2023 Datascience remote server

Libraries and programs for a generic datascience remote server to test data-science-related topics and setups.

- 1.1. How to use it
  - 1.1.1. User credentials
  - 1.1.2. R & RStudio
- 1.2. How it has been developed
  - 1.2.1. Operating System
- 1.3. Afegir repositoris extra
- 1.4. Configuració del mailer (postfix)
- 1.5. Altres paquets de sistema necessaris
- 1.6. Afegim Rstudio Server - Posit Server
  - 1.6.1. Afegim SSL letsencrypt per permetre connexions https a Rstudio
  - 1.6.2. Allow GUI connections
  - 1.6.3. R 4.x
  - 1.6.4. Allow installing packages or upgrade in R system packages
- 1.7. Set default locale as UTF-8
- 1.8. Shiny

1.1. How to use it

1.1.1. User credentials

All people will have their own credentials, sent to their email addresses.

Users will be in the sudoers group, so that you can run commands as root if your prepend those commands with sudo, as usual.

1.1.2. R & RStudio

Open the browser, and it will launch RStudio server in it by default (`http://datascience.seeds4c.org:8787`).

You have R 4.x installed.

1.2. How it has been developed

1.2.1. Operating System

Ubuntu GNU/Linux 20.04 (64 bits) lxc container + LXQt desktop.

Beware that R 4.5 cannot work as expected with RStudio in Ubuntu 20.04. Last version to work as

expected is R 4.4.x

1.3. Afegir repositoris extra

Comanda a una consola

```
user@computer:~$ sudo apt install gpg software-properties-common
```

Ara ja podem llençar les següents instruccions sobre la finestra de terminal (copiar tot i enganxar sobre la finestra del terminal, amb botó dret del ratolí i `Enganxa`, o bé amb la drecera per enganxar a les finestres de terminal `Control + Shift + V`)

Comanda a una consola

```
# Add the key for the new repo for R 4.1+ from cloud.r-project.org
sudo apt-key adv --keyserver hkp://keyserver.ubuntu.com:80 --recv-keys
E298A3A825C0D65DFD57CBB651716619E084DAB9
sudo su

# Update packages list again, just in case
apt update
```

Seguim:

Comanda a una consola

```
add-apt-repository -y 'deb https://cloud.r-project.org/bin/linux/ubuntu focal-
cran40/' # main binary packages for R 4.1+
add-apt-repository -y ppa:c2d4u.team/c2d4u4.0+ # extra binary packages for R 4.1+
from the cran2deb4ubuntu Build Team
exit
```

1.4. Configuració del mailer (postfix)

```
sudo su
apt install postfix
dpkg-reconfigure postfix # configurar com a internet site amb FQDN el domini que
```

```
ens ha proveït l'IMI que acaba amb imi.bcn
service postfix start
exit
```

1.5. Altres paquets de sistema necessaris

Comanda a una consola

```
sudo apt-get install -y curl bwidget dos2unix freeglut3 freeglut3-dev git libc6
libcairo2-dev libcurl4-gnutls-dev libgdal-dev libgeos-dev libglpk-dev libgraphviz-
dev libjq-dev libmagick++-dev libmpfr-dev libproj-dev libprotobuf-dev libssh2-1-dev
libssl-dev libudunits2-dev libnode-dev libx11-dev libxml2 libxml2-dev libxt-dev
pandoc protobuf-compiler r-cran-rjava r-base-core texlive-lang-spanish texlive-
latex-extra unaccent xvfb libssh2-1-dev libudunits2-dev apt-transport-https alien
pigz corkscrew libdbi-perl sendemail libharfbuzz-dev libfribidi-dev cmake fail2ban
```

1.6. Afegim Rstudio Server - Posit Server

Afegim RStudio server, després d'afegir els paquets de sistema extres necessaris. Info treta de: https://posit.co/download/rstudio-server/[1]

```
sudo apt-get install gdebi-core
wget
https://s3.amazonaws.com/rstudio-ide-build/server/bionic/amd64/rstudio-server-2022.
07.2-576-amd64.deb # For Ubuntu 18/20
# wget
https://s3.amazonaws.com/rstudio-ide-build/server/jammy/amd64/rstudio-server-2022.0
7.2-576-amd64.deb # For Ubuntu 22
sudo gdebi rstudio-server-2022.07.2-576-amd64.deb
```

Nota:
Les versions posteriors, com les de 2022.12.*, 2023.03.*, semblen tenir algun problema amb la connexió des de les màquines del Consorci d'Educació de Barcelona (CEB): potser algun tipus de connexió blocada des del firewall de la infraestructura informàtica del CEB/CTTI/Generalitat de Catalunya, potser. El que hem vist és que la versió 2022.07.2* si que permet les connexions com de costum.
I cal tenir present que per a R 4.3.x, la versió de Rstudio-server 2022.07.2* sembla ser massa vella, i reporta un warning:

> *R graphics engine version 16 is not supported by this version of RStudio. The Plots tab will be disabled until a newer version of RStudio is installed.*

Les versions anteriors a la darrera disponible, es poden consultar i descarregar des de: https://docs.posit.co/previous-versions/rstudio/[2]

Nota 2:
Actualitzar de R 4.4.x a R 4.5.x fa fallar l'RStudio server que hi havia instal.lat (de 2024).
Actualitzar RSTudio a 2026.x fa fallar la instal.lació per que necessita libssl 3.x., tot i que a repositoris d'ubuntu 20.04 només hi ha libssl 1.1.
Una opicó semblaria ser instal.lar libssl 3.x a partir del codi font, seguint aquestes insrtuccions https://orcacore.com/install-openssl-3-ubuntu-20-04/[3] però no ha resolt el problema.
Així que he forçat a mantenir la versió d'R a 4.4.x amb:

```
sudo apt-get install r-base=4.4.3-1.2004.0 r-recommended=4.4.3-1.2004.0  r-base-dev=4.4.3-1.2004.0 r-base-html=4.4.3-1.2004.0  r-base-core=4.4.3-1.2004.0
```

1.6.1. Afegim SSL letsencrypt per permetre connexions https a Rstudio

Tret de:
https://adisarid.github.io/post/2020-03-06-setup_rstudio_server_with_ssl/[4]

1.6.1.1. Install Let’s Encrypt and get certficates

Install the following software on your linux server

```
sudo apt update
sudo apt install letsencrypt
sudo apt install nginx
```

Update your nginx configuration as preperation for obtaining the let’s encrypt certificate. This step is needed because when requesting a certificate from let’s encrypt, the let’s encrypt server will try

to authenticate your server.

Use

```
sudo nano /etc/nginx/sites-enabled/default
```

And add the following (replace datascience.seeds4c.org with your domain):

```
server {
        listen 80;
        listen [::]:80;
        root /var/www/datascience.seeds4c.org/html;

        # Add index.php to the list if you are using PHP
        index index.html index.htm index.nginx-debian.html;

        server_name datascience.seeds4c.org;
}
```

Get your SSL certificates using the following line, just replace datascience.seeds4c.org with your subdomain.

```
letsencrypt certonly -a webroot --webroot-
path=/var/www/datascience.seeds4c.org/html/ -d datascience.seeds4c.org
```

> *IMPORTANT NOTES:*
> *- Congratulations! Your certificate and chain have been saved at: /etc/letsencrypt/live/datascience.seeds4c.org/fullchain.pem*
> *Your key file has been saved at: /etc/letsencrypt/live/datascience.seeds4c.org/privkey.pem*
> *Your cert will expire on 2023-08-14. To obtain a new or tweaked version of this certificate in the future, simply run certbot again. To non-interactively renew *all* of your certificates, run "certbot renew"*
> *- Your account credentials have been saved in your Certbot configuration directory at /etc/letsencrypt. You should make a*

> *secure backup of this folder now. This configuration directory will also contain certificates and private keys obtained by Certbot so making regular backups of this folder is ideal.*
> *- If you like Certbot, please consider supporting our work by:*
>
> *Donating to ISRG / Let's Encrypt: https://letsencrypt.org/donate*[5]
> *Donating to EFF: https://eff.org/donate-le*[6]

Update your nginx settings again

```
sudo nano /etc/nginx/sites-enabled/default
```

To have the following setup (remember to replace datascience.seeds4c.org with your domain):

```
map $http_upgrade $connection_upgrade {
  default upgrade;
  ''      close;
}

# listens on port 80 and redirects traffic to secure alternative
server {
        listen 80 default_server;
        listen [::]:80 default_server;
        server_name datascience.seeds4c.org;
        return 301 https://datascience.seeds4c.org$request_uri;
}

server {
        # SSL configuration
        listen 443 ssl;
        ssl_certificate
/etc/letsencrypt/live/datascience.seeds4c.org/fullchain.pem;
        ssl_certificate_key
/etc/letsencrypt/live/datascience.seeds4c.org/privkey.pem;
        ssl_protocols TLSv1.2;

        ssl_ciphers
EECDH+AES128:RSA+AES128:EECDH+AES256:RSA+AES256:EECDH+3DES:RSA+3DES:!MD5;
        ssl_prefer_server_ciphers On;
        ssl_session_cache shared:SSL:128m;
        add_header Strict-Transport-Security "max-age=31557600; includeSubDomains";
        ssl_stapling on;
```

```
        ssl_stapling_verify on;

        root /var/www/datascience.seeds4c.org/html;

        server_name _;

        # Reroute traffic to shiny server (i.e., reverse proxy for port 3838)
        location /shiny/ {
           proxy_pass http://127.0.0.1:3838/;
           proxy_http_version 1.1;
           proxy_set_header Upgrade $http_upgrade;
           proxy_set_header Connection $connection_upgrade;
           rewrite ^(/shiny/[^/]+)$ $1/ permanent;
        }

        # Reroute traffic to rstudio server (i.e., reverse proxy for port 8787)
        location / {
           proxy_pass http://127.0.0.1:8787/;
           proxy_http_version 1.1;
           proxy_set_header Upgrade $http_upgrade;
           proxy_set_header Connection $connection_upgrade;
        }
}
```

We restart nginx server



```
sudo service nginx restart
```

Go ahead and browse to your domain (e.g., https://datascience.seeds4c.org[7]). Check that you're able to login properly and that all pages are secure on https.

1.6.1.2. Renew letsencrypt automatically

Your server should be working now, but since Let's Encrypt certificates only last 90 days, lets put an automatically renewal process in place.



```
sudo nano /opt/renewCerts.sh
```

Paste the following text:



```
#!/bin/sh
# This script renews all the Let's Encrypt certificates with a validity < 30 days
if ! letsencrypt renew > /var/log/letsencrypt/renew.log 2>&1 ; then
    echo Automated renewal failed:
    cat /var/log/letsencrypt/renew.log
    exit 1
fi
nginx -t && nginx -s reload
```

Make sure the script is owned and executable by root:

```
chown root.root /opt/renewCerts.sh
chmod u+x /opt/renewCerts.sh
```

Add it to cron for auto execution:

```
sudo crontab -e
```

Add

```
@weekly /opt/renewCerts.sh
```

All should be set!

Go ahead and browse to your domain (e.g., https://datascience.seeds4c.org[8]). Check that you're able to login properly and that all pages are secure on https.

1.6.2. Allow GUI connections

With X2Go (https://wiki.x2go.org[9]) you can do so, from computers using GNU/Linux, Mac OSX or MS Windows

```
sudo add-apt-repository ppa:x2go/stable
```

```
sudo apt-get update
sudo apt-get install x2goserver x2goserver-xsession
sudo apt-get install --no-install-recommends lxqt
sudo apt-get install x2golxdebindings
```

Connect with X2Go client to server `datascience.seeds4c.org` , choosing as a session:

- `LXQt`

Launch parcellite and kupfer. Change parcellite to store 250 entries. And set kupfer to launch automatically on user login.

For demostration purposes, a full lubuntu desktop can be added in this virtual machine for hgiher usability when connecting through X2go and to have usual default programs there as if it was a desktop computer.

```
sudo apt-get install lubuntu-desktop
```

1.6.3. R 4.x

We add these repos to use the latest R versions released

Obrim un terminal de sistema, i iniciem una consola de R a dins escrivint `R<enter>`

Comanda a un terminal de R

```
if (!require("pacman")) install.packages("pacman"); require("pacman")
if (!require("renv")) install.packages("renv"); require("renv")
if (!require("devtools")) install.packages("devtools"); require("devtools")
```

Comandes i paquets lubuntu 22.04:

```
sudo apt-get install -y  bwidget dos2unix freeglut3 freeglut3-dev git libc6
libcairo2-dev libcurl4-gnutls-dev libgdal-dev libgeos-dev libglpk-dev libgraphviz-
dev libjq-dev libmagick++-dev libmpfr-dev libproj-dev libprotobuf-dev libssh2-1-dev
libssl-dev libudunits2-dev libnode-dev libx11-dev libxml2 libxml2-dev libxt-dev
pandoc protobuf-compiler r-recommended subversion texlive-lang-spanish texlive-
latex-extra texmaker tk-dev tk-table unaccent xvfb libssh2-1-dev libudunits2-dev
```

```
gigolo filezilla openjdk-8-jre libglpk-dev cargo libgeos-dev libgdal-dev librsvg2-
dev libmagick++-dev libcairo2-dev libharfbuzz-dev libfribidi-dev libsodium-dev
#sudo R CMD javareconf
```

Paquets de CRAN: posar dins de la comanda:


```
if (!require(pacman)) {install.packages("pacman")}; library("pacman")
p_load("tidyverse", "caTools", "bitops", "httpuv", "devtools", "rpivotTable", "DT",
"shiny", "magick", "rvg", "addinslist", "ff", "sparklyr", "data.table", "rio",
"radiant", "CRANsearcher", "rJava", "knitr", "rmarkdown", "webshot", "magick",
"rsvg", "sf", "leaflet", "htmlwidgets", "arrow", "renv", "readxl", "writexl", "gt",
"janitor", "fst", "bookdown", "learnr", "datos")
```

Rstudio Addins: CRANsearcher, addinslist

1.6.4. Allow installing packages or upgrade in R system packages



```
sudo chmod 777 /usr/lib/R/site-library /usr/lib/R/site-library/* -R
sudo chmod 777 /usr/local/lib/R/site-library /usr/local/lib/R/site-library/* -R
sudo chmod 777 /usr/lib/R/library /usr/lib/R/library/* -R
sudo chmod 777 /usr/share/R/doc/html/* -R
```

1.7. Set default locale as UTF-8



```
sudo apt install locales
sudo dpkg-reconfigure locales
```

He escollit els locales UTF-8 i ISO-8859-1 i ISO-8859-15 per a català i castellà, posant com a locale per omissió a `es_ES.UTF-8`



```
user@datascience:~$ sudo dpkg-reconfigure locales
Generating locales (this might take a while)...
  ca_ES.ISO-8859-1... done
```

```
    ca_ES.UTF-8... done
    ca_ES.ISO-8859-15@euro... done
    es_ES.ISO-8859-1... done
    es_ES.UTF-8... done
    es_ES.ISO-8859-15@euro... done
    en_US.ISO-8859-1... done
    en_US.ISO-8859-15... done
    en_US.UTF-8... done
  Generation complete.
  user@datascience:~$
```

1.8. Shiny

Shiny apps are exposed at http://datascience.seeds4c.org:3838/[10]

Example shiny app in development: http://datascience.seeds4c.org:3838/climate-shelters/[11]

See also:

- Shiny Server
  - Installation: https://posit.co/download/shiny-server/[12]
  - Documentation: https://docs.posit.co/shiny-server/[13]
- Shiny Apps: https://shiny.posit.co/r/[14]

Alias names for this page:
datascience remote server 2023 | data science remote server | datascience remote server | datascienceremoteserver | 2023 datascience server

---

[1] https://posit.co/download/rstudio-server/

[2] https://docs.posit.co/previous-versions/rstudio/

[3] https://orcacore.com/install-openssl-3-ubuntu-20-04/

[4] https://adisarid.github.io/post/2020-03-06-setup_rstudio_server_with_ssl/

[5] https://letsencrypt.org/donate

[6] https://eff.org/donate-le

[7] https://datascience.seeds4c.org

[8] https://datascience.seeds4c.org

[9] https://wiki.x2go.org

[10] http://datascience.seeds4c.org:3838/

[11] http://datascience.seeds4c.org:3838/climate-shelters/

[12] https://posit.co/download/shiny-server/

[13] https://docs.posit.co/shiny-server/

[14] https://shiny.posit.co/r/