



INSTITUTE FOR  
RESEARCH ON  
INNOVATION & SCIENCE

# IRIS VDE User Guide

**Created by the IRIS Research Support Team**  
**Last updated: October 2023**  
**Contact: [irisdatarequests@umich.edu](mailto:irisdatarequests@umich.edu)**

# Table of Contents

<b>IRIS Research Support Team</b>	<b>5</b>
Our staff and hours	5
IRIS helpdesk ticketing system	5
Service level agreement	5
<b>Accessing the IRIS VDE</b>	<b>7</b>
VDE access using Windows	7
VDE access using Mac	10
<b>Working with the IRIS UMETRICS Data in the VDE</b>	<b>13</b>
VDE desktop icons	13
How to access the IRIS UMETRICS dataset	14
Via SQL server database	14
Acquiring data in CSV format	16
Resources available in the Public folder	17
Saving your work in the VDE	18
Changing your password via Windows Security	19
Logging off from a VDE session	20
Importing files into the VDE	21
SFTP connection steps	21
Exporting files from the VDE	22
Preparing data outputs for disclosure review	23
Researcher Checklist for Disclosure Review Requests	25
<b>Getting Started with Jupyter Notebook</b>	<b>26</b>
How to change the default browser to Firefox	26
How to launch Jupyter Notebook in the VDE	28
Running an ipynb file	29
Renaming a Jupyter notebook	31
Keyboard shortcuts	32
Cells	32
Running code	34
Saving your work	35
Shutting down the Jupyter Notebook server	36
Troubleshooting	37
Restarting the kernel	37
“_xsrf” argument missing from post	38
<b>Instructions for Using Git in the VDE</b>	<b>39</b>
Setting up an account	39
Creating and accessing your Git repository	40

**Publishing Your Research with IRIS UMETRICS Data** 41

- Citations for IRIS UMETRICS data 41
- How to meet data sharing requirements 41
  - Data availability statement 41
  - Data access for replication purposes 41

**VDE Account Renewal Process** 42

## Welcome to the community of IRIS researchers

We at IRIS look forward to working with you as you delve into the IRIS UMETRICS dataset via the virtual data enclave (VDE) to meet your research goals – in some documentation you may also see the VDE referred to as the VDI, or Virtual Data Infrastructure. IRIS Enclave, VDE and VDI are used interchangeably but we will be using VDE throughout this user guide.

The VDE is a virtual machine launched from your own desktop but operating on a remote server, allowing IRIS researchers to access the restricted use data without being physically present at the IRIS offices. In order to help prevent unauthorized disclosure, researchers are not allowed to transfer files directly to or from their personal computer when working within the VDE. Researchers are encouraged to do as much of their work as they are able to within the VDE, and to that end, the VDE is regularly updated with statistical packages, programs, and other tools according to user needs and requests. IRIS recognizes that researchers need to access their analysis outside of the VDE on occasion and this document outlines the indirect transfer methods that are employed to achieve this while minimizing the risk of accidental or deliberate exposure of confidential information.

We ask all VDE users to follow best practices in restricted-data use as outlined in the [IRIS VDE Acceptable Use Policy](#). Thank you for helping us maintain the security and confidentiality of our data by following these rules. If you have any questions regarding practices or procedures for working with IRIS data in the VDE, please do not hesitate to contact us at [irisdatarequests@umich.edu](mailto:irisdatarequests@umich.edu).

Sincerely,

The IRIS Research Support Team

# IRIS Research Support Team

## Our staff and hours

Elissa Irhamy, Data Science Specialist, 2338 Perry  
[elairhamy@umich.edu](mailto:elairhamy@umich.edu)

Natsuko Nicholls, Research Manager, 2333 Perry  
[hayashin@umich.edu](mailto:hayashin@umich.edu)

**We are located in the IRIS offices at:** 330 Packard Street, Ann Arbor, Michigan, 48104

**Our business hours are:** Monday through Friday from 8:00 a.m. to 5:00 p.m. (EST). Exceptions to these hours include closures for holidays (or, on the rare occasion, inclement weather or other unplanned events), which follow those of the University of Michigan.

## IRIS helpdesk ticketing system

**For any VDE or research related questions or requests, please use the e-mail**

[irisdatarequests@umich.edu](mailto:irisdatarequests@umich.edu) **to generate a helpdesk ticket.** As a ticket is reviewed and addressed by the IRIS Research Support Team, the status will be changed and you will receive email notifications accordingly.

### *Status notifications:*

A “**Request**” status notification is an automated email generated by the help ticketing system to alert the user that a request was successfully submitted, but does not indicate that a member of the IRIS Research Support Team has seen the request. The “**In Process**” notification includes a response from the IRIS team, and a notification of “**Completed**” status is assigned once a ticket has been addressed and resolved.

## Service level agreement

In response to your research data requests, we at IRIS always seek to provide prompt and helpful service that includes technical assistance, data access and provision support, disclosure reviews, user provided data import, regular VDE updates, etc. Help tickets will be reviewed within 1-3 business days and the

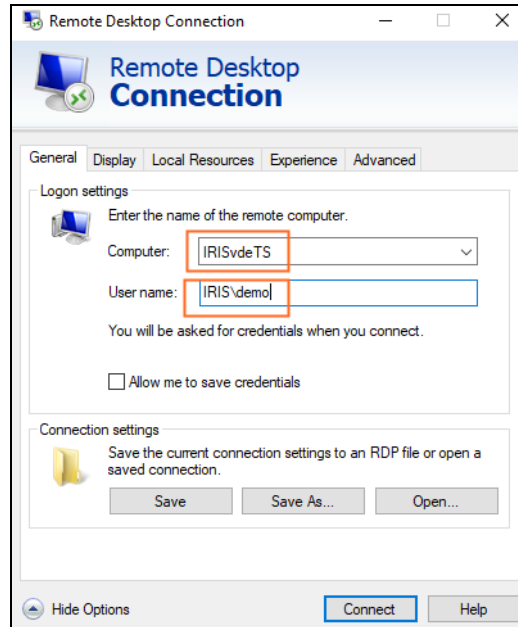
assigned IRIS team member will respond within 4-5 business days from the date of the original user request.

**A note on urgent requests:** As a general rule, help desk tickets are serviced during IRIS business hours, Monday through Friday from 8:00 a.m. to 5:00 p.m. Due to the often unpredictable nature of research, we understand that there may be occasions when you have an urgent request. **Please be sure to specify the time frame in which you need a request addressed if it is time sensitive, but be aware that IRIS does not guarantee that a researcher's deadline will be met in such cases.** The service level agreement stated above applies to all cases whether a deadline is stated by the researcher or not, so please plan accordingly.

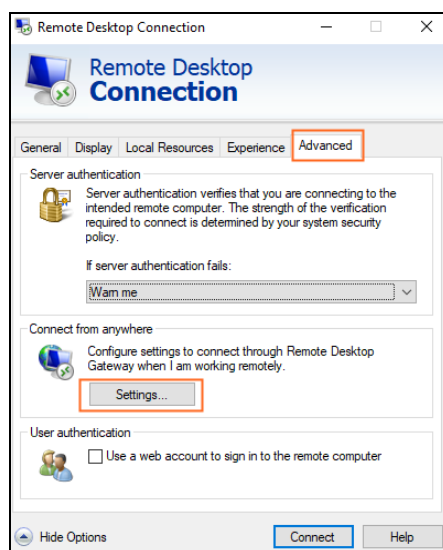
# Accessing the IRIS VDE

## VDE access using Windows

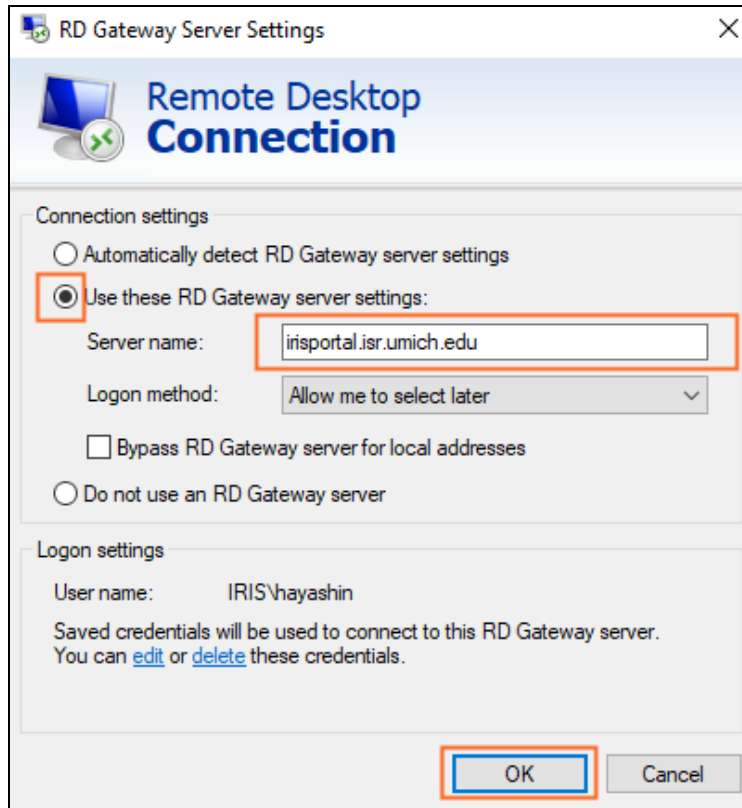
1. Open the Remote Desktop Connection application.
2. For your Computer, enter: IRISvdeTS. Enter your username and password provided by IRIS. Make sure to add "IRIS\" when typing your user name. Your username and password remain the same.



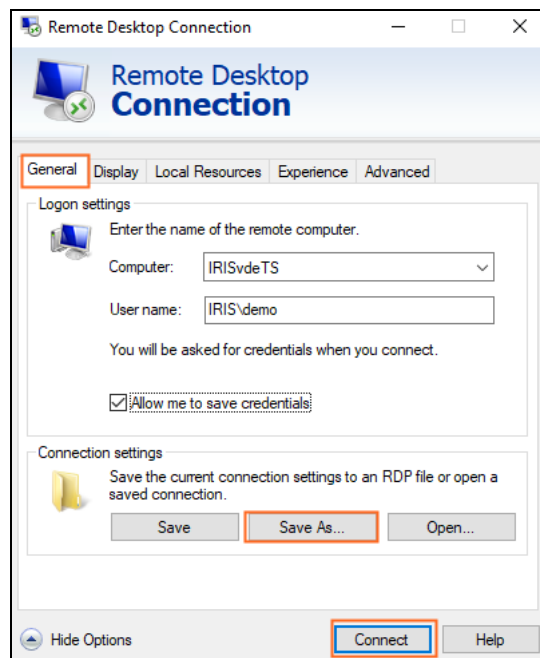
3. Click on the "Advanced" tab and select "Settings..." in the "Connect from anywhere" section.



4. Select “Use these RD Gateway server settings:” and input the server name: irisportal.isr.umich.edu. Leave the Logon method as “Allow me to select later”. Click “OK” to go back to “General tab”

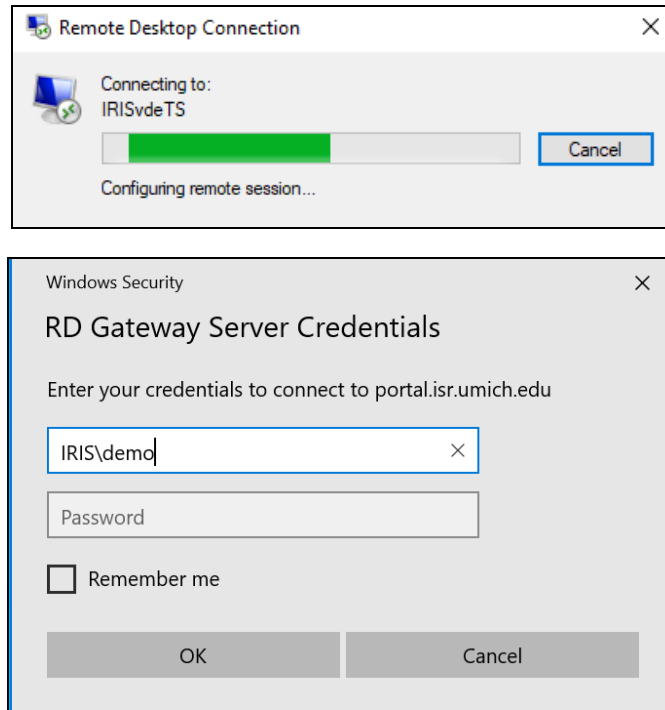


5. Click “Connect” – you may want to save the connection on your local machine.

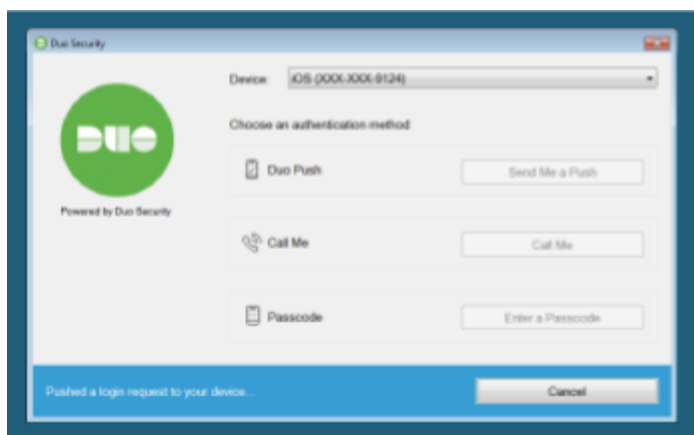




- Once you are asked to enter your credentials, use your username and password that you just used in Step 2. Again, make sure to add "IRIS" when typing your user name.



- Once your username and password are accepted, you will see your virtual session starting up and the DUO two-factor authentication prompt (have your mobile device on hand during this process). Shortly you will be presented with screens that look like these:



Computer screen



Smartphone screen

At the top of this window there is a drop-down box labeled “Device.” If you have chosen to have more than one device associated with your IRIS Duo login (i.e., your smartphone and a desk phone), you can switch the active device here. This window also gives you multiple options for authenticating. We recommend **DUO Push** (default) as it is the easiest, and most reliable. If you need to receive a phone call, you can click **Call Me** and DUO will call you and ask you to push a button on your phone’s dial pad to authenticate. If you do not have cellular or WiFi connectivity, you can click the **Passcode** button. This will bring up a text box for you to enter a code. You can get this code by opening the DUO app on your phone, and clicking the key next to “University of Michigan Institute for Social Research.” This code will generate whether you have any network connectivity or not.

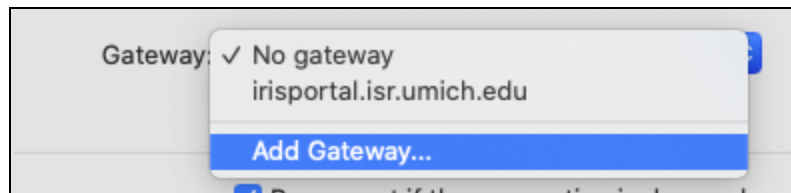
8. Click on “OK” to Warning to VDE Users. After several seconds, you will be directed to one of our IRIS Remote Workstations. This may take longer if you are connecting to VDE for the first time.

## VDE access using Mac

**Prior to following the steps below, make sure you have Microsoft Remote Desktop version 10.7.10 installed.** Microsoft Remote Desktop does not come pre-installed on a Mac, though most users will have it as part of their Office installation. Otherwise you can download it through the App Store.

1. Add the PC Name IRISvdeTS and under “Gateway” select “Add Gateway...”

The screenshot shows the "Add PC" dialog box in Microsoft Remote Desktop. The "PC name" field is filled with "IRISvdeTS". The "User account" dropdown menu is set to "Ask when required". The "General" tab is selected, showing "Friendly name" as "Optional", "Group" as "Saved PCs", and "Gateway" as "No gateway". There are checkboxes for "Bypass for local addresses", "Reconnect if the connection is dropped", "Connect to an admin session", and "Swap mouse buttons". "Cancel" and "Add" buttons are at the bottom.



2. In the “Add a Gateway” screen, enter the Gateway name as irisportal.isr.umich.edu. Select “Add” and “Add” again on the “Add PC” screen.

A screenshot of the "Add a Gateway" dialog box. It has a title bar "Add a Gateway". Below the title bar are three input fields: "Gateway name:" with the value "irisportal.isr.umich.edu", "Friendly name:" with the placeholder "Friendly name", and "User account:" with a dropdown menu showing "Use PC User Account". At the bottom right are two buttons: "Cancel" and "Add".

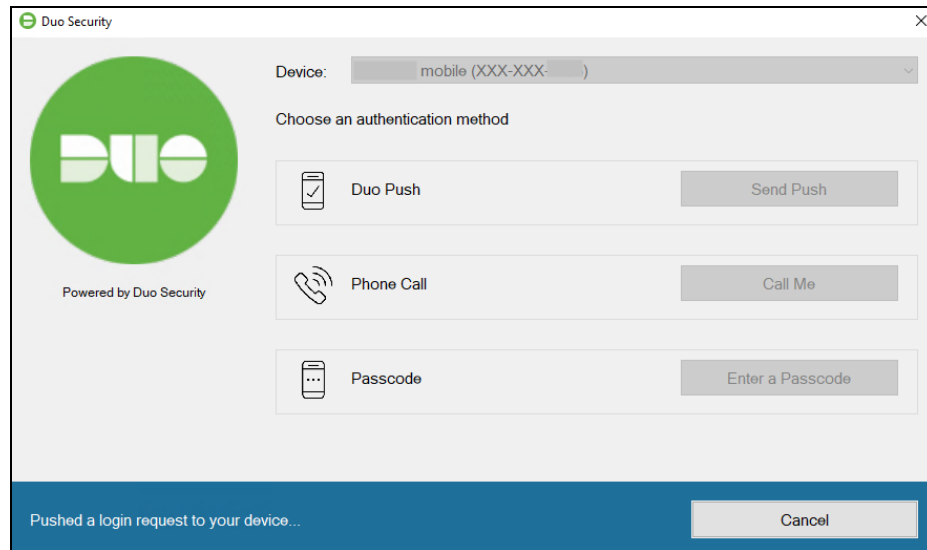
3. You will be asked to Enter Your User Account. Enter your username and password. Make sure to put “IRIS\” in front of your username.

A screenshot of the "Enter Your User Account" dialog box. It has a title bar "Enter Your User Account". Below the title bar is a message: "This user account will be used to connect to IRISvdeTS (remote PC) and irisportal.isr.umich.edu (gateway).". Below the message are two input fields: "Username:" with the value "IRIS\demo" and "Password:". Below the password field is a checkbox labeled "Show password" which is unchecked. At the bottom right are two buttons: "Cancel" and "Continue".

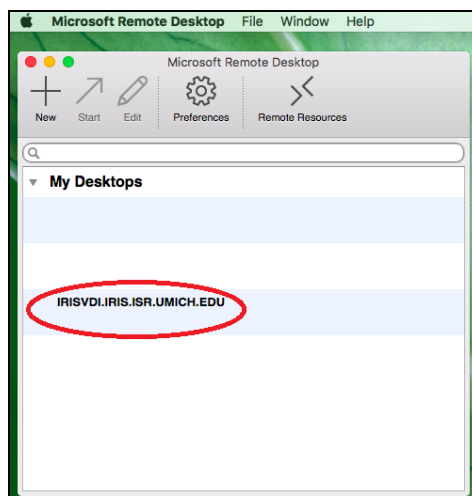
4. Click Continue and do so again on the next screen.

A screenshot of a certificate warning dialog box. It has a title bar. On the left is a yellow padlock icon with a red 'X' over it. To the right of the icon is the text: "You are connecting to the RDP host 'IRISvdeTS'. The certificate couldn't be verified back to a root certificate. Your connection may not be secure. Do you want to continue?". At the bottom are three buttons: "Show Certificate", "Cancel", and "Continue".

- Once your username and password are accepted, you will see your virtual workstation starting up and the DUO two-factor authentication prompt (have your mobile device on hand during this process).



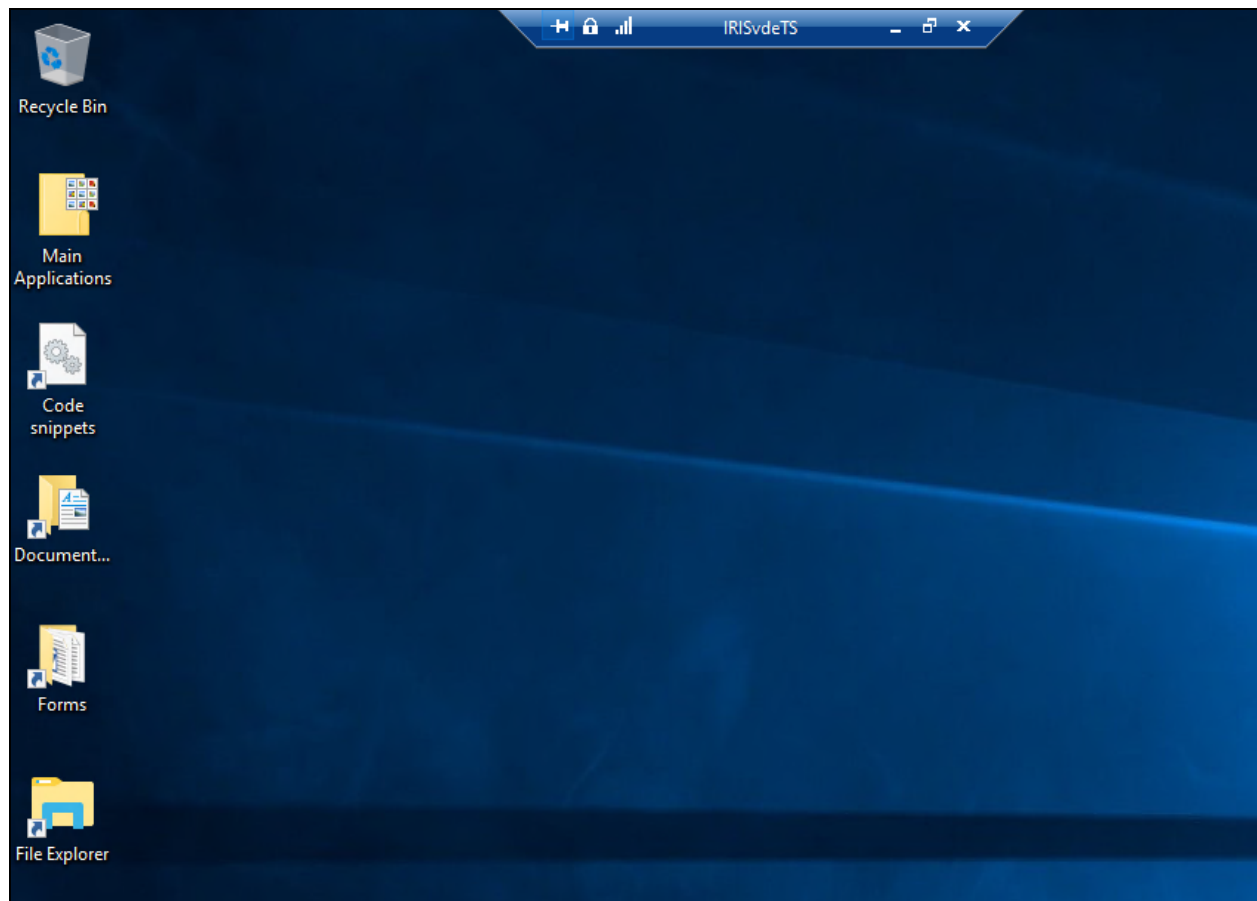
Click on “OK” to the “Warning to VDE Users.” After several seconds, you will be directed to your virtual VDE session. This may take longer if you are connecting to VDE for the first time.



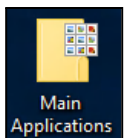
Once this initial connection is made, you will be able to access the IRIS VDE via your Microsoft Remote Desktop application going forward. **PLEASE NOTE:** Some Mac users have experienced that a saved RDP link stops working. If you experience difficulty accessing VDE from a Mac machine, try to refresh the RDP link by repeating the steps above.

# Working with the IRIS UMETRICS Data in the VDE

## VDE desktop icons



*VDE Remote Desktop*



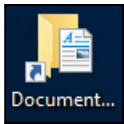
### **Main Applications**

- This is located on the desktop for quick access to VDE applications, e.g., Stata, Jupyter Notebook, Microsoft products, etc.



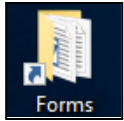
### **Code Snippets**

- Example and starter code resides here, including examples for setting up database connections.



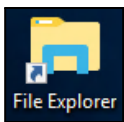
## Documentation

- Documentation is available for every major and supplemental data release to date. This includes both release summary documentation and data dictionaries.
- You'll also find the user guide stored here, if needed.



## Forms

- Use these forms for placing requests such as disclosure review, Census RDC file transfer, or exporting replication data / code required by journals for article submission
- The disclosure review request memo template is also downloadable from our website



## Workspace

- Use the file explorer shortcut and launch notebooks, code or documents from your home or directories in network location, U: drive or N: drive.

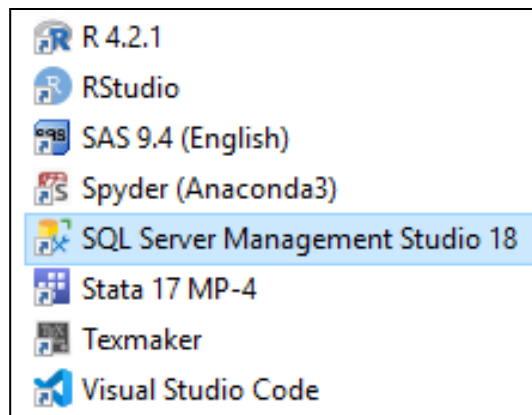
# How to access the IRIS UMETRICS dataset

## Via SQL server database

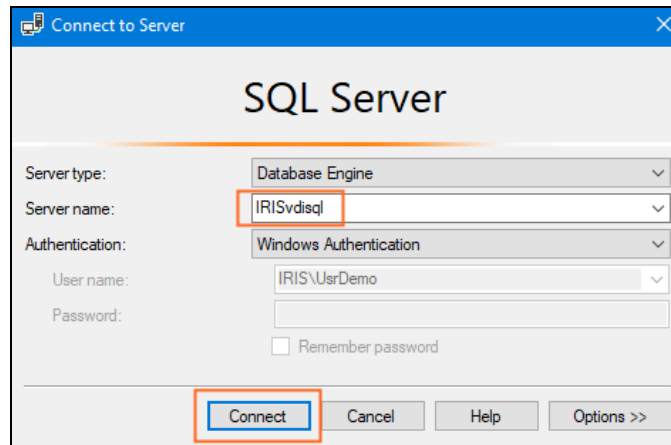
SQL Server Management Studio is installed in the VDE and allows you to connect to the IRIS SQL research database and access data files (tables/views), run queries, and export data.

Below is a step by step guide on how to connect to the database using SQL Server Management Studio.

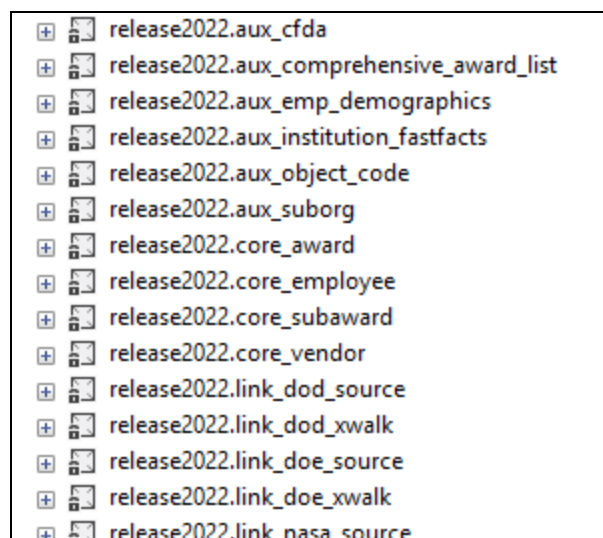
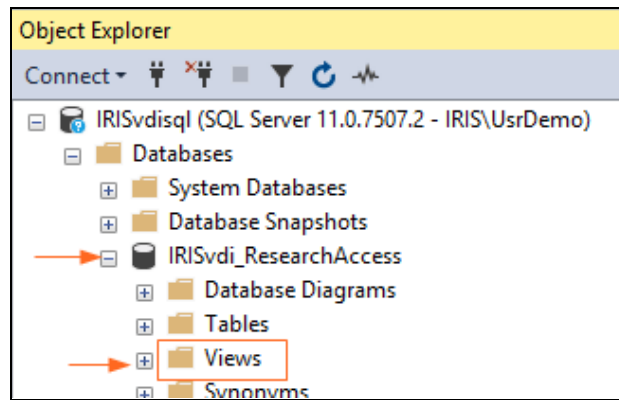
1. Click on the "Main Applications" folder on your desktop then double click on SQL Server Management Studio 18 to start the program.



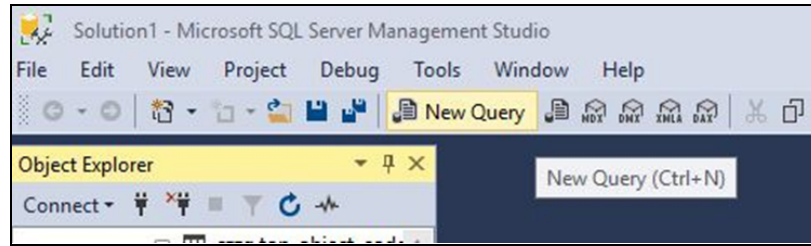
- Configure the connection to the database. The server type is “Database Engine”, the server name is “IRISvdisql” and the authentication is “Windows Authentication”. Select “Connect”.



- To explore existing IRIS UMETRICS data, you can use the Object Explorer pane on the left hand side of the screen. Click the database name “IRISvdi\_ResearchAccess” to access tables/views. Release data files are stored in “Views” and organized by schema, e.g., release2019, release2020, and release2022.

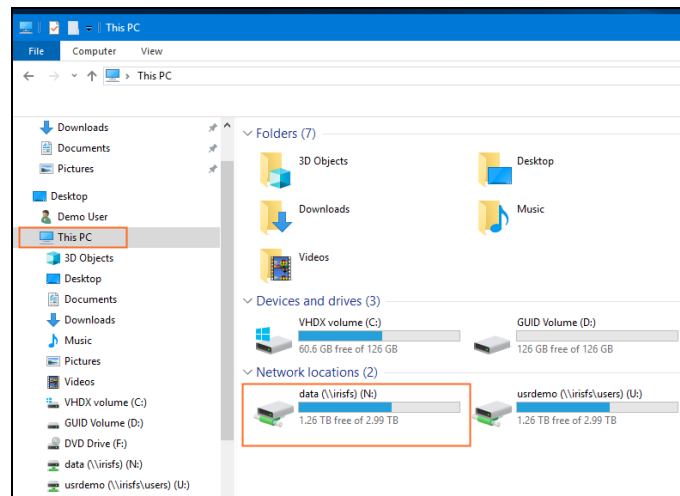


4. For queries and other SQL tasks, click “New Query” from the taskbar, or use a shortcut Ctrl + N to open a new query window.

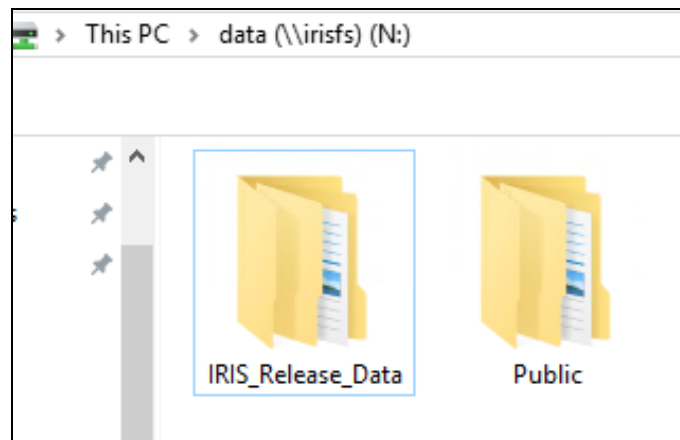


## Acquiring data in CSV format

1. Click on the “File Explorer” on your desktop and then click on “This PC” to access the N drive.



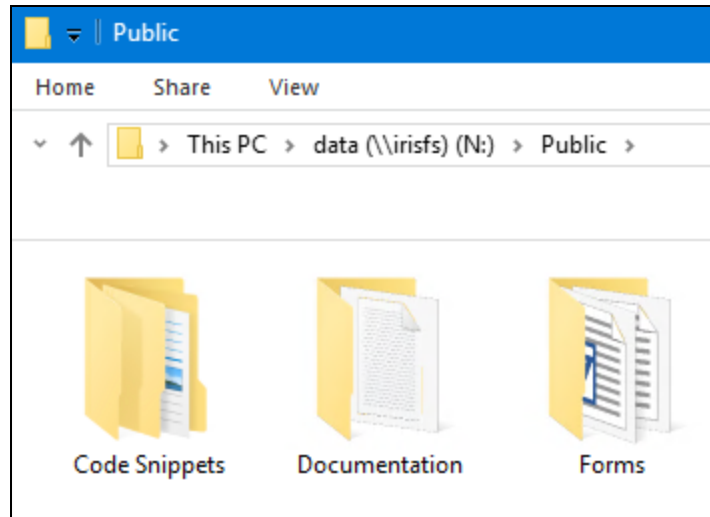
2. Click on the folder named “IRIS Release Data” to access csv files. Files are organized by release year.





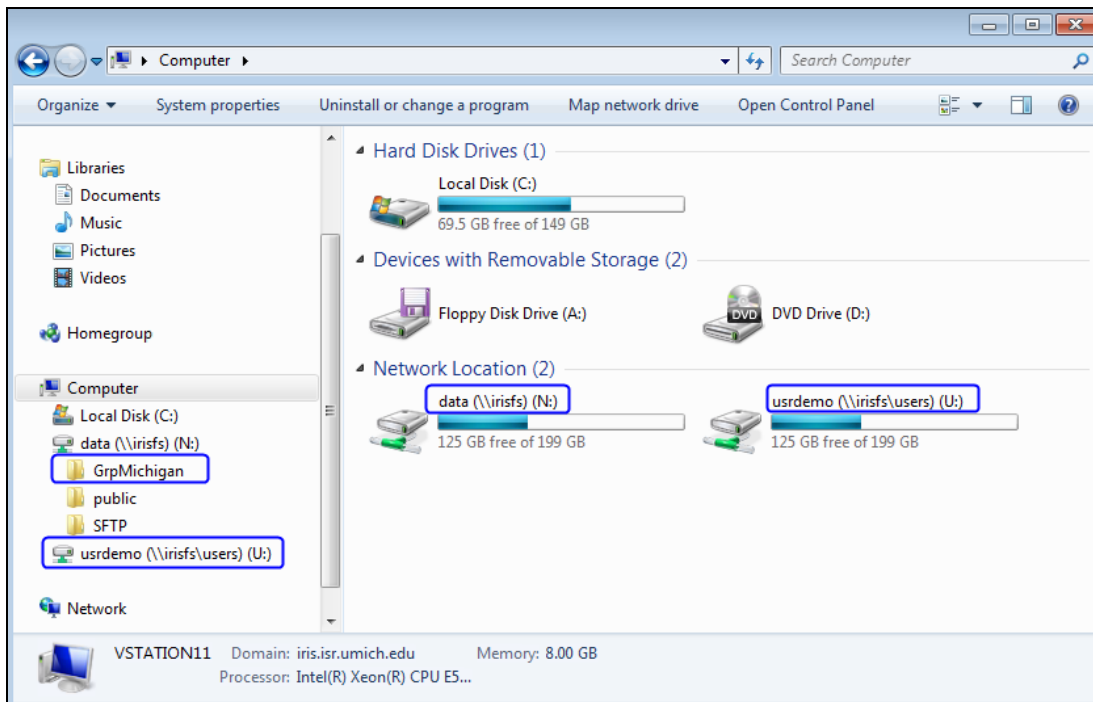
## Resources available in the Public folder

IRIS releases the de-identified data for research purposes annually (biennial since 2020) along with detailed documentation, including data dictionaries. Documentation from the current and previous releases is stored in the *Documentation* folder within the Public folder in the N drive in the VDE. In the same folder, you will also find other useful resources in Code Snippets, such as code for establishing database connections or performing data linkage, and Forms, which has the Disclosure Review Request Memo template, among others. For ease of use, links to these folders are available as desktop shortcuts.



## Saving your work in the VDE

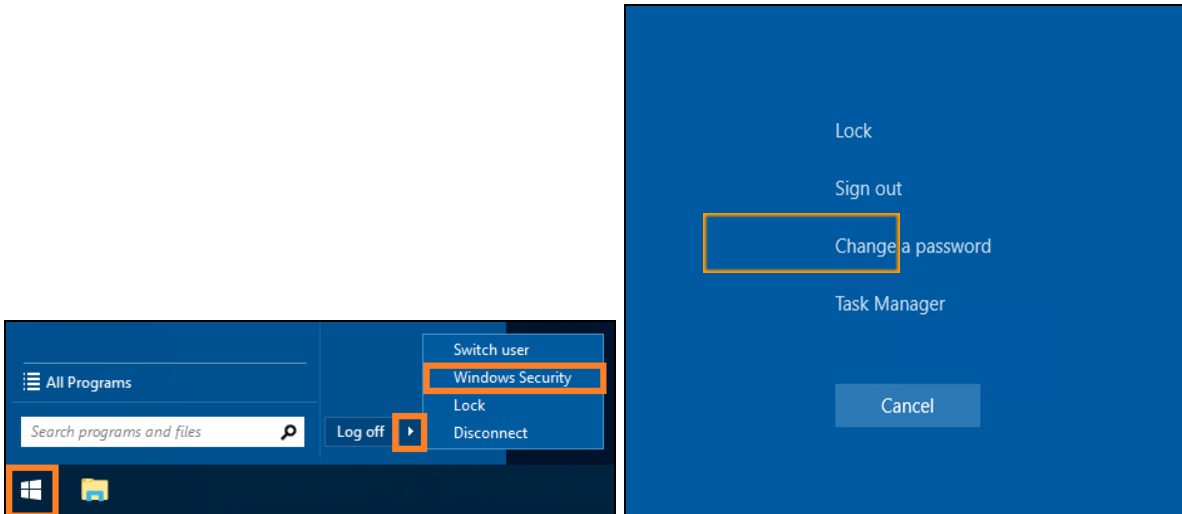
When working with the IRIS data within the VDE, you need to make sure to save your work within the IRIS file server, either within your home folder (U:\ ) or a shared group folder (N:\ ); **otherwise, all work will be lost after a current VDE connection session is terminated.** To locate these folders open Windows Explorer and view the items located under “Network Location” (example shown below).



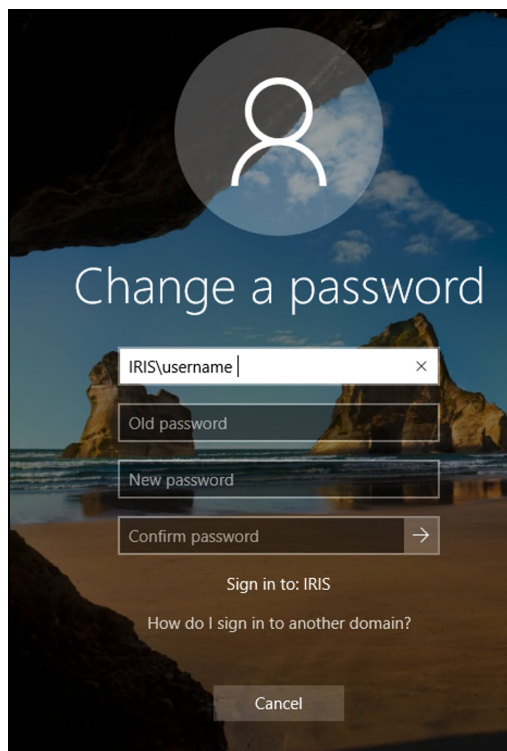
All users are set up with a personal home folder (on the U: drive) upon account creation that will have the same naming as the user login. In addition, research teams will generally have a shared group folder located on the N: drive with a name that begins with “Grp” (e.g., “GrpMichigan”). If you need a shared folder set up, please contact [irisdatarequests@umich.edu](mailto:irisdatarequests@umich.edu).

## Changing your password via Windows Security

1. Left click the Windows Start button on the bottom left of the screen.
2. Left click on the right facing arrow next to the log off button.
3. Left click on Windows Security.
4. Select Change a password.



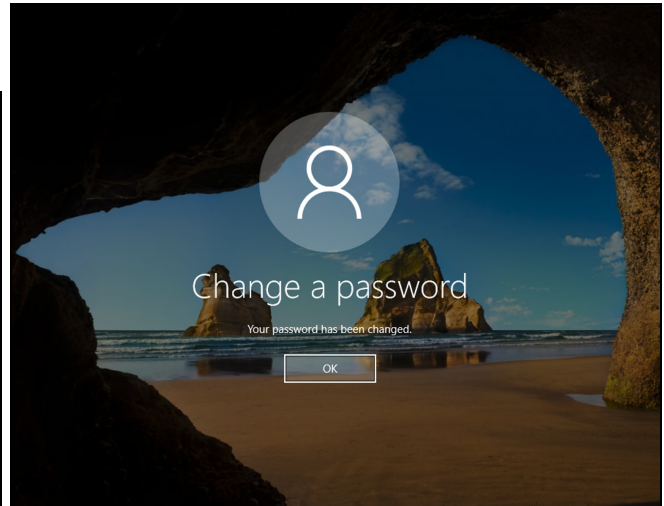
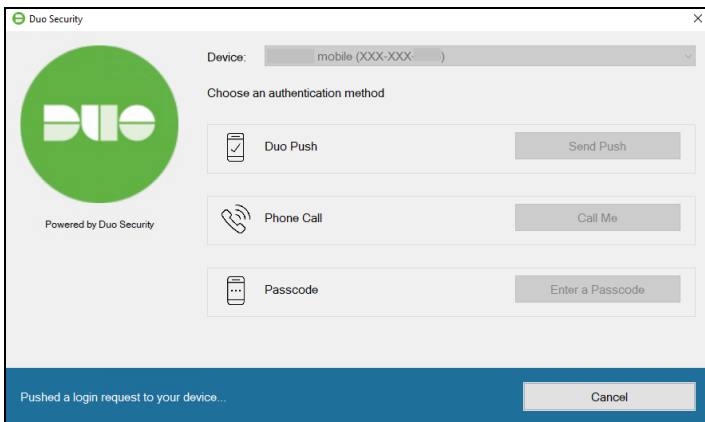
5. Enter your username and old password accompanied by your new password.



**\*\* Note that your new password must be a minimum of 15 characters including all the following:**

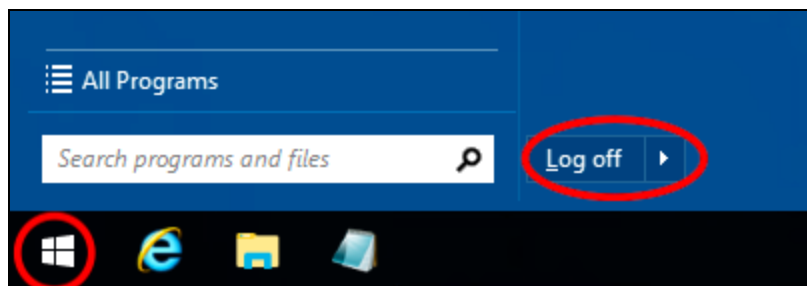
- A mixture of both uppercase and lowercase letters
- At least one number
- At least one special character, e.g., ! @ # ? ]

6. Upon changing your password, you will receive a DUO push to your mobile device confirming the password change. Once you accept the push, your password will be changed.



## Logging off from a VDE session

Once you complete your work in VDE, make sure to log off from your VDE session. If you just close the window, the session remains open until it is automatically terminated after 7 days of idling time. Instead, navigate to the Windows Start icon in the lower left hand corner of your screen and select "Log off" to end the session. This frees up system resources for other users.



## Importing files into the VDE

Any data which could potentially re-identify IRIS UMETRICS data will be prohibited from importation. This qualifier is subject to review by IRIS staff. If a researcher needs to bring data files into the VDE, the files should be uploaded via FTP (see instructions below) to the designated temporary storage area (“dropbox”) along with the import request memo (found on the IRIS [website](#)) for review by IRIS. Once the files and completed memo are uploaded, the researcher should email [irisdatarequests@umich.edu](mailto:irisdatarequests@umich.edu), to request a review of these materials. IRIS will collect your files from the “incoming” Dropbox folder to evaluate them. If approved, IRIS will move the files to the location you specify in the memo (e.g., personal or shared folder) and notify you when this transfer is complete.

## SFTP connection steps

To import files to the VDE, you can connect and upload files to the IRIS SFTP server using tools for Windows or macOS. For example, you could use FileZilla for Windows or macOS, or a system specific client such as WinSCP for Windows, as described below. Imported files should be placed in the “incoming” Dropbox folder. You may follow similar connection steps for exporting files (after disclosure review) from the “outgoing” Dropbox folder in the VDE.

1. On WinSCP, you will need to login to access the IRIS FTP server (host name: [irisdropbox.isr.umich.edu](http://irisdropbox.isr.umich.edu)). Use the same credentials given for VDE access.
2. Upon entering your IRIS login credentials, you will automatically get a DUO push on your phone. Check your DUO mobile app on your phone to respond to "Are you logging in to IRIS Dropbox SFTP?" You can tap Approve.
3. If a warning window with the message “Continue connecting to an unknown server and add its host key to a cache” pops up, click Yes. You may see this message only once in your first connection attempt.

The image shows a 'Session' dialog box with the following fields and controls:

- File protocol:** A dropdown menu set to 'SFTP'.
- Host name:** A text input field containing 'irisdropbox.isr.umich.edu'.
- Port number:** A spinner box set to '22'.
- User name:** A text input field containing 'IRIS\demo'.
- Password:** A text input field with masked characters (dots).
- Buttons:** 'Save', 'Cancel', 'Advanced...', 'Login' (with a green plus icon), 'Close', and 'Help'.

4. Once you are logged in, you can see the "incoming" and "outgoing" folders that you can access for file import and export. Select the file(s) that you wish to transfer to the dropbox (from the left panel) and drag and drop the file(s) into the destination folder (incoming) in the right panel.

## Exporting files from the VDE

Research results generated through the analysis of IRIS data can **only** be exported from the VDE after a disclosure review by IRIS staff for compliance with the [IRIS Restricted Data Use Agreement](#) and [IRIS VDE Acceptable Use Policy](#).

### What Researchers may **usually** export

- Statistical summary information including frequency tabulations, magnitude tabulations, means, variances, correlation coefficients, and regression coefficients.
  - allowed only if it does not permit the identification of any individuals and organizations
  - aggregated observations should include at least 3 IRIS member institutions

## What Researchers may **never** export

- Raw data or microdata files or analysis output containing individually identifying information
- Program/code that includes information on database connection (e.g., database name, connection credentials) or embeds identifiable information

## Preparing data outputs for disclosure review

If a researcher needs to export analysis results from the VDE, the researcher should first review the information on what may/may not be exported and common examples of issues that will cause a disclosure review request to be declined. Disclosure proofing is intended to prevent the identification of a particular individual or organization. For example, in the employee data, an employee earnings measure can only be released if the number of cases is high enough that no single person can be identified. Disclosure proofing also involves comparison of successive output files, so that a small change in the number of cases from one file to the next does not result in the identification of any one individual. Applying these tests can sometimes mean that certain data points must be excluded from your outputs. The disclosure review process influences potential outputs in several ways:

### **1. Individual Data**

Information on individual employees can only be released if it is based on a sufficient number of cases to ensure that no single person can be identified in the data.

### **2. Vendor and Subaward Data**

Information on business establishments can only be released when it is based on a large enough number of establishments and when the concentration rate is low. A concentration rate would be high if a small number of vendors accounted for a large percentage of spending in a particular category. Some by-county or by-NAICS code measures cannot be released because of these issues.

### **3. Secondary Disclosure and Implicit Samples**

Effective disclosure proofing requires attention to the possibility that data in a single output, or data disclosed in more than one output can be used to infer information about individuals or organizations. For instance, if we were to approve the release of a file with data that differed from the data in a prior disclosure release by only one or two data points, comparing files would allow you to identify information about those data points, compromising privacy and confidentiality.

Likewise, the size and (for businesses) concentration of unreported categories must be taken into account. For example, if we break information about individuals into three categories and release data on two of them along with information about the total number of individuals on which the data are based, the size of the omitted category can be easily inferred.

Concerns about secondary disclosure and implicit samples can sometimes lead to suppression of information for categories that would pass disclosure when reported alone. Suppose, for instance, we wish to report information on an industry, 1, and on its four sub industries 1.a, 1.b, 1.c, and 1.d. If 1.d is too small (or too concentrated) to pass disclosure, its suppression from the report will also require suppression of 1.a, 1.b, or 1.c because those three data points along with total information about 1 would allow information about the suppressed category (1.d) to be inferred. This is why effective privacy protection requires that IRIS also scrutinize contextual information such as the number of cases associated with a particular report or figure and the match rates for individuals and organizations.

### Common reasons for disclosure review denial

**Problem:** File includes organization names

**Solution:** Replace names with preferred anonymized designations (e.g., University A or Vendor A) or with the numeric organization id (e.g., vendor ID, subawardee ID).

**Problem:** Disclosure review lacks description

**Solution:** Send a description of the file(s) for which you are requesting disclosure review. Be sure to follow the Disclosure Review Request Memo template to provide accurate and complete details.

**Problem:** Frequency count in cells/figures is too small

**Solution:** Suppress cells/numbers in figures for which the frequency counts are smaller than 10.

**Problem:** Risk of secondary disclosure

**Solution:** Do not make disclosure requests for files that are very similar or contain overlapping information such that the risk of secondary disclosure is high.



## Researcher Checklist for Disclosure Review Requests

### **Step 1: Prepare files for the disclosure review request.**

Copy the file(s) into a single folder named “Disclosure\_Review\_MMDDYY,” replacing the “MMDDYY” with the date of your request.

### **Step 2: Write a Disclosure Review Request memo (found on IRIS [website](#)).**

This document is essential as it provides information necessary to the reviewer and establishes a reference for your own research group to use when creating future disclosure requests. Please save it in the same folder as the files for which you are requesting review. A memo must always be included with a disclosure review request or it will be denied regardless of the content of the files requested. IRIS will only review complete requests.

### **Step 3: Email the IRIS Research Support Team via [irisdatarequests@umich.edu](mailto:irisdatarequests@umich.edu).**

Please note the location of the folder to be reviewed.

### **Step 4: Download approved file(s).**

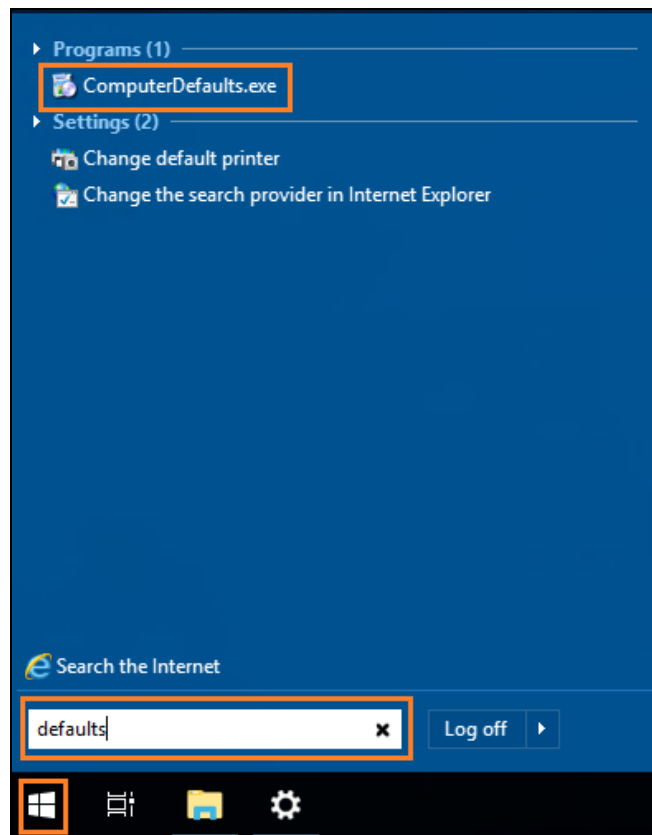
If approved, the requested output will be copied to your “outgoing” Dropbox folder for your retrieval. (See [instructions for connecting via FTP](#)) **PLEASE NOTE:** FTP folders should not be used as a permanent storage area. Files left over 7 days are subject to deletion.

# Getting Started with Jupyter Notebook

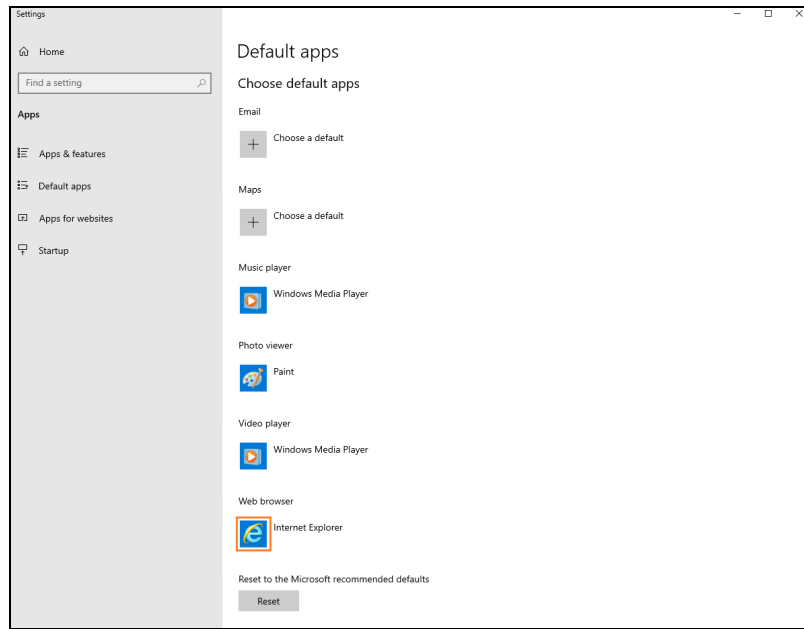
## How to change the default browser to Firefox

We recommend changing the default browser from Internet Explorer to Firefox to facilitate proper Jupyter Notebook functionality. To do this, begin by performing the following:

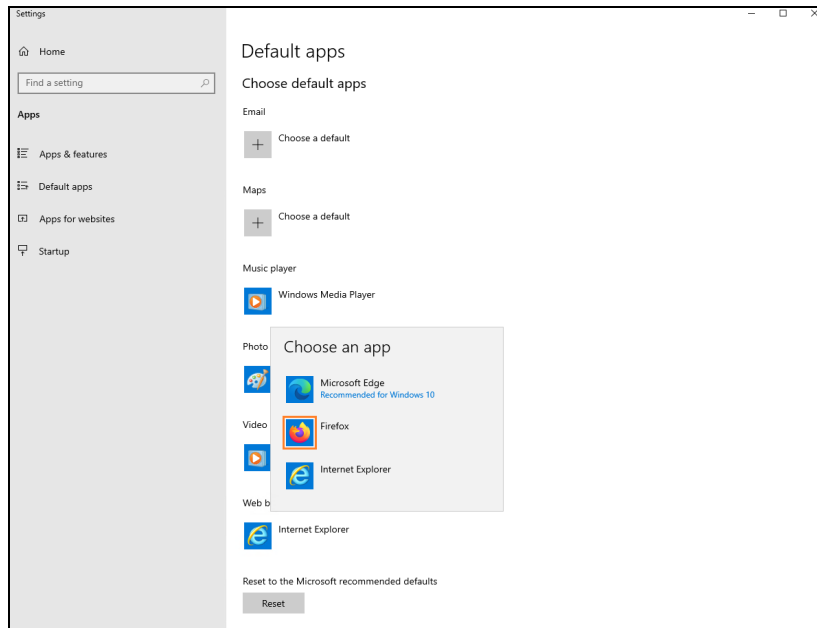
1. Left click the Windows Start button on the bottom left of the screen.
2. Type “defaults” into the “Search programs and files” text box.
3. Left click on ComputerDefaults.exe



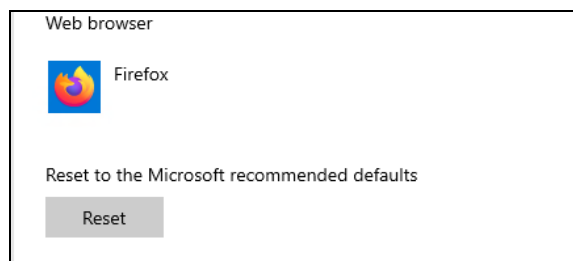
4. Under Web browser, left click on the Internet Explorer icon.



5. In the pane that pops up, left click on the Firefox icon.



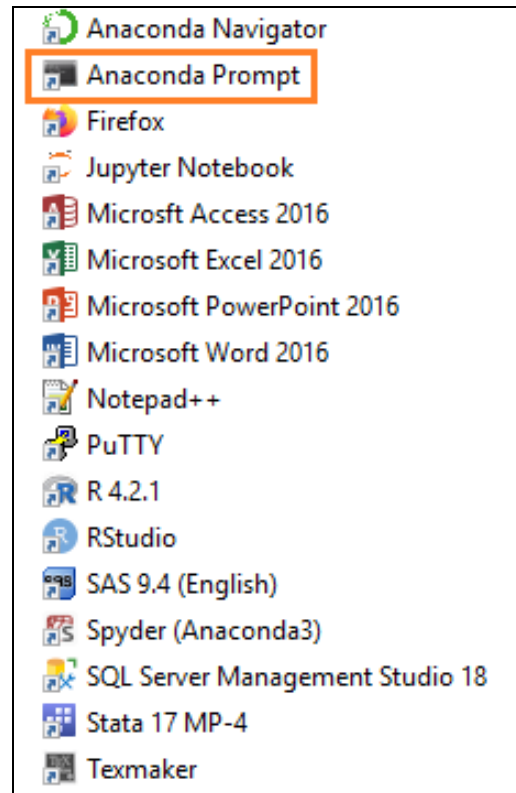
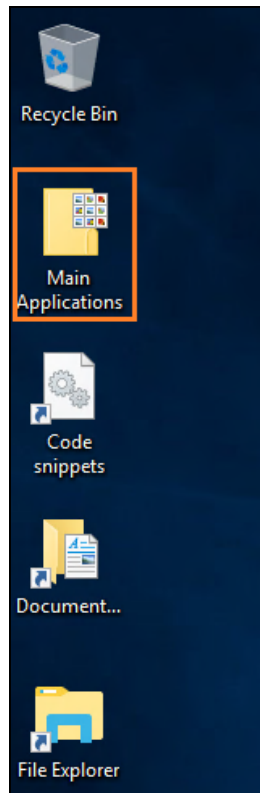
Having followed the previous steps, Firefox should now be the default Web browser, which can be verified if your settings match this screenshot:



You may instead use Microsoft Edge, if desired. Jupyter Notebook works better with either Firefox or Edge than Internet Explorer. This change will persist across sessions, thus this process only needs to be completed once.

## How to launch Jupyter Notebook in the VDE

1. Double left click the Main Applications folder located on the CDE desktop.
2. In the Main Applications folder, double click on the Anaconda Prompt shortcut.



3. Identify the path, including the drive letter, where work will be conducted.
4. For example, to work in the U: drive, follow the commands as shown in the screenshot below. These commands are:

`U:` (to change to U: drive)

`cd demo` (to change directory to a hypothetical folder on U: called "demo")

`jupyter notebook` (to launch Jupyter Notebook)

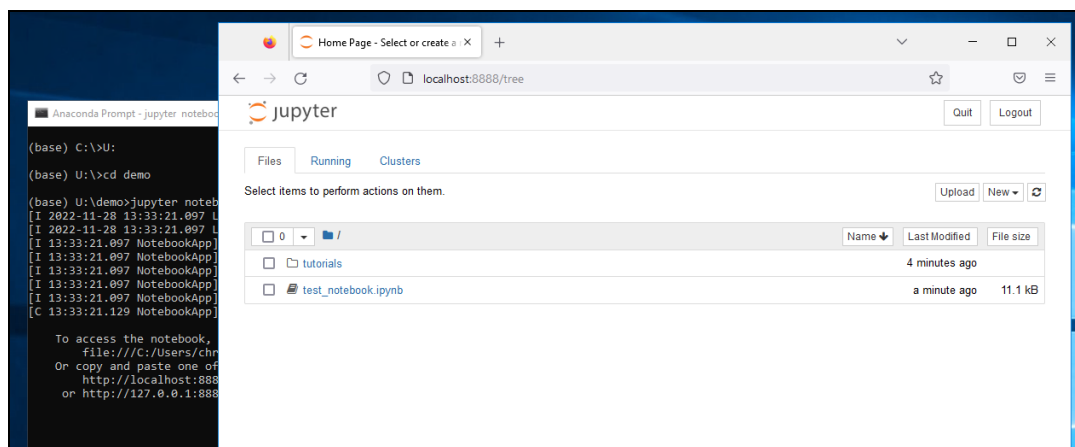
```
Anaconda Prompt
(base) C:\>U:
(base) U:\>cd demo
(base) U:\demo>jupyter notebook_
```

## Running an ipynb file

Once Jupyter Notebook launches in the chosen directory, you will see the Anaconda command prompt and a Firefox window will appear.

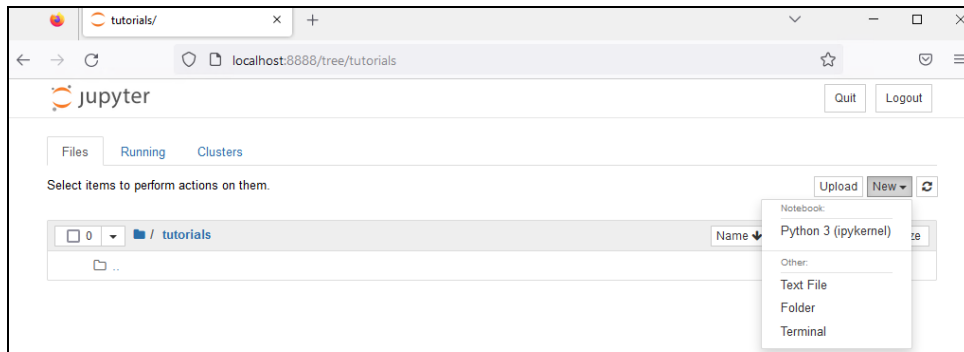
*Note: do not close the command prompt window (“X”) or it will shut down the Jupyter server and you’ll need to restart this sequence. Minimizing the prompt will not cause any issues.*

Within Jupyter, click a folder to navigate into it, or click an .ipynb file (notebook) to open it.

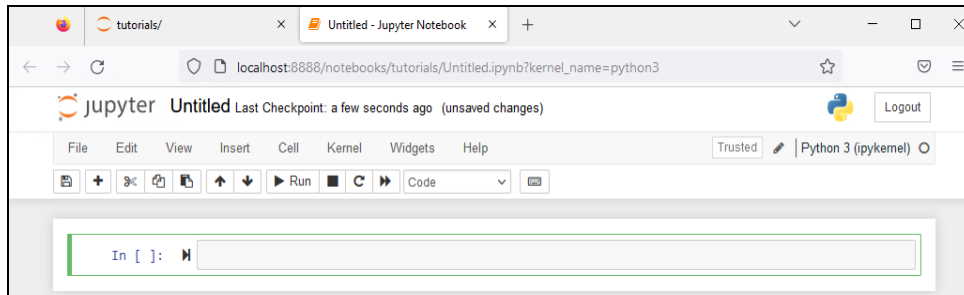


*Note: the files and folders shown in these examples don’t exist by default. There may be nothing to see until you create a notebook or directory via the “New” button.*

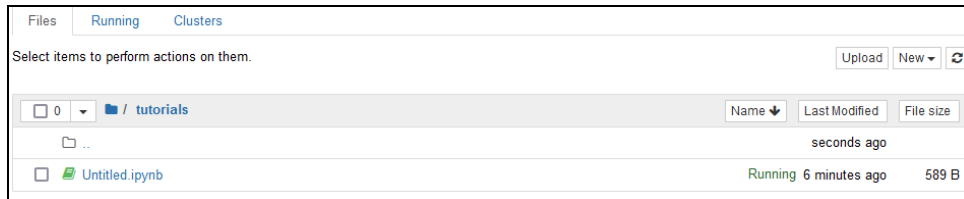
To create a new Python 3 notebook, use the dropdown menu displayed by left clicking the “New” button in the upper right.



This will automatically launch an untitled notebook in a separate tab, as shown below.

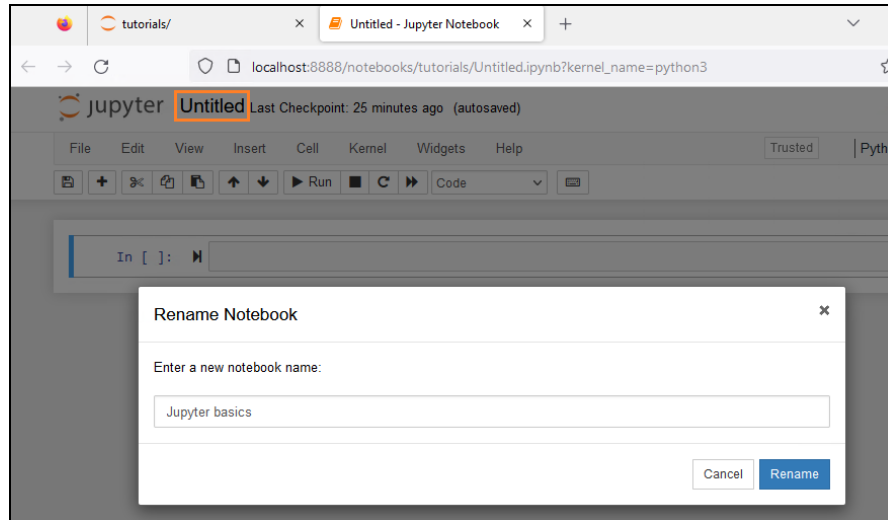


Flipping back to the first tab, the Untitled notebook will be running and the icon light green in color.

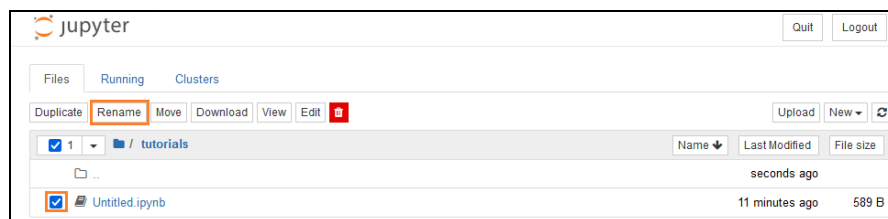


## Renaming a Jupyter notebook

To rename a notebook, left click on the title in the upper left (“Untitled”, by default). This causes a renaming window to appear.

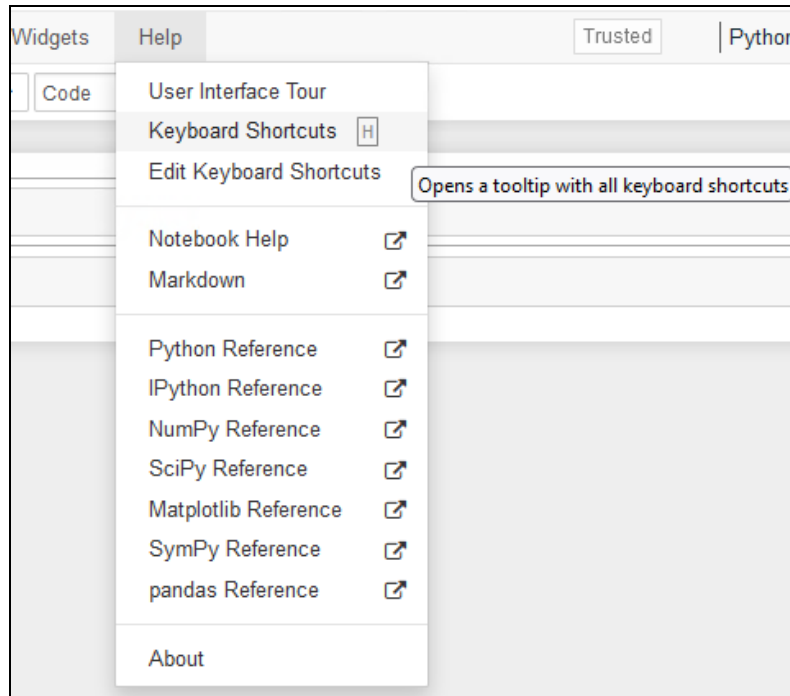


Renaming can also be done at the initial Jupyter landing page, if the notebook is not running (signified by a grey notebook icon). First, left click the checkbox next to the notebook, then left click the Rename button, and finally rename as shown above.



## Keyboard shortcuts

There are numerous options available to make common tasks easier. Press H to bring up the keyboard shortcuts tooltip or navigate to it via Help → Keyboard Shortcuts.

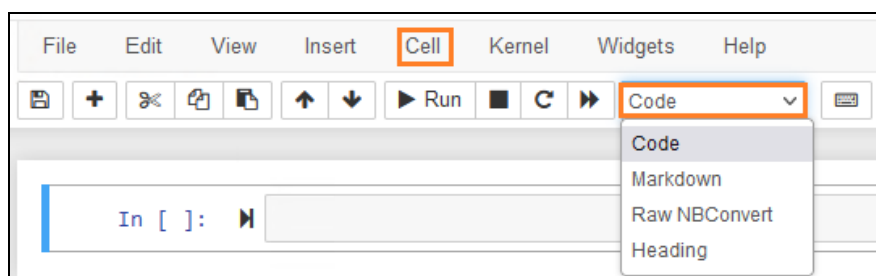


Note: the reference links shown will not work in the VDE due to lack of internet access.

## Cells

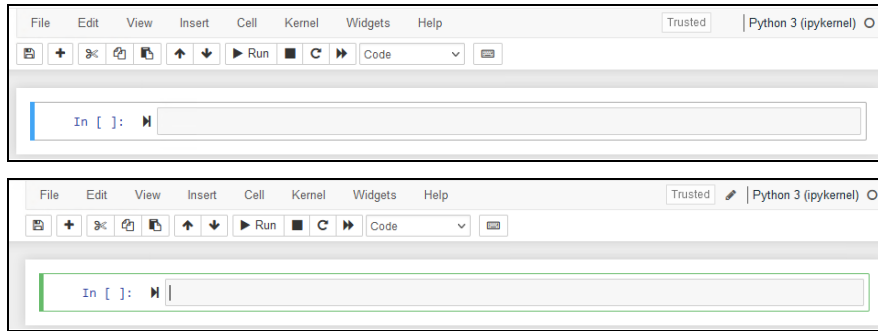
Jupyter notebooks are typically made up of primarily Code and Markdown cells. A default notebook starts with a single code cell as shown. A Code cell contains Python code and comments, whereas a Markdown cell contains text, images, and/or formatting to support documentation, presentation, etc.

Cells can be modified between these types from the Cell menu, with keyboard shortcuts, or via the dropdown menu as shown.





Jupyter cells also change appearance based on the current mode. A blue cell outline signifies command mode, whereas a green outline represents edit mode.



In command mode, keyboard shortcuts can be readily applied to add and alter cells, among other options.

In edit mode, typing will input text or code directly into the cell. Clicking on a cell will enter edit mode, and display a cursor. Hitting escape will return to command mode.

If there is no highlighting, that cell is not selected.

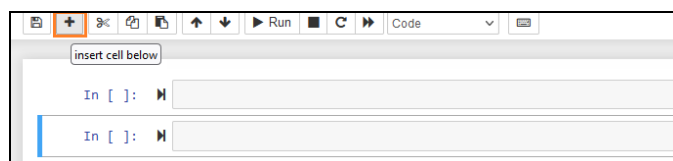


New cells can be added by:

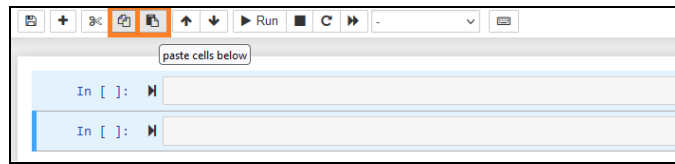
1. Double left clicking in the grey area outside the white border surrounding current cell(s).



2. Keyboard shortcuts: Press A to add a cell above the current cell, or B to add a cell below.
3. Toolbar buttons: the “+” adds a new cell below.



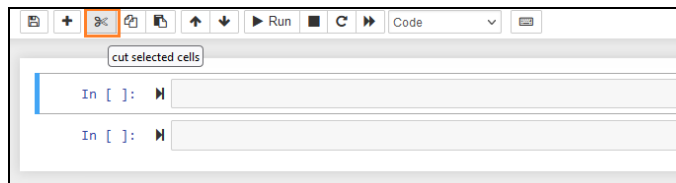
- It's also possible to copy/paste multiple cells. Hold shift to select multiple cells, then copy/paste. As is often the case, this can be done in several ways, via hotkeys or buttons.



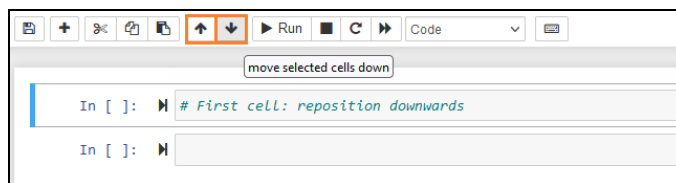
Note: the methods differ in which cell becomes highlighted.

New cells can also be removed or reordered by:

- To cut or remove the selected cell(s), hit the scissors icon or press X while in command mode. These cut cells can be then pasted elsewhere in the notebook, if desired.



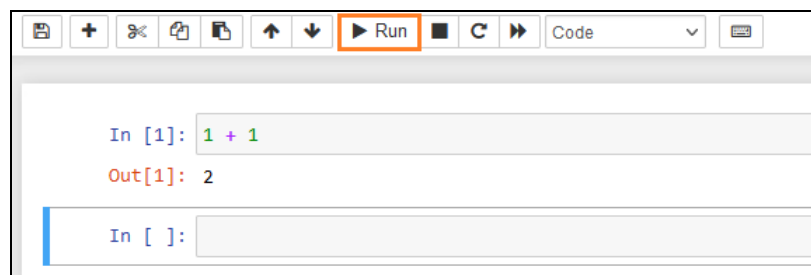
- To reposition the selected cell(s), use the up and down arrows located on the toolbar.



## Running code

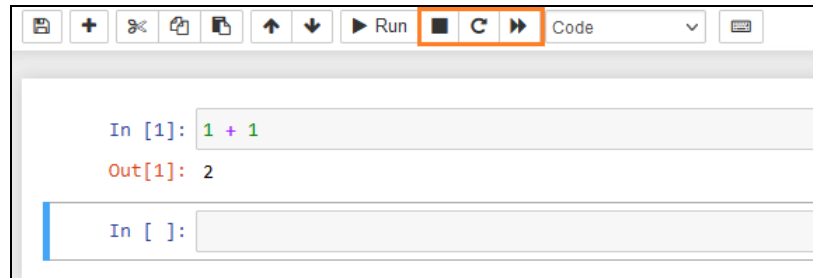
There are several ways to execute code:

- the Run button on the toolbar
- keyboard shortcuts such as Shift + Enter which runs the selected cell and progresses to the next cell.
- keyboard shortcuts such as Ctrl + Enter which runs the current cell without progressing

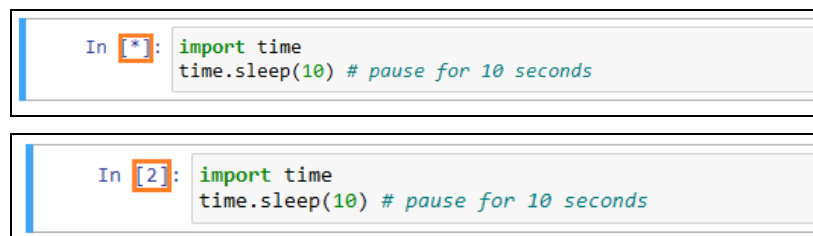


Any output of the cell (text, numbers, plots, etc) will display beneath it.

Also located nearby on the toolbar are buttons for interrupting the kernel, restarting the kernel, and restarting the kernel plus rerunning all notebook code.



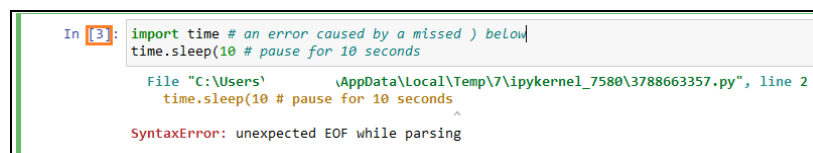
The status of a code cell can be determined by noting whether an asterisk or number is present within the square brackets to the left of the cell. An asterisk signifies that the code is currently being run, whereas a number denotes the number of cells run since the kernel was last reset. The “2” in the second image also indicates that the code finished running.



The ipykernel icon also changes from an outline to a filled grey circle while code is running.

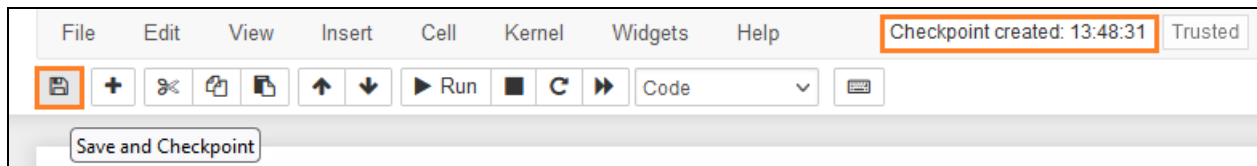


The counter increments whether code successfully executes without error or not. Restarting the notebook kernel also resets this code cell execution counter.

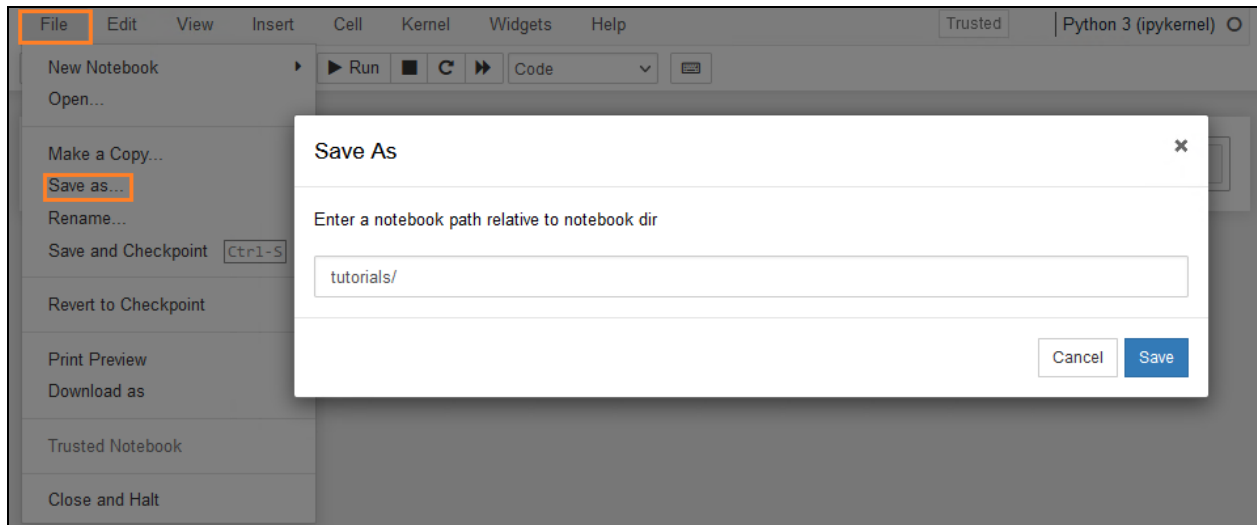


## Saving your work

To save your work in a Jupyter notebook, left click the save icon on the top left of the toolbar. You can confirm that the notebook has been saved by the checkpoint created timestamp that appears in the top right. This automatically saves the .ipynb file to the directory it was loaded from.

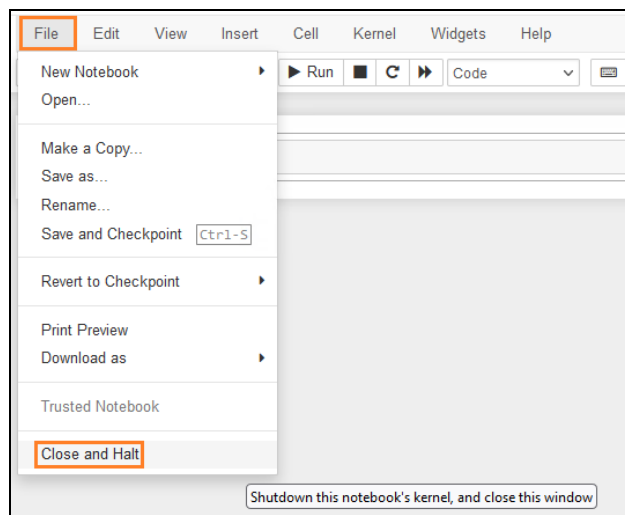


To save to a different location, use the File menu and select Save As, as shown below.

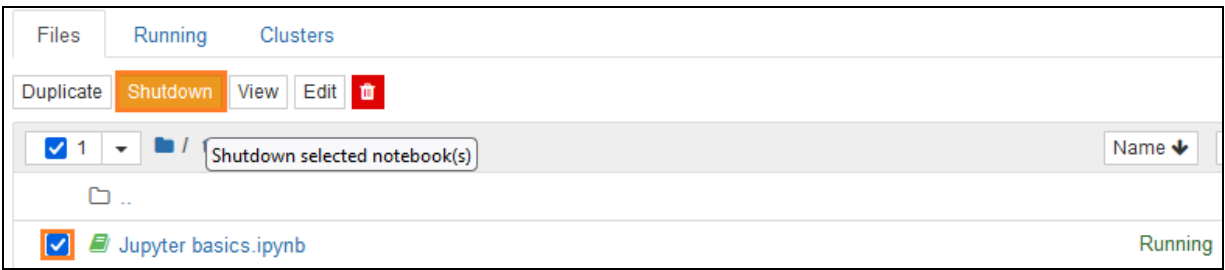


## Shutting down the Jupyter Notebook server

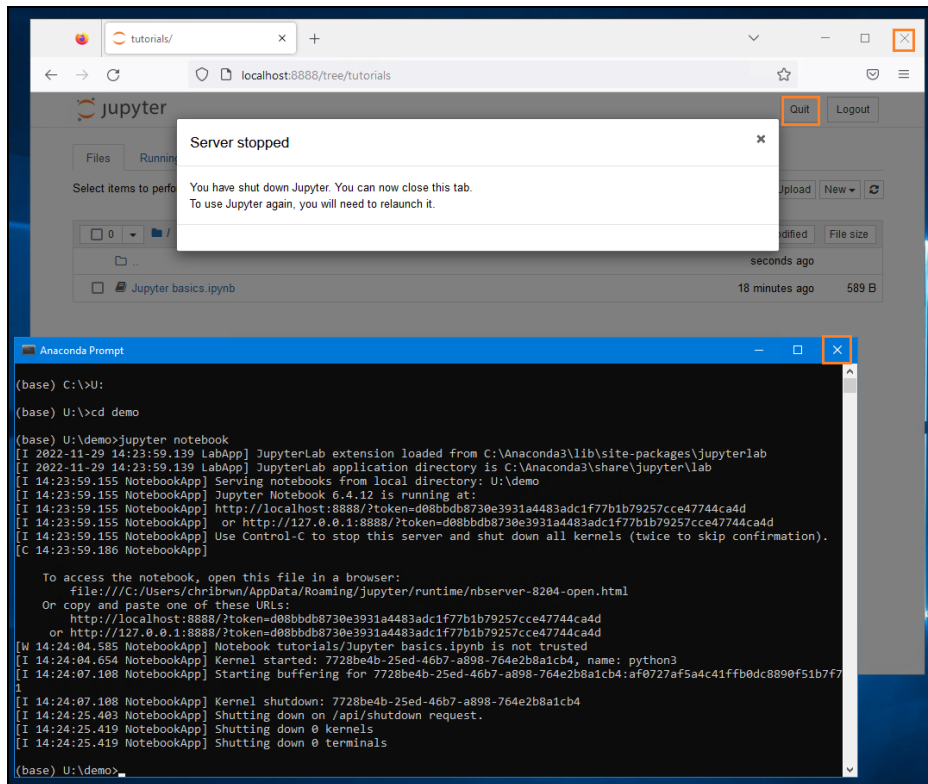
When you're finished working, exit out of Jupyter Notebook by left clicking on the File menu, then select Close and Halt. This will close the tab and return you to the initial Jupyter server page. If the notebook kernel was halted, the corresponding notebook icon will be grey.



You can also shutdown an active notebook kernel by left clicking the checkbox next to the notebook then selecting shutdown. Be careful about doing this prior to saving any work.



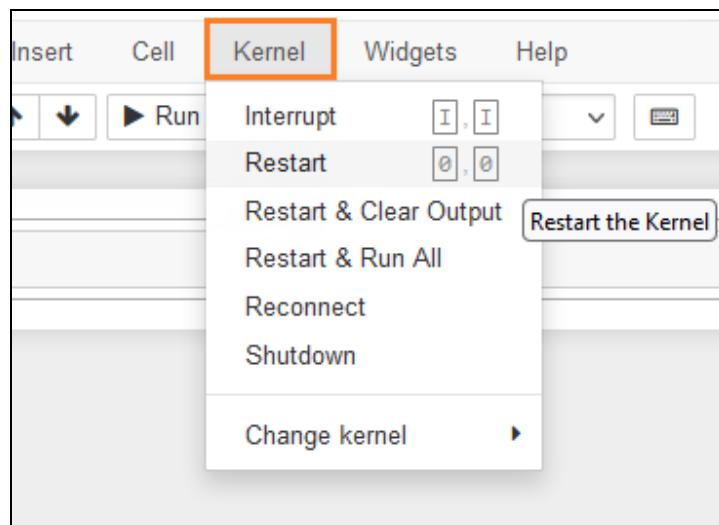
After your work is saved and the notebook kernel shutdown, left click the “Quit” button on the Jupyter server page. Then simply click the “X” on the Jupyter server page and Anaconda Prompt to finish exiting.



## Troubleshooting

### Restarting the kernel

Sometimes, while running a code cell, it may take an unusually long time to run or become stuck. It may help to restart the kernel as shown. This will reset the notebook environment, meaning that packages, data, code, etc will need to be loaded and run again.



Optionally, cell output can be cleared at the same time. It may be useful to restart the kernel periodically while developing Python code to ensure everything works as expected.

### “\_xsrif” argument missing from post

If you encounter difficulty saving your notebook and notice a message in the top right of the notebook that says “\_xsrif” *argument missing from post*, often you can recover your work by reopening the same notebook (without closing the first).

```
'_xsrif' argument missing from POST
```

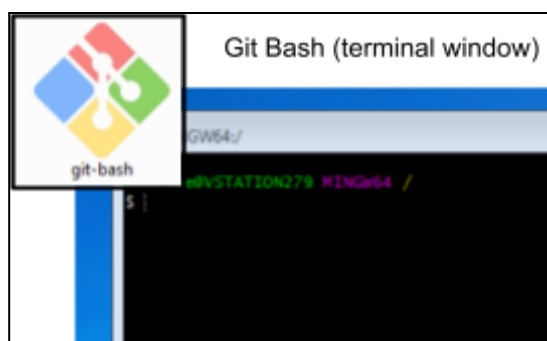
Ideally, this action refreshes the kernel used by the first notebook. After the second notebook loads, try saving your work in the first notebook. Once your work is recovered, you can safely close the second notebook and continue.

# Instructions for Using Git in the VDE

**NOTE:** for more general documentation on Git, see: <https://git-scm.com/doc>

## Setting up an account

1. Email [irisdatarequests@umich.edu](mailto:irisdatarequests@umich.edu) if you would like to create an account with Git.
  - o The default setup will be a new userspace. If requested, we can also set up a project namespace. Please specify this in your request if desired.
2. We will email you with an account name and password for using Git within the VDE.
3. Within the VDE, you may use the Git GUI (user interface) or access Git via Git Bash (terminal window).



## Creating and accessing your Git repository

1. Login to the VDE with your usual login and open Git Bash
2. Browse to your personal drive (the U:/ drive) i.e., in Git Bash use the command **cd u:**
3. To initiate a Git repository in your personal drive folder, use one of these commands:

*[for a user namespace]:*

**git clone ssh://usernamehere@irisgit:22/public/usernamehere.git**  
replacing “usernamehere” with your login name

*[for a project namespace]:*

**git clone ssh://usernamehere@irisgit:22/public/projectnamehere.git**  
replacing “usernamehere” with your login name and “projectnamehere” with your project namespace.

4. At the password prompt, use the password provided by IRIS for your Git account. This will create your Git repository folder, then you may add code files to this folder to use git within the VDE.

Please note that all Git repositories in the VDE are public within the VDE and accessible to any user in the VDE who has created a Git account.



# Publishing Your Research with IRIS UMETRICS Data

## Citations for IRIS UMETRICS data

IRIS generates a digital object identifier (DOI) for released dataset and documentation, encouraging researchers to cite the data and documentation used for your presentation, working papers, and publications. The citation is available from our website; for example, the citation for the 2022 Core Files and the Summary Documentation are:

The Institute for Research on Innovation & Science (IRIS). IRIS UMETRICS 2022 Data Release: Core Files. Ann Arbor, MI: IRIS [distributor], 2022-04-20, doi: 10.21987/09dw-vm46

Nicholls, N., Ku, R., Brown, C., and Owen-Smith, J. Summary Documentation for the IRIS UMETRICS 2022 Data Release. Ann Arbor, MI: The Institute for Research on Innovation & Science [distributor & publisher], 2022-04-20, doi: 10.21987/df2a-ha30

Researchers are encouraged to contact the IRIS Research Support Team for more information on best practices in sharing data and metadata, versioning, and data citation.

## How to meet data sharing requirements

### Data availability statement

IRIS has prepared a boilerplate language for researchers who need to include a statement of data availability at the time of article submission.

“Data cannot be shared publicly due to privacy protections, but qualified researchers can access UMETRICS in a virtual data enclave environment. These data are available from the Institute for Research on Innovation and Science under a data use agreement; see <https://iris.isr.umich.edu/>.”

### Data access for replication purposes

We understand that reviewers or researchers may want to attempt to replicate findings in a manuscript or publication, utilizing the IRIS UMETRICS research release dataset in the IRIS VDE for this non-research purpose. IRIS has developed a policy that applies to such uses.

Please note: This policy only applies to the IRIS UMETRICS research dataset and does not apply to all of the data member universities submit to IRIS. In particular, an individual wishing to attempt to replicate findings is only provided access to that specific dataset for a two-week period. No seat fees will be charged. To apply for access to the IRIS UMETRICS data for purposes other than research, submit the following materials to [irisdatarequests@umich.edu](mailto:irisdatarequests@umich.edu):

1. A statement of the purpose for which access to data is required and a link to the specific manuscript or publication (generally no more than 1 page)
2. Signature on the IRIS Non-Disclosure agreement by the individual user wishing to conduct the replication
3. Completion of the IRIS VDE Acceptable Use Policy

When the application for access to IRIS data for replication purposes has been approved you will be notified and provided with access details. We emphasize several key requirements for replication work here:

- No data downloads or image capture of data on screen are permitted.
- You will not be permitted to upload external datasets into the VDE.
- No disclosure review request can be placed; thus no findings will be removed from the VDE.

## VDE Account Renewal Process

Each project is authorized for data access in the VDE for its necessary duration. However, individual researchers are required to renew their access on an annual basis. Accounts are valid between September 1 to August 31 of the following year. Each summer, IRIS reaches out to individual account holders for renewal.