

Curso de Vigilância em Saúde apoiado por plataforma BI

Aula 01- Introdução ao R

Apresentação

Olá!

Essa é a terceira aula do Curso de Vigilância em Saúde apoiado por plataforma BI. Nesse curso faremos análises de dados temporais dos Sistemas de Informação em Saúde do Sistema Único de Saúde com apoio da linguagem R e sua interface de desenvolvimento, o RStudio.

Essas ferramentas foram escolhidas por causa do alto potencial de análise que possuem, permitindo a análise de grandes volumes de dados, além do uso e implementação de diversas funções de análise estatística e epidemiológica. Além disso, os softwares são livres e de código aberto, ou seja, permite que qualquer pessoa na comunidade inclua novas funções para a linguagem.

Nesse momento, aprenderemos a instalar a linguagem e sua interface de desenvolvimento para o uso nas próximas aulas. Desejamos sucesso nessa etapa de aprendizado.

1. O que é o R?

A linguagem R é uma ferramenta para a análise estatística e criação de visualizações. Inicialmente, o R foi criado como uma linguagem muito semelhante à antiga linguagem S e foi desenvolvida na *Bell Laboratories* por John Chambers, um estatístico da companhia, e outros colegas (1).

Como linguagem, há alguns desafios na compreensão do R, principalmente considerando que as graduações da área da saúde, em geral, não incorporam a programação como uma de suas disciplinas obrigatórias. Nesse sentido, o R é uma linguagem de alto nível orientada a objetos, o que permite seu aprendizado de forma mais rápida quando comparada a outras linguagens de programação (2).

Atualmente, lidamos com um grande volume de dados de saúde, sendo que a análise por meio de planilhas eletrônicas nem sempre é possível. São dados de notificações, atendimento ambulatorial, internação, nascimento e mortalidade, entre outros. Assim, o domínio de ferramentas de programação se torna cada vez mais essencial para a área da saúde.

Enquanto ferramenta, o R permite a criação de análises descritivas, analíticas e inferenciais, além da criação de gráficos que podem ajudar a compor uma análise de situação de saúde. O potencial para processar volumes de dados maiores que as planilhas eletrônicas também é uma vantagem. Além disso, é possível personalizar o R com qualquer função que esteja disponível no CRAN (Comprehensive R Archive Network [rede de arquivos do R, em tradução livre]).

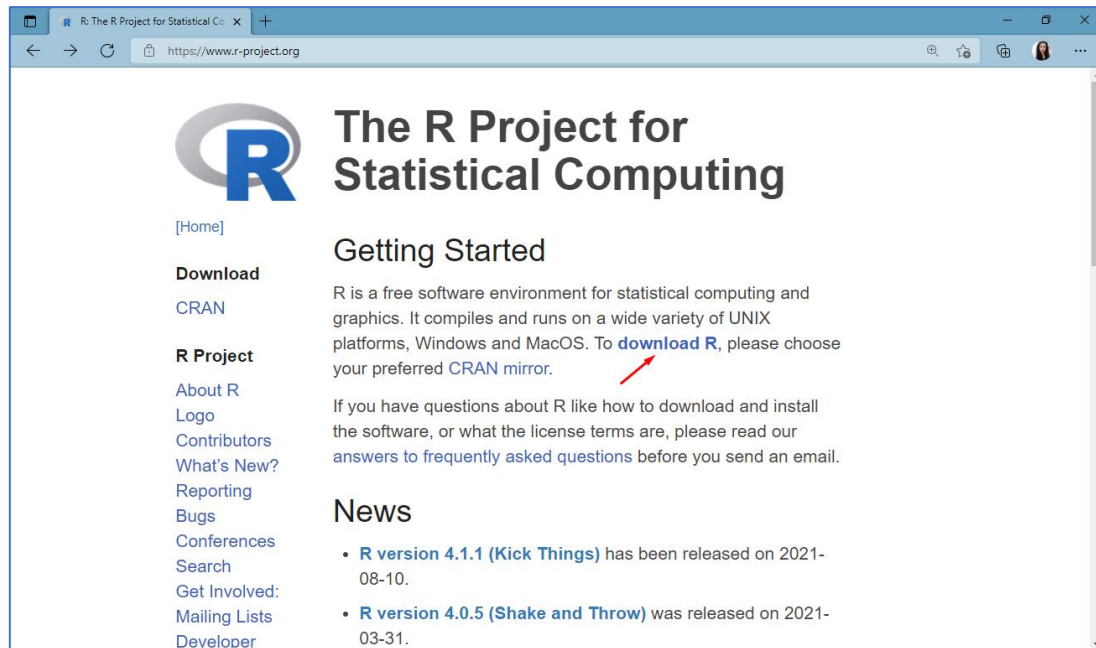
As funções do R podem ser instaladas por meio de pacotes, que são arquivos com um conjunto de funções destinadas a uma determinada temática. Além disso, é possível criar funções definindo isso por meio do R. A instalação de pacotes será abordada nessa aula, mas as funções de criação não serão abordadas nesse curso.

Para iniciar o curso, será necessário que instale o R, o Rtools e o RStudio. Todos os programas são gratuitos e serão necessários para o acompanhamento das próximas aulas.

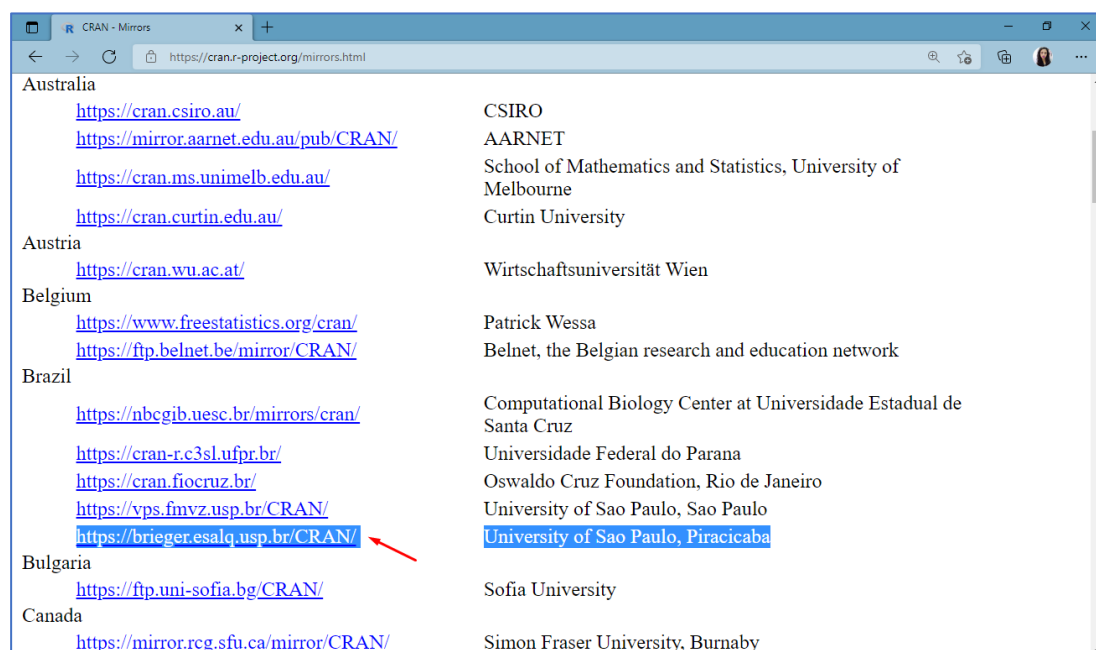
2. Instalando a linguagem R

A linguagem R está em sua 4ª versão e é possível instalar por meio do *download* de seu instalador disponível de forma gratuita por meio do *site* oficial da linguagem que possui repositórios em diversos países. Acesso o *site* aqui [R: The R Project for Statistical Computing \(r-project.org\)](https://www.R-project.org/).

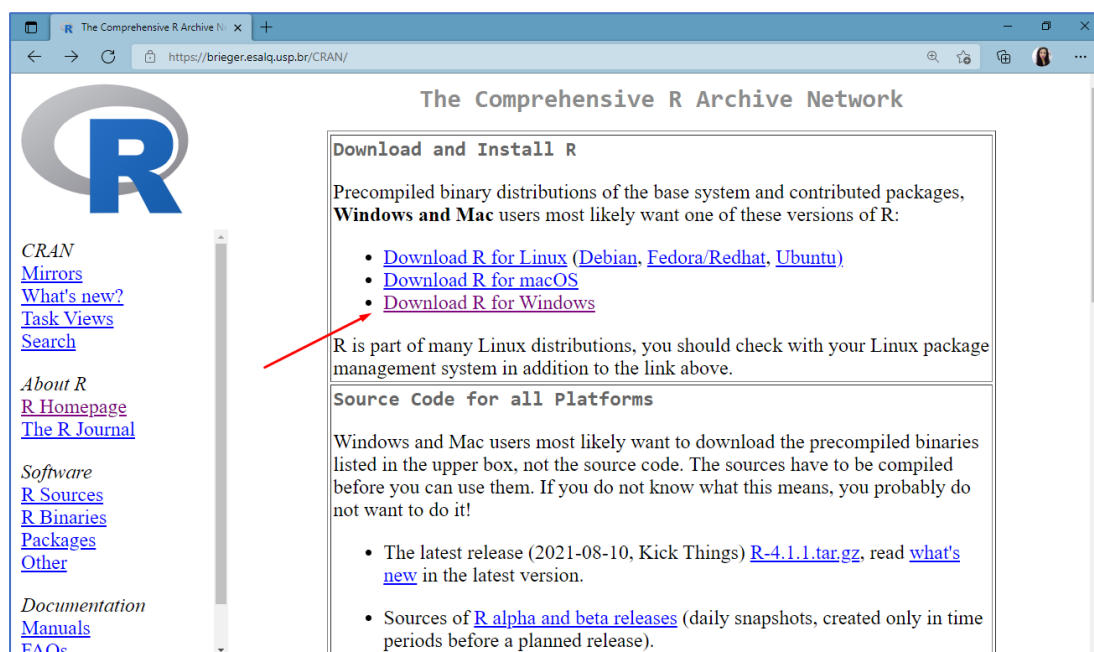
Para instalar o programa, acesse o link indicado anteriormente e faça o download da versão 4.1.1 clicando em “*download R*”, conforme a imagem (0101).



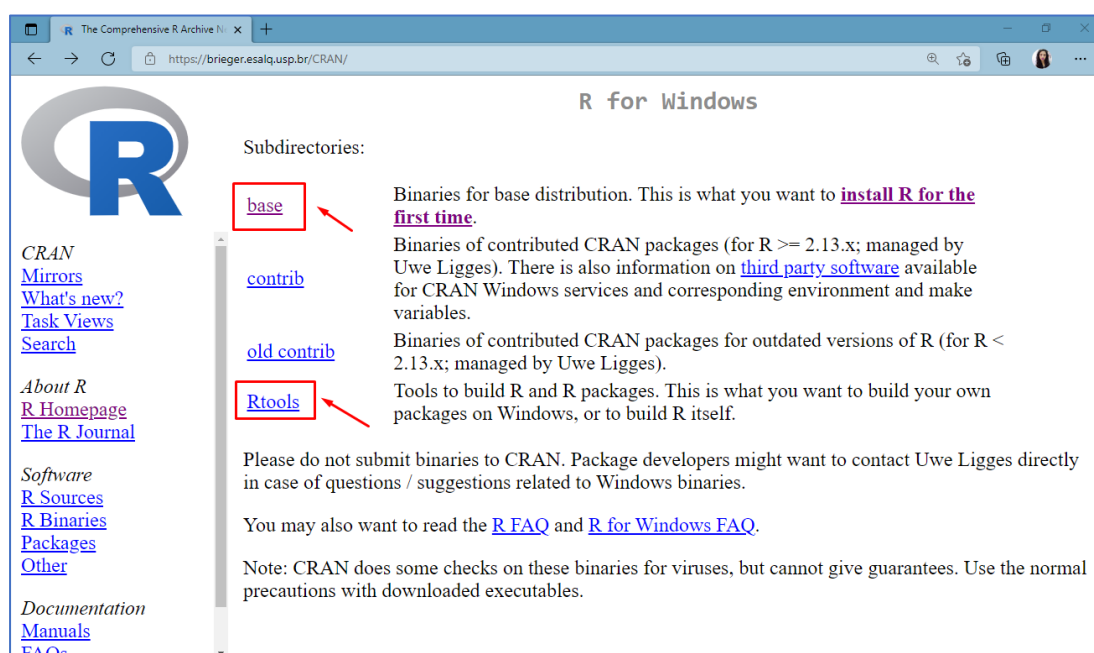
Ao clicar no link, você será direcionado para uma nova página onde estarão dispostos diversos repositórios do R em diferentes localidades do mundo. Encontre os repositórios do Brasil e selecione o último link (<https://briegeer.esalq.usp.br/CRAN/>), que é referente ao repositório da Universidade de São Paulo- Piracicaba, conforme mostra a imagem.



Uma nova página será apresentada, onde você deverá selecionar o *download* do R conforme o sistema operacional (SO) de seu computador. Nesse curso, usaremos uma distribuição para o Windows 10, portanto é possível que existam diferenças entre a forma como a linguagem será instalada em seu computador. Entretanto, não se preocupe! Todas as funções do R ensinadas nesse curso serão iguais em qualquer sistema operacional. A imagem (0103) apresenta a página de *download* do R.

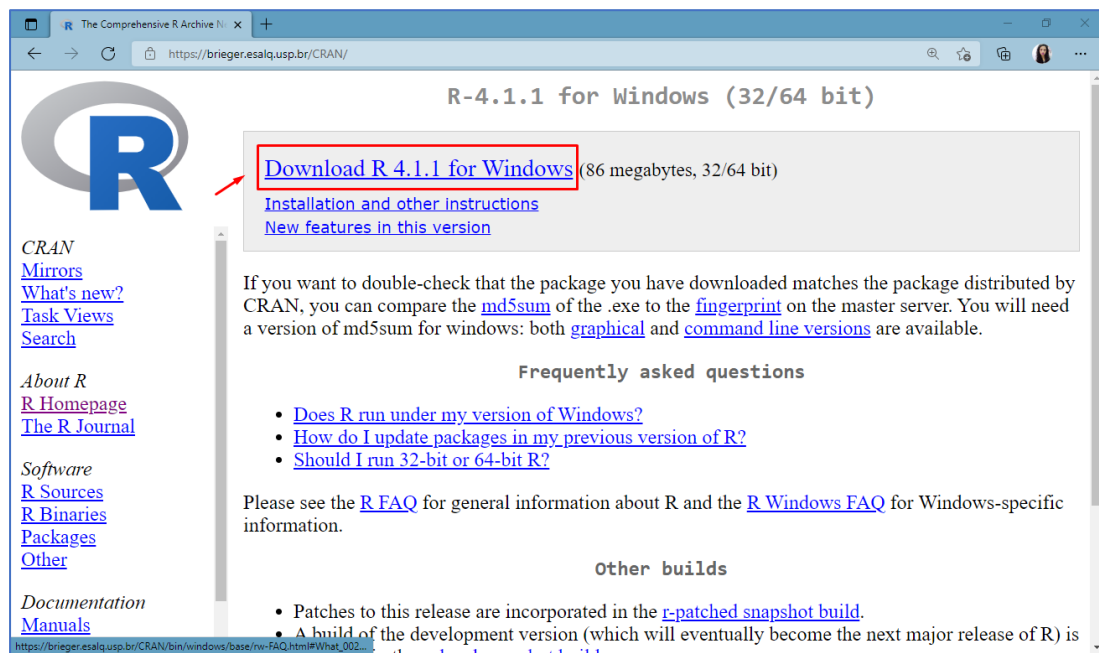


Após isso, você será direcionado para uma página onde deverá escolher qual ferramenta do R realizará o *download*, conforme a imagem (0104).

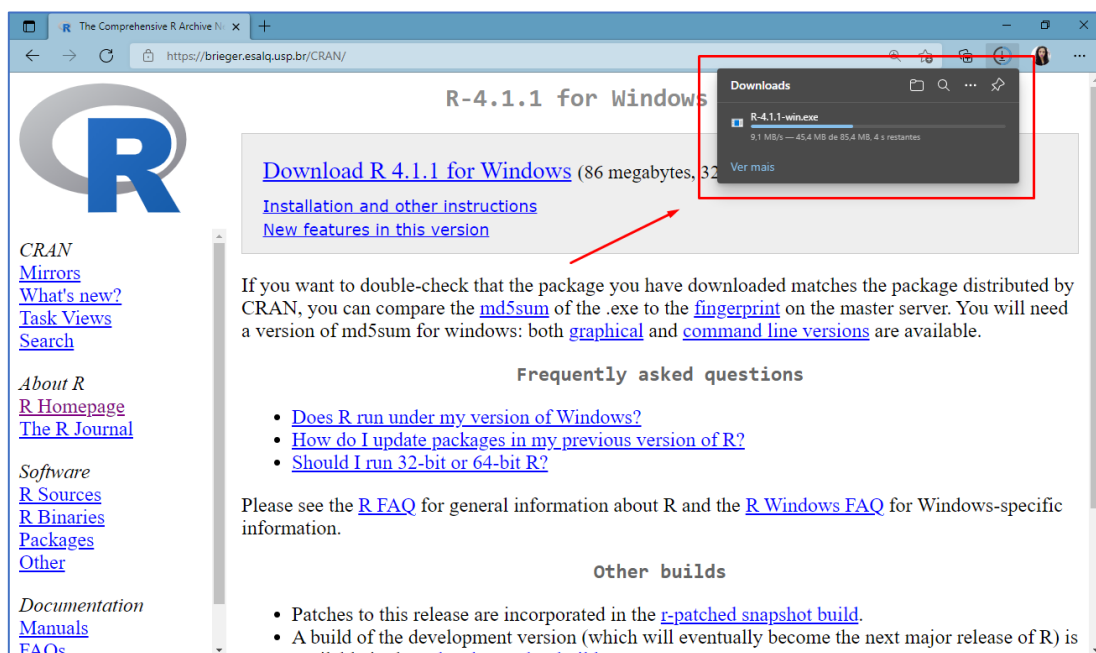


Nós realizaremos *download* do R base, que é a versão nativa do R, com todas as funções principais da linguagem e o Rtools, que será importante para a instalação de alguns pacotes que usaremos nesse curso.

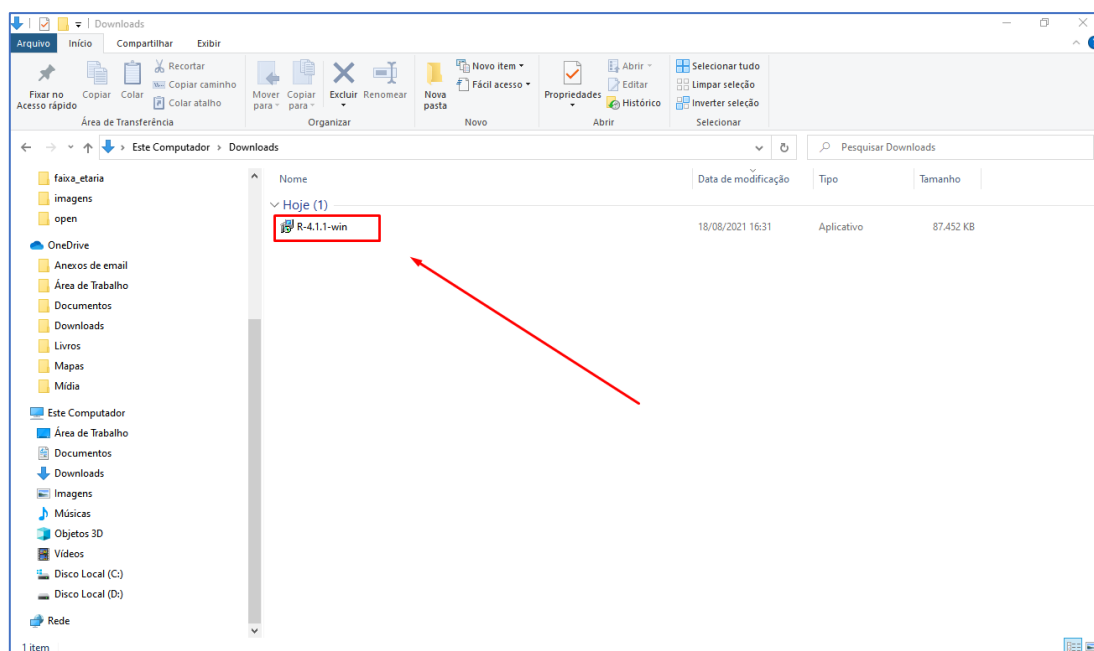
Ao clicar no em “base” você será direcionado para a página de *download* do R 4.1.1 para o Windows. Nesse curso usaremos a versão 64 bit, mas é possível baixar a versão de 32 por meio do mesmo link, conforme a imagem (0105).



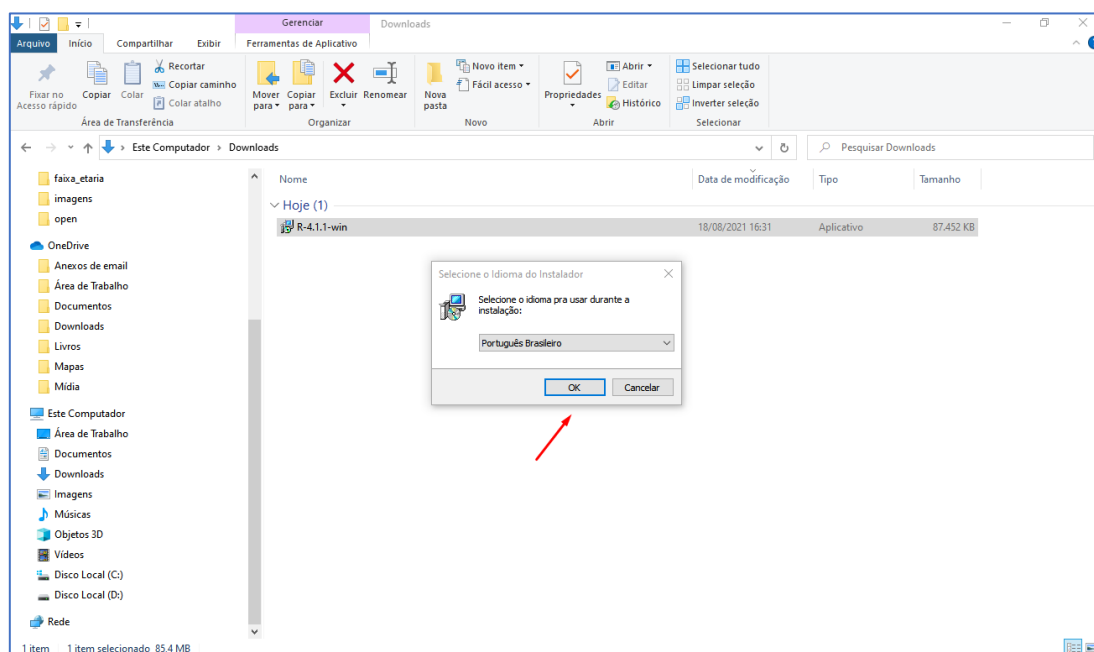
Após clicar, o seu navegador automaticamente fará o *download* do instalador da linguagem R, conforme a imagem (0106).



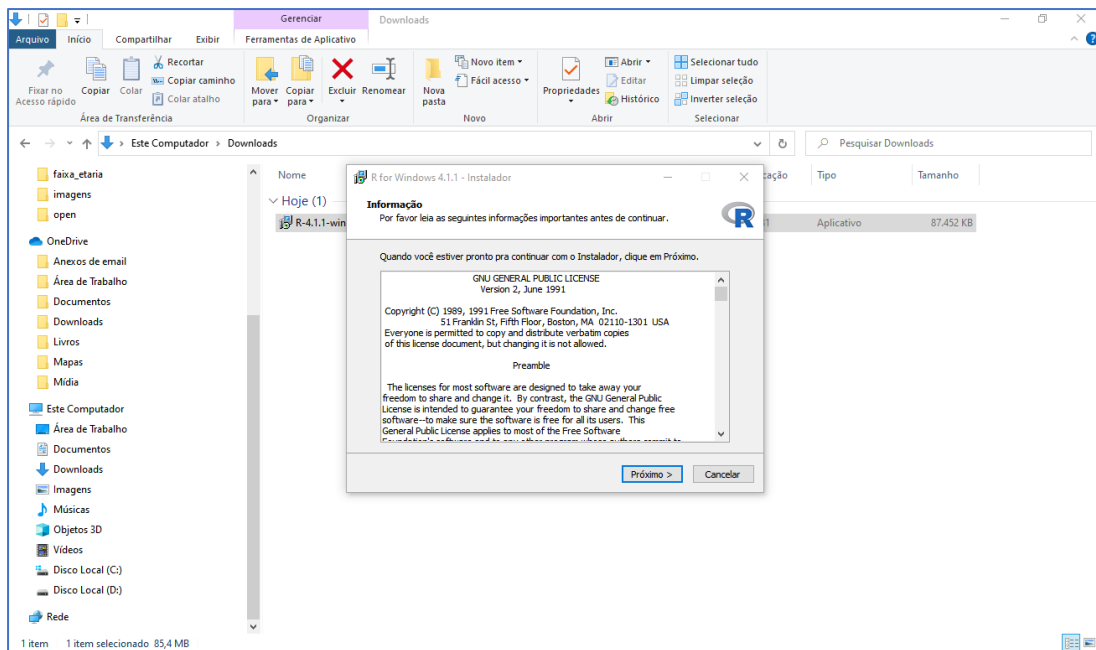
Abra a pasta em que os dados foram salvos e clique duas vezes no instalador para executar a instalação, de acordo com a imagem (0107).



Após clicar, o instalador abrirá uma janela solicitando a seleção de uma linguagem para a ferramenta. Você deverá selecionar a opção “português(Brasil)”, conforme a imagem(0108).

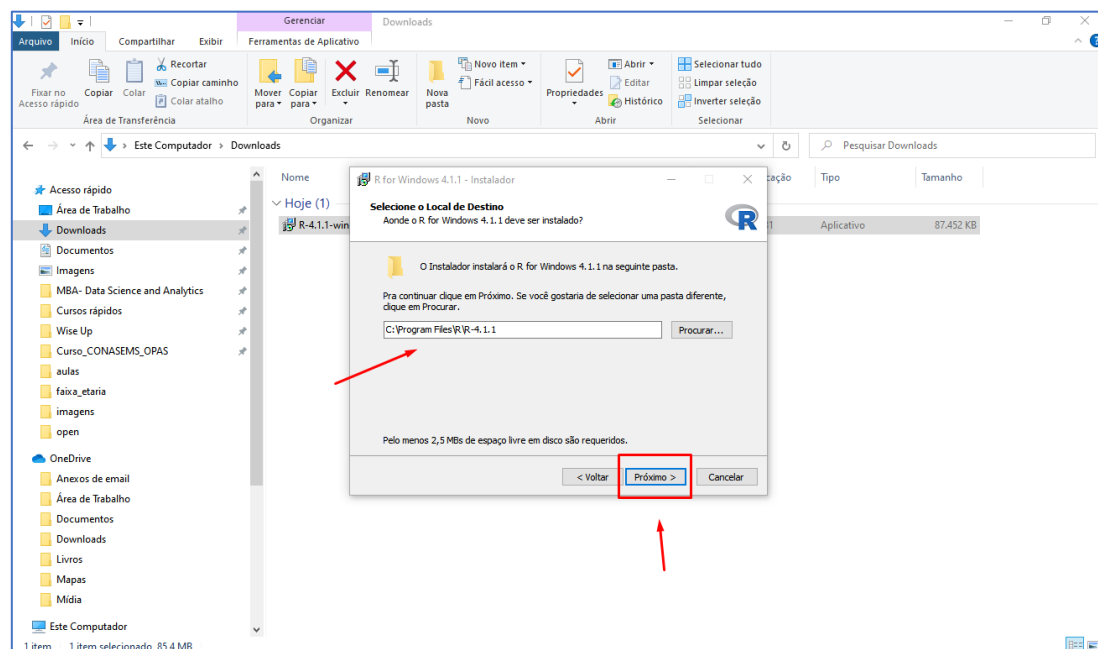


Após isso, será necessário aceitar a licença pública da linguagem. Nessa etapa, você poderá ler a licença e suas condições. Clique em “próximo” para continuar a instalação (0109).



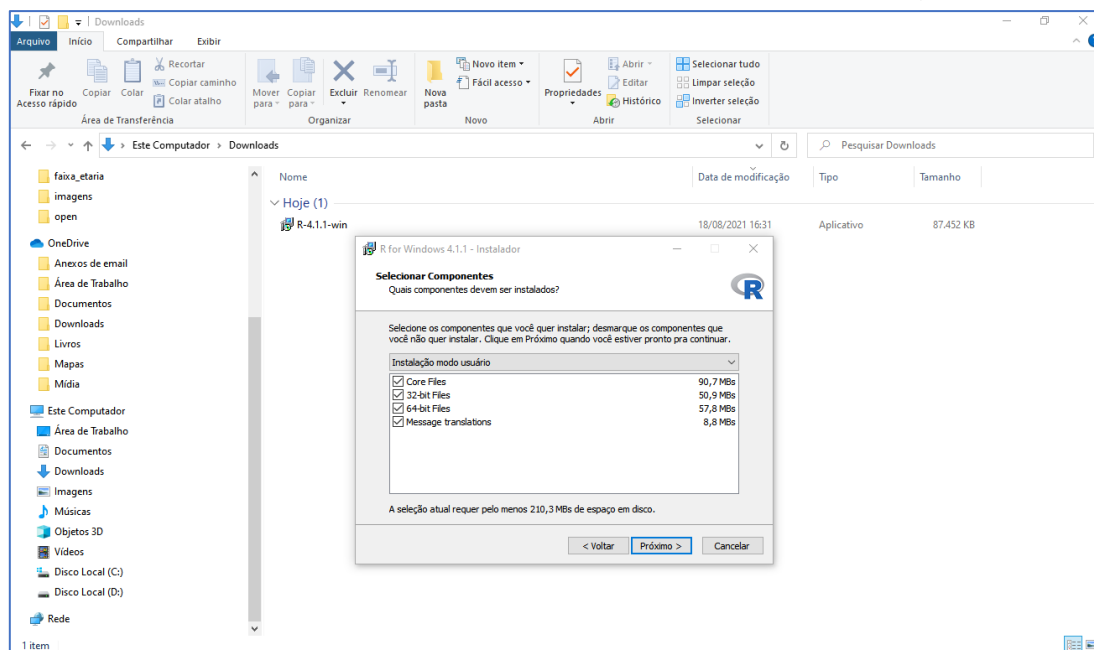
Após isso, o R automaticamente encontrará uma pasta em seu disco C:\ para instalação da linguagem, conforme a imagem (0110). Clique em “próximo” para avançar com a instalação.

Atenção! Para realizar a instalação nesse ponto, é necessário ser administrador da máquina. Caso não seja, será necessário autorizar a instalação com o *login* de administrador.

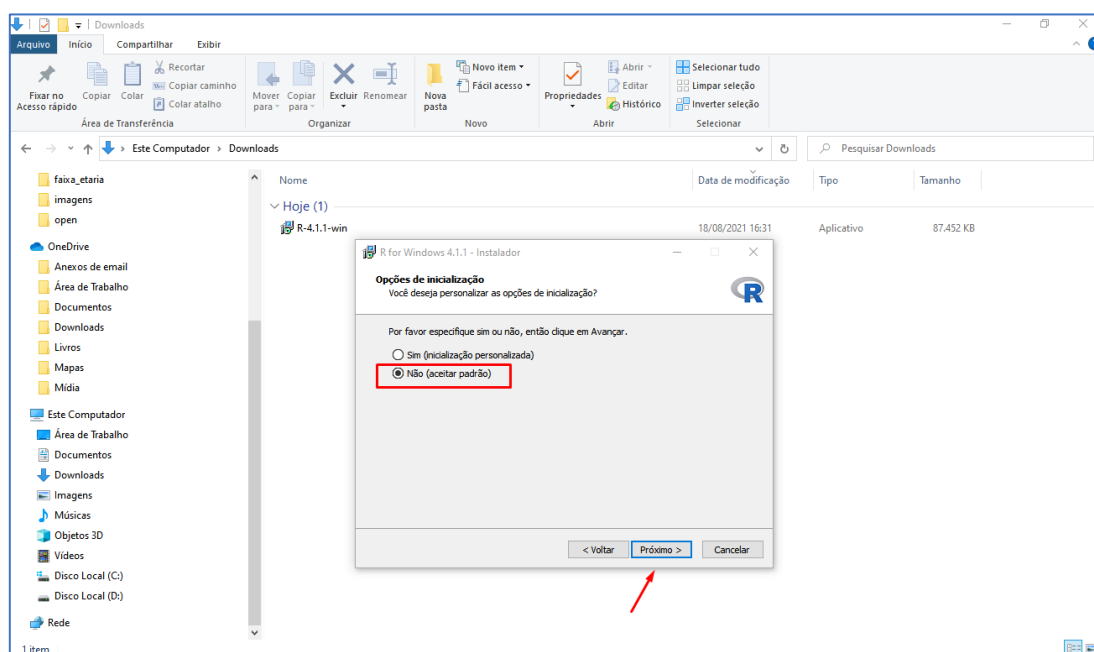


Após a seleção, o R irá apresentar os componentes que poderão ser instalados no seu computador. Se a configuração do seu Windows é de 32 bit, recomendamos que faça

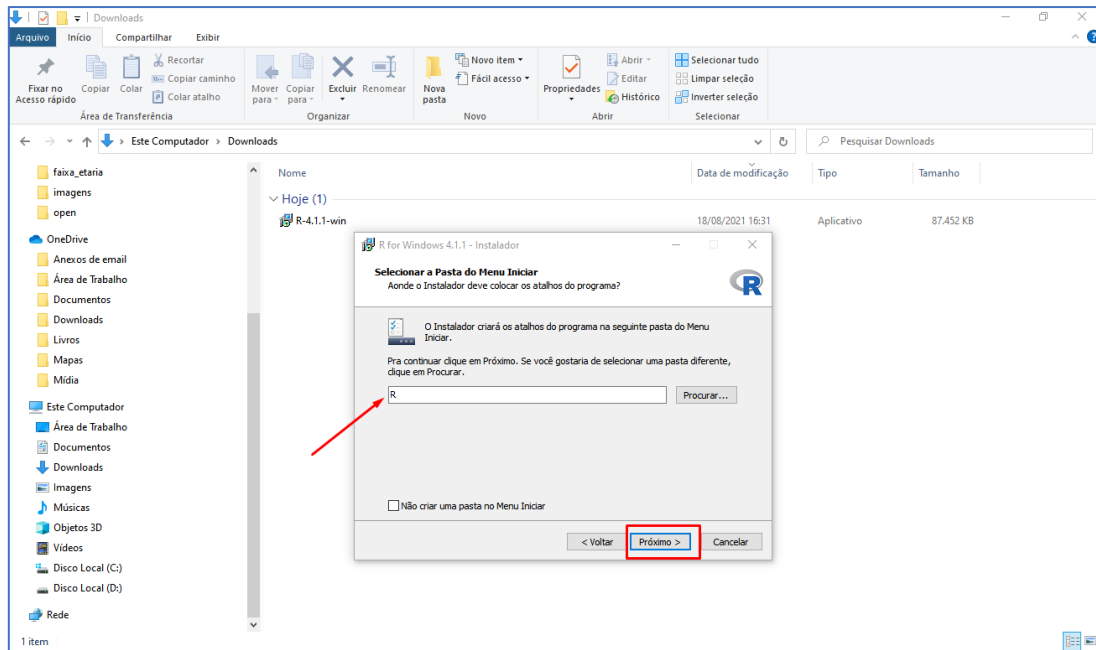
apenas a instalação desse componente, desmarcando a opção “64-bit Files”. Em nosso caso, instalaremos todos os componentes disponíveis, conforme a imagem (0111).



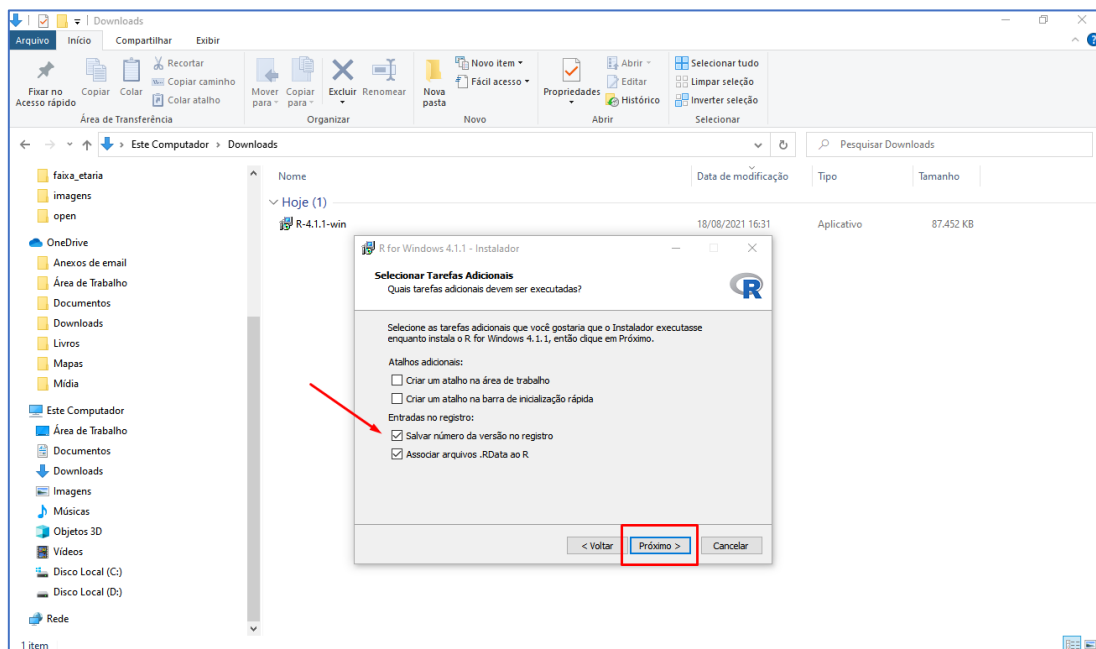
Após isso, o instalador solicitará as opções de inicialização. Recomendamos fortemente que selecione a opção “Não (aceitar padrão)”. Essa é a opção nativa do R e vai instalar a versão padrão no seu computador (0112).



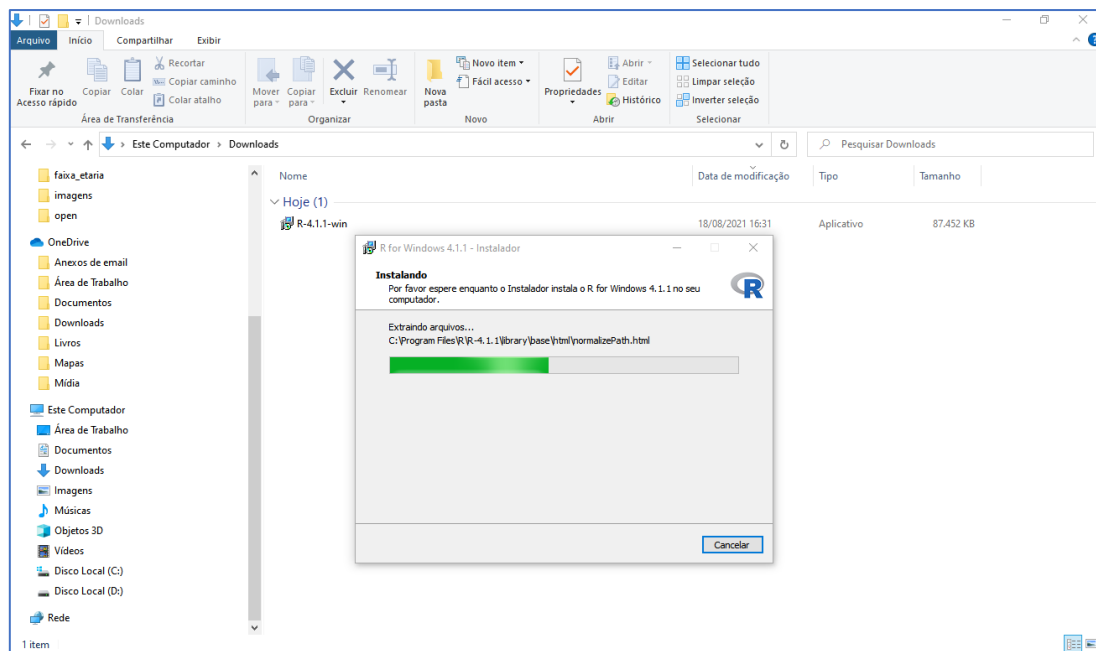
Nessa fase, o R criará uma pasta do Menu Iniciar do Windows. Você poderá aceitar o nome da pasta, mudá-lo ou optar por não criar um atalho no Menu Iniciar. Após escolher, clique em “Próximo” (0113).



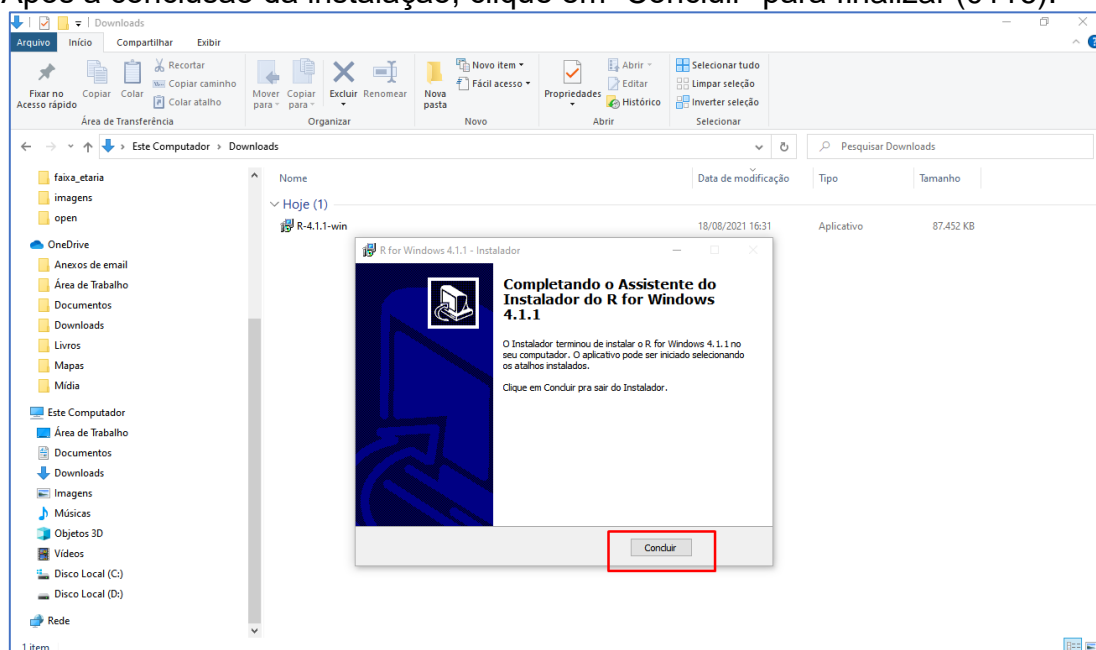
O instalador solicitará que você escolha se quer adicionar um atalho na barra de inicialização rápida e na área de trabalho. Você poderá optar por criá-las ou não. Na parte “Entradas no registro:” recomendamos que as duas opções (“Salvar número da versão no registro” e “Associar arquivos .RData ao R”) permaneçam marcadas (0114).



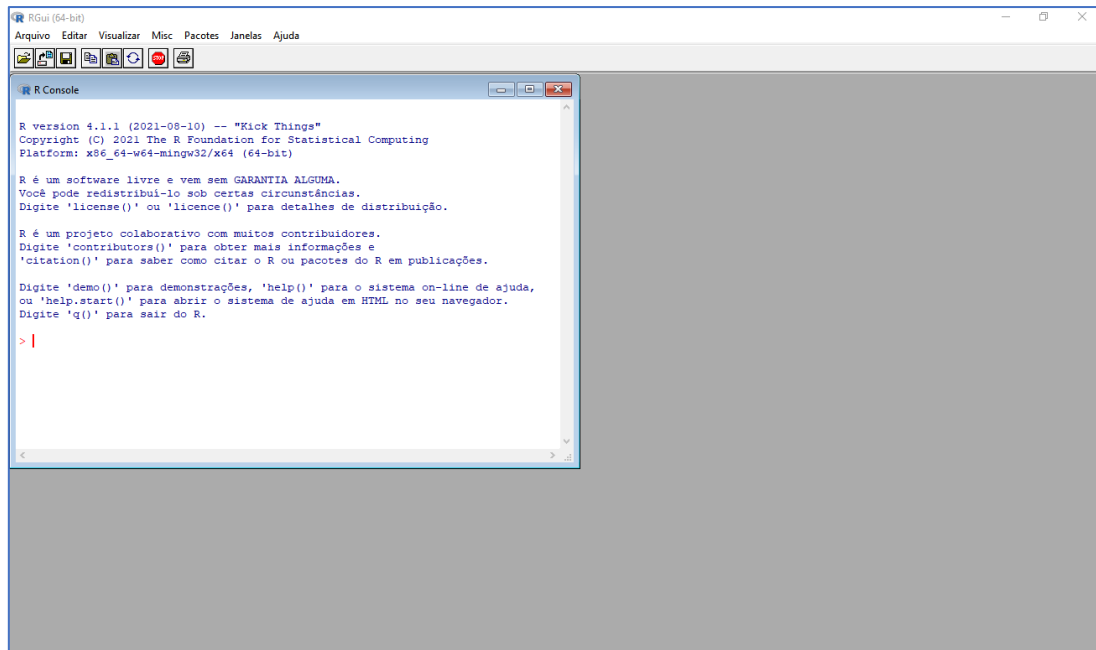
Após isso, a instalação do R será executada em seu computador (0115).



Após a conclusão da instalação, clique em “Concluir” para finalizar (0116).

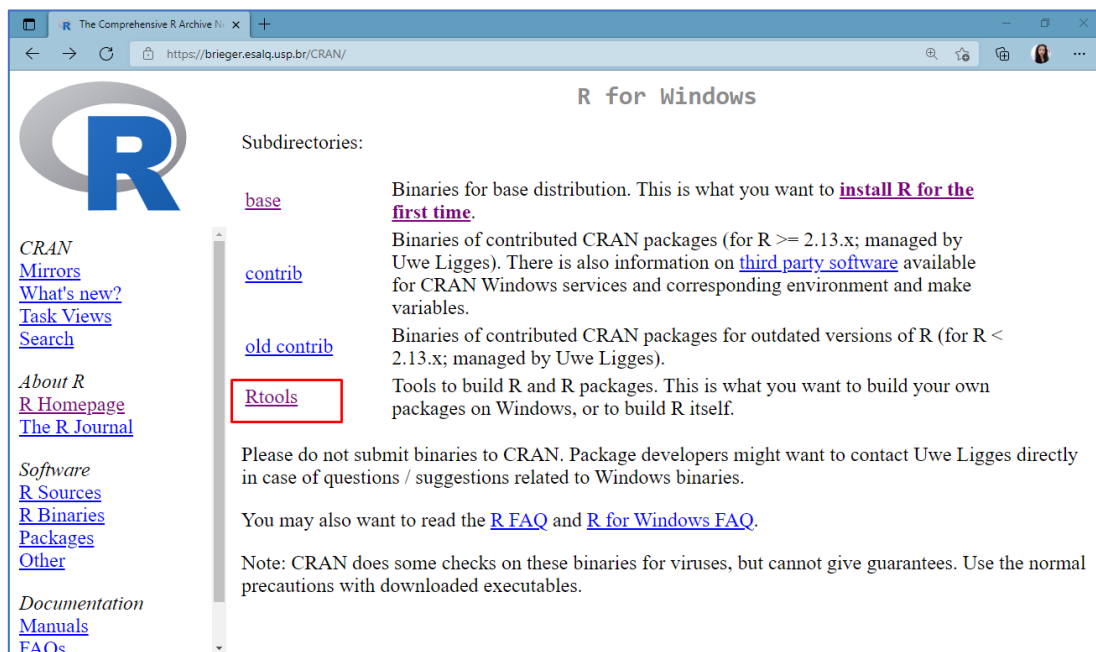


Pronto, agora você tem a linguagem R instalada em seu computador! Nesse ponto você pode acessar o console do R, que permite que você execute comandos para a linguagem(0117). Todavia, não é possível salvar os comandos por meio do RGui, sendo necessário o uso de uma interface de desenvolvimento para a criação de *scripts* e projetos.

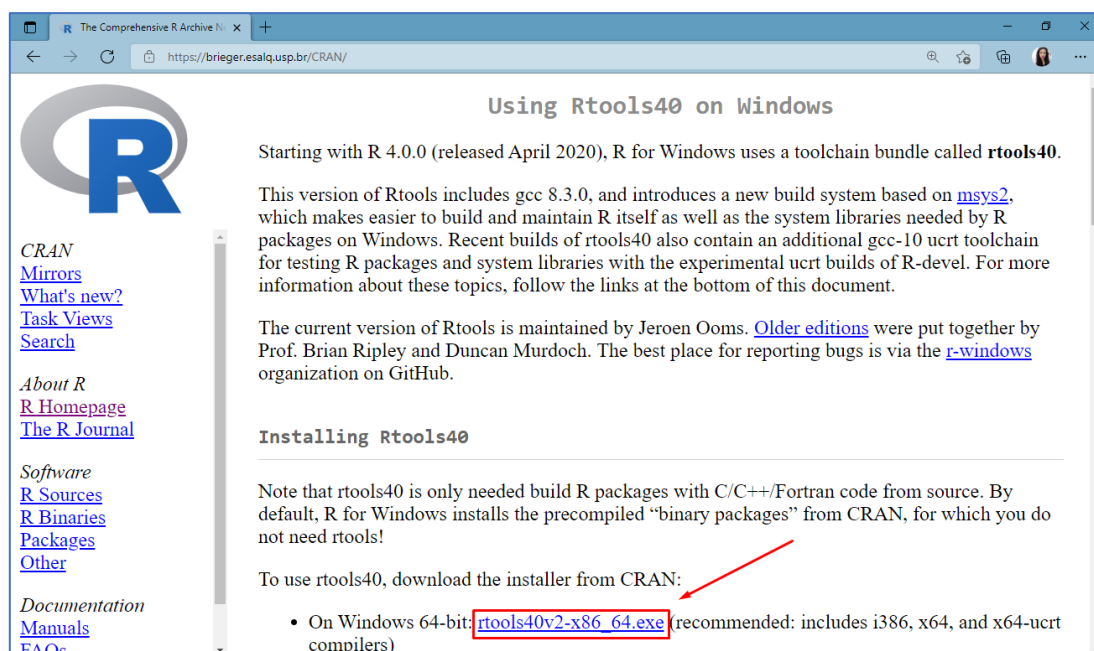


Faremos a instalação de uma interface de desenvolvimento no próximo tópico, mas antes disso voltaremos ao site do R para realizar a instalação do Rtools. O Rtools é um *software* necessário para a execução de uma série de comandos e pacotes que instalaremos no R. Ele também será útil caso queira construir seu próprio pacote de funções.

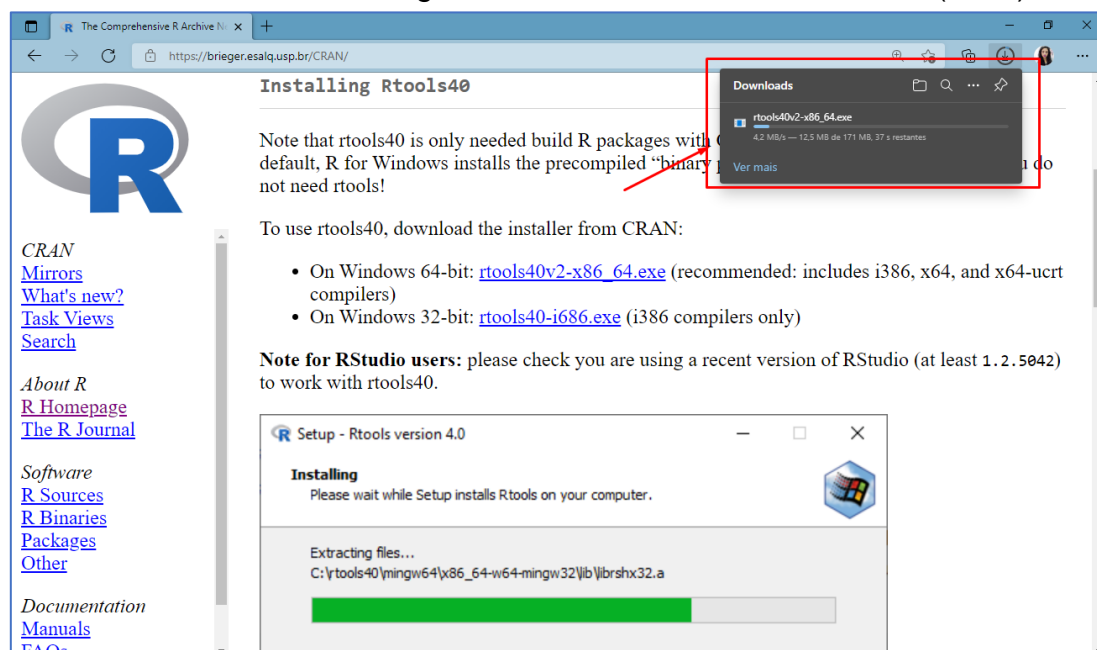
Por isso, retornaremos na página de *download* do R e realizaremos o *download* do instalador do Rtools(0118).



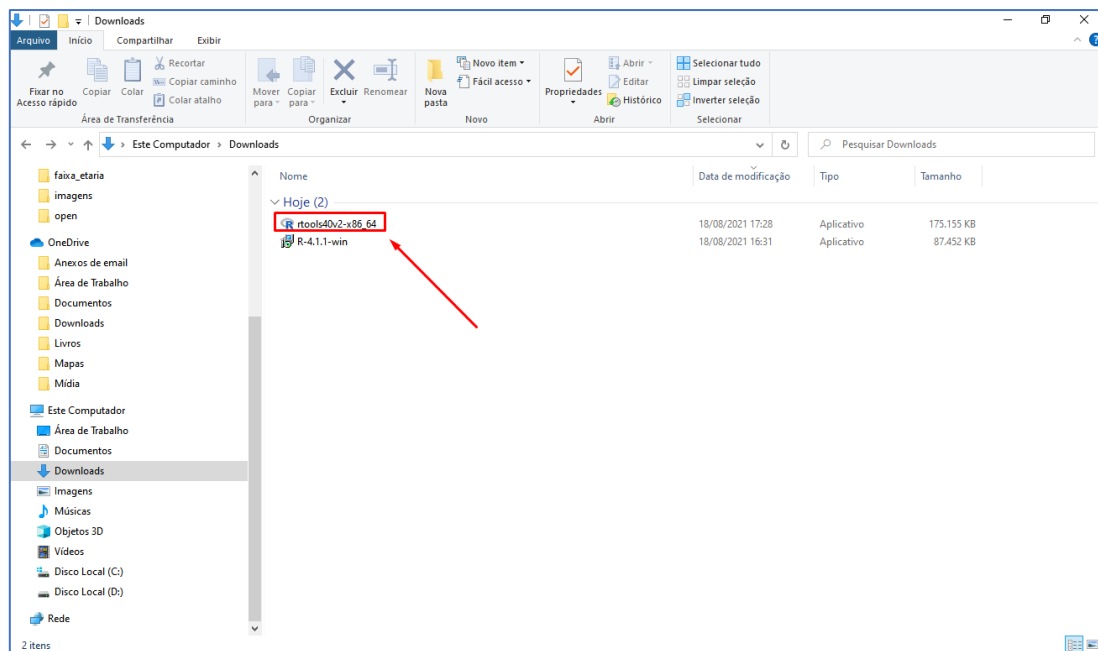
Após isso, o *site* apresentará uma tela onde serão apresentadas as opções de instalação do Rtools. Em nosso caso, instalaremos a versão Windows 64-bit. Caso seu computador possua uma configuração 32-bit, selecione essa opção (0119).



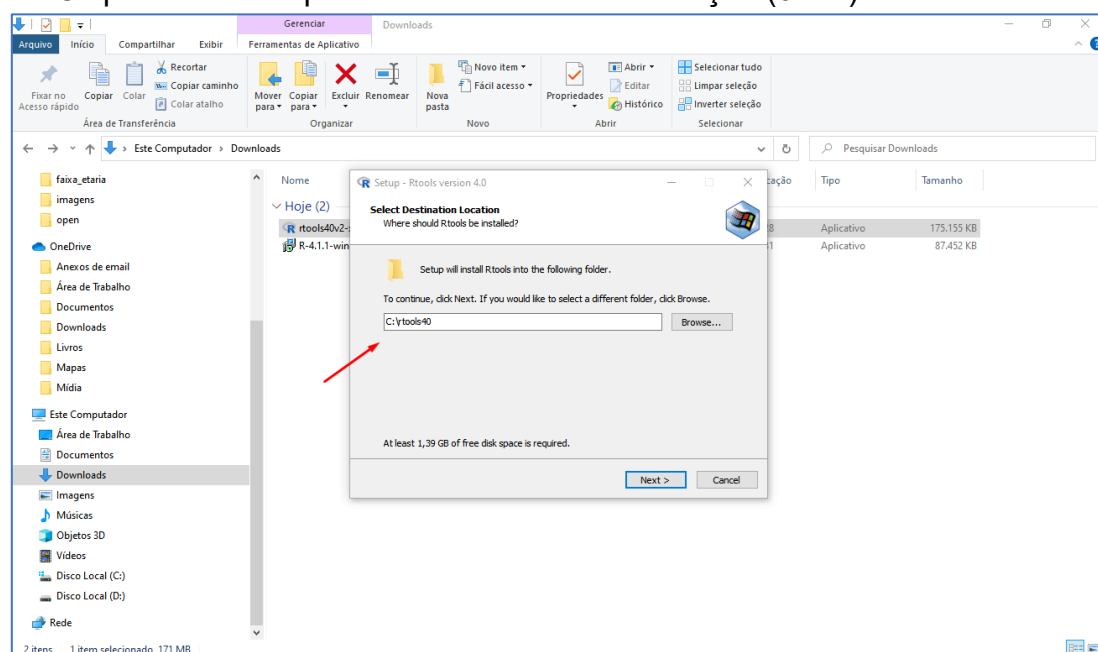
De forma automática, o navegador fará o *download* do instalador (0120).



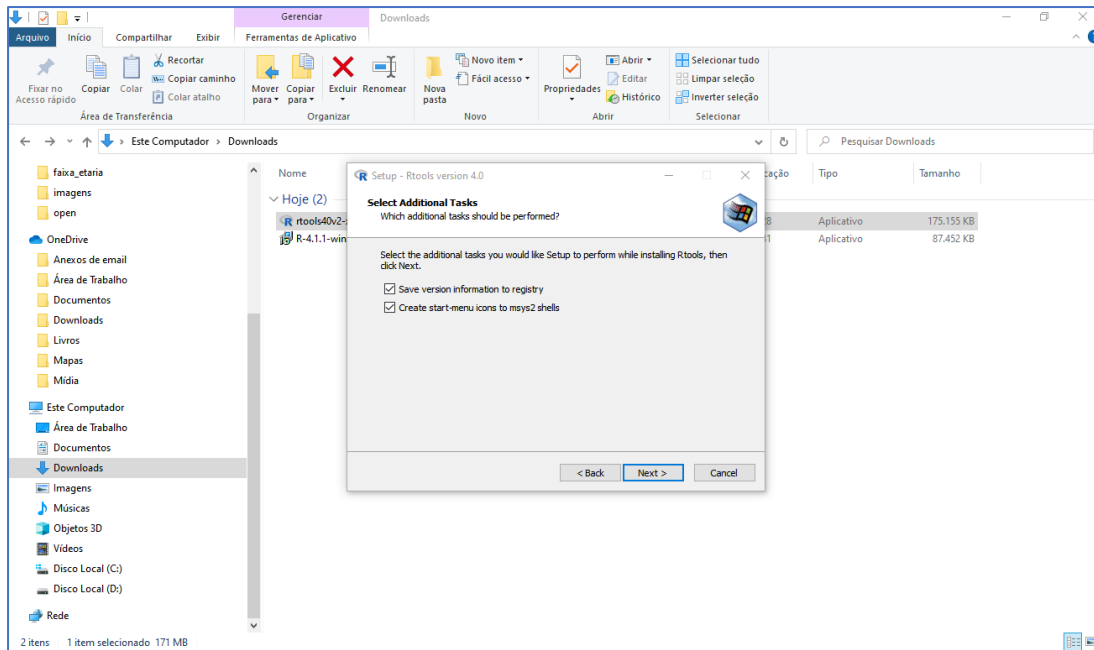
Na pasta em que o *download* foi salvo, selecione o instalador do Rtools para iniciar a instalação (0121).



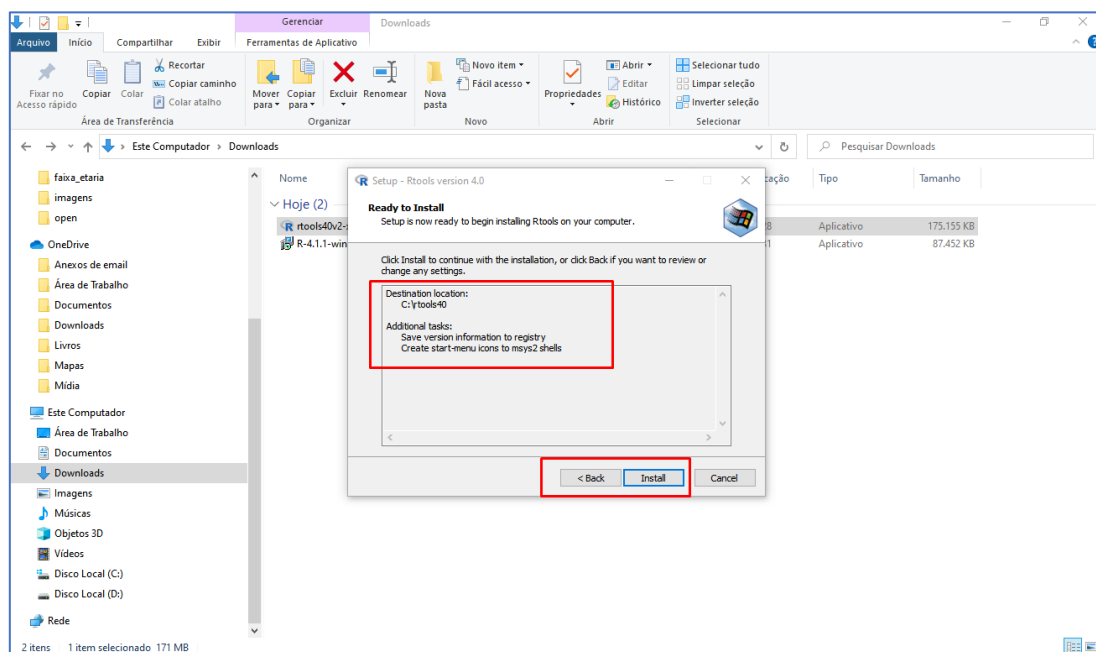
Uma nova janela será aberta, indicando a pasta do disco C:\ em que o Rtools será instalado. Clique em “Next” para continuar com a instalação (0121).



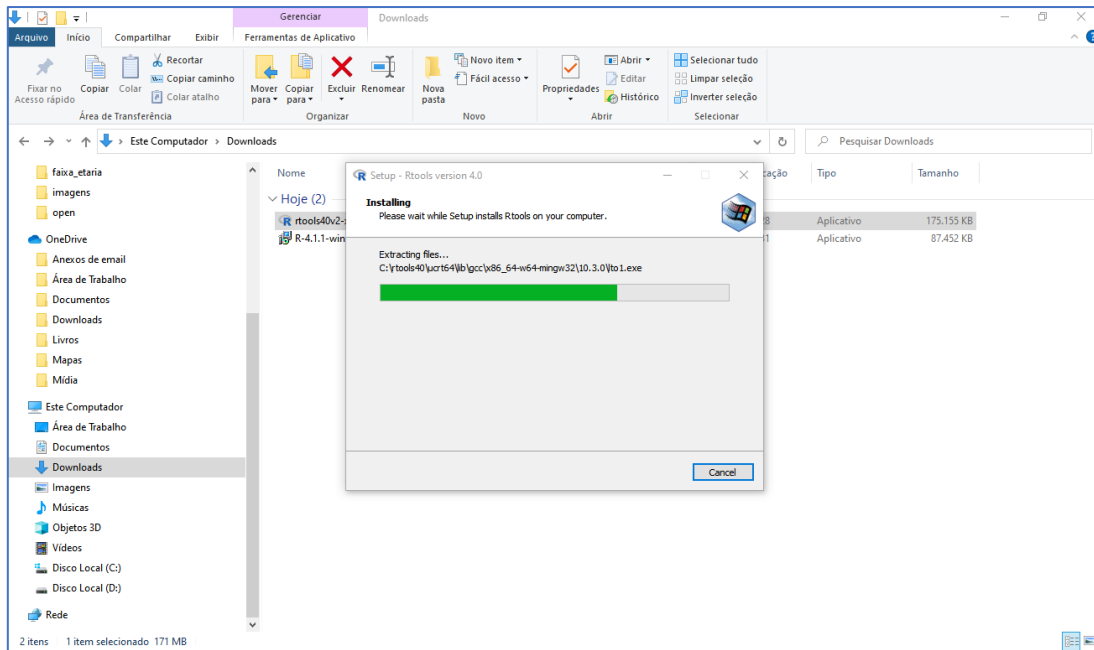
O instalador solicitará que você instale os recursos adicionais. Recomendamos que mantenha as duas caixas de seleção marcadas e clique em “Next” (0123).



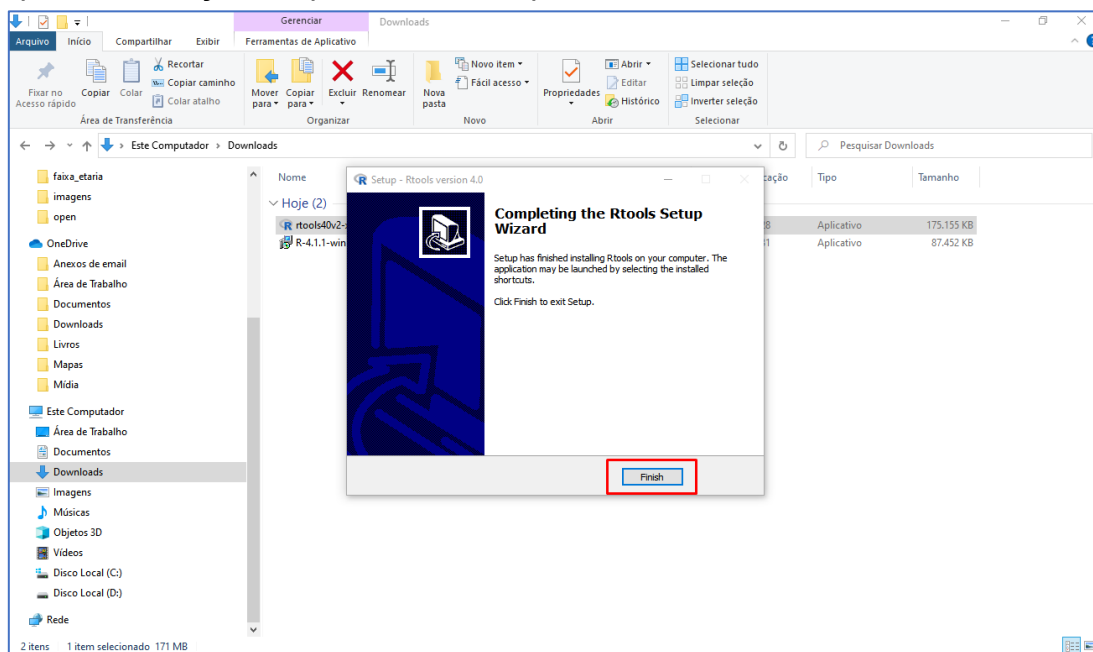
O instalador mostrará um resumo das seleções feitas e pedirá que inicie a instalação. Clique em “Install” para continuar (0124).



O instalador iniciará sua execução. Aguarde finalizar a instalação para continuar (0125).



Após a instalação, clique em “Finish” para encerrar o instalador.

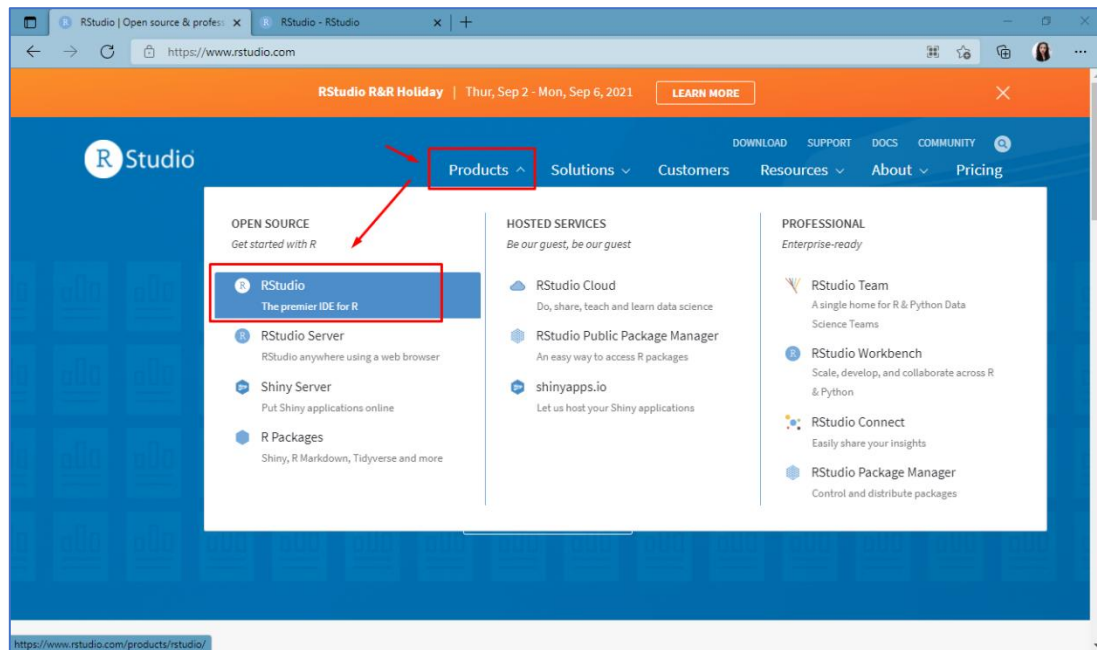


Pronto! Toda a instalação do Rtools foi finalizada e você já possui todos os elementos necessários para o uso da linguagem R. Na próxima seção instalaremos um ambiente de desenvolvimento integrado do R, que será um facilitador para a execução de códigos e permitirá que você salve *scripts* com as funções que você criará, além de permitir a criação de projetos que facilitarão o uso do R.

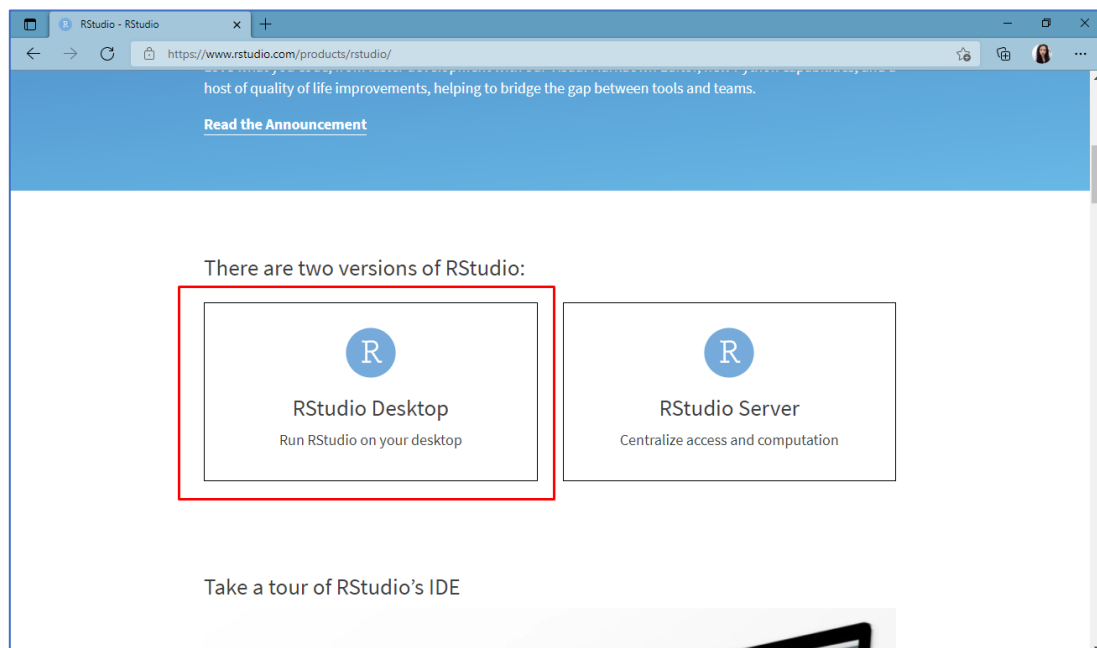
3. Instalando o ambiente de desenvolvimento integrado R Studio

O ambiente de desenvolvimento integrado RStudio está em sua versão de número 1.4. Um ambiente de desenvolvimento integrado é um *software* que permite a manipulação da linguagem por meio de comandos e execuções. Para instalar o RStudio, acesse [RStudio](https://www.rstudio.com) | [Open source & professional software for data science teams - RStudio](https://www.rstudio.com).

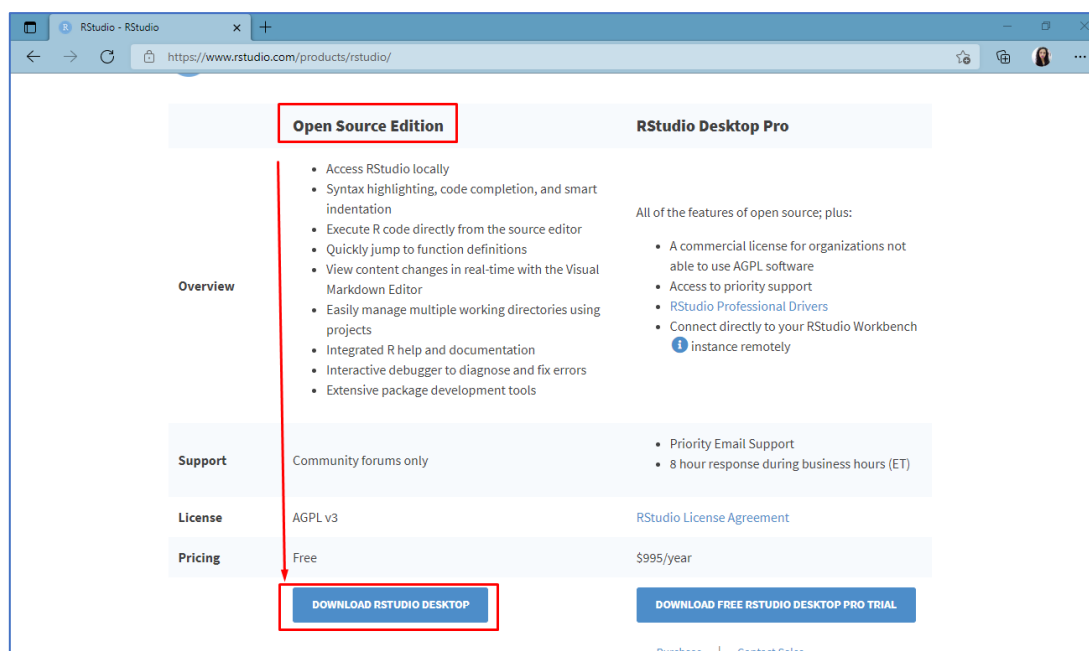
Selecione a opção “Products” e na aba “Open Source” clique em RStudio (0127).



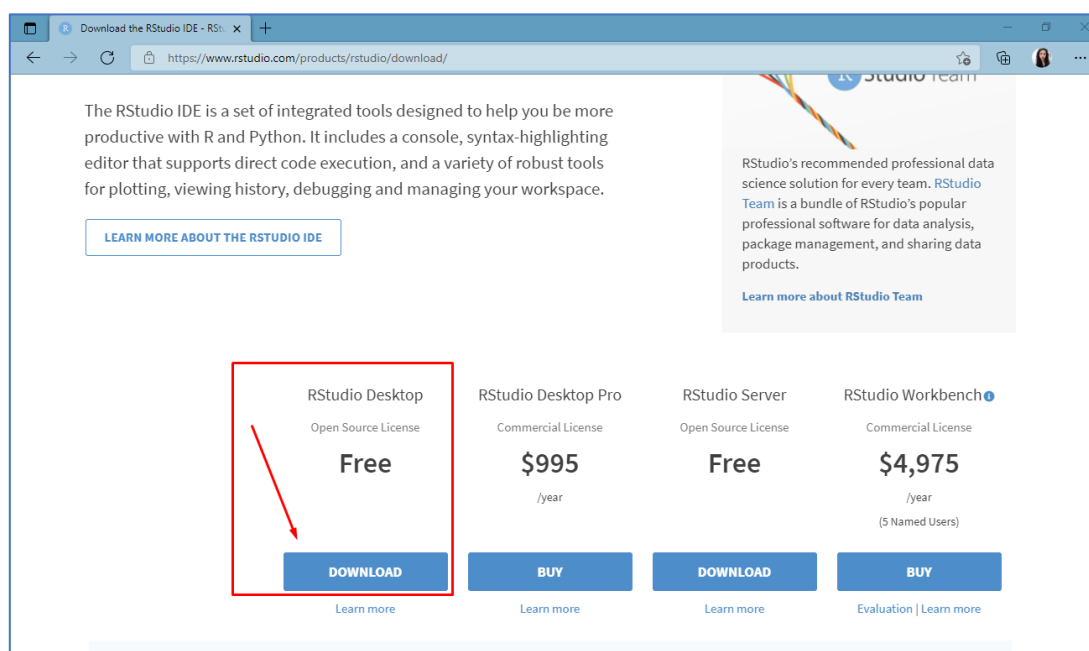
Você será direcionado a uma nova página, onde deverá selecionar a versão gratuita do RStudio (0128).



Ao clicar você será direcionado para o espaço que mostra uma visão geral do RStudio, desça até o download da “Open Source Edition” (0129).



Você será direcionado para a página de download. Selecione a versão Free (0130).



Você será direcionado para a página de *download* dos instaladores. Caso seu sistema operacional seja o Windows 64-bit, selecione a caixa em azul para instalar o programa. Se não for, selecione uma das opções na lista no canto inferior da tela.

RStudio Desktop 1.4.1717 - Release Notes

1. Install R. RStudio requires R 3.0.1+.
2. Download RStudio Desktop. Recommended for your system:

DOWNLOAD RSTUDIO FOR WINDOWS
1.4.1717 | 156.18MB
Requires Windows 10 (64-bit)

All Installers

Linux users may need to [import RStudio's public code-signing key](#) prior to installation, depending on the operating system's security policy.

RStudio requires a 64-bit operating system. If you are on a 32 bit system, you can use an [older version of RStudio](#).

OS	Download	Size	SHA-256
Windows 10	RStudio-1.4.1717.exe	156.18 MB	71b36e64

O download ocorrerá de forma automática (0132).

RStudio Desktop 1.4.1717 - Release Notes

1. Install R. RStudio requires R 3.0.1+.
2. Download RStudio Desktop. Recommended for your system:

DOWNLOAD RSTUDIO FOR WINDOWS
1.4.1717 | 156.18MB
Requires Windows 10 (64-bit)

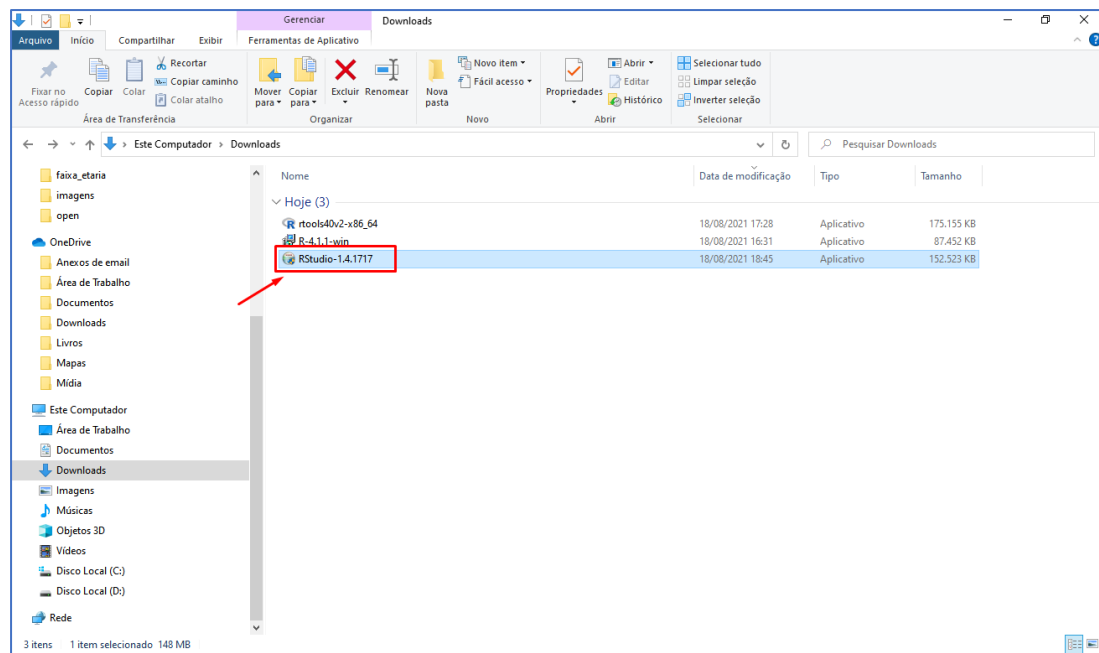
All Installers

Linux users may need to [import RStudio's public code-signing key](#) prior to installation, depending on the operating system's security policy.

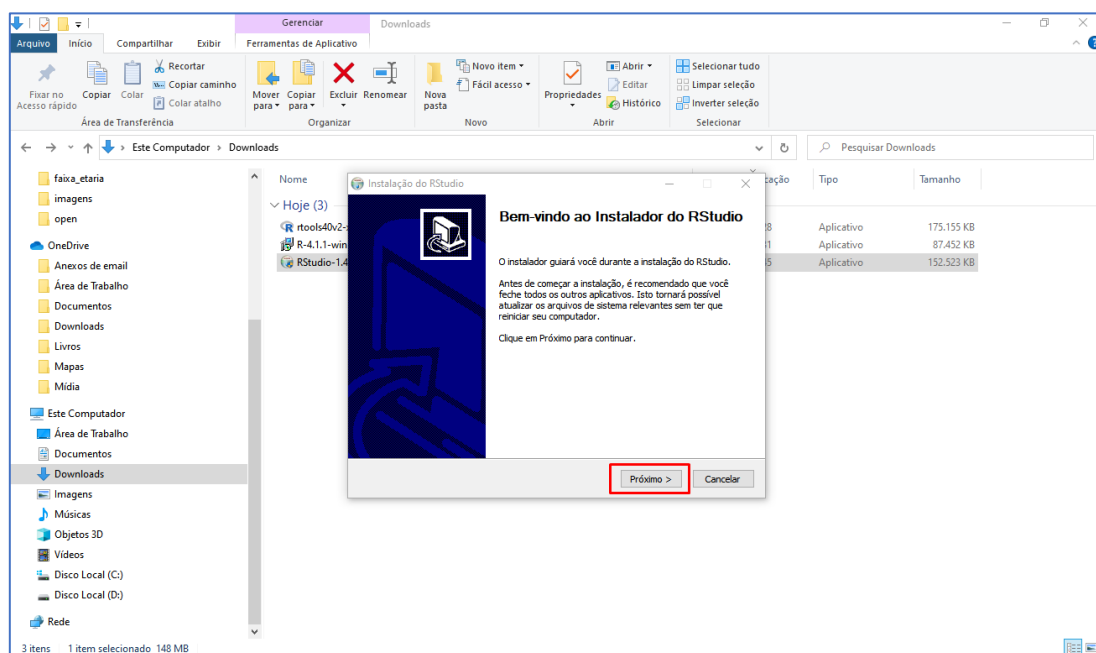
RStudio requires a 64-bit operating system. If you are on a 32 bit system, you can use an [older version of RStudio](#).

OS	Download	Size	SHA-256
Windows 10	RStudio-1.4.1717.exe	156.18 MB	71b36e64

Abra a pasta em que o instalador foi salvo e clique duas vezes para executar (0133).

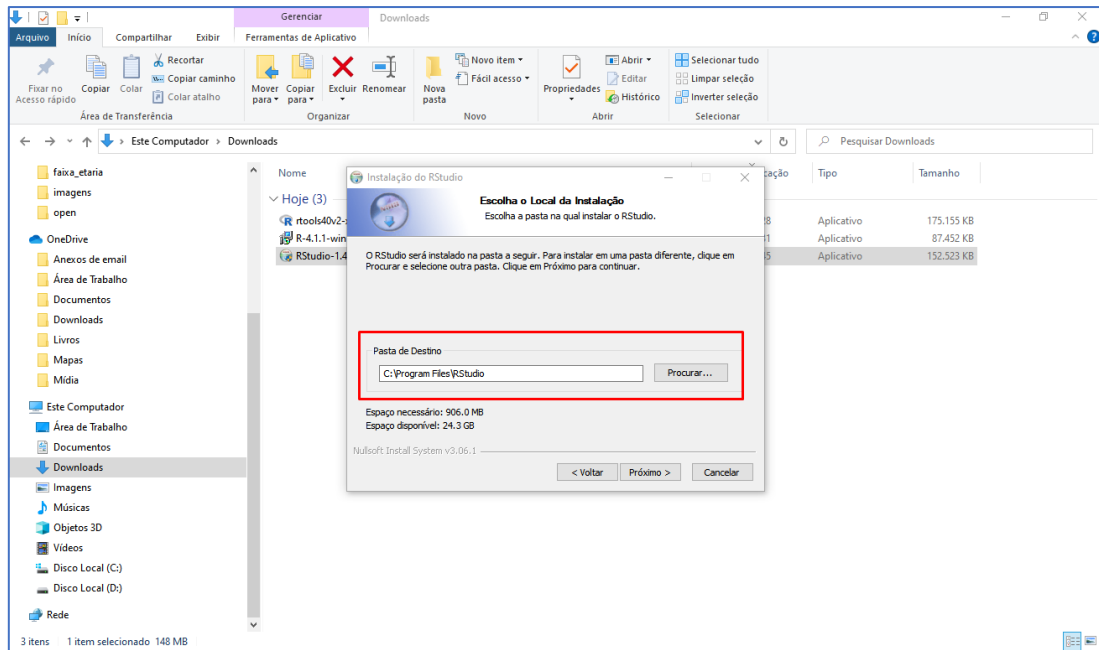


O instalador do RStudio será aberto e guiará a instalação do programa. Clique em “Próximo” para continuar (0134).

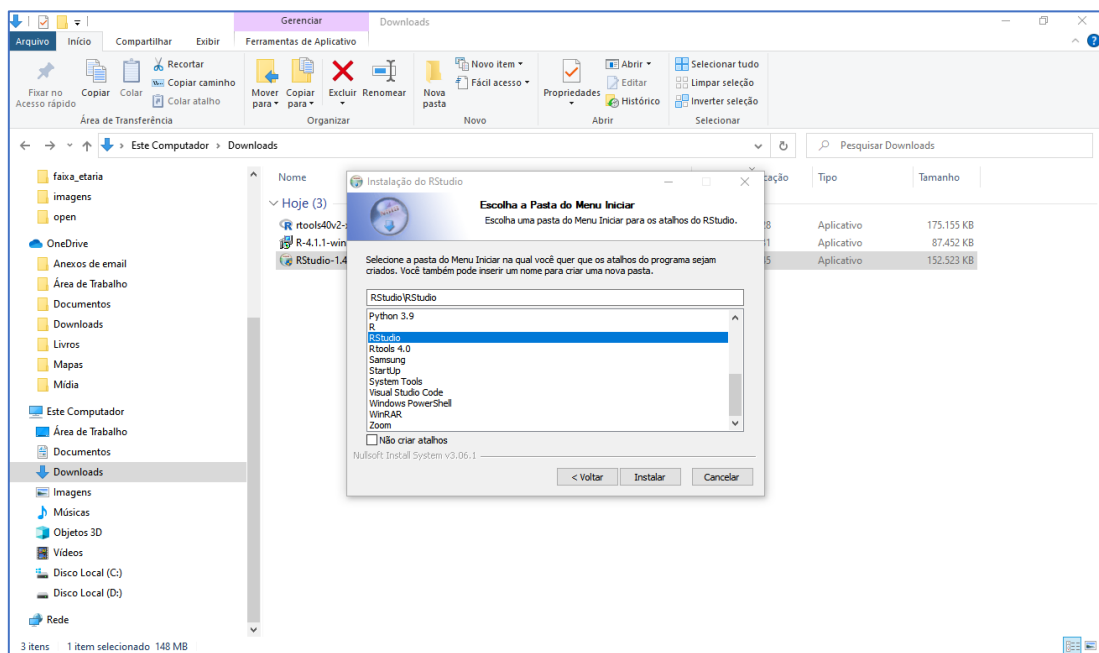


O instalador também selecionará automaticamente uma pasta de destino para a instalação do programa. Clique em “Próximo” para continuar (0135).

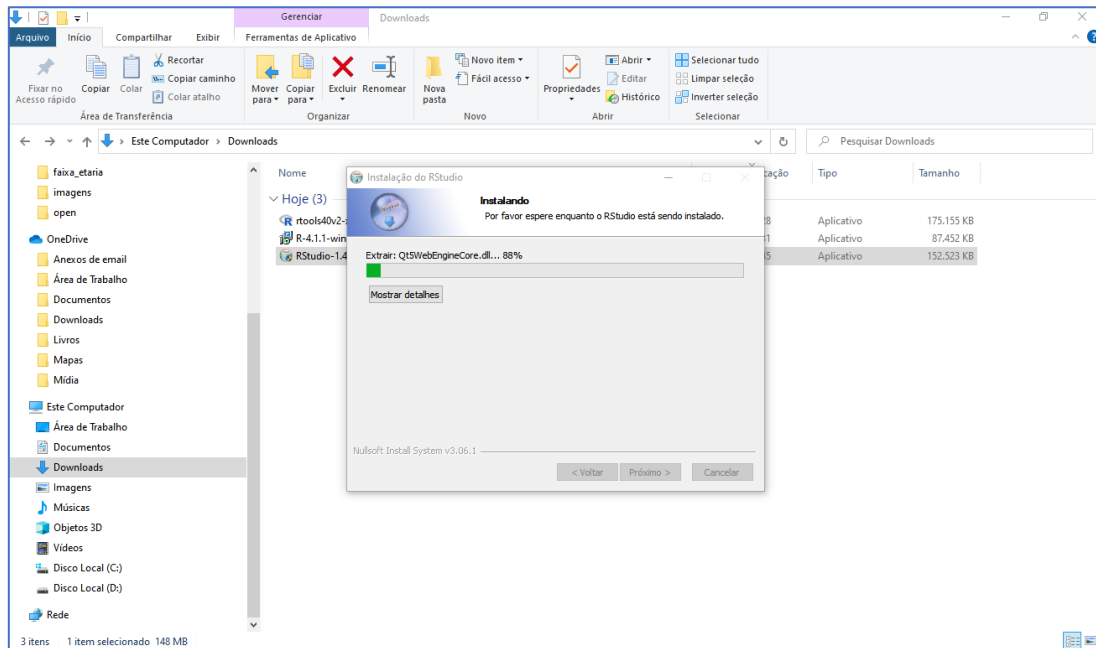
Atenção! Para realizar a instalação nesse ponto, é necessário ser administrador da máquina. Caso não seja, será necessário autorizar a instalação com o *login* de administrador.



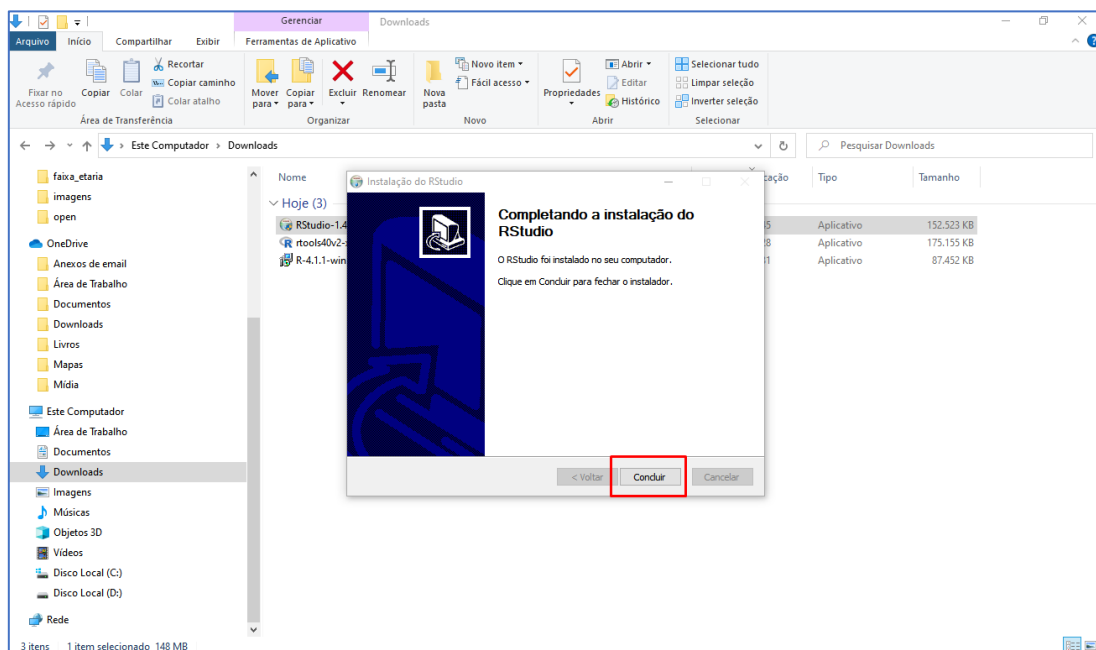
O instalador pedirá para criar um atalho do *software* no Menu Iniciar, caso você esteja usando o Windows. Clique em “Instalar” (0136).



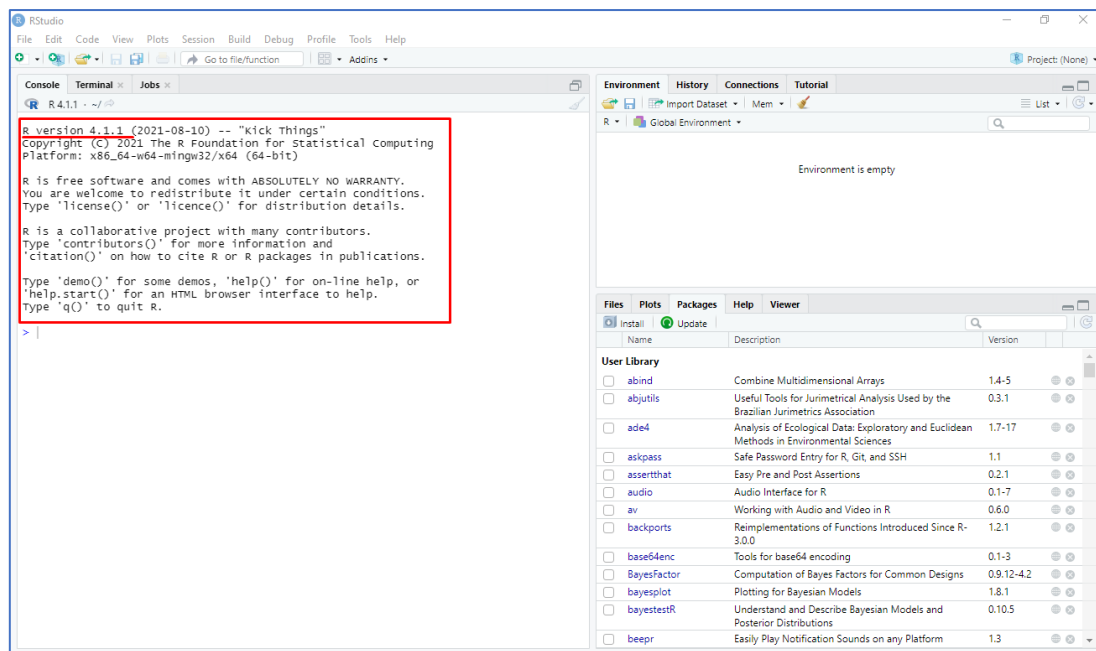
A execução da instalação será executada após isso (0137).



Após a instalação, aparecerá uma mensagem de sucesso na tela. Clique em “Concluir” para finalizar(0138).



Pronto! Você possui o RStudio instalado em seu computador. Ao abrir o programa, você verá uma janela com 3 ambientes, conforme a imagem (0139). Verifique se o RStudio identificou automaticamente a versão 4.1.1 do R *language*. Caso você tenha uma versão anterior já instalada, é possível que ele identifique a antiga versão. Se isso ocorrer, selecione a versão 4.1.1 no menu de opções.



Nas próximas aulas ensinaremos o uso do R language por meio do R Studio, por isso é importante que você tenha os dois instalados em seu computador. Recomendamos que instale as versões indicadas nesse curso para que seus comandos funcionem de acordo com o esperado.

4. CRAN

O *Comprehensive R Archive Network* é um repositório de bibliotecas, também chamadas de pacotes, e da linguagem R de uma maneira geral. Todos os pacotes aceitos para compor o CRAN são testados conforme a segurança e o funcionamento, com uma série de requisitos para sua inclusão.

Os pacotes instalados no R são todos provenientes do CRAN. É possível instalar pacotes de outros repositórios, utilizando outros meios de instalação, mas recomendamos fortemente que você utilize apenas pacotes disponíveis no CRAN. Se for necessário instalar pacotes de outras fontes, esteja seguro de que elas são seguras.

5. Pacotes

Pacotes são conjuntos de funções do R que podem ser aplicadas a contextos específicos. Se você possui uma demanda de análise, é possível que exista um pacote com a função necessária para que você utilize. Atualmente, o CRAN conta com cerca de 13 mil bibliotecas com diversas funções.

Nesse curso, instalaremos diversas bibliotecas, principalmente para o tratamento de dados, cálculo de indicadores epidemiológicos e criação de gráficos. Para acessar o CRAN, será necessário acesso à internet, pois os pacotes são baixados diretamente do repositório online.

6. Referências Bibliográficas

1. R: What is R? [Internet]. [citado 18 de agosto de 2021]. Disponível em: <https://www.r-project.org/about.html>
2. Silva HA da. Manual Básico da Linguagem R: Introdução à análise de dados com a linguagem R e o RStudio para área da saúde. 1ª edição. Henrique Alvarenga da Silva; 2019.

Atividades

1	O R é um programa de edição de textos e arquivos de planilhas.	V	
		F	X
2	O R foi criado como uma linguagem muito semelhante à antiga linguagem S e foi desenvolvida na <i>Bell Laboratories</i> por John Chambers, um estatístico da companhia	V	X
		F	
3	A linguagem R é uma ferramenta para a análise estatística e criação de visualizações.	V	X
		F	
4	Não é possível baixar o R pela internet, sendo necessário solicitar o envio de seu instalador por e-mail.	V	
		F	X
5	O R é uma linguagem de alto nível orientada a objetos	V	X
		F	
6	Sempre é possível analisar dados por planilhas eletrônicas.	V	
		F	X
7	Não é possível incluir novas funções na linguagem R que trabalha apenas com importação de dados e criação de visualizações.	V	
		F	X
8	Os pacotes permitem a criação de novas linguagens de programação fora do ambiente do R.	V	
		F	X
9	O R permite a criação de análises descritivas, analíticas e inferenciais, além da criação de gráficos que podem ajudar a compor uma análise de situação de saúde.	V	X
		F	
10	As funções do R podem ser instaladas por meio de pacotes, que são arquivos com um conjunto de funções destinadas a uma determinada temática. Além disso, é possível criar funções definindo isso por meio do R.	V	X
		F	
11	O R não está disponível para a instalação no Brasil.	V	
		F	X
12	O RStudio é o ambiente de desenvolvimento integrado para a Linguagem R.	V	X
		F	
13	É possível personalizar o R com qualquer função que esteja disponível no CRAN (Comprehensive R Archive Network [rede de arquivos do R, em tradução livre]).	V	X
		F	
14	O RStudio possui uma versão free, em que não é necessário pagar pelo <i>software</i> .	V	X
		F	

15	A instalação do R só pode ser realizada se não houver acesso à internet.	V	
		F	X