The Examples Book

Contents

Introduction **5**
How to contribute 5

Scholar **7**
Connecting to Scholar 7
Resources 10

Unix **11**
Getting started 12
Standard utilities 12
Piping & Redirection 12
Emacs 12
Nano 12
Vim 12
Writing scripts 12

SQL **13**

R **15**
Getting started 15
Variables 16
Logical operators 16
Lists & Vectors 16
Basic R functions 16
Data.frames 16
Reading & Writing data 16
Control flow 16
Apply functions 16
Writing functions 16
Plotting 16
RMarkdown 16
Tidyverse 16
Data.table's 16
SQL in R 16

Scraping . . . 16
`shiny` . . . 16

Python **17**
Getting started . . . 18
Lists & Tuples . . . 18
Dicts . . . 18
Control flow . . . 18
Writing functions . . . 18
Reading & Writing data . . . 18
`numpy` . . . 18
`scipy` . . . 18
`pandas` . . . 18
Jupyter notebooks . . . 18
Writing scripts . . . 18
Scraping . . . 18
Plotting . . . 18
Classes . . . 18

Tools **19**
Docker . . . 19
Tableau . . . 19
GitHub . . . 19
VPNs . . . 19

FAQs **21**
How do I connect to Scholar from off-campus? . . . 21
Is there an advantage to using the ThinLinc client rather than the ThinLinc web client? . . . 21
GitHub Classroom is not working – can't authorize the account. . . . 21
In Scholar, my font size looks weird or my cursor is offset. . . . 21
How do I make the ThinLinc window bigger without going to the dreaded "full screen" mode? . . . 22
I'm unable to type into the terminal in RStudio. . . . 22
I'm unable to connect to RStudio Server. . . . 22
. . . 22

Contributors **23**

Introduction

This book contains a collection of examples that students can use to reinforce topics learned in The Data Mine seminar. It is an excellent resource for students to learn what they need to know in order to solve The Data Mine projects.

How to contribute

Contributing to this book is simple: 1. Navigate to the page or section that needs to be edited 2. Click on the "Edit" button towards the upper left side of the page: 3. You'll be presented with the respective RMarkdown file. Make your modifications. 4. In the "Commit changes" box, select the radio button that says *Create a **new branch** for this commit and start a pull request.* Give your pull request a title and a detailed description. Name the new branch, and click on "Propose file change". 5. You've successfully submitted a pull request. Our team will review and merge the request shortly thereafter.

Scholar

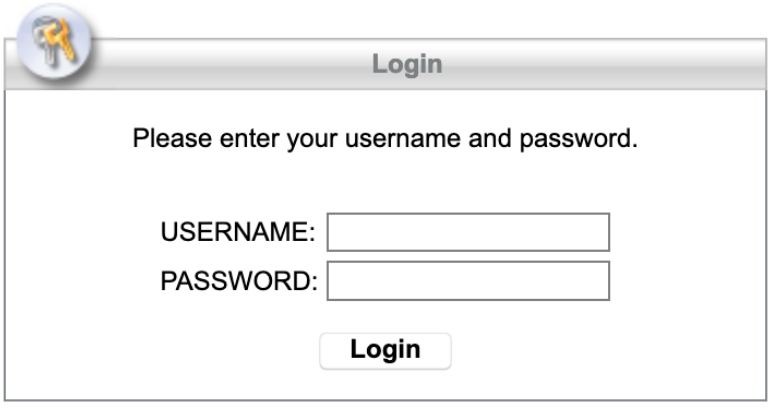
Connecting to Scholar

ThinLinc web client

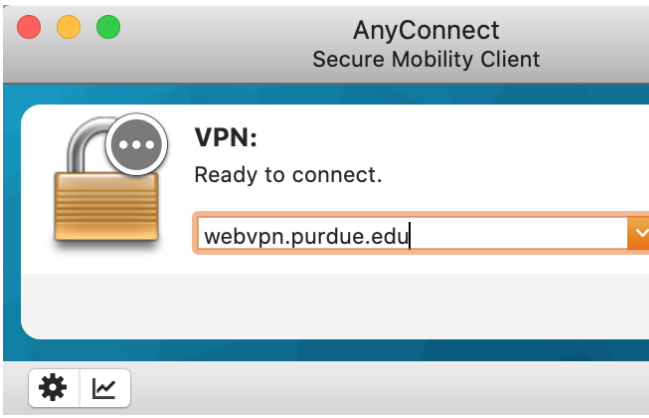
1. Open a browser and navigating to https://desktop.scholar.rcac.purdue.edu/.
2. Login with your Purdue Career Account credentials (**without** BoilerKey).
3. Congratulations, you should now be connected to Scholar using the ThinLinc web client.

ThinLinc client

1. Navigate to https://webvpn.purdue.edu/. You should see a login screen:

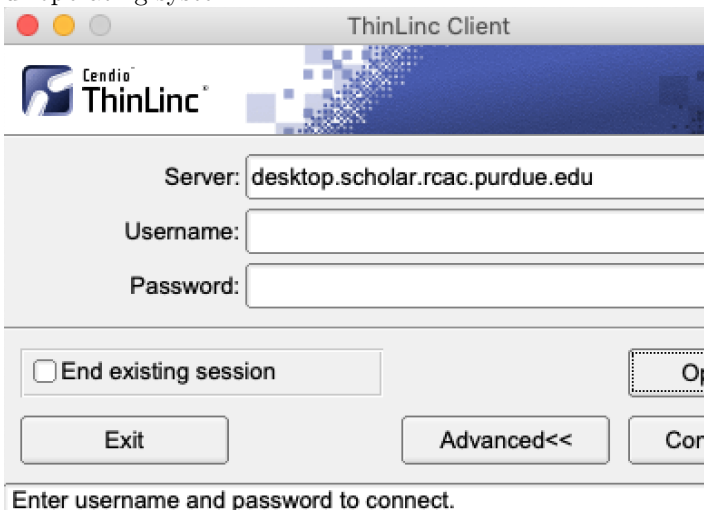


2. Enter your Purdue Career Account credentials (**with** BoilerKey).
3. Download and install the Cisco AnyConnect Secure Mobility Client.
4. Open the AnyConnect client and enter the domain for Purdue's vpn:

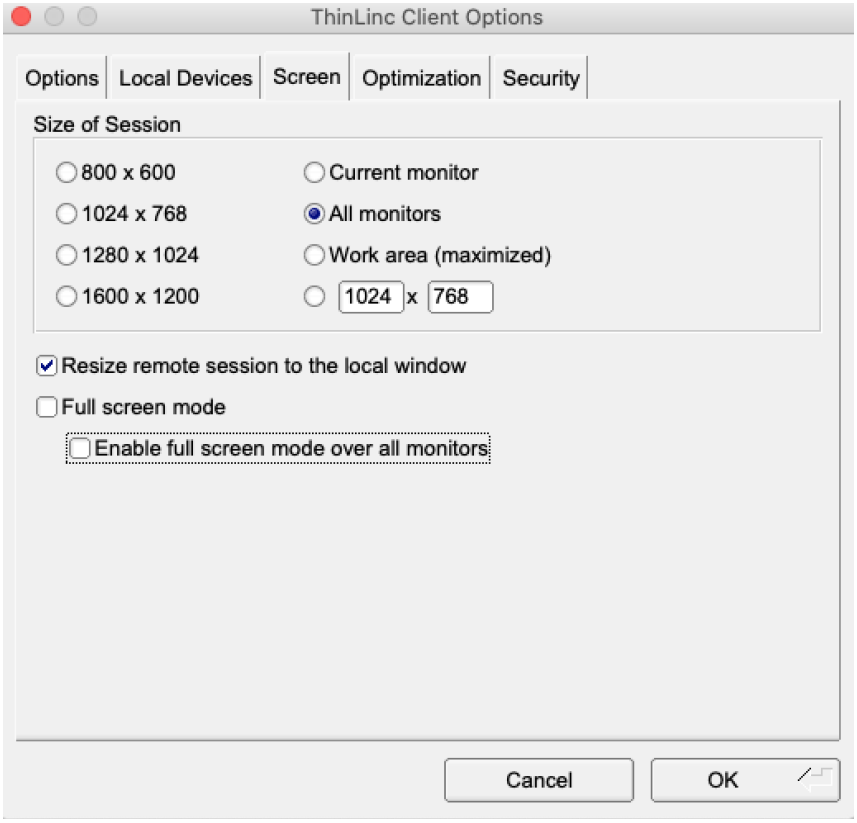


`webvpn.purdue.edu`, and click "Connect":

5. When prompted, enter your Purdue Career Account credentials (**with** BoilerKey).
6. You should be successfully connected to Purdue's VPN! You can read more about VPNs here.
7. Navigate to https://www.cendio.com/thinlinc/download, and download the ThinLinc client application for your operating system.



8. Install and launch the ThinLinc client:
9. Enter your Purdue Career Account information (**without** BoilerKey), as well as the server: `desktop.scholar.rcac.purdue.edu`.
10. Click on "Options..." and fill out the "Screen" tab as shown below:



11. Click "OK" and then "Connect". **Make sure you are connected to Purdue's VPN using AnyConnect before clicking "Connect"!**
12. If you are presented with a choice like below, click "Continue".

13. Congratulations, you are now successfully connected to Scholar using the ThinLinc client. **NOTE: If you do accidentally get stuck in full screen mode, the F8 key will help you to escape. NOTE: The very first time that you log onto Scholar, you will have an option**

of "use default config" or "one empty panel". PLEASE choose the "use default config".

SSH

Windows

MacOS

Linux

JupyterHub

1. Open a browser and navigate to https://notebook.scholar.rcac.purdue.edu/.
2. Enter your Purdue Career Account credentials (**without** BoilerKey).
3. Congratulations, you should now be able to create and run Jupyter notebooks on Scholar!

RStudio Server

1. Open a browser and navigate to https://rstudio.scholar.rcac.purdue.edu/.
2. Enter your Purdue Career Account credentials (**without** BoilerKey).
3. Congratulations, you should now be able to create and run R scripts on Scholar!

Resources

Unix

Getting started

Standard utilities

`man`

`tldr`

~ & . & ..

`cat`

`ls`

`cp`

`mv`

`pwd`

`touch`

`wc`

`ssh`

`mosh`

`scp`

`awk`

`sed`

`grep`

`ripgrep`

`find`

`fd`

`top`

`less` & `more`

`sort`

Piping & Redirection

SQL

RDBMS

SQL in R

SQL in Python

R

Getting started

How to install R (windows/mac/linux) How to install RStudio How to connect to RStudio Server on Scholar How to get help (?, help(), get function itself)

Variables

Logical operators

Lists & Vectors

Basic R functions

Data.frames

Reading & Writing data

Control flow

Apply functions

`apply`

`sapply`

`lapply`

`tapply`

Writing functions

Plotting

`ggplot2`

RMarkdown

Tidyverse

Data.table's

SQL in R

Scraping

`shiny`

Python

Getting started

Lists & Tuples

Dicts

Control flow

Writing functions

Reading & Writing data

`numpy`

`scipy`

`pandas`

Jupyter notebooks

Writing scripts

`argparse`

Scraping

Plotting

`matplotlib`

`plotly`

`plotnine`

`pygal`

`seaborn`

`bokeh`

Classes

Tools

Docker

Tableau

GitHub

VPNs

FAQs

How do I connect to Scholar from off-campus?

There are a variety of ways to connect to Scholar from off-campus. You can use the ThinLinc web client, or setup a VPN connection to Purdue's VPN, and connect using the ThinLinc client application. If you just want to use Jupyter notebooks, you can use JupyterHub. If you just want to use RStudio, you can use RStudio Server.

Is there an advantage to using the ThinLinc client rather than the ThinLinc web client?

Yes. Although it is marginally more difficult to connect with, the ThinLinc client allows the user to copy and paste directly from their native operating system. So for example, if you have an RStudio session opened on your MacBook, you can directly copy and paste code onto Scholar using the ThinLinc client. You are unable to do this via the ThinLinc web client.

GitHub Classroom is not working – can't authorize the account.

This is usually a browser issue. GitHub Classroom does not work well with Microsoft Edge or Internet Explorer. Try Firefox or Safari or Chrome.

In Scholar, my font size looks weird or my cursor is offset.

In scholar, navigate to `Tools > Global Options > Appearance`, and select the "Monospace" font, and click the "Apply" button.

How do I make the ThinLinc window bigger without going to the dreaded "full screen" mode?

See here.

I'm unable to type into the terminal in RStudio.

Try opening a new terminal, try clearing the terminal buffer, or interrupting the current terminal. All these options come from a menu that will pop up when you hit the small down arrow next to the words "Terminal 1" (it might be another number depending on how many terminals are open) which is on the left side right above the terminal in RStudio.

I'm unable to connect to RStudio Server.

Try closing it, clearing your cookies, and using the original link: https://rstudio.scholar.racac.purdue.edu/, or for ease of scrolling, try https://desktop.scholar.rcac.purdue.edu, and open up RStudio from within the ThinLinc web client.

Contributors

We are extremely thankful for all of our contributors! Get your name added to the list by making a contribution.