

Identificação de Produtos em Descrições Textuais de Compras: Uma Proposta para Portais de Transparência Pública

Eduardo Soares de Paiva

12/12/2019

Motivação



➤ Iniciativas Governamentais

- Lei Complementar 131 (Lei da Transparência)
- Lei 12.527 (lei de acesso à informação)

Motivação

- Publicação de dados não garante transparência!!!
- Dados oriundos de sistemas corporativos
 - Informações não tratadas;
 - Informações de compras descritas em formato texto.

Problema

Extrato da tela de empenho Empenho

Detalhamento do Gasto



SUBITEM	QUANTIDADE	VALOR UNITÁRIO	VALOR TOTAL	DESCRIÇÃO
MATERIAL DE EXPEDIENTE	30	164,99	4.949,70	0000000030,00000 CAIXA PAPEL PAPEL A4 75G/M2 TIPO SULFITE, PADRÃO PARA IMPRESSORAS OFÍCIO, MATERIAL PAPEL ALCALINO, COR BRANCA, EMBALAGEM RESISTENTE A UMIDADE. CAIXA COM 10 RESMAS DE 500 FOLHAS MARCA: ITEM DO PROCESSO: 00068 ITEM DE MATERIAL: 000000965

Descrição do Produto

Problema

- Publicação de dados não garante transparência!!!
- Dados oriundos de sistemas corporativos
 - Informações não tratadas

➤ Informações de compras descritas em formato texto.



Questão de Pesquisa

- **Como identificar de forma automatizada os produtos a partir das especificações textuais que são usadas para caracterizá-los nas descrições dos gastos que são apresentados nos portais de transparência pública?**



Objetivo

- Identificar os produtos mais relevantes que são comprados pela Administração Pública através da análise das descrições textuais de compras apresentadas nos portais de transparência.



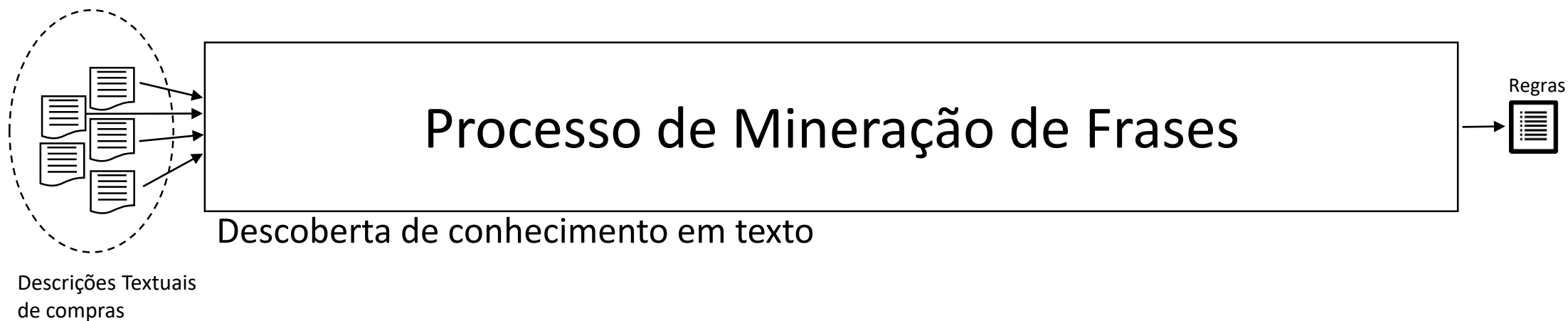
Hipótese

Se forem identificadas as sequências de tokens que mais se repetem em um determinado conjunto de descrição de compras, **então**, essas sequências de tokens caracterizarão os produtos mais relevantes desse conjunto de descrições.



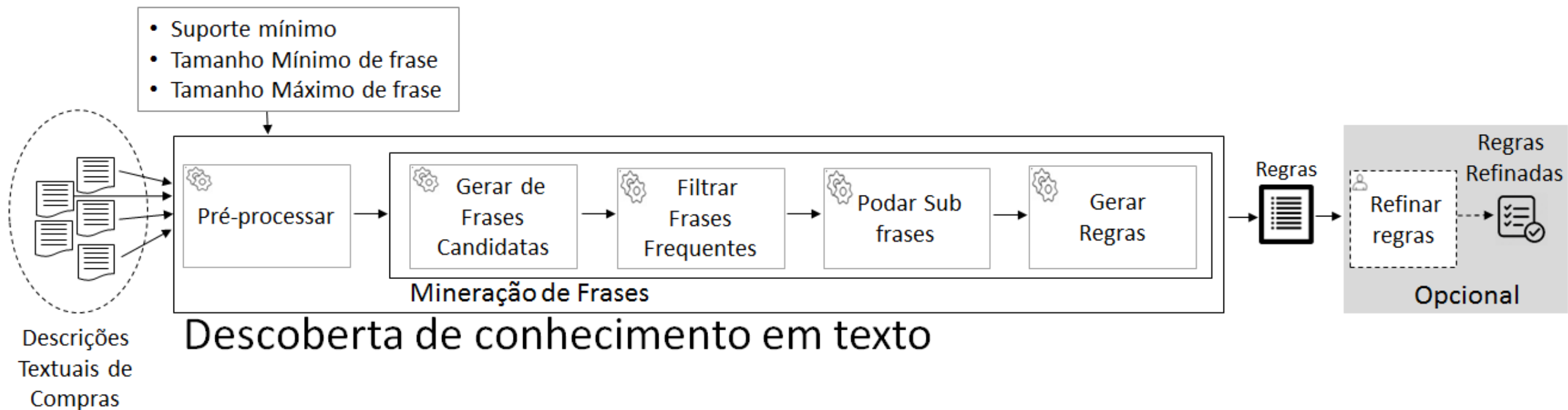
Preparação e Modelagem

➤ Proposta



Preparação e Modelagem

➤ Arquitetura



Avaliação

➤ Projeto de Avaliação

Se forem identificadas as sequências de tokens que mais se repetem em um determinado conjunto de descrição de compras,

então, essas sequências de tokens caracterizarão os produtos mais relevantes desse conjunto de descrições.

Hipótese



Análise das Regras Geradas

Análise dos Resultados Gerados

1. Introdução
2. Proposta
3. Avaliação

4. Aplicações
5. Conclusão

- 4.1. Cálculo de Preços de Referência dos Produtos Comprados Pela Administração Pública
- 4.2. Identificação de Compras com Preços Muito Acima do Esperado
- 4.3. Comparação Entre Valores Pagos em Compras Licitadas e Não Licitadas
- 4.4. Identificação de Fornecedores vendendo o mesmo produto com preços diferentes
- 4.5. Acompanhar a Tendência de Preços

Aplicações

➤ Cálculo de Preços de referência

Produto	Unidade	Preço Referência (Mediana)			
		2016	2017	2018	2019
Água Mineral	Galão 20 L	10,00	8,77	R\$ 8,90	8,47
Caneta Esferográfica	Caixa 50 Unidades	36,00	32,22	20,95	25
Caneta Marca Texto	Unidade	0,93	0,88	0,83	0,89
Carne de Boi	Kg	20,80	17,23	17,00	17,25
Carne de Frango	Kg	9,84	7,90	7,92	7,24
Cartucho Tinta Impressora	Unidade	80	79	95,22	81,00
Diesel	Litro	3,73	3,40	3,69	3,95
Gasolina	Litro	4,12	4,10	4,73	4,85
Laranja	Kg	2,48	1,98	2,25	2,41
Papel A4	Resma	14,12	14,62	14,39	14,70

1. Introdução
2. Proposta
3. Avaliação

4. Aplicações
5. Conclusão

- 4.1. Cálculo de Preços de Referência dos Produtos Comprados Pela Administração Pública
- 4.2. Identificação de Compras com Preços Muito Acima do Esperado
- 4.3. Comparação Entre Valores Pagos em Compras Licitadas e Não Licitadas
- 4.4. Identificação de Fornecedores vendendo o mesmo produto com preços diferentes
- 4.5. Acompanhar a Tendência de Preços

Aplicações

➤ Identificação de Compras com preços muito acima do esperado

Produto	Unidade	Preço de referência (2019)	Número do empenho	Valor Unitário
Água mineral	Galão 20 L	8,47	2019NE800079	70,09
			2019NE800041	320,00
Caneta Esferográfica	Caixa 50 Unidades	25,00	2019NE800044	32,87
			2019NE800029	32,87
Caneta Marca Texto	Unidade	0,89	2019NE800020	12,40
			2019NE800278	15,40
Carne de Boi	Kg	17,25	2019NE800198	53,98
			2019NE800024	48,90
Carne de Frango	Kg	7,24	2019NE800088	32,45
			2019NE800432	28,78
Cartucho Impressora Tinta	Unidade	81,00	2019NE800077	310,80
			2019NE800025	256,20
Diesel	Litro	3,95	2019NE800398	62,50
			2019NE800072	54,94
Gasolina	Litro	4,85	2019NE800473	33,33
			2019NE800211	26,96
Papel A4	Resma	14,70	2019NE801510	160,26
			2019NE800128	142,50

1. Introdução
2. Proposta
3. Avaliação

4. Aplicações
5. Conclusão

- 4.1. Cálculo de Preços de Referência dos Produtos Comprados Pela Administração Pública
- 4.2. Identificação de Compras com Preços Muito Acima do Esperado
- 4.3. Comparação Entre Valores Pagos em Compras Licitadas e Não Licitadas
- 4.4. Identificação de Fornecedores vendendo o mesmo produto com preços diferentes
- 4.5. Acompanhar a Tendência de Preços

Aplicações

➤ Comparação entre valores de compras licitadas e não licitadas

Produto	Unidade	Preço Praticado (Mediana)	
		Com Licitação	Sem Licitação
Água Mineral	Galão 20 L	R\$ 6,20	R\$ 10,51
Caneta Marca Texto	Unidade	R\$ 0,85	R\$ 12,93
Carne de Boi	Kg	R\$ 17,03	R\$ 20,05
Cartucho Tinta Impressora	Unidade	R\$ 64,45	R\$ 105,06
Diesel	Litro	R\$ 3,60	R\$ 4,03
Gasolina	Litro	R\$ 4,36	R\$ 4,59
Laranja	Kg	R\$ 1,52	R\$ 2,99
Papel A4	Resma	R\$ 14,07	R\$ 17,50

1. Introdução
2. Proposta
3. Avaliação

4. Aplicações
5. Conclusão

- 4.1. Cálculo de Preços de Referência dos Produtos Comprados Pela Administração Pública
- 4.2. Identificação de Compras com Preços Muito Acima do Esperado
- 4.3. Comparação Entre Valores Pagos em Compras Licitadas e Não Licitadas
- 4.4. Identificação de Fornecedores vendendo o mesmo produto com preços diferentes
- 4.5. Acompanhar a Tendência de Preços

Aplicações

➤ Fornecedores vendendo o mesmo produto com preços diferentes

Compra 1

Detalhamento do Gasto



BAIXAR

SUBITEM	QUANTIDADE	VALOR UNITÁRIO	VALOR TOTAL	DESCRIÇÃO
MATERIAL DE PROCESSAMENTO DE DADOS	18	94,99	1.709,82	18,00000 UNIDADE CARTUCHO TONER IMPRESSORA SAMSUNG, REFERÊNCIA CARTUCHO MLT-D203U, TIPO CARTUCHO ORIGINAL, COR TINTA PRETA MARCA: ORIGINAL MAXJET ITEM DO PROCESSO: 00001 ITEM DE MATERIAL: 000426542
MATERIAL DE PROCESSAMENTO DE DADOS	6	133,16	798,96	6,00000 CAIXA 1,00 UN CARTUCHO TINTA IMPRESSORA LEXMARK, REFERÊNCIA CARTUCHO 60F4H00, TIPO CARTUCHO ORIGINAL, COR TINTA PRETA MARCA: ORIGINAL MAXJET ITEM DO PROCESSO: 00002 ITEM DE MATERIAL: 000427092

Compra 2

Detalhamento do Gasto



BAIXAR

SUBITEM	QUANTIDADE	VALOR UNITÁRIO	VALOR TOTAL	DESCRIÇÃO
MATERIAL DE TIC - MATERIAL DE CONSUMO	4	198,99	795,96	4,00000 UNIDADE CARTUCHO TINTA IMPRESSORA LEXMARK, REFERÊNCIA CARTUCHO 60F4H00, TIPO CARTUCHO ORIGINAL, COR TINTA PRETA MARCA: ORIGINAL MAXJET ITEM DO PROCESSO: 00001 ITEM DE MATERIAL: 000427092

Preços distintos – mesmo produto e mesmo fornecedor

Produto: Cartucho de tinta Impressora Lexmark - 60F4H00

• **Compra 1**
Valor R\$ 131,16

• **Compra 2**
Valor R\$ 199,99

Variação: 52 %

1. Introdução
2. Proposta
3. Avaliação

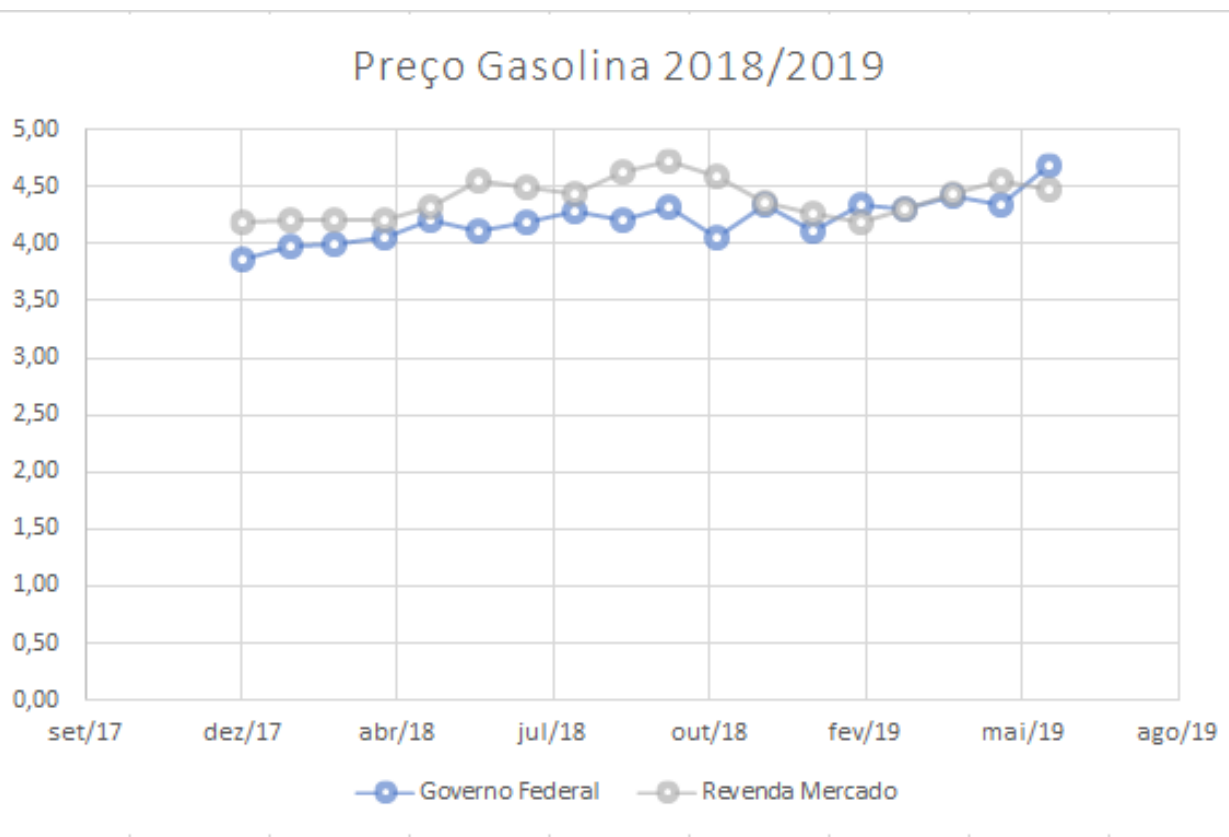
4. Aplicações
5. Conclusão

- 4.1. Cálculo de Preços de Referência dos Produtos Comprados Pela Administração Pública
- 4.2. Identificação de Compras com Preços Muito Acima do Esperado
- 4.3. Comparação Entre Valores Pagos em Compras Licitadas e Não Licitadas
- 4.4. Identificação de Fornecedores vendendo o mesmo produto com preços diferentes
- 4.5. Acompanhar a Tendência de Preços

Aplicações

➤ Acompanhamento de Tendências de Preço

Mês	Governo Federal	Revenda Mercado
jan/18	3,87	4,19
fev/18	3,98	4,21
mar/18	4,00	4,20
abr/18	4,05	4,22
mai/18	4,21	4,31
jun/18	4,12	4,55
jul/18	4,19	4,49
ago/18	4,29	4,45
set/18	4,21	4,63
out/18	4,32	4,72
nov/18	4,05	4,59
dez/18	4,35	4,37
jan/19	4,12	4,27
fev/19	4,35	4,19
mar/19	4,31	4,31
abr/19	4,42	4,44
mai/19	4,35	4,55
jun/19	4,69	4,47



Comparativo entre os preços da Gasolina comprada pelo Governo Federal e os valores de mercado

Conclusão

➤ Considerações Finais

Motivação do tema: dar maior efetividade aos dados que são publicados diariamente nos portais de transparência pública



Obrigado!

Eduardo Soares de Paiva

eduardo.paiva@cgu.gov.br

eduardo.sdepaiva@gmail.com