2020 Datascience remote server

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

1.1. How to use it

1.1.1. Connect through X2Go

intructions to come

1.1.2. 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.

1.1.3. Anaconda (and python)

Use Anaconda as usual

1.1.4. R & RStudio

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

You have R 3.6 installed.

1.2. How it has been developed

1.2.1. Operating System

Lubuntu GNU/Linux 18.04 (64 bits). Fetch iso from their website.
https://lubuntu.net/lubuntu-18-04-bionic-beaver-released/[1]
http://cdimage.ubuntu.com/lubuntu/releases/18.04/release/lubuntu-18.04-desktop-amd64.iso[2]

1.2.2. Enable Ubuntu Partners repository

First enable "partners" repos.

Contents of the updated /etc/apt/sources.list

```
## Uncomment the following two lines to add software from Canonical's
## 'partner' repository.
## This software is not part of Ubuntu, but is offered by Canonical and the
## respective vendors as a service to Ubuntu users.
deb http://archive.canonical.com/ubuntu bionic partner
deb-src http://archive.canonical.com/ubuntu bionic partner
```

1.2.3. Other repositories

```
sudo add-apt-repository -y ppa:nilarimogard/webupd8 # per a launchpadd-getkeys i
altres
sudo add-apt-repository -y ppa:utappia/stable # per a ucaresystem-core
sudo add-apt-repository -y ppa:webupd8team/java # per a java propietari (on calgui)
sudo add-apt-repository -y ppa:ubuntugis/ubuntugis-unstable # per a paquets
d'analisi geoespacial
sudo add-apt-repository -y 'deb https://cloud.r-project.org/bin/linux/ubuntu
bionic-cran35/' # main binary packages for R 3.5.x
sudo add-apt-repository -y ppa:marutter/c2d4u3.5 # extra binary packages for R
3.5.x from the usual marruter repo
```

```
# Add the key for the new repo for R 3.6.x from cloud.r-project.org
apt-key adv --keyserver hkp://keyserver.ubuntu.com:80 --recv-keys
E298A3A825C0D65DFD57CBB651716619E084DAB9 # marutter
apt-key adv --keyserver hkp://keyserver.ubuntu.com:80 --recv 089EBE08314DF160 #
ubuntugis-stable
```

Other general packages installed:

```
sudo apt install -y curl htop mc kupfer git cups-pdf parcellite
```

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

1.2.4. Add full lubuntu desktop

For demostration purposes, I'll add a full lubuntu desktop in this virtual machine

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

1.2.5. Allow GUI connections

With X2Go (https://wiki.x2go.org[3]) 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 lxde
sudo apt-get install x2golxdebindings
```

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

- `LXDE`
  or
- **Custom session**: `lxsession -e LXDE -s Lubuntu`

1.2.6. R 3.6.x

We add these repos to use the latest R versions released

Comandes i paquets lubuntu 18.04:

```
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 libv8-dev libx11-dev libxml2 libxml2-dev libxml2:i386
libxt-dev pandoc protobuf-compiler r-cran-cairodevice r-cran-devtools r-cran-
doparallel r-cran-geor r-cran-ggmap r-cran-ggplot2 r-cran-gstat r-cran-igraph r-
cran-leaflet r-cran-lme4 r-cran-mapdata r-cran-maps r-cran-misc3d r-cran-ncdf4 r-
cran-raster r-cran-rcolorbrewer r-cran-rgl r-cran-rglpk r-cran-rjags r-cran-rjava
r-cran-rmio r-cran-rmysql r-cran-roxygen2 r-cran-snow r-cran-sp r-cran-xlsx r-cran-
xml r-recommended subversion texlive-lang-spanish texlive-latex-extra texmaker tk-
dev tk-table unaccent xvfb libssh2-1-dev ucaresystem-core libudunits2-dev gigolo
filezilla openjdk-8-jre ibm-java80-jre
sudo R CMD javareconf
```

Paquets de CRAN: posar dins de la comanda:

```
library(pacman)
p_load("tidyverse", "caTools", "bitops", "httpuv", "devtools", "rpivotTable", "DT",
"shiny", "magick", "rvg", "addinslist", "ff", "sparklyr", "data.table", "rio",
"rpivotGadget", "radiant", "CRANsearcher", "regexplain", "rJava", "knitr",
"rmarkdown", "webshot", "magick", "rsvg", "sf", "leaflet", "htmlwidgets")
```

Rstudio Addins: CRANsearcher, addinslist, regexplain

Latex

- TexStudio

Paquets lubuntu 18.04: texstudio

1.2.7. RStudio

Server version, to use through browser at `http://localhost:8787` from inside the X2Go session. Or directly from your remote local computer opening a browser at: http://datascience.seeds4c.org:8787[4]

```
sudo apt-get install gdebi-core
wget
https://download2.rstudio.org/server/bionic/amd64/rstudio-server-1.3.959-amd64.deb
sudo gdebi rstudio-server-1.3.959-amd64.deb
```

1.2.8. Anaconda 3

Anaconda3

See:

- https://www.digitalocean.com/community/tutorials/how-to-install-anaconda-on-ubuntu-18-04-quickstart[5]
- https://www.digitalocean.com/community/tutorials/how-to-install-the-anaconda-python-distribution-on-ubuntu-18-04[6]
- https://www.anaconda.com/distribution/#linux[7]

1.2.8.1. Installation on /opt/py/anaconda3

We will install anaconda on a system folder as `/opt/py/`:

```
sudo mkdir /opt/py
sudo chmod 777 /opt/py
cd /tmp
curl -O https://repo.anaconda.com/archive/Anaconda3-2019.03-Linux-x86_64.sh
bash Anaconda3-2019.03-Linux-x86_64.sh
```

```
datascience@datasciencepc:/tmp$ bash Anaconda3-2019.03-Linux-x86_64.sh

Welcome to Anaconda3 2019.03

In order to continue the installation process, please review the license
agreement.
Please, press ENTER to continue
>>>
==================================
Anaconda End User License Agreement
==================================

Copyright 2015, Anaconda, Inc.

All rights reserved under the 3-clause BSD License:

Redistribution and use in source and binary forms, with or without modification,
are permitted provided that the fo
llowing conditions are met:
```

```
  * Redistributions of source code must retain the above copyright notice, this
list of conditions and the followin
g disclaimer.
  * Redistributions in binary form must reproduce the above copyright notice, this
list of conditions and the follo
wing disclaimer in the documentation and/or other materials provided with the
distribution.
  * Neither the name of Anaconda, Inc. ("Anaconda, Inc.") nor the names of its
contributors may be used to endorse
or promote products derived from this software without specific prior written
permission.

THIS SOFTWARE IS PROVIDED BY THE COPYRIGHT HOLDERS AND CONTRIBUTORS "AS IS" AND ANY
EXPRESS OR IMPLIED WARRANTIES,
INCLUDING, BUT NOT LIMITED TO, THE IMPLIED WARRANTIES OF MERCHANTABILITY AND
FITNESS FOR A PARTICULAR PURPOSE ARE D
ISCLAIMED. IN NO EVENT SHALL ANACONDA, INC. BE LIABLE FOR ANY DIRECT, INDIRECT,
INCIDENTAL, SPECIAL, EXEMPLARY, OR
CONSEQUENTIAL DAMAGES (INCLUDING, BUT NOT LIMITED TO, PROCUREMENT OF SUBSTITUTE
GOODS OR SERVICES; LOSS OF USE, DAT
A, OR PROFITS; OR BUSINESS INTERRUPTION) HOWEVER CAUSED AND ON ANY THEORY OF
LIABILITY, WHETHER IN CONTRACT, STRICT
 LIABILITY, OR TORT (INCLUDING NEGLIGENCE OR OTHERWISE) ARISING IN ANY WAY OUT OF
THE USE OF THIS SOFTWARE, EVEN IF
 ADVISED OF THE POSSIBILITY OF SUCH DAMAGE.

Notice of Third Party Software Licenses
=======================================

Anaconda Distribution contains open source software packages from third parties.
These are available on an "as is"
basis and subject to their individual license agreements. These licenses are
available in Anaconda Distribution or
at http://docs.anaconda.com/anaconda/pkg-docs. Any binary packages of these third
party tools you obtain via Anacon
da Distribution are subject to their individual licenses as well as the Anaconda
license. Anaconda, Inc. reserves t
he right to change which third party tools are provided in Anaconda Distribution.

In particular, Anaconda Distribution contains re-distributable, run-time, shared-
library files from the Intel(TM) M
ath Kernel Library ("MKL binaries"). You are specifically authorized to use the MKL
binaries with your installation
 of Anaconda Distribution. You are also authorized to redistribute the MKL binaries
with Anaconda Distribution or i
n the conda package that contains them. Use and redistribution of the MKL binaries
are subject to the licensing ter
ms located at
https://software.intel.com/en-us/license/intel-simplified-software-license. If
```

needed, instructions f
or removing the MKL binaries after installation of Anaconda Distribution are
available at http://www.anaconda.com.

Anaconda Distribution also contains cuDNN software binaries from NVIDIA Corporation
("cuDNN binaries"). You are spe
cifically authorized to use the cuDNN binaries with your installation of Anaconda
Distribution. You are also author
ized to redistribute the cuDNN binaries with an Anaconda Distribution package that
contains them. If needed, instru
ctions for removing the cuDNN binaries after installation of Anaconda Distribution
are available at http://www.anac
onda.com.

Anaconda Distribution also contains Visual Studio Code software binaries from
Microsoft Corporation ("VS Code"). Yo
u are specifically authorized to use VS Code with your installation of Anaconda
Distribution. Use of VS Code is sub
ject to the licensing terms located at https://code.visualstudio.com/License.

Cryptography Notice
==========

This distribution includes cryptographic software. The country in which you
currently reside may have restrictions
on the import, possession, use, and/or re-export to another country, of encryption
software. BEFORE using any encry
ption software, please check your country's laws, regulations and policies
concerning the import, possession, or us
e, and re-export of encryption software, to see if this is permitted. See the
Wassenaar Arrangement http://www.wass
enaar.org/ for more information.

Anaconda, Inc. has self-classified this software as Export Commodity Control Number
(ECCN) 5D992b, which includes m
ass market information security software using or performing cryptographic
functions with asymmetric algorithms. No
 license is required for export of this software to non-embargoed countries. In
addition, the Intel(TM) Math Kernel
 Library contained in Anaconda, Inc.'s software is classified by Intel(TM) as ECCN
5D992b with no license required
for export to non-embargoed countries and Microsoft's Visual Studio Code software
is classified by Microsoft as ECC
N 5D992.c with no license required for export to non-embargoed countries.

The following packages are included in this distribution that relate to
cryptography:

```
openssl
    The OpenSSL Project is a collaborative effort to develop a robust, commercial-
grade, full-featured, and Open So
urce toolkit implementing the Transport Layer Security (TLS) and Secure Sockets
Layer (SSL) protocols as well as a
full-strength general purpose cryptography library.

pycrypto
    A collection of both secure hash functions (such as SHA256 and RIPEMD160), and
various encryption algorithms (A
ES, DES, RSA, ElGamal, etc.).

pyopenssl
    A thin Python wrapper around (a subset of) the OpenSSL library.

kerberos (krb5, non-Windows platforms)
    A network authentication protocol designed to provide strong authentication for
client/server applications by u
sing secret-key cryptography.

cryptography
    A Python library which exposes cryptographic recipes and primitives.


Do you accept the license terms? [yes|no]
[no] >>> yes

Anaconda3 will now be installed into this location:
/home/gid/anaconda3

  - Press ENTER to confirm the location
  - Press CTRL-C to abort the installation
  - Or specify a different location below

[/home/datascience/anaconda3] >>> /opt/py/anaconda3
...
```

1.2.9. Python

Paquets lubuntu 18.04:

```
sudo apt install -y python-numpy python-pandas python-matplotlib python-seaborn
ipython-notebook ipython-doc
```

Alias names for this page:
datascience remote server 2020 | 2020 datascience server

---

[1] https://lubuntu.net/lubuntu-18-04-bionic-beaver-released/

[2] http://cdimage.ubuntu.com/lubuntu/releases/18.04/release/lubuntu-18.04-desktop-amd64.iso

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

[4] http://datascience.seeds4c.org:8787

[5] https://www.digitalocean.com/community/tutorials/how-to-install-anaconda-on-ubuntu-18-04-quickstart

[6] https://www.digitalocean.com/community/tutorials/how-to-install-the-anaconda-python-distribution-on-ubuntu-18-04

[7] https://www.anaconda.com/distribution/#linux