



Assembléia Legislativa de São Paulo

# Fórum São Paulo Século 21

Índice Paulista de  
Responsabilidade Social - IPRS

# Anexo Metodológico



Assembleia Legislativa de São Paulo

# Fórum São Paulo Século 21

Índice Paulista de  
Responsabilidade Social - IPRS  
**Anexo Metodológico**

## **Anexo Metodológico**

Este anexo metodológico tem por finalidade descrever os procedimentos de análise estatística que foram adotados pela Fundação Seade com vistas à classificação dos municípios do Estado de São Paulo em grupos com características distintas de riqueza, longevidade e escolaridade, mensuradas por indicadores sintéticos das principais variáveis disponíveis sobre cada um desses aspectos. Tomando por base esses indicadores sintéticos, cujos valores situam-se entre 0 e 100 e estão referidos aos anos-base de 1992 e 1997, o Fórum São Paulo Século XXI e a Assembléia Legislativa do Estado de São Paulo passam a dispor de um critério objetivo para a identificação dos municípios que mais avançaram na promoção da qualidade de vida de seus habitantes.

Embora tenham sido validados empiricamente apenas para o Estado de São Paulo, os indicadores sintéticos aqui apresentados podem ser calculados para outras unidades da Federação, pois baseiam-se, fundamentalmente, em dados estatísticos secundários, provenientes, em sua quase totalidade, de registros administrativos de atualização anual. Podem também ser utilizados em comparações internacionais, uma vez que guardam correspondência direta com o Índice de Desenvolvimento Humano – IDH e com outras medidas correlatas, divulgadas periodicamente pelo Programa das Nações Unidas para o Desenvolvimento – PNUD, das quais se diferenciam por dois aspectos fundamentais: maior desagregação territorial; e maior sensibilidade a mudanças sociais de curto prazo.

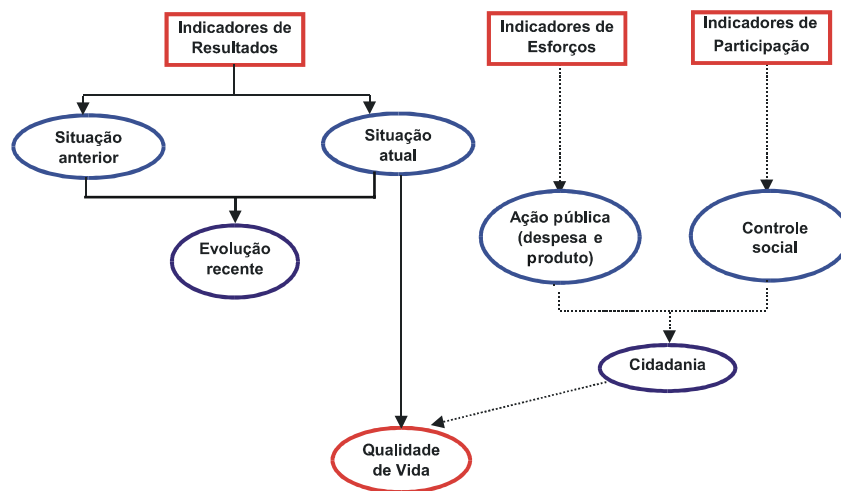
### **CONSIDERAÇÕES PRELIMINARES**

Para subsidiar os trabalhos do Fórum São Paulo Século XXI, a Fundação Seade, valendo-se de sua larga experiência no desenvolvimento de metodologias para o resumo de dados estatísticos secundários, provenientes do censo

demográfico e dos registros administrativos, propôs a criação de indicadores sintéticos para a classificação dos municípios do Estado de São Paulo segundo a qualidade de vida de seus habitantes. Amadurecida após uma longa e profunda discussão acerca das vantagens e das limitações do Índice de Desenvolvimento Humano – IDH, divulgado anualmente pelo Programa das Nações Unidas para o Desenvolvimento – PNUD, tal proposta evoluiu para a organização de um sistema de referência estatístico que, conforme pode ser visto no Quadro 1, deveria subsidiar, além da avaliação dos resultados e dos esforços empreendidos pelo poder público em prol do desenvolvimento em âmbito local, também a mensuração do grau de participação e controle da sociedade civil sobre tais ações.

**QUADRO 1**  
Síntese da metodologia proposta pela Fundação Seade

**CONCEPÇÃO**



Definido o objetivo geral do projeto, passou-se à consideração dos problemas operacionais – disponibilidade de informações para todos os municípios do Estado de São Paulo, necessidade de seleção de variáveis referentes a dois momentos distintos e garantia de atualização periódica das fontes de dados utilizadas, que

dificultavam a realização de estudo tão abrangente antes do final dos trabalhos do Fórum São Paulo Século XXI. Considerando que a solução definitiva de tais problemas deveria envolver, em ação conjunta, Legislativo, governo e representantes da sociedade civil organizada, decidiu-se realizar um estudo exploratório, a ser publicado ainda no ano 2000.

Teve início, então, o processo de levantamento das fontes de dados e seleção das variáveis capazes de representar satisfatoriamente, do ponto de vista tanto teórico como empírico, as dimensões (riqueza, longevidade e escolaridade) e os aspectos (resultados, esforços e participação) a serem mensurados. O Quadro 2 exhibe a relação das variáveis selecionadas ao final desse processo, que compreendeu, além de consultas aos parlamentares e especialistas que integravam as Comissões Temáticas do Fórum São Paulo Século XXI, a análise cuidadosa dos diagnósticos de situação publicados nos Cadernos do Fórum.

Embora fosse propósito da Fundação Seade considerar tanto os resultados como os esforços em prol do desenvolvimento em âmbito municipal, apenas três variáveis contemplam o último aspecto, por estarem sob a governabilidade do poder público local: o valor adicionado fiscal per capita, que pode ser alterado favoravelmente a partir da implementação de políticas de atração de investimento; a taxa de mortalidade perinatal, cuja redução pressupõe a melhoria da qualidade do sistema de atendimento médico-hospitalar; e a participação da rede municipal na oferta de vagas para o ensino fundamental, cujo aumento, em período recente, foi estimulado pela política do governo estadual para o setor. Outra dimensão importante, relativa ao controle social sobre a ação governamental, deixou de ser considerada com o aprofundamento desejável devido à baixa correlação entre as variáveis relativas aos conselhos e comissões municipais de emprego, de saúde e de educação e os indicadores de resultados em cada um desses aspectos. Motivos para a exclusão de outras variáveis pré-selecionadas são apontados, de forma sumária, no Quadro 3.

**Quadro 2**  
**Variáveis pré-selecionadas para a definição operacional dos indicadores sintéticos de riqueza, longevidade e escolaridade.**

Dimensões	Indicadores de Resultados	Indicadores de Esforço	Indicadores de Participação
Longevidade	Esperança de vida ao nascer Mortalidade infantil Mortalidade perinatal Mortalidade entre a população com mais de 60 anos Mortalidade entre a população com 15 a 39 anos Proporção de nascidos vivos de mães com 15 a 19 anos Óbitos por causas maldefinidas Nascidos vivos de baixo peso ao nascer Nascidos vivos de mães com 15 a 19 anos	Despesa per capita em programas de saúde e saneamento Mães que realizaram mais de seis consultas de atendimento pré-natal Redução do número de partos cesáreos Médicos por habitante Dentistas por habitante Enfermeiros por habitante Consultas médicas por habitante Abastecimento de água domiciliar Esgotamento sanitário Coleta de lixo	Existência do Conselho Municipal de Saúde Data de implantação do Conselho Municipal de Saúde Composição do Conselho Municipal de Saúde Periodicidade das reuniões do Conselho Municipal de Saúde Conferências municipais de saúde realizadas pela Prefeitura
Educação	Analfabetismo entre a população com 15 a 19 anos Frequência à creche Frequência à pré-escola Alunos do ensino fundamental com defasagem idade-série Conclusão do ensino fundamental para a população com 15 a 19 anos Frequência ao ensino médio para a população com 15 a 19 anos	Despesa per capita em programas de educação e cultura Participação da rede municipal nas matrículas em educação infantil Participação da rede municipal nas matrículas do ensino fundamental	Existência de Conselho Municipal de Educação Data de implantação do Conselho Municipal de Educação
Renda	Rendimento médio do emprego formal Consumo de energia elétrica residencial urbana Consumo de energia elétrica rural	Valor adicionado fiscal per capita	Municipal de Emprego Existência de Comissão

### QUADRO 3

#### Variáveis pré-selecionadas que deixaram de ser incorporadas à análise

Consumo anual de energia elétrica rural:

Esta variável foi substituída por uma medida global do consumo de energia elétrica no comércio, em serviços e na agricultura, que complementa a informação relativa ao consumo de energia elétrica residencial, também incorporada ao indicador sintético de riqueza municipal.

Terminais telefônicos por habitante:

Esta variável foi excluída devido à privatização das antigas empresas estatais de telefonia, o que alterou radicalmente o equilíbrio econômico entre oferta e procura, dificultando a comparação das séries históricas existentes para o setor.

Existência da Comissão Municipal de Emprego :

Fornecida pela Secretaria do Emprego e Relações do Trabalho, esta variável foi excluída por estar disponível apenas para o ano-base de 1999.

Esperança de vida ao nascer:

Componente do IDH, esta variável foi excluída por sua baixa sensibilidade a variações no curto prazo. Não obstante, serviu como base para a definição operacional do indicador sintético de longevidade, uma vez que todas as variáveis que o compuseram mostraram-se altamente correlacionadas com a esperança de vida ao nascer.

Proporção de óbitos por causas maldefinidas

Proporção de nascidos vivos com baixo peso

Proporção de nascidos vivos de mães com idade entre 15 a 19 anos:

Pré-selecionadas como indicadores do grau de cobertura do sistema de saúde, estas variáveis mostraram-se pouco correlacionadas com a esperança de vida ao nascer. Ademais, os diferentes modelos de análise fatorial que levaram em conta essas proporções apresentaram níveis de explicação inferiores ao modelo final, que considera apenas variáveis de mortalidade passíveis de controle por meio de ações de saúde pública implementadas em âmbito local.

Despesa per capita com programas de saúde e saneamento realizada pelas

Prefeituras Municipais:

Concluiu-se que não seria conveniente a utilização das informações referentes às despesas realizadas no programa de saneamento, tendo em vista que estas não são comparáveis para o conjunto dos municípios, pois apenas 44% dos sistemas municipais de saneamento são operados pelas prefeituras, enquanto os restantes são de responsabilidade da Sabesp, o que implica um diferencial significativo no gasto em saneamento realizado pelos municípios. Quanto aos dados de despesas em saúde, mudança na metodologia da PMU 97, visando aprimorar a captação desses dados, impede uma comparação imediata com anos anteriores, exigindo um período de tempo maior para a compatibilização dos dados de 1995.

Abastecimento de água

Esgotamento sanitário

Coleta de lixo:

Coletados pelo Censo Demográfico de 1991, os dados sobre cobertura dos serviços de saneamento básico estão desatualizados e não há registros administrativos comparáveis para os anos subseqüentes, devido a diferenças na metodologia de coleta de dados que é utilizada pela Sabesp e pelas empresas municipais.

Médicos, dentistas e enfermeiros por habitante registrados nos Conselhos Regionais:

Esses dados, apesar de demonstrarem a disponibilidade profissional e o potencial de cobertura para assistência à população, são insuficientes para o dimensionamento desses recursos humanos por município, uma vez que o local do registro não necessariamente corresponde à localidade de trabalho desses profissionais. Assim, a ausência de registros de profissionais em alguns municípios não significa, necessariamente, a inexistência da oferta desse recurso aos respectivos municípios.

Médicos, dentistas e enfermeiros por habitante registrados nas Prefeituras Municipais:

Os testes realizados com estes dados demonstraram a necessidade de um estudo diferenciado por porte de município, considerando-se a participação da rede municipal – e portanto do contingente de recursos humanos da Prefeitura – no sistema de saúde de cada localidade. Nos municípios menores, os recursos humanos em saúde das Prefeituras podem responder pela quase totalidade daqueles existentes na cidade, percentual que diminui conforme aumenta o tamanho dos municípios, onde há maior presença de serviços privados contratados do Sistema Único de Saúde – SUS e de outros serviços não vinculados ao sistema, como as medicinas de grupo.

Consultas médicas por habitante:

As informações atualmente disponíveis sobre produção de serviços de saúde abrangem exclusivamente aqueles produzidos no âmbito do SUS. Esse indicador dimensiona as consultas médicas realizadas segundo a localização dos serviços. Dessa forma, pessoas não residentes podem também demandar consultas, especialmente nos municípios que concentram serviços especializados de atendimento regional, o que produz distorções no indicador por município. Dificuldade adicional é encontrada na análise dos dados para os municípios mais populosos, em que há menor cobertura do SUS e, em contrapartida, maior presença das medicinas de grupo.

Mães que realizaram mais de seis consultas de atendimento pré-natal:

Pré-selecionada como indicador de esforço, esta variável foi excluída por estar disponível apenas para o ano-base de 1997.

Redução do número de partos cesáreos:

Como a alta incidência de partos cesáreos é um problema específico do Brasil, a inclusão dessa variável como componente do indicador sintético de longevidade dificultaria comparações internacionais.

Existência do Conselho Municipal de Saúde

Data de implantação do Conselho Municipal de Saúde

Composição do Conselho Municipal de Saúde

Periodicidade das reuniões do Conselho

Municipal de Saúde

Conferências Municipais de Saúde realizadas

pela Prefeitura:

Foi construído um indicador de participação social em saúde através da combinação destas variáveis. Apesar de discriminar grupos distintos de municípios, não apresentou relação com a esperança de vida ao nascer, variável que serviu como base para a definição operacional do indicador sintético de longevidade.

Frequência à creche:

A PMU coletou, em 1995, informações sobre o número de crianças de 0 a 3 e de 4 a 6 anos e, em 1997, de crianças de 0 a 6 anos, em idade simples, matriculadas nas instituições municipais e conveniadas de educação infantil. Esses dados, apesar da possibilidade de ocorrência de dupla contagem de alunos que estão matriculados na creche em um período e na pré-escola em outro e da existência de cadastro incompleto de instituições de educação infantil nas prefeituras municipais, são uma proxy possível de frequência à creche que não pôde ser testada no período de realização do projeto. O MEC passou a coletar no Censo Escolar, a partir 1997, dados de matrícula em creche. A equipe técnica da Secretaria de Estado da Educação – SEE, ao avaliar esses dados, detectou os mesmos problemas encontrados nas informações da PMU, ou seja, dupla contagem e existência de cadastro incompleto das instituições de educação infantil. O próprio MEC, reconhecendo as limitações dos dados de educação infantil, elaborou novo questionário e novo cadastro e está aplicando, até o final de novembro, o Censo da Educação Infantil, para reverter esses problemas.

Frequência à pré-escola:

Dado levantado pela Contagem Populacional de 1996, é de difícil comparação com informações de matrícula, tanto da PMU quanto do Censo Escolar, por ser captado em setembro, quando o aluno que foi matriculado, e contado pelo Censo Escolar, no mês de março pode já ter abandonado a escola. A PMU capta o dado de alunos matriculados em novembro, exigindo também testes que verifiquem a viabilidade da comparação entre essas variáveis. Posteriormente, uma alternativa a ser testada seria a substituição dessa variável pelo dado de matrícula na pré-escola, coletado pelo Censo Escolar e disponível para os anos solicitados.



Alunos do ensino fundamental com defasagem idade-série:

Dado coletado, até 1995, pela Secretaria de Estado da Educação–SEE e, a partir de 1996, pelo Censo Escolar do MEC e disponibilizado pela SEE, foi excluído devido à mudança do instrumento de coleta que, antes de 1996, perguntava a idade do aluno e, a partir desta data, o ano de nascimento, diferença que, segundo a SEE, impossibilita a comparação entre dois anos-base considerados.

Frequência ao ensino médio para a população de 15 a 19 anos:

Levantada pelo Censo Demográfico de 1991 e pela Contagem Populacional de 1996, esta variável foi substituída pelo indicador de conclusão do ensino médio para a população de 20 a 24 anos, que é uma proxy ainda mais adequada do grau de escolaridade da população.

Analfabetismo entre a população de 15 a 19 anos:

Também levantada pelo Censo Demográfico de 1991 e pela Contagem Populacional de 1996, esta variável recebeu tratamento diferenciado após análise estatística exploratória, sendo substituída por dois indicadores de instrução da população: a proporção de pessoas de 10 a 14 anos com mais de um ano de estudo e a proporção daquelas de 15 a 24 anos em igual situação.

Despesa per capita com programas de educação e cultura:

Indicador do nível de esforço do poder público em prol do desenvolvimento local, esta variável não foi utilizada pelo fato de ser ainda necessária à realização de testes estatísticos que avaliem seu poder discriminatório e sua correlação com as demais informações referentes à educação utilizadas para este trabalho.

Existência do Conselho Municipal de Educação

Data de implantação do Conselho Municipal de Educação:

Indicadores que exigem tempo maior para a realização dos testes que avaliem seu poder discriminatório e sua correlação com as demais informações referentes à educação utilizadas para este trabalho.

## **ASPECTOS OPERACIONAIS**

A metodologia de resumo de dados adotada pela Fundação Seade apóia-se em técnicas estatísticas de análise multivariada, que possibilitam reduzir o conjunto de medidas utilizadas para a quantificação de determinado fenômeno ou processo social.

A aplicação dessa metodologia, no âmbito deste estudo, visou a redução das variáveis utilizadas para a definição operacional de um conceito de desenvolvimento comparável àquele mensurado pelo Índice de Desenvolvimento Humano – IDH. Este indicador sintético considera, além do Produto Interno Bruto – PIB, dois outros aspectos ou dimensões igualmente relevantes para a mensuração do nível de desenvolvimento alcançado por uma nação: a longevidade da população, indicada pela esperança de vida ao nascer; e seu grau de escolaridade, medido por um

índice que combina a taxa de analfabetismo da população adulta com a taxa global de matrículas em escolas de níveis primário, secundário e terciário.

A reprodução pura e simples da metodologia de cálculo do IDH, tal como sugerida por participantes do Fórum São Paulo Século XXI, não foi adotada em razão de dois obstáculos fundamentais: a inexistência de informações atualizadas sobre a renda per capita dos municípios paulistas e a pouca sensibilidade deste indicador a mudanças sociais de curto prazo.

O primeiro obstáculo foi superado de forma satisfatória, pela utilização de variáveis *proxies* da renda per capita do Censo Demográfico de 1991. O último revelou-se de difícil transposição, pois implicou a seleção de variáveis bem diferentes daquelas originalmente adotadas pelo PNUD, como pode ser visto no Quadro 4.

#### QUADRO 4

Variáveis utilizadas para a mensuração do nível de desenvolvimento alcançado pelos municípios do Estado de São Paulo, segundo as mesmas dimensões, originalmente, consideradas pelo IDH

Riqueza
Consumo anual de energia elétrica no comércio, na agricultura e em serviços
Consumo anual de energia elétrica residencial
Rendimento médio do emprego formal
Valor adicionado fiscal per capita
Longevidade
Taxa de mortalidade infantil
Taxa de mortalidade perinatal
Taxa de mortalidade da população de 15 a 39 anos
Taxa de mortalidade da população de mais de 60 anos
Escolaridade
Porcentagem de pessoas de 10 a 14 anos com mais de um ano completo de estudo
Porcentagem de pessoas de 15 a 24 anos com mais de um ano completo de estudo
Porcentagem de pessoas de 15 a 19 anos que concluíram o ensino fundamental
Porcentagem de pessoas de 20 a 24 anos que concluíram o ensino médio
Porcentagem de matrículas no ensino fundamental da rede municipal no total de matrículas da rede pública

Apresenta-se, a seguir, a definição operacional das variáveis selecionadas para a composição dos indicadores sintéticos de riqueza, longevidade e escolaridade utilizados pela Fundação Seade.

- Consumo anual de energia elétrica no comércio, na agricultura e em serviços – razão entre o consumo anual de energia elétrica e o total de consumidores destes ramos de atividade. Entende-se por consumidores no comércio e serviços as unidades em que são desenvolvidas atividades comerciais ou de prestação de serviços (excluídos os serviços públicos de água, esgoto, saneamento, tração elétrica urbana e/ou ferroviária). Consumidores na agricultura englobam: unidades que desenvolvem exploração econômica da agricultura e/ou pecuária, incluídas as residências ali situadas; cooperativas de eletrificação rural; indústrias rurais situadas fora do perímetro urbano e que desenvolvem atividades de transformação e/ou beneficiamento de produtos da agricultura e/ou pecuária, com capacidade em transformadores não superior a 75 KVA; coletividades rurais; serviços públicos de irrigação; escolas agropecuárias e seu respectivo consumo.
- Consumo anual de energia elétrica residencial – razão entre o consumo anual de energia elétrica residencial e o total de consumidores residenciais. Consumidores residenciais são unidades residenciais urbanas, incluídas as instalações de uso comum de prédio ou conjunto em que predomine este tipo de unidade.
- Rendimento médio do emprego formal – razão entre a massa salarial do mês de dezembro e o número de vínculos empregatícios com contrato formal de trabalho naquele mês. Por vínculo empregatício, entende-se o número de postos de trabalho do setor formal.
- Valor adicionado fiscal per capita – razão entre o total anual do valor adicionado fiscal do município e a sua população total. O valor adicionado corresponde, para cada município, ao valor das saídas de mercadorias, acrescido do valor das prestações de serviços no seu território, deduzido o valor das entradas de mercadorias em cada ano civil. O valor adicionado é utilizado, pela Secretaria da Fazenda, como um dos critérios para a definição do Índice de Participação dos municípios no produto de arrecadação do ICMS.
- Taxa de mortalidade infantil – razão entre o total de óbitos de menores de 1 ano ocorridos no ano e o total de nascidos vivos no ano, multiplicado por mil.

- Taxa de mortalidade perinatal – razão entre o total de óbitos de menores de 6 dias e natimortos, ocorridos no ano, e o total de nascidos vivos mais o total de natimortos no ano, multiplicado por mil.
- Taxa de mortalidade da população de 15 a 39 anos – razão entre o total de óbitos entre os indivíduos de 15 a 39 anos ocorridos no ano e o total de pessoas nesta faixa etária na população, multiplicado por mil.
- Taxa de mortalidade da população de mais de 60 anos – razão entre o total de óbitos entre os indivíduos com mais de 60 anos ocorridos no ano e o total de pessoas com esta idade na população, multiplicado por mil.
- Porcentagem de pessoas de 10 a 14 anos com mais de um ano completo de estudo – razão entre o número de indivíduos de 10 a 14 anos com mais de uma série concluída e o total de pessoas nesta faixa etária.
- Porcentagem de pessoas de 15 a 24 anos com mais de um ano completo de estudo – razão entre o número de indivíduos de 15 a 24 anos com mais de uma série concluída e o total de pessoas nesta faixa etária.
- Porcentagem de pessoas de 15 a 19 anos que concluíram o ensino fundamental – razão entre o número de indivíduos de 15 a 19 que concluíram o ensino fundamental e o total de pessoas nesta faixa etária.
- Porcentagem de pessoas de 20 a 24 anos que concluíram o ensino médio – razão entre o número de indivíduos de 20 a 24 que concluíram o ensino médio e o total de pessoas nesta faixa etária.
- Porcentagem de matrículas no ensino fundamental da rede municipal no total de matrículas da rede pública – razão entre o total de matrículas iniciais na rede municipal e o total de matrículas na rede pública (estadual e municipal).

### **TRATAMENTO ESTATÍSTICO DOS DADOS**

O nível de desagregação municipal e a necessidade de se obter informações para todos os municípios do Estado de São Paulo demandaram alguns tratamentos estatísticos para as variáveis de cada dimensão de análise, apresentados a seguir.

## RIQUEZA

- Imputação para 1997 dos valores do rendimento médio do emprego formal para os 20 municípios instalados em 1997, devido ao fato de o Cadastro Rais – Relação Anual de Informações Sociais – não disponibilizar informações para estas localidades. Os valores imputados foram aqueles referentes aos municípios de origem.
- Transformação logarítmica nos dados referentes a *consumo anual de energia elétrica no comércio, agricultura e em serviços por ligação, rendimento médio do emprego formal e valor adicionado fiscal per capita*, com o objetivo de corrigir a forte assimetria das distribuições destas variáveis, minimizando assim a influência de observações “aberrantes” (muito grandes ou muito pequenas), que poderiam comprometer os resultados finais.
- Padronização das variáveis originais e, quando necessário, das variáveis transformadas na escala de 0 a 100, de modo a permitir a comparação entre os diferentes anos-base considerados e simplificar a interpretação dos dados. Para tanto, utilizou-se a seguinte padronização:  $(\text{variável} - \text{mínimo}) / (\text{máximo} - \text{mínimo})$ , em que os valores mínimo e máximo referem-se ao conjunto de observações para os dois anos-base. Tais valores encontram-se no Apêndice Estatístico.

## LONGEVIDADE

- Para o cálculo das taxas de mortalidade, que pode sofrer distorções em consequência de tamanhos populacionais, utilizou-se, para os municípios com mais de 8.000 habitantes, o total de óbitos ocorridos no período de três anos e, para aqueles com menos de 8.000 habitantes, os óbitos no período de 7 anos.
- As taxas de mortalidade para as faixas etárias de 15 a 39 anos e 60 anos e mais foram calculadas a partir das médias dos óbitos ocorridos nos períodos 1993-95 e 1997-99.

- Para as taxas de mortalidade infantil e perinatal, o procedimento acima foi utilizado para os municípios com mais de 8.000 habitantes. Para os demais, foram considerados os nascimentos e os óbitos ocorridos no período 1993-99. Portanto, para os pequenos municípios, os valores foram repetidos nos dois períodos.
- As populações aqui consideradas resultam de um modelo de projeção demográfica baseado nos resultados do Censo Demográfico de 1991 e da Contagem Populacional de 1996 (Fundação IBGE) e nos índices de crescimento calculados a partir das Estatísticas Vitais processadas na Fundação Seade. Estas estimativas referem-se a 1º de julho do ano em questão.
- Padronização das variáveis originais e, quando necessário, das variáveis transformadas na escala de 0 a 100, de modo a permitir a comparação entre os diferentes anos-base considerados e simplificar a interpretação dos dados. Para tanto, utilizou-se a seguinte padronização:  $(\text{variável} - \text{mínimo}) / (\text{máximo} - \text{mínimo})$ , em que os valores mínimo e máximo referem-se ao conjunto de observações para os dois anos-base. Tais valores encontram-se no Apêndice Estatístico.

## ESCOLARIDADE

- Imputação de valores para os 20 municípios instalados em 1997. Os valores imputados são referentes aos municípios de origem. Esta imputação decorreu da necessidade de serem obtidas informações para todos os municípios existentes em 1997.
- Padronização das variáveis originais e, quando necessário, das variáveis transformadas na escala de 0 a 100, de modo a permitir a comparação entre os diferentes anos-base considerados e simplificar a interpretação dos dados. Para tanto, utilizou-se a seguinte padronização:  $(\text{variável} - \text{mínimo}) / (\text{máximo} - \text{mínimo})$ , em que os valores mínimo e máximo referem-se ao conjunto de

observações para os dois anos-base. Tais valores encontram-se no Apêndice Estatístico.

### **USO DA ANÁLISE DE REGRESSÃO MÚLTIPLA PARA A SELEÇÃO DE VARIÁVEIS PROXIES DA RENDA FAMILIAR PER CAPITA E DA ESPERANÇA DE VIDA AO NASCER**

A necessidade de se obter um indicador de renda que pudesse ser atualizado para períodos de tempo inferiores ao do Censo Demográfico levou à exploração de registros administrativos, que fossem correlacionados com a renda familiar per capita derivada do Censo Demográfico de 1991. O Quadro 5 apresenta as variáveis pré-selecionadas, sendo que todas são referentes a 1991.

**QUADRO 5**  
Variáveis pré-selecionadas como *proxies* da renda familiar per capita

Variável	Fonte	Unidade
Rendimento médio do emprego formal	Ministério do Trabalho e Emprego – RAIS	Salários mínimos de dezembro de 1991
Valor adicionado per capita	Secretaria da Fazenda	Em 1.000 Cr\$ de 1991
Consumo anual de energia elétrica residencial por ligação	Companhia Energética de São Paulo	Em MGwh por ligação
Consumo anual de energia elétrica na agricultura por ligação	Companhia Energética de São Paulo	Em MGwh por ligação
Grau de urbanização	Censo Demográfico – IBGE	Em porcentagem
Número de terminais telefônicos	Telecomunicações de São Paulo S.A.	Nº de terminais por 100 habitantes

Um dos pré-requisitos exigidos para a seleção de variáveis era a manutenção de sua associação com a renda familiar per capita ao longo do tempo. Por isso, foram excluídos os terminais telefônicos, pois, provavelmente, a correlação com a

renda familiar per capita em 1991 não se manteria constante ao longo do tempo, devido às mudanças ocorridas no sistema de telefonia do país em 1998 e ao conseqüente e esperado aumento na oferta de telefones – fixos e celulares. O grau de urbanização foi igualmente abandonado em face de uma possível alteração na sua relação com a renda e também do fato de que a grande maioria dos municípios possui um elevado grau de urbanização.

Assim, com as quatro variáveis restantes, candidatas a compor um indicador *proxy* de renda, foi construído um modelo de regressão, tendo por finalidade mensurar a associação com a renda familiar per capita derivada do Censo Demográfico de 1991. O modelo foi construído para o total dos municípios e para cada um dos cinco estratos de municípios, definidos segundo população. Este procedimento foi adotado porque as associações da renda familiar per capita com as variáveis pré-selecionadas podem manifestar-se de formas diferentes para municípios com distintos portes populacionais.

O modelo final, apresentado na Tabela 1, mostra que apenas o rendimento médio do emprego formal e o consumo anual de energia elétrica residencial por ligação são significantes no modelo para o total dos municípios. Enquanto o valor adicionado per capita não se mostra diretamente correlacionado com a renda per capita, nem no total nem nos estratos, o consumo anual de energia elétrica por ligação, na agricultura, mostrou-se associado à renda apenas em municípios com população inferior a 5.000 habitantes e naqueles com 20.000 a 49.999 habitantes.

Apesar destes resultados, decidiu-se manter o valor adicionado per capita como um dos componentes do indicador-síntese de riqueza municipal, pois representa uma importante faceta da economia municipal, não necessariamente expressa pela renda familiar per capita. Quanto ao consumo de energia elétrica rural, optou-se por substituí-lo por uma medida global do consumo de energia no comércio, nos serviços e na agricultura, ampliando-se assim os aspectos econômicos captados, como um contraponto entre



energia residencial e não residencial (excluindo indústria). A correlação entre estas duas variáveis foi de 0,64.

A análise de regressão múltipla foi aplicada também para a seleção de variáveis *proxies* da esperança de vida aos nascer, conforme mostra a Tabela 2.

**TABELA 1**  
Associação entre renda familiar per capita e as variáveis *proxies* (1), segundo tamanho de municípios

Tamanho de municípios	Rendimento médio do emprego formal	Valor adicionado per capita por ligação	Consumo anual de energia elétrica residencial por ligação	Consumo anual de energia elétrica na agricultura por ligação	R <sup>2</sup>
Menos de 5.000 hab.					0,830
De 5.000 a 9.999 hab.					0,563
De 10.000 a 19.999 hab.					0,368
De 20.000 a 49.999 hab.					0,706
50.000 hab. e mais					0,840
<b>Total dos municípios</b>					<b>0,705</b>

Variável significativa para o modelo ao nível de 5%.

(1) Utilizou-se a análise de regressão múltipla com o método backward processada no software SPSS – Statistical Package Social Science – versão 8.0. Foram excluídos desta análise vários municípios com valores “aberrantes” nas variáveis estudadas.

**TABELA 2**  
Resultados da análise de regressão contendo esperança de vida como variável resposta (1)

Variáveis explicativas	Estimativas	Nível descritivo
Constante	84,819	0,000
Taxa de mortalidade infantil	- 0,040	0,000
Taxa de mortalidade perinatal	- 0,011	0,035
Taxa de mortalidade da população de 15 a 39 anos	- 1,643	0,000
Taxa de mortalidade da população com mais de 60 anos	- 0,227	0,000
Óbitos por causas maldefinidas	0,006	(2) 0,224
Nascidos vivos com baixo peso	0,017	(2) 0,297
Nascidos vivos de mães com menos de 20 anos	- 0,007	(2) 0,484
Coefficiente de Explicação	0,840	

(1) Utilizou-se a análise de regressão múltipla com o método backward processada no software SPSS – Statistical Package Social Science – versão 8.0.

(2) Não significativa ao nível de 5%.

## **USO DA ANÁLISE FATORIAL PARA A CONSTRUÇÃO DE INDICADORES SINTÉTICOS DE RIQUEZA, LONGEVIDADE E ESCOLARIDADE**

Os indicadores sintéticos de riqueza, longevidade e escolaridade foram obtidos por meio da análise fatorial (factor analysis), técnica estatística de análise multivariada que se aplica à identificação de fatores que apontem objetivamente para a agregação de um conjunto de medidas. Ao contrário de um indicador criado por uma composição de variáveis arbitrada pelo pesquisador, os fatores derivados da análise fatorial são dimensões abstratas, cabendo ao pesquisador verificar se tais fatores podem ser interpretados de forma coerente com a natureza dos fenômenos ou processos estudados.

As análises fatoriais foram efetuadas com as variáveis padronizadas para cada dimensão e com os dados do ano mais recente, no software SPSS – Statistical Package Social Science – versão 8.0, com extração por componentes principais e máxima verossimilhança com a resultante de um único fator. A porcentagem da variância total das variáveis originais, explicada pelos fatores derivados das análises fatoriais, é apresentada na Tabela 3. Os pesos das variáveis que compuseram os fatores derivados da análise fatorial são mostrados no Quadro 6. Note-se que tais pesos foram padronizados para somarem um.

**TABELA 3**  
Porcentagem da variância total das variáveis originais explicada pelos fatores derivados das análises fatoriais

Dimensão	Parcela da variância explicada
Riqueza municipal	61%
Longevidade	48%
Escolaridade	63%

**QUADRO 6**  
Pesos das variáveis na geração dos indicadores

Dimensão	Variáveis	Pesos
Riqueza Municipal	Rendimento médio dos postos de trabalho dos assalariados do setor privado com carteira de trabalho e dos assalariados do setor público	0,1942
	Valor adicionado fiscal per capita	0,1390
	Consumo anual per capita de energia comercial, de serviços e rural	0,2371
	Consumo anual per capita de energia residencial	0,4351
Longevidade (1)	Taxa de mortalidade infantil	0,3000
	Taxa de mortalidade perinatal	0,3000
	Taxa de mortalidade da população com mais de 60 anos	0,2000
	Taxa de mortalidade da população com 15 a 39 anos	0,2000
Escolaridade	% de pessoas de 15 a 19 anos que concluíram o ensino fundamental	0,2570
	% de pessoas de 20 a 24 anos que concluíram o ensino médio	0,2450
	% de pessoas de 10 a 14 anos com mais de um ano completo de estudo	0,2390
	% de pessoas de 15 a 24 anos com mais de um ano completo de estudo	0,2300
	% de matrículas da rede municipal do ensino fundamental no total da rede pública	0,0300

1. O indicador deve ser subtraído de 100.

### ***USO DA ANÁLISE DE AGRUPAMENTO PARA A IDENTIFICAÇÃO DOS MUNICÍPIOS DO ESTADO DE SÃO PAULO QUE APRESENTAM CARACTERÍSTICAS SIMILARES DE RIQUEZA, LONGEVIDADE E ESCOLARIDADE***

A classificação dos municípios do Estado de São Paulo em grupos com características similares de riqueza, longevidade e escolaridade foi obtida por meio da análise de agrupamentos (cluster analysis), técnica estatística de análise multivariada que se aplica à divisão de uma determinada matriz de dados em um número menor de partições, observando critérios de similaridade conjunta. Nesse tipo de análise, as distâncias entre os objetos estudados dentro do espaço multiplano constituído por eixos de todas as variáveis são calculadas e, a seguir, os objetos são agrupados conforme a proximidade entre eles.

O grupo inicial é constituído pelos dois objetos mais próximos. A partir daí, verifica-se qual o objeto que se localiza mais próximo ao centro do primeiro grupo,

formando-se um novo grupo e, assim, sucessivamente, até que todos os objetos sejam reunidos em um único grupo, formado por todos os objetos estudados.

Terminado o procedimento descrito, resta ao pesquisador a interpretação dos resultados da análise de agrupamentos. Para isso, é fundamental avaliar o comportamento das variáveis originais dentro de cada grupo, buscando identificar aquelas que mais distinguem um determinado grupo dos demais. Se a interpretação ajustar-se de forma coerente à natureza do fenômeno ou processo estudado, pode ser utilizada como regra de classificação.

Foi exatamente este o procedimento adotado pela Fundação Seade, ao aplicar a análise de agrupamentos aos indicadores setoriais do ano mais recente para os 645 municípios, utilizando-se o algoritmo k-means do software SAS versão 6.12, o que gerou cinco grupos de municípios. A partir do perfil dos cinco grupos dos indicadores setoriais, criaram-se categorias para cada um desses indicadores e a combinação destas categorias gerou os agrupamentos finais, que reproduzem os mesmos perfis iniciais, conforme mostra a Tabela 4.

**TABELA 4**

Classificação dos municípios do Estado de São Paulo em grupos com características distintas de riqueza, longevidade e escolaridade

<b>Classificação dos grupos de municípios</b>				
Riqueza municipal	Longevidade	Escolaridade		
		Baixa: até 59	Média: 60 a 69	Alta: 70 e mais
<b>Baixa – score até 49</b>	<i>Baixa: até 59</i>	5	4	4
	<i>Média: 60 a 69</i>	4	3	3
	<i>Alta: 70 e mais</i>	4	3	3
<b>Alta – score de 50 e mais</b>	<i>Baixa: até 59</i>	2	2	1
	<i>Média: 60 a 69</i>	2	1	1
	<i>Alta: 70 e mais</i>	1	1	1