

Impacto da distribuição de royalties de petróleo sobre alguns aspectos sócio-econômico nos estados produtores de petróleo do Brasil

Disraelly Hander Gurgel Pereira
Kerolayne Paiva Soares

Mestrado em Ciência da Computação
Universidade Federal do Semi-Árido – UFERSA
Universidade Estadual do Rio Grande do Norte – UERN
59625-900 - Mossoró - RN, Brasil
handerpereira@hotmail.com
kerol.paiva@live.it

Abstract. The royalties distribution of oil to states and town that produce oil come being widely discussed, mainly after the promulgation of the new oil law, law n.º 9.478/97, because there was a rapid growth in the oil generation by states. To discover patterns that help in making decision about the distribution these royalties in the producing states is relevant importance socio-economic. The methodology utilized involved a process analysis and data mining relative the distribution of royalties in the Brazilian states that produce oil and your relation with some indicators socio-economic. The search of new patterns utilizing techniques of data mining, beyond the traditional statistics tolls provides new mechanisms of knowledge discovery. The purpose of this paper is present the knowledge extracted through relative data mining of distribution of royalties by government to states that produce oil with the relation the some aspects socio-economic in the same states in the period between 2001 to 2005.

Palavras-chave: spatial data mining, oil, poverty index, investment polices, mineração de dados espaciais, petróleo, índice de pobreza, política de investimento.

1. Introdução

A Lei ordinária n.º 9.478, de 6 de agosto de 1997 estipula o fim do monopólio estatal da exploração petrolífera. Após sua promulgação, a geração de *royalties* de petróleo vem crescendo vertiginosamente, beneficiando os estados e municípios produtores. Segundo Pacheco (2003), a problemática em relação a este cenário, diz respeito ao emprego destas verbas, uma vez que deixaram de ser insignificantes dentro do fluxo dos recursos públicos. Embora a Lei n.º 9.478/97 estabeleça o modo de partilha dos *royalties* e participações especiais, não especifica os setores nos quais estes recursos devam ser empregados pelos estados e municípios.

A distribuição de recursos de *royalties* para estados e municípios afetados pela exploração e transporte visa compensá-los por eventuais efeitos deletérios de tais atividades. A Lei do Petróleo estabelece critérios para a utilização de tais recursos, os quais devem ser alocados para investimentos (Postali, 2007).

A busca de novos padrões e relacionamentos entre diferentes variáveis utilizando algoritmos de mineração de dados, além das tradicionais ferramentas estatísticas proporcionam novos mecanismos para descoberta de conhecimento (Morais et al, 2003).

O presente artigo tem como objetivo analisar os dados relativos à geração de *royalties* pelos estados produtores e alguns índices sócio-econômicos (PIB *per capita*, Índice de pobreza e Investimentos), utilizando mineração de dados espaciais na busca por informações importantes e padrões que possam ser úteis para possíveis tomadas de decisão dentro do contexto da exploração petrolífera no Brasil.

2. Mineração de Dados

A mineração de dados é uma das etapas do processo de descoberta do conhecimento em banco de dados, que é do inglês *Knowledge Discovery in Database* (KDD) e têm por objetivo extrair informações implícitas e potencialmente úteis de dados (Fayyad et al.,1996).

Segundo (Silva, 2004) o processo de KDD é interativo, iterativo, cognitivo e exploratório, envolvendo vários passos com muitas decisões sendo feitas pelo analista (que é um especialista do domínio dos dados, ou um especialista de análise de dados).

2.2 Mineração de Dados Espaciais

A mineração de dados espaciais é uma extensão da mineração de dados voltada para domínios de aplicação onde a consideração da dimensão espacial é essencial na extração de conhecimento. (Roddick et al.,2001)

3. Metodologia de Trabalho

Para este trabalho foram utilizados dados relativos à geração de *royalties* pelos estados que produzem petróleo no Brasil, que são: Alagoas, Amazonas, Bahia, Ceará, Espírito Santo, Paraná, Rio de Janeiro, Rio Grande do Norte, Sergipe e São Paulo, bem como dados sobre o PIB *per capita*, percentual de pobreza e em investimentos nesses mesmos estados.

A base de dados é relativa a um período entre 2001 e 2005. Os dados coletados sobre os *royalties* gerados pelos estados foram extraídos do site da ANP (Agência Nacional de Petróleo) e os dados sobre o PIB *per capita*, Percentual de Pobreza e as Despesas Capitais investidas foram retirados do site Ipea (Instituto de Pesquisa Econômica Aplicada). Foram excluídos da base de dados todos os estados que não são produtores de petróleo. (Figura 1)

Beneficiários	Royalties distribuídos (mil R\$)										
	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008
Total	283.704	983.600	1.867.753	2.303.290	3.183.985	4.396.378	5.042.826	6.206.086	7.703.543	7.490.613	10.936.909
Unidades da Federação	106.885	330.444	623.287	762.479	1.020.960	1.413.174	1.618.686	1.984.329	2.380.443	2.291.236	3.293.057
Alagoas	1.656	5.206	9.463	11.742	14.398	23.037	29.053	34.824	43.137	38.798	41.439
Amazonas	6.587	20.274	48.561	59.679	70.308	90.480	113.978	143.046	131.268	118.659	154.576
Bahia	12.433	33.954	58.857	70.990	77.689	114.993	129.686	148.111	166.610	152.094	203.620
Ceará	1.600	4.267	6.688	8.579	9.357	14.154	13.735	13.950	14.126	13.128	16.785
Espírito Santo	2.166	7.446	13.919	24.347	31.131	59.279	51.617	57.284	96.612	143.818	253.598
Paraná	617	1.019	2.151	1.496	660	3.017	7.503	8.688	6.477	4.744	5.404
Rio de Janeiro	55.942	190.041	367.806	461.458	671.656	907.744	1.041.661	1.318.598	1.646.732	1.563.534	2.262.774
Rio Grande do Norte	18.623	49.698	85.150	90.134	103.435	140.946	163.848	181.023	180.150	159.577	213.647
Santa Catarina	39	4	53	40	19	-	-	-	-	-	-
Sergipe	6.223	16.446	28.800	31.831	39.810	55.526	63.659	74.658	90.617	92.516	137.032,11
São Paulo	1.000	2.088	1.839	2.184	2.497	4.000	3.947	4.148	4.713	4.368	4.181

Figura 1 – Tabela com os dados relativos à distribuição de *royalties*

3.1 Pré-processamentos dos dados

Para a construção da base de dados foi feito um pré-processamento e uma seleção dos dados coletados. Os dados relativos ao PIB *per capita* e ao Índice de Pobreza continham informações referentes a todos os estados brasileiros, mas como o nosso domínio eram apenas os estados produtores de petróleo, os demais foram excluídos da base. Foi feita uma alteração também aos dados relativos ao tempo, pois como algumas tabelas possuíam apenas informações completas entre 2001 e 2005, as demais que abrangiam um período de tempo mais longo foram reduzidas.

Foi gerado um arquivo .csv, a partir da ferramenta CSVed, contendo a base de dados, pois a ferramenta vis-stamp, utilizada para minerar os dados aceita como entrada um arquivo .csv e um arquivo .shp (*shapefile*)

3.2 Mineração e análise dos dados

A ferramenta escolhida para a mineração dos dados foi o VIS-STAMP (Diasheng, 2006) (*Visualization System for Space-Time and Multivariate Patterns*), que é um software para explorar dados multivariados espaço-temporal, descobrir padrões desconhecidos e

interessantes e apresentá-los de uma forma de fácil interpretação humana, raciocínio analítico e/ou tomada de decisão. As ferramentas do software integram métodos computacionais, visual e cartográfica em conjunto para detectar e visualizar padrões nos dados multivariados espaço-temporal e isso motivou a escolha da mesma, pois o domínio em que estamos trabalhando é fortemente relacionado ao tempo.

O vis-stamp usa um mapa auto-organizável (SOM) para processar os perfis multivariados (Figura 5). Cada nó colorido representa um cluster ou agrupamento, e nós mais próximos são mais parecidos. Quanto maior o círculo colorido, maior será a quantidade de dados pertencentes a aquela *cluster*. Possi também um *Plot* de Coordenadas Paralelas (PCP) (Figura 2), para visualizar os conjuntos de dados e assim complementar a análise das características de cada cluster identificado pelo SOM.

Dispõe de uma matriz de ordenação (Figura 4). As linhas da matriz são reordenadas através do cálculo da dissimilaridade entre elas. A onde a dissimilaridade entre dois objetos espaciais pode ser considerada como a distância euclidiana entre suas fatias de atributos no tempo. A estratégia de ordenação é baseada no resultado da hierarquia de *clustering* (agrupamento).

Os mapas são ordenados numa matriz de mapas (Figura 3), organizada temporalmente, conforme foram ordenados na matriz reordenável.

Foi carregado um arquivo *shapefile*, e outro em formato *.csv*, este último contendo a base que será ligada ao *shape*.

4. Resultados e Discussão

Inicialmente, foi feita uma mineração de dados usando as seguintes variáveis: *royalties*, PIB e pobreza (Geração de *Royalties*, PIB *per capita*, Índice de Pobreza) e constatou-se uma correlação entre elas. Foi visto que com um aumento nos *royalties* incidia em um aumento no PIB e em uma diminuição na pobreza. O estado de São Paulo foi o único que não seguiu o padrão de proporcionalidade entre PIB, *royalties* e pobreza por possuir um quadro econômico anômalo em relação aos outros estados brasileiros.

Na Figura 2, as linhas coloridas representam os *clusters* formados. Pegando como exemplo a linha verde escura, que chamaremos “*cluster verde escuro*”, vemos que ela representa um grupo de estados que possuem o menor número de *royalties*, um PIB proporcionalmente baixo e um alto índice de pobreza em relação aos outros estados.

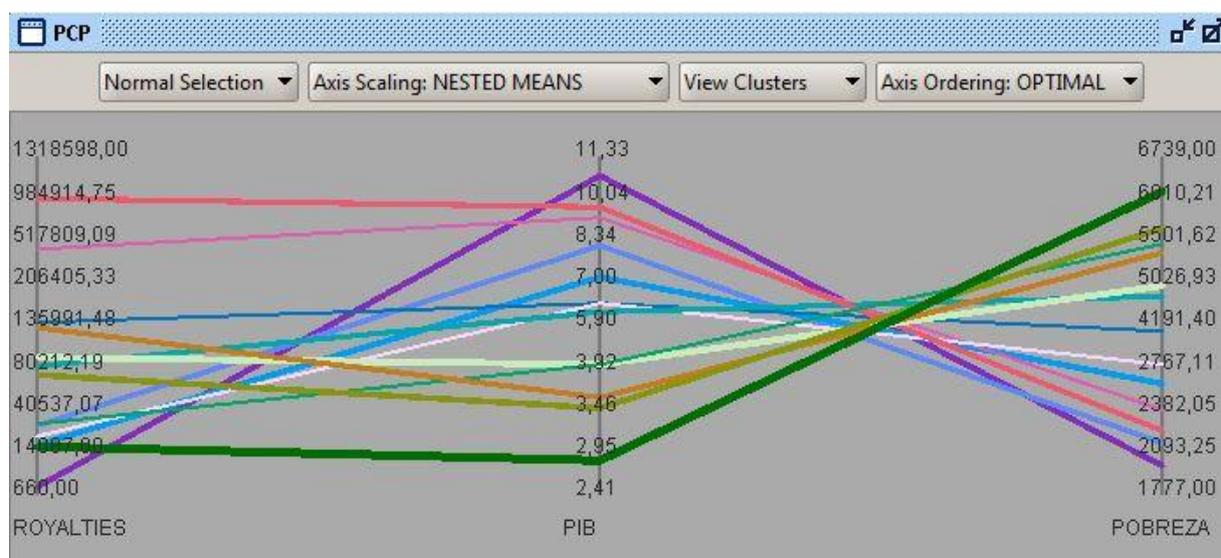


Figura 2 – PCP confrontando royalties x PIB x pobreza

Para entender que estados pertencem ao “*cluster verde escuro*” é necessário ver a Figura 3. Nota-se que entre 2001 e 2004 este *cluster* é formado pelos estados do Ceará e Alagoas. Analisando ainda a Figura 3, vemos que o estado do Ceará é representado por um verde mais claro, no último ano (2005), passando a fazer parte do “*cluster verde claro*”, ou seja, teve um aumento na sua geração de *Royalties*, seu PIB elevado e seu índice de pobreza diminuído. Podemos também observar um *cluster azul*, formado por Paraná e Espírito Santo, que possuem uma geração de *royalties* relativamente elevada, um alto PIB e um baixo índice de pobreza.

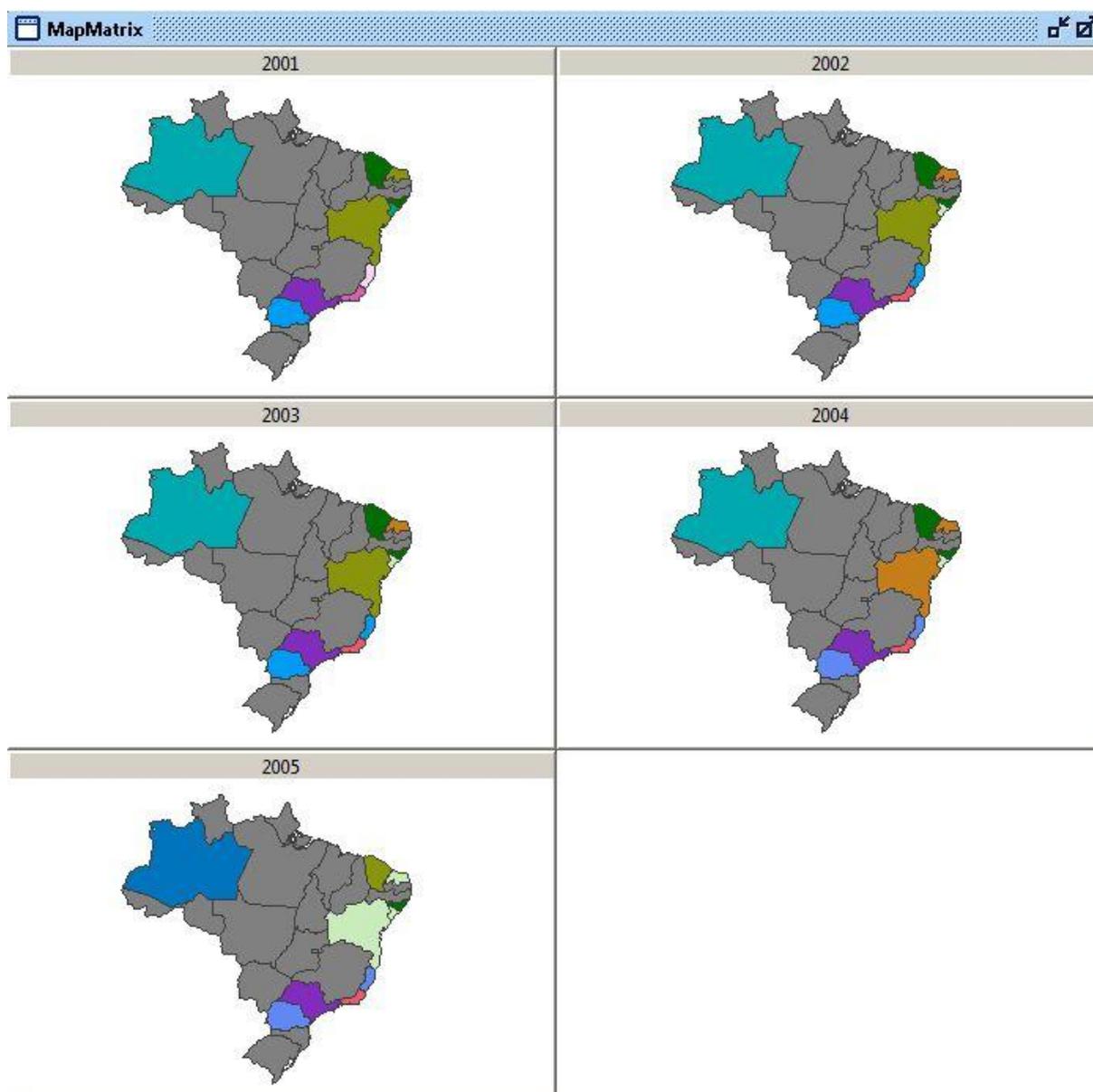


Figura 3 – Matriz de mapas representando os *clusters*



Figura 4 - Matriz reordenavel espaço-temporal

Posteriormente a base foi minerada, usando as variáveis *royalties* e investimento. Notou-se que ao longo dos anos, de 2001 a 2005, quase todos os estados e em quase todos os anos possuíam um investimento geralmente superior a quantidade de *royalties* gerados. Este padrão não se verificou para todos os estados, observando a Figura 7, vemos um cluster de cor lilás, muito maior que os outros, pois possui mais estados que os outros *clusters*.

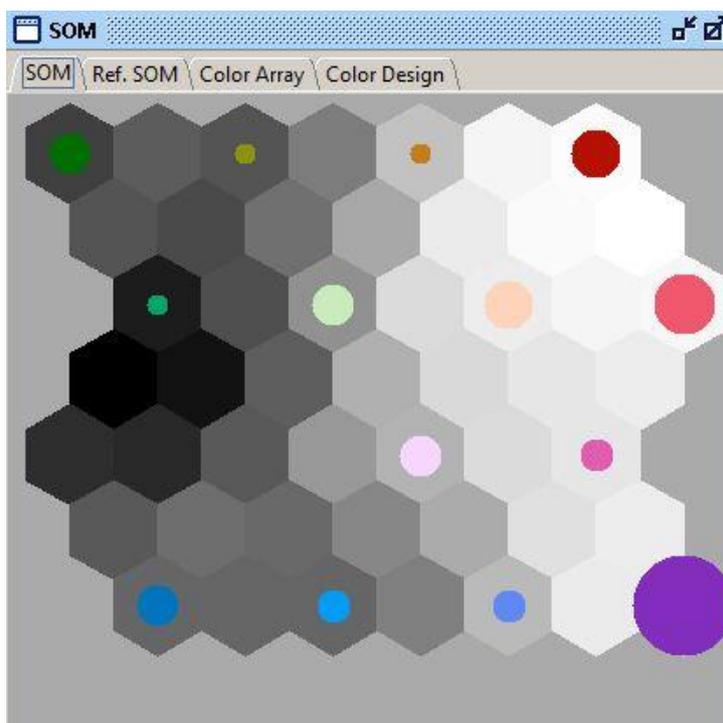


Figura 5 – SOM mostrando os *clusters* formados e suas grandezas

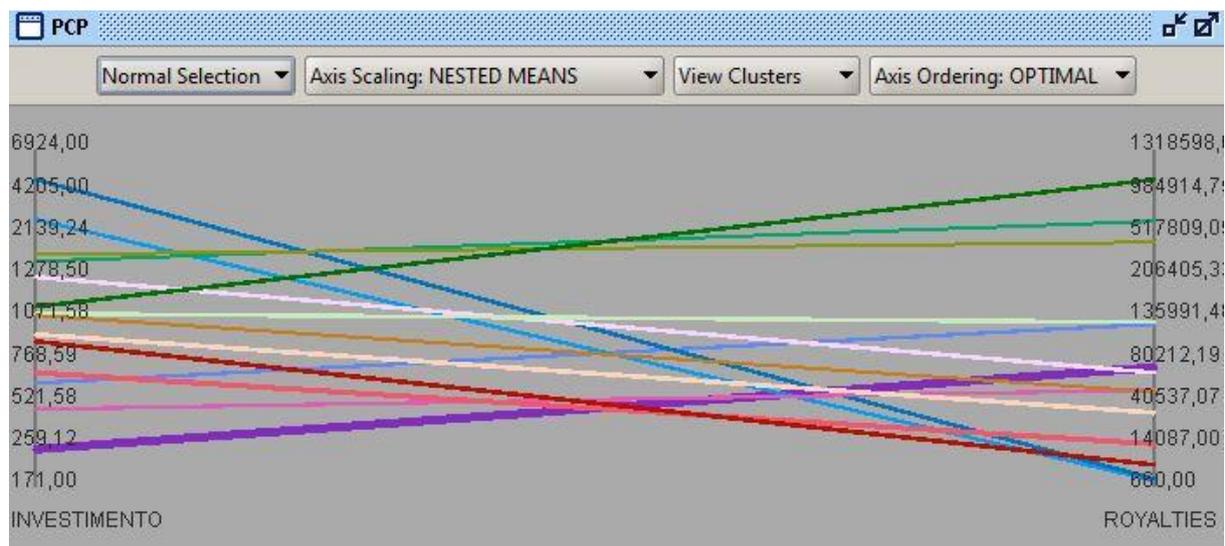


Figura 6 – PCP, resultado entre investimento x royalties

O “cluster lilás” é formado pelos estados do Rio Grande do Norte, Alagoas e Sergipe (Figura 7) e possuem os investimentos muito baixos em relação aos *royalties* gerados.

No ultimo ano (2005), Alagoas saiu do cluster formado pelo Rio Grande do Norte e Sergipe, tendo um relevante crescimento nos seus investimentos. Infelizmente para o Rio Grande do Norte e Sergipe, o quadro continuou parecido aos anos anteriores.

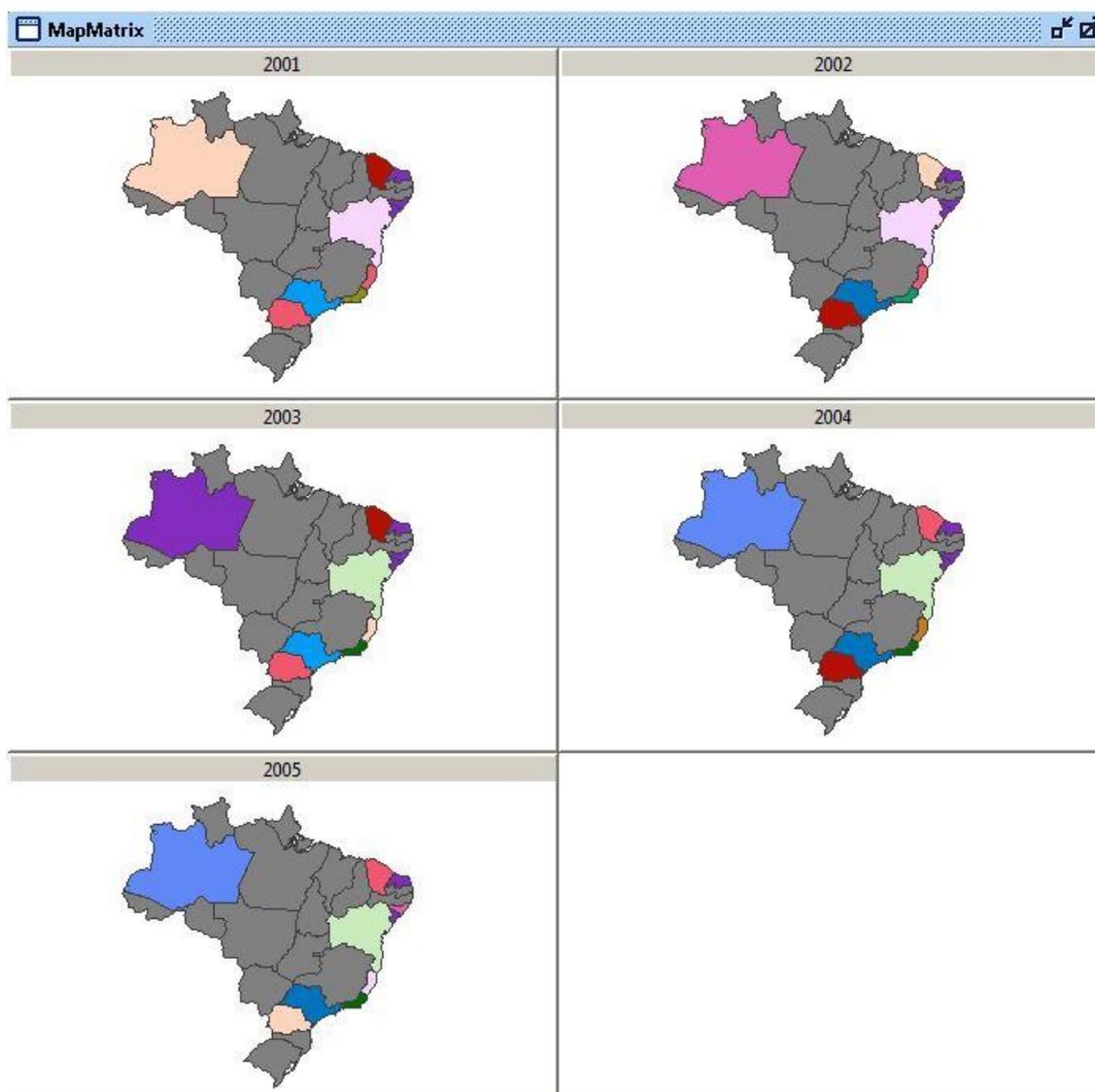


Figura 7 – Matriz de mapa em relação aos *royalties* x investimentos

5. Conclusões

Foi verificado que o Rio Grande do Norte, Alagoas e Sergipe possuem um baixo índice de capital investido em relação aos outros estados produtores. Os três estão localizados na região Nordeste, que segundo uma pesquisa divulgada pelo IBGE (Instituto Brasileiro de Pesquisa e Estatística), é a região cuja população está mais longe de superar a linha de pobreza. Segundo o Atlas do Bolso dos Brasileiros (FGV, 2009), Alagoas é o estado mais pobre do Brasil e Sergipe é o quinto mais pobre. Foi constatado também que os estados situados no Sul do país, possuem um investimento alto em relação aos seus *royalties*, assim como um PIB elevado e um índice de pobreza expressivamente menor.

Concluimos que a distribuição de *royalties*, embora contribua para a melhoria do PIB e diminuição da pobreza nos estados produtores de petróleo, nem sempre é sinônimo de desenvolvimento econômico, pois como foi visto, os índices de investimentos em alguns estados estão bem abaixo dos números de *royalties* que eles geram, resultado este que provoca uma série de interrogações a respeito de uma possível política de investimentos do dinheiro público dentro dos estados produtores.

Referências

Pacheco, C. A. G. **O Impacto dos *Royalties* de Petróleo no Desenvolvimento Econômicos dos Municípios da Região Norte Fluminense**. 2003.

ANP. Agência Nacional do Petróleo. Disponível em <<http://www.anp.gov.br>>. Acesso em 08 de nov. 2010.

Diasheng, G., Jin, C., Maceachren, M. A., Ke, L., **A Visual Inquiry System for Space-Time and Multivariate Patterns (VIS-STAMP)**, 2006.

Fayyad, U. M.; Piatetsky-Shapiro G. e Smyth P. **From Data Mining to Knowledge Discovery: An Overview**. In: *Advances in Knowledge Discovery and Data Mining*. Menlo Park, AAAI Press, 1996, pp 11-34.

Roddick, J. F.; Hornsby, K., **Eds. Temporal, Spatial, and Spatio-Temporal Data Mining**. CARBONELL, J.G.; SIEKMANN, J. *Lecture Notes in Artificial Intelligence*. Berlin: Springer, 2001.

Silva, P. S. M., **Mineração de Dados - Conceitos, Aplicações e Experimentos com Weka**, 2004.

IPEA, Instituto de Pesquisa e Econômica Aplicada. Disponível em <<http://www.ipea.gov.br/portal>>. Acessado em 10 de nov. 2010.

Legislação Federal, 1997. Disponível em <<http://www.legislacao.planalto.gov.br>>. Acessado em 10 de nov. 2010.

IBGE. Instituto Brasileiro de Geografia e Estatística. Disponível em <<http://www.ibge.gov.br>>. Acessado em 13 de nov. 2010.

FGV, Fundação Getúlio Vargas. **Atlas de Bolso do Brasileiro**, DF, 2009. Disponível em <<http://www.fgv.br>>. Acessado em 13 de nov. 2010.