Package 'tidyndr'

April 8, 2022

Title Analysis of the Nigeria National Data Repository (NDR)

Version 0.2.1

Description
The goal is to simplify routine analysis of the Nigeria National Data Repository (NDR) <https://ndr.phis3project.org.ng> using the PEPFAR Monitoring, Evaluation, and Reporting (MER) indicators (see <https://datim.zendesk.com/hc/en-us/articles/360000084446-MER-Indicator-Reference-Guides>). It is designed to import in to R patient-level line-list downloaded as 'csv' file from the front-end of the NDR.

License MIT + file LICENSE

Encoding UTF-8

LazyData true

RoxygenNote 7.1.2

Depends R (>= 3.6)

Imports dplyr (>= 1.0.3), forcats, janitor (>= 2.1.0), lubridate (>= 1.7.9.2), magrittr (>= 2.0.1), purrr (>= 0.3.4), rlang (>= 0.4.10), stringr, stats, tibble, tidyr, vroom (>= 1.3.2), tidyselect

Suggests testthat (>= 3.0.0), knitr, rmarkdown, spelling

Config/testthat/edition 3

Config/rcmdcheck/ignore-inconsequential-notes true

URL https://github.com/stephenbalogun/tidyndr

BugReports https://github.com/stephenbalogun/tidyndr/issues

VignetteBuilder knitr

Language en-US

NeedsCompilation no

Author Stephen Balogun [aut, cre] (<https://orcid.org/0000-0002-9928-3703>),
Scholastica Olanrewaju [ctb],
Oluwaseun Okunuga [ctb],
Temitope Kolade [ctb],
Geraldine Abone [ctb]

Maintainer Stephen Balogun <stephentaiyebalogun@gmail.com>

Repository CRAN

Date/Publication 2022-04-08 06:40:02 UTC

R topics documented:

`cot_cascade` *Analyse the Treatment Continuity Cascade*

Description

Generate aggregate summary of treatment continuity indicators based on a specified quarter. The indicators include `tx_curr` (previous quarter), `tx_new`, `tx_ml`, `tx_ml_dead`, `tx_ml_to`, `tx_ml_iit`, and `iit_rate`.

Usage

```
cot_cascade(
  data,
  quarter = NULL,
  ref = NULL,
  states = NULL,
  facilities = NULL,
  status = "default",
  remove_duplicates = FALSE,
  .level = "state",
  .names = NULL
)
```

Arguments

data	An NDR dataframe imported using the `read_ndr()`.
quarter	the quarter of the fiscal year for which the treatment continuity indicators should be calculated.
ref	the referenced date for the analysis. If this is not set (i.e. `NULL`) it will be assumed to be the last day of the quarter.
states	The name(s) of the State(s) of interest. The default utilizes all the states in the dataframe. If specifying more than one state, combine the states using the `c()` e.g. `c("State 1","State 2")`.
facilities	The name(s) of the facilit(ies) of interest. Default is to utilize all the facilities contained in the dataframe. If specifying more than one facility, combine the facilities using the `c()` e.g. `c("Facility 1","Facility 2")`.
status	Determines how the number of active clients is calculated. The options are to either to use the NDR current_status_28_days column or the derived current_status column ("calculated").
remove_duplicates	Boolean argument. It specifies if duplicate patient entries in the facilities should be removed or kept
.level	the level at which the aggregate summary should be performed. The options are "ip", "country", "state", "lga" and "facility".
.names	if specified, these will be used for naming of the viral load indicators instead of the default.

Value

summary of treatment continuity indicators

Examples

```
cot_cascade(ndr_example, quarter = 2, ref = "2022-02-28", .level = "state")

# Determine the treatment continuity cascade for a state at the end of quarter one of FY22
cot_cascade(ndr_example,
```

```
  quarter = 1,
  states = "Arewa"
)
```

`disaggregate` *Summarize an indicator into finer details by the specified variable*

Description

Counts the number of occurrence of an outcome based on the category of interest. It also provides "Totals" at the end of the columns or rows (or both) where appropriate.

Usage

```
disaggregate(data, by, ...)
```

Arguments

`data`	Data containing the indicator to be disaggregated.
`by`	The variable of interest to be used for the disaggregation. The options are any of: `"sex"`, `"current_age"`, `"pregnancy_status"`, `"art_duration"`, `"months_dispensed"` (of ARV), and `"age_sex"`.
`...`	Other parameters to be passed based on the by supplied. These are: 1) the `` `level` `` of disaggregation, ("ip" or "country", "state", "lga" and "facility"). 2) `` `pivot_wide` `` TRUE or FALSE, used to determine how the tabular data should be presented. The default options are "state" for by and TRUE for `pivot_wide`. value is "state".

Value

disaggregated data

Examples

```
### Disaggregate "TX_NEW" clients into age categories for each state
new_clients <- tx_new(ndr_example, from = "2021-01-01")
disaggregate(new_clients, by = "current_age") # default value of level is "state"

### Disaggregate "TX_CURR" by gender for each facility
curr_clients <- tx_curr(ndr_example)
disaggregate(curr_clients, by = "sex", level = "facility")
```

`hts_recent` *Subset Clients who had Recent Infection Test Done during a period of Interest*

Description

Subset Clients who had Recent Infection Test Done during a period of Interest

Usage

```
hts_recent(data, from = NULL, to = NULL, states = NULL, facilities = NULL)
```

Arguments

`data`	An NDR dataframe imported using the `read_ndr()`.
`from`	The start date in ISO8601 format (i.e. "yyyy-mm-dd"). The default is to start at the beginning of the current Fiscal Year (i.e. 1st of October).
`to`	The end date written in ISO8601 format (i.e. "yyyy-mm-dd"). The default is the date of analysis.
`states`	The name(s) of the State(s) of interest. The default utilizes all the states in the dataframe. If specifying more than one state, combine the states using the `c()` e.g. `c("State 1","State 2")`.
`facilities`	The name(s) of the facilit(ies) of interest. Default is to utilize all the facilities contained in the dataframe. If specifying more than one facility, combine the facilities using the `c()` e.g. `c("Facility 1","Facility 2")`.

Value

line-list of clients who had recent HIV infection test done during the period of interest

Examples

```
### Line-list of clients offered recency testing
hts_recent_clients <- hts_recent(recency_example, from = "2021-07-01", to = "2021-10-31")

hts_pos <- hts_tst_pos(recency_example) # identifies all hts_positive clients
# line-list positive clients offered recency testing in 'Okun' state
hts_recent(hts_pos, state = "Okun")
```

hts_tst_pos *Subset Newly Identified HIV Positive Clients over a Particular Period*

Description

Subset Newly Identified HIV Positive Clients over a Particular Period

Usage

```
hts_tst_pos(data, from = NULL, to = NULL, states = NULL, facilities = NULL)
```

Arguments

data	An NDR dataframe imported using the read_ndr().
from	The start date in ISO8601 format (i.e. "yyyy-mm-dd"). The default is to start at the beginning of the current Fiscal Year (i.e. 1st of October).
to	The end date written in ISO8601 format (i.e. "yyyy-mm-dd"). The default is the date of analysis.
states	The name(s) of the State(s) of interest. The default utilizes all the states in the dataframe. If specifying more than one state, combine the states using the c() e.g. c("State 1","State 2").
facilities	The name(s) of the facilit(ies) of interest. Default is to utilize all the facilities contained in the dataframe. If specifying more than one facility, combine the facilities using the c() e.g. c("Facility 1","Facility 2").

Value

line-list of hts_tst_pos clients

Examples

```
### Line-list of clients hts positives from 'Okun' and 'Abaji' states in first half of 2021
hts_pos <- hts_tst_pos(
  recency_example,
  state = c("Okun", "Abaji"),
  from = "2021-01-01",
  to = "2021-06-30"
)
```

ndr_example *Line-list of 50,000 Simulated Clients Provided in the NDR Format.*

Description

A dataset containing clients' details that were randomly generated to simulate the NDR patient line-list downloaded from the front-end of the Nigeria National Data Repository. Three additional variables are added for easy referencing during analysis. These variables are: `date_ltfu`, `appointment_date` and `current_status`.

Usage

```
ndr_example
```

Format

A data frame with 50000 rows and 52 variables:

ip Implementing Partner

state State of registration of client

lga Local Government Area where client was registered

facility Facility where the client was registered

datim_code 'DATIM CODE' of the facility of registration

sex The gender that the client identified as, "M" or "F"

patient_identifier Patient unique id

hospital_number assigned hospital reference number

date_of_birth Birth day of client, in "yyyy-dd-mm"

age_at_art_initiation Age of client at commencement of ART

current_age Age of client as at when the dataset was generated

art_start_date Date of commencement of ART

art_start_date_source The reference for the art_start_date

last_drug_pickup_date Date of last medication refill

last_drug_pickup_date_q1 Date of last medication refill in Q1

last_drug_pickup_date_q2 Date of last medication refill in Q2

last_drug_pickup_date_q3 Date of last medication refill in Q3

last_drug_pickup_date_q4 Date of last medication refill in Q4

last_regimen The combination regimen dispensed during last medication refill

last_clinic_visit_date Date of last hospital visit (for any reason)

days_of_arv_refill Number of days of medications dispensed

pregnancy_status Pregnancy status of client entered as "P", "NP" or "BF"

current_viral_load Value of the most recent viral load result of client

date_of_current_viral_load Date of most recent viral load result

current_viral_load_q1 Viral load result of client at end of Q1

date_of_current_viral_load_q1 Date of last viral load result in Q1

current_viral_load_q2 Viral load result of client at end of Q2

date_of_current_viral_load_q2 Date of last viral load result in Q2

current_viral_load_q3 Viral load result of client at end of Q3

date_of_current_viral_load_q3 Date of last viral load result in Q3

current_viral_load_q4 Viral load result of client at end of Q4

date_of_current_viral_load_q4 Date of last viral load result in Q4

current_status_28_days 28-day treatment status of client as at generation of line-list

current_status_90_days 90-day treatment status of client as at generation of line-list

current_status_q1_28_days 28-day treatment status of client as at end of Q1

current_status_q1_90_days 28-day treatment status of client as at end of Q1

current_status_q2_28_days 28-day treatment status of client as at end of Q2

current_status_q2_90_days 28-day treatment status of client as at end of Q2

current_status_q3_28_days 28-day treatment status of client as at end of Q3

current_status_q3_90_days 90-day treatment status of client as at end of Q3

current_status_q4_28_days 28-day treatment status of client as at end of Q4

current_status_q4_90_days 28-day treatment status of client as at end of Q4

patient_has_died TRUE or FALSE (or NA) indicating if patient is alive or not

patient_deceased_date Date patient known to be deceased, if dead

patient_transferred_out TRUE or FALSE (or NA) indicating if patient has been transferred out

transferred_out_date Date of transfer, if transferred out

patient_transferred_in TRUE or FALSE (or NA) indicating if patient is a transferred from other facility

transferred_in_date TRUE or FALSE (or NA) indicating date client was transferred in

x49 column missing column name and containing negligible entries

current_status TRUE or FALSE indicating whether the client is active or not based on the calculated date of LTFU

date_lost The calculated expected LTFU date based on the last_drug_pick_up_date, days_of_arv_refill and 28 days missed appointment date

appointment_date The calculated expected next date of medication refill based on the last_drug_pick_up_date and days_of_arv_refill

Note

for more information, kindly visit https://ndr.phis3project.org.ng/

read_ndr *Import NDR Line-lists into R*

Description

Import the basic NDR patient-level line lists (treatment, recent infection, and HTS) into R. Column names and types are appropriately formatted using the type argument.

Usage

```
read_ndr(path, type = "treatment", ...)
```

Arguments

path	Path to the csv file on your computer. The file path should be specified in the format "C:/users/Desktop/your file.csv" or something similar.
type	type of line list to be imported. Currently, the read_ndr() supports any of treatment, recency or hts line-lists.
...	passes other arguments supplied based on the specific NDR line-list to be read. NDR treatment line-list requires the time_stamp argument for the day of reference of the treatment line-list, and an optional quiet argument (which defaults to FALSE) denoting if R should also print the message about the new columns created. An optional cols can also be used to supply specific column type to allow the user control over the variable types should the default satisfy the user need.

Value

nicely formatted line-list

Examples

```
# Read \code{ndr_example.csv} from a path

file_path <- system.file("extdata", "ndr_example.csv", package = "tidyndr")
read_ndr(file_path, time_stamp = "2021-02-15")

# Read using a link to the NDR csv file on the internet

file_path <- "https://raw.githubusercontent.com/stephenbalogun/example_files/main/ndr_example.csv"
read_ndr(file_path, time_stamp = "2021-02-15")
```

`recency_example` *Line-list of 10,000 Simulated Clients Provided in the NDR Format.*

Description

A dataset of randomly generated HIV-1 recent infection line-list to simulate the NDR patient recency line-list.

Usage

```
recency_example
```

Format

A data frame with 10000 rows and 36 variables:

ip Implementing Partner

facility_state State of registration of recency facility

facility_lga Local Government Area where client was registered

facility Facility where the recent infection testing took place

datim_code 'DATIM CODE' of the facility of registration

client_state State of resident of client

client_lga Local Government Area where client is resident

client_code registration number assigned to client at testing point

sex The gender that the client identified as, "M" or "F"

date_of_birth Birth day of client, in "yyyy-dd-mm"

age Age of client as at when the recent infection testing took place

age_group age-band of clients based on the MER classification

client_id unique identifier assigned to client

visit_date Date when HIV testing services was offered

hts_screening_result Outcome of the first HIV testing offered

hts_result interpretation of the HTS screening

hts_confirmatory_result Outcome of second HIV testing offered

hts_tie_breaker_result Outcome of third HIV testing offered (when required)

testing_point Place where the HIV testing was conducted

index_client is the client a result of Index Testing Services

opt_out Did the client opt out of recency testing?

recency_test_name Name of testing kit used for the recent infection testing

recency_test_date Date when the recency testing was conducted

recency_number Number assigned during recency testing

control_line Was the control line visible on the Asante kit during testing?

verification_line Was the verification_line visible during testing?

longterm_line Was the longterm line visible during testing?

recency_interpretation Interpretation of the rapid testing for recency

final_recency_result The final recency outcome following Viral Load

viral_load_requested Was the viral load of the client requested?

viral_load_result The viral load count of the client

date_of_viral_load_result The date the viral load result was released

date_sample_collected The date of viral load sample collection

date_sample_sent The date the collected viral load sample was sent for assay

pcr_lab laboratory where recency viral load is analyzed

x36 column missing column name and containing negligible entries

Note

for more information, kindly visit https://ndr.phis3project.org.ng/

`recent_eligible` *Subset Clients who had Recent Infection Test Done during a period of Interest*

Description

Subset Clients who had Recent Infection Test Done during a period of Interest

Usage

```
recent_eligible(data, from = NULL, to = NULL, states = NULL, facilities = NULL)
```

Arguments

`data`	An NDR dataframe imported using the `read_ndr()`.
`from`	The start date in ISO8601 format (i.e. "yyyy-mm-dd"). The default is to start at the beginning of the current Fiscal Year (i.e. 1st of October).
`to`	The end date written in ISO8601 format (i.e. "yyyy-mm-dd"). The default is the date of analysis.
`states`	The name(s) of the State(s) of interest. The default utilizes all the states in the dataframe. If specifying more than one state, combine the states using the `c()` e.g. `c("State 1","State 2")`.
`facilities`	The name(s) of the facilit(ies) of interest. Default is to utilize all the facilities contained in the dataframe. If specifying more than one facility, combine the facilities using the `c()` e.g. `c("Facility 1","Facility 2")`.

Value

line-list of clients who had recent HIV infection test done during the period of interest

Examples

```
## Line-list all HIV positive clients who are eligible for recency testing
hts_pos <- hts_tst_pos(recency_example)

recent_eligible(hts_pos, state = "Arewa") # eligible clients in 'Arewa' state
```

---

rita_recent *Generate List of Clients who are RITA Recent*

Description

Generate List of Clients who are RITA Recent

Usage

```
rita_recent(data, from = NULL, to = NULL, states = NULL, facilities = NULL)
```

Arguments

data	An NDR dataframe imported using the `read_ndr()`.
from	The start date in ISO8601 format (i.e. "yyyy-mm-dd"). The default is to start at the beginning of the current Fiscal Year (i.e. 1st of October).
to	The end date written in ISO8601 format (i.e. "yyyy-mm-dd"). The default is the date of analysis.
states	The name(s) of the State(s) of interest. The default utilizes all the states in the dataframe. If specifying more than one state, combine the states using the `c()` e.g. `c("State 1","State 2")`.
facilities	The name(s) of the facilit(ies) of interest. Default is to utilize all the facilities contained in the dataframe. If specifying more than one facility, combine the facilities using the `c()` e.g. `c("Facility 1","Facility 2")`.

Value

Line-list of clients who are RTRI recent and have a viral load result greater or equal to 1000 copies per microliter of blood.

Examples

```
## Line-list all HIV positive clients confirmed to be RITA recent
hts_pos <- hts_tst_pos(recency_example)

rita_recent(hts_pos)
```

rita_result — *Generate List of RTRI-Positive Clients with Documented Viral Load Results*

Description

Generate List of RTRI-Positive Clients with Documented Viral Load Results

Usage

```
rita_result(data, from = NULL, to = NULL, states = NULL, facilities = NULL)
```

Arguments

data	An NDR dataframe imported using the `read_ndr()`.
from	The start date in ISO8601 format (i.e. "yyyy-mm-dd"). The default is to start at the beginning of the current Fiscal Year (i.e. 1st of October).
to	The end date written in ISO8601 format (i.e. "yyyy-mm-dd"). The default is the date of analysis.
states	The name(s) of the State(s) of interest. The default utilizes all the states in the dataframe. If specifying more than one state, combine the states using the `c()` e.g. `c("State 1","State 2")`.
facilities	The name(s) of the facilit(ies) of interest. Default is to utilize all the facilities contained in the dataframe. If specifying more than one facility, combine the facilities using the `c()` e.g. `c("Facility 1","Facility 2")`.

Value

Line-list of clients with recent infection and a viral load result

Examples

```
## Get HTS_POS clients who had a documented viral load result for recency in 2021
rita_result(recency_example, from = "2021-01-01")
```

rita_sample — *Generate List of RTRI-Positive Clients whose Viral Load Samples Have Been Collected*

Description

Generate List of RTRI-Positive Clients whose Viral Load Samples Have Been Collected

Usage

```
rita_sample(data, from = NULL, to = NULL, states = NULL, facilities = NULL)
```

Arguments

data	An NDR dataframe imported using the read_ndr().
from	The start date in ISO8601 format (i.e. "yyyy-mm-dd"). The default is to start at the beginning of the current Fiscal Year (i.e. 1st of October).
to	The end date written in ISO8601 format (i.e. "yyyy-mm-dd"). The default is the date of analysis.
states	The name(s) of the State(s) of interest. The default utilizes all the states in the dataframe. If specifying more than one state, combine the states using the c() e.g. c("State 1","State 2").
facilities	The name(s) of the facilit(ies) of interest. Default is to utilize all the facilities contained in the dataframe. If specifying more than one facility, combine the facilities using the c() e.g. c("Facility 1","Facility 2").

Value

Line-list of clients with recent infection and a viral load samples collected

Examples

```
## Get HTS_POS clients who had recency testing and viral load sample collected
sample_collected <- rita_sample(recency_example)

## Samples collected in 'Ayetoro', and 'Arewa' states

rita_sample(recency_example, states = c("Arewa", "Ayetoro"))
```

---

rtri_recent *Get List of RTRI Recent Clients*

Description

Get List of RTRI Recent Clients

Usage

```
rtri_recent(data, from = NULL, to = NULL, states = NULL, facilities = NULL)
```

Arguments

data	An NDR dataframe imported using the read_ndr().
from	The start date in ISO8601 format (i.e. "yyyy-mm-dd"). The default is to start at the beginning of the current Fiscal Year (i.e. 1st of October).
to	The end date written in ISO8601 format (i.e. "yyyy-mm-dd"). The default is the date of analysis.

states	The name(s) of the State(s) of interest. The default utilizes all the states in the dataframe. If specifying more than one state, combine the states using the c() e.g. c("State 1","State 2").
facilities	The name(s) of the facilit(ies) of interest. Default is to utilize all the facilities contained in the dataframe. If specifying more than one facility, combine the facilities using the c() e.g. c("Facility 1","Facility 2").

Value

Line-list of clients with who are RTRI Positive

Examples

```
## Line-list RTRI recent persons for 'Arewa' state for the first quarter of 2021
rtri_recent(recency_example, states = "Arewa", from = "2021-01-01", to = "2021-03-31")
```

---

summarise_ndr *Count the Number of Outcomes Based on a Specified Level*

Description

The summarise_ndr() function counts the number of occurrence of specified level for each of the supplied dataframe. It then combines the given dataframes into a single table. It also adds a "Total" roll that adds all the rows for each of the numeric columns.

Usage

```
summarise_ndr(..., level = "state", names = NULL)
```

Arguments

...	Data frames to be summarised.
level	The level at which the summary should be performed. The options are "ip" (or "country"), "state", "lga" or "facility".
names	The names to be passed to the summary columns created in the output

Value

summarise_ndr

Examples

```
new <- tx_new(ndr_example, from = "2021-03-01")
curr <- tx_curr(ndr_example)

summarise_ndr(
  new,
  curr,
  level = "state",
  names = c("tx_new", "tx_curr")
)

### summarise for only one dataframe (defaults data name when name is not specified)
summarise_ndr(
  data = new,
  level = "ip"
)
```

`summarise_recency` *Count the Number of Recency Outcomes Based on a Specified Level*

Description

The `summarise_recency()` function counts the number of occurrence of specified level for each of the supplied dataframe. It then combines the given dataframes into a single table. It also adds a "Total" roll that adds all the rows for each of the numeric columns.

Usage

```
summarise_recency(..., level = "facility_state", names = NULL)
```

Arguments

`...`	Data frames to be summarised.
`level`	The level at which the summary should be performed. The options are "ip" (or "country"), "facility_state", "facility_lga", "facility", "client_state", or "client_lga".
`names`	The names to be passed to the summary columns created in the output

Value

summary dataframe of recency indicators

Examples

```
hts_pos <- hts_tst_pos(recency_example, from = "2021-01-01") # positive clients from January 2021
hts_recent <- hts_recent(hts_pos) # positive clients from above who had recency testing done
rtri_recent <- rtri_recent(hts_recent) # hts_recent clients who were presumed recent from RTRI

summarise_recency(
```

```
  hts_pos,
  hts_recent,
  rtri_recent,
  level = "facility_state",
  names = c("positives", "recency_testing", "rtri_recent")
)

### If the `names` argument is not supplied, the names of the supplied data will be used instead
summarise_recency(
  hts_pos,
  hts_recent
)
```

---

summarize_ndr *Count the Number of Outcomes Based on a Specified Level*

Description

The summarize_ndr() function counts the number of occurrence of specified level for each of the supplied dataframe. It then combines the given dataframes into a single table. It also adds a "Total" roll that adds all the rows for each of the numeric columns.

Usage

```
summarize_ndr(..., level = "state", names = NULL)
```

Arguments

...	Dataframes to be summarized.
level	The level at which the summary should be performed. The options are "ip" (or "country"), "state", "lga" or "facility".
names	The names to be passed to the summary columns created in the output

Value

summarize_ndr

Examples

```
new <- tx_new(ndr_example, from = "2021-03-01")
curr <- tx_curr(ndr_example)

summarize_ndr(
  new,
  curr,
  level = "state",
  names = c("tx_new", "tx_curr")
)
```

```
### summarize for only one dataframe (defaults data name when name is not specified)
summarize_ndr(
  data = new,
  level = "ip"
)
```

---

summarize_recency *Count the Number of Recency Outcomes Based on a Specified Level*

Description

The summarize_recency() function counts the number of occurrence of specified level for each of the supplied dataframe. It then combines the given dataframes into a single table. It also adds a "Total" roll that adds all the rows for each of the numeric columns.

Usage

```
summarize_recency(..., level = "facility_state", names = NULL)
```

Arguments

...	Dataframes to be summarized.
level	The level at which the summary should be performed. The options are "ip" (or "country"), "facility_state", "facility_lga", "facility", "client_state", or "client_lga".
names	The names to be passed to the summary columns created in the output

Value

summary dataframe of recency indicators

Examples

```
hts_pos <- hts_tst_pos(recency_example, from = "2021-01-01") # positive clients from January 2021
hts_recent <- hts_recent(hts_pos) # positive clients from above who had recency testing done
rtri_recent <- rtri_recent(hts_recent) # hts_recent clients who were presumed recent from RTRI

summarize_recency(
  hts_pos,
  hts_recent,
  rtri_recent,
  level = "facility_state",
  names = c("positives", "recency_testing", "rtri_recent")
)

### If the `names` argument is not supplied, the names of the supplied data will be used instead
summarize_recency(
  hts_pos,
  hts_recent
```

```
)
```

`tx_appointment` *Subset Rows of Clients who have Clinic Appointment/Medication Pick-up within a Particular Period*

Description

`tx_appointment` generates the line-list of clients who have clinic appointment/medication refill for the specified state(s) and/or facilit(ies). The default is to generate the appointment list for all the states/facilities.

Usage

```
tx_appointment(
  data,
  from = NULL,
  to = NULL,
  states = NULL,
  facilities = NULL,
  active = FALSE,
  remove_duplicates = FALSE
)
```

Arguments

`data`	An NDR dataframe imported using the `read_ndr()`.
`from`	The start date in ISO8601 format (i.e. "yyyy-mm-dd"). The default is to start at the beginning of the current Fiscal Year (i.e. 1st of October).
`to`	The end date written in ISO8601 format (i.e. "yyyy-mm-dd"). The default is the date of analysis.
`states`	The name(s) of the State(s) of interest. The default utilizes all the states in the dataframe. If specifying more than one state, combine the states using the `c()` e.g. `c("State 1","State 2")`.
`facilities`	The name(s) of the facilit(ies) of interest. Default is to utilize all the facilities contained in the dataframe. If specifying more than one facility, combine the facilities using the `c()` e.g. `c("Facility 1","Facility 2")`.
`active`	TRUE or FALSE. To determine whether the appointment should be for only active patients or irrespective of their status.
`remove_duplicates`	Boolean argument. It specifies if duplicate patient entries in the facilities should be removed or kept

Value

tx_appointment

Examples

```
# Determine clients who have medication refill in Q2 of FY21
tx_appointment(ndr_example,
  from = "2021-06-01",
  to = "2021-09-30"
)

# Determine clients who have medication refill in July 2021
tx_appointment(ndr_example,
  from = "2021-07-01",
  to = "2021-07-31",
)

# Determine clients with medication refill in "Okun" state for a particular facility
tx_appointment(ndr_example,
  from = "2021-01-01",
  states = "Okun",
  facilities = "Facility1"
)
```

---

tx_curr *Subset Clients who are Currently on Treatment*

Description

`tx_curr` pulls up the line-list of clients who are active on treatment using the calculated `current_status` column. You can specify the state(s) and/or facilit(ies) of interest using the `region` or `site` arguments.

Usage

```
tx_curr(
  data,
  states = NULL,
  facilities = NULL,
  status = "default",
  remove_duplicates = FALSE
)
```

Arguments

`data`	An NDR dataframe imported using the `read_ndr()`.
`states`	The name(s) of the State(s) of interest. The default utilizes all the states in the dataframe. If specifying more than one state, combine the states using the `c()` e.g. `c("State 1","State 2")`.
`facilities`	The name(s) of the facilit(ies) of interest. Default is to utilize all the facilities contained in the dataframe. If specifying more than one facility, combine the facilities using the `c()` e.g. `c("Facility 1","Facility 2")`.

status Determines how the number of active clients is calculated. The options are to either to use the NDR current_status_28_days column or the derived current_status column ("calculated").

remove_duplicates
Boolean argument. It specifies if duplicate patient entries in the facilities should be removed or kept

Value

TX_CURR

Examples

```
# Calculated active clients using the derived current status
tx_curr(ndr_example)

# Calculate the active clients using the NDR `current_status_28_days` column
tx_curr(ndr_example, status = "default")

# generate the TX_CURR for two states (e.g. "Arewa" and "Okun" in the ndr_example file)
tx_curr(ndr_example,
  states = c("Okun", "Arewa")
)

# determine the active clients in two facilities ("Facility1", and "Facility2) in "Abaji"
tx_curr(ndr_example,
  states = "Abaji",
  facilities = c("Facility1", "Facility2")
)
```

---

tx_ml *Subset Clients who Became Inactive (IIT) Within a Given Period*

Description

`tx_ml` Generates clients who have become inactive over a specified period of time. The default is to generate all clients who became inactive in the current Fiscal Year. You can specify the period of interest (using the `from` and `to` arguments). Used together with `tx_ml_outcomes()`, generates inactive clients with a particular outcome of interest.

Usage

```
tx_ml(
  new_data,
  old_data = NULL,
  from = NULL,
  to = NULL,
  states = NULL,
  facilities = NULL,
```

```
  status = "default",
  remove_duplicates = FALSE
)
```

Arguments

new_data	The current datafame where changes in current treatment status will be checked.
old_data	The initial dataframe containing the list of clients who were previously active.
from	The start date in ISO8601 format (i.e. "yyyy-mm-dd"). The default is to start at the beginning of the current Fiscal Year (i.e. 1st of October).
to	The end date written in ISO8601 format (i.e. "yyyy-mm-dd"). The default is the date of analysis.
states	The name(s) of the State(s) of interest. The default utilizes all the states in the dataframe. If specifying more than one state, combine the states using the `c()` e.g. `c("State 1","State 2")`.
facilities	The name(s) of the facilit(ies) of interest. Default is to utilize all the facilities contained in the dataframe. If specifying more than one facility, combine the facilities using the `c()` e.g. `c("Facility 1","Facility 2")`.
status	Determines how the number of active clients is calculated. The options are to either to use the NDR current_status_28_days column or the derived current_status column ("calculated").
remove_duplicates	Boolean argument. It specifies if duplicate patient entries in the facilities should be removed or kept

Value

tx_ml

Examples

```
tx_ml(new_data = ndr_example, from = "2021-06-01")

# Find clients who were inactive in Q4 of FY21
tx_ml(
  new_data = ndr_example,
  from = "2021-07-01", to = "2021-09-30"
)

## generate line-list of `tx_ml()` using two datasets

file_path <- "https://raw.githubusercontent.com/stephenbalogun/example_files/main/ndr_example.csv"
ndr_old <- read_ndr(file_path, time_stamp = "2021-02-15")
ndr_new <- ndr_example
tx_ml(
  old_data = ndr_old,
  new_data = ndr_new,
  from = "2021-07-01",
  to = "2021-09-30"
```

```
)
```

`tx_ml_outcomes` *Subset rows of Inactive Clients with Specific Outcome*

Description

`tx_ml_outcomes` generates the line-list of clients based on the outcome of interest ("dead" or "transfer out"). It should be used after `tx_ml()`.

Usage

```
tx_ml_outcomes(data, outcome)
```

Arguments

`data`	An ndr dataframe imported using the 'read_ndr()
`outcome`	The particular outcome of interest based on options available on the NDR ("transfer out" or "dead").

Value

tx_ml_outcomes

Examples

```
tx_ml_outcomes(tx_ml(new_data = ndr_example),
  outcome = "dead"
)
```

`tx_mmd` *Subset active clients based on months of ARV Dispensed*

Description

Generates list of clients who had 3 - 6 months of ARV dispensed during the medication refill. You can specify the number of month(s) of ARV dispensed by changing the `month` argument.

Usage

```
tx_mmd(
  data,
  months = NULL,
  states = NULL,
  facilities = NULL,
  status = "default",
  remove_duplicates = FALSE
)
```

Arguments

data	An NDR dataframe imported using the read_ndr().
months	The number(s) of months of interest of ARV dispensed. The default is to subset active clients who had 3 - 6 months of ARV dispensed.
states	The name(s) of the State(s) of interest. The default utilizes all the states in the dataframe. If specifying more than one state, combine the states using the c() e.g. c("State 1","State 2").
facilities	The name(s) of the facilit(ies) of interest. Default is to utilize all the facilities contained in the dataframe. If specifying more than one facility, combine the facilities using the c() e.g. c("Facility 1","Facility 2").
status	Determines how the number of active clients is calculated. The options are to either to use the NDR current_status_28_days column or the derived current_status column ("calculated").
remove_duplicates	Boolean argument. It specifies if duplicate patient entries in the facilities should be removed or kept

Value

tx_mmd

Examples

```
tx_mmd(ndr_example)

# subset active clients who had 2 or 4 months of ARV dispensed at last encounter
tx_mmd(ndr_example,
  months = c(2, 4),
  status = "default"
)
```

---

tx_new *Subset Clients Starting ART Within a Particular Period.*

Description

Generates the line-list of clients who commenced ARV within the specified period of interest. The default is to generate the list for all clients who commenced ARV in the current Fiscal Year. You can specify the period of interest using the from and to arguments; and the state or facility of interest with the states and facilities arguments. For multiple states or facilities, use the c() to combine the names.

Usage

```
tx_new(
  data,
  from = NULL,
  to = NULL,
  states = NULL,
  facilities = NULL,
  remove_duplicates = FALSE
)
```

Arguments

data	An NDR dataframe imported using the `read_ndr()`.
from	The start date in ISO8601 format (i.e. "yyyy-mm-dd"). The default is to start at the beginning of the current Fiscal Year (i.e. 1st of October).
to	The end date written in ISO8601 format (i.e. "yyyy-mm-dd"). The default is the date of analysis.
states	The name(s) of the State(s) of interest. The default utilizes all the states in the dataframe. If specifying more than one state, combine the states using the `c()` e.g. `c("State 1","State 2")`.
facilities	The name(s) of the facilit(ies) of interest. Default is to utilize all the facilities contained in the dataframe. If specifying more than one facility, combine the facilities using the `c()` e.g. `c("Facility 1","Facility 2")`.
remove_duplicates	Boolean argument. It specifies if duplicate patient entries in the facilities should be removed or kept

Value

TX_NEW clients in the period of interest

Examples

```
tx_new(ndr_example, from = "2021-06-01", to = "2021-09-30")

# generate the TX_NEW for a specific state (Ayetoro)
tx_new(ndr_example, states = "Ayetoro")

# Determine the TX_NEW for Quarter 1 of FY21 for State 2
tx_new(ndr_example,
  from = "2021-10-01",
  to = "2021-12-31",
  states = c("Arewa", "Ayetoro")
)
```

`tx_pvls_den` *Subset Clients who have a Documented Viral Load Result*

Description

Generate the line-list of clients whose date of last viral load result is not more than one year (for adults 20 years and above) and 6 months (for pediatrics and adolescents) from the specified reference date.

Usage

```
tx_pvls_den(
  data,
  ref = NULL,
  states = NULL,
  facilities = NULL,
  status = "default",
  use_six_months = TRUE,
  remove_duplicates = FALSE
)
```

Arguments

`data`	An NDR dataframe imported using the 'read_ndr().
`ref`	Date provided in ISO8601 format ("yyyy-mm-dd"). Used to determine clients who are eligible for viral load and should have a documented result. The default is the date of analysis.
`states`	The name(s) of the State(s) of interest. The default utilizes all the states in the dataframe. If specifying more than one state, combine the states using the `c()` e.g. `c("State 1","State 2")`.
`facilities`	The name(s) of the facilit(ies) of interest. Default is to utilize all the facilities contained in the dataframe. If specifying more than one facility, combine the facilities using the `c()` e.g. `c("Facility 1","Facility 2")`.
`status`	Determines how the number of active clients is calculated. The options are to either to use the NDR current_status_28_days column or the derived current_status column ("calculated").
`use_six_months`	`TRUE` of `FALSE`. To determine if six months eligibility should be used for pediatrics and adolescents repeat viral load or the standard annual repeat.
`remove_duplicates`	Boolean argument. It specifies if duplicate patient entries in the facilities should be removed or kept

Value

tx_pvls_den

Examples

```
tx_pvls_den(ndr_example, status = "default")

# Determine clients who are virally suppressed for two state at the end of Q4
tx_pvls_den(ndr_example,
  ref = "2021-09-30",
  states = c("Okun", "Arewa")
)
```

`tx_pvls_num` *Determine Clients who are Virally Suppressed*

Description

Generate the line-list of clients whose date of last viral load result is not more than one year (for adults 20 years and above) and 6 months (for pediatrics and adolescents) from the specified reference date and are virally suppressed.

Usage

```
tx_pvls_num(
  data,
  ref = NULL,
  states = NULL,
  facilities = NULL,
  status = "default",
  n = 1000,
  use_six_months = TRUE,
  remove_duplicates = FALSE
)
```

Arguments

`data`	An NDR dataframe imported using the 'read_ndr().
`ref`	Date provided in ISO8601 format ("yyyy-mm-dd"). Used to determine clients who are eligible for viral load and should have a documented result. The default is the date of analysis.
`states`	The name(s) of the State(s) of interest. The default utilizes all the states in the dataframe. If specifying more than one state, combine the states using the `c()` e.g. `c("State 1","State 2")`.
`facilities`	The name(s) of the facilit(ies) of interest. Default is to utilize all the facilities contained in the dataframe. If specifying more than one facility, combine the facilities using the `c()` e.g. `c("Facility 1","Facility 2")`.
`status`	Determines how the number of active clients is calculated. The options are to either to use the NDR current_status_28_days column or the derived current_status column ("calculated").

n the value below which viral load result is adjudged to be suppressed.

use_six_months TRUE of FALSE. To determine if six months eligibility should be used for pediatrics and adolescents repeat viral load or the standard annual repeat.

remove_duplicates
Boolean argument. It specifies if duplicate patient entries in the facilities should be removed or kept

Value

tx_pvls_num

Examples

```
tx_pvls_num(ndr_example)

# Determine clients who are virally suppressed for a state at the end of October 2021
tx_pvls_num(ndr_example,
  ref = "2021-10-31",
  states = "Arewa"
)

# Determine clients with viral load result less than 400
tx_pvls_num(ndr_example, n = 400)
```

---

tx_regimen *Subset Clients Based on their Current ART Regimen*

Description

Generates the line-list of clients on first-line regimen who are on the choice combination regimen for their age or weight. The NDR does not currently report 'weight' so the function uses 'age' to approximate the choice-regimen for the clients.

Usage

```
tx_regimen(
  data,
  age_band = NULL,
  states = NULL,
  facilities = NULL,
  status = "default",
  remove_duplicates = FALSE
)
```

Arguments

data	An NDR dataframe imported using the 'read_ndr()
age_band	a numeric vector of length 2 c(min_age,max_age).
states	The name(s) of the State(s) of interest. The default utilizes all the states in the dataframe. If specifying more than one state, combine the states using the c() e.g. c("State 1","State 2").
facilities	The name(s) of the facilit(ies) of interest. Default is to utilize all the facilities contained in the dataframe. If specifying more than one facility, combine the facilities using the c() e.g. c("Facility 1","Facility 2").
status	Determines how the number of active clients is calculated. The options are to either to use the NDR current_status_28_days column or the derived current_status column ("calculated").
remove_duplicates	Boolean argument. It specifies if duplicate patient entries in the facilities should be removed or kept

Value

tx_regimen

Examples

```
tx_regimen(ndr_example)

tx_regimen(ndr_example,
  status = "default",
  age_band = c(0, 3)
)
```

---

tx_rtt *Subset Rows of Previously Inactive Clients Who are Now Active*

Description

Generates the line-list of clients who were inactive in the data supplied to the old_data argument but have now become active in the data supplied to the new_data argument.

Usage

```
tx_rtt(
  new_data,
  old_data,
  states = NULL,
  facilities = NULL,
  status = "default",
  remove_duplicates = FALSE
)
```

Arguments

new_data	The current datafame where changes in current treatment status will be checked.
old_data	The initial dataframe containing the list of clients who have been previously inactive.
states	The name(s) of the State(s) of interest. The default utilizes all the states in the dataframe. If specifying more than one state, combine the states using the c() e.g. c("State 1","State 2").
facilities	The name(s) of the facilit(ies) of interest. Default is to utilize all the facilities contained in the dataframe. If specifying more than one facility, combine the facilities using the c() e.g. c("Facility 1","Facility 2").
status	Determines how the number of active clients is calculated. The options are to either to use the NDR current_status_28_days column or the derived current_status column ("calculated").
remove_duplicates	Boolean argument. It specifies if duplicate patient entries in the facilities should be removed or kept

Value

tx_rtt

Examples

```
file_path <- "https://raw.githubusercontent.com/stephenbalogun/example_files/main/ndr_example.csv"
ndr_new <- read_ndr(file_path, time_stamp = "2021-02-15")
ndr_old <- ndr_example
tx_rtt(ndr_new, ndr_old)


## Determine RTT for a particular state

tx_rtt(ndr_old, ndr_new, states = "State 1")
```

---

tx_vl_eligible *Subset Clients who are Eligible for Viral Load*

Description

Generates the line-list of clients who have been (or would have been) on ARV medications for at least 6 months from the reference date. The default reference date is the date of analysis.

Usage

```
tx_vl_eligible(
  data,
  ref = NULL,
  states = NULL,
  facilities = NULL,
  status = "default",
  sample = FALSE,
  use_six_months = TRUE,
  remove_duplicates = FALSE
)
```

Arguments

data	An NDR dataframe imported using the 'read_ndr().
ref	Date provided in ISO8601 format ("yyyy-mm-dd"). Used to determine clients who are eligible for viral load and should have a documented result. The default is the date of analysis.
states	The name(s) of the State(s) of interest. The default utilizes all the states in the dataframe. If specifying more than one state, combine the states using the c() e.g. c("State 1","State 2").
facilities	The name(s) of the facilit(ies) of interest. Default is to utilize all the facilities contained in the dataframe. If specifying more than one facility, combine the facilities using the c() e.g. c("Facility 1","Facility 2").
status	Determines how the number of active clients is calculated. The options are to either to use the NDR current_status_28_days column or the derived current_status column ("calculated").
sample	Logical (TRUE or FALSE) indicating whether all clients eligible for viral load test should be filtered irrespective of their eligibility for sample collection or only those due for sample collection.
use_six_months	TRUE of FALSE. To determine if six months eligibility should be used for pediatrics and adolescents repeat viral load or the standard annual repeat.
remove_duplicates	Boolean argument. It specifies if duplicate patient entries in the facilities should be removed or kept

Value

tx_vl_eligible

Examples

```
tx_vl_eligible(ndr_example, ref = "2021-09-30")

# Determine clients who are going to be eligible for VL by the end of Q1 of FY22
tx_vl_eligible(ndr_example,
  ref = "2021-12-31"
```

```
)

# Subset clients from "Arewa" and "Okun" who are due for viral load in Q1 of FY22
tx_vl_eligible(ndr_example,
  ref = "2021-12-31",
  states = c("Arewa", "Okun"),
  sample = TRUE
)
```

`tx_vl_unsuppressed` *Determine Clients who are not Virally Suppressed*

Description

Generate the line-list of clients whose date of last viral load result is not not more than one year (for adults 20 years and above) and 6 months (for pediatrics and adolescents) from the specified reference date and are not virally suppressed.

Usage

```
tx_vl_unsuppressed(
  data,
  ref = NULL,
  states = NULL,
  facilities = NULL,
  status = "default",
  n = 1000,
  use_six_months = TRUE,
  remove_duplicates = FALSE
)
```

Arguments

`data`	An NDR dataframe imported using the 'read_ndr().
`ref`	Date provided in ISO8601 format ("yyyy-mm-dd"). Used to determine clients who are eligible for viral load and should have a documented result. The default is the date of analysis.
`states`	The name(s) of the State(s) of interest. The default utilizes all the states in the dataframe. If specifying more than one state, combine the states using the `c()` e.g. `c("State 1","State 2")`.
`facilities`	The name(s) of the facilit(ies) of interest. Default is to utilize all the facilities contained in the dataframe. If specifying more than one facility, combine the facilities using the `c()` e.g. `c("Facility 1","Facility 2")`.
`status`	Determines how the number of active clients is calculated. The options are to either to use the NDR current_status_28_days column or the derived current_status column ("calculated").

n	the value below which viral load result is adjudged to be suppressed.
use_six_months	TRUE of FALSE. To determine if six months eligibility should be used for pediatrics and adolescents repeat viral load or the standard annual repeat.
remove_duplicates	Boolean argument. It specifies if duplicate patient entries in the facilities should be removed or kept

Value

tx_vl_unsuppressed

Examples

```
tx_vl_unsuppressed(ndr_example)

# Determine clients who are virally unsuppressed for a state at the end of Q1
tx_vl_unsuppressed(ndr_example,
  ref = "2021-12-31",
  states = "Ayetoro"
)

# Determine clients with viral load result of 400 or more (low level viremia)
tx_vl_unsuppressed(ndr_example, n = 400)
```

---

vl_cascade *Analyse the Viral Load Cascade Indicators*

Description

Generate aggregate summary of viral load indicators based on a referenced date. The indicators include eligible, documented results, virally suppressed, viral load coverage, and viral load suppression rate.

Usage

```
vl_cascade(
  data,
  ref = NULL,
  states = NULL,
  facilities = NULL,
  status = "default",
  n = 1000,
  use_six_months = TRUE,
  remove_duplicates = FALSE,
  .level = "state",
  .names = NULL
)
```

Arguments

data	An NDR dataframe imported using the 'read_ndr().
ref	Date provided in ISO8601 format ("yyyy-mm-dd"). Used to determine clients who are eligible for viral load and should have a documented result. The default is the date of analysis.
states	The name(s) of the State(s) of interest. The default utilizes all the states in the dataframe. If specifying more than one state, combine the states using the c() e.g. c("State 1","State 2").
facilities	The name(s) of the facilit(ies) of interest. Default is to utilize all the facilities contained in the dataframe. If specifying more than one facility, combine the facilities using the c() e.g. c("Facility 1","Facility 2").
status	Determines how the number of active clients is calculated. The options are to either to use the NDR current_status_28_days column or the derived current_status column ("calculated").
n	the value below which viral load result is adjudged to be suppressed.
use_six_months	TRUE of FALSE. To determine if six months eligibility should be used for pediatrics and adolescents repeat viral load or the standard annual repeat.
remove_duplicates	Boolean argument. It specifies if duplicate patient entries in the facilities should be removed or kept
.level	the level at which the aggregate summary should be performed. The options are "ip", "country", "state", "lga" and "facility".
.names	if specified, these will be used for naming of the viral load indicators instead of the default.

Value

summary of viral load cascade

Examples

```
vl_cascade(ndr_example, ref = "2021-12-31", .level = "state")

# Determine the viral load cascade for a state at the end of September 2021
vl_cascade(ndr_example,
  ref = "2021-10-31",
  states = "Arewa"
)
```

Index