

Package: ocupacoesBR (via r-universe)

October 11, 2026

Title Traduz a Ocupação Declarada ao TSE em Classificações Padronizadas

Version 0.8.1

Description Traduz a ocupação registrada em fontes brasileiras para a Classificação Internacional Uniforme de Ocupações (ISCO-88 e ISCO-08) e, a partir dela, para medidas padronizadas de posição social: o índice socioeconômico ISEI (Ganzeboom e Treiman, 1996 <[doi:10.1006/ssre.1996.0010](https://doi.org/10.1006/ssre.1996.0010)>), o prestígio SIOPS (Treiman, 1977, ISBN:0-12-698750-5), o esquema de classes Erikson-Goldthorpe-Portocarero (EGP; Erikson, Goldthorpe e Portocarero, 1979 <[doi:10.2307/589632](https://doi.org/10.2307/589632)>) e um esquema de classes e estratos desenhado para o dado eleitoral. Traz ainda o ISEI-BR, uma régua estimada pelo pacote: o procedimento de Ganzeboom, De Graaf e Treiman (1992) <[doi:10.1016/0049-089X\(92\)90017-B](https://doi.org/10.1016/0049-089X(92)90017-B)> feito sobre a PNAD Contínua, em vez de importado de dado estrangeiro. As decisões de medida estão justificadas em Peixoto (2026) <[doi:10.31235/osf.io/b29kg_v1](https://doi.org/10.31235/osf.io/b29kg_v1)>. Há duas portas de entrada: o código de ocupação declarado nas candidaturas ao Tribunal Superior Eleitoral (TSE) e a Classificação Brasileira de Ocupações (CBO-2002 e CBO-94), usada na RAIS, no CAGED e no eSocial. As tabelas de conversão derivam das sintaxes publicadas do International Stratification and Mobility File (Ganzeboom e Treiman) e da tábua oficial de conversão do Ministério do Trabalho, e são geradas por script a partir dos arquivos originais, nunca transcritas manualmente. O dicionário que leva o código do TSE à ISCO-88 é, ao lado do ISEI-BR, a outra peça autoral: foi construído pelo autor, código a código, e é validado contra patrimônio e escolaridade das próprias candidaturas.

License MIT + file LICENSE

Encoding UTF-8

Language pt-BR

Roxygen list(markdown = TRUE)

Depends R (>= 4.1)
Imports stats, utils
Suggests testthat (>= 3.0.0), knitr, rmarkdown, readxl
VignetteBuilder knitr
Config/testthat/edition 3
URL <https://github.com/moraespeixoto/ocupacoesBR>,
<https://moraespeixoto.github.io/ocupacoesBR/>
BugReports <https://github.com/moraespeixoto/ocupacoesBR/issues>
LazyData true
Config/Needs/website dplyr
Config/roxygen2/version 8.1.0
Repository <https://moraespeixoto.r-universe.dev>
Date/Publication 2026-09-17 14:58:26 UTC
RemoteUrl <https://github.com/moraespeixoto/ocupacoesbr>
RemoteRef HEAD
RemoteSha b884f6bd175443161659377f8bef3bd045d8c0e6

Contents

cbo2002_concordancia	4
cbo2002_escada	5
cbo2002_familia_isco88	6
cbo2002_isco88	7
cbo2002_para_egp	8
cbo2002_para_isco	9
cbo2002_para_isco08	11
cbo2002_para_isei	12
cbo2002_para_isei_br	13
cbo2002_para_isei08	14
cbo2002_para_prestigio	15
cbo2002_para_prestigio08	15
cbo2002_para_siops	16
cbo2002_para_siops08	17
cbo94_isco88	18
cbo94_para_egp	19
cbo94_para_isco	20
cbo94_para_isco08	20
cbo94_para_isei	21
cbo94_para_prestigio	21
cbo94_para_siops	22
checa_cobertura	22
checa_cobertura_cbo2002	23

checa_cobertura_cbo94	24
checa_cobertura_cod	24
checa_perodo	25
cod_isco08	26
cod_para_egp	27
cod_para_isco	28
cod_para_isco08	29
cod_para_isei	30
cod_para_isei_br	30
cod_para_isei08	31
cod_para_prestigio	32
cod_para_prestigio08	32
cod_populacao_br	33
crosswalk_cbo2002	35
crosswalk_cbo94	36
crosswalk_cod	36
crosswalk_tse	37
isco_posicao_br	38
isco08_isco88	41
isco08_isei_br	42
isco08_medidas	44
isco08_para_isco88	45
isco08_para_isei_br	46
isco88_isco08	46
isco88_medidas	47
isco88_para_egp	48
isco88_para_isco08	50
isco88_para_isei	50
isco88_para_prestigio	51
isco88_para_siops	52
isei_retrospectivo	52
reguas	54
tse_autorrotulo_patrimonio	55
tse_codigos_autorrotulo	57
tse_diff_cadastro	58
tse_dispersao_patrimonio	59
tse_isco	61
tse_ocupacao_rotulos	62
tse_para_classe	63
tse_para_componente_alta	65
tse_para_egp	66
tse_para_estrato	68
tse_para_isco	69
tse_para_isco08	70
tse_para_isei	71
tse_para_isei_br	73
tse_para_isei08	74
tse_para_politico	75

tse_para_prestigio	76
tse_para_prestigio08	77
tse_para_rotulo	78
tse_para_siops	78
tse_para_siops08	79
tse_quebra_2002	80
tse_rotulo_para_cod	81
tse_universo	82
tse_validacao	84
tse_vigencia	86

Index	87
--------------	-----------

cbo2002_concordancia *Quao homogenea e a familia da CBO-2002?*

Description

Devolve a proporcao das ocupacoes da familia que compartilham o ISCO majoritario. Vale 1 quando a familia inteira cai num unico ISCO; abaixo disso, traduzir pela familia mistura posicoes distintas.

Usage

```
cbo2002_concordancia(cbo)
```

Arguments

cbo	Vetor de codigos da CBO-2002 (familia ou ocupacao; a ocupacao e reduzida a sua familia).
-----	--

Value

data.frame com a familia, o ISCO majoritario, quantas ocupacoes foram vistas, quantas a familia tem de fato, a cobertura, quantos ISCO distintos aparecem, a concordancia (ou NA) e a marca de empate.

Concordancia so e afirmavel sobre familia vista inteira

A tabua do MTE cobre metade da CBO-2002, e a concordancia era calculada sobre as ocupacoes **vistas**, nao sobre a familia real. O efeito: 131 familias tinham UMA unica ocupacao na tabua e reportavam concordancia = 1 — o valor que sinaliza "sem ambiguidade nenhuma" — e em 87 delas a familia de fato tem mais de uma ocupacao. Era certeza maxima onde a evidencia era minima.

Agora concordancia so recebe valor quando cobertura_familia vale 1, isto e, quando a familia foi vista por inteiro (175 das 436). Nas demais fica NA, e a proporcao entre as ocupacoes vistas segue disponivel em concordancia_vista, sem posar de diagnostico. O denominador vem do dominio oficial da CBO-2002 no Novo CAGED, 2.777 ocupacoes.

Examples

```
cbo2002_concordancia(c("5211", "4121"))
cbo2002_concordancia("3171") # vista em 1/4: concordancia NA
```

cbo2002_escada	<i>Escada hierárquica da CBO-2002 para quando a ocupação não está na tábua</i>
----------------	--

Description

Um degrau por prefixo de CBO, para subir do código de seis dígitos até um nível em que a fonte permita afirmar um ISCO. Usada por [cbo2002_para_isco\(\)](#) com `escada = TRUE`.

Usage

```
cbo2002_escada
```

Format

`data.frame` com as colunas:

prefixo prefixo da CBO-2002, de 2 a 4 dígitos.

nível quantos dígitos tem o prefixo: 4 = família, 3 = subgrupo, 2 = subgrupo principal.

isco88 ISCO-88 comum às ocupações mapeadas sob o prefixo.

n_base quantas ocupações mapeadas sustentam o degrau.

Como o degrau é apurado

Onde as ocupações mapeadas sob o prefixo não concordam num único ISCO, usa-se o **ancestral comum** delas na hierarquia da ISCO — a forma arredondada que o próprio ISMF publica: 2211 e 2212 viram 2210; 2210 e 2230 viram 2200. Só entram degraus cujo destino exista em [isco88_medidas](#), porque um ISCO sem ISEI não serve de nada. Não há destino inventado: sobe-se até onde a fonte permite afirmar, e não além.

Por que para em dois dígitos

Descer a um dígito fecharia boa parte das 283 ocupações que sobram sem rota, e cometeria exatamente a armadilha que este pacote existe para impedir: o grande grupo 9 da CBO é reparação e manutenção (ISEI ~34) e o da ISCO é o das ocupações elementares (ISEI 16 a 30). Uma cobertura maior comprada com inversão de classe não é cobertura.

Source

Agregado de [cbo2002_isco88](#) pela hierarquia da própria CBO.

Examples

```
head(cbo2002_escada)

# A escada tem três degraus – prefixo de quatro, três e dois dígitos – e é
# tentada, nessa ordem, só sobre o que sobrou NA com `escada = TRUE`.
table(cbo2002_escada$nivel)
```

cbo2002_familia_isco88

Correspondência da CBO-2002 com a ISCO-88, por família

Description

Para o microdado que publica apenas a família de quatro dígitos.

Usage

```
cbo2002_familia_isco88
```

Format

data.frame com as colunas:

familia código de quatro dígitos.

isco88 ISCO-88 **majoritário** entre as ocupações da família.

n_ocupacoes quantas ocupações da família a tábua do MTE cobre.

n_ocupacoes_cbo quantas ocupações a família tem **de fato**, pelo domínio oficial da CBO-2002.

cobertura_familia $n_ocupacoes / n_ocupacoes_cbo$. Vale 1 quando a família foi vista por inteiro — o caso de 175 das 436.

n_isco_distintos quantos ISCO diferentes aparecem na família.

concordancia proporção das ocupações que caem no ISCO majoritário, **ou NA quando** **cobertura_familia** < 1. Vale 1 quando a família é homogênea; abaixo disso, traduzir pela família mistura posições distintas.

concordancia_vista a mesma proporção calculada só sobre as ocupações vistas. Sempre preenchida, mas não é diagnóstico da família.

empate TRUE quando a moda **não é maioria** — o ISCO escolhido só venceu por ordenação. Tratar essas famílias como uma posição única é arbitrário, e concordancia sozinha não distingue "dividida" de "cara ou coroa".

Por que há duas concordâncias

A tábua do MTE cobre metade da CBO-2002, e a concordância era apurada sobre as ocupações **vistas**, não sobre a família real. O resultado é que 131 famílias tinham uma única ocupação na tábua e reportavam concordancia = 1 — o valor que sinaliza ausência total de ambiguidade — e em **87** delas a família de fato tem mais de uma ocupação. Era certeza máxima onde a evidência era mínima, e o README chegava a afirmar que 1 significava "a família inteira cai num único ISCO".

A separação resolve sem descartar informação: concordancia só é afirmada sobre família vista por inteiro; concordancia_vista guarda a proporção bruta para quem quiser inspecioná-la sabendo o que ela é.

Source

Agregado de [cbo2002_isco88](#); o denominador vem do domínio oficial da CBO-2002 na aba cbo2002ocupação do layout do Novo CAGED (2.777 ocupações), distribuído em inst/extdata/fontes/cbo2002_dominio.txt.

Examples

```
# As famílias em que a tábua empata: nelas, o argumento `empate` de
# cbo2002_para_isco() é quem decide entre NA e o ISCO majoritário.
cbo2002_familia_isco88[cbo2002_familia_isco88$empate,
                        c("familia", "isco88", "n_ocupacoes",
                          "n_isco_distintos")]
```

cbo2002_isco88

Correspondência da CBO-2002 com a ISCO-88, por ocupação

Description

Correspondência da CBO-2002 com a ISCO-88, por ocupação

Usage

```
cbo2002_isco88
```

Format

data.frame com as colunas:

cbo2002 código de seis dígitos, sem hífen, como nos microdados.

titulo título da ocupação na CBO-2002.

familia os quatro primeiros dígitos.

cbo94 código correspondente na CBO-94, sem pontuação.

isco88 código ISCO-88 de quatro dígitos.

Cobertura

A tábua do Ministério do Trabalho é uma conversão entre a CBO-2002 e a CBO-94, de modo que cobre apenas as ocupações presentes nas duas classificações. As ocupações criadas na revisão de 2002 e o grande grupo 0 (forças armadas) não têm correspondência oficial com a CIUO-88 e não aparecem aqui.

Source

Tábua de conversão CBO2002–CBO94–CIUO88 do Ministério do Trabalho, consultada família a família em http://www.mtecbo.gov.br/cbosite/pages/tabua/FiltroConversao_CB02002_CB094_CIU088.jsf

Examples

```
# A tradução ocupação a ocupação, o caminho mais curto e mais confiável:
cbo2002_isco88[cbo2002_isco88$cbo2002 == "252105", ]

# Quantas ocupações a tábua do MTE cobre, e quantos ISCO distintos alcança:
nrow(cbo2002_isco88)
length(unique(stats::na.omit(cbo2002_isco88$isco88)))
```

cbo2002_para_egp

Classe EGP a partir da CBO-2002

Description

Depende da fonte, e a diferença importa. A RAIS, o CAGED e o eSocial são registros de VINCULO EMPREGATÍCIO: toda linha é um empregado. Não há conta própria nem empregador como observação, e não existe campo com o número de subordinados. Por essas portas `conta_propria` é FALSE por construção, e o EGP sai sem IVa, IVb e V — não por omissão de quem chama, mas porque a fonte não descreve quem está fora do assalariamento. Ler a distribuição resultante como se fosse a estrutura de classes do país e o erro a evitar aqui.

Usage

```
cbo2002_para_egp(
  cbo,
  conta_propria = NULL,
  n_supervisionados = NULL,
  n_classes = 11,
  rotulo = TRUE,
  avisar = TRUE,
  empate = c("na", "moda"),
  escada = FALSE
)
```

Arguments

cbo	Vetor de codigos da CBO-2002, com 6 digitos (ocupacao) ou 4 (familia). Aceita "1111-05", "111105" ou 111105.
conta_propria	Vetor lógico opcional: a pessoa trabalha por conta própria ou é empregadora (TRUE) ou é empregada (FALSE).
n_supervisionados	Vetor numérico opcional com o número de pessoas supervisionadas.
n_classes	Número de classes do resultado: 11 (padrão), 7, 5 ou 3.
rotulo	Se TRUE (padrão), devolve rótulos — em todas as versões, inclusive nas colapsadas; se FALSE, os códigos inteiros.
avisar	Se TRUE (padrão), avisa quando conta_propria ou n_supervisionados não são fornecidos.
empate	O que fazer quando a familia nao tem ISCO majoritario: "na" (padrao) devolve NA; "moda" devolve o vencedor do desempate por menor codigo, como nas versoes anteriores.
escada	Se TRUE, sobe a hierarquia da CBO quando a ocupacao de seis digitos nao esta na tabua: tenta a familia de 4, depois o subgrupo de 3, depois o subgrupo principal de 2. Leva a cobertura do dominio oficial de 49,8% para 89,8% . Padrao FALSE, porque o resultado deixa de ser a ocupacao declarada e passa a ser o seu grupo. Use crosswalk_cbo2002() para ver nivel_usado caso a caso.

Details

Em pesquisa domiciliar que codifique em CBO, os dois campos podem existir; quando existirem, passe-os.

Value

Vetor de texto (ou inteiro, se rotulo = FALSE).

Examples

```
cbo2002_para_egp("521110", avisar = FALSE)
cbo2002_para_egp("521110", conta_propria = TRUE, n_supervisionados = 0)
```

cbo2002_para_isco	<i>Traduz a CBO-2002 em ISCO-88</i>
-------------------	-------------------------------------

Description

Converte o codigo da Classificacao Brasileira de Ocupacoes de 2002 — o que aparece na RAIS, no CAGED e no eSocial — no codigo ISCO-88 de quatro digitos, porta de entrada para o ISEI, o SIOPS e o EGP.

Usage

```
cbo2002_para_isco(cbo, empate = c("na", "moda"), escada = FALSE)
```

Arguments

cbo	Vetor de codigos da CBO-2002, com 6 digitos (ocupacao) ou 4 (familia). Aceita "1111-05", "111105" ou 111105.
empate	O que fazer quando a familia nao tem ISCO majoritario: "na" (padrao) devolve NA; "moda" devolve o vencedor do desempate por menor codigo, como nas versoes anteriores.
escada	Se TRUE, sobe a hierarquia da CBO quando a ocupacao de seis digitos nao esta na tabua: tenta a familia de 4, depois o subgrupo de 3, depois o subgrupo principal de 2. Leva a cobertura do dominio oficial de 49,8% para 89,8% . Padrao FALSE, porque o resultado deixa de ser a ocupacao declarada e passa a ser o seu grupo. Use crosswalk_cbo2002() para ver nivel_usado caso a caso.

Value

Vetor de texto com o ISCO-88 de quatro digitos, ou NA.

De onde vem a correspondencia

Da tabua oficial de conversao CBO2002–CBO94–CIUO88 do Ministerio do Trabalho, consultada familia por familia. Nao e uma equivalencia construida por nos.

O que a fonte nao cobre

A tabua e, por construcao, uma conversao entre a CBO-2002 e a CBO-94: so inclui as ocupacoes que existem nas duas. As ocupacoes criadas na revisao de 2002 — e o grande grupo 0, das forcas armadas — nao tem correspondencia oficial com a CIUO-88 e voltam NA. Isso e limite da fonte, e a funcao prefere devolver NA a inventar um destino plausivel.

Familia de quatro digitos

Muito microdado publica so a familia. Nesse caso o pacote devolve o ISCO **majoritario** entre as ocupacoes da familia. Use [cbo2002_concordancia\(\)](#) para saber se a familia e homogenea antes de confiar no valor: uma familia dividida ao meio nao deve ser tratada como uma posicao unica.

A escada hierarquica

A tabua do MTE cobre metade da CBO-2002, e o buraco NAO e aleatorio: cai sobre o que foi criado na revisao de 2002 — tecnologia da informacao em peso, pesquisadores — e sobre todo o grande grupo 0. Um vinculo de RAIS nesses codigos volta NA.

Com escada = TRUE, a funcao sobe a hierarquia da propria CBO ate achar um nivel com correspondencia. Onde as ocupacoes mapeadas sob aquele prefixo nao concordam num unico ISCO, usa-se o ancestral comum delas na hierarquia da ISCO — que e a forma arredondada que o proprio ISMF publica (2211 e 2212 viram 2210). Nao ha destino inventado: so se sobe ate onde a fonte permite afirmar. Cobertura do dominio oficial (2.777 ocupacoes):

nível usado	ocupacoes	acumulado
6 (a ocupacao)	1.384	49,8%
4 (familia)	760	77,2%
3 (subgrupo)	306	88,2%
2 (subgrupo principal)	44	89,8%
sem rota	283	—

A escada para em dois digitos de proposito. Descer a um fecharia o buraco restante e cometeria exatamente a armadilha que este pacote existe para impedir: o grande grupo 9 da CBO e reparacao e manutencao, o da ISCO e ocupacoes elementares. As 283 ocupacoes sem rota — em peso o bloco de diretores 12xx e as forcas armadas — nao tem correspondencia oficial nenhuma, e continuam NA.

A escada e a consulta por familia aplicam regras diferentes. Entrar com quatro digitos consulta `cbo2002_familia_isco88`, que traz a MODA das ocupacoes da familia. Entrar com seis e escada = TRUE exige ancestral comum na ISCO, e devolve NA quando ele nao existe acima do primeiro digito. Em oito familias as duas respostas divergem — `cbo2002_para_isco("1423")` da "2419", e `cbo2002_para_isco("142399", escada = TRUE)` da NA —, porque as ocupacoes de 1423 se espalham por 1233, 1234, 1239 e 2419. A escada e a mais conservadora das duas, e a divergencia e deliberada; quem alterna entre as duas entradas precisa saber que elas nao respondem a mesma pergunta. `tests/testthat/test-cbo.R` trava a lista das oito.

Familias empatadas

Em 19 familias a moda NAO e maioria: duas ocupacoes dividem o topo. O desempate anterior era mudo e ia sempre para o **menor** codigo ISCO — consequencia da ordenacao lexicografica de `table()` —, e como a hierarquia da ISCO e ordenada por status, o menor codigo tem o maior ISEI em 16 dos 19 casos. Isso e vies, nao ruido: nao desaparece com N.

Por isso `empate = "na"` e o padrao. Um empate 1:1 nao tem destino majoritario, e devolver NA e mais honesto que escolher por ordem alfabetica. Use `empate = "moda"` para reproduzir o comportamento antigo.

See Also

[cbo2002_para_isei\(\)](#), [cbo2002_concordancia\(\)](#), [cbo94_para_isco\(\)](#)

Examples

```
cbo2002_para_isco(c("1111-05", "225120", "5211"))
```

`cbo2002_para_isco08` *ISCO-08 a partir da CBO-2002*

Description

ISCO-08 a partir da CBO-2002

Usage

```
cbo2002_para_isco08(cbo, com_ambiguidade = FALSE)
```

Arguments

cbo Vetor de codigos da CBO-2002, com 6 digitos (ocupacao) ou 4 (familia). Aceita "1111-05", "111105" ou 111105.

com_ambiguidade Se TRUE, devolve um data.frame com o código e o número de alternativas da OIT em vez de só o código.

Value

Vetor de texto, ou um data.frame se com_ambiguidade = TRUE.

Examples

```
cbo2002_para_isco08("225120")
```

cbo2002_para_isei	<i>Indice socioeconomico ISEI a partir da CBO-2002</i>
-------------------	--

Description

Indice socioeconomico ISEI a partir da CBO-2002

Usage

```
cbo2002_para_isei(cbo, empate = c("na", "moda"), escada = FALSE)
```

Arguments

cbo Vetor de codigos da CBO-2002, com 6 digitos (ocupacao) ou 4 (familia). Aceita "1111-05", "111105" ou 111105.

empate O que fazer quando a familia nao tem ISCO majoritario: "na" (padrao) devolve NA; "moda" devolve o vencedor do desempate por menor codigo, como nas versoes anteriores.

escada Se TRUE, sobe a hierarquia da CBO quando a ocupacao de seis digitos nao esta na tabua: tenta a familia de 4, depois o subgrupo de 3, depois o subgrupo principal de 2. Leva a cobertura do dominio oficial de **49,8% para 89,8%**. Padrao FALSE, porque o resultado deixa de ser a ocupacao declarada e passa a ser o seu grupo. Use [crosswalk_cbo2002\(\)](#) para ver nivel_usado caso a caso.

Value

Vetor numerico com o escore ISEI-88.

Examples

```
cbo2002_para_isei(c("1111-05", "225120"))
```

cbo2002_para_isei_br *ISEI-BR a partir da CBO-2002*

Description

Versao para o microdado do trabalho a partir de 2003, RAIS, CAGED e eSocial. Leia [tse_para_isei_br\(\)](#) para o que a regua e e o que ela nao e.

Usage

```
cbo2002_para_isei_br(cbo, empate = c("na", "moda"), escada = FALSE)
```

Arguments

cbo	Vetor de codigos da CBO-2002, com 6 digitos (ocupacao) ou 4 (familia). Aceita "1111-05", "111105" ou 111105.
empate	O que fazer quando a familia nao tem ISCO majoritario: "na" (padrao) devolve NA; "moda" devolve o vencedor do desempate por menor codigo, como nas versoes anteriores.
escada	Se TRUE, sobe a hierarquia da CBO quando a ocupacao de seis digitos nao esta na tabua: tenta a familia de 4, depois o subgrupo de 3, depois o subgrupo principal de 2. Leva a cobertura do dominio oficial de 49,8% para 89,8% . Padrao FALSE, porque o resultado deixa de ser a ocupacao declarada e passa a ser o seu grupo. Use crosswalk_cbo2002() para ver nivel_usado caso a caso.

Value

Vetor numerico com o escore ISEI-BR, entre 10 e 90.

Duas pontes empilhadas

O caminho e CBO-2002 -> ISCO-88 (tabua do Ministerio do Trabalho) -> ISCO-08 (ponte da OIT) -> ISEI-BR, e a segunda ponte e ambigua em cerca de um terco dos codigos de origem. A ressalva e a mesma de [cbo2002_para_isei08\(\)](#) e [cbo2002_para_prestigio08\(\)](#), que ja a traziam; esta porta nao a trazia ate 05/09/2026.

Vale ainda lembrar a data. O ISEI-BR e uma regua estimada em 2025 e que a propria documentacao declara nao formar serie. Foi esse o argumento que retirou a porta pela CBO-94 na versao 0.5.1 — e o que se retirou ali foi o ATALHO, nao o caminho: `isco08_para_isei_br(cbo94_para_isco08(x))` continua funcionando, e quem escreve as duas etapas fez a escolha conscientemente, que e o que a recusa pretende garantir. O argumento nao se aplica aqui com a mesma forza — a CBO-2002 esta em uso hoje —, mas se aplica ao inicio de uma serie de RAIS que comece em 2003.

See Also

[tse_para_isei_br\(\)](#), [isco08_isei_br](#)

Examples

```
cbo2002_para_isei_br("225120")
```

cbo2002_para_isei08	<i>ISEI-08 a partir da CBO-2002</i>
---------------------	-------------------------------------

Description

ISEI-08 a partir da CBO-2002

Usage

```
cbo2002_para_isei08(cbo, empate = c("na", "moda"), escada = FALSE)
```

Arguments

cbo	Vetor de codigos da CBO-2002, com 6 digitos (ocupacao) ou 4 (familia). Aceita "1111-05", "111105" ou 111105.
empate	O que fazer quando a familia nao tem ISCO majoritario: "na" (padrao) devolve NA; "moda" devolve o vencedor do desempate por menor codigo, como nas versoes anteriores.
escada	Se TRUE, sobe a hierarquia da CBO quando a ocupacao de seis digitos nao esta na tabua: tenta a familia de 4, depois o subgrupo de 3, depois o subgrupo principal de 2. Leva a cobertura do dominio oficial de 49,8% para 89,8% . Padrao FALSE, porque o resultado deixa de ser a ocupacao declarada e passa a ser o seu grupo. Use crosswalk_cbo2002() para ver nivel_usado caso a caso.

Value

Vetor numerico com o escore ISEI-08.

Examples

```
cbo2002_para_isei08("225120")
```

cbo2002_para_prestigio

Prestigio de Treiman a partir da CBO-2002

Description

Prestigio de Treiman a partir da CBO-2002

Usage

```
cbo2002_para_prestigio(cbo, empate = c("na", "moda"), escada = FALSE)
```

Arguments

cbo	Vetor de codigos da CBO-2002, com 6 digitos (ocupacao) ou 4 (familia). Aceita "1111-05", "111105" ou 111105.
empate	O que fazer quando a familia nao tem ISCO majoritario: "na" (padrao) devolve NA; "moda" devolve o vencedor do desempate por menor codigo, como nas versoes anteriores.
escada	Se TRUE, sobe a hierarquia da CBO quando a ocupacao de seis digitos nao esta na tabua: tenta a familia de 4, depois o subgrupo de 3, depois o subgrupo principal de 2. Leva a cobertura do dominio oficial de 49,8% para 89,8% . Padrao FALSE, porque o resultado deixa de ser a ocupacao declarada e passa a ser o seu grupo. Use crosswalk_cbo2002() para ver nivel_usado caso a caso.

Value

Vetor numerico.

Examples

```
cbo2002_para_prestigio("225120")
```

cbo2002_para_prestigio08

Prestigio de Treiman ancorado na ISCO-08, a partir da CBO-2002

Description

A CBO-2002 aterrissa na ISCO-88, de modo que este escore atravessa a ponte 88 -> 08 e herda a perda dela: onde a OIT define mais de um destino, usa-se o codigo truncado. Veja [isco88_para_isco08\(\)](#).

Usage

```
cbo2002_para_prestigio08(cbo, empate = c("na", "moda"), escada = FALSE)
```

Arguments

cbo	Vetor de codigos da CBO-2002, com 6 digitos (ocupacao) ou 4 (familia). Aceita "1111-05", "111105" ou 111105.
empate	O que fazer quando a familia nao tem ISCO majoritario: "na" (padrao) devolve NA; "moda" devolve o vencedor do desempate por menor codigo, como nas versoes anteriores.
escada	Se TRUE, sobe a hierarquia da CBO quando a ocupacao de seis digitos nao esta na tabua: tenta a familia de 4, depois o subgrupo de 3, depois o subgrupo principal de 2. Leva a cobertura do dominio oficial de 49,8% para 89,8% . Padrao FALSE, porque o resultado deixa de ser a ocupacao declarada e passa a ser o seu grupo. Use crosswalk_cbo2002() para ver nivel_usado caso a caso.

Value

Vetor numerico.

Examples

```
cbo2002_para_prestigio08("225120")
```

cbo2002_para_siops	<i>Prestigio ocupacional SIOPS a partir da CBO-2002</i>
--------------------	---

Description

Prestigio ocupacional SIOPS a partir da CBO-2002

Usage

```
cbo2002_para_siops(cbo, empate = c("na", "moda"), escada = FALSE)
```

Arguments

cbo	Vetor de codigos da CBO-2002, com 6 digitos (ocupacao) ou 4 (familia). Aceita "1111-05", "111105" ou 111105.
empate	O que fazer quando a familia nao tem ISCO majoritario: "na" (padrao) devolve NA; "moda" devolve o vencedor do desempate por menor codigo, como nas versoes anteriores.
escada	Se TRUE, sobe a hierarquia da CBO quando a ocupacao de seis digitos nao esta na tabua: tenta a familia de 4, depois o subgrupo de 3, depois o subgrupo principal de 2. Leva a cobertura do dominio oficial de 49,8% para 89,8% . Padrao FALSE, porque o resultado deixa de ser a ocupacao declarada e passa a ser o seu grupo. Use crosswalk_cbo2002() para ver nivel_usado caso a caso.

Value

Vetor numerico com o escore SIOPS.

Examples

```
cbo2002_para_siops(c("1111-05", "225120"))
```

cbo2002_para_siops08 *Prestigio ancorado na ISCO-08 a partir da CBO-2002*

Description

Prestigio ancorado na ISCO-08 a partir da CBO-2002

Usage

```
cbo2002_para_siops08(cbo, empate = c("na", "moda"), escada = FALSE)
```

Arguments

cbo	Vetor de codigos da CBO-2002, com 6 digitos (ocupacao) ou 4 (familia). Aceita "1111-05", "111105" ou 111105.
empate	O que fazer quando a familia nao tem ISCO majoritario: "na" (padrao) devolve NA; "moda" devolve o vencedor do desempate por menor codigo, como nas versoes anteriores.
escada	Se TRUE, sobe a hierarquia da CBO quando a ocupacao de seis digitos nao esta na tabua: tenta a familia de 4, depois o subgrupo de 3, depois o subgrupo principal de 2. Leva a cobertura do dominio oficial de 49,8% para 89,8% . Padrao FALSE, porque o resultado deixa de ser a ocupacao declarada e passa a ser o seu grupo. Use crosswalk_cbo2002() para ver nivel_usado caso a caso.

Value

Vetor numerico com o escore de prestigio.

Examples

```
cbo2002_para_siops08("225120")
```

`cbo94_isco88`*Correspondência da CBO-94 com a ISCO-88*

Description

Para microdado anterior a 2003.

Usage

`cbo94_isco88`

Format

`data.frame` com as colunas:

cbo94 código da CBO-94, sem pontuação.

isco88 ISCO-88 correspondente.

Sem agregação

A tábua do MTE é 1:1 entre CBO-94 e CBO-2002 — cada linha traz um par —, de modo que não há maioria a apurar. Uma versão anterior desta tabela trazia uma coluna concordancia que era constante 1 por construção e sugeria um diagnóstico que não existia.

Source

Coluna CBO-94 da mesma tábua do Ministério do Trabalho.

Examples

```
head(cbo94_isco88)
```

```
# A CBO-94 chega à ISCO-88 e para aí no que toca a MEDIDAS: não há  
# cbo94_para_isei08(), cbo94_para_prestigio08() nem cbo94_para_isei_br().  
# A tradução de código existe (cbo94_para_isco08()); o que o pacote recusa é  
# pendurar escore de 2008 ou de 2025 em microdado anterior a 2003.  
nrow(cbo94_isco88)
```

cbo94_para_egp*Classe EGP a partir da CBO-94*

Description

Leia [isco88_para_egp\(\)](#) antes de usar: sem posicao no emprego e supervisao, o esquema sai degradado. A RAIS **nao** tem essas variaveis: e um registro de vinculo empregaticio, em que toda linha e um empregado. Por essa porta o EGP sai sem IVa, IVb e V por construcaao. Veja [cbo2002_para_egp\(\)](#).

Usage

```
cbo94_para_egp(  
  cbo94,  
  conta_propria = NULL,  
  n_supervisionados = NULL,  
  n_classes = 11,  
  rotulo = TRUE,  
  avisar = TRUE  
)
```

Arguments

cbo94	Vetor de codigos da CBO-94. Aceita "2-11.20" ou "21120".
conta_propria	Vetor lógico opcional: a pessoa trabalha por conta própria ou é empregadora (TRUE) ou é empregada (FALSE).
n_supervisionados	Vetor numérico opcional com o número de pessoas supervisionadas.
n_classes	Número de classes do resultado: 11 (padrão), 7, 5 ou 3.
rotulo	Se TRUE (padrão), devolve rótulos — em todas as versões, inclusive nas colapsadas; se FALSE, os códigos inteiros.
avisar	Se TRUE (padrão), avisa quando conta_propria ou n_supervisionados não são fornecidos.

Value

Vetor de texto (ou inteiro, se rotulo = FALSE).

Examples

```
cbo94_para_egp("2-11.20", avisar = FALSE)
```

cbo94_para_isco	<i>Traduz a CBO-94 em ISCO-88</i>
-----------------	-----------------------------------

Description

Para microdado anterior a 2003, que usa a classificacao de 1994 (cinco digitos). A correspondencia vem da mesma tabua oficial do Ministerio do Trabalho, pela coluna CBO-94.

Usage

```
cbo94_para_isco(cbo94)
```

Arguments

cbo94	Vetor de codigos da CBO-94. Aceita "2-11.20" ou "21120".
-------	--

Value

Vetor de texto com o ISCO-88 de quatro digitos, ou NA.

Examples

```
cbo94_para_isco(c("2-11.20", "21130"))
```

cbo94_para_isco08	<i>ISCO-08 a partir da CBO-94</i>
-------------------	-----------------------------------

Description

ISCO-08 a partir da CBO-94

Usage

```
cbo94_para_isco08(cbo94, com_ambiguidade = FALSE)
```

Arguments

cbo94	Vetor de codigos da CBO-94. Aceita "2-11.20" ou "21120".
com_ambiguidade	Se TRUE, devolve um data.frame com o código e o número de alternativas da OIT em vez de só o código.

Value

Vetor de texto com o ISCO-08, ou um data.frame se com_ambiguidade = TRUE.

Examples

```
cbo94_para_isco08("2-11.20")
```

cbo94_para_isei	<i>Indice socioeconomico ISEI a partir da CBO-94</i>
-----------------	--

Description

Indice socioeconomico ISEI a partir da CBO-94

Usage

```
cbo94_para_isei(cbo94)
```

Arguments

cbo94 Vetor de codigos da CBO-94. Aceita "2-11.20" ou "21120".

Value

Vetor numerico com o escore ISEI-88.

Examples

```
cbo94_para_isei("2-11.20")
```

cbo94_para_prestigio	<i>Prestigio de Treiman a partir da CBO-94</i>
----------------------	--

Description

Prestigio de Treiman a partir da CBO-94

Usage

```
cbo94_para_prestigio(cbo94)
```

Arguments

cbo94 Vetor de codigos da CBO-94. Aceita "2-11.20" ou "21120".

Value

Vetor numerico.

Examples

```
cbo94_para_prestigio("2-11.20")
```

cbo94_para_siops	<i>Prestigio ocupacional de Treiman a partir da CBO-94</i>
------------------	--

Description

Prestigio ocupacional de Treiman a partir da CBO-94

Usage

```
cbo94_para_siops(cbo94)
```

Arguments

cbo94	Vetor de codigos da CBO-94. Aceita "2-11.20" ou "21120".
-------	--

Value

Vetor numerico com o escore de prestigio.

Examples

```
cbo94_para_siops("2-11.20")
```

checa_cobertura	<i>Verifica se todos os códigos observados estão no dicionário</i>
-----------------	--

Description

Falha com erro se algum código de ocupação presente no seu dado não tiver entrada no dicionário do pacote, **ou se não houver código nenhum**. Chame antes de qualquer análise.

Usage

```
checa_cobertura(cod, silencioso = FALSE)
```

Arguments

cod	Vetor de códigos de ocupação observados no seu dado.
silencioso	Se TRUE, não escreve a mensagem de sucesso.

Value

Invisivelmente, TRUE. Erro se houver código fora do dicionário ou se não houver código válido algum.

Por que isto existe

O modo silencioso de errar uma medida de classe não é classificar mal um caso: é um código novo cair num rótulo residual sem que ninguém perceba. Foi o que aconteceu na primeira versão deste dicionário, em que 76 códigos reais caíam no rótulo "fora da PEA" — inclusive proprietários, que somem justamente da categoria em que mais importam.

O segundo modo silencioso é mais banal e mais comum: passar a coluna errada. dados\$CD_OCUPACAO num quadro cuja coluna se chama cd_ocupacao devolve NULL, e uma verificação ingênua aprovaria o vetor vazio. Por isso entrada vazia, NULL ou inteiramente ausente também é erro aqui.

Examples

```
checa_cobertura(c(111, 169, 257))
try(checa_cobertura(c(111, 99999)))
try(checa_cobertura(NULL))
```

checa_cobertura_cbo2002

Verifica a cobertura de códigos da CBO-2002

Description

Equivalente de [checa_cobertura\(\)](#) para a porta da CBO. Distingue o código que não existe na CBO-2002 daquele que existe mas **não tem correspondência oficial** com a CIUO-88 — este último é limite da tábua do Ministério do Trabalho, não erro do seu dado, e por isso vira aviso, não erro.

Usage

```
checa_cobertura_cbo2002(cbo, silencioso = FALSE)
```

Arguments

cbo	Vetor de códigos da CBO-2002 observados no seu dado.
silencioso	Se TRUE, não escreve a mensagem de sucesso.

Value

Invisivelmente, TRUE.

Examples

```
checa_cobertura_cbo2002(c("1111-05", "225120"))
try(checa_cobertura_cbo2002(NULL))
```

checa_cobertura_cbo94 *Verifica a cobertura de codigos da CBO-94*

Description

Equivalente de `checa_cobertura_cbo2002()` para o microdado anterior a 2003.

Usage

```
checa_cobertura_cbo94(cbo94, silencioso = FALSE)
```

Arguments

cbo94	Vetor de codigos da CBO-94 observados no seu dado.
silencioso	Se TRUE, nao escreve a mensagem de sucesso.

Value

Invisivelmente, TRUE.

Examples

```
checa_cobertura_cbo94(c("2-11.20", "21130"))
```

checa_cobertura_cod *Verifica a cobertura de codigos da COD*

Description

Verifica a cobertura de codigos da COD

Usage

```
checa_cobertura_cod(cod, silencioso = FALSE)
```

Arguments

cod	Vetor de codigos da COD observados no seu dado.
silencioso	Se TRUE, nao escreve a mensagem de sucesso.

Value

Invisivelmente, TRUE.

Examples

```
checa_cobertura_cod(c("2211", "0411"))
```

checa_periodo	<i>Verifica se o seu dado é atingido pela quebra de cadastro do TSE em 2002</i>
---------------	---

Description

O dicionário deste pacote adota o cadastro de ocupações vigente a partir de 2002. Entre 2000 e 2002 o TSE reeditou essa tabela e **reutilizou códigos** para ocupações diferentes, de modo que aplicar o dicionário a candidaturas de 1998 e 2000 classifica essas ocupações erradas. Esta função diz se, e quanto, o seu dado é atingido.

Usage

```
checa_periodo(cod, ano, avisar = TRUE)
```

Arguments

cod	Vetor de códigos de ocupação do TSE.
ano	Vetor de anos de eleição, do mesmo comprimento de cod.
avisar	Se TRUE (padrão), emite aviso quando há casos atingidos.

Value

Invisivelmente, um vetor lógico marcando as posições atingidas (código reutilizado e ano até 2000).

Como os casos foram identificados

Não pelo rótulo, e sim pelo dado. Os rótulos existem e o pacote os distribui ([tse_ocupacao_rotulos](#), [tse_para_rotulo\(\)](#)), mas um rótulo novo tanto pode ser a mesma ocupação rebatizada quanto outra ocupação no mesmo código, e sozinho não distingue as duas coisas. Para cada código presente nos dois períodos, comparou-se a proporção de candidatos com ensino superior completo. A escolaridade sobe ao longo de todo o período, então altas de 20 a 27 pontos são a tendência geral. Uma queda de 60 a 85 pontos, não: a população sob o código passou a ser outra. Ver [tse_quebra_2002](#).

O que fazer

Para análises restritas a 2002 ou depois, nada — a quebra não a atinge. Para séries que cruzam 2000/2002, o mínimo é excluir os códigos afetados do período antigo; o correto é obter os rótulos de ocupação do TSE de cada eleição e construir um dicionário com chave (código, ano).

Examples

```
checa_periodo(c("211", "211", "111"), c(2000, 2012, 2000))
```

cod_isco08

*Correspondência da COD do IBGE com a ISCO-08***Description**

A Classificação de Ocupações para Pesquisas Domiciliares é a da PNAD Contínua e do Censo Demográfico.

Usage

```
cod_isco08
```

Format

data.frame com 434 linhas:

cod grupo de base da COD, quatro dígitos.

titulo denominação oficial.

isco08 código ISCO-08 correspondente.

correspondencia identidade (428 casos), adaptacao (5) ou agregacao (1).

Por que quase tudo é identidade

A COD é construída **sobre** a ISCO-08. Não há tábua de conversão a consultar: 428 dos 434 grupos de base são o próprio código internacional. O que resta são seis adaptações brasileiras, documentadas uma a uma em [cod_para_isco08\(\)](#) — polícia e bombeiro militar, trabalhadores do sexo e pescadores.

O contraste com a perna da CBO é de desenho, não de esforço: aquela depende de uma tábua do Ministério do Trabalho que cobre metade do seu universo; esta cobre 100% do seu.

Source

Estrutura da Ocupação (COD), IBGE, distribuída em `inst/extdata/fontes/Estrutura_Ocupacao_COD.xls`.

Examples

```
head(cod_isco08)

# Quase tudo é identidade — a COD do IBGE é a ISCO-08 com outro nome. As
# poucas exceções são onde vale olhar antes de confiar.
table(cod_isco08$correspondencia)
cod_isco08[cod_isco08$correspondencia != "identidade", ]
```

cod_para_egp	Classe EGP a partir da COD
--------------	----------------------------

Description

A PNAD Continua tem posicao na ocupacao (V4012) e o numero de empregados do empregador, o que recupera a pequena burguesia e **separa IVa de IVb** — a distincao que pela porta do TSE fica indeterminada. Passe-os.

Usage

```
cod_para_egp(  
    cod,  
    conta_propria = NULL,  
    n_supervisionados = NULL,  
    n_classes = 11,  
    rotulo = TRUE,  
    avisar = TRUE  
)
```

Arguments

cod	Vetor de codigos da COD, de quatro digitos.
conta_propria	Vetor lógico opcional: a pessoa trabalha por conta própria ou é empregadora (TRUE) ou é empregada (FALSE).
n_supervisionados	Vetor numérico opcional com o número de pessoas supervisionadas.
n_classes	Número de classes do resultado: 11 (padrão), 7, 5 ou 3.
rotulo	Se TRUE (padrão), devolve rótulos — em todas as versões, inclusive nas colapsadas; se FALSE, os códigos inteiros.
avisar	Se TRUE (padrão), avisa quando conta_propria ou n_supervisionados não são fornecidos.

Details

O que ela nao tem e a supervisao exercida sobre assalariados, que e o que define V (supervisores manuais) e governa as promocoos entre I e II. Por essa porta o EGP fica mais completo que pelas outras, mas nao completo.

Value

Vetor de texto (ou inteiro, se rotulo = FALSE).

Examples

```
cod_para_egp("2211", avisar = FALSE)
```

cod_para_isco	<i>ISCO-88 a partir da COD</i>
---------------	--------------------------------

Description

Passa pela ISCO-08 e desce pela ponte reversa. E o caminho para o ISEI-88 e para o EGP a partir da PNAD Continua e do Censo.

Usage

```
cod_para_isco(cod)
```

Arguments

cod	Vetor de codigos da COD, de quatro digitos.
-----	---

Details

E tambem o ponto em que a adaptacao brasileira da COD se desfaz sem que voce peca: o IBGE aloca a policia militar e o corpo de bombeiros militar ao grande grupo 0, o das forcas armadas, e esta funcao os devolve a 5162 e 5161, que e onde os codigos 145, 232, 233 e 258 do cadastro do TSE tambem aterrissam. Quem ler o primeiro digito da COD em vez de traduzir poe o policial militar da populacao num grande grupo e o candidato policial militar em outro. O exemplo abaixo mostra os dois caminhos.

Value

Vetor de texto com o ISCO-88.

See Also

O artigo *Comparar duas fontes*, no site do pacote, que trata das oito decisoes de uma comparacao entre candidaturas e populacao e explica esta adaptacao em detalhe: <https://moraespeixoto.github.io/ocupacoesBR/articles/comparar-fontes.html>

Examples

```
cod_para_isco(c("2211", "0411"))

# a adaptacao da COD, desfeita: 0411 e policia militar
cod_para_isco08("0411") # 5412, servico protetivo na ISCO-08
cod_para_isco("0411")   # 5162, o mesmo destino dos codigos do TSE
```

cod_para_isco08

*Traduz a COD do IBGE em ISCO-08***Description**

Converte o código da Classificação de Ocupações para Pesquisas Domiciliares — a da PNAD Contínua e do Censo Demográfico — no código ISCO-08.

Usage

```
cod_para_isco08(cod)
```

Arguments

cod Vetor de códigos da COD, de quatro dígitos.

Value

Vetor de texto com o ISCO-08.

Por que esta perna e quase gratuita

A COD é construída sobre a ISCO-08, e por isso **428 dos seus 434 grupos de base são idênticos ao código internacional**. Não há tabela de conversão a consultar nem correspondência a construir: a tradução e identidade, e o que resta é decidir as seis adaptações brasileiras.

Compare com a perna da CBO, que exige a tabela do Ministério do Trabalho e cobre metade do seu universo. A diferença não é de esforço: é de desenho das classificações.

As seis adaptações

0411, 0412, 0511, 0512 polícia e bombeiro militar. Vão para 5412 e 5411, do grande grupo 5 (serviços protetivos), e não para o grande grupo 0 (forças armadas): a PM e o BM brasileiros são militarizados em estatuto, mas exercem serviço protetivo civil, e é a função que a classificação mede. **A distinção entre oficial e praca se perde** — a ISCO-08 não a tem, e no Brasil ela é um grau de status real. Quem precisar dela tem de tratá-la fora do ISEI.

5168 trabalhadores do sexo, para 5169 (serviços pessoais não classificados em outra parte).

6225 pescadores, para o **subgrupo** 6220. A COD funde numa rubrica o que a ISCO-08 reparte em quatro (aquicultura, pesca costeira, pesca de alto-mar, caca), que vão de ISEI 11 a 21. Escolher uma seria arbitrário; o subgrupo e a forma arredondada que o próprio ISMF publica.

O único buraco, e ele vem da fonte

A tradução e completa — os 434 grupos de base chegam a ISCO-08 e a ISCO-88. Mas as FORÇAS ARMADAS (0110 e 0210) ficam sem **ISEI-88, prestígio-88 e EGP**, porque o ISMF não pontua o ISCO-88 0110. As medidas ancoradas na ISCO-08 existem — `cod_para_isei08("0110")` e 60,9 (oficiais) e `("0210")`, 51,6 —, porque o `isqoisei08.sps` pontua a hierarquia militar da ISCO-08; a lacuna e só na ponte de volta a ISCO-88, e devolver NA ali é o comportamento correto, não um escore inventado.

See Also

[cod_para_isei08\(\)](#), [cod_para_isco\(\)](#), [crosswalk_cod\(\)](#)

Examples

```
cod_para_isco08(c("2211", "0411", "6225"))
```

cod_para_isei	<i>ISEI-88 a partir da COD</i>
---------------	--------------------------------

Description

Use este, e nao [cod_para_isei08\(\)](#), para comparar com dado eleitoral: o pacote e ancorado na ISCO-88 e e nela que o TSE aterrissa.

Usage

```
cod_para_isei(cod)
```

Arguments

cod Vetor de codigos da COD, de quatro digitos.

Value

Vetor numerico com o escore ISEI-88.

Examples

```
cod_para_isei(c("2211", "9629"))
```

cod_para_isei_br	<i>ISEI-BR a partir da COD</i>
------------------	--------------------------------

Description

Versao para a classificacao do IBGE, que e a da PNAD Continua e do Censo. Leia [tse_para_isei_br\(\)](#) para o que a regua e e o que ela nao e.

Usage

```
cod_para_isei_br(cod)
```

Arguments

cod Vetor de codigos da COD, de quatro digitos.

Value

Vetor numerico com o escore ISEI-BR, entre 10 e 90.

See Also

[tse_para_isei_br\(\)](#), [isco08_isei_br](#)

Examples

```
cod_para_isei_br(c("2211", "9629"))
```

<i>cod_para_isei08</i>	<i>ISEI-08 a partir da COD</i>
------------------------	--------------------------------

Description

ISEI-08 a partir da COD

Usage

```
cod_para_isei08(cod)
```

Arguments

cod Vetor de codigos da COD, de quatro digitos.

Value

Vetor numerico com o escore ISEI-08.

Examples

```
cod_para_isei08(c("2211", "9629"))
```

cod_para_prestigio	<i>Prestigio de Treiman a partir da COD</i>
--------------------	---

Description

Prestigio de Treiman a partir da COD

Usage

```
cod_para_prestigio(cod)
```

Arguments

cod	Vetor de codigos da COD, de quatro digitos.
-----	---

Value

Vetor numerico.

Examples

```
cod_para_prestigio(c("2211", "9629"))
```

cod_para_prestigio08	<i>Prestigio de Treiman ancorado na ISCO-08, a partir da COD</i>
----------------------	--

Description

Prestigio de Treiman ancorado na ISCO-08, a partir da COD

Usage

```
cod_para_prestigio08(cod)
```

Arguments

cod	Vetor de codigos da COD, de quatro digitos.
-----	---

Value

Vetor numerico.

Examples

```
cod_para_prestigio08(c("2211", "9629"))
```

cod_populacao_br	<i>A população brasileira pelos endereços da ISCO-88, por piso de elegibilidade</i>
------------------	---

Description

O denominador. A distribuição da população brasileira de 18 anos ou mais pelos códigos ISCO-88, a partir da COD da PNAD Contínua, com a parcela que **não está ocupada** como linha da própria tabela — e não como ausência dela. É o outro lado de [tse_universo](#), na mesma forma, para que comparar candidatura com população seja uma junção de duas linhas.

Usage

```
cod_populacao_br
```

Format

data.frame com 3.811 linhas:

piso o piso etário de elegibilidade: 18, 21, 30 ou 35 anos. **São pisos, não faixas** — cada um é a população inteira daquela idade em diante, e por isso se sobrepõem.

sexo "Homem", "Mulher" ou "Todos".

isco88 código ISCO-88 de quatro dígitos. NA nas duas linhas que não são ocupação.

situacao "Ocupado", "Não ocupado" ou "Ocupado sem endereço na ISCO-88" (0,01% — a COD cobre quase tudo).

pop pessoas estimadas, com o peso da PNAD Contínua.

n_pessoas pessoas distintas na amostra, atrás da estimativa. É a coluna de precisão, e **704 das 3.811 linhas têm menos de 30**.

pct percentual dentro de (piso, sexo). Somam 100 por construção.

É estimativa de amostra com peso, e é de 2025

Isto não é um censo. Cada pop é uma estimativa da PNAD Contínua dos **quatro trimestres de 2025**, com o peso V1028 dividido pelo número de trimestres, e carrega o erro amostral que n_pessoas deixa medir. Ao contrário das tábuas de conversão do pacote, que não envelhecem, **esta tabela envelhece**: uma distribuição de 2025 lida daqui a alguns anos descreve um país que mudou. Declare o ano ao usá-la.

O pacote não faz desenho amostral. Se você precisa de erro-padrão, de intervalo de confiança ou de subpopulação, o caminho é o microdado com **survey** — esta tabela dá o ponto, não a incerteza.

O piso é a decisão que mais move o denominador

Os quatro pisos são os do art. 14 §3º da Constituição: 18 anos para vereador, 21 para prefeito e deputado, 30 para governador, 35 para senador e presidente. A parcela que não está ocupada muda com o piso, e não pouco:

piso	homens	mulheres	todos
18 anos	27,0%	48,1%	38,0%
21 anos	26,0%	47,7%	37,3%
30 anos	27,8%	49,9%	39,4%
35 anos	30,3%	52,3%	42,0%

Quem compara candidatos ao Senado com "a população" sem escolher o piso erra o denominador em quatro pontos, e a diferença entre os sexos é de mais de vinte.

As duas fontes perdem quase a mesma gente, por motivos opostos

Medida sobre a população de 18 anos ou mais, a cobertura do ISEI é de **61,8%**; medida sobre as candidaturas de 2024, [tse_universo](#) registra **62,0%**. Os números quase coincidem e as razões não têm nada em comum: a população perde quem não está ocupado, o Tribunal perde quem marcou uma rubrica que não nomeia ocupação. Trocar o denominador sem dizer troca o resultado sem avisar. É o assunto do artigo *Comparar duas fontes*, no site do pacote: <https://moraespeixoto.github.io/ocupacoesBR/articles/comparar-fontes.html>

A chave passou pela ponte reversa

A PNAD classifica por COD, e a chave desta tabela é a ISCO-88: o script percorre COD -> ISCO-08 -> ISCO-88, e o segundo passo é a ponte que [isco08_para_isco88\(\)](#) documenta como voltando ao ponto de partida em apenas 69% dos códigos. Quem descer ao código de quatro dígitos deve saber disso, e olhar `n_pessoas` antes de ler a célula.

Source

Microdados da PNAD Contínua trimestral, IBGE, quatro trimestres de 2025. Gerada por `data-raw/07b_gera_populacao.R`. Os microdados **não** viajam com o pacote (212 MB por trimestre); o que entra é esta tabela agregada, sem indivíduo e sem identificador. https://ftp.ibge.gov.br/Trabalho_e_Rendimento/Pesquisa_Nacional_por_Amostra_de_Domicilios_continua/Trimestral/Microdados/

Examples

```
# O universo, declarado: quem está ocupado e quem não está.
p <- cod_populacao_br
aggregate(pct ~ situacao, p[p$piso == 18 & p$sexo == "Todos", ], sum)

# A composição da população ocupada pelo grande grupo, no piso de vereador.
o <- p[p$piso == 18 & p$sexo == "Todos" & p$situacao == "Ocupado", ]
round(tapply(o$pop, substr(o$isco88, 1, 1), sum) / sum(o$pop) * 100, 1)

# A célula que só existe porque a tabela guarda quatro dígitos: a polícia
# militar, que a COD aloca ao grande grupo 0 e cod_para_isco() devolve ao 5.
p[p$isco88 %in% c("5161", "5162") & p$piso == 18, ]
```

crosswalk_cbo2002

*Tabela completa de traducao a partir da CBO-2002***Description**

Tabela completa de traducao a partir da CBO-2002

Usage

```
crosswalk_cbo2002(cbo = NULL, empate = c("na", "moda"), escada = FALSE)
```

Arguments

cbo	Vetor opcional de codigos. Se omitido, devolve a tabua inteira.
empate	O que fazer quando a familia nao tem ISCO majoritario: "na" (padrao) devolve NA; "moda" devolve o vencedor do desempate por menor codigo, como nas versoes anteriores.
escada	Se TRUE, sobe a hierarquia da CBO para preencher as ocupacoes ausentes da tabua; veja cbo2002_para_isco() . A coluna nivel_usado sai de qualquer modo, e diz de que degrau veio o ISCO de cada linha: 6 e a propria ocupacao, 4 a familia, 3 o subgrupo, 2 o subgrupo principal.

Value

data.frame com CBO-2002, titulo, familia, CBO-94, ISCO-88, ISCO-08, ISEI, SIOPS, EGP, agregado, empate e nivel_usado. empate e TRUE quando a linha e uma familia de quatro digitos sem ISCO majoritario; com empate = "na" (padrao) o ISCO dessa linha sai NA, exatamente como em [cbo2002_para_isco\(\)](#) — a tabela nao pode devolver o que a porta recusa.

A coluna egp sai degradada, e em silêncio

O EGP não é função só da ocupação: as suas regras pedem a posição no emprego e o número de subordinados, e nenhum dos dois existe num código de CBO-2002. A coluna é calculada sem eles, e a consequência não é ruído — **IVa, IVb e V saem estruturalmente vazias**, de modo que a pequena burguesia aparece contada como classe de serviço. Isso é inversão de classe, não arredondamento.

A chamada silencia o aviso que [isco88_para_egp\(\)](#) emitiria, porque repeti-lo uma vez por linha seria inútil; a ressalva passou a viver aqui, na auditoria de 05/09/2026. Para publicar, use os colapsos de 5 ou 3 classes, onde IVa e IVb se fundem a categorias que existem, e leia [isco88_para_egp\(\)](#).

Examples

```
head(crosswalk_cbo2002())
crosswalk_cbo2002(c("1111-05", "225120"))
```

crosswalk_cbo94	<i>Tabela completa de traducao a partir da CBO-94</i>
-----------------	---

Description

Tabela completa de traducao a partir da CBO-94

Usage

```
crosswalk_cbo94(cbo94 = NULL)
```

Arguments

cbo94 Vetor opcional de codigos. Se omitido, devolve a tabua inteira.

Value

data.frame com CBO-94, ISCO-88, ISCO-08, ISEI-88, prestígio e EGP.

A coluna egp sai degradada, e em silêncio

O EGP não é função só da ocupação: as suas regras pedem a posição no emprego e o número de subordinados, e nenhum dos dois existe num código de CBO-94. A coluna é calculada sem eles, e a consequência não é ruído — **IVa, IVb e V saem estruturalmente vazias**, de modo que a pequena burguesia aparece contada como classe de serviço. Isso é inversão de classe, não arredondamento.

A chamada silencia o aviso que [isco88_para_egp\(\)](#) emitiria, porque repeti-lo uma vez por linha seria inútil; a ressalva passou a viver aqui, na auditoria de 05/09/2026. Para publicar, use os colapsos de 5 ou 3 classes, onde IVa e IVb se fundem a categorias que existem, e leia [isco88_para_egp\(\)](#).

Examples

```
head(crosswalk_cbo94())
```

crosswalk_cod	<i>Tabela completa de traducao a partir da COD</i>
---------------	--

Description

Tabela completa de traducao a partir da COD

Usage

```
crosswalk_cod(cod = NULL)
```

Arguments

cod Vetor opcional de codigos. Se omitido, devolve a tabela inteira.

Value

`data.frame` com COD, título, ISCO-08, ISCO-88, as duas versões do ISEL, o prestígio, o EGP e o tipo de correspondência.

A coluna `egp` sai degradada, e em silêncio

O EGP não é função só da ocupação: as suas regras pedem a posição no emprego e o número de subordinados, e nenhum dos dois existe num código de COD da PNAD. A coluna é calculada sem eles, e a consequência não é ruído — **IVa, IVb e V saem estruturalmente vazias**, de modo que a pequena burguesia aparece contada como classe de serviço. Isso é inversão de classe, não arredondamento.

A chamada silencia o aviso que `isco88_para_egp()` emitiria, porque repeti-lo uma vez por linha seria inútil; a ressalva passou a viver aqui, na auditoria de 05/09/2026. Para publicar, use os colapsos de 5 ou 3 classes, onde IVa e IVb se fundem a categorias que existem, e considere `isco_posicao_br` como prior empírico da posição no emprego. Compare com `crosswalk_tse()`, que é a única das quatro tabelas em que o EGP sai completo, porque o dicionário do TSE marca quem trabalha por conta própria.

Examples

```
head(crosswalk_cod())
crosswalk_cod(c("2211", "0411", "6225"))
```

crosswalk_tse

Tabela completa de tradução, do TSE a todas as medidas

Description

Devolve, num único `data.frame`, tudo o que o pacote sabe sobre cada código de ocupação do TSE. Serve para inspecionar a medida antes de aplicá-la, para publicar como material suplementar de um artigo e para auditar decisões de classificação uma a uma.

Usage

```
crosswalk_tse(cod = NULL)
```

Arguments

`cod` Vetor opcional de códigos. Se omitido, devolve o dicionário inteiro.

Value

`data.frame` com o código do TSE, ISCO-88, ISCO-08, ISEL-88, ISEL-08, SIOPS, EGP, classe, estrato, componente da classe alta e as marcas de ocupação política e de proprietário.

A coluna `egp`, e por que só aqui ela é comparável

Esta é a única das quatro tabelas de tradução em que o EGP sai com a posição no emprego informada: o dicionário do TSE marca em `tse_isco$conta_propria` quem trabalha por conta própria, e é essa marca — não proprietário — que corresponde ao `SEMP` = 2 das sintaxes do ISMF. Por isso `IVb` e `IVc` aparecem aqui e saem vazias em `crosswalk_cod()`, `crosswalk_cbo2002()` e `crosswalk_cbo94()`, onde o código não carrega posição alguma.

O que **continua** faltando é o número de subordinados: `IVa` e `IVb` não se separam, todos caem em `IVb`, e `V` fica subestimada. Para publicar, prefira os colapsos de 7 ou 5 classes.

Até 05/09/2026 esta coluna usava `proprietario`, e discordava de `tse_para_egp()` em dois códigos — o agricultor (601) e o pescador (604) saíam em `VIIb` aqui e em `IVc` na função. Como esta tabela existe para ser publicada como suplemento, o suplemento contradizia o código que produziu as estimativas. As duas saídas hoje são idênticas por teste.

Examples

```
head(crosswalk_tse())
crosswalk_tse(c(111, 169, 257))
```

isco_posicao_br	Posição na ocupação por código ISCO-88, medida na PNAD Contínua
-----------------	---

Description

A distribuição brasileira de posição no emprego — conta própria, empregador, número de empregados, setor público, posição militar — para cada código ISCO-88, apurada nos microdados da PNAD Contínua de 2025.

Usage

```
isco_posicao_br
```

Format

- `data.frame` com 319 linhas:
- isco88** código ISCO-88 de quatro dígitos.
- n_pessoas** pessoas distintas observadas (a medida honesta de precisão; veja a seção sobre o painel).
- n_obs** observações pessoa-trimestre.
- pct_conta_propria** % que trabalha por conta própria **ou** é empregadora — o `SEMP` = 2 das sintaxes do ISMF.
- pct_empregador** % que é empregadora.
- pct_emp_11mais** entre os empregadores, % com 11 ou mais empregados — o limiar que separa a ISCO 12 da 13. NA onde há menos de 25 empregadores na célula.

pct_setor_publico % empregada do setor público, inclusive empresas de economia mista (V4012 == 4). É a coluna que torna comparável, do lado da população, o vínculo público que o formulário do TSE oferece como rótulo — veja a seção "O vínculo público".

pct_militar % cuja **posição** declarada é militar (V4012 == 2). Leia a seção "Militar é duas coisas" antes de usar: esta coluna não é o grande grupo 0 da ISCO-88, e a diferença é o assunto.

grupo o grande grupo de dois dígitos.

n_pessoas_grupo, pct_conta_propria_grupo, pct_empregador_grupo, pct_setor_publico_grupo, pct_militar_grupo o mesmo, apurado no grupo de dois dígitos. Cada linha carrega a sua própria estimativa e a do grupo, para que quem cair numa célula fina possa recuar um nível sem refazer a conta — e veja, lado a lado, com que n cada uma foi apurada.

Para que serve

O EGP não é função só da ocupação: as regras do ISMF pedem a posição no emprego e a supervisão. O formulário do TSE não pergunta nenhuma das duas, e por isso o esquema sai degradado (veja [isco88_para_egp\(\)](#)). Esta tabela é o **prior empírico** dessa variável ausente: não imputa a posição de ninguém, e sim informa qual é a composição da ocupação no país.

O uso legítimo é análise de sensibilidade — rodar o EGP com e sem a posição provável e ver se a conclusão se move. O uso ilegítimo é tratar a proporção como se fosse o caso individual.

O painel rotativo

A PNAD reentrevista o mesmo domicílio por cinco trimestres. Os quatro trimestres de 2025 somam 864.870 observações de **437.880 pessoas distintas** (1,98x). Somá-los como amostras independentes inflaria o n sem acrescentar informação na mesma proporção. Por isso os pesos são divididos pelo número de trimestres — as estimativas são a média do ano civil, com os pesos somando a população e não quatro vezes ela — e a coluna de precisão é n_pessoas.

O que ela mostra, e por que isso importa

A ISCO 61 é, na definição da OIT, quem **opera a própria terra**; a 92 é o assalariado rural. O dado brasileiro separa as duas com folga:

ISCO-88	conta própria ou empregador
61 (agrícolas qualificados)	67,6%
6150 (pesca)	83,8%
92 (rurais elementares)	16,5%
todas as ocupações	29,4%

É a evidência externa que motivou separar tse_isco\$conta_propria de tse_isco\$proprietario: o agricultor familiar trabalha por conta própria sem pertencer à classe proprietária.

Como candidatos são selecionados por patrimônio, tomar estas proporções como piso — e não como estimativa central — é a leitura conservadora.

O vínculo público

O cadastro de ocupações do TSE oferece à pessoa quatro rótulos que **substituem** a ocupação em vez de a nomear — 291 OCUPANTE DE CARGO EM COMISSÃO e 296/297/298 SERVIDOR PÚBLICO FEDERAL/ESTADUAL/MUNICIPAL. Nenhum tem ISCO, e corretamente: não designam ocupação. Na PNAD Contínua as mesmas pessoas **têm** ocupação declarada, porque vínculo e ocupação são duas perguntas separadas. Medido nos quatro trimestres de 2025, o setor público é 11,7% dos ocupados com endereço na ISCO-88, e se distribui assim pelo grande grupo:

dígito	1	2	3	4	5	6	7	8	9
% do setor público	4,0	37,9	18,7	14,2	12,1	0,1	1,1	3,6	8,3

Quase quatro em dez estão no dígito 2, professores sobretudo. Uma comparação de composição entre as duas fontes que não trate disso compara um universo que exclui servidores com outro que os inclui. O artigo *Comparar duas fontes*, no site do pacote, mede a consequência: <https://moraespeixoto.github.io/ocupacoesBR/articles/comparar-fontes.html>

Esta coluna **não conserta o construto**, só o denominador: no TSE o vínculo público é uma *escolha* de rótulo, feita por quem podia ter escrito "professor"; aqui é a posição de quem *também* declarou a ocupação.

Militar é duas coisas, e elas discordam

O dicionário do IBGE define V4012 == 2 como "militar do exército, da marinha, da aeronáutica, **da polícia militar ou do corpo de bombeiros militar**". Não é o grande grupo 0 da ISCO-88. Medido aqui: entre quem declara essa posição, 39,9% cai no grande grupo 0 pelo código de ocupação e 60,1% cai no 5, o dos serviços protetivos. **A pergunta sobre posição e a pergunta sobre ocupação discordam sobre quem é militar no Brasil**, e a discordância é a mesma que a COD registra ao alocar a polícia militar e o bombeiro militar ao grande grupo 0 — adaptação que `cod_para_isco()` desfaz, devolvendo 5162 e 5161.

A consequência prática: `pct_militar` não é o tamanho das forças armadas (0,81% dos ocupados inclui o policiamento militar), e não deve ser somada nem comparada com a parcela do grande grupo 0 como se fossem a mesma coisa. Nenhuma das duas respostas é o erro da outra; são construtos diferentes.

As quatro colunas não se somam a 100

`pct_conta_propria`, `pct_setor_publico` e `pct_militar` são mutuamente exclusivas — V4012 vale 2, 4, 5 ou 6, e nunca duas ao mesmo tempo — e a soma das três nunca passa de 100. O que sobra é empregado do setor privado, trabalhador doméstico e trabalhador familiar auxiliar. `pct_empregador` **não** entra nessa soma: é subconjunto de `pct_conta_propria`.

A tabela atravessa a ponte reversa

A PNAD classifica por COD, e a chave desta tabela é a ISCO-88, de modo que `data-raw/07_gera_posicao.R` percorre COD -> ISCO-08 -> ISCO-88. O segundo passo é a ponte reversa, que `isco08_para_isco88()` documenta como voltando ao ponto de partida em apenas 69% dos códigos. A advertência existia lá e não aqui, que é onde quem usa esta tabela vai olhar (auditoria de 05/09/2026).

Na prática o dano é pequeno, porque as linhas publicadas são grupos de dois dígitos e a perda da ponte está sobretudo no quarto. Mas quem descer ao código de quatro dígitos desta tabela deve saber que a chave passou por uma tradução que não é bijetiva.

Source

Microdados da PNAD Contínua trimestral, IBGE, quatro trimestres de 2025, acessados em 28/07/2026. https://ftp.ibge.gov.br/Trabalho_e_Rendimento/Pesquisa_Nacional_por_Amostra_de_Domicilios_continua/Trimestral/Microdados/ Gerada por data-raw/07_gera_posicao.R. Os microdados **não** viajam com o pacote (212 MB por trimestre); o que entra é esta tabela agregada.

Examples

```
# Dentro do mesmo ISCO convivem quem trabalha por conta própria e quem
# emprega. É esta tabela, medida na PNAD Contínua, que separa os dois – e
# foi ela que tirou o agricultor da classe alta.
p <- isco_posicao_br
head(p[order(-p$pct_conta_propria),
      c("isco88", "n_obs", "pct_conta_propria", "pct_employador")], 5)
```

isco08_isco88

Ponte da ISCO-08 de volta para a ISCO-88

Description

Ponte da ISCO-08 de volta para a ISCO-88

Usage

```
isco08_isco88
```

Format

data.frame com 590 linhas:

isco08 código de origem.

isco88 código de destino na revisão de 1988.

A ida e a volta não se cancelam

Levar um código da ISCO-88 à ISCO-08 por [isco88_para_isco08\(\)](#) e trazê-lo de volta por [isco08_para_isco88\(\)](#) devolve o ponto de partida em **69% dos casos**. O terço restante não é defeito: a OIT reparte e funde categorias entre as revisões, e a composição das duas concordâncias não é a identidade.

Esta tabela é o que destrava o EGP e o ISEI-88 para quem entra pela COD — a PNAD Contínua e o Censo aterrisam na ISCO-08, e todo o esquema de classes do pacote está ancorado na ISCO-88.

Source

Módulo isco0888.sps do International Stratification and Mobility File. <http://www.harryganzeboom.nl/ismf/index.htm>

Examples

```
head(isco08_isco88)

# É esta tábua que permite entrar pela PNAD ou pelo Censo e chegar às
# medidas ancoradas na ISCO-88.
nrow(isco08_isco88)
```

isco08_isei_br

ISEI-BR: status ocupacional estimado na PNAD Contínua

Description

O ISEI que o pacote carregava era importado: Ganzeboom, De Graaf e Treiman (1992) escalonaram a ISCO sobre 31 conjuntos de dados de dezesseis países, levantados entre 1968 e 1982. O Brasil está entre eles, e não de passagem: a PNAD de 1973 e a de 1982 deram 15.439 dos 73.901 homens da amostra — 20,9%, a segunda maior contribuição nacional, atrás só da norte-americana.

Usage

```
isco08_isei_br
```

Format

data.frame com 590 linhas, uma para cada chave de [isco08_medidas](#), e as colunas:

isco08 código ISCO-08 de quatro dígitos.

isei_br escore, contínuo entre 10 e 90.

n_pessoas pessoas distintas na célula que forneceu o escore.

n_obs observações trimestrais nessa mesma célula.

nivel em quantos dígitos o escore foi apurado: 4 é a própria ocupação, 3, 2 e 1 são grupos cada vez mais grossos.

anos_estudo média de anos de estudo na célula, sem residualizar.

log_renda média do log do rendimento habitual, sem residualizar.

Os atributos theta, beta_direto, beta_total, parcela_mediada, n_obs e n_pessoas guardam os parâmetros da estimação e o tamanho da amostra que a produziu. Note que n_obs como atributo é a amostra de estimação inteira, e n_obs como coluna é a célula daquela linha.

Details

O que a régua importada não faz é deixar o Brasil ter ângulo próprio. Os dezesseis países entram num único escalonamento, e o dado brasileiro que o alimenta é de 1973 e 1982. Esta tabela é o mesmo procedimento feito do zero sobre a PNAD Contínua de 2025, com um ângulo estimado só aqui. Ela fecha a lacuna que vignette("validacao") declarava em aberto.

Método

A ocupação entra como variável interveniente entre escolaridade e renda, e o escore de cada ocupação é a combinação das médias de escolaridade e de renda dos seus ocupantes que minimiza o efeito direto da escolaridade sobre a renda. A amostra tem idade de 21 a 64 anos, ambos os sexos, pelo menos 30 horas semanais no trabalho principal e rendimento habitual positivo. Escolaridade e renda entram residualizadas em idade, idade ao quadrado, sexo e trimestre. O peso é dividido pelo número de trimestres, porque a PNAD é painel rotativo e a mesma pessoa reaparece.

O resíduo, que é achado e não defeito

O procedimento de 1992 supõe que a ocupação medeia integralmente a relação entre escolaridade e renda, e escolhe o ângulo onde o efeito direto zera. No Brasil ele não zera. A curva tem mínimo interior e para em 0,206, contra um efeito total de 0,494, de modo que a ocupação medeia 58,4% do efeito. Os outros 41,6% são escolaridade que paga dentro da mesma ocupação, o que se lê como heterogeneidade intraocupacional e informalidade. Adota-se o ângulo de mínimo e publica-se o resíduo.

Sensibilidade

O ordenamento não depende do ângulo exato nem das escolhas de método. A correlação de Spearman entre a escala no ângulo adotado e em mais ou menos 0,15 radianos é de 0,999. Sete especificações alternativas foram testadas, entre elas exigir 40 horas, dispensar a restrição de horas, usar renda-hora, usar rendimento efetivo, restringir a homens como o artigo de 1992 fez, e usar escolaridade em categorias: nenhuma move o ordenamento abaixo de 0,99 contra a especificação adotada, e a parcela mediada fica sempre entre 57,9% e 60,1%. Quem reproduz esses três números é data-raw/09b_sensibilidade_isei_br.R, que imprime a tabela inteira e não grava nada.

A escala não é a do ISEI, e isso importa na comparação

Os escores saem de um min-max sobre as células de estimação, para o intervalo de 10 a 90. O intervalo é o mesmo do ISEI, a **dispersão não**: o desvio padrão do ISEI-BR é da ordem de 14 a 15, contra 21 do ISEI-08, porque o min-max é governado pelas células extremas. A consequência é operacional e foi encontrada na auditoria de 05/09/2026: a diferença bruta `isei_br - iseio8` correlaciona-se a **-0,85** com o próprio `iseio8`, de modo que a lista das ocupações que "mais sobem" é, em boa parte, a lista das que estavam mais embaixo. Num gráfico das duas, a diagonal de 45 graus marca a igualdade de escore e não a de posição.

Para comparar as duas régua, padronize as duas antes de subtrair. É o que vignette("validacao") passou a fazer. Para usar o ISEI-BR sozinho — que é o uso previsto — nada disso importa, porque a escala é monotônica e qualquer análise por ordenamento é indiferente a ela.

O que se perde

A COD funde oficiais e praças de polícia e bombeiro militar, então a escala não distingue os dois. Códigos da ISCO-08 em que nenhuma COD aterrissa herdam o escore do grupo acima. Dos 407 códigos de quatro dígitos, 348 têm escore próprio e 59 são herdados por terem menos de 30 pessoas na amostra. As outras 183 linhas são as formas arredondadas de dois e três dígitos, que existem porque a ocupação declarada ao TSE é grossa e aterrissa nelas.

O que ela não é

Não é substituta do ISEI-08. Comparação internacional continua exigindo a âncora internacional, e uma série não troca de régua no meio. É um ano só, 2025, e não forma série.

Source

PNAD Contínua trimestral, microdados dos quatro trimestres de 2025, IBGE. Acesso em 05/09/2026. Gerada por `data-raw/09_gera_isei_br.R`, com a sensibilidade em `data-raw/09b_sensibilidade_isei_br.R`; os sha256 dos arquivos estão em `inst/extdata/PROVENIENCIA.yml`.

References

Ganzeboom, H. B. G., De Graaf, P. M., & Treiman, D. J. (1992). A standard international socio-economic index of occupational status. *Social Science Research*, 21(1), 1-56.

See Also

[tse_para_isei_br\(\)](#) e as irmãs por COD, CBO-2002 e ISCO-08.

Examples

```
head(isco08_isei_br)

# Concorda com a âncora internacional sem ser cópia dela.
f <- isco08_isei_br[isco08_isei_br$nivel == 4, ]
i <- isco08_medidas$isei08[match(f$isco08, isco08_medidas$isco08)]
round(stats::cor(f$isei_br, i, use = "complete.obs", method = "spearman"), 3)

# Os parâmetros da estimação viajam com a tabela.
attributes(isco08_isei_br)[c("theta", "parcela_mediada")]
```

isco08_medidas

Medidas ancoradas na ISCO-08

Description

Medidas ancoradas na ISCO-08

Usage

```
isco08_medidas
```

Format

`data.frame` com as colunas:

isco08 código ISCO-08 de quatro dígitos.

isei08 escore ISEI-08, contínuo.

siops08 escore SIOPS-08, contínuo.

Source

Módulos `isqoisei08.sps` e `isqotrei08.sps` do International Stratification and Mobility File.

<http://www.harryganzeboom.nl/ismf/index.htm>

Examples

```
head(isco08_medidas)

# A âncora é outra: um escore da ISCO-08 não é comparável com um da
# ISCO-88, e misturar as duas numa mesma série é o erro mais comum.
nrow(isco08_medidas)
summary(isco08_medidas$isei08)

# Estes escores foram estimados fora do Brasil. Para a régua estimada em
# dado brasileiro, com o mesmo método, veja [isco08_isei_br].
```

<code>isco08_para_isco88</code>	<i>Converte ISCO-08 em ISCO-88</i>
---------------------------------	------------------------------------

Description

A ponte de volta, do modulo `isco0888.sps` do ISMF.

Usage

```
isco08_para_isco88(isco08)
```

Arguments

`isco08` Vetor de codigos ISCO-08.

Value

Vetor de texto com o ISCO-88.

A ida e a volta nao se cancelam

Levar um código da ISCO-88 a ISCO-08 e trazê-lo de volta devolve o ponto de partida em apenas **69% dos casos**. Isso não é defeito de implementação: é propriedade da concordância da Organização Internacional do Trabalho, que reparte e funde categorias entre as duas revisões. Um terço dos códigos não tem volta ao mesmo lugar, e quem encadeia as duas pontes precisa saber disso.

Examples

```
isco08_para_isco88(c("2211", "5412"))
```

isco08_para_isei_br	<i>ISEI-BR a partir da ISCO-08</i>
---------------------	------------------------------------

Description

Entrada direta para quem ja tem o codigo internacional. Leia [tse_para_isei_br\(\)](#) para o que a regua e e o que ela nao e.

Usage

```
isco08_para_isei_br(isco08)
```

Arguments

isco08 Vetor de codigos ISCO-08.

Value

Vetor numerico com o escore ISEI-BR, entre 10 e 90.

See Also

[tse_para_isei_br\(\)](#), [isco08_isei_br](#)

Examples

```
isco08_para_isei_br(c("2211", "5411"))
```

isco88_isco08	<i>Ponte da ISCO-88 para a ISCO-08</i>
---------------	--

Description

Ponte da ISCO-88 para a ISCO-08

Usage

```
isco88_isco08
```

Format

`data.frame` com as colunas:

isco88 código de origem.

isco08 código de destino, já truncado como manda a sintaxe.

n_alternativas quantos destinos a Organização Internacional do Trabalho define para esse código de origem. Vale 1 quando a conversão é unívoca; acima disso, a tradução é uma escolha entre alternativas, e convém dizê-lo ao leitor.

Source

Módulo `isco8808.sps` do International Stratification and Mobility File. <http://www.harryganzeboom.nl/ismf/index.htm>

Examples

```
head(isco88_isco08)

# A ponte não é uma bijeção: 186 dos 530 códigos têm mais de um destino
# possível na ISCO-08, e a tábua escolhe um deles.
table(isco88_isco08$n_alternativas > 1)

# Peça a marca junto com a tradução quando a ambiguidade importar:
isco88_para_isco08("1229", com_ambiguidade = TRUE)
```

isco88_medidas

Medidas ancoradas na ISCO-88

Description

Medidas ancoradas na ISCO-88

Usage

`isco88_medidas`

Format

`data.frame` com as colunas:

isco88 código ISCO-88 de quatro dígitos, incluindo as formas arredondadas da hierarquia (2400 para o grande grupo 24).

isei88 escore ISEL.

siops88 escore de prestígio SIOPS, de Treiman.

egp classe EGP da tabela-base, **antes** das regras que dependem de posição na ocupação e supervisão. Não use esta coluna diretamente: use `isco88_para_egp()`, que aplica as regras.

rotulo rótulo em inglês, quando a sintaxe original o traz.

Source

Módulos `iskoisei.sps`, `iskotrei.sps`, `iskoroot.sps` e `iskolab.sps` do International Stratification and Mobility File, de Harry B. G. Ganzeboom e Donald J. Treiman. <http://www.harryganzeboom.nl/ismf/index.htm>

Examples

```
head(isco88_medidas)

# O ISEI e o prestígio ordenam parecido, mas não igual – e é na diferença
# que mora a escolha entre um e outro.
m <- isco88_medidas[stats::complete.cases(isco88_medidas[, c("isei88",
                                                             "siops88")]), ]
round(stats::cor(m$isei88, m$siops88, method = "spearman"), 3)
```

isco88_para_egp

Traduz ISCO-88 no esquema de classes EGP

Description

Implementação das sintaxes `iskopromo.sps` e `iskoegp.sps` do International Stratification and Mobility File, de Ganzeboom e Treiman.

Usage

```
isco88_para_egp(
  isco,
  conta_propria = NULL,
  n_supervisionados = NULL,
  n_classes = 11,
  rotulo = TRUE,
  avisar = TRUE
)
```

Arguments

<code>isco</code>	Vetor de códigos ISCO-88, preferencialmente com quatro dígitos. Códigos de 2 ou 3 dígitos são expandidos com zeros à DIREITA, para as formas arredondadas que as tabelas do ISMF já contêm (24 vira 2400). O grande grupo 0 (forças armadas) é a exceção: o ISMF o escreve com zero à ESQUERDA (0110), de modo que 110 é ambíguo. Nesses casos a função lê a forma hierárquica e avisa ; informe quatro dígitos para desempatar.
<code>conta_propria</code>	Vetor lógico opcional: a pessoa trabalha por conta própria ou é empregadora (TRUE) ou é empregada (FALSE).
<code>n_supervisionados</code>	Vetor numérico opcional com o número de pessoas supervisionadas.
<code>n_classes</code>	Número de classes do resultado: 11 (padrão), 7, 5 ou 3.

rotulo	Se TRUE (padrão), devolve rótulos — em todas as versões, inclusive nas colapsadas; se FALSE, os códigos inteiros.
avisar	Se TRUE (padrão), avisa quando conta_propria ou n_supervisionados não são fornecidos.

Value

Vetor de texto (ou inteiro, se rotulo = FALSE).

O que o EGP exige e o TSE não tem

O EGP não é uma função só da ocupação. As suas regras usam duas variáveis adicionais: a posição na ocupação (conta_propria) e o número de pessoas supervisionadas (n_supervisionados). Sem elas, **IVa e IVb — a pequena burguesia — ficam estruturalmente vazias**, e V (técnicos de nível inferior e supervisores manuais) sai fortemente subestimada, porque só dois códigos ISCO a produzem sem a variável de supervisão. Como o formulário do TSE pergunta apenas a ocupação, quem parte dele obtém uma versão degradada do esquema, e a função avisa quando é esse o caso.

Partindo só da ocupação, o resultado **não deve ser publicado como uma tabela EGP de onze classes**. Os colapsos de 5 e 3 classes, em que IVa e IVb se fundem a categorias que existem, são o uso defensável.

Essa não é uma limitação do pacote, e sim do dado: é exatamente o ponto que Carvalhaes (2015) levanta ao avaliar o EGP no Brasil, onde a categoria de conta própria é a que o esquema melhor capta — e é justamente a que se perde.

Os colapsos de 7, 5 e 3 classes

Vêm da tabela "The class schema" de Erikson e Goldthorpe (1992), pp. 38–39. Cuidado ao comparar com outras implementações: várias numeram IVc na sétima posição, e os vetores parecem divergir quando na verdade concordam.

References

- Erikson, R.; Goldthorpe, J. H. (1992). "The class schema", em *The Constant Flux: A Study of Class Mobility in Industrial Societies*. Oxford: Clarendon Press, pp. 38–39.
- Ganzeboom, H. B. G.; Treiman, D. J. (1996). Internationally comparable measures of occupational status for the 1988 International Standard Classification of Occupations. *Social Science Research*, 25(3), 201–239.
- Carvalhaes, F. (2015). A tipologia ocupacional Erikson-Goldthorpe-Portocarero (EGP): uma avaliação analítica e empírica. *Sociedade e Estado*, 30(3), 673–703.

Examples

```
isco88_para_egp(c("2211", "1300", "6100"), avisar = FALSE)

# com as variáveis que o esquema realmente pede, a pequena burguesia aparece
isco88_para_egp(c("5220", "5220"),
  conta_propria = c(TRUE, TRUE),
  n_supervisionados = c(0, 5))
```

isco88_para_isco08	<i>Converte ISCO-88 em ISCO-08</i>
--------------------	------------------------------------

Description

Aplica a conversão publicada no módulo `isco8808.sps` do ISMF, derivada da correspondência oficial da Organização Internacional do Trabalho.

Usage

```
isco88_para_isco08(isco88, com_ambiguidade = FALSE)
```

Arguments

<code>isco88</code>	Vetor de códigos ISCO-88.
<code>com_ambiguidade</code>	Se TRUE, devolve um <code>data.frame</code> com o código e o número de alternativas da OIT em vez de só o código.

Value

Vetor de texto com o ISCO-08, ou um `data.frame` se `com_ambiguidade = TRUE`.

A ponte é ambígua e a ambiguidade importa

A OIT define, para muitos códigos da ISCO-88, mais de um destino possível na ISCO-08. A sintaxe original guarda esse número na parte decimal e instrui a truncá-lo quando não houver informação adicional. O pacote trunca, como manda, mas **preserva a contagem** em `n_alternativas` na tabela [isco88_isco08](#), e [tse_para_isco08\(\)](#) pode devolvê-la. Cerca de um terço dos pares tem mais de uma alternativa: a conversão é uma escolha razoável, não um equivalente exato.

Examples

```
isco88_para_isco08(c("2211", "1300"))
isco88_para_isco08(c("2211", "1300"), com_ambiguidade = TRUE)
```

isco88_para_isei	<i>Índice socioeconômico ISEI a partir do ISCO-88</i>
------------------	---

Description

Índice socioeconômico ISEI a partir do ISCO-88

Usage

```
isco88_para_isei(isco88)
```

Arguments

`isco88` Vetor de codigos ISCO-88.

Value

Vetor numerico com o escore ISEI-88.

Examples

```
isco88_para_isei(c("2221", "5220", "9130"))
```

`isco88_para_prestigio` *Prestigio de Treiman a partir do ISCO-88*

Description

Prestigio de Treiman a partir do ISCO-88

Usage

```
isco88_para_prestigio(isco88)
```

Arguments

`isco88` Vetor de codigos ISCO-88.

Value

Vetor numerico.

Examples

```
isco88_para_prestigio("2221")
```

isco88_para_siops	<i>Prestigio ocupacional de Treiman (SIOPS) a partir do ISCO-88</i>
-------------------	---

Description

Prestigio ocupacional de Treiman (SIOPS) a partir do ISCO-88

Usage

```
isco88_para_siops(isco88)
```

Arguments

isco88	Vetor de codigos ISCO-88.
--------	---------------------------

Value

Vetor numerico com o escore de prestigio.

Examples

```
isco88_para_siops(c("2221", "5220", "9130"))
```

isei_retrospectivo	<i>Carrega o último escore conhecido da própria pessoa</i>
--------------------	--

Description

Registros administrativos repetidos — candidaturas do TSE, vínculos da RAIS, prontuários — trazem a mesma pessoa várias vezes, e às vezes sem ocupação classificável. Esta função preenche a lacuna com o **último escore observado da própria pessoa**, e devolve junto a marca de herança e a defasagem, para que a qualidade do preenchimento seja auditável em vez de invisível.

Usage

```
isei_retrospectivo(escore, id, tempo, cod = NULL, excluir = NULL)
```

Arguments

escore	Vetor numérico com a medida a carregar (ISEI, prestígio, o que for), com NA onde não há observação.
id	Vetor identificador da pessoa, do mesmo comprimento.
tempo	Vetor numérico ordenável — ano da eleição, data do vínculo.
cod	Vetor opcional de códigos de ocupação, usado só por excluir.
excluir	Códigos que não recebem carregamento. O padrão NULL não exclui nada; no dado do TSE, "931" (estudante e estagiário) é a escolha defensável.

Value

data.frame com escore (o valor final), herdado (lógico) e defasagem (quantas unidades de tempo desde a observação de origem).

Por que isto não é imputação estatística

Nada é modelado nem estimado. Usa-se apenas a **história ocupacional da própria pessoa**, jamais patrimônio, partido, escolaridade ou qualquer covariável — de modo que o escore resultante continua independente das variáveis com que se vai cruzá-lo. Uma imputação por covariável criaria colinearidade por desenho; esta não cria.

As três regras

só o passado usar aparições futuras descreveria a posição no momento do registro por uma ocupação que a pessoa ainda não tinha.

excluir fica de fora há posições em que a recuperação pelo passado é mínima e o apelo estaria justamente no futuro que a primeira regra proíbe — estudante e estagiário são o caso típico.

a herança é declarada herdado e defasagem saem no resultado. Um escore herdado de doze anos atrás não vale o mesmo que um de quatro, e quem analisa precisa poder condicionar ou excluir.

O que ele erra, medido

O carregamento supõe que a ocupação da pessoa não mudou. Ela muda com frequência. Entre os **684.179** pares de candidaturas consecutivas da mesma pessoa em que o ISEI foi **observado nas duas pontas** — isto é, onde dá para conferir:

52,2% mudam de código de ocupação.

45,2% mudam de escore ISEI.

18,9 pontos é a diferença absoluta média **quando** muda.

Ou seja: mesmo com defasagem zero, o escore herdado estaria errado em cerca de 45% dos casos, e por uma margem grande. E a defasagem raramente é zero — dos escores herdados, a mediana é de **4 anos**, **34,6%** vêm de 8 anos ou mais, e o máximo observado é 26.

Pior que o tamanho é a **direção**: o erro aponta sempre para o passado. Para quem ascendeu — a trajetória típica de profissionalização política — o carregamento puxa a posição para baixo. O viés não é ruído que se cancela na média; ele é sistemático e corre contra justamente a quantidade que estudos de recrutamento e oferta eleitoral querem medir.

Por isso herdado e defasagem saem no resultado, e não como cortesia: **relate quantos escores são herdados, e rode a sua análise também só com os observados**. Se a conclusão depender dos herdados, ela depende de uma suposição que este parágrafo mostra ser falsa em quase metade dos casos.

O que ele resolve, medido

Nas candidaturas ao TSE de 1998 a 2024, a cobertura do ISEI sobe de 68,5% para 76,0% entre homens e de **52,9% para 58,8% entre mulheres**. A janela para aqui de propósito, e não por desatualização: a medição exige o painel por pessoa, que identifica a mesma candidatura em eleições diferentes, e a safra de 2026 ainda não tem esse painel montado. O restante do pacote vai a 2026; este número, não. O ganho é maior entre elas porque o padrão de ausência é fortemente generificado: 14,7% das candidatas declaram posição fora da PEA, contra 1,3% dos candidatos.

Examples

```
d <- data.frame(
  pessoa = c("A", "A", "A", "B", "B"),
  ano     = c(2012, 2016, 2020, 2016, 2020),
  isei    = c(70, NA, NA, NA, 45))
isei_retrospectivo(d$isei, d$pessoa, d$ano)
```

reguas	<i>Qual régua responde a qual pergunta</i>
--------	--

Description

Uma linha por medida de posição social oferecida pelo pacote, com a pergunta que ela responde, o que ela mede e quando **não** usá-la.

Usage

```
reguas
```

Format

- `data.frame` com nove linhas e as colunas:
- medida** nome curto da régua.
- tipo** continua, categorica ou ordinal.
- ancora** a classificação em que a régua está ancorada: ISCO-88, ISCO-08, a PNAD Contínua de 2025 ou o próprio cadastro do TSE.
- escala** o intervalo ou o número de categorias. Para as réguas contínuas é o intervalo **observado** nas tabelas do pacote, calculado no momento da geração, não digitado.
- funcao_tse** a porta de entrada pela ocupação declarada ao TSE.
- outras_portas** as demais funções que chegam à mesma régua, separadas por vírgula; NA quando a medida só existe pela porta do TSE.
- pergunta** a pergunta de pesquisa que a régua responde.
- mede** o que a régua mede, em uma linha.
- quando_ao_usar** a armadilha própria daquela régua — a razão pela qual ela é a escolha errada para certas perguntas.
- fonte** de onde a régua vem.
- ver** onde ler o assunto por inteiro.

Details

As régua não são intercambiáveis, e a facilidade de trocar uma pela outra é o principal risco de usar este pacote: a troca custa uma letra no nome da função e não produz erro nenhum. Escolha pela pergunta, e diga qual escolheu. Esta tabela existe para que esse critério seja consultável — filtrável por tipo, por âncora ou por porta de entrada — e não apenas legível em prosa.

O que a tabela não diz, porque não é por régua

Quatro regras valem para todas e não cabem em nenhuma linha. NA não é zero, e o padrão de ausência é fortemente generificado — cerca de 15% das mulheres declaram posição fora da PEA, contra pouco mais de 1% dos homens —, de modo que média por grupo exige reportar a cobertura junto. O argumento ano resolve a quebra de cadastro de 2002, e sem ele uma candidatura de 1998 é traduzida pelo dicionário errado em silêncio. `checa_cobertura()` e `checa_periodo()` são `opt-in` e devem ser chamadas antes da análise. E toda tradução válida é, no mínimo, a dois dígitos: pelo primeiro dígito a correspondência inverte classes inteiras.

Source

Destilada da documentação do próprio pacote — as seções de `tse_para_isei()`, `tse_para_isei_br()`, `tse_para_egp()`, `tse_para_estrato()` e `tse_para_componente_alta()`, e a vinheta `vignette("qual-regua")`. É a única tabela do pacote que não deriva de fonte externa, e por isso não entra na auditoria de proveniência. Gerada por `data-raw/12_gera_reguas.R`; o que a mantém em dia com o pacote é o teste que amarra as funções citadas aqui aos exports reais, nas duas direções.

See Also

`vignette("qual-regua")`, `crosswalk_tse()`

Examples

```
# As régua contínuas, e por onde se entra nelas
reguas[reguas$tipo == "continua", c("medida", "ancora", "funcao_tse")]

# Antes de publicar EGP a partir do TSE, leia isto:
cat(strwrap(reguas$quando_nao_usar[reguas$medida == "EGP"], 72), sep = "\n")

# Qual régua responde a uma pergunta sobre relação de emprego?
reguas[grep("emprego", reguas$pergunta), c("medida", "funcao_tse")]
```

tse_autorrotulo_patrimonio

Patrimônio por trás dos códigos que são autodescrição

Description

Quantis do patrimônio declarado, por código de ocupação e cargo disputado, para os dez códigos de `tse_codigos_autorrotulo` mais o 131 (ADVOGADO), que entra como contraexemplo ancorado. Sustenta o argumento de `tse_para_componente_alta()` sobre a assimetria de confiabilidade entre as duas metades da classe alta.

Usage

tse_autorrotulo_patrimonio

Format

data.frame com 8 colunas:

cod_tse código de ocupação do TSE, texto.

rotulo rótulo do código, vigente após 2002.

cargo cargo disputado, fator na ordem da hierarquia e não na alfabética. O nível "TODOS" **não é um cargo**: é a linha agregada do código, e existe porque uma razão entre quantis não se recupera das células.

ancorado TRUE para o 131, cujo rótulo pressupõe inscrição na OAB; FALSE para os dez autodeclarados.

n candidaturas com patrimônio declarado positivo na célula.

p10, mediana, p90 quantis do patrimônio, em reais.

Por que esta tabela existe

Ela não acrescenta fato novo — os números já estavam em ?tse_para_componente_alta. O que ela acrescenta é que eles passam a ser **executados**. Digitados, erraram duas vezes: entraram com a agregação de bens que somava cada bem duas vezes e sobreviveram intactos à correção de 09/2026, porque ninguém revisita um número que não roda. A auditoria de 05/09/2026 os encontrou ainda dobrados. Agora a documentação lê a tabela, e uma troca de fonte se propaga sozinha.

O contraste

Os dois regimes de rótulo têm gradiente por cargo — quem disputa posto mais alto é mais rico nos dois casos. O que os separa é a **dispersão interna**: na linha "TODOS", o ADVOGADO tem a **menor** razão entre p90 e p10 de todos os códigos da tabela (47,8), e os autodeclarados vão de 55,1 a 75,6. É um único caso ancorado contra cinco autodeclarados, então isto é uma ilustração e não um teste. Mas é esse contraste, e não a diferença de mediana, que torna o 257 menos confiável que o 131 de um jeito que nenhuma tabela de escores mostra.

Corte e sigilo

Células com menos de 30 candidaturas ficam de fora (atributo n_min). O que se publica são três quantis por célula: não há candidatura, não há município, não há ano. Os atributos anos e n_obs guardam o período e a amostra.

Source

Declaração de bens das candidaturas ao TSE, 2006–2024, agregada por data-raw/10_gera_autorrotulo.R. Mesma microbase e mesma deflação de [tse_dispersao_patrimonio](#); 2026 não entra porque a declaração de bens daquela safra ainda não está fechada.

See Also

[tse_para_componente_alta\(\)](#), que interpreta esta tabela; [tse_codigos_autorrotulo](#), a lista dos códigos; [tse_dispersao_patrimonio](#), para a dispersão dentro do nível de status.

Examples

```
p <- tse_autorrotulo_patrimonio
# o gradiente do "empresário" por cargo
p[p$cod_tse == "257", c("cargo", "n", "mediana")]

# a dispersão interna, autodeclarado contra ancorado
a <- p[p$cargo == "TODOS", ]
data.frame(a["rotulo"], a["ancorado"], razao = round(a$p90 / a$p10, 1))
```

tse_codigos_autorrotulo

Códigos cujo rótulo é autodescrição, sem registro externo que o valide

Description

Os códigos de ocupação do TSE que designam **propriedade ou empreendimento** — empresário, comerciante, industrial, proprietário de estabelecimento — e que, por isso, não são ancorados em nenhum registro. Servem para rodar uma análise com e sem eles, em uma linha.

Usage

```
tse_codigos_autorrotulo
```

Format

Vetor de texto com os códigos.

Value

Não é função, e sim um vetor character de comprimento 10, exportado como constante. Cada elemento é um código de ocupação do TSE, em texto com três dígitos, cujo rótulo é autodescrição de propriedade ou empreendimento. Serve de filtro: `x %in% tse_codigos_autorrotulo` marca as candidaturas cujo score ocupacional não tem registro externo que o ancore.

Por que estes e não outros

O critério é **a priori, não empírico**: existe um registro externo que precisa ser satisfeito para que alguém use o rótulo? "Advogado" pressupõe inscrição na OAB, "médico" pressupõe CRM; "empresário" não pressupõe nada. Não é uma afirmação sobre a honestidade de quem declara, e sim sobre a existência de uma trava externa.

Este critério é conceitual porque **o dado não o produz sozinho**. A dispersão de patrimônio dentro do código, por exemplo, não separa os dois grupos: é de 76 vezes (p90/p10) entre os que declaram

empresário e de 55 entre os comerciantes, que são autorrotulos, mas de 48 entre os advogados, que não é — e o comerciante fica **entre** os outros dois. Patrimônio é disperso em toda parte. O que distingue não é a dispersão bruta, é a ausência da trava.

Estes três números saem de [tse_autorrotulo_patrimonio](#) e são conferidos contra ela em `tests/testthat/test-numeros-doc`. A versão anterior dizia 80, 47 e 54: eram os valores da agregação que somava cada bem duas vezes, corrigida em 09/2026, e sobreviveram por estarem digitados.

Como usar

Os escores destes códigos não são inválidos — são menos confiáveis do que os ancorados, e a diferença não aparece na tabela, porque um ISEI de 68 é impresso do mesmo jeito venha de onde vier. A recomendação é tratá-los como análise de sensibilidade:

```
d$isei <- tse_para_isei(d$cod)
d$isei_ancorado <- ifelse(d$cod %in% tse_codigos_autorrotulo, NA, d$isei)
# rode a sua análise com as duas colunas e relate as duas
```

São **13,3%** das candidaturas de 1998 a 2026 — o bastante para mover um resultado, e por isso o bastante para valer o teste.

Relação com `tse_isco$proprietario`

Hoje os dois conjuntos coincidem, mas respondem a perguntas diferentes, e por isso são objetos diferentes: `proprietario` é **pertença de classe** (define quem entra em "Proprietários e empregadores"); este é **confiabilidade de medida** (define de quem o escore é autodeclarado). É a mesma distinção que separou `proprietario` de `conta_propria`, e pela mesma razão: quando dois conceitos compartilham um vetor, mexer num arrasta o outro.

See Also

[tse_para_componente_alta\(\)](#), que documenta a assimetria de confiabilidade entre as duas metades da classe alta.

Examples

```
tse_codigos_autorrotulo
tse_para_rotulo(tse_codigos_autorrotulo)
```

tse_diff_cadastro

O que mudou no cadastro entre duas eleições

Description

Devolve os códigos **extintos**, os **criados** e os que **trocaram de rótulo** entre dois anos. É a função que falta a quem monta série longa: a reutilização de código é o problema mais citado, mas o mais frequente é a troca de inventário. Entre 2000 e 2002 o TSE aposentou 21 códigos, criou 69 e renomeou 36.

Usage

```
tse_diff_cadastro(ano1, ano2)
```

Arguments

ano1, ano2 Anos de eleição a comparar.

Value

data.frame com cod_tse, mudanca (extinto, criado ou rotulo), e os rótulos dos dois anos.

O que conta como "em vigor"

A comparação usa a **vigência** registrada em [tse_ocupacao_rotulos](#), não a presença de candidaturas naquela eleição específica. Um código que existe no cadastro mas não teve candidato num ano continua em vigor — o contrário faria uma ocupação rara "sumir e voltar" a cada eleição.

Por isso os números diferem de uma contagem por período. Comparando *qualquer ano até 2000* com *qualquer ano a partir de 2002*, são 13 códigos extintos (12,4% das candidaturas do período antigo) e 122 criados (33,2% do novo). Comparando as eleições de 2000 e 2002 uma com a outra, são 21 e 69. As duas leituras estão certas e respondem a perguntas diferentes.

Ausência não é revogação, e 2026 mostra isso

tse_diff_cadastro(2024, 2026) devolve 46 códigos como **extintos** e nenhum criado. Os 46 não foram extintos. A vigência desta tabela se constrói do rótulo **observado**, de modo que um código sem nenhum candidato numa eleição registra ausência — e ausência é o que se espera de um código raro quando a safra é pequena. A de 2026 é pequena duas vezes: é geral, sem prefeito nem vereador, e está **aberta**, com o Tribunal ainda julgando registros. Trocar de par não resolve: tse_diff_cadastro(2022, 2026), que compara duas gerais, devolve 47 extintos, e não menos.

A metade que se sustenta é a outra: **nenhum código criado**, nos dois pares. Essa é robusta ao tamanho da safra, porque um código novo apareceria ainda que uma vez só. Leia a contagem de extintos de uma safra aberta como o que ela é — quem ainda não apareceu —, e refaça a comparação quando o cadastro de 2026 fechar.

Examples

```
d <- tse_diff_cadastro(2000, 2002)
table(d$mudanca)
```

tse_dispersao_patrimonio

Dispersão do patrimônio dentro de cada nível de status

Description

Somatórios que reproduzem, sem microdado, a correlação entre o índice de status e o patrimônio declarado **no nível da candidatura** — o número que quantifica o que uma escala de posição ocupacional não explica.

Usage

```
tse_dispersao_patrimonio
```

Format

data.frame com 29 linhas e 6 colunas:

isei88 escore ISEI-88, que é discreto: cada código do dicionário tem um valor só, e a tabela agrupa por ele.

n candidaturas com escore e com patrimônio declarado positivo.

soma_log soma do logaritmo natural do patrimônio no grupo.

soma_log2 soma dos quadrados desses logaritmos.

media_log média do log do patrimônio, arredondada.

sd_log desvio padrão do log do patrimônio DENTRO do grupo.

Por que somatórios e não os dados

O coeficiente de Pearson é função apenas de n , $\sum x$, $\sum y$, $\sum x^2$, $\sum y^2$ e $\sum xy$. Como o ISEI é constante dentro do grupo, essas seis quantidades se recuperam das colunas acima, e o coeficiente sai **exato** — não aproximado. O que a tabela descarta é tudo o que o coeficiente não usa, isto é, a candidatura individual. É o que permite publicar o número mantendo a política de não distribuir microdado de patrimônio.

O contraste que a tabela existe para sustentar

No nível da **ocupação**, o ISEI correlaciona-se a 0,681 com o log da mediana de patrimônio (veja [tse_validacao](#)). No nível da **candidatura**, a 0,207. A queda não é defeito de medida: é a definição operacional do que uma escala de posição ocupacional faz. E o "dentro" é exato, não aproximado: o ISEI é constante na ocupação, logo explica *zero* da variância intraocupacional. O que se mede é o quanto fica *entre* níveis de status — $\eta^2 = 0,077$, contra $r^2 = 0,043$ da relação linear. A coluna **sd_log** mostra o fenômeno diretamente: o desvio padrão do log do patrimônio dentro de um mesmo nível de status é de cerca de 1,7, ou seja, um fator de cinco a seis por desvio padrão — dois desvios já são trinta vezes. A consequência prática está em [tse_para_isei\(\)](#): **não use o ISEI como proxy de renda ou de patrimônio individual.**

A unidade é a candidatura, não a pessoa

As 1.103.019 observações são candidaturas, de 767.015 pessoas distintas entre 2006 e 2024 — a mesma pessoa entra até seis vezes. A documentação chamava isso de "nível do indivíduo" até 05/09/2026, o que convidava a tratar as observações como independentes. Para este número não há consequência, porque ele é descritivo e não vem com erro padrão. Mas não peça um intervalo de confiança a ele sem antes decidir o que fazer com a repetição.

Source

Declaração de bens das candidaturas ao TSE, 2006–2024, agregada por `data-raw/08_gera_dispersao.R`. O patrimônio é deflacionado a reais de outubro de 2024; 2026 não entra, pela razão exposta em [tse_validacao](#).

See Also

[tse_validacao](#), para a validação no nível da ocupação.

Examples

```
# A correlação individual, recuperada dos somatórios (exata):
d <- tse_dispersao_patrimonio
N <- sum(d$n); sx <- sum(d$isei88 * d$n); sy <- sum(d$soma_log)
sxx <- sum(d$isei88^2 * d$n); syy <- sum(d$soma_log2)
sxy <- sum(d$isei88 * d$soma_log)
round((N * sxy - sx * sy) /
      sqrt((N * sxx - sx^2) * (N * syy - sy^2)), 3)

# A dispersão dentro do nível de status, que é o mesmo fato visto de perto:
summary(d$sd_log)
```

tse_isco

*Dicionário de ocupações do TSE***Description**

Uma linha por código de ocupação (CD_OCUPACAO) das candidaturas.

Usage

```
tse_isco
```

Format

data.frame com as colunas:

cod_tse código de ocupação do TSE, como texto.

isco88 ISCO-88 de quatro dígitos; NA quando o código não designa ocupação (não informada, fora da PEA, vínculo sem função).

nivel número de dígitos da classificação de origem — 2, 3 ou 4. A maioria é a dois dígitos; desce-se a três ou quatro só onde dois fundiriam posições distantes demais, como médico e enfermeiro.

classe esquema de doze categorias para o dado eleitoral.

estrato classe alta, média, populares, ou uma das duas residuais.

componente_alta partição da classe alta em proprietária, credenciada e dirigentes; NA fora dela.

politico a ocupação é o próprio mandato ou cargo político.

proprietario o código nomeia explicitamente um proprietário ou empregador. Define a pertença à classe "Proprietários e empregadores" no esquema de classes.

conta_propria a pessoa trabalha por conta própria — o SEMPL = 2 que as sintaxes do ISMF exigem para o EGP. É um **superconjunto** de proprietario: o agricultor (601) e o pescador (604) trabalham por conta própria sem pertencerem à classe proprietária. As duas marcas eram uma só até julho de 2026, e enquanto foram, marcar o agricultor como conta própria — o que o EGP exige para chegar a IVc — o promovia junto à classe alta, que o patrimônio não sustenta. A evidência que separou as duas está em [isco_posicao_br](#).

Source

Construído a partir dos microdados de candidaturas do Tribunal Superior Eleitoral e da correspondência com a ISCO-88.

Examples

```
# O agricultor trabalha por conta própria sem ser proprietário: as duas
# marcas eram uma só até julho de 2026, e enquanto foram, o EGP o promovia
# a IVc junto à classe alta.
tse_isco[tse_isco$cod_tse == "601", ]

# Quantos códigos o dicionário reconhece como o próprio mandato:
tse_isco$cod_tse[tse_isco$politico]
```

tse_ocupacao_rotulos *Vigências do cadastro de ocupações do TSE, 1998–2026*

Description

Uma linha por **vigência**: o período em que um código carregou um dado rótulo. Um código que nunca mudou de nome tem uma linha; um que mudou tem uma por período. É a diferença entre um dicionário e a história de um cadastro.

Usage

```
tse_ocupacao_rotulos
```

Format

data.frame com 334 linhas e as colunas:

cod_tse código de ocupação.

de, ate primeiro e último ano de eleição em que a vigência valeu.

rotulo o DS_OCUPACAO como o TSE o escreveu, forma modal do período.

n candidaturas na vigência.

Vigência não é presença

Um código sem candidato numa eleição continua na vigência: o contrário faria uma ocupação rara "sumir e voltar" a cada pleito. A vigência se interrompe quando o **rótulo** muda, não quando a frequência cai a zero.

Source

DS_OCUPACAO dos arquivos consulta_cand do TSE, 1998–2026 (3.369.244 candidaturas, campo 100% preenchido nas 15 eleições). A safra de 2026 é a geração de 01/09/2026, 12:31, e é **aberta**: o prazo de registro encerrou em 15/08/2026, mas o Tribunal ainda julga e publica candidaturas, de modo que o n de 2026 há de crescer. O rótulo, que é o que esta tabela guarda, não depende disso.

Examples

```
# 0 que as candidaturas mais declaram, em 1998--2026 somados:
r <- tse_ocupacao_rotulos
head(r[order(-r$n), c("cod_tse", "rotulo", "n")], 5)

# Os códigos vigentes na safra em curso:
sum(r$ate == 2026)

# Um código pode ter mais de uma vigência, com rótulos diferentes:
r[r$cod_tse == "215", ]
```

tse_para_classe	<i>Classe social da ocupação declarada ao TSE</i>
-----------------	---

Description

Esquema de doze categorias desenhado para o dado eleitoral brasileiro. Não é o EGP: parte do ISCO como o EGP parte, mas trata como categorias próprias duas situações que o registro eleitoral cria e que os esquemas canônicos não preveem — proprietários e empregadores nomeados pelo próprio código de ocupação, e o vínculo público declarado sem função especificada. Para o EGP canônico, veja [tse_para_egp\(\)](#).

Usage

```
tse_para_classe(cod, superior = NULL, ano = NULL)
```

Arguments

cod	Vetor de códigos de ocupação do TSE (numérico ou texto).
superior	Vetor lógico opcional, do mesmo comprimento de cod, indicando ensino superior completo. Quando fornecido, a categoria residual "Vínculo público não especificado" — que sozinha reúne cerca de 9% das candidaturas, do faxineiro ao secretário de finanças — é dividida pela escolaridade declarada.
ano	Vetor opcional de anos de eleição, do mesmo comprimento de cod. Com ele, as candidaturas cujo código o TSE reutilizou depois voltam NA com aviso, em vez de traduzidas pelo dicionário errado: o código 215 designava um cargo de direção até 2000 e passou a designar artista plástico. Sem ele, o comportamento é o de sempre. Veja tse_vigencia() .

Value

Vetor de texto com a classe.

O agricultor, e por que ele não tem classe própria

Os códigos 601 (agricultor) e 604 (pescador) ficam em "Trabalhadores rurais", portanto nas classes populares, embora `tse_para_egp()` os classifique como IVc: proprietário rural. A aparente discordância entre as duas medidas é **artefato de granularidade**, e o registro é útil porque a classe própria chegou a ser criada, em 29/07/2026, e foi revertida no mesmo dia.

O esquema relacional separa IVc de VIIb nas **onze** e nas **sete** classes. Nos dois colapsos mais fortes de Erikson e Goldthorpe, o agricultor e o assalariado rural voltam a ser a mesma coisa:

n_classes	601 agricultor	606 assalariado rural
11	IVc	VIIb
7	IVc	VIIb
5	IVc+VIIb	IVc+VIIb
3	Agrícolas	Agrícolas

A linha de sete classes estava ausente desta tabela até 05/09/2026, e a ausência importava: ela é o colapso mais usado para publicar, e nela a distinção **sobrevive**. Comparar um esquema de onze classes com uma partição em três estratos é comparar resoluções diferentes, não encontrar divergência. Só nas resoluções equivalentes à do estrato o próprio EGP funde os dois.

O dado externo concorda. O ISEI do agricultor, 23, está **dentro** da faixa das classes populares, que vai de 16 a 43, e corresponde ao percentil 5 do dicionário. O patrimônio mediano declarado, de R\$ 122.395, fica dentro da faixa das classes populares, cujo máximo é R\$ 126.801 – só um dos sessenta demais códigos do estrato declara mais: põe o agricultor no topo da classe popular, não fora dela.

A distinção entre agricultura familiar e proletariado rural é real, e continua disponível onde Erikson e Goldthorpe a puseram: em `tse_para_egp(n_classes = 11)`. O que não se justifica é transportá-la para a partição em estratos.

Não há equivalente pela COD, e o motivo é do esquema

`cod_para_classe()` não existe, e não é omissão. Quatro das categorias deste esquema — "Vínculo público não especificado", "Fora da PEA por posição", "Inativo com trajetória" e "Não informado" — são **rubricas do formulário do TSE**, e não posições da ISCO. Elas existem porque o cadastro do Tribunal oferece à pessoa rótulos que *substituem* a ocupação em vez de a nomear. A PNAD Contínua e o Censo não têm nada disso: lá, vínculo e ocupação são perguntas separadas, e quem é servidor público **também** declara a ocupação. Traduzir o esquema para a COD exigiria inventar do lado da população categorias que só existem por causa de um formulário.

A consequência prática é direta: **este esquema não atravessa para a população**, e quem precisa comparar composição de classe entre as duas fontes não tem régua pronta. O caminho é comparar pelo ISEI, que é literalmente o mesmo construto dos dois lados, ou construir uma régua própria declarando as categorias que só existem de um lado. O tamanho do problema está medido em [tse_universo](https://moraespeixoto.github.io/ocupacoesBR/articles/comparar-fontes.html) e o artigo *Comparar duas fontes* desenvolve as alternativas: <https://moraespeixoto.github.io/ocupacoesBR/articles/comparar-fontes.html>

Examples

```
tse_para_classe(c(111, 169, 291))
tse_para_classe(c(291, 291), superior = c(TRUE, FALSE))
```

tse_para_componente_alta

Partição da classe alta em proprietária, credenciada e dirigente

Description

Separa a classe alta nas duas espécies de recurso que a compõem — o capital e a credencial — além das ocupações propriamente políticas. Devolve NA fora da classe alta.

Usage

```
tse_para_componente_alta(cod, ano = NULL)
```

Arguments

cod	Vetor de códigos de ocupação do TSE (numérico ou texto).
ano	Vetor opcional de anos de eleição, do mesmo comprimento de cod. Com ele, as candidaturas cujo código o TSE reutilizou depois voltam NA com aviso, em vez de traduzidas pelo dicionário errado: o código 215 designava um cargo de direção até 2000 e passou a designar artista plástico. Sem ele, o comportamento é o de sempre. Veja tse_vigencia() .

Value

Vetor de texto, ou NA fora da classe alta.

As duas metades não têm a mesma confiabilidade

A partição é simétrica no desenho e **assimétrica na medição**, e quem publica o contraste precisa dizê-lo.

A **alta credenciada** é ancorada em registro externo. "Advogado" não é autodescrição: pressupõe inscrição na OAB. Por isso o rótulo se comporta de modo estável — entre os que declaram o código 131, a proporção com ensino superior varia pouco mais de um ponto entre vereador e presidente (98,6% a 100%).

A **alta proprietária** não tem âncora nenhuma. "Empresário" (código 257) é autodeclaração: nenhum registro precisa existir para que alguém se descreva assim. E o mesmo rótulo cobre posições materialmente incomparáveis — entre os que declaram 257, o patrimônio mediano varia por um fator de **25** conforme o cargo disputado:

cargo	patrimônio mediano de quem declara 257
Vereador	R\$ 185.000
Prefeito	R\$ 812.544
Senador	R\$ 4.678.498

Tudo isso sob um **único** ISEI (68) e uma única classe. Dentro do código, os extremos de patrimônio distam 76 vezes (p10 R\$ 19,8 mil, p90 R\$ 1,50 milhão). A tabela acima e esses dois quantis

saem de [tse_autorrotulo_patrimonio](#), e não estão digitados aqui — a versão anterior deles estava, e sobreviveu errada à correção do patrimônio de 09/2026.

O 257 não é uma ocupação medida com erro: é uma **mistura** de duas populações — o microempregado e o capitalista — sob um rótulo só. Nenhum escore único está certo para as duas, razão pela qual mudá-lo de valor não resolve (veja [tse_codigos_autorrotulo](#)).

Consequência: um gráfico que compare os dois componentes está comparando uma quantidade ancorada em registro externo a uma quantidade autodeclarada que agrega posições muito distantes. Teoricamente a partição é boa (é Bourdieu e Wright operacionalizados), mas Wright insistia em *employment relations*, não em rótulos.

O que o dado NÃO mostra, e vale dizer para que ninguém repita: não há evidência de que as pessoas troquem de rótulo conforme o cargo. A frequência do 257 não cresce com a importância do posto — tem pico em prefeito (10,6%) e cai em senador (7,2%) e governador (6,9%). Quem cresce monotonicamente é "advogado" (1,6% a 14,8%), que é o caso ancorado, e "comerciante" **cai** de 6,2% a 0%. O gradiente de patrimônio acima é consistente tanto com recrutamento seletivo quanto com relabeling, e estes dados não separam as duas hipóteses.

Se a sua conclusão depende do tamanho relativo dos dois componentes, rode-a também sem [tse_codigos_autorrotulo](#) e relate as duas versões.

Examples

```
tse_para_componente_alta(c(257, 111, 274, 411))
```

tse_para_egp

Classe EGP da ocupação declarada ao TSE

Description

Atalho de [tse_para_isco\(\)](#) seguido de [isco88_para_egp\(\)](#). Leia [isco88_para_egp\(\)](#) antes de usar: partindo só do TSE, o EGP sai incompleto, porque o formulário não pergunta supervisão.

Usage

```
tse_para_egp(
  cod,
  conta_propria = NULL,
  n_supervisionados = NULL,
  n_classes = 11,
  rotulo = TRUE,
  avisar = TRUE,
  usa_proprietario = TRUE,
  ano = NULL
)
```

Arguments

<code>cod</code>	Vetor de códigos de ocupação do TSE (numérico ou texto).
<code>conta_propria</code>	Vetor lógico opcional: a pessoa trabalha por conta própria ou é empregadora (TRUE) ou é empregada (FALSE).
<code>n_supervisionados</code>	Vetor numérico opcional com o número de pessoas supervisionadas.
<code>n_classes</code>	Número de classes do resultado: 11 (padrão), 7, 5 ou 3.
<code>rotulo</code>	Se TRUE (padrão), devolve rótulos — em todas as versões, inclusive nas colapsadas; se FALSE, os códigos inteiros.
<code>avisar</code>	Se TRUE (padrão), avisa quando <code>conta_propria</code> ou <code>n_supervisionados</code> não são fornecidos.
<code>usa_proprietario</code>	Se TRUE (padrão) e <code>conta_propria</code> não for informado, usa a marca <code>proprietario</code> do dicionário como posição no emprego. Informar <code>conta_propria</code> explicitamente tem precedência.
<code>ano</code>	Vetor opcional de anos de eleição, do mesmo comprimento de <code>cod</code> . Com ele, as candidaturas cujo código o TSE reutilizou depois voltam NA com aviso, em vez de traduzidas pelo dicionário errado: o código 215 designava um cargo de direção até 2000 e passou a designar artista plástico. Sem ele, o comportamento é o de sempre. Veja tse_vigencia() .

Value

Vetor de texto (ou inteiro, se `rotulo = FALSE`).

A posição no emprego que o dicionário conhece

O TSE não pergunta posição na ocupação, mas alguns dos seus códigos **nomeiam** no próprio rótulo quem trabalha por conta própria — comerciante, empresário, pecuarista, proprietário de estabelecimento, e também o agricultor e o pescador. O dicionário marca essa condição em `tse_isco$conta_propria`, e é ela, não `proprietario`, que corresponde ao SEMPL = 2 das sintaxes do ISMF. As duas marcas são distintas desde 07/2026: `proprietario` é o subconjunto que pertence à classe proprietária, e o agricultor (601) e o pescador (604) estão fora dele sem deixarem de trabalhar por conta própria. Enquanto eram uma coisa só, os dois caíam em VIIb, o fundo do esquema, ao lado do assalariado rural.

Por padrão (`usa_proprietario = TRUE` — o nome do argumento é anterior à separação das duas marcas e foi mantido por compatibilidade) a função a utiliza. São doze os códigos marcados, e sete deles mudam de classe por causa dela. Cinco — 169, 902, 903, 904 e 905 — saíam em II, a classe de serviço assalariada, o que é uma inversão: a pequena burguesia contada como classe de serviço. Os outros dois — 601 e 604 — saíam em VIIb, ao lado do assalariado rural.

O que **continua** indeterminado é a divisão entre IVa (conta própria com empregados) e IVb (sem), que exige o número de subordinados. Todos caem em IVb. Para publicar, prefira `n_classes = 7` ou 5, onde as duas se fundem em IVab.

Examples

```
# 169 é COMERCIANTE: pequena burguesia, não classe de serviço
tse_para_egp(c(111, 169, 601), avisar = FALSE)
tse_para_egp(169, usa_proprietario = FALSE, avisar = FALSE)
```

tse_para_estrato	<i>Estrato social da ocupação declarada ao TSE</i>
------------------	--

Description

Agrega as classes de [tse_para_classe\(\)](#) em classe alta, classe média e classes populares, mantendo à parte as duas categorias residuais (vínculo público sem função e fora da PEA), que não são estratos e não devem ser silenciosamente somadas a nenhum deles.

Usage

```
tse_para_estrato(cod, ano = NULL)
```

Arguments

cod	Vetor de códigos de ocupação do TSE (numérico ou texto).
ano	Vetor opcional de anos de eleição, do mesmo comprimento de cod. Com ele, as candidaturas cujo código o TSE reutilizou depois voltam NA com aviso, em vez de traduzidas pelo dicionário errado: o código 215 designava um cargo de direção até 2000 e passou a designar artista plástico. Sem ele, o comportamento é o de sempre. Veja tse_vigencia() .

Value

Vetor de texto com o estrato.

O estrato e o ISEI se cruzam por desenho

O estrato é uma régua de **posição** — de onde vem o sustento: de um diploma, de um patrimônio, de um salário, da terra. O ISEI ([tse_para_isei\(\)](#)) é uma régua de **status**, que ordena ocupações pela eficiência com que convertem escolaridade em renda. Que discordem em alguns pontos não é defeito: é a razão de haver duas. Um esquema de classes que apenas reordenasse o ISEI em fatias seria redundante, e a distinção entre classe e status é justamente o que Weber propôs e Goldthorpe endureceu.

Medidas sobre os 258 códigos do dicionário, as duas régua concordam quase inteiramente — os intervalos interquartis não se tocam:

estrato	n	mín	mediana	máx
Classe alta	88	43	68	90
Classe média	69	45	50	55
Classes populares	101	16	33	43

Nenhum código de classe média tem ISEI acima da mediana da classe alta. Quatro dos 88 códigos de classe alta ficam abaixo da mediana da classe média, e cada um tem explicação própria:

- **113 (enfermeiro), ISEI 43** — é a anomalia de gênero da própria escala, documentada em `tse_para_isei()`: enfermagem recebe menos que o técnico de enfermagem (48) e que o escriturário (45). Aqui é o ISEI que erra, não a classe. Enfermagem é bacharelado regulado, e a própria ISCO-88 a põe no grande grupo 2.
- **234, 602 e 901 (produtor agropecuário, pecuarista, proprietário agrícola), ISEI 43** — são classe alta por propriedade, com status baixo. É exatamente o contraste que `tse_para_componente_alta()` existe para exibir: capital econômico e capital cultural não se acompanham.

A leitura prática: uma discordância entre as duas réguas é achado a interpretar, não erro a corrigir. Quem publicar as duas no mesmo gráfico precisa dizer qual pergunta cada uma responde — veja vignette("qual-regua").

Examples

```
tse_para_estrato(c(111, 169, 291, 931))
```

tse_para_isco

Traduz a ocupação declarada ao TSE em ISCO-88

Description

Converte o código de ocupação da candidatura (CD_OCUPACAO) no código da Classificação Internacional Uniforme de Ocupações de 1988, com quatro dígitos.

Usage

```
tse_para_isco(cod, ano = NULL)
```

Arguments

cod	Vetor de códigos de ocupação do TSE (numérico ou texto).
ano	Vetor opcional de anos de eleição, do mesmo comprimento de cod. Com ele, as candidaturas cujo código o TSE reutilizou depois voltam NA com aviso, em vez de traduzidas pelo dicionário errado: o código 215 designava um cargo de direção até 2000 e passou a designar artista plástico. Sem ele, o comportamento é o de sempre. Veja <code>tse_vigencia()</code> .

Value

Vetor de texto com o ISCO-88 de quatro dígitos. NA para códigos que não designam ocupação (não informada, fora da PEA, vínculo sem função).

A armadilha do primeiro dígito

Traduzir ocupação pelo primeiro dígito **inverte classes inteiras**. O grande grupo 9 da CBO brasileira é reparação e manutenção (mecânicos, ISEI em torno de 34); o grande grupo 9 da ISCO é o das ocupações elementares (serventes, ISEI entre 16 e 30). Toda tradução válida é, no mínimo, a dois dígitos — que é o nível em que este pacote opera, descendo a três ou quatro onde dois fundiriam posições distantes demais (médico e enfermeiro, por exemplo).

See Also

[tse_para_isei\(\)](#), [tse_para_classe\(\)](#), [crosswalk_tse\(\)](#)

Examples

```
tse_para_isco(c(111, 169, 257))
```

tse_para_isco08	<i>ISCO-08 da ocupação declarada ao TSE</i>
-----------------	---

Description

ISCO-08 da ocupação declarada ao TSE

Usage

```
tse_para_isco08(cod, com_ambiguidade = FALSE, ano = NULL)
```

Arguments

cod	Vetor de códigos de ocupação do TSE (numérico ou texto).
com_ambiguidade	Se TRUE, devolve um <code>data.frame</code> com o código e o número de alternativas da OIT em vez de só o código.
ano	Vetor opcional de anos de eleição, do mesmo comprimento de <code>cod</code> . Com ele, as candidaturas cujo código o TSE reutilizou depois voltam NA com aviso, em vez de traduzidas pelo dicionário errado: o código 215 designava um cargo de direção até 2000 e passou a designar artista plástico. Sem ele, o comportamento é o de sempre. Veja tse_vigencia() .

Value

Vetor de texto, ou um `data.frame` se `com_ambiguidade = TRUE`.

Não é a mera composição das duas etapas

Em 21 dos 275 códigos o resultado difere de `isco88_para_isco08(tse_para_isco(cod))`, porque o código do TSE carrega informação que a ponte genérica não vê: a marca de proprietário — o SEMPL do ISMF, que aciona a promoção de `iskopromo.sps` — e um rótulo mais fino do que o código ISCO-88 agregado a que o dicionário o associa. A ponte `isco88_para_isco08()` permanece intacta e continua devolvendo o destino oficial da OIT para quem entra pela ISCO-88. Veja `NEWS.md` da versão 0.2.1 para a lista dos 21, com valor antigo, valor novo e fonte de cada um.

Examples

```
tse_para_isco08(c(111, 169))
# o proprietário rural não é rebaixado a trabalhador agrícola:
tse_para_isco08(234)
isco88_para_isco08("1311")
```

tse_para_isei	<i>Índice socioeconômico ISEI da ocupação declarada ao TSE</i>
---------------	--

Description

Devolve o escore ISEI (International Socio-Economic Index of occupational status) de Ganzeboom, De Graaf e Treiman, ancorado na ISCO-88.

Usage

```
tse_para_isei(cod, ano = NULL)
```

Arguments

cod	Vetor de códigos de ocupação do TSE (numérico ou texto).
ano	Vetor opcional de anos de eleição, do mesmo comprimento de cod. Com ele, as candidaturas cujo código o TSE reutilizou depois voltam NA com aviso, em vez de traduzidas pelo dicionário errado: o código 215 designava um cargo de direção até 2000 e passou a designar artista plástico. Sem ele, o comportamento é o de sempre. Veja <code>tse_vigencia()</code> .

Value

Vetor numérico com o escore ISEI-88, ou NA. **Nem todo escore tem a mesma procedência.** Um ISEI de 68 é impresso do mesmo jeito venha de "advogado", que pressupõe inscrição na OAB, ou de "empresário", que não pressupõe registro nenhum e cobre desde o microempreendedor até o capitalista. Veja `tse_codigos_autorrotulo` para rodar a sua análise com e sem esses códigos — são 13,3% das candidaturas.

Anomalias conhecidas da escala

O ISEI não é constante da natureza: é o escalonamento que maximiza o efeito indireto da escolaridade sobre a renda **via ocupação**, estimado por Ganzeboom, De Graaf e Treiman (1992) sobre dados internacionais dos anos 1970–80. Duas consequências para quem usa a versão ancorada na ISCO-88.

A escala é enviesada contra ocupações femininas. Regressão no nível do código, ponderada por número de candidaturas (168 códigos com $n \geq 500$, dos quais 28 são majoritariamente femininos):

$$\text{ISEI} = 38,6 + 0,469 \times (\% \text{ com superior}) - 6,17 \times (\text{código majoritariamente feminino})$$

(ep 3,03; $p = 0,043$)

A credencial constante, um código feminino recebe 6,2 pontos de ISEI a menos. E não é que essas ocupações tenham menos escolaridade — têm mais: 29,9% de superior contra 24,1%, com ISEI médio de 46,0 contra 49,0.

O sinal é robusto; o p não. A especificação acima pondera por número de candidaturas, o que muda a população-alvo de códigos para candidaturas — é escolha defensável, mas o erro-padrão de 1m ponderado não a descreve, e o ISEI é atributo determinístico do código, não estimativa com variância inversa a n . Trocando a especificação: com erro-padrão robusto (HC3) o coeficiente fica em -6,2 com $p = 0,14$; sem ponderação, -5,3 com $p = 0,012$; com `pct_mulher` contínuo em vez da dicotomia, -0,09 por ponto percentual com $p = 0,08$. O sinal e a ordem de grandeza sobrevivem a todas; o $p = 0,043$ é da primeira. Use o intervalo de especificações, não o p pontual.

E "viés" pede uma premissa que convém explicitar. O ISEI pondera escolaridade e renda, e ocupações femininas rendem menos com credencial igual: a escala registra isso fielmente. Chamar de viés supõe que status *deveria* seguir a credencial. Se o que se quer medir é credencial, é viés; se é retorno, é fidelidade. A recomendação abaixo vale nos dois casos, porque a ISCO-08 reestima o retorno sobre uma população que inclui mulheres.

O caso emblemático é a enfermagem, que a ISCO-88 põe em 2230 com **ISEI 43** — abaixo dos escriturários (4100, ISEI 45), apesar de ser profissão universitária. A âncora da ISCO-08 corrige boa parte disso: o mesmo posto recebe **68,7**.

Dá a recomendação prática: **para análise de gênero, prefira `tse_para_isei08()`**. Ganzeboom rebalanceou saúde e cuidado na revisão de 2008. Isto qualifica a defesa de que a ISCO-88 é "a escolha conservadora": ela está certa no sinal, mas é conservadora *porque* a escala é enviesada contra ocupações femininas, o que não é a mesma coisa que ser robusta.

A relação com o critério externo não é linear nem monótona. Veja `vignette("validacao")`: a correlação entre ISEI e patrimônio é de 0,681 no nível da ocupação e de apenas 0,207 no da candidatura, valor que `tse_dispersao_patrimonio` permite recalcular. Uma medida de posição ocupacional explica a variância *entre* ocupações e quase nada *dentro* de cada uma — por isso **não use o ISEI como proxy de renda individual**.

References

Ganzeboom, H. B. G.; De Graaf, P. M.; Treiman, D. J. (1992). A standard international socio-economic index of occupational status. *Social Science Research*, 21(1), 1–56.

See Also

`tse_para_isei08()`, para a âncora da ISCO-08; `isco_posicao_br`, para a posição na ocupação; `tse_codigos_autorrotulo`, para a sensibilidade aos rótulos sem âncora; `vignette("qual-regua")`.

Examples

```
tse_para_isei(c(111, 169, 257))
```

tse_para_isei_br	<i>ISEI-BR: status ocupacional estimado em dado brasileiro</i>
------------------	--

Description

Escore de status ocupacional obtido refazendo o procedimento de Ganzeboom, De Graaf e Treiman (1992) sobre a PNAD Continua de 2025, em vez de importar os escores calculados por eles em dado estrangeiro. Esta funcao entra pela ocupacao declarada ao TSE; as irmas entram pela COD ([cod_para_isei_br\(\)](#)), pela CBO-2002 ([cbo2002_para_isei_br\(\)](#)) e pela propria ISCO-08 ([isco08_para_isei_br\(\)](#)). Nao ha porta pela CBO-94, e a razao esta em [cbo94_isco88](#).

Usage

```
tse_para_isei_br(cod, ano = NULL)
```

Arguments

cod	Vetor de códigos de ocupação do TSE (numérico ou texto).
ano	Vetor opcional de anos de eleição, do mesmo comprimento de cod. Com ele, as candidaturas cujo código o TSE reutilizou depois voltam NA com aviso, em vez de traduzidas pelo dicionário errado: o código 215 designava um cargo de direção até 2000 e passou a designar artista plástico. Sem ele, o comportamento é o de sempre. Veja tse_vigencia() .

Value

Vetor numerico com o escore ISEI-BR, entre 10 e 90.

O que esta regua e e o que ela nao e

O ISEI trata a ocupacao como o meio pelo qual a escolaridade se converte em renda, e da a cada ocupacao o escore que melhor cumpre esse papel de intermediario. O ISEI-88 e o ISEI-08 que este pacote tambem oferece foram estimados assim, mas em dado de outros paises. O ISEI-BR e o mesmo metodo estimado aqui, sobre 655.787 observacoes de 339.181 pessoas.

Ele **nao substitui** o ISEI-08. Comparacao internacional continua exigindo a ancora internacional, e uma serie nao pode trocar de regua no meio. O ISEI-BR responde a outra pergunta: como o mercado de trabalho brasileiro de hoje, e nao um escalonamento unico de dezesseis paises com dado brasileiro de 1973 e 1982, ordena as ocupacoes.

Duas coisas precisam ser ditas por quem usa. A primeira e que a mediacao nao e completa: no Brasil a ocupacao explica cerca de 58% do efeito da escolaridade sobre a renda, e o resto e escolaridade que paga dentro da mesma ocupacao. Os valores exatos estao nos atributos de [isco08_isei_br](#). A segunda e que a escala vem de um ano so, 2025, e nao forma serie.

Codigos finos com menos de 30 pessoas na amostra herdaram o escore do grupo acima. A coluna nivel de [isco08_isei_br](#) diz quais.

References

Ganzeboom, H. B. G., De Graaf, P. M., & Treiman, D. J. (1992). A standard international socio-economic index of occupational status. *Social Science Research*, 21(1), 1-56.

See Also

[isco08_isei_br](#) para a tabela e o metodo; [tse_para_isei08\(\)](#) para a ancora internacional; `vignette("validacao")` para a comparacao das duas contra o patrimonio dos candidatos.

Examples

```
data.frame(cod    = c(111, 113, 601),
           rotulo  = tse_para_rotulo(c(111, 113, 601)),
           isei08  = tse_para_isei08(c(111, 113, 601)),
           isei_br = tse_para_isei_br(c(111, 113, 601)))
```

tse_para_isei08	<i>ISEI-08 da ocupação declarada ao TSE</i>
-----------------	---

Description

Escore ISEI ancorado na ISCO-08, obtido pela ponte a partir do ISCO-88.

Usage

```
tse_para_isei08(cod, ano = NULL)
```

Arguments

cod	Vetor de códigos de ocupação do TSE (numérico ou texto).
ano	Vetor opcional de anos de eleição, do mesmo comprimento de cod. Com ele, as candidaturas cujo código o TSE reutilizou depois voltam NA com aviso, em vez de traduzidas pelo dicionário errado: o código 215 designava um cargo de direção até 2000 e passou a designar artista plástico. Sem ele, o comportamento é o de sempre. Veja tse_vigencia() .

Value

Vetor numérico com o escore ISEI-08.

Quando não usar

Para comparar candidaturas entre si, prefira [tse_para_isei\(\)](#), na ISCO-88: o pacote é ancorado nela e a ponte introduz erro. O ISEI-08 serve para juntar o dado eleitoral a fontes que já classificam por ISCO-08 — a PNAD Contínua, via COD, é o caso típico. As duas réguas correlacionam-se fortemente no agregado ($isei08 = 1,266 \times isei88 - 12,72$, $r = 0,97$, $n = 258$), e a maior parte das diferenças de 5 a 15 pontos é o próprio reescalonamento, não erro. Esses coeficientes valem

para a ponte **corrigida** por `.corrige_isco08_tse()`; sem ela, seriam 1,212 e -9,80, e foram esses os que esta seção publicou até a auditoria de 05/09/2026. Mas alguns deslocamentos individuais são grandes **e reais**: o enfermeiro (código 113) sobe 26 pontos, de 43 para 68,7 — não por reclassificação, já que a ISCO-88 o punha em 2230, no grande grupo dos profissionais, e o pacote o leva ao agregado 2220 da ISCO-08, mas porque o ISEI-08 foi reestimado sobre o ISSP de 2002-2007, com os dois sexos, e reavalia as ocupações de cuidado; o vendedor e o comerciante (411, 170) caem 13, de 43 para 29,7, porque a revisão reavaliou o grupo 52 inteiro. Uma série que troque de âncora no meio mede essas duas coisas como se fossem mobilidade.

Examples

```
data.frame(cod = c(111, 113), isei88 = tse_para_isei(c(111, 113)),
           isei08 = tse_para_isei08(c(111, 113)))
```

tse_para_politico	<i>A ocupação declarada é um mandato ou cargo político?</i>
-------------------	---

Description

Marca as ocupações que são o próprio ofício político (vereador, deputado, prefeito, governador e afins). Útil para separar quem já vive da política de quem chega a ela de outra ocupação.

Usage

```
tse_para_politico(cod, ano = NULL)
```

Arguments

cod	Vetor de códigos de ocupação do TSE (numérico ou texto).
ano	Vetor opcional de anos de eleição, do mesmo comprimento de cod. Com ele, as candidaturas cujo código o TSE reutilizou depois voltam NA com aviso, em vez de traduzidas pelo dicionário errado: o código 215 designava um cargo de direção até 2000 e passou a designar artista plástico. Sem ele, o comportamento é o de sempre. Veja tse_vigencia() .

Value

Vetor lógico.

Examples

```
tse_para_politico(c(274, 169))
```

tse_para_prestigio	<i>Prestigio ocupacional de Treiman</i>
--------------------	---

Description

Devolve o escore da Standard International Occupational Prestige Scale, de Donald Treiman. É uma regua de PRESTIGIO — o quanto uma ocupação é socialmente estimada —, distinta do ISEI, que mede posição socioeconômica.

Usage

```
tse_para_prestigio(cod, ano = NULL)
```

Arguments

cod	Vetor de códigos de ocupação do TSE (numérico ou texto).
ano	Vetor opcional de anos de eleição, do mesmo comprimento de cod. Com ele, as candidaturas cujo código o TSE reutilizou depois voltam NA com aviso, em vez de traduzidas pelo dicionário errado: o código 215 designava um cargo de direção até 2000 e passou a designar artista plástico. Sem ele, o comportamento é o de sempre. Veja tse_vigencia() .

Value

Vetor numérico com o escore de prestígio.

Quando o pressuposto quebra

A escala e a média de estudos de prestígio de cerca de 60 países levantados nos anos 1960 e 1970. Aplica-la a dado recente pressupõe que a ordem de prestígio é invariante no tempo e no espaço — que é a tese estrutural de Treiman (1977), e não um fato dado. Ela quebra onde a ocupação mudou de posição desde então: bancário, professor, policial, ocupações de tecnologia.

Sobre o nome

Estas funções substituem `tse_para_siops()` e companhia. No Brasil a sigla SIOPS designa o Sistema de Informações sobre Orcamentos Públicos em Saúde, e a colisão é certa num pacote em português. As formas antigas seguem funcionando como aliases depreciados.

References

Treiman, D. J. (1977). *Occupational Prestige in Comparative Perspective*. New York: Academic Press.

Examples

```
tse_para_prestigio(c(111, 169, 257))
```

tse_para_prestigio08 *Prestigio de Treiman ancorado na ISCO-08, a partir do TSE*

Description

Prestigio de Treiman ancorado na ISCO-08, a partir do TSE

Usage

```
tse_para_prestigio08(cod, ano = NULL)
```

Arguments

cod	Vetor de códigos de ocupação do TSE (numérico ou texto).
ano	Vetor opcional de anos de eleição, do mesmo comprimento de cod. Com ele, as candidaturas cujo código o TSE reutilizou depois voltam NA com aviso, em vez de traduzidas pelo dicionário errado: o código 215 designava um cargo de direção até 2000 e passou a designar artista plástico. Sem ele, o comportamento é o de sempre. Veja tse_vigencia() .

Value

Vetor numerico.

O argumento ano, e por que ele chegou tarde aqui

A versao 0.2.0 acrescentou ano as portas ancoradas na ISCO-08 e nomeou tres delas: tse_para_isco08(), tse_para_isei08() e tse_para_siops08(). Esta ficou de fora por descuido, de modo que o nome PREFERIDO da medida de prestigio nao mascarava vigencia enquanto o seu proprio alias depreciado mascarava. Quem montasse serie pela porta recomendada recebia, calado, o escore do cadastro errado nos codigos reutilizados em 2002. Corrigido em 0.3.0; a assinatura agora e a mesma de tse_para_siops08().

Examples

```
tse_para_prestigio08(c(111, 169))  
tse_para_prestigio08("215", ano = 2000)    # reutilizado: NA com aviso
```

tse_para_rotulo	<i>Rótulo da ocupação, como o TSE o escreveu</i>
-----------------	--

Description

Rótulo da ocupação, como o TSE o escreveu

Usage

```
tse_para_rotulo(cod, ano = NULL)
```

Arguments

cod	Vetor de códigos de ocupação do TSE.
ano	Vetor opcional de anos de eleição, do mesmo comprimento de cod. Com ele, devolve o rótulo vigente naquele ano; sem ele, o mais recente.

Value

Vetor de texto com o rótulo, ou NA.

Examples

```
tse_para_rotulo(c(215, 215), c(2000, 2020))
tse_para_rotulo(c(111, 169, 257))
```

tse_para_siops	<i>Prestígio ocupacional SIOPS da ocupação declarada ao TSE</i>
----------------	---

Description

Devolve o escore SIOPS (Standard International Occupational Prestige Scale), de Treiman, ancorado na ISCO-88. É uma régua de prestígio, distinta do ISEI, que mede posição socioeconômica.

Usage

```
tse_para_siops(cod, ano = NULL)
```

Arguments

cod	Vetor de códigos de ocupação do TSE (numérico ou texto).
ano	Vetor opcional de anos de eleição, do mesmo comprimento de cod. Com ele, as candidaturas cujo código o TSE reutilizou depois voltam NA com aviso, em vez de traduzidas pelo dicionário errado: o código 215 designava um cargo de direção até 2000 e passou a designar artista plástico. Sem ele, o comportamento é o de sempre. Veja tse_vigencia() .

Value

Vetor numérico com o escore SIOPS, ou NA.

Examples

```
tse_para_siops(c(111, 169, 257))
```

tse_para_siops08	<i>SIOPS-08 da ocupação declarada ao TSE</i>
------------------	--

Description

SIOPS-08 da ocupação declarada ao TSE

Usage

```
tse_para_siops08(cod, ano = NULL)
```

Arguments

cod	Vetor de códigos de ocupação do TSE (numérico ou texto).
ano	Vetor opcional de anos de eleição, do mesmo comprimento de cod. Com ele, as candidaturas cujo código o TSE reutilizou depois voltam NA com aviso, em vez de traduzidas pelo dicionário errado: o código 215 designava um cargo de direção até 2000 e passou a designar artista plástico. Sem ele, o comportamento é o de sempre. Veja tse_vigencia() .

Value

Vetor numérico com o escore SIOPS ancorado na ISCO-08.

Examples

```
tse_para_siops08(c(111, 169))
```

tse_quebra_2002	<i>Códigos de ocupação do TSE que mudaram de nome ou de sentido em 2002</i>
-----------------	---

Description

O TSE reeditou a tabela de ocupações entre as eleições de 2000 e 2002. Nem toda mudança de rótulo é problema, e é por isso que esta tabela tem a coluna tipo: só reutilizado invalida a tradução. Use `checa_periodo()` para saber se o seu dado é atingido, e `tse_vigencia()` para ver a história de um código.

Usage

tse_quebra_2002

Format

data.frame com 44 linhas e as colunas:

cod_tse o código.

rotulo_ate_2000, rotulo_apos_2002 como o TSE o chamava antes e depois.

ultimo_ano_antigo, primeiro_ano_novo as eleições comparadas.

n_ate_2000 candidaturas com esse código até 2000, somando 1998 e 2000. Até 05/09/2026 esta coluna trazia apenas a eleição de 2000.

similaridade sobreposição de palavras entre os dois rótulos (0 a 1).

pct_superior_ate_2000, pct_superior_apos_2002, delta_pp proporção com ensino superior completo em cada período, e a diferença.

delta_vs_tendencia delta_pp menos a tendência geral do período (+7,6 pp). NA onde há menos de 30 candidaturas em algum dos lados, que é pouco para o sinal significar coisa alguma.

tipo reutilizado, renomeado, redefinido ou refinado.

Por que tipo é um julgamento, e não uma fórmula

Há dois sinais disponíveis, e **nenhum dos dois basta**:

O rótulo 601 foi de "TRABALHADOR AGRÍCOLA" para "AGRICULTOR": muda todo o léxico e é o mesmo ofício — a população sob o código nem se move (+2,2 pp, abaixo da tendência). Tratá-lo como reutilização mandaria descartar 52.090 candidaturas válidas.

O 215 foi de "OCUPANTE DE CARGO DE DIREÇÃO E ASSESSORAMENTO SUPERIOR" para "ARTISTA PLÁSTICO": reutilização inequívoca que **não move a escolaridade** (+6,5 pp), porque um DAS e um artista plástico têm perfil de diploma parecido. Foi por isso que a versão anterior desta tabela, que só olhava escolaridade, não o via — e ela documentava 3 códigos onde há 7.

Quatro das sete reutilizações são invisíveis ao sinal de escolaridade. A distinção é semântica, e nenhuma métrica automática a alcança. São 44 casos: o tipo foi julgado um a um, e o julgamento é **auditável na própria tabela** — os dois rótulos viajam ao lado das duas evidências.

O que fazer com cada tipo

reutilizado (7 códigos, 1.755 candidaturas) o código passou a designar outra ocupação. Traduzir o período antigo pelo dicionário é erro; exclua ou reclassifique.

renomeado (4) mesma ocupação, nome novo. Não é problema — está aqui para que ninguém a confunda com reutilização ao comparar rótulos.

redefinido (15) o escopo mudou. Cautela.

refinado (18) mesmo posto, rótulo mais preciso.

Source

Rótulos e escolaridade das candidaturas de 1998 a 2026. Os valores de `pct_superior_*` reproduzem exatamente os que a versão anterior desta tabela trazia digitados à mão — que era a única tabela do pacote não gerada por script, e deixou de ser.

Examples

```
# Os códigos reutilizados: o mesmo número, outra ocupação. Ler qualquer um
# deles como série contínua de 1998 a hoje produz uma trajetória que nunca
# existiu.
tse_quebra_2002[tse_quebra_2002$tipo == "reutilizado",
               c("cod_tse", "rotulo_ate_2000", "rotulo_apos_2002")]

# A que tipo pertence cada um dos 44:
table(tse_quebra_2002$tipo)
```

tse_rotulo_para_cod	<i>Encontra o código a partir do rótulo da ocupação</i>
---------------------	---

Description

Serve a quem recebe a ocupação como texto e não como código — o caso de quem baixa `br_tse_eleicoes.candidatos` no `basedosdados`. A busca ignora caixa, acento e pontuação.

Usage

```
tse_rotulo_para_cod(rotulo, exato = TRUE)
```

Arguments

<code>rotulo</code>	Vetor de rótulos de ocupação.
<code>exato</code>	Se <code>TRUE</code> (padrão), só casa o rótulo inteiro; se <code>FALSE</code> , casa por conteúdo e pode devolver mais de um código por rótulo.

Value

`data.frame` com `rotulo` (o que se procurou), `cod_tse` e o rótulo canônico encontrado. Sempre um `data.frame`, com zero linhas quando não há o que procurar — o tipo do retorno não depende do conteúdo do argumento.

Examples

```
tse_rotulo_para_cod("Agricultor")
tse_rotulo_para_cod("advogado", exato = FALSE)
```

tse_universo	<i>O universo das candidaturas: quem o formulário deixa de fora, ano a ano</i>
--------------	--

Description

A decomposição do resíduo das candidaturas brasileiras, por safra e por sexo. [checa_cobertura\(\)](#) diz *quanto* de um vetor recebe *escore*; esta tabela diz *de que é feito* o que não recebe — e mostra que a resposta muda com o tempo e com o sexo de quem se candidata.

Usage

```
tse_universo
```

Format

`data.frame` com 315 linhas:

ano a safra eleitoral, de 1998 a 2026.

sexo "Homem", "Mulher" ou "Todos". A linha "Todos" existe para que quem não se interessa por sexo não tenha de somar nada — e porque a soma correta não é óbvia quando o sexo falta em alguma safra.

rubrica a categoria do universo. Sete valores, na ordem em que a tabela os apresenta: "Com *escore*", "Ocupação sem *escore*", "Vínculo público", "Fora da força de trabalho", "Inativo com trajetória", "Outros (999)" e "Não informada".

n candidaturas na célula.

pct percentual dentro de (ano, sexo). Somam 100 por construção.

As sete rubricas não são a mesma coisa

Descartá-las juntas como "sem *escore*" trata como ruído o que é, em quase todos os casos, informação:

- **Com *escore*** — o código traduz a um ISCO-88 com ISEI, com ano = passado. É a base classificável.
- **Ocupação sem *escore*** — o rótulo nomeia uma ocupação e o *escore* não existe. O caso é 295 MEMBRO DAS FORÇAS ARMADAS: o ISCO-88 0110 não tem ISEI, e NA é a resposta certa.
- **Vínculo público** — 291 OCUPANTE DE CARGO EM COMISSÃO e 296/297/298 SERVIDOR PÚBLICO FEDERAL/ESTADUAL/MUNICIPAL. A pessoa **tem** ocupação e escolheu não a declarar. É a rubrica que quebra a comparação com a população, onde as mesmas pessoas declaram ocupação — veja `pct_setor_publico` em [isco_posicao_br](#).

- **Fora da força de trabalho** — 581 DONA DE CASA e 931 ESTUDANTE, BOLSISTA, ESTAGIÁRIO E ASSEMELHADOS. Posição declarada, não omissão.
- **Inativo com trajetória** — aposentado, pensionista, militar reformado, capitalista de ativos financeiros. Há ocupação passada, e `isei_retrospectivo()` recupera parte dela.
- **Outros (999)** — recusa de nomear. É a única rubrica que é ausência pura de informação, e é a que mais cresce.
- **Não informada** — códigos -4, 0 e 949, ausência de registro.

O buraco cresce, e não é o mesmo buraco

Duas leituras que a tabela permite e que uma taxa de cobertura única esconde. A primeira: **a não declaração vai a zero a partir de 2006**, o que indica preenchimento obrigatório do campo — e a recusa de nomear assume o lugar dela. A rubrica 999 sai de 13,5% em 1998 e chega a 21,9% em 2024. A cobertura de 2024 (62,0%) é mais baixa que a de 2006 (73,2%) **sem que a tradução tenha piorado**: o que mudou foi o que se declara.

A segunda: **a diferença de cobertura entre os sexos está quase toda numa linha**. De 2004 em diante, Fora da força de trabalho é 1,2% das candidaturas de homens e 14,6% das de mulheres. Descartar o resíduo apaga mulheres em proporção muito maior, e isso inverte o sinal de proporções de classe calculadas sobre a base cheia.

O escore sai do pacote, com o ano

A rubrica Com escore é definida por `tse_para_isei(cod, ano = ano)`, e o ano não é decorativo: sem ele, os sete códigos reaproveitados depois de 2002 receberiam o escore da ocupação errada, e a tabela mediria a cobertura de um erro. Veja [tse_quebra_2002](#) e [checa_periodo\(\)](#).

Source

Microdados de candidaturas do Tribunal Superior Eleitoral, 1998 a 2026, pela microbase descrita em `data-raw/05a_microbase_2026.R`. Gerada por `data-raw/13_gera_universo.R`. Os microdados **não** viajam com o pacote; o que entra é esta tabela agregada, sem indivíduo e sem identificador.

Examples

```
# A cobertura do escore, por safra – e o que a substitui quando ela cai.
u <- tse_universo
u[u$sexo == "Todos" & u$rubrica %in% c("Com escore", "Outros (999)"),
  c("ano", "rubrica", "pct")]

# A linha que explica a diferença de cobertura entre os sexos.
u[u$ano == 2024 & u$rubrica == "Fora da força de trabalho", ]
```

tse_validacao	<i>Critério externo para aferir a medida: patrimônio e escolaridade por ocupação</i>
---------------	--

Description

Agregado por ocupação de duas variáveis que o TSE coleta e que **não entram na construção da medida em momento nenhum**: o patrimônio declarado na candidatura e o grau de instrução. É contra ele que a vinheta vignette("validacao") afere o ISEI.

Usage

tse_validacao

Format

data.frame com 220 linhas e as colunas:

cod_tse código de ocupação.

n candidaturas com esse código, 1998–2026.

pct_superior % com ensino superior completo.

pct_mulher % de mulheres.

n_com_bens candidaturas com patrimônio declarado maior que zero.

mediana_patrimonio mediana do patrimônio declarado, em reais. NA **onde** n_com_bens < 200 — ver a seção sobre o piso.

Por que agregado, e por que este piso

A unidade é a **ocupação**, não a candidatura, porque é nesse nível que uma medida de posição ocupacional é definida, e porque um agregado de 220 linhas não é microdado, não identifica ninguém e pode viajar com o pacote.

Entram ocupações com pelo menos 200 candidaturas. A mediana de patrimônio exige um segundo piso, de 200 declarações de bens, porque mediana apoiada em poucas declarações é ruidosa: sem ele a correlação com o ISEI cai de 0,681 para cerca de 0,63, não porque a medida piore, mas porque o critério externo fica instável. Este 0,63 é o único número desta página que não se recalcula a partir do pacote — as medianas abaixo do piso são NA aqui —, e sai de `data-raw/05_gera_validacao.R`.

O segundo piso zera a mediana; não descarta a linha, e a diferença importa. Até 29/07/2026 ele descartava a linha inteira, o que amputava do conjunto 44 ocupações cuja escolaridade e cuja composição por gênero estão perfeitamente medidas e que apenas carecem de declarações de bens. A consequência era que a regressão de gênero documentada em `tse_para_isei()` não se reproduzia a partir do dado publicado. Hoje reproduz. Quem correlacionar com patrimônio deve filtrar `!is.na(mediana_patrimonio)`; quem usar escolaridade ou gênero tem as 220 linhas à disposição.

Os códigos reutilizados entram só na vigência nova

Sete códigos foram reaproveitados para ocupação diferente depois de 2002 (veja [tse_quebra_2002](#)). Até a auditoria de 05/09/2026 este conjunto os agregava sobre a série inteira, de modo que `pct_superior` e `pct_mulher` misturavam duas populações: o código 214 saía com 30,6% de ensino superior porque metade da massa era DELEGADO DE POLÍCIA, quando ESCULTOR E PINTOR tem 2,3%; o 521 saía com 42,3% de mulheres onde a GOVERNANTA tem 97,2%. Hoje cada um desses códigos entra apenas a partir do seu `primeiro_ano_novo`, que é o mesmo corte que o argumento `ano` = das funções de tradução aplica. Um deles deixou de alcançar o piso de 200 candidaturas dentro da própria vigência e saiu do conjunto — por isso 220 linhas, e não 221.

O efeito sobre as correlações é imperceptível (a de escolaridade não se move na terceira casa), porque são 7 códigos em 220 e cerca de 1.800 candidaturas em 3,37 milhões. A correção não é pelo tamanho: é porque quem toma este conjunto como a tabela descritiva por ocupação — que é para isso que ele é publicado — lia cinco linhas materialmente erradas.

As correlações que este conjunto sustenta

Contra a escolaridade, sobre as 207 ocupações com ISEI: $r = 0,765$ (Spearman 0,816). Contra o logaritmo da mediana de patrimônio, sobre as 165 que também têm mediana: $r = 0,681$ (Spearman 0,695), ambas **não ponderadas** — cada ocupação conta uma, que é o certo para aferir uma régua de ocupações. No nível da **candidatura** a correlação com patrimônio é de apenas 0,207, e o contraste é o resultado, não um defeito. A comparação estritamente equivalente, na mesma unidade e com o mesmo peso, é com a correlação ecológica de 0,748; veja [tse_dispersao_patrimonio](#).

O patrimônio de 2026 não entra, e por quê

A safra de 2026 contribui para `n`, `pct_superior` e `pct_mulher`, e **se abstém de `n_com_bens` e da mediana de patrimônio**. A razão é de unidade, não de qualidade do dado: o patrimônio está deflacionado a reais de outubro de 2024, pelo número-índice do IPCA do mês da eleição, e outubro de 2026 ainda não aconteceu. Deflacionar por um mês que não é o da eleição poria na coluna um valor que a definição da coluna desmente.

A assimetria não é nova: `n` sempre cobriu um período mais largo do que `n_com_bens`, porque o TSE só publica declaração de bens a partir de 2006 e as candidaturas de 1998 a 2004 já entravam nessa mesma condição. Quando o IPCA de outubro de 2026 existir, a coluna sai de graça.

Source

Declaração de bens e grau de instrução das candidaturas ao TSE, 1998–2026 (patrimônio: 2006–2024; ver a seção acima).

Examples

```
# As ocupações com mais candidaturas no conjunto de validação:
head(tse_validacao[order(-tse_validacao$n), ], 5)

# A correlação que sustenta a medida: o ISEI atribuído por tradução contra
# a escolaridade declarada, que o pacote nunca viu ao construir a régua.
v <- tse_validacao
v$isei <- suppressWarnings(tse_para_isei(v$cod_tse))
round(stats::cor(v$isei, v$pct_superior, use = "complete.obs"), 3)
```

`tse_vigencia`*Em que eleições cada código esteve em vigor, e com que nome*

Description

Uma linha por **vigência**: um código que nunca mudou de nome tem uma; um que mudou tem uma por período.

Usage

```
tse_vigencia(cod = NULL)
```

Arguments

`cod` Vetor opcional de códigos. Se omitido, devolve a tabela inteira.

Value

`data.frame` com `cod_tse`, `de`, `ate`, `rotulo` e `n`.

Examples

```
tse_vigencia(215)
```

Index

* datasets

cbo2002_escada, [5](#)
cbo2002_familia_isco88, [6](#)
cbo2002_isco88, [7](#)
cbo94_isco88, [18](#)
cod_isco08, [26](#)
cod_populacao_br, [33](#)
isco08_isco88, [41](#)
isco08_isei_br, [42](#)
isco08_medidas, [44](#)
isco88_isco08, [46](#)
isco88_medidas, [47](#)
isco_posicao_br, [38](#)
reguas, [54](#)
tse_autorrotulo_patrimonio, [55](#)
tse_dispersao_patrimonio, [59](#)
tse_isco, [61](#)
tse_ocupacao_rotulos, [62](#)
tse_quebra_2002, [80](#)
tse_universo, [82](#)
tse_validacao, [84](#)

cbo2002_concordancia, [4](#)
cbo2002_concordancia(), [10](#), [11](#)
cbo2002_escada, [5](#)
cbo2002_familia_isco88, [6](#)
cbo2002_isco88, [5](#), [7](#), [7](#)
cbo2002_para_egp, [8](#)
cbo2002_para_egp(), [19](#)
cbo2002_para_isco, [9](#)
cbo2002_para_isco(), [5](#), [35](#)
cbo2002_para_isco08, [11](#)
cbo2002_para_isei, [12](#)
cbo2002_para_isei(), [11](#)
cbo2002_para_isei08, [14](#)
cbo2002_para_isei08(), [13](#)
cbo2002_para_isei_br, [13](#)
cbo2002_para_isei_br(), [73](#)
cbo2002_para_prestigio, [15](#)
cbo2002_para_prestigio08, [15](#)

cbo2002_para_prestigio08(), [13](#)
cbo2002_para_siops, [16](#)
cbo2002_para_siops08, [17](#)
cbo94_isco88, [18](#), [73](#)
cbo94_para_egp, [19](#)
cbo94_para_isco, [20](#)
cbo94_para_isco(), [11](#)
cbo94_para_isco08, [20](#)
cbo94_para_isei, [21](#)
cbo94_para_prestigio, [21](#)
cbo94_para_siops, [22](#)
checa_cobertura, [22](#)
checa_cobertura(), [23](#), [55](#), [82](#)
checa_cobertura_cbo2002, [23](#)
checa_cobertura_cbo2002(), [24](#)
checa_cobertura_cbo94, [24](#)
checa_cobertura_cod, [24](#)
checa_periodo, [25](#)
checa_periodo(), [55](#), [80](#), [83](#)
cod_isco08, [26](#)
cod_para_egp, [27](#)
cod_para_isco, [28](#)
cod_para_isco(), [30](#), [40](#)
cod_para_isco08, [29](#)
cod_para_isco08(), [26](#)
cod_para_isei, [30](#)
cod_para_isei08, [31](#)
cod_para_isei08(), [30](#)
cod_para_isei_br, [30](#)
cod_para_isei_br(), [73](#)
cod_para_prestigio, [32](#)
cod_para_prestigio08, [32](#)
cod_populacao_br, [33](#)
crosswalk_cbo2002, [35](#)
crosswalk_cbo2002(), [9](#), [10](#), [12–17](#), [38](#)
crosswalk_cbo94, [36](#)
crosswalk_cbo94(), [38](#)
crosswalk_cod, [36](#)
crosswalk_cod(), [30](#), [38](#)

crosswalk_tse, 37
crosswalk_tse(), 37, 55, 70

isco08_isco88, 41
isco08_isei_br, 14, 31, 42, 46, 73, 74
isco08_medidas, 42, 44
isco08_para_isco88, 45
isco08_para_isco88(), 34, 40, 41
isco08_para_isei_br, 46
isco08_para_isei_br(), 73
isco88_isco08, 46, 50
isco88_medidas, 5, 47
isco88_para_egp, 48
isco88_para_egp(), 19, 35–37, 39, 47, 66
isco88_para_isco08, 50
isco88_para_isco08(), 15, 41, 71
isco88_para_isei, 50
isco88_para_prestigio, 51
isco88_para_siops, 52
isco_posicao_br, 37, 38, 61, 72, 82
isei_retrospectivo, 52
isei_retrospectivo(), 83

reguas, 54

tse_autorrotulo_patrimonio, 55, 58, 66
tse_codigos_autorrotulo, 55, 57, 57, 66, 71, 72
tse_diff_cadastro, 58
tse_dispersao_patrimonio, 56, 57, 59, 72, 85
tse_isco, 61
tse_ocupacao_rotulos, 25, 59, 62
tse_para_classe, 63
tse_para_classe(), 68, 70
tse_para_componente_alta, 65
tse_para_componente_alta(), 55, 57, 58, 69
tse_para_egp, 66
tse_para_egp(), 38, 55, 63, 64
tse_para_estrato, 68
tse_para_estrato(), 55
tse_para_isco, 69
tse_para_isco(), 66
tse_para_isco08, 70
tse_para_isco08(), 50
tse_para_isei, 71
tse_para_isei(), 55, 60, 68–70, 74, 84
tse_para_isei08, 74
tse_para_isei08(), 72, 74
tse_para_isei_br, 73
tse_para_isei_br(), 13, 14, 30, 31, 44, 46, 55
tse_para_politico, 75
tse_para_prestigio, 76
tse_para_prestigio08, 77
tse_para_rotulo, 78
tse_para_rotulo(), 25
tse_para_siops, 78
tse_para_siops08, 79
tse_quebra_2002, 25, 80, 83, 85
tse_rotulo_para_cod, 81
tse_universo, 33, 34, 64, 82
tse_validacao, 60, 61, 84
tse_vigencia, 86
tse_vigencia(), 63, 65, 67–71, 73–80