

**CENTRO DE ESTATÍSTICA APLICADA – CEA – USP**

**RELATÓRIO DE ANÁLISE ESTATÍSTICA**

**TÍTULO:** Relatório de análise estatística sobre o projeto: “Avaliação e pesquisa: Investigando as dificuldades em Matemática no Ensino Fundamental da Rede Municipal da cidade de São Paulo – 2<sup>a</sup> fase”

**PESQUISADOR:** Vinícius de Macedo Santos

**INSTITUIÇÃO:** Faculdade de Educação – USP

**FINALIDADE:** Projeto vinculado ao Programa Melhoria da Escola Pública da FAPESP

**RESPONSÁVEIS PELA ANÁLISE:** Viviana Giampaoli

Marcos Nascimento Magalhães

Fábio Carriel Fonseca

Natália Ferreira Anorozo

**REFERÊNCIA DESTE TRABALHO:**

GIAMPAOLI, V.; MAGALHÃES, M. N.; FONSECA, F. C. e ANOROZO, N. F., **Relatório de análise estatística sobre o projeto: “Avaliação e pesquisa: Investigando as dificuldades em Matemática no Ensino Fundamental da Rede Municipal da cidade de São Paulo – 2<sup>a</sup> fase”**. São Paulo, IME-USP, 2008. (RAE – CEA – 08P27)

## FICHA TÉCNICA

### PROGRAMAS COMPUTACIONAIS UTILIZADOS

Microsoft Excel for Windows (versão 2003)

Microsoft Word for Windows (versão 2003)

Minitab 14.

R 2.8.0

### TÉCNICAS ESTATÍSTICAS UTILIZADAS

Análise Descritiva Unidimensional (03:010)

Análise Descritiva Multidimensional (03:020)

Outros (07:990)

### ÁREA DE APLICAÇÃO

Educação (14:990).

### REFERÊNCIAS BIBLIOGRÁFICAS

BUSSAB, W.O.; MORETTIN, P.A. (2005). **Estatística Básica**. 5 ed. Ed. Saraiva. 280 p.

Pinheiro, J.C., Bates, D. M. (2000). **Mixed-Effects Models in S and S-Plus**. 1.ed.

Springer 526p.

Goldstein, H. (2003). **Multilevel Stastical Models**. 3. ed, Arnold. 253 p.

<http://atlasmunicipal.prefeitura.sp.gov.br/Login/Login.aspx>

<http://arqs.portaleducacao.prefeitura.sp.gov.br/ProvaSaoPaulo/livro.pdf>

[http://www.inep.gov.br/informativo/2007/ed\\_172.htm](http://www.inep.gov.br/informativo/2007/ed_172.htm)

## ÍNDICE

<b>Resumo.....</b>	<b>4</b>
<b>1. Introdução.....</b>	<b>5</b>
<b>2. Descrição do Estudo .....</b>	<b>6</b>
<b>3. Descrição das Variáveis.....</b>	<b>7</b>
<b>4. Análise Descritiva.....</b>	<b>8</b>
<b>5. Análise Inferencial.....</b>	<b>15</b>
<b>6. Conclusões.....</b>	<b>20</b>
<b>Apêndice A.....</b>	<b>21</b>
<b>Apêndice B.....</b>	<b>27</b>
<b>Apêndice C.....</b>	<b>37</b>
<b>Apêndice D.....</b>	<b>42</b>
<b>Apêndice E.....</b>	<b>48</b>

## Resumo

Os índices de reprovação na rede municipal de ensino da cidade de São Paulo são cada vez mais alarmantes, notando-se um ponto de “afunilamento” no aproveitamento dos alunos na passagem do primeiro para o segundo ciclo do ensino fundamental na rede municipal de ensino de São Paulo. Acredita-se que a Matemática seja a principal causa deste problema.

Foi considerada uma amostra de 531 alunos de classe regulares de 20 escolas e 15 subprefeituras da cidade de São Paulo, que se submeteram a uma prova de Matemática (Avaliação Diagnóstico) de 27 questões no total.

O objetivo era comparar resultados da Avaliação de Diagnóstico do estudo com os resultados da macro-avaliação Prova São Paulo e verificar a possível influência de variáveis socioeconômicas no aprendizado de Matemática dos alunos da quarta série do Ensino Fundamental da rede de ensino do município de São Paulo.

Devido às limitações técnicas, não foi possível obter dados socioeconômicos dos alunos. Dessa forma, foram consideradas variáveis socioeconômicas do ano 2000 relativas às subprefeituras da cidade de São Paulo na qual a escola estava inserida

Em razão da característica da amostra, na qual os alunos estão agrupados por conglomerados (escolas), construiu-se um modelo de regressão que considerou o agrupamento dos alunos por subprefeitura (modelo multinível), sendo a variável resposta a nota do aluno na Avaliação Diagnóstico. As variáveis explicativas consideradas inicialmente foram: nota do aluno na P.S.P., idade do aluno, sexo do aluno (0 para meninos e 1 para meninas) e 10 variáveis socioeconômicas relativas à subprefeitura

Após a seleção das variáveis explicativas, verificou-se que apenas a nota na P.S.P. e a idade do aluno são significativas. O modelo obteve um ajuste razoável. Os índices socioeconômicos da subprefeitura em que a escola está inserida não influenciam o desempenho médio da escola ou do aluno. Talvez sejam necessárias observações socioeconômicas em nível de aluno para que seja constatada alguma influência no seu

desempenho na Avaliação Diagnóstico.

## **1. Introdução**

Segundo a Secretaria Municipal de Educação (SME) da Prefeitura do Município de São Paulo (PMSP), os índices de reprovação em Matemática são cada vez mais alarmantes, sendo consideravelmente superiores na passagem do primeiro para o segundo ciclo do ensino fundamental.

Este projeto visa dar continuidade a uma fase inicial, desenvolvida no segundo semestre de 2007, cujo objetivo era levantar as dificuldades verificadas no processo de ensino e aprendizagem de Matemática ao final do primeiro ciclo do ensino fundamental na rede de ensino da PMSP através de uma Avaliação Diagnóstico. Aplicada a uma amostra de alunos do quarto ano do referido ciclo, essa Avaliação contém questões referentes a cinco domínios de conteúdo.

A segunda fase do projeto tem por objetivo realizar um estudo comparativo entre os resultados da Avaliação Diagnóstico, analisados na fase inicial do projeto, e os resultados da Prova São Paulo para os alunos do quarto ano do Ciclo 1 do ensino fundamental, considerando-se a natureza de cada uma das avaliações, bem como verificar a possível influência de dados referentes ao perfil sócio-econômico dos alunos no desempenho em Matemática dos mesmos. A Prova São Paulo é uma macro-Avaliação externa elaborada e aplicada nesse caso pela empresa Herkenhoff & Prates em novembro de 2007. Segue a matriz do Saeb (Sistema de Avaliação da Educação Básica) e foi realizada utilizando os mesmos parâmetros e valores do sistema do MEC (Ministério da Educação) e formulada segundo a Teoria de Resposta ao Item (TRI). Este fato torna mais consistente a comparação dos resultados ano a ano devido ao fato de permitir que os mesmos conteúdos sejam avaliados em edições diferentes dos testes, inclusive, possibilita comparar provas com itens diferentes, porém de igual dificuldade. As provas foram estruturadas com base nas Matrizes de Competências e Habilidades de Referência para a Avaliação. Para cada ano/ciclo as provas apresentaram aos alunos cerca de 40 itens de Matemática. As provas foram estruturadas com itens de resposta escrita pelos

alunos e itens de múltipla escolha. Os resultados de todas as escolas foram colocados na escala nacional do Saeb.

## **2. Descrição do estudo**

Na primeira fase do projeto uma Avaliação Diagnóstico foi aplicada a uma amostra de 1328 alunos da quarta série do ensino fundamental, selecionados a partir de 26 escolas da rede municipal de ensino da cidade de São Paulo. As escolas foram distribuídas pelas 13 coordenadorias de ensino nas quais a Secretaria Municipal de Educação (SME) divide o município. Os alunos foram selecionados a partir de dois tipos de classes: as classes regulares e as classes participantes do Programa Intensivo do Ciclo I (PIC). Nas classes PIC, destinada aos alunos com histórico de reprovação, espera-se um ambiente de estudo diferenciado, caracterizado por maior tempo em sala de aula e mais atenção do professor aos alunos.

As crianças participantes do estudo realizaram uma Avaliação Diagnóstico contendo 27 questões, entre dissertativas e de múltipla escolha, relativas a cinco dos denominados domínios de conteúdo: Números (oito questões), Operações (oito questões), Tratamento da Informação (uma questão), Grandezas e Medidas (cinco questões) e Espaço e Forma (cinco questões). O aproveitamento do aluno em cada questão foi mensurado em uma escala qualitativa com quatro níveis: acerto (A), acerto parcial (AP), erro (E) e não-resposta (NR). A Avaliação foi realizada em um único dia, com duração máxima de 1h40min, sempre na mesma escola onde a criança estuda, sendo exatamente a mesma Avaliação para todas as escolas.

A seleção das crianças participantes do estudo foi feita através de uma amostragem probabilística em dois estágios: no primeiro estágio, selecionou-se a escola e, no segundo, os alunos da quarta série. O tamanho da amostra foi fixado inicialmente em 1350 crianças em 30 escolas, que era a capacidade disponível para a aplicação e correção das provas. A partir de cada uma das 13 coordenadorias de ensino da cidade foi selecionado um número de unidades de ensino proporcional ao número de alunos em cada coordenadoria; definido o número de escolas em cada coordenadoria, as unidades

escolares foram selecionadas de forma eqüiprovável. Em seguida, o número de crianças em cada escola foi determinado de forma proporcional ao total de alunos da respectiva unidade de ensino, tanto para os alunos de classes regulares quanto para os alunos de classes PIC; na ausência de alunos de classes PIC em uma unidade escolar, a amostra foi recolhida nas escolas restantes com este tipo de classe.

Nesta segunda fase do estudo, a fim de comparar a Avaliação Diagnóstico com a Prova São Paulo, foram disponibilizados pela Secretaria Municipal de Ensino os dados dos resultados obtidos na Prova São Paulo de 605 dos 1328 alunos que participaram da Avaliação Diagnóstico inicial, pertencentes a 20 das 26 escolas da amostra original. Complicações de caráter burocrático impediram a obtenção das informações para a totalidade dos alunos. Com o intuito de comparações mais precisas, foram excluídos os alunos das classes PIC, restando 536 alunos de classes regulares. Juntamente com a Prova São Paulo foi aplicado um questionário sócio-econômico para ser respondido pelos responsáveis dos alunos que realizaram a prova. Entretanto, não foi possível a disponibilidade dos dados de tal questionário. Assim, para atender ao objetivo do pesquisador de verificar a influência das variáveis socioeconômicas no aprendizado de matemática dos alunos, foram obtidos dados sócio-econômicos do *Atlas do Trabalho e Desenvolvimento da Cidade de São Paulo* das 31 subprefeituras de São Paulo referentes ao ano 2000, destacando que a amostra do presente estudo contempla 15 subprefeituras.

Para comparação entre subprefeituras obteve-se a nota média de cada subprefeitura, tanto para a Avaliação quanto para a Prova São Paulo. No caso de haver subprefeituras com mais de uma escola, a média foi calculada ponderando-se cada escola pelo número de alunos.

### **3. Descrição das variáveis**

#### **Variáveis da Avaliação Diagnóstico:**

- Subprefeitura:** subprefeitura de ensino a qual pertence o aluno.
- Pontuação:** O aproveitamento em cada questão recebeu um dos seguintes conceitos ou notas:

Acerto (A) – 1 ponto;

Acerto Parcial (AP) – 0,5 ponto;

Erro (E) – 0 ;

Não-resposta (NR) – 0.

- Sexo:** Sexo do aluno, assumindo o valor 0 para alunos do sexo masculino e 1 para valores do sexo feminino.

#### **Variável Prova São Paulo:**

- Escore:** Mede a nota geral do aluno na Prova São Paulo, calculado através de Modelos da Teoria de Resposta ao Item.

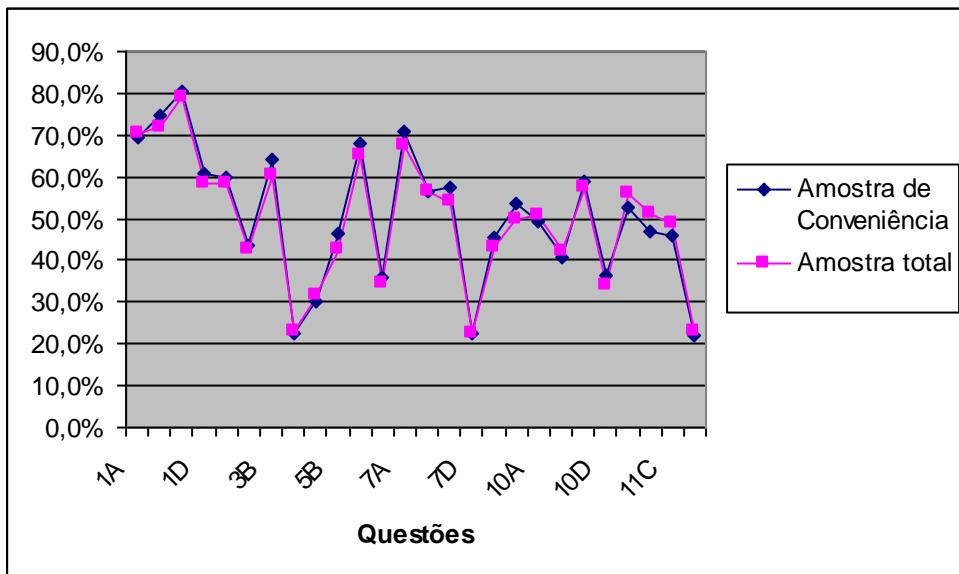
#### **Variáveis socioeconômicas:**

- Índice de Desenvolvimento Humano Municipal (IDH);
- Taxa de alfabetização;
- Média de moradores por domicílio;
- Renda média dos responsáveis pelo domicílio;
- Renda per capita;
- Índice de Gini;
- Média de anos de estudo dos responsáveis pelo domicílio;
- Percentual de adolescentes do sexo feminino entre 10 e 14 anos com filhos;
- Percentual de adolescentes do sexo feminino entre 15 e 17 anos com filhos;
- Percentual de crianças em domicílios com renda per capita menor que R\$37,75.

#### **4. Análise Descritiva**

Devido ao fato de não se ter a amostra total, mas sim uma sub-amostra da mesma, foi construído o Gráfico 1 apresentado a seguir, com a porcentagem de acertos de cada uma das 27 questões da Avaliação Diagnóstico para a amostra total e para a sub-amostra. Os Gráficos B.1, B.2 e B.3, apresentados no Apêndice B, com a porcentagem de erros, respostas em branco e acertos parciais, complementam a avaliação da sub-amostra apresentada. De agora em diante, denominaremos a sub-amostra de amostra de conveniência.

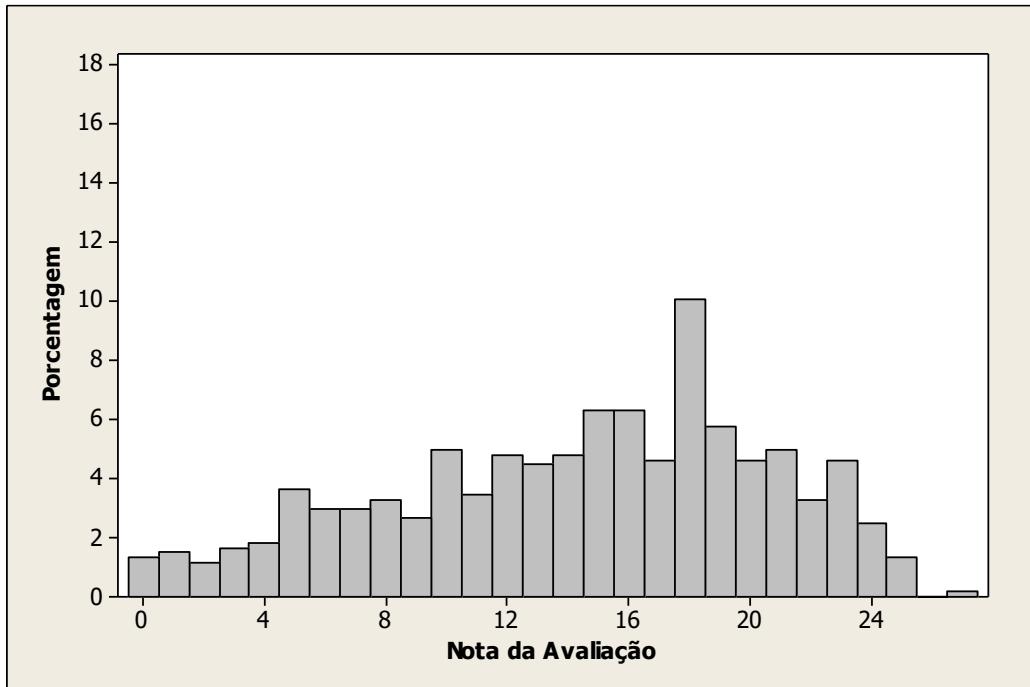
**Gráfico 1 – Porcentagem de acertos para as 2 amostras.**



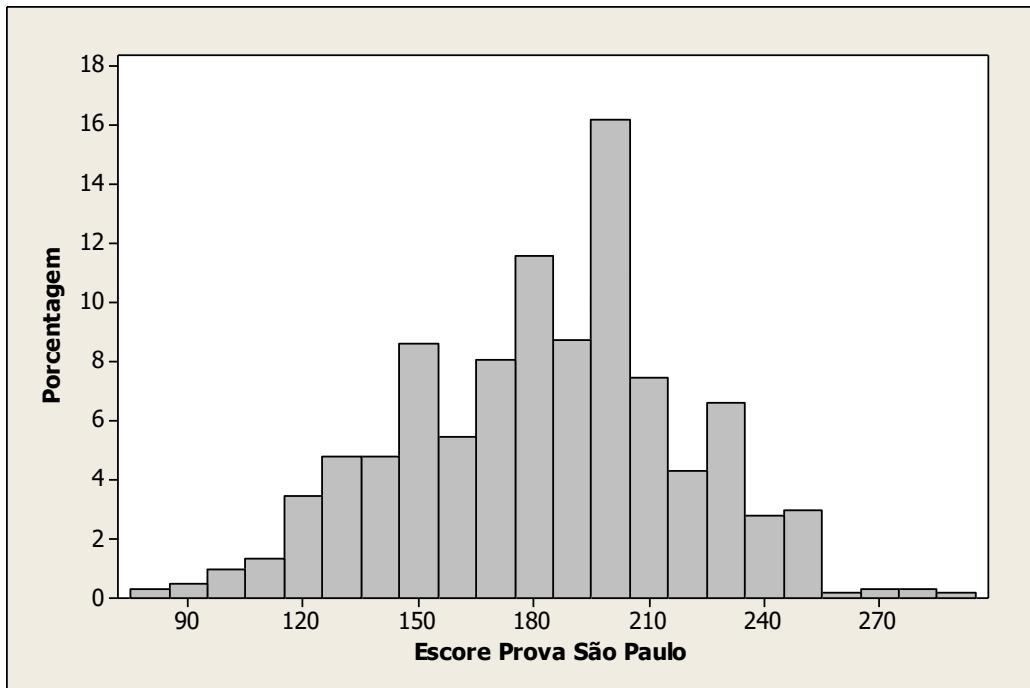
Observando-se os Gráficos 1, B.1, B.2 e B.3 nota-se uma grande proximidade entre as porcentagens de acertos da amostra total e da amostra de conveniência para todos os tipos de resposta, sugerindo que a amostra de conveniência também é adequada para se fazer inferências sobre a população.

O Gráfico 2 e o Gráfico 3 apresentados a seguir, indicam uma distribuição assimétrica das notas da Avaliação e dos escores da Prova São Paulo. Entretanto, nota-se que a Prova São Paulo possui assimetria um pouco mais acentuada (Coeficiente de Assimetria= - 0,38, Tabela 1).

**Gráfico 2 – Histograma das freqüências relativas das notas da Avaliação Diagnóstico.**



**Gráfico 3 – Histograma da das freqüências relativas das notas da Prova São Paulo.**



A Tabela 1 abaixo apresenta as medidas descritivas das notas para ambas as provas.

**Tabela 1 -** Medidas descritivas das notas da Avaliação Diagnóstico e da Prova São

Paulo.

Variável	IIº de Observações	Média	Desvio Padrão	Variância	Mínimo	Coef de Var
Prova São Paulo	605	182,33	36,31	1318,27	81,60	19,91%
Variável	Q1	Mediana	Q3	Máximo	Assimetria	Q3 - Q1
Prova São Paulo	155,39	185,11	205,26	289,29	-0,16	49,87
Avaliação	9,50	15,00	19,00	26,50	-0,38	9,50

Chama a atenção o valor grande do Coeficiente de Variação da Avaliação Diagnóstico (43,52%) em relação à Prova São Paulo (19,91%).

A nota da Avaliação pertence ao intervalo [0;27]. A mediana da nota da Avaliação apresentada na Tabela 1 indica que 50% dos alunos obtiveram uma nota inferior ou igual a 15 e o Terceiro Quartil (Q3) indica que 75% dos alunos obtiveram uma nota inferior a 19.

No Apêndice A, na Tabela A.1, são apresentadas as medidas descritivas de ambas as provas para cada uma das subprefeituras. As medianas mais altas foram obtidas na Vila Prudente / Sapopemba (206,74 para a Prova São Paulo e 18,25 para a Avaliação Diagnóstico). Nota-se que a subprefeitura Vila Prudente foi a que obteve a maior média das notas da Avaliação Diagnóstico (17,50), porém esta nota é inferior ao terceiro quartil das notas dos alunos (19,00), e foi também a que obteve a maior média dos escores da Prova São Paulo (206,49), praticamente igual ao terceiro quartil das notas dos alunos (205,26). Nota-se também que a menor média das notas da Avaliação Diagnóstico foi obtida pela subprefeitura Itaquera (10,7), nota próxima ao primeiro quartil das notas dos alunos (9,5), porém a que obteve menor média dos escores da Prova São Paulo foi a subprefeitura Butantã (166,79), superior ao primeiro quartil (155,39).

Por outro lado, a subprefeitura de Butantã apresentou a mediana mínima na Prova São Paulo (166,89), enquanto que na Avaliação foi Itaquera(10,25).

No geral, os coeficientes de assimetria são maiores na Avaliação Diagnóstico do que na Prova São Paulo.

Observando a Tabela A.2 e o Gráfico B.11, nota-se que na subprefeitura Tiradentes foram avaliados poucos alunos (apenas 6 alunos) e que a subprefeitura com maior número de alunos avaliados foi Jabaquara (73 alunos).

Estas diferenças podem sinalizar uma perda de informação ao considerar como unidade de estudo a subprefeitura no lugar do aluno.

**Tabela 2** - Medidas descritivas dos indicadores socioeconômicos.

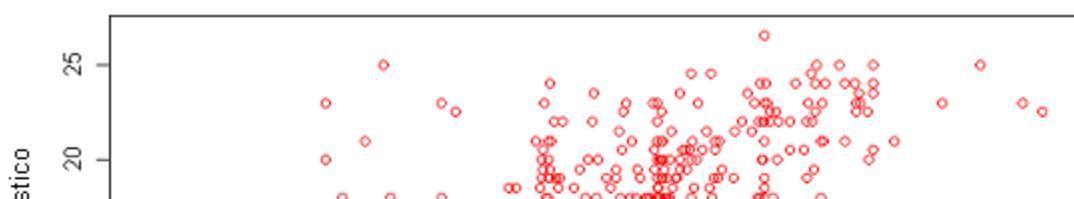
Variável	IIº de Observações	Média	Desvio Padrão	Variância	Mínimo	Coef de Var
IDH	15,00	0,81	0,04	0,00	0,76	0,05
Taxa de Alfabetização	15,00	94,28	1,26	1,60	92,58	0,01
nº de moradores	15,00	3,63	0,16	0,02	3,35	0,04
Variável	Q1	Mediana	Q3	Máximo	Assimetria	Q3 - Q1
IDH	0,78	0,80	0,86	0,89	0,52	0,08
Taxa de Alfabetização	92,95	94,22	95,42	96,35	0,05	2,47
nº de moradores	3,51	3,66	3,76	3,83	-0,43	0,25

Observando-se a Tabela 2, nota-se que o IDH obteve valores considerados de médios (entre 0,5 e 0,76) e altos (superiores a 0,76), taxas de alfabetização altas (superiores a 92,58%) e que o número de moradores por domicílio foi entre 3 e 4.

Os Gráficos B.4 e B.5 reforçam a idéia de assimetria para as notas de ambas as provas. Observa-se que a variável escore da Prova São Paulo apresenta dois pontos atípicos.

O Gráfico 4 a seguir apresenta a dispersão do escore da Prova São Paulo versus a nota da Avaliação Diagnóstico, indicando uma aparente correlação linear positiva entre as variáveis.

**Gráfico 4** – Gráfico de dispersão da nota Avaliação Diagnóstico versus o escore da Prova São Paulo por aluno.

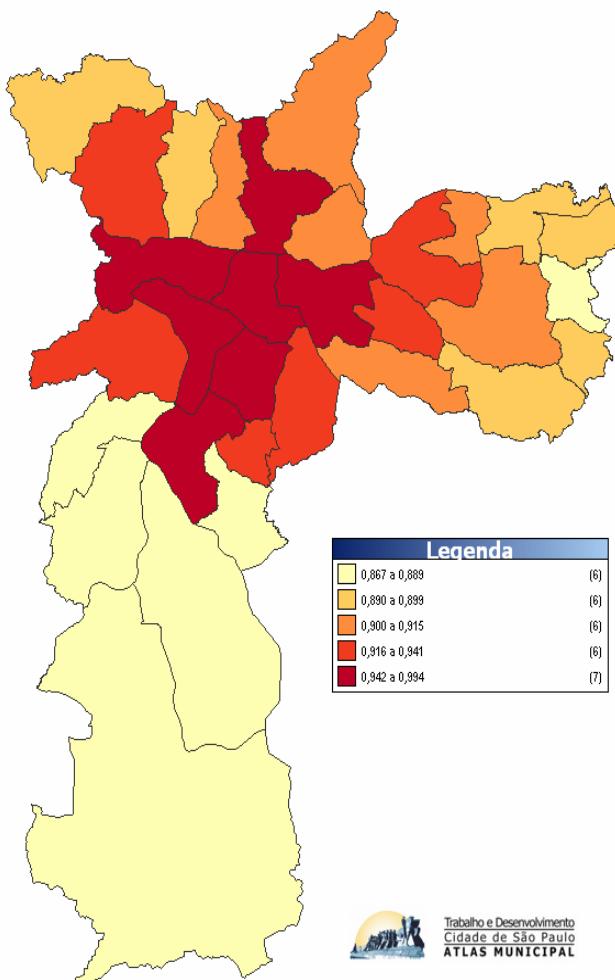


O Gráfico B.6, também evidencia uma correlação linear positiva entre as duas provas, sendo que a subprefeitura de Itaquera mostra-se um provável ponto atípico.

Os gráficos de dispersão B.7 e B.8 indicam a inexistência de correlação entre as variáveis Índice de Desenvolvimento Humano (IDH) e desempenho médio de cada subprefeitura para ambas as provas. A mesma conclusão pode ser obtida para a variável Taxa de Alfabetização (Gráficos B.9 e B.10).

O Mapa 1 a seguir mostra o IDH de todas as subprefeituras do município de São Paulo. Ao observar o mapa, nota-se que as subprefeituras da região central do município pertencem à maior faixa de IDH (0,942 a 0,944) e que as coordenadorias do sul do município pertencem à menor faixa de IDH (0,867 a 0,899). Pela Tabela 2, também percebe-se que o IDH possui uma assimetria à esquerda (0,52), com a mediana mais próxima do 1º quartil em comparação ao 3º.

**Mapa 1** – Mapa temático do IDH de todas as subprefeituras do município do São Paulo.



No Apêndice C são apresentados mapas temáticos obtidos do Atlas do Trabalho e Desenvolvimento da Cidade de São Paulo das 31 subprefeituras de São Paulo referentes ao ano 2000 (último ano disponível).

Para os Mapas C.1 e C.3, que mostram respectivamente, a Taxa de Alfabetização e a Renda per Capita média de todas as subprefeituras do município de São Paulo, pode ser observado um comportamento muito semelhante ao do Mapa 1. As subprefeituras da região central do município possuem os maiores valores observados das duas variáveis socioeconômicas em questão, ocorrendo o inverso para a zona Sul da cidade.

Já o Mapa C.4 possui um comportamento inverso ao dos mapas C.1 a C.3: para a maioria das subprefeituras: quanto maior o IDH e a renda per capita, menor o número

de moradores por domicílio

O Mapa C.2 mostra a localização das 15 subprefeituras utilizadas no estudo, destacadas em vermelho. A tabela junto ao mapa mostra a média das notas da Avaliação Diagnóstico e a média dos escores da Prova São Paulo de cada uma dessas 15 subprefeituras.

## 5. Análise Inferencial

A análise inferencial é a etapa onde estende-se os resultados da amostra para toda a população de interesse usando modelos estatísticos.

Para esta finalidade, foram ajustados Modelos Lineares Hierárquicos (multiníveis) aos dados com o objetivo de explicar a variável nota da Avaliação Diagnóstico em função das demais variáveis consideradas no estudo.

Selecionamos este tipo de modelo devido ao fato de ele levar em conta a hierarquia existente entre os dados. Nesse caso, podemos observar 3 níveis de hierarquia: alunos, agrupados em escolas, que estão agrupadas em subprefeituras. Optamos pelo agrupamento em 2 níveis, com os alunos agrupados em subprefeituras, devido ao fato de que as variáveis socioeconômicas (A.4) a serem incluídas no modelo referirem-se à subprefeitura, e não a escola, e pelo fato de que, das 15 subprefeituras consideradas no estudo, 10 delas contêm apenas uma escola, as 5 outras contém 2 escolas cada.

Assim, ao considerar a hierarquia presente nos dados, o modelo leva em conta as variações entre as subprefeituras, podendo incorporar diferentes aspectos e a possível correlação entre as notas de seus alunos, permitindo uma maior facilidade na interpretação de seus parâmetros.

Na etapa inicial da análise foi considerado o modelo E.1 apresentado em detalhes no Apêndice E, que considera em nível de subprefeituras todas as variáveis socioeconômicas. Porém, depois de um processo de seleção de variáveis, estas resultaram estatisticamente não significativas. O modelo final selecionado considera a relação linear da nota da Avaliação Diagnóstico com a nota da Prova São Paulo e a idade do aluno centradas nos valores 100 e 10, respectivamente.

A tabela a seguir apresenta os resultados do ajuste do modelo selecionado.

**Tabela 3 – Ajuste do modelo selecionado.**

	<b>Coeficiente</b>	<b>Desvio Padrão</b>	<b>Graus de Liberdade</b>	<b>Valor t</b>	<b>p-valor</b>
<b>Intercepto</b>	6,365	1,037	514	6,136	<0,000
<b>Escore - 100</b>	0,096	0,013	514	7,359	<0,000
<b>Idade - 10</b>	-0,694	0,287	514	-2,416	0,016

A Tabela 3 acima mostra os coeficientes das variáveis do modelo final, bem como o desvio padrão dos mesmos, os graus de liberdade, o valor da estatística t e o nível descritivo. Observando-se os níveis descritivos das variáveis, nota-se que todas são significantes para o modelo, pois o nível descritivo das mesmas é próximo de zero. Entretanto, faz-se aqui a ressalva de que, para alguns autores, o coeficiente da variável Idade possui um p-valor relativamente alto quando se considera o tamanho da amostra, o que levaria à conclusão de que a Idade não é significativa para o modelo. Apesar disso, optamos por incluir a variável Idade no modelo final.

Da Tabela 3 podemos perceber que, como o coeficiente do escore da Prova São Paulo é positivo (0,096), espera-se que, entre dois alunos com a mesma idade, aquele com nota mais alta na Prova São Paulo terá nota mais alta na Avaliação Diagnóstico. Por outro lado, como o coeficiente associado à idade é negativo, espera-se que a nota da Avaliação Diagnóstico diminua à medida que a idade aumenta.

A Tabela A.3 também apresenta o desvio padrão dos efeitos aleatórios do intercepto do modelo, do escore e do resíduo, bem como a correlação existente entre os efeitos aleatórios do intercepto e do escore.

A necessidade da inclusão de um efeito aleatório associado à variável escore foi testada por meio da estatística da razão de verossimilhanças (Apêndice E.4)

O Gráfico B.12 mostra os *boxplots* dos resíduos padronizados do modelo por subprefeitura e indicam um ajuste razoável.

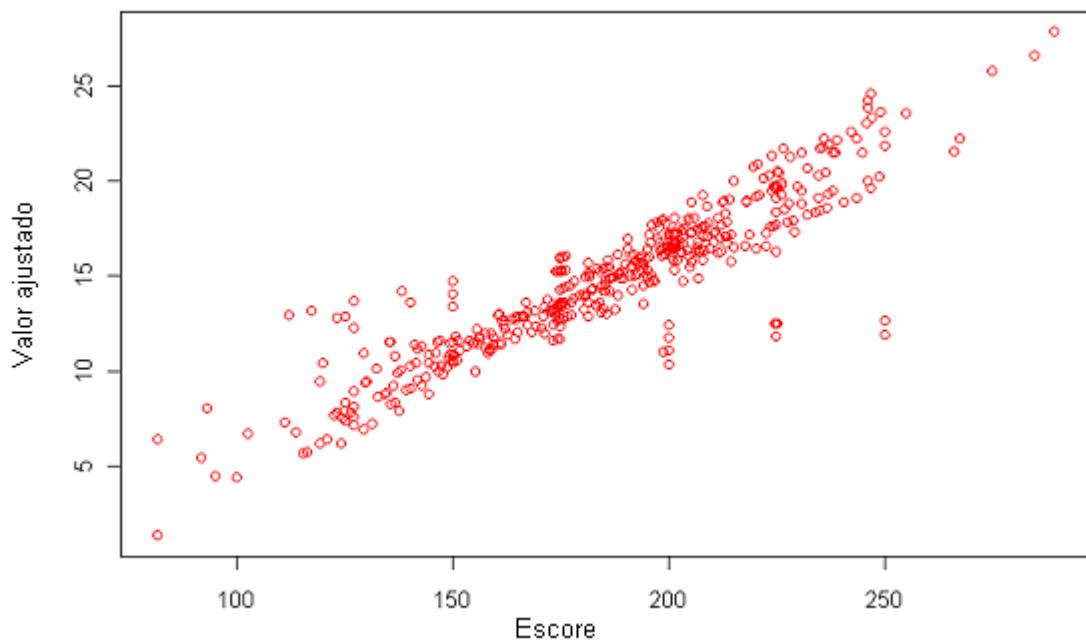
O Gráfico B.13 indica que os resíduos padronizados do modelo seguem aproximadamente a distribuição normal padrão, pois os pontos aproximam-se de uma reta crescente. O mesmo pode ser observado para o Gráfico B.14, onde são

apresentados também gráficos QQ-plot, porém por subprefeitura. Para todas as subprefeituras, nota-se que os pontos aproximam-se de uma reta crescente.

O Gráfico B.15 mostra os resíduos padronizados por subprefeitura. Nota-se que para todas as subprefeituras os resíduos distribuem-se aleatoriamente ao redor do zero, com exceção dos resíduos das subprefeituras Itaquera e Socorro, cujo comportamento apresenta-se bastante atípico.

O Gráfico B.16 mostra os gráficos QQ-plot dos efeitos aleatórios do intercepto e da variável escore, nota-se que os pontos aproximam-se pouco de uma reta, indicando que a distribuição dos efeitos aleatórios aproxima-se razoavelmente da distribuição normal.

**Gráfico 5 –** Gráfico de dispersão do escore da Prova São Paulo versus a nota da Avaliação Diagnóstico ajustada pelo modelo.



O Gráfico 5 apresentado anteriormente indica um bom ajuste do modelo, pois os pontos aproximam-se consideravelmente mais de uma reta crescente do que os pontos do Gráfico 4 apresentado anteriormente. Ou seja, a variável Nota da Avaliação Diagnóstico pode ser melhor explicada pela variável escore da Prova São Paulo, levando-se em conta também a idade dos alunos.

O Gráfico B.17 da dispersão da nota ajustada pelo modelo versus a idade dos alunos mostra que a idade dos alunos varia de 8 a 13 anos, porém a grande maioria dos alunos tem de 9 a 11 anos, nota-se também que com o aumento da idade parece haver uma diminuição da nota ajustada pelo modelo, como é esperado.

**Tabela 4** - Tabela comparativa das notas médias da Avaliação Diagnóstico por subprefeitura

Subprefeitura	Valores Observados	Valores Ajustados
Aricanduva	15,83	15,71
Butantã	13,03	12,88
Campo Limpo	13,06	13,37
Casa Verde/Cachoeirinha	15,08	15,00
Cidade Tiradentes	14,66	14,30
Freguesia/Brasilândia	14,70	14,75
Ipiranga	16,00	15,54
Itaim Paulista	15,27	15,27
Itaquera	10,70	11,81
Jabaquara	14,40	14,47
M' Boi Mirim	15,27	15,36
São Mateus	14,13	14,01
São Miguel	16,23	16,32
Socorro	16,23	15,64
V. Prudente/Sapopemba	17,50	17,61

Na tabela acima, observamos que as notas médias observadas das subprefeituras são muito próximas das notas médias ajustadas pelo modelo. A subprefeitura de Itaquera apresenta as notas de Avaliação Diagnóstico mínimas ajustadas e observadas (10,70 e 11,81 respectivamente). Já a subprefeitura de Vila Prudente / Sapopemba obteve as notas máximas 17,5 e 17,61. Desta análise podemos concluir que o modelo proposto é razoável.

Para fins de comparação, ajustamos um modelo padronizando as variáveis de forma que ambas as notas estejam no intervalo [0,1], sem centralizar (ver Apêndice E, Modelo E.3). Na tabela a seguir estão os valores estimados para esse modelo. As conclusões são as mesmas do modelo anterior, destacando que os valores da Prova São

Paulo são sempre inferiores à Avaliação Diagnóstico. Porém, preferimos o modelo (E.2) na escala original das mesmas, uma vez que a Prova São Paulo, segundo os especialistas, possui uma interpretação bem definida e que se perde ao realizar a padronização para o intervalo [0,1].

Para verificar se de fato as variáveis socioeconômicas não influenciam o desempenho dos alunos na Avaliação Diagnóstico, foi ajustado um modelo para explicar a Nota da Avaliação Diagnóstico em função apenas das variáveis socioeconômicas, excluindo-se a variável escore da Prova São Paulo, e notou-se um ajuste ruim, nenhuma delas mostrou ser significativas para o modelo.

**Tabela 5** - Tabela comparativa das notas médias da Avaliação das subprefeituras no modelo padronizado para a escala [0,1].

Subprefeitura	Observado	Ajustado	Nota SP
Aricanduva	0,586	0,588	0,396
Butantã	0,483	0,468	0,334
Campo Limpo	0,484	0,526	0,362
Casa Verde/Cachoeirinha	0,558	0,554	0,382
Cidade Tiradentes	0,543	0,540	0,369
Freguesia/Brasilândia	0,545	0,540	0,371
Ipiranga	0,593	0,539	0,368
Itaim Paulista	0,566	0,557	0,374
Itaquera	0,396	0,563	0,383
Jabaquara	0,533	0,533	0,364
M' Boi Mirim	0,566	0,552	0,368
São Mateus	0,523	0,511	0,354
São Miguel	0,601	0,583	0,395
Socorro	0,601	0,523	0,360
V, Prudente/Sapopemba	0,648	0,620	0,413

## 6. Conclusões

Os resultados sugerem que a amostra de conveniência e a amostra original

apresentam o mesmo comportamento. Entretanto, faz-se a ressalva que talvez o comportamento da amostra de conveniência para cada subprefeitura tomada isoladamente pode não se aproximar tanto do comportamento da amostra total.

A partir da Análise inferencial, pode-se concluir que nenhuma das variáveis socioeconômicas utilizadas no estudo das subprefeituras influenciam o desempenho médio dos alunos na Avaliação Diagnóstico. Talvez se tais variáveis fossem observadas em nível de aluno, e referentes ao ano em que foram aplicadas a Avaliação Diagnóstico e a Prova São Paulo (2007), obteríamos resultados mais confiáveis.

O ajuste do modelo foi bom e mostrou que de fato a nota da Avaliação Diagnóstico está associada ao Escore da Prova São Paulo positivamente, bem como está associada também à variável idade.

## **Tabelas - Apêndice A**

**Tabela A.1 –** Tabela das medidas descritivas das notas da Avaliação Diagnóstico e da Prova São Paulo por subprefeitura para os alunos regulares.

### Aricanduva

Variável	Nº de Observações	Média	Desvio Padrão	Variância	Mínimo	Coef de Var
<b>Prova São Paulo</b>	49,00	195,80	38,62	1.491,68	93,17	0,20
<b>Avaliação</b>	49,00	15,48	6,43	41,37	0,00	0,42
Variável	Q1	Mediana	Q3	Máximo	Assimetria	Q3 - Q1
<b>Prova São Paulo</b>	181,14	201,27	228,16	266,00	-0,63	47,02
<b>Avaliação</b>	10,00	16,50	21,00	26,50	-0,34	11,00

### Butantã

Variável	Nº de Observações	Média	Desvio Padrão	Variância	Mínimo	Coef de Var
<b>Prova São Paulo</b>	16,00	166,79	34,50	1.190,31	81,60	0,21
<b>Avaliação</b>	16,00	13,03	4,11	16,92	6,50	0,32
Variável	Q1	Mediana	Q3	Máximo	Assimetria	Q3 - Q1
<b>Prova São Paulo</b>	146,14	166,89	195,03	217,86	0,92	48,89
<b>Avaliação</b>	9,63	12,75	17,50	19,00	0,05	7,87

### Campo Limpo

Variável	Nº de Observações	Média	Desvio Padrão	Variância	Mínimo	Coef de Var
<b>Prova São Paulo</b>	55,00	181,03	33,54	1.124,93	91,92	0,19
<b>Avaliação</b>	55,00	13,06	5,87	34,42	0,00	0,45
Variável	Q1	Mediana	Q3	Máximo	Assimetria	Q3 - Q1
<b>Prova São Paulo</b>	162,30	180,61	203,37	267,34	-0,08	41,07
<b>Avaliação</b>	8,50	13,50	17,00	24,00	-0,35	8,50

### Casa Verde / Cachoeirinha

Variável	Nº de Observações	Média	Desvio Padrão	Variância	Mínimo	Coef de Var
Prova São Paulo Avaliação	20,00	191,00	28,33	802,59	129,93	0,15
Variável	Q1	Mediana	Q3	Máximo	Assimetria	Q3 - Q1
Prova São Paulo Avaliação	177,35	196,58	206,22	238,18	-0,82	28,87
	9,63	16,50	20,00	22,00	-0,64	10,37

### Cidade Tiradentes

Variável	Nº de Observações	Média	Desvio Padrão	Variância	Mínimo	Coef de Var
Prova São Paulo Avaliação	6,00	184,60	27,90	778,41	141,10	0,15
Variável	Q1	Mediana	Q3	Máximo	Assimetria	Q3 - Q1
Prova São Paulo Avaliação	162,10	185,30	209,80	220,40	-0,41	47,70
	11,00	14,50	18,25	22,00	0,24	7,25

### Freguesia / Brasilândia

Variável	Nº de Observações	Média	Desvio Padrão	Variância	Mínimo	Coef de Var
Prova São Paulo Avaliação	31,00	180,92	43,27	1.872,29	95,27	0,24
Variável	Q1	Mediana	Q3	Máximo	Assimetria	Q3 - Q1
Prova São Paulo Avaliação	146,09	186,35	204,90	284,00	-0,16	58,81
	8,50	16,00	19,50	23,00	-0,61	11,00

### Ipiranga

Variável	Nº de Observações	Média	Desvio Padrão	Variância	Mínimo	Coef de Var
Prova São Paulo Avaliação	25,00	183,89	33,01	1.089,66	125,34	0,18
Variável	Q1	Mediana	Q3	Máximo	Assimetria	Q3 - Q1
Prova São Paulo Avaliação	155,42	190,51	206,67	246,16	-0,09	51,25
	11,75	18,00	20,25	24,50	-0,83	8,50

**Itaim Paulista****São Mateus**

Variável Variável	Nº de Nº de Observações	Média Média	Desvio Desvio Padrão	Variância Variância	Mínimo Mínimo	Coef de Coef de Var
Prova São Paulo	38,00	17618128	24,88	833,87	16109	0,82
Avaliação Variável	3Q10 Q1	Mediana Mediana	Q3 5,56 Q3	Máximo Máximo	Assimetria Assimetria	Q3 - Q39 Q3 - Q1
Prova São Paulo	160,92	173,06	198,50	224,07	0,15	34,00
Avaliação	11,00	15,00	18,00	23,00	-0,61	7,00

**Itaquera****São Miguel**

Variável Variável	Nº de Nº de Observações	Média Média	Desvio Desvio Padrão	Variância Variância	Mínimo Mínimo	Coef de Coef de Var
Prova São Paulo	26,00	19714570	36,22	943,89	125025	0,83
Avaliação Variável	2Q10 Q1	Mediana Mediana	Q3 5,33 Q3	Máximo Máximo	Assimetria Assimetria	Q3 - Q33 Q3 - Q1
Prova São Paulo	175,60	200,26	212,06	223,00	-0,07	49,25
Avaliação	14,75	16,75	19,00	25,00	-0,93	4,25

**Jabaquara****Socorro**

Variável Variável	Nº de Nº de Observações	Média Média	Desvio Desvio Padrão	Variância Variância	Mínimo Mínimo	Coef de Coef de Var
Prova São Paulo	66,00	18011040	36,68	926,26	102507	0,42
Avaliação Variável	6Q10 Q1	Mediana Mediana	Q3 4,93 Q3	Máximo Máximo	Assimetria Assimetria	Q3 - Q31 Q3 - Q1
Prova São Paulo	167,39	176,94	208,54	225,50	-0,66	32,25
Avaliação	13,38	17,50	19,00	24,00	-0,82	5,63

**M' Boi Mirim****Vila Prudente / Sapopemba**

Variável Variável	Nº de Nº de Observações	Média Média	Desvio Desvio Padrão	Variância Variância	Mínimo Mínimo	Coef de Coef de Var
Prova São Paulo	32,00	206,29	34,86 6,89	1.275472	131046	0,45
Avaliação Variável	3Q10 Q1	Mediana Mediana	Q3 5,88 Q3	Máximo Máximo	Assimetria Assimetria	Q3 - Q34 Q3 - Q1
Prova São Paulo	185,28	206,04	225,25	229,09	-0,66	50,00
Avaliação	13,75	18,25	22,38	25,00	-0,81	8,63

**Tabela A.2 – Tabela do número de alunos regulares por subprefeitura.**

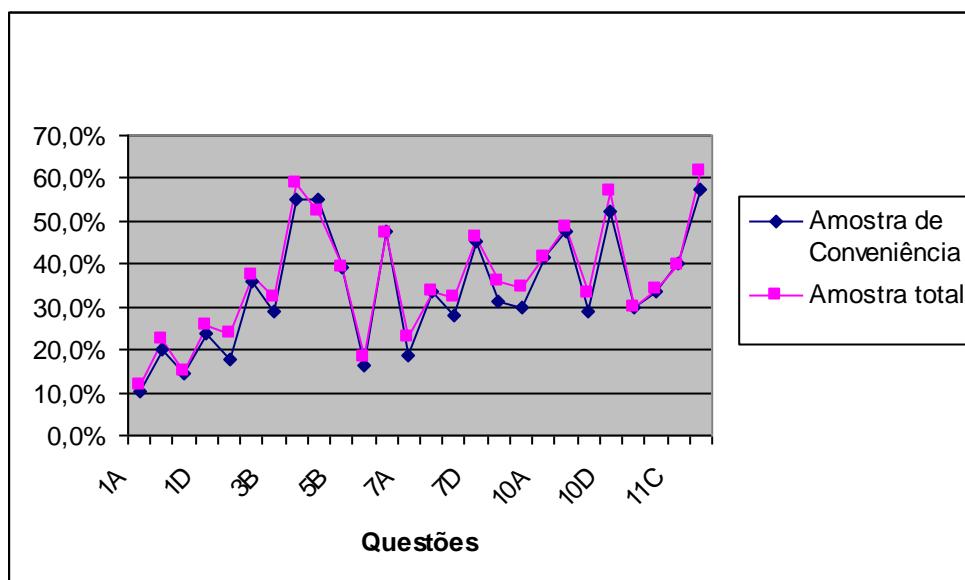
<b>Subprefeitura</b>	<b>Número de alunos regulares</b>
Aricanduva	49
Butantã	16
Campo Limpo	55
Casa Verde / Cachoeirinha	20
Cidade Tiradentes	6
Freguesia / Brasilândia	31
Ipiranga	25
Itaim Paulista	43
Itaquera	30
Jabaquara	73
M' Boi Mirim	33
São Mateus	31
São Miguel	26
Socorro	66
Vila Prudente / Sapopemba	32

**Tabela A.3 – Tabela do desvio padrão dos efeitos aleatórios e da correlação entre os mesmos.**

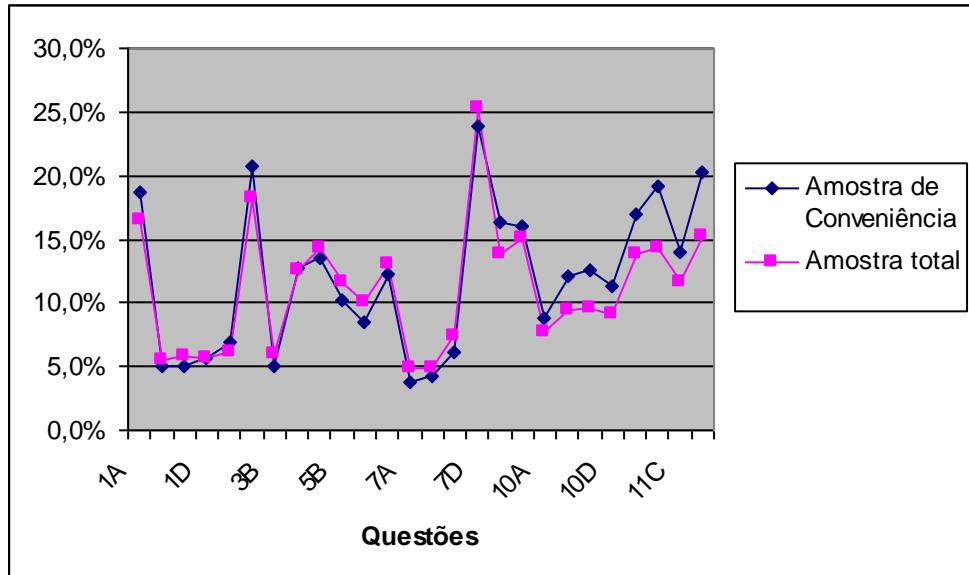
	<b>Desvio Padrão</b>	<b>Correlação</b>
<b>Intercepto</b>	3,287	-0,954
<b>Escore</b>	0,043	
<b>Resíduo</b>	4,439	

## **Gráficos - Apêndice B**

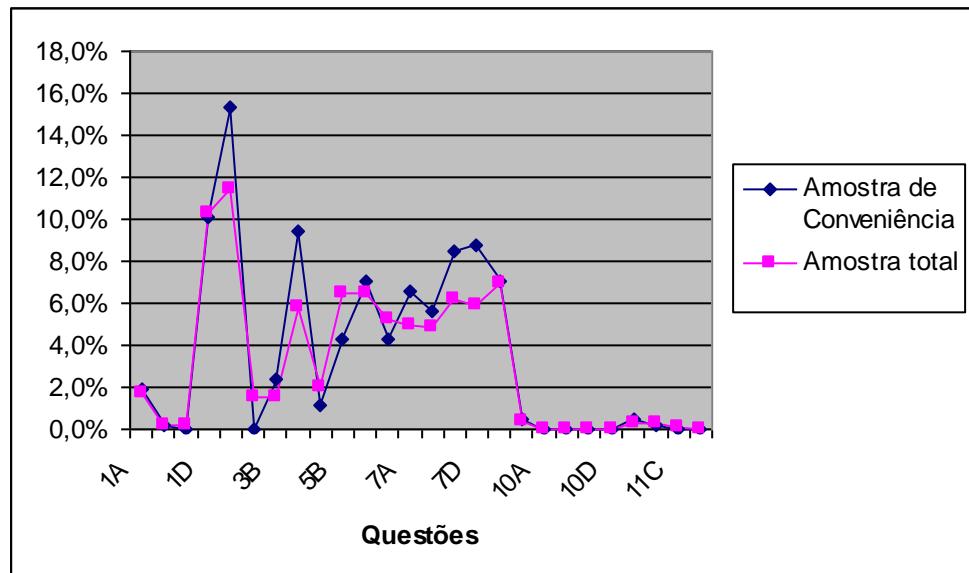
**Gráfico B.1 – Porcentagem de erros para as 2 amostras.**



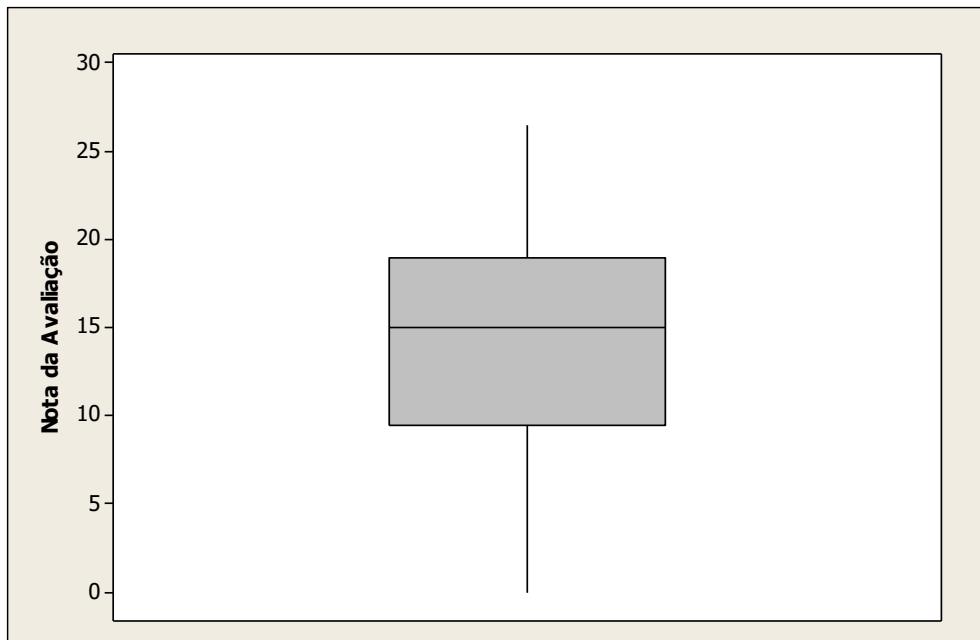
**Gráfico B.2 – Porcentagem de respostas em branco para as 2 amostras.**



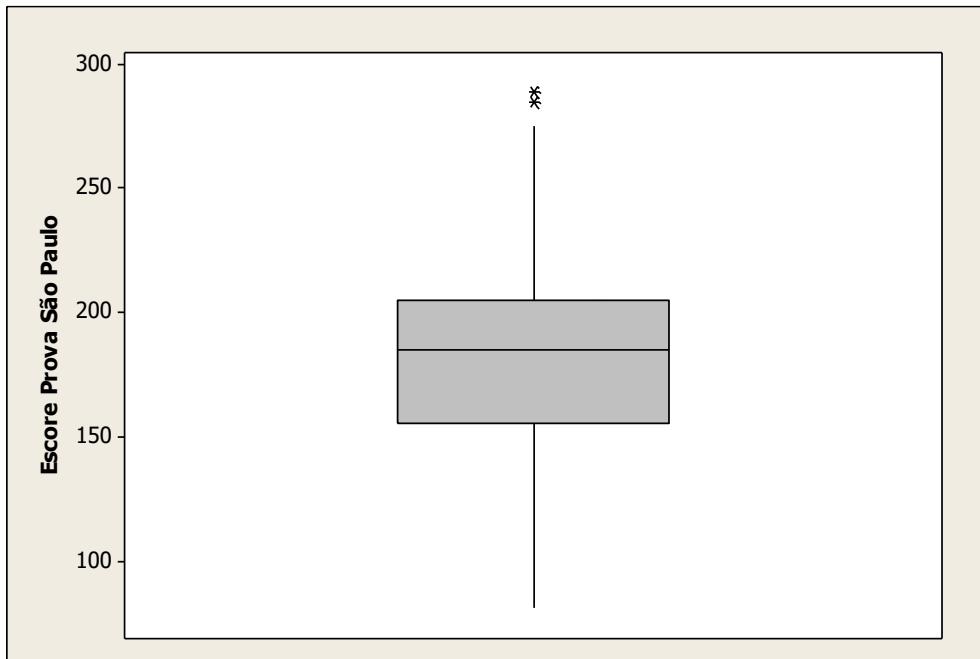
**Gráfico B.3 – Porcentagem de acertos parciais para as 2 amostras.**



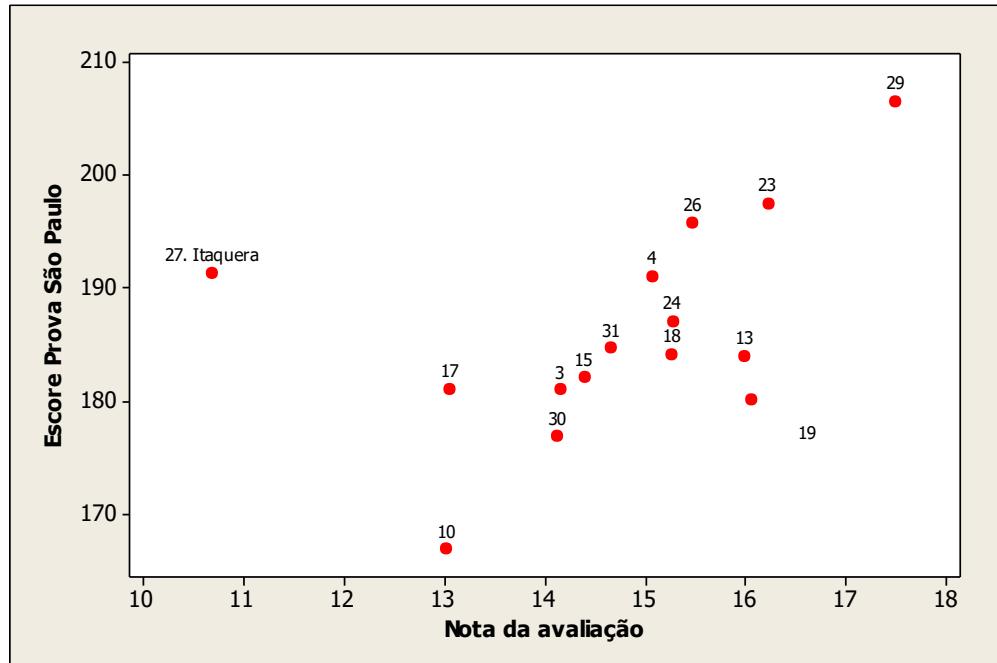
**Gráfico B.4 – Boxplot das notas da Avaliação Diagnóstico.**



**Gráfico B.5** – Boxplot do escore da Prova São Paulo.

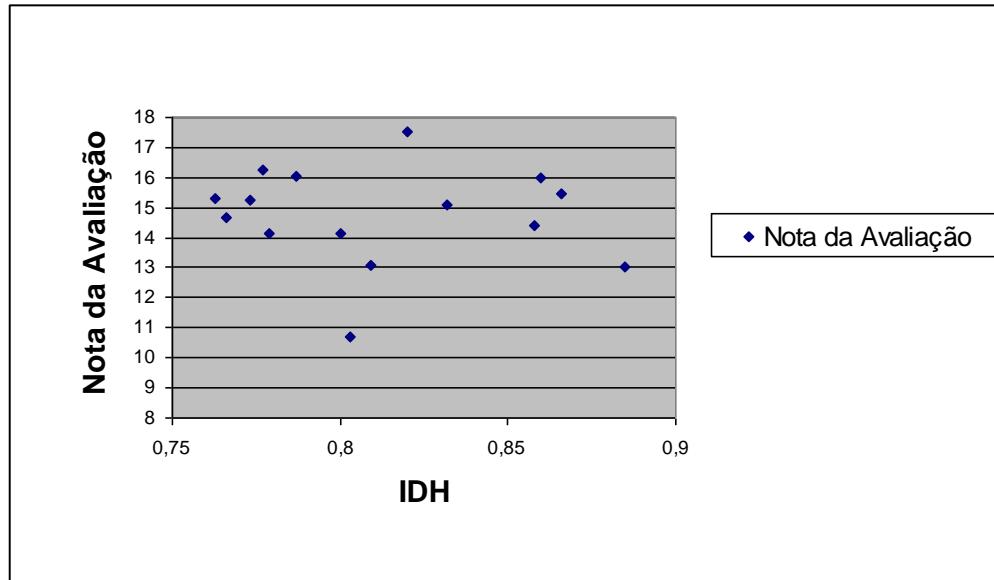


**Gráfico B.6** – Gráfico de dispersão do Escore Médio da Prova São Paulo por nota média na Avaliação Diagnóstico em cada subprefeitura (vide tabela na página seguinte).

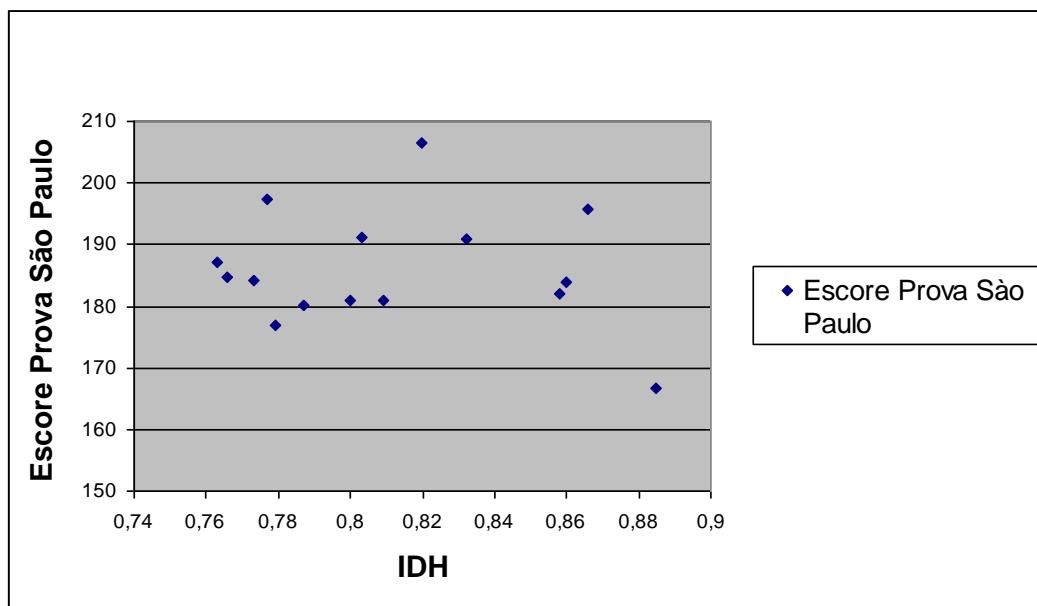


Subprefeitura	Avaliação	Prova São Paulo
01. Perus	-	-
02. Pirituba	-	-
<b>03. Freguesia do Ó</b>	<b>14,16</b>	<b>180,92</b>
<b>04. Casa Verde</b>	<b>15,08</b>	<b>191,00</b>
05. Santana/Tucuruvi	-	-
06. Jaçanã/Tremembé	-	-
07. Vila Maria/Vila Guilherme	-	-
08. Lapa	-	-
09. Sé	-	-
<b>10. Butantã</b>	<b>13,03</b>	<b>166,79</b>
11. Pinheiros	-	-
12. Vila Mariana	-	-
<b>13. Vila Prudente</b>	<b>16,00</b>	<b>183,89</b>
14. Santo Amaro	-	-
<b>15. Jabaquara</b>	<b>14,40</b>	<b>182,04</b>
16. Cidade Ademar	-	-
<b>17. Campo Limpo</b>	<b>13,06</b>	<b>181,03</b>
<b>18. M'Boi Mirim</b>	<b>15,27</b>	<b>184,10</b>
<b>19. Capela Do Socorro</b>	<b>16,06</b>	<b>180,10</b>
20. Parelheiros	-	-
21. Penha	-	-
22. Engenho Matarazzo	-	-
<b>23. São Miguel</b>	<b>16,23</b>	<b>197,45</b>
<b>24. Iaim Paulista</b>	<b>15,28</b>	<b>187,00</b>
25. Mooca	-	-
<b>26. Antônio Duarte</b>	<b>15,48</b>	<b>195,80</b>
<b>27. Itaquera</b>	<b>10,70</b>	<b>191,29</b>
28. Guiaianazes	-	-
<b>29. Vila Prudente</b>	<b>17,50</b>	<b>206,49</b>
<b>30. São Mateus</b>	<b>14,13</b>	<b>176,81</b>
<b>31. Cidade Tiradentes</b>	<b>14,67</b>	<b>184,61</b>

**Gráfico B.7 – Gráfico de dispersão da média das notas dos alunos na Avaliação Diagnóstico por IDH em cada subprefeitura.**

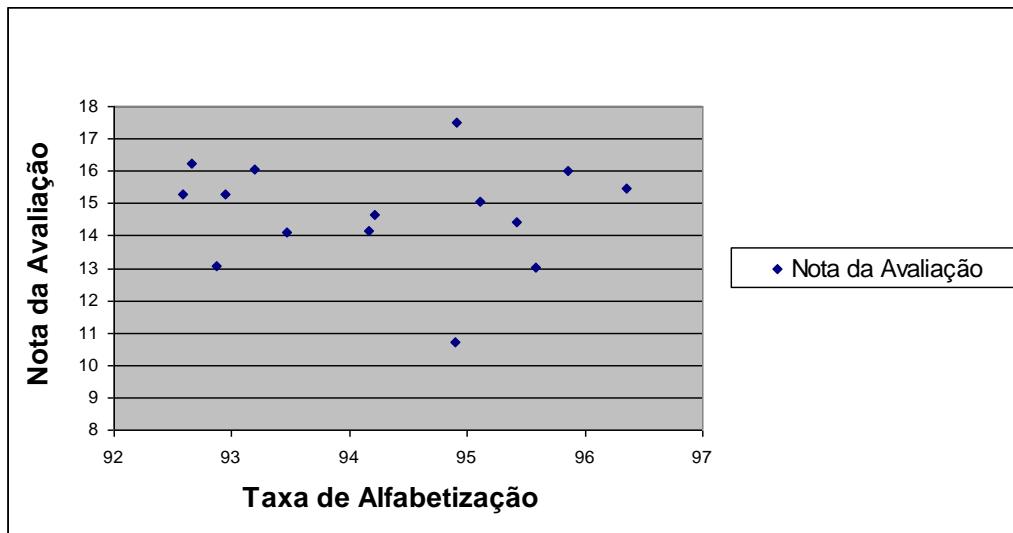


**Gráfico B.8 – Gráfico de dispersão da média dos escores dos alunos na Prova São Paulo por IDH em cada subprefeitura.**

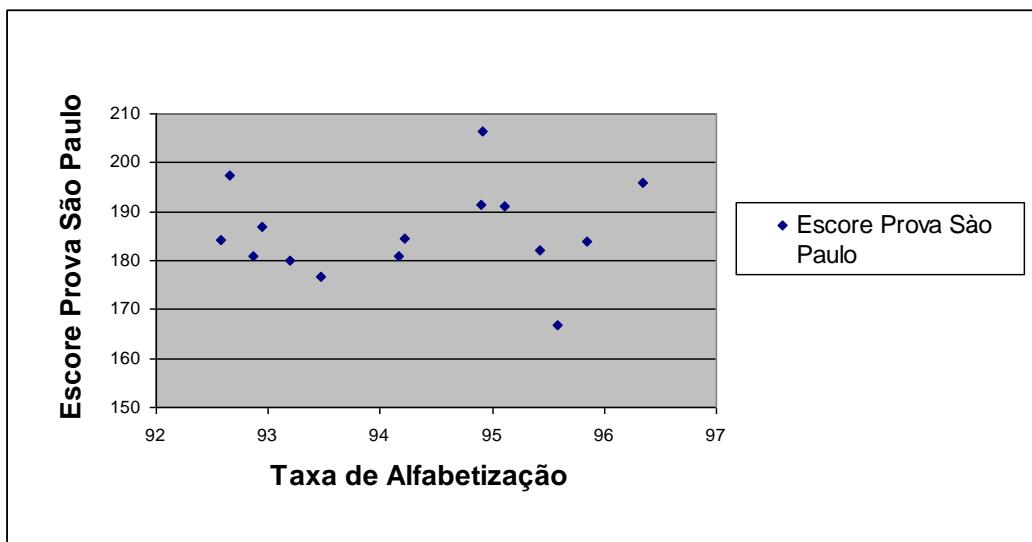


**Gráfico B.9 – Gráfico de dispersão da média das notas dos alunos na Avaliação**

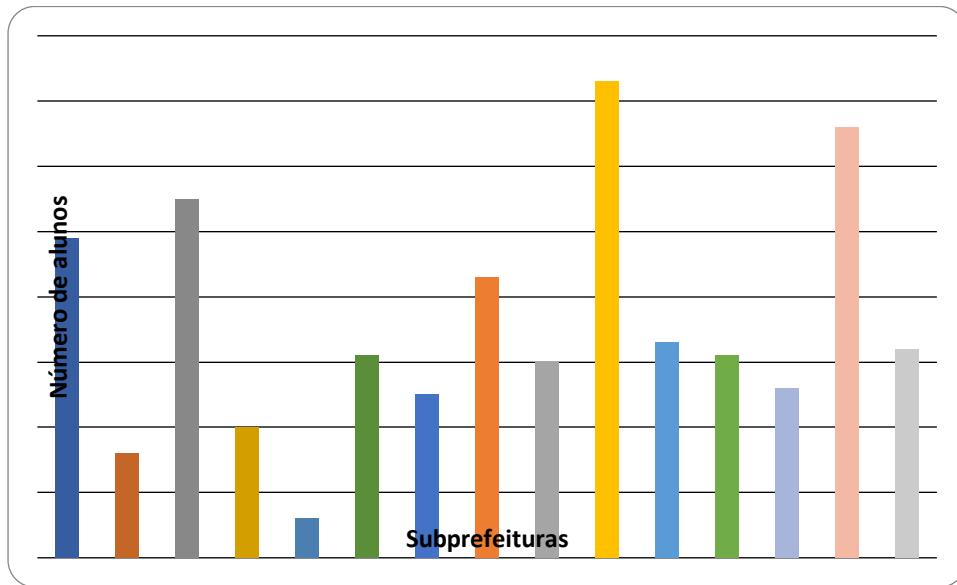
Diagnóstico por Taxa de Alfabetização em cada subprefeitura.



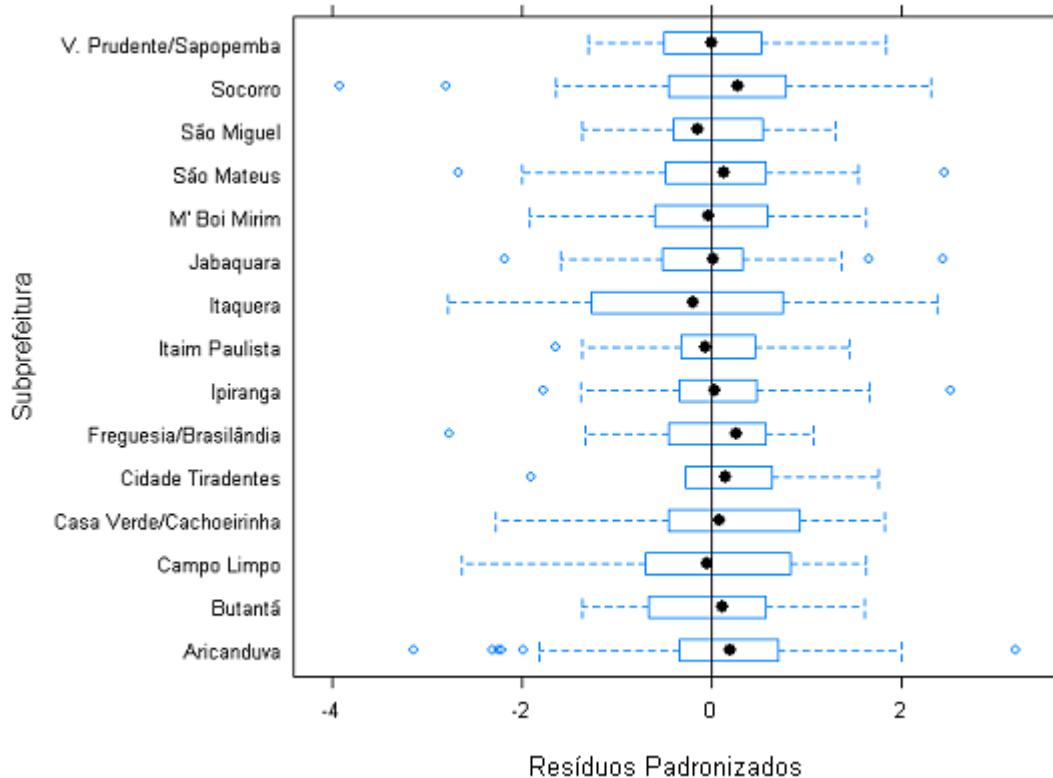
**Gráfico B.10** – Gráfico de dispersão da média dos escores dos alunos na Prova São Paulo por Taxa de Alfabetização em cada subprefeitura.



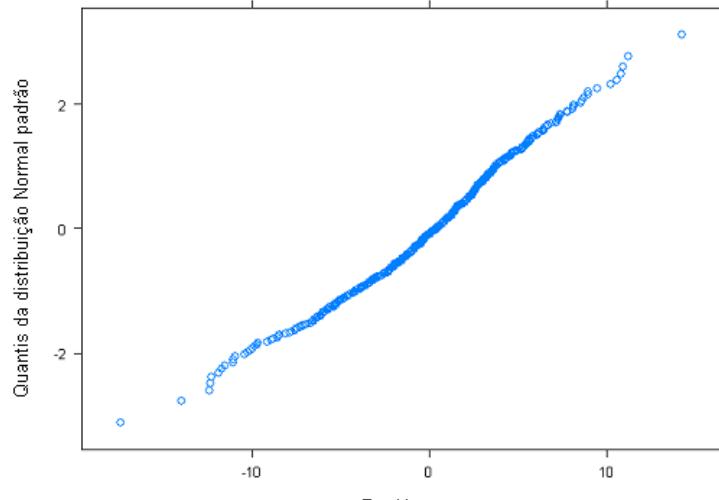
**Gráfico B.11 – Distribuição do número de alunos regulares por subprefeitura.**



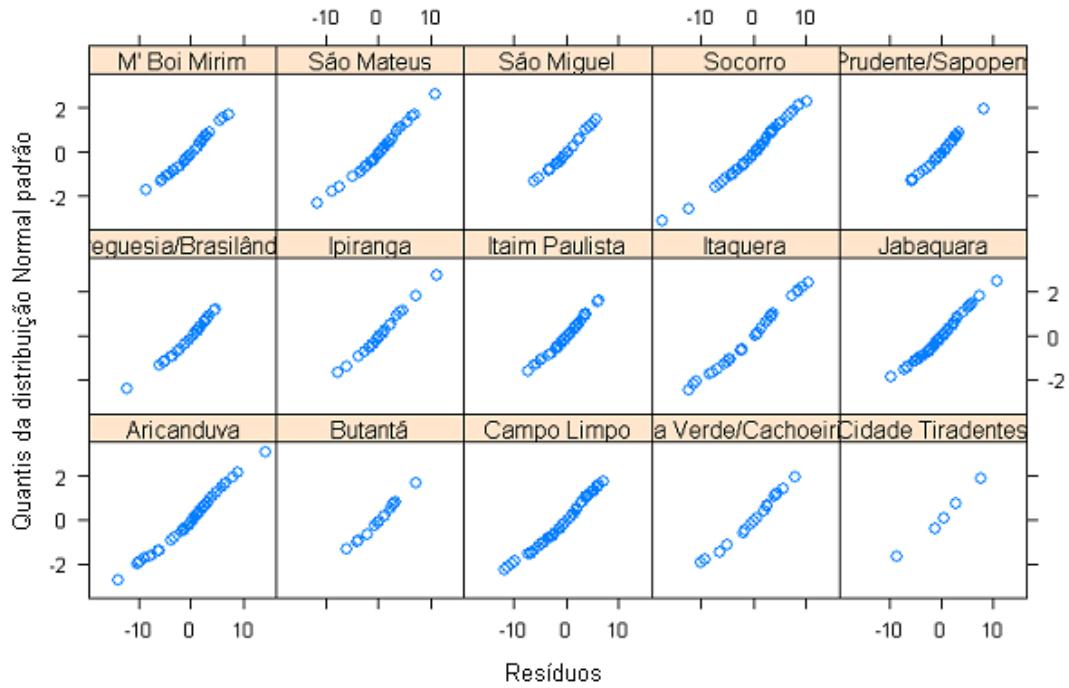
**Gráfico B.12 – Boxplots dos resíduos padronizados por subprefeitura.**



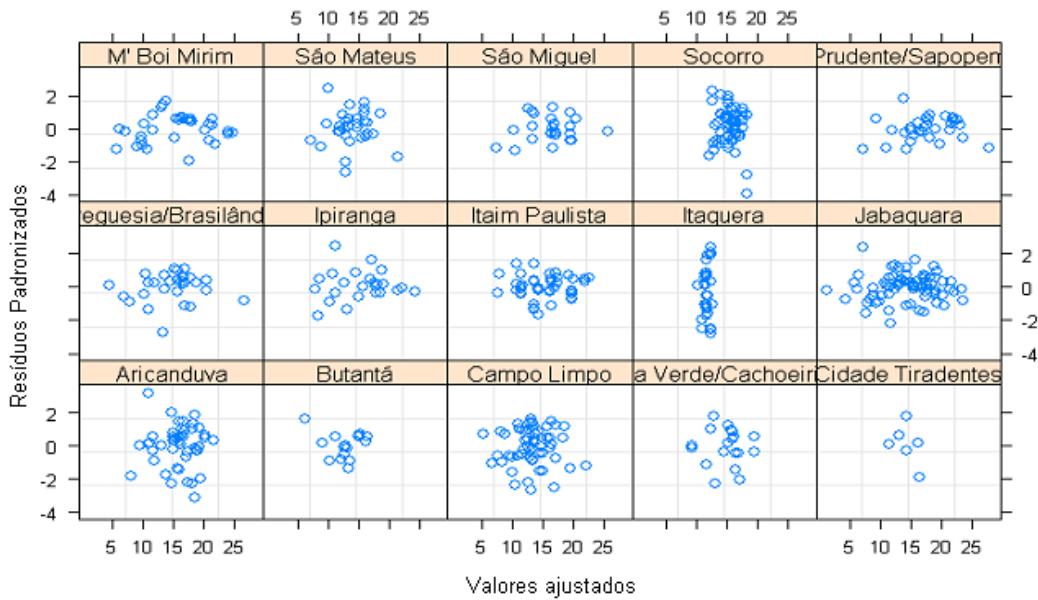
**Gráfico B.13 – Gráfico QQ-plot dos resíduos do modelo.**



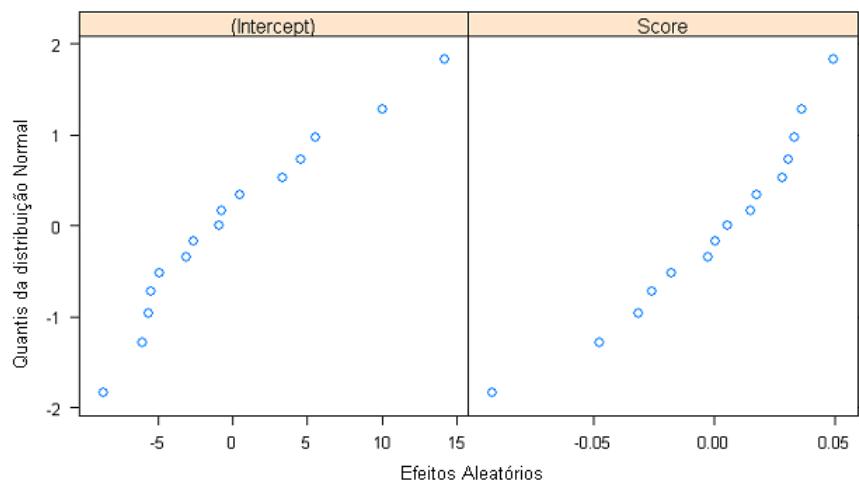
**Gráfico B.14**



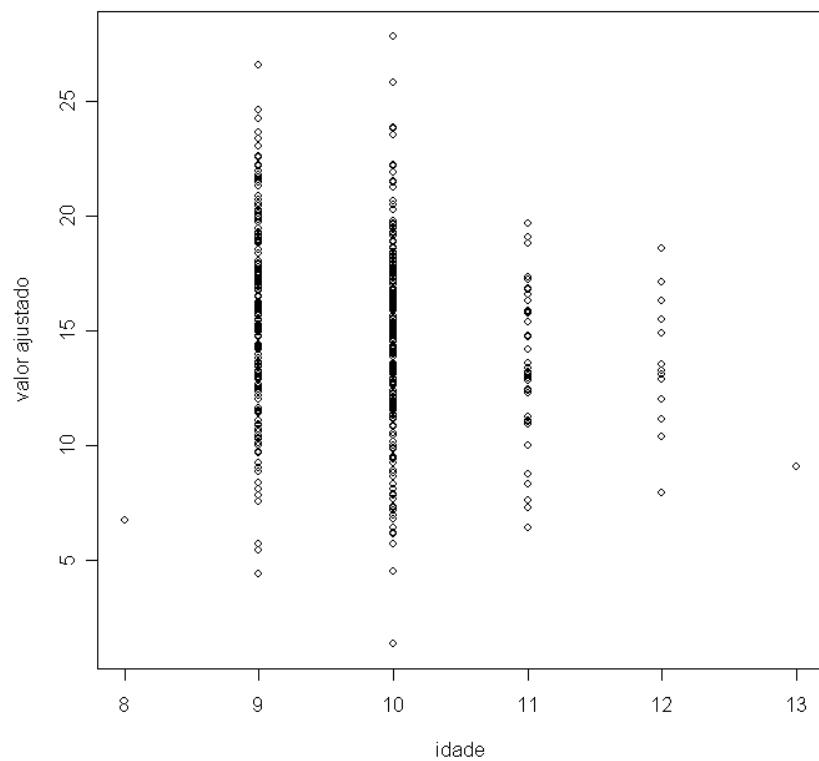
**Gráfico B.15 – Gráfico dos resíduos padronizados por subprefeitura.**



**Gráfico B.16 - Gráfico QQ – plot dos efeitos aleatórios do intercepto e da variável escore.**

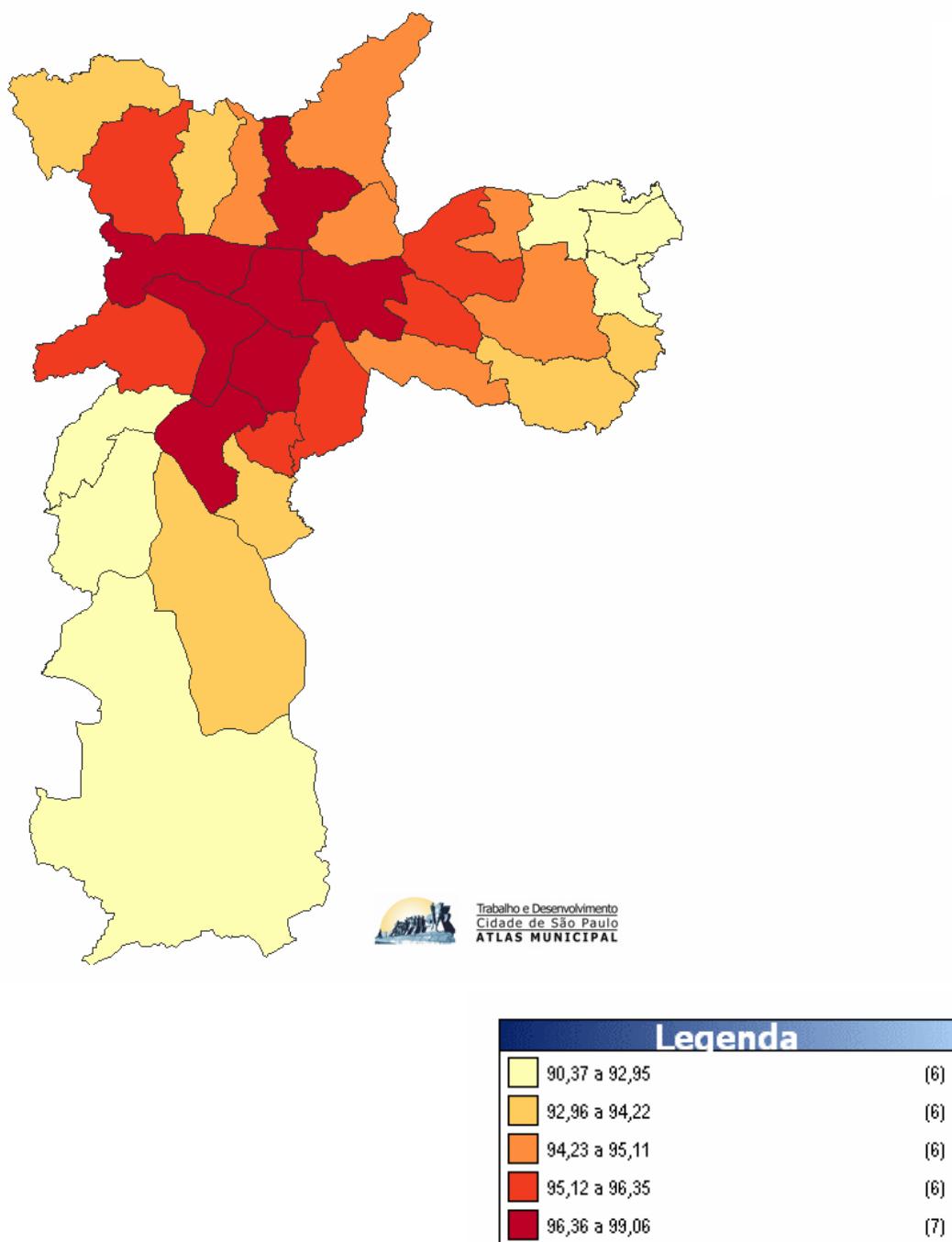


**Gráfico B.17 – Gráfico de dispersão da nota ajustada pelo modelo versus a idade.**



## Mapas – Apêndice C

**Mapa C.1** – Mapa temático da Taxa de alfabetização de todas as subprefeituras do município de São Paulo.

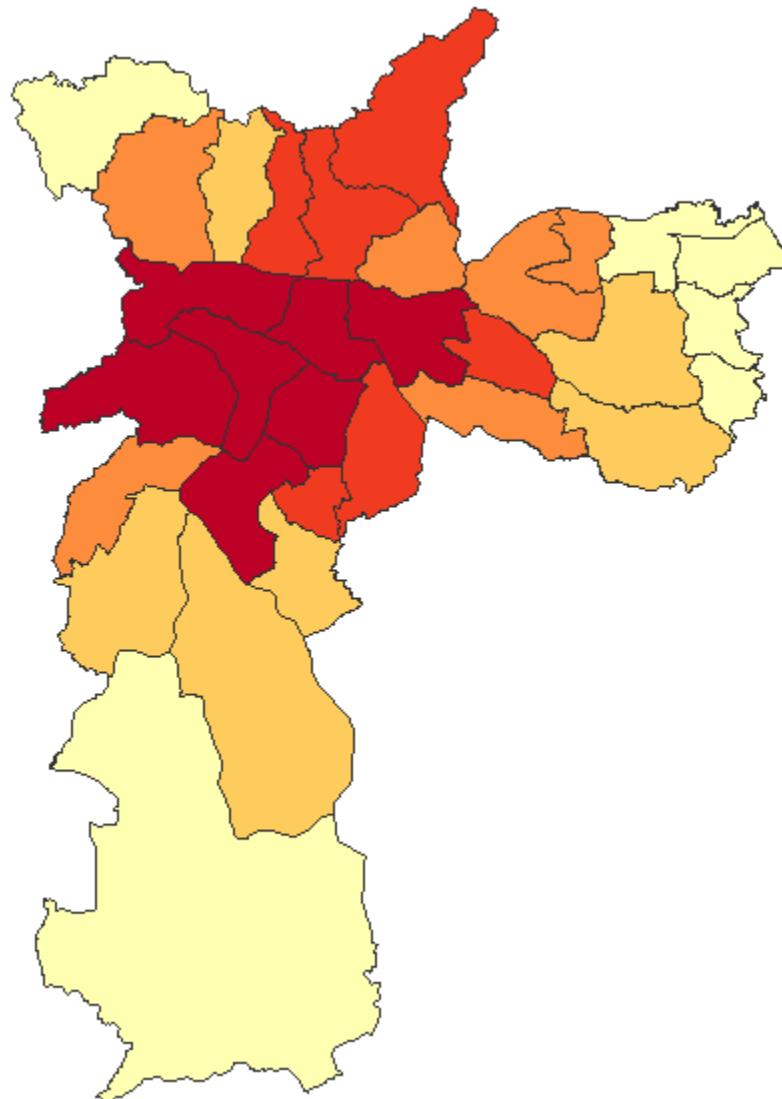


**Mapa C.2** – Mapa de indicação das 15 subprefeituras do município do São Paulo pertencentes no estudo (em vermelho).



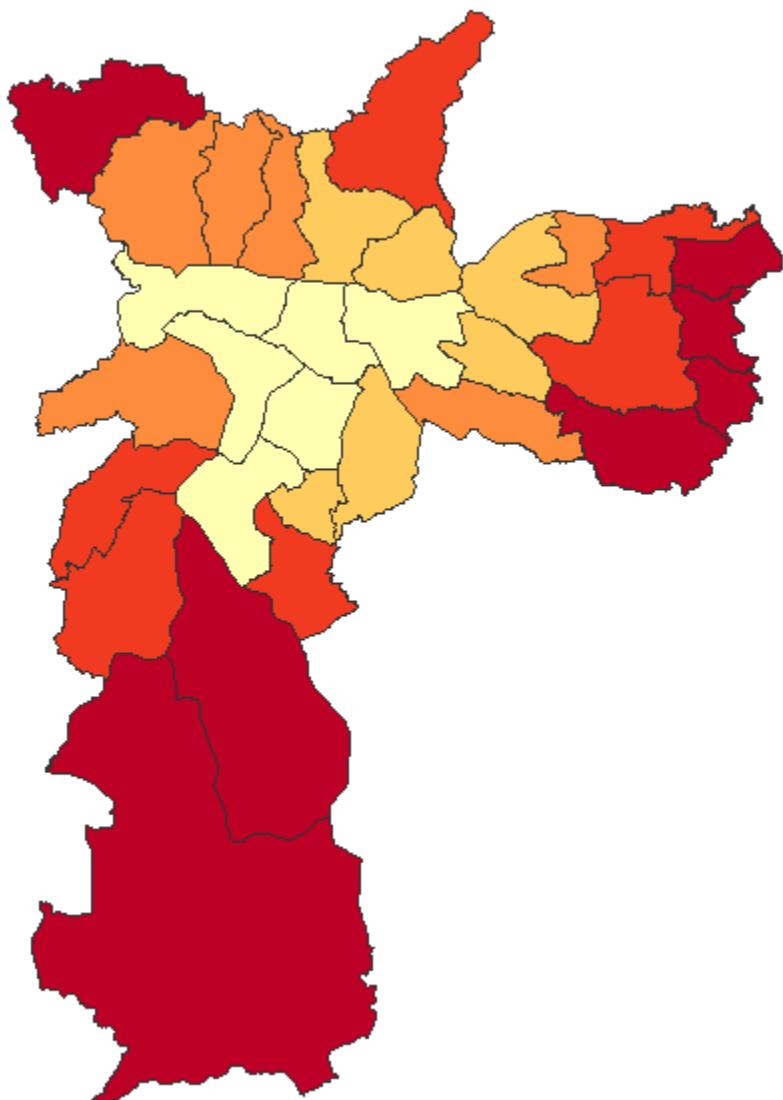
Subprefeitura	Avaliação	Prova São Paulo
01. Perus	-	-
02. Pirituba	-	-
<b>03. Freguesia do Ó</b>	<b>14,16</b>	<b>180,92</b>
<b>04. Casa Verde</b>	<b>15,08</b>	<b>191,00</b>
05. Santana/Tucuruvi	-	-
06. Jaçanã/Tremembé	-	-
07. Vila Maria/Vila Guilherme	-	-
08. Lapa	-	-
09. Sé	-	-
<b>10. Butantã</b>	<b>13,03</b>	<b>166,79</b>
11. Pinheiros	-	-
12. Vila Mariana	-	-
<b>13. Ipiranga</b>	<b>16,00</b>	<b>183,89</b>
14. Santo Amaro	-	-
<b>15. Jabaquara</b>	<b>14,40</b>	<b>182,04</b>
16. Cidade Ademar	-	-
<b>17. Campo Limpo</b>	<b>13,06</b>	<b>181,03</b>
<b>18. M'Boi Mirim</b>	<b>15,27</b>	<b>184,10</b>
<b>19. Capela Do Socorro</b>	<b>16,06</b>	<b>180,10</b>
20. Parelheiros	-	-
21. Penha	-	-
22. Ermo/Ino Matarazzo	-	-
<b>23. São Miguel</b>	<b>16,23</b>	<b>197,45</b>
<b>24. Iaim Paulista</b>	<b>15,28</b>	<b>187,00</b>
25. Mooca	-	-
<b>26. Aricanduva</b>	<b>15,48</b>	<b>195,80</b>
<b>27. Itaquera</b>	<b>10,70</b>	<b>191,29</b>
28. Guaiamazés	-	-
<b>29. Vila Prudente</b>	<b>17,50</b>	<b>206,49</b>
<b>30. São Mateus</b>	<b>14,13</b>	<b>176,81</b>
<b>31. Cidade Tiradentes</b>	<b>14,67</b>	<b>184,61</b>

**Mapa C.3 – Mapa temático da Renda Per Capita de todas as subprefeituras do município de São Paulo.**



<b>Legenda</b>	
218,93 a 269,97	(6)
269,98 a 367,89	(6)
367,90 a 478,76	(6)
478,77 a 894,77	(6)
894,78 a 2.405,17	(7)

**Mapa C.4 – Mapa temático de número de moradores por domicílio de todas as subprefeituras do município do São Paulo.**



<b>Legenda</b>	
2,55 a 3,20	(6)
3,21 a 3,45	(6)
3,46 a 3,64	(6)
3,65 a 3,75	(6)
3,76 a 3,93	(7)

## **Apêndice D - Avaliação Diagnóstico**

**Projeto de Avaliação e Pesquisa: Investigando as dificuldades em Matemática no Ensino Fundamental da Rede Municipal de São Paulo (Processo FAPESP 2006/50608-1)**  
 Secretaria Municipal de Educação de São Paulo / Faculdade de Educação da Universidade de São Paulo / Fundação de Amparo à Pesquisa do Estado de São Paulo  
 Coordenador do projeto: Prof. Dr. Vinícius de Macedo Santos

N.º \_\_\_\_\_ Classe \_\_\_\_\_ Idade \_\_\_\_\_ anos Sexo: ( ) Masculino ( ) Feminino

**1. OBSERVE OS NÚMEROS NO QUADRO.**

1 099	588	1 900
998	134	300
		989

- A) FAÇA UM CÍRCULO NO NÚMERO NOVECENTOS E OITENTA E NOVE.  
 B) QUAL É O NÚMERO MAIOR? \_\_\_\_\_  
 C) QUAL É O NÚMERO MENOR? \_\_\_\_\_  
 D) PREENCHA OS ESPAÇOS ABAIXO COM OS NÚMEROS DO QUADRO, DO MENOR PARA O MAIOR:

_____	_____	_____	_____	_____	_____	_____
-------	-------	-------	-------	-------	-------	-------

**2. MARINA, JOÃO, ANDRÉ E PAULO COLECCIONAM FIGURINHAS.**

VEJA A QUANTIDADE DE FIGURINHAS QUE CADA UM POSSUI:

JOÃO TEM 223 FIGURINHAS

MARINA TEM 349 FIGURINHAS

ANDRÉ TEM 110 FIGURINHAS

PAULO TEM 500 FIGURINHAS

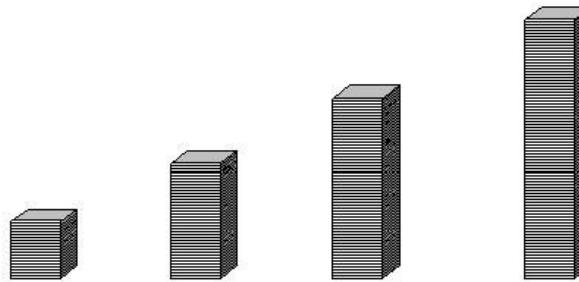
LIGUE O NOME DE CADA UM DELES AO SEU BOLO DE FIGURINHAS:

MARINA

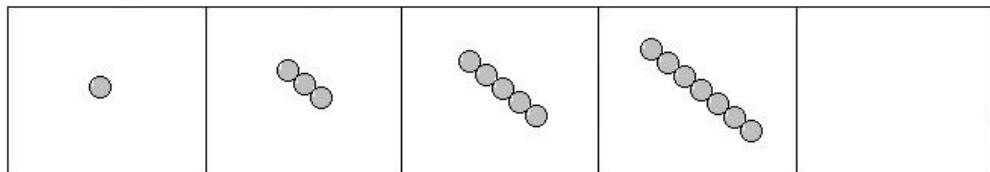
ANDRÉ

PAULO

JOÃO



**3.** A) DESENHE O PRÓXIMO GRUPO DE BOLINHAS DE ACORDO COM A SEQÜÊNCIA:



B) COMPLETE AS SEQÜÊNCIAS

- $0, 3, 6, 9, 12, \underline{\quad}, \underline{\quad}, \underline{\quad}$

- $206, 200, 194, 188, \underline{\quad}, \underline{\quad}, \underline{\quad}$

**4.** UM PROFESSOR DE EDUCAÇÃO FÍSICA RESOLVEU FAZER UMA PESQUISA SOBRE OS ESPORTES PREFERIDOS DOS ALUNOS DO 4.º ANO DE SUA ESCOLA E PEDIU AJUDA A UM ALUNO, PAULO, PARA FAZER AS ANOTAÇÕES.

VEJA COMO PAULO REGISTROU AS INFORMAÇÕES.

BASQUETE          |

FUTEBOL            |

HANDEBOL         |

VÔLEI            |

AJUDE PAULO A COMPLETAR A TABELA COM O NÚMERO DE ALUNOS QUE PRATICAM CADA ESPORTE.

**TABELA. ESPORTE PREFERIDO DOS ALUNOS**

TIPO DE ESPORTE	NÚMERO DE ALUNOS

**5.**

A)

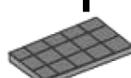


QUANTAS LATAS DE ÓLEO PODEM SER COMPRADAS COM R\$ 10,00 SEM SOBRAR TROCO?

B)

**CHOCOLATE**  
R\$ 3,00

TENHO 5 MOEDAS DE R\$ 0,50. POSSO COMPRAR O CHOCOLATE?  
POR QUÊ?



**6. RESOLVA OS PROBLEMAS E ANOTE COMO ENCONTROU OS RESULTADOS.**

A) MARCOS E SUA IRMÃ ALICE ESTÃO JOGANDO *VIDEOGAME*. MARCOS FEZ 138 PONTOS NA 1<sup>a</sup> FASE DO JOGO E 212 PONTOS NA 2<sup>a</sup> FASE. QUANTOS PONTOS MARCOS JÁ ACUMULOU?

B) NA 4<sup>a</sup> FASE, ALICE ACUMULOU 513 PONTOS, ENQUANTO MARCOS ACUMULOU 465. QUANTOS PONTOS MARCOS PRECISARIA FAZER PARA EMPATAR COM A SUA IRMÃ?

4

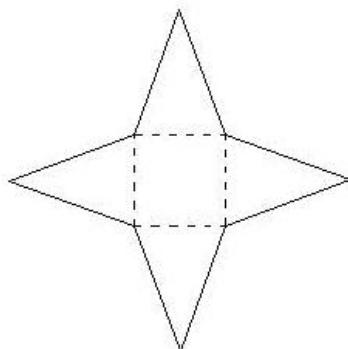
**7. EFETUE:**

A) 
$$\begin{array}{r} 1345 \\ + 738 \\ \hline \end{array}$$

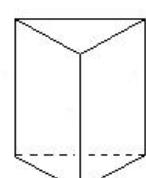
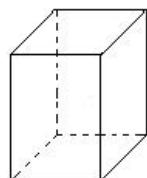
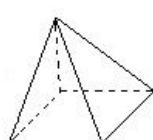
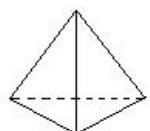
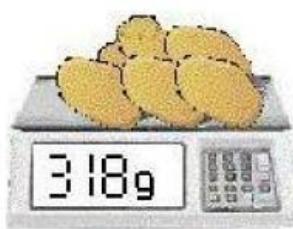
B) 
$$\begin{array}{r} 926 \\ - 384 \\ \hline \end{array}$$

C) 
$$\begin{array}{r} 132 \\ \times 4 \\ \hline \end{array}$$

D) 
$$\begin{array}{r} 87 \\ \boxed{3} \end{array}$$

**8. OBSERVE A FIGURA DE UMA CAIXA DESMONTADA.**

FAÇA UM "X" NA CAIXA QUE ELA REPRESENTA.

**9.**

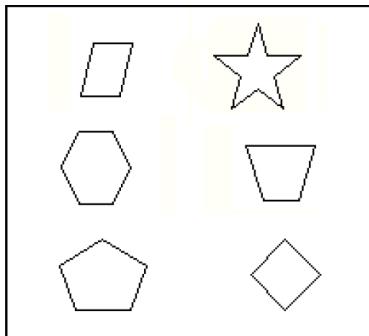
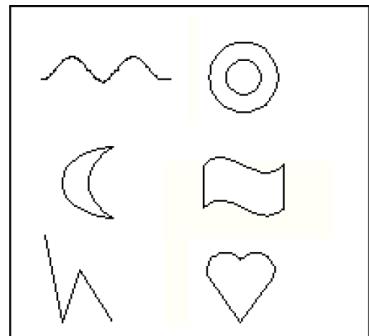
DONA MARIA LEVOU AS BATATAS DAS DUAS BALANÇAS. QUANTO ELA LEVOU?

- MENOS DE MEIO QUILO
- MAIS DE MEIO QUILO
- MEIO QUILO

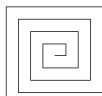
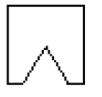
5

**10.** OBSERVE AS FIGURAS, PENSE NAS COISAS REAIS QUE ELAS REPRESENTAM E MARQUE COM UM “X” A MEDIDA MAIS PROVÁVEL EM CADA CASO.

- |  |                                    |                                     |
|--|------------------------------------|-------------------------------------|
| A)  | ( <input type="checkbox"/> ) 70 cm | ( <input type="checkbox"/> ) 7 cm   |
| B)  | ( <input type="checkbox"/> ) 10 ℥  | ( <input type="checkbox"/> ) 100 ℥  |
| C)  | ( <input type="checkbox"/> ) 20 kg | ( <input type="checkbox"/> ) 200 kg |
| D)  | ( <input type="checkbox"/> ) 40 m  | ( <input type="checkbox"/> ) 4 m    |

**11.****GRUPO A****GRUPO B**

EM QUAL GRUPO VOCÊ COLOCARIA CADA UMA DAS SEGUINTE FIGURAS?

GRUPO A (  )GRUPO B (  )GRUPO A (  )GRUPO B (  )GRUPO A (  )GRUPO B (  )GRUPO A (  )GRUPO B (  )

## **Modelos – Apêndice E**

**Modelo E.1** - O nível 1 (aluno) do modelo inicial é apresentado a seguir :

$$Y_{ij} = \beta_{0j} + \beta_{1j}(X_{1ij} - 100) + \beta_{2j}(X_{2ij} - 10) + \beta_{3j}S_{ij} + \sum_{k=1}^{10} \delta_{kj}W_{kj} + e_{ij}$$

Onde:

$j = 1, 2, \dots, 15$  subprefeituras,

$i = 1, 2, \dots$  alunos da subprefeitura  $j$ ,

$Y_{ij}$  : Nota da Avaliação Diagnóstico do aluno  $i$  da subprefeitura  $j$

$X_{1ij}$  : Escore da Prova São Paulo do aluno  $i$  da subprefeitura  $j$ ;

$X_{2ij}$  : Idade do aluno  $i$  da subprefeitura  $j$ ;

$W_{kj}$  : Valor assumido pela variável socioeconômica  $k$  da subprefeitura  $j$ ;

$S_{ij}$  : Variável indicadora do Sexo do aluno  $i$  da escola  $j$ . 0 para meninos e 1 para meninas;

$e_{ij}$  : Erro aleatório associado ao aluno  $i$  da escola  $j$ .

Ao ajustar o modelo inicial no software R, apenas as variáveis explicativas idade e escore mostraram ser significativas para explicar a variável nota (nível descriptivo menor que 0,05), as variáveis socioeconômicas e a variável sexo foram excluídas do modelo, apresentando um nível descriptivo alto (consideravelmente maior que 0,05).

Assim, o nível 1 do modelo final foi o seguinte :

$$Y_{ij} = \beta_{0j} + \beta_{1j}(X_{1ij} - 100) + \beta_{2j}(X_{2ij} - 10) + e_{ij}$$

Onde:

$Y_{ij}$  : Nota da Avaliação Diagnóstico do aluno  $i$  da subprefeitura  $j$

$X_{1ij}$  : Escore da Prova São Paulo do aluno  $i$  da subprefeitura  $j$ ;

$X_{2ij}$  : Idade do aluno  $i$  da subprefeitura  $j$ ;

$\beta_{0j}$  : Nota esperada da Avaliação Diagnóstico da subprefeitura  $j$  para alunos

com 10 anos de idade e que obtiveram escore 100 na Prova São Paulo;

$\beta_{1j}$ : Variação esperada na nota da Avaliação Diagnóstico da subprefeitura  $j$  por unidade de variação do escore obtido na Prova São Paulo, fixando a idade do aluno;

$\beta_{2j}$ : Variação esperada na nota da Avaliação Diagnóstico da subprefeitura  $j$  por unidade de variação da idade dos alunos, fixando o escore obtido na Prova São Paulo

$e_{ij}$ : Erro aleatório associado ao aluno  $i$  da escola  $j$

Para o nível 2 (subprefeitura) formulamos o modelo apresentado a seguir:

$$\beta_{0j} = \gamma_{00} + u_{0j}$$

$$\beta_{1j} = \gamma_{10} + u_{1j}$$

Em que:

$\gamma_{00}$ : nota esperada dos alunos que obtiveram escore 100 na Prova São Paulo e com 10 anos na população de subprefeituras

$\gamma_{10}$ : valor esperado das inclinações referentes à variável Escore na população de subprefeituras

$u_{0j}$ : efeito aleatório da subprefeitura  $j$  no intercepto

$u_{1j}$ : efeito aleatório da subprefeitura  $j$  na inclinação

O modelo é ajustado sob as seguintes suposições:

- $e_{ij}$  seguem uma distribuição normal padrão, com média 0 e variância  $\sigma^2$  constante
- Os erros são independentes para alunos da mesma subprefeitura
- $u_{0j} \sim N(0, \tau_{00})$  i.i.d.  
 $u_{1j} \sim N(0, \tau_{11})$  i.i.d.  
 $Cov(u_{0j}, u_{1j}) = \tau_{01}$
- $\beta_{0j} \sim N(\gamma_{00}, \tau_{00})$  i.i.d.  
 $\beta_{1j} \sim N(\gamma_{10}, \tau_{11})$  i.i.d.  
 $Cov(\beta_{0j}, \beta_{1j}) = \tau_{01}$

Onde:

$\tau_{00}$  : variância populacional dos interceptos;

$\tau_{11}$  : variância populacional das inclinações;

$\tau_{01}$  : covariância entre  $\beta_{0j}$  e  $\beta_{1j}$

- Os erros e efeitos aleatórios são independentes.

$$\text{Corr}(Y_{ij}, Y_{i'j}) = \frac{\tau_{00} + \tau_{11} \times X_{1ij} \times X_{1i'j} + \tau_{01} \times (X_{1ij} + X_{1i'j})}{\sqrt{\sigma^2 + \tau_{00} + \tau_{11} \times X_{1ij}^2 + 2 \times \tau_{01} \times X_{1ij}} \times \sqrt{\sigma^2 + \tau_{00} + \tau_{11} \times X_{1i'j}^2 + 2 \times \tau_{01} \times X_{1i'j}}}$$

### Modelo E.2 – Modelo final

$$Y_{ij} = 6,365 + 0,096(X_{1ij} - 100) - 0,694(X_{2ij} - 10) + e_{ij}$$

### Modelo E.3 – Modelo final

$$Y_{ij} / 27 = 0,139 + 1,771 \times X_{1ij} / 500 - 0,026 \times X_{2ij} + e_{ij}$$

### Teste E.4 – Teste Sobre as componentes de variância

$$H_0: \tau_{11} = 0, \tau_{01} = 0$$

$$H_1: \tau_{11} > 0, \tau_{01} \neq 0$$

$$LR \sim 0.5\chi_1^2 + 0.5\chi_2^2$$

Para o modelo considerando as Subprefeituras:

$$LR_{MS} = 26,1977 \text{ p - valor } \approx 0$$

Logo, existem os efeitos aleatórios na inclinação dado que existem efeitos aleatórios nos interceptos.