

Package ‘datos’

October 13, 2022

Title Traduce al Español Varios Conjuntos de Datos de Práctica

Version 0.5.0

Description Provee una versión traducida de los siguientes conjuntos de datos: 'airlines', 'airports', 'AwardsManagers', 'babynames', 'Batting', 'credit_data', 'diamonds', 'faithful', 'fueleconomy', 'Fielding', 'flights', 'gapminder', 'gss_cat', 'iris', 'Managers', 'mpg', 'mtcars', 'atmos', 'palmerpenguins', 'People', 'Pitching', 'planes', 'presidential', 'table1', 'table2', 'table3', 'table4a', 'table4b', 'table5', 'vehicles', 'weather', 'who'. English: It provides a Spanish translated version of the datasets listed above.

License CC0

URL <https://github.com/cienciadedatos/datos>

BugReports <https://github.com/cienciadedatos/datos/issues>

Depends R (>= 3.5.0)

Imports babynames, dplyr, forcats, fueleconomy, gapminder, ggplot2, Lahman, nasaweather, nycflights13, palmerpenguins, modeldata (>= 1.0.0), rlang, tibble, tidyr, yaml

Suggests covr, testthat (>= 2.1.0)

ByteCompile true

Encoding UTF-8

Language es

LazyData true

RoxygenNote 7.2.1

NeedsCompilation no

Author Riva Quiroga [aut, cre] (<<https://orcid.org/0000-0002-1147-4135>>),
Edgar Ruiz [aut],
Mauricio Vargas [aut],
Mauro Lepore [aut] (<<https://orcid.org/0000-0002-1986-7988>>),
Rayna Harris [ctb],
Daniela Vasquez [ctb],
Joshua Kunst [ctb]

Maintainer Riva Quiroga <riva.quiroga@uc.cl>

Repository CRAN

Date/Publication 2022-08-29 14:40:05 UTC

R topics documented:

aerolineas	3
aeropuertos	3
atmosfera	4
aviones	4
bateadores	5
capitanes	6
clima	7
comunes	8
datos-stringr	8
datos_credito	9
diamantes	10
dirigentes	10
encuesta	11
fiel	12
fildeo	12
flores	13
jardineros	13
lanzadores	14
millas	15
mtautos	16
nombres	17
oms	17
países	20
personas	21
pinguinos	22
premios_capitanes	23
premios_dirigentes	23
presidencial	24
representantes	24
salarios	25
tabla1	25
tabla2	26
tabla3	26
tabla4a	27
tabla4b	27
tabla5	28
vehiculos	28
vuelos	29

Index

30

aerolíneas	<i>Nombres de aerolíneas</i>
------------	------------------------------

Description

Nombres de aerolíneas y su respectivo código carrier de dos dígitos.

Usage

aerolíneas

Format

Un data.frame con 16 filas y 2 columnas

aerolinea abreviación de dos caracteres del nombre de la aerolínea

nombre nombre completo de la aerolínea

aeropuertos	<i>Datos de aeropuertos</i>
-------------	-----------------------------

Description

Metadatos útiles de los aeropuertos

Usage

aeropuertos

Format

Un data.frame con 1.458 filas y 8 columnas

codigo_aeropuerto código FAA del aeropuerto

nombre nombre del aeropuerto

latitud coordenada de ubicación del aeropuerto

longitud coordenada de ubicación del aeropuerto

altura altura medida en pies

zona_horaria diferencia horaria respecto del meridiano de Greenwich (GMT)

horario_verano horario legal de acuerdo a las políticas de ahorro de energía. A = Zona Horaria de Ahorro de Luz de Día (DST) estándar de los EEUU: comienza el segundo Sábado de Marzo y termina el primer Sábado de Noviembre. U = Desconocido. N = No tiene zona horaria especial.

zona_horaria_iana zona horaria IANA, que se determina de acuerdo al servicio web GeoNames

 atmosfera

Datos atmosféricos

Description

Datos atmosféricos

Usage

atmosfera

Format

Un data.frame con 41472 filas y 11 columnas

latitud Ubicación de la medida. Grilla espacial de 24 por 24 desde longitud 113.80 a 56.20 y desde latitud 36.2N a 21.2S.

longitud Ubicación de la medida. Grilla espacial de 24 por 24 desde longitud 113.80 a 56.20 y desde latitud 36.2N a 21.2S.

anio 72 puntos en el tiempo. Una vez al mes desde enero de 1995 hasta diciembre de 2000.

mes 72 puntos en el tiempo. Una vez al mes desde enero de 1995 hasta diciembre de 2000.

temp_superficie Temperatura media mensual basada en la energía emitida desde la Tierra en condiciones de cielo despejado (en K).

temperatura Temperatura media mensual del aire cerca de la superficie de la Tierra (en K).

presion Presión media mensual en una ubicación determinada de la superficie de la Tierra (en mb).

ozono Cantidad media mensual de ozono en la columna atmosférica (en Dobsons).

nube_baja Porcentaje de la media mensual de cielo cubierto por nubes con una presión superior mayor a 680 mb o aproximadamente menor que 3.24 km.

nube_media Porcentaje de la media mensual de cielo cubierto por nubes con una presión superior entre 440-680 mb o aproximadamente menor que 3.24 a 6.5 km.

nube_alta Porcentaje de la media mensual de cielo cubierto por nubes con una presión superior menor o igual a 440 mb o aproximadamente mayor que 6.5 km.

 aviones

Datos de aviones

Description

Metadatos de los aviones y sus códigos de cola en el registro de aviación de la FAA. American Airways (AA) y Envoy Air (MQ) reportan número de flota en lugar de número de cola, por lo que no es posible trazarlos.

Usage

aviones

Format

Un data.frame con 3.322 filas y 9 columnas

codigoCola código de cola**anio** año de fabricación**tipo** tipo de avión**fabricante** fabricante del avión**modelo** modelo del avión**motores** cantidad de motores**asientos** cantidad de asientos**velocidad** velocidad crucero promedio medida en millas por hora**tipo_motor** tipo de motor del avión

bateadores

*Tabla de estadísticas de bateo***Description**

Estadísticas de bateadores

Format

Un data frame con 105.861 filas y 22 columnas

id_jugador ID del jugador**id_anio** Año**orden Equipos** Orden en que el jugador se movió entre equipos dentro de la misma temporada**id_equipo** ID del equipo (factor)**id_liga** ID de la liga (factor con niveles AA, AL, FL, NL, PL, UA)**juegos** Número de juegos en los que participó oficialmente el jugador**al_bate** Turnos al bate**carreras** Carreras anotadas**golpes** Cantidad de veces en la que el jugador alcanzó una base debido a un golpe, sin error de la defensa**dobles** Hits en los que el bateador llegó a segunda base**triples** Golpes en los que el bateador llegó a tercera base**cuadrangulares** Cuadrangulares - jonrones

carreras_empujadas Carreras empujadas
bases_robadas Bases robadas
atrapado_robando Atrapado (out) en intento de robo
base_bolas Base por bolas
ponches Ponches
base_intencional Base por bolas intencional
golpeado Golpeado por lanzamiento
toque_sacrificio Toque de sacrificio
elavado_sacrificio Elevado de sacrificio
doble_matanza Batazos para doble matanza (doble play)

 capitanes

Tabla de capitanes

Description

Información de los equipos que dirigieron y algunas estadísticas básicas de los equipos en cada año.

Format

Un data frame con 3.504 filas y 10 columnas

id_jugador ID del jugador

anio Año

id_equipo ID equipo (factor)

id_liga ID liga (factor)

en_temporada Cero si fue capitán del equipo al equipo toda la temporada. En otro caso denota el orden de entrada en la temporada (uno si fue el primer capitán del equipo que entró esa temporada, dos si fue el segundo capitán del equipo que entró esa temporada, etc)

juegos Juegos dirigidos

juegos_ganados Juegos ganados

juegos_perdidos Juegos perdidos

posicion Posición del equipo en la clasificación final del año

jugador_representado Tiene valor "S" (sí) para los jugadores estuvieron de capitán del equipo durante la temporada y "N" (no) en caso contrario (factor).

clima

Datos de clima

Description

Datos meteorológicos por hora para LGA, JFK y EWR.

Usage

clima

Format

Un data.frame con 3,322 líneas y 9 variables

origen Estación de clima

anio año de la fecha de registro

mes mes de la fecha de registro

dia día de la fecha de registro

hora hora de la fecha de registro

temperatura temperatura medida en grados Fahrenheit

punto_rocio temperatura medida en grados Fahrenheit en la que, dada la presión atmosférica y humedad, condensa el vapor de agua

humedad humedad relativa

direccion_viento dirección del viento medida en grados

velocidad_viento velocidad del viento medida en millas por hora

velocidad_rafaga velocidad de las ráfagas de viento medida en millas por hora

precipitacion precipitación de lluvia medida en pulgadas

presion presión atmosférica al nivel del mar medida en milibares

visibilidad visibilidad medida en millas

fecha_hora fecha y hora de registro en formato POSIXct

comunes

Modelos comunes de vehículos

Description

Estos modelos tienen al menos 10 años de datos. Este dataset es adecuado para efectuar un join que restrinja el número de observaciones a un nivel manejable.

Usage

comunes

Format

Un data.frame con 347 filas y 4 columnas

fabricante Fabricante

modelo Modelo

total_modelos Número total de modelos

total_modelos_año Número total de modelos-año

datos-stringr

Vectores de caracteres dentro del paquete stringr

Description

Muestra de vectores de caracteres para practicar la manipulación de cadenas (strings)

Usage

oraciones

frutas

palabras

Format

Un vector

An object of class character of length 50.

An object of class character of length 25.

An object of class character of length 1000.

Details

oraciones - es una selección y adaptación al español de los datos de práctica contenidos en stringr::sentences.

fruta - es una selección y adaptación al español de los datos de práctica contenidos en stringr::fruit.

palabras - es una selección de las 1000 palabras más frecuentes del Corpus de Referencia del Español Actual (CREA) de la RAE, disponible en github.com/rivaquioga/frecuencias-crea

Examples

```
length(oraciones)
oraciones[1:5]
```

```
length(frutas)
frutas[1:5]
```

```
length(palabras)
palabras[1:5]
```

datos_credito

Datos de crédito

Description

Datos asociados a créditos de consumo.

Format

Un data.frame con 14 filas y 4454 columnas

Estado estado del crédito

Antigüedad Antigüedad laboral

Vivienda tipo de propiedad de la vivienda

Plazo Plazo del crédito en meses

Edad edad del cliente

EstadoCivil estado civil

Registros existencia de registros previos

Trabajo tipo de trabajo

Gastos cantidad o monto de gastos

Ingresos cantidad o monto de ingresos

Activos cantidad o monto de activos

Deuda cantidad o monto de deudas

Cantidad cantidad solicitada de préstamo

Precio precio del crédito

diamantes	<i>Precio de 50.000 diamantes</i>
-----------	-----------------------------------

Description

Un conjunto de datos que contiene los precios de casi 54.000 diamantes.

Usage

diamantes

Format

Un data.frame con 53.940 filas y 10 columnas

precio Precio en dólares americanos (\$326-\$18,823)

quilate Peso del diamante (0.2-5.01)

corte Calidad del corte (Regular, Bueno, Muy bueno, Premium, Ideal)

color Color del diamante, de J (peor) a D (mejor)

claridad Medida de que tan claro es el diamante (I1 (peor), SI1, SI2, VS1, VS2, VVS1, VVS2, IF (mejor))

profundidad Porcentaje de la profundidad total = $z / \text{mean}(x, y) = 2 * z / (x + y)$ (43-79)

tabla Ancho de la parte superior del diamante con relación a su punto más ancho (43-95)

x Largo en milímetros

y Ancho en milímetros

z Profundidad en milímetros

dirigentes	<i>Tabla de dirigentes</i>
------------	----------------------------

Description

Información de los equipos que dirigieron y algunas estadísticas básicas de los equipos en cada temporada.

Format

Un data frame con 3.504 filas y 10 columnas

id_jugador ID del jugador

anio Año

id_equipo ID equipo (factor)

id_liga ID liga (factor)

orden_temporada Orden (secuencia) en la dirigencia de equipos durante una temporada. 0 si fue dirigente del equipo durante toda una temporada. En otro caso denota el orden de entrada a un equipo en una temporada (1 si fue el primer dirigente del equipo que entró en esa temporada, 2 si fue el segundo dirigente del equipo que entró en esa temporada, etc.)

juegos_dirigidos Juegos dirigidos

juegos_ganados Juegos ganados

juegos_perdidos Juegos perdidos

posicion Posición del equipo en la clasificación final de la temporada regular

jugador_dirigente Tiene valor 'S' (sí) para los jugadores que hayan sido jugadores y dirigentes del equipo durante una temporada y 'N' (no) en caso contrario (factor).

encuesta

Muestra de variables categóricas de una encuesta social

Description

Muestra de variables categóricas de la Encuesta Social General de EE.UU.

Usage

encuesta

Format

Un data.frame con 21.473 filas y 9 columnas

anio año de la encuesta, 2000-2014

estado_civil estado civil

edad edad. Edad máxima truncada a 89 años

raza categorías raciales según la encuesta

ingreso ingreso reportado por los encuestados (US\$)

partido partido político con el que se identifica

religion religión

denominacion denominación religiosa a la que pertenece

horas_tv horas de televisión que ve al día

fiel	<i>Datos del géiser Viejo Fiel (Old Faithful)</i>
------	---

Description

Datos de tiempo de duración y espera entre erupciones del géiser Viejo Fiel (Old Faithful), ubicado en el Parque Nacional de Yellowstone, en Wyoming, Estados Unidos.

Usage

fiel

Format

Un data frame con 272 observaciones y 2 columnas

erupciones Tiempo de erupciones, en minutos

espera Minutos entre erupciones

fildeo	<i>Tabla de Fildeo</i>
--------	------------------------

Description

Estadísticas de jardineros

Format

Un data frame con 140.921 filas y 18 columnas

id_jugador ID del jugador

id_anio Año

orden_equipos Orden en que el jugador se movió entre equipos dentro de la misma temporada

id_equipo ID del equipo (factor)

id_liga ID de la liga (factor con niveles AA, AL, FL, NL, PL, UA)

posicion Posición del jugador a la defensiva

juegos_jugados Juegos jugados

juegos_iniciados Juegos iniciados

InnOuts Tiempo jugado por el jugador, expresado en outs

outs_realizados Outs realizados por el jugador

asistencias Asistencias realizadas por el jugador

errores Errores realizados por el jugador

- doble_matanzas** Doble matanzas en las que el jugador ha estado envuelto defensivamente
- bolas_pasadas** Bolas dejadas pasar (aplicable a los receptores)
- lanzamientos_desviados** Lanzamientos desviados (aplicable a los receptores)
- bases_robadas** Bases robadas por oponente (aplicable a los receptores)
- atrapados_robando** Oponentes atrapados en intento de robo (aplicable a los receptores)
- clasificacion_zona** Porcentaje de pelotas atrapadas por un fildeador, en su zona defensiva "típica"

 flores

Datos sobre la flor Iris de Edgar Anderson

Description

El set de datos sobre la flor Iris entrega las medidas en centímetros del largo y ancho de los sépalos y los pétalos de 50 flores de 3 especies de Iris (setosa, versicolor y virginica).

Usage

flores

Format

Un dataframe con 150 casos (filas) y 5 variables (columnas)

Largo.Sepalo Largo del sépalo

Ancho.Sepalo Ancho del sépalo

Largo.Petalo Largo del pétalo

Ancho.Petalo Ancho del pétalo

Especie A qué especie de la flor Iris corresponde (setosa, versicolor, virginica)

 jardineros

Tabla de estadísticas de jardineros

Description

Estadísticas de jardineros

Format

Un data frame con 140.921 filas y 18 columnas

id_jugador ID del jugador

id_anio Año

ordenEquipos Orden en que el jugador se movió entre equipos dentro de la misma temporada

id_equipo ID del equipo (factor)

id_liga ID de la liga (factor con niveles AA, AL, FL, NL, PL, UA)

posicion Posición del jugador a la defensiva

juegos_jugados Juegos jugados

juegos_iniciados Juegos iniciados

inn_outs Tiempo jugado por el jugador, expresado en outs

outs_realizados Outs realizados por el jugador

asistencias Asistencias realizadas por el jugador

errores Errores realizados por el jugador

doble_matanzas Doble matanzas en las que el jugador ha estado envuelto defensivamente

bolas_pasadas Bolas dejadas pasar (aplicable a los receptores)

lanzamientos_desviados Lanzamientos desviados (aplicable a los receptores)

bases_robadas Bases robadas por oponente (aplicable a los receptores)

atrapados_robando Oponentes atrapados en intento de robo (aplicable a los receptores)

clasificacion_zona Porcentaje de pelotas atrapadas por un fildeador, en su zona defensiva 'típica'

lanzadores

Tabla de estadísticas de lanzadores

Description

Estadísticas de lanzadores

Format

Un data frame con 46.699 filas y 30 columnas

id_jugador ID del jugador

id_anio Año

ordenEquipos Orden en que el jugador se movió entre equipos dentro de la misma temporada

id_equipo ID del equipo (factor)

id_liga ID de la liga (factor con niveles AA, AL, FL, NL, PL, UA)

juegos_ganados Juegos jugados ganados

juegos_perdidos Juegos jugados perdidos

juegos_jugados Juegos jugados
juegos_iniciados Juegos jugados iniciados
juegos_completados Número de juegos completos (9 innings lanzados) que lanzó el jugador
blanqueadas Número de blanqueos (juegos completos sin permitir carreras) que lanzó el jugador
juegos_salvados Juegos salvados
IPouts Outs propinados al equipo contrario (Igual a innings lanzados x 3)
hits Hits permitidos del oponente
carreras_ganadas Carreras limpias recibidas
cuadrangulares Cuadrangulares recibidos
BB Base por bolas cedidas al oponente
ponches Ponches propinados al oponente
promedio_bateo_rival Promedio de bateo del rival
promedio_carreras_ganadas Promedio de carreras limpias permitidas (normalizada a 9 innings lanzados)
IBB Base por bolas intencionales cedidas al oponente
lanzamientos_desviados Lanzamientos desviados lanzados por el lanzador
HBP Bateadores golpeador por el lanzador
BK Balks (movimiento ilegal del cuerpo realizado por el lanzador)
BFP Bateadores a los que el lanzador se ha enfrentado
juegos_finalizados Juegos en los que el lanzador finalizó el juego
carreras Carreras recibidas (sucias y limpias)
sacrificios_golpeados Toques de sacrificio que el oponente le hizo al lanzador
vuelos_sacrificio Elevado (fly) de sacrificio que el oponente le hizo al lanzador
doble_matanza Doble matanza inducidas por el lanzador

 millas

Datos de economía de combustible de 1999 y 2008 para 38 modelos populares de automóviles

Description

Este conjunto de datos contiene un subconjunto de los datos de economía de combustible que la Agencia de Protección Medioambiental (EPA) pone a disposición en <http://fueleconomy.gov>. Contiene solo modelos que tuvieron una nueva versión cada año entre 1999 y 2008, lo que fue utilizado como un proxy de la popularidad del modelo.

Usage

millas

Format

Un data.frame con 234 filas y 11 columnas

fabricante fabricante

modelo nombre del modelo

cilindrada tamaño del motor del automóvil, en litros

año año de fabricación

cilindros número de cilindros

transmision tipo de transmisión

traccion tipo de tracción (d = delantera, t = trasera, 4 = 4 ruedas)

ciudad millas por galón de combustible en ciudad

autopista millas por galón de combustible en autopista

combustible tipo de combustible (p = premium, r = regular, e = etanol, d = diesel, g = gas natural comprimido)

clase tipo de auto

mtautos

Pruebas de ruta de automóviles de Motor Trend

Description

Los datos fueron extraídos de la revista Motor Trend de Estados Unidos de 1974, y tiene datos de consumo de combustible y 10 aspectos de diseño y rendimiento para 32 automóviles (modelos de 1973-1974).

Usage

mtautos

Format

Un data.frame con 32 filas y 12 columnas

millas millas por galón de Estados Unidos

cilindros número de cilindros

cilindrada suma del volumen útil de todos los cilindros del motor en pulgadas cúbicas

caballos caballos de fuerza brutos

eje relación del eje de transmisión trasero

peso peso (1000 libras)

velocidad tiempo en recorrer 1/4 de milla

forma forma del motor (en V o en línea)

transmision tipo de transmisión (0 = automático, 1 = manual)

cambios número de cambios de la caja de cambios

carburadores número de carburadores

nombres	<i>Nombres de bebés</i>
---------	-------------------------

Description

Datos completos sobre nombres de bebés provistos por la Social Security Administration de EE.UU. Incluye todos los nombres con al menos 5 usos.

Usage

nombres

Format

Un data.frame con 1924665 observaciones y 5 variables

anio Año del registro

sexo Sexo del bebé

nombre Nombre

FALSE Cantidad de ocurrencias

prop Proporción

oms	<i>Datos de tuberculosis de la Organización Mundial de la Salud</i>
-----	---

Description

Subconjunto de datos del informe anual de tuberculosis de la Organización Mundial de la Salud

Usage

oms

Format

Un data.frame con 7.240 filas y 60 columnas

pais nombre del país

iso2 código ISO de dos dígitos del país

iso3 código ISO de tres dígitos del país

anio año

nuevos_fpp_h014 casos nuevos, diagnóstico frotis pulmonar positivo (fpp), hombres (h) de 0 a 14 años (014)

- nuevos_fpp_h1524** casos nuevos, diagnóstico frotis pulmonar positivo (fpp), hombres (h) de 15 a 24 años (1524)
- nuevos_fpp_h2534** casos nuevos, diagnóstico frotis pulmonar positivo (fpp), hombres (h) de 25 a 34 años (2534)
- nuevos_fpp_h3544** casos nuevos, diagnóstico frotis pulmonar positivo (fpp), hombres (h) de 35 a 44 años (3544)
- nuevos_fpp_h4554** casos nuevos, diagnóstico frotis pulmonar positivo (fpp), hombres (h) de 45 a 54 años (4554)
- nuevos_fpp_h5564** casos nuevos, diagnóstico frotis pulmonar positivo (fpp), hombres (h) de 55 a 64 años (5564)
- nuevos_fpp_h65** casos nuevos, diagnóstico frotis pulmonar positivo (fpp), hombres (h) de 65 o más años
- nuevos_fpp_m014** casos nuevos, diagnóstico frotis pulmonar positivo (fpp), mujeres (m) de 0 a 14 años (014)
- nuevos_fpp_m1524** casos nuevos, diagnóstico frotis pulmonar positivo (fpp), mujeres (m) de 15 a 24 años (1524)
- nuevos_fpp_m2534** casos nuevos, diagnóstico frotis pulmonar positivo (fpp), mujeres (m) de 25 a 34 años (2534)
- nuevos_fpp_m3544** casos nuevos, diagnóstico frotis pulmonar positivo (fpp), mujeres (m) de 35 a 44 años (3544)
- nuevos_fpp_m4554** casos nuevos, diagnóstico frotis pulmonar positivo (fpp), mujeres (m) de 45 a 54 años (4554)
- nuevos_fpp_m5564** casos nuevos, diagnóstico frotis pulmonar positivo (fpp), mujeres (m) de 55 a 64 años (5564)
- nuevos_fpp_m65** casos nuevos, diagnóstico frotis pulmonar positivo (fpp), mujeres (m) de 65 o más años
- nuevos_fpn_h014** casos nuevos, diagnóstico frotis pulmonar negativo (fpn), hombres (h) de 0 a 14 años (014)
- nuevos_fpn_h1524** casos nuevos, diagnóstico frotis pulmonar negativo (fpn), hombres (h) de 15 a 24 años (1524)
- nuevos_fpn_h2534** casos nuevos, diagnóstico frotis pulmonar negativo (fpn), hombres (h) de 25 a 34 años (2534)
- nuevos_fpn_h3544** casos nuevos, diagnóstico frotis pulmonar negativo (fpn), hombres (h) de 35 a 44 años (3544)
- nuevos_fpn_h4554** casos nuevos, diagnóstico frotis pulmonar negativo (fpn), hombres (h) de 45 a 54 años (4554)
- nuevos_fpn_h5564** casos nuevos, diagnóstico frotis pulmonar negativo (fpn), hombres (h) de 55 a 64 años (5564)
- nuevos_fpn_h65** casos nuevos, diagnóstico frotis pulmonar negativo (fpn), hombres (h) de 65 o más años
- nuevos_fpn_m014** casos nuevos, diagnóstico frotis pulmonar negativo (fpn), mujeres (m) de 0 a 14 años (014)

nuevos_fpn_m1524 casos nuevos, diagnóstico frotis pulmonar negativo (fpn), mujeres (m) de 15 a 24 años (1524)

nuevos_fpn_m2534 casos nuevos, diagnóstico frotis pulmonar negativo (fpn), mujeres (m) de 25 a 34 años (2534)

nuevos_fpn_m3544 casos nuevos, diagnóstico frotis pulmonar negativo (fpn), mujeres (m) de 35 a 44 años (3544)

nuevos_fpn_m4554 casos nuevos, diagnóstico frotis pulmonar negativo (fpn), mujeres (m) de 45 a 54 años (4554)

nuevos_fpn_m5564 casos nuevos, diagnóstico frotis pulmonar negativo (fpn), mujeres (m) de 55 a 64 años (5564)

nuevos_fpn_m65 casos nuevos, diagnóstico frotis pulmonar negativo (fpn), mujeres (m) de 65 o más años

nuevos_ep_h014 casos nuevos, diagnóstico extrapulmonar (ep), hombres (h) de 0 a 14 años (014)

nuevos_ep_h1524 casos nuevos, diagnóstico extrapulmonar (ep), hombres (h) de 15 a 24 años (1524)

nuevos_ep_h2534 casos nuevos, diagnóstico extrapulmonar (ep), hombres (h) de 25 a 34 años (2534)

nuevos_ep_h3544 casos nuevos, diagnóstico extrapulmonar (ep), hombres (h) de 35 a 44 años (3544)

nuevos_ep_h4554 casos nuevos, diagnóstico extrapulmonar (ep), hombres (h) de 45 a 54 años (4554)

nuevos_ep_h5564 casos nuevos, diagnóstico extrapulmonar (ep), hombres (h) de 55 a 64 años (5564)

nuevos_ep_h65 casos nuevos, diagnóstico extrapulmonar (ep), hombres (h) de 65 o más años

nuevos_ep_m014 casos nuevos, diagnóstico extrapulmonar (ep), mujeres (m) de 0 a 14 años (014)

nuevos_ep_m1524 casos nuevos, diagnóstico extrapulmonar (ep), mujeres (m) de 15 a 24 años (1524)

nuevos_ep_m2534 casos nuevos, diagnóstico extrapulmonar (ep), mujeres (m) de 25 a 34 años (2534)

nuevos_ep_m3544 casos nuevos, diagnóstico extrapulmonar (ep), mujeres (m) de 35 a 44 años (3544)

nuevos_ep_m4554 casos nuevos, diagnóstico extrapulmonar (ep), mujeres (m) de 45 a 54 años (4554)

nuevos_ep_m5564 casos nuevos, diagnóstico extrapulmonar (ep), mujeres (m) de 55 a 64 años (5564)

nuevos_ep_m65 casos nuevos, diagnóstico extrapulmonar (ep), mujeres (m) de 65 o más años

nuevosrecaida_h014 nuevas recaídas, hombres (h) de 0 a 14 años (014)

nuevosrecaida_h1524 nuevas recaídas, hombres (h) de 15 a 24 años (1524)

nuevosrecaida_h2534 nuevas recaídas, hombres (h) de 25 a 34 años (2534)

nuevosrecaida_h3544 nuevas recaídas, hombres (h) de 35 a 44 años (3544)

nuevosrecaida_h4554 nuevas recaídas, hombres (h) de 45 a 54 años (4554)

nuevosrecaida_h5564 nuevas recaídas, hombres (h) de 55 a 64 años (5564)
nuevosrecaida_h65 nuevas recaídas, hombres (h) de 65 o más años
nuevosrecaida_m014 nuevas recaídas, mujeres (m) de 0 a 14 años (014)
nuevosrecaida_m1524 nuevas recaídas, mujeres (m) de 15 a 24 años (1524)
nuevosrecaida_m2534 nuevas recaídas, mujeres (m) de 25 a 34 años (2534)
nuevosrecaida_m3544 nuevas recaídas, mujeres (m) de 35 a 44 años (3544)
nuevosrecaida_m4554 nuevas recaídas, mujeres (m) de 45 a 54 años (4554)
nuevosrecaida_m5564 nuevas recaídas, mujeres (m) de 55 a 64 años (5564)
nuevosrecaida_m65 nuevas recaídas, mujeres (m) de 65 o más años

países

Datos de Gapminder

Description

Extracto de datos de Gapminder sobre expectativa de vida, PIB per cápita y población, según país

Usage

países

Format

Un data.frame de 1704 filas y 6 columnas

pais países incluidos (factor con 142 niveles)

continente continentes (factor con 5 niveles)

anio desde 1952 a 2007, datos cada 5 años

esperanza_de_vida esperanza de vida al nacer, en años

poblacion población

pib_per_capita PIB per cápita (en dólares americanos, ajustados según inflación)

Source

<http://www.gapminder.org/data/>

personas

*Tabla de personas***Description**

Nombres de jugadores, fecha de nacimiento e información bibliográfica. Este archivo se usa para obtener detalles acerca de los jugadores listados en bateadores, lanzadores y otros archivos en que los jugadores están identificados por `id_jugador`.

Format

Un data frame con 19.617 filas y 26 columnas

id_jugador ID del jugador

anio_nacimiento Año de nacimiento

mes_nacimiento Mes de nacimiento

día_nacimiento Día de nacimiento

pais_nacimiento País de nacimiento

estado_nacimiento Estado de nacimiento

ciudad_nacimiento Ciudad de nacimiento

anio_fallecimiento Año de fallecimiento

mes_fallecimiento Mes de fallecimiento

día_fallecimiento Día de fallecimiento

pais_fallecimiento País de fallecimiento

estado_fallecimiento Estado de fallecimiento

ciudad_fallecimiento Ciudad de fallecimiento

nombre Nombre

apellido Apellido

nombre_completo Nombre y apellido

peso Peso en libras

estatura Estatura en pulgadas

bateador Si como bateador es diestro (D) o zurdo (Z)

lanzador Si como lanzador es diestro (D) o zurdo (Z)

fecha_primer_juego Fecha primer juego en primera división

fecha_ultimo_juego Fecha último juego en primera división

id_retrosheet Código de jugador usado en Retrosheet (<http://www.retrosheet.org/>)

id_brw Código de jugador usado en Baseball Reference website (<http://www.baseball-reference.com/>)

fecha_fallecimiento Fecha de fallecimiento en formato yyyy-mm-dd

fecha_nacimiento Fecha de nacimiento en formato yyyy-mm-dd

pingüinos

Medidas de tamaño de pingüinos adultos en busca de comida cerca de la estación Palmer en la Antártica.

Description

Incluye medidas para: especie de pingüino, isla dentro del Archipiélago de Palmer, tamaño (largo de aleta, masa corporal, dimensiones de pico) y sexo.

Usage

pingüinos

Format

Un tibble con 344 filas y 8 variables

especie un factor que indica la especie de pingüino (Adelia, Barbijo y Papúa)

isla un factor que indica la isla en el archipiélago de Palmer (Biscoe, Dream o Torgersen)

largo_pico_mm un número que indica el largo del pico (milímetros)

alto_pico_mm un número que indica el alto del pico (milímetros)

largo_aleta_mm un entero que indica el largo de la aleta (milímetros)

masa_corporal_g un entero que indica la masa corporal (gramos)

sexo un factor que indica el sexo del pingüino (hembra, macho)

año un entero que indica el año estudiado (2007, 2008 o 2009)

Source

*** Pingüinos Adelia: Palmer Station Antarctica LTER and K. Gorman. 2020. Structural size measurements and isotopic signatures of foraging among adult male and female Adélie penguins (*Pygoscelis adeliae*) nesting along the Palmer Archipelago near Palmer Station, 2007-2009 ver 5. Environmental Data Initiative <https://doi.org/10.6073/pasta/98b16d7d563f265cb52372c8ca99e60f>. *** Pingüinos Papúa: Palmer Station Antarctica LTER and K. Gorman. 2020. Structural size measurements and isotopic signatures of foraging among adult male and female Gentoo penguin (*Pygoscelis papua*) nesting along the Palmer Archipelago near Palmer Station, 2007-2009 ver 5. Environmental Data Initiative <https://doi.org/10.6073/pasta/7fca67fb28d56ee2ffa3d9370ebda689>. *** Pingüinos Barbijo: Palmer Station Antarctica LTER and K. Gorman. 2020. Structural size measurements and isotopic signatures of foraging among adult male and female Chinstrap penguin (*Pygoscelis antarcticus*) nesting along the Palmer Archipelago near Palmer Station, 2007-2009 ver 6. Environmental Data Initiative <https://doi.org/10.6073/pasta/c14dfcfada8ea13a17536e73eb6f9e>. *** Publicado originalmente en: Gorman KB, Williams TD, Fraser WR (2014) Ecological Sexual Dimorphism and Environmental Variability within a Community of Antarctic Penguins (Genus *Pygoscelis*). PLoS ONE 9(3): e90081. doi:10.1371/journal.pone.0090081

premios_capitanes *Tabla de premios de los capitanes*

Description

Información de los premios de los capitanes de equipo

Format

Un data frame con 179 filas y 6 columnas

id_jugador ID del jugador

nombre_premio Nombre del premio recibido

anio Año

id_liga ID liga (factor)

empate Tiene valor "S" (sí) para los premios que corresponden a un empate y "N" (no) en caso contrario.

notas Comentarios sobre el premio

premios_dirigentes *Tabla de premios de los dirigentes*

Description

Información de los premios de los dirigentes de equipo

Format

Un data frame con 179 filas y 6 columnas

id_dirigente ID del dirigente

nombre_premio Nombre del premio recibido por el dirigente

anio Año

id_liga ID liga (factor)

empate Tiene valor "S" (sí) para los premios que corresponden a un empate y "N" (no) en caso contrario.

notas Comentarios sobre el premio

presidencial	<i>Periodos de 11 presidentes, desde Eisenhower a Obama</i>
--------------	---

Description

Nombre, fecha de inicio y término de su mandato y partido político de 11 presidentes de EE.UU., desde Eisenhower a Obama.

Format

Un data.frame con 11 filas y 4 columnas

nombre Nombre del presidente

inicio Fecha de inicio de su periodo presidencial

fin Fecha de término de su periodo presidencial

partido Partido político al que pertenece

representantes	<i>Tabla de representantes (bateadores)</i>
----------------	---

Description

Información de los representantes de los bateadores

Format

Un data frame con 3.504 filas y 10 columnas

id_jugador ID del jugador

anio Año

id_equipo ID equipo (factor)

id_liga ID liga (factor)

en_temporada Cero si representó al equipo toda la temporada. En otro caso denota el orden de entrada en la temporada (uno si fue el primer representante que entró esa temporada, dos si fue el segundo representante que entró esa temporada, etc)

partidos Partidos dirigidos

partidos_ganados Partidos ganados

partidos_perdidos Partidos perdidos

posicion Posición del equipo en la clasificación final del año

jugador_representado Tiene valor "S" (sí) para los jugadores que tienen representante en la temporada y "N" (no) en caso contrario (factor).

salarios	<i>Tabla de salarios</i>
----------	--------------------------

Description

Información de los salarios de los jugadores.

Format

Un data frame con 26.428 filas y 5 columnas

anio Año

id_equipo ID equipo (factor)

id_liga ID liga (factor)

id_jugador ID del jugador

salario Salario en dólares

tabla1	<i>Registros de tuberculosis de la Organización Mundial de la Salud (1era variante)</i>
--------	---

Description

Contiene observaciones de casos registrados en Afganistán, Brasil y China en los años 1999 y 2000.

Format

Un tibble con 6 filas y 4 columnas

pais Nombre del país

anio Año correspondiente a los valores registrados

casos Número de casos registrados

poblacion Población del país

tabla2	<i>Registros de tuberculosis de la Organización Mundial de la Salud (2da variante)</i>
--------	--

Description

Contiene observaciones de casos registrados en Afganistán, Brasil y China en los años 1999 y 2000.

Format

Un tibble con 12 filas y 4 columnas

país Nombre del país

año Año correspondiente a los valores registrados

tipo Categoría a la que pertenece la observación

cuenta Cantidad de observaciones para cada tipo

tabla3	<i>Registros de tuberculosis de la Organización Mundial de la Salud (3ra variante)</i>
--------	--

Description

Contiene observaciones de casos registrados en Afganistán, Brasil y China en los años 1999 y 2000.

Format

Un tibble con 6 filas y 3 columnas

país Nombre del país

año Año correspondiente a los valores registrados

tasa Tasa de casos de tuberculosis en relación a la población total

tabla4a	<i>Registros de tuberculosis de la Organización Mundial de la Salud (variante 4a)</i>
---------	---

Description

Contiene observaciones de casos registrados en Afganistán, Brasil y China en los años 1999 y 2000.

Format

Un tibble con 3 filas y 3 columnas

pais Nombre del país

1999 Casos registrados el año 1999.

2000 Casos registrados el año 2000.

tabla4b	<i>Registros de tuberculosis de la Organización Mundial de la Salud (variante 4b)</i>
---------	---

Description

Contiene observaciones de casos registrados en Afganistán, Brasil y China en los años 1999 y 2000.

Format

Un tibble con 3 filas y 3 columnas

pais Nombre del país

1999 Población en el año 1999.

2000 Población en el año 2000.

tabla5	<i>Registros de tuberculosis de la Organización Mundial de la Salud (5ta variante)</i>
--------	--

Description

Contiene observaciones de casos registrados en Afganistán, Brasil y China en los años 1999 y 2000.

Format

Un tibble con 6 filas y 4 columnas

país Nombre del país

siglo Siglo correspondiente a los valores registrados

año Año correspondiente a los valores registrados

tasa Tasa de casos de tuberculosis con relación a la población total

vehiculos	<i>Datos de economía de combustible</i>
-----------	---

Description

Datos de economía de combustible de la Agencia de Protección Medioambiental (EPA) de EE.UU., 1985-2015. Contiene una selección de variables y no considera vehículos con datos incompletos.

Usage

vehiculos

Format

Un data.frame con 33442 filas y 12 columnas

id Identificador único EPA

fabricante Fabricante

modelo Nombre del modelo

año año de fabricación

clase tipo de auto

transmision tipo de transmisión

traccion tipo de tracción

cilindros número de cilindros

motor tamaño del motor del automóvil, en litros

combustible tipo de combustible

autopista millas por galón de combustible en autopista

ciudad millas por galón de combustible en ciudad

vuelos	<i>Datos de vuelos</i>
--------	------------------------

Description

Datos temporales de todos los vuelos que despegaron de Nueva York (JFK, LGA o EWR) durante 2013.

Usage

vuelos

Format

Un data.frame con 336.776 filas y 19 columnas

anio año de la fecha de salida

mes mes de la fecha de salida

dia día de la fecha de salida

horario_salida horario efectivo de salida del vuelo (formato HHMM o HMM), hora local

salida_programada horario programado para la salida (formato HHMM o HMM), hora local

atraso_salida atraso de la salida en minutos. Valores negativos indican salida adelantada

horario_llegada horario efectivo de llegada del vuelo (formato HHMM o HMM), hora local

llegada_programada horario programado para la llegada (formato HHMM o HMM), hora local

atraso_llegada atraso de la llegada en minutos. Valores negativos indican llegada adelantada

aerolinea abreviación de dos letras de la aerolínea. Ver 'aerolineas' para obtener el nombre

vuelo número de vuelo

codigoCola código de cola del avión

origen origen del vuelo. Ver 'aeropuertos' para metadatos adicionales

destino destino del vuelo. Ver 'aeropuertos' para metadatos adicionales

tiempo_vuelo cantidad de tiempo en aire, en minutos

distancia distancia entre aeropuertos, en millas

hora hora del horario programado para la salida

minuto minutos del horario programado para la salida

fecha_hora fecha y horario programados del vuelo en formato POSIXct

Index

* datasets

- aerolineas, 3
- aeropuertos, 3
- atmosfera, 4
- aviones, 4
- bateadores, 5
- capitanes, 6
- clima, 7
- comunes, 8
- datos-stringr, 8
- datos_credito, 9
- diamantes, 10
- dirigentes, 10
- encuesta, 11
- fiel, 12
- fildeo, 12
- flores, 13
- jardineros, 13
- lanzadores, 14
- millas, 15
- mtautos, 16
- nombres, 17
- oms, 17
- países, 20
- personas, 21
- pinguinos, 22
- premios_capitanes, 23
- premios_dirigentes, 23
- presidencial, 24
- representantes, 24
- salarios, 25
- tabla1, 25
- tabla2, 26
- tabla3, 26
- tabla4a, 27
- tabla4b, 27
- tabla5, 28
- vehiculos, 28
- vuelos, 29

- aerolineas, 3
- aeropuertos, 3
- atmosfera, 4
- aviones, 4

- bateadores, 5

- capitanes, 6
- clima, 7
- comunes, 8

- datos-stringr, 8
- datos_credito, 9
- diamantes, 10
- dirigentes, 10

- encuesta, 11

- fiel, 12
- fildeo, 12
- flores, 13
- frutas (datos-stringr), 8

- jardineros, 13

- lanzadores, 14

- millas, 15
- mtautos, 16

- nombres, 17

- oms, 17
- oraciones (datos-stringr), 8

- países, 20
- palabras (datos-stringr), 8
- personas, 21
- pinguinos, 22
- premios_capitanes, 23
- premios_dirigentes, 23
- presidencial, 24

representantes, [24](#)

salarios, [25](#)

tabla1, [25](#)

tabla2, [26](#)

tabla3, [26](#)

tabla4a, [27](#)

tabla4b, [27](#)

tabla5, [28](#)

vehiculos, [28](#)

vuelos, [29](#)