

## ARTIGO

# Pequeno manual para aplicação do modelo Profluxo

Flávia Vitor Longo dos Santos<sup>1</sup> 

### RESUMO

O modelo Profluxo, elaborado no final dos anos de 1980 para resolver o problema da falta de informação nos recenseamentos escolares, visava aprimorar os modelos de análise de fluxo escolar com base na fonte de dados demográficas. Desde então a ferramenta foi aplicada em diversos trabalhos, especialmente aqueles interessados no diagnóstico e na correção do fluxo escolar. Este texto, orientado para profissionais com formação na área de ciências sociais e humanidades, traz um passo a passo didático da aplicação desse modelo.

### PALAVRAS-CHAVE

modelo Profluxo; PNAD; fluxo escolar; políticas públicas de educação; texto didático.

---

<sup>1</sup>Universidade Estadual de Campinas, Campinas, SP, Brasil.

## *A SMALL GUIDE FOR APPLICATION OF THE PROFLUXO MODEL*

### ABSTRACT

The Profluxo model, developed in the late 1980s to solve the problem of lack of information on school censuses, aimed to improve models of school flow analysis from demographic datasources. Since then, the tool has been applied in several works, especially those interested in the diagnosis and correction of school flow. This text, aimed at professionals trained in Social Sciences and Humanities, provides a didactic step-by-step of this model's application.

### KEYWORDS

Profluxo model; PNAD; grade transitions; educational public policies; didactic paper.

## *PEQUEÑO MANUAL DE APLICACIÓN DEL MODELO PROFLUXO*

### RESUMEN

El modelo Profluxo, desarrollado a fines de la década de 1980 para resolver el problema de la falta de información de los censos escolares, tenía como objetivo mejorar los modelos de análisis del flujo escolar a partir de una fuente de datos demográficos. Desde entonces, la herramienta se ha aplicado en varios trabajos, especialmente aquellos interesados en el diagnóstico y la corrección del flujo escolar. Este texto, dirigido a profesionales formados en el área de las ciencias sociales y las humanidades, presenta una didáctica paso a paso de la aplicación del modelo.

### PALABRAS CLAVE

modelo Profluxo; PNAD; flujo escolar; políticas de educación pública; texto didáctico.

## INTRODUÇÃO

Este texto procura expor o passo a passo da elaboração e algumas das possibilidades de análise dos resultados do modelo Profluxo (Fletcher e Ribeiro, 1989). Por se tratar de um modelo criado há cerca de três décadas, uma série de estudos já utilizou essa ferramenta (Ribeiro, 1991; Costa, 1993; Cunha e Ascama, 2000; Santos, 2006; Tavares, Faria e Lima, 2012), inclusive para fins didáticos (Golgher, 2004). O objetivo aqui, no entanto, é dialogar com estudantes de graduação, professores, pesquisadores e técnicos de secretarias públicas que não estejam familiarizados com notações matemáticas. Contudo, algum conhecimento sobre organização de banco de dados e manipulação de variáveis (objetos que guardam determinados tipos de dados) é desejável.

Por que retomar um modelo criado há tanto tempo? Um dos pontos fortes do Profluxo é indicar, como seu próprio nome sugere, o fluxo de todas as pessoas, e não somente de estudantes, ao longo de um sistema escolar seriado. No caso do Brasil, pode-se recuperar esse fluxo nas séries da educação básica (ensinos fundamental e médio) e do ensino superior. Portanto, essa ferramenta possibilita conhecer, ainda que de forma aproximada, a trajetória escolar de todas as pessoas que tenham alguma vez ingressado nas séries escolares, do ensino fundamental ao ensino superior. E por que conhecer essas trajetórias é importante? Porque podemos identificar uma série de indicadores e eventos, tais como: a idade nas quais as pessoas evadem da escola, as transições entre séries que são mais desafiantes para a permanência no sistema escolar e a proporção (estoque) de pessoas com escolaridade incompleta. Esses indicadores podem funcionar como subsídios para o planejamento de políticas públicas de educação e/ou para a elaboração de estratégias e soluções relacionados à oferta de vagas escolares, por exemplo.

Antes de entrarmos nas etapas da construção do modelo, é necessário esclarecer alguns aspectos da natureza dos dados que alimentam essa ferramenta de análise e, especialmente para aqueles cuja formação deu-se na área de Humanidades e/ou Ciências Sociais, algumas palavras sobre modelagem. Afinal, por quais motivos modelamos?

## CONSIDERAÇÕES SOBRE MODELAGEM

“Modelo” vem do latim *modulus* e pode ser traduzido como “medida” ou “padrão”. Na língua portuguesa costumamos utilizar a palavra modelo no sentido de paradigma. Paradigma, por sua vez, deriva da palavra grega *paradeigma*, advinda de *paradeiknynai*, que significa “mostrar, representar”. Nas chamadas modelagens de dados, o que se busca é encontrar um modelo explicativo que mostre ou, ainda, que represente uma dada realidade observada.

Por se tratar de uma representação, um modelo não descreve em sua inteireza a realidade. Muitas vezes essa representação é reduzida, acomodada e ajustada de forma que se consiga cumprir seu objetivo: o de representar, da melhor maneira possível, aspectos de interesse da realidade observada. Para exemplificar, vejamos a tirinha do cartunista argentino Quino e sua personagem Mafalda (Figura 1):



Fonte: Quino (2003, p. 104).

Figura 1 – Mafalda e o princípio de modelagem.

Todos os globos terrestres que conhecemos são modelos, representações de uma realidade: nosso planeta Terra. O planeta “real” só pode ser visto em sua completa magnitude por pouquíssimos seres humanos, e a maioria de nós apenas viu imagens, fotos e maquetes que representam o planeta.

Anteriormente mencionei que um modelo busca representar da melhor maneira possível algo de interesse da realidade. Essa “melhor maneira possível”, no entanto, não é única. Ela depende do objetivo pelo qual um modelo é elaborado, da qualidade e da quantidade de dados disponíveis. Em toda modelagem é preciso considerar quais informações ou detalhes podem ser excluídos sem que isso traga prejuízo para a representação que se almeja. Pensemos novamente no exemplo do globo terrestre: não seria possível inserir na maquete miniaturizada a representação de cada árvore ou montanha existentes na superfície da Terra. Menos possível ainda seria inserir cada peixe, cada pessoa ou cada pedrinha na beira dos rios. Ou seja: nos modelos, há perda de informações. Quanto mais informações puderem ser inseridas, mais acurado, mais ajustado será o modelo. Contudo, um bom modelo não necessita, obrigatoriamente, de uma infinidade de dados/informações para ser construído. Ele precisa ser minimamente capaz de representar e de explicar uma dada parte da realidade, um dado fenômeno observado. Quando vemos um globo terrestre — um objeto em formato esférico, com contornos separando continentes de oceanos, ainda que sem a informação sobre peixes e pedrinhas —, sabemos que se trata de uma miniatura do nosso planeta.

Vamos aplicar um raciocínio semelhante à modelagem de dados. Ainda que “dados” sejam entidades mais abstratas que o planeta Terra, é possível manipulá-los para criar representações da realidade. Mesmo considerando que há perdas ao se elaborar modelos, as modelagens são úteis para uma série de aspectos da vida humana e da vida em sociedade. Por exemplo, há modelos preditivos, que a partir de uma série de eventos observados são capazes de projetar cenários futuros (pense nos modelos de clima ou de projeções populacionais) e assim contribuir para o planejamento de ações de diversas esferas públicas. No campo da estatística há um aforismo: “Todos os modelos estão errados, mas alguns são úteis” (George E. P. Box, estatístico inglês, 1919–2013). O Profluxo, como veremos, é um modelo simples capaz de oferecer diversos indicadores, mas com limitações.

O modelo Profluxo pertence à classe dos modelos descritivos, portanto trata de oferecer um retrato, um diagnóstico do que se observa. Essa ferramenta visa modelar — reconstruir — o fluxo de pessoas ao longo das séries do sistema escolar. A partir de observações, tais como o número de pessoas que concluiu cada série e o número de pessoas que frequenta determinada série, essa ferramenta indica a permanência e a transição das pessoas ao longo das séries de um sistema escolar.

Hoje, o contexto de produção de dados sobre escolares é muito diferente daquele em que o Profluxo foi criado. Pode-se argumentar que essa ferramenta já não é importante para a finalidade a qual ele foi criado, uma vez que há indicadores de transição de série computados e publicados pelo Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira, o INEP — principal instituição responsável pelos dados escolares do país. Porém, a principal virtude desse modelo é a de nos permitir conhecer a trajetória escolar de todas as pessoas, em todas as idades, mesmo que já não sejam estudantes inseridos no sistema escolar. E isso é especialmente necessário para conhecermos o contexto e o perfil daqueles que estão à margem dos processos de escolarização, daqueles que tiveram seu direito à educação negado durante as fases de infância e juventude, daqueles invisíveis ou não prioritários para as políticas públicas de educação.

A seguir, vamos recuperar muito rapidamente a história desse modelo e indicar quais são os dados necessários para sua construção.

## O PROFLUXO E OS DADOS SOBRE ESCOLARIDADE

O modelo Profluxo (Fletcher e Ribeiro, 1989) é um aprimoramento de modelos de fluxo escolar (Thonstad, 1980) e foi elaborado em uma época em que os dados sobre os estudantes e sua sobrevivência no sistema escolar eram inconsistentes. O Censo Escolar, embora tenha tido seus primeiros dados levantados em 1932, em virtude do histórico de rupturas que marcou as instituições governamentais responsáveis por esse levantamento estatístico, só passou a ter dados mais robustos a partir da segunda metade da década de 1990<sup>1</sup>.

Portanto, para se conhecer o fluxo de estudantes e suas taxas de transição<sup>2</sup>, foi necessário recorrer a outras fontes de dados. O matemático Philip Ralph Fletcher e o engenheiro e físico Sérgio Costa Ribeiro empregaram informações sobre escolaridade das pessoas contidas na Pesquisa Nacional por Amostra de Domicílios (PNAD), elaborada pelo Instituto Brasileiro de Geografia e Estatística (IBGE) para a construção do modelo. Foi a partir da combinação de informações sobre idade

1 No endereço eletrônico do INEP estão disponíveis os microdados de 1995 a 2019.

2 As taxas de fluxo ou transição são compostas pelas taxas de promoção, repetência e evasão. Um estudante promovido é aquele que se matriculou, no ano letivo seguinte, na série seguinte àquela em que fora aprovado. Um estudante repetente matricula-se, no ano letivo seguinte, na mesma série em que fora reprovado. Já um aluno evadido é aquele que reprovou ou abandonou a escola e não se matriculou no sistema escolar no ano letivo seguinte (INEP, 2014).

e sobre acesso às séries escolares que os autores conseguiram identificar quantos progrediam, quantos repetiam e quantos evadiam do sistema escolar (Fletcher e Ribeiro 1989). Um dos primeiros trabalhos referenciados que aplicou o modelo Profluxo foi o de Ribeiro (1991). Utilizando dados da PNAD de 1982, o autor criticou os dados do então Ministério da Educação e Cultura (MEC) e indicou que o maior problema da educação, à época, não era o abandono escolar, mas sim as elevadas taxas de repetência.

A pesquisa utilizada como fonte desses dois trabalhos (Fletcher e Ribeiro, 1989; Ribeiro, 1991), a PNAD, é uma pesquisa amostral que foi a campo entre 1967 e 2015, cujos últimos resultados foram publicados em 2016. Sua principal unidade de análise é o domicílio, o qual deve ser representativo da área em que está localizado. Além de características de infraestrutura do domicílio, a pesquisa levanta informações sobre seus residentes. Seus resultados podem ser lidos segundo os níveis nacional, estaduais e metropolitanos.

Durante os anos em que a PNAD foi levantada, seus dados captaram uma diversidade de temas, sobretudo os associados ao mercado de trabalho (Jannuzzi, 2017). O levantamento de características de escolaridade da população manteve-se ao longo dos anos, havendo também anos da pesquisa em que esse assunto fora aprofundado. Em 2004 e 2006, houve suplementos sobre aspectos complementares de educação; em 2007, perguntas sobre educação de jovens e adultos e educação profissional; e, em 2014, o suplemento contemplou quesitos sobre educação e qualificação profissional. Em 2016, a PNAD deu espaço a uma nova pesquisa, a PNAD Contínua (PNAD-C), com nova metodologia, e que permanece captando informações sobre a escolaridade da população. Em virtude de mudanças metodológicas, os resultados da PNAD e da PNAD-C não são comparáveis.

A PNAD permite recuperar as seguintes informações, por idade:

- número de pessoas que frequentavam escola;
- série/curso que frequentavam;
- pessoas que não frequentavam a escola e nunca frequentaram;
- pessoas que não frequentavam a escola, mas frequentaram anteriormente — e, entre essas pessoas, qual a última série/curso concluída com aprovação.

A combinação dessas variáveis permite conhecer também os *anos de estudo* completos que cada pessoa tem, esteja ela frequentando escola ou não. Outras pesquisas que permitam reconstituir os anos de estudo das pessoas também podem ser usadas como entrada de dados de alimentação do Profluxo. Esse é o caso de algumas rodadas do próprio Censo Demográfico (IBGE) e da Pesquisa de Demografia e Saúde (*Demographic Health Survey* — DHS/USAID), por exemplo.

A fim de manter o propósito didático e ilustrativo deste texto, vamos empregar dados da PNAD 2009, para o estado de São Paulo, sem expandir a amostra. Isto é, sem multiplicar o número de casos observados pelo fator de ponderação — cujo produto indicaria a ordem de grandeza observada em toda a

população (universo). Para fins analíticos, o ideal é sempre aplicar esse fator de ponderação, cuja variável corresponde ao “peso” que cada pessoa na amostra tem para representar outras no universo. Aqui, vamos trabalhar com os valores encontrados na amostra, pois sua ordem de grandeza é suficiente para compreender o funcionamento do Profluxo, além de facilitar a reprodução dos resultados sem maiores dificuldades.

Para o propósito deste texto, além da justificativa do uso dos valores da amostra em vez dos valores expandidos, é preciso explicar minimamente a lógica de funcionamento dos bancos de dados.

Em linhas gerais, um banco de dados é uma forma de organizar um conjunto de características quantitativas e/ou qualitativas coletadas sobre determinada unidade de análise. No caso da PNAD, há informações sobre domicílios e sobre pessoas. Vamos trabalhar com a unidade “pessoas”, cujo conjunto de dados coletados permite conhecer características sobre a escolaridade. Os bancos de dados são organizados em linhas e colunas. Geralmente, cada linha representa um caso observado — pessoas entrevistadas na pesquisa — cujas características pesquisadas estão dispostas em colunas (variáveis). São como tabelas, muitas vezes, com milhões de linhas e centenas de colunas. As bases de dados podem ser lidas e manipuladas por programas computacionais específicos para essa finalidade, sendo exemplos de *softwares* pagos o Microsoft Excel, o Stata, o SAS e o SPSS (*Statistic Package for Social Sciences*) — este bastante utilizado por cientistas sociais; e programas computacionais gratuitos, como o R, que possui código aberto (*open-source*) e grande comunidade ativa de usuários. No caso da PNAD, encontram-se no endereço eletrônico do próprio IBGE diferentes sintaxes prontas (sintaxes são instruções para o programa ler o arquivo de banco de dados).

Da PNAD, vamos então trabalhar com as seguintes variáveis para alimentar o Profluxo:

- frequenta escola ou creche (resposta: “sim”);
- série que frequenta;
- curso que frequenta;
- anos de estudo (todas as pessoas).

A variável “frequenta escola ou creche” deriva da pergunta “Frequenta escola ou creche?”. Caso a resposta seja afirmativa, ela compreende todas as pessoas inseridas no sistema escolar, da educação básica ao ensino superior, incluindo-se pós-graduação e cursos não seriados, como educação infantil e cursos pré-vestibulares. Ainda no caso afirmativo, investiga-se qual série/curso é frequentado, se o curso é seriado, e, nos anos mais recentes da PNAD, qual o tipo da rede de ensino.

Já a variável “anos de estudo” é uma variável derivada, isto é, calculada posteriormente pelo IBGE a partir das informações prestadas sobre escolaridade de todos os entrevistados. Essa variável resulta da correspondência entre o nível de escolaridade no momento da pesquisa e a quantidade de anos de estudo completos, cuja contagem é iniciada a partir da conclusão, com aprovação, da

primeira série/primeiro ano do ensino fundamental. Importante lembrar que desde 2007 iniciou-se a transição gradual da mudança da duração do ensino fundamental, de 8 para 9 anos (Brasil, 2006). Por isso, o IBGE ajustou o cálculo dos anos de estudo de modo que a variável fosse comparável com rodadas anteriores da pesquisa:

Na contagem dos anos de estudo para o ensino fundamental com duração de 9 anos, a primeira série concluída com aprovação foi enquadrada em menos de 1 ano de estudo, a segunda série, em 1 ano de estudo, e assim, sucessivamente, até a nona série, classificada em 8 anos de estudo (IBGE, 2015, p. 40).

Essa opção metodológica do IBGE não interfere substancialmente na variável “anos de estudo” das primeiras pesquisas após a mudança na duração do ensino fundamental, quando uma menor parte dos estudantes nesse nível de ensino encontrava-se matriculada no ensino fundamental com duração de 9 anos. No entanto, conforme avançam os anos, mais alunos passam a compor o novo sistema. Desse modo, se estiver trabalhando com anos mais recentes, vale a pena avaliar qual decisão metodológica tomar para harmonizar as categorias entre os anos escolares concluídos com aprovação e os anos de estudo: se suprimir os que cursam o primeiro ano escolar ou se somar as categorias do último ano do ensino fundamental, isto é, oitava série e nono ano.

Uma vez de posse das informações sobre escolaridade das pessoas, sigamos à lógica, à construção e à análise do Profluxo, descritas na próxima seção.

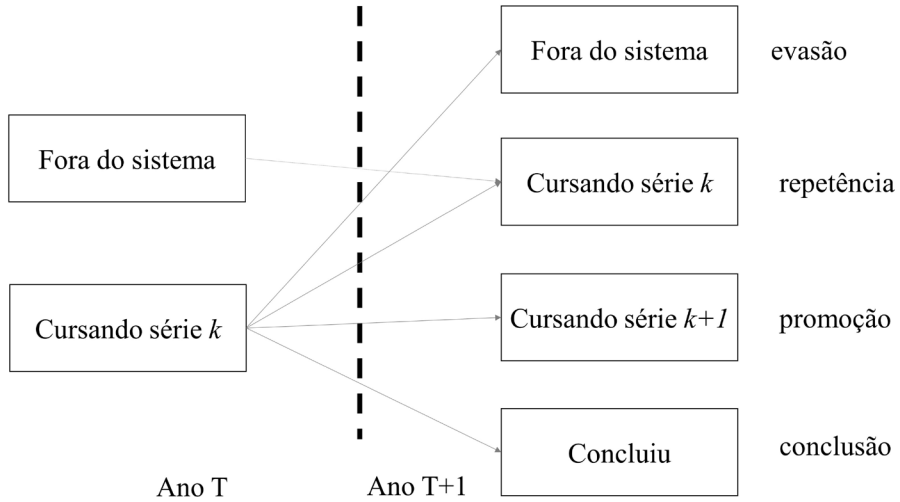
## LÓGICA, CONSTRUÇÃO E ANÁLISE DO PROFLUXO

A lógica desse modelo consiste em identificar o fluxo de pessoas no sistema escolar a partir de um único ponto no tempo, ou seja, de dados captados em um único ano. Isso é possível em razão do pressuposto que considera haver relativa estabilidade nos níveis de transições entre anos consecutivos e pouca ou nenhuma variação no tamanho das coortes sucessivas (Fletcher e Ribeiro, 1989). Uma coorte é, por definição, um conjunto de pessoas que compartilham um evento comum em um dado período. O dado reconstruído pelo Profluxo é, portanto, transversal<sup>3</sup>. Isso significa dizer que se trata de uma fotografia do fluxo observado em dado momento. Desse modo, estamos obrigados a lidar com um “pseudo” fluxo de todas as pessoas que estão inseridas ou passaram pelo ensino seriado (Fletcher e Ribeiro, 1989; Golgher, 2004). As transições possíveis dentro do sistema escolar estão dispostas na Figura 2:

---

3 O fluxo real de estudantes, ou seja, uma análise longitudinal, de acompanhamento das pessoas ao longo das séries é possível, porém somente utilizando os dados do censo escolar e em condições bastante específicas: é preciso solicitar ao INEP acesso para trabalhar com esses dados.





Fonte: Adaptado de Santos (2006).

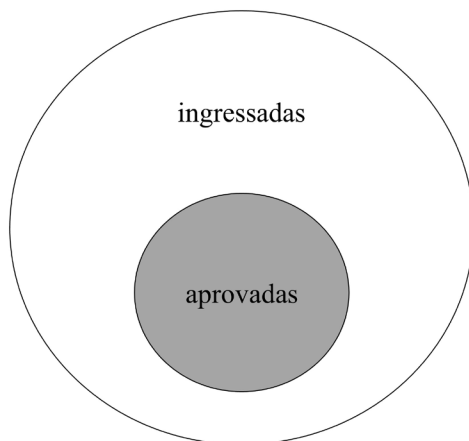
**Figura 2 – Fluxos/ transições em um sistema escolar seriado.**

Recuperar essa lógica e as transições de séries é importante para compreender dois conceitos necessários para a construção do Profluxo: a condição de aprovação e de ingresso das pessoas em cada série.

As pessoas *aprovadas* são aquelas que *concluíram com aprovação* determinada série/ano escolar. Chamemos determinada série/ano escolar de série  $k$ . Essa série  $k$  pode ser, por exemplo, o primeiro ano (antiga primeira série) do ensino fundamental. Então, nesse exemplo, são consideradas aprovadas no primeiro ano do ensino fundamental todas as pessoas que concluíram com aprovação essa série.

Já as pessoas *ingressadas* são aquelas que *concluíram com aprovação* determinada série  $k$  (se aprovadas, necessariamente tiveram que ingressar) e as pessoas que ingressaram, *mas não concluíram* a referida série  $k$ . Esse estado de “não conclusão” pode acontecer porque a pessoa encontrava-se frequentando a série  $k$  no ano da pesquisa ou porque a pessoa já não frequentava a escola e não concluiu essa série em algum momento do passado. A PNAD tem a informação das pessoas que estão cursando determinada série (ingressadas), mas não identifica se a pessoa que já não frequenta o sistema escolar e concluiu determinada série chegou a se matricular/ingressar no ano letivo seguinte, ou seja, na série  $k+1$ . A relação entre pessoas aprovadas e pessoas ingressadas em determinada série  $k$  pode ser visualizada na Figura 3:

Para se obter a proporção de aprovados, é preciso ter a distribuição do número absoluto de pessoas — dentro e fora do sistema escolar —, segundo a última série concluída com aprovação, por idade simples. A informação sobre



Fonte: Elaboração própria.

**Figura 3 – Conjunto de pessoas ingressadas e de pessoas aprovadas em determinada série  $k$ .**

a última série concluída com aprovação será disposta em linha, chamada “ $i$ ”<sup>4</sup>. Já a informação sobre as idades das pessoas será disposta em colunas, “ $j$ ”. Portanto, o conjunto dessas informações será apresentado em uma série de matrizes (tabelas).

Golgher (2004) obtém a matriz de aprovados a partir da variável derivada “anos de estudo” utilizando-se do seguinte raciocínio: *foi aprovado na série “ $i$ ” quem tem pelo menos o equivalente a “ $i$ ” anos de estudo*. Vamos ao passo a passo da montagem das matrizes e do modelo Profluxo:

## 1º PASSO: MATRIZ DE DISTRIBUIÇÃO DE PESSOAS POR IDADE SEGUNDO ANOS DE ESTUDO

Para conhecermos a proporção de aprovados em cada série segundo a idade, o primeiro passo é obter a distribuição do número de pessoas por idade, segundo a quantidade de anos de estudo completos até o ano da pesquisa. Essa distribuição pode ser representada em forma de tabela, conforme exemplo dado pela Tabela 1 [matriz  $E(i,j)$ ].

## 2º PASSO: MATRIZ DE APROVADOS

É a partir da matriz  $E(i,j)$  que trilhamos o segundo passo para obter a proporção de pessoas aprovadas. Para isso, precisamos gerar uma segunda matriz com os números absolutos de pessoas aprovadas em cada série, segundo a idade. É nessa segunda matriz que consideramos o raciocínio de que o indivíduo tenha “pelo menos  $x$  anos de estudo”.

4 Embora na língua portuguesa tenhamos a associar, no caso desse modelo, “ $i$ ” à “idade”, mantive a notação tal como geralmente se encontra nas notações de matrizes: “ $i$ ”, de *index*, para linha e “ $j$ ” para coluna.

Tabela 1 – Distribuição de anos de estudo (exatos), por idade simples. São Paulo, 2009.

Anos de estudo (exatos)	Idade simples													
	6	7	8	9	10	11	12	13	14	15	16	17	18	19
0	542	478	221	45	16	6	7	5	7	8	8	9	12	9
1	19	118	319	221	44	13	5	2	2	4	1	1	1	0
2	0	9	119	361	195	55	20	5	4	0	1	1	1	0
3	0	0	8	89	322	201	66	24	5	3	4	1	4	1
4	0	0	0	4	91	321	223	66	38	20	6	4	3	2
5	0	0	0	0	17	67	346	201	72	14	16	9	8	12
6	0	0	0	0	0	7	73	326	160	70	46	31	13	14
7	0	0	0	0	0	0	9	64	370	142	56	32	27	28
8	0	0	0	0	0	0	0	11	51	350	185	77	74	51
9	0	0	0	0	0	0	0	0	8	65	306	146	81	51
10	0	0	0	0	0	0	0	0	0	7	45	298	153	71
11	0	0	0	0	0	0	0	0	0	0	3	43	318	376
12	0	0	0	0	0	0	0	0	0	0	0	2	14	73
13	0	0	0	0	0	0	0	0	0	0	0	0	0	15
14	0	0	0	0	0	0	0	0	0	0	0	0	0	5
15 ou +	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Total de pessoas	561	605	667	720	685	670	749	704	717	683	677	654	709	708

Fonte: IBGE/PNAD, 2009, dados da amostra.

Para obter esse cumulativo de anos de estudo, basta somar, em cada coluna da matriz  $E(i, j)$ , as células correspondentes à quantidade de anos de estudo subsequentes à célula de referência. Vamos chamar essa nova matriz de aprovados de  $A(i, j)$ , expressa pela seguinte notação:

$$A(i, j) = \sum_{k=i}^{15} E(k, j) ,$$

em que:

- $A(i, j)$ : matriz do número de aprovados em cada série  $i$ , segundo a idade  $j$ ;
- $E(i, j)$ : matriz do número de anos de estudo  $i$ , segundo a idade  $j$ .

Vejamos a seguir como ocorre esse procedimento. A Tabela 2 contém a distribuição de pessoas por idade, segundo a quantidade de anos de estudo (*pele menos*). Observe que há três células destacadas em escala de cinza na coluna referentes às pessoas de 7 anos. Vamos utilizar essas células marcadas como exemplos: A célula em cinza mais escuro refere-se às pessoas de 7 anos com 0 anos de estudo. O número “605” corresponde ao total de pessoas nessa idade observado na amostra.

**Tabela 2 – Distribuição de anos de estudo (pelo menos), por idade simples. São Paulo, 2009.**

Anos de estudo (pelo menos)	Idade simples													
	6	7	8	9	10	11	12	13	14	15	16	17	18	19
0	561	605	667	720	685	670	749	704	717	683	677	654	709	708
1	19	127	446	675	669	664	742	699	710	675	669	645	697	699
2	0	9	127	454	625	651	737	697	708	671	668	644	696	699
3	0	0	8	93	430	596	717	692	704	671	667	643	695	699
4	0	0	0	4	108	395	651	668	699	668	663	642	691	698
5	0	0	0	0	17	74	428	602	661	648	657	638	688	696
6	0	0	0	0	0	7	82	401	589	634	641	629	680	684
7	0	0	0	0	0	0	9	75	429	564	595	598	667	670
8	0	0	0	0	0	0	0	11	59	422	539	566	640	642
9	0	0	0	0	0	0	0	0	8	72	354	489	566	591
10	0	0	0	0	0	0	0	0	0	7	48	343	485	540
11	0	0	0	0	0	0	0	0	0	0	3	45	332	469
12	0	0	0	0	0	0	0	0	0	0	0	2	14	93
13	0	0	0	0	0	0	0	0	0	0	0	0	0	20
14	0	0	0	0	0	0	0	0	0	0	0	0	0	5
15 ou +	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Total de pessoas na idade	561	605	667	720	685	670	749	704	717	683	677	654	709	708

Fonte: IBGE/PNAD, 2009, dados da amostra.

Isso acontece porque, na idade de 7 anos, todas as pessoas têm ou tiveram, em algum momento, 0 anos de estudo. Ou seja, a linha referente a 0 anos de estudo será sempre igual ao total de pessoas observado em determinada idade ou grupo etário.

Na célula correspondente às pessoas de 7 anos com *pelo menos* 1 ano de estudo, encontramos o valor “127”. Faz-se a seguinte leitura: das 605 crianças de 7 anos, 127 já haviam sido aprovadas na primeira série/primeiro ano do ensino fundamental. Esse valor é obtido do somatório de pessoas que tinham 1 ano de estudo completo com aqueles que haviam concluído mais que 1 ano de estudo (isto é, 118+9). Ou seja, compõem esse valor as pessoas que tinham *pelo menos, no mínimo*, 1 ano de estudo.

Seguindo a lógica do raciocínio, podemos concluir que o próximo valor destacado, “9”, encontrado para pessoas de 7 anos, indica que, dentro da população de 605 crianças, 9 já haviam sido aprovadas na segunda série/segundo ano do ensino fundamental.

Se somarmos os valores obtidos de cada coluna de  $A(i,j)$  (Tabela 2), teremos um valor irreal, inflacionado, dado que se trata da soma dos cumulativos. Portanto, contaremos uma mesma pessoa várias vezes, de acordo com a quantidade de

anos de estudo obtidos. Essa soma não tem um significado prático. Por essa razão, mantemos na linha referente ao total de pessoas na idade o valor real observado no banco de dados.

Os valores em  $A(i, j)$  servem para compor o cálculo da proporção de aprovados, expressa por uma matriz denominada  $P_A(i, j)$ , cujo cálculo é apresentado no próximo passo.

### 3º PASSO: MATRIZ DE PROPORÇÃO DE APROVADOS

A proporção de aprovados  $P_A(i, j)$  é obtida dividindo-se o número de pessoas aprovadas em cada uma das séries (tem pelo menos  $x$  anos de estudo), segundo a idade, pelo total de pessoas naquela mesma idade, isto é:

$$P_A(i, j) = \frac{1}{\sum_{k=1}^{15} E(k, j)} A(i, j),$$

em que:

- $A(i, j)$ : matriz do cumulativo de aprovados em cada série  $i$ , segundo a idade  $j$ ;
- $E(i, j)$ : matriz do número de anos de estudo  $i$ , segundo a idade  $j$ ;
- $P_A(i, j)$ : matriz da proporção do número de aprovados na série  $i$  e idade  $j$ .

A matriz  $P_A(i, j)$  resultante em nosso exemplo é apresentada na Tabela 3.

**Tabela 3 – Proporção de pessoas aprovadas em cada série, segundo a idade (%). São Paulo, 2009.**

Anos de estudo	Nível escolar correspondente	Idade simples															
		6	7	8	9	10	11	12	13	14	15	16	17	18	19		
1	Fundamental	1ª série	3	21	67	94	98	99	99	99	99	99	99	99	99	98	99
2		2ª série	0	1	19	63	91	97	98	99	99	98	99	98	98	98	99
3		3ª série	0	0	1	13	63	89	96	98	98	98	99	98	98	98	99
4		4ª série	0	0	0	1	16	59	87	95	97	98	98	98	97	97	99
5		5ª série	0	0	0	0	2	11	57	86	92	95	97	98	97	98	98
6		6ª série	0	0	0	0	0	1	11	57	82	93	95	96	96	96	97
7		7ª série	0	0	0	0	0	0	1	11	60	83	88	91	94	94	95
8		8ª série	0	0	0	0	0	0	0	2	8	62	80	87	90	90	91
9	Médio	1º ano	0	0	0	0	0	0	0	0	1	11	52	75	80	83	
10		2º ano	0	0	0	0	0	0	0	0	0	1	7	52	68	76	
11		3º ano	0	0	0	0	0	0	0	0	0	0	0	7	47	66	
12	Superior	1º ano	0	0	0	0	0	0	0	0	0	0	0	0	2	13	
13		2º ano	0	0	0	0	0	0	0	0	0	0	0	0	0	3	
14		3º ano	0	0	0	0	0	0	0	0	0	0	0	0	0	1	
15		4º ano ou +	0	0	0	0	0	0	0	0	0	0	0	0	0	0	

Fonte: IBGE/PNAD 2009, dados da amostra.

Uma vez que encontramos a proporção de pessoas aprovadas, falta agora calcular a proporção de pessoas ingressadas em cada série do sistema escolar, mesmo que a pessoa já não frequente uma instituição de ensino. O número de pessoas ingressadas em cada série, por idade, é obtido por meio da soma dos valores em  $A(i, j)$ , com o número de pessoas que frequentava determinada série, segundo a idade. Essa soma irá resultar em outra matriz,  $F(i, j)$ , obtida no passo seguinte.

#### 4º PASSO: MATRIZ DE PESSOAS QUE FREQUENTAVAM ESCOLA

A matriz  $F(i, j)$  é dada pela distribuição das pessoas que frequentavam a escola, segundo a série que cursavam e a idade que tinham no momento em que a pesquisa foi levantada. Como a PNAD separa os quesitos série e curso frequentado, é preciso combinar as variáveis “curso que frequenta” e “série que frequenta”. Recordamos que a PNAD diferencia o ensino fundamental de duração de 8 e 9 anos a partir de 2007. Se estiver trabalhando com pesquisas em que há diferenciação da duração do ensino fundamental, basta seguir o procedimento de harmonização mencionado anteriormente. Nessa parte do exercício, como 76% dos que frequentavam o ensino fundamental estavam matriculados em curso de duração de 8 anos, optamos por harmonizar somando os que cursavam o nono ano (residual em 2009: 97 pessoas em uma amostra de 2.785 ou 3,3% dos que frequentavam o fundamental de 9 anos) àqueles que cursavam a oitava série.

Trata-se de construir uma tabela com três camadas de informação, dado que será preciso acrescentar a distribuição das pessoas de acordo com a série/curso, segundo a idade. Essa é uma tarefa que pode ser facilmente executada pelos *softwares* estatísticos: é preciso indicar para o programa utilizado na extração dos microdados que se deseja combinar as variáveis “série que frequenta” e “curso que frequenta” (linha) e “idade” (coluna). A tabela/matriz obtida  $F(i, j)$  será parecida com a Tabela 4.

#### 5º PASSO: MATRIZ DE INGRESSADOS

Somando-se os números em  $A(i, j)$  com a distribuição dos que frequentam escola segundo série e idade, temos o número de pessoas ingressadas em cada série, em cada idade  $I(i, j)$ :

$$I(i, j) = A(i, j) + F(i, j),$$

em que:

- $I(i, j)$ : matriz do número de ingressados em cada série  $i$ , segundo a idade  $j$ ;
- $F(i, j)$ : matriz do número de pessoas que frequentam a série  $i$ , na idade  $j$ .

A matriz resultante é apresentada na Tabela 5.

#### 6º PASSO: MATRIZ DA PROPORÇÃO DE INGRESSADOS

E do mesmo modo que calculamos a proporção de aprovados, podemos calcular a proporção de ingressados ( $P_1(i, j)$ ): divide-se o número de pessoas ingressadas em cada série, em cada idade, pelo total de pessoas naquela idade. O resultado é apresentado na Tabela 6.

**Tabela 4 – Número de pessoas que frequentam escola, segundo série e idade. São Paulo, 2009.**

Curso	Série	Idade simples													
		6	7	8	9	10	11	12	13	14	15	16	17	18	19
Regular / EJA Fundamental	1ª	285	262	39	9	3	1	0	2	2	2	3	1	1	1
	2ª	65	272	336	59	12	5	3	1	0	1	3	0	0	0
	3ª	3	41	247	404	69	24	9	3	3	2	0	1	1	0
	4ª	0	1	32	211	390	99	33	10	5	3	3	0	1	1
	5ª	0	0	0	23	173	380	123	44	18	5	3	2	4	0
	6ª	0	0	0	0	23	136	398	115	50	18	7	2	1	0
	7ª	0	0	0	0	3	21	155	375	119	36	21	7	3	2
	8ª	0	0	0	0	0	1	16	128	373	148	58	24	6	5
Regular / EJA Médio	1º								11	51	345	168	51	29	6
	2º								0	7	62	300	133	50	15
	3º								0	0	5	39	289	128	43
Superior	1º											1	9	90	59
	2º											0	1	11	69
	3º											0	0	0	10
	4º+											0	0	0	5
Total de pessoas que frequentam escola		353	576	654	706	673	667	737	689	628	627	606	520	325	216

EJA: Educação de Jovens e Adultos.

Fonte: IBGE/PNAD 2009, dados da amostra.

**Tabela 5 – Número de pessoas ingressadas em cada série escolar, segundo a idade. São Paulo, 2009.**

Curso	Série	Idade simples													
		6	7	8	9	10	11	12	13	14	15	16	17	18	19
Regular / EJA Fundamental	1ª	304	389	485	684	672	665	742	701	712	677	672	646	698	700
	2ª	65	281	463	513	637	656	740	698	708	672	671	644	696	699
	3ª	3	41	255	497	499	620	726	695	707	673	667	644	696	699
	4ª	0	1	32	215	498	494	684	678	704	671	666	642	692	699
	5ª	0	0	0	23	190	454	551	646	679	653	660	640	692	696
	6ª	0	0	0	0	23	143	480	516	639	652	648	631	681	684
	7ª	0	0	0	0	3	21	164	450	548	600	616	605	670	672
	8ª	0	0	0	0	0	1	16	139	432	570	597	590	646	647
Regular / EJA Médio	1º	0	0	0	0	0	0	0	11	59	417	522	540	595	597
	2º	0	0	0	0	0	0	0	0	7	69	348	476	535	555
	3º	0	0	0	0	0	0	0	0	0	5	42	334	460	512
Superior	1º	0	0	0	0	0	0	0	0	0	0	1	11	104	152
	2º	0	0	0	0	0	0	0	0	0	0	0	1	11	89
	3º	0	0	0	0	0	0	0	0	0	0	0	0	0	15
	4º+	0	0	0	0	0	0	0	0	0	0	0	0	0	5
Total de pessoas na idade		561	605	667	720	685	670	749	704	717	683	677	654	709	708

EJA: Educação de Jovens e Adultos.

Fonte: IBGE/PNAD 2009, dados da amostra.

**Tabela 6 – Proporção de ingressados em cada série, segundo a idade (%). São Paulo, 2009.**

Curso	Série	Idade simples													
		6	7	8	9	10	11	12	13	14	15	16	17	18	19
Regular / EJA Fundamental	1ª	54	64	73	95	98	99	99	100	99	99	99	99	98	99
	2ª	12	46	69	71	93	98	99	99	99	98	99	98	98	99
	3ª	1	7	38	69	73	93	97	99	99	99	99	98	98	99
	4ª	0	0	5	30	73	74	91	96	98	98	98	98	98	99
	5ª	0	0	0	3	28	68	74	92	95	96	97	98	98	98
	6ª	0	0	0	0	3	21	64	73	89	95	96	96	96	97
	7ª	0	0	0	0	0	3	22	64	76	88	91	93	94	95
	8ª	0	0	0	0	0	0	2	20	60	83	88	90	91	91
Regular / EJA Médio	1º	0	0	0	0	0	0	0	2	8	61	77	83	84	84
	2º	0	0	0	0	0	0	0	0	1	10	51	73	75	78
	3º	0	0	0	0	0	0	0	0	0	1	6	51	65	72
Superior	1º	0	0	0	0	0	0	0	0	0	0	0	2	15	21
	2º	0	0	0	0	0	0	0	0	0	0	0	0	2	13
	3º	0	0	0	0	0	0	0	0	0	0	0	0	0	2
	4º+	0	0	0	0	0	0	0	0	0	0	0	0	0	1

EJA: Educação de Jovens e Adultos.

Fonte: IBGE/PNAD 2009, dados da amostra.

Resumindo, os passos são:

- obter a distribuição do número de pessoas, em cada idade, que foram aprovadas em determinada série. Na PNAD, isso pode ser feito diretamente por meio da variável “anos de estudo”;
- obter a distribuição cumulativa de pessoas aprovadas em cada série, em cada idade;
- obter a proporção de pessoas aprovadas, dividindo-se o número de pessoas aprovadas, em cada série, segundo a idade, pelo total de pessoas naquela idade;
- obter a distribuição de pessoas que frequentam escola, segundo idade e série que frequentam;
- somar a matriz de aprovados (2º passo) com a matriz de pessoas que frequentam escola, segundo idade e série (4º passo);
- obter a proporção de ingressados, dividindo-se o número de ingressados, em cada série, segundo a idade, pelo total de pessoas naquela idade.

## ANALISANDO AS SAÍDAS DO PROFLUXO

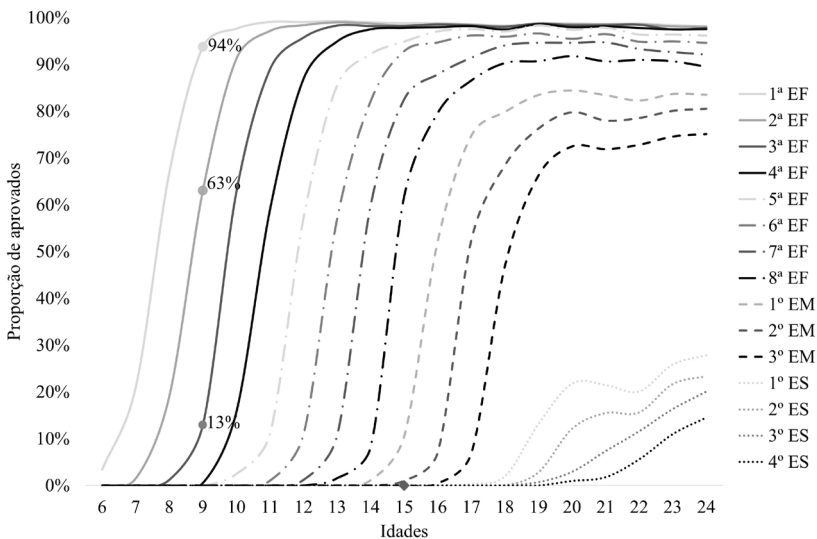
As proporções de aprovados [matriz  $P_A(i,j)$ ; Tabela 3] e de ingressados [matriz  $P_I(i,j)$ ; Tabela 6] são as duas principais saídas do modelo Profluxo. Como há muita informação nas tabelas, fica mais fácil visualizá-las e analisá-las por meio de gráficos.



A melhor forma de apresentar graficamente essas informações é por meio um gráfico de linhas. As linhas representam a proporção de aprovados e de ingressados em cada série, segundo as idades. Embora um gráfico de linhas seja adequado para apresentação de séries temporais de dados, no caso do Profluxo, esse tipo de gráfico é o que melhor permite observar os dados. Se optássemos por um gráfico de colunas, haveria muitas sobreposições, o que dificultaria ler os resultados.

Embora o gráfico de linhas apresente os dados de forma mais clara em comparação com as tabelas, corremos o risco de tentar interpretá-las como linhas de coortes — ou seja, como se fossem um mesmo grupo etário observado ao longo do tempo. As linhas *não* são representações de coortes. As linhas representam as proporções e deve-se ler os gráficos segundo as idades. Por exemplo: qual a proporção de pessoas na idade  $j$  que havia sido aprovada na série  $i$ ?

Utilizando a matriz da proporção de aprovados (Tabela 3) construímos o Gráfico 1. Os pontos destacados são os valores das proporções encontradas para as



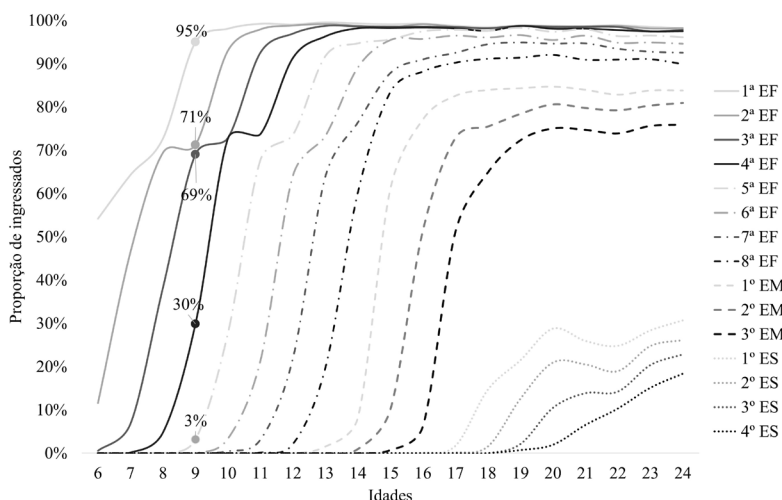
EF: ensino fundamental; EM: ensino médio; ES: ensino superior.  
Fonte: IBGE/PNAD 2009, dados da amostra.

**Gráfico 1 – Proporção de aprovados em cada série, segundo a idade. São Paulo, 2009.**

crianças de 9 anos residentes no estado de São Paulo em 2009. No exemplo, 94% haviam sido aprovadas na primeira série/primeiro ano do ensino fundamental, 63% na segunda série/segundo ano e 13% na terceira série/terceiro ano. De modo complementar, podemos afirmar que 6% (ou 100-94%) das crianças nessa idade haviam sequer concluído a primeira série/primeiro ano do fundamental, ainda que tivessem anos de vida suficientes para ter completado essa série (Brasil, 1996). Esses 6% podem representar as crianças que nunca entraram na escola, que repetiram o primeiro ano — e, portanto, não foram aprovadas — ou ainda que saíram da escola antes da conclusão do ano letivo.

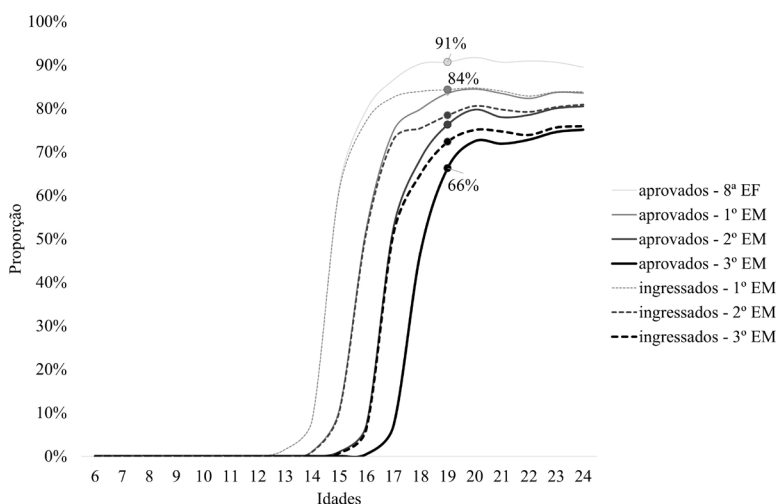
E a partir da matriz da proporção de ingressados (Tabela 6), obtemos o Gráfico 2, cujos pontos destacados refletem os valores das proporções entre as crianças de 9 anos que ingressaram em cada uma das séries do ensino fundamental:

A esta altura o leitor deve estar se questionando se todo esse trabalho se resume a conhecer a proporção de pessoas aprovadas e ingressadas em cada série escolar. O potencial do Proflixo se revela quando apresentamos as informações da proporção de aprovados e de ingressados em um mesmo gráfico (Gráfico 3).



EF: ensino fundamental; EM: ensino médio; ES: ensino superior.  
 Fonte: IBGE/PNAD 2009, dados da amostra.

**Gráfico 2 – Proporção de ingressados em cada série, segundo a idade. São Paulo, 2009.**



EF: ensino fundamental; EM: ensino médio.  
 Fonte: IBGE/PNAD 2009, dados da amostra.

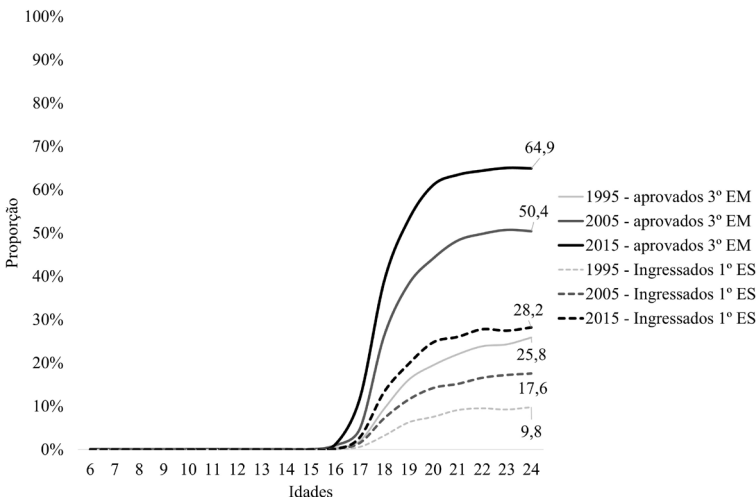
**Gráfico 3 – Proporção de aprovados e de ingressados em cada série, segundo a idade. São Paulo, 2009.**

No Gráfico 3 selecionamos para exemplo somente a transição do ensino fundamental para o médio e as séries correspondentes a essa última etapa da educação básica. Os ingressados são representados pelas linhas tracejadas e os aprovados pelas linhas contínuas. A diferença entre os ingressados na série  $k$  e os aprovados na mesma série  $k$  indica a proporção de pessoas que não foram aprovadas nessa série, seja porque ainda estão cursando, seja porque repetiram ou ainda porque abandonaram. Já a diferença entre os que foram aprovados na série  $k$  e os ingressados na série  $k+1$  indica evasão, uma vez que tendo credenciais para seguir nos estudos, a pessoa não se matriculou na série seguinte.

Observe os pontos correspondentes à idade de 19 anos no Gráfico 3: 91% das pessoas nessa idade haviam concluído o último ano do ensino fundamental; 84% ingressaram no primeiro ano do ensino médio; 83% foram aprovadas nesse primeiro ano; 78% ingressaram na série seguinte (segundo ano do ensino médio) e 76% foram aprovadas nessa série; 72% ingressaram no último ano do ensino médio e apenas 66% haviam concluído a educação básica. Recordamos que aos 19 anos, teoricamente e em termos legais, as pessoas já possuem idade suficiente para ter concluído a última série do ensino médio.

O que notamos, além dos baixos percentuais de ingresso e de aprovação nas idades esperadas, é a dificuldade de sobrevivência à transição de séries. Voltemos ao Gráfico 3, desta vez observando a linha de aprovados no último ano do ensino fundamental e a linha de ingressados no primeiro ano do ensino médio. A partir dos 19–20 anos, a distância entre essas linhas permanece aproximadamente constante, o que sugere haver uma dificuldade de transitar do ensino fundamental para o ensino médio, evidenciando um dos principais “gargalos” da educação básica brasileira.

Vejam outros exemplos e possibilidades de aplicação do Profluxo. O Gráfico 4 ilustra a proporção de pessoas que foram aprovadas no último ano do ensino



EM: ensino médio; ES: ensino superior.

\*Até 2004 a PNAD não abrangia as zonas rurais dos estados de Rondônia, Acre, Amazonas, Roraima, Pará e Amapá. Portanto, os dados de 1995 não incluem as pessoas residentes nessas regiões.

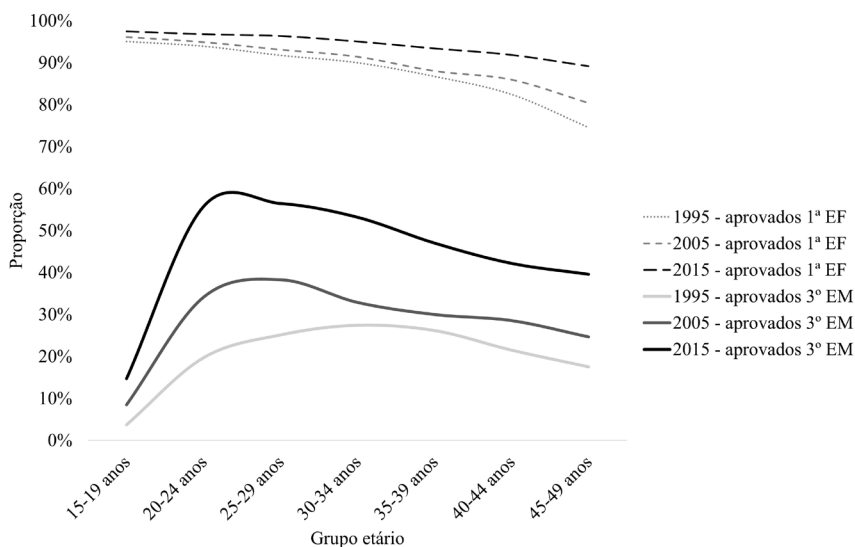
Fonte: IBGE/PNAD 1995, 2005 e 2015, dados da amostra expandida.

**Gráfico 4 – Proporção de aprovados no último ano do ensino médio e proporção de ingressados no primeiro ano do ensino superior. Brasil, 1995\*, 2005, 2015.**

médio e a proporção de pessoas que ingressaram no primeiro ano do ensino superior, no Brasil, nos anos de 1995, 2005 e 2015. Temos, portanto, a transição entre educação básica e ensino superior em três pontos no tempo:

No Gráfico 4, observamos que uma maior proporção de jovens em 2015 completou o ensino médio em comparação com os jovens de 1995 e 2005 (representados pelas linhas contínuas). Em 1995, 25,8% das pessoas nessa idade haviam concluído o ensino médio — indicando igualmente que três quartos dessa população não tinham a educação básica completa naquele ano. A situação melhorou para os jovens ao longo das décadas seguintes, pois, em 2005, metade deles havia concluído o ensino médio; em 2015, passou a quase 65%. Notamos ainda que, ao longo dos anos, houve aumento na proporção daqueles que ingressaram no ensino superior: em 1995, menos de 10% dos jovens de 24 anos havia ingressado no ensino superior, em 2015, eram 30%.

O Profluxo também nos permite sondar indiretamente as condições de escolarização das coortes e algo sobre a sobrevivência das pessoas às séries escolares. Vejamos nos Gráficos 5 a 9 a proporção de pessoas aprovadas na primeira série do ensino fundamental e a proporção de pessoas aprovadas no último ano do ensino médio, nos anos de 1995, 2005 e 2015, nas cinco grandes regiões brasileiras. Essa série de gráficos representa a entrada e a saída da educação básica. Optamos por apresentar os resultados das regiões em gráficos separados, pois em uma única representação seriam 30 linhas em escala de cinza, o que inviabilizaria a leitura dos resultados.

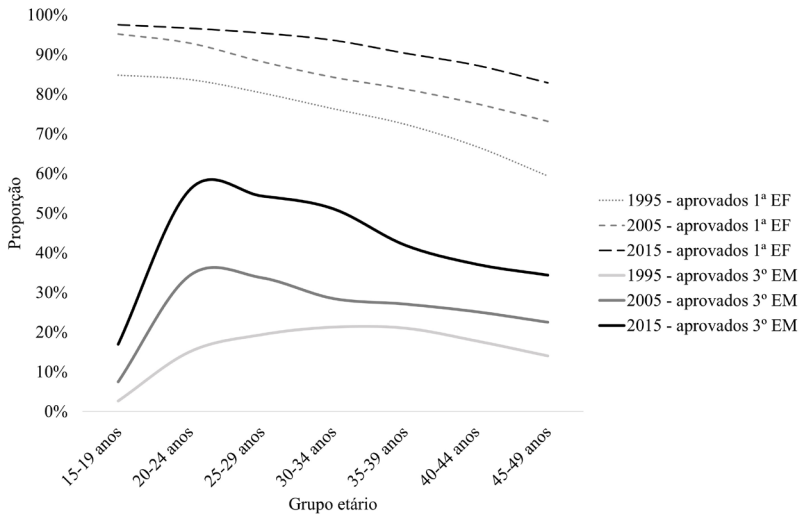


EF: ensino fundamental; EM: ensino médio.

\*\*Não foram computadas pela PNAD as pessoas residentes em áreas rurais.

Fonte: IBGE/PNAD 1995, 2005 e 2015. Dados da amostra expandida.

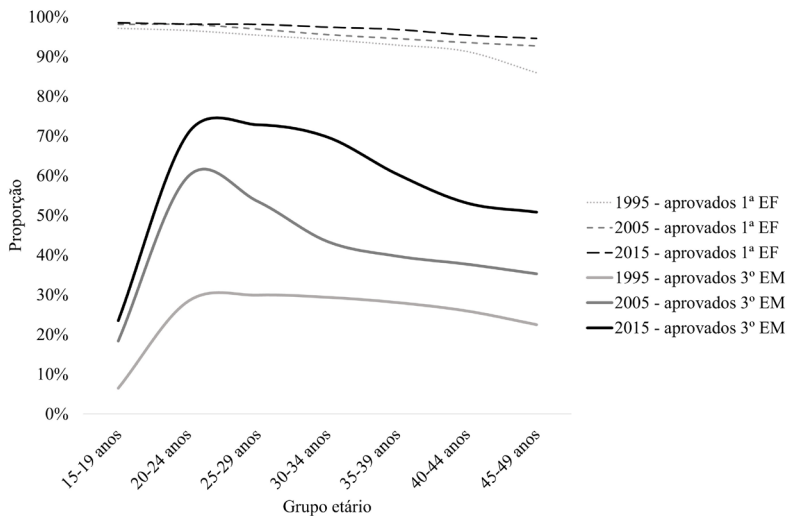
**Gráfico 5 – Proporção de aprovados na primeira série do ensino fundamental e no último ano do ensino médio, segundo grupo etário. Brasil, Região Norte, 1995\*\*, 2005 e 2015.**



EF: ensino fundamental; EM: ensino médio.

Fonte: IBGE/PNAD 1995, 2005 e 2015. Dados da amostra expandida.

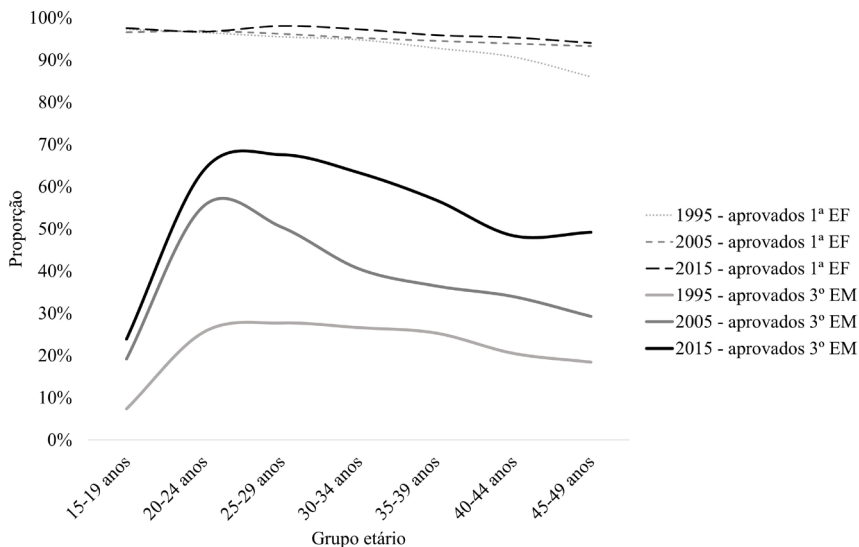
**Gráfico 6 – Proporção de aprovados na primeira série do ensino fundamental e no último ano do ensino médio, segundo grupo etário. Brasil, Região Nordeste, 1995, 2005 e 2015.**



EF: ensino fundamental; EM: ensino médio.

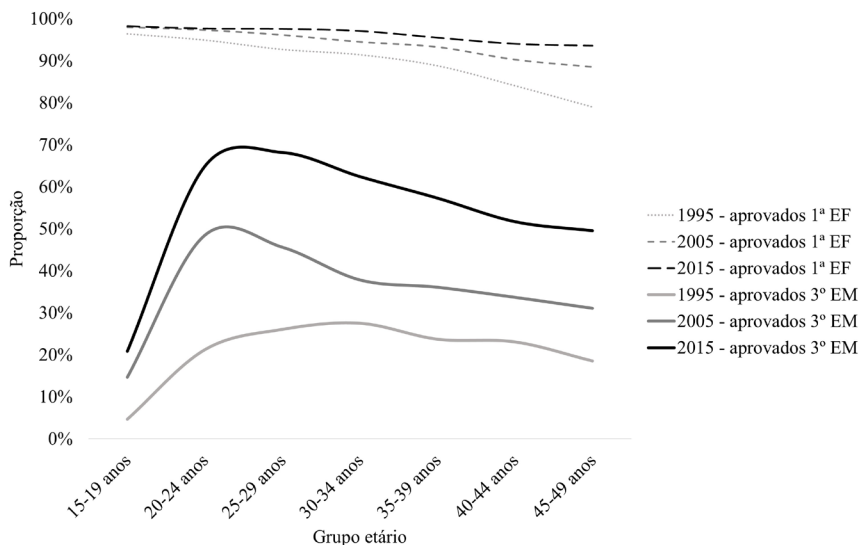
Fonte: IBGE/PNAD 1995, 2005 e 2015. Dados da amostra expandida.

**Gráfico 7 – Proporção de aprovados na primeira série do ensino fundamental e no último ano do ensino médio, segundo grupo etário. Brasil, Região Sudeste, 1995, 2005 e 2015.**



EF: ensino fundamental; EM: ensino médio.  
 Fonte: IBGE/PNAD 1995, 2005 e 2015. Dados da amostra expandida.

**Gráfico 8 – Proporção de aprovados na primeira série do ensino fundamental e no último ano do ensino médio, segundo grupo etário. Região Sul, 1995, 2005 e 2015.**



EF: ensino fundamental; EM: ensino médio.  
 Fonte: IBGE/PNAD 1995, 2005 e 2015. Dados da amostra expandida.

**Gráfico 9 – Proporção de aprovados na primeira série do ensino fundamental e no último ano do ensino médio, segundo grupo etário. Brasil, Região Centro-Oeste, 1995, 2005 e 2015.**

Os resultados das cinco regiões, nos três anos selecionados, nos permitem afirmar que:

- o grupo etário de 20 a 24 anos é o que apresenta maiores proporções de conclusão da educação básica em 2005 e 2015. Em 1995, isso só se observa discretamente na região Sudeste;
- apesar do aumento progressivo das proporções dos que concluem o ensino médio ao longo dos anos, uma parcela significativa das populações segue sem a educação básica completa;
- há um efeito da escolarização dos grupos etários mais velhos em todas as regiões, especialmente no Nordeste, evidenciando que as coortes mais velhas teriam tido maior dificuldade de acesso e de permanência escolar;
- as regiões Sul e Sudeste são as que apresentam as maiores proporções de pessoas que chegaram a ingressar no ensino fundamental nos anos observados;
- em 2015, nas regiões Sul, Sudeste e Centro-Oeste, mais da metade da população nos grupos etários maiores de 24 anos tinha ensino médio completo.

Além da comparação entre regiões, outros critérios podem ser empregados para aplicação do Profluxo, por exemplo, comparando-se as populações segundo cor, sexo, residência em área urbana/rural ou grupos de renda.

## COMENTÁRIOS FINAIS

Neste trabalho visitamos rapidamente noções de modelagem e de banco de dados para apresentar, de forma didática, a construção e algumas das possibilidades do modelo Profluxo. Essa ferramenta fornece insumos não apenas para exercícios com bases de dados ou pesquisas acadêmicas. O conhecimento desses indicadores de fluxo são subsídios para a formulação de políticas mais acuradas. Ao utilizar uma fonte de dados que não é orientada para fins de pesquisa educacional, tal como a PNAD, lidamos com uma série de limitações, mas também temos a possibilidade de olhar para as populações dentro e fora do sistema escolar. Esse olhar mais ampliado é o que permite considerar pessoas normalmente fora do campo de visão das políticas voltadas para a educação básica, especialmente nas etapas do ensino fundamental e médio. O fato de jovens e adultos não estarem mais na idade escolar não significa que estejam para sempre condenados a terem a educação básica incompleta. Nesse sentido, o Profluxo permite um ajuste de lentes de modo que se possa focalizar também essa população, permanecendo como um modelo útil para diagnóstico e auxílio às políticas públicas de educação.

## REFERÊNCIAS

BRASIL. Lei n. 9.394, de 20 de dezembro de 1996. Estabelece as diretrizes e bases da educação nacional. **Diário Oficial da União**, Brasília, DF, 23 dez. 1996.

- BRASIL. Lei n. 11.274, de 6 de fevereiro de 2006. Altera a redação dos arts. 29, 30, 32 e 87 da Lei n. 9.394, de 20 de dezembro de 1996, que estabelece as diretrizes e bases da educação nacional, dispondo sobre a duração de 9 (nove) anos para o ensino fundamental, com matrícula obrigatória a partir dos 6 (seis) anos de idade. **Diário Oficial da União**, Brasília, DF, 7 fev. 2006. Disponível em: [http://www.planalto.gov.br/ccivil\\_03/\\_ato2004-2006/2006/lei/l11274.htm](http://www.planalto.gov.br/ccivil_03/_ato2004-2006/2006/lei/l11274.htm). Acesso em: 23 jun. 2020.
- COSTA, L. B. Uma análise demográfica do desempenho escolar. **São Paulo em Perspectiva**, São Paulo, v. 7, n. 1, p. 58–69, 1993.
- CUNHA, J. M. P.; ASCAMA, M. O. O. Indicadores do fluxo escolar: modelo de profluxo e projeções de demanda escolar. **Textos NEPO 38**, Campinas, v. 38, p. 33–65, 2000.
- FLETCHER, P.; RIBEIRO, S. C. **Modeling education system performance with demographic data**: in introduction to the PROFUXO Model. Paris: UNESCO, 1989.
- GOLGHER, A. B. Modelo profluxo e indicadores derivados. In: RIOS-NETO, E. L. G.; RIANI, J. L. R. (orgs.). **Introdução à demografia da educação**. Rio de Janeiro: ABEP, 2004. p. 159–208. Disponível em: <http://www.abep.org.br/publicacoes/index.php/livros/article/view/156>. Acesso em: 23 jun. 2020.
- IBGE — Instituto Brasileiro de Geografia e Estatística. **Pesquisa Nacional por Amostra de Domicílios**: microdados da amostra. Rio de Janeiro: 2009.
- IBGE — Instituto Brasileiro de Geografia e Estatística. **Pesquisa Nacional por Amostra de Domicílios**: notas metodológicas. Rio de Janeiro: 2015.
- INEP — Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira. Coordenação-Geral de Sistemas Integrados de Informações. **Dicionário de indicadores educacionais**: formas de cálculo. Brasília, DF: INEP, 2004.
- JANNUZZI, P. M. **Indicadores sociais no Brasil**: conceitos, fontes de dados e aplicações. Campinas: Alínea, 2017.
- QUINO, J. L. **Toda Mafalda**. São Paulo: Martins Fontes, 2003.
- RIBEIRO, S. C. A pedagogia da repetência. **Estudos Avançados**, São Paulo, v. 5, n. 12, p. 7–21, 1991. Disponível em: [https://www.scielo.br/scielo.php?script=sci\\_arttext&pid=S0103-40141991000200002](https://www.scielo.br/scielo.php?script=sci_arttext&pid=S0103-40141991000200002). Acesso em: 23 jun. 2020. <https://doi.org/10.1590/S0103-40141991000200002>
- SANTOS, A. R. **Estimação de indicadores educacionais por sexo e raça/cor através da metodologia profluxo**. 2006. 152 f. Dissertação (Mestrado em Estatística) — Escola Nacional de Ciências Estatísticas, Rio de Janeiro, 2006. Disponível em: [http://www.dominiopublico.gov.br/pesquisa/DetalheObraForm.do?select\\_action=&co\\_obra=128446](http://www.dominiopublico.gov.br/pesquisa/DetalheObraForm.do?select_action=&co_obra=128446). Acesso em: 23 jun. 2020.
- TAVARES JR., F.; FARIA, V. B.; LIMA, M. A. Indicadores de fluxo escolar e políticas educacionais: avaliação das últimas décadas. **Estudos em Avaliação Educacional**, São Paulo, v. 25, n. 52, p. 48–67, 2012. Disponível em: <http://publicacoes.fcc.org.br/ojs/index.php/eae/article/view/1929>. Acesso em: 23 jun. 2020.
- THONSTAD, T. Analysing and projecting school enrolment in developing countries: a manual of methodology. **Statistical Reports and Studies**. Paris: UNESCO, 1980. (Statistical Reports and Studies, 24).



## SOBRE A AUTORA

FLÁVIA VITOR LONGO DOS SANTOS é doutora em demografia pela Universidade Estadual de Campinas (Unicamp).  
*E-mail:* longo.fla@gmail.com

**Conflitos de interesse:** A autora declara que não possui nenhum interesse comercial ou associativo que represente conflito de interesses em relação ao manuscrito.

**Financiamento:** Este artigo é um subproduto da tese de doutorado intitulada Demografia e políticas educacionais: um olhar sobre a proposta da reorganização escolar paulista de 2015, de autoria de Flávia Vitor Longo dos Santos, defendida junto ao Programa de Pós-Graduação em Demografia/UNICAMP. Financiamento: Coordenação de Aperfeiçoamento de Pessoal de Nível Superior — Brasil (CAPES), processo n. 150615.

*Recebido em 24 de junho de 2020*  
*Aprovado em 26 de novembro de 2020*

