

Comparing health facility location data available for Saint Louis, Senegal

Andy South @afrimapr, Mark Herringer @sharehealthdata, Lamine Ndiaye @lamineyasey

DRAFT 2022-03-17

This is a data-driven document [open-source code here](#)

Funded by COVIDaction Vaccine Data Co-Lab - [Healthsites.io](#) Emergency health data validation.

Initial comparison of data provided by COUS [Centre des Opérations d'Urgence Sanitaire](#) & DPRS [Direction de la Planification, de la Recherche et des Statistiques](#) for Saint Louis with OSM Senegal survey conducted in February 2022.

Summary The aim is to see if we can develop re-useable code to help with the process of improving health facility location data by comparing and combining data from different sources.

Prior to an in person workshop COUS and DPRS provided data.

COUS provided data for 9 facilities in Saint Louis in a table with **9 rows** and **130 columns** including coordinates and attributes such as the numbers of beds and staff.

DPRS provided the names of facilities and regions for Saint Louis in a file with **374 rows** and **1 column**. Due to the format of the data, it is not straightforward automatically to compare these with the COUS data.

OSM Senegal surveyed facilities in person in Saint Louis, providing a file with **394 rows** and recording attribute data in **289 columns**.

source	num facilities	columns of info	coords?
1. COUS	9	130	coords
3. DPRS	<374	1	no
2. OSM-Senegal survey	394	289	coords

Agreement on a strict naming convention for facilities would make it much easier to compare and combine these data.

COUS data

COUS provided an Excel file with two sheets.

COUS Sheet2 for Saint Louis has nine facilities **9 rows** and **130 columns** including coordinates and attributes such as the numbers of beds and staff.

Names of facilities.

cous_names
EPS1 de Richard Toll
Centre de Sante de Dagana
Centre de Sante de Saint Louis
Centre de Sante Amadou Malick Gaye
Centre de Sante de Pete
Centre de santé de Richard Toll
HOPITAL REGIONAL DE SAINT LOUIS
CHR NDioum
Centre de santé de Ross Bethio

DPRS data

DPRS provided the names of facilities and regions for the whole country in a file with **374 rows and 1 column**.

Looking at the first 9 rows of the DPRS data.

dprs_names
[-] RM Saint-Louis
[-] Dagana
CS Dagana
DPC de Dagana
PM MAC Dagana
PS 4e Et 5e Secteur
[-] PS Bokhol
Case de Geum Yalla
PS Bouteyni

We can see that sometimes rows indicate regions, e.g. [-] RM Saint-Louis and [-] Dagana. Also sometimes a row with a [-] indicates a facility that has control over others e.g. [-] PS Bokhol. This makes it difficult to analyse the data automatically because it is difficult to detect whether a row refers to a facility or not. It may be that if a row has a [-] and does not have PS, CS or Case then it is not a facility.

We can compare the COUS & DPRS tables above to check whether the facility names in the COUS data appear exactly the same as in the DPRS data. We would need to be able to write code to find matching names to be able to write a reproducible pipeline for comparing data.

An initial search by eye shows the mismatches between naming in the two datasets. There is at least one facility in the COUS data that seems not present in the DPRS list (Centre de Sante Amadou Malick Gaye).

See section below for Comparison of facility naming between DPRS and the new OSM-Senegal survey.

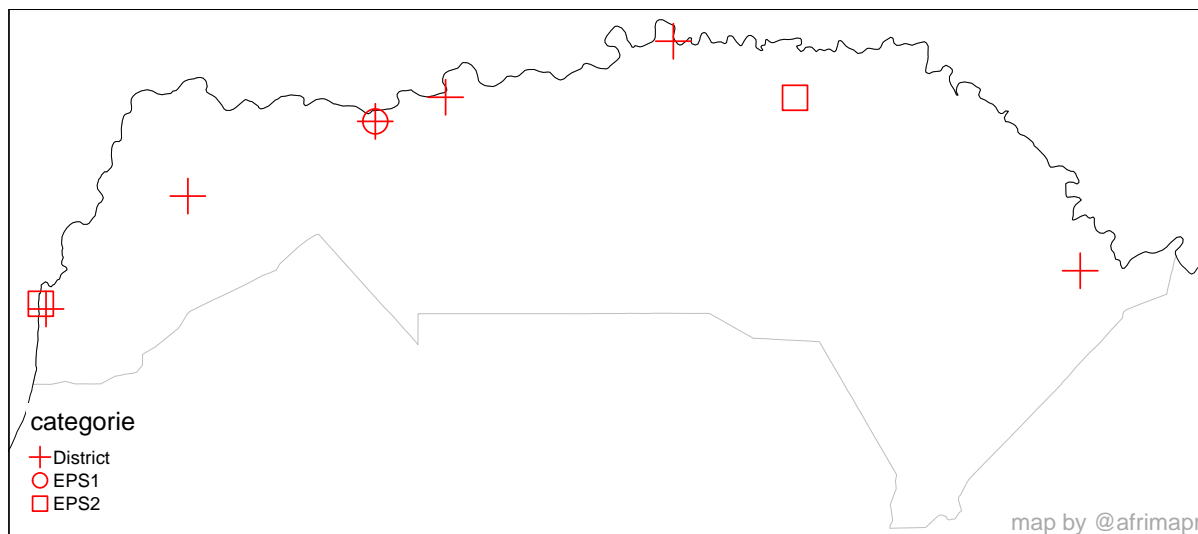
cous_names	dprs_names
EPS1 de Richard Toll	EPS 1 Richard Toll
Centre de Sante de Dagana	CS Dagana
Centre de Sante de Saint Louis	CS Saint - Louis
Centre de Sante Amadou Malick Gaye	?
Centre de Sante de Pete	[-] CS Pete
Centre de santé de Richard Toll	CS Richard Toll
HOPITAL REGIONAL DE SAINT LOUIS	EPS2 de Saint Louis
CHR NDioum	?Antenne de Ndioum

cous_names	dprs_names
Centre de santé de Ross Bethio	[-] CS Ross Bethio

Categories of COUS health facilities Saint Louis

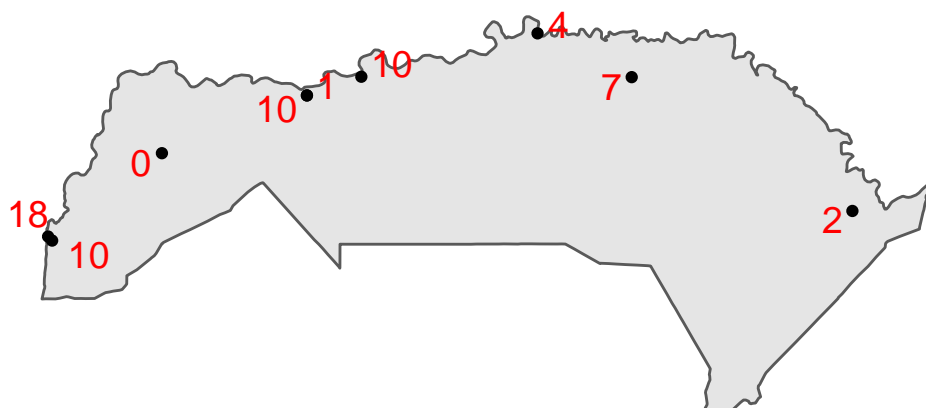
The coordinates provided in the COUS data allow these to be plotted on maps. (Note that one coordinate was missing a negative sign that incorrectly placed it in Chad, it was corrected for the maps here).

COUS health facilities Saint Louis 2022

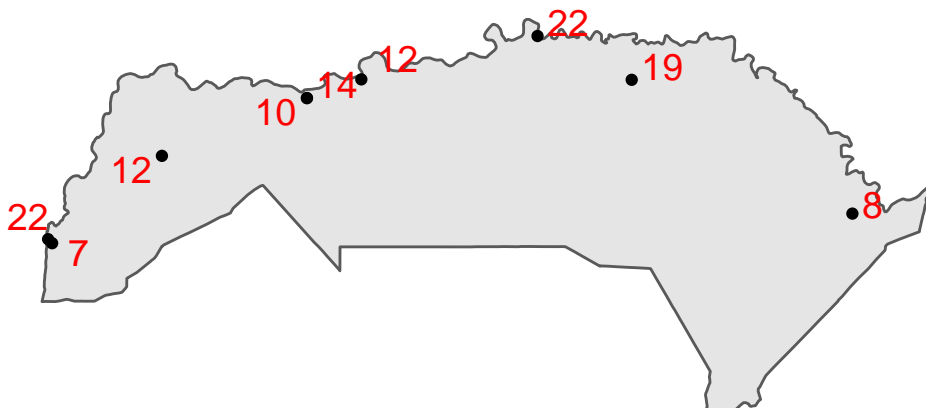


Numbers of beds (lits) in different wards from COUS data

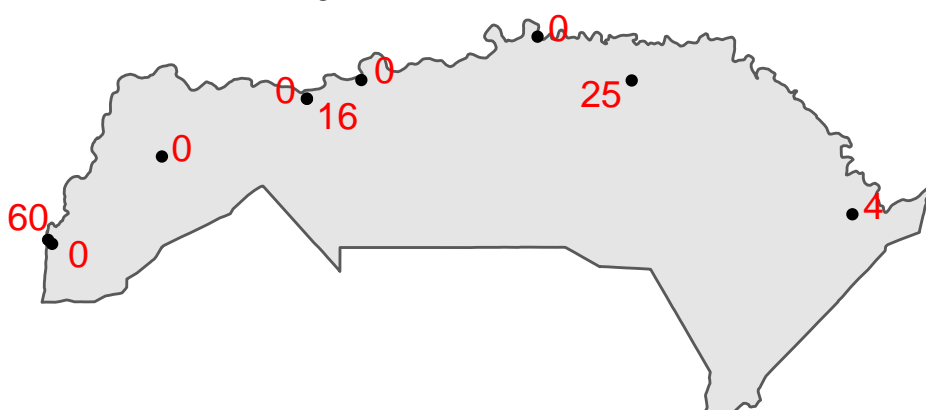
COUS lits_en_sau



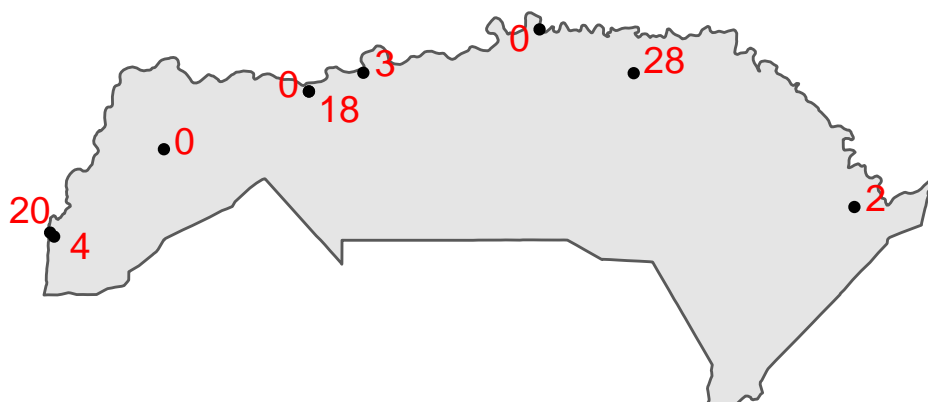
COUS lits_en_medecine



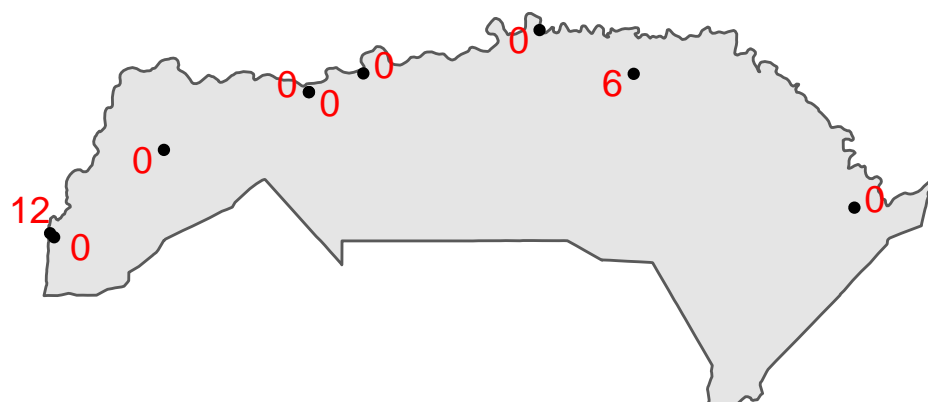
COUS lits_en_chirurgie



COUS lits_en_pediatrie



COUS lits_en_reanimation



OSM-Senegal survey

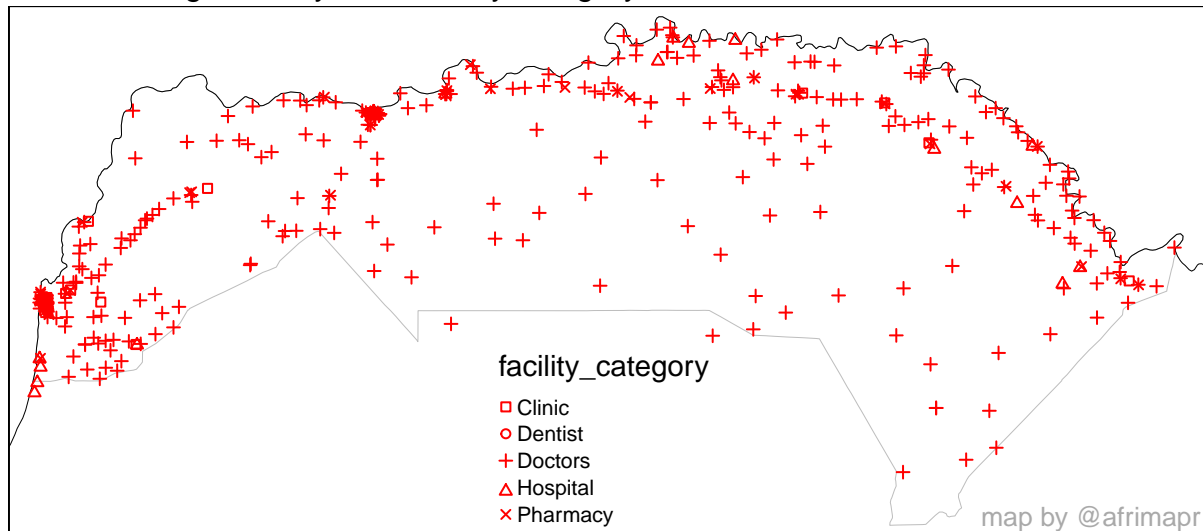
OSM-Senegal surveyed facilities in person in Saint Louis in February 2022, providing a file with 394 rows and recording attribute data in 289 columns. Here are some maps summarising the facilities surveyed.

Categories of facilities in OSM-Senegal survey

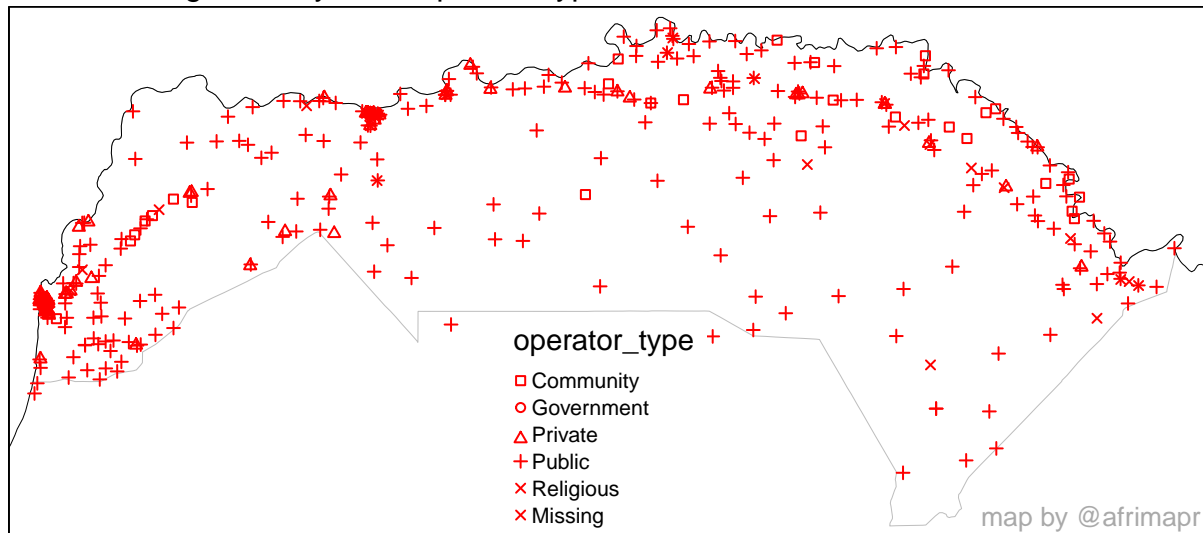
Senegal health facility levels are converted into OSM categories of Hospital, Clinic, Doctor as follows :

Senegal level	Senegal type	Senegal abbreviation	OSM category
1	Hopital	HP	Hospital
1	Etablissement public de Santé 1 (départementale)	EPS 1	Hospital
1	Etablissement public de Santé 2 (régionale)	EPS 2	Hospital
1	Etablissement public de Santé 3 (nationale)	EPS 3	Hospital
2	Centre de Santé	CS	Clinic
2	District de Sanitaire	DS	Clinic
3	Poste de Santé	PS	Doctors
3	Case de Santé	Case de	Doctors

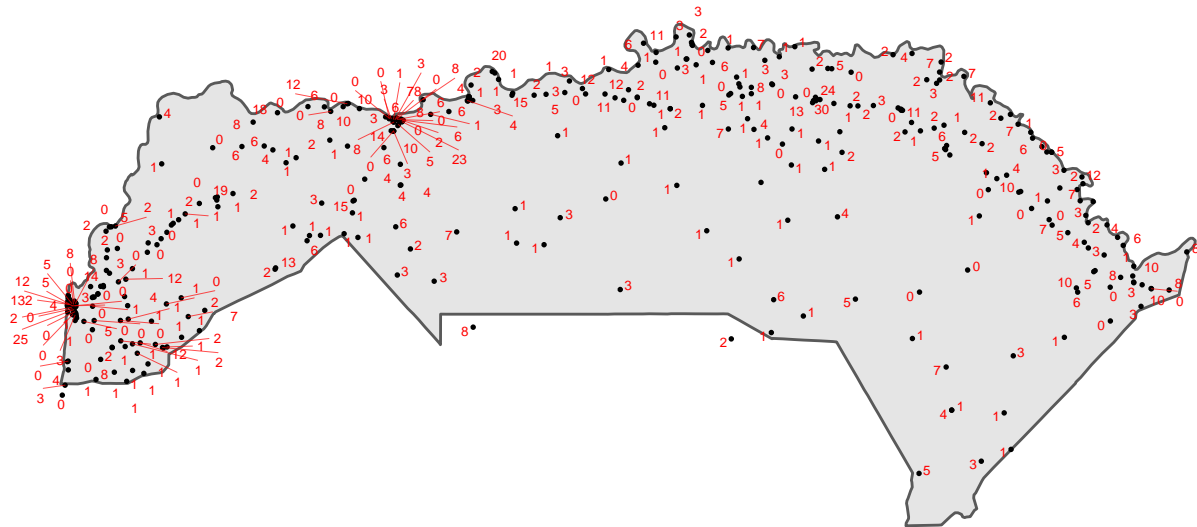
OSM–Senegal survey 2022 facility category



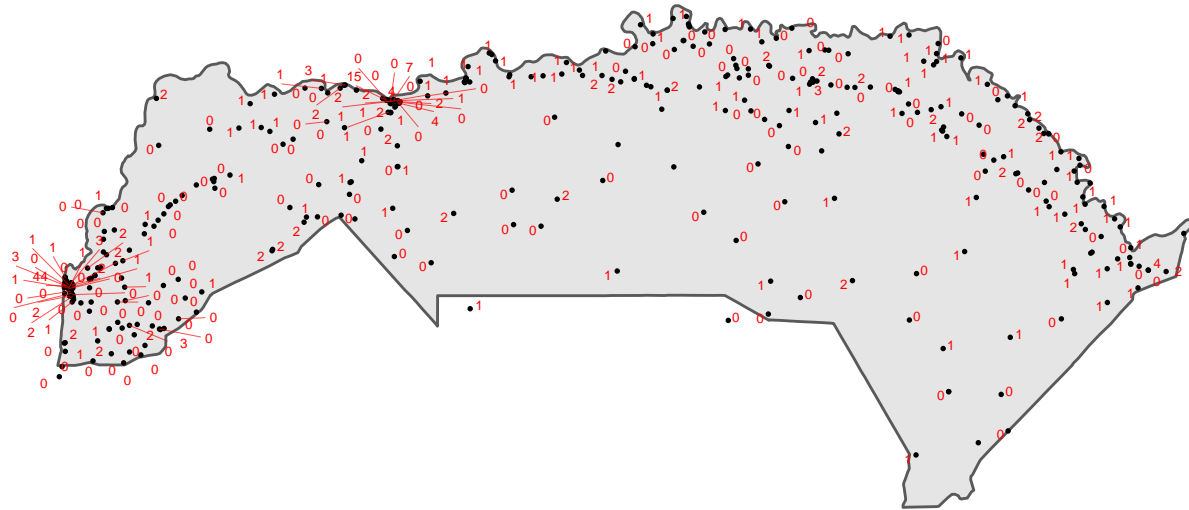
OSM–Senegal survey 2022 operator type



OSM number_of_beds



OSM number_of_nurses



Comparison of facility naming between DPRS and the new OSM-Senegal survey

Can we automatically match facility names in the OSM-Senegal survey data with their equivalent names used by DPRS ?

The DPRS data mostly uses abbreviations for facility categories.

Senegal level	OSM naming	DPRS abbreviation
1	Hopital	HP
1	Établissement public de santé	EPS
2	Centre de Santé	CS
2	District de Sanitaire	DS
3	Poste de Santé	PS
3	Case de Santé	Case de

TODO what is PM ?

There are also differences in upper/lower case, accents, use of 'de' & spaces.

We can try the following to match the names

1. expanding the abbreviations from DPRS
2. remove accents
3. convert to lower case
4. ensure just single spaces
5. remove 'de'