Centre for Affordable Housing Finance in Africa

fsd Kenya

71point4 NUMBERS TO LIFE

A HOME FOR HUMAN POTENTIAL Reall

KNBS KENYA NATIONAL BUREAU OF STATISTICS *Keeping you informed*

KENYA AFFORDABLE HOUSING DATA PROJECT

Server Setup

Agenda

1) Server specifications
2) Software installation

Server specifications

Server specifications

> **Setup was tested on an AWS machine with the following specifications:**

- AMI: Ubuntu Server 20.04 LTS (HVM), SSD Volume Type [ami-0fb653ca2d3203ac1 (64-bit x86)]
- AWS server type: t2.2xlarge
  - CPUs: 8
  - Memory: 32GB
- OS: Ubuntu 20.04 with Linux kernel 5.11
- users and privileges
  - this slide deck assumes a user called `knbs` exists and that this user has `sudo` rights -

```
sudo adduser knbs
sudo usermod -aG sudo knbs
su knbs
cd ~
```

Centre for Affordable Housing Finance in Africa

fsd Kenya

71point4 NUMBERS TO LIFE

Reall A HOME FOR HUMAN POTENTIAL

KNBS KENYA NATIONAL BUREAU OF STATISTICS *Keeping you informed*

Preamble

> **The following software is required for the pipeline:**

- MariaDB - opensource version of MySQL database management system (used for data curation)
- R and RStudio - software for ETL (preparing data and populating indicators)

The slides below provide instructions for installing this software on the server and setting up users.

> **Everything is done via the command line**

MariaDB installation

MariaDB installation

- The code chunk below will install `mariadb` on the server and take you through a series of prompts where you can make some changes to your MariaDB installation's security options

```
sudo apt update
sudo apt install mariadb-server
sudo mysql_secure_installation
```

- The first prompt will ask you to enter the current database root password. Since you have not set one up yet, press ENTER to indicate "none"

```
# NOTE: RUNNING ALL PARTS OF THIS SCRIPT IS RECOMMENDED FOR ALL MariaDB
#       SERVERS IN PRODUCTION USE!  PLEASE READ EACH STEP CAREFULLY!
#
# In order to log into MariaDB to secure it, we'll need the current
# password for the root user.  If you've just installed MariaDB, and
# you haven't set the root password yet, the password will be blank,
# so you should just press enter here.
#
# Enter current password for root (enter for none):
```

MariaDB installation

- Be sure to enter "n". You can always add a password to the root user later for extra security.

```
# OK, successfully used password, moving on ...
#
# Setting the root password ensures that nobody can log into the MariaDB
# root user without the proper authorisation.
#
# Set root password? [Y/n] N
```

- For the rest, you can say "Y"

```
# By default, a MariaDB installation has an anonymous user, allowing anyone
# to log into MariaDB without having to have a user account created for
# them.  This is intended only for testing, and to make the installation
# go a bit smoother.  You should remove them before moving into a
# production environment.
#
# Remove anonymous users? [Y/n] Y
#  ... Success!
```

MariaDB installation

- For the rest, you can say "Y"

```
#
# Normally, root should only be allowed to connect from 'localhost'.  This
# ensures that someone cannot guess at the root password from the network.
#
# Disallow root login remotely? [Y/n] Y
#  ... Success!
#
# By default, MariaDB comes with a database named 'test' that anyone can
# access.  This is also intended only for testing, and should be removed
# before moving into a production environment.
#
# Remove test database and access to it? [Y/n] Y
#  - Dropping test database ...
#  ... Success!
#  - Removing privileges on test database ...
#  ... Success!
```

MariaDB installation

- For the rest, you can say "Y"

```
#
# Reloading the privilege tables will ensure that all changes made so far
# will take effect immediately.
#
# Reload privilege tables now? [Y/n] Y
#  ... Success!
#
# Cleaning up...
#
# All done!  If you've completed all of the above steps, your MariaDB
# installation should now be secure.
#
# Thanks for using MariaDB!
```

MariaDB installation

- Testing MariaDB - the code below should show that MariaDB is running

```
sudo systemctl status mariadb

# ● mariadb.service - MariaDB 10.3.34 database server
#      Loaded: loaded (/lib/systemd/system/mariadb.service; enabled; vendor preset: enabled)
#      Active: active (running) since Thu 2022-04-07 09:17:42 UTC; 1min 11s ago
#        Docs: man:mysqld(8)
#              https://mariadb.com/kb/en/library/systemd/
#    Main PID: 2450 (mysqld)
#      Status: "Taking your SQL requests now..."
#       Tasks: 31 (limit: 38511)
#      Memory: 63.7M
#      CGroup: /system.slice/mariadb.service
#              └─2450 /usr/sbin/mysqld
```

Setup MySQL

- Next we need to create a user and profile for MySQL
- This is done via the `~/.my.cnf` configuration file
    - We use `vim` for editing text files, but you could use another text editor (e.g. `nano`) if you prefer

```
vim /home/knbs/.my.cnf
```

Input the following in the file and save it - this creates a user called `knbs` with a password (change from default provided below) and chooses `less` as the terminal pager for query results.

```
[client]
user = knbs
password = knbs_cahf_fsdk_reall_71point4_housing # put a different password of your choosing here
pager = less -S
```

Setup MySQL

- We now need to create this user in `mysql` too
- First launch `mysql` as root (recall we didn't set a password for root)

```
sudo mysql -u root
```

- Then create the `knbs` user - this user will have all privileges on the system

```
CREATE USER 'knbs'@localhost IDENTIFIED BY 'knbs_fsdk_reall_71point4_housing';
GRANT ALL PRIVILEGES ON *.* TO 'knbs'@localhost IDENTIFIED BY 'knbs_cahf_fsdk_reall_71point4_housing';
FLUSH PRIVILEGES;
```

- At this point you should be able to login in to `mysql` as the `knbs` user:

```
mysql
mysql -u knbs -p
```

Centre for Affordable Housing Finance in Africa

fsd Kenya

71point4 NUMBERS TO LIFE

Reall A HOME FOR HUMAN POTENTIAL

KNBS KENYA NATIONAL BUREAU OF STATISTICS Keeping you informed

R-base installation

R-base installation

- We first install base R using the code below
    - Note that for ubuntu 20.04 LTS the repository is `focal_cran40`, for Ubuntu 18 it is `bionic40`

```
sudo apt-key adv --keyserver keyserver.ubuntu.com --recv-keys E298A3A825C0D65DFD57CBB651716619E084DAB9
sudo add-apt-repository 'deb https://cloud.r-project.org/bin/linux/ubuntu focal-cran40/'
sudo apt update
sudo apt install -y r-base
```

- Launch R to see if installation was successfull by typing `R` in the terminal

```
R
#
# R version 4.1.3 (2022-03-10) -- "One Push-Up"
# Copyright (C) 2022 The R Foundation for Statistical Computing
# Platform: x86_64-pc-linux-gnu (64-bit)
```

RStudio server installation

RStudio server installation

- Instructions are available on Rstudio's website

```
sudo apt-get install gdebi-core
mkdir downloads && cd downloads
wget https://download2.rstudio.org/server/bionic/amd64/rstudio-server-2021.09.1-372-amd64.deb
sudo gdebi rstudio-server-2021.09.1-372-amd64.deb
```

- The installation should end showing that `rstudio-server` is up and running

```
# ● rstudio-server.service - RStudio Server
#      Loaded: loaded (/lib/systemd/system/rstudio-server.service; enabled; vendor preset: enabled)
#      Active: active (running) since Thu 2022-04-07 09:59:01 UTC; 1s ago
#     Process: 13927 ExecStart=/usr/lib/rstudio-server/bin/rserver (code=exited, status=0/SUCCESS)
#    Main PID: 13928 (rserver)
#       Tasks: 3 (limit: 38511)
#      Memory: 2.5M
#      CGroup: /system.slice/rstudio-server.service
#              └─13928 /usr/lib/rstudio-server/bin/rserver
```

RStudio server installation

- Next we need to configure `rstudio-server` to accept connections from specific ports. See this guide for more information
  - Note the port can be changed, default is just `8787`
  - By default `rstudio-server` binds to `0.0.0.0` this can also be changed
  - Changes only take effect after restarting the service: `sudo rstudio-server restart`
- These configurations are made in `/etc/rstudio/rserver.conf`

```
sudo vim /etc/rstudio/rserver.conf
```

- Input the following into the file:

```
www-port=8787
```

RStudio server installation

- A few useful dependencies to install include:

```
sudo apt-get install -y default-jre
sudo apt-get install -y default-jdk
sudo R CMD javareconf

sudo apt install libcurl4-openssl-dev
sudo apt install libxml2-dev
sudo apt install libssl-dev
sudo apt install libmariadbclient-dev
sudo apt install libpq-dev
sudo apt install libsodium-dev
```

RStudio server installation

- You should be able to login to RStudio by going to the ipaddress of the server with port 8787
  - e.g. `197.85.3.47:8787`
  - Login using the `knbs` user credentials (the ubuntu user name and password, not the `mysql` password as per `/home/knbs/.my.cnf`)
- Then install the following additional packages through the RStudio console

```
install.packages("rJava", Ncpus = 4)
install.packages("tidyverse", Ncpus = 4)
install.packages("RMySQL", Ncpus = 4)
install.packages("RPostgres", Ncpus = 4)
install.packages("glue", Ncpus = 4)
install.packages("plumber", Ncpus = 4)
install.packages("vroom", Ncpus = 4)
install.packages("renv", Ncpus = 4)
install.packages("lubridate", Ncpus = 4)
```

Centre for Affordable Housing Finance in Africa

fsd Kenya

71point4 NUMBERS TO LIFE

A HOME FOR HUMAN POTENTIAL Reall

KNBS KENYA NATIONAL BUREAU OF STATISTICS *Keeping you informed*

Fin