

Servidor de investigación

Instructivo básico de uso

Introducción

Existe en nuestra facultad un servidor especialmente dedicado para cómputo, que a diferencia de otros servidores que la facultad dispone, no tiene como finalidad guardar información institucional o personal, ni disponer de información académica ni de otro tipo de que pudiera tener carácter sensible; sólo sirve a los fines de cómputos de alta performance. Pero sí está conectado a la red local de la FCE y es accesible desde cualquier máquina de escritorio que esté conectada por red cableada a dicha red local. Esta máquina, que llamaremos “server de investigación” o simplemente “server”, difiere de una PC de escritorio en varios aspectos. En sí no es distinta de una máquina común de no ser por la gran cantidad de núcleos de su procesador, 56 en total, que están disponibles en todo momento a toda investigadora que posea una cuenta, y de mucha más memoria ram, 64 gigas, que una máquina usual de escritorio.

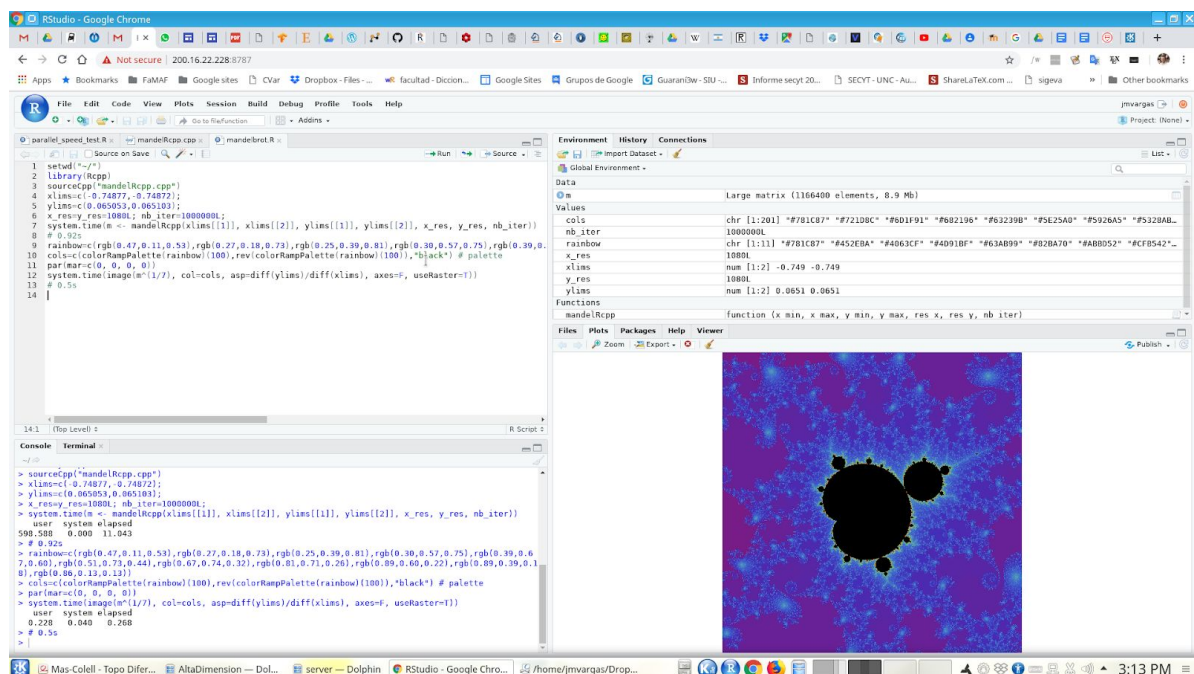
En este servidor, si tiene cuenta para acceder, dispone de un directorio exclusivamente para su uso personal que puede almacenar las bases de datos que necesita procesar y en el cual puede guardar las salidas para su posterior análisis. Es crucial, entonces, entender cómo se puede acceder a esta facilidad, mover archivos desde y hacia el servidor, y usarla desde su escritorio en la FCE.

Cómo conectarse al servidor de investigación

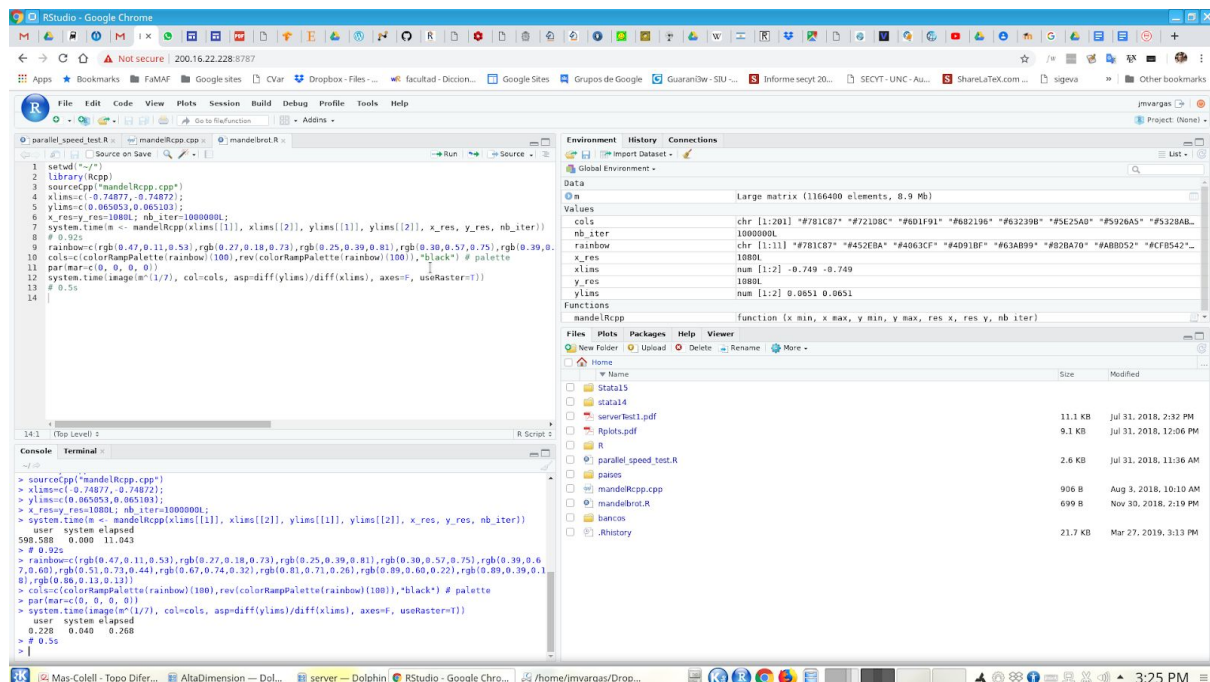
Para acceder, se debe tener una cuenta en el servidor. Si Usted no tiene una cuenta, se debe dirigir a las oficinas de administración de la Prosecretaría de Informática (ex Cómputos) de nuestra facultad para que se le provea de una.

Si ya dispone de una cuenta en el servidor, puede acceder al mismo de tres formas:

1) **Para acceder a RStudio** vía un navegador web como chrome o firefox, colocando en la ventana de direcciones <http://investigacion.eco.unc.edu.ar:8787/auth-sign-in> después de hacer *login* con su usuario y clave, debería ver en su navegador algo como esto

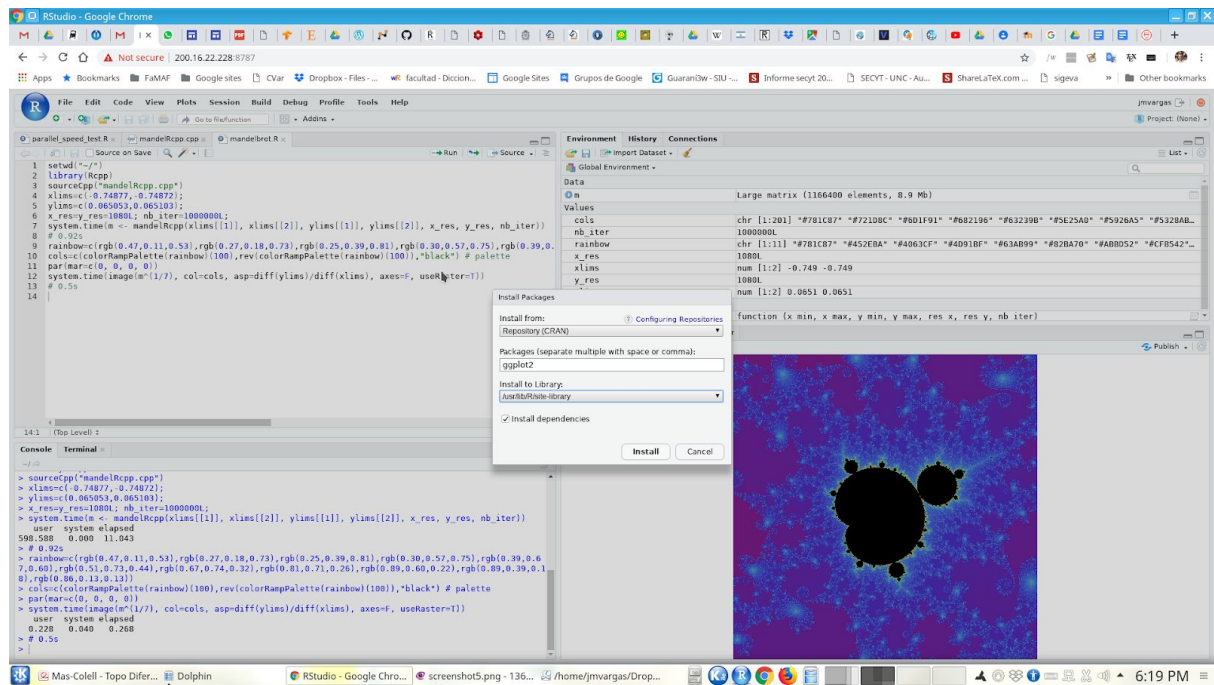


Ese es el entorno de programación RStudio para R (y otros lenguajes que se pueden usar en funciones con R como C, C++, fortran, python y más) como se ve desde un típico escritorio, sin importar el sistema operativo que usted tenga en su máquina de escritorio. En la ventana inferior derecha de RStudio, aparece su directorio personal en el server:



Desde los controles en la solapa “files” en esa ventana, puede crear directorios, borrar archivos, moverlos dentro de su directorio (sin salir del server) y otras utilidades más. Para mover una base de datos, u otro archivo, desde su máquina personal a su directorio en el server, existe una conexión samba que describiré en la próxima sección.

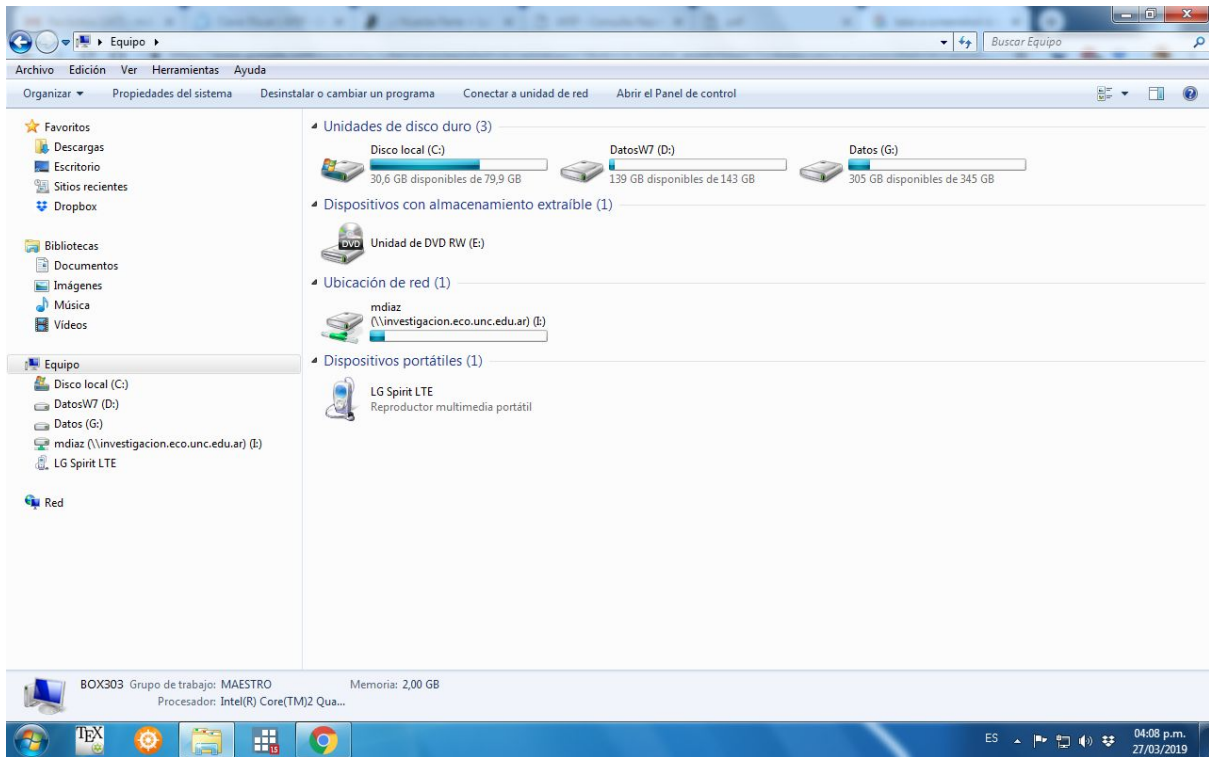
Si necesita instalar librerías de R, use desde RStudio, “Tools”, “Install libraries”, y, muy importante, use el directorio `/usr/lib/R/site-library` que se encuentra en las opciones desplegadas de “Install to Library”, como se ve en la siguiente captura de pantalla



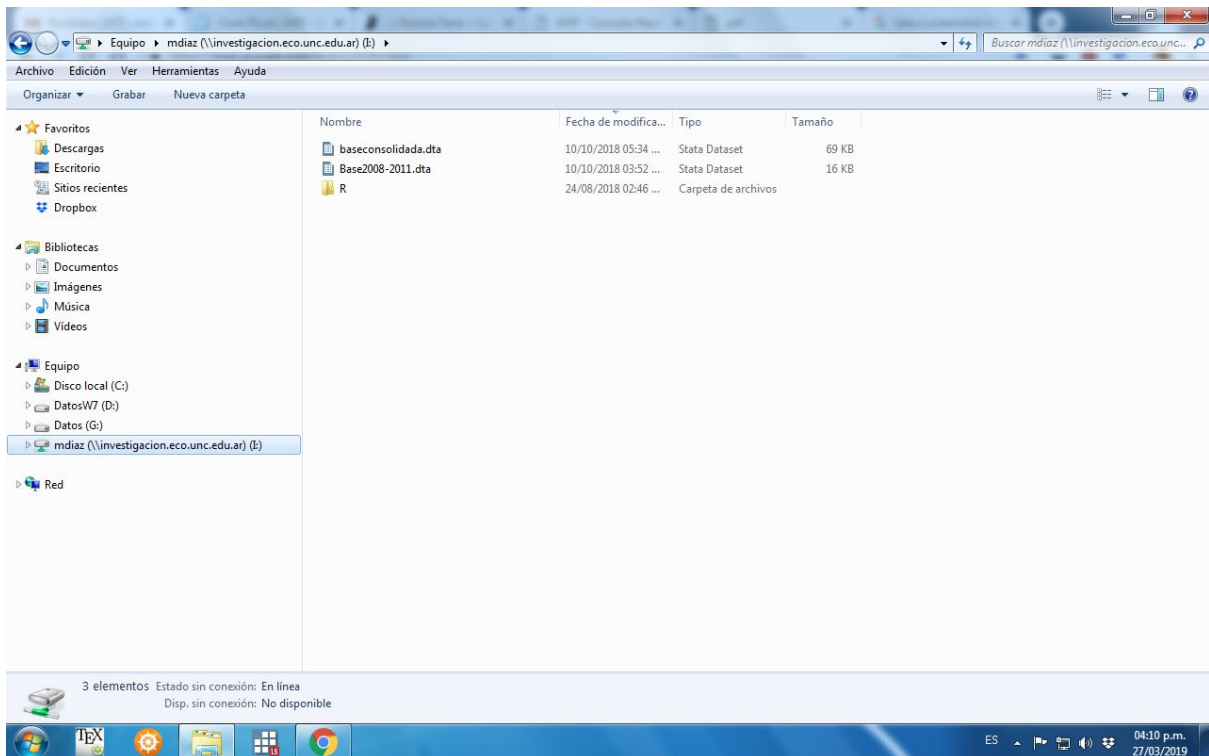
Esto tiene la ventaja de instalar las librerías para todos los o las usuarios, en vez de repetir librerías en cada usuario. Se ahorra espacio en el servidor.

2) **Conexión samba (smb protocol)**. Si tiene una PC de escritorio con windows como sistema operativo y tiene una cuenta en el servidor, es posible que ya disponga de una conexión especial desde el explorador de windows creada por la administración de sistemas de la facultad. Si Usted no la tiene, deberá solicitarla en las oficinas de sistemas. Asumiendo que dispone de esta conexión, debería ver un directorio que aparece en su explorador como si estuviese en el disco rígido de su máquina, pero en realidad se encuentra

en el server; es su directorio personal en el server.



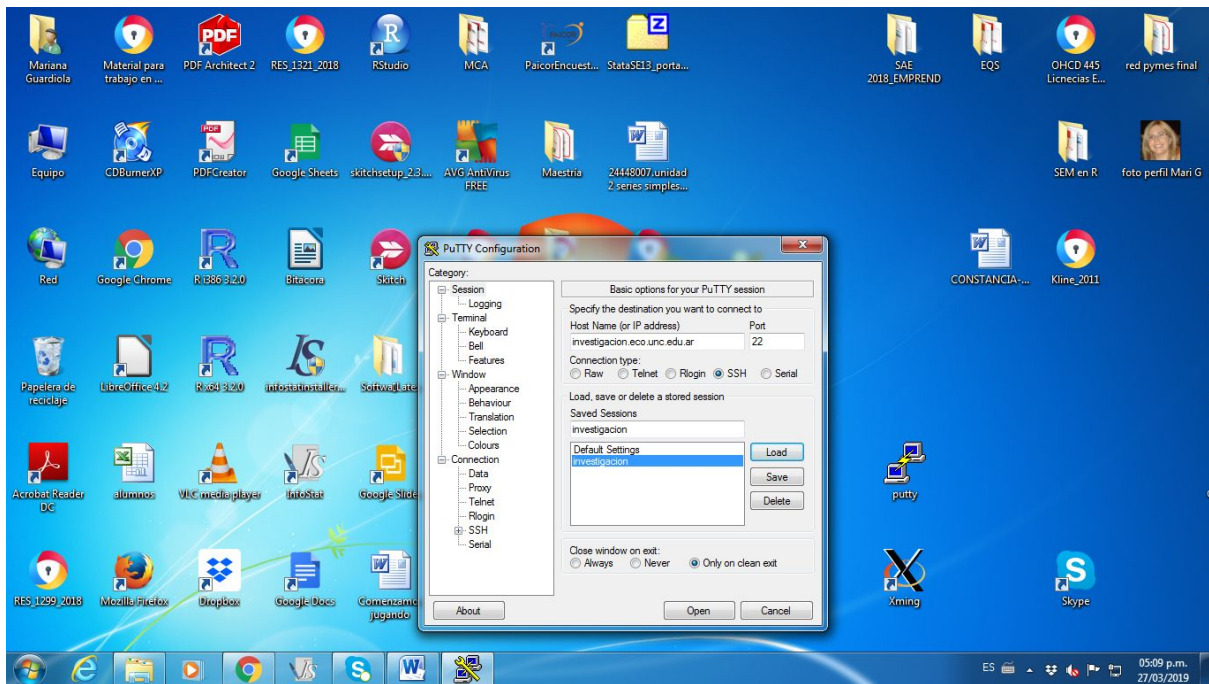
Se puede ver ese directorio en “Ubicación de red (1)”. Si accede haciendo doble clic, verá su directorio en el server



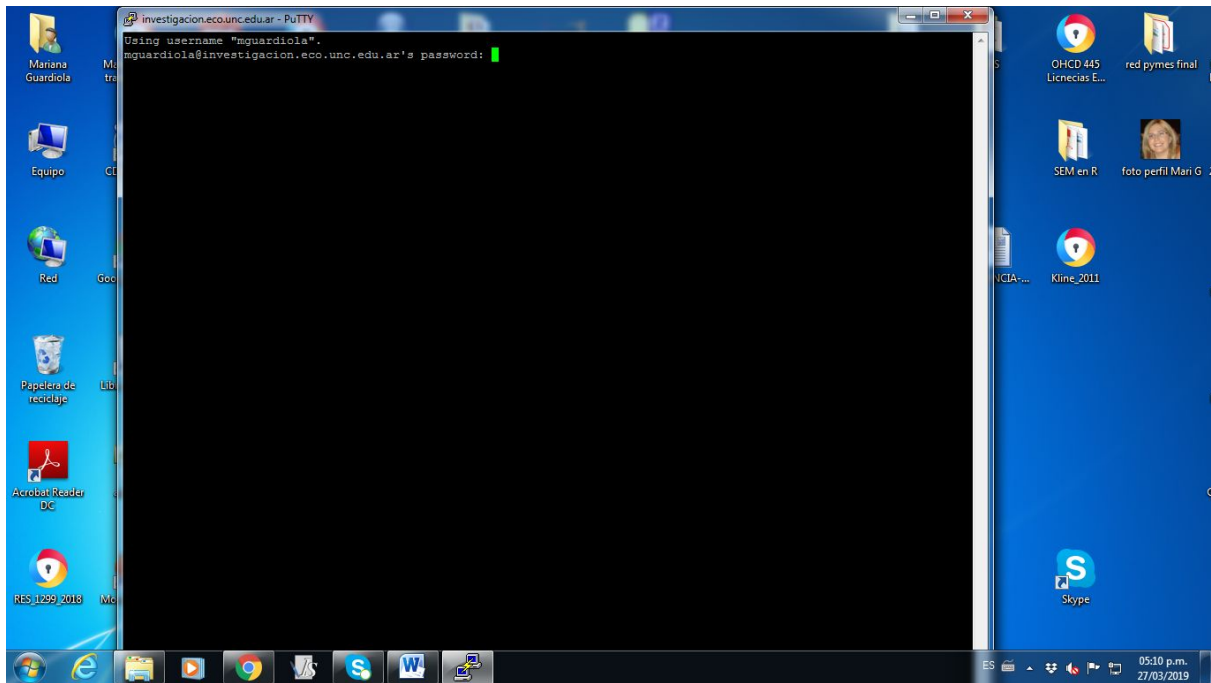
Desde el explorer de windows puede entonces mover, copiar, crear o borrar archivos. Puede copiar un directorio de su disco a ese directorio. Un pedido especial desde la

administración del servidor: Usted podría potencialmente copiar y mover al servidor archivos que no se ajusten al propósito de esta facilidad, por ello contamos con confianza en su responsabilidad.

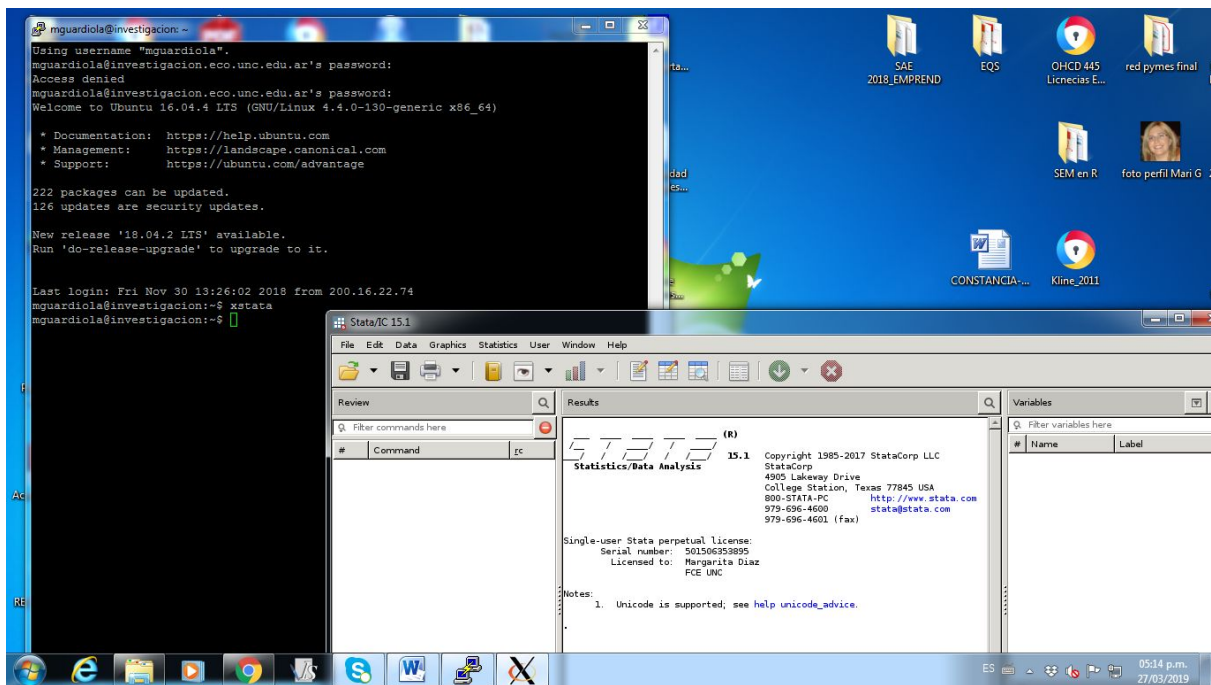
3) **Conexión vía ssh (*secure shell protocol*) para usar STATA15.** Existe una tercera forma de conexión al servidor por medio del protocolo *ssh* de linux, que es el sistema operativo en el server. Si Usted tiene una cuenta en el servidor, puede acceder al mismo usando líneas de comando desde una terminal *tty*. Si su máquina de escritorio tiene windows, dispone de una terminal *tty* a través del programa PuTTY, que seguramente esté ya instalado, y de Xming que debería estar residente. Proceda de esta manera: en el escritorio hay un icono para el programa Putty, haga doble click, y verá esto



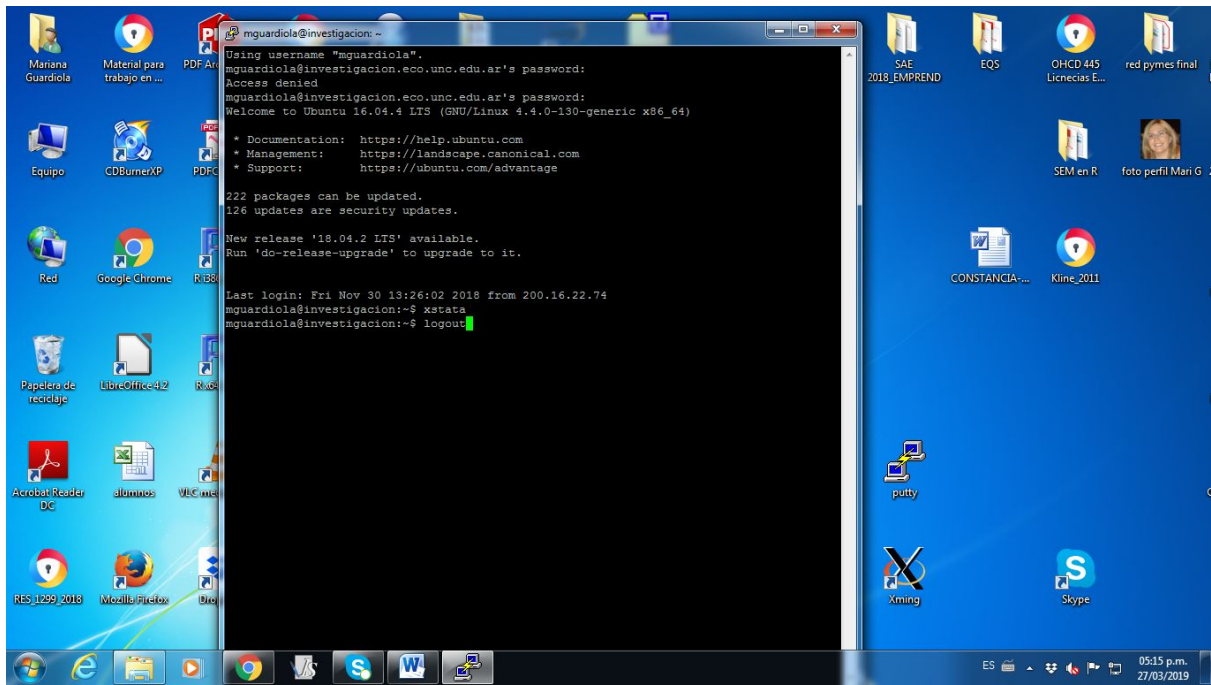
Seleccione “investigación”, haga clic en “Load”, luego en “Open”. Y verá esta pantalla



Esta es una terminal tty. Escriba (aunque no aparezca lo que escriba en pantalla) su usuario, enter, luego su password, enter. Debería ver la siguiente pantalla



En la terminal, escriba después del símbolo de prompt \$
xstata
y se abrirá stata15, como se ve en la captura de pantalla de arriba.
Para cerrar stata, ciérrelo desde su ventana, File, Close. Finalmente, desde el símbolo de prompt \$, escriba logout, luego enter. Debería verse así antes de dar enter



Cuando da enter, no sólo cierra su sesión en el server, lo cual es una buena práctica por seguridad para la red local, sino que también cierra la terminal de Putty.

Consideraciones finales

El alcance de este instructivo es que Usted pueda acceder al servidor desde su máquina de escritorio, mover archivos desde y hacia el servidor, usar R a través de RStudio, y usar la última versión (legal, multinúcleo) de STATA15. Existen muchísimas operaciones y procedimientos posibles de cómputo a través de programación en lenguajes que el server tiene ya instalados como C, C++, fortran, Python y Julia, que son posibles desde una terminal pero que van mucho más allá de este instructivo. Para disponer de toda la potencia del servidor es necesario que los programas en R o Python hagan uso explícito en el código de paralelización. En ese respecto, está instalado, por ejemplo, la librería OpenMP para C++.

STATA15 aún cuando es multicore, no todas sus subrutinas están preparadas para varios procesadores; sólo algunas. Por esto la velocidad de procesamiento es similar al de una máquina de escritorio. Por supuesto, aún en ese caso, la gran ventaja es que se paga una licencia en lugar de muchas individuales, optimizando los recursos de la facultad.

Si su código R hace uso de procesamiento en paralelo, por ejemplo, a través de las librerías doParallel o parallelR, por experimentación directa, el beneficio marginal de usar más de ocho núcleos es negligible. Por eso, para que otras usuarias puedan tener recursos, no use más de ocho en su código. Pero si Usted implementa funciones en C o C++, puede usar OpenMP y 56 núcleos (y ocupar todos los recursos). En ese caso ninguna máquina de escritorio puede igualar esa capacidad de cómputo.

Esta facilidad instalada, disponible a todas las investigadoras de la FCE, abre posibilidades de procesamiento estadístico que no son posibles desde una máquina de escritorio, ni siquiera esperando días o semanas, como modelos georeferenciados bayesianos para no muy grandes bases de datos por MCMC o INLA.