Package: faperu (via r-universe)

November 27, 2026

Title Access and Design-Based Analysis of Peruvian Official Surveys and Censuses

Version 0.4.1

Description One catalog, one download function and a few helpers to read, weight, locate and analyse ten Peruvian official data sources: ENAHO, ENDES, ENA, EPEN, ENAPRES, ENUT, ENARES and the Annual Economic Survey from INEI, the School Census from MINEDU (ESCALE) and the 2017 National Censuses. Files are cached and checked with SHA-256, and new survey years are detected from the INEI microdata portal. Design-based estimates of means, proportions and totals come with standard errors, confidence intervals and the coefficient-of-variation quality bands used by INEI; seven ready indicators reproduce INEI's published figures; estimates from several surveys can be linked by district and year and drawn as maps and tables. This package is independent and is not affiliated with INEI or MINEDU.

License MIT + file LICENSE

Encoding UTF-8

Depends R (>= 4.1)

Imports curl (>= 5.0.0), digest, foreign, graphics, haven (>= 2.4.0), stats, survey, tidyselect (>= 1.1.0), tools, utils, yaml

Suggests data.table, ggplot2, grDevices, grid, knitr, redatamx, rmarkdown, sf, testthat (>= 3.0.0), withr

VignetteBuilder knitr

Config/testthat/edition 3

URL https://faperu.ftorres.org/, https://faperu.ftorres.org/docs/, https://github.com/Andre031222/peru-official-surveys-r

BugReports https://github.com/Andre031222/peru-official-surveys-r/issues

Roxygen list(markdown = TRUE)

Config/roxygen2/version 8.0.0

Config/pak/sysreqs make libssl-dev libx11-dev zlib1g-dev

Repository https://andre031222.r-universe.dev

Date/Publication 2026-11-27 17:02:51 UTC

RemoteUrl https://github.com/Andre031222/peru-official-surveys-r

RemoteRef main

RemoteSha 025e353d64f19abb3aca86d8045bf4df3d429139

Contents

estimate-methods . . . 2
fa_cache . . . 3
fa_catalog . . . 4
fa_check_updates . . . 5
fa_design . . . 6
fa_districts . . . 7
fa_download . . . 8
fa_estimate . . . 9
fa_geometry . . . 10
fa_indicator . . . 11
fa_indicators . . . 12
fa_link . . . 12
fa_map . . . 14
fa_quality . . . 15
fa_read . . . 16
fa_redatam . . . 17
fa_surveys . . . 18
fa_table . . . 19
fa_ubigeo . . . 20
fa_variables . . . 21
fa_years . . . 22
faperu-methods . . . 23

Index **24**

`estimate-methods` *Print, summarise and plot estimate tables*

Description

`fa_estimate()` and `fa_indicator()` return data frames of class `fa_estimate`, and `fa_link()` one of class `fa_link`. They print as a formatted table (the one `fa_table()` builds), `summary()` counts the estimates in each INEI quality band, and `plot()` draws each estimate with its confidence interval, coloured by quality. They are still data frames; `as.data.frame()` drops the class.

Usage

```
## S3 method for class 'fa_estimate'
print(x, n = 30L, percent = NULL, ...)

## S3 method for class 'fa_estimate'
summary(object, ...)

## S3 method for class 'fa_estimate'
plot(x, percent = NULL, main = NULL, ...)

## S3 method for class 'fa_link'
print(x, n = 30L, ...)
```

Arguments

x, object	A table from `fa_estimate()`, `fa_indicator()` or `fa_link()`.
n	Rows to show.
percent	Show proportions as percentages.
...	Passed on to `print.data.frame()` or `graphics::plot()`.
main	Plot title; by default the variable or indicator.

Value

`print()` and `plot()` return `x` invisibly. `summary()` returns a data frame with one row per quality band: `quality`, `n` and the share of estimates.

Examples

```
library(survey)
data(api, package = "survey")
d <- svydesign(~dnum, weights = ~pw, data = apiclus1, fpc = ~fpc)
d$variables$high <- d$variables$api00 > 700
e <- fa_estimate(d, "high", type = "prop", by = "stype")
e
summary(e)
plot(e)
```

---

fa_cache *Cache folder*

Description

Without arguments, returns the current cache folder. With `path`, changes it for the session. The default comes from the `FAPERU_CACHE` environment variable or, if unset, from `tools::R_user_dir()`.

Usage

```
fa_cache(path = NULL)
```

Arguments

path	New folder, or NULL to only query.

Details

If you already have the microdata on disk with the same layout (<SURVEY>/<year>/<file>.zip), point the cache there and fa_download() will reuse it.

Value

The cache path (invisibly when it is changed).

Examples

```
fa_cache()
```

---

fa_catalog *Modules of a source*

Description

Lists year, module code, name and available formats. When a copy was already downloaded and verified, it also gives its relative path, size and SHA-256.

Usage

```
fa_catalog(survey, years = NULL)
```

Arguments

survey	Short name, for example "enaho". See fa_surveys().
years	Years to keep. NULL keeps all.

Details

When you ask for a year that the catalog does not have, faperu looks it up at the source the first time: on the INEI microdata portal for INEI surveys (so a year published after this version, or an old one such as ENDES 2009, still works) and in the ESCALE folder for the School Census, whose catalog only lists the latest year. That needs a connection; the modules found are kept for the session, and their series is the last part of the portal survey name.

Value

A data frame with one row per module, of class fa_catalog; summary() counts the modules of each year (see faperu-methods).

Examples

```
head(fa_catalog("enaho", 2024))
```

`fa_check_updates` *Look for new years*

Description

Asks the source which years are published and compares them with the catalog. INEI surveys are checked on the INEI microdata portal and the School Census on the ESCALE folder listing. The 2017 Census is not checked: the 2025 Census will be added by hand once its full data are out.

Usage

```
fa_check_updates(survey = NULL)
```

Arguments

`survey`	One or more short names. `NULL` checks all.

Details

A source that cannot be reached does not stop the others; it shows up with `status = "unreachable"`.

Value

A data frame with `survey`, `year`, `new` (not in the catalog) and `status`, of class `fa_updates` (see faperu-methods).

Examples

```
## Not run:
fa_check_updates("ena")

## End(Not run)
```

`fa_design` *Survey design*

Description

Builds a `survey::svydesign()` object with the weight, strata and cluster variables that the catalog declares for that source. The first declared name found in the data is used, so the same call works for ENDES women (`V005`) and households (`HV005`). Any of them can be overridden, and a stratum can be the crossing of two columns, written `"CCDD*AREA"`.

Usage

```
fa_design(
  data,
  survey,
  weight = NULL,
  strata = NULL,
  cluster = NULL,
  territory = NULL
)
```

Arguments

`data`	A data frame, usually from `fa_read()`.
`survey`	Short name, for example `"enaho"`. See `fa_surveys()`.
`weight, strata, cluster`	Variable names to use instead of the catalog defaults.
`territory`	Optional. Departments, provinces or districts to keep: INEI codes of 2, 4 or 6 digits (`"21"`, `"2101"`, `"210101"`) or department names (`"Puno"`, `"Lima"`), several at once if needed. The data need a `ubigeo` column or `ccdd`, `ccpp` and `ccdi`.

Details

The names were checked in INEI's technical sheets (fichas tecnicas) and in the headers of 2024 and 2025 files; the `diseno` entry of each YAML in `inst/catalog/` lists them with a note and the document behind them.

- **EPEN**: `FAC300_ANUAL` (or the quarterly and Lima weights), `CONGLOMERADO`, and department by `ESTRATO` as strata. Rows without weight (people under 14) are dropped with a message.
- **ENAPRES**: `FACTOR`, `CONGLOMERADO`, department by area. Its `ESTRATO` column is socioeconomic, not a design stratum.
- **ENUT** and **ENARES**: INEI publishes no strata, so there are none and a warning says so. ENARES weights depend on the questionnaire, and in the school questionnaires the cluster is the school.

- **Annual Economic Survey**: FACTOR_EXP, one stage. Firms included with certainty (FACTOR_EXP == 1) form their own stratum with no variance. The real strata are not published, so standard errors are approximate.

Where a stratum has a single cluster, as in ENDES, survey stops unless you set options(survey.lonely.psu = "adjust"); faperu warns but does not change the option for you.

The ENA has thousands of fine strata, most with one cluster. Collapsing them is not official, but if you need it: ena$estrato_c <- paste0(ena$CCDD, substr(ena$ESTRATO, 7, 9)) and fa_design(ena, "ena", strata = "estrato_c").

Territory first, the right way. To study one department, province or district, pass it as territory instead of filtering the rows yourself. Filtering the data before the design drops the strata and clusters of the rest of the country and gives wrong standard errors. With territory, faperu builds the design on the whole sample and then keeps only that territory with survey::subset(), so the estimates are those of the territory and their precision is correct.

Value

A survey.design2 object.

Examples

```
## Not run:
s <- fa_read("enaho", 2024, "34", file = "Sumaria-2024\\.")
d <- fa_design(s, "enaho")
survey::svymean(~I(pobreza < 3), d)

# Only Puno, with correct standard errors
puno <- fa_design(s, "enaho", territory = "Puno")
survey::svymean(~I(pobreza < 3), puno)

## End(Not run)
```

---

fa_districts *Districts valid in a year*

Description

The official INEI district list for one year. It comes from the National Registry of Municipalities (RENAMU) 2004-2025, with three corrections, because RENAMU lists a new district a year or two after its law: a district starts in the year of its creation law when INEI's 2025 national directory gives the date; the 2017 list is the 2017 Census frame (1874 districts); and Santa Rosa de Loreto, missing from RENAMU 2025, is added from MINEDU's 2025 school register. Sources are listed in system.file("ubigeo", "README.md", package = "faperu").

Usage

```
fa_districts(year)
```

Arguments

year	A year between 2004 and 2025.

Value

A data frame with ubigeo, distrito, provincia, departamento, desde, hasta, where desde comes from (desde_fuente), the first year in RENAMU and the creation law.

Examples

```
nrow(fa_districts(2017))
```

---

fa_download *Download modules*

Description

Fetches the requested modules into the cache. Before downloading anything it looks in the cache: a file that is already there and whose SHA-256 matches the catalog is reused as is.

Usage

```
fa_download(
  survey,
  year,
  module = NULL,
  format = "auto",
  series = NULL,
  cache = fa_cache(),
  overwrite = FALSE,
  period = NULL
)
```

Arguments

survey	Short name, for example "enaho". See fa_surveys().
year	One year.
module	Module codes as shown by fa_catalog(), or their names (e.g. "00_Padron" in the School Census). NULL means all modules of that year.
format	"auto" uses the format of the verified copy, or else the first of STATA, SPSS, CSV and DBF that the module has. Stata comes first because it is read exactly, with labels, and quickly; INEI's CSV files can have line breaks inside text answers. You can also ask for one of those formats.
series	For sources with several series, e.g. "panel" in ENAHO.
cache	Cache folder, see fa_cache().
overwrite	Download again even if the file is cached.
period	For sources with several files per year, the period shown by fa_catalog(), e.g. "Anual - (Ene-Dic)" in EPEN. A regular expression.

Details

The files that are missing are downloaded together, a few at a time (`options(faperu.connections = 4)`). A transfer that breaks is retried and resumed from where it stopped, up to `options(faperu.retries = 3)` times; an unfinished file is kept as `.part` so the next call picks it up. Every file is checked before it enters the cache: it must be a readable zip, not an error page. The INEI portal only serves whole modules, so there is no way to download single variables; download the module once and read the columns you need with `fa_read(cols = )`.

Value

Invisibly, a data frame with one row per module, a `status` column (`"cached"`, `"downloaded"` or `"changed"`, when the source now serves a file that differs from the catalog), the size in `bytes` and the download time in `seconds` (0 for cached files), of class `fa_download` (see faperu-methods).

Examples

```
## Not run:
fa_download("ena", 2024, "1893")

## End(Not run)
```

`fa_estimate` *Design-based estimate with precision*

Description

Estimates a mean, a proportion or a total from a survey design (for example one built with `fa_design()`), overall or by domain, and adds the standard error, a confidence interval, the coefficient of variation and the INEI quality band (`fa_quality()`).

Usage

```
fa_estimate(
  design,
  var,
  type = c("mean", "prop", "total"),
  by = "national",
  level = 0.95,
  na.rm = TRUE,
  min_n = 30
)
```

Arguments

`design`	A `survey.design` object.
`var`	Name of the variable to estimate. For `type = "prop"` it must be logical or coded 0/1.

type	"mean", "prop" (a mean of a 0/1 variable, kept between 0 and 1) or "total".
by	Domain: "national", "department", "area", "district", or the name of any column in the design. Departments and districts are read from a ubigeo column or from ccdd, ccpp and ccdi; area from an area column (1 urban, 2 rural) or ENDES hv025.
level	Confidence level.
na.rm	Drop missing values of var.
min_n	Fewest unweighted cases for a figure to be rated by its CV. Below it the quality is "referencial" whatever the CV, because a standard error from a handful of cases is itself unreliable. Set to 0 to rate by the CV alone.

Details

Intervals are normal (Wald) intervals; for proportions they are cut at 0 and 1. Rows without a domain (for example the ENA's agricultural firms, which carry no location) count in national figures but are left out of a breakdown, with a message. If the option survey.lonely.psu is left at its default ("fail"), strata with a single primary unit are handled with "adjust" during the call.

Value

A data frame with one row per domain: by, domain, name (department name when by = "department"), var, type, estimate, se, ci_low, ci_high, cv (percent), n (unweighted cases) and quality, of class fa_estimate (see estimate-methods).

Examples

```
library(survey)
data(api, package = "survey")
d <- svydesign(~dnum, weights = ~pw, data = apiclus1, fpc = ~fpc)
d$variables$high <- d$variables$api00 > 700
fa_estimate(d, "high", type = "prop", by = "stype")
```

---

fa_geometry *INEI boundaries for maps*

Description

Downloads the department or district boundaries published by INEI in its spatial data infrastructure (IDE INEI, layers ig_departamento and ig_distrito) and keeps them in the cache, so the download happens once. The files are large (about 9 MB and 47 MB) and are never part of the package.

Usage

```
fa_geometry(
  level = c("department", "district"),
  cache = fa_cache(),
  refresh = FALSE
)
```

Arguments

level	"department" or "district".
cache	Cache folder; the files go to <cache>/GEO.
refresh	Download again even if the file is cached.

Value

An sf object with code (two-digit department code or six-digit ubigeo), name and the geometry.

Source

https://ide.inei.gob.pe/, WFS service at https://geoespacial.inei.gob.pe/geoserver/Interoperabilidad/ow

Examples

```
## Not run:
dep <- fa_geometry("department")

## End(Not run)
```

---

fa_indicator *Compute a ready indicator*

Description

Reads the INEI files with faperu, builds the design, restricts it to the indicator's population and calls fa_estimate(). Where INEI published the figure, it is returned next to the estimate.

Usage

```
fa_indicator(id, year, by = "national", cache = fa_cache(), level = 0.95)
```

Arguments

id	Indicator id, see fa_indicators().
year	Survey year.
by	Domain passed to fa_estimate(): "national", "area", "department", ... ENA has no urban/rural split.
cache	Folder with the microdata, as in fa_cache(). Missing files are downloaded there.
level	Confidence level.

Value

The fa_estimate() table with the columns indicator, survey, year, official (the published value, NA if none) and source.

Examples

```
## Not run:
fa_indicator("pobreza_monetaria", 2024, by = "department")

## End(Not run)
```

`fa_indicators` *List the ready indicators*

Description

List the ready indicators

Usage

```
fa_indicators()
```

Details

Milk yield per cow from ENA is not included: INEI publishes no ENA figure to check it against (MIDAGRI's monthly figures come from administrative records).

Value

A data frame with the indicator `id`, the `survey`, a `label`, the years with an official figure to compare (`checked_years`) and the `source` of that figure.

Examples

```
fa_indicators()
```

`fa_link` *Join estimates from several surveys*

Description

Puts side by side estimate tables from different surveys (from `fa_estimate()` or `fa_indicator()`) for the same departments or districts. District codes are first moved to the 2017 Census frame with `fa_ubigeo()`, so a district created after 2017 lines up with the district it came from.

Usage

```
fa_link(
  ...,
  level = c("department", "district"),
  base = 2017,
  by_year = NULL,
  all = FALSE
)
```

Arguments

`...`	Estimate tables, named: the names become column suffixes, e.g. `fa_link(poverty = p, women = w)`.
`level`	`"department"` or `"district"`. Each table must have been estimated at that level.
`base`	Census year of the district frame (2017).
`by_year`	Join on year as well. By default, only when every table has a `year` column (fa_indicator() adds it); tables from fa_estimate() are joined by area alone. Set to `FALSE` to cross different years on purpose.
`all`	Keep areas that are missing in some tables (`TRUE`) or only those present in all (`FALSE`).

Details

Warns when a table has estimates with CV above 15% ("referencial" for INEI), when a district split from several 2017 districts ("multi_parent") and cannot be placed, and when several districts fall on the same 2017 district: then the estimate of the 2017 district itself is kept, because survey estimates of different areas cannot be added without their totals.

Value

A data frame with the key (`department` or `ubigeo_base`), its `name`, `year` when joined by year, and for each table the columns `estimate_<name>`, `se_<name>`, `cv_<name>`, `quality_<name>` and `n_<name>` (plus `status_<name>` for districts), of class `fa_link` (see estimate-methods).

Examples

```
a <- data.frame(by = "department", domain = c("01", "02"), estimate = c(.3, .4),
                se = .01, cv = c(3, 2.5), n = 100, quality = "muy buena", year = 2024)
b <- transform(a, estimate = c(.5, .6))
fa_link(poverty = a, other = b)
```

`fa_map` *Choropleth map of estimates*

Description

Draws estimates by department or district on the INEI boundaries (`fa_geometry()`), ready for a paper: UTM zone 18S projection, a legend by classes, department names, a north arrow, a scale bar in kilometres and, around Peru, the Pacific Ocean, the neighbouring countries and a locator inset of South America. Areas without an estimate are grey; estimates that are only a reference (INEI "referencial": CV above 15% or few cases) get an asterisk and a dashed outline.

Usage

```
fa_map(
  x,
  value = "estimate",
  level = c("department", "district"),
  key = NULL,
  base = NULL,
  percent = TRUE,
  title = NULL,
  geometry = NULL,
  cache = fa_cache(),
  breaks = "quantile",
  n = 5,
  palette = "YlOrRd",
  labels = "name",
  context = TRUE,
  legend = NULL,
  quality = "quality"
)
```

Arguments

`x`	A table with one row per area, e.g. from `fa_estimate()`, `fa_indicator()` or `fa_link()`.
`value`	Column to map.
`level`	`"department"` or `"district"`.
`key`	Column with the area code. By default `department`, `ubigeo_base`, `ubigeo` or `domain`, the first one found.
`base`	With `level = "district"`, the census year of the codes in `x` (e.g. 2017 after `fa_link()`). The district boundaries are then merged into the districts of that year; districts that split from several of them stay grey.
`percent`	Show the values in percent.
`title`	Map title.

geometry	Boundaries to use instead of downloading them, an sf object with code and name as returned by fa_geometry().
cache	Cache folder for the boundaries.
breaks	How to group the values into classes: "quantile" (each class with about the same number of areas), "pretty" (round limits), a numeric vector of limits, or "continuous" for a colour gradient without classes.
n	Number of classes.
palette	A sequential palette from grDevices::hcl.colors(), such as "YlOrRd", "Teal", "Purples" or "Viridis", light for low values and dark for high ones.
labels	"name" (department names), "code" (INEI codes) or "none". Districts are never labelled.
context	Draw the ocean, the neighbouring countries, the locator inset, the north arrow and the scale bar.
legend	Legend title.
quality	Column with the INEI quality band (quality in fa_estimate()); NULL to ignore it.

Value

A ggplot object.

Examples

```
## Not run:
p <- fa_indicator("pobreza_monetaria", 2024, by = "department")
fa_map(p, title = "Monetary poverty 2024", legend = "Poor people")
fa_map(p, breaks = c(0, .1, .2, .3, .4, .5), palette = "Purples")

## End(Not run)
```

---

fa_quality *INEI quality band for a coefficient of variation*

Description

INEI rates the precision of a survey estimate by its coefficient of variation (CV): below 5% is very good, 5% to 10% good, 10% to 15% acceptable, and above 15% the estimate is only a reference ("referencial") and should be read with caution.

Usage

```
fa_quality(cv)
```

Arguments

cv	Coefficient of variation in percent.

Value

A character vector: "muy buena", "buena", "aceptable" or "referencial", NA where cv is missing.

References

INEI, notes on precision of the estimates in its publications from ENAHO and ENDES (e.g. "Peru: Evolucion de los Indicadores de Empleo e Ingreso por Departamento", Lib. 1690, chap. 9).

Examples

```
fa_quality(c(3, 7.5, 12, 20))
```

---

fa_read *Read a module*

Description

Downloads the module if needed (see fa_download()), unzips it next to the zip file and reads one data file. SPSS and Stata files keep their variable and value labels (through haven). CSV files from INEI come in Latin-1 and are converted to UTF-8; the separator (, , or ; with decimal commas as in the ENAHO 2025) is detected, and code columns (UBIGEO, CCDD, CCPP, CCDI, IRUC, NSEGM) are kept as text so no leading zero is lost. DBF files (ESCALE before 2017 and the INEI DBF format) are read with the foreign package.

Usage

```
fa_read(
  survey,
  year,
  module,
  file = NULL,
  format = "auto",
  series = NULL,
  cache = fa_cache(),
  lower = TRUE,
  period = NULL,
  cols = NULL
)
```

Arguments

survey	Short name, for example "enaho". See fa_surveys().
year	One year.
module	One module code.
file	Part of the file name to read, when the module has several data files (e.g. "Sumaria-2024"). A regular expression.

format	"auto" uses the format of the verified copy, or else the first of STATA, SPSS, CSV and DBF that the module has. Stata comes first because it is read exactly, with labels, and quickly; INEI's CSV files can have line breaks inside text answers. You can also ask for one of those formats.
series	For sources with several series, e.g. "panel" in ENAHO.
cache	Cache folder, see fa_cache().
lower	Turn variable names to lower case, so SPSS and Stata files look the same.
period	For sources with several files per year, the period shown by fa_catalog(), e.g. "Anual - (Ene-Dic)" in EPEN. A regular expression.
cols	Variables to read, in any case. NULL reads them all.

Details

With cols only those variables are read. The module is still downloaded whole, because INEI serves nothing smaller, but reading a few columns of a wide file is several times faster and uses a fraction of the memory. fa_variables() lists the names. CSV files are read with data.table when it is installed and with base R otherwise; options(faperu.csv_engine = "base") forces base R.

For the 2017 Census use fa_redatam().

Value

A data frame (a tibble for SPSS and Stata files).

Examples

```
## Not run:
s <- fa_read("enaho", 2024, "34", file = "Sumaria-2024\\.")
s <- fa_read("enaho", 2024, "34", file = "Sumaria-2024\\.",
             cols = c("conglome", "vivienda", "hogar", "pobreza", "factor07"))

## End(Not run)
```

---

fa_redatam *Query the 2017 Census in REDATAM format*

Description

INEI does not publish the 2017 Census microdata as a download, only a web REDATAM query and tables. CELADE (ECLAC) distributes the full census database in REDATAM format at redatam.org. fa_redatam() downloads it once into the cache (678 MB, SHA-256 checked), unzips it and runs a REDATAM query on it with the redatamx package.

Usage

```
fa_redatam(query = NULL, cache = fa_cache())
```

Arguments

query	A REDATAM query. NULL lists the variables of every entity instead.
cache	Cache folder, see fa_cache().

Details

redatamx is only suggested: it downloads the REDATAM engine when it is installed. Install it with `install.packages("redatamx")`.

Queries use REDATAM syntax, for example `"freq PERSONA.C5P2"` (sex) or `"freq DEPARTAM.CCDD by PERSONA.C5P2"`. The entities are DEPARTAM, PROVINCI, DISTRITO, VIVIENDA, HOGAR and PERSONA.

Value

A data frame: the table returned by `redatamx::redatam_query()`, or with `query = NULL` one row per variable with its `entity`.

Examples

```
## Not run:
fa_redatam("freq PERSONA.C5P2")   # 14 450 757 men and 14 931 127 women
v <- fa_redatam()                  # what is there

## End(Not run)
```

---

fa_surveys *The ten data sources covered by faperu*

Description

The ten data sources covered by faperu

Usage

```
fa_surveys()
```

Value

A data frame with one row per source, of class `fa_surveys` (see faperu-methods).

Examples

```
fa_surveys()
```

`fa_table` *Table of estimates for a paper*

Description

Formats an estimate table (`fa_estimate()` or `fa_indicator()`) the way it goes in a paper: one row per area with the estimate, its confidence interval, the CV and the number of cases. Estimates with CV above 15% (INEI "referencial") are marked with an asterisk and a note.

Usage

```
fa_table(
  x,
  digits = 1,
  percent = TRUE,
  format = c("data.frame", "markdown", "latex"),
  caption = NULL
)
```

Arguments

`x`	An estimate table.
`digits`	Decimals for the estimate and the interval.
`percent`	Show proportions in percent.
`format`	`"data.frame"`, `"markdown"` or `"latex"`.
`caption`	Caption for markdown and LaTeX.

Value

A data frame, or a character vector with the lines of the table (print it with `cat(..., sep = "\\n")`).

Examples

```
x <- data.frame(name = c("Lima", "Ica"), estimate = c(.27, .06), se = c(.012, .009),
                ci_low = c(.246, .042), ci_high = c(.294, .078), cv = c(4.5, 15.6),
                n = c(5000, 900))
fa_table(x)
cat(fa_table(x, format = "markdown"), sep = "\n")
```

`fa_ubigeo` *Standard and harmonised ubigeo columns*

Description

Adds a six-digit `ubigeo` with its `department`, `province` and `department_name`. It takes an existing ubigeo column, or builds one from INEI's separate codes (`CCDD`, `CCPP`, `CCDI`), as in the ENA. Codes are always text padded to six digits, because INEI's CSV files drop the leading zero (`10101` is `010101`).

Usage

```
fa_ubigeo(data, var = NULL, year = NULL, base = NULL)
```

Arguments

`data`	A data frame.
`var`	Name of the ubigeo column. If `NULL`, it looks for `ubigeo` and then for `ccdd`, `ccpp` and `ccdi` (any case).
`year`	Year the data were collected. Needed for `base`, and used to flag codes that did not exist that year. Years outside 2004-2025 are moved to the nearest end of the table, with a warning.
`base`	Year whose coding you want, usually `2017`.

Details

The codes are INEI's. RENIEC and SUNAT number districts differently (INEI puts Callao in `07`, RENIEC in `24`), so do not mix them.

With `year` and `base`, it also translates the codes from the coding used in `year` to the coding used in `base`. The natural base is `2017`: the 1874 districts of the 2017 Census, the frame of every INEI survey since 2018. The translation follows, and only follows:

- the recodings INEI made when it created the provinces of Datem del Maranon (2006) and Putumayo (2015), and the Bagua and La Peca swap of 2007, as seen in RENAMU 2004-2025;
- for districts created after 2017, the 2017 district that holds their schools in MINEDU's 2025 school register. This crosswalk has not been checked law by law yet.

Nothing else is guessed. Districts created between 2005 and 2017 have no known parent, and Lambras and Santa Lucia have two, so they get `NA`. Sources and method are in `system.file("ubigeo", "README.md", package = "faperu")`.

`ubigeo_status` is one of:

- `"same"`: the code means the same district in both years.
- `"recoded"`: same district, new code.
- `"recoded_partial"`: same name, new code, but the new province also got new districts at the time, so the old district may have been larger.

- "area_changed": same code, but a district was carved out of it between the two years (from the 2017 crosswalk).
- "to_parent": a district created after base, given the code of the district it came from.
- "multi_parent": created after base out of two districts; NA.
- "created_after_base": created after base, parent unknown; NA.
- "legacy": an old code INEI surveys used that matches no district (ENAHO's 120699); NA.
- "unknown": not a district in year; NA.

Value

data with the new columns. With base, also ubigeo_base and ubigeo_status.

Examples

```
fa_ubigeo(data.frame(UBIGEO = c("210101", "150101")))
fa_ubigeo(data.frame(CCDD = "21", CCPP = "01", CCDI = "01"))
fa_ubigeo(data.frame(ubigeo = c("130112", "160109")), year = 2024, base = 2017)
```

---

fa_variables *Variables of a module*

Description

Lists the variables of the data files in a module, with their labels when the file has them (SPSS and Stata). Only the file headers are read, so it is quick even for large modules. Use it to choose cols in fa_read().

Usage

```
fa_variables(
  survey,
  year,
  module,
  file = NULL,
  format = "auto",
  series = NULL,
  cache = fa_cache(),
  period = NULL
)
```

Arguments

survey	Short name, for example "enaho". See fa_surveys().
year	One year.
module	One module code.

file	Part of the file name to read, when the module has several data files (e.g. "Sumaria-2024"). A regular expression.
format	"auto" uses the format of the verified copy, or else the first of STATA, SPSS, CSV and DBF that the module has. Stata comes first because it is read exactly, with labels, and quickly; INEI's CSV files can have line breaks inside text answers. You can also ask for one of those formats.
series	For sources with several series, e.g. "panel" in ENAHO.
cache	Cache folder, see fa_cache().
period	For sources with several files per year, the period shown by fa_catalog(), e.g. "Anual - (Ene-Dic)" in EPEN. A regular expression.

Value

A data frame with file, name, label and type, one row per variable and file.

Examples

```
## Not run:
fa_variables("enaho", 2024, "34")

## End(Not run)
```

---

fa_years *Published years of a source*

Description

Published years of a source

Usage

```
fa_years(survey)
```

Arguments

survey	Short name, for example "enaho". See fa_surveys().

Value

An integer vector.

Examples

```
fa_years("ena")
```

faperu-methods *Print and summarise faperu tables*

Description

`fa_surveys()`, `fa_catalog()`, `fa_download()` and `fa_check_updates()` return data frames with an extra class, so they print as a short report. They are still data frames: subset, filter or save them as usual, or use `as.data.frame()` to drop the class.

Usage

```
## S3 method for class 'fa_surveys'
print(x, ...)

## S3 method for class 'fa_catalog'
print(x, n = 10L, ...)

## S3 method for class 'fa_catalog'
summary(object, ...)

## S3 method for class 'fa_download'
print(x, ...)

## S3 method for class 'fa_updates'
print(x, ...)
```

Arguments

`x`, `object`	A table returned by one of those functions.
`...`	Passed on to `print.data.frame()`.
`n`	Rows to show.

Value

`print()` returns `x` invisibly. `summary()` of a catalog returns a data frame with one row per year: modules, series and formats.

Examples

```
fa_surveys()
summary(fa_catalog("enut"))
```

Index

as.data.frame(), *2*, *23*

estimate-methods, 2, *10*, *13*

fa_cache, 3
fa_cache(), *8*, *11*, *17*, *18*, *22*
fa_catalog, 4
fa_catalog(), *8*, *17*, *22*, *23*
fa_check_updates, 5
fa_check_updates(), *23*
fa_design, 6
fa_design(), *9*
fa_districts, 7
fa_download, 8
fa_download(), *4*, *16*, *23*
fa_estimate, 9
fa_estimate(), *2*, *3*, *11*, *12*, *14*, *15*, *19*
fa_geometry, 10
fa_geometry(), *14*, *15*
fa_indicator, 11
fa_indicator(), *2*, *3*, *12*, *14*, *19*
fa_indicators, 12
fa_indicators(), *11*
fa_link, 12
fa_link(), *2*, *3*, *14*
fa_map, 14
fa_quality, 15
fa_quality(), *9*
fa_read, 16
fa_read(), *6*, *21*
fa_redatam, 17
fa_redatam(), *17*
fa_surveys, 18
fa_surveys(), *4*, *6*, *8*, *16*, *21–23*
fa_table, 19
fa_table(), *2*
fa_ubigeo, 20
fa_ubigeo(), *12*
fa_variables, 21
fa_variables(), *17*
fa_years, 22
faperu-methods, *4*, *5*, *9*, *18*, 23

graphics::plot(), *3*
grDevices::hcl.colors(), *15*

plot.fa_estimate *(*estimate-methods*)*, 2
print.data.frame(), *3*, *23*
print.fa_catalog *(*faperu-methods*)*, 23
print.fa_download *(*faperu-methods*)*, 23
print.fa_estimate *(*estimate-methods*)*, 2
print.fa_link *(*estimate-methods*)*, 2
print.fa_surveys *(*faperu-methods*)*, 23
print.fa_updates *(*faperu-methods*)*, 23

summary.fa_catalog *(*faperu-methods*)*, 23
summary.fa_estimate *(*estimate-methods*)*, 2
survey::subset(), *7*
survey::svydesign(), *6*

tools::R_user_dir(), *3*