

SOEP Remote Guide

Inhaltsverzeichnis

1 Vorwort	2
2 Vorbereitung auf die Remote-Nutzung	3
2.1 Empfohlene Vorbereitungsmaterialien	3
2.2 Zentrale Arbeitsschritte im Überblick	4
3 Nutzungsregeln an den externen Arbeitsplätzen	5
4 Nutzung der virtuellen Infrastruktur	6

Table of Contents

1 Preface	13
2 Preparing for Remote Use	14
2.1 Recommended Preparatory Materials	14
2.2 Key Steps at a Glance	15
3 Rules of Use at the External Workplaces	16
4 Using the Virtual Infrastructure	17

1 Vorwort

Liebe Nutzerinnen und Nutzer,

um den Einstieg in die Arbeit mit den SOEP-Daten im Rahmen des Remote Access von einem Thin-Client aus möglichst reibungslos zu gestalten, stellen wir Ihnen in der vorliegenden Handreichung die wesentlichen Informationen zu dieser Zugriffsmethode zusammen. Die Handreichung umfasst sowohl organisatorische und administrative Hinweise zu den externen Arbeitsplätzen als auch eine praxisorientierte Anleitung zur Nutzung unserer virtuellen Desktop-Infrastruktur für den Datenzugriff von einem Gastarbeitsplatz.

Da Sie bei der Remote-Nutzung der SOEP-Daten nicht bei uns vor Ort sind, ist es uns ein besonderes Anliegen, dass Sie sich bei Fragen oder Problemen – sowohl im Hinblick auf die Daten als auch auf den Remote Access – jederzeit an uns wenden können. Für Rückfragen steht Ihnen unser Community Management gerne zur Verfügung:

E-Mail: soepmail@diw.de

Telefon: +49 30 89789-292

Über Rückmeldungen und Anregungen freuen wir uns jederzeit.

Ihr SOEP-Team

2 Vorbereitung auf die Remote-Nutzung

Sofern Sie die SOEP-Daten an einer externen Institution analysieren, erfolgt die Datenanalyse in der Regel an Terminals innerhalb eines Datensicherheitsraums. In diese Räume dürfen keine Laptops, Mobiltelefone oder sonstigen elektronischen Geräte mitgeführt werden, über die Daten nach außen gelangen könnten. Wir empfehlen daher, sich bereits im Vorfeld mit der Struktur und den Inhalten der SOEP-Daten vertraut zu machen.

Hinweis: Über SOEP Remote können SOEP-Daten mit regionalen Indikatoren bis auf Gemeindeebene analysiert werden. Für Analysen mit exakten Haushaltskoordinaten ist weiterhin eine Nutzung vor Ort bei uns erforderlich. Die hier gegebenen Hinweise sind auch für die Nutzung vor Ort relevant

2.1 Empfohlene Vorbereitungsmaterialien

Hierfür empfehlen wir in erster Linie den [SOEP Companion](#). Anhand zahlreicher praxisnaher Beispiele – inklusive Stata und R-Skripte – vermittelt dieser bereits im Vorfeld einen Einblick in den Aufbau der SOEP-Daten, etwa hinsichtlich der Zusammenführung der Datensätze, der Verortung einzelner Variablen, der Arbeit mit regionalen Indikatoren sowie des Exports von Analyseergebnissen. Insbesondere die Abschnitte „[Data Structure of SOEP-Core](#)“ und „[Working with SOEP Data](#)“ bieten hilfreiche Hinweise für Forschende, die erstmals mit den Daten arbeiten.

Darüber hinaus finden Sie unter [paneldata.org](#) eine umfassende Dokumentation der SOEP-Daten. Diese ermöglicht es, gezielt nach Variablen und Erhebungsinstrumenten zu recherchieren sowie mithilfe eines Baskets die ausgewählten Variablen als csv Tabelle zu exportieren.

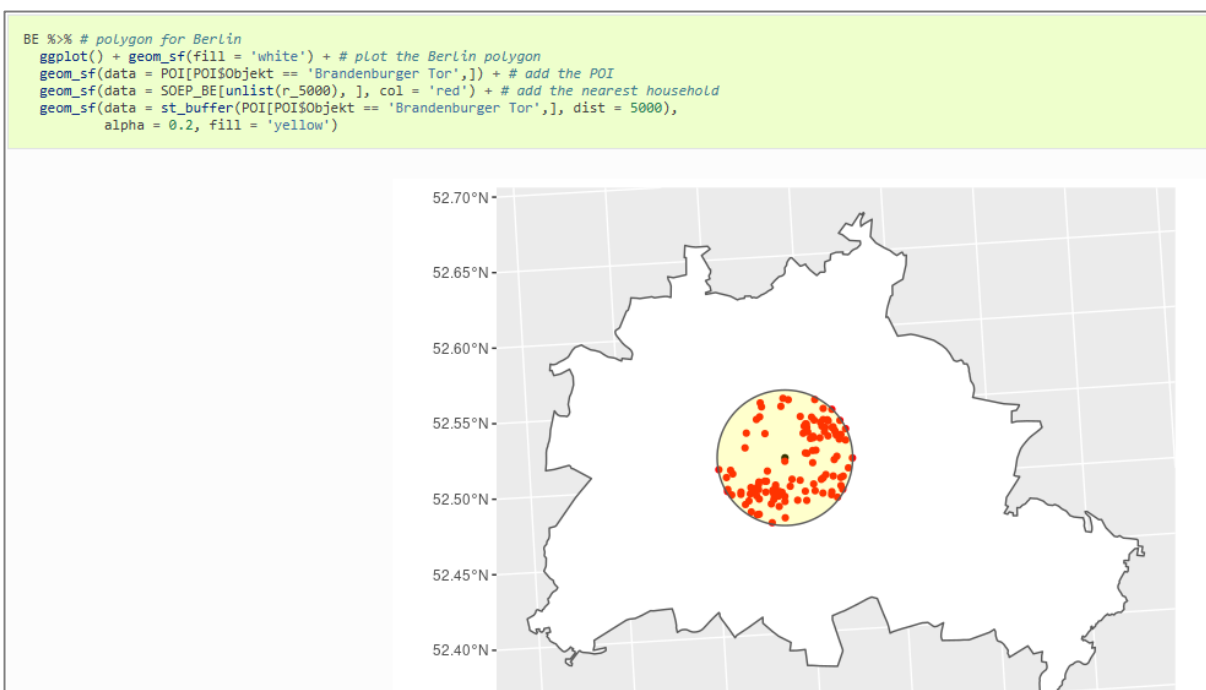


Abbildung 1: Beispiele im SOEP Companion

2.2 Zentrale Arbeitsschritte im Überblick

Zu den zentralen Arbeitsschritten bei der Nutzung der SOEP-Daten zählen unter anderem:

- **Arbeitsverzeichnis:** Die Datenanalyse erfolgt auf einem Server, auf dem Ihnen ein eigenes Arbeitsverzeichnis unter der Pfadstruktur `/home/USER/work/` zur Verfügung steht. Innerhalb dieses Verzeichnisses verfügen Sie über Lese- und Schreibrechte. „USER“ müssen Sie mit ihrem tatsächlichen Nutzernamen ersetzen.

Hinweis: Jeder Nutzer verfügt über einen Speicherplatz von 10 GB; dies ist bei der Planung der Datenanalyse zu berücksichtigen.

- **SOEP-Daten auf dem Server:** Zum Einlesen der SOEP-Daten stehen Ihnen R und Stata zur Verfügung. Die Daten befinden sich in den Verzeichnissen `~/soep-data/`.
- **R-Packages oder Stata ado-files importieren:** Sie können auf dem Server keine R-Packages oder Stata ado-files selbständig installieren. Bitte teilen sie unserem Community Management mindestens eine Woche vor Ihrem Aufenthalt mit, welche Erweiterungen sie benötigen.
- **Skripte oder externe Daten importieren:** Sollen im Vorfeld Skripte oder externe Daten in das Arbeitsverzeichnis importiert werden, sind diese **spätestens zwei Werktage vor dem geplanten Termin** an uns zu übermitteln. Die entsprechenden Vorgaben sowie die notwendigen Informationen hierzu erhalten Sie von unserem Community Management. Importierte Daten stehen anschließend im Verzeichnis `/home/USER/transfer/import/` zur Verfügung.
- **Daten exportieren:** Ein selbstständiger Export von Daten oder Ergebnissen ist zu keinem Zeitpunkt möglich. Jeder Export ist vorab bei uns zu beantragen, damit eine entsprechende Ergebniskontrolle (Output-Kontrolle) durchgeführt werden kann. Nähere Informationen hierzu stellt Ihnen unser Community Management zur Verfügung. Eine Anleitung finden sie im SOEP Companion im Abschnitt [„Instructions for exporting from Hauser to user“](#)

3 Nutzungsregeln an den externen Arbeitsplätzen

Um die Sicherheit der Forschungsdaten stets gewährleisten zu können, gelten an den Gastwissenschaftsarbeitsplätzen sowie innerhalb der Datensicherheitsräume Sicherheitskriterien, die von den Nutzenden einzuhalten sind. Diese können sich zwischen den einzelnen Institutionen bzw. Standorten im Detail unterscheiden, umfassen jedoch in der Regel die folgenden Grundsätze:

1. Vor dem erstmaligen Zugang erfolgt eine Identitätskontrolle; ein amtlicher Lichtbildausweis ist daher mitzuführen. Der Zutritt zu den Räumlichkeiten ist ausschließlich autorisierten Nutzenden gestattet. Dies setzt sowohl einen gültigen Nutzungsvertrag mit dem SOEP als auch eine mit der bereitstellenden Institution vereinbarte Terminbuchung voraus.
2. Elektronische Geräte wie Laptops, Smartphones oder Tablets dürfen nicht in den Datensicherheitsraum mitgeführt werden. In der Regel wird den Nutzenden für die Aufbewahrung persönlicher Gegenstände ein Schließfach zur Verfügung gestellt.
3. Die jeweiligen Öffnungszeiten der Institution sind einzuhalten.
4. Jeglicher Versuch, Forschungsdaten außerhalb des Datensicherheitsraums zugänglich zu machen oder anderen Personen innerhalb des Datensicherheitsraums Zugriff auf die Daten zu gewähren, ist untersagt.
5. Die Einhaltung dieser Vorgaben wird im Datensicherheitsraum in regelmäßigen Abständen kontrolliert.
6. In der Regel unterzeichnen Nutzende vor Ort eine Nutzungsvereinbarung des GWAP, mit der die Einhaltung der genannten Sicherheitskriterien bestätigt wird.

Hinweis: Ein Verstoß gegen die vorgenannten Regelungen führt zum sofortigen und dauerhaften Ausschluss von der Nutzung des betreffenden sowie sämtlicher weiterer Gastwissenschaftsarbeitsplätze, an denen SOEP-Daten verfügbar sind.

Neben den hier definierten Regelungen können zusätzliche Kriterien durch die Institution umgesetzt werden, die den Gastwissenschaftsarbeitsplatz bereitstellt. Diese können auch vorsehen, dass vor dem eigentlichen Termin eine Einweisung oder Schulung stattfindet; hierüber werden Sie bereits im Rahmen der Terminvereinbarung informiert.

4 Nutzung der virtuellen Infrastruktur

Technischer Hintergrund: Für den Fernzugriff wird ein sogenanntes Remote-Desktop-System verwendet. Dabei erhalten Sie Zugriff auf einen virtuellen Desktop, über den Sie auf den SOEP-Server zugreifen und Ihre Analysen durchführen. Der virtuelle Desktop selbst enthält keine nutzbaren Funktionen oder Daten, sondern führt lediglich Befehle auf unserem Server aus – er fungiert als sogenannter Jump-Host. In der Praxis starten Sie auf dem Desktop verschiedene Programme, die jedoch auf unserem Server ausgeführt werden; Sie sehen und steuern also ausschließlich, was auf dem Server geschieht. Das bedeutet auch, dass mit einer geringen Latenz zu rechnen ist.

4.1 Anmeldung und virtueller Desktop

Der Zugang zum virtuellen Desktop erfolgt in wenigen Schritten:

1. Anmeldung über Omnissa Horizon: Starten Sie Omnissa Horizon. Symbol und Bezeichnung können je nach Standort und Terminal variieren; fragen Sie im Zweifel das Personal vor Ort nach dem korrekten Icon.

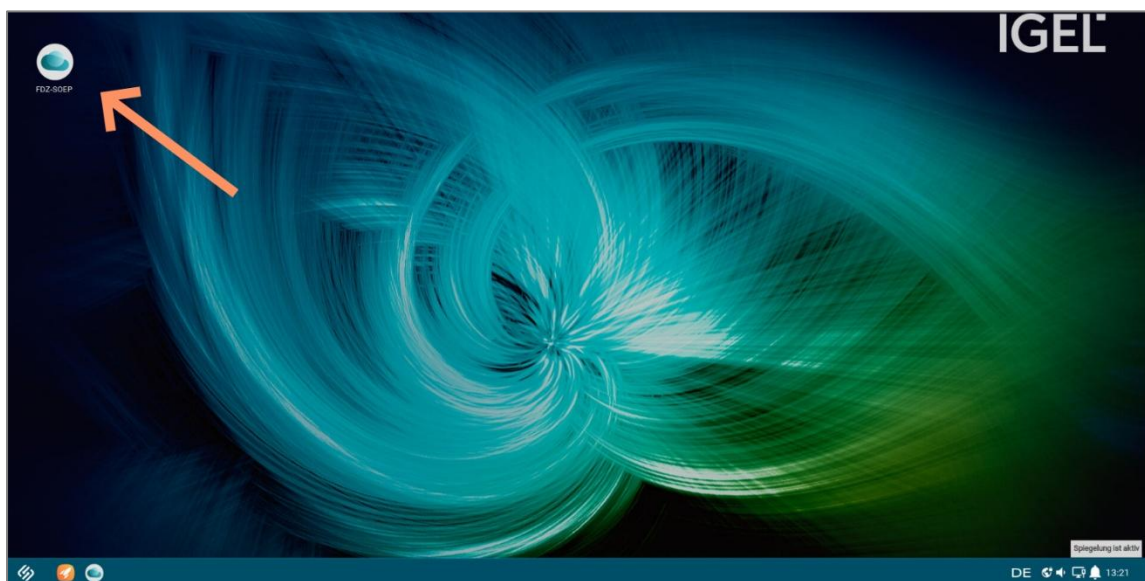


Abbildung 2: Desktop mit Horizon Anwendung

2. Wählen Sie den vordefinierten Zielserver „vdi.soep.de“ aus.



Abbildung 3: Zielserver auswählen

3. Geben Sie den Benutzernamen und das Passwort ein, dass Ihnen vom SOEP Community Management mitgeteilt wurde.

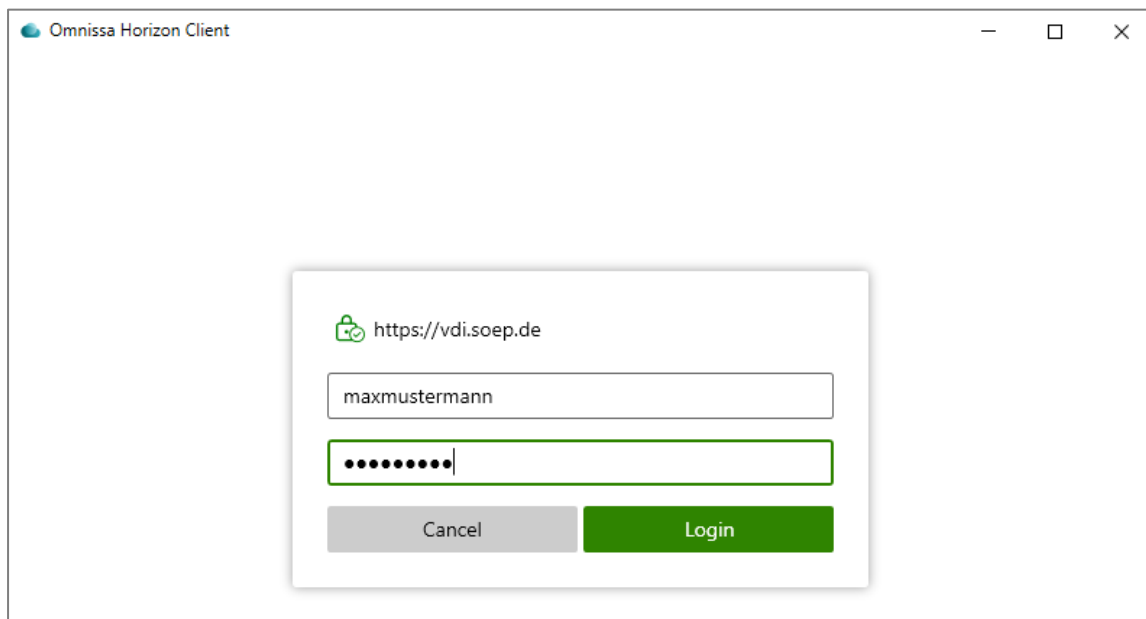


Abbildung 4: Benutzerdaten eingeben

4. Desktop-Pool auswählen: Wählen Sie nach der Anmeldung den angezeigten Desktop-Pool aus (z. B. FDZ SOEP).

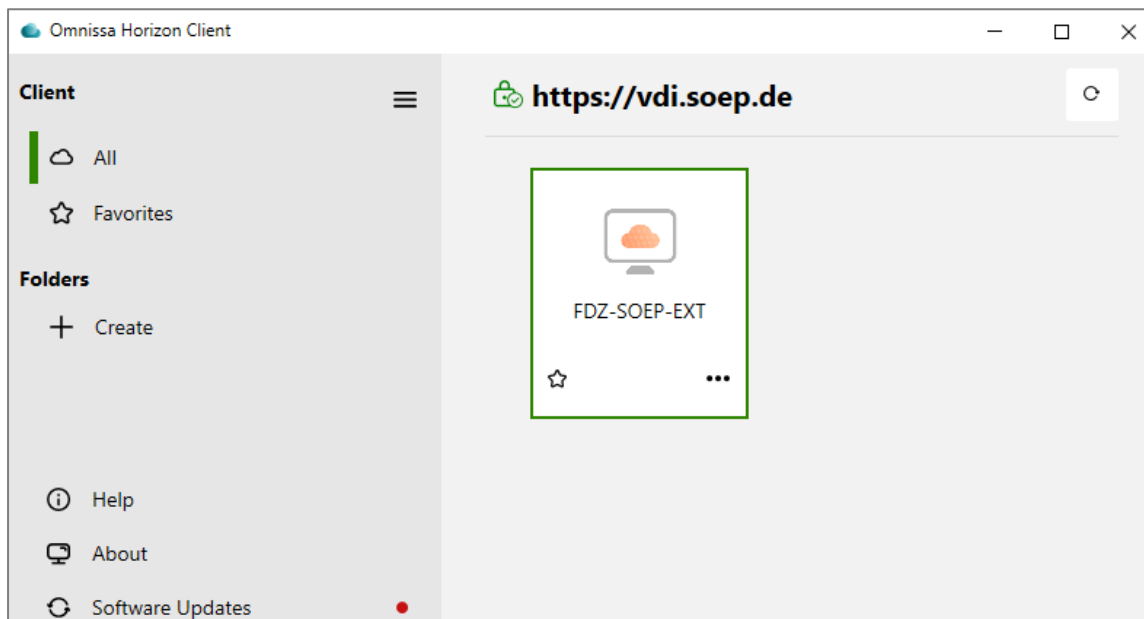


Abbildung 5: Auswahl Desktop-Pool

5. Sie befinden sich nun auf dem virtuellen Desktop, über den Sie Ihre Analysen auf dem SOEP-Server durchführen. Durch Klick auf die Symbole starten Sie die einzelnen Anwendungen auf dem SOEP-Server (siehe Abschnitt 4.2).

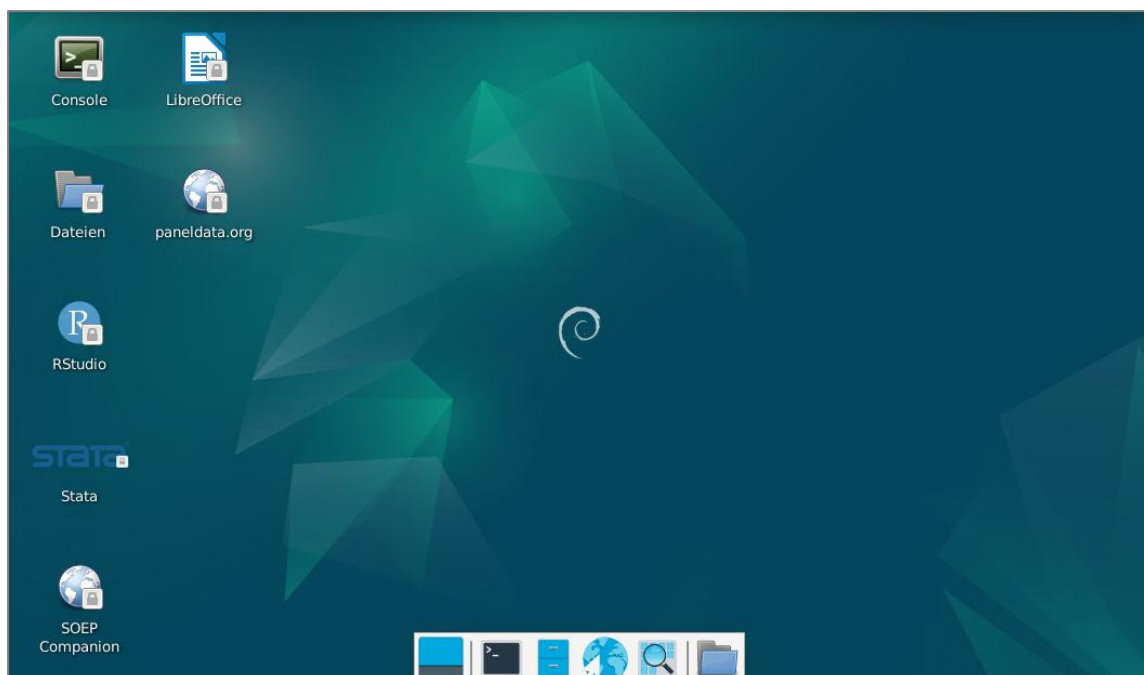


Abbildung 6: Virtueller Desktop

4.2 Verfügbare Anwendungen

Beim erstmaligen Ausführen einer Anwendung (oder nach einer erneuten Anmeldung) öffnet sich eine Konsole, in der Sie Ihr Passwort erneut eingeben und mit Enter bestätigen. Verwenden Sie dasselbe Passwort wie bei der Anmeldung.

Hinweis: Aus Sicherheitsgründen wird die Passworteingabe nicht angezeigt – die Eingabe erfolgt „blind“. Das System registriert Ihre Eingabe dennoch.

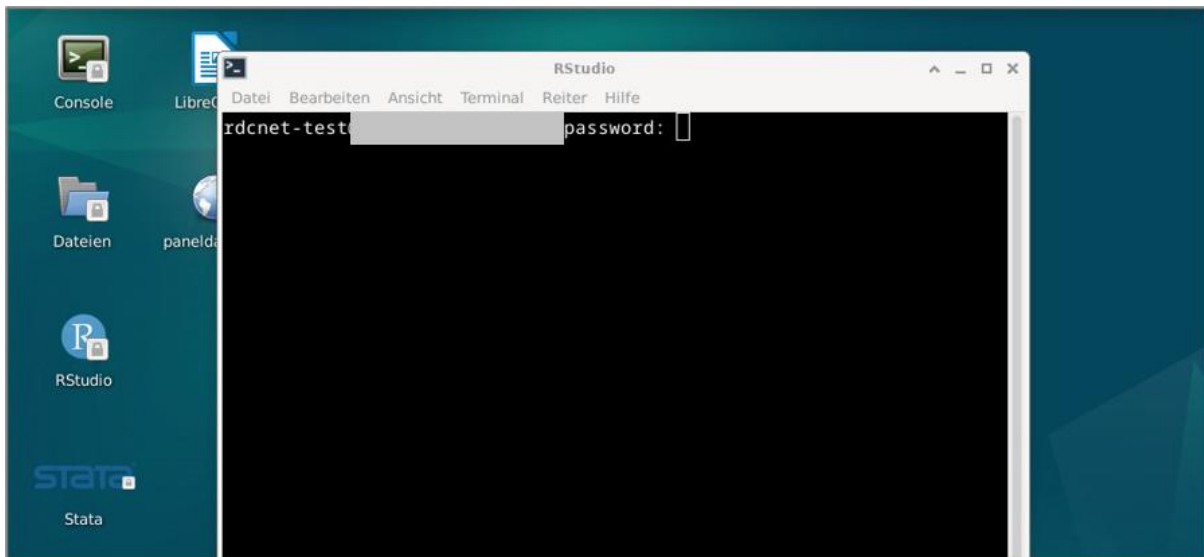
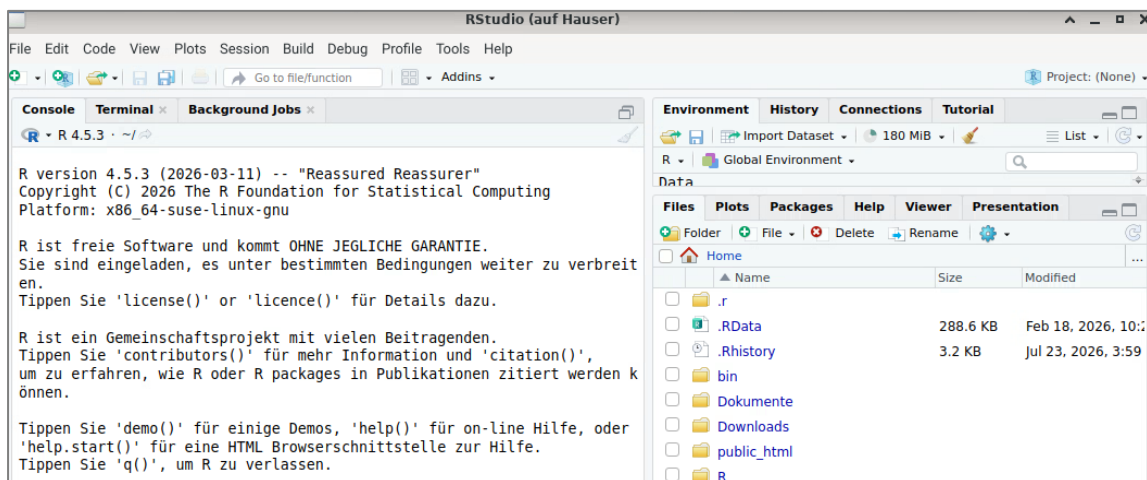


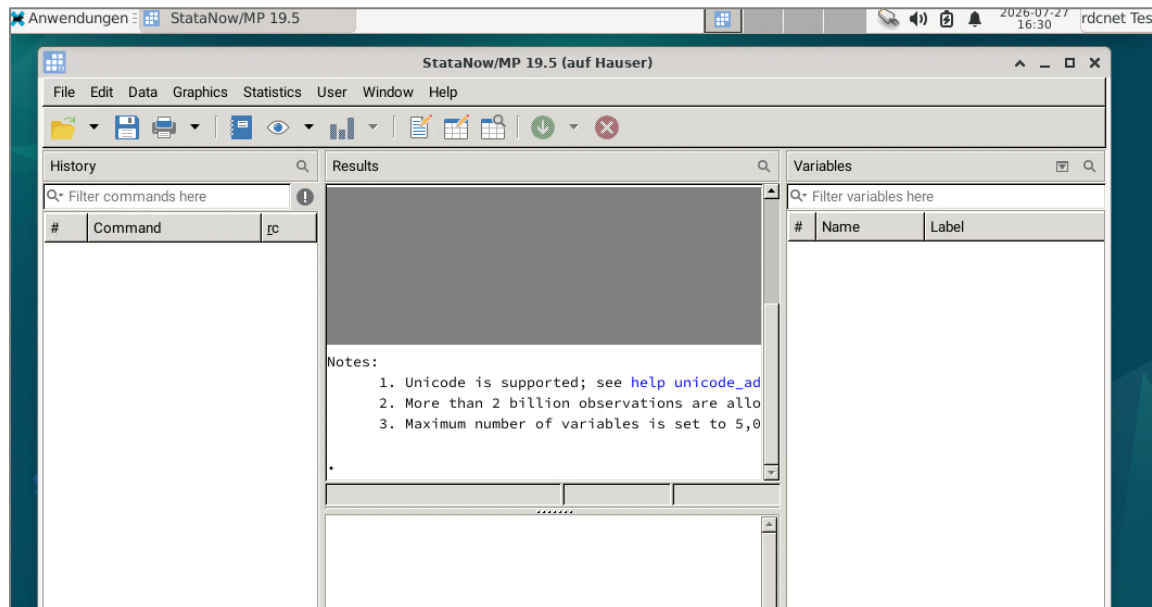
Abbildung 7: Passworteingabe beim Starten von Anwendungen

Auf dem virtuellen Desktop stehen Ihnen die folgenden Anwendungen zur Verfügung:

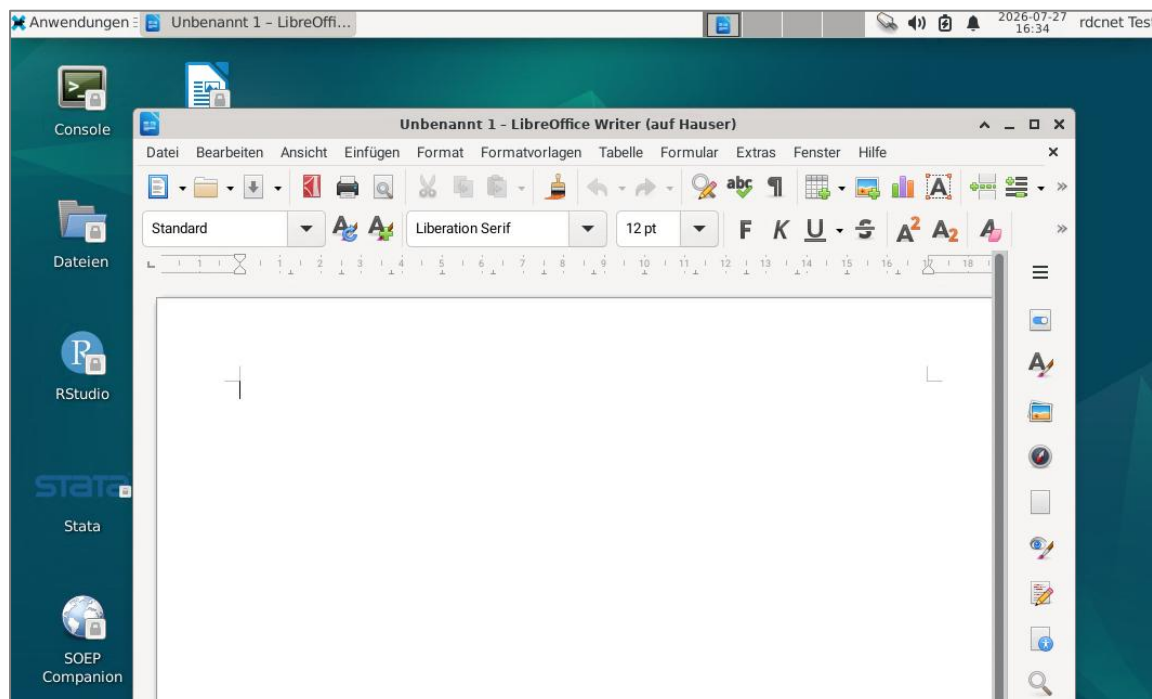
RStudio – Falls Sie Pakete benötigen, die sich nicht auf unserem Server befinden, wenden Sie sich bitte an das SOEP Community Management.



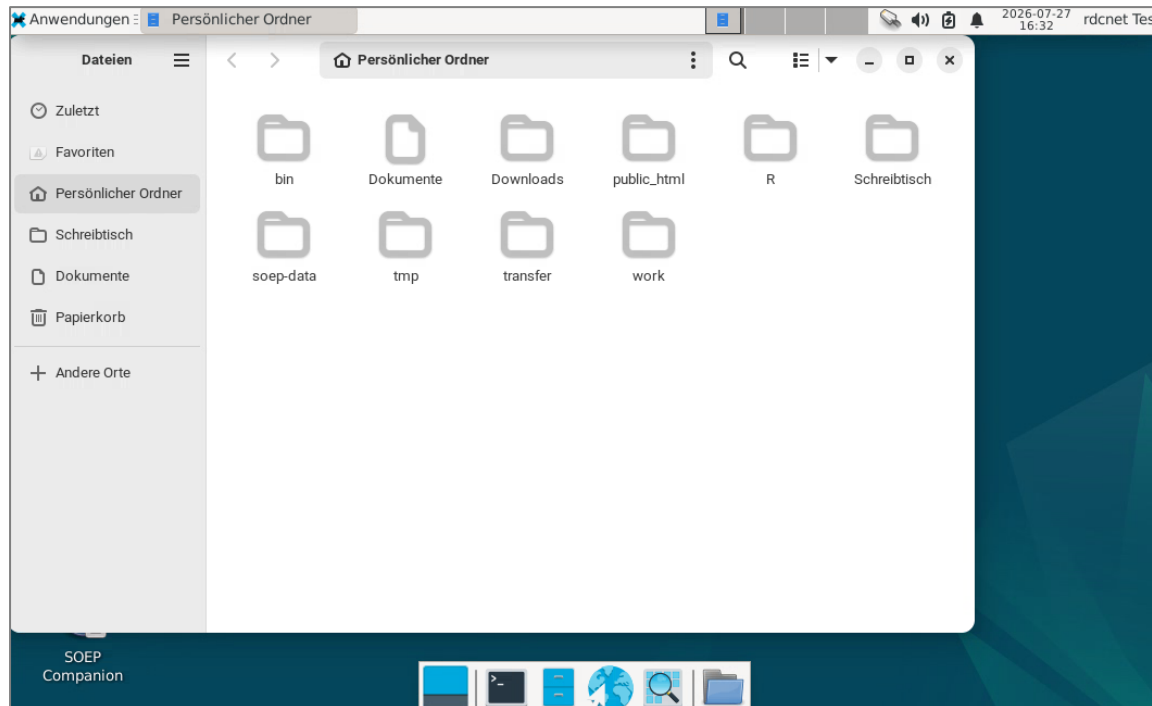
Stata MP – Falls Sie Pakete benötigen, die sich nicht auf unserem Server befinden, wenden Sie sich bitte an das SOEP Community Management.



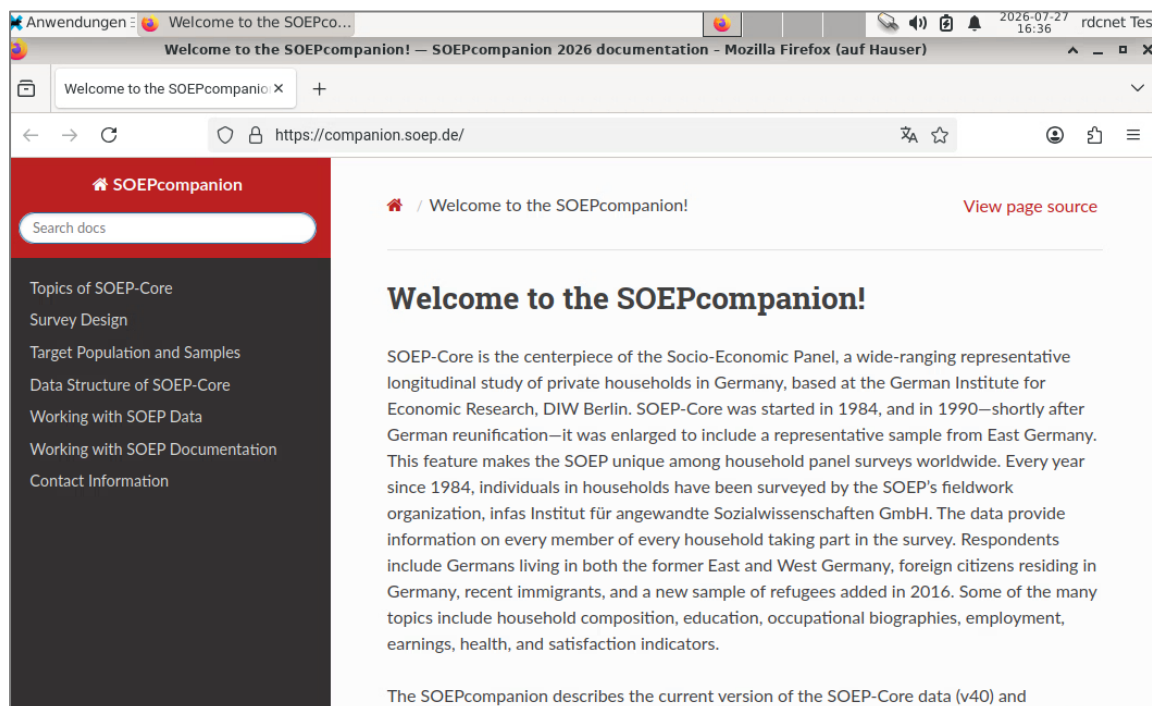
LibreOffice – Textverarbeitung, z. B. für Notizen oder Dokumentation.



Datei-Explorer – zum Durchsuchen Ihres Arbeitsverzeichnisses. Beim ersten Start benötigt der Datei-Explorer einen Moment zur Initialisierung; bitte warten Sie kurz.



SOEP Companion & paneldata.org – über Firefox haben Sie Zugriff auf companion.soep.de und paneldata.org. Andere Websites sind gesperrt.





Console – über die Konsole können die Anwendungen ebenfalls per Befehl gestartet werden. Die entsprechenden Befehle lauten:

BEFEHL	ANWENDUNG
rstudio	RStudio öffnen
xstata-mp	Stata öffnen
lowriter	LibreOffice (Textverarbeitung) öffnen
nautilus	Datei-Explorer öffnen
firefox	Firefox öffnen (Zugriff auf paneldata.org und companion.soep.de)

1 Preface

Dear users,

To make getting started with the SOEP data via remote access at a thin-client as smooth as possible, this guide brings together the key information on this method of access. It covers both organisational and administrative information on the external workplaces (also called “guest researcher workplaces”) and practical instructions for using our virtual desktop infrastructure to access the data.

Since you are not on site with us when working with the SOEP data remotely, it is particularly important to us that you can contact us at any time with questions or problems – whether these concern the data or remote access itself. Our Community Management team will be glad to help:

Email: soepmail@diw.de

Phone: +49 30 89789-292

We always welcome your feedback and suggestions.

Your SOEP Team

2 Preparing for Remote Use

If you analyse the SOEP data at an external institution, the analysis is usually carried out at terminals within a secure data room. Laptops, mobile phones and other electronic devices that could be used to transfer data out of the room may not be taken into these rooms. We therefore recommend familiarising yourself with the structure and content of the SOEP data in advance.

Note: SOEP Remote allows SOEP data to be analysed with regional indicators down to municipality level. Analyses using exact household coordinates must still be carried out on site with us. The guidance provided here is also relevant for on-site use.

2.1 Recommended Preparatory Materials

We primarily recommend the [SOEP Companion](#). Using numerous practical examples – including Stata and R scripts – it provides an early insight into the structure of the SOEP data, for example on merging datasets, locating individual variables, working with regional indicators and exporting analysis results. The section [“Working with SOEP Data”](#) and [„Data Structure of SOEP-Core“](#) offer helpful guidance for researchers working with the data for the first time.

In addition, paneldata.org provides comprehensive documentation of the SOEP data. It allows you to search specifically for variables and survey instruments as well as to export the selected variables as a CSV table using a basket.

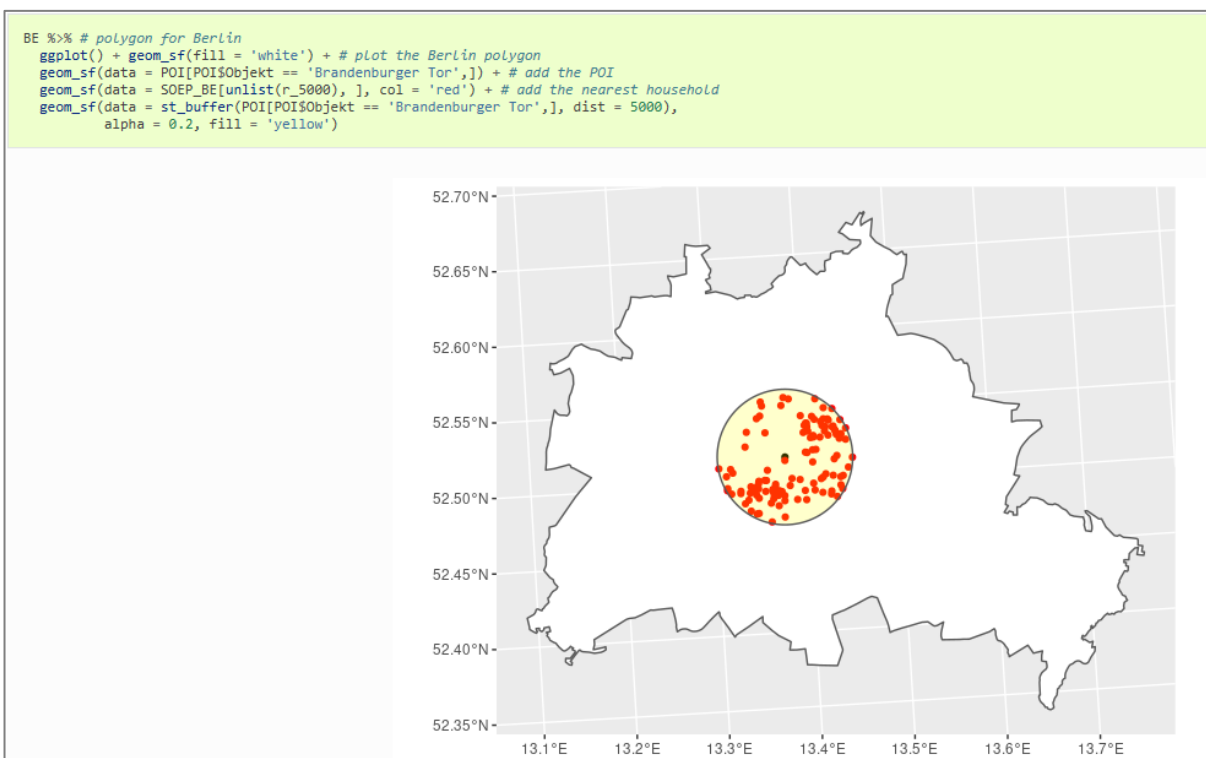


Figure 1: Examples in the SOEP Companion

2.2 Key Steps at a Glance

The key steps when working with the SOEP data include:

- **Working directory:** The analysis is carried out on a server on which you have your own working directory under the path `/home/USER/work/`. You have read and write permissions within this directory. You must replace "USER" with your actual username.

Note: Each user has 10 GB of storage space; please take this into account when planning your analysis.

- **Importing R packages or Stata ado-files:** You cannot install R packages or Stata ado-files on the server yourself. Please inform our Community Management at least one week before your stay which extensions you require.
- **Importing SOEP data:** R and Stata are available for reading in the SOEP data. The data are located in the directories `~/soep-data/`.
- **Importing scripts or external data:** If scripts or external data are to be imported into your working directory beforehand, they must be sent to us no later than two working days before your scheduled appointment. You will receive the relevant requirements and all necessary information from our Community Management team. Imported data are then available in the directory `/home/USER/transfer/import/`.
- **Exporting data:** You may not export data or results yourself at any time. Every export must be requested from us in advance so that the necessary output control can be carried out. Our Community Management team will provide you with further information. You can find instructions in the SOEP Companion in the section ["Instructions for exporting from Hauser to user"](#).

3 Rules of Use at the External Workplaces

To always ensure the security of the research data, users must comply with the security criteria that apply at the guest researcher workplaces and within the secure data rooms. These may differ in detail between individual institutions and locations, but generally comprise the following principles:

1. Before your first visit, your identity will be verified; please therefore bring an official photo ID. Access to the premises is granted to authorised users only. This requires both a valid data use agreement with the SOEP and an appointment booked with the workplace providing institution.
2. Electronic devices such as laptops, smartphones or tablets may not be taken into the secure data room. Users are generally provided with a locker for storing personal belongings.
3. The opening hours of the respective institution must be observed.
4. Any attempt to make research data accessible outside the secure data room, or to give other people inside the secure data room access to the data, is prohibited.
5. Compliance with these requirements is checked at regular intervals in the secure data room.
6. Users generally sign a user agreement on site, confirming that they will comply with the security criteria set out above.

Note: Any breach of the above rules will result in immediate and permanent exclusion from the guest researcher workplace concerned and from all other workplaces at which SOEP data are available.

In addition to the rules set out here, the institution providing the guest researcher workplace may implement further criteria. These may also provide for an induction or training session to take place before your actual appointment; you will be informed of this when the appointment is arranged.

4 Using the Virtual Infrastructure

Technical background: Remote access is provided via a remote desktop system. You are given access to a virtual desktop through which you connect to the SOEP server and run your analyses. The virtual desktop itself contains no usable functions or data; it merely executes commands on our server – it acts as a so-called jump host. In practice, you start various applications on the desktop, but these run on our server; you can therefore only see and control what happens on the server. This also means that a slight latency may be noticeable.

4.1 Login on the Virtual Desktop

Access to the virtual desktop takes just a few steps:

1. Log in via Omnisia Horizon: Start Omnisia Horizon. The icon and its label may vary depending on the location and terminal; if in doubt, ask the staff on site which icon should be used.

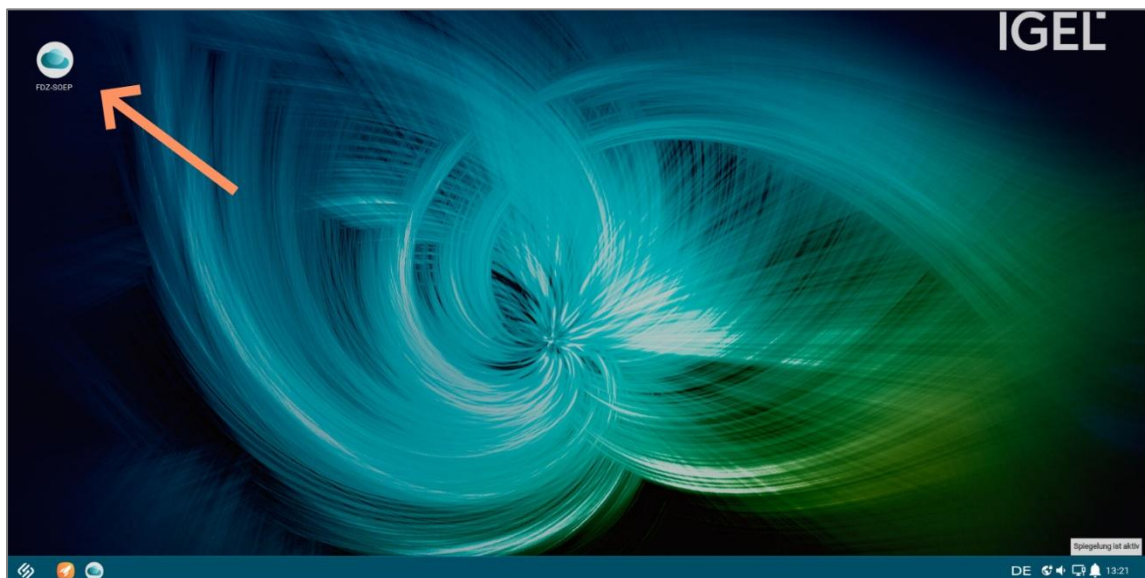


Figure 2: Desktop with the Horizon application

2. Select the predefined target server “vdi.soep.de”.



3. Enter the username and password provided to you by the SOEP Community Management.

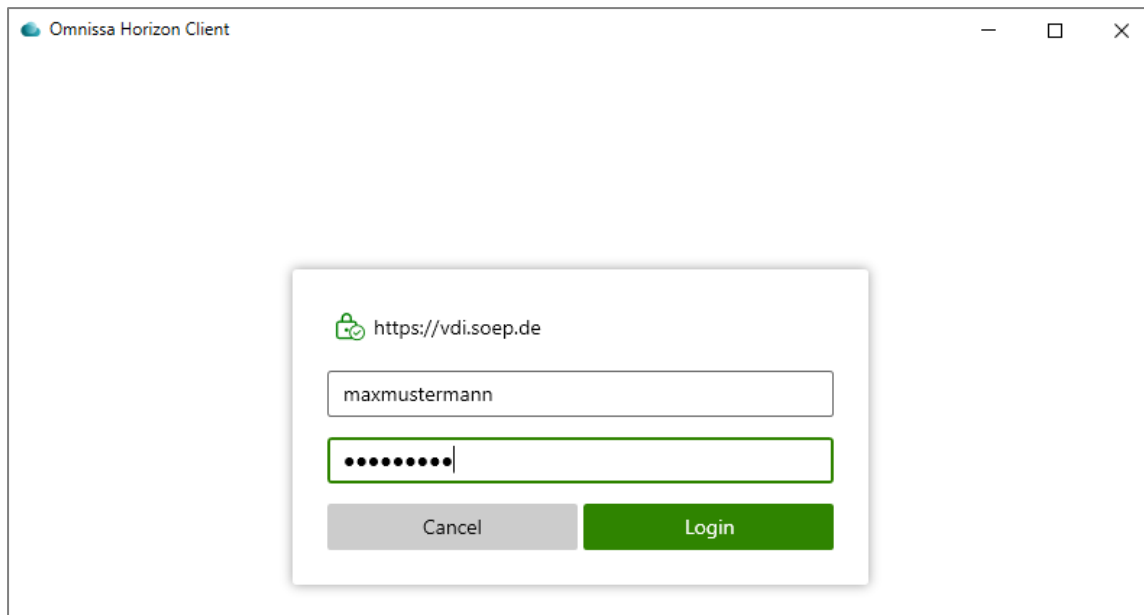


Figure 3: Entering your credentials

4. Select the desktop pool: After logging in, select the desktop pool shown (e.g. FDZ SOEP).

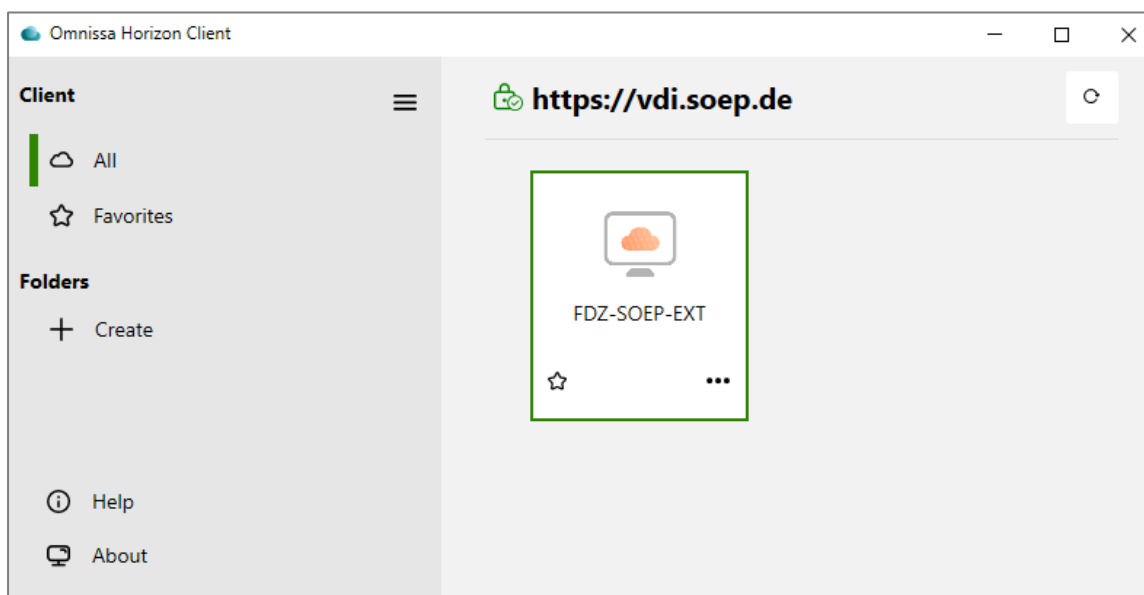


Figure 4: Selecting the desktop pool

5. You are now on the virtual desktop, through which you run your analyses on the SOEP server. Click the icons to start the individual applications (see Section 4.2).

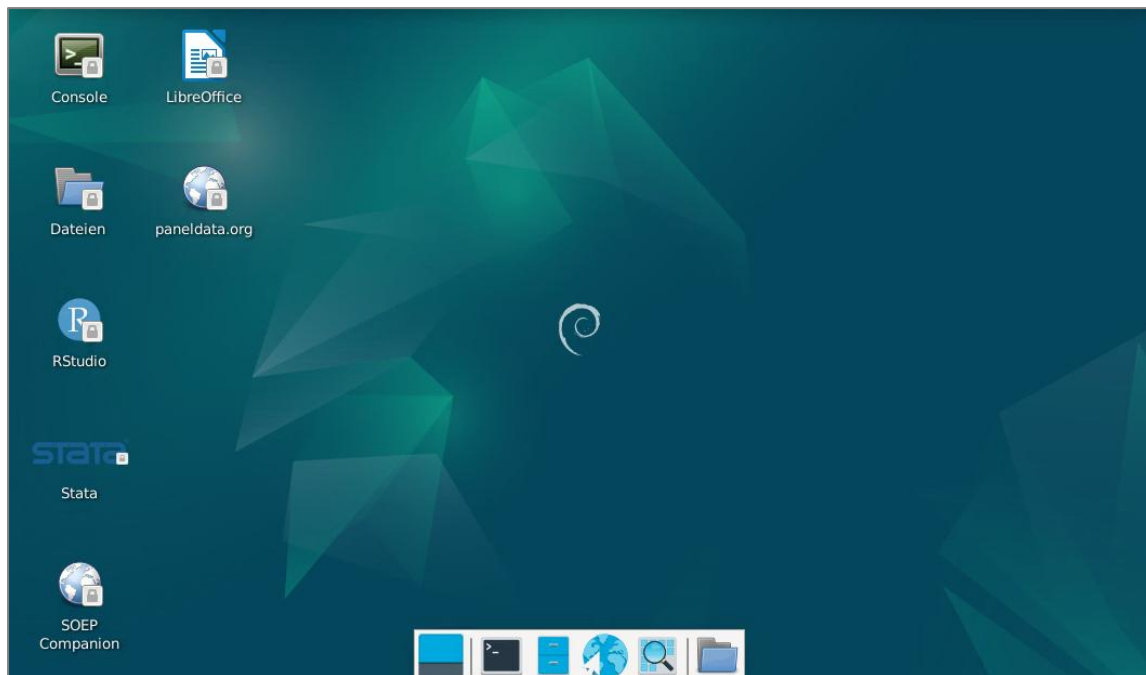


Figure 5: The virtual desktop

4.2 Available Applications

The first time you start an application (or after logging in again), a console opens in which you re-enter your password and confirm with Enter. Use the same password as for logging in.

Note: For security reasons, the password is **not shown** as you type – entry is “blind”. The system registers your input nonetheless.

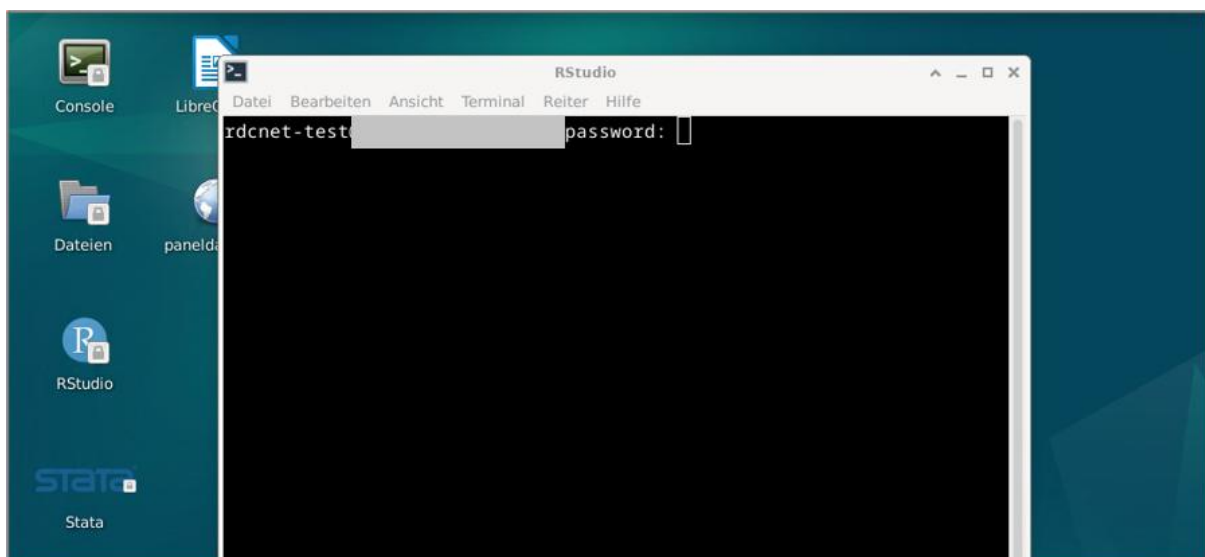
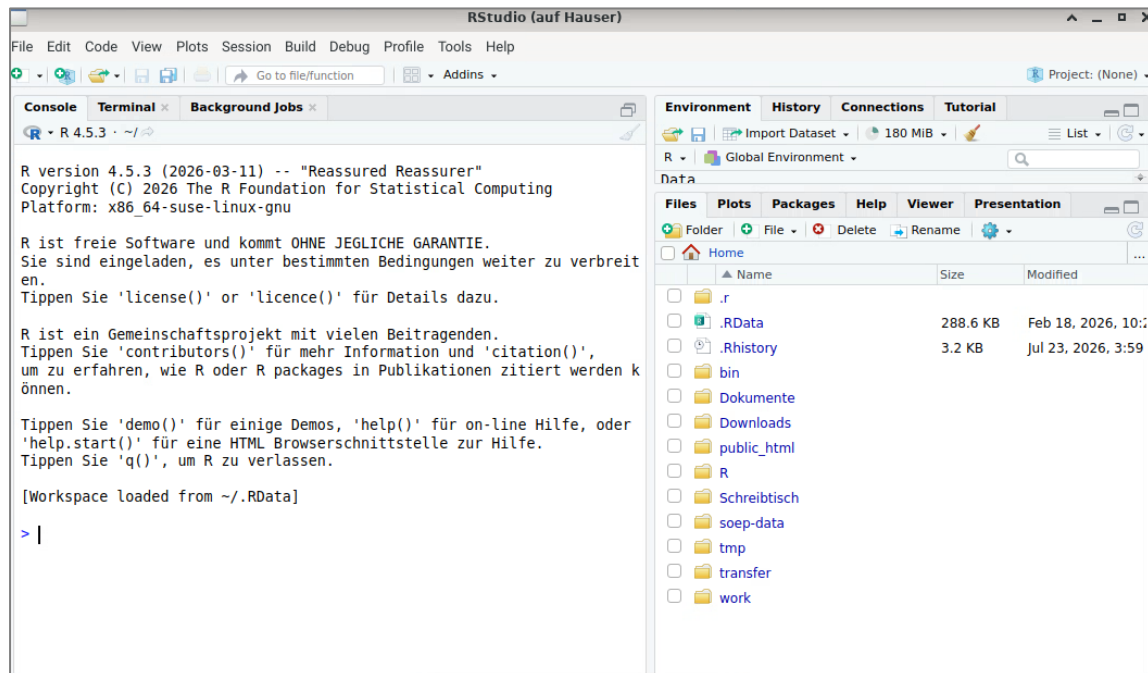


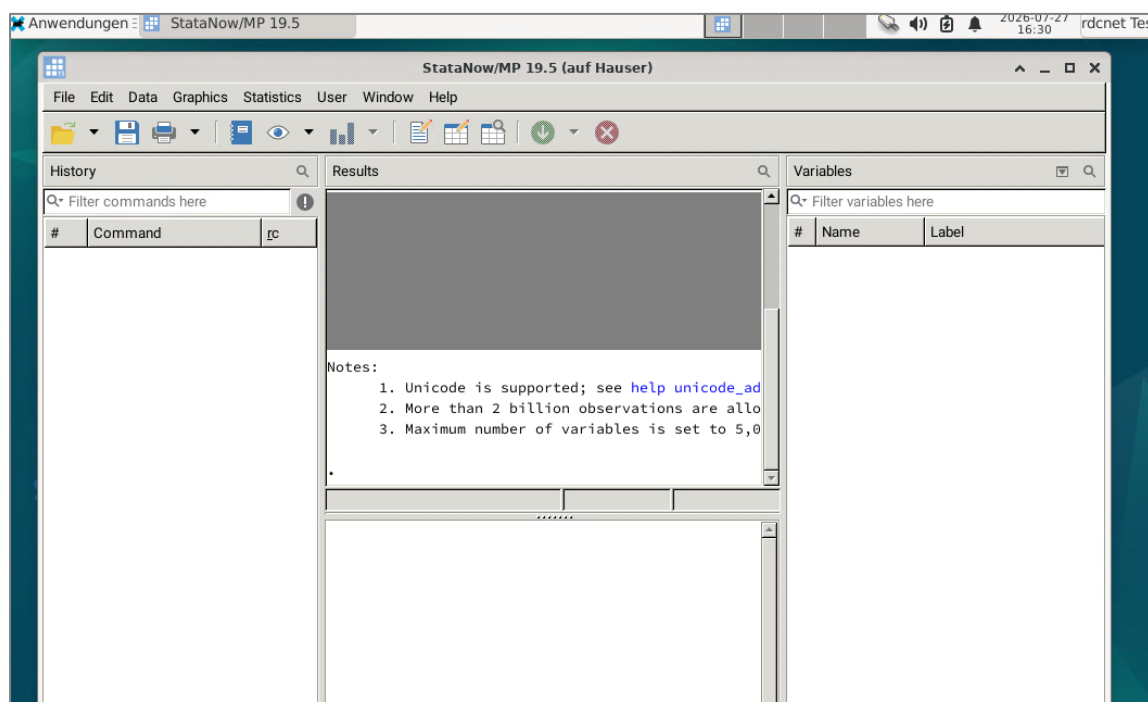
Figure 6: Password entry when starting an application

The following applications are available on the virtual desktop:

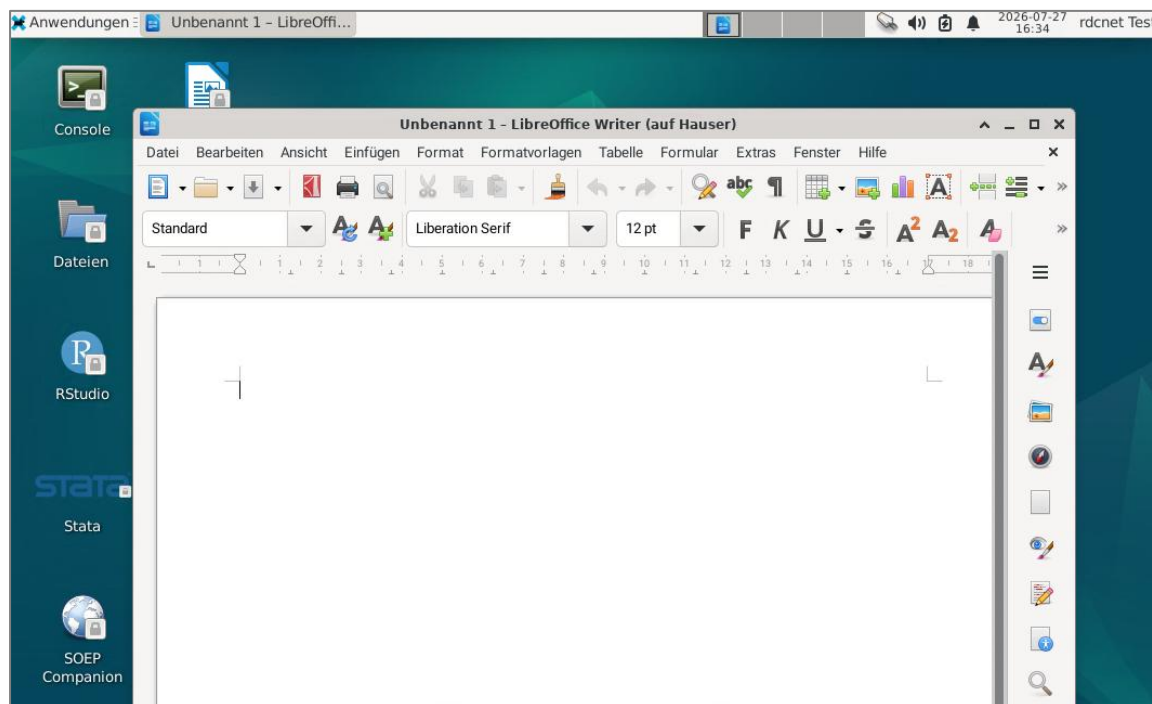
RStudio – If you need packages that are not available on our server, please contact the SOEP Community Management team.



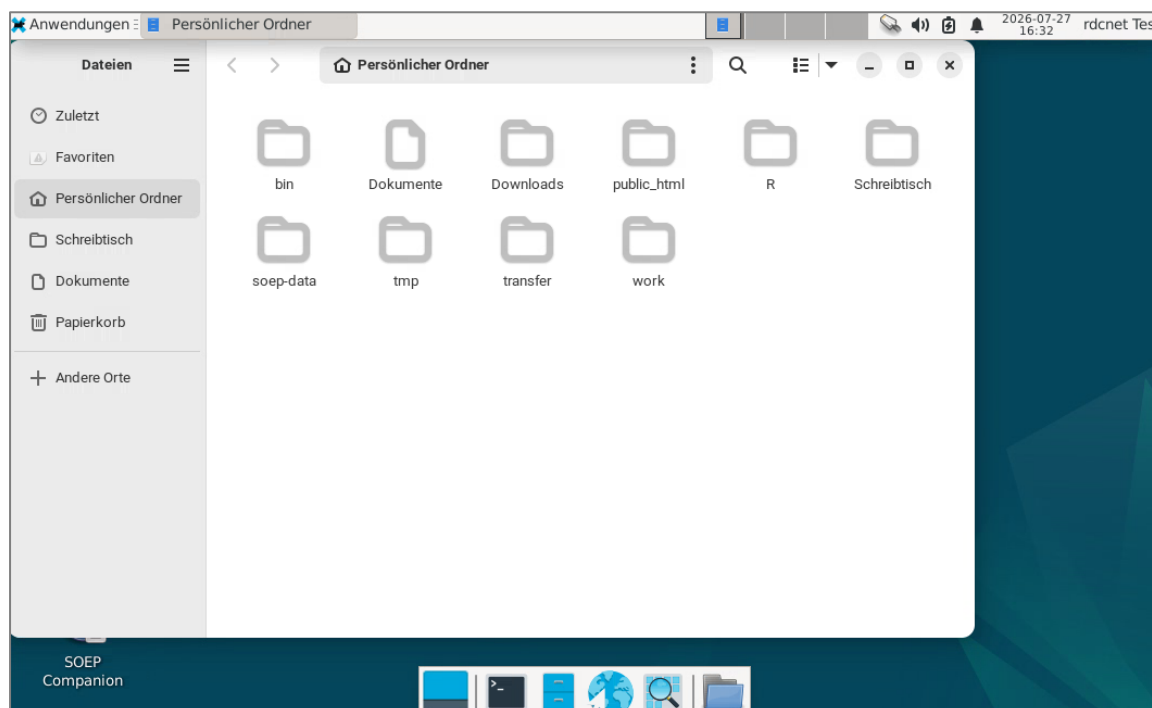
Stata MP – If you need packages that are not available on our server, please contact the SOEP Community Management team.



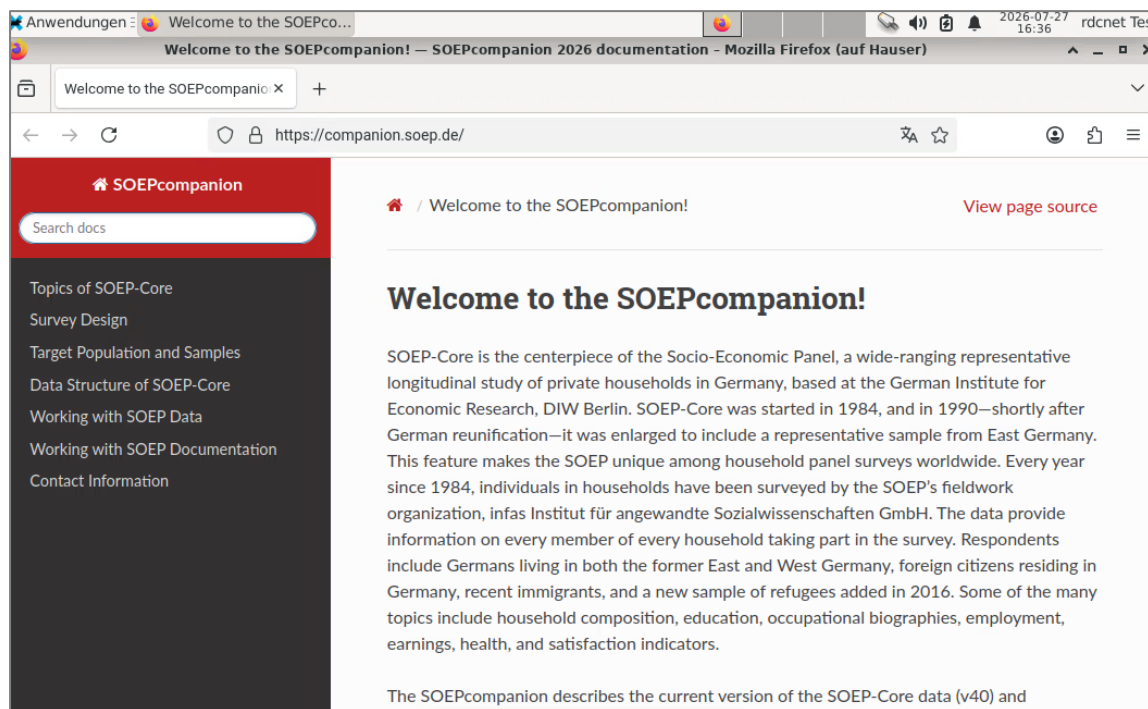
LibreOffice – word processing, e.g. for notes or documentation.



File Explorer – for browsing your working directory. On first start, the file explorer needs a moment to initialise; please wait briefly.



SOEP Companion & paneldata.org – via Firefox you have access to companion.soep.de and paneldata.org. Other websites are blocked.



Console – the console can also be used to start the applications via commands. The relevant commands are:

COMMAND	APPLICATION
rstudio	Open RStudio
xstata-mp	Open Stata
lowriter	Open LibreOffice (word processor)
nautilus	Open the file explorer
firefox	Open Firefox (access to paneldata.org and companion.soep.de)