Introduction & Computing Resources

2020 DSS Bootcamp

Shawn Santo

08-03-20

Supplementary materials

Companion videos

1. Computing resources

Welcome

Welcome to the Department of Statistical Science (DSS). The following slides, slide decks, and exercises will serve to

- give you an understanding of the computing resources available to you within the DSS and Duke University;
- inform you on the best way to get help with your computing needs within the DSS and Duke University;
- introduce you to R, Python, and version control with Git/GitHub;
- highlight the importance of reproducible research and how the aforementioned software can help.

Duke Computing Resources

Duke NetID

Your Duke NetID is the electronic key to making many Duke resources work.
Make sure your Duke Account has been setup.

All of the following can be accomplished at https://idms-web.oit.duke.edu/portal/:

- Changing your password
- Changing your challenge questions
- Setting up multi-factor authentication

Duke email

With your NetID and password, you can access your email on the web at http://mail.duke.edu

Your Duke email is not permanent; your account expires once you leave Duke.

Duke WiFi

Duke network connections:

- Dukeblue:
  - 24-hour access to secure (encrypted) wireless throughout Duke's residence halls, academic and administrative buildings
- DukeOpen:
  - unencrypted wireless access for devices such as gaming systems, or other devices
- Eduroam (education roaming):
  - secure (encrypted) wireless access using your Duke NetID and password
  - To use eduroam at a participating institution, configure your machine ahead of time while at Duke - https://dukeblue.duke.edu/eduroam/.

Duke VPN

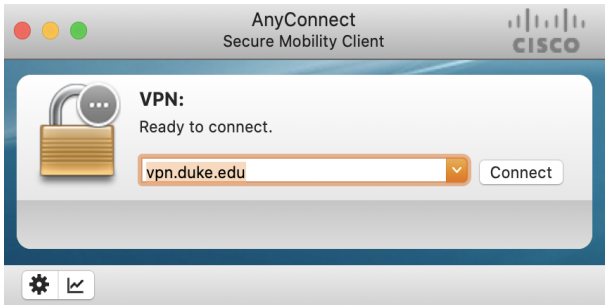
- Duke's virtual private network (VPN) allows you to create a secure connection from your computer to Duke over a public network while working remotely. This will be necessary for you to use, if you want to access certain Duke and DSS resources off campus.

- Instructions to get started with the VPN are available on the next slide. For more information on Duke's VPN visit https://oit.duke.edu/what-we-do/services/vpn.

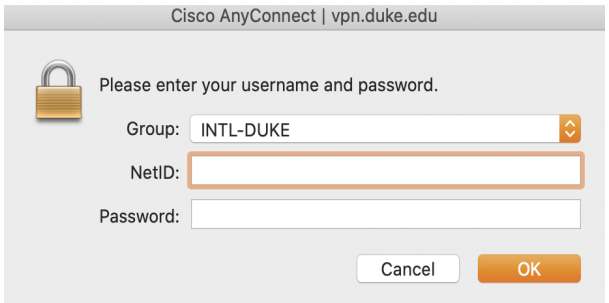
Duke VPN set-up

1. Download and install the free Cisco AnyConnect VPN software

2. Launch Cisco AnyConnect on your machine and enter `vpn.duke.edu` in the box. Click `Connect`.



3. Another dialog box will appear. Choose `INTL-DUKE` from the dropdown menu for `GROUP`, and enter your Duke NetID and password. Click `OK` and you should connect.



Software

- Duke offers software for download to students, faculty, and staff through https://software.duke.edu/
- Duke negotiates with vendors to make software available to the Duke community for discounted rates or, in many cases, for free. If you have any questions, comments or suggestions, please e-mail the site-license office at site@duke.edu.
- Some free software relevant to you as students:
    - Microsoft Office
    - MATLAB & Simulink
    - SAS Education Analytical Suite
    - VPN Client
    - Mathematica
    - Tableau
    - Adobe Creative Cloud

Virtual computing

Duke OIT offers virtual software containers and semester-long virtual machines.

- Virtual Software Containers – Students and instructors can reserve personal computer environments running applications such as RStudio, the Eclipse IDE, Jupyter Notebooks, Matlab, Octave, Mathematica, and others for a semester. These are run through your web browser; no software download is required. Two containers you will find most useful are:
  - `RStudio - statistics application with Rmarkdown and knitr support`
  - `Jupyter - interactive data science and scientific computing notebooks`
- Virtual Machine (VM) – Your Duke VM is like having a second computer that lives with Duke. You can log into and use your VM from your own machine. It allows you to access specialized software without installing it on your own computer, host your own server for development projects and coursework, or customize your own environment to use for the semester.
  - Run Windows or Linux VMs
  - Computing resources are light: 2 processors and 2 GB

Duke Compute Cluster

The Duke Compute Cluster (DCC) consists of machines that the University has provided for community use and that researchers have purchased to conduct their research. You will need to be granted access before use.

- Runs on Linux - Red Hat Enterprise Linux 7
- Offers over 200 compute nodes and over 20,000 cores
- Most nodes are purchased by labs and departments. The DSS has three nodes.
- Uses the SLURM job management system

The DCC User Guide will help you get up and running. They also host workshops to help new users.

Getting help with Duke resources

Duke Office of Information Technology (OIT) manages Duke's technology infrastructure and application support.

- https://oit.duke.edu/help
- Live chat, 24 hours a day, Monday - Thursday; chat is available on a limited basis Fridays and Sundays
- Walk up hours are available at the Link in Perkins Library.

DSS Computing Resources

RStudio Pro

The DSS has an RStudio Pro license that will allow you to run an instance of RStudio in your browser while harnessing the computing power of a remote multiprocessor server.

To access RStudio Pro:

1. If off campus, use the VPN to create a secure connection from your computer to Duke. If you are on campus, be sure you are connected to the Dukeblue network.

2. Navigate to one of

    - http://pawn.stat.duke.edu:8787
    - http://rook.stat.duke.edu:8787 (MS students)
    - http://knight.stat.duke.edu:8787 (PhD students)

    You'll have access to all three, but we try to spread things out to prevent overload.

3. Log-in with your Duke NetID and password.

Other resources

- Local file system on DSS servers
    - Connect via SSH to
        - pawn.stat.duke.edu
        - rook.stat.duke.edu
        - knight.stat.duke.edu
        - monster.stat.duke.edu
- RStudio Connect
    - a publishing platform for the work you do in R and Python
    - share Shiny applications, R Markdown reports, Plumber APIs, dashboards, Jupyter Notebooks, interactive Python content, and more
    - if you need access, ask and I can help get you set-up
- Email aliases: https://stat.duke.edu/resources/email/aliases
- Archived resources page: https://stat.duke.edu/resources/

Getting help with DSS resources

The best way to get help with DSS computing resources is to email stat-help@duke.edu. One of our great IT staff members will get back with you ASAP.

Zoyia Melton - Senior IT Analyst

- Phone: 919-684-5419
- Location: 027 Old Chemistry
- Office Hours: TBD

Frisco Rose - Assistant Research Professor in the Department of Mechanical Engineering and Materials Science

- Phone: 919-660-5496
- Location: 026 Old Chem
- Office Hours: TBD

References

1. DCC User Guide | Research Computing. (2020). Retrieved from https://rc.duke.edu/dcc/dcc-user-guide/

2. Software Licensing. (2020). Retrieved from https://software.duke.edu/

3. VPN. (2016). Retrieved from https://oit.duke.edu/what-we-do/services/vpn