36} = 118,4722$$

Vemos acima que o pico de vendas ocorre no mês de dezembro ($\hat{\rho}_{DEZ} = 1,76$), enquanto que as vendas mais baixas ocorrem em julho ($\hat{\rho}_{JUL} = 0,12$).

No exemplo acima estamos supondo que a sazonalidade da série é constante ao longo do tempo. Nada impede, por outro lado, que consideremos a sazonalidade variável no tempo, ou seja, ao invés de termos: $\rho_{JAN} \ldots \rho_{DEZ}$, podemos ter $\rho_{JAN}(t) \ldots \rho_{DEZ}(t)$.

14

Há vários métodos para a estimação da sazonalidade, sendo que os mais usuais são:

a) Método de regressão;

b) Método de médias móveis.

Para maiores informações ver Morettin e Toloi(1981 a).

Quando dividimos ou subtraímos dos dados originais, dependendo do modelo em questão, as constantes sazonais, obtemos uma série livre de sazonalidade ou, conforme usado na literatura "Série Dessazonalizada". Assim temos matematicamente:

(i) MODELO ADITIVO

$$Z_t - S_t = T_t + S_t + C_t + e_t - S_t$$

$$= T_t + C_t + e_t$$

(ii) MODELO MULTIPLICATIVO

$$\frac{Z_t}{S_t} = \frac{T_t \times S_t \times C_t \times e_t}{S_t} = T_t \times C_t \times e_t$$

1.4 - COMPONENTE CÍCLICA

As variações cíclicas se caracterizam por movimentos ascendentes e descendentes, em sucessivos períodos de tempo, causando uma ondulação, no comportamento, de razoável amplitude. Se diferenciam das variações sazonais por se estenderem em intervalos de tempo mais largos, em geral mais de 2 anos.

Grauger e Newbold (1977) alertam que não há evidên- 15
cias de existir, em séries macroeconômicas modernas, outras 16 componentes periódicas além da sazonal. Por este motivo e por geralmente trabalharmos com uma quantidade de valores que não nos permite identificar um ciclo, é que alguns métodos de decomposição consideram uma série temporal composta somente dos 3 elementos:

1 - Tendência - Ciclo (TC_t)

2 - Sazonalidade (S_t)

3 - Componente Aleatório (e_t)

1.5 - COMPONENTE ALEATÓRIA

A componente aleatória pode ser definida como sendo a influência que o acaso exerce sobre o comportamento da variável ao longo do tempo.

Os movimentos irregulares podem ser causados por acontecimentos esporádicos como uma guerra, um terremoto, uma 17
seca, etc... Estas flutuações são imprevisíveis, mas podem ser identificadas, numa série, quando desta eliminamos as demais componentes. Em termos analíticos, temos, para os dois modelos:

(i) MODELO ADITIVO

$$e_t = Z_t - TC_t - S_t$$

(ii) MODELO MULTIPLICATIVO

$$e_t = \frac{Z_t}{TC_t \times S_t}$$

2) DECOMPOSIÇÃO POR MÉDIA MÓVEL

Os métodos de decomposição por média móvel sempre foram muito utilizados por não apresentarem cálculos complexos e de difícil compreensão. Estes métodos serviram de base para a criação da decomposição Census II, que veremos na próxima seção.

A técnica de média móvel consiste no cálculo, a cada período, da média aritimética das "r" observações mais recentes, isto é, a observação mais antiga é substituída por uma mais recente.

Assim se r = 12

$$M_{6,5} = (Z_1 + Z_2 + \ldots + Z_{12})/12 \qquad (2.1)$$

$$M_{7,6} = (Z_2 + Z_3 + \ldots + Z_{13})/12 \qquad (2.2)$$

18

Onde M_i é a média no instante I.

As equações (2.1) e (2.2) não estão centradas, isto é, a média calculada em (2.1) está entre 6 e 7 não se relacionando com nenhum período. Isto ocorre, porque o "lag" da média móvel (r = 12) é par. Podemos solucionar o problema efetuando uma nova média móvel, agora de lag 2. Teremos, assim, a média de (2.1) e (2.2) colocada no momento 7. Este procedimento é mostrado abaixo: 19

$$M_7'' = \frac{M_{6,5} + M_{7,5}}{2} \qquad (2.3)$$

Colocando (2.1) e (2.2) na equação (2.3) teremos:

$$M_7'' = \left(\frac{Z_1 + Z_2 + \ldots + Z_{12}}{12} + \frac{Z_2 + Z_3 + \ldots + Z_{13}}{12}\right) \div 2 \qquad (2.4)$$

Desenvolvendo (2.4)

$$M''_7 = (Z_1 + 2Z_2 + 2Z_3 + \ldots + 2Z_{11} + 2Z_{12} + Z_{13}) \div 24$$

A média móvel não elimina as flutuações muito acentuadas da série, mas reduz substancialmente a amplitude das variações dos dados originais.

Os passos para se decompor uma série utilizando média móvel são apresentados abaixo:

a) Para a série, Z_t, é calculada a média móvel centrada cujo comprimento "r" é igual ao comprimento da sazonalidade. O objetivo desta média móvel é eliminar a sazonalidade e a aleatoriedade.

Como ilustração foram utilizados dados, apresentados por Chao, Lincoln L. (1975), de unidades habitacionais construídas nos Estados Unidos do 3º trimestre de 1964 ao segundo trimestre de 1972 (em mil unidades). (Vide tabela 2.1 na página seguinte)

TABELA (2.1)

ANO POR TRIMESTRE (1)		DADOS ORIGINAIS Z (2)	MÉDIA MÓVEL DE QUATRO TRIMESTRES (r=4) (3)	MÉDIA MÓVEL CENTRADA (r=2) (4)	PERCENTAGEM ((2) : (4)) x 100 (5)
1964:	3	398			
	4	352			
			372		
1965:	1	283		371	76,3
			370		
	2	454		369	123,0
			368		
	3	392		367	106,8
			366		
	4	345		359	96,1
			351		
1966:	1	274		338	81,1
			325		
	2	392		309	126,9
			292		
	3	290		285	101,8
			277		
	4	210		276	76,1
			275		
1967:	1	218		287	76,0
			298		
	2	382		315	121,3
			331		
	3	382		341	112,0
			350		
	4	340		359	94,7
			368		
1968:	1	298		373	79,9
			378		
	2	452		382	118,3
			386		
	3	423		391	108,2
			396		
	4	372		398	93,5
			400		
1969:	1	336		395	85,1
			391		
	2	468		383	122,2
			375		
	3	387		366	105,7
			357		
	4	309		349	88,5
			340		
1970:	1	264		343	77,0
			345		
	2	399		356	112,1
			367		
	3	408		383	106,5
			398		
	4	396		424	93,4
			449		
1971:	1	389		471	82,6
			492		
	2	604		507	119,1
			521		
	3	579		537	107,8
			552		
	4	513		559	91,8
			566		
1972:	1	510			
	2	661			

Quando suavizamos os dados originais por uma média móvel obtemos como resultado uma aproximação da tendência--ciclo (coluna (4)). Supondo que o modelo é multiplicativo, quando dividimos os dados originais (coluna (2)) pela tendência-ciclo (coluna (4)) temos aproximadamente a sazonalidade com uma componente aleatória (coluna (5)).

b) A próxima operação consiste em isolar o fator sazonal eliminando as variações irregulares. Achamos a média para um mesmo trimestre de anos diferentes. Estas médias seriam uma estimação dos índices sazonais e sua soma deve dar 400. Como o total é 397.685 devemos fazer um ajuste multiplicando cada média trimestral por (400 ÷ 397,685).

TABELA (2.2)

ANO	TRIMESTRES				TOTAL
	1	2	3	4	
1964					
1965	76,3	123,0	106,8	96,1	
1966	81,1	126,9	101,8	76,1	
1967	76,0	121,3	112,0	94,7	
1968	79,9	118,3	108,2	93,5	
1969	85,1	122,2	105,7	88,5	
1970	77,0	112,1	106,5	93,4	
1971	82,6	119,1	107,8	91,8	
1972					
TOTAL	558,0	842,9	748,8	634,1	
MÉDIA	79,714	120,414	106,971	90,586	397,685
IND.SAZONAL	80,18	121,12	107,59	91,11	400

Se quisermos eliminar a sazonalidade da série, bastaria dividir os valores originais pelo índice sazonal. Para o ano de 1966 teríamos:

TABELA (2.3)

TRIMESTRE	VALOR DAS VENDAS	ÍNDICE SAZONAL	VALOR DAS VENDAS DESSAZONALIZADO
1	274	80,18	342
2	392	121,12	324
3	290	107,59	270
4	210	91,11	230

c) Identificar a forma apropriada da tendência (linear, exponecial,...) e calcular estes valores para cada período, T_t.

Podemos ver, através do programa "ajuste", que a função que melhor representa a tendência é uma função potência da forma:

$$T_t = at^b$$

Onde os coeficientes "a" e "b" serão aproximadamente 288, 106 e 0,1054, respectivamente, como é mostrado a seguir.

Efetuando a operação:

$$T_t = 288{,}106 \times t^{0{,}1054} \quad \text{para} \quad t = 1, 2, \ldots, 32$$

Obtemos a tendência da série que é mostrada na tabela (2.4).

LINE 246546 COLUNA 1 INPUT

1 AJUSTE SERIE

DISPERCAO PERCENTUAL DA RETA, POLIN. 2. E 3. GRAU, EXPON. 1. E 2. GRAU E FUNCAO
POTENCIA

2434.220532 3754.822335 7849.717514 3327.681655 3676.690466 1956.50295

A FUNCAO QUE MELHOR SE AJUSTA AOS DADOS E A FUNCAO POTENCIA

COEFICIENTES DA FUNCAO POTENCIA

288.1064327 0.1054024005

PROJECAO DA TENDENCIA PARA OS PROXIMOS DOZE VALORES

416.4933244 417.8059154 419.0844134 420.3306408 421.5462756 422.7328658
423.8918429 425.0245334 426.1321688 427.2158945 428.276778 429.3158154

TABELA (2.4)

TEMPO (t) (1)	ANO POR TRIMESTRE (2)	DADOS ORIGINAIS (3)	VALOR DA TENDÊNCIA (4)
1	1964: 3	398	288,106
2	4	352	309,943
3	1965: 1	283	323,477
4	2	454	333,435
5	3	392	341,371
6	4	345	347,994
7	1966: 1	274	353,695
8	2	392	358,708
9	3	290	363,189
10	4	210	367,245
11	1967: 1	218	370,953
12	2	382	375,370
13	3	382	377,542
14	4	340	380,503
15	1968: 1	298	383,280
16	2	452	385,896
17	3	423	388,370
18	4	372	390,717
19	1969: 1	336	392,949
20	2	468	395,080
21	3	387	397,117
22	4	309	399,069
23	1970: 1	264	400,943
24	2	399	402,745
25	3	408	404,482
26	4	396	406,158
27	1971: 1	389	407,777
28	2	604	409,343
29	3	579	410,859
30	4	513	412,330
31	1972: 1	510	413,758
32	2	661	415,145

d) Separar a tendência, obtida no passo (3), do valor obtido no passo (1), isto é, isolar a componente cíclica.

Se dividirmos a coluna (4) da tabela (2.1) (aproximação T x C) pela coluna (4) da tabela (2.4) (tendência) temos como resultado a componente cíclica. Como no passo 1 perdemos 2 valores do início da série e 2 valores do fim, efetuamos a divisão não considerando os correspondentes valores da tendência.

TABELA (2.5) CÁLCULO PARA IDENTIFICAR A COMPONENTE CÍCLICA

(1) ANO POR TRIMESTRE	(2) T x C	(3) T	(2) ÷ (3) COMP. CÍCLICA
1964: 3	-	-	-
4	-	-	-
1965: 1	371	323,477	1,147
2	369	333,435	1,108
3	367	341,371	1,076
4	359	347,994	1,030
1966: 1	338	353,695	0,956
2	309	358,708	0,860
3	285	363,189	0,783
4	276	367,245	0,752
1967: 1	287	370,953	0,772
2	315	375,370	0,839
3	341	377,542	0,902
4	359	380,503	0,944
1968: 1	373	383,280	0,974
2	382	385,896	0,991
3	391	388,370	1,007
4	398	390,717	0,018
1969: 1	395	392,949	0,006
2	383	395,080	0,969
3	366	397,117	0,922
4	349	399,069	0,873
1970: 1	343	400,943	0,854
2	356	402,745	0,884
3	383	404,482	0,945
4	424	406,158	1,043
1971: 1	471	407,777	1,154
2	507	409,343	1,238
3	537	410,859	1,305
4	559	412,330	1,355
1972: 1	-	-	-
2	-	-	-

e) Retirados dos dados originais a sazonalidade, a tendência e o ciclo o que resta é apenas a componente aleatória.

A tabela (2.6) mostra as quatro componentes que formam o valor original.

TABELA (2.6)

ANO POR TRIMESTRE (1)		DADOS ORIGINAIS (2)	T_t (3)	S_t (4)	C_t (5)	E_t (6)
1964:	3	398	288,106	107,59	-	-
	4	352	309,943	91,11	-	-
1965:	1	283	323,477	80,18	114,7	95,14
	2	454	333,435	121,12	110,8	101,48
	3	392	341,371	107,59	107,6	99,18
	4	345	347,994	91,11	103,0	105,62
1966:	1	274	353,695	80,18	95,6	101,10
	2	392	358,708	121,12	86,0	104,95
	3	290	363,189	107,59	78,3	94,74
	4	210	367,245	91,11	75,2	83,44
1967:	1	218	370,953	80,18	77,2	94,90
	2	382	375,370	121,12	83,9	100,36
	3	382	377,542	107,59	90,2	104,27
	4	340	380,503	91,11	94,4	103,88
1968:	1	298	383,280	80,18	97,4	99,61
	2	452	385,896	121,12	99,1	97,63
	3	423	388,370	107,59	100,7	100,55
	4	372	390,717	91,11	101,8	102,65
1969:	1	336	392,949	80,18	100,6	106,02
	2	468	395,080	121,12	96,9	100,92
	3	387	397,117	107,59	92,2	98,28
	4	309	399,069	91,11	87,3	97,35
1970:	1	264	400,943	80,18	85,4	96,17
	2	399	402,745	121,12	88,4	92,57
	3	408	404,482	107,59	94,5	99,17
	4	396	406,158	91,11	104,3	102,60
1971:	1	389	407,777	80,18	115,4	103,09
	2	604	409,843	121,12	123,8	98,43
	3	579	410,859	107,59	130,5	100,33
	4	513	412,330	91,11	135,5	100,79
1972:	1	510	413,758	80,18	-	-
	2	661	415,145	121,12	-	-

f) Para fazermos previsões levamos em consideração apenas a tendência e a sazonalidade. Desconsideramos a componente cíclica pelos motivos apresentados na seção (1.4) e por ser, no nosso caso, de difícil previsão. A aleatoriedade, obviamente, só pode ser detectada após a ocorrência do dado.

TABELA (2.7) PREVISÃO PARA OS PRÓXIMOS 4 TRIMESTRES

TEMPO (t)	ANO POR TRIMESTRE	VALOR DA TENDÊNCIA (T_t)	ÍNDICE SAZONAL (S_t)	PREVISÕES ($\hat{Z}_t$)
33	1972: 3	416,493	107,59	448,105
34	4	417,806	91,11	380,663
35	1973: 1	419,084	80,18	336,022
36	2	420,331	121,12	509,105

As previsões acima foram calculadas pela seguinte fórmula:

$$\hat{Z}_t = 288,106 \times t^{0,1054} \times \rho_t$$

Onde:

$\hat{Z}_t$ é o valor a ser previsto no tempo t

t é o tempo para o qual faremos a previsão

ρ_t é o índice sazonal correspondente ao tempo t

É importante salientar que este é um método muito simples, mas um método que serve de base para outros bem mais sofisticados e precisos.

3) MÉTODO DE DECOMPOSIÇÃO CENSUS II

O método de decomposição conhecido por Census II, como já foi dito na apresentação, foi desenvolvido por Julius Shiskin na "Bureau of the Census of the U.S. Departament of Commerce" por volta de 1955. Desde então, apareceram várias versões deste método, que são, até hoje, muito utilizadas nos escritórios, agências governamentais e em diversas empresas nos EUA.

Census II é baseado na decomposição clássica por média móvel, mas possui um número de passos e uma sofisticação muito superior, como veremos a seguir.

Para ilustrar a aplicação do Census II, utilizaremos as médias mensais dos depósitos à vista de jan/77 a dez/82, de uma Agência "X" do Banco Sul Brasileiro S.A., localizada em uma região agropecuária do Rio Grande do Sul.

MÉDIAS MENSAIS DE DEPÓSITOS À VISTA DA AGÊNCIA BANCÁRIA "X"

TABELA (3.1)

Em Cr$ 1.000.000

ANO	JAN	FEV	MAR	ABR	MAIO	JUN	JUL	AGO	SET	OUT	NOV	DEZ
1977	20	20	19	23	26	30	27	23	22	28	28	27
1978	29	30	30	33	39	42	40	37	36	38	41	49
1979	53	54	58	69	79	94	86	82	82	83	81	101
1980	112	126	133	139	151	160	158	147	140	133	134	143
1981	146	139	152	161	185	226	199	170	144	151	160	194
1982	312	205	209	237	274	296	310	267	260	276	292	334

3.1 - AJUSTAMENTO DOS DIAS ÚTEIS

Dependendo do tipo de dado com que estamos trabalhando, é importante fazer um ajuste de dias úteis, pois estes são diferentes para os meses de cada ano. No caso dos depósitos não levamos em conta esta diferença, mas o processo será ilustrado para o leitor utilizá-lo quando necessário.

O primeiro passo é determinar o nº de dias úteis de cada mês para todos os anos de interesse e calcular a média para os doze meses. A tabela (3.2) mostra o número de dias úteis de 1977 a 1982.

TABELA (3.2) NÚMERO DE DIAS ÚTEIS

ANO	JAN	FEV	MAR	ABR	MAIO	JUN	JUL	AGO	SET	OUT	NOV	DEZ
1977	21	20	23	20	22	22	21	23	21	20	21	22
1978	22	20	23	19	23	22	21	23	20	21	21	20
1979	22	20	22	21	23	21	22	23	19	22	21	20
1980	22	19	21	19	21	20	23	21	22	23	20	21
1981	21	20	21	19	20	21	23	21	21	21	20	21
1982	20	17	23	19	21	21	22	22	21	20	20	22
MÉDIAS	21,33	19,33	22,17	19,50	21,67	21,17	22,00	22,17	20,67	21,17	20,50	21,00

Calcula-se um coeficiente de ajuste dividindo o número de dias úteis pela média do mês correspondente. Este coeficiente divide o dado original obtendo-se, assim, o dado ajustado aos dias úteis. Este método fica bem claro quando observamos a tabela (3.3), onde é apresentado um exemplo para o mês de maio.

TABELA (3.3) CÁLCULO DO AJUSTE DE DIAS ÚTEIS PARA MAIO

ANO	(1) DIAS ÚTEIS	(2) COEF. DE AJUSTAMENTO	(3) DADOS DE MAIO	DADOS AJUSTADOS (3) ÷ (2)
1977	22	22/21,67 = 1,0152	26	25,61
1978	23	23/21,67 = 1,0614	39	36,74
1979	23	23/21,67 = 1,0614	79	74,43
1980	21	21/21,67 = 0,9691	151	155,81
1981	20	20/21,67 = 0,9229	185	200,46
1982	21	21/21,67 = 0,9691	274	282,74
	21,67			

O mesmo processo deve ser aplicado aos outros onze meses do ano.

3.2 - AJUSTAMENTO SAZONAL PRELIMINAR

Na segunda fase do Census II é separada preliminarmente a sazonalidade da tendência-ciclo, isolando a aleatoriedade. Usamos como exemplo os dados das médias mensais, de depósito à vista, da agência do Banco Sul Brasileiro. Assim sendo, temos os seguintes passos nesta fase:

3.2.1 - CÁLCULO DE UMA MÉDIA MÓVEL CENTRADA DE 12 MESES

A média móvel centrada de 12 meses (tamanho da sazonalidade) aplicada aos dados originais eliminará, em grande parte, a sazonalidade e a aleatoriedade que existe na série. Os cálculos são mostrados na tabela (3.4).

TABELA (3.4) CÁLCULO DA SAZONALIDADE PRELIMINAR

	(1) DADOS ORIGINAIS		(2) MÉDIA MÓVEL CENTRADA DE 12 MESES	$\frac{(1)}{(2)}$	
1977:	JAN	20	–	–	
	FEV	20	–	–	
	MAR	19	–	–	
	ABR	23	–	–	
	MAIO	26	–	–	
	JUN	30	–	–	
	JUL	27	24,542	110,02	
	AGO	23	25,333	90,79	
	SET	22	26,208	83,94	
	OUT	28	27,083	103,38	
	NOV	25	28,041	89,15	
	DEZ	27	29,082	92,84	
1978:	JAN	29	29,083	96,27	
	FEV	30	30,125	96,00	
	MAR	30	.	.	(Veja os valores na tabela 3.4.a.)
	ABR	33	.	.	
	MAIO	39	.	.	
	JUN	42	.	.	

RAZÃO – DADOS ORIGINAIS ÷ MÉDIA MÓVEL CENTRADA DE 12 MESES (R_t)

ANO	JAN	FEV	MAR	ABR	MAIO	JUN	JUL	AGO	SET	OUT	NOV	DEZ
1977	–	–	–	–	–	–	110,02	90,79	83,94	103,38	89,15	92,84
1978	96,27	96,00	92,54	98,75	113,04	116,40	105,26	92,50	85,38	84,76	85,42	94,53
1979	94,78	90,44	91,34	102,54	111,53	125,89	108,46	96,76	90,23	85,64	78,77	93,02
1980	97,96	104,96	106,26	107,20	112,72	116,01	111,99	102,77	96,97	91,04	90,29	93,72
1981	92,97	87,03	94,51	99,54	135,50	135,50	115,92	95,95	78,99	80,39	82,17	96,36
1982	101,50	94,23	92,31	100,23	110,86	114,51	–	–	–	–	–	–

Matematicamente:

$$Z_t = TC_t \times S_t \times e_t$$

$$M_t = TC_t$$

$$\frac{Z_t}{M_t} = R_t = \frac{TC_t \times S_t \times e_t}{TC_t} = S_t \times e_t$$

Onde: Z_t = dados originais

M_t = média móvel centrada

Os valores de R_t são formados por uma sazonalidade e uma componente aleatória e são mostradas na tabela 3.4.a , como podemos notar, durante estas operações foram perdidos seis valores do início e seis valores no fim da série.

3.2.2 - SUBSTITUIÇÃO DOS VALORES EXTREMOS (ELIMINAÇÃO DOS OUTLIERS)

A razão (R_t) da tabela (3.4) pode possuir junto de si eventos raros, isto é, valores anormais que chamamos de "OUTLIERS". A próxima tarefa do Census II é excluir estes valores, mas antes eliminando a aleatoriedade. Este processo tem dois estágios:

a) Calcula-se uma média móvel dupla 3 x 3 da razão R_t, para cada mês isoladamente, com o objetivo de suavizar um pouco esta série, isto é, eliminar parte da aleatoriedade. Como este cálculo causa a perda de duas observações no início e duas no fim, Census II estima quatro valores (2 para o início e 2 para o fim) antes de executar a média móvel dupla. A tabela (3.5) mostra esta operação para o mês de fevereiro.

TABELA (3.5) CÁLCULO DA MÉDIA MÓVEL DUPLA 3 x 3 PARA FEVEREIRO

ANO	R_t	INTRODUÇÃO DE 2 VALORES NO INÍCIO E 2 NO FIM		MÉDIA MÓVEL DUPLA 3 x 3
1977	-	$\frac{96,00 + 90,44}{2} =$	93,22	-
			93,22	-
1978	96,00		96,00	94,83
1979	90,44		102,96	94,83
1980	104,96		86,03	95,56
1981	87,03		94,23	93,40
1982	94,23	$\frac{87,03 + 94,23}{2} =$	90,63	-
			90,63	-

b) No segundo estágio obtem-se limites de controle. Calcula-se a diferença entre a razão R_t e a média móvel dupla, eleva-se ao quadrado e acha-se a média. A raiz quadrada desta média chamamos de erro padrão. Nossos limites, inferior e superior, serão o valor da média móvel dupla menos dois erros padrões e mais dois erros padrões respectivamente. Verifica-se quais os valores da razão R_t que estão fora do intervalo e substitui-se pela média dos valores vizinhos. Se a observação outliers não possuir um dos vizinhos substituí-se esta pela média dos três dados posteriores ou dos anteriores. As tabelas (3.6) e (3.7) exemplificam este passo.

TABELA (3.6) CÁLCULO DO ERRO PADRÃO PARA O MÊS DE FEVEREIRO

ANOS	(1) R_t (TAB.3.4)	(2) MÉDIA MÓVEL DUPLA 3x3	(3) DIFERENÇA (1) - (2)	(4) DIFERENÇA AO QUADRADO
1977	-	-	-	-
1978	96,00	94,83	1,17	1,3689
1979	90,44	94,83	-4,39	19,2721
1980	104,96	95,56	9,40	88,3600
1981	87,03	93,40	-6,37	40,5769
1982	94,23	92,63	1,60	2,5600
				152,1379

$$\text{MÉDIA}(4) = \frac{152{,}1379}{5} = 30{,}4276$$

$$\text{EP} = \sqrt{30{,}4276} = 5{,}5161$$

TABELA (3.7)

ANOS	MÉDIA MÓVEL DUPLA 3x3	$\pm$ E. P.	R_t	
1977	-	-		
1978	94,83	±11,0322	96,00	Não é um outliers
1979	94,83	±11,0322	90,44	Não é um outliers
1980	95,56	±11,0322	104,96	Não é um outliers
1981	93,40	±11,0322	97,03	Não é um outliers
1982	92,63	±11,0322	94,23	Não é um outliers

23

3.2.3 - FATOR SAZONAL PRELIMINAR

Após serem substituídos os valores anormais, da razão R_t da tabela (3.4), temos que efetuar alguns ajustes para calcularmos os fatores sazonais preliminares.

Os ajustes a serem executados são os seguintes:

a) Os seis meses iniciais e os seis meses finais da razão R_t (ver tabela (3.4)) foram perdidos, quando calculada a média móvel centrada de 12 meses. Estas observações são consideradas como sendo o correspondente valor do ano seguinte ou do ano anterior. (Ver tabela 3.8)

TABELA (3.8) ESTIMAÇÃO DOS VALORES PERDIDOS NA MÉDIA MÓVEL CENTRADA DE 12 MESES

ANOS	JAN	FEV	MAR	ABR	MAIO	JUN	JUL	AGO	SET	OUT	NOV	DEZ
1977	96,27	96,00	92,54	98,75	113,04	116,40	110,02	90,79	83,84	103,38	39,15	92,84
	↑	↑	↑	↑	↑	↑						
1978	96,27	96,00	92,53	98,75	113,04	116,40	105,26	92,50	85,38	84,76	85,42	94,53
1979	.	.										
	.	.										
1980	.	.										
1981	92,97	87,03	94,51	99,54	113,09	135,50	115,92	95,95	78,99	80,39	82,17	96,36
							↓	↓	↓	↓	↓	↓
1982	101,50	94,23	92,31	100,23	110,86	114,51	115,92	95,95	78,99	80,39	82,17	96,36

b) As razões são normalizadas, passando a somar, em cada ano, 1.200. Esta normalização é feita dividindo cada valor pela média do ano correspondente. Este processo é ilustrado na tabela (3.9)

TABELA (3.9) NORMALIZAÇÃO DAS RAZÕES R_t

ANO	JAN	FEV	MAR	ABR	MAIO	JUN	JUL	AGO	SET	OUT	NOV	DEZ	TOTAL
1978	96,27	96,00	98,54	98,75	113,04	116,40	105,26	92,50	85,38	84,76	85,42	95,53	1.160,85
	↓					↓							MÉDIA = $\frac{1.160,85}{12}$
	$\frac{96,27}{96,74}$					$\frac{116,40}{96,74}$							= 96,74
	↓					↓							
	99,52	99,24	95,66	102,08	116,85	120,33	108,81	95,62	88,26	87,62	88,30	97,72	1.200
													MÉDIA = $\frac{1.200}{12}$ = 100

O próximo passo é a eliminação da aleatoriedade,ainda existente na razão R_t (tabela 3.4), através de uma média móvel dupla 3x3. Esta é análoga à mostrada na tabela (3.5) exceto que os dados sofreram as modificações descritas acima.

TABELA (3.10) - FATOR SAZONAL PRELIMINAR

ANOS	JAN	FEV	MAR	ABR	MAIO	JUN	JUL	AGO	SET	OUT	NOV	DEZ
1977	98,33	97,59	94,55	101,47	115,49	120,17	110,48	94,26	87,18	96,27	88,53	95,69
1978	98,04	97,62	95,61	102,51	114,93	120,85	110,17	95,95	89,04	92,58	87,19	95,51
1979	97,16	96,43	96,68	103,41	114,12	123,44	111,10	97,73	89,91	82,27	85,50	95,24
1980	96,26	96,33	98,04	103,66	113,45	123,69	112,84	98,88	88,32	86,32	85,20	95,46
1981	98,10	94,67	97,48	103,23	113,89	126,21	115,96	98,96	85,35	84,50	84,78	96,85
1982	99,95	94,60	96,72	102,91	114,39	125,54	118,05	98,80	82,68	83,31	84,86	98,19

Como a razão R_t era composta apenas por uma sazonalidade e uma aleatoriedade, ao executarmos a média móvel dupla 3x3 obteremos fatores sazonais preliminares isolados. Dividindo os dados originais por estes fatores teremos como resultado somente tendência-ciclo e componente aleatória. Os valores dessazonalizados preliminarmente (ver tabela (3.11)) servirão de base para o cálculo das componentes finais.

Matematicamente temos:

$$PS_t = \frac{Z_t}{S_t} = \frac{TC_t \times S_t \times e_t}{S_t} = TC_t \times e_t \qquad (3.2.1)$$

Onde PS_t são os dados originais dessazonalizados.

TABELA (3.11) SÉRIE DESSAZONALIZADA PRELIMINARMENTE

ANOS	JAN	FEV	MAR	ABR	MAIO	JUN	JUL	AGO	SET	OUT	NOV	DEZ
1977	20,34	20,49	20,10	22,67	22,51	24,96	24,44	24,40	25,24	29,09	28,24	28,22
1978	29,58	30,73	31,38	32,19	33,93	34,75	36,31	38,56	40,43	41,04	47,02	51,30
1979	54,55	56,00	59,99	66,72	69,33	76,15	77,40	83,90	91,20	92,98	94,74	106,05
1980	115,16	130,80	137,67	134,09	133,10	129,36	140,02	148,67	157,50	157,07	157,28	149,80
1981	148,83	146,82	155,93	155,96	162,43	179,07	171,61	171,78	168,71	178,69	188,73	200,30
1982	212,10	216,70	216,08	230,30	239,54	235,78	262,61	270,24	314,48	331,29	334,10	340,15

3.3 - AJUSTAMENTO DA SAZONALIDADE FINAL

Neste estágio, do Census II, calculamos novas médias móveis sobre a série ajustada a sazonalidade preliminar, visando eliminar algum efeito sazonal e irregular não detectado. Estes cálculos seguem uma seqüência de passos similares ao aplicado na seção anterior.

3.3.1 - ISOLANDO A TENDÊNCIA-CICLO

Removemos a aleatoriedade existente nos dados dessazonalizados aplicando uma média móvel de "SPENCER" de lag 15. 24
Esta média móvel equivale a uma média móvel quádrupla 5x5x4x4 e se resume na fórmula abaixo (demonstração em MAKRIDAKIS e WHEELWRIGHT, 1978):

$$M_I'''' = (-3X_{I-7} - 6X_{I-6} - 5X_{I-5} + 3X_{I-4} + 21X_{I-3} +$$

$$+ 46X_{I-2} + 67X_{I-1} + 74X_I + 67X_{I+1} + 46X_{I+2} +$$

$$+ 21X_{I+3} + 3X_{I+4} - 5X_{I+5} - 6X_{I+6} - 3X_{I+7})$$

$$\div 360$$

Para suprir a perda, quando da aplicação de "SPENCER", de sete dados no início da série e sete no fim, Census II acrescenta sete valores iniciais (média dos quatro primeiros) e sete valores finais (média dos quatro últimos). Este procedimento é mostrado na tabela (3.12).

TABELA (3.12) CÁLCULO DA MÉDIA MÓVEL DE SPENCER COM LAG 15

	1		20,90 x (-3÷360)	
	2		20,90 x (-6÷360)	
VALORES ADICIONAIS	3		20,90 x (-5÷360)	
	4	$\frac{83,60}{4}$ = 20,90	20,90 x (3÷360)	
	5		20,90 x (20÷360)	
	6		20,90 x (46÷360)	
	7		20,90 x (67÷360)	
1977	JAN		20,34 x (74÷360)	18,256
	FEV		20,49 x (67÷360)	
	MAR	83,60	20,10 x (46÷360)	
	ABR		22,67 x (21÷360)	
	MAIO		22,51 x (3÷360)	
	JUN		24,96 x (-5÷360)	
	JUL		24,44 x (-6÷360)	
	AGO		24,40 x (-3÷360)	
	SET		25,24	
	OUT		29,09	
	.			
	.			
	.			

TABELA (3.13) VALORES OBTIDOS PELA MÉDIA MÓVEL DE SPENCER COM LAG 15

ANOS	JAN	FEV	MAR	ABR	MAIO	JUN	JUL	AGO	SET	OUT	NOV	DEZ
1977	18,26	18,45	18,88	19,50	20,26	21,03	21,80	22,58	23,37	24,16	24,93	25,68
1978	26,37	27,09	27,91	28,82	29,82	30,97	32,29	33,93	35,97	38,46	41,28	44,34
1979	47,58	50,93	54,43	58,19	62,12	66,20	70,21	74,11	78,24	83,14	89,26	96,41
1980	103,85	110,35	115,11	118,14	120,36	122,95	126,50	130,61	134,19	136,21	136,23	135,15
1981	134,32	134,91	137,41	141,36	145,54	148,90	151,39	153,60	156,68	161,49	168,23	175,92
1982	183,55	190,28	196,04	201,80	209,35	220,32	235,12	252,90	270,94	286,22	296,37	300,92

Quando os dados originais são divididos pelos dados da tabela (3.13) (aproximação da tendência-ciclo) restam apenas a sazonalidade e a aleatoriedade. Matematicamente, temos:

$$M'_t = TC_t$$

$$FSE_t = \frac{Z_t}{M'_t} = \frac{TC_t \times S_t \times e_t}{TC_t} = S_t \times e_t \quad (3.3.1)$$

Onde M'_t é a média móvel de spencer com lag 15 e FSE_t a sazonalidade e aleatoriedade final.

3.3.2 - AJUSTES NA SAZONALIDADE E ALEATORIEDADE FINAL

A sazonalidade e a aleatoriedade final são submetidas a tratamentos idênticos aos apresentados nas tabelas (3.5), (3.6), (3,7) e (3.9). Em resumo os passos são os seguintes:

a) Análise e Substituição dos "OUTLIERS"

Localizamos os outliers através do cálculo de uma média móvel dupla 3x3 e do erro padrão, para cada mês, exatamente com foi feito nas tabelas (3.5) e (3.6), exceto que os dados são os da saída da equação (3.3.1). Estes "outliers" são substituídos pela média dos valores vizinhos.

b) Estimação dos Valores Desaparecidos

Não precisamos fazer uma estimação para valores per didos, pois estas observações já foram acrescentadas quando aplicamos a Fórmula de Spencer de lag 15.

c) Normalização dos Valores

Os valores devem ser ajustados para que a soma de cada ano seja 1.200, exatamente como fizemos na tabela (3.9).

Na tabela (3.14) apresentamos a sazonalidade e a aleatoriedade final, já submetidas aos acertos descritos acima. Os valores na parte inferior da tabela são as médias para cada mês, equivalendo ao índice sazonal calculado no método clássico de decomposição.

TABELA (3.14) SAZONALIDADE E ALEATORIEDADE FINAL

ANOS	JAN	FEV	MAR	ABR	MAIO	JUN	JUL	AGO	SET	OUT	NOV	DEZ
1977	98,74	97,69	90,70	106,27	115,67	128,58	111,63	91,80	84,83	89,00	90,33	94,74
1978	97,68	98,36	95,50	101,74	116,17	120,50	110,05	96,89	88,91	87,77	88,25	98,18
1979	99,39	94,60	95,07	105,80	113,40	126,68	109,29	98,73	93,53	89,08	80,97	93,46
1980	95,55	101,16	102,37	104,24	111,16	115,30	110,66	99,72	92,44	86,51	87,14	93,75
1981	96,77	91,72	98,47	101,38	113,15	135,10	117,02	98,52	81,81	83,23	84,67	98,17
1982	102,52	95,62	94,64	104,25	116,19	119,27	117,03	93,71	85,18	85,60	87,46	98,52
MÉDIA - INDICES SAZONAIS												
	98,44	96,53	96,12	103,95	114,29	124,23	112,62	96,56	87,78	86,87	86,47	96,13

3.3.3 - FATOR SAZONAL FINAL

Os fatores sazonais finais são derivados da aplicação de uma média móvel dupla 3x3 nos dados da tabela (3.14). Procedimento é idêntico ao exposto na tabela (3.5) e apresenta como resultado os fatores sazonais finais. Estes fatores são projetados, para doze meses, multiplicando-se o fator do último ano por três, subtraindo o do ano anterior e dividindo o resultado por dois. A previsão para maio é calculada como abaixo:

$$[(114,45 \times 3) - 113,58] \div 2 = 114,88$$

TABELA (3.15) FATOR SAZONAL FINAL

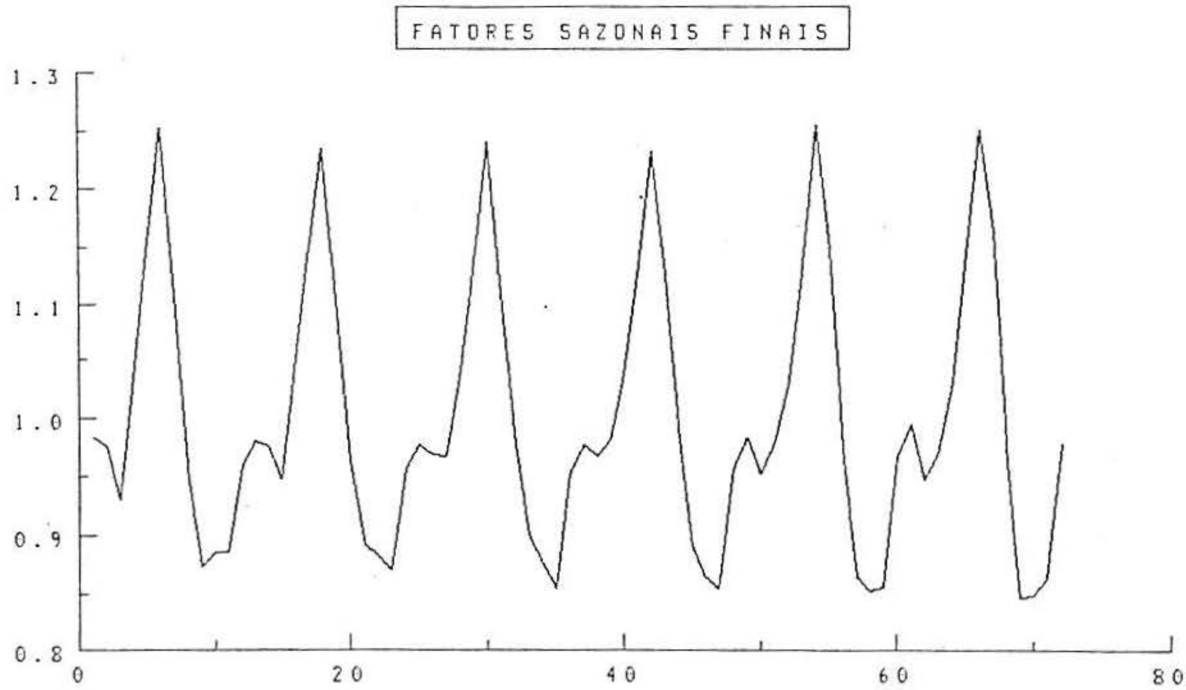
ANOS	JAN	FEV	MAR	ABR	MAIO	JUN	JUL	AGO	SET	OUT	NOV	DEZ
1977	98,74	97,61	93,05	104,46	115,61	125,23	110,76	94,55	87,38	88,53	88,48	85,94
1978	98,12	97,65	94,83	104,18	114,86	123,54	110,39	96,20	89,19	88,27	87,09	95,68
1979	97,79	96,92	96,69	104,11	113,74	123,92	110,88	97,75	89,99	87,56	85,41	95,24
1980	97,69	96,68	98,26	103,67	113,21	123,25	112,41	98,25	89,12	86,39	85,38	95,69
1981	98,39	95,22	97,89	103,30	113,58	125,37	114,75	97,47	86,41	85,27	85,58	96,76
1982	99,51	94,72	96,09	103,13	114,45	124,99	116,32	96,25	84,68	84,78	86,34	97,85
PREVISÃO DOS FATORES SAZONAIS PARA UM ANO												
	100,07	94,37	96,53	103,05	114,88	124,80	117,10	95,64	83,81	84,54	86,71	98,40

Matematicamente, este passo equivale a calcular a esperança para os valores da tabela (3.14), removendo, assim, qualquer aleatoriedade ainda presente nos dados. Portanto, temos:

$$FS'_t = E(S_t x\ e_t) = S_t$$

Onde FS'_t é o fator sazonal final para o período t.

O gráfico da tabela (3.15) revela claramente o movimento sazonal, dos depósitos à vista, nesta agência pertencente a uma região agropecuária. Podemos notar que os meses de maio, junho e julho recebem um maior peso, por ser é poca de comercialização das safras.



3.3.4 - SÉRIE AJUSTADA A SAZONALIDADE FINAL

A série ajustada a sazonalidade final é achada dividindo-se os dados originais pelos fatores sazonais finais da tabela (3.15). Esta divisão elimina completamente a sazonalidade de série, restando apenas a tendência-ciclo e a aleatoriedade (tabela (3.16)). Matematicamente temos:

$$\frac{Z_t}{E(S_t e_t)} = \frac{S_t \times TC_t \times e_t}{S_t} = TC_t \times e_t \quad (3.3.2)$$

TABELA (3.16) SÉRIE DESSAZONALIZADA

ANOS	JAN	FEV	MAR	ABR	MAIO	JUN	JUL	AGO	SET	OUT	NOV	DEZ
1977	20,33	20,49	20,42	22,02	22,49	23,96	24,38	24,33	25,18	31,63	28,25	28,14
1978	29,56	30,72	31,63	31,68	33,95	34,00	36,24	38,46	40,36	43,05	47,08	51,21
1979	54,20	55,72	59,99	66,28	69,46	75,85	77,56	83,89	51,12	94,79	94,84	106,05
1980	114,65	130,32	135,36	134,07	133,38	129,82	110,56	149,62	157,09	153,95	156,95	149,45
1981	148,39	145,97	155,27	155,85	162,88	180,27	173,42	174,40	166,65	177,09	186,96	200,50
1982	213,04	216,42	215,50	229,80	239,41	236,82	266,51	277,40	307,06	325,55	338,21	341,34

3.3.5 - CÁLCULO DA TENDÊNCIA-CICLO E ALEATORIEDADE

Na fase final do Census II é eliminada a aleatoriedade existente na tabela (3.16), restando apenas a tendência-ciclo. Isto é feito através do cálculo de uma média móvel de Spencer com lag 15 idêntica a descrita na tabela (3.12).

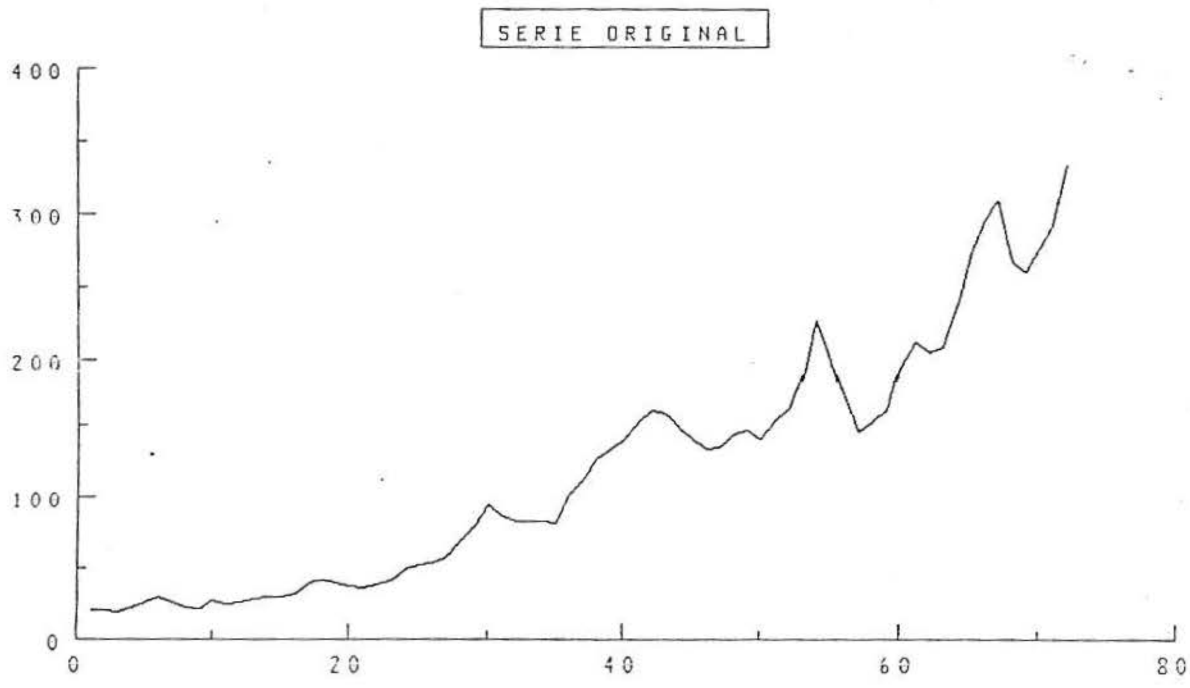
Matematicamente dizemos que:

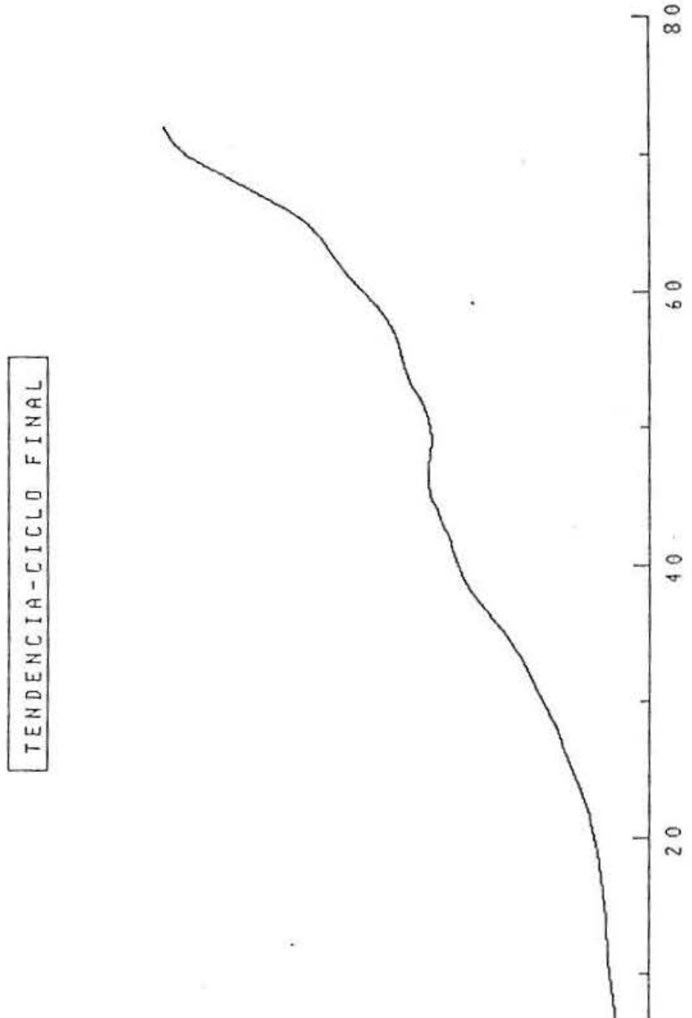
$$E(TC_t \times e_t) = TC_t \qquad (3.3.3)$$

TABELA (3.17) ESTIMATIVA FINAL DA TENDÊNCIA-CICLO

ANOS	JAN	FEV	MAR	ABR	MAIO	JUN	JUL	AGO	SET	OUT	NOV	DEZ
1977	20,50	20,68	21,06	21,64	22,41	23,32	24,37	25,54	26,69	27,70	28,57	29,25
1978	29,82	30,44	31,20	32,12	33,22	34,56	36,80	38,26	40,76	43,63	46,74	50,00
1979	53,44	57,06	60,94	65,17	69,70	74,37	79,02	83,56	88,32	93,83	100,59	108,43
1980	116,57	123,76	129,16	132,75	135,47	138,56	142,60	147,13	150,96	153,00	152,80	151,38
1981	150,33	151,06	154,16	159,06	164,24	168,29	170,99	173,06	175,98	180,94	188,34	197,05
1982	205,76	213,55	220,37	227,38	236,45	249,09	265,40	284,29	303,14	318,92	329,43	334,23

Quando observamos os gráficos abaixo, constatamos que tivemos sucesso na eliminação da sazonalidade e da aleatoriedade e que, realmente, isolamos a tendência-ciclo.




TENDENCIA-CICLO FINAL
0
20
40
60
80
0
100
200
300
400

Finalmente identificaremos a aleatoriedade presente, nos dados, dividindo a equação (3.3.2) pela equação (3.3.3).

$$\frac{TC_t \times e_t}{TC_t} = e_t$$

TABELA (3.18) ESTIMATIVA FINAL DA COMPONENTE ALEATÓRIA

ANOS	JAN	FEV	MAR	ABR	MAIO	JUN	JUL	AGO	SET	OUT	NOV	DEZ
1977	99,85	99,07	96,96	101,76	100,35	102,73	100,01	95,24	94,32	114,16	98,89	96,21
1978	99,12	100,93	101,38	98,62	102,21	98,36	100,10	100,52	99,02	98,66	100,73	102,42
1979	101,42	97,65	98,44	101,69	99,64	101,99	98,15	100,39	103,18	101,02	94,28	97,80
1980	98,36	105,30	104,80	101,00	98,45	93,69	98,57	101,69	104,06	100,62	102,72	98,72
1981	98,72	96,63	100,72	97,98	99,17	107,11	101,42	100,77	94,70	97,87	99,27	101,75
1982	103,54	101,35	97,79	101,06	101,25	95,07	100,42	97,58	101,29	102,08	102,67	102,13

4) CONSIDERAÇÕES FINAIS

Existem muitas versões do método de decomposição Census II. Foi descrita na seção anterior a versão X-9 que possui, complementarmente, alguns testes não estatísticos, despresados neste trabalho.

Como foi salientado, na apresentação, a finalidade deste estudo é obter maiores detalhes sobre o comportamento de algumas séries econômicas, para estimar seus valores futuros. Esta estimação pode ser feita em três estágios:

a) Dessazonalização das séries através do método Census II.

b) Projeção das séries dessazonalizadas ($Z'_t = TC_t \times e_t$) através dos modelos linear e quadrático, com estimação dos coeficientes utilizando a técnica de amortecimento exponencial de BROWN (maiores detalhes em FARIOLI, Wàlter E. (1983), SOUZA, Reinaldo (b)).

c) Retorno da sazonalidade para uma melhor adequação das previsões à realidade.

A utilização deste processo tem fornecido bons resultados, principalmente quando temos um número expressivo de dados, permitindo assim uma perfeita identificação da sazonalidade.

Se fizermos previsões para os próximos quatro valores, da série do exemplo desenvolvido na seção 3, obteremos:

MESES	PREVISTO	REALIZADO
JAN/83	359	349
FEV/83	351	353
MAR/83	372	384
ABR/83	411	410

Podemos notar que os valores previstos estão próximos dos realizados, possuindo, é claro, uma pequena diferença causada pela componente aleatória. 26

A literatura sobre séries temporais apresenta muitos métodos mais potentes que o apresentado aqui, mas este tem como vantagem o baixo custo e a automaticidade, que permite a análise e previsão, em um tempo mínimo, de um grande número de séries.

APÊNDICE

Apresentamos, nas próximas páginas, a listagem dos programas principais utilizados neste trabalho.

O programa "CENSUS II" apresenta adaptações a realidade da instituição onde foi desenvolvido, realizando, além dos passos descritos na seção 3, previsões através de amorteci - mentos exponenciais de Brown. 27

Tanto o programa "ajuste", quanto o "CENSUS 2" foram elaborados em linguagem APL.

```
[33]  →FINTO
[34] FUI:MX4←((ρX5)÷12),12)ρX5
[35]  X4←MX4÷Q(1ΦρMX4)ρ(+/MX4)÷12
[36]  SIN←(+/[1] X4)÷1↑ρX4
[37] FINTO:AUXILM3(,X4)
[38]  MEDMOV3 X1L
[39]  FSAF←MA3
[40]  MFSAF←(((⌊(ρFSAF)÷12)),12)ρFSAF
[41]  →((1|(ρFSAF)÷12)=0)/CAI

LINE 248698    COLUMN 1                                   HOLDING
[42]  PROJS FSAF
[43]  →CUA
[44] CAI:PROJSAZ←((MFSAF[((ρFSAF)÷12);]×3)-MFSAF[(((ρFSAF)÷12)-1);])÷2
[45] CUA:FSAS←DADOS÷FSAF
[46]  SPENCEL FSAS
[47]  TENDC←SP
[48]  ALEA←FSAS÷SP
[49]  PADRAO ALEA
[50]  O2←2ρ0
[51]  Q←0
[52]  2 AELQB FSAS
[53]  ALL← 12 1 ρ(,PROJSAZ)×(,PREV)
[54]  3 AELQB FSAS
[55]  ALL1← 12 1 ρ(,PROJSAZ)×(,PREV)
[56]  →(O2[1]≤O2[2])/PTQP
[57] ⍝INTERVALO DE PROJECAO PARA 12 VALORES (ALIS. 3.G)
[58]  LSUP←2↓8↑PREVIS←,(ALL1+O2[2])
[59] ⍝O LIMITE INFERIOR E SIMPLESMENTE O VALOR PREVISTO
[60]  LINF←2↓8↑PREVII←,ALL1
[61]  LIM←2↓8↑PREVIV←PREVII-(O2[2])
[62]  →0
[63] ⍝INTERVALO DE PROJECAO PARA 12 VALORES (ALIS. 2.G)

LINE 248720    COLUMN 1                                   INPUT
[64] PTQP:LSUP←2↓8↑PREVIS←,ALL+O2[1]
[65]  LINF←2↓8↑PREVII←,ALL
[66]  LIM←2↓8↑PREVIV←PREVII-(O2[1])
     ∇
```

```
LINE 248654    COLUMN 1                                                    HOLDING
      ∇CENSUS2[⎕]∇
      ∇ CENSUS2 X;DADOS
[1]   DADOS←X
[2]   12 MEDMOV X
[3]   X1←100×(6↓((ρX)−6)↑X)÷(,MM)
[4]   AUXILM3 X1
[5]   MEDMOV3 X1L
[6]   QD←(X1−MA3)*2
[7]   PRESD QD
[8]   INTCONF X1
[9]   C←ρN
[10]  X2←N[7],N[8],N[9],N[10],N[11],N[12],N,N[C−11],N[C−10],N[C−9],N[C−8],N[C−7
      ],N[C−6]
[11]  →((1|(ρX2)÷12)=0)/PULA
[12]  DFAL X2
[13]  →SALTA
[14]  MX3←(((ρX3)÷12),12)ρX3←X2,(12−12×1|(ρX2)÷12)ρ0
[15]  →SALTA
[16] PULA:MX3←(((ρX2)÷12),12)ρX2
[17]  X4←,MX3←⍉(1⌽ρMX3)ρ(+/MX3)÷12
[18] SALTA:AUXILM3 X4
[19]  MEDMOV3 X1L

LINE 248676    COLUMN 1                                                    HOLDING
[20]  ASP←DADOS÷MA3
[21]  SPENCEL ASP
[22]  FSIR←DADOS÷SP
[23]  AUXILM3 FSIR
[24]  MEDMOV3 X1L
[25]  QD1←(FSIR−MA3)*2
[26]  PRESD QD1
[27]  INTCONF FSIR
[28]  X5←N
[29]  →((1|((ρX5)÷12))=0)/FUI
[30]  DFAL X5
[31]  →FINTO
[32]  MX4←(((ρX6)÷12),12)ρX6←X5,(12−12×1|(ρX5)÷12)ρ0
```

```
LINE 248724    COLUMN 1                                          HOLDING
      ∇AJUSTE[⎕]∇
      ∇ X AJUSTE Y
[1]    YY←Y
[2]    LOGY←10⍟Y
[3]   UM:L←(ρY)−1
[4]    S←(ρY)↑(¯1×L)↑Y
[5]    Y←L↑(S−Y)
[6]    M←+/Y÷ρY
[7]    V←⌈/((Y−M)*2)*0.5
[8]    →((ρY)<((ρX)−1))/11
[9]    DM1←100×((V÷M)*2)*0.5
[10]   →UM
[11]   →((ρY)=((ρX)−3))/14
[12]   DM2←100×((V÷M)*2)*0.5
[13]   →UM
[14]   DM3←100×((V÷M)*2)*0.5
[15]   L←(ρLOGY)−1
[16]   S←(ρLOGY)↑(¯1×L)↑LOGY
[17]   G←L↑(S−LOGY)
[18]   M←+/G÷ρG
[19]   V←⌈/((G−M)*2)*0.5
[20]   DM4←100×((V÷M)*2)*0.5

LINE 248746    COLUMN 1                                          HOLDING
[21]   S0←+/(LOGY≤0)
[22]   →(S0>0)/30
[23]   LOLOGY←10⍟LOGY
[24]   L←(ρLOLOGY)−1
[25]   S←(ρLOLOGY)↑(¯1×L)↑LOLOGY
[26]   H←L↑(S−LOLOGY)
[27]   M←+/H÷ρH
[28]   V←⌈/((H−M)*2)*0.5
[29]   DM5←100×((V÷M)*2)*0.5
[30]   LOGX←10⍟X
[31]   L←(ρLOGX)−1
[32]   S←(ρLOGX)↑(¯1×L)↑LOGX
[33]   K←L↑(S−LOGX)
```

```
[34]  LOXY←G÷K
[35]  M←+/LOXY÷ρLOXY
[36]  V←Γ/((LOXY-M)*2)*0.5
[37]  DM6←100×((V÷M)*2)*0.5
[38]  →(S0>0)/50
[39]  'DISPERCAO PERCENTUAL DA RETA, POLIN. 2. E 3. GRAU, EXPON. 1. E 2. GRAU E
       FUNCAO POTENCIA'
[40]  DM1,DM2,DM3,DM4,DM5,DM6
[41]  N←L/DM1,DM2,DM3,DM4,DM5,DM6

LINE 248768    COLUMN 1                                              HOLDING
[42]  →(N=DM5)/44
[43]  →DOIS
[44]  'A FUNCAO QUE MAIS SE AJUSTA AOS DADOS E UMA EXPONENCIAL DO 2. GRAU'
[45]  C←LOLOGY⌹X∘.* 0 1
[46]  D←10*C
[47]  'COEFICIENTES DA FUNCAO EXPONENCIAL DE 2. GRAU'
[48]  D
[49]  →QUATRO
[50]  'DISPERSAO PERCENTUAL DA RETA ,POLIN. 2. E 3. GRAU,EXPONENCIAL 1. GRAU E
      FUNCAO POTENCIA'
[51]  DM1,DM2,DM3,DM4,DM6
[52]  N←L/DM1,DM2,DM3,DM4,DM6
[53] DOIS:→(N=DM1)/58
[54]  →(N=DM2)/63
[55]  →(N=DM3)/68
[56]  →(N=DM4)/73
[57]  →(N=DM6)/78
[58]  'A FUNCAO QUE MAIS SE AJUSTA AOS DADOS E UMA RETA'
[59]  COEF←YY⌹X∘.* 0 1
[60]  'COEFICIENTES DA RETA'
[61]  COEF
[62]  →QUATRO
```

```
LINE 248790    COLUNA 1                                                    HOLDING
[63]  'A FUNCAO QUE MAIS SE AJUSTA AOS DADOS E UMA POLINOMIAL 2. GRAU'
[64]  COEF←Y⌹X∘.* 0 1 2
[65]  'COEFICIENTES DO POLINOMIO DE 2. GRAU'
[66]  COEF
[67]  →QUATRO
[68]  'A FUNCAO QUE MELHOR SE AJUSTA E UMA POLINOMIAL DO 3. GRAU'
[69]  COEF←Y⌹X∘.* 0 1 2 3
[70]  'COEFICIENTES DO POLINOMIO DE 3. GRAU'
[71]  COEF
[72]  →QUATRO
[73]  'A FUNCAO QUE MELHOR SE AJUSTA E UMA EXPONENCIAL DO 1. GRAU'
[74]  C←10*(LOGY⌹X∘.* 0 1)
[75]  'COEFICIENTES DA EXPONENCIAL DE 1. GRAU'
[76]  C
[77]  →QUATRO
[78]  'A FUNCAO QUE MELHOR SE AJUSTA AOS DADOS E A FUNCAO POTENCIA'
[79]  C←LOGY⌹LOGX∘.* 0 1
[80]  C[1]←10*C[1]
[81]  'COEFICIENTES DA FUNCAO POTENCIA'
[82]  C
[83] QUATRO:V← 1 12 ⍴(X[2]−X[1])
[84]  P←+\V
```

```
LINE 248812   COLUMN 1                                        HOLDING
[85]  P1←X[ρX]+P
[86]  →(N=DM1)/92
[87]  →(N=DM2)/94
[88]  →(N=DM3)/96
[89]  →(N=DM4)/98
[90]  →(N=DM5)/102
[91]  →(N=DM6)/100
[92]  P2←(P1∘.* 0 1)+.×COEF
[93]  →CINCO
[94]  P2←(P1∘.* 0 1 2)+.×COEF
[95]  →CINCO
[96]  P2←(P1∘.* 0 1 2 3)+.×COEF
[97]  →CINCO
[98]  P2←C[1]×(C[2]*P1)
[99]  →CINCO
[100]  P2←C[1]×(P1*C[2])
[101]  →CINCO
[102]  P2←10*(D[1]×(D[2]*P1))
[103] CINCO:'PROJECAO DA TENDENCIA PARA OS PROXIMOS DOZE VALORES'
[104]  P2
[105] TRES.
      ∇
```

BIBLIOGRAFIA

[1] CAMPOS, Andrade Marcília. (1978). Alguns Modelos para Série de Tempo. Universidade Federal de Pernambuco.

(1975)
[2] CHAO, Lincoln L.. Estadistica para Las Ciências Administrativas. Bogotá, MC Graw-Hill. ~~1975.~~

[3] FARIOLI, Walter E.. (1983). Discussão de Três Padrões Típicos de Comportamento para Séries Temporais: Constante, Linear e Quadrático. Relatório de Estágio. Universidade Federal do Rio Grande do Sul.

[4] FISCHER, Sérgio. (1982). Séries Univariantes de Tempo-Metodologia de Box & Jenkins. Fundação de Economia e Estatística.

(1978)
[5] MAKRIDAKIS, Spyros & WHEELWRIGHT, Steven C.. Forecasting, Methods and Applications. local, editora.

[6] MORETTIN, P.A. & TOLOI, C.M.C.. (1981a). Modelos para Previsão de Séries Temporais. Vol. 1. 13º Colóquio Brasileiro de Matemática, Poços de Caldas (MG). Junho 1981.

[7] SOUZA, Reinaldo Castro. (a). Artigo da Revista de econometria de Novembro de 81 - "Metodologias para a Análise e Previsão de Séries Temporais Univariadas e Multivariadas".

data

[8] SOUZA, Reinaldo Castro. (b). *Métodos Automáticos de A
mortecimento Exponencial para Previsão de Séries Tem
rais*. PUC, Rio de Janeiro.