

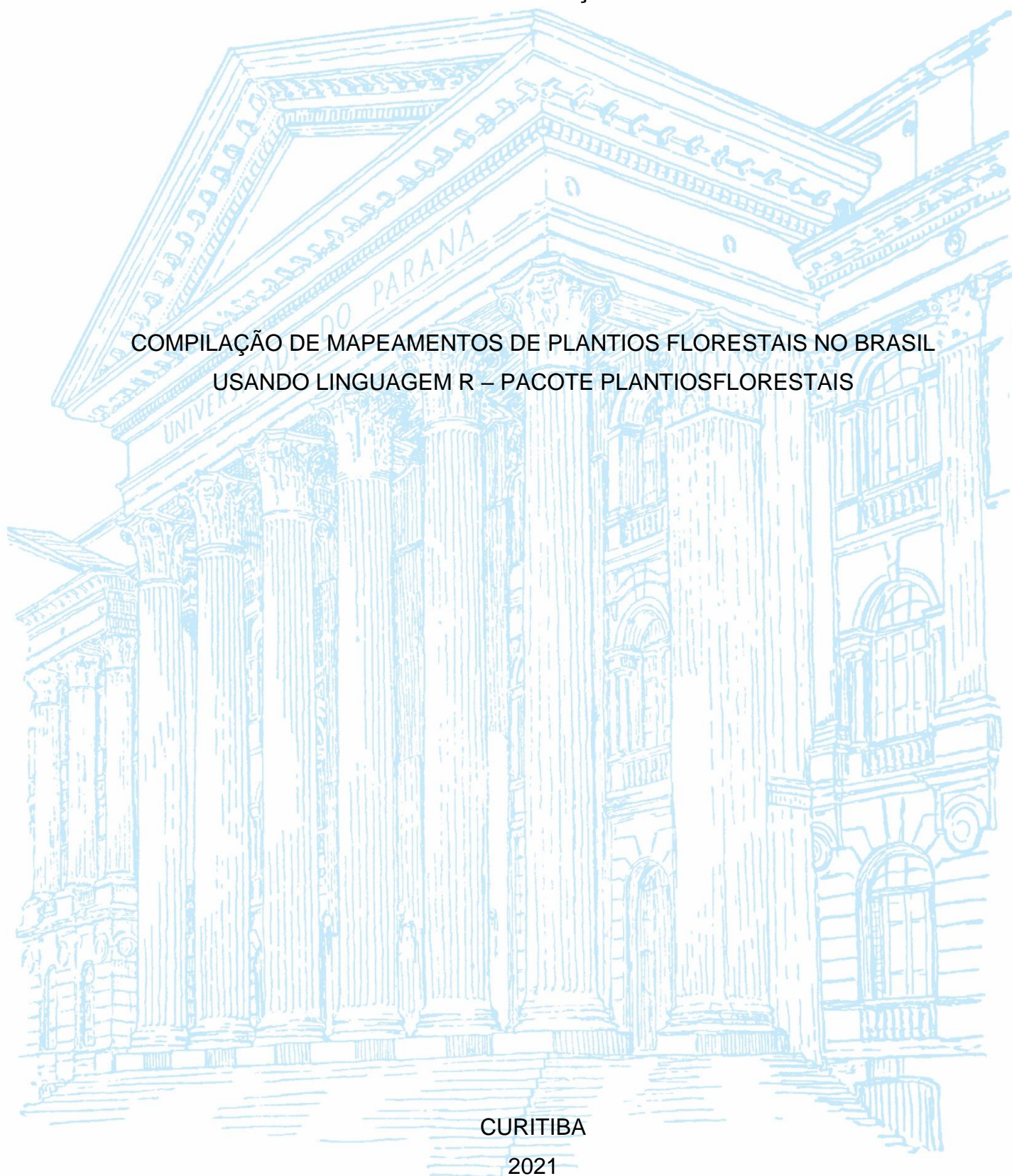
UNIVERSIDADE FEDERAL DO PARANÁ

MAYKON GABRIEL GONÇALVES PEDRO

COMPILAÇÃO DE MAPEAMENTOS DE PLANTIOS FLORESTAIS NO BRASIL
USANDO LINGUAGEM R – PACOTE PLANTIOSFLORESTAIS

CURITIBA

2021



MAYKON GABRIEL GONÇALVES PEDRO

COMPILAÇÃO DE MAPEAMENTOS DE PLANTIOS FLORESTAIS NO BRASIL
USANDO LINGUAGEM R – PACOTE PLANTIOSFLORESTAIS

Trabalho de Conclusão de Curso apresentado ao curso de Engenharia Florestal, Setor de Ciências Agrárias, Universidade Federal do Paraná, como requisito parcial à obtenção do título de Bacharel em Engenharia Florestal.

Orientador: Prof. Dr. Allan Libanio Pelissari

CURITIBA

2021



MINISTÉRIO DA EDUCAÇÃO
UNIVERSIDADE FEDERAL DO PARANÁ
SETOR DE CIÊNCIAS AGRÁRIAS
CURSO DE ENGENHARIA FLORESTAL

PARECER

Defesa nº 246

A Banca Examinadora, instituída pelo Colegiado do Curso de Engenharia Florestal do Setor de Ciências Agrárias, da Universidade Federal do Paraná, após arguir o Acadêmico **MAYKON GABRIEL GONÇALVES PEDRO** em relação ao seu Trabalho de Conclusão de Curso intitulado “**COMPILAÇÃO DE MAPEAMENTOS DE PLANTIOS FLORESTAIS NO BRASIL USANDO LINGUAGEM R – PACOTE PLANTIOSFLORESTAIS**”, é de parecer favorável à **APROVAÇÃO** na Disciplina ENGF010 - Trabalho de Conclusão de Curso em Engenharia Florestal, condicionada a entrega da versão final corrigida.

Prof. Dr. Luan Demarco Fiorentin
1º. Avaliador

Prof.^a Dr.^a Carla Krulikowski Rodrigues
2º. Avaliadora

Prof. Dr. Allan Libanio Pelissari
Orientador - Presidente da Banca

Curitiba, 10 de agosto de 2021.

Prof. Dr. Allan Libanio Pelissari
Vice-Coordenador do Curso de Engenharia Florestal em exercício

AGRADECIMENTOS

Agradeço primeiramente à universidade pública e a todos que fazem dela possível. Os programas de cota e auxílio aos estudantes desamparados fazem toda a diferença na vida daqueles que precisam. Sem isso, dificilmente eu teria conseguido chegar nesta fase de finalização do ensino superior.

Agradeço aos bons professores que tive, por terem me ensinado muito.

Agradeço ao meu professor e orientador Allan, por me auxiliar nesta caminhada do mundo florestal, e, principalmente, por ter me recomendado e incentivado a aprender a Linguagem R.

Agradeço à comunidade envolvida com o R e ao mundo *opensource*, o material que usei para construir os códigos desenvolvidos neste trabalho foram, em sua grande maioria, disponibilizados gratuitamente por pessoas dispostas a compartilhar conhecimento.

Agradeço aos meus pais, Valdecir Pedro e Maria da Gloria, pela educação que me proporcionaram e pelos caminhos que me ajudaram a chegar.

Agradeço, por último, à minha companheira, Vitoria, por ser todos os dias uma das razões para eu ser positivo e otimista em relação ao futuro deste mundo.

“O inferno não são os outros, pequena Halla. Eles são o paraíso, porque um homem sozinho é apenas um animal. A humanidade começa nos que te rodeiam, e não exatamente em ti. Ser-se a pessoa implica a tua mãe, as nossas pessoas, um desconhecido ou a sua expectativa. Sem ninguém no presente nem no futuro, o indivíduo pensa tão sem razão quanto pensam os peixes. Dura pelo engenho que tiver e perece como um atributo indiferenciado do planeta. Parece como uma coisa qualquer”

Valter Hugo Mãe, em *A Desumanização* (2013).

RESUMO

Mapeamentos florestais são um requisito indispensável em políticas públicas e privadas que visam o setor florestal. Considerando o método de distribuição dos relatórios que disponibilizam dados sobre área e localidade dos plantios de espécies florestais, geralmente distribuídos por meio de arquivos em extensão PDF, se faz necessário uma maneira simples e prática de consulta aos dados referentes, facilitando o acesso e a análise envolvendo tais dados. Partindo desse princípio, empregou-se a linguagem de programação R para construir um pacote - uma maneira efetiva de compartilhamento de códigos, dados, funções e documentações – utilizando dados obtidos desses mapeamentos. Esse pacote consolidou os dados pertencentes aos mapeamentos florestais de onze relatórios distintos, com abordagem tanto municipal quanto estadual. As bases consolidadas foram divididas em duas, i) uma para os mapeamentos de municípios e ii) outra para os mapeamentos de estados. O pacote desenvolvido possui sete funções que tornam seu uso acessível para iniciantes da Linguagem R, sendo quatro delas relativas à exploração dos dados e uma destinada à confecção de gráficos, ao passo que as duas últimas funções são aplicadas para exportação das bases de dados para arquivos em formatos CSV e XLSX. Pode-se concluir que o pacote forneceu uma forma simplificada de acesso e exploração dos dados de área dos mapeamentos, entretanto, não substitui a leitura e a consulta dos relatórios originais, visto que dispõem de informações sobre a metodologia empregada nos levantamentos de área e, principalmente, contextualização dos números e grandezas dessas áreas dentro do setor florestal.

Palavras-chave: Relatórios de áreas florestais. Extração e limpeza de dados. Manipulação de PDF. Desenvolvimento de pacotes em R. Desenvolvimento de funções em R.

ABSTRACT

Forest mapping is a crucial requirement for public and private policies aimed at the forest sector. Considering the method of distribution of reports that provide data on the area and location of plantation forest species, generally shared through PDF extension files, a simple and practical way of consulting the related data is necessary in order to ease access and analysis involving such data. Based on this principle, the R programming language is used to develop a package - an effective way of sharing codes, data, functions, and documentation – using data obtained from these mappings. This package consolidates data from forest mappings of eleven different reports with different levels of detail each regarding Brazilian states and municipalities. The consolidated databases were divided into two, i) one for mapping municipalities and ii) the other for mapping states. The package built has seven functions that make its use accessible to R Language beginners. Four of these functions relate to data analysis, one for graph making, and the other two are intended to export databases to CSV and XLSX format files. It can be implied that the package provides a simplified way to access and analyze area data from mappings. However, it does not replace the inspection of original reports, as they bear information on the methodology used in area surveys and, mainly, contextualization of the numbers and magnitudes of these areas within the forest sector.

Keywords: Forest area reports. Data extraction and cleaning. PDF Handling. Development of packages in R. Development of functions in R.

LISTA DE FIGURAS

FIGURA 1 - FLUXOGRAMA: METODOLOGIA GERAL.....	21
FIGURA 2 - CONSOLE DA LINGUAGEM R	22
FIGURA 3 - INTERFACE DO RSTUDIO	23
FIGURA 4 - MAPEAMENTO HISTÓRICO IBGE - DADOS DISPONÍVEIS NO SNIF	26
FIGURA 5 - MAPEAMENTO IBGE/PEVS 2018 - RELATÓRIO DO SFB (2019).....	26
FIGURA 6 - MAPEAMENTO HISTÓRICO IBÁ - DADOS DISPONÍVEIS NO SNIF ..	27
FIGURA 7 - MAPEAMENTO IBÁ (EUCALIPTO) - RELATÓRIO ANUAL 2020	28
FIGURA 8 - MAPEAMENTO IFPR/SFB – MAPEAMENTO DOS PLANTIOS FLORESTAIS DO ESTADO DO PARANÁ 2015.....	29
FIGURA 9 - MAPEAMENTO APRE e UFPR - ESTUDO SETORIAL 2020	30
FIGURA 10 - MAPEAMENTO ACR e UDESC-CAV - ANUÁRIO ESTATÍSTICO 2019	30
FIGURA 11 - MAPEAMENTO AGEFLOR - A INDÚSTRIA DE BASE FLORESTAL NO RIO GRANDE DO SUL 2017 – SÉRIE HISTÓRICA.....	32
FIGURA 12 - MAPEAMENTO AGEFLOR - A INDÚSTRIA DE BASE FLORESTAL NO RIO GRANDE DO SUL 2017 – VINTE MUNICÍPIOS COM MAIOR ÁREA PLANTADA	32
FIGURA 13 - MAPEAMENTO AGEFLOR – O SETOR DE BASE FLORESTAL NO RIO GRANDE DO SUL 2020 – ÁREA PLANTADA POR COREDE...33	33
FIGURA 14 - MAPEAMENTO FAMATO/IMEA - DIAGNÓSTICO DE FLORESTAS PLANTADAS NO ESTADO DO MATO GROSSO 2013	33
FIGURA 15 - FLUXOGRAMA: ETAPAS DA EXTRACÇÃO E LIMPEZA DE DADOS..	35
FIGURA 16 - COMANDOS: INSTALAÇÃO DO PACOTE	37
FIGURA 17 - ORGANIZAÇÃO DO PACOTE	38
FIGURA 18 - COMANDOS: ACESSO ÀS BASES.....	39
FIGURA 19 - PRIMEIRAS LINHAS DA BASE DE MUNICÍPIOS	40
FIGURA 20 - ESTRUTURA DE DADOS DA BASE DE MUNICÍPIOS	40
FIGURA 21 - DOCUMENTAÇÃO: BASE DE DADOS DE MUNICÍPIOS.....	41
FIGURA 22 - DOCUMENTAÇÃO: BASE DE DADOS DE ESTADOS.....	42
FIGURA 23 - FUNÇÃO: MAPEAMENTOS DISPONÍVEIS.....	43
FIGURA 24 - FUNÇÃO: GÊNEROS DE PLANTIOS DISPONÍVEIS	44

FIGURA 25 - FUNÇÃO: MAPEAMENTO DISPONÍVEL EM ÁREA PARA UMA DETERMINADA UF45

FIGURA 26 - FUNÇÃO: SÉRIES HISTÓRICAS DISPONÍVEIS.....46

FIGURA 27 - COMANDOS: EXPORTAR XLSX.....49

FIGURA 28 - COMANDOS: EXPORTAR CSV.....49

LISTA DE GRÁFICOS

GRÁFICO 1 - FUNÇÃO: PLOTAR MAPEAMENTO IBÁ - BRASIL	47
GRÁFICO 2 - FUNÇÃO: PLOTAR MAPEAMENTO IBÁ - PARANÁ	48

LISTA DE QUADROS

QUADRO 1 - MAPEAMENTOS UTILIZADOS.....	25
--	----

Lista de Tabelas

TABELA 1 - PACOTES UTILIZADOS	24
TABELA 2 - FUNÇÕES POR SCRIPT	43

LISTA DE ABREVIATURAS OU SIGLAS

ACR	- Associação Catarinense de Empresas Florestais
AFUBRA	- Associação dos Fumicultores do Brasil
AGEFLOR	- Associação Gaúcha de Empresas Florestais
APRE	- Associação Paranaense de Empresas de Base Florestal
CAV	- Centro de Ciências Agroveterinárias
FAMATO	- Federação da Agricultura e Pecuária do Estado do Mato Grosso
FEPAM	- Fundação Estadual de Proteção Ambiental
IBÁ	- Indústria Brasileira de Árvores
IBGE	- Instituto Brasileiro de Geografia e Estatística
IDE	- <i>Integrated Development Environment</i>
IFPR	- Instituto Federal do Paraná
IMEA	- Instituto Mato-Grossense de Economia e Agropecuária
PEVS	- Produção da Extração Vegetal e da Silvicultura
SEAB	- Secretaria Estadual de Agricultura e Abastecimento
SEMA	- Secretaria Estadual do Meio Ambiente
SFB	- Serviço Florestal Brasileiro
SNIF	- Serviço Nacional de Informações Florestais
UDESC	- Universidade do Estado de Santa Catarina
UFPR	- Universidade Federal do Paraná

SUMÁRIO

AGRADECIMENTOS	4
1 INTRODUÇÃO	16
1.1 OBJETIVOS	16
1.1.1 Objetivo geral	17
1.1.2 Objetivos específicos	17
2 REVISÃO DE LITERATURA	18
2.1 MAPEAMENTOS DE PLANTIOS FLORESTAIS	18
2.2 LINGUAGEM R E IDE	18
2.3 GIT E GITHUB	19
3 MATERIAL E MÉTODOS	21
3.1 METODOLOGIA GERAL	21
3.2 LINGUAGEM R E IDE	21
3.2.1 Instalação do R e do Rstudio	21
3.3 PACOTES R UTILIZADOS	23
3.4 FONTE DE DADOS	24
3.4.1 IBGE e PEVS	25
3.4.2 IBÁ	26
3.4.3 SFB e IFPR	28
3.4.4 APRE e UFPR	29
3.4.5 ACR e UDESC-CAV	30
3.4.6 AGEFLOR	31
3.4.7 FAMATO	33
3.5 EXTRAÇÃO E ORGANIZAÇÃO DOS DADOS	34
3.6 CRIAÇÃO E DESENVOLVIMENTOS DE PACOTES	35
4 RESULTADOS E DISCUSSÃO	37
4.1 PACOTE R: PLANTIOSFLORESTAIS	37
4.1.1 Organização dos arquivos	37
4.1.2 Acessar mapeamentos	39
4.1.3 Funções	42
4.1.3.1 Mapeamentos disponíveis	43
4.1.3.2 Gêneros de plantios disponíveis	44
4.1.3.3 Mapeamento disponível por unidade federativa em área	44

4.1.3.4 Obter séries históricas	45
4.1.3.5 Plotar série histórica do IBÁ	46
4.1.3.6 Exportar bases em extensão XLSX.....	48
4.1.3.7 Exportar bases em extensão CSV.....	49
5 CONSIDERAÇÕES FINAIS	50
REFERÊNCIAS.....	51

1 INTRODUÇÃO

O setor florestal cresce cada vez mais, com perspectivas de aumento nos próximos anos e previsões de investimento de até R\$ 32,6 bilhões até 2023. Ademais, a aplicação desses recursos buscará atender a um maior consumo de matéria-prima e maior necessidade por florestas plantadas (IBÁ, 2019).

Considerando essa necessidade, o conhecimento das áreas e localidades das florestas plantadas torna-se um requisito para políticas públicas e privadas que planejam beneficiar-se com o setor (EISFELD et al., 2018). Nesse ponto, os mapeamentos de plantios florestais e os relatórios oficiais de entidades e organizações tem por característica consolidar essas informações.

Os relatórios publicados com as informações de área tendem a ser compartilhados em formatos que dificultam a manipulação dos dados, justamente por terem um foco mais amplo ao abordarem assuntos gerais do setor florestal. Esses resultados são, em sua grande maioria, disponibilizados publicamente por meio de relatórios em formato PDF. Usando isso como motivação, este trabalho se propôs a disponibilizar, de maneira simples e prática, as informações de área e localidade de dados referentes aos plantios de espécies florestais no Brasil.

A linguagem R foi o meio escolhido para facilitar o acesso aos dados, pensando em praticidade, reprodutibilidade e alcance. É uma linguagem de destaque no mundo da programação, chegando ao décimo segundo lugar em 2021 no índice TIOBE (TIOBE, 2021), companhia de *software* que realiza uma classificação das linguagens de programação mais usadas no mundo. Além disso, o R tem uma longa e positiva história na estatística, sendo uma das primeiras linguagens de programação que permitiu aos usuários realizarem análises estatísticas por meio de softwares gratuitos, com uma grande gama de pacotes (bibliotecas) disponíveis que favorecem essas análises.

Usando a linguagem de programação escolhida, o trabalho aborda o desenvolvimento de um pacote em R, que pode ser entendido como um compilado de códigos, dados, documentação e testes que pode ser compartilhado de maneira descomplicada entre a comunidade de usuários (WICKHAM et al., 2019).

1.1 OBJETIVOS

1.1.1 Objetivo geral

Construir um pacote em linguagem R para distribuir os dados de mapeamentos públicos de área de plantios florestais disponíveis no Brasil.

1.1.2 Objetivos específicos

- Extrair dados dos mapeamentos públicos de plantios florestais, geralmente disponibilizados em formato PDF;
- Realizar a limpeza e organização dos dados, tornando-os parte de uma base de dados estruturada; e
- Construir e desenvolver um pacote em R para disponibilizar o acesso rápido e descomplicado aos dados dos mapeamentos de plantios florestais.

2 REVISÃO DE LITERATURA

2.1 MAPEAMENTOS DE PLANTIOS FLORESTAIS

Os mapeamentos de plantios florestais têm como objetivo levantar o uso e a cobertura de solos que estão sendo utilizados para fins silviculturais, isto é, com cultivos de florestas. Pode ser realizado com diversas metodologias e geralmente terá um nível de acurácia para qualificar as informações levantadas (NICOLETTI et al., 2019).

Coutinho et al. (2017) categorizam o mapeamento de florestas plantadas como fundamental para guiar o planejamento de nível estratégico, pois gera informações que incentivam políticas públicas voltadas para o setor florestal. Além disso, permitem diferentes iniciativas no meio privado, como atração de novos investidores e planejamento de consumo de madeira, aumento de oferta, direcionamento de metas de expansão, verificação de aptidão local, entre outros.

Quanto à forma de obtenção dos dados de área cultivadas com florestas, Brandelero (2007) resume em três formas distintas, sendo elas: uso de imagens de satélite nas técnicas de sensoriamento remoto; fotogrametria aérea; ou amostragem direta no campo. As imagens de satélite possibilitam uma vantagem por serem capazes de fornecer informações de cobertura para grandes áreas, com custos relativamente menores se comparados à uma amostragem em campo, por exemplo; e são geralmente as mais utilizadas para mapeamentos de regiões extensas (CARDOSO, 2017).

2.2 LINGUAGEM R E IDE

R é uma linguagem de programação de código aberto (*open source*) criada e desenvolvida no departamento de Estatística da Universidade de Auckland, Nova Zelândia, em 1993, pelos estatísticos Ross Ihaka e Robert Gentleman. A linguagem é constantemente melhorada e atualizada por uma comunidade voluntária de colaboradores do mundo todo (R CORE TEAM, 2021).

A linguagem R foi construída por seus desenvolvedores pensando em uma semântica similar a outras linguagem que ambos estavam acostumados, nesse caso, a linguagem *Scheme*. Além disso, pensando nas estruturas de dados, decidiram desenvolver a sintaxe baseada em uma linguagem focada em estatística:

na época, a linguagem S. O S pode ser considerado, de certa forma, um antecessor ao R, pois os estatísticos atuavam fortemente com a mesmo (IHAKA et al., 1996). Tendo essas características, o R foi aceito de maneira natural na comunidade estatística.

O R durante muito tempo teve destaque por ser um dos poucos softwares livres e gratuitos para realizar processamentos e análises estatísticas. Atualmente, com o crescimento do mesmo tanto em comunidade quanto em potencial, a linguagem se difundiu na Ciência de Dados em geral (MUECHEN, 2012). Sua utilização é pautada no desenvolvimento de diversos produtos, como: análise de dados e estatísticas, relatórios automáticos, gráficos e visualização de dados, apresentações de slides, modelagem e *machine learning*, *dashboards*, entre outros.

Uma das formas mais práticas e recomendadas de se compartilhar códigos e análises no R é por meio de pacotes. Wickham et al. (2019) definem um pacote como uma junção de códigos, dados, documentações e testes que qualquer pessoa pode facilmente compartilhar com outros. Os pacotes são criados pela comunidade e são distribuídos por duas principais formas: pelo CRAN, sigla para “*Comprehensive R Archive Network*”, local de armazenamento público gratuito para consolidação e distribuição dos pacotes em R; e pelo GitHub, *síte* para armazenagem, versionamento e compartilhamento de códigos (WICKHAM et al., 2019).

Como toda linguagem de programação, o R também dispõe de uma *Integrated Development Environment* (IDE), ou em tradução livre: Ambiente de Desenvolvimento Integrado. Basicamente, a IDE é um editor de códigos de uma ou mais linguagens de programação, geralmente provido de diversas ferramentas que tem como objetivo tornar o trabalho do programador menos complicado, fornecendo diferentes meios para isso (SANTOS, 2007). O *software* RStudio é uma das IDEs mais utilizadas para editar códigos em R, tendo como principal característica um ambiente mais amigável para os iniciantes na linguagem, além do mesmo ambiente possuir ferramentas extremamente produtivas para os usuários mais avançados (WICKHAM et al., 2019).

2.3 GIT E GITHUB

O Git é um sistema de controle de versões, utilizado para desenvolvimento de *software* ou códigos, controlando as diferentes etapas de realização do projeto, fornecendo um versionamento prático e eficaz (GIT-FAQ, 2021). Os diretórios de trabalhos do Git são denominados como “repositórios”, um local que permite acompanhamento total das diferentes versões, além de não depender de acesso a um servidor central. O *software* Git foi desenvolvido por Linus Torvalds, criador do Linux, inicialmente para uso e desenvolvimento do *kernel* Linux, e tem como característica ser *open source*, como a Linguagem R (CHACON et al., 2014).

Uma das maiores vantagens do Git é sua capacidade de facilitar a reprodutibilidade científica, como levantado por Perkel (2018). Com o histórico completo dos códigos e etapas realizadas, é completamente factível não só analisar o resultado do projeto, mas sim toda sua construção e desenvolvimento, assim permitindo testes e reconstruções do código realizadas por outros que não seus autores.

Pensando na reprodutibilidade com Git, é necessária uma forma de compartilhar facilmente os repositórios criados pelo usuário. Para isso, o site GitHub é uma das maiores recomendações da comunidade R, segundo Wickham (2015). O GitHub é um website que fornece aos programadores e desenvolvedores uma maneira simples e prática de armazenar e compartilhar seus códigos e projetos, além de permitir o trabalho em equipe e contribuições de melhoria para o repositório advindas de qualquer lugar do mundo por todos aqueles dispostos a contribuir (CHACON et al., 2014).

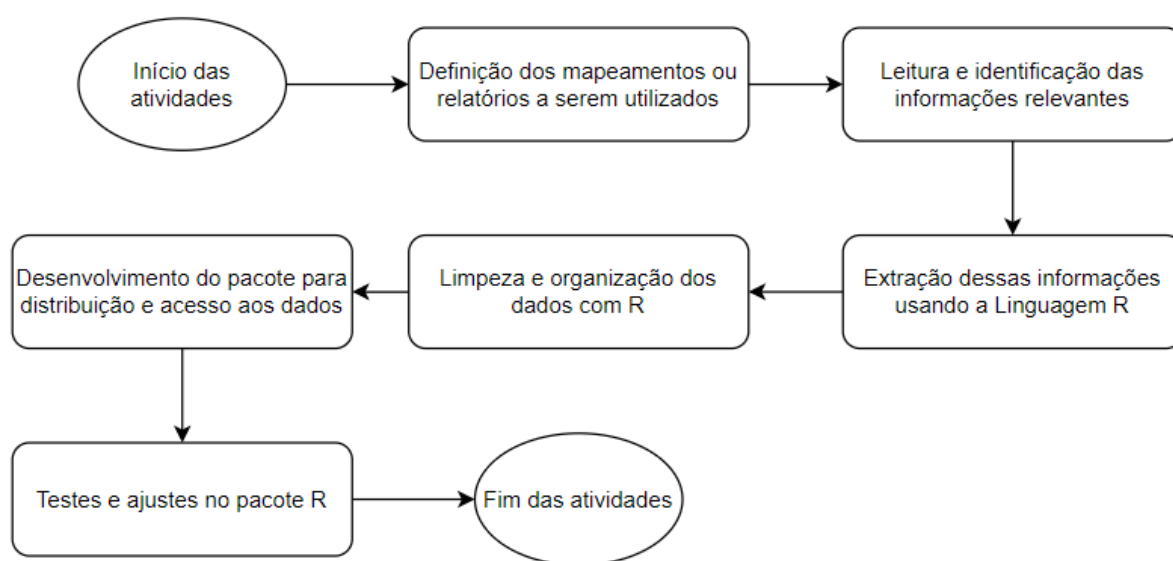
O GitHub é amplamente utilizado pela comunidade R, pois permite a distribuição de pacotes sem necessariamente estarem dentro dos padrões requisitados pelo R-Core Team. Em conjunto com essa vantagem está o fato de permitir que diferentes usuários reportem ao autor do pacote possíveis *bugs*, ou, em alguns casos, os usuários podem sugerir diretamente a mudança para o autor, cabendo ao “mantenedor” do projeto apenas aceitar essa alteração (WICKHAM, 2015).

3 MATERIAL E MÉTODOS

3.1 METODOLOGIA GERAL

O fluxograma apresentado na FIGURA 1 resume as etapas aplicadas para a realização deste trabalho.

FIGURA 1 - FLUXOGRAMA: METODOLOGIA GERAL



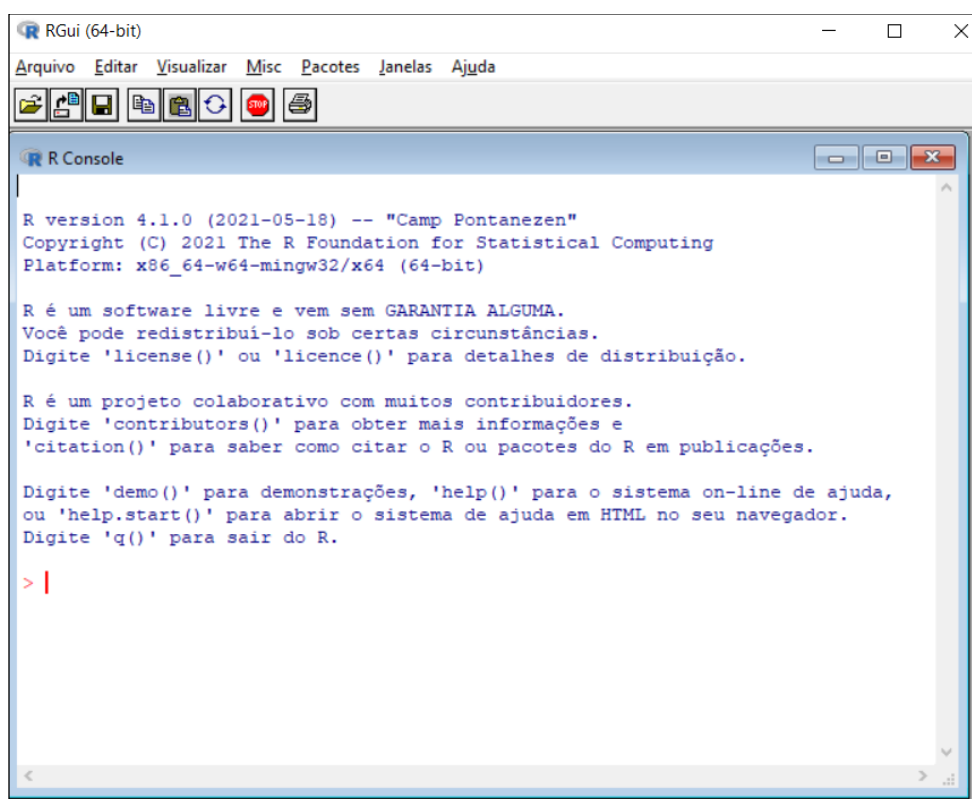
FONTE: O autor (2021).

3.2 LINGUAGEM R E IDE

3.2.1 Instalação do R e do Rstudio

O *download* e posterior instalação do software que permite o uso da Linguagem R pode ser realizado pelo *site* oficial da linguagem: <https://cran.r-project.org/>. Nesse caso, o instalador utilizado foi o executável obtido pelo site, na seção: “Download R for Windows”. Esse procedimento pode variar de acordo com o sistema operacional do usuário. No momento da realização deste trabalho, o R se encontra em sua versão 4.1.0 (FIGURA 2).

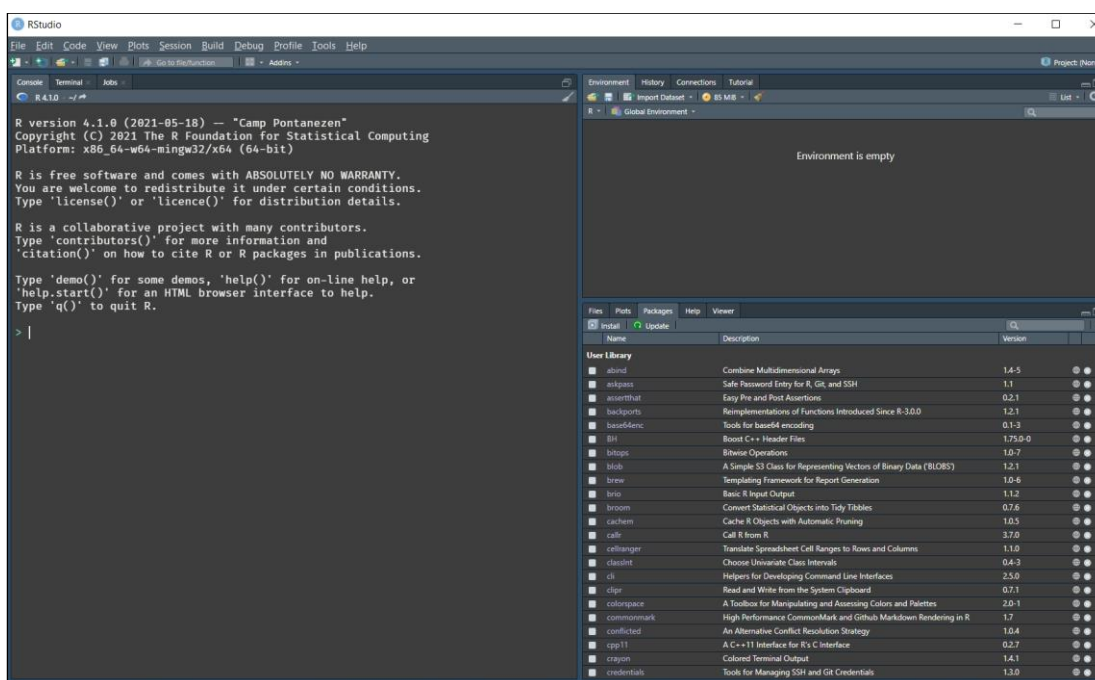
FIGURA 2 - CONSOLE DA LINGUAGEM R



FONTE: O autor (2021).

O ambiente de desenvolvimento integrado (IDE) RStudio pode ser obtido pelo site: <https://www.rstudio.com/products/rstudio/download/#download>. Para este trabalho, o executável utilizado foi o arquivo advindo da seção “*Download RStudio For Windows*” na versão 1.4.1717 (FIGURA 3).

FIGURA 3 - INTERFACE DO RSTUDIO



FONTE: O autor (2021).

3.3 PACOTES R UTILIZADOS

Dentro do R tem-se pacotes básicos que já são instalados em conjunto com a linguagem, porém, pode-se instalar pacotes adicionais desenvolvidos pela comunidade, assim aumentando a gama de utilizações que podem ser efetuadas com o R.

Para o desenvolvimento dos processos envolvidos neste trabalho, foram utilizados pacotes que auxiliaram em diferentes tarefas, como: extração automática dos dados, limpeza e organização de tabelas, sumarização de resultados, manipulação de *dataframes*, desenvolvimento de pacotes, entre outros. Os pacotes R utilizados no trabalho se encontram na TABELA 1.

TABELA 1 - PACOTES UTILIZADOS

Pacote	Objetivo de utilização
<i>pacman</i>	Organização para instalação de outros pacotes
<i>tidyverse</i>	Manipular, transformar e visualizar dados
<i>pdftools</i>	Extração de dados e páginas de PDFs
<i>tabulizer</i>	Extração de tabelas de PDFs
<i>janitor</i>	Limpeza de tabelas
<i>tesseract</i>	Processamento OCR
<i>Rcpp</i>	Necessário para não ocorrer problemas com o <i>tesseract</i>
<i>magick</i>	Leitura de imagens para o processamento OCR
<i>geobr</i>	Gerar bases dos estados e municípios do Brasil
<i>usethis</i>	Auxílio no desenvolvimento de pacotes
<i>devtools</i>	Criação e desenvolvimento de pacotes
<i>ggtext</i>	Inserir rótulos nos gráficos
<i>glue</i>	Transformar títulos e subtítulos dos gráficos em <i>markdown</i>
<i>sysfonts</i>	Importar fonte de texto específica para uso nos gráficos
<i>showtext</i>	Exibir fonte importada nos gráficos
<i>writexl</i>	Exportar arquivos do R em planilhas excel

FONTE: O autor (2021)

3.4 FONTE DE DADOS

Todos os dados utilizados no presente trabalho foram obtidos diretamente dos mapeamentos e relatórios oficiais de entidades públicas e privadas, cada qual descrita nos tópicos seguintes, sendo todos os arquivos distribuídos de maneira aberta e gratuita para acesso ao público interessado, desde que citada a fonte dos dados. O QUADRO 1 resume os mapeamentos e relatórios utilizados, com o nome da organização responsável e a que nível de abrangência ele se propôs a abordar, além de informar qual o estado ou região mapeada.

QUADRO 1 - MAPEAMENTOS UTILIZADOS

Responsável	Citação	Relatório	Nível de abrangência	Região/Estado	Ano-base
IBGE	SNIF (2018)	Histórico disponível no SNIF	Municipal	Brasil	2014-2016
SFB	SFB (2019)	Florestas do Brasil em resumo 2019	Estadual	Brasil	2018
IBÁ	SNIF (2018)	Histórico disponível no SNIF	Estadual	Brasil	2006-2016
IBÁ	IBÁ (2020)	IBÁ - Relatório Anual 2020	Estadual	Brasil	2009-2019
APRE	APRE (2020)	APRE - Estudo Setorial 2020	Estadual	Paraná	2019
IFPR	IFPR (2015)	Mapeamentos dos plantios florestais do estado do Paraná	Municipal	Paraná	2014
ACR	ACR (2019)	ACR - Anuário Estatístico de Base Florestal para o estado de Santa Catarina 2019	Estadual/Regional	Santa Catarina	2018
AGEFLOR	AGEFLOR (2017)	AGEFLOR - A indústria de base florestal no Rio Grande do Sul 2017	Estadual/Regional	Rio Grande do Sul	2006-2016
AGEFLOR	AGEFLOR (2020)	AGEFLOR - O setor de base florestal no Rio Grande do Sul 2020	Estadual/Regional	Rio Grande do Sul	2010-2016, 2019
FAMATO	FAMATO (2013)	Diagnóstico de Florestas Plantadas do Estado do Mato Grosso	Municipal	Mato Grosso	2012

FONTE: O autor (2021)

3.4.1 IBGE e PEVS

Do Instituto Brasileiro de Geografia e Estatística (IBGE) foi usado o mapeamento histórico disponível no Portal Brasileiro de Dados Abertos, na seção reservada ao Serviço Nacional de Informações Florestais (SNIF). Os dados compreendem o período de 2014 até 2016, em nível municipal para o país todo. A base de dados disponibilizada pelo SNIF encontra-se em formato CSV. A FIGURA 4 apresenta as dez primeiras linhas dessa base.

FIGURA 4 - MAPEAMENTO HISTÓRICO IBGE - DADOS DISPONÍVEIS NO SNIF

1	Ano (data)	País	Região	Estado	Estado sigla	Município (Municípios)	Município Estado	Latitude	Longitude	Espécie florestal	Área (ha)
2	31/12/2014	Brasil	Norte	Rondônia	RO	Alta Floresta D'Oeste	Alta Floresta D'Oeste (RO)	-12,47	-62,27	Eucalipto	0
3	31/12/2015	Brasil	Norte	Rondônia	RO	Alta Floresta D'Oeste	Alta Floresta D'Oeste (RO)	-12,47	-62,27	Eucalipto	0
4	31/12/2016	Brasil	Norte	Rondônia	RO	Alta Floresta D'Oeste	Alta Floresta D'Oeste (RO)	-12,47	-62,27	Eucalipto	0
5	31/12/2014	Brasil	Norte	Rondônia	RO	Alta Floresta D'Oeste	Alta Floresta D'Oeste (RO)	-12,47	-62,27	Pinus	0
6	31/12/2015	Brasil	Norte	Rondônia	RO	Alta Floresta D'Oeste	Alta Floresta D'Oeste (RO)	-12,47	-62,27	Pinus	0
7	31/12/2016	Brasil	Norte	Rondônia	RO	Alta Floresta D'Oeste	Alta Floresta D'Oeste (RO)	-12,47	-62,27	Pinus	0
8	31/12/2014	Brasil	Norte	Rondônia	RO	Alta Floresta D'Oeste	Alta Floresta D'Oeste (RO)	-12,47	-62,27	Outras espécies	250
9	31/12/2015	Brasil	Norte	Rondônia	RO	Alta Floresta D'Oeste	Alta Floresta D'Oeste (RO)	-12,47	-62,27	Outras espécies	2425
10	31/12/2016	Brasil	Norte	Rondônia	RO	Alta Floresta D'Oeste	Alta Floresta D'Oeste (RO)	-12,47	-62,27	Outras espécies	3316

FONTE: IBGE (2021).

Para os dados da pesquisa de Produção da Extração Vegetal e da Silvicultura (PEVS), pesquisa realizada pelo IBGE, foram usados os dados do levantamento de área de florestas plantadas por gênero (ano-base 2018) disponibilizados no relatório “Florestas do Brasil em resumo 2019” do Serviço Florestal Brasileiro (SFB). A tabela possui dados em nível estadual. As primeiras dez linhas podem ser conferidas na FIGURA 5.

FIGURA 5 - MAPEAMENTO IBGE/PEVS 2018 - RELATÓRIO DO SFB (2019)

Estado	Eucalipto (ha)	Pinus (ha)	Outras espécies (ha)	Total (ha)
MG	1.912.194	36.405	4.996	1.953.595
PR	670.954	896.242	22.571	1.589.767
MS	1.117.740	5.252	-	1.122.992
SP	883.828	194.639	3.801	1.082.268
RS	593.597	272.779	146.166	1.012.542
SC	353.824	610.944	30.138	994.906
BA	567.003	575	-	567.578
ES	269.526	2.491	375	272.392
MT	189.296	-	74.115	263.411
MA	235.655	-	9.511	245.166

FONTE: IBGE/PEVS (2019)

3.4.2 IBÁ

Os dados disponíveis da Indústria Brasileira de Árvores (IBÁ) se dividem em duas séries históricas: uma base disponível no Portal Brasileiro de Dados Abertos pelo SNIF, aborda o período de 2006 a 2016, em nível estadual; e outra série histórica que pode ser visualizada nas páginas finais do Relatório Anual 2020 do próprio IBÁ, que aborda o ano de 2009 até 2019, também em nível estadual.

A base armazenada no SNIF está em formato CSV e pode ser conferida na FIGURA 6 (primeiras linhas da tabela).

FIGURA 6 - MAPEAMENTO HISTÓRICO IBÁ - DADOS DISPONÍVEIS NO SNIF

1	Ano	Cultura	Espécie	Estado	Área (ha)
2	31/12/2006	Eucalipto	Eucalipto	Minas Gerais	1181429
3	31/12/2006	Eucalipto	Eucalipto	São Paulo	915841
4	31/12/2006	Eucalipto	Eucalipto	Mato Grosso do Sul	119319
5	31/12/2006	Eucalipto	Eucalipto	Bahia	540172
6	31/12/2006	Eucalipto	Eucalipto	Rio Grande do Sul	184245
7	31/12/2006	Eucalipto	Eucalipto	Espírito Santo	207800
8	31/12/2006	Eucalipto	Eucalipto	Maranhão	93285
9	31/12/2006	Eucalipto	Eucalipto	Paraná	121908
10	31/12/2006	Eucalipto	Eucalipto	Mato Grosso	113770

FONTE: IBÁ (2021).

O relatório anual 2020 é disponibilizado no portal do IBÁ, o material apresenta quatro tabelas com histórico, sendo uma para cada gênero (*Pinus*, *Eucalyptus* e outros) e uma com a área total. Este trabalho se propôs a extrair as três tabelas iniciais com os gêneros e consolidar as mesmas em uma só, gerando assim a tabela de área total dividida por gênero. A FIGURA 7 apresenta as dez primeiras linhas do histórico de eucalipto presente no relatório. As outras duas que foram extraídas seguem o mesmo padrão.

FIGURA 7 - MAPEAMENTO IBÁ (EUCALIPTO) - RELATÓRIO ANUAL 2020

Estado State	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018*	2019
Minas Gerais	1.300.000	1.400.000	1.401.787	1.438.971	1.404.429	1.400.232	1.395.032	1.390.032	1.381.652	1.977.410	1.920.329
São Paulo	1.029.670	1.044.813	1.031.677	1.041.695	1.010.444	976.186	976.613	946.124	937.138	1.040.284	1.215.901
Mato Grosso do Sul	290.890	378.195	475.528	587.310	699.128	803.699	826.031	877.795	901.734	1.093.805	1.124.637
Bahia	628.440	631.464	607.440	605.464	623.971	630.808	614.390	612.199	608.781	585.269	588.035
Rio Grande do Sul	271.980	273.042	280.198	284.701	316.446	309.125	308.515	308.178	309.602	426.371	456.001
Espírito Santo	204.570	203.885	197.512	203.349	221.559	228.781	227.222	233.760	234.082	225.520	225.055
Paraná	157.920	161.422	188.153	197.835	200.473	224.089	285.125	294.050	295.520	255.955	266.473
Maranhão	137.360	151.403	165.717	173.324	209.249	211.334	210.496	221.859	228.801	200.612	199.911
Mato Grosso	147.378	150.646	175.592	184.628	187.090	187.090	185.219	185.219	181.515	187.947	188.605
Pará	139.720	148.656	151.378	159.657	159.657	125.110	130.431	133.996	135.843	151.888	154.402

FONTE: IBÁ/FGV/Pöyry (2020).

3.4.3 SFB e IFPR

O mapeamento realizado pelo Instituto de Florestas do Paraná, em conjunto com o Serviço Florestal Brasileiro, visa representar as áreas de florestas plantadas de pinus e eucalipto do Paraná. O nível de abrangência é municipal, tendo dados para todas as cidades do estado. O relatório foi publicado em 2015, com ano-base 2014. A forma de disponibilização do relatório é PDF, o mesmo pode ser encontrado no *site* da APRE.

O relatório é um dos mais completos utilizados neste trabalho. O arquivo original possui várias tabelas, cada uma representando uma regional do Paraná, com seus respectivos municípios. Um adicional em relação aos outros é que foi levantada a área em situação de corte (recém-colhida). Essa informação foi consolidada dentro dos gêneros de plantio, resultando na coluna "Corte". A FIGURA 8 apresenta uma parte da tabela que resume os dados totais do relatório por regional.

FIGURA 8 - MAPEAMENTO IFPR/SFB – MAPEAMENTO DOS PLANTIOS FLORESTAIS DO ESTADO DO PARANÁ 2015

Região	Núcleo Regional	Área (ha)			Total	%
		Corte*	Eucalipto	Pinus		
Centro-Oeste	Campo Mourão	188	9.339	1.401	10.927	1,02%
	Sub-Total	188	9.339	1.401	10.927	1,02%
Centro-Sul	Curitiba	23.153	16.597	159.648	199.398	18,70%
	Guarapuava	9.592	14.037	50.870	74.499	6,99%
	Irati	8.554	9.925	39.751	58.231	5,46%
	Laranjeiras do Sul	3.565	8.229	12.241	24.035	2,25%
	Pato Branco	2.948	6.913	39.732	49.594	4,65%
	Ponta Grossa	10.045	143.849	239.448	393.342	36,88%
	União da Vitória	8.042	11.137	71.844	91.022	8,53%
	Sub-Total	65.900	210.687	613.535	890.121	83,46%

FONTE: SFB/IFPR (2015).

3.4.4 APRE e UFPR

No Estudo Setorial de 2020, realizado pela Associação Paranaense de Empresas de Base Florestal (APRE), foram levantadas as áreas de florestas plantadas por região do SEAB e núcleo regional no estado do Paraná para os gêneros de *Pinus*, *Eucalyptus* e áreas colhidas. O mapeamento foi desenvolvido em conjunto com a Universidade Federal do Paraná (UFPR), com ano-base 2019. O arquivo de acesso é disponibilizado em PDF no *site* da APRE. A tabela com os dados é semelhante à encontrada no mapeamento do IFPR. A FIGURA 9 apresenta um recorte dos dados originais.

FIGURA 9 - MAPEAMENTO APRE e UFPR - ESTUDO SETORIAL 2020

Região	Núcleo Regional	Área (ha)				%
		Corte*	Eucalipto	Pinus	Total	
Centro-Oeste	Campo Mourão	560,64	8.403,80	1.143,26	10.108,70	1,00%
	Subtotal	560,64	8.403,80	1.143,26	10.107,70	1,00%
Centro-Sul	Curitiba	17.329,82	7.543,26	171.037,04	195.910,12	19,42%
	Guarapuava	6.939,91	11.258,27	55.301,59	73.499,77	7,28%
	Irati	5.264,87	5.189,54	41.385,54	51.839,96	5,14%
	Laranjeiras do Sul	2.764,64	4.589,11	14.356,24	21.709,99	2,15%
	Pato Branco	3.368,82	2.375,45	45.736,52	51.480,79	5,10%
	Ponta Grossa	11.415,47	117.499,56	249.602,37	378.517,39	37,51%
	União da Vitória	4.739,11	7.238,85	67.158,33	79.136,29	7,84%
	Subtotal	51.822,64	155.694,04	644.577,63	852.094,31	84,45%

FONTE: APRE/UFPR (2020)

3.4.5 ACR e UDESC-CAV

A Associação Catarinense de Empresas Florestais (ACR), em conjunto com o Centro de Ciências Agroveterinárias da Universidade do Estado de Santa Catarina (UDESC-CAV), realizou um mapeamento das florestas plantadas de pinus e eucalipto por macrorregião do estado de Santa Catarina em 2018. Os dados referentes a esse trabalho foram publicados no Anuário Estatístico de Base Florestal para o Estado de Santa Catarina 2019, da ACR. A FIGURA 10 apresenta a tabela completa contida no relatório, que pode ser acessada em PDF no *site* da ACR.

FIGURA 10 - MAPEAMENTO ACR e UDESC-CAV - ANUÁRIO ESTATÍSTICO 2019

Região	Área Plantada (ha)			% da Área TOTAL
	Pinus	Eucalyptus	TOTAL	
Serrana	242.338	27.525	269.863	33%
Oeste Catarinense	161.592	97.034	258.626	31%
Norte Catarinense	98.338	34.939	133.277	16%
Vale do Itajaí	29.066	53.006	82.073	10%
Grande Florianópolis	15.131	23.469	38.600	5%
Sul Catarinense	7.136	39.282	46.418	6%
TOTAL	553.602	275.255	828.857	100%

FONTE: ACR/UDESC-CAV (2019).

3.4.6 AGEFLOR

Para os dados referentes ao estado do Rio Grande do Sul, os mapeamentos encontrados se referem aos relatórios publicados anualmente pela Associação Gaúcha de Empresas Florestais (AGEFLOR), sendo disponibilizados em PDF no *site* da associação. Os dois relatórios usados foram: A indústria de base florestal no RS 2017 e o Setor de base florestal no RS 2020.

Os relatórios publicados pela AGEFLOR são compostos pelas seguintes tabelas com áreas de florestas plantadas: série histórica por gênero de plantio; detalhe de área por gênero para o ano-base do relatório por corede (regional no Rio Grande do Sul); três tabelas com os vinte municípios com maior área plantada por gênero; e uma tabela com os vinte municípios com maior área plantada. Dessa forma, no relatório de 2017 unicamente, apesar de existirem dados em nível municipal dentro do arquivo, a área total desses municípios não corresponde ao total de florestas plantadas encontradas no estado, pois as tabelas de municípios englobam apenas aos vinte primeiros municípios com maior área plantada.

Os gêneros abordados pela AGEFLOR são: *Eucalyptus*, *Pinus* e *Acacia*. O levantamento foi efetuado em conjunto com algumas organizações do estado, como a: Fundação Estadual de Proteção Ambiental (FEPAM), Secretaria Estadual do Meio Ambiente (SEMA), Associação dos Fumicultores do Brasil (AFUBRA), Codex, e a RDK Logs. As FIGURAS 11 a 13 mostram como os dados são disponibilizados originalmente dentro do relatório.

FIGURA 11 - MAPEAMENTO AGEFLOR - A INDÚSTRIA DE BASE FLORESTAL NO RIO GRANDE DO SUL 2017 – SÉRIE HISTÓRICA

ANO	ÁREA PLANTADA (1.000 HA)			
	EUCALIPTO	PINUS	ACÁCIA	TOTAL
2006	184,2	181,4	142,4	508,0
2007	222,2	182,4	159,0	563,6
2008	277,3	173,2	188,3	638,8
2009	272,0	171,2	139,1	582,3
2010	273,0	169,0	89,9	531,9
2011	280,2	164,8	89,1	534,1
2012	284,7	164,8	90,2	539,7
2013	316,4	164,2	88,8	569,4
2014	309,1	184,6	103,6	597,3
2015	308,5	184,6	100,0	593,1
2016	426,7	264,6	89,6	780,9

FONTE: AFUBRA/AGEFLOR, FEPAM, RDK Logs, SEMA (2017).

FIGURA 12 - MAPEAMENTO AGEFLOR - A INDÚSTRIA DE BASE FLORESTAL NO RIO GRANDE DO SUL 2017 – VINTE MUNICÍPIOS COM MAIOR ÁREA PLANTADA

MUNICÍPIO	ÁREA PLANTADA (HA)	MUNICÍPIO	ÁREA PLANTADA (HA)
Encruzilhada do Sul	59.957	Bagé	15.666
São Francisco de Paula	42.957	Mostardas	15.618
Piratini	36.262	São José dos Ausentes	15.328
Cambará do Sul	29.776	Pantano Grande	15.153
Canguçu	22.473	São Gabriel	14.524
Butiá	21.721	São José do Norte	14.172
Cachoeira do Sul	19.199	Rio Pardo	11.791
Bom Jesus	17.078	Pinheiro Machado	11.306
Jaquirana	16.734	Dom Feliciano	10.979
Triunfo	16.508	Rio Grande	10.729

FONTE: AFUBRA/AGEFLOR, FEPAM, RDK Logs, SEMA (2017).

FIGURA 13 - MAPEAMENTO AGEFLOR – O SETOR DE BASE FLORESTAL NO RIO GRANDE DO SUL 2020 – ÁREA PLANTADA POR COREDE

COREDE	PINUS	EUCALIPTO	ACÁCIA	TOTAL
Sul	54.346	89.497	28.221	172.064
Centro Sul	12.044	106.584	8.909	127.536
Vale do Rio Pardo	19.502	80.351	16.157	116.010
Hortênsias	84.836	1.665	41	86.543
Vale do Taquari	468	64.067	690	65.226
Campos de Cima da Serra	56.547	2.075	0	58.623
Fronteira Oeste	1.669	54.606	0	56.275
Litoral	32.641	17.612	0	50.253
Metropolitano Delta do Jacuí	365	38.210	10.186	48.761
Campanha	286	40.411	6.614	47.311

FONTE: AGEFLOR, FEPAM, Codex, RDK Logs (2020)

3.4.7 FAMATO

A publicação do Diagnóstico de florestas plantadas do Estado do Mato Grosso de 2013, realizada pela Federação da Agricultura e Pecuária do Estado do Mato Grosso (FAMATO), em conjunto com o Instituto Mato-Grossense de Economia e Agropecuária (IMEA), disponibilizou dados de área das florestas plantadas por município no estado do Mato Grosso para o ano de 2012. Os plantios identificados foram de eucalipto e teca.

O diagnóstico pode ser encontrado no *site* da Associação de Reflorestadores do Mato Grosso (AREFLORESTA), disponibilizado no formato PDF. A FIGURA 14 traz as primeiras dez linhas da tabela de dados contida no relatório.

FIGURA 14 - MAPEAMENTO FAMATO/IMEA - DIAGNÓSTICO DE FLORESTAS PLANTADAS NO ESTADO DO MATO GROSSO 2013

Região	Município	Área de Eucalipto (ha)	Área de Teca (ha)	Total (ha)
Nordeste	ÁGUA BOA	302,17	5.241,23	5.543,41
Norte	ALTA FLORESTA		3.724,70	3.724,70
Sudeste	ALTO ARAGUAIA	11.105,71	415,76	11.521,47
Sudeste	ALTO GARÇAS	951,30	85,58	1.036,88
Sudeste	ALTO TAQUARI	928,78		928,78
Sudeste	ARAGUAINHA	82,23		82,23
Centro-Sul	ARENÁPOLIS		272,81	272,81
Centro-Sul	BARRA DO BUGRES		7.564,63	7.564,63
Sudeste	BARRA DO GARÇAS	3.765,38	392,79	4.158,17
Nordeste	BOM JESUS DO ARAGUAIA	82,56		82,56

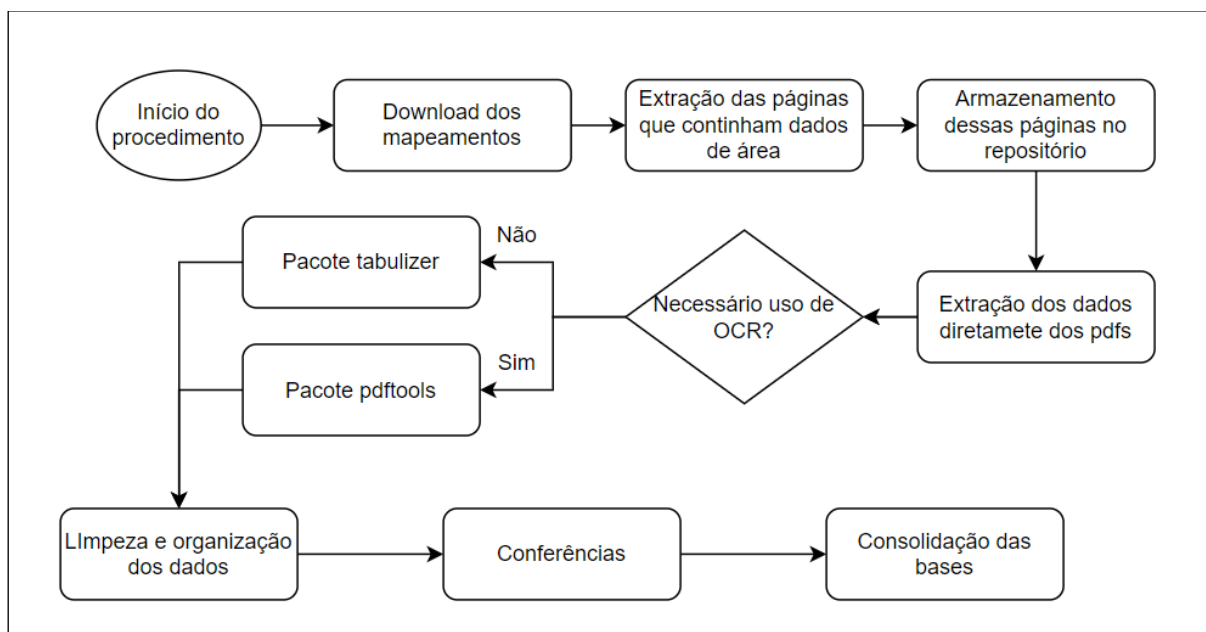
FONTE: IMEA (2013).

3.5 EXTRAÇÃO E ORGANIZAÇÃO DOS DADOS

A extração dos dados seguiu uma rotina semelhante entre os mapeamentos e relatórios utilizados. Em sua grande maioria, os dados dos arquivos em PDF foram extraídos usando a função **tabulizer::extract_tables**, do pacote *tabulizer*, comando que tem como objetivo extrair em formato tabular uma tabela de um PDF. O requisito é que a tabela esteja efetivamente formatada como *tabela* dentro do arquivo, e não como uma imagem.

Para os casos onde o pacote *tabulizer* não pode ser utilizado, por razões de baixa capacidade de extração ou devido a tabela com os dados ser uma imagem, foi utilizado a função **pdf_tools::pdf_ocr_text**, do pacote *pdf_tools*, que emprega o algoritmo *Optimal Character Recognition* (OCR) para identificar, por meio de inteligência artificial, quais são os caracteres textuais contidos em uma imagem. O resultado desse procedimento gera um texto, onde os dados de área acabam sendo misturados com o texto da página. A limpeza de dados realizada nesses casos foi extensa e pode ser observada no script “/data-raw/06-ext-fax-mapeamento-ageflor”, do repositório de extração e faxina de dados, no seguinte link: <https://github.com/maykongpedro/2021-07-04-extracao-mapeamentos-plantios-florestais>. A FIGURA 15 resume as etapas de extração em um esquema visual de fácil compreensão:

FIGURA 15 - FLUXOGRAMA: ETAPAS DA EXTRAÇÃO E LIMPEZA DE DADOS



FONTE: O autor (2021).

As bases consolidadas foram divididas em duas: uma que contém os mapeamentos com detalhe de área por município e outra com os mapeamentos a nível estadual/regional. Para as bases com município, foram adicionados os códigos únicos para cada município de acordo com o IBGE, facilitando assim possíveis consultas e análises. As duas bases finais serviram como *input* no pacote final desenvolvido como resultado deste trabalho.

Todo o código gerado nessa etapa de extração e limpeza pode ser consultado no repositório do autor organizado para esse processo, como citado anteriormente. As etapas de download, extração e organização dos dados podem ser reproduzidas por qualquer usuário interessado.

3.6 CRIAÇÃO E DESENVOLVIMENTOS DE PACOTES

Para criação de pacotes na linguagem R, foram utilizados dois pacotes principais que auxiliam nesse processo: *devtools* e *usethis*. O *devtools* é a base do desenvolvimento dentro do R, pois com ele é possível criar pacotes e ferramentas com a linguagem. O comando **`devtools::create_package("NOME_DO_PACOTE")`** gera um pacote novo na máquina do usuário, após isso, basta seguir as boas práticas de desenvolvimento consolidadas por Wickham

(2015), em seu livro “*R packages*“. O *usethis* fornece recursos para organizar e manipular o pacote, criando scripts e pastas dentro do repositório.

Os pacotes criados pelo *devtools* podem ou não ser enviados ao CRAN, sendo a organização responsável por distribuir pacotes estáveis à comunidade R. Para este trabalho, foi optado pelo desenvolvimento de um pacote com documentação e funções em português, o que contraria um dos pré-requisitos para submissão ao CRAN, para enviar seria necessário existir duas versões, uma em inglês e outra em português. Outro quesito é o tempo necessário de resposta, visto que o pacote era a última etapa deste estudo, a aprovação e correções exigidas para uma submissão não seriam possíveis de serem realizadas dentro do prazo. Assim sendo, após o seu desenvolvimento, o pacote foi hospedado no GitHub, local onde os pacotes podem ser facilmente instalados pelo comando: **`devtools::install_github("nomedousuario/nomedopacote")`**.

4 RESULTADOS E DISCUSSÃO

4.1 PACOTE R: PLANTIOSFLORESTAIS

O pacote foi definido como “plantiosflorestais”, em que seu repositório oficial se encontra no GitHub do autor deste trabalho pelo seguinte endereço: <https://github.com/maykongpedro/plantiosflorestais>. O pacote pode ser instalado utilizando as seguintes linhas de comando (FIGURA 16), com o auxílio do pacote *devtools*.

FIGURA 16 - COMANDOS: INSTALAÇÃO DO PACOTE

```
1. # install.packages("devtools")  
2. devtools::install_github("maykongpedro/plantiosflorestais")
```

FONTE: O autor (2020)

4.1.1 Organização dos arquivos

O repositório foi organizado visando as boas práticas de desenvolvimento, tendo a distribuição de arquivos e pastas exibidas na FIGURA 17.

FIGURA 17 - ORGANIZAÇÃO DO PACOTE

```

+-- data
|   +-- mapeamentos_estados.rda
|   \-- mapeamentos_municipios.rda
+-- data-raw
|   +-- mapeamentos_gerais.rds
|   +-- mapeamentos_municipios.rds
|   \-- obter_dados.R
+-- DESCRIPTION
+-- LICENSE.md
+-- man
|   +-- exportar_csv.Rd
|   +-- exportar_xlsx.Rd
|   +-- generos_plantios_disponiveis.Rd
|   +-- mapeamentos_disponiveis.Rd
|   +-- mapeamentos_estados.Rd
|   +-- mapeamentos_municipios.Rd
|   +-- mapeamento_existente_uf.Rd
|   +-- pipe.Rd
|   +-- plotar_historico_iba.Rd
|   \-- serie_historicas_disponiveis.Rd
+-- NAMESPACE
+-- plantiosflorestais.Rproj
+-- R
|   +-- data.R
|   +-- explorar_mapeamentos.R
|   +-- exportar_dados.R
|   +-- graficos.R
|   \-- utils-pipe.R
+-- README.md
\-- README.Rmd

```

FONTE: O autor (2021).

Os itens em azul na FIGURA 17 representam as pastas do projeto, enquanto os verdes expressam os scripts. Dentro da pasta “data” foram armazenadas as bases oficiais usadas pelo pacote. Na pasta “data-raw” constam as duas bases importadas do projeto de extração e limpeza de dados, o script “obter_dados.R” contém os códigos utilizados para importação. Os arquivos “.rd” contidos na pasta “man” são referentes à documentação das funções existentes no pacote. Eles são gerados automaticamente ao se rodar a função **devtools::check()** ao longo do desenvolvimento do projeto, desde que a função esteja corretamente padronizada, seguindo as recomendações de Wickham (2015), do livro “R packages”. Na pasta “R” estão presentes os scripts que armazenam as funções que serão utilizadas pelo usuário final, em que todas as funções precisam ser documentadas e organizadas.

Essa estrutura organizacional só pode ser encontrada diretamente no repositório oficial. Uma vez que o usuário instale o pacote, é importante apenas fazer o devido uso das funções nele contidas.

4.1.2 Acessar mapeamentos

Conforme descrito no tópico de [3.5 EXTRAÇÃO E ORGANIZAÇÃO DOS DADOS](#), os mapeamentos foram divididos em duas bases principais. Uma contém os relatórios que apresentam dados em nível municipal e a outra possui dados em níveis regional ou estadual. Ambas podem ser acessas diretamente pelo pacote por meio dos comandos demonstrados na FIGURA 18.

FIGURA 18 - COMANDOS: ACESSO ÀS BASES

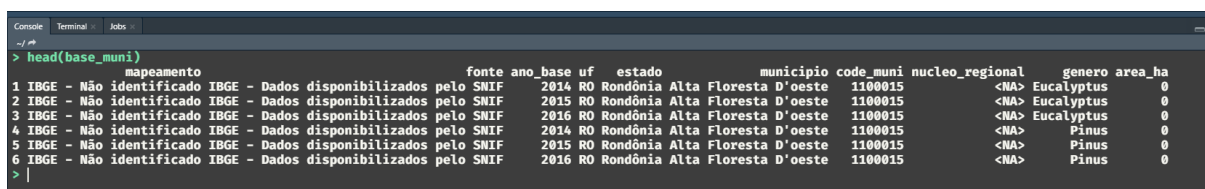
```
1. # Acessar base de municípios
2. plantiosflorestais::mapeamentos_municipios
3.
4. # Acessar base de estados
5. plantiosflorestais::mapeamentos_estados
```

FONTE: O autor (2021).

O retorno desses comandos são duas tabelas de dados do R, conhecidas como *tibbles*, que podem ser alocadas em vetores que facilitem o uso do usuário, como no código: **base_muni <- plantiosflorestais::mapeamentos_municipios**, onde o vetor “base_muni” recebe os mapeamentos de municípios. Assim, o usuário pode manipular a base sem precisar acessar novamente o comando do pacote.

Para ter uma ideia geral dos dados, o comando **head(base_muni)** retorna as primeiras linhas da base de mapeamentos de municípios (FIGURA 19). Outro comando interessante para entender a base é o **str(base_muni)**, que retorna a estrutura de dados das colunas existentes na base, como demonstrado na FIGURA 20.

FIGURA 19 - PRIMEIRAS LINHAS DA BASE DE MUNICÍPIOS

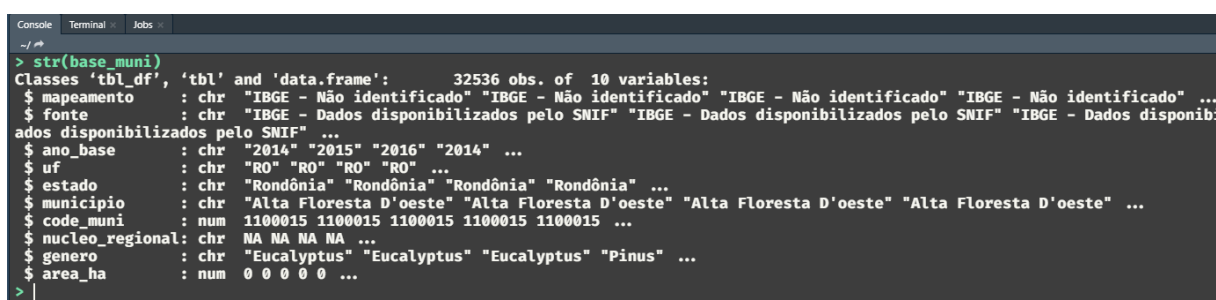


```
> head(base_muni)
```

	mapeamento	fonte	ano_base	uf	estado	municipio	code_muni	nucleo_regional	genero	area_ha
1	IBGE - Não identificado	IBGE - Dados disponibilizados pelo SNIF	2014	RO	Rondônia	Alta Floresta D'oeste	1100015	<NA>	Eucalyptus	0
2	IBGE - Não identificado	IBGE - Dados disponibilizados pelo SNIF	2015	RO	Rondônia	Alta Floresta D'oeste	1100015	<NA>	Eucalyptus	0
3	IBGE - Não identificado	IBGE - Dados disponibilizados pelo SNIF	2016	RO	Rondônia	Alta Floresta D'oeste	1100015	<NA>	Eucalyptus	0
4	IBGE - Não identificado	IBGE - Dados disponibilizados pelo SNIF	2014	RO	Rondônia	Alta Floresta D'oeste	1100015	<NA>	Pinus	0
5	IBGE - Não identificado	IBGE - Dados disponibilizados pelo SNIF	2015	RO	Rondônia	Alta Floresta D'oeste	1100015	<NA>	Pinus	0
6	IBGE - Não identificado	IBGE - Dados disponibilizados pelo SNIF	2016	RO	Rondônia	Alta Floresta D'oeste	1100015	<NA>	Pinus	0

FONTE: O autor (2021).

FIGURA 20 - ESTRUTURA DE DADOS DA BASE DE MUNICÍPIOS



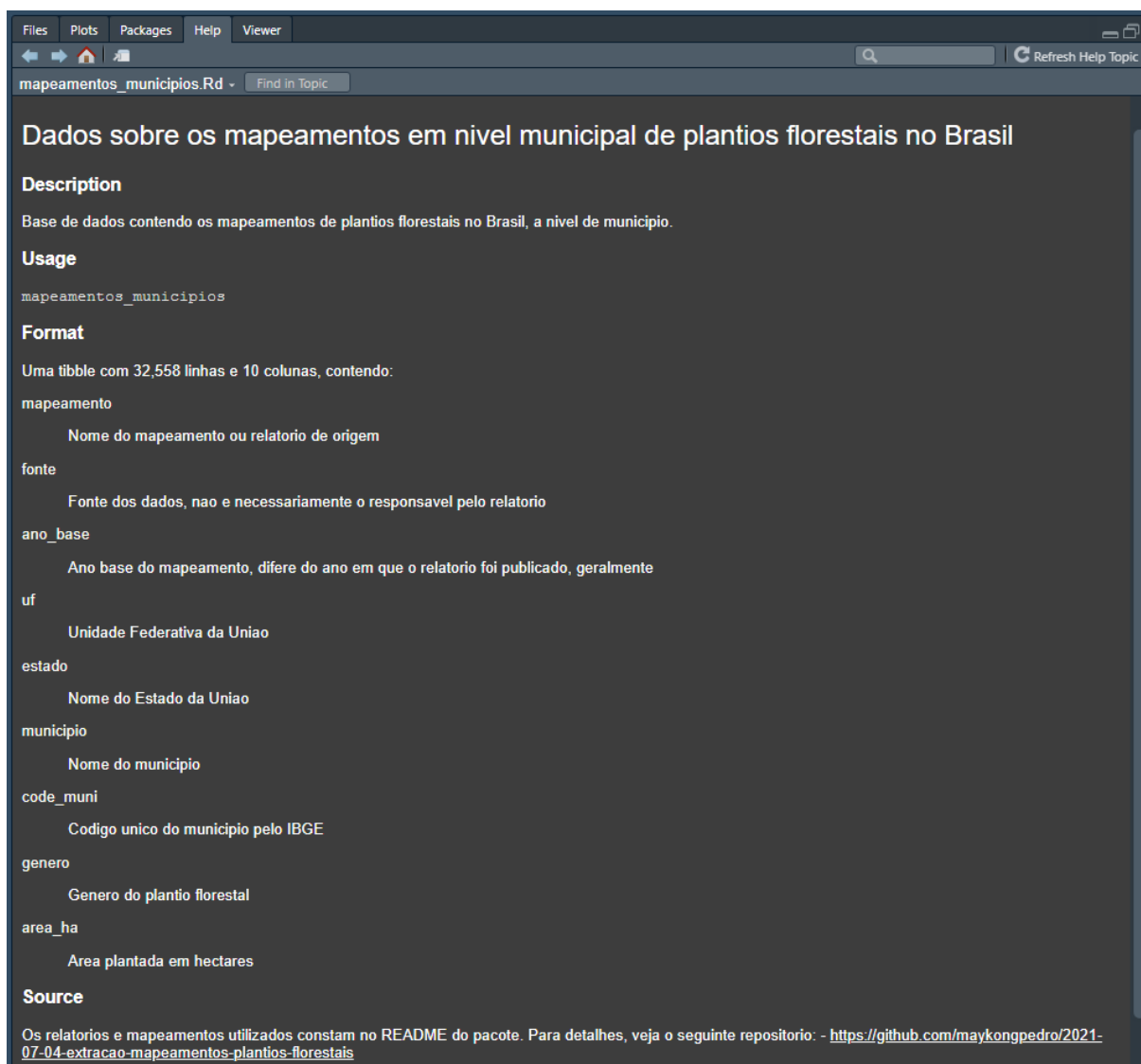
```
> str(base_muni)
```

```
Classes 'tbl_df', 'tbl' and 'data.frame': 32536 obs. of 10 variables:
 $ mapeamento : chr "IBGE - Não identificado" "IBGE - Não identificado" "IBGE - Não identificado" "IBGE - Não identificado" ...
 $ fonte : chr "IBGE - Dados disponibilizados pelo SNIF" "IBGE - Dados disponibilizados pelo SNIF" "IBGE - Dados disponibilizados pelo SNIF" "IBGE - Dados disponibilizados pelo SNIF" ...
 $ ano_base : chr "2014" "2015" "2016" "2014" ...
 $ uf : chr "RO" "RO" "RO" "RO" ...
 $ estado : chr "Rondônia" "Rondônia" "Rondônia" "Rondônia" ...
 $ municipio : chr "Alta Floresta D'oeste" "Alta Floresta D'oeste" "Alta Floresta D'oeste" "Alta Floresta D'oeste" ...
 $ code_muni : num 1100015 1100015 1100015 1100015 ...
 $ nucleo_regional: chr NA NA NA NA ...
 $ genero : chr "Eucalyptus" "Eucalyptus" "Eucalyptus" "Pinus" ...
 $ area_ha : num 0 0 0 0 ...
```

FONTE: O autor (2021).

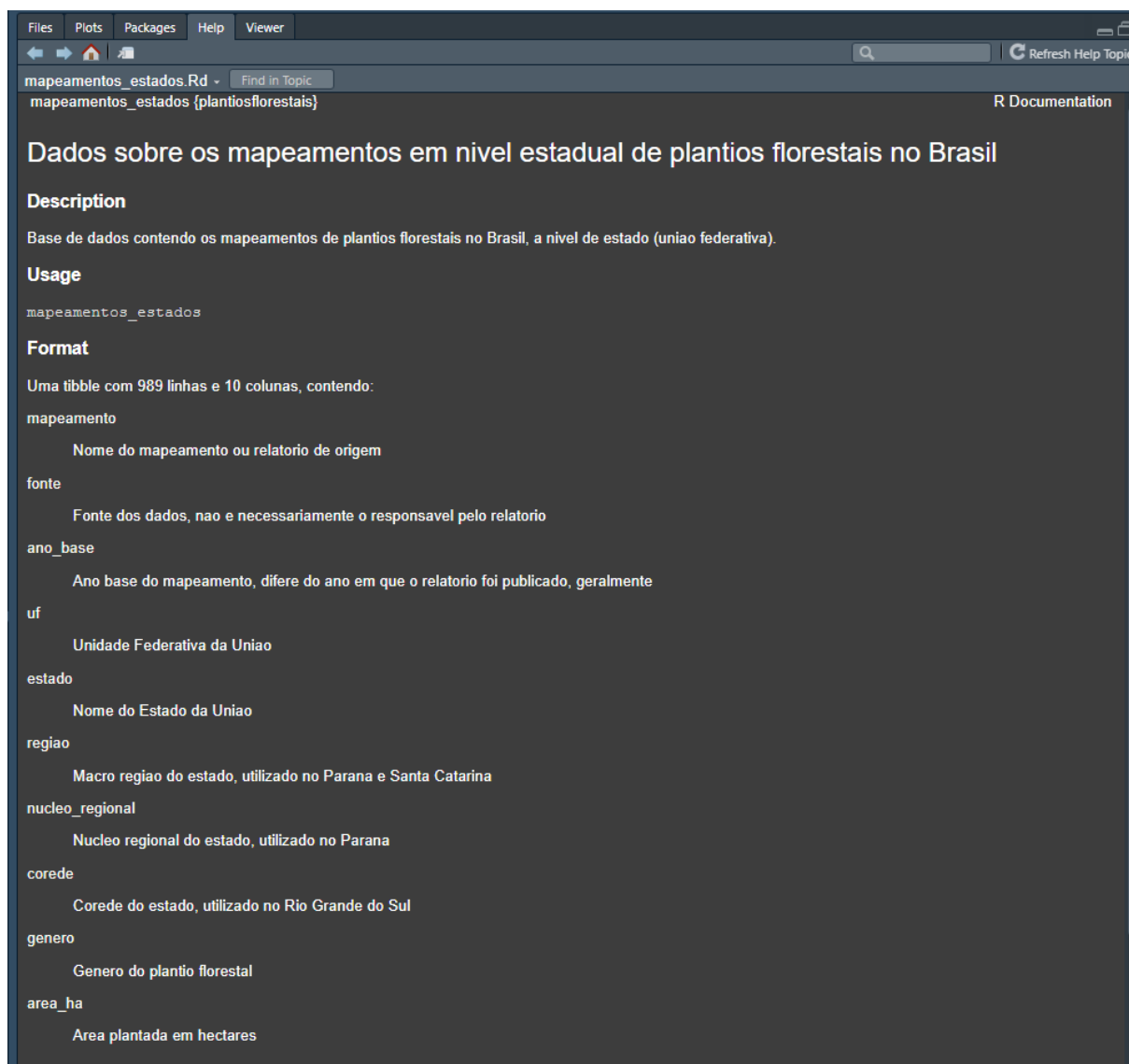
A documentação das duas bases são armazenadas no script “data.R”, dentro da pasta “R” do pacote. Um detalhe da documentação é que como ela é padronizada para o inglês, o texto não pode receber acentuação, então não é possível colocar as palavras em português gramaticalmente correto. Para carregar as documentações de forma rápida e prática, basta colocar um ponto de interrogação na frente dos comandos apresentados na FIGURA 18, ficando, por exemplo, da seguinte maneira: **?plantiosflorestais::mapeamentos_municipios**. O resultado desse comando é exibido no painel *Help* do RStudio, que pode ser verificado na FIGURA 21. A FIGURA 22 apresenta a documentação da base de mapeamentos estaduais.

FIGURA 21 - DOCUMENTAÇÃO: BASE DE DADOS DE MUNICÍPIOS



FONTE: O autor (2021).

FIGURA 22 - DOCUMENTAÇÃO: BASE DE DADOS DE ESTADOS



FONTE: O autor (2021).

4.1.3 Funções

O pacote conta com sete funções simples que visam facilitar o uso de usuários iniciantes na Linguagem R. Todas as funções são devidamente documentadas, e o *Help* pode ser acessado da mesma maneira que a documentação das bases, usando o ponto de interrogação antes da função. Elas foram armazenadas em três diferentes scripts que constam na pasta “R”. A TABELA 2 resume cada script e as respectivas funções contidas nele.

TABELA 2 - FUNÇÕES POR SCRIPT

Script	Funções
explorar_mapeamentos	mapeamentos_disponiveis() mapeamento_disponivel_uf() generos_plantios_disponiveis() serie_historicas_disponiveis()
exportar_dados	exportar_xlsx() exportar_csv()
graficos	plotar_historico_iba()

FONTE: O autor (2021).

As funções podem ser usadas utilizando diretamente a sintaxe do comando em R **nome_do_pacote::nome_da_função()**, ou, se aplicado o comando **library("nome_do_pacote")** para carregar o pacote na sessão ativa, basta acessar no script ou no console o nome da função desejada: **nome_da_função()**. Os tópicos a seguir trazem uma descrição das funcionalidades de cada uma das funções desenvolvidas.

4.1.3.1 Mapeamentos disponíveis

Para descobrir quais os mapeamentos existentes no pacote sem necessidade de manipulações adicionais de base, basta acessar a função: **mapeamentos_disponiveis ()**, cujo resultado exibido no console é apresentado na FIGURA 23.

FIGURA 23 - FUNÇÃO: MAPEAMENTOS DISPONÍVEIS

	base	mapeamento	fonte
1	mapeamentos_municipios	IBGE - Não identificado	IBGE - Dados disponibilizados pelo S~
2	mapeamentos_municipios	Famato - Diagnóstico de florestas plantadas do Estado de Mato Grosso	Imea
3	mapeamentos_municipios	IFPR - Mapeamento dos Plantios Florestais do Estado do Paraná	IFPR e SFB
4	mapeamentos_municipios	AGEFLOR - A indústria de base florestal no Rio Grande do Sul 2017	AFUBRA, AGEFLOR, FEPAM, RDK e SEMA
5	mapeamentos_municipios	AGEFLOR - O setor de base florestal no Rio Grande do Sul 2020	Fepam, Codex, RDK e AGEFLOR
6	mapeamentos_estados	IBA - Relatório Anual 2020	Pöyry e IBA
7	mapeamentos_estados	IBA - Relatório Anual 2020	FGV e IBA
8	mapeamentos_estados	IBA - Não identificado	IBA - Dados disponibilizados pelo SN~
9	mapeamentos_estados	IBGE e PEVS - 2018	PEVS/IBGE
10	mapeamentos_estados	APRE - Estudo Setorial 2020	UFPR e APRE
11	mapeamentos_estados	AGEFLOR - A indústria de base florestal no Rio Grande do Sul 2017	AFUBRA, AGEFLOR, FEPAM, RDK e SEMA
12	mapeamentos_estados	AGEFLOR - O setor de base florestal no Rio Grande do Sul 2020	Fepam, Codex, RDK e AGEFLOR
13	mapeamentos_estados	ACR - Anuário Estatístico 2019	UDESC-CAV e ACR

FONTE: O autor (2021).

O retorno é uma tabela (*tibble*) com três colunas, informando, respectivamente, a qual das duas bases o mapeamento se refere, qual o nome do relatório ou mapeamento e qual a fonte dos dados contidos no mesmo.

4.1.3.2 Gêneros de plantios disponíveis

Pode-se considerar um complemento à função anterior, o comando **generos_plantios_disponiveis()** retorna quais os gêneros dos plantios florestais existentes nas bases utilizadas. Para essa função, pode ser declarado um argumento booleano (*TRUE* ou *FALSE*), onde o usuário decide se deve ser exibido os gêneros existentes por nome de mapeamento ou apenas os gêneros, sem definir qual o mapeamento. O comando tem como padrão o argumento *FALSE*, para exibir somente o resultado simplificado. A FIGURA 24 exibe os dois resultados possíveis, sendo o resultado padrão e depois o resultado declarando o argumento como *TRUE*.

FIGURA 24 - FUNÇÃO: GÊNEROS DE PLANTIOS DISPONÍVEIS

```

C:/Users/mayko/OneDrive/Linguagem R/06. Projetos/plantiosflorestais/
> plantiosflorestais::generos_plantios_disponiveis()
# A tibble: 6 x 1
  genero
  <chr>
1 Eucalyptus
2 Pinus
3 Outros
4 Tectona
5 Corte
6 Acácia
> plantiosflorestais::generos_plantios_disponiveis(exibir_nome_mapeamento = TRUE)
# A tibble: 28 x 2
  mapeamento                                genero
  <chr>                                <chr>
1 IBGE - Não identificado                  Eucalyptus
2 IBGE - Não identificado                  Pinus
3 IBGE - Não identificado                  Outros
4 Famato - Diagnóstico de florestas plantadas do Estado de Mato Grosso - 2013 Eucalyptus
5 Famato - Diagnóstico de florestas plantadas do Estado de Mato Grosso - 2013 Tectona
6 IFPR - Mapeamento dos Plantios Florestais do Estado do Paraná               Corte
7 IFPR - Mapeamento dos Plantios Florestais do Estado do Paraná               Eucalyptus
8 IFPR - Mapeamento dos Plantios Florestais do Estado do Paraná               Pinus
9 AGEFLOR - A indústria de base florestal no Rio Grande do Sul 2017          Pinus
10 AGEFLOR - A indústria de base florestal no Rio Grande do Sul 2017          Eucalyptus
# ... with 18 more rows
> |

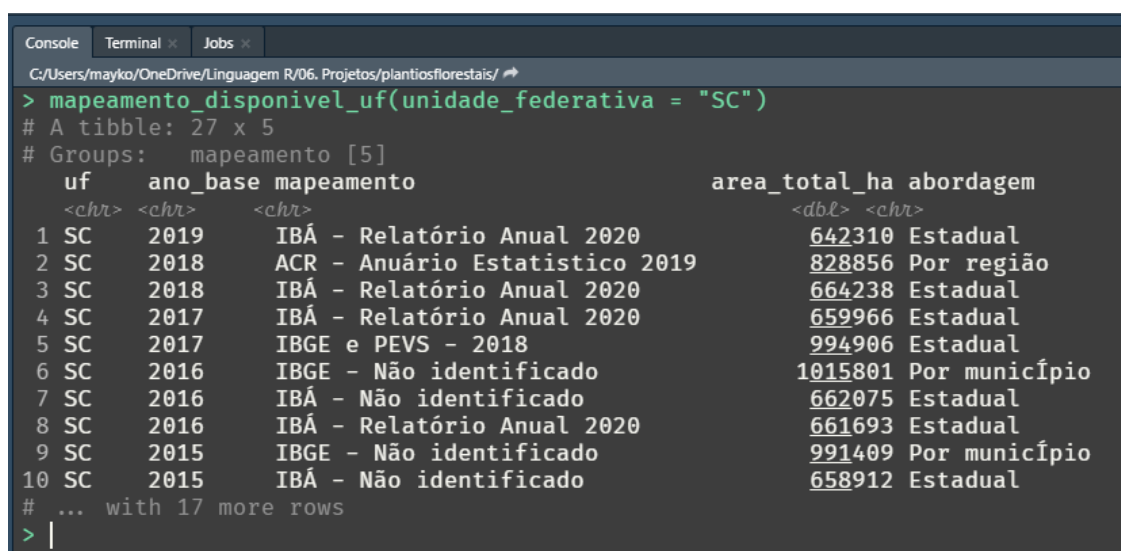
```

FONTE: O autor (2021).

4.1.3.3 Mapeamento disponível por unidade federativa em área

Tem como objetivo exibir a área dos mapeamentos existentes nas bases para uma determinada Unidade Federativa (UF) do Brasil para cada ano-base. O argumento necessário é a definição da sigla da UF escolhida em caracteres caixa alta. Caso o usuário não defina uma UF, o padrão da função é UF igual a PR, ou seja, ela vai resultar na área dos mapeamentos disponíveis para o estado do Paraná. Para acessá-la basta usar o comando **mapeamento_disponivel_uf("UF")**. A FIGURA 25 apresenta o resultado para o estado de Santa Catarina, uma *tibble* com quatro colunas: ano base, nome do mapeamento, área total em hectares e nível de abordagem do mapeamento.

FIGURA 25 - FUNÇÃO: MAPEAMENTO DISPONÍVEL EM ÁREA PARA UMA DETERMINADA UF



```

C:/Users/mayko/OneDrive/Linguagem R/06. Projetos/plantiosflorestais/
> mapeamento_disponivel_uf(unidade_federativa = "SC")
# A tibble: 27 x 5
# Groups:   mapeamento [5]
  uf      ano_base mapeamento      area_total_ha abordagem
  <chr>   <chr>    <chr>          <dbl>   <chr>
1 SC     2019      IBÁ - Relatório Anual 2020    642310 Estadual
2 SC     2018      ACR - Anuário Estatístico 2019    828856 Por região
3 SC     2018      IBÁ - Relatório Anual 2020    664238 Estadual
4 SC     2017      IBÁ - Relatório Anual 2020    659966 Estadual
5 SC     2017      IBGE e PEVS - 2018          994906 Estadual
6 SC     2016      IBGE - Não identificado    1015801 Por município
7 SC     2016      IBÁ - Não identificado     662075 Estadual
8 SC     2016      IBÁ - Relatório Anual 2020     661693 Estadual
9 SC     2015      IBGE - Não identificado     991409 Por município
10 SC    2015      IBÁ - Não identificado     658912 Estadual
# ... with 17 more rows
> |

```

FONTE: O autor (2021).

4.1.3.4 Obter séries históricas

Essa função pode ser acessada pelo comando **series_historicas_disponiveis()**, a qual tem como resultado uma *tibble* que resume a área por ano-base e mapeamento apenas dos relatórios que possuem alguma série histórica, isto é, aqueles que possuem mais de um ano-base em seus dados. A intenção é facilitar ao usuário a obtenção rápida de números históricos, visando a análise de progressão das florestas plantadas. O resultado exibido no console do R pode ser observado na FIGURA 26.

FIGURA 26 - FUNÇÃO: SÉRIES HISTÓRICAS DISPONÍVEIS

```

Console Terminal Jobs
C:/Users/mayko/OneDrive/Linguagem R/06. Projetos/plantiosflorestais/
> plantiosflorestais::series_historicas_disponiveis()
# A tibble: 43 x 4
# Groups:   ano_base [14]
   base      mapeamento      ano_base area_total_ha
  <chr>      <chr>      <chr>      <dbl>
1 mapeamentos_municip~ IBGE - Não identificado      2014      9366741
2 mapeamentos_municip~ IBGE - Não identificado      2015      9937947
3 mapeamentos_municip~ IBGE - Não identificado      2016     10023076
4 mapeamentos_estados AGEFLOR - A indústria de base florestal no Rio Grande do~ 2006      508000
5 mapeamentos_estados AGEFLOR - A indústria de base florestal no Rio Grande do~ 2007      563600
6 mapeamentos_estados AGEFLOR - A indústria de base florestal no Rio Grande do~ 2008      638800
7 mapeamentos_estados AGEFLOR - A indústria de base florestal no Rio Grande do~ 2009      582300
8 mapeamentos_estados AGEFLOR - A indústria de base florestal no Rio Grande do~ 2010      531900
9 mapeamentos_estados AGEFLOR - A indústria de base florestal no Rio Grande do~ 2011      534100
10 mapeamentos_estados AGEFLOR - A indústria de base florestal no Rio Grande do~ 2012      539700
# ... with 33 more rows
> |

```

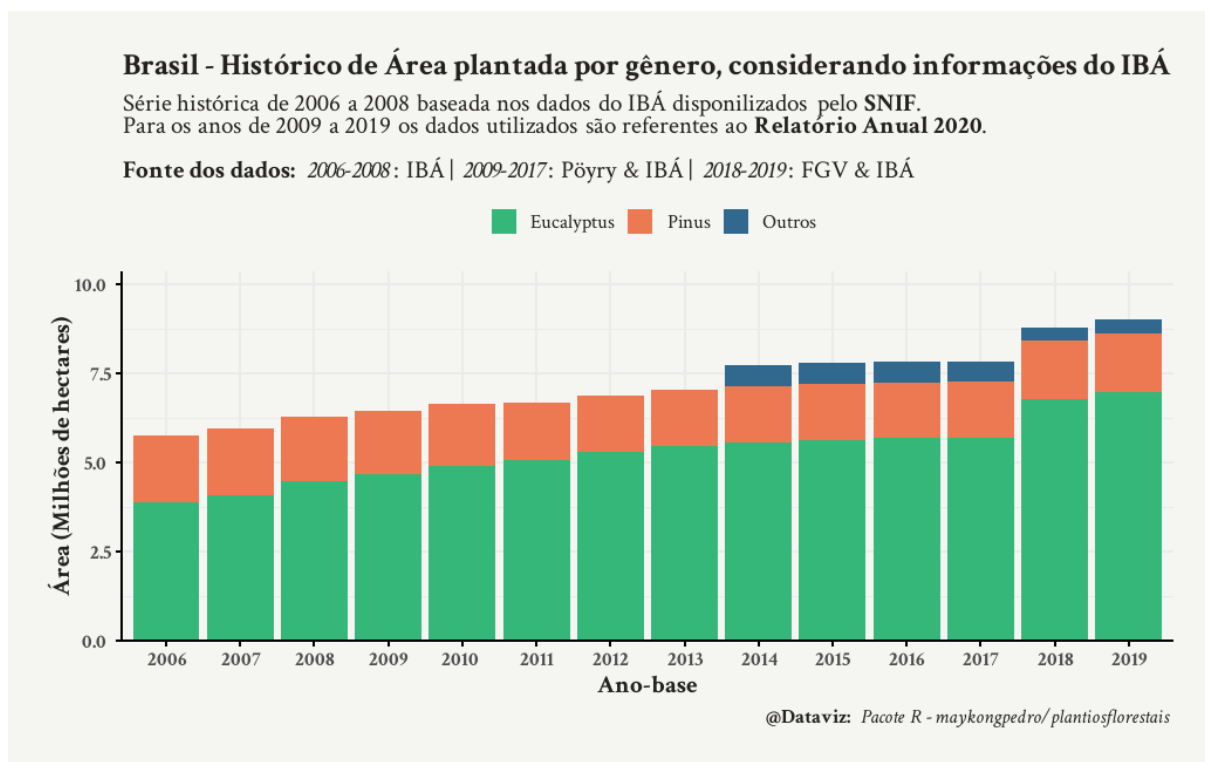
FONTE: O autor (2021).

4.1.3.5 Plotar série histórica do IBÁ

Única função do pacote que retorna um gráfico como resultado, ou seja, um objeto *plot*. O objetivo dessa função é plotar um gráfico de colunas agrupadas por gênero, com área em hectares no eixo y, por ano base das florestas plantadas no eixo x, com os dados da Indústria Brasileira de Árvores (IBÁ), tanto dos disponibilizados no SNIF quanto dos que constam no Relatório Anual 2020.

É utilizada pelo comando **plotar_historico_iba()**, e possui dois argumentos opcionais: um no formato de texto (*string*), chamado de “abrangencia_uf”; e outro lógico (*booleano*), com o nome “exibir_rotulos”. O primeiro define o nível de abrangência desejado para o gráfico, em outras palavras, qual o estado do Brasil o *plot* deve mostrar, é preciso preencher com uma unidade federativa, como “PR”, por exemplo. Se não for definida nenhuma UF, o gráfico é plotado com os dados do Brasil como um todo. Já o segundo argumento define se o gráfico deve exibir os rótulos totais para cada coluna agrupada, e deve ser preenchido com *TRUE* ou *FALSE*, tendo como padrão o *FALSE*, para não colocar os rótulos no gráfico. Não há necessidade de declarar os dois argumentos caso se deseje o GRÁFICO 1, que é demonstrado a seguir:

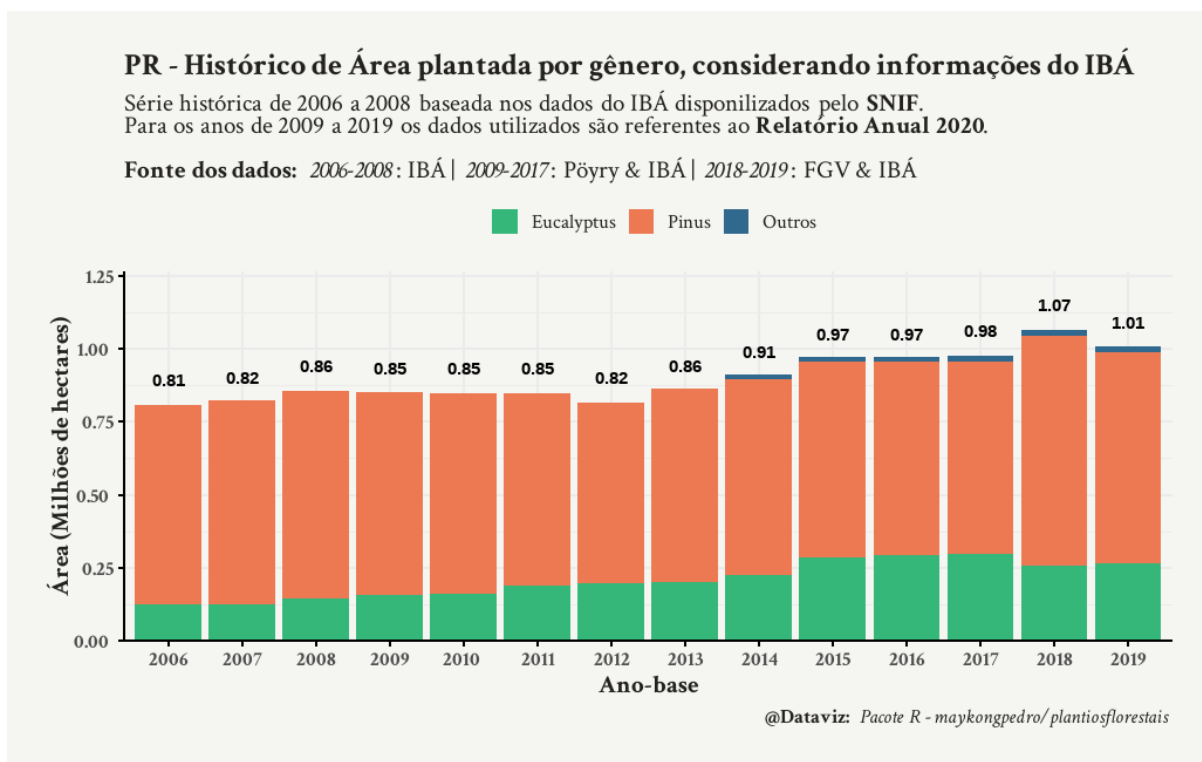
GRÁFICO 1 - FUNÇÃO: PLOTAR MAPEAMENTO IBÁ - BRASIL



FONTE: IBÁ (2020), visualização gerada pelo autor (2021).

O GRÁFICO 2 exibe o resultado para um estado em específico, onde o argumento “abrangencia_uf” foi definido como “PR”, e o “exibir_rotulos” como *TRUE*.

GRÁFICO 2 - FUNÇÃO: PLOTAR MAPEAMENTO IBÁ - PARANÁ



FONTE: IBÁ (2020), visualização gerada pelo autor (2021).

4.1.3.6 Exportar bases em extensão XLSX

A função **exportar_xlsx()** exporta as duas bases contidas no pacote em um arquivo excel, com duas planilhas, uma referente à base com mapeamentos de municípios e outra com a base dos estados. A função tem como argumento obrigatório o “caminho_para_salvar_base”, que consiste no caminho da pasta onde o usuário quer salvar o arquivo xlsx. Esse caminho deve ser necessariamente válido, terminando com uma barra, como no seguinte exemplo: **exportar_xlsx(“C:/Users/mayko/Documents/”)**. Se o usuário quiser salvar no projeto ativo basta definir o caminho como “./”. A FIGURA 27 apresenta os comandos para as duas situações exemplificadas. O retorno da função é invisível, após rodar o console não irá printar nada, contudo, o arquivo será salvo no endereço escolhido.

FIGURA 27 - COMANDOS: EXPORTAR XLSX

```

1. # Salvar em alguma pasta do computador
2. plantiosflorestais::exportar_xlsx("C:/Users/mayko/Documents/")
3.
4. # Salvar no projeto ativo
5. plantiosflorestais::exportar_xlsx("./")
6.
7. # Salvar no projeto ativo, em uma pasta chamada 'dados'
8. plantiosflorestais::exportar_xlsx("./dados/")

```

FONTE: O autor (2021).

4.1.3.7 Exportar bases em extensão CSV

A exportação das bases nessa função é muito semelhante à anterior, a diferença é que a **exportar_csv()** gera dois arquivos distintos em extensão CSV, separados por “;” (ponto e vírgula), com *encoding* UTF-8. Como não é possível aplicar o conceito de “planilhas” dentro dessa extensão, visto que ela é necessariamente um arquivo de texto, no momento da exportação é gerado um arquivo para a base de municípios e outro para a base de estados, ambos no mesmo endereço passado como argumento. A FIGURA 28 resume alguns exemplos do uso da função.

FIGURA 28 - COMANDOS: EXPORTAR CSV

```

1. # Salvar em alguma pasta do computador
2. plantiosflorestais::exportar_csv("C:/Users/mayko/Downloads/")
3.
4. # Salvar no projeto ativo
5. plantiosflorestais::exportar_csv("./")
6.
7. # Salvar no projeto ativo, em uma pasta chamada 'dados'
8. plantiosflorestais::exportar_xlsx("./dados/")

```

FONTE: O autor (2021).

5 CONSIDERAÇÕES FINAIS

O pacote desenvolvido no presente trabalho permite acesso rápido e fácil aos dados de área dos mapeamentos florestais disponíveis no país. Além disso, estando em um ambiente de análise de dados, como a linguagem R, as bases abrem espaço para novas explorações e usos analíticos.

As funções existentes permitem ao usuário obter respostas sobre os dados com maior facilidade, possibilitando também a exportação para ambientes mais amigáveis, como é o caso de planilhas, como *software Microsoft Excel*, onde esses dados podem ser utilizados para criação de dashboards em *PowerBI*, por exemplo. Todas são funções que podem ser replicadas com o conhecimento de outros pacotes disponíveis para a linguagem, entretanto, dentro do contexto de simplificar a entrada de novos usuários, elas podem atuar de forma satisfatória.

Cabe ressaltar que o uso dos dados consolidados nas bases trabalhadas neste projeto não substitui a leitura nem a consulta dos relatórios ou mapeamentos originais. Cada levantamento tem sua peculiaridade e metodologia, e esses detalhes podem ser consultados nos arquivos originais. Ademais, a grande parte deles traz informações riquíssimas sobre o setor florestal, agregando muito no conhecimento do leitor.

Como próximos passos, pode-se citar possíveis extensões de plotagens para outros mapeamentos e eventuais adições advindas da comunidade, além da possibilidade de submissão ao CRAN. Afinal, o pacote está disponível em código aberto no GitHub, permitindo a contribuição de quem estiver interessado.

REFERÊNCIAS

Associação Catarinense de Empresas Florestais (ACR). **Anuário Estatístico de Base Florestal para o estado de Santa Catarina 2019**. Lages, 2019. Disponível em: http://www.acr.org.br/uploads/biblioteca/Anuario_ACR_2019_atualizado.pdf. Acesso em: 17 jun. 2021.

Associação Gaúcha de Empresas Florestais (AGEFLOR). **O Setor de Base Florestal no Rio Grande do Sul 2020**. Porto Alegre, 2020. Disponível em: <http://www.ageflor.com.br/noticias/wp-content/uploads/2020/12/O-Setor-de-Base-Florestal-no-Rio-Grande-do-Sul-2020-ano-base-2019.pdf>. Acesso em: 17 jun. 2021.

Associação Paranaense de Empresas de Base Florestal (APRE). **Estudo Setorial APRE**. Curitiba, 2020. Disponível em: <https://apreflorestas.com.br/publicacoes/estudo-setorial-apre-2020-2/> Acesso em: 17 jun. 2021.

BRANDELERO, C. **Aplicabilidade da tecnologia móvel em atividades de silvicultura de precisão: mapeamento, inventário e geoestatística florestais**. 2007. 116 f. Dissertação (Mestrado em Geomática) - Área de Concentração em Tecnologia da Geoinformação, Universidade Federal de Santa Maria, Santa Maria (RS), 2007. Disponível em: <https://repositorio.ufsm.br/bitstream/handle/1/9595/Catize%20Brandelero.pdf?sequence=1&isAllowed=y>. Acesso em: 25 jun. 2021.

CARDOSO, F. B. **Contribuições para o mapeamento das florestas plantadas de Santa Catarina**. 2017. 112 f. Dissertação (Mestrado em Engenharia Florestal) - Centro de Ciências Tecnológicas, Universidade Regional de Blumenau, Blumenau (SC), 2017. Disponível em: https://bu.furb.br/docs/DS/2017/362205_1_1.pdf. Acesso em: 25 jun. 2021.

CHACON, S.; STRAUB, B. **Pro Git**. Apress, 2014. Disponível em: <https://git-scm.com/book/en/v2>. Acesso em: 24 jul. 2021.

COUTINHO, V. M. *et al.* MAPEAMENTO DAS ÁREAS PLANTADAS DE EUCALYPTUS SPP. NO ESTADO DO PARANÁ. **Biofix Scientific journal**, Curitiba, v. 2 n.1, abril, 2017. Disponível em: <https://revistas.ufpr.br/biofix/article/view/51222/32004>. Acesso em: 25 jun. 2021.

EISFELD, R. L. *et al.* Mapeamento das áreas de Pinus spp. no estado do Paraná. **Advences in Forestry Science**, Cuiabá, v.5, n.3, p.403-409, 2018. Disponível em: <https://periodicoscientificos.ufmt.br/ojs/index.php/afor/article/view/5788>. Acesso em: 25 jun. 2021.

Federação da Agricultura e Pecuária do Estado de Mato Grosso (FAMATO). **Diagnóstico de Florestas Plantadas do Estado de Mato Grosso**. Instituto Mato-

Grossense de Economia Agropecuária (IMEA) – Cuiabá: 2013. Disponível em: <http://www.arefloresta.org.br/uploads/downloads/00072201414739.pdf>. Acesso em: 17 jun. 2021.

Git FAQ. 2021. Disponível em: https://git.wiki.kernel.org/index.php/GitFaq#What_is_Git.3F. Acesso em: 22 jul. 2021.

IHAKA, R.; GENTLEMAN, R. **R: A language for data analysis and graphics**. Journal of Computational and Graphical Statistics. 1996. Disponível em: <https://www.stat.auckland.ac.nz/~ihaka/downloads/R-paper.pdf>. Acesso em: 17 ago.2021.

Indústria Brasileira de Árvores (IBÁ). **Relatório Anual IBÁ**. São Paulo, 2020. Disponível em: <https://www.iba.org/datafiles/publicacoes/relatorios/relatorio-iba-2020.pdf>. Acesso em: 15 jun.2021.

Indústria Brasileira de Árvores (IBÁ). **Setor florestal investe R\$ 32,6 bilhões até 2023**. São Paulo, 2019. Disponível em: <https://www.iba.org/setor-florestal-investe-r-32-6-bilhoes-ate-2023>. Acesso em: 29 jul.2021.

Instituto de Florestas do Paraná (IFPR). **Mapeamento dos plantios florestais do estado do Paraná**. Curitiba, 2015. Disponível em: <https://apreflorestas.com.br/publicacoes/ifpr-e-sfb-mapeamento-dos-plantios-florestais-do-estado-do-parana/>. Acesso em: 17 jun. 2021.

MUENCHEN, R. A. **The Popularity of Data Science Software**. 2012. Disponível em: <http://r4stats.com/articles/popularity/>. Acesso em: 22 jul. 2021.

PERKEL, J. **When it comes to reproducible science, Git is code for success**. Nature index. 2018. Disponível em: <https://www.natureindex.com/news-blog/when-it-comes-to-reproducible-science-git-is-code-for-success>. Acesso em: 22 jul. 2021.

NICOLETTI, A. L. *et al.* Acurácia de mapeamento de florestas nativas e plantadas em Santa Catarina, Sul do Brasil. In: Simpósio Brasileiro de Sensoriamento Remoto, 19., 2019, Santos. **Anais**. Santos: INPE – SANTOS – SP, 2019 p. 700-703. Disponível em: <http://marte2.sid.inpe.br/col/sid.inpe.br/marte2/2019/09.06.13.58/doc/97345.pdf>. Acesso em: 23 jul. 2021.

R Core Team. **R: A language and environment for statistical computing**. R Foundation for Statistical Computing. Vienna, Austria, 2020. Disponível em: <https://www.R-project.org/>.

R Core Team. **A Free Software Project**. Disponível em: https://cran.r-project.org/doc/html/interface98-paper/paper_2.html. Vienna, Austria, 2021. Acesso em: 22 jul. 2021.

SANTOS, K. A. **Os IDE's (Ambientes de Desenvolvimento Integrado) como ferramentas de trabalho em informática**. Universidade Federal de Santa Maria (UFSM). 2007. Disponível em: <http://www-usr.inf.ufsm.br/~alexks/elc1020/artigo-elc1020-alexks.pdf>. Acesso em: 23 jul. 2021.

Serviço Florestal Brasileiro (SFB). **Florestas do Brasil em resumo**. Brasília, 2019. Disponível em: http://www.acr.org.br/uploads/biblioteca/Florestas_Brasil_2019_Portugues.pdf. Acesso em: 15 jun. 2021.

Sistema Nacional de Informações Florestais (SNIF). **Florestas Plantadas - Iba - 2006-2016**. 2018. Disponível em: [Sistema Nacional de Informações Florestais - SNIF - Florestas Plantadas- Iba -2006-2016.csv - Portal Brasileiro de Dados Abertos](#). Acesso em: 15 jun. 2021.

Sistema Nacional de Informações Florestais (SNIF). **Florestas Plantadas - IBGE - 2014-2016**. 2018 Disponível em: [Sistema Nacional de Informações Florestais - SNIF - Florestas Plantadas - IBGE - 2014-2016.csv - Portal Brasileiro de Dados Abertos](#). Acesso em: 15 jun. 2021.

TIOBE software BV. **TIOBE index**. 2021. Disponível em: [index | TIOBE - The Software Quality Company](#). Acesso em 29 jul.2021.

WICKHAM, H. **R Packages: Organize, Test, Document, and Share Your Code (English Edition)**. O'Reilly Media, 2015. Disponível em: [Welcome | R for Data Science \(had.co.nz\)](#). Acesso em: 20 jun. 2021.

WICKHAM, H.; GROLEMUND, G. **R for Data Science**. Alta Books, 2019. Disponível em: [Welcome! | R Packages \(r-pkgs.org\)](#). Acesso em: 05 jul.2021.