

**UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL  
INSTITUTO DE LETRAS  
PROGRAMA DE PÓS-GRADUAÇÃO EM LETRAS**

**THALES PANTALEONI NICOLETTI**

**HUMANIDADES DIGITAIS: MAPAS E GRÁFICOS SOBRE O ROMANCE  
BRASILEIRO DO SÉCULO XIX**

**PORTO ALEGRE**

**2023**

**THALES PANTALEONI NICOLETTI**

**HUMANIDADES DIGITAIS: MAPAS E GRÁFICOS SOBRE O ROMANCE  
BRASILEIRO DO SÉCULO XIX**

Dissertação de Mestrado em Letras, área de Estudos de Literatura, apresentada como requisito parcial para a obtenção do título de Mestre pelo Programa de Pós-Graduação em Letras da Universidade Federal do Rio Grande do Sul.

Orientador: Prof. Dr. Luís Augusto Fischer

**Porto Alegre**

**2023**

### CIP - Catalogação na Publicação

Pantaleoni Nicoletti, Thales  
HUMANIDADES DIGITAIS: MAPAS E GRÁFICOS SOBRE O  
ROMANCE BRASILEIRO DO SÉCULO XIX / Thales Pantaleoni  
Nicoletti. -- 2023.  
122 f.  
Orientador: Luís Augusto Fischer.

Dissertação (Mestrado) -- Universidade Federal do  
Rio Grande do Sul, Instituto de Letras, Programa de  
Pós-Graduação em Letras, Porto Alegre, BR-RS, 2023.

1. Humanidades Digitais. 2. Mapas interativos. 3.  
Romance brasileiro do século XIX. I. Fischer, Luís  
Augusto, orient. II. Título.

**Thales Pantaleoni Nicoletti**

**Humanidades Digitais: mapas e gráficos sobre o romance brasileiro do século XIX**

Dissertação de Mestrado em Letras, área de Estudos de Literatura, apresentada como requisito parcial para a obtenção do título de Mestre pelo Programa de Pós-Graduação em Letras da Universidade Federal do Rio Grande do Sul.

Porto Alegre, 14 de Junho de 2023.

Resultado: Aprovado

**BANCA EXAMINADORA**

---

**Prof. Dr. Thiago Lima Nicodemo**  
Instituto de Filosofia e Ciências Humanas  
Universidade Estadual de Campinas

---

**Prof. Dr. Andre Dutra Boucinhas**  
Colégio Santo Inácio

---

**Prof<sup>a</sup>. Dra. Mariana da Silva Lima**  
Centro Federal de Educação Tecnológica Celso)

*À Polyana, minha companheira de vida, esposa e  
melhor amiga, que nunca deixou de me apoiar, seja  
qual fosse a aventura.*

## AGRADECIMENTOS

Em primeiro lugar, um agradecimento especial a todos os professores que me inspiraram durante a minha trajetória escolar e acadêmica. Em especial, à professora Nara Odi, que marcou o meu ensino médio no Colégio Santa Inês, em Porto Alegre; e ao professor Luís Augusto Fischer, que me acompanhou desde o começo da graduação até o término deste mestrado na UFRGS. É com muita honra que divido a profissão de professor com vocês.

Em segundo lugar, gostaria de agradecer à minha mãe, Carla, e ao meu pai, Atila, que sempre me incentivaram a estudar e alcançar os meus objetivos. Das reuniões de professores aos cursos de francês, vocês foram incansáveis nessa tarefa. Não poderia deixar de mencionar o meu irmão, Nicolas, que, do momento em que me segurou no colo no hospital, nunca mais me soltou. Sem vocês, nada disso teria sido possível; sou eternamente grato por cada segundo de afeto, dedicação e esforço que vocês dedicaram a mim.

Ainda, um agradecimento muito especial à minha esposa e melhor amiga, Polyana Araújo da Silva, que desde o dia em que nos conhecemos, aceitou dividir a vida dela comigo. Este mestrado não teria sido possível sem ela, uma vez que todos os questionamentos que tive, as dúvidas que surgiram ou as sugestões que apareceram foram discutidas com ela antes de aparecerem nesta dissertação. A sua clareza de raciocínio e a sua sensibilidade só me fizeram crescer na vida.

Muito obrigado a todos vocês.

## RESUMO

Esta dissertação tem dois objetivos principais: 1) encorajar pesquisadores e pesquisadoras a trabalharem com Humanidades Digitais no estudo da literatura e 2) exemplificar essa abordagem de pesquisa mediante a análise de cerca de 150 romances brasileiros publicados no século XIX. Em síntese, as Humanidades Digitais correspondem a um método de pesquisa em que se aplicam *softwares* e linguagem de programação no estudo das ciências humanas, tais como literatura e história. Nesse sentido, o primeiro objetivo consistiu em discutir essa abordagem metodológica, a qual permite que se quantifiquem dados que, muito dificilmente, poderiam ser obtidos sem o auxílio de linguagem computacional. O segundo objetivo, por sua vez, compreendeu não só uma análise da geografia do romance brasileiro, como também a quantificação de alguns assuntos importantes para o Brasil do século XIX. Em conjunto, tratou-se da confecção de uma série de mapas interativos com a ocorrência de todos municípios citados nos enredos das narrativas e de um banco de dados com a presença de várias temáticas importantes para o Brasil oitocentista, quais sejam: escravidão, abolição, índios, sistemas de governo, classes sociais, sertão e nação. O uso de *Python*, um tipo de linguagem de programação, possibilitou que cerca de 150 romances fossem analisados em cerca de três meses, da feitura dos códigos à leitura das obras. A utilização desse método de pesquisa pode resultar na obtenção de dados mais objetivos, os quais podem servir para comprovar ou desmentir uma tese ou ideia nos estudos literários.

**Palavras-chave:** Humanidades digitais; Literatura brasileira; Análise quantitativa; Mapas interativos; Python.

## ABSTRACT

This dissertation has two main objectives: 1) to encourage researchers to work with Digital Humanities in the study of literature and 2) to exemplify this methodological approach through the analysis of approximately 150 Brazilian novels published in the 19th century. In summary, the Digital Humanities correspond to a research method in which software and programming languages are applied to the study of human sciences, such as literature and history. In this sense, our first objective consisted in discussing this methodological approach, which basically allows quantifying data that could hardly be obtained without the aid of computational language. Our second objective comprised not only the geographical analysis of the Brazilian novel, but also the quantification of some important themes that marked this country during the 19th century. Jointly, these two analyses generated not only a series of interactive maps with the occurrence of all municipalities mentioned in the novels, but also a database with the occurrence of several important themes for nineteenth-century Brazil. These themes are slavery, abolition, indigenous people, wilderness (“*sertão*”), social classes and nation. The use of Python, a type of programming language, made possible that about 150 novels were analysed in approximately three months, from the making of the codes to the reading of the novels. The use of this research method can result in the making of more objective data, which can serve to prove or disprove a thesis or idea in literary studies.

**Keywords:** Digital humanities; Brazilian literature; Quantitative analysis; Interactive maps; Python.



## LISTA DE FIGURAS

Figura 1	Demonstração do código.....	25
Figura 2	Demonstração do código.....	25
Figura 3	Demonstração do código.....	26
Figura 4	Demonstração do código.....	26
Figura 5	Demonstração do código.....	31
Figura 6	Demonstração do código.....	31
Figura 7	Demonstração do código.....	32
Figura 8	Demonstração do código.....	32
Figura 9	Demonstração do código.....	33
Figura 10	Demonstração do código.....	34
Figura 11	Demonstração do código.....	35
Figura 12	Demonstração do código.....	36
Figura 13	Demonstração do código.....	37
Figura 14	Demonstração do código.....	38
Figura 15	Demonstração do código.....	38
Figura 16	Demonstração do código.....	39
Figura 17	Demonstração do código.....	40

## LISTA DE GRÁFICOS

Gráfico 1	Quantidade de romances brasileiros analisados nesta pesquisa.....	44
Gráfico 2	Porcentagem da área do território brasileiro citada nos romances analisados.....	48
Gráfico 3	Algumas das cidades mais citadas nos romances do século XIX.....	82
Gráfico 4	Menções a Olinda e Recife no século XIX.....	83
Gráfico 5	População de alguns municípios ao longo do século XIX.....	84
Gráfico 6	População escravizada no Brasil ao longo do século XIX.....	87
Gráfico 7	Quantidade de romances em que a temática “escravidão” aparece.....	88
Gráfico 8	Quantidade média de menções à ideia de escravidão em romances do século XIX (dados preliminares).....	89
Gráfico 9	Quantidade média de menções à ideia de escravidão em romances do século XIX, exceto <i>As Vítimas-Algozes</i> (dados preliminares).....	90
Gráfico 10	Quantidade média de menções à ideia de escravidão em romances do século XIX, com alterações (dados finais).....	91
Gráfico 11	População escravizada no Brasil ao longo do século XIX.....	91
Gráfico 12	Quantidade de romances em que a temática “abolição” aparece.....	93
Gráfico 13	Quantidade média de menções à ideia de abolição em romances do século XIX.....	95
Gráfico 14	Comparação entre as médias de menções à escravidão e à abolição (com os dados finais sobre a escravidão).....	97
Gráfico 15	Quantidade de romances em que a temática “monarquia/parlamento” aparece.....	98
Gráfico 16	Quantidade de romances em que a temática “presidencialismo/federação/república” aparece.....	99
Gráfico 17	Quantidade média de menções à ideia de governo em romances do século XIX.....	101
Gráfico 18	Quantidade de romances em que a temática “estrato social superior” aparece.....	103
Gráfico 19	Quantidade de romances em que a temática “estrato social inferior” aparece.....	104

Gráfico 20	Quantidade média de menções à ideia de estratos sociais sociais em romances do século XIX.....	104
Gráfico 21	Quantidade de romances em que a temática “sertão” aparece.....	107
Gráfico 22	Quantidade média de menções à ideia de sertão em romances do século XIX .....	108
Gráfico 23	Quantidade de romances em que a temática “índios” aparece .....	110
Gráfico 24	Quantidade média de menções à ideia de índio em romances do século XIX (dados preliminares) .....	111
Gráfico 25	Quantidade média de menções à ideia de índio em romances do século XIX, exceto O Guarani (dados preliminares) .....	112
Gráfico 26	Quantidade média de menções à ideia de índio em romances do século XIX, com alterações (dados finais).....	112
Gráfico 27	Quantidade de romances em que a temática “nação” aparece .....	115
Gráfico 28	Quantidade média de menções à ideia de nação em romances do século XIX .....	116
Gráfico 29	Médias de todos os tópicos analisados, com as devidas alterações (escravidão e índios) .....	118
Gráfico 30	Médias de alguns tópicos analisados, com as devidas alterações (índios) .....	119
Gráfico 31	Médias de alguns dos tópicos analisados.....	120
Gráfico 32	Médias dos tópicos referentes à sociedade .....	121
Gráfico 33	Médias dos tópicos referentes ao estado.....	121

## LISTA DE MAPAS

Mapa 1	Série histórica completa (1840 - 1900).....	45
Mapa 2	Romances de José de Alencar.....	46
Mapa 3	Romances de Machado de Assis.....	47
Mapa 4	Série histórica completa (1840 - 1900).....	49
Mapa 5	Densidade populacional (1872).....	50
Mapa 6	Série histórica completa (1840 - 1900).....	51
Mapa 7	Economia brasileira (século XIX).....	51
Mapa 8	Série histórica completa (1840 - 1900).....	52
Mapa 9	As ferrovias construídas até 1898.....	53
Mapa 10	Série histórica por grupos (1840-1900).....	54
Mapa 11	Citações da cidade do Rio de Janeiro (1840 - 1900).....	55
Mapa 12	Detalhamento das citações à cidade do Rio de Janeiro (1840 - 1900).....	55
Mapa 13	Onda de 1840.....	57
Mapa 14	Onda de 1850.....	59
Mapa 15	Ocorrências do Norte no Romance <i>Simá</i> .....	60
Mapa 16	Onda de 1860.....	61
Mapa 17	Estado de São Paulo na Onda de 1850.....	62
Mapa 18	Estado de São Paulo na Onda de 1860.....	62
Mapa 19	Onda de 1870.....	65
Mapa 20	Área de fronteira na Onda de 1860.....	66
Mapa 21	Área de fronteira na Onda de 1870.....	66
Mapa 22	Estado de São Paulo na Onda de 1860.....	67
Mapa 23	Estado de São Paulo na Onda de 1870.....	68
Mapa 24	Onda de 1880.....	69
Mapa 25	Estado de São Paulo na Onda de 1870.....	70
Mapa 26	Estado de São Paulo na Onda de 1880.....	70
Mapa 27	O sudeste e o centro-oeste na Onda de 1870.....	71
Mapa 28	O sudeste e o centro-oeste na Onda de 1880.....	71
Mapa 29	Onda de 1890.....	74
Mapa 30	Estado de Minas Gerais na Onda de 1880.....	77
Mapa 31	Estado de Minas Gerais na Onda de 1890.....	77

Mapa 32	Nordeste na Onda de 1880.....	78
Mapa 33	Nordeste na Onda de 1890.....	79
Mapa 34	Onda de 1850.....	79
Mapa 35	Onda de 1860.....	80
Mapa 36	Onda de 1870.....	80
Mapa 37	Onda de 1880.....	80
Mapa 38	Onda de 1890.....	81

## SUMÁRIO

<b>APRESENTAÇÃO</b> .....	<b>13</b>
<b>1 INTRODUÇÃO</b> .....	<b>17</b>
1.1 OBJETIVOS DA PESQUISA.....	17
1.2 A DEFINIÇÃO, OS PRESSUPOSTOS E AS VANTAGENS DAS HUMANIDADES DIGITAIS .....	20
1.3 PROGRAMAÇÃO: A PARTE TÉCNICA DAS HUMANIDADES DIGITAIS.....	22
1.4 PENSAR COMO UM PROGRAMADOR É A PARTE MAIS DIFÍCIL DE SE TRABALHAR COM HUMANIDADES DIGITAIS .....	23
<b>2 DELIMITAÇÃO E METODOLOGIA</b> .....	<b>27</b>
2.1 A CURADORIA DOS TEXTOS .....	27
2.2 PRÉ-PROCESSAMENTO DOS DADOS .....	29
2.3 O FUNCIONAMENTO DE CÓDIGOS EM LINGUAGEM PYTHON.....	33
2.4 O CÓDIGO DA SEÇÃO 1: O LABORATÓRIO DE ANÁLISE DE DADOS (LAD) E A CRIAÇÃO DOS MAPAS .....	35
2.5 O CÓDIGO DA SEÇÃO 2: A EXPERIÊNCIA MAIS BEM SUCEDIDA.....	39
<b>3 RESULTADOS DA PESQUISA</b> .....	<b>42</b>
3.1 ALGUNS APONTAMENTOS PRELIMINARES SOBRE AS SEÇÕES 1 E 2.....	42
3.2 SEÇÃO 1: TERRITÓRIO, MUNICÍPIOS E MAPAS. ....	45
<b>3.2.1 Série histórica completa (1840 - 1899)</b> .....	<b>45</b>
<b>3.2.2 Mapas e gráficos: década a década</b> .....	<b>55</b>
3.2.2.1 Década de 1840 .....	56
3.2.2.2 Década de 1850 .....	58
3.2.2.3 Década de 1860 .....	61
3.2.2.4 Década de 1870 .....	64
3.2.2.5 Década de 1880 .....	68
3.2.2.6 Década de 1890 .....	73
<b>3.2.3 Cidades importantes em função do tempo</b> .....	<b>81</b>
3.3 SEÇÃO 2: A PRESENÇA DE TEMAS IMPORTANTES NOS ROMANCES .....	85
<b>3.3.1 Escravidão</b> .....	<b>86</b>

3.3.2	Abolição.....	92
3.3.3	Sistemas de governo .....	97
3.3.4	Estratos sociais.....	102
3.3.5	Sertão .....	106
3.3.6	Índios .....	110
3.3.7	Nação .....	114
3.3.8	Síntese .....	118
4	CONCLUSÃO .....	123
	REFERÊNCIAS .....	126
	GLOSSÁRIO .....	129
	APÊNDICE - LISTA DOS ROMANCES ANALISADOS .....	130

## APRESENTAÇÃO

Em 2017, um pouco antes de me graduar em Letras na Universidade Federal do Rio Grande do Sul (UFRGS), decidi estudar para o Concurso de Admissão à Carreira de Diplomata, também chamado de CACD. Além do prestígio e da responsabilidade da profissão de diplomata, considerei que essa escolha me permitiria aprofundar meus conhecimentos em história, línguas estrangeiras, direito, economia e literatura, algumas das disciplinas que mais me chamaram a atenção durante a graduação. Como todo aluno entusiasmado por essa interdisciplinaridade, textos como “O Direito à Literatura”, de Antonio Candido, ou “Shakespeare e a economia”, de Gustavo H.B. Franco, sempre me interessaram.

De 2018 a 2021, dediquei todo o meu tempo a estudar para o CACD. Foram anos de muito aprendizado, dedicação e frustrações. Segundo alguns professores especializados nessa preparação, o concurso é tão exigente que é como se os candidatos precisassem ser mestres em todas as disciplinas avaliadas para que sejam aprovados no certame. Embora essa afirmação seja discutível, creio que o nível de comprometimento de um candidato ao CACD e o de um aluno de mestrado são parecidos, sobretudo porque se aprende muito sobre si mesmo durante ambos os percursos. Por se tratar de um exame muito tradicional, estudar para o CACD significa seguir um método de estudo também bastante convencional, o qual consiste em ler textos, analisar obras e reformular ideias, em síntese.

Na quinta e última vez em que participei do certame, finalmente obtive nota para ser aprovado. Como havia poucas vagas disponíveis naquele ano, no entanto, não fui convocado para ingressar como Terceiro Secretário no Instituto Rio Branco. Ao contrário do que ocorreu após as tentativas anteriores, não tive vontade de prestar o concurso mais uma vez. Tratava-se mais de uma rejeição à forma de estudo que ao conteúdo exigido, o qual nunca deixou de me interessar. Na verdade, senti que já tinha alcançado meu objetivo: havia aprofundado meus conhecimentos nas áreas que mais me despertaram curiosidade desde que tinha ingressado na universidade. Assim, embora não tenha me tornado diplomata, tenho certeza de que esses anos de estudo foram fundamentais para a minha vida pessoal e intelectual.

Em 2021, aproximadamente cinco anos depois da minha cerimônia de graduação, decidi continuar meus estudos universitários. Como já disse anteriormente, nunca deixei de me interessar por línguas, literatura e história. Assim, quando me deparei com a oportunidade de ingressar no Programa de Pós-Graduação da Letras da UFRGS, sabia que tomava a decisão certa. Entretanto, também sabia que, depois de anos de estudos para o Itamaraty, a forma



tradicional de pesquisar as humanidades não me satisfazia mais: não queria apenas ler e analisar obras literárias e teóricas a fim de formular respostas para alguma pergunta. Esses procedimentos são, sem dúvidas, importantes para toda e qualquer pesquisa, mas eu havia operado daquela forma por muito tempo ao estudar para o CACD. Minha pesquisa de mestrado começou, portanto, mais pelo método que pelo conteúdo.

Como fazia alguns anos que eu estava afastado do mundo universitário, quis me informar sobre áreas de destaque nas Humanidades. Após acessar alguns sites, sobretudo de universidades britânicas, deparei com o termo “*Computational Linguistics*”, uma novidade para mim. Embora também se use o termo “Linguística de Corpus” em língua portuguesa, parece-me que a expressão “Linguística Computacional” seja mais adequada. Entendi que essa área de estudos proporciona um novo método (computação) para analisar um mesmo conteúdo (linguística). Inicialmente, no entanto, fiquei com a impressão de que a aplicação da computação se limitava apenas à linguística. Assim, tive que pesquisar um pouco mais para encontrar a área que também compreendesse o estudo da literatura, ou seja: as *Digital Humanities* ou as Humanidades Digitais.

Essa área de estudos era a novidade da qual eu precisava, justamente para que eu pudesse utilizar um método de pesquisa diferente do tradicional. Além disso, como eu precisaria aprender a programar em linguagem de computação, a ideia de desenvolver tal habilidade me despertou curiosidade novamente. De fato, aprender a programar em Python foi uma das coisas mais interessantes e desafiadoras que já fiz na vida. É bem verdade que cada linha de código demanda algumas horas de trabalho, mas a sensação de se executar, com um comando apenas, uma tarefa que demoraria semanas, talvez meses, para ser realizada é extremamente satisfatória.

Como se pode ver, defini o tema da minha pesquisa apenas depois de encontrar o método com o qual trabalharia. Na verdade, selecionar um assunto me parecia mais simples, já que, desde a iniciação científica, sempre pesquisei a história da literatura brasileira, com ênfase nos processos históricos que nela se manifestam. No meu TCC, também orientado pelo professor Luís Augusto Fischer, estudei como se representava o capitalismo na obra de Aluísio Azevedo, mais especificamente em três romances desse autor: *O Mulato*, *Casa de Pensão* e *O Cortiço*. Naquele então, no entanto, nem imaginava poder aumentar exponencialmente o escopo da minha pesquisa. Não surpreende que um dos objetivos deste mestrado tenha sido, justamente, o de ampliar muito a quantidade de obras que normalmente são examinadas em um trabalho sobre literatura.

Além de apresentar os resultados da minha pesquisa, também espero cumprir outro objetivo importante com esta dissertação: facilitar o percurso formativo daqueles pesquisadores que queiram trabalhar com Humanidades Digitais, notadamente na área da literatura<sup>1</sup>. Embora existam muitos cursos sobre programação, poucos tratam de literatura, de modo que cabe ao pesquisador aprender a usar a programação para operar nessa área. Por essa razão, decidi redigir este trabalho como se estivesse escrevendo um manual, um livro-texto sobre o assunto. Assim como acontece com todo o bom livro-texto, espero ser didático, claro e sucinto nas minhas análises. Por fim, esse trabalho não é, vale ressaltar, um manual de como aprender a programar. O que me proponho é exemplificar e explicar o que um pesquisador interessado em Humanidades Digitais pode fazer.

Finalmente, após a defesa desta dissertação, os membros da banca recomendaram que eu explicitasse e aprofundasse alguns aspectos importantes para este trabalho, sobretudo em função das novidades que ele traz. Em primeiro lugar, gostaria de reiterar que esta dissertação de mestrado representa um trabalho inicial, uma espécie de tomada de consciência do que pode ser realizado ao se analisar literatura por meio da ciência da computação. Em outras palavras, tratou-se de entender as potencialidades das Humanidades Digitais. Como a banca muito corretamente assinalou, esta dissertação poderia ter consistido no aprofundamento de um ponto específico dentre todos os temas abordados nesta pesquisa. De fato, eu poderia ter dedicado uma análise mais extensa ao tópico “escravidão”, por exemplo. Desta forma, não só as discussões epistêmicas em relação às Humanidades Digitais, como também a própria escolha dos vocábulos escolhidos para a análise dos temas poderiam ter sido melhor discutidas. A continuação deste trabalho, que vai se materializar em um futuro doutorado, certamente servirá para aprofundar as discussões epistêmicas sobre as Humanidades Digitais, uma vez que a digitalização de um texto acarreta uma série de consequências para o pesquisador e seu objeto de estudo. Gostaria de agradecer imensamente os membros da banca pelos aportes teóricos e metodológicos que pude receber durante a defesa deste mestrado.

---

<sup>1</sup> É claro que muitos outros pesquisadores já trabalharam nessa área, mas ainda há muito a ser feito. Gostaria de mencionar, especificamente, o trabalho de Henrique Cardoni, que trabalhou com a quantificação da frequência de palavras a respeito do rock do Rio Grande do Sul nos anos 1960, 70 e 80.

# 1 INTRODUÇÃO

## 1.1 OBJETIVOS DA PESQUISA

O objetivo desta pesquisa foi examinar cerca de 150 romances a fim de 1) estudar a geografia do romance oitocentista brasileiro e 2) verificar a presença de algumas temáticas importantes nos romances daquela época. Para tratar desses pontos, dois códigos em linguagem Python foram desenvolvidos: o primeiro quantificou a menção a municípios nos enredos dos romances; o segundo verificou a presença de temas como escravidão, abolição, monarquia e república nas obras. Além disso, mapas e gráficos foram criados com o objetivo de apresentar os dados de maneira interativa, os quais podem ser acessados autonomamente por todos os leitores. Os links para os mapas podem ser encontrados no final do trabalho.

É importante observar, no entanto, que nem todos os dados obtidos durante a pesquisa foram usados para as análises desta dissertação, sobretudo em razão do elevado volume de material que foi gerado para os fins de uma dissertação de mestrado. De toda a forma, ainda que não tenham sido explicitamente apresentados no corpo do trabalho, reiteramos que mesmo esses resultados estão disponíveis online.

De início, ainda vale ressaltar que, em razão da grande quantidade de textos analisados, este trabalho não teria sido exequível sem o uso de Python, uma das várias linguagens de programação que existem. Nesta introdução, é suficiente dizer que as Humanidades Digitais podem ser caracterizadas como a área de pesquisa que permite estudar as ciências humanas por meio de linguagem de programação ou *software*. Embora tivesse sido possível ler centenas de romances manualmente, eu provavelmente não teria conseguido concluir este mestrado em 2 anos. Com efeito, se multiplicarmos os 150 romances por um número aproximado de 200 páginas por obra, podemos considerar que 30.000 páginas foram examinadas ao longo desta pesquisa. Em capítulo posterior, ainda tratarei da parte técnica da pesquisa de maneira detalhada.

Quando comecei a pensar nesta pesquisa, tinha a pretensão de trabalhar com todos os romances brasileiros publicados entre os anos de 1830 e 1900. No entanto, em função de problemas técnicos dos quais também tratarei oportunamente, não tive acesso a todas as obras, de modo que precisei reduzir a quantidade de textos examinados. De toda forma, creio ter feito a curadoria de um número representativo relativamente ao número total de obras publicadas naquele século. Em outras palavras, espero ter conseguido trabalhar com o que o

crítico Franco Moretti denominou o “*great unread*”, ou seja: os 99% das obras que não fazem parte do cânone literário de um país.<sup>2</sup>

Do ponto de vista teórico, dois livros foram particularmente importantes para a formulação das perguntas que embasaram esta dissertação: *Duas formações, uma história*, de Luís Augusto Fischer, e *Distant reading*, de Franco Moretti. Não obstante ambos os livros tratem de história da literatura, Fischer atenta para a necessidade de se rediscutir a história da literatura brasileira, ao passo que Moretti sugere o uso de métodos mais objetivos, tais como mapas, gráficos e tabelas, para se estudar essa disciplina. O que eu tentei fazer foi combinar as duas ideias, isto é: estudar um gênero literário, mais precisamente o romance oitocentista brasileiro, por meio de gráficos, tabelas e mapas, os quais desenvolvi mediante linguagem de programação. Como se pode notar, a novidade deste trabalho está, sobretudo, no uso das Humanidades Digitais como método de pesquisa literária.

O professor Luís Augusto Fischer argumenta que estudar história da literatura significa, também, entender a espacialidade das formas literárias, isto é: o “onde” também é tão importante quanto o “quando”. De fato, as formas literárias que se desenvolveram na Zona da Mata nordestina são diferentes das que se desenvolveram no Pampa gaúcho, como exemplificam a gauchesca e o cordel, respectivamente. Com efeito, em *Duas formações, uma história*, Fischer chama a atenção para a extensão territorial do Brasil, a qual dificulta qualquer generalização a respeito da história da literatura do País. Ao falar sobre uma nova história da literatura, Fischer pergunta:

E como fica o mundo amazônico? Será ele uma outra coisa no horizonte, uma outra formação histórica existente dentro do Brasil, distinta o suficiente das outras duas, Plantation e Sertão? Ouro Preto, Belo Horizonte, Goiás Velho e Goiânia, mais Brasília e outras partes, são claramente nascidas no sertão; mas qual o estatuto de Porto Alegre? E Fortaleza e Olinda/Recife? (FISCHER, 2021, p. 341)

Para a minha pesquisa, essa preocupação com aspectos geográficos é essencial, como se poderá ver nos mapas que foram desenvolvidos. Com base nessa indagação, dispus-me a encontrar e quantificar todas as menções a municípios nos romances brasileiros do século XIX, o que possibilitou entender a maneira como os autores e o público leitor imaginavam o território nacional ao longo desse século.

A obra de Franco Moretti, por sua vez, é marcada pela tentativa de propor novos métodos de estudo para a história da literatura, sobretudo ao desenvolver o conceito de *distant*

---

<sup>2</sup> O conceito de “*the great unread*” foi desenvolvido por Franco Moretti em um artigo chamado “Conjectures on World Literature”.

*reading*. Com efeito, em “*Distant reading*”, o crítico italiano afirma que, ao se analisar várias obras, outras ferramentas analíticas devem ser usadas, uma vez que não se pode almejar o mesmo tipo de conhecimento que se tem mediante leituras individuais:

Knowing two hundred novels is already difficult. Twenty thousand? How can we do it, what does 'knowledge' mean, in this new scenario? One thing for sure: it cannot mean the very close reading of very few texts-secularized theology, really ('canon!') (...). A larger literary history requires other skills: sampling; statistics; work with series, tides, concordances, incipits-and perhaps also the 'trees' that I discuss in this essay. (MORETTI, 2015, p. 67)

Para explicitar as novidades de seu método de pesquisa, Moretti diferencia dois tipos de leituras: *close reading* e *distant reading*. A leitura em *close reading* permite que se estudem aspectos sutis, únicos ou particulares de uma obra, tais como a construção psicológica dos personagens ou a interpretação de uma frase ou parágrafo, por exemplo. A leitura em *distant reading*, a mais comum no trabalho do crítico italiano, possibilita um outro tipo de conhecimento sobre a literatura. A ideia de analisar a presença de temas importantes para a história do Brasil com base no romance do século XIX decorre justamente desse “outro tipo de conhecimento” de que trata Moretti. De fato, não se trata de interpretar passagens individuais ou explicar o excepcionalismo de alguma obra, mas de observar aspectos macroestruturais do romance brasileiro do século XIX.

Ainda que de maneira sucinta, considero importante ressaltar que as Humanidades Digitais permitem outro tipo de indagação sobre a história da literatura: o trabalho com o *big data* (grande quantidade de dados). Em *Exploring Big Historical Data*, os professores Shawn Graham, Ian Milligan e Scott Weingart argumentam que essa área de estudos possibilita considerar mais dados e informações para a reconstrução do passado, objetivo principal de qualquer historiador. Os autores, no entanto, alertam que:

“Digital history does not offer direct truths, but only new ways of interpreting and understanding traces of the past. More traces, yes, but still traces: brief shadows of things that were”. (GRAHAM, 2016, p. 33)

Como se pode perceber, as ferramentas de análise digital são apenas mais um recurso de que dispomos para estudar a história. Apesar de breve, esse aviso é importante para que se entendam os limites das Humanidades Digitais. Esse aspecto será tratado em mais detalhes oportunamente.

Em resumo, o que se propõe neste trabalho é materialmente simples e tecnicamente complexo. A simplicidade está justamente nas duas questões que me propus a examinar: 1)

entender e dispor em mapa a geografia do romance oitocentista brasileiro e 2) quantificar a presença de temas importantes para o século XIX nos romances. A complexidade, por sua vez, está na quantidade de obras e na velocidade com que encontrei as informações: quase 150 romances (o mais perto que cheguei do “*great unread*”) examinados em cerca de 3 meses, dos primeiros códigos ao exame efetivo dos textos via Python. Dessa forma, espero que esse trabalho possa ajudar e inspirar pesquisas ulteriores, na medida em que sirva como fonte de dados primários e como exemplo de pesquisa em Humanidades Digitais na área da literatura.

## 1.2 A definição, os pressupostos e as vantagens das Humanidades Digitais

*Digital Humanities* ou Humanidades Digitais é a área de pesquisa em que se estudam as ciências humanas por meio de linguagem de programação ou *software*. A distinção entre linguagem de programação e *software* ainda não é importante, de modo que podem ser entendidos como sinônimos por enquanto. Dessa forma, as Humanidades Digitais possibilitam a intersecção entre a ciência da computação e a história, geografia ou literatura, por exemplo. Em síntese, o pesquisador trabalha com uma quantidade tão grande de dados que provavelmente não conseguiria analisar sem o auxílio de *software*. Embora essa área de pesquisa ainda seja pouco conhecida no Brasil, já existem vários centros, instituições e universidades em que se trabalha com isso, sobretudo na Europa Ocidental e nos Estados Unidos.

Existem, basicamente, duas formas de se trabalhar com as Humanidades Digitais. A primeira consiste em usar dados já conhecidos e, por meio de *software*, realçar um aspecto difícil de ser apreendido manualmente<sup>3</sup>. O *software*, como se pode intuir, não é usado para obter os dados, mas somente para analisá-los. Por exemplo: a fim de estudar a história do Brasil, o pesquisador poderia acessar quantos votos cada presidente recebeu nas eleições da Primeira República e, mediante linguagem de programação, examinar a existência de alguma relação entre a quantidade de votos e as circunscrições eleitorais. A segunda forma consiste em usar o *software* para obter dados que ainda não existem. Um pesquisador poderia quantificar e organizar a ocorrência dos discursos direto, indireto e indireto livre nos romances brasileiros do século XIX, por exemplo. Como esse dado ainda não existe, tratar-se-

---

<sup>3</sup> Neste trabalho, a palavra “manualmente” deve ser entendida como “sem o uso de *softwares*/linguagem computacional”.

ia de uma pesquisa cujo objetivo principal é a própria obtenção dos dados. Evidentemente, nada impede que também se tente analisá-los em um segundo momento.

Uma das vantagens de se trabalhar com as Humanidades Digitais é que qualquer aspecto quantificável pode se tornar um dado. Conseqüentemente, pressupõe-se que mesmo materiais tão subjetivos como textos literários possam ser objetificados de alguma forma. Suponhamos que o pesquisador desejasse observar quais são as palavras mais usadas na poesia setecentista brasileira. Em teoria, existem duas possibilidades para se realizar essa tarefa: ele pode ou ler manualmente todos os textos, anotando em uma lista as palavras mais frequentes; ou desenvolver um código que execute essa tarefa automaticamente. Como é extremamente improvável que alguém se disponha a fazer esse trabalho manualmente, apenas o uso do código pode efetivamente transformar a frequência das palavras em um dado.

Um outro benefício das Humanidades Digitais é a diminuição do tempo necessário para se completar uma tarefa com maior precisão. De fato, como acabei de mostrar, praticamente qualquer operação que se pretenda realizar por meio de *software* pode ser feita manualmente. Embora possa levar muito tempo, é possível ler todos os relatórios do Conselho de Estado do Império do Brasil com o objetivo de contabilizar quantas vezes a palavra “soberania” foi usada, por exemplo. Por meio de um *software*, no entanto, pode-se levar a cabo a mesma tarefa em minutos ou horas. Além disso, a leitura manual de vários documentos pode não ser tão eficiente quanto o exame realizado via código. Com efeito, ao contrário de seres humanos, *softwares* não cansam nem erram, conquanto tenham sido escritos corretamente, como veremos em um próximo capítulo.

Como se pode ver, a linguagem de computação permite que o pesquisador não só economize tempo, como também o use para tarefas mais interpretativas ou criativas. A respeito do exemplo anterior - a quantificação da palavra “soberania” nos relatórios do Conselho de Estado -, a mera quantificação dessa palavra não significa nada, em princípio. A interpretação analítica dos dados é, e provavelmente sempre será, competência exclusiva do pesquisador-programador. Os códigos apenas facilitam a execução das tarefas preponderantemente manuais.

As Humanidades Digitais são fundamentais não só na formulação de novas respostas a discussões antigas, como também na proposição de novos questionamentos para os estudos sobre literatura. Essa área permite ampliar a maneira como, tradicionalmente, se provam hipóteses nas ciências humanas, uma vez que se fornecem dados mais objetivos para corroborar uma tese ou ideia. De todas as disciplinas das humanidades, parece-me que a área dos estudos literários é a que tem sido, historicamente, menos influenciada pelas

Humanidades Digitais. Mais um exemplo: a fim de estudar o romantismo, poder-se-ia quantificar os temas mais presentes nas poesias do mal-do-século. Assim, seria possível corroborar teses sobre literatura de maneira mais objetiva, assim como já se faz mais frequentemente na história, geografia ou ciências sociais.

Por fim, é importante notar que trabalhar com Humanidades Digitais não pode jamais significar o esvaziamento de um dos aspectos que marca a literatura: a subjetividade. Com efeito, a programação é apenas um método possível para se estudar literatura. No já citado *Exploring Big Historical Data*, trata-se das dificuldades de se transformar as informações disponíveis em narrativas históricas:

Even with massive arrays of data, historians do not simply cut and paste findings from computer databases; such an approach would be evocative of the “scissors and paste” model noted by historian R.G. Colling wood. Having more data is not a bad thing. With more data, in practical terms, there is arguably a higher likelihood that historical narratives will be closer in accordance with past events, as we have more traces to base them on. (GRAHAM, 2016, p. 32)

Nesse sentido, as narrativas históricas, que organizam logicamente fatos e dados, sempre precisarão da elaboração ulterior do pesquisador-programador. O que as ferramentas digitais permitem é a obtenção de mais informações para que se possa verificar e comprovar melhor interpretações e pontos de vista sobre a área estudada.

### 1.3 Programação: a parte técnica das Humanidades Digitais

A pesquisa em Humanidades Digitais depende de conhecimento em programação. Embora se possa pensar que as linguagens de programação são como toda e qualquer linguagem, programar é uma tarefa extremamente difícil, sobretudo para um estudante de ciências humanas, que normalmente não desenvolve esse tipo de competência na universidade. Por um lado, é verdade que existem alguns aspectos em comum entre as línguas naturais e as linguagens de programação, notadamente a existência de símbolos finitos que permitem expressar praticamente qualquer ideia. Por outro lado, contrariamente às línguas naturais, as linguagens de programação não permitem nenhum tipo de subjetividade. Essa diferença é essencial, uma vez que o programador precisa entender que nada pode ficar subentendido ao se escrever um código. Se o programador não for completamente objetivo e claro, como veremos em capítulo posterior, o código não realizará a função para a qual foi designado.



Neste momento, é importante diferenciar *machine learning* dos módulos computacionais já existentes. Também chamados de bibliotecas (*libraries*, em inglês), esses módulos foram desenvolvidos por outros programadores e estão disponíveis para o uso de qualquer um. Normalmente, essas bibliotecas já apresentam as funcionalidades necessárias para se trabalhar com Humanidades Digitais. O desenvolvimento de *machine learning*, por outro lado, demanda não só muita experiência em programação, como também computadores extremamente potentes (muita memória RAM), uma vez que se objetiva treinar a máquina para observar e emular aspectos mais sutis do comportamento humano. Um exemplo atual de *machine learning* é o ChatGPT, um *software* cuja finalidade é, em síntese, alcançar um nível de comunicação comparável ao de um ser humano.

As bibliotecas já existentes possibilitam executar tarefas importantes para o pesquisador de literatura, tais como buscar palavras específicas ou as mais comuns, quantificar a ocorrência de alguma expressão ou identificar o uso de uma sintaxe particular em um texto. O *machine learning*, por sua vez, permite outro tipo de automação, uma vez que possibilita a criação de códigos para lidar com problemas complexos de linguagem, como estabelecer campos semânticos para a obra de algum autor, por exemplo. De maneira simplificada, podemos afirmar que aspectos formais podem ser tratados pelo uso de módulos já existentes, ao passo que aspectos semânticos demandam o uso de *machine learning*.

Embora programar seja difícil, existe uma grande comunidade de programadores na internet. Assim, pode-se pedir ajuda a pessoas mais experientes que, normalmente, encontram a resposta para o problema, contanto que quem tenha perguntado conheça um pouco de linguagem de programação, a fim de poder descrever o código que não está funcionando. Os sites<sup>4</sup> são, em geral, gratuitos, e esses programadores não só propõem soluções para a dificuldade, como também explicam o funcionamento dos códigos. Na verdade, pode-se aprender muito dessa forma. Vale notar, ainda, que esses materiais são normalmente escritos em língua inglesa, uma vez que as próprias bibliotecas foram redigidas nessa língua.

#### 1.4 Pensar como um programador é a parte mais difícil de se trabalhar com Humanidades Digitais

Em essência, programar é dar instruções para que um computador execute um determinado comando. Os códigos, por sua vez, são não somente a materialização dessas

---

<sup>4</sup> O mais importante é, sem dúvida, o *stackoverflow*, cujo site é <https://stackoverflow.com/>.

instruções, como também a única linguagem que um computador compreende. Ora, dar instruções a uma máquina é uma tarefa complexa, uma vez que computadores não toleram nenhum tipo de ambiguidade ou imprecisão na comunicação. Dessa forma, a parte mais difícil de se trabalhar com Humanidades Digitais não é efetivamente escrever os códigos, mas desenvolver o raciocínio de programador. Como trabalhei com Python, tratarei do desenvolvimento de uma visão *pythoniana* a seguir.

Programadores experientes costumam dizer que linguagens de programação nunca erram, porquanto apenas executam os códigos desenvolvidos por pessoas. Com efeito, ao trabalhar com Python, o pesquisador de literatura precisa conhecer não só literatura e linguística, mas também como manipular dados e informações nessa linguagem de programação. Como cada projeto é composto por várias etapas - que podem compreender a confecção de uma tabela, mapa, *tuple*<sup>5</sup>, base de dados, por exemplo - o pesquisador também precisa conhecer a melhor forma de manipular os dados específicos a cada projeto. Assim, precisa conhecer com quais tipos de arquivo os resultados obtidos podem ser armazenados: *txt*, *excel* ou *csv*, por exemplo.

Para se desenvolver uma visão *pythoniana*, é preciso entender que os códigos de programação representam a explicitação máxima de uma linguagem artificial. Até começar a programar, eu não tinha consciência do quão potente e eficiente o cérebro humano é. Com efeito, a efetividade da comunicação humana depende de muitas informações implícitas, o que pode ser entendido a partir do exemplo seguinte: durante o café da manhã com sua filha, o pai diz “passa a manteiga no pão, filha”. Depois de rapidamente processar a instrução do pai, a filha 1) pega o pacote de pão, 2) abre o pacote, 3) retira o pão de dentro do pacote, 4) põe o pão no prato, 5) abre o pote de manteiga, 6) pega a faca com uma mão e passa na manteiga, 7) pega o pão com a outra mão e, finalmente, 8) passa a manteiga no pão. Ainda que não se tenham explicitado todas as ações necessárias, podemos afirmar que muitas informações estavam implícitas na instrução do pai (“passa a manteiga no pão”). Em síntese, o que impede a filha de pegar o pote de manteiga fechado e esfregá-lo no pacote de pão, também fechado, é a capacidade que seu cérebro tem de processar a linguagem humana. Uma linguagem de programação não opera dessa forma.

A capacidade que um computador tem de interpretar informações implícitas é extremamente limitada, sobretudo quando comparada à humana. Assim, convém que a primeira lição de um programador iniciante seja entender que um computador executa somente as instruções que lhe são explicitamente dadas. Na prática, significa afirmar que o

---

<sup>5</sup> Uma forma de armazenar dados em Python.

programador não deve culpar o computador pelo insucesso de um código, uma vez que o computador apenas realizou o que lhe foi pedido<sup>6</sup>. Para exemplificar as consequências dessa primeira lição, criei um código muito simples cujo único objetivo é o de verificar a existência da palavra “ateneu” no romance *O Ateneu*. Para realizar essa tarefa, informei ao Python que se (if) aquela palavra constasse (in) no romance (*texto\_tokenized*), a frase “O python encontrou essa palavra no texto” deveria aparecer como o resultado da operação. De, foi exatamente o que aconteceu:

Figura 1 — Demonstração do código.

```
1 if "ateneu" in texto_tokenized:
2     print('O Python encontrou essa palavra no texto')
O Python encontrou essa palavra no texto
```

Fonte: Elaborado pelo autor (2023)

Como já se podia esperar, a palavra “ateneu” aparece várias vezes naquele romance. Podemos perguntar o que aconteceria caso se mudasse a palavra “ateneu” para “artefato”, por exemplo, ou qualquer outra palavra aleatória. Para um ser humano, a resposta é mais ou menos imediata: se a palavra “artefato” constasse no romance, o Python deveria mostrar a frase precedente no resultado da operação; do contrário, deveria apresentar algum resultado como “O Python não encontrou essa palavra no texto”. Vejamos o que acontece na prática:

Figura 2 — Demonstração do código.

```
1 if "artefato" in texto_tokenized:
2     print('O Python encontrou essa palavra no texto')
```

Fonte: Elaborado pelo autor (2023)

Como se pode ver acima, não acontece nada: o Python simplesmente executa o código e não mostra nenhum resultado. A razão é simples: o programador não instruiu o computador a executar alguma ação na ausência de uma palavra no texto. Foi exatamente isso que a linguagem de programação fez. Para que ele dispusesse outra mensagem, o programador precisaria instruí-lo, notadamente. Nesse sentido, ao adicionar a linha *else*, eu instruí o Python a me fornecer outra informação na ausência de uma palavra no texto. Ao buscar uma palavra que existe, vejamos o que acontece:

<sup>6</sup> É claro que eventuais omissões podem aparecer na prática, mas o problema pode estar no formato de um arquivo ou em outras incompatibilidades, por exemplo.

Figura 3 — Demonstração do código.

```
1 if "ateneu" in texto_tokenized:
2     print('O Python encontrou essa palavra no texto')
3 else:
4     print('O Python não encontrou essa palavra no texto')
```

O Python encontrou essa palavra no texto

Fonte: Elaborado pelo autor (2023)

O resultado é o mesmo que anteriormente, uma vez que a palavra “ateneu”, de fato, consta no romance. Em seguida, ao buscar uma palavra que não está no texto (“artefato”), o resultado é diferente, porquanto agora também a instrução foi modificada:

Figura 4 — Demonstração do código.

```
1 if "artefato" in texto_tokenized:
2     print('O Python encontrou essa palavra no texto')
3 else:
4     print('O Python não encontrou essa palavra no texto')
```

O Python não encontrou essa palavra no texto

Fonte: Elaborado pelo autor (2023)

Espero que tenha ficado suficientemente claro que, ao desenvolver códigos, o pesquisador/programador deve ter em mente a necessidade de explicitar ao máximo as instruções fornecidas ao computador. Do contrário, os códigos ou não funcionarão ou funcionarão parcialmente. Em síntese, a lição mais importante é que o computador só executa o que lhe foi solicitado, nada mais, nada menos.

## 2 DELIMITAÇÃO E METODOLOGIA

### 2.1 A curadoria dos textos

A despeito da pretensão inicial de examinar todos os romances brasileiros publicados no século XIX, cerca de 150 romances foram analisados ao longo deste trabalho. Na verdade, o número exato de obras efetivamente estudadas foi 140. Como não se sabe ao certo quantos romances foram publicados ao longo do século XIX, o meu primeiro objetivo era encontrar a maior quantidade de textos possível, ou seja: quantificar uma espécie de “*great unread*” do romance brasileiro do século XIX, para usar a expressão consagrada por Franco Moretti. O site da Biblioteca Digital de Literatura de Países Lusófonos (FLUC)<sup>7</sup>, sob curadoria da Universidade Federal de Santa Catarina, foi a principal ferramenta de busca utilizada para realizar a curadoria dos textos. Parece-me que esse site contém o mais completo acervo digital de obras disponível em língua portuguesa, já que integra uma série de outros repositórios, com especial menção ao da Biblioteca Brasileira Guita e José Mindlin, organizado pela USP.

Infelizmente, ainda preciso dizer que não mantive o registro de todas as operações de busca realizadas durante a curadoria dos textos. Desta forma, o que vou relatar a partir dos próximos parágrafos é a história aproximada de como 140 romances acabaram sendo contabilizados nesta dissertação. Ademais, o processo que descreverei não foi feito de uma só vez, de maneira linear. Não creio, no entanto, que essas eventuais omissões desmereçam o mérito da pesquisa, na medida em que todos esses romances foram, de fato, considerados nos resultados finais.

O site da FLUC foi o ponto de partida para a curadoria dos textos, na medida em que a ferramenta de busca permite filtrar por critérios como gênero literário, idioma original, ano de publicação e local da editora, por exemplo. Ao selecionar as opções que satisfaziam os critérios da minha pesquisa (ou seja: romance/novela, português, século XIX e todos, respectivamente), obtive 358 obras para a consulta. O problema é que nem todas essas obras estavam digitalizadas, isto é: o banco de dados do site também apresentava resultados de obras cujos textos não estavam disponíveis para *download*. Como o acesso aos textos dos romances era imprescindível para a minha pesquisa, não podia trabalhar com todos aqueles resultados. A curadoria tinha que continuar, portanto.

---

<sup>7</sup> O link para esse site é: <https://www.literaturabrasileira.ufsc.br/busca/documento/>

A fim de encontrar as obras que pudessem ser consultadas, ativei a opção “apenas obras digitalizadas”, o que restringiu de 358 para 253 o número de textos disponíveis. Conseqüentemente, podemos constatar que 105 obras estão catalogadas, mas indisponíveis para *download*. De toda forma, 253 romances ainda compunham uma quantidade representativa de textos. Infelizmente, havia ainda um outro problema: como o filtro “local da editora” tinha sido selecionado para mostrar tanto editoras nacionais quanto estrangeiras, havia vários autores portugueses ou eventuais estrangeiros que tivessem escrito em língua portuguesa incluídos nessas obras. Essencialmente, o problema é que não se podia filtrar pela nacionalidade do autor, o que dificultava a curadoria dos textos para este trabalho, já que o objetivo era estudar apenas autores brasileiros.

Quando me dei conta desse problema, procedi a uma leitura manual de cada resultado obtido no site, com o objetivo de ver se o autor era ou não brasileiro. Foi nesse momento que comecei a buscar em vários acervos digitais diferentes e, conseqüentemente, acabei perdendo o controle do que constava e do que não constava especificamente no site da FLUC. De toda a forma, ao selecionar a opção “editoras portuguesas” na ferramenta de busca, 145 obras foram disponibilizadas para *download*. Ao subtrair 253 de 145, restam 108 romances: 32 romances a menos que os 140 efetivamente analisados. Assim, podemos supor que esses 33 romances ou dizem respeito a autores brasileiros cujas obras foram publicadas em Portugal ou a romances que consegui em outros arquivos, pois não estavam disponíveis no site da FLUC. A tabela que apresenta todos os romances examinados está anexada no final deste trabalho.

Uma das principais questões metodológicas deste trabalho foi a quantidade de obras examinadas, uma vez que se pretendeu analisar o “*great unread*” do romance oitocentista brasileiro. Ora, quanto maior for o número de romances examinados, mais relevantes tendem a ser as análises. Dessa forma, se considerarmos 358 como a quantidade total de romances brasileiros publicados no século XIX, os 140 romances examinados representam aproximadamente 40% de todos os romances da época. Se, no entanto, considerarmos 253 como a quantidade total, os 140 romances passam a representar 55%. Apesar de expressivos, esses números já indicam que esta pesquisa ainda pode continuar no futuro, quer por mim em um doutorado, quer por outro pesquisador.

## 2.2 Pré-processamento dos dados

Como já havia mencionado anteriormente, utilizei Python (linguagem de programação) e Jupyter Notebook (ambiente de execução de códigos) para gerar e processar os dados. Em resumo, uma linguagem de programação é um método padronizado mediante o qual se escrevem os códigos, ao passo que um ambiente de execução é o local virtual no qual esses códigos podem operar, isto é: a interface em que os resultados aparecem para o programador. Ademais, o Python e o Jupyter Notebook foram escolhidos em razão da simplicidade e interatividade desses dois *softwares*. Com efeito, para programadores inexperientes, sugiro fortemente que se trabalhe com a combinação dessas duas ferramentas.

Neste momento, convém diferenciar um *software* de uma linguagem de programação. Um *software* é um programa cujas funcionalidades são bem delimitadas. O *Qgis*, por exemplo, é um *software* que permite trabalhar com mapas e coordenadas de maneira simples e intuitiva. Uma linguagem de programação, por sua vez, não tem funcionalidade bem definida, uma vez que é um conjunto de normas sintáticas que permite escrever códigos para executar as mais variadas tarefas. Com efeito, o Python serve para realizar infinitas funções, de cálculos aritméticos a desenvolvimento de aplicativos. Em verdade, um software é composto pela combinação de vários códigos. É importante dizer que tanto a linguagem de programação Python quanto *softwares* como *Convertio* e *Qgis* foram usados neste trabalho.

Depois da curadoria dos textos, começa a etapa de pré-processamento dos dados, uma vez que muito dificilmente se pode trabalhar com os textos sem realizar nenhuma modificação nos arquivos. Para a minha pesquisa, a primeira operação de pré-processamento consistiu na conversão do formato original dos arquivos para o formato “txt”: o mesmo usado no “Bloco de notas” do *Windows*, por exemplo. Esse formato de arquivo permite manipular os textos de maneira fácil e eficiente em Python. De fato, trabalhar com arquivos em formato PDF requer bibliotecas cujo funcionamento é significativamente mais complexo.

Como quase nenhum dos 140 romances estava originalmente em formato “txt”, precisei usar o *Convertio*<sup>8</sup>, um *software* que modifica a formatação dos arquivos. A má qualidade de alguns PDFs dificultou ou impossibilitou, em alguns casos, a conversão em “txt”, de modo que não pude trabalhar com alguns romances. De toda a forma, todos os romances que constam nesta pesquisa estavam em boas condições de leitura e, repito, foram contabilizados nos resultados deste trabalho.

---

<sup>8</sup> O link deste site é: <https://convertio.co/pt/pdf-txt/>

A segunda operação de pré-processamento se restringiu à segunda parte da pesquisa: a análise de alguns temas importantes para o Brasil oitocentista. Com efeito, para que o código funcionasse eficientemente, precisei 1) converter todos os textos para letra minúscula, 2) retirar as pontuações, 3) eliminar palavras gramaticais (*stopwords*) e 4) separar o texto em palavras (*tokenizer*). Cada uma dessas operações tem uma função específica, conforme veremos a seguir.

A primeira operação é necessária na medida em que a linguagem Python diferencia letras maiúsculas de minúsculas. Como essa distinção não era relevante para a minha pesquisa, todos os romances foram convertidos para a letra minúscula a fim de uniformizá-los. Do contrário, ao pesquisar pela palavra “escravidão”, por exemplo, o código não mostraria a ocorrência da mesma palavra em letra maiúscula, ou seja: “Escravidão”. A segunda e a terceira operação aumentam a velocidade do processamento, uma vez que se elimina um aspecto que não é importante para a tarefa: a pontuação e as palavras gramaticais (preposições, artigos e pronomes, por exemplo). A quarta é não só a mais importante como também a mais complexa: o Python, por ser uma linguagem de programação, não entende a linguagem da mesma maneira que os humanos. Assim, para comparar uma lista de palavras e um texto, o Python necessita que o texto também tenha sido transformado em uma espécie de grande lista de palavras (*tokens*). É justamente por essa razão que, no exemplo do romance *O Ateneu*, o código continha uma linha “`texto_tokenized`”.

Para exemplificar como funciona o pré-processamento de textos, optei por trabalhar com o romance *Senhora*, de José de Alencar. Antes de mais nada, convém testar se o Python consegue encontrar e ler o arquivo do romance em formato “`txt`”. Abaixo, disponho somente os resultados obtidos, sem os códigos que os geraram, já que ainda não expliquei como se podem escrever códigos em Python. O que se vê abaixo é justamente o ambiente de execução dos códigos.



Figura 5 — Demonstração do código.

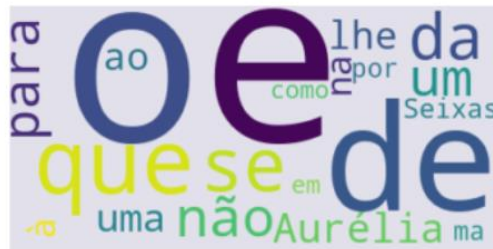
```
In [2]: print(text)
senhora
O Preço

I
Há anos raiou no céu fluminense uma nova estrela.
Desde o momento de sua ascensão ninguém lhe disputou o cetro; foi proclamada a rainha dos salões.
Tornou-se a deusa dos bailes; a musa dos poetas e o ídolo dos noivos em disponibilidade.
Era rica e formosa.
Duas opulências, que se realçam como a flor em vaso de alabastro; dois esplendores que se refletem, como o raio de sol no prisma do diamante.
Quem não se recorda da Aurélia Camargo, que atravessou o firmamento da Corte como brilhante meteoro, e apagou-se de repente no meio do deslumbramento que produzira o seu -fulgor?
Tinha ela dezoito anos quando apareceu a primeira vez na sociedade. Não a conheciam; e logo buscaram todos com avidéz informações acerca da grande novidade do dia.
Dizia-se muita coisa que não repetirei agora, pois a seu tempo saberemos a verdade, sem os comentários malévolos de que usam vesti-la os noveleiros.
```

Fonte: Elaborado pelo autor (2023)

De fato, o código consegue ler as primeiras linhas do romance em exame. Agora, suponhamos que eu não conheça o romance e que o meu objetivo seja ter uma ideia do tema principal dessa obra. Essa tarefa pode ser feita pela análise das palavras mais utilizadas, notadamente. Se não fizermos as 4 operações de pré-processamento, obteremos o seguinte mapa, que mostra as 20 palavras mais frequentes nesse romance:

Figura 6 — Demonstração do código.



Fonte: Elaborado pelo autor (2023)

Como se pode observar, as palavras que mais apareceram no romance foram preposições, artigos e conjunções, o que não nos diz nada sobre o tema principal de *Senhora*. É por essa razão que as *stopwords* devem ser excluídas. Na verdade, salvo exceções, todo texto em língua portuguesa será composto majoritariamente por esse tipo de palavra. Abaixo, também apresento alguns resultados do banco de dados de *stopwords* da língua portuguesa, para que se possa entender do que se trata na prática:



### 2.3 O funcionamento de códigos em linguagem Python

Em um capítulo anterior, tentei explicar as várias formas de se trabalhar com as Humanidades Digitais, com o objetivo de mostrar algumas das possibilidades desse método de pesquisa. Neste capítulo, o objetivo é explicitar as ferramentas utilizadas para realizar esta pesquisa em particular. Os dados obtidos neste trabalho foram coletados e processados por meio de códigos escritos em Python. Além disso, vale ressaltar que não trabalhei com inteligência artificial (*machine learning*) e que usei programas como *Excel* e *Qgis*. Tratou-se, assim, de uma pesquisa primária com o uso de bibliotecas já existentes, tais como *Regex*, *Pandas* e *Folium*.

Apesar de já ter explicado e exemplificado algumas de suas utilidades, ainda não tínhamos tratado do funcionamento das bibliotecas, que é relativamente simples. Lembro, no entanto, que o objetivo deste trabalho não é ensinar técnicas ou métodos avançados de programação. As breves explicações a seguir servem apenas para ajudar o programador inexperiente a iniciar sua pesquisa.

De início, para usar uma biblioteca, devemos primeiramente instalá-la. Há várias maneiras de se fazer isso: a mais fácil, no entanto, é digitar *pip install*, seguido pelo nome da biblioteca, como ilustra a figura abaixo:

Figura 9 — Demonstração do código.

```
In [1]: 1 pip install Regex
Requirement already satisfied: Regex in c:\users\tpant\anaconda3\lib\site-packages (2022.8.17)
Note: you may need to restart the kernel to use updated packages.
```

Fonte: Elaborado pelo autor (2023)

Como se pode ler no ambiente de execução do código (“*requirement already satisfied*”), a biblioteca *Regex* já tinha sido instalada no meu computador. Se eu ainda não a tivesse instalado, teria recebido a mensagem de que a biblioteca foi instalada com sucesso ou que algum erro tinha sido encontrado. De toda forma, a instalação de bibliotecas não costuma causar grandes dificuldades. Caso se encontre algum erro, no entanto, uma solução pode ser rapidamente encontrada na internet.

Após a instalação, a biblioteca precisa ser evocada (*import*, em inglês) pela primeira vez para poder ser utilizada. É deste jeito que se importa uma biblioteca: digita-se *import*, seguido do nome que os programadores da biblioteca lhe deram. Para o *Regex*, por exemplo, o

nome é *re*. Essa informação também pode ser encontrada na internet. No exemplo abaixo, a biblioteca *Regex* é importada com sucesso:

Figura 10 — Demonstração do código.

```
In [2]: 1 import re
```

Fonte: Elaborado pelo autor (2023)

Ao contrário do comando que instala a biblioteca, o comando *import* só é bem sucedido se não houver nenhuma mensagem de erro. No caso do exemplo, foi, portanto, bem sucedido. Depois da importação, é possível utilizar todas as funções da biblioteca.

Para explicar os códigos desenvolvidos para esta pesquisa, é importante notar que o capítulo dedicado ao desenvolvimento desta dissertação consistirá em duas seções, que se diferenciam tanto pela área temática quanto pelos códigos usados. A seção 1 será dedicada ao estudo dos municípios, ao passo que a seção 2, à quantificação de temas importantes do século XIX. No sub-capítulo precedente, tratei do pré-processamento dos dados, isto é: das etapas anteriores à análise dos textos. Nos parágrafos seguintes, abordarei justamente as diferenças entre os códigos da primeira e segunda seções. Em síntese, o objetivo é exemplificar as linhas de código desenvolvidas, na medida em que essa explicação seja importante para entender como cheguei aos resultados que serão apresentados oportunamente.

Inicialmente, posso afirmar que os códigos que executam as tarefas da primeira seção são consideravelmente menos eficientes que os da segunda. Com efeito, à medida que aprendia a programar, fui conseguindo aperfeiçoá-los. Assim, quando comecei a tratar dos assuntos da segunda seção, já tinha adquirido um pouco mais de experiência. Além disso, conforme já discutimos, praticamente qualquer tarefa que pode ser feita por um *software* também pode ser realizada manualmente. Da mesma forma, códigos menos eficientes também funcionam, ainda que o pesquisador precise eventualmente suplantar as ineficiências do código de forma manual.

Finalmente, para tratar dos dados da seção 1, desenvolvi o que chamei de Laboratório de Análise de Dados (LDA), uma vez que o código não era eficiente o bastante para executar todas as tarefas necessárias de maneira automática. Para trabalhar com a seção 2, por outro lado, consegui desenvolver linhas de códigos mais eficientes, de modo que não eu não precisei realizar muitas operações de maneira manual.

## 2.4 O código da Seção 1: o Laboratório de Análise de Dados (LAD) e a criação dos mapas

O Laboratório de Análise de Dados (LAD) representou a possibilidade de se trabalhar com as Humanidades Digitais desde o início da pesquisa. Do contrário, eu só teria introduzido a programação depois de algum tempo de experiência, o que poderia ter atrasado esta dissertação. A respeito de sua definição, um Laboratório de Análise de Dados consiste em códigos que, embora não executem automaticamente a tarefa desejada, facilitam enormemente a consecução dela.

Para quantificar as ocorrências de municípios, tarefa da seção 1, o código precisava, antes de mais nada, identificar os nomes dos municípios nos arquivos dos romances. Para realizar essa tarefa, foram utilizadas as bibliotecas *Regex*, *Pandas* e *CSV*. O problema, no entanto, é que nomes de municípios podem coincidir com nomes de pessoas, bairros, rios ou santos, por exemplo. Dessa forma, se tivesse escrito um código para realizar essa tarefa de forma imediata, teria obtido muitas ocorrências erradas. Foi nesse momento que desenvolvi o LAD, cuja finalidade era verificar, caso a caso, a utilidade das ocorrências para a pesquisa. Abaixo, apresento um *print* de alguns resultados obtidos, com o intuito de mostrar como o código funcionava:

Figura 11 — Demonstração do código.

```
<re.Match object; span=(51, 65), match='Rio de Janeiro'> tudo isto encheu o primeiro mês de minha estada no Rio de Janeiro. D
epois desse tributo pago à

<re.Match object; span=(26, 32), match='Glória'> Depois da festa da Glória tinha-a encontrado algumas vezes, mas sem l
he falar. Lembro-me
```

Fonte: Elaborado pelo autor (2023)

Ao examinar o romance *Lucíola*, obtive dados como os apresentados acima. Da maneira como foi escrito, o meu código dispôs a palavra “Glória” como um resultado válido, apesar de não se tratar de um município brasileiro no contexto em que ocorre. O que aconteceu é simples: o código, ao verificar o nome de todos os municípios brasileiros, relacionou a ocorrência “Glória”, no romance, com o município de Glória, na Bahia. De fato, o meu código não eliminava esses resultados incorretos de maneira automática. Por esse motivo, precisei verificar os resultados manualmente, o que foi feito 3 vezes, para minorar eventuais erros e omissões de minha parte. Precisamos ressaltar, no entanto, que trabalhar com essa quantidade de dados é sempre desafiador e que nenhum código é infalível, sobretudo quando o programador é inexperiente.

De maneira detalhada, o código comparava os textos dos romances com uma lista composta tanto pelos municípios brasileiros conforme dados de 2020 quanto pelos municípios constantes no Censo de 1872, o qual preserva as ortografias daquela época. Esse último aspecto do Censo é importante, na medida em que alguns romances examinados ainda conservavam a ortografia original da época. Ademais, a verificação manual dos resultados acabou permitindo que se pudessem contabilizar ocorrências como “São Domingos” ou “Santa Catharina”, os quais representam atualmente Niterói e Florianópolis, respectivamente. De fato, precisei pesquisar a evolução do nome de alguns municípios, para não deixar de quantificar nenhum dado. Além disso, ao trabalhar com o Laboratório de Análises de Dados, também consegui contabilizar romances em que nenhuma cidade aparecesse, mas cujas histórias se passavam inequivocamente em algum município específico. O caso do Rio de Janeiro foi o mais recorrente, como se pode ver com a análise do romance *Ressurreição*, de Machado de Assis:

Figura 12 — Demonstração do código.

```
<re.Match object; span=(10, 17), match='Ouvidor'> Na Rua do Ouvidor encontrou o Doutor Meneses, jovem advogado com quem entre
tinha

<re.Match object; span=(56, 67), match='Laranjeiras'> prestou atenção. as 3 horas separaram-se, Félix para as Laranjeiras, Me
neses para o

<re.Match object; span=(9, 20), match='Laranjeiras'> carta as Laranjeiras, justamente na ocasião em que Félix acabava de ler
outra carta de

<re.Match object; span=(38, 43), match='Viana'> Pareceu-lhe reconhecer Livia, irmã de Viana. Com as faces avermelhadas e o se
io
```

Fonte: Elaborado pelo autor (2023)

Apesar de não mencionar o nome da cidade textualmente, Machado de Assis refere-se claramente ao Rio de Janeiro quando fala em "Laranjeiras" e “Ouvidor”. Além disso, a palavra “Viana” - a última ocorrência na imagem -, é apenas um nome próprio, não se tratando de um município no contexto em que ocorre, portanto.

Depois de ter adquirido mais experiência em programação, posso afirmar que há formas mais eficientes e automáticas de se gerar os mesmos resultados obtidos por meio de um LAD. Também é verdade, no entanto, que se trataria de um código muito mais complexo, uma vez que existem várias variáveis que o código teria que levar em conta para dispor os resultados corretamente. Dessa forma, creio que o benefício de se ter trabalhado com um LAD superou as desvantagens, sobretudo em se tratando de um programador inexperiente, vale ressaltar.

Após realizar a primeira tarefa da seção 1 - a verificação dos municípios que tinham sido citados nos romances - era preciso dispor esses dados em um mapa do Brasil, de modo a espacializar as informações. Na verdade, sem os mapas, essa quantidade de dados não passaria de um grande aglomerado de nomes de cidades e de romances. Era muito importante, portanto, visualizar os resultados de maneira a ressaltar algum aspecto ainda desconhecido. A biblioteca utilizada para realizar essa tarefa foi o *Folium*. De fácil utilização, essa biblioteca permite que se criem mapas a partir de informações contidas em tabelas. A tabela abaixo apresenta todos os dados que obtive antes de criar um mapa:

Figura 13 — Demonstração do código.

nome_do_romance	autor	cidade_natal	ano_de_publicação	ocorrências	ocorrências_match	NOME_MUNICIPIO	LONGITUDE	LATITUDE
0	Ubirajara	José de Alencar	Fortaleza	1874	[aratuba]	aratuba	aratuba	-39.047555 -4.416931
1	Iracema	José de Alencar	Fortaleza	1865	[rio de janeiro, jaguaribe, são paulo, ipu, ca...]	camocim	camocim	-40.847528 -2.900967
2	Iracema	José de Alencar	Fortaleza	1865	[rio de janeiro, jaguaribe, são paulo, ipu, ca...]	rio de janeiro	rio de janeiro	-43.227875 -22.876652
3	Iracema	José de Alencar	Fortaleza	1865	[rio de janeiro, jaguaribe, são paulo, ipu, ca...]	ipu	ipu	-40.718726 -4.313432
4	Iracema	José de Alencar	Fortaleza	1865	[rio de janeiro, jaguaribe, são paulo, ipu, ca...]	uruburetama	uruburetama	-39.508178 -3.623017
...	...	...	...	...	...	...	...	...
673	O martírio do Tiradentes	Joaquim Norberto	Rio de Janeiro	1882	[rio de janeiro, ouro preto, varginha]	rio de janeiro	rio de janeiro	-43.227875 -22.876652
674	O martírio do Tiradentes	Joaquim Norberto	Rio de Janeiro	1882	[rio de janeiro, ouro preto, varginha]	ouro preto	ouro preto	-43.506448 -20.386776
675	Rosa	Joaquim Manuel de Macedo	Itaboraí	1849	[rio de janeiro, itaboraí]	rio de janeiro	rio de janeiro	-43.227875 -22.876652
676	Rosa	Joaquim Manuel de Macedo	Itaboraí	1849	[rio de janeiro, itaboraí]	itaboraí	itaboraí	-42.860343 -22.745863
677	Uma paixão romântica	Joaquim Manuel de Macedo	Itaboraí	1860	[rio de janeiro]	rio de janeiro	rio de janeiro	-43.227875 -22.876652

Fonte: Elaborado pelo autor (2023)

Como se pode intuir, o próximo código tinha que ler as informações acima a fim de dispô-las em um mapa. Além de mostrar as cidades mencionadas, o código também precisava apresentar o nome do romance, o ano de publicação da obra e o nome do autor no *display* do mapa, conforme ilustra a figura abaixo:

Figura 14 — Demonstração do código.

```

map_tempo = folium.Map(location=[-14.235004, -51.92528],
                        zoom_start=4)

feature_ea = folium.FeatureGroup(name='Romances (1840 - 1849)')
feature_pr = folium.FeatureGroup(name='Romances (1850 - 1859)')
feature_sr = folium.FeatureGroup(name='Romances (1860 - 1869)')
feature_ea1 = folium.FeatureGroup(name='Romances (1870 - 1879)')
feature_pr2 = folium.FeatureGroup(name='Romances (1880 - 1889)')
feature_sr3 = folium.FeatureGroup(name='Romances (1890 - 1900)')

for i, v in res.iterrows():
    popup = """
    Cidade : <b>%s</b><br>
    Autor : <b>%s</b><br>
    Nome do romance : <b>%s</b><br>
    Ano de publicação da obra : <b>%s</b><br>
    """ % (v['ocorrências_match'], v['autor'], v['nome_do_romance'], v['ano_de_publicação'])

    if v['ano_de_publicação'] in range(1840,1850):
        folium.CircleMarker(location=[v['LATITUDE'], v['LONGITUDE']],
                            radius=4,
                            tooltip=popup,
                            color='#FFBA00',
                            fill_color='#FFBA00',
                            fill=True).add_to(feature_ea)

```

Fonte: Elaborado pelo autor (2023)

As primeiras linhas criaram legendas em função do tempo. As restantes instruíram o computador a dispor as cidades mencionadas (“ocorrências\_match”), o autor, o nome do romance e o ano de publicação no mapa do Brasil (coordenadas geográficas -14.235004 e -51.92528). Em seguida, o código deve ser entendido desta forma: se (*if*) o ano de publicação estiver entre os anos de 1840 e 1850 - informação constante na coluna “ano\_de\_publicação” - devem-se mostrar as informações sobre a latitude e a longitude constantes na tabela, como mostra o resultado abaixo.

Figura 15 — Demonstração do código.



Fonte: Elaborado pelo autor (2023)



Embora seja possível observar apenas o dado referente à cidade de Triunfo, no Rio Grande do Sul, todas as ocorrências podem ser consultadas da forma como se mostra no mapa acima. Com efeito, basta que o usuário escolha quais décadas quer observar e passe o cursor no ponto indicado no mapa.

## 2.5 O código da seção 2: a experiência mais bem sucedida

Como já afirmei, os códigos da seção 2 são bem mais eficientes dos que os que mostrei no subcapítulo imediatamente anterior. Em se tratando de programação, eficiência muitas vezes equivale à automação. Para obter os dados da seção 1, era necessário aplicar o código a cada romance individualmente, ao passo que, para recolher as informações da seção 2, bastava apenas um comando para que o código realizasse a operação em todos os romances. Essa maior automação não significa, no entanto, que os códigos não apresentem limitações ou problemas, conforme veremos no final deste subcapítulo.

Em síntese, desenvolvi um código que criava um tabela com todos os romances (em forma de arquivo), seguidos dos textos pré-processados, isto é: letra minúscula, sem pontuações, sem palavras gramaticais (*stopwords*) e em *tokens*. Para realizar essa tarefa, as bibliotecas *Pandas*, *Nltk*, *String* e *Spacy* foram utilizadas. A figura abaixo é o *print* dos cinco primeiros resultados obtidos. O *display* que se observa foi gerado pelo *Pandas*, biblioteca muito eficiente para se trabalhar com banco de dados obtidos a partir de arquivos CSV. A figura abaixo mostra os resultados.

Figura 16 — Demonstração do código.

	file_name	text
0	acaiaca.txt	[acaiaca, morro, santo, antônio, cuja, encosta...
1	afilhada.txt	[afilhada, i, desembargador, osório, pereira, ...
2	alfarrabios.txt	[alfarrábios, garatuja, primeira, serie, chron...
3	alma_lazaro.txt	[alma, lázaro, alma, penada, triste, irrisão, ...
4	amor_esposo.txt	[amor, esposo, declive, austral, serra, carioc...
5	amor_que_mata.txt	[amor, mata, fresca, tarde, agosto, carro, des...

Fonte: Elaborado pelo autor (2023)

A coluna “file\_name” dispõe os nomes dos arquivos em formato “txt”, ao passo que a coluna “text” mostra as primeiras palavras dos romances. Em regra, o *Jupyter Notebook* não apresenta todos os resultados no ambiente do código, uma vez que essa ação consumiria muita

memória do computador. As reticências indicam que há mais dados naquela linha. Nesse caso, trata-se do texto inteiro de cada romance.

Depois de organizar os arquivos e pré-processar os textos, a próxima tarefa era aplicar o código principal nos textos. Em outras palavras: efetivamente examinar os romances. O código funcionava da seguinte forma: buscava em cada um dos arquivos as palavras constantes em uma lista. Para o exemplo abaixo, utilizou-se a lista com as palavras 'república', 'repúblicas', 'federação', 'federações', 'federalismo', 'federalismos', 'presidencialismo', 'presidencialismos'. O resultado foi o seguinte:

Figura 17 — Demonstração do código.

file_name	text	text_string	menções_política	quantidade_menções_política	nome_do_romance	autor	ano_de_publicação
acaiaca.txt	[acaiaca, morro, santo, antônio, cuja, encosta...	acaiaca morro santo antônio cuja encosta orien...	[]	0	Acaiaca	Joaquim Felício dos Santos	1866
afilhada.txt	[afilhada, i, desembargador, osório, pereira, ...	afilhada i desembargador osório pereira goís d...	[república, república, república]	3	A afilhada	Manuel de Oliveira Paiva	1889
alfarrábios.txt	[alfarrábios, garatuja, primeira, serie, chron...	alfarrábios garatuja primeira serie chronicas ...	[]	0	Alfarrábios	José de Alencar	1873
alma_lazaro.txt	[alma, lázaro, alma, penada, triste, irrisão, ...	alma lázaro alma penada triste irrisão glória ...	[república]	1	A alma do Lázaro	José de Alencar	1872
amor_esposo.txt	[amor, esposo, declive, austral, serra, carioc...	amor esposo declive austral serra carioca n lo...	[]	0	Amor de esposo	Pedro Américo	1886

Fonte: Elaborado pelo autor (2023)

É isto que se pode fazer com o recurso à programação: em aproximadamente 2 minutos (o tempo de processamento da máquina), o código examinou 140 romances, mostrou quais palavras da lista constavam nos romances (coluna “menções\_política”) e contabilizou essas menções (coluna “quantidade\_menções\_política”). Além disso, o código adicionava 3 colunas suplementares à tabela: o nome do romance, que não se confunde com o nome do arquivo do romance, o nome do autor e o ano de publicação da obra.

Ainda que esse código tenha sido eficiente, duas operações ainda foram realizadas à mão. A primeira diz respeito à alteração das palavras da lista, que servia como parâmetro para a busca. As listas foram formuladas manualmente em função dos vários testes que fiz até encontrar as palavras mais convenientes para a confecção dos temas que foram quantificados. No exemplo acima, mostrei os dados obtidos a partir de uma lista com menções à política e a sistemas de governo, mas também examinei outros conjuntos de palavras, como se verá nos resultados do trabalho. A segunda operação, por sua vez, diz respeito à criação dos gráficos: embora tenha esboçado modelos em Python, preferi utilizar o *Excel* para criar os gráficos da

Seção 2. A razão é simples: os gráficos do *Excel* são mais organizados e têm melhor resolução.

Apesar de eficientes, os códigos apresentam ao menos duas limitações importantes: 1) só permitem a busca de sintagmas compostos por um só vocábulo e 2) quantificam, erroneamente, ocorrências compostas por um mesmo radical. O primeiro problema não gerou muitas perdas para a pesquisa, presumivelmente: o código conseguia encontrar as ocorrências de “escravo”, mas não as de “escravo de ganho” no romance, por exemplo. Assim, resolvi o problema ao não usar sintagmas compostos por mais de uma palavra. O segundo problema, no entanto, me impediu de trabalhar com algumas palavras. À guisa de exemplo, não pude quantificar a palavra "servo", pois o código também mostrava a ocorrência "conservador", assim como a busca pela palavra "serva" gerava a ocorrência “conservar”, o que não corresponde a um resultado esperado.

### 3 RESULTADOS DA PESQUISA

#### 3.1 Alguns apontamentos preliminares sobre as seções 1 e 2

Antes de discutir os resultados da pesquisa, queremos ressaltar alguns aspectos em comum relativamente às duas próximas seções. Em síntese, o objetivo é entender os pressupostos teóricos que embasaram esta dissertação. Na verdade, vários desses pontos serão aprofundados na medida em que forem mencionados durante as análises. Inicialmente, lembramos que o conceito de romance usado é muito simples: trata-se da classificação do já citado site da Universidade Federal de Santa Catarina (FLUC). Similarmente, como a ferramenta de busca da FLUC não diferencia os gêneros romance e novela (o filtro é “romance/novela”), os dois termos também são intercambiáveis para os propósitos da pesquisa.

Também ressaltamos que a pesquisa é baseada na ideia de “engenharia reversa”, ou seja: no entendimento de que é possível reconstruir o passado por meio da análise das formas literárias. Criado por Franco Moretti, esse conceito permite entender a literatura como parte essencial da história, uma vez que as formas literárias não se desenvolvem em qualquer lugar ou a qualquer tempo. Com efeito, assim como documentos oficiais, jornais e relatos pessoais, as formas literárias engendram processos históricos que podem ser recuperados mediante a análise formal das obras. Ao falar sobre os tipos de evidências apresentadas por obras literárias, Moretti assim exemplifica:

Eis a possível contribuição dela para o conhecimento histórico: ao compreendermos a opacidade das alusões de Ibsen ao passado, ou a semântica oblíqua dos adjetivos vitorianos, ou até (à primeira vista uma tarefa nada agradável) a função do gerúndio em *Robinson Crusóé*, ingressamos em um reino de sombras onde o passado recobra sua voz e ainda fala conosco. (MORETTI, 2013, p. 23).

Dessa forma, defendemos que a análise da geografia do romance e o exame da frequência de temas importantes para a época podem ajudar a reconstruir alguns dos processos históricos que marcaram o Brasil do século XIX. Em outras palavras, significa dizer que a menção a uma cidade ou a um tema não só não é aleatória, como também pode ser entendida como a manifestação de alguma dinâmica mais profunda em curso na sociedade. Em síntese, trata-se de entender o romance como uma mediação possível entre a realidade e a literatura, quer formal, quer materialmente.

Igualmente, é importante explicitar as razões pelas quais optamos por trabalhar especificamente com a forma literária romance. A razão é relativamente simples: por ser um gênero literário inovador e relativamente popular, o romance é uma das formas literárias essenciais para se entender o século XIX. De fato, o romance permitiu que histórias de pessoas, dramas e vidas comuns fossem representados na literatura de maneira séria<sup>10</sup>. Segundo o historiador da literatura Otto Maria Carpeaux,

“Depois, [romance] será o espelho do nosso mundo, dos nossos países, das nossas cidades e ruas, das nossas casas, dos dramas que se passam em nossos apartamentos e quartos.” (CARPEAUX, 1956, p. 1717)

O romance é normalmente considerado como a forma literária que mais contribuiu para a construção da sociedade moderna dos séculos XVIII e XIX. Como argumenta a professora Catherine Gallagher, as bases dessa sociedade dependiam fortemente da crença em algumas instituições cuja existência e concretude não eram claras para a população em geral. Segundo a já citada professora, a ficção contribuiu para que se consolidassem algumas práticas sociais importantes durante aqueles séculos:

The same suspension of literal truth claims helped even common people to accept paper money: too wise to believe that the treasury held enough specie to cover all of their paper at once, they instead understood that the credit they advanced collectively obviated the need to hoard precious metals privately. So the government, too, relied on the imaginative sophistication of its people and financed a vast military and imperial enterprise by selling national debt bonds. (MORETTI, 1999, p. 3)

De fato, a ficção permitiu não só que o grande público aceitasse o uso da moeda fiduciária (*paper money*), como também que os governos financiassem suas dívidas nacionais mediante a compra e venda de títulos públicos. Conforme pretendemos argumentar depois, a construção imaginária do território brasileiro também dependeu da literatura, notadamente o romance do século XIX.

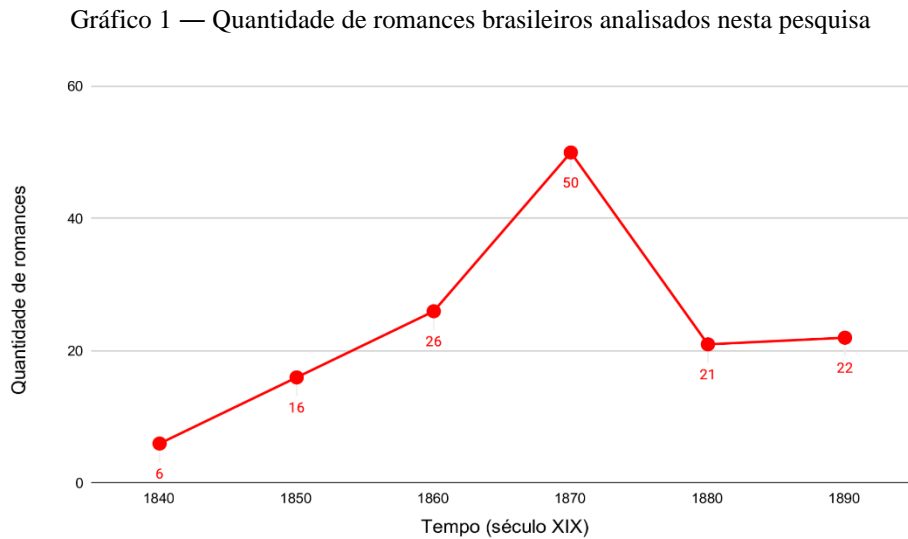
Gostaríamos ainda de chamar a atenção do leitor para um aspecto metodológico importante: em se tratando de estudos literários, examinar 150 romances significa trabalhar com *big data*. É bem verdade que, quanto maior o volume de dados disponíveis, mais acuradas costumam ser as interpretações, uma vez que se distingue melhor a tendência do acaso. No entanto, convém observar que não conseguimos examinar todos os romances do século XIX. Assim, todas as análises desta pesquisa devem ser consideradas de maneira

---

<sup>10</sup> “Sério” no sentido de não rebaixado. Em outras palavras, a vida das pessoas comuns também passou a ser importante para a produção literária.

preliminar, sobretudo em função da novidade do método utilizado para a obtenção dos dados. Na verdade, embora tenhamos tentado minimizar ao máximo eventuais omissões e erros, é preciso admitir que alguma citação (para a seção 1) ou palavra (para a seção 2) pode ter escapado à quantificação e leitura dos códigos.<sup>11</sup>

Por fim, ainda queremos mostrar que, como a quantidade de romances aos quais tivemos acesso varia significativamente ao longo dos anos, devemos ter cuidado ao comparar as décadas. Todos os romances examinados foram publicados entre 1843 e 1900, embora o objetivo fosse estudar todo o século XIX. A razão para essa discrepância é meramente factual, uma vez que o romance só se desenvolve a partir da segunda metade do século XIX no Brasil, como veremos a seguir. O gráfico abaixo apresenta a relação entre a quantidade de romances e o tempo.



No eixo vertical: quantidade de romances publicados; no eixo horizontal: tempo medido em décadas. Fontes: dados da pesquisa

A diferença entre as décadas é tal que apenas 6 romances foram publicados em 1840, ao passo que 22 obras foram publicadas em 1890. Além disso, já podemos afirmar que é a partir de 1850 que o romance brasileiro se desenvolve efetivamente. Em síntese, de 1850 a 1870, o crescimento no número de romances é constante; de 1870 a 1880, há uma diminuição, que é seguida por um valor constante na década de 1890. Também podemos asseverar que a década de 1870 é fundamental para se entender o romance oitocentista brasileiro.

<sup>11</sup> Vale sempre lembrar que eu não sou um programador experiente ainda, apesar de ter ganhado muita experiência ao longo da pesquisa.

### 3.2 Seção 1: território, municípios e mapas.

#### 3.2.1 Série histórica completa (1840 - 1899)

Embora todos os mapas da seção 1 estejam disponíveis online, decidimos mostrar os resultados mais interessantes no corpo do trabalho, a fim de facilitar a leitura desta dissertação. Começamos do mapa mais para o menos abrangente. O primeiro mapa apresenta a série histórica completa de todas as cidades citadas nos enredos dos romances. Não se trata da cidade na qual o autor nasceu ou na qual os livros foram publicados, mas de toda e qualquer cidade mencionada nos romances analisados. Além disso, ressaltamos que cada cor no mapa corresponde a uma década, recurso que pode ser melhor aproveitado no material disponibilizado online.

Mapa 1 — Série histórica completa (1840 - 1900)<sup>12</sup>



Mapa da série completa. Legenda: romances 1840-1849 (amarelo); romances 1850-1859 (azul); romances 1860-1869 (vermelho); romances 1870-1879 (verde); romances 1880-1889 (preto); romances 1890-1900 (roxo).

Fontes: dados da pesquisa

O mapa 1 é interessante na medida em que mostra as cidades mais relevantes para a construção do imaginário criativo dos autores ao longo do século XIX. Essa primeira análise diz respeito à importância de certas cidades para os romancistas e as suas obras. Com efeito, a

<sup>12</sup> Não consegui encontrar um mapa interativo do Brasil com as dimensões territoriais do País ao longo do século XIX. Assim, como usei dados do século XIX e o formato atual do Brasil, algumas áreas do mapa não deveriam ser apresentadas como brasileiras, como o Acre, por exemplo.

menção a uma cidade significa que o autor não só projeta sua imaginação até aquela localidade, como também a transforma em lugar de narração. Trata-se, nesse sentido, da formação de seu imaginário criativo. Para um país enorme como o Brasil, a mera citação já é notável, uma vez que o autor tinha conhecimento da existência da cidade, mais ou menos aprofundadamente.

Uma consequência importante dessa primeira análise é que o imaginário criativo de um autor pode ser mais ou menos abrangente, a depender dos municípios mencionados em sua obra. A fim de exemplificar essa afirmação, comparamos a obra de José de Alencar com a de Machado de Assis. Esse paralelo é interessante, na medida em que ambos os autores têm obras de tamanho similar. Vejamos a comparação:

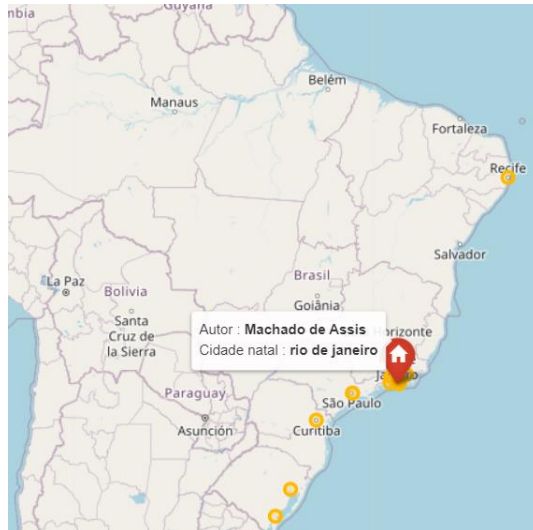
Mapa 2 — Romances de José de Alencar



Em azul, as cidades citadas nos romances.  
Fontes: dados da pesquisa



Mapa 3 — Romances de Machado de Assis



Em amarelo, as cidades citadas nos romances.

Fontes: dados da pesquisa

Conforme se pode notar, o imaginário criativo de José de Alencar é mais abrangente que o de Machado de Assis. De fato, Alencar situou seus romances no Rio Grande do Sul, no interior de São Paulo, no Centro-Oeste e em algumas partes do Nordeste brasileiro também. Embora tivesse conhecimento da existência de outras cidades brasileiras, Machado de Assis acabou por circunscrever o enredo de seus romances em algumas cidades litorâneas apenas. Nesse momento, reiteramos que o leitor pode acessar os mapas interativos, escolhendo e comparando os autores da maneira que melhor lhe convier.

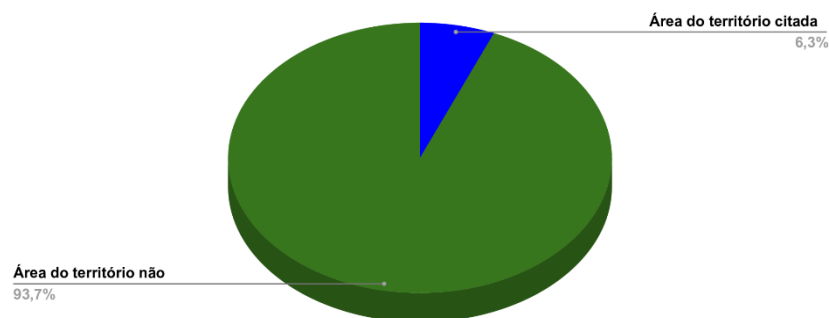
O mapa 1 ainda permite afirmar que a construção simbólica do território brasileiro também dependeu do romance. Trata-se de uma representação simbólica na medida em que, no imaginário dos leitores, o território brasileiro compreendia mais ou menos aquelas distâncias e limites ao final do século XIX. Compreender o que significava fazer parte de um país do tamanho do Brasil naquela época requeria muita imaginação e elaboração por parte da população. Ora, as grandes distâncias e a lentidão dos meios de comunicação tornavam difícil a movimentação pelo território e, conseqüentemente, o conhecimento da geografia do País. Dessa forma, assim como a ficção foi importante para que as pessoas aceitassem o uso da moeda fiduciária (conforme argumenta Catherine Gallagher), o romance foi essencial para a construção imaginária do território nacional.

Como a construção simbólica de um território também decorre de outros símbolos nacionais (bandeira e hino, por exemplo), quero chamar de “ficcionalização do território” o processo mediante o qual o romance possibilitou a incorporação do território brasileiro ao

imaginário da população. Em outras palavras, território ficcionalizado é o território que serviu como lugar de narração no romance. Com efeito, à medida que os leitores liam as obras, algumas cidades se tornaram mais familiares e apareciam mais frequentemente que outras. Como se viu anteriormente, o público leitor de Machado de Assis não estava habituado a imaginar um Brasil que não estivesse circunscrito ao litoral, por exemplo. Nesse sentido, o mapa da série completa também permite acompanhar a trajetória espacial do romance brasileiro no século XIX.

Sem a devida atenção, no entanto, o mapa 1 permitiria assumir que praticamente todos os municípios brasileiros foram mencionados no final do século XIX. Na verdade, menos de 10% do território do Brasil apareceu no romance daquela época, como se pode observar nas estimativas apresentadas no gráfico abaixo. Assim, podemos afirmar que, embora tenha começado no século XIX, o processo de ficcionalização do território brasileiro deve ter continuado nos séculos seguintes, notadamente a partir de 1930.

Gráfico 2 — Porcentagem da área do território brasileiro citada nos romances analisados



Fontes: dados da pesquisa

O gráfico 2<sup>13</sup> indica uma estimativa da porcentagem de território brasileiro que efetivamente aparece nos romances daquele século. Embora essa análise considere apenas os municípios citados, é muito improvável que esse número seja maior que 10%. Dessa forma, argumentamos que os leitores brasileiros estavam habituados a ler romances cujos enredos se passavam em não mais que 10% do território nacional durante o século XIX. Ademais, seria interessante coletar os resultados para os romances do século XX a fim de entender se essa

<sup>13</sup> Esse gráfico é muito aproximativo. É somente para que se tenha a dimensão desse aspecto, pois combinei um estimativa oficial do tamanho do território realizada em 1889 com a área atual da soma de todos os municípios, respectivamente: 8.337.218.000 km<sup>2</sup> e 560.429.360. De toda a forma, provavelmente esse número não passa de 10%, ou seja: apenas 1 décimo do território brasileiro foi citado.

porcentagem aumenta com o passar do tempo. Preliminarmente, parece razoável pensar que a resposta é positiva.

A análise conjunta do mapa 1 e do gráfico 2 ainda corrobora a ideia de que, até a década de 1930, o território brasileiro era composto por fundos territoriais, isto é: áreas que ainda não haviam sido efetivamente ocupadas pela população brasileira. Na verdade, trata-se de um outro termo para sertão. Segundo Adma Hamam de Figueiredo,

(...) até o início do Século XX, em termos geográficos e políticos, o Brasil não constituía, plenamente, um território, sendo mais bem caracterizado como um conjunto disperso de regiões fragilmente articuladas em meio a vastos fundos territoriais, genericamente associados aos sertões ou às extensas áreas de fronteira, que permaneciam, em grande parte, desconhecidos e precariamente revelados pela cartografia, em meados da década de 1930. (FIGUEIREDO, 1999, p. 9)

Essa dinâmica é particularmente importante para países territorialmente extensos, como Brasil, Austrália e EUA. Para o Brasil, especificamente, significa dizer que nem todo o território nacional estava sob o controle do Estado. O mais interessante, no entanto, é que, se compararmos o mapa 1 com o mapa seguinte - que apresenta a densidade populacional do território brasileiro em 1872 - percebemos que os fundos territoriais também aparecem no mapa dos romances. Vale notar que o conceito de densidade diz respeito à quantidade de habitantes por quilômetro quadrado. Vejamos essa comparação:

Mapa 4 — Série histórica completa (1840 - 1900)



Fontes: dados da pesquisa

Mapa 5 — Densidade populacional (1872)<sup>14</sup>

Legenda: número de habitantes por km<sup>2</sup>.  
Fontes: IBGE

De fato, praticamente as mesmas áreas que não são citadas nos romances são caracterizadas por fraca densidade populacional. Em resumo, trata-se dos Estados do Tocantins, do Amazonas, bem como da parte norte do Estado do Mato Grosso e do interior dos Estados do Pará, de Santa Catarina e do Paraná. Dessa forma, podemos argumentar que a trajetória espacial do romance coincidiu, de maneira geral, com a efetiva ocupação territorial do Brasil. Ademais, também se pode notar que os municípios litorâneos são mais mencionados que os interioranos, uma vez que são não só as áreas mais conhecidas pelo público leitor, como também aquelas com mais força política e demográfica naquela época, tais como Rio de Janeiro e Salvador.

O mapa da série histórica (mapa 1) também permite observar semelhanças entre a geografia do romance e a economia brasileira do século XIX. De início, precisamos notar que a circulação de pessoas de uma localidade a outra também depende das dinâmicas econômicas que possibilitam essas relações. Conforme sintetiza Marcus Antônio Croce, a economia brasileira já contava com um mercado interno razoavelmente relevante em meados do século XIX<sup>15</sup>:

“A economia brasileira no século XIX foi marcada pelo fato de o país ser um centro periférico agroexportador. Porém, essa economia voltada para o mercado externo permitiu

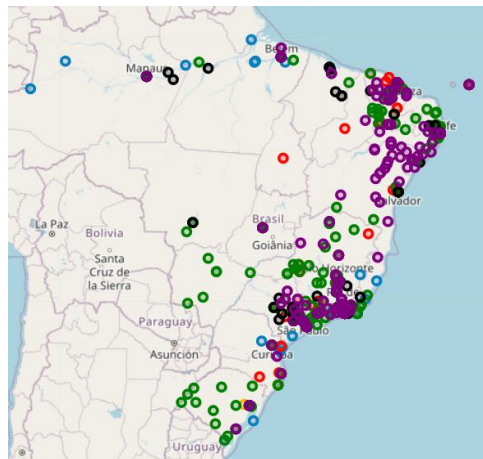
<sup>14</sup> Organizado pelo IBGE, esse mapa indica a densidade populacional no Brasil em 1872, isto é: “Habitantes por unidade de superfície, expressa pela expressão hab/km<sup>2</sup>. A densidade demográfica é uma medida da distribuição espacial da população e permite o estudo da concentração ou dispersão dessa população no espaço geográfico considerado.” Essa informação pode ser acessada pelo seguinte link: (<https://seriesestatisticas.ibge.gov.br/series.aspx?no=10&op=0&vcodigo=POP117&t=densidade-demografica>)

<sup>15</sup> Para mais detalhes sobre a economia brasileira no século XIX, ver *História da Riqueza no Brasil* (2017), de Jorge Caleira.

que fossem estabelecidas no Brasil condições para que se fortalecesse o seu mercado interno.” (A Economia do Brasil no século XIX, página 2)

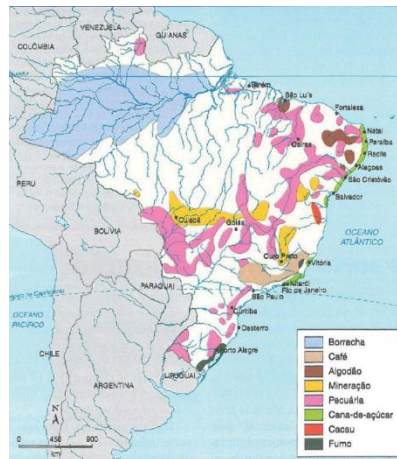
Para a nossa análise, é mais interessante observarmos essas dinâmicas de maneira espacializada, no entanto. Assim, propomos a comparação entre o já citado mapa 1 e o mapa abaixo (mapa 5), que apresenta as principais atividades econômicas no Brasil no século XIX.

Mapa 6 — Série histórica completa (1840 - 1900)



Fontes: dados da pesquisa

Mapa 7 — Economia brasileira (século XIX)



Fonte: (COSTA e MELLO, 2008, p. 411)

De fato, existe uma forte semelhança entre as cidades que foram citadas nos romances e as áreas mais importantes para a economia brasileira no século XIX. Em outras palavras, as manchas em ambos os mapas coincidem, de maneira geral. Em síntese, os romances parecem refratar o que acontecia na economia nessa época, ou seja: espaços geográficos economicamente relevantes recebem uma maior atenção por parte dos romancistas. Com base

nessas análises, ainda podemos constatar que, quanto mais importante for uma cidade, maiores são as chances de que ela seja mencionada nos romances.

Além do mercado interno, também o mercado externo foi muito importante para o País durante o século XIX, sobretudo em função do comércio internacional. Em síntese, as divisas geradas via balança comercial superavitária permitiam o pagamento dos encargos financeiros do Brasil a seus credores internacionais. A partir da década de 1840, o produto mais exportado pelo Brasil passou a ser o café, cuja exportação dependia do escoamento da produção via estrada de ferro, internamente, e portos, externamente. Assim, podemos considerar que as estradas de ferro também representam, no Brasil do século XIX, um indício de crescimento econômico de atividades ligadas ao mercado externo. Propomos mais uma comparação: o já citado mapa 1 com o mapa seguinte, o qual apresenta todas as ferrovias construídas até 1898.

Mapa 8 — Série histórica completa (1840 - 1900)



Fontes: dados da pesquisa

Mapa 9 — As ferrovias construídas até 1898



Fontes: acervo pessoal de Marcelo Werther

Inicialmente, podemos notar a concentração de estradas de ferro e, conseqüentemente, a maior pujança econômica das regiões Sul e Sudeste relativamente às demais. Embora o geógrafo Milton Santos tenha criado o conceito de região concentrada para tratar do Brasil do século XX, o processo de acumulação econômica daquelas regiões começou ainda no século XIX. Ainda segundo o geógrafo, em razão da desigualdade do desenvolvimento econômico das diversas regiões do País, o processo de modernização que marcou o Brasil do século XIX e XX se concentrou nas regiões Sul e Sudeste, que se industrializaram e se urbanizaram mais rapidamente. Assim, embora os mapas 1 e 6 não coincidam perfeitamente, podemos reafirmar que as áreas economicamente mais importantes para o Brasil do século XIX também contam com forte representatividade na geografia do romance brasileiro. O Rio Grande do Sul, por exemplo, aparece de forma parecida nos dois mapas, assim como os Estados de São Paulo e Rio de Janeiro, notadamente.

Outra informação que o mapa 1 não mostra muito satisfatoriamente é a frequência com a qual os municípios foram citados ao longo daquela época. De fato, não se consegue ver quantas vezes uma cidade foi mencionada, uma vez que as várias cores que marcam a legenda se sobrepõem no mapa 1. Dessa forma, para ressaltar a informação sobre a frequência, apresentamos o mapa seguinte, o qual dispõe os mesmos valores, porém de maneira agregada.

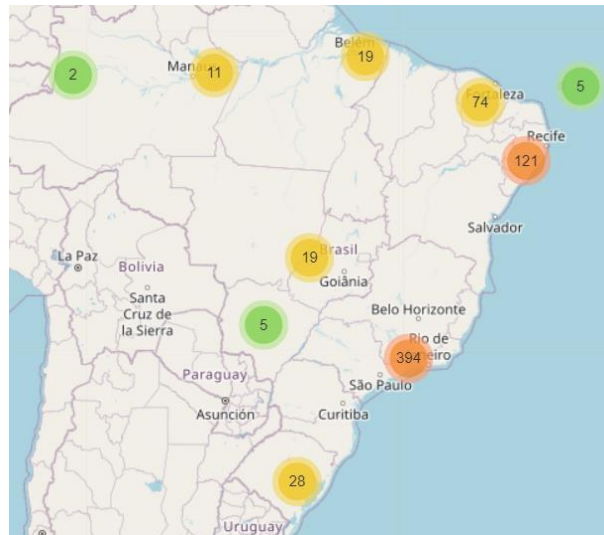
Trata-se de um mapa que agrupa as ocorrências de cidades próximas em um grande grupo. Esses agrupamentos apenas aproximadamente podem ser interpretados segundo a divisão das regiões brasileiras (Sul, Sudeste, Centro-Oeste, por exemplo), uma vez que a divisão verificada no mapa é organizada segundo critérios de proximidade determinados pelo própria biblioteca *Folium*. Mais uma vez, convidamos o leitor a acessar os dados em seu



computador de maneira autônoma. Ao clicar nesses aglomerados, o leitor acessa todas as citações de maneira detalhada e interativa, conforme mostraremos a seguir.

De toda a forma, o mapa seguinte permite visualizar melhor a representatividade de cada região, ainda que de forma aproximada. De fato, se mediante o mapa 1 podemos ressaltar as distâncias percorridas, o mapa a seguir permite entender a força imaginária ou o nível de polarização de cada cidade.

Mapa 10 — Série histórica por grupos (1840-1900)



Legenda: grupos com até 10 ocorrências (verde); grupos com até 100 ocorrências (amarelo); grupos com mais de 100 ocorrências (laranja). Fontes: dados da pesquisa

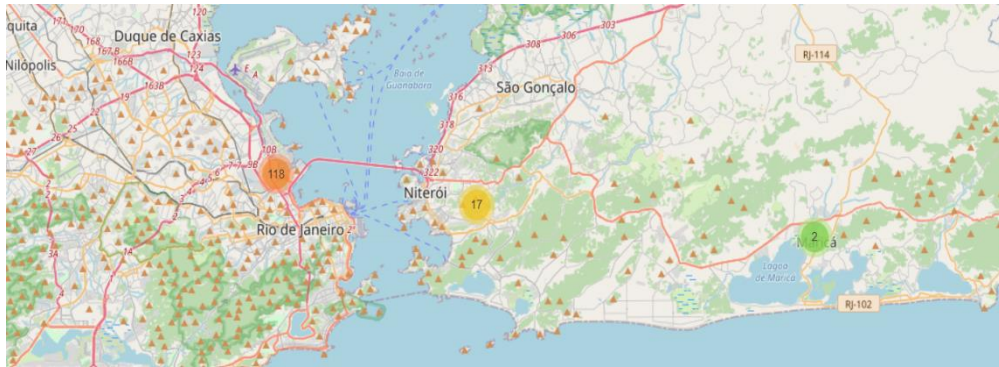
O mapa 7 indica que a quantidade de menções às regiões Sul, Sudeste e Nordeste é superior à das outras regiões. Esse mapa também permite entender que, apesar de aparecerem na série histórica, cidades como Manaus, Santarém (PA) e Macapá têm representatividade muito menor que outras, uma vez que foram citadas poucas vezes ao longo do século. Além disso, a área que corresponde, aproximadamente, ao Sudeste representa, individualmente, 394 das 678 menções, o que mais uma vez denota não só a concentração econômica, como também a polarização dessa região para o imaginário do País.

À guisa de exemplo, abaixo apresentamos duas imagens que dispõem dos dados sobre a cidade do Rio de Janeiro. A primeira imagem mostra o total de citações à cidade; a segunda individualiza os dados de cada citação contida nos grupos (autor, nome do romance e ano de publicação da obra). Como se poderá ver abaixo, a cidade do Rio de Janeiro foi citada 118 vezes, informação verificável pelo agrupamento de cor laranja da primeira imagem. Ao clicar nesse grupo de citações, as informações contidas na segunda imagem aparecem. Para



exemplificar, escolhemos a ocorrência do Rio de Janeiro contida no romance *História de um voluntário da Pátria*, publicado em 1869 por Vicente Felix de Castro.

Mapa 11 — Citações da cidade do Rio de Janeiro (1840 - 1900)



Legenda: grupos com até 10 ocorrências (verde); grupos com até 100 ocorrências (amarelo); grupos com mais de 100 ocorrências (laranja).  
Fontes: dados da pesquisa

Mapa 12 — Detalhamento das citações à cidade do Rio de Janeiro (1840 - 1900)



Legenda: grupos com até 10 ocorrências (verde); grupos com até 100 ocorrências (amarelo); grupos com mais de 100 ocorrências (laranja).  
Fontes: dados da pesquisa

Novamente, convidamos o leitor a acessar os mapas interativos a fim de explorar o banco de dados da maneira que mais lhe convier.

### 3.2.2 Mapas e gráficos: década a década

Neste capítulo, apresentaremos os dados individualizados para cada década examinada na pesquisa. Cada mapa é também acompanhado de um gráfico em forma de pizza, o qual mostra a frequência das menções às cidades. Assim, será possível contemplar as distâncias e

as frequências dos dados sobre a geografia do romance brasileiro. O objetivo é propor hipóteses que expliquem as ocorrências dos municípios citados ao longo do tempo. Além disso, o método que adotamos para analisar os dados passa pelo entendimento de que os mapas devem ser interpretados como ondas. A fim de explicar o conceito de ondas, Franco Moretti observa:

The tree describes the passage from unity to diversity: one tree, with many branches: from Indo-European, to dozens of different languages. The wave is the opposite: it observes uniformity engulfing an initial diversity: Hollywood films conquering one market after another (or English swallowing language after language). Trees need geographical discontinuity (in order to branch off from each other, languages must first be separated in space, just like animal species); waves dislike barriers, and thrive on geographical continuity (from the viewpoint of a wave, the ideal world is a pond). (MORETTI, 2015, p. 60)

Dessa forma, cada onda representa uma espécie de fotografia do momento analisado. Tal método é conveniente na medida em que uma análise comparativa nos parece ser uma alternativa interessante para se examinar o que ocorreu de uma década a outra. Em outras palavras, consideramos que as tendências, as continuidades e as novidades podem ser mais facilmente verificáveis dessa maneira.

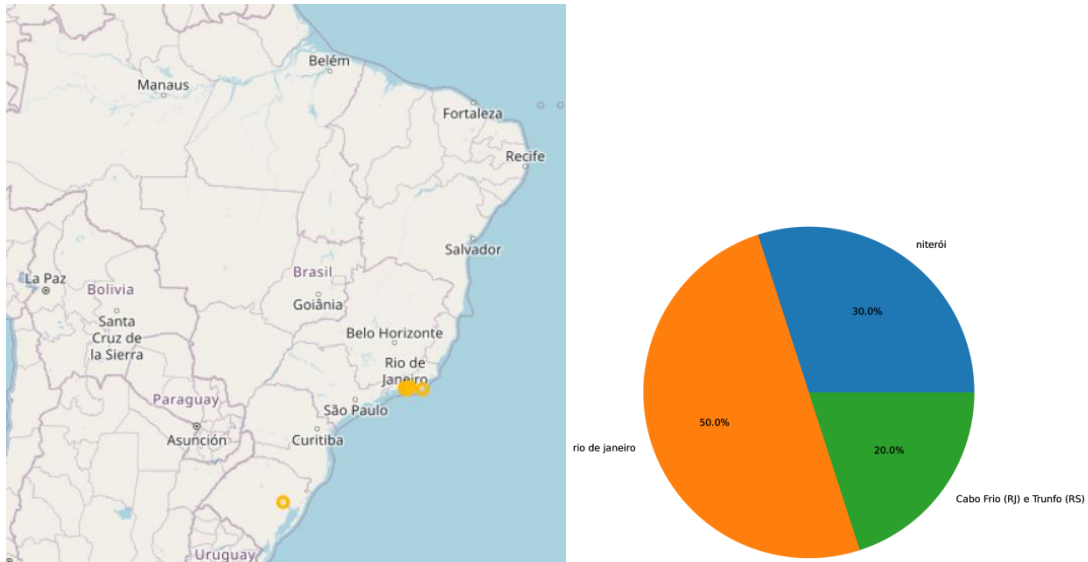
A despeito de já ter aparecido no subcapítulo anterior, a diferença entre os conceitos “representação” e “frequência” é ainda mais importante para os próximos subcapítulos. Em síntese, trabalhar com “representação” significa analisar a área geográfica mencionada nos romances. Trata-se da distância do território alcançado pelos romances em uma determinada década. Trabalhar com “frequência”, por sua vez, significa examinar a quantidade de menções que uma certa localidade recebe. Em outras palavras, o mais importante é o percentual de citações a uma cidade.

### 3.2.2.1 Década de 1840

Embora a década de 1840 seja fundamental para a história do Brasil, poucos romances foram publicados ao longo desses anos, de modo que os dados para a geografia do romance são relativamente escassos. Ainda assim, podemos tentar esboçar algumas análises. De início, no entanto, é importante entender o contexto no qual essas obras vieram à tona. É justamente ao longo dos anos 1840 que a maioria das contestações à centralização do poder político no Rio de Janeiro cessa. Além disso, o fracasso dessas revoltas e guerras civis (Sabinada e

Farroupilha, para citar alguns exemplos) representa a consolidação e o fortalecimento do Estado imperial brasileiro. Vejamos como é a geografia do romance nessa década:

Mapa 13 — Onda de 1840



Fontes: dados da pesquisa

À exceção de algumas cidades litorâneas, o mapa está quase todo por ser preenchido. O vazio territorial<sup>16</sup> é enorme, conseqüentemente. Nos seis romances publicados nesse período, somente quatro cidades foram mencionadas: Rio de Janeiro, Cabo Frio (RJ), Triunfo (RS) e Niterói (RJ). Como se pode constatar, a representação simbólica do território brasileiro ainda é extremamente limitada. Assim, queremos argumentar que ainda não se pode falar de processo de ficcionalização do território brasileiro nessa década, quer pelo baixíssimo número de obras publicadas, quer pela diminuta distribuição geográfica das poucas cidades mencionadas nas narrativas.

A escassez dos dados provavelmente contribuiu para a enorme polarização do território nacional por parte do Rio de Janeiro, cidade que representa 50% das citações na década em análise. De toda a forma, precisamos lidar com os dados que efetivamente obtivemos. Ademais, essa cidade nunca mais alcançaria tamanho nível de polarização em décadas posteriores, conforme veremos ao longo deste subcapítulo.

Em geografia urbana, utiliza-se o conceito de polarização para explicar o nível de atratividade relativa de uma localidade. Por exemplo, quanto maior for a economia de uma cidade, mais polarizadora e atrativa ela tende a ser. Dessa forma, podemos afirmar que, além

<sup>16</sup> Ainda delimitaremos melhor esse conceito. Nesse momento, basta entender que, por “vazio territorial”, queremos falar sobre as partes do território que ainda não foram citadas.

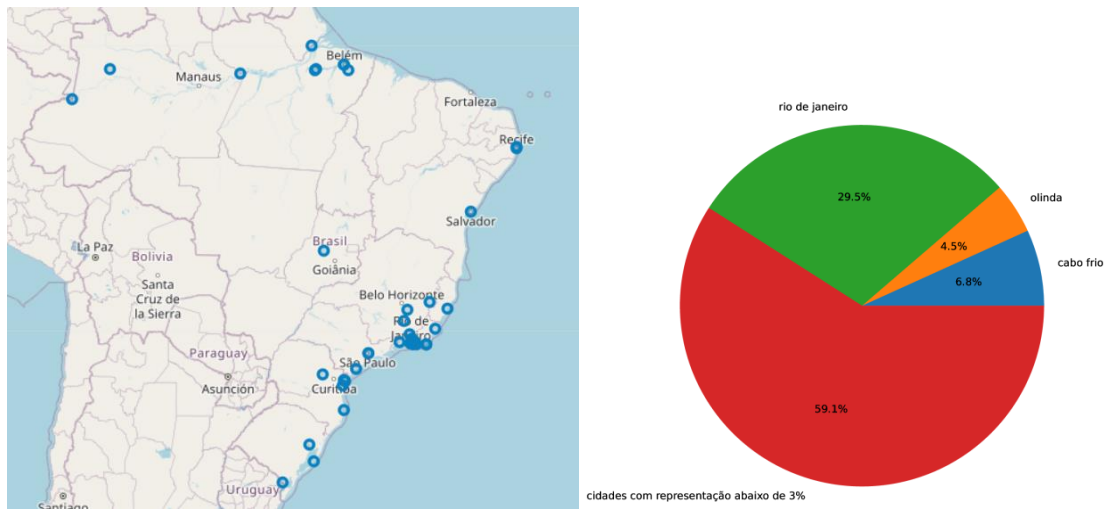
de ser o centro do poder imperial e de ter a economia mais importante da época, o Rio de Janeiro também atraía a imaginação dos escritores e leitores de maneira absolutamente superior a qualquer outra cidade, na década em exame.

A Onda de 1840 ainda sugere que Niterói, Cabo Frio e Triunfo também eram cidades representativas naquele momento. Como o romance brasileiro efetivamente começou na cidade do Rio de Janeiro, parece razoável afirmar que Niterói e Cabo Frio também fossem localidades relevantes nessa década, justamente em razão dessa proximidade geográfica à Corte imperial. A presença de Triunfo (10%), no entanto, não parece indicar a importância dessa cidade, uma vez que ela nunca mais apareceu entre as localidades mais citadas, em nenhuma das décadas subsequentes. Assim, trata-se de uma menção acidental, porquanto nenhum outro dado (histórico, geográfico, demográfico, por exemplo) ajuda a explicar ou contextualizar essa ocorrência.

#### 3.2.2.2 Década de 1850

A década de 1850 é caracterizada pela vitória do projeto político defendido pelo partido conservador brasileiro. Em síntese, a centralização do poder no Rio de Janeiro, a manutenção da escravidão e a monarquia parlamentar passaram a representar as bases do Brasil até a década de 1880, aproximadamente. Conseguimos analisar 16 romances publicados nessa década, *corpus* que já permite análises um pouco mais interessantes. Conforme argumentamos a seguir, é na Onda de 1850 que efetivamente se inicia o processo de ficcionalização do território nacional.

Mapa 14 — Onda de 1850



Fontes: dados da pesquisa

A maior quantidade de romances publicados e a melhor distribuição geográfica das cidades mencionadas permitem afirmar que a Onda de 1850 é o primeiro alargamento da geografia do romance brasileiro. Pela primeira vez, aparecem cidades do Rio Grande do Sul, de São Paulo, de Santa Catarina, do Paraná, de Goiás e do Amazonas, por exemplo. O Rio de Janeiro ainda é a cidade mais polarizadora do romance, embora já comece a perder representatividade. Cabo Frio e Olinda também aparecem com bastante força, ainda que nem combinadas atinjam a porcentagem do Rio de Janeiro.

Nesse momento, é importante notar que o gráfico da década de 1850 contém uma legenda (“cidades com representação abaixo de 3%”) que não existia nos dados anteriores. Trata-se de um mecanismo que permite individualizar as cidades mais representativas de cada década. Esse valor, que pode ser 3% ou 3,5%, varia em função do período examinado. Na ausência de um valor mínimo, os gráficos perdiam o objetivo, já que muitas cidades eram citadas uma única vez, mas apareciam com porcentagem maior que 1% por razões matemáticas.

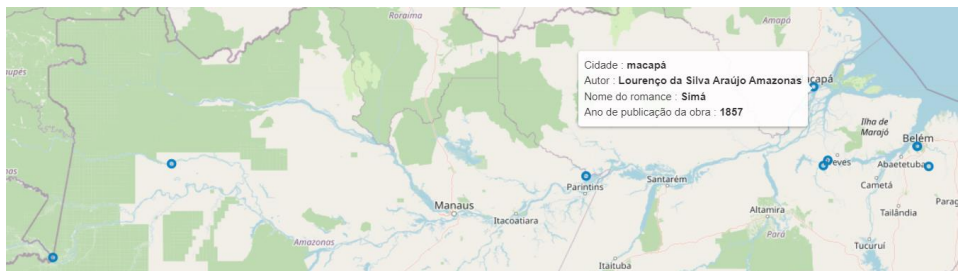
Na década de 1850, quase 60% das cidades mencionadas sequer atingiram o valor mínimo de 3,5%. Por um lado, essas cidades, individualmente, polarizam muito pouco o imaginário da época. Como tendência geral, podemos constatar que quanto maior for essa porcentagem, menor é a polarização das cidades mais importantes. Na década de 1840, por exemplo, a porcentagem dessas cidades era zero, e o Rio de Janeiro representava 50% das ocorrências, isto é: altamente polarizador. Nos anos 1850, por sua vez, 59,1% das cidades citadas não alcançaram aquele limite, o que resultou em menor polarização por parte do Rio de Janeiro. Por outro lado, muitas cidades secundárias - no sentido de não aparecerem com

muita frequência - são mencionadas já na década de 1850. Trata-se de uma expansão significativa em relação à década anterior.

Outro ponto importante é a presença de cidades amazônicas, o que não tinha acontecido na década anterior. Argumentamos que essas ocorrências decorrem mais da consciência individual de um autor, que de um processo orgânico do romance brasileiro da época. Em outras palavras: não fosse por Lourenço da Silva Araújo Amazonas, a região amazônica provavelmente não teria aparecido no mapa do romance nesse momento. Por “processo orgânico”, entendemos um desenvolvimento mais geral e coletivo do romance, como vamos desenvolver a seguir.

Em 1857, Lourenço da Silva Araújo Amazonas publicou o romance *Simá*. Apesar de tratar de temas históricos relacionados à região, Lourenço da Silva estava ciente dos problemas geopolíticos da Amazônia, uma vez que foi o Comandante Militar da Comarca do Alto Amazonas em meados do século XIX. Nessa época, o principal problema internacional da região era a exigência, por parte de países como Reino Unido e França, que o Brasil permitisse a navegação internacional dos rios amazônicos, o que só foi permitido no final da década de 1860.

Mapa 15 — Ocorrências do Norte no Romance *Simá*



Fontes: dados da pesquisa

Em regra, as potências europeias pressionavam o Brasil por meios diplomáticos. Os norte-americanos, no entanto, cogitaram o uso da força para controlar a Amazônia. De fato, entre 1850 e 1860, os fazendeiros do Sul dos Estados Unidos da América (EUA) e o Departamento de Estado discutiram uma estratégia militar para colonizar a região. Em resumo, a ideia era estender as plantações de algodão e, conseqüentemente, o sistema escravista de produção para o norte do Brasil, uma vez que já se previa o fim da escravidão nos EUA.<sup>17</sup> Assim, reiteramos o argumento de que a presença da região amazônica no

<sup>17</sup> Para conhecer melhor essa história, sugiro o livro *The Deepest South: The United States, Brazil, and the African Slave Trade*, de Gerald Horne.

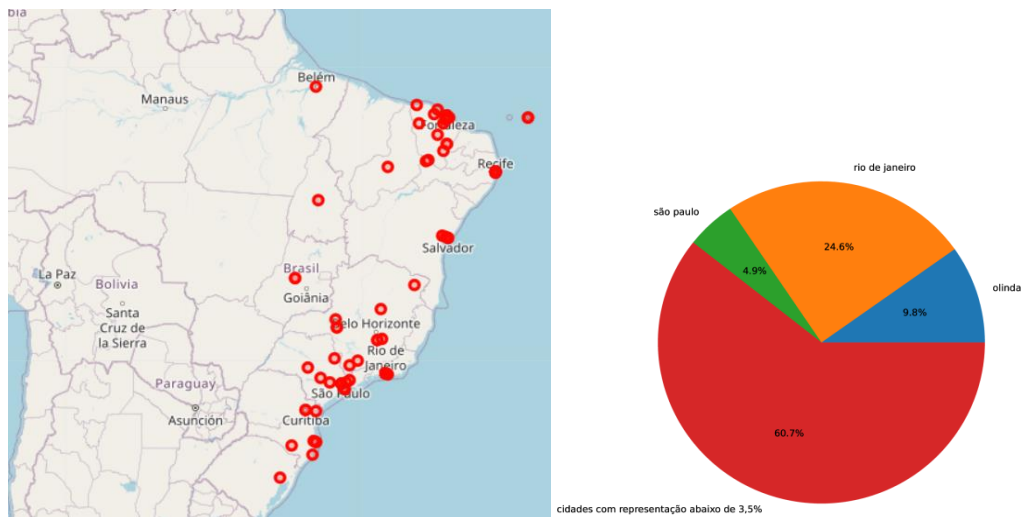


romance dos anos 1850 pode ser explicada mais por Lourenço da Silva Araújo Amazonas, que por um reconhecimento coletivo a respeito da importância da região para a ficcionalização do território brasileiro.

### 3.2.2.3 Década de 1860

A fim de examinar a década de 1860, ao menos dois acontecimentos históricos precisam ser mencionados: o início do alargamento da produção de café em direção ao Oeste paulista e os efeitos da Lei de Terras e do Código Comercial, ambos editados em 1850, para o ordenamento territorial brasileiro. Embora a Guerra do Paraguai tenha marcado profundamente os anos 1860, as consequências do conflito para a geografia do romance só podem ser observadas na década seguinte. Assim, a Onda de 1860 marca o início da interiorização do romance, notadamente através dos Estados de São Paulo e do Ceará.

Mapa 16 — Onda de 1860



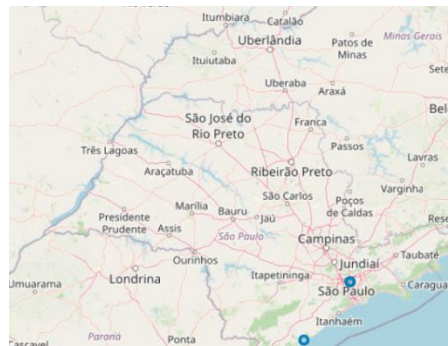
Fontes: dados da pesquisa

Na Onda de 1860, continuamos a acompanhar a diminuição da centralidade da cidade do Rio de Janeiro. Talvez o mais interessante, no entanto, seja observar a ascensão de São Paulo (a terceira cidade mais citada), que ocorreu nessa mesma década. Embora tenha polarizado o romance junto ao Rio e à Olinda, São Paulo ainda era amplamente considerada como um “burgo de estudantes” ou uma “cidade acadêmica” nesse momento, segundo relata

parte da historiografia a respeito da cidade<sup>18</sup>. Desse modo, é notável que essa cidade, demograficamente inferior a Recife e ainda economicamente menos relevante que o Rio de Janeiro, tenha polarizado tão fortemente o imaginário nacional já na década de 1860.

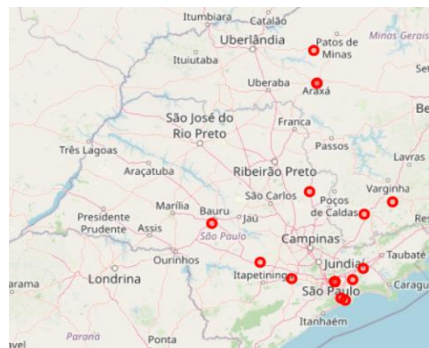
Outro aspecto importante da Onda de 1860 é que a interiorização do romance ocorre justamente em direção ao Estado ou à Província de São Paulo, como eram denominadas essas circunscrições políticas no Brasil imperial. Nessa década, o cultivo do café começava a adentrar esse Estado, o que implicaria a construção de estradas de ferro e a formação de bancos. A comparação das Ondas de 1850 e de 1860 permite verificar esse começo de interiorização:

Mapa 17 — Estado de São Paulo na Onda de 1850



Fontes: dados da pesquisa

Mapa 18 — Estado de São Paulo na Onda de 1860



Fontes: dados da pesquisa

Além de São Paulo, também o Ceará é responsável pela interiorização do romance na Onda de 1860. Esse estado, que simplesmente não aparece nas Ondas anteriores, é muito mencionado nos anos 1860, o que pode ser explicado por razões econômicas e geográficas.

<sup>18</sup> Ver a obra *A cidade de São Paulo no século XIX*, de Odilon Nogueira de Matos.



Com efeito, a Guerra de Secessão dos Estados Unidos acarretou uma diminuição do comércio de algodão entre os EUA e o Reino Unido. Esse conflito militar acabou por beneficiar o Ceará, uma vez que os britânicos começaram a importar mais algodão produzido no Brasil. O resultado foi não só um aumento da produção e da exportação de algodão cearense em direção ao Reino Unido, como também um aumento do comércio entre a capital portuária (Fortaleza) e o interior do estado do Ceará, como se pode ler no excerto abaixo.

Durante a safra, o comércio de Fortaleza apresentava grande movimento de mercadorias. Suas vias urbanas eram preenchidas por comboieiros, mascates e caixeiros viajantes que transportavam as mercadorias do interior para capital e vice-versa. Em 1866, a importância da cidade de Fortaleza já estava consolidada como maior polo econômico da província. (DE MOURA CUNHA, 2020, p. 17)

Nesse sentido, podemos afirmar que a dinâmica demográfica que levou à interiorização do Ceará também se manifestou no romance. Dessa forma, à medida que o interior do Estado se tornava mais importante econômica e demograficamente, mais menções a essa área ocorriam no romance da época.

Ainda a respeito do Nordeste, podemos observar que Olinda se mantém entre as cidades mais citadas da década (9,8%). Apesar de economicamente decadente, a cidade continuou a ser um dos centros culturais mais importantes do País durante o século XIX. Assim, essa pujança cultural pode explicar a força de Olinda na geografia do romance. Além disso, a forte presença de uma cidade nordestina permite argumentar que o processo de consolidação da região sudeste ainda não estava completo na década de 1860. Nas próximas décadas, no entanto, veremos uma contínua tendência de queda para todas as cidades nordestinas.

Outra característica importante da Onda de 1860 é que o Norte do Brasil aparece muito pouco. De fato, as menções à região amazônica quase desaparecem nesse momento (apenas Belém é citada). Ora, ao tratar do período precedente, defendemos que as menções a cidades amazônicas podiam ser explicadas pela consciência individual de um autor. Dessa forma, à medida que se tornava claro que a Amazônia pertencia, inequivocamente, ao Brasil, também diminuíram as menções a cidades da região, conforme se pode confirmar no mapa 13. A abertura à navegação internacional dos rios amazônicos, decretada pelo Brasil no final da década de 1860, não só põe termo a essa disputa geopolítica, como também corrobora a nossa análise.

Por fim, outro aspecto importante é o leve aumento percentual das cidades que, individualmente, não atingiram 3,5% de representatividade. De 1840 a 1860, a porcentagem

dessas cidades aumenta de maneira estável, com uma certa tendência de estabilização. Como argumentamos anteriormente, esse número indica que existe uma menor concentração geográfica do romance da época.

#### 3.2.2.4 Década de 1870

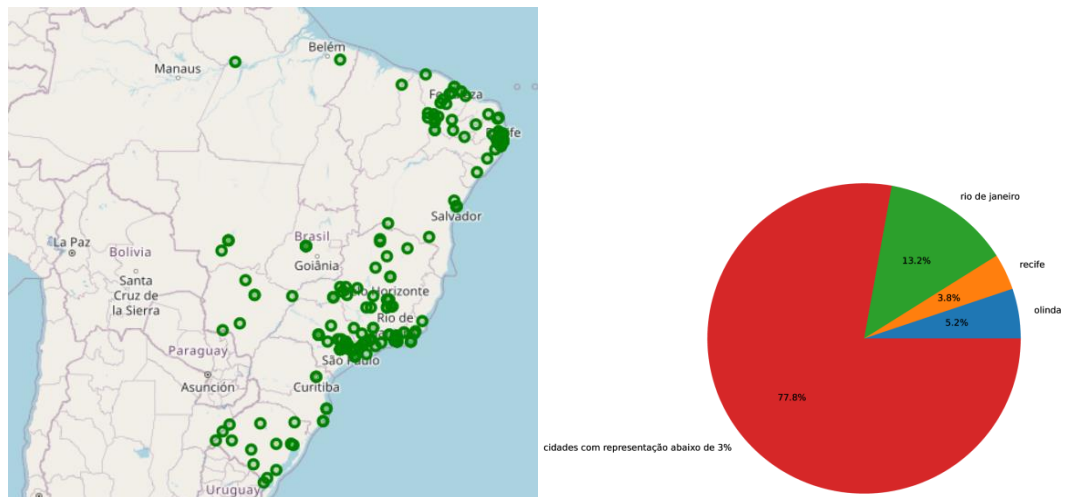
A década de 1870 marca o início de um movimento de contestação política e ideológica a dois aspectos fundamentais na história do Brasil: a escravidão e o “parlamentarismo às avessas”<sup>19</sup>. Em síntese, tratou-se da conscientização de que a escravidão não era sustentável a longo prazo e de que a monarquia, apesar de ter adotado práticas parlamentaristas desde a década de 1840, ainda podia contrariar a maioria eleita, caso lhe conviesse. Para a geografia do romance, tratou-se do período mais representativo da série histórica, uma vez que 35% dos romances examinados neste trabalho foram publicados nessa década.

A Onda de 1870 assinala uma significativa expansão da geografia do romance brasileiro, já que não só a região central e o interior de São Paulo aparecem com forte representatividade, como também o interior do Rio Grande do Sul aparece pela primeira vez. Apesar desse alargamento, no entanto, o vazio territorial permanece, porquanto várias cidades ainda não existem para o romance nacional, conforme exemplificam regiões como o interior do Pará, os Estados do Tocantins e do Maranhão, a maioria do Estado do Amazonas, o Norte do Mato Grosso, o interior de Curitiba e o interior de Santa Catarina. Vejamos o mapa do romance desse momento.

---

<sup>19</sup> Expressão cunhada por José Murilo de Carvalho, “parlamentarismo às avessas” se refere a maneira como os gabinetes eram constituídos durante o Período Imperial. Em resumo, contrariamente ao parlamentarismo britânico, no qual a maioria legislativa nomeia o Primeiro Ministro, cabia ao Imperador nomear o Presidente do Conselho de Ministros (o Primeiro Ministro).

Mapa 19 — Onda de 1870



Fontes: dados da pesquisa

Na verdade, não surpreende que a Onda de 1870 seja a de maior intensidade para a geografia do romance brasileiro. Queremos argumentar que essa ampliação territorial resulta de dois acontecimentos históricos importantes: as novas ideias da Geração de 1870 e as consequências da Guerra do Paraguai. Segundo Ângela Alonso, para contestar a ordem imperial, os intelectuais da chamada Geração de 1870 precisaram, primeiramente, se conscientizar das práticas sociais, políticas públicas e valores adotados pela elite imperial. Em *Ideias em Movimento*, lemos que

O sentido das manifestações intelectuais da geração 1870, procurei mostrar, é precisamente o contrário da "evasão", do "alheamento", da "indiferença" em relação à realidade nacional. Os membros do movimento intelectual desenvolveram interpretações críticas acerca dos principais problemas brasileiros e buscaram instrumentos para intervir politicamente. Trata-se de um pensamento engajado, que analisou e contestou o status quo monárquico. (ALONSO, 2005, p. 337)

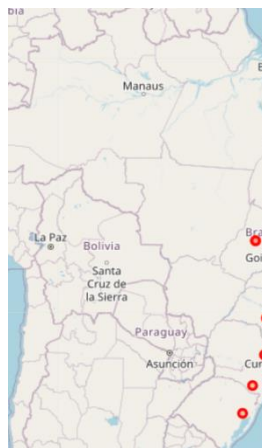
De modo geral, os romancistas da época faziam parte ou da elite dirigente ou de camadas sociais confortáveis, como advogados, médicos e funcionários públicos. Dessa forma, parece razoável supor que também os romancistas se conscientizaram dos “principais problemas brasileiros”. Nessa época, as fronteiras internacionais do Brasil representavam um desses problemas. Por ter envolvido toda a sociedade brasileira, a Guerra do Paraguai gerou um novo entendimento sobre o território, sobretudo a respeito das fronteiras nacionais. De fato, a ideia da pátria se fortaleceu, como se pode ler no excerto seguinte:

Talvez pela primeira vez um sentido positivo de ‘pátria’ tenha começado a se desenvolver, em especial no início da guerra, com os batalhões de voluntários e as

primeiras vitórias. A bandeira começou a ser hasteada de modo sistemático, o imperador virou o líder da nação empenhado em conseguir apoio dos dois partidos, e surgiram os heróis nacionais, Caxias, Osório, Barroso. (SCHWARCZ, 2015, p.299)

Assim, o que pode explicar a ocorrência de cidades fronteiriças na Onda de 1870 é a conscientização de que o território brasileiro não se limitava somente ao litoral. Em outras palavras, também os romancistas começaram a escrever sobre regiões mais afastadas da costa leste, notadamente as fronteiras internacionais do Brasil no Centro-Oeste e no Sul do País. Podemos comparar dois gráficos a seguir, os quais focam nas áreas de fronteiras:

Mapa 20 — Área de fronteira na Onda de 1860



Fontes: dados da pesquisa

Mapa 21 — Área de fronteira na Onda de 1870

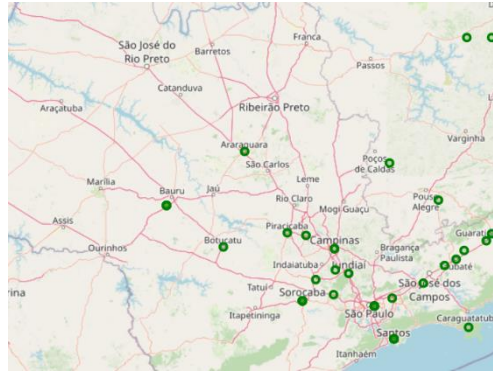


Fontes: dados da pesquisa

Apesar desse notável alargamento do qual estamos tratando, podemos constatar a quase ausência de cidades do Norte do Brasil. Com efeito, recordamos que, na década anterior, a região também não apareceu, o que sugere uma falta de interesse dos romancistas



Mapa 23 — Estado de São Paulo na Onda de 1870

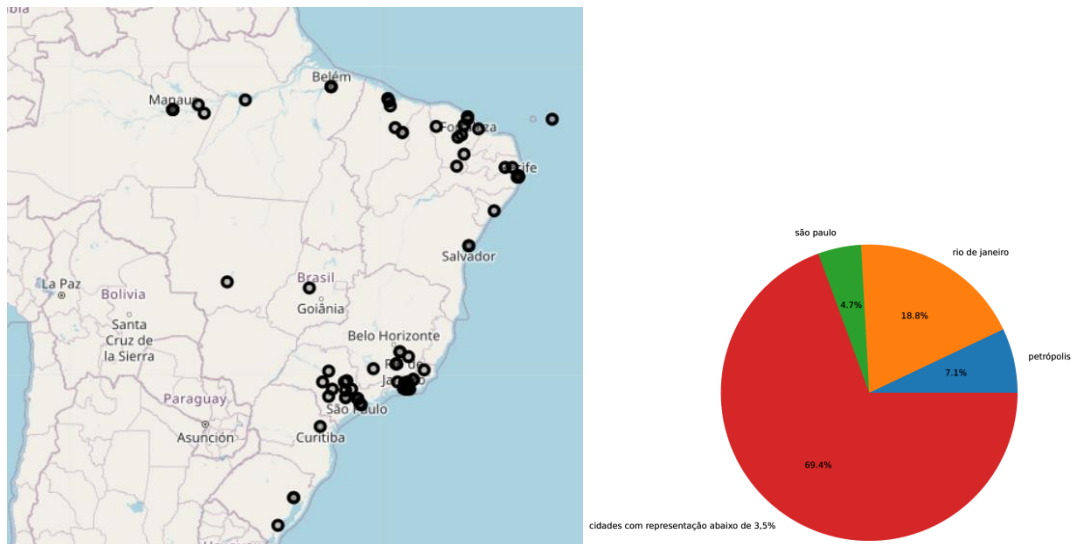


Fontes: dados da pesquisa

### 3.2.2.5 Década de 1880

A década de 1880 também é marcada por importantes transformações econômicas, políticas e sociais no Brasil. É nessa década que algumas das ideias preconizadas pela Geração de 1870 tiveram êxito, como exemplificam a abolição da escravidão, a adoção do pan-americanismo ou a Proclamação da República. Além disso, a economia paulista se tornou mais dinâmica que a carioca, e a região amazônica passou pelo que a historiografia costuma chamar de “boom da borracha”. Ao captar essas dinâmicas, a Onda de 1880 é caracterizada por uma alta concentração de menções ao sudeste brasileiro (Rio de Janeiro, Petrópolis e São Paulo são notadamente as três cidades mais recorrentes nessa década) e pelo reaparecimento da região amazônica no mapa. Por fim, queremos lembrar que o número de romances publicados na década de 1880 é significativamente menor que na década anterior. Dessa forma, assim como em diversos outros momentos desta pesquisa, as análises apresentadas a seguir precisam ser entendidas de forma preliminar, na medida em que faltam dados para corroborar interpretações mais assertivas.

Mapa 24 — Onda de 1880

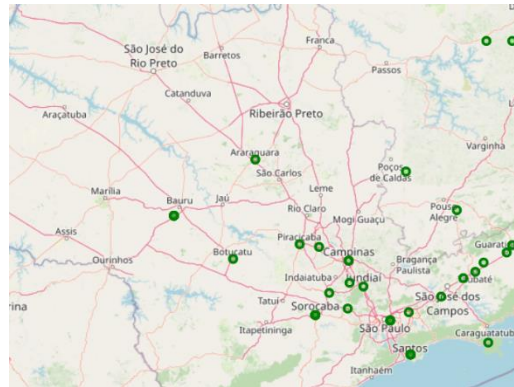


Fontes: dados da pesquisa

Embora as menções ao Rio de Janeiro tenham efetivamente aumentado de 1870 a 1880, continuamos a observar a tendência de enfraquecimento dessa cidade na geografia do romance. Na verdade, por se tratar de uma pequena diferença percentual (de 13,2% para 18,8%), podemos perceber mais uma estabilização do que uma real diminuição de sua frequência. De toda a forma, esse último aspecto será mais facilmente perceptível a partir dos dados da década de 1890, como veremos em breve.

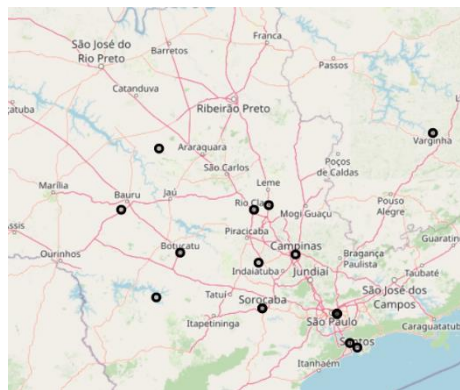
Também chama a atenção que, contrariamente à década de 1870, a relativa queda da cidade do Rio de Janeiro seja acompanhada pela maior presença tanto da cidade quanto do Estado de São Paulo. De fato, São Paulo volta a ser uma das cidades mais mencionadas da década, com mais ou menos a porcentagem que tinha em 1860 (mais precisamente, 4,9% em 1860; 4,7% em 1880). Além disso, podemos notar que, à medida que a cidade é mais citada, o interior do estado perde representatividade, conforme podemos visualizar nos mapas abaixo, os quais permitem comparar as duas décadas em exame.

Mapa 25 — Estado de São Paulo na Onda de 1870



Fontes: dados do trabalho

Mapa 26 — Estado de São Paulo na Onda de 1880

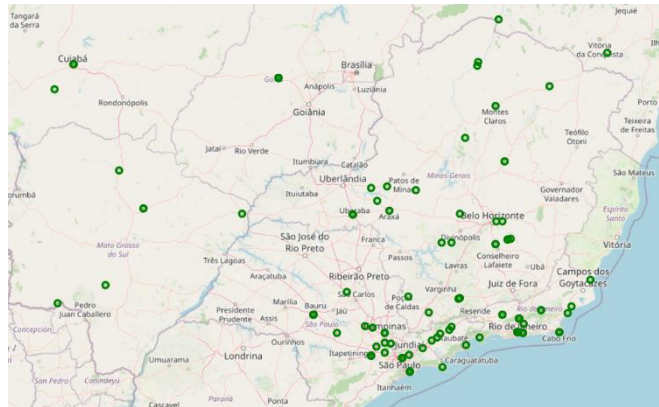


Fontes: dados do trabalho

Comparada à década precedente, a Onda de 1880 também é caracterizada por uma menor presença do interior do Brasil, o que contrasta com a tendência de expansão observada desde a década de 1850. À guisa de exemplo, o sudeste e o centro-oeste do País têm pouca presença na geografia do romance dos anos 1880, ao passo que essas mesmas áreas apareceram nitidamente nos anos 1870. Mais especificamente, significativas porções dos estados do Mato Grosso do Sul e de Minas Gerais simplesmente desapareceram dos dados de 1880, conforme se poderá ver a seguir.

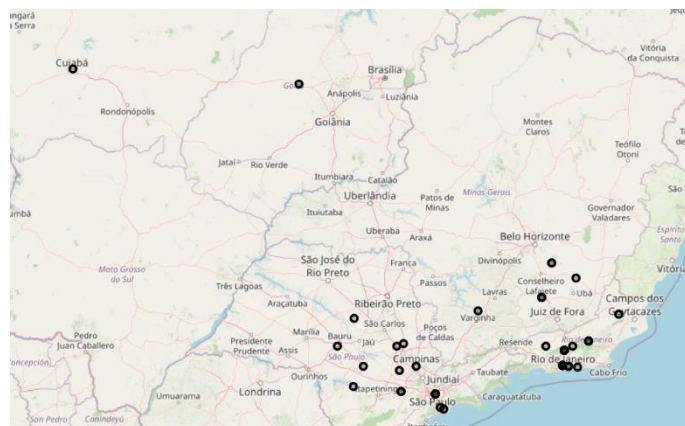


Mapa 27 — O sudeste e o centro-oeste na Onda de 1870



Fontes: dados do trabalho

Mapa 28 — O sudeste e o centro-oeste na Onda de 1880



Fontes: dados do trabalho

Como se pode ver, o alargamento territorial alcançado na década de 1870 não foi mantido na década imediatamente posterior. Como resultado, a geografia do romance de 1880 se parece mais com a de 1860 que com a de 1870. Nesse sentido, a característica principal da Onda de 1880 é a retomada da alta polarização por parte de algumas cidades, notadamente Rio de Janeiro e São Paulo, em detrimento de uma maior expansão territorial da geografia do romance.

Embora não seja simples explicar as razões para essa nova concentração geográfica, podemos tentar propor uma explicação de maneira preliminar. A consolidação da economia paulista e a importância política e cultural do Rio de Janeiro, sem dúvida, contribuíram para esse novo movimento de concentração. É como se depois da expansão geográfica ocorrida em 1870, as dinâmicas socioeconômicas se voltassem aos centros urbanos que efetivamente polarizavam o Brasil da época. Como queremos argumentar neste trabalho, o romance

também capta esses movimentos. Lilia M. Schwarcz e Heloisa M. Starling corroboram assim afirmam:

Importa lembrar, ainda, o intenso movimento de migração interna, resultado da lenta desmontagem do sistema escravocrata. No período que vai de 1872 a 1900, foi na Região Nordeste que houve maior perda populacional, como consequência do comércio interno de escravos que despovoou a economia do açúcar e do algodão, e reforçou a densidade dos estados cafeeiros. Castigados pelas secas de 1870 e 1880, grupos de migrantes dirigiram-se para o Rio de Janeiro, que funcionava como chamariz cultural, além de se apresentar como provedor de empregos em geral, e mais especificamente para o funcionalismo público e estatal. (SCHWARCZ, 2015, p.326)

Um outro aspecto interessante da Onda de 1880 é o ressurgimento da região Norte na geografia do romance. Diferentemente do que ocorreu em 1850, no entanto, as menções à Amazônia na década em análise parecem fazer parte de um movimento mais orgânico de conscientização a respeito da importância da região para o Brasil. Em outras palavras, vários são os autores que citaram a Amazônia em 1880, o que contrasta com a década de 1850, na qual apenas Lourenço da Silva Araújo Amazonas a mencionou, conforme já discutimos. Em síntese, argumentamos que esse ressurgimento pode ser explicado pelas consequências do “boom da borracha” na região, conforme vamos desenvolver a seguir.

O Teatro Amazonas, inaugurado em 1896, pode ser considerado o ápice do chamado “ciclo da borracha”, termo usado por grande parte da historiografia brasileira para denominar esse acontecimento. Esse ciclo econômico, que ocorreu no Norte do Brasil, entre 1880 e 1900, é caracterizado pela preponderância da borracha como o produto mais representativo das exportações internacionais para essa região. Tal dinamismo econômico acarretou a migração de milhares de nordestinos para o Norte do Brasil, os quais fugiam da seca em direção de uma região que crescia economicamente. Segundo João Antônio de Paula,

Além do mais, é possível dizer que o crescimento da produção brasileira de borracha deveu-se, do lado da oferta, ao significativo processo da imigração para a Amazônia, sobretudo de nordestinos, em duas grandes ondas. A primeira está relacionada às grandes secas de 1877-1880, que determinaram a transferência da produção nordestina, em particular de cearenses, para a Amazônia. Estima-se que cerca de 500 mil pessoas foram obrigadas a abandonar as áreas devastadas pela seca. (SCHWARCZ, 2012, p.186)

Dessa forma, podemos argumentar que as consequências geográficas e políticas do “boom da borracha” podem explicar a forte presença da região nos romances dos anos 1880. De fato, o professor Carlos José de Farias Pontes afirma que a dinâmica econômica que marcou a região resultou em um desenvolvimento muito expressivo de Belém e Manaus:

O *boom gumífero* propiciou às capitais das províncias do Grão-Pará e Amazonas, respectivamente, Belém e Manaus, um desenvolvimento urbano sem precedentes na história. A riqueza promovida pela borracha permitiu que essas cidades participassem do espírito eufórico da Belle Époque. A Belle Époque (bela época) refere-se ao triunfo burguês no momento em que se notava grande desenvolvimento material e tecnológico. O comércio expandia-se pelo mundo e integrava-se com regiões de vários pontos do planeta.” (PONTES, 2014, p. 13)

Por fim, a respeito das localidades que não atingiram nem 3,5% das citações, convém notar que a quantidade diminuiu nessa década, apesar de ainda ter permanecido bastante alta (69,4%). Conseqüentemente, podemos afirmar que o número de cidades brasileiras que aparecem, mas que, individualmente, não polarizam a geografia do romance permanece alta, tendência que pode ser observada desde a década de 1850. Em outras palavras: a despeito da forte polarização por parte de São Paulo, Rio de Janeiro e Petrópolis<sup>20</sup> nos anos 1880, muitas outras cidades foram citadas de maneira mais ou menos residual.

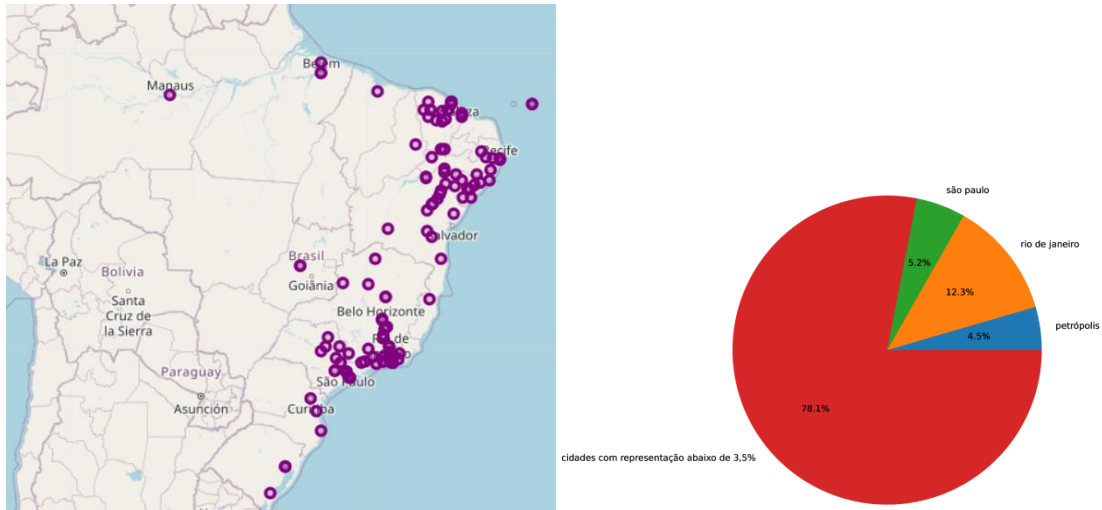
#### 3.2.2.6 Década de 1890

A década de 1890 é marcada não só pelas conseqüências da Proclamação da República, como também pelas repercussões do fim da escravidão e pela imigração maciça de europeus para o Brasil. A despeito de terem ocorrido durante a década de 1880, esses três acontecimentos-chave para a história do País repercutiram mais fortemente ao longo dos anos 1890. Desse modo, podemos perceber tanto continuidades quanto novidades no romance dessa época: por um lado, a Onda de 1890 representa a consolidação de dinâmicas que já estavam em curso na década anterior, tais como a preponderância do sudeste e a fraca presença da região Norte na geografia do romance brasileiro; por outro lado, a Onda de 1890 corresponde não só à interiorização do Estado de Minas Gerais, como também ao alargamento da representação do Nordeste brasileiro, duas dinâmicas ausentes na década de 1880, conforme podemos ver nos dados abaixo.

---

<sup>20</sup> Trataremos do caso de Petrópolis no próximo subcapítulo.

Mapa 29 — Onda de 1890



Fontes: dados da pesquisa

De início, precisamos examinar mais detidamente o sudeste brasileiro, que se mantém como a região mais mencionada nos romances. De fato, a despeito de exibirem porcentagens diferentes de uma década a outra, São Paulo, Rio de Janeiro e Petrópolis permanecem como as cidades mais citadas nos anos 1890, o que apenas confirma uma tendência verificada na década anterior. Ademais, São Paulo atinge o seu ápice (5,2%) e o Rio de Janeiro, o seu mínimo (12,3%) nesta última quadra histórica. Também é preciso notar que, apesar de efetivamente perder presença em termos percentuais, o Rio de Janeiro continua a ser a cidade mais polarizadora do romance nesse momento. Dessa forma, argumentamos que foi efetivamente durante os anos 1890 que o sudeste se consolidou como a região concentrada do romance brasileiro<sup>21</sup>.

Em um primeiro momento, essa última afirmação pode parecer óbvia, sobretudo se considerarmos que o Rio de Janeiro sempre teve um lugar de destaque na geografia do romance brasileiro. Ao examinar o intervalo 1850-1890, no entanto, constatamos que é somente no final do século XIX que três cidades situadas no sudeste figuram entre as mais citadas. Senão, vejamos: nos anos 1850, foram Rio de Janeiro, Olinda e Cabo Frio; nos 1860, foram Rio de Janeiro, São Paulo e Olinda e nos 1870, foram Rio de Janeiro, Recife e Olinda. Vale notar, ainda, que optamos por desconsiderar a década de 1840 nessa análise, na medida em que o processo de ficcionalização do território brasileiro - conforme definimos anteriormente - começou somente em 1850.

<sup>21</sup> Apenas para lembrar que, por “região concentrada”, entendemos a região que mais polarizou o imaginário dos romances.

Ainda a respeito do sudeste, chama a atenção a presença de Petrópolis como uma das cidades mais frequentes ao longo da década em exame. Na verdade, essa posição de destaque também confirma uma tendência verificada nos anos 1880. As razões pelas quais tal cidade figura entre as localidades mais mencionadas no final do século XIX parecem ser razoavelmente evidentes. Em síntese, Petrópolis acabou por polarizar o imaginário do romance na medida em que era uma cidade fundamental para a elite brasileira, quer turística, quer politicamente. Conforme afirma Roberta dos Santos Gregório, da segunda metade até o final do século XIX,

... a cidade se desenvolvia rapidamente, com forte tendência aristocrática, por força da presença do Imperador e de sua corte, nas temporadas do verão petropolitano. Nobres, políticos, diplomatas, grandes senhores e todo seu *entourage*, ricos negociantes e a intelectualidade da época se transferia para Petrópolis, durante um semestre a cada ano. Palacetes eram construídos para morada dessa gente abonada. Quem não tinha moradia se hospedava em hotéis e casas de família. E a cidade assumia um aspecto elegante (GREGÓRIO, 2023, p. 6)

Como se pode perceber, Petrópolis foi uma cidade fundamental para o Império desde a década de 1840. Assim, uma questão interessante é entender por que ela não se destacou tão fortemente antes de 1880. Apesar de realmente não ter aparecido em todas as décadas analisadas, Petrópolis já tinha sido citada nos anos 1850 e 1870. Além disso, é preciso ressaltar que a cidade se tornou ainda mais importante no final da década de 1880, o que também ajuda a explicar a alta frequência com a qual foi citada nos romances a partir dessa década. De fato, é o que afirma a já citada pesquisadora no mesmo artigo:

(...) o Imperador, já doente, passou mais tempo em Petrópolis do que no Rio de Janeiro, o que tornou a cidade num centro político, pois foi a capital do Brasil durante os três anos que antecederam o fim do Império. GREGÓRIO, 2023, p. 7)

A conclusão é que o sudeste polarizava não só a economia e a política, como também o imaginário do Brasil nessa época. Uma maneira de medir a importância do sudeste, tanto social quanto economicamente, é observar os dados a respeito da imigração internacional para o estado de São Paulo. De 1890 a 1929, desembarcaram aproximadamente 3.500.000 de estrangeiros no Brasil, dos quais cerca de 2.000.000 no estado de São Paulo, o que representa 57% do total dos estrangeiros que vieram para o País.<sup>22</sup> Assim, considerando essa enorme presença do sudeste na geografia do romance, podemos afirmar que a Onda de 1890 parece

---

<sup>22</sup> Ver o artigo *Participação dos imigrantes italianos no desenvolvimento de Minas Gerais*, de Domingos A Giroletti.

ter efetivamente captado algumas das dinâmicas que estavam em curso no Brasil. Segundo Lilia Schwarcz e Heloisa Starling:

A população concentrou-se em algumas grandes cidades. O Rio de Janeiro seria o coração da República, São Paulo a cabeça, e anos depois viria Belo Horizonte, a cidade que a República imaginou à sua imagem e semelhança, e que se tornou a primeira urbe planejada do Brasil. Os três centros controlariam recursos e mostrariam como o eixo econômico do país estava definitivamente virado para a região sudeste. (SCHWARCZ, 2015, p. 326)

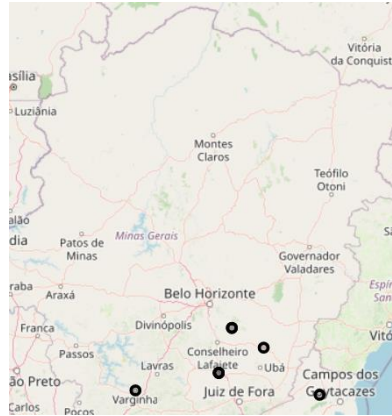
Como se pôde ler no trecho acima, Rio de Janeiro, São Paulo e Belo Horizonte eram as três cidades mais importantes no final do século XIX. Apesar da inegável importância histórica de Belo Horizonte, no entanto, a cidade nem sequer é citada no romance da década de 1890. A explicação para isso parece ser relativamente simples. Como a inauguração de BH ocorreu no final de 1890 - mais precisamente em 1897 - os romancistas devem ter optado por não situar seus romances em uma cidade em construção. A ausência de Belo Horizonte não significa, entretanto, que o Estado de Minas Gerais não fosse importante para a geografia do romance da época, conforme veremos a seguir.

Apesar da ausência de Belo Horizonte, a interiorização de Minas Gerais pode ser percebida na geografia do romance de 1890. Esse aspecto é fundamental, uma vez que, na década anterior, Minas Gerais esteve praticamente ausente do romance nacional. A preocupação com a posição geográfica da capital, que se manifestou em várias discussões nacionais e regionais, quer pela sociedade civil, quer pela elite dirigente<sup>23</sup>, pode explicar a presença do interior desse Estado nos romances. De fato, uma das razões que embasaram a mudança da capital de Minas Gerais, até então Ouro Preto, para Belo Horizonte foi justamente a maior centralidade de BH.<sup>24</sup> Dessa forma, não surpreende que se possa notar uma grande diferença entre o interior desse estado nas décadas de 1880 e 1890, conforme se pode ver nas imagens abaixo.

<sup>23</sup> A respeito das discussões sobre a nova capital de Minas Gerais, ver *A fundação de Belo Horizonte: ordem, progresso e higiene, mas não para todos*, de Patrícia Capanema Alvares Fernandes.

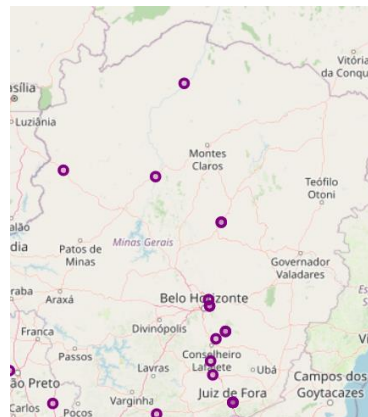
<sup>24</sup> Conforme assinala Patrícia Capanema Alvares Fernandes, “Três atributos essenciais para a nova capital podem ser extraídos dessa proposta. Primeiramente, centralidade era uma característica importante.” (AIVARES, 2010, p. 5)

Mapa 30 — Estado de Minas Gerais na Onda de 1880



Fontes: dados da pesquisa

Mapa 31 — Estado de Minas Gerais na Onda de 1890



Fontes: dados da pesquisa

Os resultados obtidos para a região norte correspondem a um outro aspecto de continuidade a respeito da década anterior. De fato, embora tenha perdido tanto frequência quanto representatividade em relação aos anos 1880, a região Norte continuou a ser citada de maneira mais orgânica nos anos 1890, o que nos parece indicar a manutenção de um movimento de maior consciência acerca da importância da região para o Brasil. No entanto, se considerarmos que o “boom da borracha” se intensifica na década em exame, também é possível afirmar que a representação da região norte ainda é subestimada no romance. Em outras palavras, poderíamos esperar uma presença muito maior da região, sobretudo em razão da forte dinâmica econômica, migratória e urbana pela qual passou nesse momento, o que não se manifestou nos enredos dos romances da época. Na verdade, essa divergência reitera que o romance não é um espelho acrítico e imediato da realidade. Em síntese, podemos dizer que o Norte do País nunca polarizou a geografia do romance nacional com a mesma intensidade que o sudeste, nem sequer quando a região passava por profundas transformações de toda a sorte.

Nessa quadra histórica, a posição ocupada pelo Nordeste é uma novidade em relação à década de 1880. Desde os anos 1850, o Nordeste aparece com bastante regularidade na geografia do romance brasileiro, o que não surpreende, uma vez que se trata de uma região fundamental para a construção da identidade brasileira. De 1850 a 1870, podemos perceber o crescimento da área ficcionalizada nessa região, com destaque para os estados do Ceará e de Pernambuco. A interiorização desse estado, no entanto, demorou um pouco mais para ocorrer. A despeito da extensa ampliação do interior do Nordeste, alcançada ainda durante a década de 1870, é apenas nos anos 1890 que podemos constatar a interiorização da Bahia no romance. A despeito da relativa decadência econômica e política do Nordeste nessa década, é interessante poder observar a interiorização de um dos estados mais importantes do Brasil, conforme podemos constatar nas imagens abaixo.

Mapa 32 — Nordeste na Onda de 1880



Fontes: dados da pesquisa



Mapa 33 — Nordeste na Onda de 1890



Fontes: dados da pesquisa

O último aspecto do qual queremos tratar é a conclusão do processo de expansão territorial que vinha se manifestando no romance desde a década de 1850. De fato, as Ondas de 1850, 1860 e 1870 mostraram uma ampliação significativa do território ficcionalizado no romance nacional. Em outras palavras, ao longo desse intervalo temporal, podemos observar que cada vez mais cidades passaram a servir como locais para os enredos das obras. Na década de 1880, no entanto, essa tendência parece chegar ao fim. É ao observar a década de 1890 que podemos confirmar o término desse processo, no entanto. A fim de facilitar essa constatação, dispomos abaixo os mapas do intervalo 1850-1890:

Mapa 34 — Onda de 1850



Fontes: dados da pesquisa

Mapa 35 — Onda de 1860



Fontes: dados da pesquisa

Mapa 36 — Onda de 1870



Fontes: dados da pesquisa

Mapa 37 — Onda de 1880



Fontes: dados da pesquisa

Mapa 38 — Onda de 1890



Fontes: dados da pesquisa

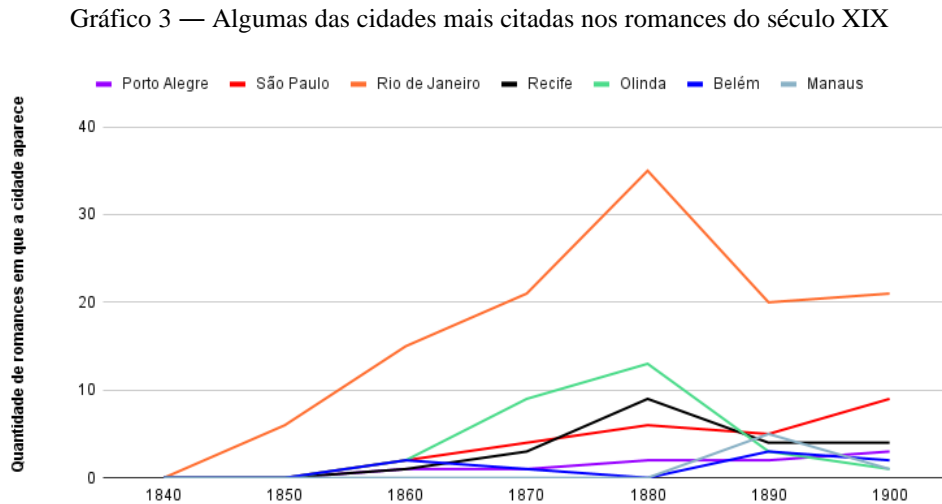
A constatação de que a expansão territorial não se manteve nas últimas décadas do século XIX resulta em uma consequência importante e uma pergunta intrigante, no mínimo. A consequência é que o grande vazio territorial do qual falamos anteriormente não é preenchido ao final do século, contrariamente ao que se podia prever em um primeiro momento. Nesse sentido, o romance nacional é incompleto, uma vez que não consegue abarcar toda a extensão territorial do Brasil. De fato, muitas áreas geográficas permanecem sem ser ficcionalizadas nos romances. Em síntese, trata-se dos já citados Estados do Tocantins, do Amazonas, bem como da parte norte do Estado do Mato Grosso e do interior dos Estados do Pará, de Santa Catarina e do Paraná. Não podemos esquecer que essas áreas representam cerca de 90% do território brasileiro, conforme as estimativas apresentadas anteriormente neste trabalho.

A explicação para o término dessa expansão parece ser uma tarefa impossível nesse momento da pesquisa. De fato, para os fins desta dissertação, trata-se do término de um processo que começou em 1840 e terminou em 1900, uma vez que estamos estudando apenas romances publicados durante o século XIX. É muito provável, no entanto, que o processo de ficcionalização do território brasileiro se aprofunde ao longo do século XX, sobretudo a partir da década de 1930, como também já argumentamos anteriormente. Mais uma vez, podemos trabalhar essas questões em um doutorado.

### 3.2.3 Cidades importantes em função do tempo

Na subseção anterior, apresentamos os dados 1) em formato de mapa, o que nos permitiu visualizar o processo de ficcionalização do território brasileiro, e 2) em formato de gráfico, o que nos possibilitou examinar a frequência (o nível de polarização) com a qual as

idades apareceram ao longo do tempo. Nesta subseção, decidimos apresentar um gráfico com a frequência de algumas das cidades mais citadas (localizadas no eixo vertical) em função do tempo (localizado no eixo horizontal). O objetivo é comparar, em números absolutos, as ocorrências das cidades de Recife, Porto Alegre, Rio de Janeiro, Manaus, Belém, São Paulo e Olinda, como se pode ver no gráfico abaixo.



Eixo vertical: quantidades de romances em que a cidade aparece; eixo horizontal: tempo medido em décadas.  
Fontes: dados da pesquisa

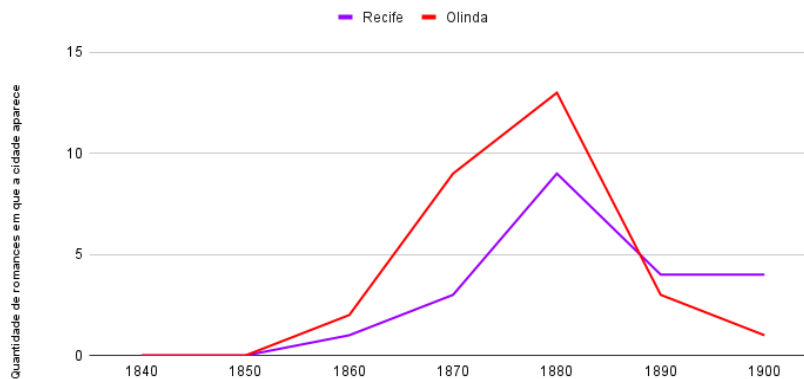
O Rio de Janeiro é efetivamente a cidade mais citada no romance oitocentista brasileiro. A ascensão dessa cidade ocorreu de maneira estável até a década de 1880, quando se pode perceber um declínio, seguido de uma outra estabilização, que permanece até o final do século. É na década de 1870 que se verifica o ápice do Rio de Janeiro como a cidade mais polarizadora do romance. Além disso, a partir de 1880, a cidade passou a perder força na geografia do romance, apesar de continuar a ser a cidade mais representativa durante todo o século em exame. Nesse sentido, a capital do Império representa o principal eixo simbólico do romance brasileiro.

A respeito de São Paulo, podemos notar que houve um crescimento de representatividade a partir de 1870, estabilização em 1880 e novo crescimento em 1890. Outro aspecto interessante é que, à medida que o Rio de Janeiro se enfraqueceu, São Paulo se fortaleceu na geografia do romance, notadamente a partir dos anos 1890. Queremos chamar esse movimento de “mudança do eixo simbólico do romance nacional”, que consiste justamente no crescimento da importância cultural de São Paulo em detrimento da posição histórica ocupada pelo Rio de Janeiro. De fato, embora não se complete no século XIX, essa mudança de eixo já é uma tendência a partir do final do período examinado.

A crescente importância econômica de São Paulo certamente influenciou a paulatina mudança desse eixo de representação. Na década de 1880, São Paulo passou a ser a província mais rica do Império, conforme já discutimos anteriormente. Podemos afirmar, portanto, que a crescente presença da cidade e do estado de São Paulo na geografia do romance coincide, também, com a superação da economia carioca pela paulista e com as consequências demográficas e políticas dessa suplantação. Se tivéssemos analisado os romances do século XX, provavelmente teríamos observado a conclusão desse processo. Novamente, convém ressaltar a oportunidade de continuar essa pesquisa em um doutorado.

Também é interessante notar uma diminuição da frequência de Olinda e de Recife a partir da década de 1870, justamente quando atingiram o auge de sua representação na geografia do romance. Na verdade, à medida que estagnava econômica e politicamente, o Nordeste também perdia relevância para a ficcionalização do território nacional. Além disso, desde ao menos a Guerra dos Mascates (1710), Olinda e Recife disputavam a polarização da Província de Pernambuco, em particular, e do Nordeste, em geral. No gráfico abaixo, podemos verificar que essa rivalidade também se manifestou na representação de ambas as cidades no romance nacional.

Gráfico 4 — Menções a Olinda e Recife no século XIX



Eixo vertical: quantidades de romances em que a cidade aparece; eixo horizontal: tempo medido em décadas.  
Fontes: dados da pesquisa

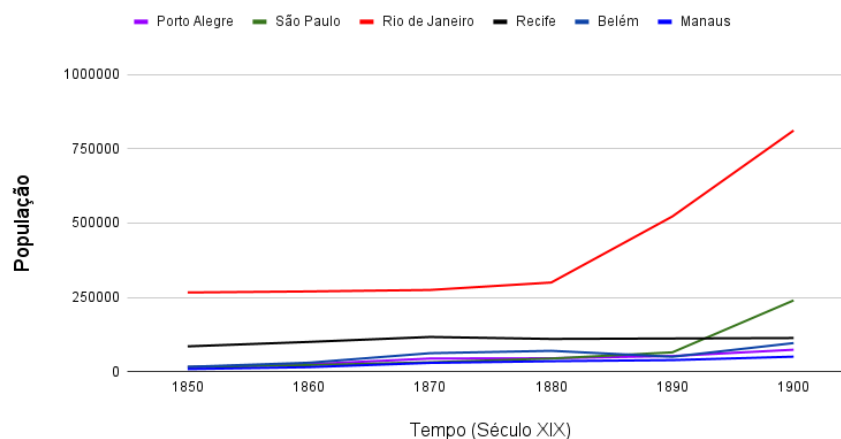
Segundo assinala o gráfico acima, Olinda, a antiga capital de Pernambuco, foi mais mencionada até o fim da década de 1870. A partir dos anos 1880, Recife começa a polarizar mais a região Nordeste. Como a historiografia relata, Recife ultrapassa a economia de Olinda nessa mesma época, conforme se pode ler no excerto abaixo:

Como vimos, Olinda entra em decadência durante longo período, e isso vai se refletir no orgulho e nos sentimentos identitários dos olindenses. Este declínio estende-se durante todo o século XIX (...) Recife é o novo centro do poder. Também foram iniciadas as primeiras ferrovias, em 1858, com a primeira linha para o São Francisco, que favoreceu Recife, por seu papel propulsor do desenvolvimento regional, facilitando as comunicações e escoamento da produção. Os dados do primeiro censo oficial,<sup>37</sup> em 1872, registraram 116.671 habitantes em Recife, que já se transformava em núcleo de influência de toda a região, englobando, além de Pernambuco, a Paraíba, o Rio Grande do Norte e Alagoas. (VASCONCELOS DO NASCIMENTO, 2008, p. 200)

A força simbólica de uma cidade depende de vários fatores, tais como a economia, política ou demografia de uma localidade. Nesse momento, queremos tratar justamente das relações entre a demografia e a representação das cidades, tal como dispõe o gráfico abaixo, que apresenta justamente a população dos municípios em análise. Para dispor os municípios, optamos por uma lógica geográfica, do mais ao sul ao mais ao norte do Brasil.

Antes de propor interpretações, no entanto, é preciso lembrar que censos populacionais apenas atestam o que ocorreu no período antecedente à coleta dos dados. Por exemplo, um aumento populacional verificado em 1860 ocorreu, na verdade, durante a década de 1850. Assim, há sempre uma diferença temporal entre a realidade demográfica e a constatação (o registro) dessa realidade.

Gráfico 5 — População de alguns municípios ao longo do século XIX



Eixo vertical: população de Porto Alegre, São Paulo, Rio de Janeiro, Recife, Belém e Manaus; eixo horizontal: tempo medido em décadas.  
Fontes: dados da pesquisa

Inicialmente, precisamos observar que a associação entre a demografia e a representatividade das cidades no romance não acarreta uma relação direta de causa e consequência. A dinâmica do Rio de Janeiro exemplifica esse aspecto: a população cresce

significativamente a partir da década de 1880, ao passo que a representatividade da cidade diminui nessa mesma época. Similarmente, apesar de apresentarem população mais ou menos constante ao longo do período, Recife, Porto Alegre, Manaus e Belém têm frequências muito díspares no romance.

A dinâmica de São Paulo, no entanto, parece indicar uma correspondência maior entre a representatividade simbólica e a demografia da cidade. À medida que a população cresce a partir da década de 1880 (o censo capta essa dinâmica na década de 1890), a representação de São Paulo no romance também aumenta. Dessa forma, considerando que dados demográficos não se convertem em resultados convergentes no romance, podemos entender o romance como um lugar fictício no qual mudanças sociais, geográficas e econômicas também se manifestam de alguma maneira.

### 3.3 Seção 2: a presença de temas importantes nos romances

Neste capítulo, abordaremos a presença de alguns temas no romance brasileiro do século XIX. Em síntese, trata-se de quantificar e qualificar as menções a temas<sup>25</sup> importantes de acordo com os romances publicados ao longo desse período. O objetivo foi entender a frequência e a intensidade com as quais algumas temáticas reconhecidamente importantes se manifestaram ao longo do romance oitocentista. Os tópicos eleitos foram: escravidão, abolição, nação, formas de governo (república e monarquia), sertão, indígenas e estratos sociais.

Antes de efetivamente mostrar os resultados obtidos, convém explicar como os gráficos desta seção foram organizados. Em geral, cada temática examinada é composta por dois tipos de gráficos: 1) gráficos que mostram quantos romances citaram e quantos não citaram cada tópico e 2) gráficos que apresentam a quantidade média das menções a cada tópico. Como já se pode perceber, os gráficos têm propósitos diferentes.

Os gráficos quantitativos permitem entender a popularidade ou a frequência de um determinado tópico nos diversos romances analisados; os gráficos qualitativos, a intensidade ou a relevância com a qual um determinado tema é aludido ao longo do tempo. Para os gráficos de tipo quantitativo, é igualmente importante notar que a quantidade de menções ao tema é irrelevante. Em outras palavras, basta que o assunto seja citado uma só vez para que se

---

<sup>25</sup> Vale notar que, para este trabalho, os termos “tópico”, “temática”, “tema” ou “assunto” devem ser entendidos como sinônimos e serão utilizados intercambiavelmente, a menos que alguma especificação seja feita.

considere que o tema foi abordado no romance. Esperamos, ainda, que essas explicações fiquem mais claras no momento em que forem postas à prática nas análises.

A respeito dos gráficos de tipo qualitativo, ainda convém observar que todas as palavras utilizadas para a busca se encontram nas legendas de cada gráfico. Além disso, ressaltamos que esses gráficos indicam a média das menções a respeito de cada tema, porquanto a quantidade de romances publicados varia muito ao longo do tempo, como também já mencionamos. Assim, a média foi usada não só como uma forma de atenuar essa variação, como também de facilitar a comparação entre as décadas examinadas. Por fim, o tempo é medido por décadas, de modo que a marcação 1850 no eixo horizontal significa, na verdade, toda a década de 1850.

Também é importante notar que, como não se analisou nenhum romance anterior a 1840, o valor “zero” foi atribuído a essa década em todos os gráficos de intensidade. Dessa forma, podemos comparar a importância de cada temática de maneira relativa. Em outras palavras: como todos os temas começam com a mesma frequência (zero), é possível contrastar a magnitude de cada assunto mais objetivamente. Além de apresentar os dados obtidos para cada assunto analisado, o nosso objetivo foi interpretar e analisar as razões para essa ou aquela frequência.

### **3.3.1 Escravidão**

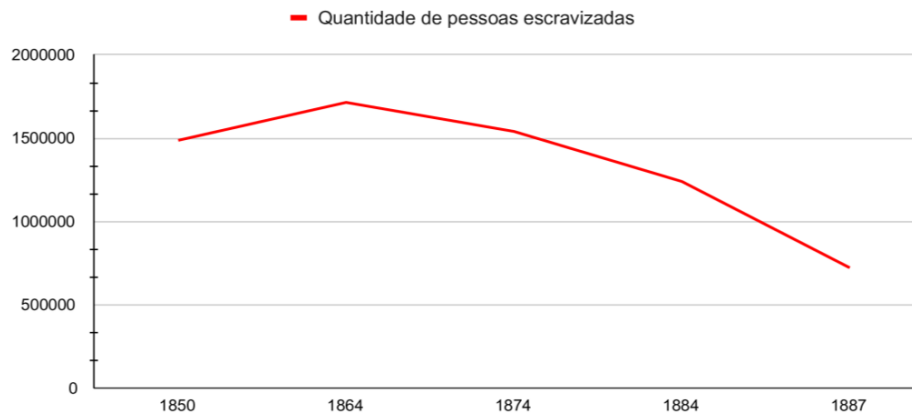
Escravidão é o primeiro tópico que examinaremos. Antes de mostrar os resultados obtidos, no entanto, apresentamos um gráfico com a quantidade de pessoas escravizadas no Brasil entre 1850 e 1887. Esses dados e, conseqüentemente, essas datas, foram escolhidos por se tratarem de informações oficiais disponibilizadas pelo Instituto Brasileiro de Geografia e Estatística (IBGE). Embora não coincidam perfeitamente com o período em análise, o intervalo temporal investigado pelo IBGE satisfaz os nossos objetivos neste momento. Além disso, é importante observar que não pretendemos discutir a exatidão desse tipo de dado (demográfico, histórico, econômico, por exemplo), uma vez que o consideramos como uma representação possível, ainda que imperfeita, da realidade.<sup>26</sup>

---

<sup>26</sup> Apesar de reconhecer que não existe neutralidade na linguagem, defendemos que dados históricos, geográficos e econômicos não são marcados pelos aspectos estéticos que caracterizam a literatura, por exemplo. Por essa razão e por não se tratar de um trabalho de revisão historiográfica, consideramos esses dados como uma representação, ainda que imperfeita, da realidade.



Gráfico 6 — População escravizada no Brasil ao longo do século XIX

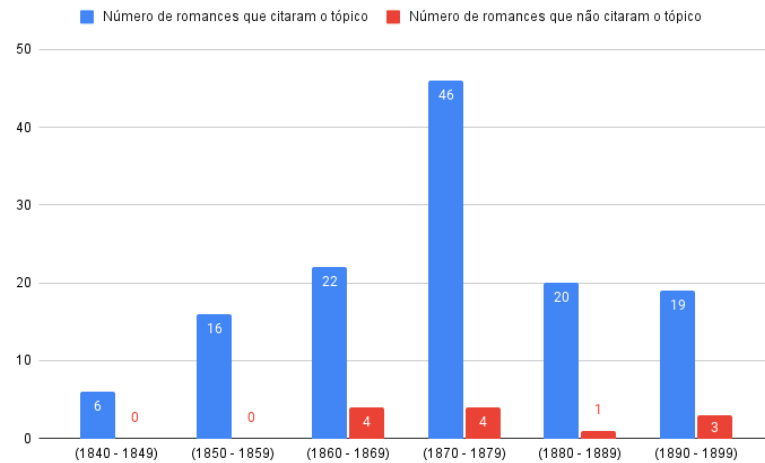


Fontes: REIS, J. J. *Presença Negra: conflitos e encontros*. In: INSTITUTO BRASILEIRO DE GEOGRAFIA E ESTATÍSTICA. *Brasil: 500 anos de povoamento* Rio de Janeiro, 2000. p. 91

O gráfico 6 indica que a população escravizada começou a decrescer a partir da segunda metade da década de 1860 no Brasil. Com efeito, ao menos duas mudanças na legislação brasileira possibilitaram essa diminuição. A primeira foi a Lei Eusébio de Queiroz que, editada em 1850, dificultou sobremaneira o tráfico de pessoas para o Brasil. Em síntese, essa lei não só equiparou o tráfico à pirataria, como também dispôs que cabia à Auditoria da Marina e ao Conselho de Estado julgar essa prática, que já era ilegal desde a Lei Feijo, editada em 1831. A segunda foi o Decreto n 3.3310, editado em 1864, que determinou a emancipação definitiva de todos os africanos livres ilegalmente introduzidos em território brasileiro desde 1831. Além disso, a partir de 1864, seguem várias leis que marcam a abolição gradual da escravidão no Brasil.

Após essa breve retomada histórica, uma primeira análise consiste em observar a frequência com a qual o tópico “escravidão” foi mencionado nos romances da época. Esses primeiros dados correspondem, portanto, a uma apreciação quantitativa do assunto, conforme a tipologia dos gráficos estabelecida anteriormente. Assim, o gráfico a seguir mostra a quantidade de romances que citaram alguma palavra relacionada à temática em exame. Como o regime servil representava uma das bases sociais, econômicas e políticas do Brasil oitocentista, podemos imaginar que o tópico apareceu em quase todos os romances daquela época, o que pode ser efetivamente confirmado pelos dados a seguir.

Gráfico 7 — Quantidade de romances em que a temática “escravidão” aparece



Palavras pesquisadas: 'escravatura', 'escravaturas', 'escravidão', 'escravo', 'escrava', 'escravos', 'escravas', 'escravinho', 'escravinhos', 'escravinha', 'escravinhas'.

Fonte: dados da pesquisa

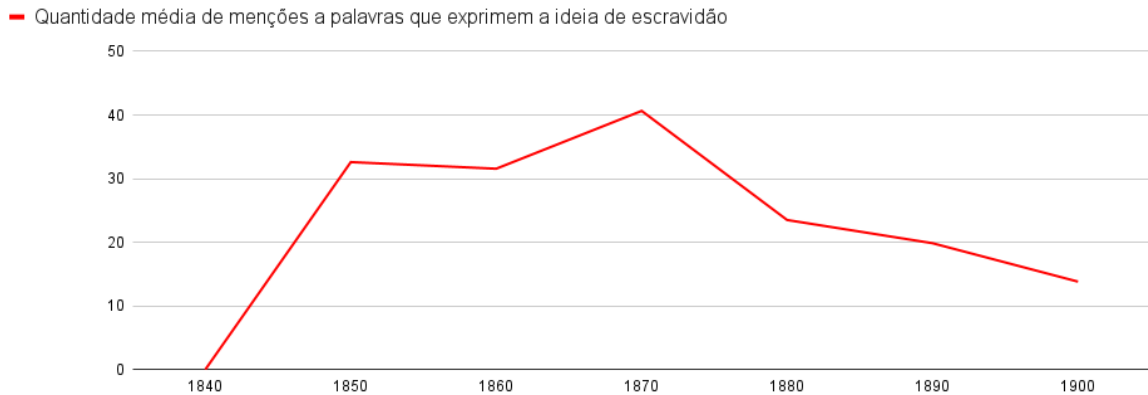
De fato, os dados acima atestam que o tópico “escravidão” era extremamente recorrente no romance brasileiro, porquanto 90% dos romances analisados citaram a temática. Na verdade, trata-se de um dos assuntos mais frequentemente abordados no romance oitocentista. Ao longo de todo o período estudado, a quantidade de romances em que o assunto foi aludido é muito superior à quantidade em que não se verifica nenhuma alusão ao tema. Nesse sentido, podemos dizer que os enredos da maioria dos romances examinados - 90%, mais especificamente - contavam com menções, alusões ou discussões acerca da escravidão.

Também chama a atenção que o gráfico do IBGE e o gráfico acima, respectivamente os gráficos 6 e 7, não sinalizem nenhuma correlação direta ou indireta entre os seus dados. Em outras palavras, embora o número de pessoas escravizadas decresça a partir de 1860, não se verifica nenhum aumento ou diminuição particular na quantidade de romances que mencionam o tópico. Na verdade, parece que a temática “escravidão” é tão onipresente que o tema aparece em quase todos os romances, independentemente da década ou das dinâmicas demográficas da população cativa.

Após ter examinado a recorrência do tema, podemos prosseguir a nossa análise ao comparar o primeiro gráfico (número 6) com o gráfico seguinte, que apresenta a quantidade média de menções à ideia de escravidão no romance. Convém mais uma vez lembrar que a análise dos tópicos não permite determinar se o assunto é tratado positivamente ou negativamente.

Nesse momento, trata-se apenas de compreender a intensidade ou a força com a qual um tema apareceu nos romances ao longo do século, sem juízo de valor.

Gráfico 8 — Quantidade média de menções à ideia de escravidão em romances do século XIX (dados preliminares)



Palavras pesquisadas: 'escravatura', 'escravaturas', 'escravidão', 'escravo', 'escrava', 'escravos', 'escravas', 'escravinho', 'escravinhos', 'escravinha', 'escravinhas'.

Fonte: dados da pesquisa

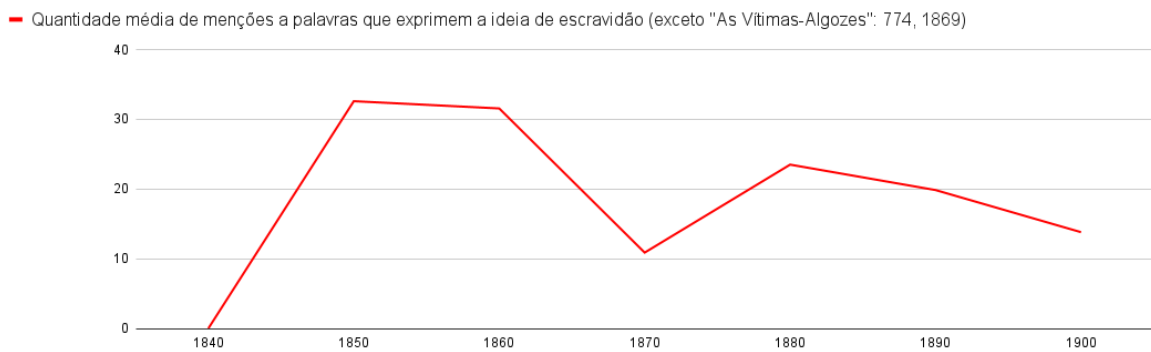
O gráfico 8 indica que a quantidade de menções à escravidão cresceu durante a década de 1860. Em um primeiro momento, poderíamos argumentar que os gráficos 6 e 8 mostram dinâmicas opostas: nessa mesma década, o tópico se fortaleceu à medida que o número de pessoas escravizadas decrescia. Há, no entanto, um aspecto que ainda precisa ser considerado: o romance *As Vítimas Algozes*, que sozinho representa 72% das citações na década em exame. Publicado em 1869, esse romance menciona 774 palavras relacionadas à escravidão, ao passo que nenhum outro romance dessa década contém mais que 55 menções. Trata-se, portanto, de uma anomalia, ou seja: um valor atípico em um padrão mais ou menos uniforme.

Antes de dar prosseguimento às análises, portanto, precisamos tratar de uma questão metodológica importante: dados anômalos. Lidar com anomalias é uma tarefa delicada para o pesquisador de Humanidades Digitais. Com efeito, a manutenção ou a exclusão de um dado desse tipo pode acarretar consequências importantes para as análises de uma pesquisa. Há, certamente, várias maneiras de resolver esse tipo de problema. Como este trabalho também pode servir de referência para futuros pesquisadores, disponho-me a exemplificar duas: a mera exclusão da anomalia e a adequação da anomalia aos demais dados da série.

De início, é preciso reconhecer que a exclusão de uma anomalia implica, também, a perda de um dado para a pesquisa. Quando se trabalha com séries históricas, essa exclusão pode resultar em interpretações diametralmente opostas umas às outras. Ademais, permitir

que um único dado reverta uma tendência, por exemplo, significa valorizar mais a exceção que a regra, o que não parece adequado, notadamente quando se analisam séries e processos históricos, como ocorre neste trabalho. Nesse sentido, a fim de exemplificar as consequências dessa primeira alternativa, as ocorrências encontradas no romance *As Vítimas-Algozes* foram suprimidas. O resultado dessa primeira operação é apresentado no gráfico a seguir.

Gráfico 9 — Quantidade média de menções à ideia de escravidão em romances do século XIX, exceto *As Vítimas-Algozes* (dados preliminares)



Palavras pesquisadas: 'escravatura', 'escravaturas', 'escravidão', 'escravo', 'escrava', 'escravos', 'escravas', 'escravinho', 'escravinhos', 'escravinha', 'escravinhas'.

Fonte: dados da pesquisa

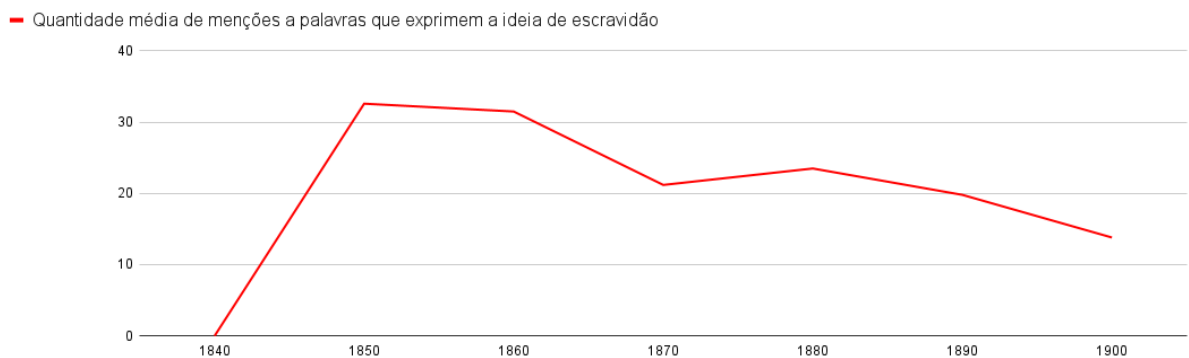
Como se pode ver, a exclusão daquele romance não só muda a configuração do gráfico, como também permite uma outra interpretação a respeito da série histórica. Segundo o gráfico 9, a temática da escravidão teria decrescido durante a década de 1860, crescido a partir de 1870 e praticamente se estabilizado a partir de 1880. Consequentemente, poderíamos afirmar que o tema se enfraquece ao mesmo tempo em que a quantidade de pessoas escravizadas também decresce no Brasil, o que resulta em uma conclusão consideravelmente diferente da anterior. Na verdade, essa operação nos permitiu encontrar um resultado que aproxima ainda mais a trajetória do romance à situação dos escravos no Brasil, tal como relatada pela historiografia.

Por um lado, essa primeira solução assinala que a anomalia dos dados de *As Vítimas Algozes* realmente escondia uma tendência de declínio do tema na década de 1860, o que é essencial para corroborar os dados históricos com os quais estamos trabalhando. Por outro lado, precisamos admitir que a reversão da tendência é muito acentuada, como mostra o formato do intervalo 1860-1870. Ora, apesar de incomum, um dado anômalo ainda é um dado que efetivamente existiu. Como já tínhamos mencionado, uma outra alternativa é transformar

o dado atípico em um dado similar a outros que já tenham aparecido no conjunto que está sendo analisado.

Para esse caso particular, a segunda alternativa consistiu em usar os dados de outro romance como base: *A Escrava Isaura*, que contém 257 menções à temática da escravidão. De fato, à exceção de *As Vítimas-Algozes*, cujos dados são, vale lembrar, anômalos, trata-se do romance com o maior número de ocorrências do tema. Dessa forma, atribuímos 257 menções àquele romance, o que resultou no gráfico a seguir. Para facilitar a comparação, decidimos reproduzir novamente o gráfico com os dados históricos sobre a escravidão.

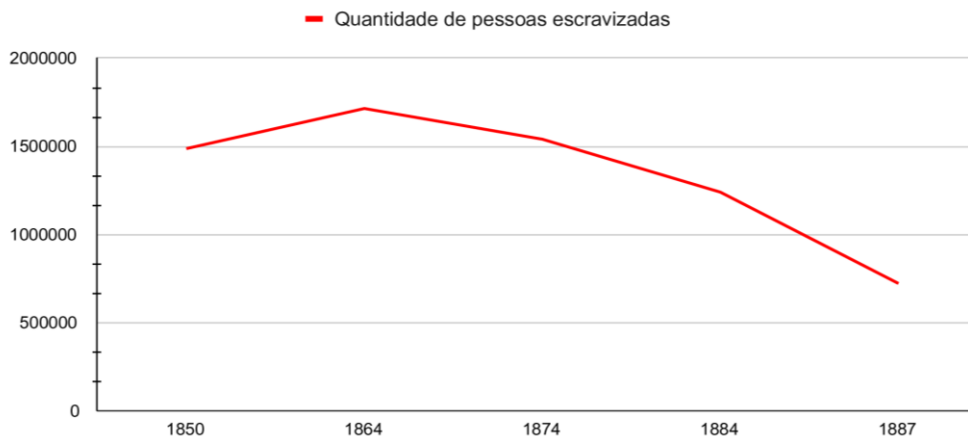
Gráfico 10 — Quantidade média de menções à ideia de escravidão em romances do século XIX, com alterações (dados finais)



Palavras pesquisadas: 'escravatura', 'escravaturas', 'escravidão', 'escravo', 'escrava', 'escravos', 'escravas', 'escravinho', 'escravinhos', 'escravinha', 'escravinhas'.

Fonte: dados da pesquisa

Gráfico 11 — População escravizada no Brasil ao longo do século XIX



Fontes: REIS, J. J. *Presença Negra: conflitos e encontros*. In: INSTITUTO BRASILEIRO DE GEOGRAFIA E ESTATÍSTICA. *Brasil: 500 anos de povoamento Rio de Janeiro*, 2000. p.91

Finalmente, podemos comparar o gráfico da população escravizada (gráfico 6) com o gráfico acima (gráfico 10), o qual apresenta uma solução satisfatória à anomalia encontrada. Podemos constatar que, de 1850 a 1860, a quantidade de pessoas escravizadas cresceu levemente, ao passo que o tema da escravidão permaneceu estável no romance. De 1860 a 1870, observamos uma dinâmica convergente: tanto a população escravizada quanto as menções à temática diminuíram. Por fim, de 1870 a 1890, a tendência é parecida em ambos os gráficos, isto é: enquanto a quantidade de pessoas escravizadas continua a diminuir, as menções ao tema apresentam um leve aumento, ainda que rapidamente seguido de um declínio até o final do século.

Ao comparar os dados desses dois últimos gráficos, podemos afirmar que o tópico “escravidão” foi intensamente aludido ao longo de todo o século XIX. De fato, além de aparecer em quase todos os romances analisados na pesquisa, o tema é abordado abundantemente nos enredos das narrativas. Além disso, é já na década de 1840 que o tópico aparece com mais expressividade no romance nacional. Dessa forma, podemos argumentar que os dados obtidos indicam uma correlação entre o número de pessoas escravizadas e a intensidade com a qual o tema “escravidão” era tratado nos romances. Em síntese, a diminuição do número de escravizados correspondeu a uma menor expressividade do tema nos romances. Apesar de não se tratar de uma relação de causa e efeito, parece existir uma coincidência interessante entre os dois dados.

### **3.3.2 Abolição**

Abolição é o segundo tópico que examinaremos. Embora se trate de um assunto bastante estudado pela historiografia, gostaríamos de ressaltar alguns pontos de maneira introdutória. De início, conforme indicam vários estudos, as alforrias (também chamadas “manumissões”) eram muito mais comuns no sistema escravista brasileiro que no norte-americano ou cubano.<sup>27</sup> Essa particularidade pode ter contribuído para que o instituto da escravidão, notadamente do ponto de vista jurídico, tivesse um estatuto menos objetivo no Brasil. Em outras palavras, a situação jurídica de uma pessoa escravizada era mais facilmente alterável no Brasil do que nos Estados Unidos ou em Cuba, por exemplo. Da mesma forma, é

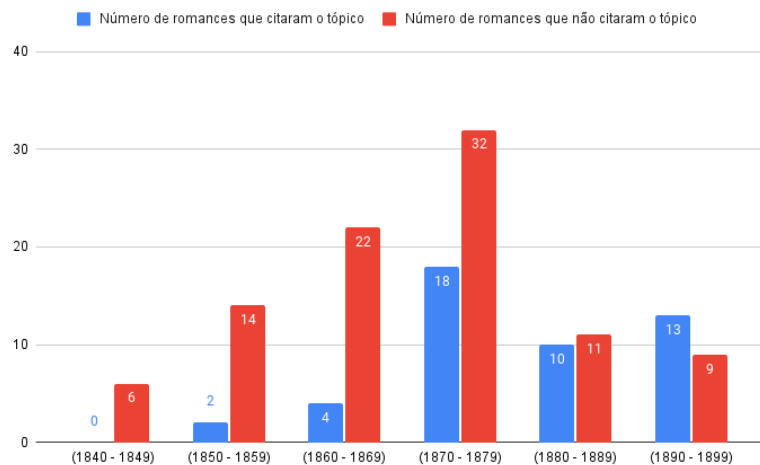
---

<sup>27</sup> Para mais detalhes sobre essa questão, ver o livro *Visões da Liberdade*, de Sidney Chalhoub.

preciso lembrar que o regime servil não só era aceito por quase toda a sociedade brasileira, o que inclui os próprios escravizados, como também compreendia todo o território nacional.<sup>28</sup>

A fim de analisar o tópico “abolição”, decidimos adotar o mesmo procedimento utilizado no tópico anterior. Consequentemente, começamos por apresentar o gráfico abaixo, que mostra a quantidade de romances em que a ideia de “abolição” foi mencionada. Em razão da extensão do instituto da escravidão no Brasil, quer geográfica, quer ideologicamente, poderíamos supor que o tema fosse constar em praticamente todos os romances publicandos na época. Os dados obtidos sugerem uma análise um pouco mais sutil, no entanto, como discutiremos abaixo.

Gráfico 12 — Quantidade de romances em que a temática “abolição” aparece



Palavras pesquisadas: 'abolição', 'aboliconismo', 'aboliconismos', 'aboliconista', 'aboliconistas', 'emancipação', 'manumissão', 'manumissões', 'alforria', 'alforriado', 'alforriada', 'alforriados', 'alforriadas'. Fonte: dados da pesquisa

Inicialmente, chama a atenção que, de 1840 a 1869, quase nenhum romance mencione o tema “abolição”. Com efeito, dos 48 romances publicados entre essas décadas, apenas 6 citam o tópico, o que representa 12,5% das narrativas. Dessa forma, podemos argumentar que o assunto foi muito pouco abordado durante essas décadas, uma vez que não existia nenhum partido político ou entidade da sociedade civil que defendesse, aberta e organizadamente, a abolição da escravidão no Brasil. Nesse sentido, os dados obtidos nos romances correspondem à maneira como os brasileiros, de forma geral, pensavam sobre o tema na primeira metade do século XIX.

<sup>28</sup> Em *Uma nova história da riqueza no Brasil* (2017), Jorge Caldeira corrobora essa ideia ao constatar que cada unidade de produção (famílias, empreendedores, fazendas, trabalhadores) contava, em média, com 5 escravos.

Talvez ainda mais interessante seja notar que, a partir da década de 1870, vários romances começaram a mencionar a temática em exame. Ocorre que, justamente nessa década, o movimento abolicionista ganhou força no Brasil. Segundo José Murilo de Carvalho, “a escravidão estava tão enraizada na sociedade brasileira que não foi colocada seriamente em questão até o final da guerra do Paraguai.”<sup>29</sup> Dessa forma, é possível argumentar que o início do movimento abolicionista brasileiro ajuda a explicar o crescimento de romances que passaram a mencionar o tema, notadamente a partir dos anos 1870. De fato, 44% dos romances publicados entre 1870 e 1900 citaram a temática, o que representa um aumento significativo em comparação ao intervalo 1840-1870. Em outras palavras, o tópico se tornou mais popular e recorrente nos romances, em razão da crescente oposição popular e política a respeito da escravidão.

Como também podemos observar no gráfico 13, as décadas de 1880 e 1890 correspondem a uma ainda maior proporção de romances que aludem à temática “abolição”. Nesse sentido, podemos dizer que o crescimento verificado é resultado das consequências práticas do movimento abolicionista brasileiro. Com efeito, a partir dos anos 1880, os abolicionistas começaram a receber mais apoio da sociedade brasileira, conforme exemplificam a criação da Sociedade Brasileira contra a Escravidão, fundada em 1880, e da Confederação Abolicionista, instituída em 1883. Além da criação dessas instituições, várias revoltas populares irromperam pelo país a partir dos anos 1880:

A partir da década de 1880 (...), o abolicionismo tomava novamente as ruas e os jornais: *Jornal do Commercio*, *A Onda*, *A Abolição*, *Oitenta e Nove*, *A Redenção*, *A Vida Semanária*, *Vila da Redenção*, *A Liberdade*, *O Alliot*, *A Gazeta da Tarde*, *A Terra da Redenção*, *O Amigo do Escravo*, *A luta*, *O Federalista*, bem como dezenas de panfletos e pasquins. Nessa época não era incomum assistir a procissões, participar de rituais, cerimônias emocionais nos teatros da corte ou de manifestações pelo fim da escravidão, que perdia em eficácia e aceitação. (SCHWARCZ, 2015, p. 307).

Em síntese, podemos concluir que as dinâmicas do movimento abolicionista - das primeiras manifestações à abolição da escravatura - ajudam a explicar o aumento de romances que aludem a esse tópico ao longo do século XIX. Para que se tenha uma ideia, 54% dos romances abordaram o tema durante o intervalo 1880-1890. É claro, no entanto, que existem muitos outros fatores que não foram considerados nessa análise, de modo que qualquer conclusão deve ser entendida de maneira preliminar. De toda a forma, é interessante que os dados e as datas coincidam de tal maneira.

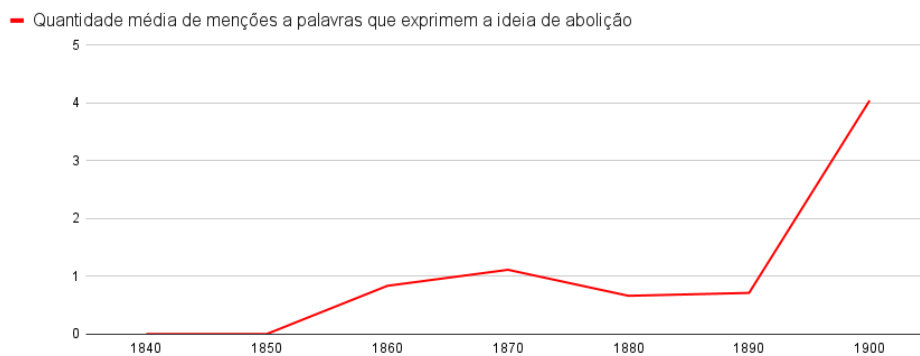
---

<sup>29</sup> CARVALHO, José Murilo de. *Cidadania no Brasil. O longo Caminho*. 3ª ed. Rio de Janeiro: Civilização Brasileira, 2002. Página 45.



Após ter analisado a recorrência do tema nas narrativas, podemos prosseguir com o exame da intensidade com que o tópico foi mencionado ao longo do tempo. Conforme já havíamos discutido, este segundo procedimento diz respeito a uma apreciação qualitativa dos dados. Assim, apresentamos o gráfico abaixo, o qual indica a quantidade média de menções à ideia de abolição nos romances analisados.

Gráfico 13 — Quantidade média de menções à ideia de abolição em romances do século XIX



Palavras pesquisadas: 'abolição', 'aboliconismo', 'aboliconismos', 'aboliconista', 'aboliconistas', 'emancipação', 'manumissão', 'manumissões', 'alforria', 'alforriado', 'alforriada', 'alforriados', 'alforriadas'. Fonte: dados da pesquisa

Uma primeira constatação importante é que o tema foi citado com uma intensidade baixa, mas estável entre as décadas de 1850 e 1880. Na verdade, os dados acima sugerem que, apesar de ter aparecido nos romances com bastante frequência, sobretudo de 1870 a 1900, conforme vimos anteriormente, o tema não tinha muito espaço nos diálogos, encontros<sup>30</sup> e enredos das narrativas. Além disso, é importante notar que estabilidade, nesse caso, significa uma citação por romance em média, o que é muito pouco, sobretudo quando comparado ao assunto imediatamente precedente.

O perfil conservador do público leitor e dos romancistas da época talvez possa explicar a baixa intensidade do tema no intervalo 1850-1880. De fato, é possível argumentar que, como esses indivíduos geralmente pertenciam a camadas confortáveis da sociedade, existia uma forte tendência a evitar esse tipo de assunto nas obras. Na verdade, podemos dizer que se tratava mais de não enfatizar muito a discussão de temas politicamente sensíveis ou impopulares para essas camadas sociais. Em síntese, embora o tema efetivamente aparecesse

<sup>30</sup> Segundo Franco Moretti, “enchimento” é “um novo modo, autenticamente secular, de conceber o sentido da vida: disperso em meio a incontáveis acontecimentos miúdos, precário, entremeado com a indiferença ou os egoísmos mesquinhos do mundo (...)” (MORETTI, 2013, p.82)

nas obras, os romancistas não dedicavam muitas partes do enredo para mencionar o tema “abolição”.

Conforme já ressaltamos, as discussões a respeito da abolição da escravatura começaram na década de 1870 no Brasil. O gráfico 11, no entanto, indica que o tema foi pouco aludido durante esses anos. De fato, a despeito das contestações políticas e civis contra o instituto da escravidão, a temática só se fortalece no romance a partir de 1890, justamente um ano após da Lei Áurea. Dessa forma, queremos argumentar que, como a temática era politicamente sensível, ao menos uma década a mais foi necessária para que a abolição se tornasse um assunto mais expressivo no romance nacional. A ampla aceitação do instituto da escravidão no Brasil pode ajudar a explicar esse descompasso. É novamente Murilo José de Carvalho que afirma:

Significa que os valores da escravidão eram aceitos por quase toda a sociedade. Mesmo os escravos, embora lutassem pela própria liberdade, embora repudiassem sua escravidão, uma vez libertos admitiam escravizar os outros. Que os senhores achassem normal ou necessária a escravidão, pode entender-se. Que libertos o fizessem, é matéria para reflexão. (CARVALHO, 2002, p. 32)

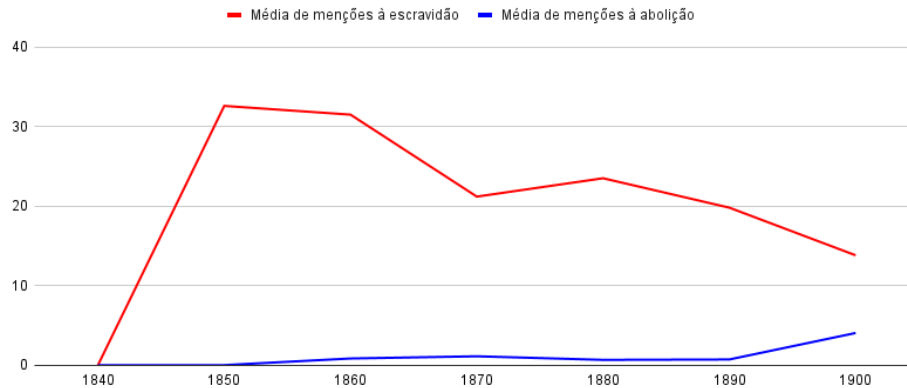
Como grande parte da população brasileira aceitava a existência da escravidão - tácita ou explicitamente - demorou algum tempo para que a ideia de abolição fosse abordada mais detalhadamente no romance. Caso esteja correta, essa argumentação pode explicar não só a baixa intensidade do tema, como também a massiva frequência do tema “escravidão” no romance brasileiro. De fato, o gênero romance não só acolhe, como também é resultado das opções ideológicas de uma sociedade ou de um escritor. A literatura produzida em uma sociedade escravista e desigual deve forçosamente incorporar esses aspectos, quer na escolha de temas, quer na maneira de abordá-los. Ao falar sobre o processo de adaptação do romance no Brasil, o professor Rodrigo do Prado Bittencourt corrobora esse ponto:

*A adaptação não se dá apenas na forma romanesca, mas na ideologia das obras. Mesmo com a presença de semelhanças, há que se pensar que havia muitas diferenças entre as realidades sociais de Portugal ou Brasil e as da França, centro da produção romanesca da época e país hegemônico culturalmente, em relação aos outros dois. Assim, basta ver que um romance que se passa numa cidade grande e desenvolvida, com inúmeras formas diferentes de trabalho e amplo operariado era possível na França ou na Inglaterra, mas não no Brasil. Bem como a presença de escravos nos romances brasileiros não se justificaria num romance francês.”* (BITTENCOURT, 2010, p.1)

Por fim, gostaríamos de mostrar o gráfico abaixo, que permite comparar a intensidade com a qual os tópicos “escravidão” e “abolição” apareceram nas narrativas examinadas. O

aspecto mais interessante é a existência de uma correlação importante: na década de 1890, as menções à escravidão diminuem ao mesmo tempo em que as menções à abolição aumentam.

Gráfico 14 — Comparação entre as médias de menções à escravidão e à abolição (com os dados finais sobre a escravidão)



Palavras pesquisadas: 'escravatura', 'escravaturas', 'escravidão', 'escravo', 'escrava', 'escravos', 'escravas', 'escravinho', 'escravinhos', 'escravinha', 'escravinhas', 'abolição', 'aboliconismo', 'aboliconismos', 'aboliconista', 'aboliconistas', 'emancipação', 'manumissão', 'manumissões', 'alforria', 'alforriado', 'alforriada', 'alforriados', 'alforriadas'.

Fonte: dados da pesquisa

O gráfico 12 indica que o tema da abolição nunca foi tão aludido quanto o da escravidão, a despeito da importância de ambas temáticas no Brasil. De fato, ainda que não tenhamos conseguido explicar completamente as razões disso, os dados sugerem que, efetivamente, se falava muito mais sobre escravidão que sobre abolição no romance brasileiro do século XIX.

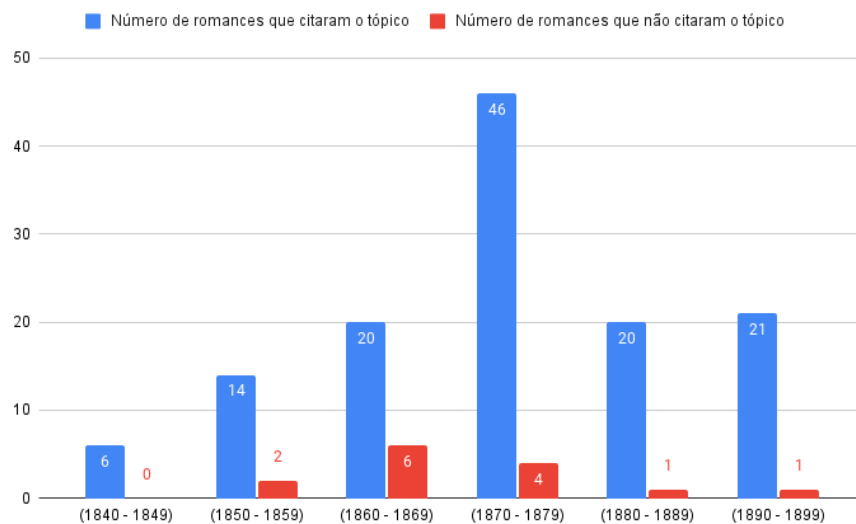
### 3.3.3 Sistemas de governo

Sistemas de governo<sup>31</sup> é o terceiro tópico que analisaremos. Diferentemente das temáticas examinadas anteriormente, no entanto, esse assunto se divide em dois sub-temas historicamente correlatos: 1) monarquia e parlamento e 2) presidencialismo, federalismo e república. Com efeito, discussões acerca de qual sistema de governo adotar também marcaram o Brasil no século XIX, sobretudo durante o período imediatamente posterior à Independência e ao fim da Guerra do Paraguai.

<sup>31</sup> Como o objetivo não é estudar ciência política, as diferenças entre sistema e forma de governo são irrelevantes para este trabalho. Dessa forma, os termos “sistema de governo” e “formas de governo” devem compreender tanto monarquia e república, quanto parlamentarismo e presidencialismo.

Diferentemente do que ocorreu na maior parte dos países do continente americano, a primeira forma de governo adotada no Brasil foi a monarquia, o que já indica a importância do tema para a história institucional e política do País. Na verdade, segundo José Murilo de Carvalho<sup>32</sup>, a monarquia representou o regime político que permitiu a manutenção da união territorial, política e social do Brasil recém independente. Após essa breve contextualização, podemos examinar a quantidade de romances que mencionaram a temática “monarquia/parlamento” ao longo do tempo. Para esse fim, apresentamos o gráfico a seguir.

Gráfico 15 — Quantidade de romances em que a temática “monarquia/parlamento” aparece



Palavras pesquisadas: 'monarquia', 'monarquias', 'coroa', 'império', 'impérios', 'imperador', 'imperadores', 'imperatriz', 'imperatrizes', 'parlamento', 'parlamentos', 'parlamentarismo', 'parlamentarismos'.

Fonte: dados da pesquisa

O primeiro aspecto interessante é que, de 1840 a 1900, quase todos os romances analisados - 90%, mas especificamente - mencionaram o tópico em exame. Nesse sentido, os dados acima correspondem à realidade histórica verificada na época, uma vez que o regime monárquico e os debates parlamentares faziam parte do cotidiano dos brasileiros, sobretudo das camadas mais confortáveis da sociedade, notadamente a partir dos anos 1850<sup>33</sup>. Dessa forma, podemos argumentar que os romances, por meio de seus enredos, diálogos e tramas, também ecoavam algumas das dinâmicas políticas da época.

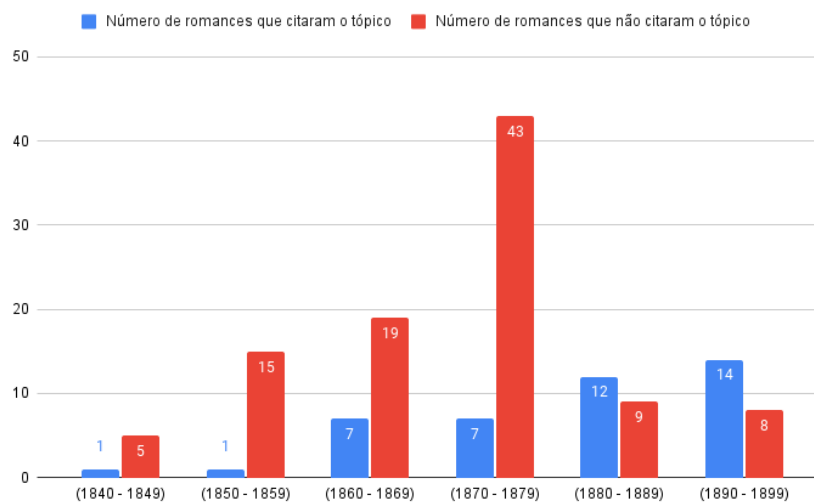
A despeito da força da monarquia brasileira, as interferências do Imperador no funcionamento do Parlamento, notadamente por ocasião do episódio conhecido como Golpe

<sup>32</sup> Para mais detalhes, ver CARVALHO, José Murilo de. Cidadania no Brasil. O longo Caminho.

<sup>33</sup> É a partir de 1850 que o parlamento nacional começa a funcionar mais regularmente no Brasil, uma vez que o Conselho de Ministros foi instituído em 1847.

de 1868<sup>34</sup>, e os altos custos humanos e financeiros da Guerra do Paraguai acarretaram debates acerca do sistema de governo adotado na época. Para que se tenha uma ideia da amplitude dessas discussões, Joaquim Nabuco chegou a discutir a proposta de uma Monarquia Federalista em uma sessão parlamentar em 1883. Dessa forma, também precisamos examinar os dados a respeito do tópico “presidencialismo/federação/república”, uma vez que as contestações havidas na época indicavam a adoção de um governo similar àquele dos Estados Unidos ou da Argentina, por exemplo. Assim como procedemos com o subtópico anterior, apresentamos um gráfico que indica a quantidade de romances que citaram o tema “presidencialismo/federação/república” ao longo do tempo.

Gráfico 16 — Quantidade de romances em que a temática “presidencialismo/federação/república” aparece



Palavras pesquisadas: 'república', 'repúblicas', 'federação', 'federações', 'federalismo', 'federalismos', 'presidencialismo', 'presidencialismos'.

Fonte: dados da pesquisa

Inicialmente, chama a atenção que, da década de 1840 a de 1870, poucos romances - apenas 19%, mais especificamente - mencionem o tópico em análise. Esses dados sugerem que, mais uma vez, os romances acabaram por refratar dinâmicas políticas importantes para a época, na medida em que o regime monárquico ainda se mantinha sólido naquele momento. Em outras palavras, o fato de que opiniões contrárias à monarquia brasileira eram minoritárias até a década de 1870 também pode ser percebido nas narrativas.

<sup>34</sup> Em síntese, o Imperador desrespeitou a maioria parlamentar em 1868. “(...) o imperador chamou ao poder o Partido Conservador quando a Câmara era dominada por progressistas e liberais. A mudança era constitucional, mas no ambiente reformista da época foi acusada de golpe de Estado.” (SCHWARCZ, 2012, p. 106)

De maneira geral, as bases institucionais e os valores que fundamentavam a monarquia brasileira só começam a ser questionadas a partir da década de 1870, após a Guerra do Paraguai. Segundo o historiador José Murilo de Carvalho:

Os últimos 14 anos da monarquia foram marcados pela erradicação da escravidão, pelo surgimento dos militares como ator político e pelo crescimento do movimento republicano. Todos esses fatores contribuíam para a perda de legitimidade da monarquia perante os setores influentes da sociedade. (SCHWARCZ, 2012, p. 117)

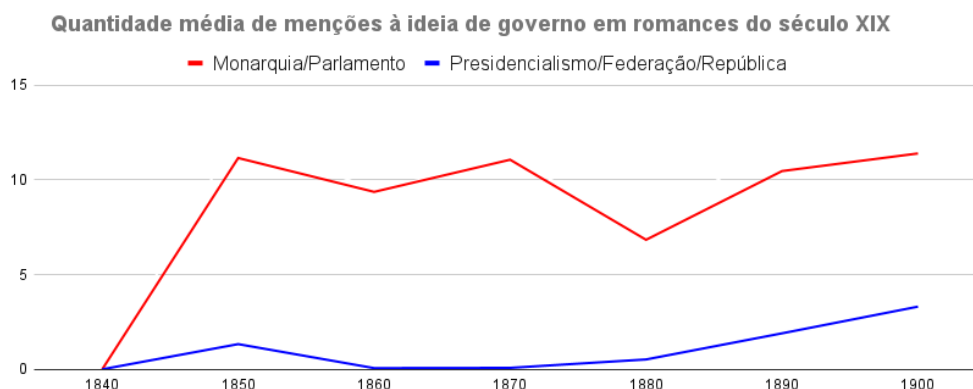
A despeito desse começo de descontentamento com a monarquia, apenas 14% dos romances abordaram o tópico nessa década, o que não parece condizer com a importância do tema para a época. Ocorre que as contestações à monarquia a favor da república só se tornaram oposição política na prática a partir da década de 1880. É justamente nessa década que se verifica um aumento muito significativo no número de romances que fizeram alusão à temática, mais especificamente 60% dos romances publicados entre as décadas de 1880 e 1900.

Ao tratar do período posterior à publicação do Manifesto Republicano (1870) e à criação do Partido Republicano (1873), Ângela Alonso corrobora a nossa interpretação. Segundo a historiadora, embora os intelectuais da Geração de 1870 tenham formulado propostas para contestar as “instituições e os valores saquaremas”, foi somente na década de 1880 que essas ideias se tornaram oposição política na prática:

Três associações procuraram arregimentar a miríade de pequenos grupos pró-reformas sob seu comando ao longo dos anos 1880: o Partido Republicano, a Confederação Abolicionista e a Sociedade Positivista. Cada uma destas estruturas de mobilização privilegiava um ponto nevrálgico da mobilização: a contestação à monarquia, à escravidão, à tradição imperial. (ALONSO, 2005, p.265)

Por fim, após ter discutido os resultados quantitativos, podemos prosseguir com o exame dos dados qualitativos, conforme fizemos com os demais tópicos analisados. Assim, apresentamos o gráfico a seguir, que mostra a quantidade média de menções aos dois tópicos (linhas vermelha e azul) em análise. Em síntese, trata-se de entender a intensidade com a qual as temáticas foram abordadas nos romances.

Gráfico 17 — Quantidade média de menções à ideia de governo em romances do século XIX



Palavras pesquisadas: 'monarquia', 'monarquias', 'coroa', 'império', 'impérios', 'imperador', 'imperadores', 'imperatriz', 'imperatrizes', 'parlamento', 'parlamentos', 'parlamentarismo', 'parlamentarismos', 'república', 'repúblicas', 'federação', 'federações', 'federalismo', 'federalismos', 'presidencialismo', 'presidencialismos'.

Fonte: dados da pesquisa.

O gráfico 14 indica que o tópico “monarquia” foi aludido mais intensamente que a temática “república” durante todo o período analisado. Tal resultado não surpreende, uma vez que se tratava da forma de governo que o Brasil adotou até 1889. De fato, as discussões acerca de presidencialismo, república e federação só crescem a partir de 1880, conforme discutimos anteriormente. É interessante notar, porém, que a intensidade do tópico “república” aumenta consideravelmente a partir dos anos 1880.

Por um lado, os dados assinalam que, além de ter aparecido nos romances com bastante frequência ao longo do século XIX (resultado da análise quantitativa), o tópico “monarquia” tinha altíssima representatividade nas narrativas da época. Com efeito, trata-se de uma das temáticas mais expressivas desta pesquisa. Por outro lado, os dados permitem ponderar que o tema “república” apareceu com bem menos intensidade, embora mostre um aumento expressivo a partir da década de 1880.

Também chama a atenção que, mesmo após a Proclamação da República (1889), menções à monarquia e ao parlamento não sejam incomuns nos romances. Podemos argumentar que, como a adoção do regime republicano não resolveu os problemas institucionais e sociais do País, o confronto político e ideológico a respeito da escolha do sistema de governo continuou a marcar a história do Brasil, ao menos até a década de 1890. O historiador Renato Lessa cunhou o termo “Entropia Republicana” (1889 - 1894) para denominar esse período histórico, o qual é caracterizado por incertezas acerca do futuro da recém instituída república. Segundo ele,

No correr na análise, a ideia de entropia será aplicada ao quadro que resultou da ruptura dos canais tradicionais de integração entre polis, demos e governo, tal como definidos pela ordem imperial. O caráter entrópico da reputar foi agravado pela legislação errática dos anos iniciais da República que, ao tentar superar o vazio institucional provocado pela queda do Império, produziu, tal como será revelado, ainda mais incerteza. (LESSA, 2015, p. 50-1.)

Uma das explicações possíveis para a contínua frequência do tema é a relevância e contemporaneidade do assunto para a elite brasileira. É certo que, para poder comprovar ou refutar essa argumentação, precisaríamos examinar os romances do século XX, com o objetivo de entender se as menções diminuem ou aumentam. Novamente, podemos ver que essa pesquisa pode avançar muito mais, notadamente em uma pesquisa de doutorado.

### **3.3.4 Estratos sociais**

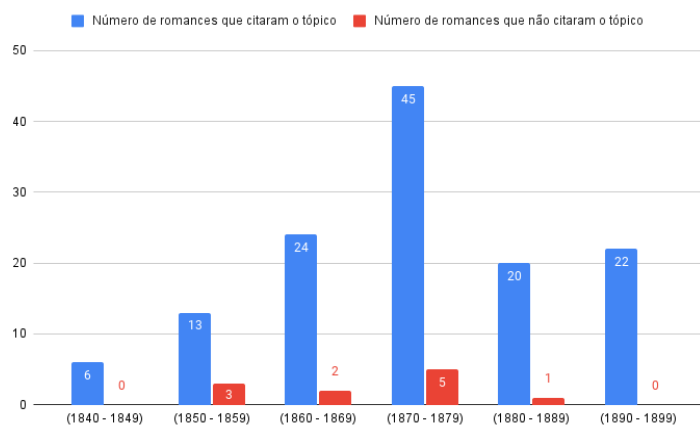
“Estratos sociais” é o quarto tópico que analisaremos. Em síntese, o objetivo é entender como algumas dinâmicas socioeconômicas apareceram nos romances. Apesar de frequentemente utilizados na área das ciências sociais, os termos “burguesia”, “aristocracia” e “proletariado” não foram incluídos na quantificação dos dados do tópico em análise. A razão é simples: os conceitos acima não parecem descrever satisfatoriamente o Brasil oitocentista, uma vez que foram formulados para explicar a realidade europeia daquele século. De fato, embora possam caracterizar alguns segmentos da sociedade brasileira, tais conceitos não levam a categoria “escravo” em consideração, o que impossibilita um entendimento mais holístico das dinâmicas sociais do Brasil oitocentista.

Uma consequência importante em se considerar aqueles termos como inadequados para esta pesquisa consistiu no desenvolvimento de uma outra categoria de análise. A fim de tentar resolver esse problema, elaboramos a categoria “Estratos sociais”, cuja finalidade é quantificar profissões e ofícios mencionados nos romances. De fato, podemos argumentar que profissões e ofícios representam categorias sociais mais facilmente definidas, porquanto vocábulos como “deputado”, “professor” ou “costureira”, por exemplo, têm sentido e realidade mais concreta que “burguês”, “aristocrata” e “proletariado”. Por fim, por razões metodológicas, decidimos dividir a categoria “estratos sociais” em dois subtópicos: 1) “estrato social superior”, que corresponde a profissões como “médico”, “advogado” e “engenheiro” e 2) “estrato social inferior”, o qual contabiliza ofícios como “operário”, “jornaleiro” e “quitandeira”.



Após esses esclarecimentos preliminares, podemos proceder com a apreciação quantitativa do tópico “estrato social superior”, tal como definido acima. O objetivo do é justamente mostrar a frequência com a qual essa temática é aludida nos romances.

Gráfico 18 — Quantidade de romances em que a temática “estrato social superior” aparece



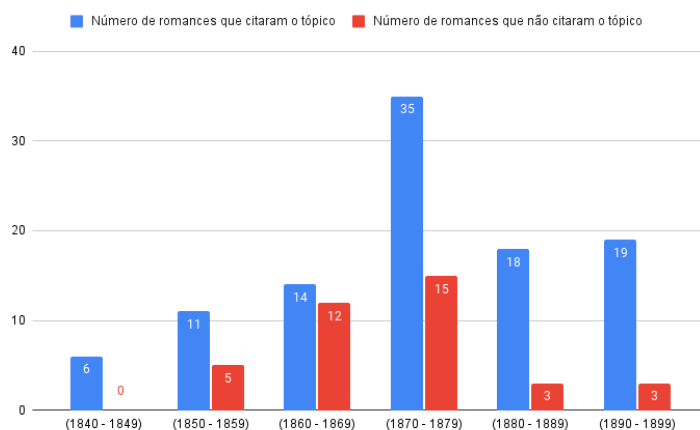
Palavras procuradas: 'elite', 'elites', 'conselheiro', 'conselheiros', 'ministro', 'ministros', 'deputado', 'deputados', 'congressista', 'congressistas', 'médico', 'médicos', 'advogado', 'advogados', 'engenheiro', 'engenheiros', 'professor', 'professores'.

Fonte: dados da pesquisa

O gráfico 22 indica que a maioria dos romances - mais especificamente 92% das narrativas - menciona o tópico ao longo de todo o período. Conseqüentemente, os dados sugerem que profissões como “médico”, “advogado” e “deputado” foram frequentemente aludidas nos romances brasileiros do século XIX. Em outras palavras, significa dizer que o enredo de quase todo romance conta com menções às elites ou às classes mais confortáveis e abastadas da sociedade.

A próxima análise consistirá em entender a dinâmica do segundo subtópico apresentado acima, qual seja: “estrato social inferior”. Apesar de ter sido citado por uma quantidade menor de romances, o tópico agora em exame foi mencionado em aproximadamente 73% dos romances publicados ao longo do século XIX, conforme se pode ver no gráfico a seguir.

Gráfico 19 — Quantidade de romances em que a temática “estrato social inferior” aparece



Palavras procuradas: 'moleque', 'moleques', 'fãmulos', 'fãmulos', 'forro', 'operário', 'operários', 'jornaleiro', 'jornaleiros', 'cozinheiro', 'cozinheira', 'costureira', 'costureiras', 'barbeiro', 'barbeiros', 'quitandeiro', 'quitandeiros', 'quitandeira', 'quitandeiras', 'doceira', 'doceiro', 'doceiras', 'doceiros'.

Fonte: dados da pesquisa

Além da análise qualitativa de ambos os dados, também precisamos examinar a intensidade com a qual os tópicos foram citados nos romances. Nesse sentido, conforme procedemos com as demais temáticas, apresentamos o gráfico a seguir, que mostra a quantidade média de menções a essas temáticas.

Gráfico 20 — Quantidade média de menções à ideia de estratos sociais sociais em romances do século XIX



Palavras procuradas: 'elite', 'elites', 'conselheiro', 'conselheiros', 'ministro', 'ministros', 'deputado', 'deputados', 'congressista', 'congressistas', 'médico', 'médicos', 'advogado', 'advogados', 'engenheiro', 'engenheiros', 'professor', 'professores', 'moleque', 'moleques', 'fãmulos', 'fãmulos', 'forro', 'operário', 'operários', 'jornaleiro', 'jornaleiros', 'cozinheiro', 'cozinheira', 'costureira', 'costureiras', 'barbeiro', 'barbeiros', 'quitandeiro', 'quitandeiros', 'quitandeira', 'quitandeiras', 'doceira', 'doceiro', 'doceiras', 'doceiros'.

Fonte: dados da pesquisa

Inicialmente, podemos constatar que os estratos sociais superiores aparecem com mais expressividade no romance brasileiro, em qualquer década. Em outras palavras, podemos afirmar que os enredos narrativos se passam sobretudo nas camadas mais confortáveis da

sociedade. Dessa forma, trata-se mais da relação entre médicos, deputados e advogados que de quitandeiras, moleques e costureiras, por exemplo. Também é interessante perceber que, apesar das variações, nunca os estratos inferiores superaram a representatividade dos superiores nas narrativas. Nesse sentido, o déficit de representatividade das camadas menos abastadas da sociedade é enorme.

Outro aspecto interessante que podemos observar é um aumento e uma relativa estabilização da presença dos estratos sociais inferiores a partir do fim da década de 1870. Embora a Proclamação da República (1889) não tenha resultado na ampliação da cidadania, é possível argumentar que as reflexões propostas pela Geração de 1870 acarretaram mais e não menos consciência acerca da desigualdade que marca a história do Brasil. De fato, o romance produzido entre 1870 e 1900 é mais crítico a respeito dos problemas sociais brasileiros, ao menos em regra. Ao afirmar que o movimento intelectual de 1870 criticou o “*modus operandi* da sociedade brasileira”, Ângela Alonso afirma que a Geração de 1870

Negou a "desigualdade natural" entre grupos sociais e étnicos como base legítima da hierarquia social e da restrição da política a um estamento - os cidadãos proprietários. (...) Apontou os fundamentos socioeconômicos e as instituições saquaremas como herança colonial e obstáculos ao desenvolvimento do país. As práticas clientelistas alimentadas pela lógica estamental, a patronagem no preenchimento dos cargos públicos, estariam na base da impossibilidade de uma burocracia autônoma em relação aos partidos e aos estamentos senhoriais que foi, por sua vez, vista por todos os grupos como causa da inoperância do aparelho de Estado e do estreitamento da carreira pública. (ALONSO, 2005, p. 244)

Como os romancistas faziam parte desse movimento intelectual, também eles se deram conta das desigualdades do Brasil oitocentista. Assim, os estratos sociais inferiores passam a ganhar mais expressividade para os romancistas dessas décadas. O resultado disso foi um romance mais heterogêneo em termos de representatividade social, justamente a partir de 1870. Com efeito, os dados obtidos corroboram com as teses históricas formuladas por Angela Alonso. Ao tratar da prosa científica e do combate à retórica romântica própria do Segundo Reinado, a já citada historiadora afirma que os romancistas

Criaram novos tipos nacionais, incluindo a gama de alijados na auto-imagem do regime monárquico. Esses romances, minuciosamente descritivos, sociológicos, rompiam com a estetização da sociedade imperial que os romances de Alencar tinham nutrido. Processo que aliás todos os contemporâneos bem perceberam (Machado de Assis, 1879). O desvelamento das "patologias" da sociedade imperial se exprime nos vários ataques que a geração 1870 lançou contra a maior expressão do universo cultural saquarema, José de Alencar. (ALONSO, 2005, p. 242)

Por fim, os dados desse tópico são interessantes na medida em que nos ajudam a comprovar uma tese. Esse é justamente um dos objetivos das Humanidades Digitais, conforme vimos na introdução desta pesquisa. De fato, mediante critérios objetivos de análise e quantificação, podemos argumentar que o romance brasileiro apresenta uma preferência por retratar mais a elite que o povo durante o século XIX.

### 3.3.5 Sertão

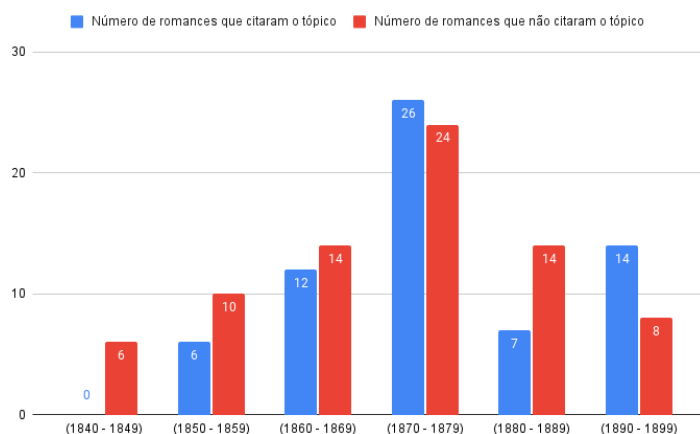
Sertão é o quinto tópico que examinaremos. De início, precisamos ressaltar que a palavra “sertão” deve ser entendida como “lugar geográfico oposto a litoral”. Em *Duas formações, uma história*, o professor Luís Augusto Fischer argumenta que “o sertão está em toda parte, a partir dos limites da *plantation*”. Conseqüentemente, tudo o que estiver distante do litoral pode ser considerado “sertão”. Dessa forma, contrariamente às suas acepções correntes, os termos “sertão” e “sertanejo”, ambas palavras examinadas neste tópico, devem ser entendidos dessa maneira.

Até a segunda metade do século XX, a maioria da população brasileira vivia no meio rural, embora o processo de urbanização já tivesse começado no fim do século XIX, conforme afirmam alguns geógrafos.<sup>35</sup> Na verdade, a demografia do sertão é essencial para se entender as dinâmicas territoriais do Brasil oitocentista. Como a presença do Estado nacional naquelas áreas era fraca, inúmeros eram conflitos fundiários que ocorriam no interior do País. Assim, considerando a importância do tema para a história do Brasil, esperávamos que o assunto fosse recorrente nos romances da época, o que de fato acontece, como revela o gráfico abaixo.

---

<sup>35</sup> Considerando várias opiniões divergentes, Milton Santos estima que menos de 10% da população brasileira vivia em cidades no final do século XIX. Para mais detalhes sobre essa discussão, recomendo o livro *Urbanização brasileira*, de Milton Santos.

Gráfico 21 — Quantidade de romances em que a temática “sertão” aparece



Palavras procuradas: 'sertão', 'sertanejo', 'sertaneja', 'sertanejos', 'sertanejas', 'sertanista', 'sertanistas'.

Fonte: dados da pesquisa

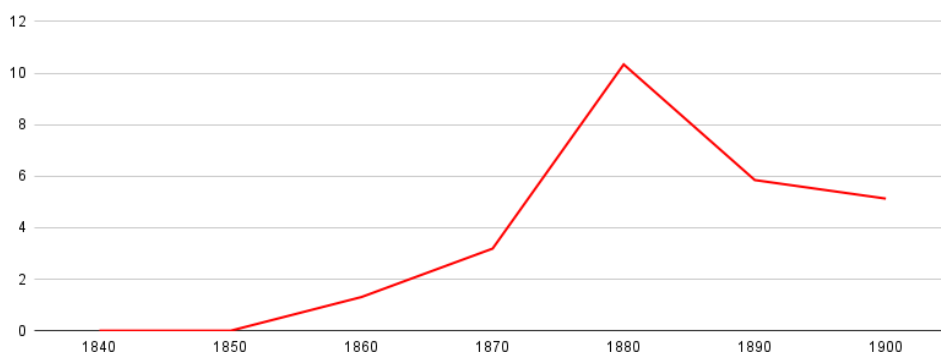
Os dados acima permitem a formulação de algumas análises interessantes. Inicialmente, chama a atenção que aproximadamente 47% dos romances tenham citado o tema de 1840 a 1900. Em outras palavras, a despeito da alta concentração do público leitor nas cidades litorâneas do Brasil do século XIX<sup>36</sup>, o tópico é razoavelmente recorrente ao longo do período. Além disso, podemos observar que a quantidade de romances que citam a temática começou a aumentar a partir da década de 1850. De fato, o contraste entre as décadas de 1840 e 1850 é significativo, porquanto passa de 0% para 37% o número de obras em que o tema aparece. É somente a partir da década de 1860, entretanto, que o tópico adquire certa popularidade no romance brasileiro, como veremos a seguir.

De 1860 a 1900, metade dos romances analisados aludiram ao tópico em discussão. Dessa forma, a despeito das temáticas urbanas, que marcam muitos dos romances brasileiros da época, podemos afirmar que o sertão também é frequentemente mencionado ao longo do século XIX. Em termos percentuais, a década de 1890 representa o auge do tema no romance, porquanto 63% das obras mencionaram o tópico.

Após analisar os dados quantitativos, podemos apreciar os dados quantitativos a respeito desse tópico, conforme viemos procedendo até agora. Recordamos que o objetivo é examinar a intensidade do tópico ao longo do século. Assim, o gráfico a seguir mostra a média de menções ao tema durante todo o período examinado.

<sup>36</sup> Embora os dados sobre o letramento no Brasil do século XIX sejam escassos, parece razoável afirmar que o nível de letramento no litoral fosse maior do que no interior, sobretudo considerando o Censo de 1872.

Gráfico 22 — Quantidade média de menções à ideia de sertão em romances do século XIX



Palavras procuradas: 'sertão', 'sertanejo', 'sertaneja', 'sertanejos', 'sertanejas', 'sertanista', 'sertanistas'.  
Fonte: dados da pesquisa

Embora sua história ainda seja pouco conhecida, pode-se afirmar que o sertão se tornou um tema muito importante e controverso para o Brasil no século XIX. Em síntese, a elite imperial visava a instituir uma nova ordem nessas áreas, as quais não eram controladas pelo Estado nacional, na prática. De fato, o gráfico 21 indica que o tema se torna mais presente justamente quando a política territorial se modifica, notadamente a partir da edição da Lei de Terras, em 1850. Uma das formas de levar a cabo essa nova política territorial consistiu no estabelecimento de colônias militares e presídios no interior do Brasil, conforme explica o professor Carlos Henrique Ferreira Leite:

Entre estas e outras ações, esteve a criação de colônias militares estabelecidas estrategicamente nas regiões fronteiriças e interioranas do território brasileiro (...). O estabelecimento de uma rede de colônias militares e presídios foi, portanto, um dos meios de penetração, de domínio e de povoamento de uma grande extensão de terras no Brasil independente. Com isso, as colônias militares tornavam-se um instrumento da centralização política do império brasileiro [...] (FERREIRA LEITE, 2020, p. 4)

Essa operação deve forçosamente ter chamado a atenção da população brasileira naquela época. Na verdade, a despeito do adjetivo “militares”, essas colônias logo precisaram envolver as populações locais nas atividades para as quais foram designadas, uma vez que o número de oficiais militares era insuficiente para cumprir com os objetivos estabelecidos pelo Império. Com efeito, além dos oficiais de baixa patente, as colônias passaram a ser compostas não só por colonos livres e pobres, como também por escravos e africanos livres. O já citado Carlos Henrique Ferreira Leite afirma:

Atraídos pela possibilidade de se tornarem proprietários de um lote de terra, estes homens enfrentariam as dificuldades pertinentes a uma região de matas densas e desconhecidas sob um regulamento rígido e estranho a sua realidade, além do

desinteresse das autoridades em atender os seus direitos e necessidades, como no caso das concessões dos títulos de posse. (FERREIRA LEITE, 2020, p. 8)

Assim, argumentamos que as menções à temática “sertão” no romance podem ser explicadas pela importância desse projeto centralizador capitaneado pelo Estado nacional e conduzido por alguns setores da população brasileira. Além disso, após a edição da Lei de Terras (1850), marco da nova política territorial do império, a situação fundiária também se complicou para as populações indígenas e tradicionais (gaúchos, caboclos, por exemplo)<sup>37</sup>, as quais não contavam com meios para adquirir terras. O resultado foi mais confronto, agora entre indígenas e fazendeiros, como mostram os já citados Higor Mozart Geraldo Santos e Maria Isabel de Jesus Chrysostomo:

“As falas do Diretor Geral dos Índios em 1875 e do presidente da Província em 1882 manifestavam preocupações com as intermináveis disputas entre os “selvagens e civilizados” pela posse da terra e para implementar projetos de infraestrutura nas áreas de conflito.” (POLÍTICA TERRITORIAL NOS SERTÕES DOS ÍNDIOS, SÉCULO XIX, página 61)

Se compararmos as décadas de 1840 e 1880, perceberemos que os baixos níveis do começo do século não mais se verificam no final da série histórica. De fato, apesar de menos relevante, a temática permaneceu frequente nas décadas seguintes, o que mais uma vez não surpreende, na medida em que os conflitos fundiários continuaram a acontecer durante todo o século XIX. Como a concentração fundiária era muito alta, não é difícil imaginar a frequência e magnitude desses confrontos. A esse respeito, o professor Lucas Guedes Vilas Boas afirma:

No ano de 1850 foi instituída a Lei de Terras, que transformou a terra em mercadoria, privilegiando as elites e oligarquias agrárias do país. (...) Todavia, em um país cujo acesso à terra depende da renda desde 1850, após a promulgação da Lei de Terras, a maioria dos residentes no campo não consegue adquirir seu pedaço de terra. Deste modo, a disputa pela terra engendra uma luta de classes: de um lado os camponeses, e de outro, os fazendeiros. (DE MOURA CUNHA, 2020, p. 38)

Em síntese, a temática “sertão” ganhou relevância a partir da década de 1850, justamente quando a Lei de Terras foi editada e regulamentada. Ademais, como os conflitos fundiários continuaram ao longo do século XIX, o tema se manteve presente no romance, apesar de ter perdido importância a partir da década de 1880. De fato, mesmo no final do século, o controle do sertão preocupa a elite dirigente, sobretudo pela questão indígena, não por acaso assunto do próximo subcapítulo.

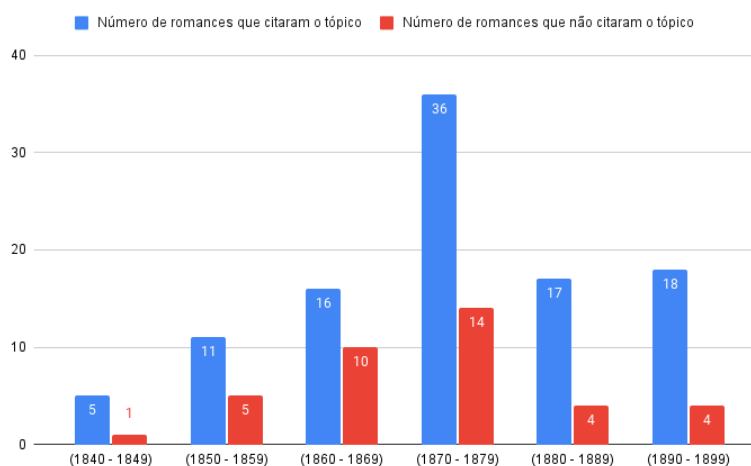
---

<sup>37</sup> Na verdade, todos que não detinham o título de posse de terras precisaram enfrentar essa situação.

### 3.3.6 Índios

Índios é o penúltimo tópico que abordaremos nesta pesquisa. Antes de mais nada, vale recordar que a população indígena não cessou de diminuir no Brasil ao longo do século XIX. Segundo estimativas da FUNAI, a população indígena brasileira, que era composta de 360 mil pessoas em 1825, decresceu a 200 mil em 1940<sup>38</sup>. Dessa forma, pode-se constatar a contradição entre o apreço estético de intelectuais e literatos pelas populações autóctones - o que na literatura brasileira recebeu o nome de “indianismo” - e esse contínuo extermínio das populações originárias. Em síntese, considerando a importância do tema e, mais especificamente, do indianismo para a história e romance brasileiros, esperávamos que o assunto fosse aparecer muito frequentemente. O gráfico abaixo, que permite a apreciação dos dados quantitativos, corrobora essa suposição.

Gráfico 23 — Quantidade de romances em que a temática “índios” aparece



(Palavras procuradas: 'índio', 'índios', 'índia', 'índias', 'indígena', 'indígenas', 'tribo', 'tribos', 'autóctone', 'autóctones', 'aborígene', 'aborígenes', 'indianismos', 'indianismo', 'indigenista', 'indigenistas'. Fonte: dados da pesquisa)

O gráfico 22 indica que aproximadamente 74% dos romances analisados citaram o tópico ao longo do século XIX. Trata-se de um assunto extremamente recorrente nos romances, portanto. Ademais, no “momento áureo” do indianismo brasileiro, de 1840 a 1860<sup>39</sup>, o tema foi citado por aproximadamente 67% das obras, o que também é bastante expressivo. Entretanto, devemos recordar que esse tópico não se refere, somente, a

<sup>38</sup> Os dados estão no link seguinte: <https://www.gov.br/funai/pt-br/atuacao/povos-indigenas/quem-sao>

<sup>39</sup> Para mais detalhes, ver Antonio Candido, *Formação da Literatura brasileira*, página 336.

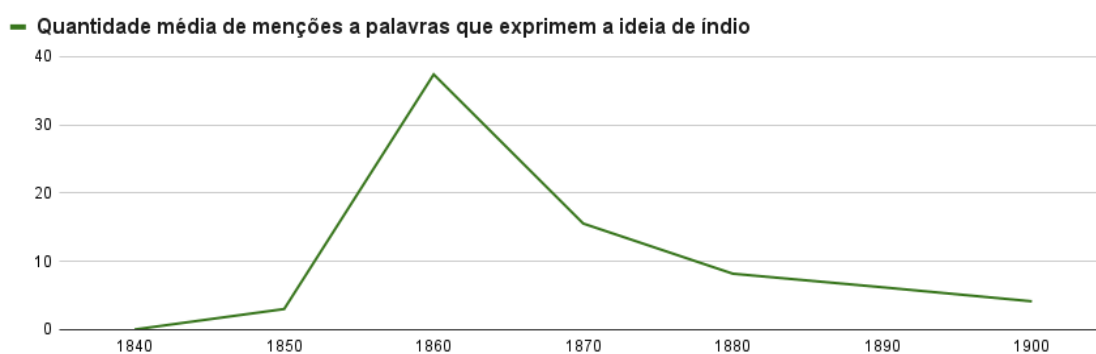


indianismo, mas a índios de maneira mais geral. De toda forma, a conclusão é que o tópico era extremamente popular nos romances brasileiros.

Uma outra análise possível é comprar as temáticas “sertão” e “índios”, uma vez que as populações originárias estavam, geralmente, presentes mais nos sertões que nas cidades. O resultado dessa comparação indica que apenas 48% das obras fizeram alusão ao tópico “sertão” durante todo o século analisado. Conseqüentemente, podemos afirmar que o tópico “índios” era mais popular que a temática “sertão”.

Após apreciar os dados quantitativos, podemos proceder com as análises qualitativas, tal como fizemos com os demais tópicos abordados. Assim, apresentamos o gráfico abaixo, que mostra a quantidade média de menções à temática em análise.

Gráfico 24 — Quantidade média de menções à ideia de índio em romances do século XIX (dados preliminares)

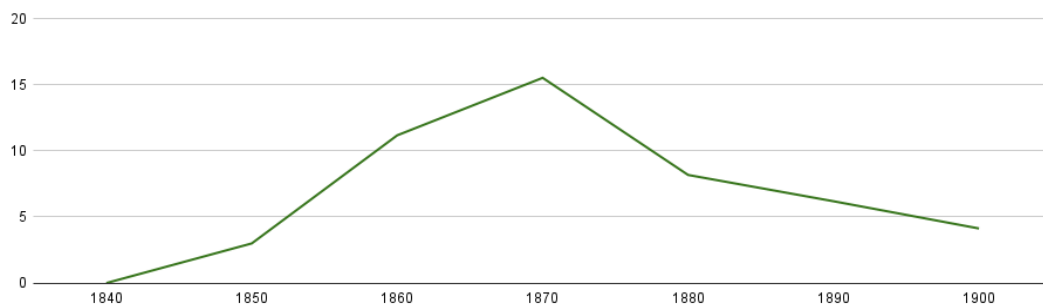


Palavras procuradas: 'índio', 'índios', 'índia', 'índias', 'indígena', 'indígenas', 'tribo', 'tribos', 'autóctone', 'autóctones', 'aborígene', 'aborígenes', 'indianismos', 'indianismo', 'indigenista', 'indigenistas'.

Fonte: dados da pesquisa

Em primeiro lugar, é preciso reconhecer que o tópico “índios” também apresentou uma anomalia. Com efeito, o romance *O Guarani*, individualmente, contém 451 menções à ideia de índios, o que representa 71% das citações da década de 1850. Conseqüentemente, os dados precisavam de alteração similar àquela feita para o tema “escravidão”. Como já havíamos indicado algumas formas de lidar com esse tipo de problema, mencionaremos, brevemente, os procedimentos utilizados para organizar a análise dos dados desse tópico. A primeira operação consistiu em excluir os dados anômalos, o que resultou no gráfico a seguir:

Gráfico 25 — Quantidade média de menções à ideia de índio em romances do século XIX, exceto *O Guarani* (dados preliminares)

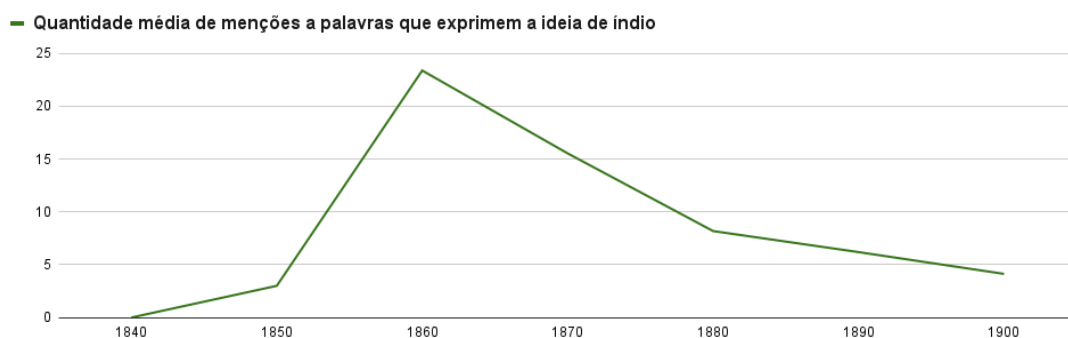


Palavras procuradas: 'índio', 'índios', 'índia', 'índias', 'indígena', 'indígenas', 'tribo', 'tribos', 'autóctone', 'autóctones', 'aborígene', 'aborígenes', 'indianismos', 'indianismo', 'indigenista', 'indigenistas'.

Fonte: dados da pesquisa

A mera exclusão do “romance problemático” não parece ser, no entanto, a melhor solução para lidar com um dado anômalo, conforme já havíamos discutido. Semelhantemente aos procedimentos aplicados no tópico “escravidão”, decidimos adequar os dados de *O Guarani* a uma realidade mais próxima aos demais dados da década de 1850. Vejamos: em toda a série histórica, o máximo de menções foi 195 no romance *Os Índios do Jaguaribe*. Dessa forma, atribuímos 195 citações a *O Guarani*, o que resultou no gráfico a seguir:

Gráfico 26 — Quantidade média de menções à ideia de índio em romances do século XIX, com alterações (dados finais)



Palavras procuradas: 'índio', 'índios', 'índia', 'índias', 'indígena', 'indígenas', 'tribo', 'tribos', 'autóctone', 'autóctones', 'aborígene', 'aborígenes', 'indianismos', 'indianismo', 'indigenista', 'indigenistas'.

Fonte: dados da pesquisa

O gráfico 25 consiste em uma alternativa ao problema encontrado. Assim, podemos continuar as análises. De início, chama a atenção que o tópico “índios” tenha sido citado mais intensamente entre as décadas de 1850 e 1860. A preocupação com o território nacional, notadamente o sertão, está intrinsecamente relacionada com a questão indígena. Dessa forma, não surpreende que o auge do tema coincida com o que os professores Higor Mozart Geraldo

Santos e Maria Isabel de Jesus Chrysostomo chamam “a nova ordem imaginada para o sertão”, a qual é concebida na metade do século XIX. É assim que os professores a resumem:

(...) criação de aldeamentos e destacamentos militares, distribuídos em pontos considerados estratégicos (em geral seguindo o curso dos rios); concessão de terras e títulos honoríficos; criação de ermidas (mais tarde transformadas em freguesias e vilas); instalação de estradas; definição e distribuição de órgãos administrativos voltados para gestão do patrimônio indígena; e, finalmente, mudança do regime de terras. Dentre estas práticas, merece destaque a transformação do patrimônio indígena em terras devolutas. Considerada uma das principais reformas do Estado, ela definiu uma estratégia territorial de importante significado: a apropriação material e simbólica das terras, legitimando a figura do Estado territorial. (MOZART GERALDO, 2016, p. 52)

Como se pode ver, essa “nova ordem” compreendeu mudanças significativas a respeito da política indigenista brasileira. Em síntese, podemos afirmar que os regulamentos editados pelo Império (em 1845 e 1854) acabaram por impedir que os povos autóctones adquirissem terras, ao menos na prática. Embora o discurso oficial defendesse um convívio amigável entre as povos, o resultado para os índios foi catastrófico. Conforme explica a professora Maria Regina Celestino Almeida, tratou-se de uma política contraditória, no mínimo:

Apesar das divergências, predominava a proposta de incorporar os índios ao Império como cidadãos civilizados para servir ao novo Estado na condição de trabalhadores eficientes. “O Regulamento de 1845 decretara o direito dos índios à terra nas aldeias, considerando, no entanto, a possibilidade de extingui-las conforme seu estado de decadência, e o regulamento de 1854 estabeleceu para os índios o usufruto temporário das terras, até que atingissem o ‘estado de civilização’, quando o governo imperial poderia incluí-los no pleno gozo dos direitos de todos os cidadãos. (ALMEIDA, 2013, p. 43)

A ambiguidade consistia em reconhecer o direito à terra aos indígenas, mas aceitar que o governo extinguisse esse direito em algumas circunstâncias. Em um primeiro momento, portanto, as aldeias indígenas, efetivamente, representaram a possibilidade de aquisição de terras. Ocorre que, a partir da década de 1860, essas aldeias começam a ser paulatinamente dissolvidas pelas autoridades imperiais<sup>40</sup>, uma vez que não foi possível compatibilizar os interesses do governo, de um lado, e o dos índios, do outro. Interessantemente, também é nessa década que as menções aos indígenas começam a diminuir no romance, conforme se pode ver no gráfico 25. Segundo a professora Patrícia Melo:

---

<sup>40</sup> POLÍTICA INDIGENISTA NO SÉCULO XIX: CONSEQUÊNCIAS ATUAIS, de Maria José Barboza

Ao fazer um balanço do processo de extinção dos aldeamentos no século XIX, Almeida sublinhou que a insistência nesse discurso da decadência, recorrente nas falas oficiais, demonstrava a forte preocupação do Estado em dar continuidade a uma política indigenista assimilacionista e o esforço para demonstrar o decaimento das aldeias fazia parte desse processo, garantindo, assim, as condições objetivas de sua execução. O desaparecimento dos “índios”, no limite, garantia a liberação legal de suas terras. (MELO, 2022, p. 35)

Ainda, contrariamente ao assunto anterior, a temática “índios” quase desapareceu no final do século XIX. Na verdade, a década de 1890 significa, praticamente, um retorno aos níveis dos anos 1840. Embora os confrontos tenham continuado, o tema parece ter perdido relevância, na medida em que a estratégia levada a cabo pela elite imperial funcionou: as terras permaneceram na posse dos fazendeiros, e os índios ficaram ainda mais marginalizados na sociedade. De fato, o discurso conciliatório que fundamentou a Lei de Terras praticamente desapareceu ao final do século, conforme resume o professor Manuela Carneiro da Cunha:

Aos poucos, porém, o poder local ganha terreno: a partir de 1875, as Câmaras Municipais passam a poder vender aos foreiros as terras das aldeias extintas, e a poder "usá-las para fundação de vilas, povoações, ou mesmo logradouros públicos" (Decreto 2672 de 20/10/1875). Em 1887, as terras das aldeias extintas reverterem ao domínio das províncias e as Câmaras Municipais passam a poder aforá-las (Lei 3348 de 20/10/1887, art. 8, par. 3; 12/12/1887 e 4/4/1888). (...) Cada passo é uma pequena burla, e o produto final, resultante desses passos mesquinhos, é uma expropriação total. (CARNEIRO, 2012, p.16)

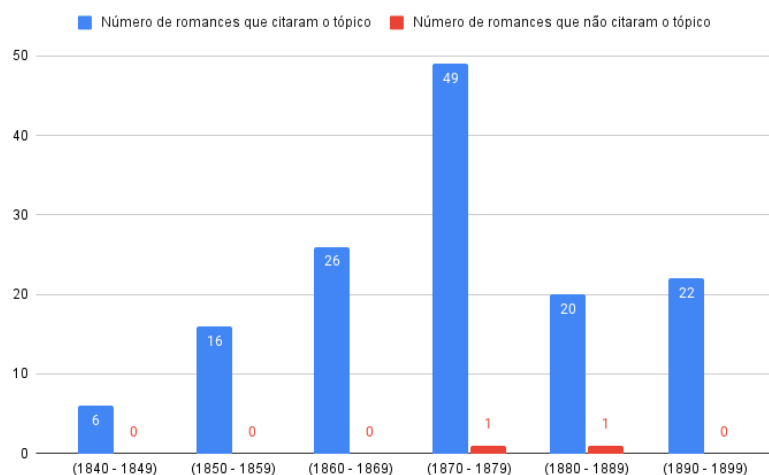
Por fim, a comparação entre os dados quantitativos e qualitativos, gráficos 22 e 25, respectivamente, permite afirmar que o tópico foi citado na maioria dos romances, mas com menor intensidade na década de 1870. Ademais, embora perca expressividade no intervalo 188-1890, o tema continuou a aparecer regularmente na maioria dos romances da época, conforme indica o gráfico 22.

### 3.3.7 Nação

“Nação” é o último tópico que examinaremos na pesquisa. Embora pertençam a um campo semântico parecido, nação e nacionalismo são, indubitavelmente, conceitos diferentes. Senão, vejamos: desenvolvido nos séculos XVIII e XIX, nacionalismo é uma corrente de pensamento político e ideológico segundo a qual um país é dotado de cultura, costumes, língua e território próprios; já a ideia de nação, conceito menos politizado e mais objetivo, diz respeito a um território soberano dotado de um sistema político que governa uma sociedade. Consequentemente, podemos afirmar que o conceito de nação é mais abrangente que o de

nacionalismo. Como ambos os conceitos foram fundamentais para o Brasil oitocentista, consideramos tanto “nação” quanto “nacionalismo” para a quantificação dos dados. O gráfico a seguir permite uma apreciação quantitativa desses dados.

Gráfico 27 — Quantidade de romances em que a temática “nação” aparece



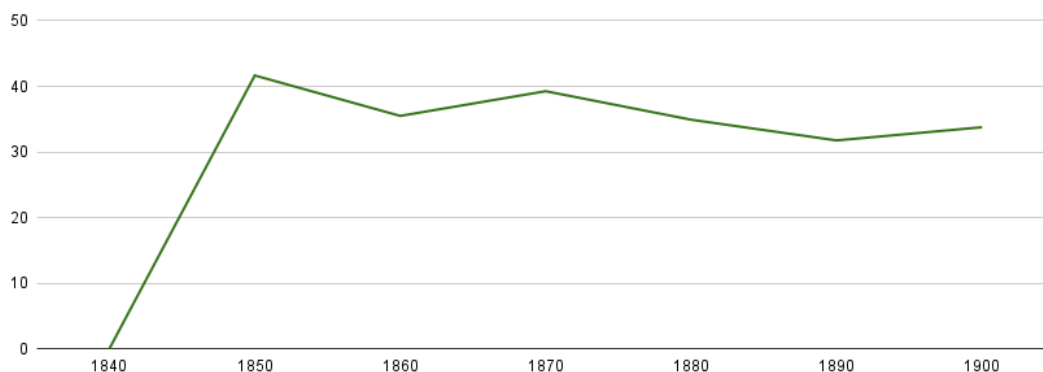
Palavras procuradas: 'nação', 'nações', 'nacional', 'nacionais', 'nacionalismo', 'nacionalismos', 'nacionalidade', 'pátria', 'pátrio', 'pátrios', 'pátrias'.

Fonte: dados da pesquisa

A despeito de sua linearidade, o gráfico 27 é extremamente interessante, na medida em que permite entender como essa temática era popular no romance oitocentista brasileiro. Praticamente todos os romances - 98%, mais especificamente - fizeram alusão ao tópico em análise. De toda a forma, os dados não surpreendem, porquanto a consolidação do Brasil como estado independente coincidiu com a gênese do romance no País. Mais uma vez, convém ressaltar que esta pesquisa pode continuar em um doutorado, na medida em que a análise dos romances publicados no século XX poderia ensejar comparações interessantes também.

Nesse momento, podemos continuar as análises com a apreciação dos dados quantitativos acerca do tópico. A despeito de também se tratar de um gráfico extremamente linear, outras interpretações podem ser apresentadas. Vale recordar que o objetivo passa a ser o estudo da intensidade com a qual o tema foi mencionado ao longo do tempo.

Gráfico 28 — Quantidade média de menções à ideia de nação em romances do século XIX



Palavras procuradas: 'nação', 'nações', 'nacional', 'nacionais', 'nacionalismo', 'nacionalismos', 'nacionalidade', 'pátria', 'pátrio', 'pátrios', 'pátrias'.

Fonte: dados da pesquisa

Em síntese, o gráfico 26 assinala uma trajetória de ascensão e continuidade: crescimento de 1840 a 1850, seguido de uma estabilização em altos níveis dessa década em diante. De fato, diferentemente de outros tópicos examinados neste trabalho, a temática “nação” apresenta um crescimento significativo já a partir da primeira década examinada. Na verdade, tal dinâmica não surpreende, na medida em que a gênese do romance brasileiro ocorreu nessa mesma década, justamente quando o Estado nacional tentava consolidar sua independência e manter a unidade do País.

A formação da identidade nacional foi muito importante para a consolidação do Brasil como País independente. Com efeito, embora tenha sido conduzido com sucesso, o processo de independência não assegurou a unidade do território nacional. De 1830 a 1850, várias foram as revoltas internas que puseram em perigo essa unidade. Para que se tenha uma ideia dos desafios dessa época, José Murilo de Carvalho chama esse período de “Brasil em perigo”<sup>41</sup>. Embora a repressão militar tenha sido essencial para pôr fim às revoltas, outros aspectos foram importantes para deter esses movimentos centrífugos ou separatistas. Do ponto de vista institucional, o Instituto Histórico e Geográfico brasileiro em 1838 foi o mais relevante<sup>42</sup>; do ponto de vista estético, foi o romance, como veremos a seguir.

Conforme discutimos na introdução deste trabalho, o gênero romance também contribuiu para que as pessoas se adaptassem à modernidade. De fato, por ser o gênero literário de maior circulação na Europa do século XIX, essa forma estética acabou por ajudar

<sup>41</sup> Página 90, Brasil Nação, Vol II

<sup>42</sup> Essa instituição deveria estudar “não só o passado da nova nação, mas sua geografia, seus componentes étnicos, sua fisionomia contemporânea. Um programa nacionalista que não se esgotasse, por’pem, na expressão desse ideal e explorasse cientificamente nosso espaço físico e cultural.” (SCHWARCZ, 2012, p.233)

as pessoas a se habituarem às novidades da época, tais como o dinheiro em papel, o casamento por amor e a ideia de nação. Segundo Franco Moretti, a forma simbólica mais apropriada à ideia de nação é justamente o romance.

And again: village, court, city, valley, universe can all be visually represented - in paintings, for instance: but the nation-state? Well, the nation-state... found the novel. And viceversa: the novel found the nation-state. And being the only symbolic form that could represent it, it became an essential component of our modern culture.

(...)

How much of a problem the nation-state initially was: an unexpected coercion, quite unlike previous power relations; a wider, more abstract, more enigmatic dominion - that needed a new symbolic form in order to be understood. And here, Austen's novelistic geography shows all its intelligence. (MORETTI, 1999, p. 17)

Similarmente ao que ocorreu na Europa, queremos argumentar que a ficção (o romance, notadamente) também contribuiu para a consolidação de uma ideia de nação brasileira, sobretudo a partir da década de 1850. De fato, os dados do tópico “nação” podem ajudar a comprovar essa ideia de maneira objetiva. Ora, muitos teóricos da literatura afirmaram que o romance dessa época tem uma vocação quase pedagógica a respeito da ideia de nação. Segundo Antonio Candido, por exemplo:

Por isso mesmo, o nosso romance tem fome de espaço e uma ânsia topográfica de apalpar todo o país. (...) Assim, o que vai se formando e permanecendo na imaginação do leitor é um Brasil colorido e multiforme, que a criação artística sobrepõe à realidade geográfica e social. Esta vocação ecológica se manifesta por uma conquista progressiva de território. (...) Literatura extensiva, como se vê, esgotando regiões literárias e deixando pouca terra para os sucessores, num romance descritivo e de costumes como é o nosso. (CANDIDO, 2000, p. 433)

Desde o período posterior à Independência, intelectuais, políticos e escritores discutiam os significados de brasilidade ou de “nação brasileira”. Conforme indica o gráfico 26, esse assunto também foi abordado no romance. Ademais, embora não tenhamos dados quantitativos muito precisos sobre a circulação do romance no território nacional, a historiadora Marlyse Meyer afirma que o folhetim também teve muito sucesso no Brasil:

(...) O folhetim, portanto, instala-se no jornal e espalha-se em volume barato pelas bibliotecas, onde, já o dissemos, é espantosa sua ocorrência. Muito embora o estudo de tiragem e público da imprensa brasileira ainda esteja por ser feito, o simples exame das modificações havidas no jornal leva a crer que, como na França, sua prosperidade esteve ligada diretamente ao sucesso e, portanto, à publicação do folhetim. E tal sucesso mostra igualmente, guardadas as proporções, a existência no Brasil de um público consumidor de novelas já suficiente para constituir em elemento favorável de venda de jornal (MEYER, 1996, p.59)

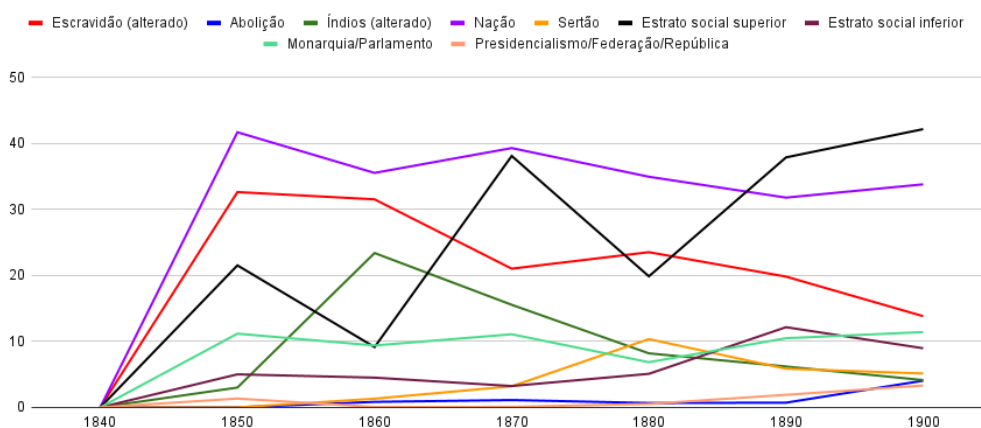
Dessa forma, podemos afirmar que os dados obtidos com o tópico “nação” corroboram a afirmação de que o romance oitocentista também contribuiu para a consolidação da ideia de nação no Brasil.

### 3.3.8 Síntese

O objetivo deste subcapítulo é propor algumas sínteses a respeito de todos os tópicos analisados ao longo desta pesquisa. Assim, foram desenvolvidos novos gráficos, nos quais se mostram as várias temáticas de diferentes maneiras, com base em diferentes configurações. Em síntese, os dados serão rearranjados e reorganizados, de modo a realçar um ou outro aspecto. Por fim, trata-se de uma comparação qualitativa desses conjuntos de dados.

O primeiro gráfico apresenta a quantidade média de menções a todos os assuntos examinados. O objetivo é visualizar a intensidade com a qual todos os temas apareceram ao longo do século XIX. Para a temática “escravidão” e “índios”, foram utilizados os dados obtidos após a resolução das anomalias, conforme se pode ver no gráfico a seguir.

Gráfico 29 — Médias de todos os tópicos analisados, com as devidas alterações (escravidão e índios)



Fonte: dados da pesquisa

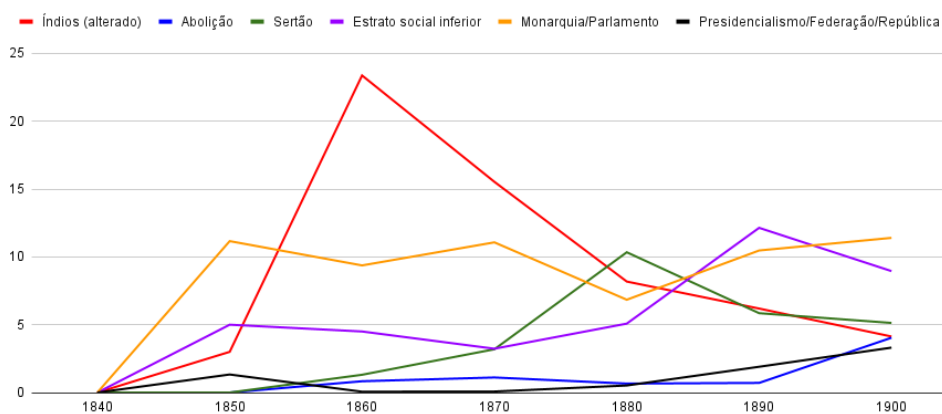
Esses dados podem ser divididos em dois grandes grupos: o das temáticas muito expressivas (“nação”, “escravidão” e “estrato social superior”) e o das temáticas pouco expressivas (“sertão”, “sistemas de governo”, “abolição” e “estrato social inferior”). A exceção é notadamente a temática “índios”, tópico que ganha muita relevância durante um curto período de tempo, o intervalo 1850-1860, notadamente. Com base nessa divisão, podemos propor algumas interpretações.



Por um lado, podemos argumentar que as temáticas “nação”, “escravidão” e “estrato social superior” caracterizam o romance brasileiro, na medida em que são apresentadas não só explícita como também intensamente nas obras. Assim, constatamos a existência de uma consonância entre a importância desses temas para a história do Brasil e a sua representação, explícita, no romance nacional. Por outro lado, podemos considerar que “sertão”, “sistemas de governo”, “abolição” e “estrato social inferior” configuram as temáticas menos relevantes da época, já que não aparecem com tanta intensidade nas obras. Entretanto, a menor expressividade desses temas também lhes confere importância, na medida em que a omissão dos tópicos pode sugerir um apagamento deliberado por parte dos romancistas. Em outras palavras, a consonância é, nesse caso, implícita, mas não menos importante.

Uma outra configuração possível é a exclusão dos tópicos “nação”, “escravidão” e “estrato social superior” desses dados. Apesar dessas supressões, no entanto, podemos constatar que “sistemas de governo”, “abolição” e “sertão” permanecem temas pouco expressivos ao longo da série histórica, conforme indica o gráfico abaixo.

Gráfico 30 — Médias de alguns tópicos analisados, com as devidas alterações (índios)

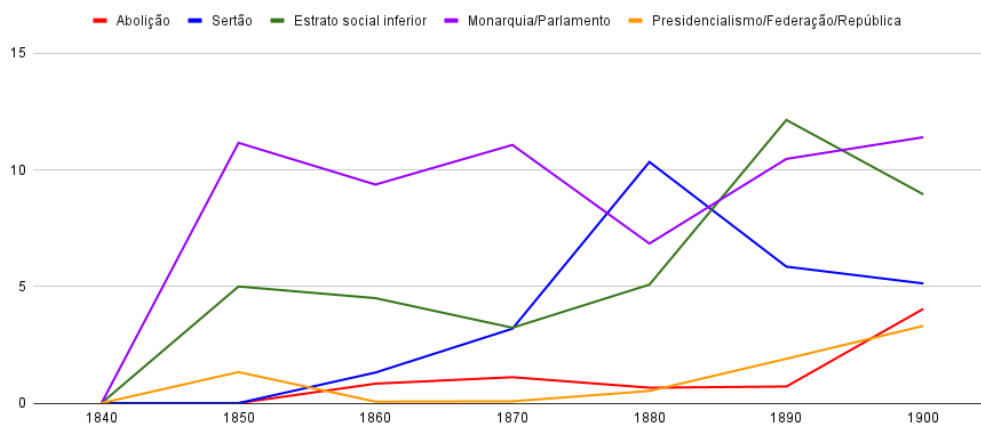


Fonte: dados da pesquisa

De fato, nenhum desses seis tópicos alcançou a quantidade média de 14 menções por década, o que efetivamente ratifica a pouca expressividade desses tópicos para os romances. À guisa de comparação, basta recordar que algumas temáticas alcançaram a média de 30 ou 40 menções, como exemplificam “escravidão” e “nação”, respectivamente. Tal conclusão não significa afirmar que esses assuntos não sejam importantes para a história do Brasil, conforme já ponderamos.

Ainda, uma outra configuração possível consiste em excluir o tópico “índios”, conforme se pode ver no gráfico abaixo. A razão é relativamente simples: por apresentar um pico de frequência no intervalo 1850-1860, esse tópico pode estar efetivamente dificultando a análise do gráfico 29.

Gráfico 31 — Médias de alguns dos tópicos analisados

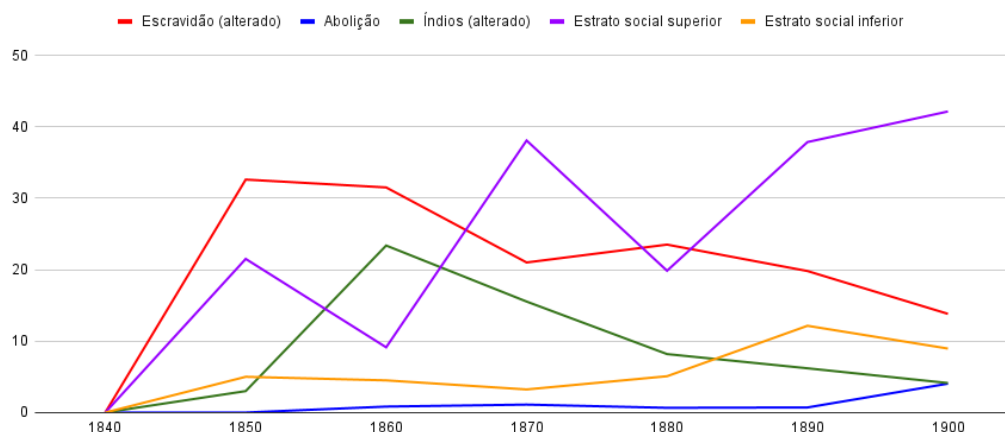


Fonte: dados da pesquisa

Considerando apenas as temáticas apresentadas no gráfico acima, podemos afirmar que “abolição” é, sem dúvida, o assunto menos frequente em todo o romance oitocentista. Similarmente, podemos notar que o assunto “monarquia/parlamento” é relativamente recorrente ao longo de todo o século, já que apresenta trajetória constante, sobretudo a partir da década de 1850. Além disso, outra trajetória que impressiona é a do “estrato social inferior”, uma vez que se trata do assunto mais importante no final da década de 1880.

Existem, certamente, diversas outras maneiras de agrupar esses dados. O último rearranjo que propomos consiste em separar os dados entre 1) temas referentes à sociedade e 2) temas referentes ao Estado. Nesse sentido, “Escravidão”, “abolição”, “índios”, “estrato social superior” e “estrato social inferior” compõem o conjunto de temas a respeito da sociedade, ao passo que “nação”, “sertão”, “monarquia/parlamento” e “presidencialismo/federação/república” dizem respeito ao segundo conjunto.

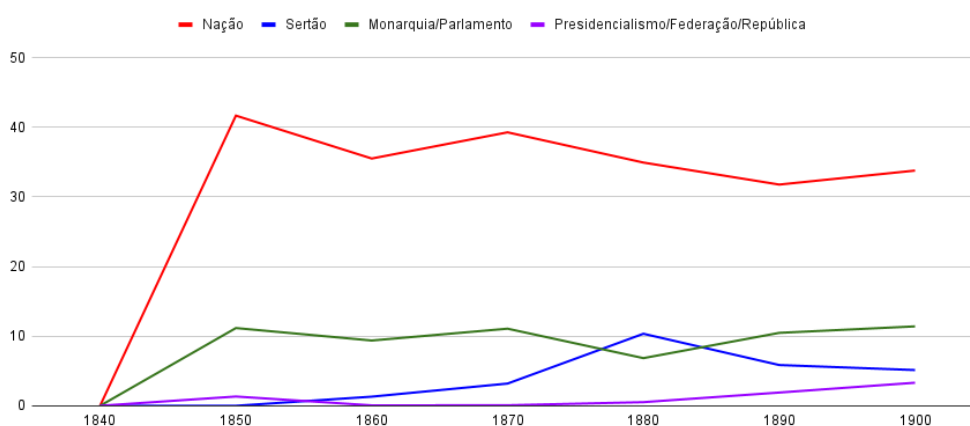
Gráfico 32 — Médias dos tópicos referentes à sociedade



Fonte: dados da pesquisa

O gráfico 31 indica que os assuntos “escravidão” e “estrato social superior” são absolutamente importantes para o romance oitocentista. Também podemos perceber que o assunto “índios”, quando analisado sob essa perspectiva, aparece relativamente bastante ao longo do século examinado. De fato, é apenas a partir da década de 1880 que a temática “estrato social inferior” passa a ser mais recorrente que a dos “índios”, por exemplo. Mais uma vez, podemos constatar a absoluta raridade do tema “abolição” nos romances da época.

Gráfico 33 — Médias dos tópicos referentes ao estado



Fonte: dados da pesquisa

O gráfico 32 assinala a importância do assunto “nação” no romance brasileiro. Com efeito, tal tema se mantém relevante ao longo de todo o intervalo temporal analisado. Além disso, podemos perceber que a temática “monarquia/parlamento” é regularmente mencionada

no romance, contrariamente a “presidencialismo/federação/república”, a despeito da importância do republicanismo brasileiro no final do século XIX.

## 4 CONCLUSÃO

Esta dissertação de mestrado tinha dois objetivos principais: 1) encorajar pesquisadores e pesquisadoras da área da literatura a se interessarem pelas Humanidades Digitais e 2) exemplificar esse método de pesquisa mediante a análise de alguns romances brasileiros publicados no século XIX. Essa ordem não é aleatória, porquanto o primeiro objetivo é mais importante que o segundo. Nesse sentido, como as interpretações formuladas não foram propriamente inéditas, a finalidade principal da pesquisa foi, sobretudo, incentivar o uso de *softwares* e linguagem de programação na área dos estudos literários.

Embora este número venha crescendo nos últimos anos, ainda são poucos os pesquisadores que trabalham com Humanidades Digitais no Brasil. A velocidade e a eficiência representam alguns dos aspectos mais interessantes a respeito desse método de pesquisa. Para esta dissertação, vale lembrar que a utilização da linguagem Python permitiu que 140 romances fossem examinados em cerca de três meses, da feitura dos códigos à análise dos romances. Além disso, recordamos que, quanto mais experiente for o programador, menor ainda tende a ser esse tempo.

Quando comparada a outras disciplinas, a metodologia de pesquisa dos estudos literários se transformou pouco desde a segunda metade do século XX. Em contraste, podem-se constatar grandes transformações na metodologia de pesquisa de várias ciências humanas, tais como a sociologia, a história e a psicanálise, sobretudo nos últimos 15 anos. Em síntese, essas transformações consistiram na aplicação da ciência de dados à disciplina trabalhada. Na área da psicanálise, um exemplo interessante é o livro *Sonhos confinados: O que sonham os brasileiros em tempos de pandemia*, cuja proposta combina análise dos sonhos com ciência de dados.<sup>43</sup>

Ao longo desta dissertação, tentamos exemplificar o uso de códigos e *softwares* mediante algumas análises a respeito do romance brasileiro do século XIX. Mais especificamente, o objetivo foi visualizar a geografia do romance nacional e quantificar a ocorrência de alguns temas importantes para o Brasil do século XIX nesses romances. Para esse fim, confeccionamos não só alguns mapas com a ocorrência dos municípios citados nas narrativas examinadas, como também um banco de dados que quantificou a presença de

---

<sup>43</sup> “(...) sem cair em generalizações apressadas nem na mera coleção de casos que serviriam apenas à ilustração de uma teoria engessada e imune ao novo, fomos obrigados a inventar. Para tanto, fizemos uso de várias ferramentas, incluindo *softwares* de processamento de linguagem natural, que nos ajudaram a detectar recorrências e co-ocorrências de palavras (...)” *Sonhos confinados: O que sonham os brasileiros em tempos de pandemia*, página 39, 2021.

tópicos importantes para aquele século, conforme discutimos exaustivamente em capítulos precedentes.

Por definição, trabalhar com dados significa isolar alguma particularidade da disciplina que se estuda. Na verdade, trata-se de mais de uma necessidade inerente a esse tipo de método que de uma opção à disposição do pesquisador. Duas razões ajudam a explicar essa característica: 1) poucos aspectos de uma determinada disciplina podem ser examinados quantitativamente e 2) tal medição, mesmo quando possível, precisa ser razoavelmente objetiva para ser válida. Dessa forma, é preciso observar que esta dissertação se concentrou em uma pequena parte das inúmeras possibilidades em se trabalhar com os romances e romancistas brasileiros do século XIX, menos por ignorância do pesquisador, que por imposição do método utilizado.

Uma outra vantagem em se trabalhar com Humanidades Digitais é a possibilidade de comprovar teses e ideias com base em dados e informações mais objetivas. Com efeito, os dados obtidos nesta pesquisa podem contribuir à comprovação de alguma generalização a respeito do romance brasileiro do século XIX, por exemplo. Essa opção foi viabilizada, na medida em que se trabalhou com aspectos que podiam ser tanto isolados quanto quantificados. Além disso, os resultados foram apresentados em forma de gráficos e mapas, o que permitiu outro tipo de visualização e análise dos dados encontrados.

A quantificação dos tópicos consistiu na busca de algumas palavras-chave nos romances examinados. Para o tópico “nação”, por exemplo, foram escolhidos os vocábulos “nação”, “nações”, “nacional”, “nacionais”, “nacionalismo”, “nacionalismos”, “nacionalidade”, “pátria”, “pátrio”, “pátrios”, “pátrias”. Ora, reduzir essa temática à ocorrência dessas palavras específicas, estabelecidas de forma mais ou menos arbitrária, é justamente a simplificação de que se tratou anteriormente. Consequentemente, significa afirmar que, se esses mesmos vocábulos não constassem no romance, considerava-se que a temática não tinha sido mencionada ao longo da narrativa.

Essa limitação da ciência de dados não pode ser ignorada, sobretudo quando se trabalha com literatura. De fato, existem inúmeras outras formas de abordar o tópico “nação”, apenas para continuar com o exemplo acima. Basta pensar em metáforas para que se tenha uma ideia de quantas possibilidades podem ter sido ignoradas nesta pesquisa, de maneira não proposital, vale reforçar. Dessa forma, precisamos considerar que a ausência da menção às palavras em exame não deve significar, forçosamente, um silêncio com relação ao assunto analisado.

Por fim, uma última consideração é necessária: não se pretendeu discutir a consciência ou a intencionalidade dos autores nesta pesquisa. Em outras palavras, a citação de uma cidade ou a menção a um tema não deve significar, necessariamente, que o autor ou a autora tinha consciência disso. Em síntese, as análises se concentraram mais em dinâmicas estruturais, as quais podem ser generalizadas para todo um período ou década.

## REFERÊNCIAS

- A GIROLETTI, D. PARTICIPAÇÃO DOS IMIGRANTES ITALIANOS NO DESENVOLVIMENTO DE MINAS GERAIS. *Imigração e Sociedade: Fontes e Acervos da Imigração Italiana no Brasil.*, p. p.328-85, 2015.
- ALMEIDA, M. R. C. Os índios na história do Brasil no século XIX: da invisibilidade ao protagonismo. *Revista História Hoje, [S. l.]*, v. 1, n. 2, p. 21–39, 2013. DOI: 10.20949/rhhj.v1i2.39. Disponível em: <https://rhhj.anpuh.org/RHHJ/article/view/39>. Acesso em: 12 abr. 2023.
- ALONSO, A. *Ideias em movimento.* [s.l: s.n.].
- ANTÔNIO CÂNDIDO. *Formação da literatura brasileira : (momentos decisivos).* Belo Horizonte: Editora Itatiaia, 2000.
- ASPECTOS HISTÓRICOS E SOCIAIS DA FORMAÇÃO DO ROMANCE EM PAÍSES DA PERIFERIA DO CAPITALISMO – LL Journal. Disponível em: <https://lljournal.commons.gc.cuny.edu/2010-2-doprado-texto/>. Acesso em: 12 abr. 2023.
- BARBOZA, M.; MESTRA; HISTÓRIA -UFPE, E. POLÍTICA INDIGENISTA NO SÉCULO XIX: CONSEQUÊNCIAS ATUAIS. [s.l: s.n.]. Disponível em: [https://www.encontro2020.pe.anpuh.org/resources/anais/22/anpuh-pe-eeh2020/1601743351\\_ARQUIVO\\_11dfaab7f3cb8ebd775efc61fd7c061f.pdf](https://www.encontro2020.pe.anpuh.org/resources/anais/22/anpuh-pe-eeh2020/1601743351_ARQUIVO_11dfaab7f3cb8ebd775efc61fd7c061f.pdf). Acesso em: 12 abr. 2023.
- BOAS, L. G. V. CONSIDERAÇÕES SOBRE A CONCENTRAÇÃO FUNDIÁRIA NO BRASIL. *Revista Georaguia*, v. 8, n. 1, 17 jul. 2018.
- CALDEIRA, J. *História da riqueza no Brasil.* [s.l.] Sextante, 2017.
- CARNEIRO, M. *Índios no Brasil: história, direitos e cidadania.* São Paulo, SP: Claro Enigma, 2012.
- CARVALHO, José Murilo de. *Cidadania no Brasil. O longo Caminho.* 3ª ed. Rio de Janeiro: Civilização Brasileira, 2002
- CHALHOUB, S. *Visões da liberdade : uma história das últimas décadas da escravidão na Corte.* São Paulo, Sp: Companhia De Bolso, 2011.
- CHRYSOSTOMO, M. I. DE J.; SANTOS, H. M. G. TERRITORIAL POLITICS IN NATIVE-BRAZILIAN BACKWOODS, XIX CENTURY. *MERCATOR*, v. 15, n. 01, p. 55–71, 26 mar. 2016.



CROCE, M. A Economia do Brasil no século XIX. [s.l: s.n.]. Disponível em: <[https://www.abphe.org.br/arquivos/2015\\_marcus\\_antonio\\_croce\\_a-economia-do-brasil-no-seculo-xix.pdf](https://www.abphe.org.br/arquivos/2015_marcus_antonio_croce_a-economia-do-brasil-no-seculo-xix.pdf)>.

DE MOURA CUNHA, G. H. O ALGODÃO NA ECONOMIA DA PROVÍNCIA DO CEARÁ DURANTE O SÉCULO XIX: ALGUMAS CONSIDERAÇÕES SOBRE A SUA IMPORTÂNCIA. Revista de Desenvolvimento Econômico, v. RDE - Ano XXII – V. 3 - N. 47, dez. 2020.

DEAN, Warren. A Industrialização de São Paulo. 2a ed. São Paulo: Difel Editorial S.A., 1971.

DOS, R.; GREGÓRIO, S. “PETRÓPOLIS -CIDADE IMPERIAL?” DA REPRESENTAÇÃO DO ESPAÇO AOS ESPAÇOS DE REPRESENTAÇÃO. [s.l: s.n.]. Disponível em: <<http://observatoriogeograficoamericalatina.org.mx/egal12/Geografiasocioeconomica/Geografiaurbana/277.pdf>>. Acesso em: 12 abr. 2023.

DUNKER, C. et al. Sonhos confinados. [s.l.] Autêntica Editora, 2021.

FERNANDES, P. C. A. A fundação de Belo Horizonte: ordem, progresso e higiene, mas não para todos. Cadernos Metrôpole, v. 23, n. 52, p. 1061–1084, dez. 2021.

FERREIRA LEITE, C. H. HISTÓRIA AGRÁRIA E BRASIL NO SÉCULO XIX: CONTRIBUIÇÕES E REFLEXÕES PARA O ESTUDO DA COLÔNIA MILITAR DO JATAÍ (1851-1897). XIX Encontro de História da Anpuh-Rio, p. 21–25, set. 2020.

FRANCO MORETTI. Atlas of the European novel, 1800-1900. London: Verso, 1999.

FRANCO MORETTI. Distant reading. London: Verso, 2015.

FRANCO MORETTI. The bourgeois : between literature and history. London ; Brooklyn, Ny: Verso, 2013.

FRANCO MORETTI. The novel. Princeton, N.J.: Princeton University Press, 2007.

GRAHAM, S.; MILLIGAN, I.; WEINGART, S. Exploring big historical data : the historian's microscope. London: Imperial College Press, 2016.

HAMAM DE FIGUEIREDO, A. Formação territorial\*. [s.l: s.n.]. Disponível em: <[https://biblioteca.ibge.gov.br/visualizacao/livros/liv97884\\_cap1.pdf](https://biblioteca.ibge.gov.br/visualizacao/livros/liv97884_cap1.pdf)>.

História da literatura ocidental (vol. 107) - 3a ed. Disponível em: <<https://livraria.senado.leg.br/historia-da-literatura-ocidental-vol-107-3a-ed>>. Acesso em: 12 abr. 2023.

HORNE, G. The deepest south : the United States, Brazil, and the African slave trade. New York: New York University Press, 2007.

LESSA, R.. A invenção Republicana. op.cit., 2015

LILIA MORITZ SCHWARCZ. História do Brasil nação : 1808-2010. 2, A construção nacional : 1830-1889. Rio De Janeiro: Objetiva, 2012.

LILIA MORITZ SCHWARCZ; MURGEL, M. Brasil : Uma Biografia. São Paulo, Sp: Companhia Das Letras, 2015.

LUÍS AUGUSTO FISCHER. Duas formações, uma história : das ideias fora do lugar ao perspectivismo ameríndio. Porto Alegre, Rs: Arquipélago, 2021.

MATOS, O. N. DE. A cidade de São Paulo no século XIX. Revista de História, v. 10, n. 21-22, p. 89, 7 jun. 1955.

MELO, P. ENTRE “ABUSOS, USURPAÇÕES E DESACERTOS”: TERRAS INDÍGENAS NO RIO DE JANEIRO, SÉCULO XIX. Almanack, n. 30, 2022.

MEYER, M. Folhetim: Uma história. São Paulo: Companhia das Letras, 1996.

PONTES, C. J. F. O PRIMEIRO CICLO DA BORRACHA NO ACRE: DA FORMAÇÃO DOS SERINGAIS AO GRANDE COLAPSO. South American Journal of Basic Education, Technical and Technological, v. 1, n. 1, 10 dez. 2014.

REIS, J. J. Presença Negra: conflitos e encontros. In: INSTITUTO BRASILEIRO DE GEOGRAFIA E ESTATÍSTICA. Brasil: 500 anos de povoamento Rio de Janeiro, 2000.

SANTOS, M. A urbanização brasileira. São Paulo: Edusp, 2013.

UNIVERSIDADE FEDERAL DA BAHIA FACULDADE DE FILOSOFIA E CIÊNCIAS HUMANAS PROGRAMA DE PÓS-GRADUAÇÃO EM HISTÓRIA SOCIAL ELIANE MARIA VASCONCELOS DO NASCIMENTO OLINDA: UMA LEITURA HISTÓRICA E PSICANALÍTICA DA MEMÓRIA SOBRE A CIDADE. [s.l: s.n.]. Disponível em: <<https://repositorio.ufba.br/bitstream/ri/11230/1/Tese%20Eliane%20Nascimento1.pdf>>. Acesso em: 15 fev. 2023.

## GLOSSÁRIO

Mapa completo:

([https://mapasseculoxix.github.io/Humanidades\\_Digitais/Mapa\\_tempo\\_completo.html](https://mapasseculoxix.github.io/Humanidades_Digitais/Mapa_tempo_completo.html))

Mapa completo (densidade):

([https://mapasseculoxix.github.io/Humanidades\\_Digitais/map\\_tempo\\_densidade\\_completo.html](https://mapasseculoxix.github.io/Humanidades_Digitais/map_tempo_densidade_completo.html))

Mapa completo (autores e obras):

([https://mapasseculoxix.github.io/Humanidades\\_Digitais/Mapa\\_todos\\_autores\\_completo.html](https://mapasseculoxix.github.io/Humanidades_Digitais/Mapa_todos_autores_completo.html))

## APÊNDICE - LISTA DOS ROMANCES ANALISADOS

nome_do_romance	autor	cidade_natal	ano_de_publicação
Ubirajara	José de Alencar	Fortaleza	1874
Iracema	José de Alencar	Fortaleza	1865
Guerra dos mascates	José de Alencar	Fortaleza	1871
O guarani	José de Alencar	Fortaleza	1857
Gonzaga ou a conjuração do Tiradentes	Teixeira e Sousa	Cabo Frio	1851
Lourenço de Mendonça	Moreira de Azevedo	Itaboraí	1868
No tempo do rei	Moreira de Azevedo	Itaboraí	1899
Os franceses no Rio de Janeiro	Moreira de Azevedo	Itaboraí	1870
As minas de prata	José de Alencar	Fortaleza	1862
As Mulheres de Mantilha	Joaquim Manuel de Macedo	Itaboraí	1870
Memórias da Rua do Ouvidor	Joaquim Manuel de Macedo	Itaboraí	1878
A retirada de Laguna	Visconde de Taunay	Rio de Janeiro	1874
Acaiaca	Joaquim Felício dos Santos	Serro	1866
Adelaide de Sargans	Caetano Alves de Sousa Filgueiras	Salvador	1869
O Índio Afonso	Bernardo Guimarães	Ouro Preto	1872
O vaqueano	Apolinário Porto-Alegre	Porto Alegre	1872
Narcisa de Vilar	Ana Luísa de Azevedo Castro	São Francisco do Sul	1858
A providência	Teixeira e Sousa	Cabo Frio	1853
Amor de esposo	Pedro Américo	Areias	1886
O matuto	Franklin Távora	Baturité	1878
Til	José de Alencar	Fortaleza	1871
Sonhos de ouro	José de Alencar	Fortaleza	1872
Senhora	José de Alencar	Fortaleza	1875
O tronco do ipê	José de Alencar	Fortaleza	1871
O sertanejo	José de Alencar	Fortaleza	1875
O gaúcho	José de Alencar	Fortaleza	1870
Lucíola	José de Alencar	Fortaleza	1862
Alfarrábios	José de Alencar	Fortaleza	1873
A pata da gazela	José de Alencar	Fortaleza	1870
A alma do Lázaro	José de Alencar	Fortaleza	1872
Cinco Minutos	José de Alencar	Fortaleza	1856
A viuvinha	José de Alencar	Fortaleza	1857
Diva	José de Alencar	Fortaleza	1864
Encarnação	José de Alencar	Fortaleza	1893
Dom Casmurro	Machado de Assis	Rio de Janeiro	1899
Memórias Póstumas de Brás Cubas	Machado de Assis	Rio de Janeiro	1880
Quincas Borba	Machado de Assis	Rio de Janeiro	1886
Ressurreição	Machado de Assis	Rio de Janeiro	1872
A mão e a luva	Machado de Assis	Rio de Janeiro	1874
Helena	Machado de Assis	Rio de Janeiro	1876
Iaiá Garcia	Machado de Assis	Rio de Janeiro	1878
A luneta mágica	Joaquim Manuel de Macedo	Itaboraí	1869
A moreninha	Joaquim Manuel de Macedo	Itaboraí	1844
As Vítimas-Algozes	Joaquim Manuel de Macedo	Itaboraí	1869

O Moço Louro	Joaquim Manuel de Macedo	Itaboraí	1845
Os Dois Amores	Joaquim Manuel de Macedo	Itaboraí	1848
A carne	Júlio Ribeiro	Sabará	1888
A conquista	Coelho Neto	Caxias	1899
Inverno em flor	Coelho Neto	Caxias	1897
A escrava Isaura	Bernardo Guimarães	Ouro Preto	1875
O Ermitão de Muquém	Bernardo Guimarães	Ouro Preto	1869
O garimpeiro	Bernardo Guimarães	Ouro Preto	1872
O Seminarista	Bernardo Guimarães	Ouro Preto	1872
Inocência	Visconde de Taunay	Rio de Janeiro	1872
A normalista	Adolfo Caminha	Aracati	1893
Bom-Crioulo	Adolfo Caminha	Aracati	1895
Tentação	Adolfo Caminha	Aracati	1896
A rainha do ignoto	Emília Freitas	Jaguaruana	1899
Flor de Sangue	Valentim Magalhães	Rio de Janeiro	1897
O Cortiço	Aluísio Azevedo	São Luís	1890
O homem	Aluísio Azevedo	São Luís	1887
O mulato	Aluísio Azevedo	São Luís	1881
Memórias de um Condenado ou A Condessa Vésper	Aluísio Azevedo	São Luís	1882
A casa de pensão	Aluísio Azevedo	São Luís	1884
Filomena Borges	Aluísio Azevedo	São Luís	1884
Mistério da Tijuca ou Girândola de Amores	Aluísio Azevedo	São Luís	1882
O coruja	Aluísio Azevedo	São Luís	1890
O livro de uma sogra	Aluísio Azevedo	São Luís	1895
Hóspede	Pardal Mallet	Bagé	1887
Maurício ou Os paulistas em S. João de El-Rei	Bernardo Guimarães	Ouro Preto	1877
Memórias de um pobre diabo	Bruno Seabra	Belém (PA)	1868
Miragem	Coelho Neto	Caxias	1895
Mocidade morta	Gonzaga Duque	Rio de Janeiro	1899
Mota Coqueiro ou a pena de morte	José do Patrocínio	Campos dos Goytacazes	1877
Os retirantes	José do Patrocínio	Campos dos Goytacazes	1889
O cabeleira	Franklin Távora	Baturité	1876
O Sacrifício	Franklin Távora	Baturité	1879
O missionário	Inglês de Souza	Óbidos	1888
O morto	Coelho Neto	Caxias	1898
Rose-Castle	Virgílio Várzea	Florianópolis	1893
Uma tragédia no Amazonas	Raul Pompeia	Angra dos Reis	1880
Úrsula	Maria Firmina dos Reis	São Luís	1859
O Ateneu	Raul Pompeia	Angra dos Reis	1888
A baronesa de amor	Joaquim Manuel de Macedo	Itaboraí	1876
A corveta Diana	Barão de Teffé	Itaguaí	1873
A divina pastora	José Antônio do Vale Caldre Fião	Porto Alegre	1847
A espelunca	Germano Hasslocher	Santa Cruz do Sul	1889
A ilha maldita	Bernardo Guimarães	Ouro Preto	1879
A misteriosa	Joaquim Manuel de Macedo	Itaboraí	1872
A mocidade de Trajano	Visconde de Taunay	Rio de Janeiro	1871

A namoradeira	Joaquim Manuel de Macedo	Itaboraí	1870
Amor que mata	Visconti Coaracy	Niterói	1873
As fatalidades de dois jovens	Teixeira e Sousa	Cabo Frio	1856
Casamento e mortalha no céu se talha	Júlio César Leal	Bahia	1876
Dalmo	Lúis Ramos Figueira	Angra dos Reis	1863
Dedicação de uma amiga	Dionísia Gonçalves Pinto	Nísia Floresta	1850
Elzira - a morta virgem	Pedro Ribeiro Viana	Salvador	1883
Favos e travos	Rosendo Moniz	Salvador	1872
História de um voluntário da pátria	Vicente Félix de Castro	Silveiras	1869
Lágrimas do coração	Visconde de Taunay	Rio de Janeiro	1872
Lar	Pardal Mallet	Bagé	1888
Manuel de Moraes	João Manuel Pereira da Silva	Nova Iguaçu	1866
Maricota e o padre Chico	João Salomé de Queiroga	Serro	1871
Simá	Lourenço da Silva Araújo Amazonas	Bahia	1857
Memórias do sobrinho de meu tio	Joaquim Manuel de Macedo	Itaboraí	1867
O cacaulista	Inglês de Souza	Óbidos	1876
O rei dos jagunços	Manuel Benício	Pernambuco	1899
O Rio do Quarto	Joaquim Manuel de Macedo	Itaboraí	1869
Os mistérios da roça	Vicente Félix de Castro	Silveiras	1861
Os quatro pontos cardeais; A misteriosa	Joaquim Manuel de Macedo	Itaboraí	1872
Paulo	Bruno Seabra	Belém (PA)	1861
Rosaura	Bernardo Guimarães	Ouro Preto	1883
Um casamento no arrabalde	Franklin Távora	Baturité	1869
Um noivo a duas noivas	Joaquim Manuel de Macedo	Itaboraí	1871
Vicentina	Joaquim Manuel de Macedo	Itaboraí	1854
Januário Garcia ou as sete orelhas	Joaquim Norberto	Rio de Janeiro	1852
Estrychinina	Souza/Totta/Azurenha	Porto Alegre	1897
Memórias de um sargento de milícias	Manuel Antônio de Almeida	Rio de Janeiro	1853
A carteira do meu tio	Joaquim Manuel de Macedo	Itaboraí	1855
A Massambu	Duarte Schutel	Florianópolis	1860
Lupe	Conde Afonso Celso	Ouro Preto	1894
Luizinha	Araripe Júnior	Fortaleza	1878
A afilhada	Manuel de Oliveira Paiva	Fortaleza	1889
A Capital Federal	Coelho Neto	Caxias	1893
A família Medeiros	Júlia Lopes de Almeida	Rio de Janeiro	1891
A sonâmbula da Itapuca	Leonel de Alencar	Rio de Janeiro	1861
Nina	Joaquim Manuel de Macedo	Itaboraí	1870
A cruz de fogo	José da Rocha Leão	Rio Grande	1862
Remédio para matar paixões	A. Augusto de Pinho	Bósnia-Herzegovina	1879
O forasteiro	Joaquim Manuel de Macedo	Itaboraí	1855
Marabá	Salvador de Mendonça	Itaboraí	1875
Maria ou A menina roubada	Teixeira e Sousa	Cabo Frio	1852
O filho do pescador	Teixeira e Sousa	Cabo Frio	1843
Jovens interessantes	Paulina Filadélfia	Desconhecido	1865
Padre Belchior de Pontes (incompleto)	Júlio Ribeiro	Sabará	1874
Os índios do Jaguaribe (incompleto)	Franklin Távora	Baturité	1862
Lourenço	Franklin Távora	Baturité	1878
O martírio do Tiradentes	Joaquim Norberto	Rio de Janeiro	1882
Rosa	Joaquim Manuel de Macedo	Itaboraí	1849
Uma paixão romântica	Joaquim Manuel de Macedo	Itaboraí	1860