

UNIVERSIDADE FEDERAL DE SÃO CARLOS
CENTRO DE CIÊNCIAS E TECNOLOGIAS PARA A SUSTENTABILIDADE
CAMPUS DE SOROCABA
CURSO DE GRADUAÇÃO EM CIÊNCIAS ECONÔMICAS

RAPHAEL ROBERTO DE GÓES REIS

**AVALIAÇÃO DO DESEMPENHO DOS MUNICÍPIOS DA REGIÃO
ADMINISTRATIVA DE SOROCABA UTILIZANDO TÉCNICAS DE ANÁLISE
MULTIVARIADA**

Sorocaba
2013

UNIVERSIDADE FEDERAL DE SÃO CARLOS
CENTRO DE CIÊNCIAS E TECNOLOGIAS PARA A SUSTENTABILIDADE
CAMPUS DE SOROCABA
CURSO DE GRADUAÇÃO EM CIÊNCIAS ECONÔMICAS

RAPHAEL ROBERTO DE GÓES REIS

**AVALIAÇÃO DO DESEMPENHO DOS MUNICÍPIOS DA REGIÃO
ADMINISTRATIVA DE SOROCABA UTILIZANDO TÉCNICAS DE ANÁLISE
MULTIVARIADA**

Trabalho de Conclusão de Curso apresentado ao Centro de Ciências e Tecnologias para a Sustentabilidade da Universidade Federal de São Carlos, *campus* Sorocaba, para a obtenção do título de Bacharel em Ciências Econômicas.

Orientação: Prof. Dr. Alexandre Lopes Gomes

Sorocaba
2013

Ficha catalográfica

Reis, Raphael

Avaliação do desempenho dos municípios da Região Administrativa de Sorocaba Utilizando Técnicas de Análise Multivariada / Raphael Roberto de Góes Reis. -- Sorocaba, 2013
53 f. : il. ; 28 cm

Trabalho de Conclusão do Curso de Bacharelado em Ciências Econômicas - UFSCar, *Campus* Sorocaba, 2013.

Orientador: Prof. Dr. Alexandre Lopes Gomes

Banca examinadora: Prof. Dr. Geraldo Edmundo Silva Junior, Prof. Dr. José César Cruz Junior

Bibliografia

1. Região Administrativa de Sorocaba. 2. Análise Multivariada. 3. Avaliação de desempenho municipal. I. Avaliação do desempenho dos municípios da Região Administrativa de Sorocaba Utilizando Técnicas de Análise Multivariada. II. Sorocaba-Universidade Federal de São Carlos.

CDD 330

RAPHAEL ROBERTO DE GÓES REIS

**AVALIAÇÃO DO DESEMPENHO DOS MUNICÍPIOS DA REGIÃO
ADMINISTRATIVA DE SOROCABA UTILIZANDO TÉCNICAS DE ANÁLISE
MULTIVARIADA**

Trabalho de Conclusão de Curso
apresentado ao Centro de Ciências e
Tecnologias para a Sustentabilidade da
Universidade Federal de São Carlos,
campus Sorocaba, para a obtenção do
título de Bacharel em Ciências
Econômicas.

Orientador

Dr. Alexandre Lopes Gomes
Universidade Federal de São Carlos – Campus Sorocaba

Examinador

Dr. Geraldo Edmundo Silva Junior
Universidade Federal de São Carlos – Campus Sorocaba

Examinador

Dr. José César Cruz Junior
Universidade Federal de São Carlos – Campus Sorocaba

AGRADECIMENTOS

Agradeço primeiramente a Deus, pela vida e por permitir que eu pudesse realizar este curso de graduação.

Agradeço a meus pais pelo apoio e pelos ensinamentos, que me ajudaram muito em minhas conquistas.

A todos os meus familiares, por não deixarem de acreditar em mim em nenhum momento.

A todos os meus amigos e colegas, os quais sempre se mostraram dispostos a me ajudar em minhas dificuldades e dividir comigo minhas alegrias.

A todos os meus professores, em especial ao professor Alexandre Lopes Gomes, que com paciência e bom humor me ensinaram muito durante toda a minha formação.

Finalmente, agradeço à UFSCar – Campus Sorocaba e a todos do curso de Economia, os quais me proporcionaram um ambiente acolhedor e experiências que pretendo jamais esquecer em minha vida.

RESUMO

REIS, Raphael. *Avaliação do desempenho dos municípios da Região Administrativa de Sorocaba utilizando técnicas de Análise Multivariada*. 2013. 53 f. Trabalho de Conclusão de Curso (Graduação em Ciências Econômicas) – Centro de Ciências e Tecnologias para Sustentabilidade, Universidade Federal de São Carlos, Sorocaba, 2013.

A Região Administrativa de Sorocaba (RAS), localizada em uma área de 41.077 km² na porção sudoeste do estado de São Paulo, é constituída por 79 municípios que recebem influencia econômica e política, em maior ou menor grau, do município de Sorocaba. Com a necessidade crescente de um fortalecimento regional ante o cenário macroeconômico atual, os municípios da região têm buscado se integrar economicamente seja através de atividades comerciais inter-município seja através de tentativas políticas de montar centros de dinamismo econômico. No entanto, mesmo com os dados agregados indicando que a RAS apresentando evolução tanto em variáveis econômicas quanto sociais e estruturais, o que se verifica é que a heterogeneidade dos municípios que nela estão contidos pode se tornar um obstáculo para o desenvolvimento municipal e regional. Diante desse cenário regional, torna-se importante a identificação dos municípios mais e menos desenvolvidos em relação a variáveis inseridas nas dimensões de Educação, Saúde, Emprego, Infra-estrutura, Política, entre outras. Para isso, é proposto um estudo conjunto de diversas variáveis em cada uma das referidas dimensões e a classificação dos municípios da Região Administrativa de Sorocaba através de métodos de Análise Multivariada. A Análise Multivariada permite com que se agrupe um determinado número de variáveis em um número menor de variáveis denominadas fatores, permitindo a ordenação dos municípios da RAS quanto ao seu desempenho nas distintas áreas consideradas importantes para o desenvolvimento municipal. Os resultados do modelo indicaram que a análise multivariada é viável para as variáveis selecionadas, o que possibilitou a criação de quatro fatores denominados: Urbanização e Atividade Econômica; Desempenho Setorial e Arrecadação Fiscal; Infraestrutura Urbana e; Condições e Desempenho Sociais. Os municípios da RAS foram hierarquizados através do escore fatorial para cada um dos fatores e, com isso, foram identificados padrões regionais de municípios com maiores e com menores desempenhos, formando os denominados *clusters*. A união dos desempenhos de cada município para cada fator permitiu a criação de um escore de desempenho que abrange todas as dimensões consideradas importantes para o desenvolvimento municipal, podendo ser utilizado como *proxy* deste. Dessa forma, foram também hierarquizados os municípios da RAS quanto a seu desenvolvimento e identificados *clusters* de municípios mais e menos desenvolvidos.

Palavras-chave: Região Administrativa de Sorocaba. Análise Multivariada. Avaliação de desempenho municipal

ABSTRACT

The Administrative Region of Sorocaba (ARS), located in an area of 41.077 km² in the southwestern portion of the state of São Paulo, is composed of 79 municipalities. These municipalities receive economic and political influence from the city of Sorocaba, in greater or lesser degree. With the growing need for regional strengthening, given the current macroeconomic environment, the municipalities of the region have sought to integrate economically whether through commercial activities between them or through political attempts to set up centers of economic dynamism. However, even with the aggregated data indicating that the ARS presents developments in both social, economic and structural variables, it is notorious that the heterogeneity of the municipalities of the region might become an obstacle to the municipal and regional development. Given this regional scenario, it becomes important to identify the most and least developed municipalities in relation to variables included in the dimensions of Education, Health, Employment, Infrastructure, Policy, among others. For this reason, it is proposed the study of several variables in each of these dimensions, and the classification of municipalities of the ARS through Multivariate Analysis. Multivariate Analysis allows to group a number of variables into a smaller number of variables called factors, allowing the ordination of municipalities of the ARS on their performance into the different areas considered important for municipal development. The model results indicated that multivariate analysis is feasible for the selected variables, allowing the creation of four factors called: Urbanization and Economic Activity, Sector Performance and Tax Revenue, Urban Infrastructure; Conditions and Social Performance. The municipalities of ARS were ranked by factor score for each factor and, thus, were identified regional patterns of cities with larger and smaller performances, forming the so-called *clusters*. The union of the performances of each municipality for each factor allowed the creation of a performance score that encompasses all dimensions considered important for municipal development. This performance score can be used as proxy of municipal development. Thereby, the municipalities of ARS was also hierarchized based on its development and it was identified *clusters* of the more and the less developed cities.

Keywords: Administrative Region of Sorocaba. Multivariate Analysis. Evaluation of Municipal Performance

LISTA DE FIGURAS

Figura 1 – Dimensões a serem consideradas para analisar o Desempenho Municipal.....	10
Figura 2 – Composição da População da RAS.....	12
Figura 3 – Participação do PIB da RAS no PIB de São Paulo.....	13
Figura 4 – PIB per capita da RAS deflacionado pelo IGP-2000.....	13
Figura 5 – Despesas da RAS com Educação – Deflacionadas com base IGP-2000.....	14
Figura 6 – Despesas da RAS com Saúde - Deflacionadas com base IGP 2000	16
Figura 7 – Saneamento Básico da RAS	17
Figura 8 – Rotação fatorial ortogonal.....	25
Figura 9 – Análise Scree Plot dos Autovalores.....	33
Figura 10 – Desempenho dos Municípios da RAS em Urbanização e Atividade Econômica.....	37
Figura 11 – Desempenho dos Municípios da RAS em Desempenho Setorial e Arrecadação Fiscal.....	39
Figura 12 – Desempenho dos Municípios da RAS em Infraestrutura Urbana.....	40
Figura 13 – Desempenho dos Municípios da RAS em Condições e Desempenho Sociais	42
Figura 14 – Desempenho dos Municípios da RAS em Desenvolvimento Municipal.....	45

LISTA DE TABELAS

Tabela 1 – Taxa de Analfabetismo e Taxa de Aprovação Escolar da RAS.....	15
Tabela 2 – Taxa de Mortalidade Infantil e Índice de Envelhecimento da RAS.....	16
Tabela 3 – Análise Descritiva das Variáveis Seleccionadas.....	27
Tabela 4 – Componentes principais para as variáveis seleccionadas.....	32
Tabela 5 – Cargas fatoriais com fatores rotacionados.....	34
Tabela 6 – Divisão dos Escores Fatoriais em categorias de desempenho.....	36

SUMÁRIO

1. INTRODUÇÃO	9
2. BREVE REVISÃO BIBLIOGRÁFICA	19
3. METODOLOGIA	20
4. FONTE DE DADOS E SOFTWARE	26
5. RESULTADOS	26
5.1. ANÁLISE DESCRITIVA DOS DADOS.....	26
5.2. DISCUSSÃO DO MODELO.....	29
6. CONSIDERAÇÕES FINAIS	46
7. REFERÊNCIAS BIBLIOGRÁFICAS	49
ANEXOS	50

1. INTRODUÇÃO

A partir da elaboração da Constituição de 1988, ficou definida para os municípios brasileiros a classificação de entes federativos. A Constituinte trouxe, em decorrência disso, a ampliação das possibilidades de receita tributária aos governos locais, bem como das responsabilidades referentes aos serviços públicos prestados pelos municípios. Dessa forma, criou-se no Brasil um aparente processo de descentralização do poder público, ou seja, as demandas municipais ou regionais passaram a ser mais direcionadas para o poder público local ao invés de entidades estaduais ou federais.

Diante desse fator descentralizador, diversos autores, tais quais Alves et. al. (2008), Silva et. al. (2005) e Lemos et.al. (2001), voltaram seus estudos para questões pertinentes à esfera municipal de governo e às aglomerações regionais. O principal objetivo desses trabalhos foi o de encontrar uma maneira de estudar determinadas regiões brasileiras levando em consideração as questões políticas, sociais e econômicas de conjuntos de municípios, visando obter modelos que auxiliem na orientação de estudos de desempenho e desenvolvimento municipal.

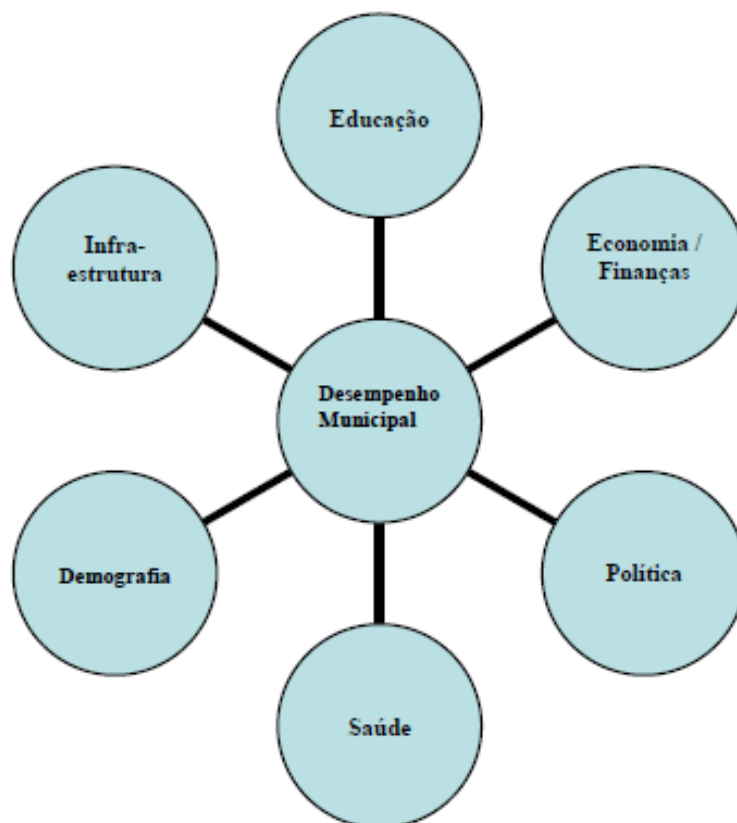
O conceito de desenvolvimento municipal, segundo Buarque (1999), está associado a um processo endógeno registrado em unidades municipais e agrupamentos humanos, capaz de promover o dinamismo econômico e a melhoria da qualidade de vida da população. Dessa forma, qualquer tentativa de classificar os desempenhos municipais, referente aos respectivos desenvolvimentos dos municípios, tem de contemplar as dimensões econômicas e demográficas dos municípios estudados.

Por outro lado, Castels e Borja (1996) também expressam a importância das dimensões políticas e sociais para o desenvolvimento municipal e, conseqüentemente, para a avaliação do desempenho municipal. Segundo eles, as experiências bem-sucedidas de desenvolvimento local decorrem, quase sempre, de um ambiente político e social favorável, expresso por uma mobilização, e, principalmente, de convergência importante dos atores sociais do município ou comunidade em torno de determinadas prioridades e orientações básicas de desenvolvimento.

Alves et. al. (2008), por sua vez, une essas dimensões em seu estudo de desempenho regional e recomenda o uso de diversas variáveis para que outras análises semelhantes possam ser realizadas. Segundo ele, os estudos de desempenho municipal

devem estar ligados, no mínimo, às dimensões de Educação, Saúde, Demográfica, Infraestrutura, Econômica/ Financeira e Política. A Figura 1 mostra a interdependência entre o desempenho municipal e essas dimensões.

Figura 1 – Dimensões a serem consideradas para analisar o Desempenho Municipal



Fonte: Alves et. al. (2008)

Uma das vantagens dos estudos de desempenho municipal é a possibilidade de se estabelecer medidas de comparação entre desempenhos locais. A determinação de uma medida de comparação do desempenho municipal, dentre muitas outras aplicações, é importante para identificar a existência de discrepâncias entre municípios próximos, havendo maior interesse de estudo se esses municípios se encontrarem numa mesma região. Dessa forma, a comparação entre desempenhos municipais dentro de uma determinada região pode permitir que órgãos municipais ou intermunicipais tenham um maior embasamento para seus planejamentos de médio e longo prazo e para uma futura construção de políticas públicas visando o desenvolvimento.

Um dos métodos mais utilizados para esse tipo de comparação entre desempenhos municipais é a chamada Análise Multivariada, que consiste no estudo de diversas variáveis simultaneamente para um determinado conjunto de dados. Esse método permite com que, dentre diversas variáveis, as variáveis mais semelhantes entre si sejam agrupadas em fatores, auxiliando, dessa forma, na classificação e interpretação dos dados selecionados.

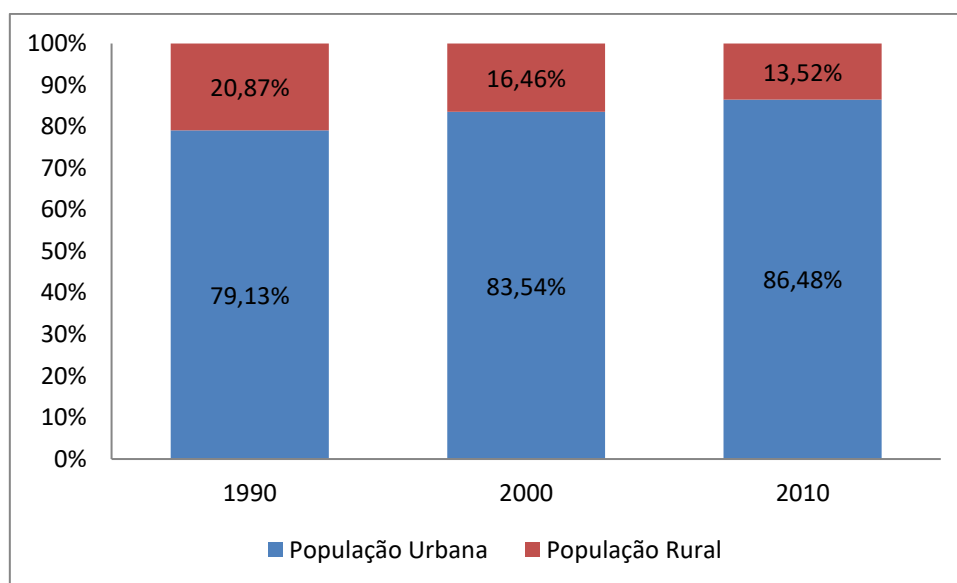
Dada a intenção deste estudo de verificar o desempenho de determinados municípios agrupados em um formato regional, optou-se por selecionar a Região Administrativa de Sorocaba para tal verificação.

A Região Administrativa de Sorocaba (RAS), localizada em uma área de 41.077 km² na porção sudoeste do estado de São Paulo e cortada em grande parte pelo eixo Raposo Tavares - Castello Branco, é constituída por 79 municípios que recebem influencia econômica e política, em maior ou menor grau, do município de Sorocaba.

Atualmente a RAS caracteriza-se como a terceira área de maior concentração da população paulista com aproximadamente 2,8 milhões de habitantes em 2010, apenas superada pela região de Campinas e pela região metropolitana de São Paulo. Dentre os municípios com maior concentração populacional (SEADE, 2010) se encontram a própria Sorocaba, com aproximadamente 586 mil habitantes, Itu, com aproximadamente 154 mil e Itapetininga com 144 mil. A taxa de crescimento da RAS, quando considerados os anos de 2000 a 2011, se encontra em aproximadamente 1,30% anual, havendo nesse período um aumento de aproximadamente 15,21% no total de habitantes.

Do total da população da RAS, pode-se verificar que a grande maioria habita a zona urbana em detrimento da zona rural (SEADE, 2010). Esse fenômeno de urbanização tem sido verificado em grande parte dos municípios da RAS, podendo-se perceber que inclusive em municípios mais voltados à produção agropecuária, como por exemplo, Ibiúna e Piedade, existe um aumento considerável da população residente na zona urbana. A Figura 2 mostra o aumento da população da zona urbana da RAS.

Figura 2 - Composição da População da RAS

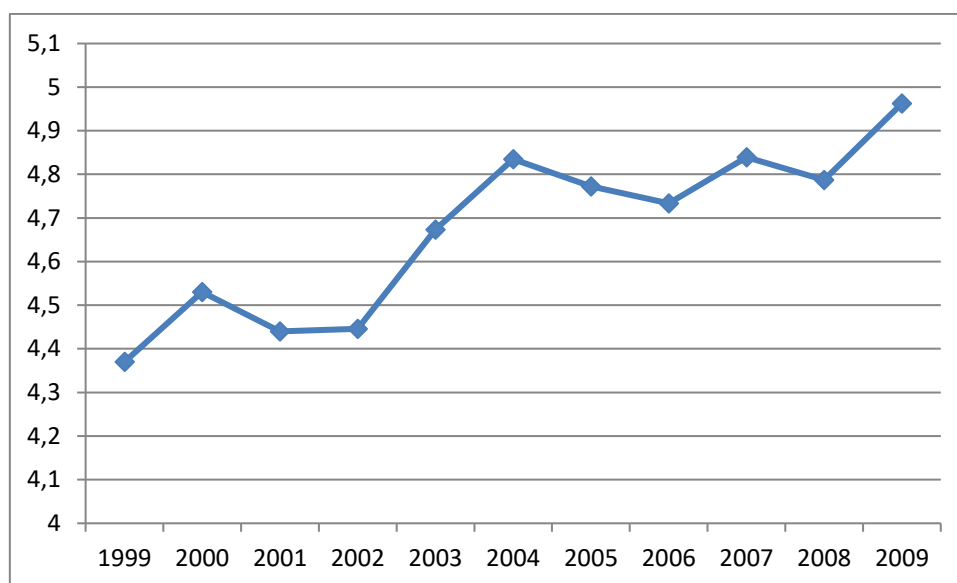


Fonte: Elaboração Própria

Como é possível se verificar através da Figura 2, o percentual de habitantes da zona urbana da RAS passou de 79,13% em 1990 para 86,48% em 2010, ilustrando o já referido fenômeno de urbanização de diversos municípios da RAS.

Em relação ao Produto Interno Bruto (PIB) da RAS, verifica-se que este tem se elevado de forma bastante expressiva nos últimos anos. De 1999 a 2009, a participação do PIB da RAS em relação ao total do estado de São Paulo passou de 4,37% para 4,96%, de acordo com a Figura 3.

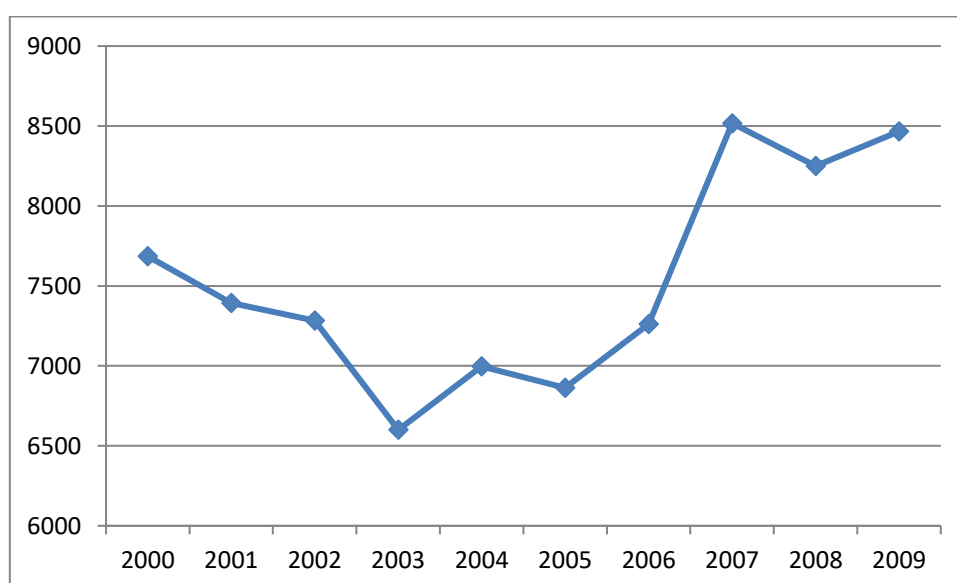
Figura 3 – Participação do PIB da RAS no PIB de São Paulo



Fonte: Elaboração Própria

O PIB per capita da RAS, deflacionado pelo IGP-DI com base em 2000, também mostra crescimento significativo, principalmente a partir do ano de 2003. Como se pode observar na Figura 4, o PIB per capita da RAS passou de R\$ 6.600 anuais em 2003 para R\$ 8.467 em 2009 um aumento da ordem de 28%.

Figura 4 – PIB per capita da RAS deflacionado pelo IGP-2000

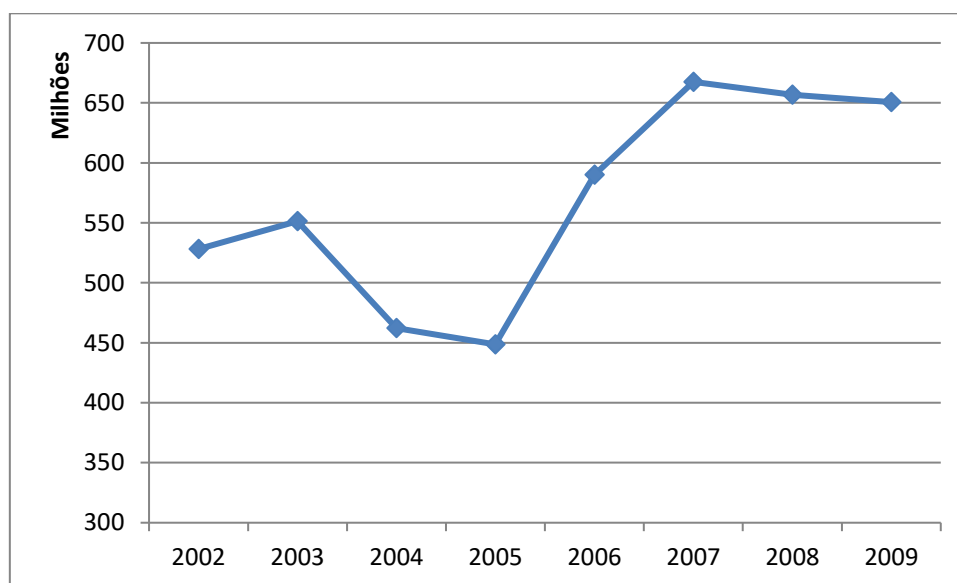


Fonte: Elaboração Própria

Em relação à Educação, os municípios que compõem a RAS apresentaram no período de 2002 a 2009 um aumento significativo no que diz respeito às despesas municipais, bem como melhora dos indicadores que expressam a qualidade do ensino nos municípios.

A Figura 5 mostra a evolução das despesas municipais da RAS, deflacionadas segundo o IGP com base em 2000. Deve-se atentar para o fato do cálculo tomar como base para as despesas totais o somatório apenas dos valores disponíveis para os municípios da RAS, sendo desconsiderados os valores indisponíveis para alguns dos anos da série.

Figura 5 – Despesas da RAS com Educação – Deflacionadas com base IGP-2000



Fonte: Elaboração Própria

Como se pode perceber pela Figura 5, as despesas reais com Educação na RAS, após um período de declínio entre os anos de 2003 e 2005, apresentaram tendência crescente de 2005 a 2007 (aumento de aproximadamente 48.9%), seguidas por um período sem grandes variações até o ano de 2009.

Quanto aos indicadores de desempenho da Educação pode-se verificar que estes também apresentaram crescimento nos períodos contemplados pelos dados. Conforme se pode perceber pela Tabela 1, a Taxa Média de Analfabetismo decresceu de 12,76% da população considerada em 1991 para 8,1%, em 2000. De 2000 a 2010 percebe-se também evolução nas Taxas de Aprovação Escolar da RAS, dado que a taxa correspondente ao Ensino Fundamental passou de 91,79% dos alunos para 94,4% e a taxa correspondente ao Ensino Médio passou de 83,79% dos alunos para 87,2%.

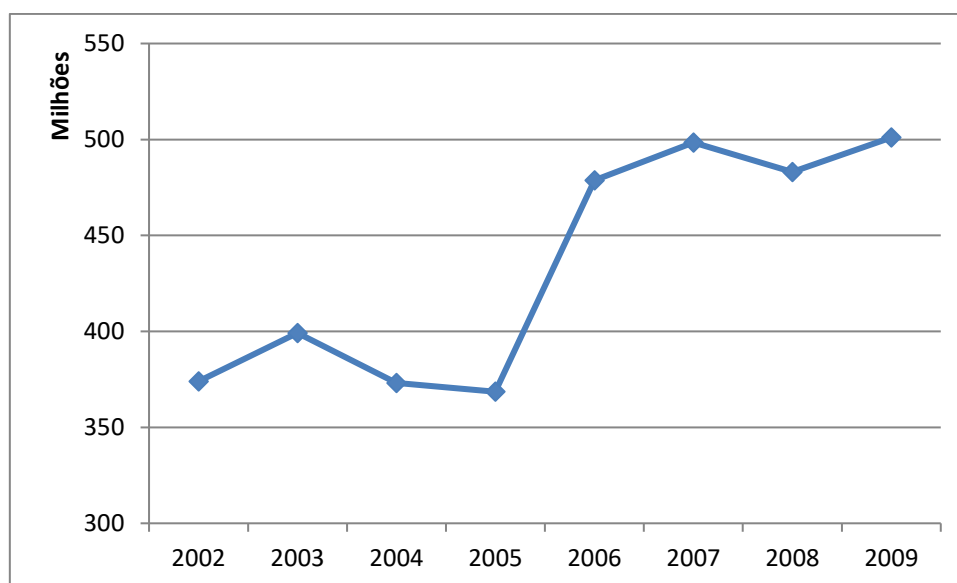
Tabela 1 – Taxa de Analfabetismo e Taxa de Aprovação Escolar da RAS

	1991	2000	2002	2003	2004	2005	2007	2008	2009	2010
Taxa de Analfabetismo da População de 15 Anos e Mais (Em %)	12,76	8,1	-	-	-	-	-	-	-	-
Taxa de Aprovação do Ensino Fundamental Total (Em %)	91,79	91,26	92,4	92,7	91,8	92,4	92,8	93,6	93,9	94,4
Taxa de Aprovação do Ensino Médio Total (Em %)	85,26	83,79	85,6	84,4	82,5	82,9	83,3	85,4	85,7	87,2

Fonte: Elaboração Própria

Quanto à Saúde da RAS pode-se encontrar resultado semelhante ao que ocorre nas despesas e indicadores de Educação, indicando aumento significativo de ambos. A Figura 6 mostra a evolução das despesas reais com Saúde e a Tabela 2 mostra a evolução dos indicadores de desempenho considerados.

Figura 6 - Despesas da RAS com Saúde - Deflacionadas com base IGP 2000



Fonte: Elaboração Própria

Como se percebe pela Figura 6, as despesas reais da RAS com Saúde passaram a assumir um comportamento crescente a partir de 2005, passando de aproximadamente R\$ 368 milhões em 2005 para R\$ 501 milhões em 2009, o que representou um aumento de aproximadamente 36% por cento.

Tabela 2 – Taxa de Mortalidade Infantil e Índice de Envelhecimento da RAS

	2001	2002	2003	2004	2005	2006	2007	2008
Taxa de Mortalidade Infantil (Por mil nascidos vivos)	19,87	15,95	17,37	14,68	14,89	15,59	14,65	14,11
Índice de Envelhecimento (Em %)	33,13	34,71	36,36	38,08	39,89	41,78	43,77	45,85

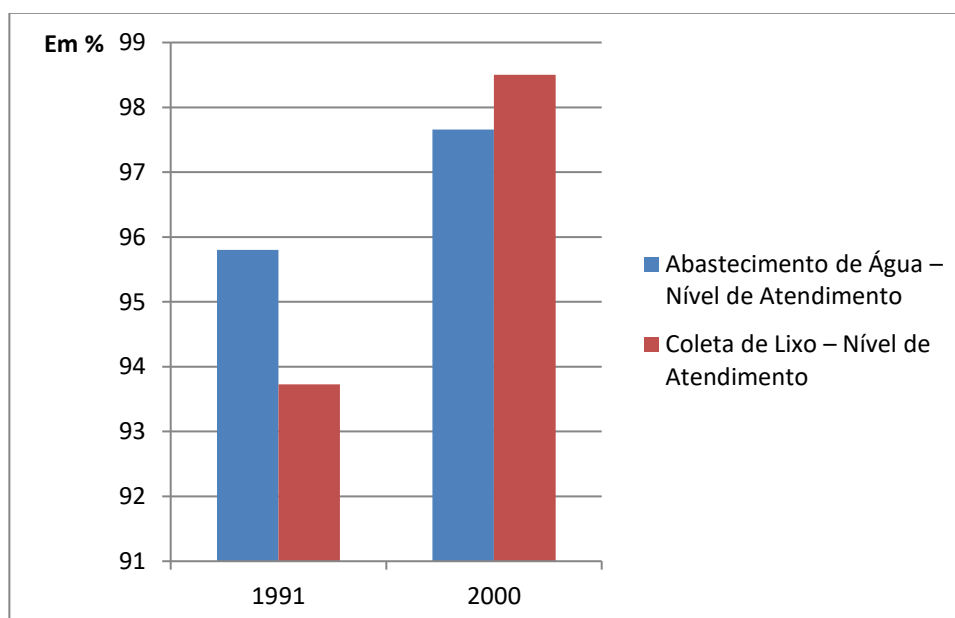
Fonte: Elaboração Própria

Os indicadores de desempenho da Tabela 2, por sua vez, também apresentaram resultados favoráveis bastante relevantes no período de 2001 a 2008. A Taxa de Mortalidade Infantil passou de aproximadamente 20 por mil nascidos vivos em 2001 para apenas 14 em 2008, ou seja, uma queda de aproximadamente 30%. O Índice de

Envelhecimento passou, por sua vez, de aproximadamente 33% de idosos em relação a crianças para aproximadamente 45%. Esses indicadores mostram que, aparentemente, as despesas da RAS na área da saúde têm gerado melhoras significativas para a população que vive na região.

Quanto à Infraestrutura da RAS, pode-se perceber que esta também se destaca pelos elevados e crescentes valores nos dados considerados. A Figura 7 ilustra a evolução dos dados de Saneamento Básico com Água Encanada e Coleta de Lixo entre os anos de 1991 e 2000, período mais recente contemplado pela base de dados utilizada.

Figura 7 – Saneamento Básico da RAS



Fonte: Elaboração Própria

Como se pode perceber pela Figura 7, de 1991 a 2000 o Abastecimento de Água aumentou de 95,8 para 97,66% dos domicílios atendidos da RAS. No caso da Coleta de Lixo, esta passou de 93,73 para 98,5% da dos domicílios atendidos da RAS. Isso representa que há uma tendência de melhora nos níveis de atendimento relacionados ao Saneamento Básico dos municípios, o qual tende a contemplar quase que a totalidade dos domicílios da RAS.

É possível perceber através dos dados, portanto, que de maneira geral, a Região Administrativa de Sorocaba tem apresentado através dos últimos anos resultados favoráveis no que diz respeito às dimensões econômicas, estruturais, demográficas, educacionais e de saúde. No entanto, o que se verifica é que, por ser formada por 79 municípios distintos, os resultados positivos estudados de maneira agregada não representam necessariamente a realidade de todos os municípios. Municípios menores e mais carentes de recursos, ou até mesmo de mão-de-obra qualificada para gerir esses recursos, podem apresentar baixo desempenho frente a outros municípios da RAS. Isso, portanto, pode fazer com que haja uma grande discrepância entre os desempenhos de municípios próximos geograficamente, prejudicando toda a região administrativa no que se refere a desenvolvimento socioeconômico presente e futuro.

Dessa forma, este estudo tem por objetivo avaliar o desempenho dos municípios da Região Administrativa de Sorocaba e verificar se existe algum padrão regional entre eles. Para isto serão usadas técnicas de análise multivariada, contemplando as diversas dimensões consideradas importantes para o desenvolvimento municipal e permitindo com que o desempenho dos municípios seja avaliado não apenas por uma variável, mas por um conjunto destas, de forma que esse desempenho seja mais bem interpretado.

Este trabalho está dividido em mais seis capítulos, além desta Introdução. O Capítulo 2 apresentará uma revisão bibliográfica com estudos que se valeram também da Análise Multivariada, buscando melhor entendimento sobre a abrangência da técnica e sobre os procedimentos a serem realizados na verificação de dados municipais. O Capítulo 3 descreve a metodologia utilizada para a análise multivariada dos dados. O Capítulo 4 descreve a base de dados utilizada para pesquisa e o software utilizado para os cálculos estatísticos. No Capítulo 5 são selecionadas variáveis que representem as dimensões de Educação, Saúde, Demográfica, Infraestrutura, Economia/ Finanças e Política. A partir dessas variáveis são calculadas suas estatísticas descritivas, importantes para conhecimento prévio dos dados a serem analisados. Também nesse capítulo são descritos os resultados obtidos na análise fatorial, a fim de que se identifiquem padrões regionais a partir desses resultados. O Capítulo 6 apresenta considerações finais a respeito do método utilizado e dos resultados por ele gerados na análise. O Capítulo 7 mostra as referências bibliográficas utilizadas para a composição deste estudo.

2. BREVE REVISÃO BIBLIOGRÁFICA

Segundo Costa (2006) apud Klefens (2009), a análise multivariada é um conjunto de técnicas que são adequadas para situações onde várias variáveis correlacionadas estão envolvidas. Dentre essas técnicas encontram-se a Análise dos Componentes Principais (ACP) e a Análise Fatorial (AF). A técnica dos componentes principais reduz a dimensão de uma matriz de dados originais através de combinações lineares, facilitando a interpretação desses dados. A análise fatorial, por sua vez, é o nome dado a uma classe de métodos estatísticos paramétricos (e não paramétricos) multivariados que correspondem a um grande número de métodos e técnicas que utilizam simultaneamente todas as variáveis do conjunto na interpretação do inter-relacionamento das variáveis observadas.

A Análise de Componentes Principais tem por objetivo descrever os dados contidos num quadro indivíduos-variáveis numéricas: p variáveis serão mediadas com n indivíduos (BOUROCHE, 1982). Ou seja, ela tem por objetivo reduzir o número de variáveis de um conjunto de dados e explicar a estrutura de variâncias-covariâncias por poucas combinações lineares das variáveis originais (ARTES, 1998 apud KLEFENS, 2009).

A Análise Fatorial, por sua vez, tem por objetivo descrever a ação de um conjunto grande de variáveis em termos de alguns fatores comuns, isto é, estuda uma quantidade de variáveis originais usando um número menor de fatores obtidos a partir dessas variáveis originais. Assume também que cada variável original pode ser escrita como uma combinação linear desses fatores mais um termo residual (KLEFENS, 2009).

Embora as duas análises tenham algumas diferenças entre si, os resultados propostos por elas geralmente são bastante similares. Segundo Bezerra (2012), tanto o método de extração de fatores por Análise Fatorial Comum quanto por Análise dos Componentes Principais podem ser utilizados, mas deve-se atentar para o fato da ACP levar em conta a variância total dos dados e a AF apenas a variância comum entre eles.

A Análise dos Componentes Principais está presente em diversos trabalhos na área de economia regional.

Tolosa (1977) apud Haddad (1989), a partir das observações de vários indicadores socioeconômicos relativos a itens do bem-estar da população, cria três índices de pobreza urbana para as 95 principais cidades e aglomerados urbanos brasileiros. A extração de fatores pelo método das componentes principais foi utilizada em seu trabalho para reduzir todo o conjunto de dados relativos às características sociais dos locais estudados, ponderando as variáveis utilizadas e permitindo a criação dos índices de pobreza.

Silva et. al. (2005) utiliza a análise dos componentes principais para estudar a dinâmica agropecuária e urbanização do estado de Minas Gerais. Para identificar o padrão regional de modernização do setor primário, as componentes principais são utilizadas na identificação da variância de dados de qualidade de vida da população, infraestrutura urbana e de condições de oferta de mão-de-obra.

Lemos et.al. (2001) utiliza a ACP em dados dos centros urbanos, tais quais: renda fundiária, produtividade do trabalho, verticalização urbana, entre outros. Seu objetivo é estudar a dinâmica urbana das regiões metropolitanas brasileiras através do comportamento de fatores aglomerativos e desaglomerativos, evidenciados pela ACP.

Silveira et. al. (2008) utiliza a análise fatorial com extração de componentes principais para estudar variáveis relacionadas à qualidade de vida e criar um Índice Relativo de Qualidade de Vida (IRQV) para a Região Norte do país, ordenando e classificando os municípios de acordo com o IRQV obtido.

Portanto, pode-se inferir através dos já referidos autores, que para se cumprir o objetivo de avaliar o desempenho dos municípios da RAS nas dimensões explicitadas na Figura 1, é possível a utilização da análise fatorial com extração dos componentes principais. Esse método permite não só a construção de fatores, mas também a classificação de unidades municipais em relação a esses fatores.

3. METODOLOGIA

O método da extração de componentes principais, como já destacado, tem sido bastante utilizado em análise regional e urbana, com a finalidade classificatória de

regiões ou de municípios por meio da criação de índices ou critérios que permitem sua hierarquização.

Segundo Bezerra (2012, p.74), a análise fatorial é uma técnica estatística que busca, através da avaliação de um conjunto de variáveis, a identificação de dimensões e variabilidade comuns existentes em um conjunto de fenômenos, sendo o intuito desvendar estruturas existentes, mas não observáveis diretamente. Cada uma dessas dimensões de variabilidade comum recebe o nome de fator.

A aplicação do método de determinação de fatores inicia-se com o cálculo da matriz S (no caso do uso das variáveis originais) ou a matriz R (no caso do uso das variáveis padronizadas). O objetivo para ambas as matrizes é o mesmo: verificar se as variáveis estão correlacionadas umas em relação às outras. Caso não estejam, deve-se aplicar o teste do Kaiser-Meyer-Olkin (KMO), ou fazer um teste que verifique se as correlações entre as variáveis são significativas, ou não, para verificar se é possível proceder a análise dos dados aplicando esta técnica (VICINI, 2005).

Em análises voltadas à Economia Regional, é mais comum o uso da matriz de correlação R. A matriz de correlação R é utilizada quando se necessita de uma padronização dos dados, evitando-se problemas como a influência da magnitude das variáveis (SOUZA, 2000, apud VICINI, 2005).

Considerando-se X_1, X_2, \dots, X_p , as variáveis originais, a estimativa da matriz de correlação (que é igual à estimativa da matriz de variância-covariância entre as variáveis padronizadas Z_1, Z_2, \dots, Z_p) é denotada por R, da seguinte forma:

$$R = \begin{bmatrix} 1 & r_{12} & \dots & r_{1p} \\ r_{21} & 1 & \dots & r_{2p} \\ \dots & \dots & \dots & \dots \\ r_{p1} & r_{p2} & \dots & 1 \end{bmatrix} \quad (3.1)$$

na qual:

$$r_{j\bar{j}} = r(X_j, X_{\bar{j}}) = C\hat{o}v(Z_j, Z_{\bar{j}}) = \frac{C\hat{o}v(X_j, X_{\bar{j}})}{\sqrt{Var(X_j), Var(X_{\bar{j}})}} \quad (3.2)$$

A determinação da matriz de correlação muitas vezes já é suficiente para, através de uma observação preliminar dos valores obtidos, prosseguir com a extração de fatores dos dados. No entanto, é recomendável a utilização do teste KMO e de Bartlett para verificar o nível de adequação dos dados para a análise fatorial.

Segundo Vicini (2005), a medida de KMO é a razão da soma dos quadrados das correlações de todas as variáveis dividida por essa mesma soma, acrescida da soma dos quadrados das correlações parciais de todas as variáveis. Para encontrar o valor do KMO, utiliza-se a expressão:

$$KMO = \frac{\sum_i \sum_j r_{ij}^2}{\sum_i \sum_j r_{ij}^2 + \sum_i \sum_j a_{ij}^2} \quad (3.3)$$

Onde:

r_{ij} = é o coeficiente de correlação observado entre as variáveis i e j .

a_{ij} = é o coeficiente de correlação parcial entre as mesmas variáveis, que é, simultaneamente, uma estimativa das correlações entre os fatores.

Para a interpretação do KMO consideram-se os valores: 0,90-1,00: excelente; 0,80-0,90: ótimo; 0,70-0,80: bom; 0,60-0,70: regular; 0,50-0,60: ruim; 0,00-0,50: inadequado.

O *Barlett Test of Sphericity (BTS)*, por sua vez, testa a hipótese de que a matriz de correlação é uma matriz identidade, ou seja, que não há correlação entre as variáveis. A hipótese nula poderá ser rejeitada caso o α adotado for igual a 5% e o valor encontrado for inferior ao valor de α . O teste de Bartlett na aplicação da ACP pressupõe que se rejeite a hipótese nula $H_0 = P = I$ (PEREIRA, 2001, p. 124 e 125, apud VICINI, 2005).

Após a confirmação da viabilidade da análise fatorial nos dados, pode-se iniciar o método da extração de fatores para agrupamento de variáveis.

Por avaliar a correlação existente entre um grande número de variáveis e identificar a possibilidade de essas variáveis serem agrupadas em um número menor de variáveis latentes, pode-se dizer que a análise fatorial possibilita o agrupamento de i variáveis ($X_1, X_2, X_3, \dots, X_i$) em um número menor de j fatores ($F_1, F_2, F_3, \dots, F_j$).

Dessa forma, o modelo matemático para a análise é o seguinte:

$$X_i = a_{i1}F_1 + a_{i2}F_2 + a_{i3}F_3 + \dots + a_{ij}F_j + e_i \quad (3.5)$$

Onde X_i são as variáveis padronizadas, a_i são as cargas fatoriais, F_j são os fatores comuns não relacionados entre si e e_i é um erro que representa a parcela de variação da variável i que é exclusiva dela e não pode ser explicada por um fator nem por outra variável do conjunto analisado. As cargas fatoriais, por sua vez, são valores que medem o grau de correlação entre a variável original e os fatores.

Dessa forma, os fatores podem ser estimados por uma combinação linear das variáveis originais. Assim, tem-se:

$$F_j = \omega_{j1}X_1 + \omega_{j2}X_2 + \omega_{j3}X_3 + \dots + \omega_{ji}X_i \quad (3.6)$$

Onde F_j são os fatores comuns não relacionados, ω_{ji} são os coeficientes dos escores fatoriais e X_i são as variáveis originais envolvidas no estudo.

O escore fatorial é um número resultante da multiplicação dos coeficientes (ω_{ji}) pelo valor das variáveis originais. Quando existe mais de um fator, o escore fatorial representa as coordenadas da variável em relação aos eixos, que são os fatores. Esses escores podem depois ser hierarquizados, mostrando o grau de identificação de cada unidade estudada com o conjunto de variáveis reunido em cada fator.

Segundo Bezerra (2012, p.81), o método mais comum de extração dos fatores é o método da ACP, pelo qual se procura uma combinação linear entre as variáveis, de forma que o máximo de variância seja explicado por essa combinação. Em seguida,

retira-se a variância já explicada no passo anterior e busca-se uma nova combinação linear entre as variáveis que explique a maior quantidade de variância restante, e assim por diante. Este procedimento resulta em fatores ortogonais, ou seja, não correlacionados entre si.

Quanto à escolha do número de fatores, existem diversos métodos para que essa escolha seja realizada de maneira a melhor sumarizar os dados e permanecer fiel aos maiores percentuais de variação dos fatores. Os dois métodos mais utilizados para determinar o número de fatores do estudo são o critério de Kaiser e o critério scree plot.

Pelo critério de Kaiser, apenas os fatores com autovalores acima de 1 são considerados. O autovalor corresponde a quanto o fator consegue explicar da variância, ou seja, quanto da variância total dos dados pode ser associado ao fator.

O critério scree plot consiste na análise gráfica dos autovalores, seguindo o raciocínio de que grande parcela de variância será explicada pelos primeiros fatores e que entre eles haverá sempre uma diferença significativa. Quando essa diferença se torna pequena (suavização da curva) este ponto determina o número de fatores a serem considerados.

Definidos os números de fatores a serem extraídos e encontrados as cargas e os escores fatoriais, o próximo passo é a interpretação dos resultados obtidos. No entanto, algumas técnicas podem ser utilizadas para aumentar o poder de explicação da análise fatorial antes da interpretação. Uma dessas técnicas é a rotação dos fatores.

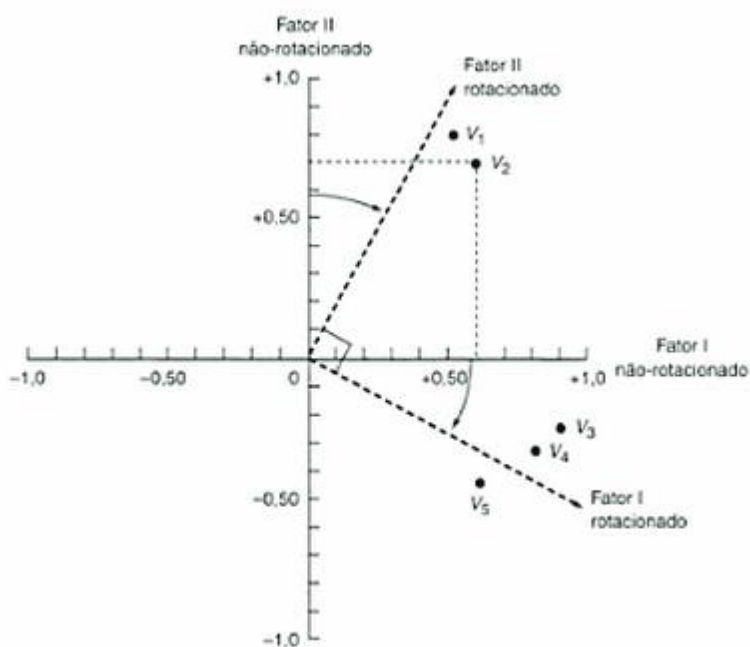
Como são comuns os casos em que numa primeira extração (não-rotacionada) os fatores estejam relacionados à grande maioria das variáveis, a rotação dos eixos fatoriais coloca os fatores em posições em que serão associadas só às variáveis relacionadas distintamente a um fator (VICINI, 2005).

O método de rotação mais utilizado é a o da rotação *varimax*. A rotação *varimax* é um tipo de rotação ortogonal que tem como característica o fato de minimizar a ocorrência de uma variável possuir altas cargas fatoriais para diferentes fatores. A rotação *varimax* é possível pois as cargas fatoriais podem ser representadas como pontos entre eixos (fatores) e estes eixos podem ser girados sem alterar a distância entre os pontos. Contudo, cabe observar que as coordenadas do ponto em relação aos eixos são alteradas, ou seja, as cargas fatoriais são alteradas na rotação. Além disso, conforme

Pereira (2001) apud Vicini (2005), “a rotação da matriz não afeta a inércia (comunalidades) das variáveis nem a percentagem de variações explicadas pelos fatores”.

HAIR et.al. (2005) ilustra o procedimento realizado quando há a rotação dos eixos na análise fatorial num modelo com apenas cinco variáveis como pontos e dois fatores como eixos, sendo que estes sofrem uma rotação ortogonal. Essa ilustração se encontra na Figura 8, abaixo.

Figura 8 – Rotação fatorial ortogonal



Fonte: HAIR et.al. (2005)

Como se pode perceber pela Figura 8, o método de rotação ortogonal dos eixos apenas aumenta o poder de explicação dos fatores ao aproximar para cada um dos fatores escolhidos as cargas fatoriais obtidas para as variáveis de estudo mais relacionadas a ele.

4. FONTE DE DADOS E SOFTWARE

Os dados coletados foram provenientes da Fundação Sistema Estadual de Análise de Dados (SEADE) e são utilizados para uma avaliação do desempenho de cada município por se encaixarem nas dimensões especificadas pela Figura 1: Educação, Saúde, Demográfica, Infraestrutura, Economia/ Finanças e Política.

Devido à indisponibilidade de dados para alguns dos municípios considerados, os dados foram coletados para três anos distintos, sendo eles os anos de 2000, 2003 e 2006. Como a análise multivariada apenas considera um cenário estático, ou seja, não considera a evolução dos dados numa série temporal, cada variável escolhida teve de ser coletada em apenas um desses anos. O critério de seleção do ano correspondente a cada variável é o de disponibilidade de dados para a maior quantidade de municípios.

Os dados foram tabulados no Excel e o software utilizado para a análise multivariada dos dados foi o STATISTICA v.10.

5. RESULTADOS

5.1 ANÁLISE DESCRITIVA DOS DADOS

A análise estatística descritiva dos dados tem sua importância no fato de mostrar ao pesquisador um panorama geral dos dados disponíveis para cada uma das variáveis selecionadas. Segundo Rodrigues e Paulo (2012, p.13), na análise descritiva de uma distribuição estatística é fundamental, além da determinação de uma medida de tendência central, conhecer a dispersão dos dados.

A Tabela 3 mostra as variáveis selecionadas para compor a base de dados, juntamente com suas respectivas estatísticas descritivas tanto de tendência central como de dispersão (média, valores mínimos e máximos assumidos e desvio-padrão).

Tabela 3 – Análise Descritiva das Variáveis Seleccionadas

Variáveis	Ano	Média	Mínimo	Máximo	Desv.Pad
Saneamento - Abastecimento de Água – Nível de Atendimento (Em %)	2000	96,26	80,9100	99,9	4,06
Saneamento - Coleta de Lixo – Nível de Atendimento (Em %)	2000	97,27	76,8100	100,0	4,24
Justiça e Segurança - Total de Ocorrências de Crimes	2006	1520,83	39,0000	25545,0	3204,30
Finanças Públicas Municipais - Total de Despesas Municipais per capita – Educação (Em reais de 2011)	2006	544,84	224,9239	1190,8	199,34
Finanças Públicas Municipais - Total de Despesas Municipais per capita – Saúde (Em reais de 2011)	2006	365,08	93,0786	760,0	119,98
Saneamento - Domicílios Urbanos Servidos por Rede Pública de Coleta de Esgotos (Em %)	2003	83,82	33,0000	100,0	15,14
População e Estatísticas Vitais - Índice de Envelhecimento (Em %)	2006	42,18	21,9500	69,5	10,11
Finanças Públicas Municipais - Arrecadação de ISS per capita (Em reais de 2011)	2006	77,76	2,9862	708,9	95,65
População e Estatísticas Vitais - Óbitos por Agressões	2006	5,37	0,0000	92,0	11,94
População e Estatísticas Vitais - População Urbana	2000	26051,87	668,0000	485520,0	59129,72
Finanças Públicas Municipais - Total da Receita Tributária Municipal per capita (Em reais de 2011)	2006	191,04	33,6999	852,4	148,38
Trabalho - Rendimento Médio no Total de Vínculos Empregatícios (Em reais correntes)	2006	843,35	531,6700	2337,5	265,34
Educação - Taxa de Analfabetismo da População de 15 Anos e Mais (Em %)	2000	11,30	4,6600	21,5	3,83
População e Estatísticas Vitais - Taxa de Mortalidade Infantil (Por mil nascidos vivos)	2006	19,26	5,7100	79,6	12,84
Trabalho - Total de Vínculos Empregatícios	2006	6868,22	145.0000	130707,0	16023,47
Finanças Públicas Municipais - Receita Municipal por Transferências Correntes per capita (Em reais de 2011)	2006	1551,96	828.3969	3681,2	571,79
Produto e Renda - Valor Adicionado Fiscal per capita – Total (Em reais de 2011)	2006	10014,50	727.1273	111553,6	12820,91
Produto e Renda - PIB (Em milhões de reais correntes)	2006	480,97	11.6700	10161,8	1230,48

Fonte: Elaboração Própria

Além da constatação do comportamento dos dados em cada variável, é importante também que se verifiquem as correlações entre cada uma das variáveis. Como a análise multivariada leva em consideração essas correlações, uma observação em seus valores pode ser fundamental para decidir prosseguir ou não com a metodologia. O Quadro 1 mostra a matriz de correlações entre as variáveis. Os valores destacados são os indicadores de correlação que se mostraram significativos durante o cálculo da matriz.

Quadro 1 – Matriz de Correlações entre as variáveis selecionadas

	Água	Lixo	Crimes	D. Educ	D. Saúde	Esgoto	Envelhec.	ISS	Óbitos	Pop. Urb.	Rec. Trib.	Rend. Méd.	Analfab.	Mortalid.	Empregos	Transfer	VAF	PIB
Abast. Água	1,00	0,43	0,11	-0,28	-0,07	0,38	0,22	0,00	0,01	0,14	0,00	0,03	-0,24	-0,11	0,11	-0,17	0,06	0,12
Coleta Lixo		1,00	0,11	0,00	-0,07	0,30	0,22	0,16	0,06	0,12	0,23	0,15	-0,51	-0,20	0,11	-0,16	0,13	0,12
Crimes			1,00	-0,30	-0,08	0,16	0,02	0,09	0,94	0,99	0,20	0,33	-0,44	-0,15	0,98	-0,34	0,11	0,96
Despesas Educação				1,00	0,57	-0,19	-0,34	0,58	-0,27	-	0,46	0,30	0,17	0,10	-0,27	0,77	0,39	-
Despesas Saúde					1,00	-0,02	-0,08	0,45	-0,09	-	0,35	0,18	0,22	0,27	-0,05	0,82	0,33	0,03
Coleta Esgoto						1,00	0,25	-	0,08	0,18	-0,14	-0,06	-0,12	-0,02	0,15	-0,02	-0,02	0,16
Índice Envelhecimento							1,00	-	0,01	0,03	0,06	-0,22	-0,28	-0,11	0,04	-0,14	-0,18	0,01
ISS								1,00	0,09	0,10	0,83	0,70	-0,34	-0,05	0,14	0,47	0,82	0,23
Óbitos por Agressões									1,00	0,94	0,22	0,35	-0,42	-0,13	0,94	-0,33	0,10	0,92
Pop. Urbana										1,00	0,21	0,33	-0,45	-0,14	0,99	-0,32	0,11	0,97
Receita Tributária											1,00	0,53	-0,51	-0,16	0,25	0,24	0,59	0,30
Rendimento Médio												1,00	-0,43	-0,05	0,36	0,21	0,80	0,43
Taxa Analfabetismo													1,00	0,37	-0,45	0,38	-0,33	-
Taxa Mortalidade														1,00	-0,15	0,37	-0,06	-
Vínculos Empregatícios															1,00	-0,29	0,16	0,98
Transferências Correntes																1,00	0,37	-
VAF																	1,00	0,26
PIB																		1,00

Fonte: Elaboração Própria

Como se pode perceber pelo Quadro 1, grande parte das variáveis mostrou correlação parcial acima de 0,5 com outras variáveis consideradas. Essas variáveis que apresentaram correlação elevada entre elas já indicam o modo de como alguns dos fatores irão se comportar. No entanto, algumas variáveis não apresentaram correlação elevada (positiva ou negativa) com outras. Isso pode influenciar a continuidade da análise, fazendo-se necessário, portanto, a aplicação de testes que verifiquem a viabilidade da análise fatorial com extração de componentes principais.

5.2 DISCUSSÃO DO MODELO

Primeiramente, antes da extração das componentes principais dos dados, é preciso determinar se a análise fatorial é adequada ou não para o caso a ser estudado. Isso pode ser feito através do teste de Kaisen – Meyer – Olkin (KMO) e do teste de esfericidade de Bartlett (BTS – *Bartlett Test of Sphericity*).

Anteriormente à extração das componentes principais dos dados, foram realizados os testes KMO e de Bartlett para verificar a adequação da análise fatorial ao caso a ser verificado. O Quadro 2 mostra os resultados dos testes.

Quadro 2 – Teste KMO e de Bartlett

KMO	Valor	0.753
Teste de Bartlett	Qui-Quadrado Aproximado	1598.428
	Sig.	.000

Fonte: Elaboração Própria

Como se pode perceber pelo Quadro 2, o teste KMO assumiu o valor de 0.753, indicando que o ajuste dos dados à análise fatorial é considerado bom. Segundo esse teste, portanto, a análise fatorial é adequada para os dados considerados.

No caso do teste de Bartlett, foi obtido um valor de 1598.428 e nível de significância de $p = 0.000$. Isso demonstra que foi rejeitada a hipótese de que a matriz

de correlação dos dados seja uma matriz identidade, ou seja, há correlação suficiente entre as variáveis para que seja viabilizada uma análise fatorial dos dados.

Uma vez que se determina viável a realização de uma análise fatorial dos dados, deve-se selecionar o número de componentes principais que devem agrupar todas as variáveis selecionadas, visando traduzir no menor número possível de componentes a maior parte da variação dos dados.

Segundo o Critério de Kaiser apud Vicini (2005), deve-se considerar na análise fatorial apenas as componentes cujos valores próprios (eigenvalues ou autovalores) sejam superiores a 1. Isso permite com que sejam mantidas no modelo apenas as combinações lineares capazes de explicar, no mínimo, a quantidade de variância de uma variável original padronizada (MINGOTI, 2005 apud ALVES et.al., 2008).

Cabe aqui destacar que alguns dos municípios da RAS apresentaram ausência de dados para a análise, entrando no modelo como valores faltantes (*missingvalues*). Segundo Rodrigues e Paulo (2012, p.37), é possível utilizar métodos de atribuição para substituir um dado faltante, tais quais a substituição pela média dos dados ou a atribuição por regressão. Neste estudo optou-se por escolher o método de substituição por um caso fora da amostra. Ou seja, para os dados faltantes para determinada variável optou-se por substituir esses dados por outros disponíveis para períodos posteriores. Isso permite com que ainda possam ser considerados os desempenhos de municípios com dados faltantes, apenas com a ressalva da diferença de datas consideradas.

Mesmo com o método de substituição de dados faltantes, cinco municípios não apresentaram dados também no período subsequente ao analisado, tendo sido desconsiderados na extração dos fatores e dos escores. São eles: Alambari, Barra do Chapéu, Itaóca, Jumirim e Ribeira.

A Tabela 4 mostra os autovalores e o percentual da variância explicada por cada componente principal.

Tabela 4 – Componentes principais para as variáveis selecionadas

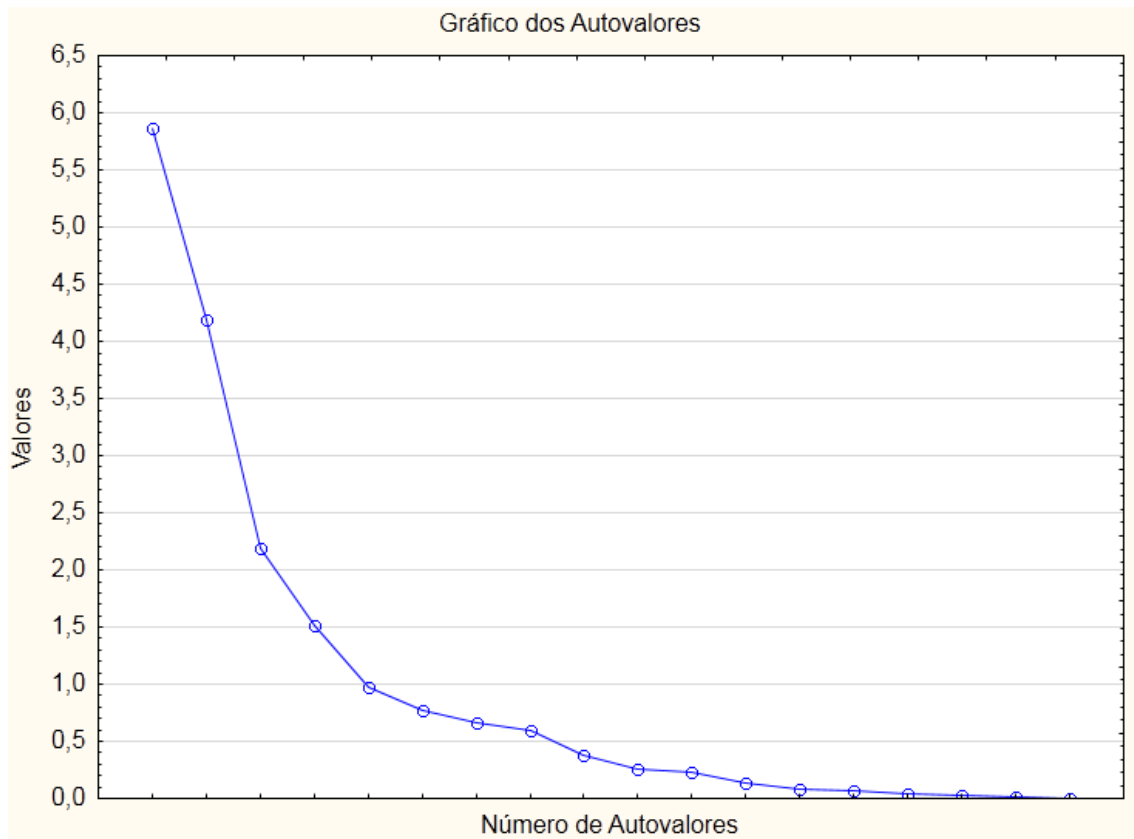
Número de Componentes Principais	Autovalores	% da variância explicada	Autovalores acumulados	% da variância explicada acumulada
1	5.860544	32.55858	5.86054	32.5586
2	4.196561	23.31423	10.05711	55.8728
3	2.190647	12.17026	12.24775	68.0431
4	1.515165	8.41759	13.76292	76.4607
5	0.977284	5.42936	14.74020	81.8900
6	0.766219	4.25677	15.50642	86.1468
7	0.663037	3.68354	16.16946	89.8303
8	0.594543	3.30302	16.76400	93.1333
9	0.374439	2.08022	17.13844	95.2136
10	0.263419	1.46344	17.40186	96.6770
11	0.224934	1.24964	17.62679	97.9266
12	0.140557	0.78087	17.76735	98.7075
13	0.078229	0.43461	17.84558	99.1421
14	0.064177	0.35654	17.90976	99.4986
15	0.046721	0.25956	17.95648	99.7582
16	0.026624	0.14791	17.98310	99.9061
17	0.009715	0.05397	17.99282	99.9601
18	0.007184	0.03991	18.00000	100.0000

Fonte: Elaboração própria

Como se pode verificar na Tabela 4, considerando todas as variáveis selecionadas puderam ser encontrados 18 fatores que explicam 100% da variância dos dados originais. Entretanto, seguindo o Critério de Kaiser, apenas os quatro primeiro fatores devem ser considerados. Esses quatro fatores conseguem explicar em torno de 76.46% da variância total dos dados, o que é um valor adequado segundo o critério citado por Vicini (2005) de que se deve considerar escolher os componentes que conseguem sintetizar uma variância acumulada em torno de 70%.

Outro método muito utilizado para a escolha do número de componentes a serem utilizados é o método gráfico ou scree plot. A Figura 9 mostra a seleção do número de componentes principais através do método scree plot.

Figura 9 – Análise Scree Plot dos Autovalores



Fonte: Elaboração Própria

Como é possível perceber pela Figura 9, através do método de seleção scree plot também devem ser utilizados quatro fatores na análise, embora a utilização de um quinto fator também fosse plausível em alguns casos. Isso é percebido quando se verifica que a partir do quarto fator a variância total dos dados é cada vez menos explicada e o gráfico de autovalores começa a se estabilizar. A partir do quinto fator, dessa forma, há muito pouca representatividade dos fatores quanto à variação dos dados, devendo ser optado pela utilização de apenas quatro ou cinco fatores.

A partir da escolha dos quatro primeiro fatores como representativos, é necessário que se efetue uma análise para identificar as variáveis que mais contribuem para cada um dos fatores. Essa contribuição pode ser verificada pelos carregamentos fatoriais (Factor Loadings) de cada variável, correspondentes à proporção de variação

da variável que é explicada pelo fator, ou também, o quanto cada variável colabora na formação de cada componente (SAAD, 2009).

Para melhor se avaliar as cargas fatoriais do modelo, o conjunto de dados do estudo foi submetido a uma rotação ortogonal dos eixos de referência (rotação *varimax*), mantendo os fatores independentes entre si e pesos fatoriais altos associados a algumas variáveis para um fator, mas baixo para outros. Segundo Almeida (2008) apud Klefens (2009) esse tipo de rotação facilita a leitura e define mais claramente os grupos de variáveis que fazem parte de um fator estudado. Klefens (2009) destaca que essa rotação mantém as propriedades estatísticas inalteradas, mesmo alterando a matriz de cargas fatoriais.

A Tabela 5 mostra as cargas fatoriais obtidas após a rotação *varimax* dos eixos dos dados, sendo destacadas as cargas que apresentaram valores superiores a 0.5.

Tabela 5 – Cargas fatoriais com fatores rotacionados

	Fator 1	Fator 2	Fator 3	Fator 4
Abastecimento de Água – Nível de Atendimento (Em %)	0.063870	0.009149	0.741900	-0.094639
Coleta de Lixo – Nível de Atendimento (Em %)	0.013642	0.306531	0.668423	-0.245999
Total de Ocorrências de Crimes	0.979603	0.045962	0.047300	-0.108092
Total de Despesas Municipais per capita – Educação	-0.294048	0.566749	-0.235802	0.529778
Total de Despesas Municipais per capita – Saúde	-0.018064	0.318114	0.031386	0.814978
Domicílios Servidos por Rede Pública de Coleta de Esgotos (Em %)	0.168605	-0.183011	0.718688	0.165537
Índice de Envelhecimento (Em %)	-0.012148	-0.127055	0.581968	-0.166599
Arrecadação de ISS per capita	0.076616	0.914954	-0.001147	0.238923
Óbitos por Agressões	0.954623	0.058276	-0.047323	-0.119657
População Urbana	0.983722	0.052700	0.075563	-0.086244
Total da Receita Tributária Municipal per capita	0.163128	0.841820	0.052783	0.044947
Rendimento Médio no Total de Vínculos Empregatícios	0.336054	0.763965	-0.068980	0.061029
Taxa de Analfabetismo da População de 15 Anos e Mais (Em %)	-0.357444	-0.515334	-0.355655	0.524137
Taxa de Mortalidade Infantil (Por mil nascidos vivos)	-0.041820	-0.248665	-0.066909	0.617845
Total de Vínculos Empregatícios	0.980490	0.095979	0.056113	-0.080442
Receita Municipal por Transferências Correntes per capita	-0.269114	0.333657	-0.058233	0.858612
Valor Adicionado Fiscal per capita – Total	0.109878	0.837152	0.011748	0.166833
PIB	0.969522	0.171107	0.062328	-0.007802

Fonte: Elaboração Própria

Como é possível verificar, através da rotação *varimax* dos eixos cada um dos fatores foi capaz de explicar bem a variação de um determinado conjunto de variáveis. Ao relacionar essas variáveis a um único fator, é possível agrupá-las e denominar cada um dos fatores em função do tipo de cada variável que o representa.

O Fator 1 se mostrou positivamente mais relacionado com as variáveis: Total de Ocorrências de Crimes, Óbitos por Agressões, População Urbana, Total de Vínculos Empregatícios e PIB. Esse fator, portanto, está associado a municípios maiores, mais urbanizados, com maiores índices de violência. Esse fator foi denominado Urbanização e Atividade Econômica.

O Fator 2 se mostrou positivamente mais relacionado com as variáveis: Despesas Municipais per capita com Educação, Arrecadação de ISS per capita, Total da Receita Tributária Municipal per capita, Rendimento Médio no Total de Vínculos Empregatícios, Valor Adicionado Fiscal per capita. Como a variável de Despesas Municipais per capita com Educação teve uma carga fatorial também elevada no Fator 4, optou-se por conveniência de análise agrupá-la nesse fator, ao invés do Fator 2. Dessa forma, pode-se inferir que o Fator 2 está relacionado ao nível de atividade dos setores econômicos dos municípios, refletindo também em suas arrecadações fiscais. Esse fator foi denominado Desempenho Setorial e Arrecadação Fiscal.

O Fator 3 se mostrou positivamente mais relacionado com as variáveis: Abastecimento de Água, Coleta de Lixo, Domicílios Servidos por Rede Pública de Coleta de Esgotos e Índice de Envelhecimento. Esse fator, portanto, está associado a variáveis que expressam as condições de vida dos habitantes dos municípios, refletindo a infraestrutura municipal vigente. Esse fator foi denominado Infraestrutura Urbana.

O Fator 4 se mostrou positivamente mais relacionado com as variáveis: Total de Despesas Municipais per capita com Educação, Total de Despesas Municipais per capita com Saúde, Taxa de Analfabetismo, Taxa de Mortalidade Infantil e Receita Municipal por Transferências Correntes per capita. Esse fator, portanto, está associado a variáveis que expressam os gastos dos municípios com questões sociais e os resultados/produtos desses gastos para seus níveis de educação e saúde. Esse fator foi denominado Condições e Desempenho Sociais.

Dadas as disposições das variáveis em relação aos fatores definidos, pode-se definir um método de classificação dos municípios baseado no desempenho de cada município em cada fator. A classificação dos municípios, portanto, pode ser feita através dos escores fatoriais, que fornecem a estrutura de correlação e a influência dos fatores sobre as observações.

Segundo Haddad (1989), os escores fatoriais encontrados numa análise das componentes principais devem ser apenas interpretados numa escala ordinal, apenas indicando a posição relativa dos municípios hierarquizados. Dessa forma, municípios com escores fatoriais muito elevados próximos a municípios com escores fatoriais menores apenas indicam que há superioridade entre um e outro município, não o valor dessa superioridade.

No ANEXO 1 se encontram os valores dos escores obtidos para cada município em cada fator.

Existem diversos modos de se ordenar e categorizar os escores fatoriais encontrados para os municípios em cada fator. Neste trabalho optou-se pelo simples método de divisão dos escores em 5 categorias através de seus percentis, delimitando grupos de municípios em cada fator nas categorias: Muito Baixo, Baixo, Médio, Alto e Muito Alto. A Tabela 6 mostra as delimitações de cargas fatoriais correspondentes a cada categoria para cada um dos fatores encontrados.

Tabela 6 – Divisão dos Escores Fatoriais em categorias de desempenho

	Fator 1	Fator 2	Fator 3	Fator 4
Muito Alto	7.5210 até 0.0945	5.9589 até 0.4848	1.5571 até 0.7958	2.9030 até 0.9489
Alto	0.0944 até -0.1698	0.4847 até -0.0233	0.7957 até 0.3771	0.9488 até -0.0290
Médio	-0.1697 até -0.2889	-0.0232 até -0.2846	0.3770 até -0.0165	-0.0289 até -0.3715
Baixo	-0.2888 até -0.4028	-0.2845 até -0.5880	-0.0164 até -0.7725	-0.3714 até -0.8209
Muito Baixo	-0.4827 até -0.7235	-0.5879 até -1.7616	-0.7724 até -3.5738	-0.8208 até -1.5712

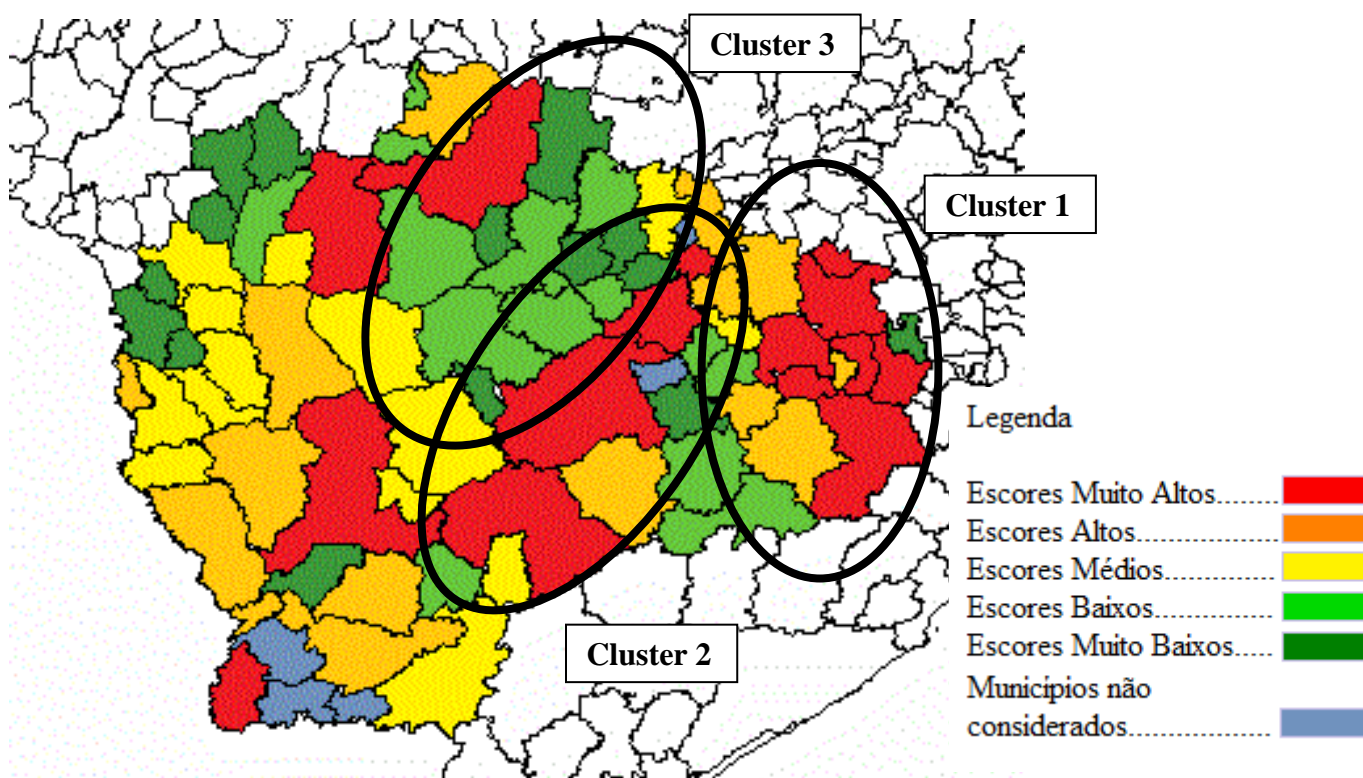
Fonte: Elaboração Própria

A partir dessa divisão em grupos, pode-se visualizar em um mapa da RAS o comportamento de municípios próximos geograficamente em relação a cada fator, havendo possibilidade de também se identificar visualmente padrões regionais (*clusters*) de desempenho em cada um dos fatores. No ANEXO 2 está ilustrada a divisão

territorial dos municípios da RAS, onde podem ser acompanhadas as análises realizadas para cada município que for citado.

A Figura 10 mostra os municípios da RAS subdivididos nas categorias propostas para o Fator Urbanização e Atividade Econômica.

Figura 10 – Desempenho dos Municípios da RAS em Urbanização e Atividade Econômica



Fonte: Elaboração Própria

Como se pode verificar através da Figura 10, existem alguns padrões regionais que podem ser identificados quanto ao fator de Urbanização e Atividade Econômica dos municípios. Primeiramente, pode-se identificar um *cluster* de municípios com escores altos e muito altos próximos ao município de Sorocaba, denominado *Cluster 1*. Os municípios de Sorocaba, Votorantim, Salto de Pirapora, Piedade, Ibiúna, Alumínio, Mairinque, São Roque, Itu, Salto e Porto Feliz se encontram nesse *cluster*. Além disso, pode-se verificar a presença de outro *cluster* de escores altos e muito altos próximos ao

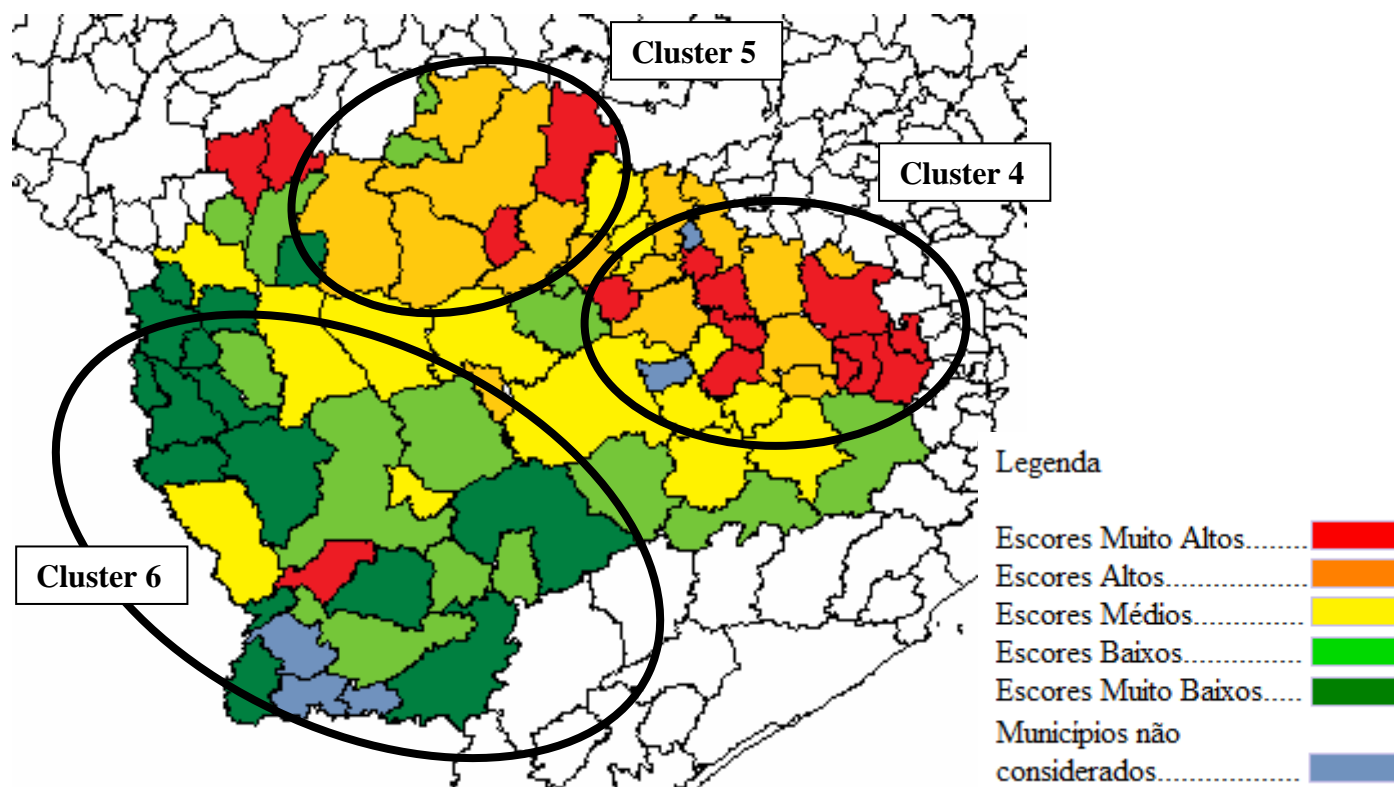
município de Itapetininga, composto por Tatuí, Capão Bonito, Cerquilha, São Miguel Arcanjo, Boituva e Tietê, denominado *Cluster 2*.

Tanto o *Cluster 1* quanto o *Cluster 2* mostram que os municípios próximos de Sorocaba e Itapetininga têm seu nível geral de urbanização e de atividade econômica bastante acima da média dos outros municípios da RAS. Nesse caso, pode-se inferir que o desempenho desses municípios quanto às questões econômicas é relativamente elevado, influenciando positivamente, inclusive, no nível de empregos gerados. Porém esse alto desempenho relativo implica em aumento no grau de urbanização municipal, que por sua vez gera problemas comumente urbanos como o aumento da violência local.

Deve-se destacar também o *Cluster 3*, de escores baixos e muito baixos encontrado entre os municípios de Guareí, Quadras, Cesário Lange, Porangaba, Pereiras, Conchas, Anhambí e Bofete. Esses municípios, embora próximos geograficamente de outros como Itapetininga e Tatuí, apresentaram desempenho bastante abaixo da média quanto a sua Urbanização e Atividade Econômica. A pequena estrutura urbana nesses municípios, além de afetar seu desempenho em relação às variáveis econômicas, também contribui para que os índices de violência sejam baixos, o que ajuda a explicar os valores muito baixos dos escores nessa localidade.

A Figura 11 mostra os municípios da RAS subdivididos nas categorias propostas para o Fator Desempenho Setorial e Arrecadação Fiscal.

Figura 11 - Desempenho dos Municípios da RAS em Desempenho Setorial e Arrecadação Fiscal



Fonte: Elaboração Própria

Como se pode verificar na Figura 11, novamente foi encontrado um *cluster* de escores altos e muito altos entre Sorocaba e os municípios adjacentes, denominado *Cluster 4*. Entre os municípios encontrados nesse *cluster* estão a própria Sorocaba, Votorantim, Alumínio, Mairinque, São Roque, Araçariguama, Araçoiaba da Serra, Iperó, Boituva, Porto Feliz, Itu e Salto. O alto desempenho desses municípios no fator indica que os mesmos apresentaram, no geral, valores de rendimentos médios acima da média regional dado seu alto índice relativo de atividade econômica (descrito pelo Fator 1). O desempenho relativamente elevado entre os setores básicos da economia desses municípios também pode ter sido responsável por um aumento relativo do nível de tributação municipal, implicando também em elevadas receitas fiscais.

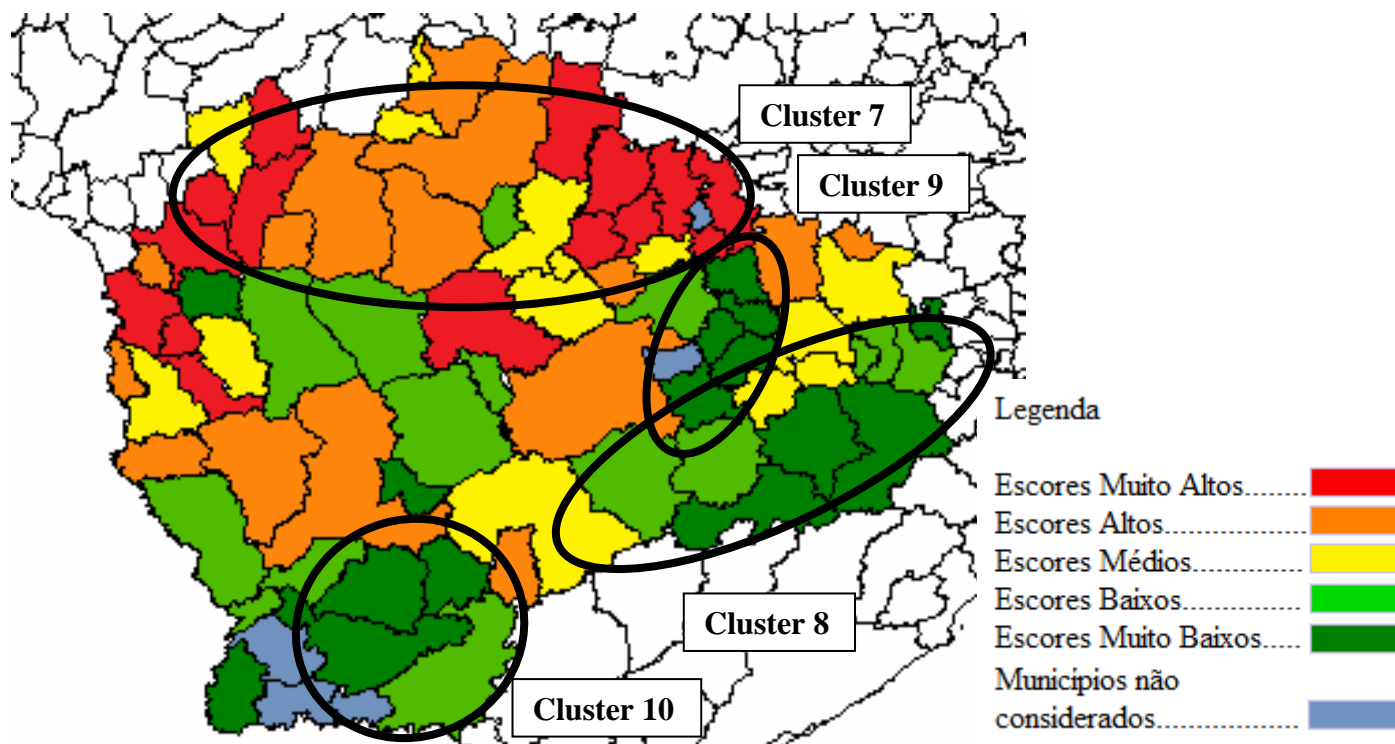
Um *Cluster 5*, de municípios com escores altos e muito altos também pode ser notado entre os municípios de Anhembi, Bofete, Pardinho, Itatinga, Botucatu, São

Manuel e Avaré. Embora o nível de urbanização desses municípios não seja tão expressivo, como é detalhado pelo Fator 1, os mesmos apresentaram desempenhos setoriais e de arrecadação fiscal relativamente altos principalmente em função dos dados de financiamento público, per capita, acima da média. A principal inferência que se pode atingir com esse *cluster* é que seus municípios possuem um reduzido número de habitantes e uma elevada arrecadação fiscal, indicando que esta deve ser revertida em benefícios estruturais ou sociais para os municípios. Caso esses benefícios não sejam verificados em seus desempenhos nos próximos fatores, os municípios do *Cluster 5* podem estar gerando receitas muito superiores a seus gastos, havendo espaço para uma melhor alocação de recursos públicos.

A presença de um *Cluster 6*, de valores baixos e muito baixos, também deve ser destacada em diversos dos municípios da porção sudoeste da RAS. Entre eles se encontram: Itaberá, Riversul, Itaporanga, Coronel Macedo, Barão de Antonina, Taguaí, Fartura, Tejupá, Sarutaiá, Buri, Itapeva, Capão Bonito, Ribeirão Grande, Guapiara, Ribeirão Branco, Apiaí, Iporanga, Itapirapuã Paulista e Bonsucesso do Itararé. Quase a totalidade dos municípios considerados nesse *cluster* é formada por municípios pouco urbanizados e com uma produção agrícola ainda percentualmente muito participativa em sua economia. Isso pode ser o motivo dos mesmos assumirem escores baixos tanto quanto ao desempenho de setores mais dinâmicos de sua economia quanto de sua arrecadação fiscal, sendo que esse impacto assumiu proporções regionais ao englobar diversos municípios adjacentes.

A Figura 12 mostra os municípios da RAS subdivididos nas categorias propostas para o Fator Infraestrutura Urbana.

Figura 12 - Desempenho dos Municípios da RAS em Infraestrutura Urbana



Fonte: Elaboração Própria

Pode-se verificar na Figura 12 que, diferentemente do que ocorre nos outros dois fatores, Sorocaba e os municípios próximos (Votorantim, Salto de Pirapora e Itu) apresentaram apenas escores médios de desempenho na categoria Infraestrutura Urbana. Os escores altos e muito altos para esse fator praticamente se concentraram no porção norte da RAS. Essa concentração de escores foi formada entre os municípios de Tietê, Laranjal Paulista, Pereiras, Porangaba, Conchas, Anhembí, Botucatu, Itatinga, São Manuel, Avaré, Arandu, Cerqueira César, Iaras e Manduri, formando o *Cluster 7* nessa localidade. Esse comportamento se deve a duas causas principais.

A primeira delas é que Sorocaba e os municípios adjacentes possuem um grau de urbanização bastante elevado em comparação aos outros municípios da RAS. Portanto, mesmo possuindo percentuais de água encanada, coleta de lixo e esgoto elevados, municípios menos urbanizados, como os do *Cluster 7*, alcançam os maiores valores percentuais por possuírem um número menor de domicílios a serem atendidos pelo sistema de saneamento básico municipal.

Outra possível explicação é a ocorrência de valores elevados para os índices de envelhecimento dos municípios que obtiveram escores altos e muito altos. Pelo fato de alguns desses municípios serem relativamente menos urbanizados e apresentarem uma baixa taxa de natalidade, o percentual de idosos em relação a jovens se torna bastante elevado. Dessa forma, em algumas situações esse índice de envelhecimento não diz respeito à qualidade da saúde e do saneamento básico municipal que propiciaram um maior percentual de idosos, mas sim a um componente cultural e indiretamente urbanístico dos municípios.

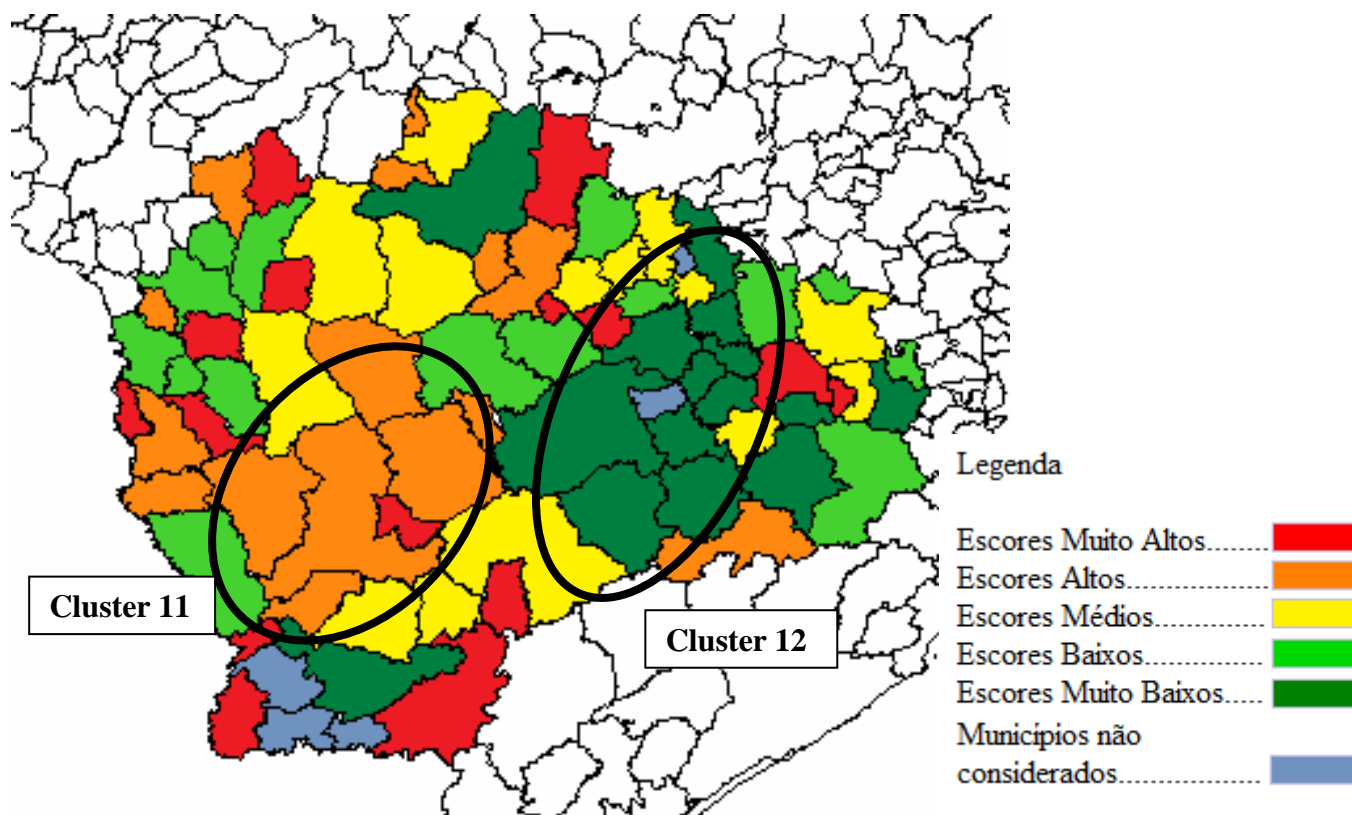
O fato de que alguns dos municípios do *Cluster 7* terem mostrado desempenho alto ou muito alto no fator que abrange a Arrecadação Fiscal, além dos municípios citados no *Cluster 6* da Figura 11, também deve ser considerado na explicação desse *cluster*. Municípios com maiores arrecadações per capita tem maior capacidade de investir em Infraestrutura Urbana, o que deve implicar em melhores índices de saneamento básico e melhorias nas condições de vida da população.

Ainda na Figura 12, três *clusters* de escores muito baixos nesse fator podem ser encontrados entre os municípios da RAS. O *Cluster 8* tem como centro os municípios de Tapiraí, Piedade e Ibiúna e pode influenciar negativamente os municípios de Pilar do Sul, São Miguel Arcanjo, Alumínio, Mairinque, São Roque e Araçariguama. O *Cluster 9* é composto por escores muito baixos dos municípios de Sarapuí, Araçoiaba da Serra, Iperó e Boituva, podendo influenciar negativamente também Tatuí. O *Cluster 10* envolve os municípios de Apiaí, Bonsucesso do Itararé, Ribeirão Branco e Guapiara, podendo também influenciar negativamente Iporanga.

É interessante observar que tanto o *Cluster 8* quanto o *Cluster 9* estão formados nos arredores dos municípios mais urbanizados e mais influentes economicamente de Sorocaba e Itapetininga. A explicação desses *clusters* de baixos desempenhos em Infraestrutura Urbana, portanto, pode ser o fato dos habitantes dos municípios menores não dependerem tanto da infraestrutura disponibilizada por seus próprios municípios, mas sim dependerem indiretamente da infraestrutura disponibilizada pelos centros de dinamismo mais próximos.

A Figura 13 mostra os municípios da RAS subdivididos nas categorias propostas para o Fator Condições e Desempenho Sociais.

Figura 13 - Desempenho dos Municípios da RAS em Condições e Desempenho Sociais



Fonte: Elaboração Própria

Como se pode observar na Figura 13, foram selecionados os dois *clusters* mais evidentes na RAS para o fator Condições e Desempenho Sociais. O *Cluster 11* é composto por municípios com escores altos e muito altos e o *Cluster 12* por municípios com escores muito baixos. A análise desse fator requer maior cuidado, uma vez que o mesmo sofre uma limitação no modelo. Essa limitação se dá pelo fato de que nele estão combinadas tanto variáveis que possuem relação direta com o desenvolvimento municipal, quanto variáveis que possuem relação inversa com o desenvolvimento municipal. Dessa forma, municípios com, por exemplo, taxas médias mortalidade elevadas podem estar no mesmo patamar de desempenho do fator que outros que apresentem, por exemplo, despesas per capita com saúde elevadas.

Os municípios pertencentes ao *Cluster 11* de escores altos e muito altos são os municípios de Buri, Campina do Monte Alegre, Coronel Macedo, Itaberá, Itapeva, Itaporanga, Nova Campina, Paranapanema, Riversul e Taquarivaí. O motivo desses

municípios formarem um *cluster* nesse fator pode ser decorrente ou do fato de apresentarem uma combinação de valores acima da média para os dados de despesas com saúde e educação e receitas por transferências, ou do fato de apresentarem combinação de valores acima da média para os dados de taxa de mortalidade infantil e taxa de analfabetismo. Fazendo uma análise em cada um dos casos, pode-se perceber que esses municípios, no geral, possuem valores acima da média tanto no que se refere às despesas com saúde e educação, quanto no que se refere aos indicadores de analfabetismo e mortalidade infantil. Não se pode inferir por esse fator, portanto, que tais municípios são eficientes na transformação de seus gastos em melhorias sociais, apenas pode-se dizer que eles possuem um conjunto acima da média tanto de gastos per capita (positivamente relacionados ao desenvolvimento) quanto de indicadores de saúde (negativamente associados ao desenvolvimento).

Os municípios pertencentes ao *Cluster* 11 de escores baixos e muito baixos são os municípios de Itapetininga, São Miguel Arcanjo, Pilar do Sul, Piedade, Sarapuí, Araçoiaba da Serra, Capela do Alto, Iperó, Tatuí, Boituva, Votorantim e Tietê. É possível perceber que esses municípios, ao serem analisados individualmente, apresentam geralmente valores abaixo da média para os gastos com saúde e educação e receitas por transferência, mas também apresentam valores abaixo da média para os índices de saúde e educação. Nesse caso, pode-se perceber nitidamente a limitação deste fator na captação das reais condições e desempenhos sociais dos municípios, uma vez que os municípios que conseguem manter seus índices de analfabetismo e mortalidade infantil baixos, com poucos gastos per capita, deveria ter seu desempenho considerado alto.

Cabe destacar que, embora não sejam aqui considerados um *cluster*, os municípios de Sorocaba e Alumínio assumiram escores muito altos nesse fator. O fato de Sorocaba apresentar elevado desempenho quanto às condições sociais atesta a validade dos escores obtidos para esse fator, mesmo com a limitação anteriormente referida. Sorocaba apresenta valores abaixo da média para os indicadores de analfabetismo e mortalidade infantil. Alumínio, por sua vez, possui valores acima da média para as receitas por transferências e despesas com saúde e educação.

Finalmente, para poder se classificar os municípios da RAS quanto ao seu desenvolvimento, deve-se contemplar o desempenho dos municípios analisados em

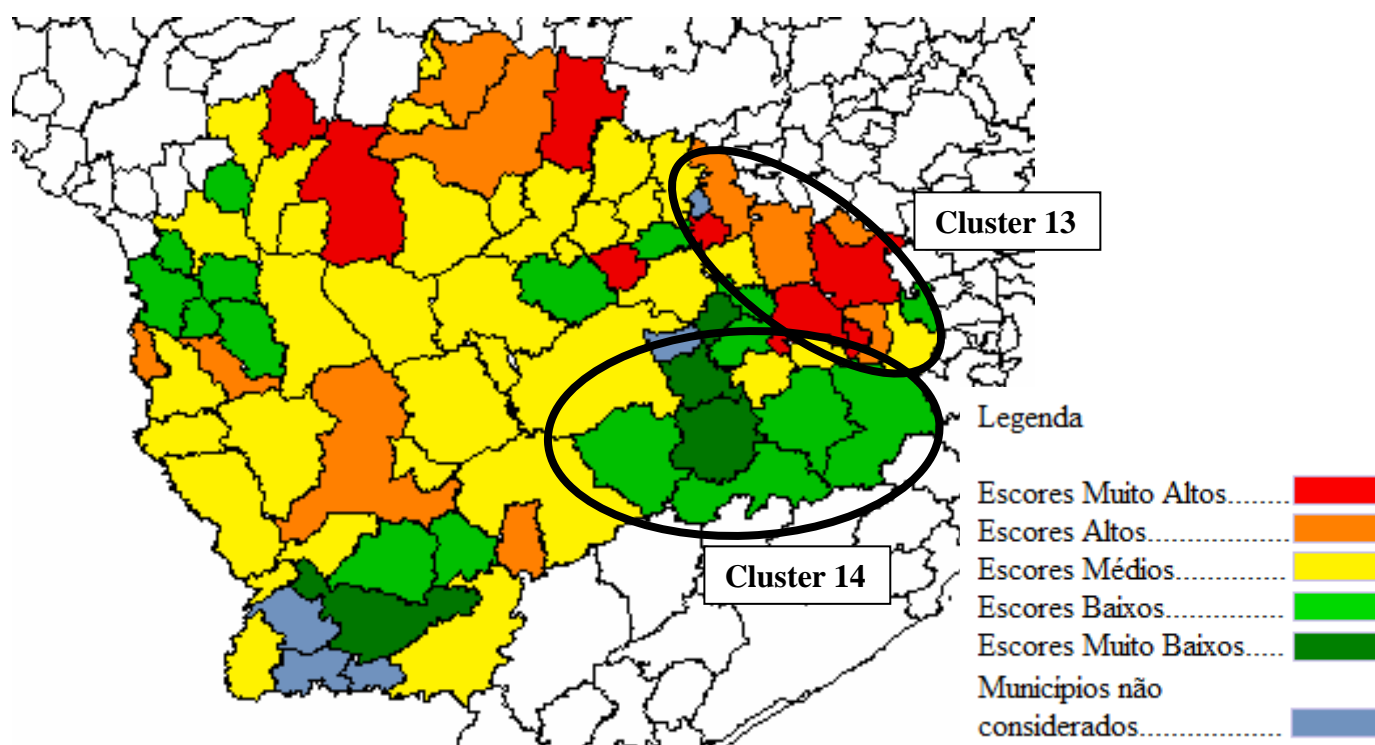
cada um dos fatores. Embora na prática os fatores possuam importâncias diferentes, tanto para os municípios como para seus habitantes, para unir todos os fatores será considerado que todos os fatores terão pesos iguais para o desenvolvimento. Isso implica que, por exemplo, municípios com escores muito baixos nos fatores 1 e 2 possam ser compensados por escores muito altos nos fatores 3 e 4 ou vice-versa.

Portanto, municípios com escores muito altos em algum dos fatores serão atribuídos de valor 2, municípios com escores altos de valor 1, municípios de escores médios de valor 0, municípios de escores baixos de valor -1 e municípios de escores muito baixos de valor -2.

Com a união de todos os fatores, através dos valores acima estabelecidos, os municípios de valor final acima de 3 são considerados de desempenho muito alto, os de valor final entre 2 e 3 são considerados de desempenho alto, os de valor final entre -1 e 1 são considerados de desempenho médio, os de valor final entre -2 e -3 são considerados de desempenho baixo e os de valor final abaixo de -3 são considerados de desempenho muito baixo.

A Figura 14 mostra os municípios da RAS subdivididos em categorias de desenvolvimento municipal, considerando todos os quatro fatores encontrados na análise fatorial.

Figura 14 – Desempenho dos Municípios da RAS em Desenvolvimento Municipal



Fonte: Elaboração Própria

Assim como o esperado, esse método de classificação do desenvolvimento municipal, pela atribuição de valores comparativos a cada um dos escores fatoriais encontrados, permitiu com que houvesse uma compensação entre altos e baixos desempenhos nos fatores para grande parte dos municípios. A vantagem desse método é que quando determinados municípios assumem valores altos ou baixos para o desenvolvimento, mesmo após a compensação entre fatores, pode-se corroborar o entendimento de que esse nível de desenvolvimento é realmente factível.

Dois *clusters* podem ser destacados na Figura 14 quanto ao desempenho municipal.

O *Cluster* 13, associado a desempenhos altos e muito altos, se encontra entre os municípios de Sorocaba, Alumínio, Mairinque, Itu, Salto, Porto Feliz, Tietê e Cerquilha. É possível perceber nesse *cluster* que a presença de municípios de destaque

regional, como Sorocaba, Itu e Alumínio, fez com que os municípios próximos recebessem influência direta de seus desenvolvimentos e se desenvolvessem conjuntamente. Todos os municípios desse *cluster* são bastante urbanizados, possuem uma atividade econômica que privilegia os setores industriais e de serviços e possuem destaque em seu desempenho com arrecadação fiscal, mesmo que alguns deles não possuam desempenho acima da média com infraestrutura urbana ou com condições sociais.

O *Cluster* 14, associado a desempenhos baixos e muito baixos, se encontra entre os municípios de Iperó, Araçoiaba da Serra, Capela do Alto, Sarapuí, Pilar do Sul, São Miguel Arcanjo, Tapiraí, Piedade e Ibiúna. Esses municípios possuem, em geral, níveis de urbanização e de atividade econômica relativamente menores do que seus municípios adjacentes, principalmente os mais próximos a Sorocaba. Esse fato pode ter refletido diretamente no baixo desempenho desses municípios também quanto à sua Infraestrutura Urbana e Condições Sociais.

6. CONSIDERAÇÕES FINAIS

Em suma, observou-se ao longo deste trabalho que a aplicação da Análise Multivariada na Região Administrativa de Sorocaba é um método simples, porém de muita utilidade para o estudo abrangente das características municipais e para a classificação de municípios quanto ao seu desenvolvimento, levando em consideração diversas variáveis simultaneamente.

Verificou-se através da aplicação do método que, dentre todas as variáveis selecionadas, algumas mostraram níveis de correlação mais significativos com outras, podendo ser agrupadas em fatores que sintetizam o comportamento dos dados e auxiliam da compreensão do desempenho de cada município em relação a eles.

As variáveis de Total de Ocorrências de Crimes, Óbitos por Agressões, População Urbana, Total de Vínculos Empregatícios e PIB, puderam ser sintetizadas em um fator denominado Urbanização e Atividade Econômica. Nesse fator observou-se um padrão de concentração regional, no qual assumem valores concentrados altos e muito altos os municípios de: Sorocaba, Votorantim, Salto de Pirapora, Piedade, Ibiúna, Alumínio, Mairinque, São Roque, Itu, Salto, Porto Feliz, Itapetininga, Tatuí, Capão

Bonito, Cerquilha, São Miguel Arcanjo, Boituva e Tietê. Esses padrões indicam, portanto que a porção leste da RAS é mais urbanizada e possui um nível de atividade econômica relativamente maior do que a média de toda a região. Padrões de valores baixos e muito baixos de desempenho nesse fator foram encontrados para os municípios de: Guareí, Quadras, Cesário Lange, Porangaba, Pereiras, Conchas, Anhambí e Bofete.

As variáveis de Arrecadação de ISS per capita, Total da Receita Tributária Municipal per capita, Rendimento Médio no Total de Vínculos Empregatícios e Valor Adicionado Fiscal per capita, puderam ser sintetizadas em um fator denominado Desempenho Setorial e Arrecadação Fiscal. Nesse fator observou-se um padrão de concentração regional, no qual assumem valores concentrados altos e muito altos os municípios de: Sorocaba, Votorantim, Alumínio, Mairinque, São Roque, Araçariguama, Araçoiaba da Serra, Iperó, Boituva, Porto Feliz, Itu e Salto. Outros municípios com esse mesmo padrão são os de: Anhambí, Bofete, Pardinho, Itatinga, Botucatu, São Manuel e Avaré. Padrões de valores baixos e muito baixos de desempenho nesse fator foram encontrados para os municípios de: Itaberá, Riversul, Itaporanga, Coronel Macedo, Barão de Antonina, Taguaí, Fartura, Tejuapá e Sarutaiá.

As variáveis de Abastecimento de Água, Coleta de Lixo, Domicílios Servidos por Rede Pública de Coleta de Esgotos e Índice de Envelhecimento, puderam ser sintetizadas em um fator denominado Infraestrutura Urbana. Nesse fator observou-se um padrão de concentração regional, no qual assumem valores concentrados altos e muito altos os municípios de: Tietê, Laranjal Paulista, Pereiras, Porangaba, Conchas, Anhambí, Botucatu, Itatinga, São Manuel, Avaré, Arandu, Cerqueira César, Iaras e Manduri. Padrões de valores baixos e muito baixos de desempenho nesse fator foram encontrados para os municípios de: Tapiraí, Piedade, Ibiúna, Pilar do Sul, São Miguel Arcanjo, Alumínio, Mairinque, São Roque e Araçariguama. Outros municípios com esse mesmo padrão são os de: Sarapuí, Araçoiaba da Serra, Iperó, Boituva e Tatuí. Outro *cluster* também pode ser identificado entre os municípios de: Apiaí, Bonsucesso do Itararé, Ribeirão Branco, Guapiara e Iporanga.

As variáveis de Total de Despesas Municipais per capita com Educação, Total de Despesas Municipais per capita com Saúde, Taxa de Analfabetismo, Taxa de Mortalidade Infantil e Receita Municipal por Transferências Correntes per capita, puderam ser sintetizadas em um fator denominado Condições e Desempenho Sociais. Nesse fator observou-se um padrão de concentração regional, no qual assumem valores concentrados altos e muito altos os municípios de: Buri, Campina do Monte Alegre,

Coronel Macedo, Itaberá, Itapeva, Itaporanga, Nova Campina, Paranapanema, Riversul e Taquarivaí. Padrões de valores baixos e muito baixos de desempenho nesse fator foram encontrados para os municípios de: Itapetininga, São Miguel Arcanjo, Pilar do Sul, Piedade, Sarapuí, Araçoiaba da Serra, Capela do Alto, Iperó, Tatuí, Boituva, Votorantim e Tietê.

7. REFERÊNCIAS BIBLIOGRÁFICAS

- ALMEIDA, T. L. Departamento de Matemática. Arquivos de aulas. FURG. Disponível em: <<http://www.dmat.furg.br/ taba/posanafat.htm>>. Acesso em: junho 2012
- ALVES, M.R.F; CAVENAGHI, S.M e ALVES, J. E.D. Avaliação do desempenho de municípios utilizando técnicas de Análise Multivariada. Rio de Janeiro, 2008.
- BEZERRA, F. A. Análise Fatorial. In: CORRAR, L. J. ; PAULO, E. ; DIAS FILHO, J. M. Análise multivariada para os cursos de administração, ciências contábeis e economia. Ed. Atlas. 1ª ed. 4ª reimpr. 2012.
- BOUROCHE, J. M.; SAPORTA, G. Análise de dados. Rio de Janeiro: Zahar, 1982.
- BUARQUE, Sérgio C. Metodologia de planejamento do desenvolvimento local e municipal sustentável. Brasília, 1998.
- CASTELS, Manuel; e BORJA, Jordi. “As cidades como atores políticos”. *Novos estudos CEBRAP – Dossiê Cidades* (mimeo.), no 45, julho de 1996.
- CORRAR, L. J. ; PAULO, E. ; DIAS FILHO, J. M. Análise multivariada para os cursos de administração, ciências contábeis e economia. Ed. Atlas. 1ª ed. 4ª reimpr. 2012.
- HADDAD, Paulo Roberto. (org.) Economia Regional: teoria e métodos de análise. Fortaleza: BNB, 1989. 694p.
- HAIR JR, J.F.; ANDERSON, R.E.; TATHAM, R.L.; BLACK, W.C. Análise multivariada de dados. 5ª ed. Reimpr. 2007. Porto Alegre: Bookman, 2005.
- KLEFENS, P.C.O. O Biplot na análise fatorial multivariada. 2009. 102 p. Escola Superior de Agricultura “Luiz de Queiroz” Universidade de São Paulo. Dissertação de Mestrado. Disponível em <<http://www.teses.usp.br>> Acesso em: maio de 2012.
- LEMOS, Mauro B. et al. A dinâmica da competitividade das regiões metropolitanas brasileiras. In: ENCONTRO NACIONAL DE ECONOMIA, 29, 2001, Salvador. Anais... Salvador: ANPEC, 2001. (Disponível em CD-Rom).

MINGOTI, S. A. Análise de dados através de métodos de estatística multivariada: uma abordagem aplicada. Belo Horizonte: Editora UFMG. 2005. 295 p.

PEREIRA, J. C. R. Análise de dados qualitativos: estratégias metodológicas para as ciências da saúde, humanas e sociais. São Paulo: Edusp, 2001.

RENCHER, A.C., Methods of multivariate analysis. John Wiley & Sons, 1995.

RODRIGUES, A.; PAULO, E. Introdução à Análise Multivariada. In: CORRAR, L. J. ; PAULO, E. ; DIAS FILHO, J. M. Análise multivariada para os cursos de administração, ciências contábeis e economia. Ed. Atlas. 1ª ed. 4ª reimpr. 2012.

SAAD, D. Aplicação de técnicas estatísticas multivariadas em dados de cerâmica vermelha produzida na Região Central do Rio Grande do Sul, 2009. Dissertação de mestrado, pg 55. Acesso em: novembro de 2012

SEADE – Sistema Estadual de Análise de Dados. Disponível em <www.seade.gov.br>. Acesso em:

SILVA, H.; RESENDE, A.; ROSA, C.; SIMÕES, R. 2005. Dinâmica Agropecuária e urbanização: uma análise multivariada para Minas Gerais, 1995-2000. In: ENCONTRO NACIONAL DE ECONOMIA, XXXII, Natal, 2005. Anais... Natal, p. 1-20.

SILVEIRA, B. C.; SILVA, R. G.; CARVALHO, L. A. Índice relativo de qualidade de vida na região norte: Uma aplicação da análise fatorial. Revista Brasileira de Gestão e Desenvolvimento Regional. v. 4. Taubaté, 2008.

TOLOSA, H. C. Dimensões e Causas da Pobreza Urbana. *Estudos Econômicos* vol. 7, nº 1 jan./abr. 1977.

VICINI, L. Análise Multivariada: da Teoria à Prática. Santa Maria: UFSM, CCNE, Caderno Didático, 215p., 2005.

ANEXO 1 – Escores fatoriais obtidos para cada município

	Fator 1	Fator 2	Fator 3	Fator 4
3500550 - Águas de Santa Bárbara	-0,723513	2,06732	0,16814	-0,00897
3503109 - Arandu	-0,173408	-0,79501	0,51982	1,41473
3504503 - Avaré	0,524235	0,06759	0,60876	-0,36990
3505005 - Barão de Antonina	-0,164894	-1,00263	0,41866	1,98603
3511409 - Cerqueira César	-0,304530	-0,38483	1,16113	-0,59534
3512605 - Coronel Macedo	-0,250509	-1,22117	1,01959	1,02921
3515400 - Fartura	-0,430603	-0,60450	1,15172	-0,71108
3519253 - Iaras	-0,447485	0,57756	0,85520	1,16932
3521804 - Itaí	-0,143635	-0,12728	-0,25266	-0,26607
3522802 - Itaporanga	-0,205211	-1,18014	0,30868	0,09021
3528601 - Manduri	-0,457754	-0,39512	1,27125	-0,67666
3535804 - Paranapanema	-0,267365	-0,03100	-0,13881	0,58542
3538808 - Piraju	-0,241518	-0,18423	1,22128	-0,76904
3551207 - Sarutaiá	-0,403664	-0,78344	0,69007	0,72708
3553005 - Taguaí	-0,442949	-0,65854	0,91712	-0,69352
3553807 - Taquarituba	-0,262421	-0,47511	0,29160	-0,57438
3554201 - Tejuapá	-0,234715	-1,03682	-1,08823	0,99496
3502309 - Anhembi	-0,530399	0,86167	0,99635	1,10748
3503604 - Areiópolis	-0,394417	-0,43011	0,31184	0,21924
3506904 - Bofete	-0,400641	0,11902	0,15527	-0,02469
3507506 - Botucatu	1,369317	0,27904	0,56239	-1,23123
3512308 - Conchas	-0,293133	-0,19828	1,05969	-0,38167
3523503 - Itatinga	-0,383698	0,43687	0,44709	-0,32334
3526407 - Laranjal Paulista	-0,214962	0,18194	1,02251	-0,35165
3536109 - Pardinho	-0,627833	1,10490	-0,21792	0,46460
3537503 - Pereiras	-0,429349	-0,28452	1,26295	-0,22391
3540507 - Porangaba	-0,551042	0,17441	1,55716	-0,25007
3541059 - Pratânia	-0,358077	-0,36941	0,32210	0,78401
3550100 - São Manuel	0,045970	0,00130	0,65404	-0,29734
3554656 - Torre de Pedra	-0,394941	-0,42924	0,30709	1,63176
3502200 - Angatuba	-0,298274	-0,21222	0,92106	-0,43108
3507001 - Boituva	0,086756	1,04273	-1,47038	-1,01708
3509452 - Campina do Monte Alegre	-0,479984	0,03758	-0,15510	0,11837
3510302 - Capela do Alto	-0,362535	-0,20920	-0,79259	-0,91447
3511508 - Cerquilha	0,160000	0,51464	1,02652	-0,33740
3511607 - Cesário Lange	-0,474745	0,14701	0,04552	-0,80657
3518503 - Guareí	-0,402271	-0,53861	0,15902	-0,80819
3522307 - Itapetininga	1,147794	-0,22834	0,38731	-0,88574
3541653 - Quadra	-0,340755	0,87387	0,74073	1,95587
3550209 - São Miguel Arcanjo	-0,146735	-0,44220	-0,19365	-0,87954
3551108 - Sarapuí	-0,481748	-0,26649	-1,09808	-1,33487
3554003 - Tatuí	0,887038	0,33803	-0,33128	-1,48542
3502705 - Apiaí	-0,166874	-0,44219	-1,04993	-1,06356

	Fator 1	Fator 2	Fator 3	Fator 4
3507159 - Bom Sucesso de Itararé	-0,113113	-0,62327	-0,30091	2,26726
3508009 - Buri	-0,253419	-0,32256	-0,16747	0,33658
3510203 - Capão Bonito	0,106231	-0,60435	0,15015	-0,36192
3517604 - Guapiara	-0,309993	-0,57718	-1,43102	-0,07981
3521200 - Iporanga	-0,244961	-0,69571	-0,04751	2,90304
3521705 - Itaberá	-0,098708	-0,75171	0,57880	0,07110
3522406 - Itapeva	0,623023	-0,28467	0,38558	0,44649
3522653 - Itapirapuã Paulista	0,212959	-1,76167	-3,57389	1,94747
3523206 - Itararé	0,007098	-0,05657	-0,03208	-0,66128
3532827 - Nova Campina	-0,473613	1,10636	-0,75912	0,21904
3543006 - Ribeirão Branco	-0,123245	-1,15605	-2,19302	-0,04671
3543253 - Ribeirão Grande	-0,254280	-0,32095	0,61824	1,47676
3543501 - Riversul	-0,272143	-1,40236	0,78110	0,92347
3553856- Taquarivaí	-0,211909	-0,17848	-2,53916	1,00570
3501152 - Alumínio	-0,162089	5,95896	-0,14586	2,30247
3502754 - Araçariguama	-0,506397	2,54093	-1,55219	-0,37194
3502903 - Araçoiaba da Serra	-0,331977	0,62136	-1,74108	-1,23579
3519709 - Ibiúna	0,620159	-0,32773	-2,36369	-0,58167
3521002 - Iperó	-0,201911	0,56905	-0,98826	-0,84013
3523909 - Itu	1,787089	0,55094	0,34320	-0,19984
3528403 - Mairinque	0,164327	0,78354	-0,52968	-0,32500
3537800 - Piedade	-0,168908	-0,11556	-0,97724	-1,57128
3537909 - Pilar do Sul	-0,359978	-0,24594	-0,19118	-1,08854
3540606 - Porto Feliz	0,070455	0,14628	0,62091	-0,69984
3545308 - Salto de Pirapora	0,013119	-0,04136	0,25718	-0,15232
3545209 - Salto	0,967288	0,48281	0,56175	-0,44084
3550605 - São Roque	0,376336	0,48789	-0,46327	-0,92688
3552205 - Sorocaba	7,521065	-0,02145	0,21344	0,98708
3553500 - Tapiraí	-0,303793	-0,33874	-1,40073	0,23962
3554508 - Tietê	-0,025750	0,40977	0,81794	-1,08123
3557006 - Votorantim	0,612518	0,27360	0,31612	-1,02657

ANEXO 2 – Divisão Territorial dos Municípios da Região
Administrativa de Sorocaba

