Package 'NeuroDataSets'

November 23, 2025

Type Package

Title A Comprehensive Collection of Neuroscience and Brain-Related Datasets

Version 0.3.0

Maintainer Renzo Caceres Rossi <arenzocaceresrossi@gmail.com>

Description Offers a rich and diverse collection of datasets focused on the brain, nervous system, and related disorders.
The package includes clinical, experimental, neuroimaging, behavioral, cognitive, and simulated data on conditions such as
Parkinson's disease, Alzheimer's disease, dementia, epilepsy, schizophrenia, autism spectrum disorder, attention deficit, hyperactivity disorder,
Tourette's syndrome, traumatic brain injury, gliomas, migraines, headaches, sleep disorders, concussions, encephalitis,
subarachnoid hemorrhage, and mental health conditions. Datasets cover structural and functional brain data, cross-sectional and longitudinal
MRI imaging studies, neurotransmission, gene expression, cognitive performance, intelligence metrics, sleep deprivation effects, treatment outcomes,
brain-body relationships across species, neurological injury patterns, and acupuncture interventions. Data sources include peer-reviewed studies,
clinical trials, military health records, sports injury databases, and international comparative studies.
Designed for researchers, neuroscientists, clinicians, psychologists, data scientists, and students, this package facilitates exploratory data analysis,
statistical modeling, and hypothesis testing in neuroscience and neuroepidemiology.

License GPL-3

Language en

URL https://github.com/lightbluetitan/neurodatasets,
https://lightbluetitan.github.io/neurodatasets/

BugReports https://github.com/lightbluetitan/neurodatasets/issues

Encoding UTF-8

LazyData true

Suggests ggplot2, testthat (>= 3.0.0), dplyr, knitr, rmarkdown

Depends R (>= 4.1.0)

Imports utils

RoxygenNote 7.3.2

Config/testthat/edition 3

VignetteBuilder knitr

NeedsCompilation no

Author Renzo Caceres Rossi [aut, cre] (ORCID:
<https://orcid.org/0009-0005-0744-854X>)

Repository CRAN

Date/Publication 2025-11-23 13:30:19 UTC

Contents

`aba_phenotype_data_df` *Allen Brain Atlas Phenotype Data*

Description

This dataset, aba_phenotype_data_df, is a data frame containing brain tissue phenotype measurements from the Allen Brain Atlas Aging, Dementia, and TBI study. The data includes immunohistochemistry markers for microglia and astrocytes across 377 brain samples, intended for correlation analyses with expression data.

Usage

```
data(aba_phenotype_data_df)
```

Format

A data frame with 377 observations and 4 variables:

structure_acronym.x Character: Brain structure acronym

ihc_iba1_ffpe Numeric: IBA1 immunohistochemistry measurement (microglia marker)

ihc_gfap_ffpe Numeric: GFAP immunohistochemistry measurement (astrocyte marker)

id Character: Sample identification code

Details

The dataset name has been kept as 'aba_phenotype_data_df' to avoid confusion with other datasets in the R ecosystem. This naming convention helps distinguish this dataset as part of the NeuroDataSets package. The suffix 'df' indicates that the dataset is a data frame. The original content has not been modified.

Source

Data taken from the BRETIGEA package version 1.0.3. Original data from: Allen Brain Atlas Aging, Dementia, and TBI study.

---

ability_intelligence_list

Ability and Intelligence Tests

Description

This dataset, ability_intelligence_list, is a list containing psychometric data from six cognitive tests administered to 112 individuals. The list includes a covariance matrix, variable means, and observation count for tests measuring various intellectual abilities.

Usage

```
data(ability_intelligence_list)
```

Format

A list with 3 components:

cov Numeric matrix [6×6]: Test score covariance matrix

center Numeric vector [6]: Variable means

n.obs Numeric: Number of observations (112)

Details

The dataset name has been kept as 'ability_intelligence_list' to avoid confusion with other datasets in the R ecosystem. This naming convention helps distinguish this dataset as part of the NeuroDataSets package. The suffix 'list' indicates that the dataset is a list object. The original content has not been modified.

Source

Data taken from the educationR package version 0.10

`acupuncture_df` *Acupuncture Therapy for Chronic Headache*

Description

This dataset, acupuncture_df, is a data frame from a randomized controlled trial (RCT) evaluating the effectiveness of acupuncture therapy for chronic headaches. The primary outcome was the headache severity score, measured using a 6-item Likert-type scale at the one-year follow-up. The dataset includes group allocation, baseline headache score, one-year follow-up score, and the corresponding change score. Some observations may contain missing values due to omitted cases recorded in the dataset attributes.

Usage

```
data(acupuncture_df)
```

Format

A data frame with 301 observations and 4 variables:

group Group assignment (integer)

pk1 Baseline headache severity score (numeric)

pk5 Headache severity score at one-year follow-up (numeric)

change Change in headache severity score (numeric)

Details

The dataset name has been kept as acupuncture_df to avoid confusion with other datasets in the R ecosystem. This naming convention helps distinguish this dataset as part of the NeuroDataSets package and assists users in identifying its specific characteristics. The suffix df indicates that the dataset is a data frame. The original content has not been modified in any way.

Source

Data taken from the R4HCR package version 0.1

ADHD_df *ADHD Symptom Checklist for Children Aged 6–8 Years*

Description

This dataset, ADHD_df, is a data frame containing ADHD symptom ratings for 355 children aged 6 to 8 years from the Children's Attention Project (CAP) cohort (Silk et al. 2019). The sample consists of 146 children diagnosed with ADHD and 209 without a diagnosis. Symptoms were assessed through structured interviews with parents using the NIMH Diagnostic Interview Schedule for Children IV (DISC-IV) (Shaffer et al. 2000). The checklist includes 18 items: 9 Inattentive (I) and 9 Hyperactive/Impulsive (HI). Each symptom item is binary coded (1 = present, 0 = absent), providing a comprehensive assessment of ADHD symptomatology in young children.

Usage

```
data(ADHD_df)
```

Format

A data frame with 355 observations and 19 variables:

group Group indicator (integer: 1 = ADHD diagnosis, 0 = no diagnosis)

avoid Avoids tasks requiring sustained mental effort (integer: 0 or 1)

closeatt Fails to give close attention to details (integer: 0 or 1)

distract Easily distracted by extraneous stimuli (integer: 0 or 1)

forget Forgetful in daily activities (integer: 0 or 1)

instruct Fails to follow through on instructions (integer: 0 or 1)

listen Does not seem to listen when spoken to directly (integer: 0 or 1)

loses Loses things necessary for tasks or activities (integer: 0 or 1)

org Difficulty organizing tasks and activities (integer: 0 or 1)

susatt Difficulty sustaining attention in tasks or play (integer: 0 or 1)

blurts Blurts out answers before questions are completed (integer: 0 or 1)

fidget Fidgets with hands or feet or squirms in seat (integer: 0 or 1)

interrupt Interrupts or intrudes on others (integer: 0 or 1)

motor Acts as if driven by a motor (integer: 0 or 1)

quiet Difficulty playing or engaging quietly in leisure activities (integer: 0 or 1)

runs Runs about or climbs excessively in inappropriate situations (integer: 0 or 1)

seat Leaves seat in situations when remaining seated is expected (integer: 0 or 1)

talks Talks excessively (integer: 0 or 1)

turn Difficulty waiting turn (integer: 0 or 1)

Details

The dataset name has been kept as ADHD_df to avoid confusion with other datasets in the R ecosystem. This naming convention helps distinguish this dataset as part of the NeuroDataSets package and assists users in identifying its specific characteristics. The suffix df indicates that the dataset is a data frame. The original content has not been modified.

Source

Data taken from the bgms package version 0.1.6.1

---

adolescent_mental_health_df

Adolescent Mental Health Study

Description

This dataset, adolescent_mental_health_df, is a data frame containing mental health assessments from the National Longitudinal Study of Adolescent Health. The data includes depression and anxiety measures for 4,344 students in grades 7-12 from a cross-sectional sample analyzed by Warne (2014).

Usage

```
data(adolescent_mental_health_df)
```

Format

A data frame with 4,344 observations and 3 variables:

grade Ordered factor with 6 levels: School grade (7-12)

depression Integer: Depression symptom score

anxiety Integer: Anxiety symptom score

Details

The dataset name has been kept as 'adolescent_mental_health_df' to avoid confusion with other datasets in the R ecosystem. This naming convention helps distinguish this dataset as part of the NeuroDataSets package. The suffix 'df' indicates that the dataset is a data frame. The original content has not been modified.

Source

Data taken from the heplots package version 1.7.4. Original analysis: Warne, R.T. (2014) A primer on Multivariate Analysis of Variance (MANOVA) for Behavioral Scientists. *Practical Assessment, Research & Evaluation*, 19(1).

`AD_biomarkers_tbl_df` *Alzheimer's Biomarkers*

Description

This dataset, AD_biomarkers_tbl_df, is a tibble containing clinical data from 333 patients in a study of Alzheimer's disease biomarkers. The study included patients with mild cognitive impairment and healthy controls, with measurements of demographic characteristics, apolipoprotein E genotype, protein biomarkers (including Abeta, Tau, and pTau), and clinical dementia scores.

Usage

```
data(AD_biomarkers_tbl_df)
```

Format

A tibble with 333 observations and 131 variables:

age Numeric: Patient age

male Numeric: Indicator for male gender (1 = male, 0 = female)

Genotype Factor: Apolipoprotein E genotype

Class Factor: Clinical classification (e.g., healthy, impaired)

Ab_42 Numeric: Amyloid-beta 42 protein measurement

tau Numeric: Tau protein measurement

p_tau Numeric: Phosphorylated Tau protein measurement

[131 additional biomarker variables] Numeric measurements of various proteins and biomarkers

Details

The dataset name has been kept as 'AD_biomarkers_tbl_df' to avoid confusion with other datasets in the R ecosystem. This naming convention helps distinguish this dataset as part of the NeuroDataSets package. The suffix 'tbl_df' indicates that the dataset is a tibble. The original content has not been modified.

Source

Data taken from the modeldata package version 1.4.0. Original study: Craig-Schapiro R, Kuhn M, Xiong C, et al. (2011) Multiplexed Immunoassay Panel Identifies Novel CSF Biomarkers for Alzheimer's Disease Diagnosis and Prognosis. PLoS ONE 6(4): e18850.

alzheimer_smoking_df *Smoking and Alzheimer's Disease*

Description

This dataset, alzheimer_smoking_df, is a data frame containing case-control data from a study examining the association between smoking and Alzheimer's disease. The study included 538 participants with information on smoking status, disease classification, and gender.

Usage

```
data(alzheimer_smoking_df)
```

Format

A data frame with 538 observations and 3 variables:

smoking Factor: Smoking status of participants (4 levels)

disease Factor: Disease classification including Alzheimer's diagnosis (3 levels)

gender Factor: Participant's gender (2 levels)

Details

The dataset name has been kept as 'alzheimer_smoking_df' to avoid confusion with other datasets in the R ecosystem. This naming convention helps distinguish this dataset as part of the NeuroDataSets package. The suffix 'df' indicates that the dataset is a data frame. The original content has not been modified.

Source

Data taken from the coin package version 1.4-3. Original study: Salib, E. and Hillier, V. (1997). A case-control study of smoking and Alzheimer's disease. International Journal of Geriatric Psychiatry 12: 295-300.

ASD_risks_tbl_df *Autism Spectrum Disorder (ASD) Risk Factors in Children*

Description

This dataset, ASD_risks_tbl_df, is a tibble containing information on various clinical, behavioral, genetic, and developmental factors associated with the risk of developing Autism Spectrum Disorder (ASD) traits in children. The dataset consists of 1,985 observations and 28 variables, including the Autism Spectrum Quotient items (A1–A10), Social Responsiveness Scale, Qchat-10 score, Childhood Autism Rating Scale, and multiple indicators related to speech, learning, genetics, mental health, developmental delays, behavioral issues, demographics, and family history. The final column indicates whether the child is expected to develop ASD traits in the future (0 or 1).

Usage

```
data(ASD_risks_tbl_df)
```

Format

A tibble with 1,985 observations and 28 variables:

CASE_NO_PATIENT'S Patient case identifier (numeric)

A1 Autism Spectrum Quotient item A1 (numeric)

A2 Autism Spectrum Quotient item A2 (numeric)

A3 Autism Spectrum Quotient item A3 (numeric)

A4 Autism Spectrum Quotient item A4 (numeric)

A5 Autism Spectrum Quotient item A5 (numeric)

A6 Autism Spectrum Quotient item A6 (numeric)

A7 Autism Spectrum Quotient item A7 (numeric)

A8 Autism Spectrum Quotient item A8 (numeric)

A9 Autism Spectrum Quotient item A9 (numeric)

A10_Autism_Spectrum_Quotient Autism Spectrum Quotient item A10 (numeric)

Social_Responsiveness_Scale Social Responsiveness Scale score (numeric)

Age_Years Age in years (numeric)

Qchat_10_Score Q-CHAT-10 score (numeric)

Speech Delay/Language Disorder Indicator of speech delay or language disorder (character)

Learning disorder Indicator of learning disorder (character)

Genetic_Disorders Presence of genetic disorders (character)

Depression Presence of depression (character)

Global developmental delay/intellectual disability Indicator of global developmental delay or intellectual disability (character)

Social/Behavioural Issues Presence of social or behavioral issues (character)

Childhood Autism Rating Scale Childhood Autism Rating Scale score (numeric)

Anxiety_disorder Presence of anxiety disorder (character)

Sex Sex of the participant (character)

Ethnicity Ethnicity of the participant (character)

Jaundice History of jaundice (character)

Family_mem_with_ASD Indicator of family member with ASD (character)

Who_completed_the_test Relationship of the respondent who completed the test (character)

ASD_traits Indicator of whether the child is expected to develop ASD traits (character)

Details

The dataset name has been kept as ASD_risks_tbl_df to avoid confusion with other datasets in the R ecosystem. This naming convention helps distinguish this dataset as part of the NeuroDataSets package and assists users in identifying its specific characteristics. The suffix tbl_df indicates that the dataset is a tibble (a modern data frame). The original content has not been modified in any way. Variable names and values are provided exactly as they appear in the source.

Source

Data taken from Kaggle: https://www.kaggle.com/datasets/uppulurimadhuri/dataset

---

bilingual_brains_df *Brain Structure in Bilingual Humans*

Description

This dataset, bilingual_brains_df, is a data frame containing measurements of second language proficiency scores and gray matter density in the left inferior parietal region from 22 observations.

Usage

```
data(bilingual_brains_df)
```

Format

A data frame with 22 observations and 2 variables:

proficiency Numeric vector representing second language proficiency scores (summary of reading, writing, and speech)

greymatter Numeric vector representing density of gray matter in the left inferior parietal region

Details

The dataset name has been kept as 'bilingual_brains_df' to avoid confusion with other datasets in the R ecosystem. This naming convention helps distinguish this dataset as part of the NeuroDataSets package and assists users in identifying its specific characteristics. The suffix 'df' indicates that the dataset is a data frame. The original content has not been modified in any way.

Source

Data taken from the abd package version 0.2-8

`blood_brain_barrier_df` *Blood-Brain Barrier*

Description

This dataset, blood_brain_barrier_df, is a data frame containing experimental measurements from a rat study investigating sugar-infusion methods for temporary blood-brain barrier disruption. The barrier's protective function was assessed through multiple biological markers.

Usage

```
data(blood_brain_barrier_df)
```

Format

A data frame with 34 observations and 9 variables:

Brain Integer: Brain tissue measurement (units?)

Liver Integer: Liver tissue measurement (units?)

Time Numeric: Experimental time measurement (hours)

Treatment Factor with 2 levels: Experimental treatment groups

Days Integer: Observation period (days)

Sex Factor with 2 levels: Animal sex (Male/Female)

Weight Integer: Subject weight (grams)

Loss Numeric: Physiological loss measurement

Tumor Integer: Tumor presence indicator (0/1)

Details

The dataset name has been kept as 'blood_brain_barrier_df' to avoid confusion with other datasets in the R ecosystem. This naming convention helps distinguish this dataset as part of the NeuroDataSets package. The suffix 'df' indicates that the dataset is a data frame. The original content has not been modified.

Source

Data taken from the Sleuth3 package version 1.0-6. Original reference: Ramsey, F.L. and Schafer, D.W. (2013) *The Statistical Sleuth: A Course in Methods of Data Analysis* (3rd ed), Cengage Learning.

brainexpression_df *Proteolipid Protein 1 Gene Expression in Brain Tissue*

Description

This dataset, brainexpression_df, is a data frame containing expression levels of the proteolipid protein 1 gene (PLP1) in 45 individuals across three groups. The dataset includes group classifications and corresponding PLP1 expression measurements, making it useful for comparative gene expression analysis and studying differences in myelin-related protein expression across populations.

Usage

```
data(brainexpression_df)
```

Format

A data frame with 45 observations and 2 variables:

group Group classification (factor with 3 levels)

PLP1.expression Expression level of the proteolipid protein 1 gene (numeric)

Details

The dataset name has been kept as brainexpression_df to avoid confusion with other datasets in the R ecosystem. This naming convention helps distinguish this dataset as part of the NeuroDataSets package and assists users in identifying its specific characteristics. The suffix df indicates that the dataset is a data frame. The original content has not been modified.

Source

Data taken from the abd package version 0.2-8

brains_cognitive_matrix *BRAiNS Cohort Cognitive States Matrix*

Description

This dataset, brains_cognitive_matrix, is a matrix containing the states and covariates of 649 participants enrolled in the BRAiNS cohort at the University of Kentucky's Alzheimer's Disease Research Center. The data includes longitudinal cognitive assessments and various health covariates across multiple visits.

Usage

```
data(brains_cognitive_matrix)
```

Format

A matrix with 6240 observations and 13 variables:

ID Integer: Participant identification number

visitno Integer: Visit number

prstate Integer: Previous cognitive state

custate Integer: Current cognitive state

bagec Integer: Baseline age centered

famhx Integer: Family history of dementia (0 = No, 1 = Yes)

HBP Integer: History of high blood pressure (0 = No, 1 = Yes)

apoe4 Integer: APOE ε_4 allele carrier status (0 = Non-carrier, 1 = Carrier)

smk1 Integer: Smoking status indicator 1

smk2 Integer: Smoking status indicator 2

smk3 Integer: Smoking status indicator 3

lowed Integer: Low education indicator (0 = No, 1 = Yes)

headinj Integer: History of head injury (0 = No, 1 = Yes)

Details

The dataset name has been kept as `brains_cognitive_matrix` to avoid confusion with other datasets in the R ecosystem. This naming convention helps distinguish this dataset as part of the **NeuroDataSets** package. The suffix `matrix` indicates that the dataset is a matrix. The original content has not been modified.

Source

Data taken from the **RRMLRfMC** package version 0.4.0. Original study: University of Kentucky's Alzheimer's Disease Research Center BRAiNS cohort.

`brainvolume_df` *Meta-Analysis on Human Brain Volume and Intelligence*

Description

This dataset, brainvolume_df, is a data frame containing 83 empirical studies included in the meta-analysis by Pietschnig, Penke, Wicherts, Zeiler, and Voracek (2015), which examined the association between human brain volume and intelligence as measured by full-scale IQ. The dataset includes study identifiers, publication year, correlation coefficients, Fisher's z-transformed values, standard errors, sample sizes, sex composition, and mean participant age. These data provide a comprehensive resource for investigating population-level relationships between brain volume and cognitive ability.

Usage

```
data(brainvolume_df)
```

Format

A data frame with 83 observations and 8 variables:

study_name Study identifier (character)

year Year of publication (integer)

r Correlation coefficient between brain volume and intelligence (numeric)

z Fisher's z-transformed correlation (numeric)

z_se Standard error of the Fisher's z value (numeric)

n Sample size (integer)

sex Sex composition of the sample (factor with 4 levels)

mean_age Mean age of participants (numeric)

Details

The dataset name has been kept as brainvolume_df to avoid confusion with other datasets in the R ecosystem. This naming convention helps distinguish this dataset as part of the NeuroDataSets package and assists users in identifying its specific characteristics. The suffix df indicates that the dataset is a data frame. The original content has not been modified in any way.

Source

Data taken from the metaviz package version 0.3.1

---

brain_litter_mammals_df

Mammal Brain Size and Litter Size Relationship

Description

This dataset, `brain_litter_mammals_df`, is a data frame comparing relative brain weights between 96 mammalian species divided by reproductive strategy: 51 species with small litters (< 2 offspring) and 45 species with large litters (≥ 2 offspring).

Usage

```
data(brain_litter_mammals_df)
```

Format

A data frame with 96 observations and 2 variables:

BrainSize Numeric: Relative brain weight measurement (encephalization quotient or similar metric)

LitterSize Factor with 2 levels: Reproductive strategy ("Small" (< 2) and "Large" (≥ 2) litter sizes)

Details

The dataset name has been kept as `brain_litter_mammals_df` to avoid confusion with other datasets in the R ecosystem. This naming convention helps distinguish this dataset as part of the **NeuroDataSets** package. The suffix `df` indicates that the dataset is a data frame. The original content has not been modified.

Source

Data taken from the **Sleuth3** package version 1.0-6. Original reference: Ramsey, F.L. and Schafer, D.W. (2002) *The Statistical Sleuth: A Course in Methods of Data Analysis* (2nd ed), Duxbury.

`brain_size_iq_df` *Brain Size and IQ Study Data*

Description

This dataset, brain_size_iq_df, is a data frame containing neurocognitive measurements from a study examining relationships between brain size, gender, and intelligence. The data include 40 right-handed psychology students with no neurological history, selected based on extreme Scholastic Aptitude Test scores.

Usage

```
data(brain_size_iq_df)
```

Format

A data frame with 40 observations and 7 variables:

ID Numeric: Participant identification number

GENDER Factor with 2 levels: Participant's gender (Male/Female)

FSIQ Numeric: Full Scale IQ score

VIQ Numeric: Verbal IQ score

PIQ Numeric: Performance IQ score

MRI Numeric: Brain size measurement from MRI (in cubic cm)

IQDI Factor with 2 levels: IQ group classification (High/Low)

Details

The dataset name has been kept as 'brain_size_iq_df' to avoid confusion with other datasets in the R ecosystem. This naming convention helps distinguish this dataset as part of the NeuroDataSets package. The suffix 'df' indicates that the dataset is a data frame. The original content has not been modified.

Source

Data taken from the sur package version 1.0.4. Original study: Willerman, L., Schultz, R., Rutledge, J.N. and Bigler, E.D. (1991) In Vivo Brain Size and Intelligence. *Intelligence*, 15, 223-228.

---

brain_string_players_df

Brain Activity in String Players

Description

This dataset, brain_string_players_df, is a data frame containing neurophysiological measurements from a study of 15 violin and other string instrument players. The data examines the relationship between years of musical practice and measured brain activity levels in relevant cortical regions.

Usage

```
data(brain_string_players_df)
```

Format

A data frame with 15 observations and 2 variables:

Years Integer: Years of musical practice

Activity Numeric: Brain activity measurement (likely fMRI or similar neuroimaging units)

Details

The dataset name has been kept as 'brain_string_players_df' to avoid confusion with other datasets in the R ecosystem. This naming convention helps distinguish this dataset as part of the NeuroDataSets package. The suffix 'df' indicates that the dataset is a data frame. The original content has not been modified.

Source

Data taken from the Sleuth3 package version 1.0-6. Original reference: Ramsey, F.L. and Schafer, D.W. (2013) *The Statistical Sleuth: A Course in Methods of Data Analysis* (3rd ed), Cengage Learning.

`cerebellar_age_df` *Repeated Measurements of Age and Cerebellar Volume*

Description

This dataset, cerebellar_age_df, is a data frame containing repeated measurements of age and adjusted volume of cerebellar hemispheres from 72 participants. Each participant was measured on two occasions (Time), resulting in a total of 144 observations. The measurements were captured from Figure 8, Cerebellar Hemispheres (lower right) of Raz et al. (2005). The dataset includes participant identifiers, measurement time, age, and cerebellar hemisphere volume. Some observations may contain missing values.

Usage

```
data(cerebellar_age_df)
```

Format

A data frame with 144 observations and 4 variables:

Participant Participant ID (integer)

Time Measurement occasion (integer)

Age Age of the participant (numeric)

Volume Adjusted cerebellar hemisphere volume (numeric)

Details

The dataset name has been kept as cerebellar_age_df to avoid confusion with other datasets in the R ecosystem. This naming convention helps distinguish this dataset as part of the NeuroDataSets package and assists users in identifying its specific characteristics. The suffix df indicates that the dataset is a data frame. The original content has not been modified in any way.

Source

Data taken from the rmcorr package version 0.7.0

chimpbrains_df *Brodmann's Area 44 Asymmetry in Chimpanzees*

Description

This dataset, chimpbrains_df, is a data frame containing measurements of asymmetry in Brodmann's area 44 for 20 chimpanzees. Brodmann's area 44 is a brain region associated with language processing in humans and is located in the inferior frontal gyrus. The dataset includes individual identifiers, sex, and asymmetry measurements, providing insights into neural lateralization patterns in non-human primates. This data can be useful for comparative neuroanatomy studies and understanding the evolution of language-related brain structures.

Usage

```
data(chimpbrains_df)
```

Format

A data frame with 20 observations and 3 variables:

name Individual chimpanzee identifier (factor with 20 levels)

sex Sex of the chimpanzee (factor with 2 levels: "F" = female, "M" = male)

asymmetry Asymmetry measurement of Brodmann's area 44 (numeric)

Details

The dataset name has been kept as chimpbrains_df to avoid confusion with other datasets in the R ecosystem. This naming convention helps distinguish this dataset as part of the NeuroDataSets package and assists users in identifying its specific characteristics. The suffix df indicates that the dataset is a data frame. The original content has not been modified.

Source

Data taken from the abd package version 0.2-8

cocaine_dopamine_df *Effects of Cocaine on Dopamine Receptors*

Description

This dataset, cocaine_dopamine_df, is a data frame containing measurements of dopamine receptor blockade and perceived high levels from 34 human subjects as determined by PET scans.

Usage

```
data(cocaine_dopamine_df)
```

Format

A data frame with 34 observations and 2 variables:

percent.blocked Integer vector representing percent of dopamine receptors blocked

high Integer vector representing perceived level of high from PET scans

Details

The dataset name has been kept as 'cocaine_dopamine_df' to avoid confusion with other datasets in the R ecosystem. This naming convention helps distinguish this dataset as part of the NeuroDataSets package and assists users in identifying its specific characteristics. The suffix 'df' indicates that the dataset is a data frame. The original content has not been modified in any way.

Source

Data taken from the abd package version 0.2-8

`DA_schizophrenia_tbl_df`

DBH in Schizophrenia

Description

This dataset, 'DA_schizophrenia_tbl_df', is a tibble containing measurements of dopamine β-hydroxylase (DBH) activity in 25 schizophrenic patients treated with antipsychotic medication. The data compares DBH levels between patient groups.

Usage

```
data(DA_schizophrenia_tbl_df)
```

Format

A tibble with 25 observations and 2 variables:

dbh Integer: Dopamine β-hydroxylase activity level (nmol/(mL·hr))

group Character: Treatment/patient group classification

Details

The dataset name has been changed to `DA_schizophrenia_tbl_df` to provide a shorter, neuroscience-standard abbreviation where "DA" refers to dopamine. This naming convention maintains clarity and consistency within the **NeuroDataSets** package. The suffix `tbl_df` indicates that the dataset is a tibble. The original content has not been modified.

Source

Data taken from the **BSDA** package version 1.2.2

dementia_df *Dementia Scores Dataset*

Description

This dataset, dementia_df, is a data frame containing information related to dementia assessment. The data includes dementia scores along with demographic variables such as age and sex, as well as study identifiers. The dataset consists of 1,000 observations across 4 variables and was originally sourced from the PBImisc package. This dataset can be useful for analyzing patterns in dementia scores across different demographic groups and studies.

Usage

```
data(dementia_df)
```

Format

A data frame with 1,000 observations and 4 variables:

demscore Dementia score (integer)

age Age group of the participant (factor with 2 levels)

sex Sex of the participant (factor with 2 levels)

study Study identifier (factor with 10 levels)

Details

The dataset name has been kept as dementia_df to avoid confusion with other datasets in the R ecosystem. This naming convention helps distinguish this dataset as part of the NeuroDataSets package and assists users in identifying its specific characteristics. The suffix df indicates that the dataset is a data frame. The original content has not been modified.

Source

Data taken from the PBImisc package version 1.0

encephalitis_df *Cases of Herpes Encephalitis in Bavaria and Saxony*

Description

This dataset, encephalitis_df, is a data frame containing reported cases of herpes encephalitis in children from the regions of Bavaria and Lower Saxony. The data were collected between 1980 and 1993 as part of a study investigating the occurrence of herpes encephalitis in pediatric populations. The dataset includes the year of observation, regional identifiers, and the corresponding case counts, providing valuable information for epidemiological and public health research.

Usage

```
data(encephalitis_df)
```

Format

A data frame with 26 observations and 3 variables:

year Year of recorded cases (integer)

country Regional identifier (integer)

count Number of reported herpes encephalitis cases (integer)

Details

The dataset name has been kept as encephalitis_df to avoid confusion with other datasets in the R ecosystem. This naming convention helps distinguish this dataset as part of the NeuroDataSets package and assists users in identifying its specific characteristics. The suffix df indicates that the dataset is a data frame. The original content has not been modified in any way.

Source

Data taken from the catdata package version 1.2.4

---

epilepsy_drug_qol_df *SANAD Epilepsy Drug Treatment Quality of Life Study*

Description

This dataset, epilepsy_drug_qol_df, is a data frame containing quality of life measurements from the SANAD randomized controlled trial comparing carbamazepine and lamotrigine in 544 epilepsy patients. QoL assessments were collected at baseline, 3 months, 1 year and 2 years using validated instruments.

Usage

```
data(epilepsy_drug_qol_df)
```

Format

A data frame with 1,852 observations and 9 variables:

id Integer: Patient identification number

with.time Numeric: Time to withdrawal/discontinuation (days)

trt Factor with 2 levels: Treatment group (carbamazepine/lamotrigine)

with.status Integer: Withdrawal status indicator

time Numeric: Assessment time point (days since baseline)

anxiety Numeric: Anxiety score (from QoL measure)

depress Numeric: Depression score (from QoL measure)

aep Numeric: Adverse effects profile score

with.status2 Numeric: Alternative withdrawal coding

Details

The dataset name has been kept as 'epilepsy_drug_qol_df' to avoid confusion with other datasets in the R ecosystem. This naming convention helps distinguish this dataset as part of the NeuroDataSets package. The suffix 'df' indicates that the dataset is a data frame. The original content has not been modified.

Source

Data taken from the joineRML package version 0.4.7. Original study: Marson, A.G., et al. (2007) The SANAD study of effectiveness of carbamazepine, gabapentin, lamotrigine, oxcarbazepine, or topiramate for treatment of partial epilepsy: an unblinded randomised controlled trial. *The Lancet*, 369(9566), 1000-1015.

---

epilepsy_drug_trial_df

Epileptic Seizures Clinical Drug Trial

Description

This dataset, epilepsy_drug_trial_df, is a data frame containing seizure counts from a clinical trial of anti-epileptic medication. The data includes seizure frequency measurements along with treatment indicators and patient covariates for 295 observations.

Usage

```
data(epilepsy_drug_trial_df)
```

Format

A data frame with 295 observations and 6 variables:

seizures Numeric: Count of epileptic seizures

id Integer: Patient identification number

treat Numeric: Treatment indicator

expind Numeric: Exposure period indicator

timeadj Numeric: Adjusted time period

age Numeric: Patient age in years

Details

The dataset name has been kept as 'epilepsy_drug_trial_df' to avoid confusion with other datasets in the R ecosystem. This naming convention helps distinguish this dataset as part of the NeuroDataSets package. The suffix 'df' indicates that the dataset is a data frame. The original content has not been modified.

Source

Data taken from the faraway package version 1.0.9

`epilepsy_RCT_tbl_df` *Epilepsy Treatment Randomized Controlled Trial*

Description

This dataset, epilepsy_RCT_tbl_df, is a tibble containing data from a randomized controlled trial of progabide for epilepsy treatment. The trial recorded seizure counts for 59 patients at baseline and four follow-up visits.

Usage

```
data(epilepsy_RCT_tbl_df)
```

Format

A tibble with 59 observations and 8 variables:

id Integer: Patient identification number

treat Factor with 2 levels: Treatment group (progabide/control)

base Integer: Baseline seizure count

age Integer: Patient age in years

y1 Integer: Seizure count at first follow-up

y2 Integer: Seizure count at second follow-up

y3 Integer: Seizure count at third follow-up

y4 Integer: Seizure count at fourth follow-up

Details

The dataset name has been kept as 'epilepsy_RCT_tbl_df' to avoid confusion with other datasets in the R ecosystem. This naming convention helps distinguish this dataset as part of the NeuroDataSets package. The suffix 'tbl_df' indicates that the dataset is a tibble. The original content has not been modified.

Source

Data taken from the pubh package version 2.0.0

gm_expected_patterns_tbl_df
Patterns of Gray Matter in Schizophrenia

Description

This dataset, gm_expected_patterns_tbl_df, is a tibble containing expected patterns of gray matter in schizophrenia derived from large-scale meta-analyses by the ENIGMA consortium. It includes data from multiple neurological and psychiatric conditions for comparison.

Usage

```
data(gm_expected_patterns_tbl_df)
```

Format

A tibble with 33 observations and 16 variables:

GM Character vector indicating gray matter regions

SSD Numeric vector of expected patterns for schizophrenia spectrum disorder

MDD Numeric vector of expected patterns for major depressive disorder

AD_ADNI Numeric vector of expected patterns for Alzheimer's disease (ADNI cohort)

AD_ADNIOSYRIX Numeric vector of expected patterns for Alzheimer's disease (ADNI+OSYRIX cohort)

BD Numeric vector of expected patterns for bipolar disorder

PD Numeric vector of expected patterns for Parkinson's disease

Diabetes Numeric vector of expected patterns for diabetes

HighBP Numeric vector of expected patterns for high blood pressure

HighLipids Numeric vector of expected patterns for high lipids

MET Numeric vector of expected patterns for metabolic syndrome

DS_22q Numeric vector of expected patterns for 22q11.2 deletion syndrome

Suicide Numeric vector of expected patterns for suicide

OCD_pediatric Numeric vector of expected patterns for pediatric OCD

OCD_adult Numeric vector of expected patterns for adult OCD

AN Numeric vector of expected patterns for anorexia nervosa

Details

The dataset name has been kept as 'gm_expected_patterns_tbl_df' to avoid confusion with other datasets in the R ecosystem. This naming convention helps distinguish this dataset as part of the NeuroDataSets package and assists users in identifying its specific characteristics. The suffix 'tbl_df' indicates that the dataset is a tibble. The original content has not been modified in any way.

Source

Data taken from the RVIpkg package version 0.3.2.

`guineapig_neuro_df` *Guinea Pig Neural Data*

Description

This dataset, guineapig_neuro_df, is a data frame containing measurements of spontaneous current amplitudes recorded from individual brain cells in adult guinea pigs. The study investigated whether synaptic transmission occurs in quantal units, which would manifest as multimodal amplitude distributions with regularly spaced peaks.

Usage

```
data(guineapig_neuro_df)
```

Format

A data frame with 346 observations and 1 variable:

y Numeric: Peak amplitude of spontaneous synaptic currents (pA or similar units)

Details

The dataset name has been updated to 'guineapig_neuro_df' for clarity and brevity while preserving consistency with other datasets in the NeuroDataSets package. The suffix 'df' indicates that the dataset is a standard data frame.

Source

Data taken from the boot package version 1.3-31. Original study: Paulsen, O. and Heggelund, P. (1994) The quantal size at retinogeniculate synapses determined from spontaneous and evoked EPSCs in guinea-pig thalamic slices. *Journal of Physiology*, 480, 505–511.

hippocampus_lesions_df *Memory and the Hippocampus*

Description

This dataset, hippocampus_lesions_df, is a data frame containing measurements of spatial memory scores and percent lesion of the hippocampus from 57 observations.

Usage

```
data(hippocampus_lesions_df)
```

Format

A data frame with 57 observations and 2 variables:

lesion Numeric vector representing percent lesion of the hippocampus

memory Numeric vector representing spatial memory scores

Details

The dataset name has been kept as 'hippocampus_lesions_df' to avoid confusion with other datasets in the R ecosystem. This naming convention helps distinguish this dataset as part of the NeuroDataSets package and assists users in identifying its specific characteristics. The suffix 'df' indicates that the dataset is a data frame. The original content has not been modified in any way.

Source

Data taken from the abd package version 0.2-8

iq_country_tbl_df *Average Global IQ per Country*

Description

This dataset, iq_country_tbl_df, is a tibble containing information on the average intelligence quotient (IQ) of countries around the world. In addition to average IQ scores, the dataset includes several socioeconomic and demographic indicators such as literacy rate, number of Nobel Prizes won collectively by each country, Human Development Index (HDI, 2021), mean years of schooling (2021), gross national income (GNI, 2021), and population estimates for 2023. These variables provide a broad context for understanding cognitive performance at the country level.

Usage

```
data(iq_country_tbl_df)
```

Format

A tibble with 193 observations and 10 variables:

Rank Global ranking based on average IQ (numeric)

Country Name of the country (character)

Average IQ Estimated average IQ score of the population (numeric)

Continent Continent to which the country belongs (character)

Literacy Rate Literacy rate of the population (numeric)

Nobel Prices Total number of Nobel Prizes won collectively by the country (numeric)

HDI (2021) Human Development Index for the year 2021 (numeric)

Mean years of schooling - 2021 Average years of schooling in 2021 (numeric)

GNI - 2021 Gross national income for 2021 (numeric)

Population - 2023 Estimated population in 2023 (character)

Details

The dataset name has been kept as iq_country_tbl_df to avoid confusion with other datasets in the R ecosystem. This naming convention helps distinguish this dataset as part of the NeuroDataSets package and assists users in identifying its specific characteristics. The suffix tbl_df indicates that the dataset is a tibble (a modern data frame). The original content has not been modified in any way. Variable names and values are provided exactly as they appear in the source.

Source

Data taken from Kaggle: https://www.kaggle.com/datasets/mlippo/average-global-iq-per-country-with-other

---

mammals_brain_body_df *Mammal Brain and Body Size*

Description

This dataset, mammals_brain_body_df, is a data frame containing comparative neuroanatomical and life history data for 96 mammalian species. The data examine the relationship between brain size, body size, and reproductive characteristics across different mammal species.

Usage

```
data(mammals_brain_body_df)
```

Format

A data frame with 96 observations and 5 variables:

Species Factor with 96 levels: Mammalian species names

Brain Numeric: Brain weight (grams)

Body Numeric: Body weight (kilograms)

Gestation Integer: Gestation period (days)

Litter Numeric: Average litter size

Details

The dataset name has been kept as 'mammals_brain_body_df' to avoid confusion with other datasets in the R ecosystem. This naming convention helps distinguish this dataset as part of the NeuroDataSets package. The suffix 'df' indicates that the dataset is a data frame. The original content has not been modified.

Source

Data taken from the Sleuth3 package version 1.0-6. Original study: Allison, T. and Cicchetti, D.V. (1976) Sleep in Mammals: Ecological and Constitutional Correlates. *Science*, 194, 732-734.

`markers_brain_df` *Cross-Species Brain Cell Marker Genes*

Description

This dataset, markers_brain_df, is a data frame containing the top 1,000 marker genes for each of six major brain cell types (astrocytes, endothelial cells, microglia, neurons, oligodendrocytes, and OPCs) identified through meta-analysis of both human and mouse brain gene expression data.

Usage

```
data(markers_brain_df)
```

Format

A data frame with 6,000 observations and 2 variables:

markers Character: Gene symbol for cell-type specific marker (human/mouse orthologs)

cell Character: Cell type classification (astrocytes/endothelial/microglia/neurons/oligodendrocytes/OPCs)

Details

The dataset name has been kept as 'markers_brain_df' to avoid confusion with other datasets in the R ecosystem. This naming convention helps distinguish this dataset as part of the NeuroDataSets package. The suffix 'df' indicates that the dataset is a data frame. The original content has not been modified.

Source

Data taken from the BRETIGEA package version 1.0.3. Derived from: Meta-analysis of human and mouse brain cell-type specific gene expression datasets.

`markers_human_brain_df`

Human Brain Cell Marker Genes

Description

This dataset, markers_human_brain_df, is a data frame containing the top 1,000 marker genes for each of six major brain cell types (astrocytes, endothelial cells, microglia, neurons, oligodendrocytes, and OPCs) identified through meta-analysis of human brain gene expression data.

Usage

```
data(markers_human_brain_df)
```

Format

A data frame with 5,500 observations and 2 variables:

markers Character: Gene symbol for cell-type specific marker

cell Character: Cell type classification (astrocytes/endothelial/microglia/neurons/oligodendrocytes/OPCs)

Details

The dataset name has been kept as 'markers_human_brain_df' to avoid confusion with other datasets in the R ecosystem. This naming convention helps distinguish this dataset as part of the NeuroDataSets package. The suffix 'df' indicates that the dataset is a data frame. The original content has not been modified.

Source

Data taken from the BRETIGEA package version 1.0.3.

markers_mouse_brain_df *Mouse Brain Cell Marker Genes*

Description

This dataset, markers_mouse_brain_df, is a data frame containing the top 1,000 marker genes for each of six major brain cell types (astrocytes, endothelial cells, microglia, neurons, oligodendrocytes, and OPCs) identified through meta-analysis of mouse brain gene expression data.

Usage

```
data(markers_mouse_brain_df)
```

Format

A data frame with 5,430 observations and 2 variables:

markers Character: Gene symbol for cell-type specific marker

cell Character: Cell type classification (astrocytes/endothelial/microglia/neurons/oligodendrocytes/OPCs)

Details

The dataset name has been kept as 'markers_mouse_brain_df' to avoid confusion with other datasets in the R ecosystem. This naming convention helps distinguish this dataset as part of the NeuroDataSets package. The suffix 'df' indicates that the dataset is a data frame. The original content has not been modified.

Source

Data taken from the BRETIGEA package version 1.0.3. Original study: Mckenzie AT, Wang M, Hauberg ME, et al. (2018) Brain Cell Type Specific Gene Expression and Co-expression Network Architectures. *Scientific Reports*, 8(1), 8868.

migraines_df *Effects of Transcranial Magnetic Stimulation on Migraine Headaches*

Description

This dataset, migraines_df, is a data frame containing data on the effects of transcranial magnetic stimulation (TMS) on migraine headaches. The dataset includes two groups along with counts of participants who reported improvement ("Yes"), no improvement ("No"), and the total number of trials. These data are useful for evaluating the potential therapeutic impact of TMS on migraine symptoms.

Usage

```
data(migraines_df)
```

Format

A data frame with 2 observations and 4 variables:

Group Group indicator (factor with 2 levels)

Yes Number of participants reporting improvement (integer)

No Number of participants reporting no improvement (integer)

Trials Total number of trials (integer)

Details

The dataset name has been kept as migraines_df to avoid confusion with other datasets in the R ecosystem. This naming convention helps distinguish this dataset as part of the NeuroDataSets package and assists users in identifying its specific characteristics. The suffix df indicates that the dataset is a data frame. The original content has not been modified in any way.

Source

Data taken from the Stat2Data package version 2.0.0

---

migraine_treatment_df *Migraine Headache Treatment*

Description

This dataset, migraine_treatment_df, is a data frame containing clinical data on 4,152 migraine treatment cases collected by Tammy Kostecki-Dillon. The data includes treatment details, headache characteristics, and patient demographics.

Usage

```
data(migraine_treatment_df)
```

Format

A data frame with 4,152 observations and 9 variables:

id Integer: Patient identification number

time Integer: Time measurement (likely days or hours)

dos Integer: Treatment dosage

hatype Factor with 3 levels: Headache type classification

age Integer: Patient age in years

airq Numeric: Air quality index measurement

medication Factor with 3 levels: Medication type

headache Factor with 2 levels: Headache presence/severity

sex Factor with 2 levels: Patient sex

Details

The dataset name has been kept as 'migraine_treatment_df' to avoid confusion with other datasets in the R ecosystem. This naming convention helps distinguish this dataset as part of the NeuroDataSets package. The suffix 'df' indicates that the dataset is a data frame. The original content has not been modified.

Source

Data taken from the carData package version 3.0-5. Original collection: Kostecki-Dillon, T. (Year not specified) Migraine Treatment Study.

---

migrane_dose_df *Migraine Dose–Response Trial Data*

Description

This dataset, migrane_dose_df, is a data frame obtained from a randomized, placebo-controlled dose–response clinical trial for the treatment of acute migraine (clinicaltrials.gov identifier NCT00712725). The primary endpoint was "pain freedom at 2 hours postdose," measured as a binary outcome. The dataset includes dose levels, the number of participants achieving pain freedom, and the total number of treated participants at each dose level. These data are useful for dose–response modeling and clinical trial analysis in migraine research.

Usage

```
data(migrane_dose_df)
```

Format

A data frame with 8 observations and 3 variables:

dose Dose level administered (numeric)

painfree Number of participants who achieved pain freedom at 2 hours postdose (integer)

ntrt Total number of treated participants at the corresponding dose level (integer)

Details

The dataset name has been kept as migrane_dose_df to avoid confusion with other datasets in the R ecosystem. This naming convention helps distinguish this dataset as part of the NeuroDataSets package and assists users in identifying its specific characteristics. The suffix df indicates that the dataset is a data frame. The original content has not been modified in any way.

Source

Data taken from the DoseFinding package version 1.4-1

`neanderthal_brains_df` *Cranial Capacity in Neanderthals and Modern Humans*

Description

This dataset, neanderthal_brains_df, is a data frame containing measurements of brain size (lnbrain) and body mass (lnmass) from 39 specimens of Neanderthals and early modern humans, identified by species.

Usage

```
data(neanderthal_brains_df)
```

Format

A data frame with 39 observations and 3 variables:

ln.mass Numeric vector representing natural logarithm of body mass

ln.brain Numeric vector representing natural logarithm of brain size

species Factor indicating species with 2 levels (Neanderthals and early modern humans)

Details

The dataset name has been kept as 'neanderthal_brains_df' to avoid confusion with other datasets in the R ecosystem. This naming convention helps distinguish this dataset as part of the NeuroDataSets package and assists users in identifying its specific characteristics. The suffix 'df' indicates that the dataset is a data frame. The original content has not been modified in any way.

Source

Data taken from the abd package version 0.2-8

NeuroDataSets — *NeuroDataSets: A Comprehensive Collection of Neuroscience and Brain-Related Datasets*

Description

This package provides a diverse collection of datasets focused on the brain, nervous system, and related disorders. The package includes clinical, experimental, neuroimaging, behavioral, and cognitive data on conditions including Parkinson's, Alzheimer's, epilepsy, schizophrenia, autism, ADHD, Tourette's, TBI, brain tumors, migraines, sleep disorders, and mental health.

Details

NeuroDataSets: A Comprehensive Collection of Neuroscience and Brain-Related Datasets

A Comprehensive Collection of Neuroscience and Brain-Related Datasets.

Author(s)

Maintainer: Renzo Caceres Rossi <arenzocaceresrossi@gmail.com>

See Also

Useful links:

- https://github.com/lightbluetitan/neurodatasets

neurodeg_dose_df — *Simulated Neurodegenerative Disease Dose Data*

Description

This dataset, neurodeg_dose_df, is a data frame containing simulated longitudinal data from a Phase 2 clinical study of a potential treatment for a neurodegenerative disease. The disease state is assessed using a functional scale, where smaller values indicate more severe neurodeterioration. The primary goal of the drug is to slow disease progression, which is quantified through the linear slope of the functional scale over time. The dataset includes repeated measurements for multiple individuals across varying dose levels, allowing investigation of dose–response relationships in disease progression.

Usage

```
data(neurodeg_dose_df)
```

Format

A data frame with 1250 observations and 4 variables:

resp Measured value of the functional scale (numeric)

id Participant identifier (integer)

dose Dose level administered (numeric)

time Measurement time point (numeric)

Details

The dataset name has been kept as neurodeg_dose_df to avoid confusion with other datasets in the R ecosystem. This naming convention helps distinguish this dataset as part of the NeuroDataSets package and assists users in identifying its specific characteristics. The suffix df indicates that the dataset is a data frame. The original content has not been modified in any way.

Source

Data taken from the DoseFinding package version 1.4-1

---

neuro_pointprocess_matrix

Neurophysiological Point Process Data

Description

This dataset, neuro_pointprocess_matrix, is a matrix containing times of observed neuronal firing in windows of 250ms surrounding stimulus application in human subjects. Each row represents an experimental replication (469 total replicates), with values indicating spike times relative to stimulus onset.

Usage

```
data(neuro_pointprocess_matrix)
```

Format

A numeric matrix with 469 observations (rows) and 6 variables (columns):

[,1:6] Numeric: Spike times (milliseconds) relative to stimulus onset, with NA representing no spike in that trial window

Details

The dataset name has been kept as 'neuro_pointprocess_matrix' to avoid confusion with other datasets in the R ecosystem. This naming convention helps distinguish this dataset as part of the NeuroDataSets package. The suffix 'matrix' indicates that the dataset is a matrix. The original content has not been modified.

Source

Data taken from the boot package version 1.3-31. Original collection: Dr. S.J. Boniface, Neurophysiology Unit, Radcliffe Infirmary, Oxford.

`nfl_concussions_tbl_df`

Concussions in the National Football League (2012–2014)

Description

This dataset, nfl_concussions_tbl_df, is a tibble containing detailed information on concussion injuries that occurred in the National Football League (NFL) from 2012 to 2014. The dataset includes hundreds of recorded concussion cases, capturing information such as player identity, team, game, date of injury, position, whether the injury occurred during pre-season, and multiple injury-related details including weeks injured, games missed, and reported injury type.

Usage

```
data(nfl_concussions_tbl_df)
```

Format

A tibble with 392 observations and 18 variables:

ID Unique identifier for each concussion record (character)

Player Name of the player who sustained the concussion (character)

Team Team of the injured player (character)

Game Game in which the injury occurred (character)

Date Date of the concussion incident (character)

Opposing Team Opponent team during the game (character)

Position Player's position (character)

Pre-Season Injury? Indicates if the injury occurred during pre-season (character)

Winning Team? Indicates if the player's team won the game (character)

Week of Injury Week number of the season when the injury occurred (numeric)

Season NFL season year associated with the injury (character)

Weeks Injured Number of weeks the player was injured (numeric)

Games Missed Number of games missed due to the concussion (numeric)

Unknown Injury? Indicates if the injury type was unknown (character)

Reported Injury Type Reported type of concussion injury (character)

Total Snaps Total snaps played before injury (numeric)

Play Time After Injury Playtime after injury occurred (character)

Average Playtime Before Injury Average playtime before injury (character)

Details

The dataset name has been kept as nfl_concussions_tbl_df to avoid confusion with other datasets in the R ecosystem. This naming convention helps distinguish this dataset as part of the NeuroDataSets package and assists users in identifying its specific characteristics. The suffix tbl_df indicates that the dataset is a tibble (a modern data frame). The original content has not been modified in any way. Variable names and values are provided exactly as they appear in the source.

Source

Data taken from Kaggle: https://www.kaggle.com/datasets/rishidamarla/concussions-in-the-nfl-20122014

OASIS_cross_tbl_df *Cross-sectional Brain MRI Data Across Adult Lifespan*

Description

This dataset, OASIS_cross_tbl_df, is a tibble containing a cross-sectional collection of MRI data from 436 individuals aged 18 to 96, obtained as part of the Open Access Series of Imaging Studies (OASIS). For each subject, 3 or 4 T1-weighted MRI scans acquired during a single scanning session are included. All participants are right-handed and include both men and women. Among the subjects over the age of 60, 100 have been clinically diagnosed with very mild to moderate Alzheimer's disease (AD).

Usage

```
data(OASIS_cross_tbl_df)
```

Format

A tibble with 436 observations and 12 variables:

ID Subject identifier (character)

M/F Sex of the participant (character)

Hand Handedness of the participant (character)

Age Age in years (numeric)

Educ Years of education (numeric)

SES Socioeconomic status score (numeric)

MMSE Mini-Mental State Examination score (numeric)

CDR Clinical Dementia Rating score (numeric)

eTIV Estimated total intracranial volume (numeric)

nWBV Normalized whole-brain volume (numeric)

ASF Atlas scaling factor (numeric)

Delay Inter-scan interval in days (character)

Details

The dataset name has been kept as OASIS_cross_tbl_df to avoid confusion with other datasets in the R ecosystem. This naming convention helps distinguish this dataset as part of the NeuroDataSets package and assists users in identifying its specific characteristics. The suffix tbl_df indicates that the dataset is a tibble (a modern data frame). The original content has not been modified in any way. Variable names and values are provided exactly as they appear in the source.

Source

Data taken from Kaggle: https://www.kaggle.com/datasets/jboysen/mri-and-alzheimers

---

OASIS_long_tbl_df *Longitudinal MRI Data in Nondemented and Demented Older Adults*

Description

This dataset, OASIS_long_tbl_df, is a tibble containing a longitudinal collection of MRI data from 150 subjects aged 60 to 96, obtained as part of the Open Access Series of Imaging Studies (OASIS). Each participant completed two or more MRI sessions, with visits spaced at least one year apart, resulting in a total of 373 imaging sessions. The dataset includes both nondemented and demented older adults and provides comprehensive demographic, clinical, and neuroimaging measures for each visit.

Usage

```
data(OASIS_long_tbl_df)
```

Format

A tibble with 373 observations and 15 variables:

Subject ID Unique identifier for each subject (character)

MRI ID Identifier for each MRI session (character)

Group Clinical group classification (character)

Visit Visit number for longitudinal assessment (numeric)

MR Delay Time in days between MRI sessions (numeric)

M/F Sex of the participant (character)

Hand Handedness of the participant (character)

Age Age in years at the time of the visit (numeric)

EDUC Years of education (numeric)

SES Socioeconomic status score (numeric)

MMSE Mini-Mental State Examination score (numeric)

CDR Clinical Dementia Rating score (numeric)

eTIV Estimated total intracranial volume (numeric)

nWBV Normalized whole-brain volume (numeric)

ASF Atlas scaling factor (numeric)

Details

The dataset name has been kept as OASIS_long_tbl_df to avoid confusion with other datasets in the R ecosystem. This naming convention helps distinguish this dataset as part of the NeuroDataSets package and assists users in identifying its specific characteristics. The suffix tbl_df indicates that the dataset is a tibble (a modern data frame). The original content has not been modified in any way. Variable names and values are provided exactly as they appear in the source.

Source

Data taken from Kaggle: https://www.kaggle.com/datasets/jboysen/mri-and-alzheimers

---

parkinsons_dopamine_list

Dopamine Agonists as Adjunct Therapy in Parkinson's

Description

This dataset, parkinsons_dopamine_list, is a list containing information from 7 studies investigating the mean lost work-time reduction in patients given 4 dopamine agonists and placebo as adjunct therapy for Parkinson's disease. There is placebo and four active drugs coded 2 to 5.

Usage

```
data(parkinsons_dopamine_list)
```

Format

A list with 5 components:

Outcomes Numeric vector containing the outcomes (mean lost work-time reduction)

SE Numeric vector containing standard errors for the outcomes

Treat Character vector indicating the treatment (placebo or drug codes 2-5)

Study Numeric vector indicating the study number (1-7)

Treat.order Character vector showing the treatment order (placebo and drugs 2-5)

Details

The dataset name has been kept as 'parkinsons_dopamine_list' to avoid confusion with other datasets in the R ecosystem. This naming convention helps distinguish this dataset as part of the NeuroDataSets package and assists users in identifying its specific characteristics. The suffix 'list' indicates that the dataset is a list. The original content has not been modified in any way.

Source

Data taken from the bnma package version 1.6.0.

`pediatric_glioma_tbl_df` *Pediatric High-Grade Glioma Clinical Dataset*

Description

This dataset, pediatric_glioma_tbl_df, is a tibble containing comprehensive clinical and tumor characteristics for 57 pediatric patients with high-grade glioma. The data includes 22 variables covering demographic, symptomatic, pathological, treatment, and outcome measures.

Usage

```
data(pediatric_glioma_tbl_df)
```

Format

A tibble with 57 observations and 22 variables:

Age Numeric: Patient age in years

Gender Character: Patient gender

Headache Character: Headache presence/characteristics

Epilepsy Character: Epilepsy status

Hemparesis Character: Hemiparesis presence

increaseICT Character: Increased intracranial pressure indicators

Pathology Character: Tumor pathology classification

Pathology_Grade Numeric: WHO tumor grade (III-IV)

Thalamic_extension Character: Thalamic involvement

Bil_extension Character: Bilateral extension

Post_extension Character: Posterior fossa extension

BrainStem_extension Character: Brainstem involvement

MultiFocality Character: Multifocal tumor presence

Midlineshift Character: Midline shift presence

Edema Character: Peritumoral edema characteristics

Approx_Tumor_Vol Numeric: Estimated tumor volume (cm^3)

ExtentofSurgicalresection Character: Surgical resection extent

Shunt Character: Ventricular shunt presence

ResidualsizeonMRI Character: Post-surgical residual tumor

Neurostate Character: Neurological status

PSBeforeRT Numeric: Performance status pre-radiotherapy

Died Character: Mortality outcome

Details

The dataset name has been kept as 'pediatric_glioma_tbl_df' to avoid confusion with other datasets in the R ecosystem. This naming convention helps distinguish this dataset as part of the NeuroDataSets package. The suffix 'tbl_df' indicates that the dataset is a tibble. The original content has not been modified.

Source

Kaggle dataset: Pediatric High-Grade Glioma Dataset. URL: https://www.kaggle.com/datasets/amraam/pediatric-high-grade-glioma-dataset

---

psych_neurocog_df *Psychotic Cognition*

Description

This dataset, psych_neurocog_df, is a data frame containing comprehensive neurocognitive assessments from a study comparing performance patterns in schizophrenia, schizoaffective disorder, and controls. The data includes 242 observations across multiple cognitive domains using a psychosis-specific neurocognitive battery.

Usage

```
data(psych_neurocog_df)
```

Format

A data frame with 242 observations and 10 variables:

Dx Factor with 3 levels: Diagnostic group (Schizophrenia/Schizoaffective/Control)

Speed Integer: Processing speed score

Attention Integer: Attention/vigilance score

Memory Integer: Working memory score

Verbal Integer: Verbal learning score

Visual Integer: Visual learning score

ProbSolv Integer: Problem solving score

SocialCog Integer: Social cognition score

Age Integer: Participant age in years

Sex Factor with 2 levels: Participant sex

Details

The dataset name has been updated to 'psych_neurocog_df' for brevity and clarity, while maintaining consistency with the naming style of the NeuroDataSets package. The suffix 'df' indicates that the dataset is a data frame.

Source

Data taken from the heplots package version 1.7.4. Original research: Hartman, L.I. (2016) Schizophrenia and Schizoaffective Disorder: One Condition or Two? Unpublished PhD dissertation, York University.

`SAHemorrhage_df` *Subarachnoid Hemorrhage Clinical and Laboratory Data*

Description

This dataset, SAHemorrhage_df, is a data frame containing clinical and laboratory variables from 113 patients diagnosed with aneurysmal subarachnoid hemorrhage. The dataset includes functional outcomes, demographic information, clinical severity scores, and biomarker measurements. These data provide valuable information for studying neurological prognosis, biomarker associations, and clinical patterns in patients with subarachnoid hemorrhage.

Usage

```
data(SAHemorrhage_df)
```

Format

A data frame with 113 observations and 7 variables:

gos6 Glasgow Outcome Scale at 6 months (ordered factor with 5 levels)

outcome Clinical outcome classification (factor with 2 levels)

gender Gender of the patient (factor with 2 levels)

age Age of the patient (integer)

wfns WFNS clinical grade (ordered factor with 5 levels)

s100b S100B biomarker level (numeric)

ndka Nucleoside diphosphate kinase A level (numeric)

Details

The dataset name has been kept as SAHemorrhage_df to avoid confusion with other datasets in the R ecosystem. This naming convention helps distinguish this dataset as part of the NeuroDataSets package and assists users in identifying its specific characteristics. The suffix df indicates that the dataset is a data frame. The original content has not been modified in any way.

Source

Data taken from the reportROC package version 3.6

`sleep_deprivation_tbl_df` *Sleep Deprivation and Cognitive Performance Data*

Description

This dataset, sleep_deprivation_tbl_df, is a tibble containing data from a 2024 study conducted in the Middle East that investigated the effects of sleep deprivation on cognitive performance and emotional regulation. The dataset includes 60 participants from diverse backgrounds and captures detailed information on sleep duration, sleep quality, daytime sleepiness, cognitive performance metrics (reaction times and memory accuracy), and emotional stability. Additionally, the dataset records demographic and lifestyle factors such as age, gender, BMI, caffeine intake, physical activity level, and stress level.

Usage

```
data(sleep_deprivation_tbl_df)
```

Format

A tibble with 60 observations and 14 variables:

Participant_ID Unique identifier for each participant (character)

Sleep_Hours Average hours of sleep per night (numeric)

Sleep_Quality_Score Self-reported sleep quality score (numeric)

Daytime_Sleepiness Level of daytime sleepiness (numeric)

Stroop_Task_Reaction_Time Reaction time on the Stroop cognitive task (numeric)

N_Back_Accuracy Accuracy score on the N-Back working memory task (numeric)

Emotion_Regulation_Score Score reflecting emotional regulation ability (numeric)

PVT_Reaction_Time Reaction time on the Psychomotor Vigilance Task (numeric)

Age Age of the participant in years (numeric)

Gender Gender of the participant (character)

BMI Body Mass Index (numeric)

Caffeine_Intake Daily caffeine intake (numeric)

Physical_Activity_Level Self-reported physical activity level (numeric)

Stress_Level Self-reported stress level (numeric)

Details

The dataset name has been kept as sleep_deprivation_tbl_df to avoid confusion with other datasets in the R ecosystem. This naming convention helps distinguish this dataset as part of the NeuroDataSets package and assists users in identifying its specific characteristics. The suffix tbl_df indicates that the dataset is a tibble (a modern data frame). The original content has not been modified in any way. Variable names and values are provided exactly as they appear in the source.

Source

Data taken from Kaggle: https://www.kaggle.com/datasets/sacramentotechnology/sleep-deprivation-and-cogn

---

`sleep_disorder_df` *Transient Sleep Disorder Polysomnography Scoring Data*

Description

This dataset, sleep_disorder_df, is a data frame containing polysomnographic (PSG) measurements from a clinical study designed to compare automated and semi-automated scoring methods used in the diagnosis of transient sleep disorders. The study included 82 patients who were administered a sleep-inducing drug (Zolpidem 10 mg). The primary measure of interest is the latency to persistent sleep (LPS), defined as the time from lights out to the beginning of 10 consecutive minutes of uninterrupted sleep. LPS was measured using three different scoring methods: manual, automated, and partial.

Usage

```
data(sleep_disorder_df)
```

Format

A data frame with 82 observations and 3 variables:

manual Latency to persistent sleep measured using manual scoring (numeric)

automated Latency to persistent sleep measured using automated scoring (numeric)

partial Latency to persistent sleep measured using semi-automated (partial) scoring (numeric)

Details

The dataset name has been kept as sleep_disorder_df to avoid confusion with other datasets in the R ecosystem. This naming convention helps distinguish this dataset as part of the NeuroDataSets package and assists users in identifying its specific characteristics. The suffix df indicates that the dataset is a data frame. The original content has not been modified in any way.

Source

Data taken from the MVT package version 0.3-81

sleep_performance_df *Sleep and Learning Performance*

Description

This dataset, sleep_performance_df, is a data frame containing measurements of the increase in slow-wave sleep and corresponding improvements in spatial learning tasks from 10 human subjects.

Usage

```
data(sleep_performance_df)
```

Format

A data frame with 10 observations and 2 variables:

sleep Integer vector representing increase in slow-wave sleep (units)

improvement Integer vector representing improvement in spatial learning tasks (units)

Details

The dataset name has been kept as 'sleep_performance_df' to avoid confusion with other datasets in the R ecosystem. This naming convention helps distinguish this dataset as part of the NeuroDataSets package and assists users in identifying its specific characteristics. The suffix 'df' indicates that the dataset is a data frame. The original content has not been modified in any way.

Source

Data taken from the abd package version 0.2-8

subcortical_patterns_tbl_df *Patterns of Subcortical Structures*

Description

This dataset, subcortical_patterns_tbl_df, is a tibble containing expected patterns of subcortical structures in schizophrenia derived from large-scale meta-analyses by the ENIGMA consortium. It includes data from multiple neurological and psychiatric conditions for comparison.

Usage

```
data(subcortical_patterns_tbl_df)
```

Format

A tibble with 8 observations and 16 variables:

Subcortical Character vector indicating subcortical regions

SSD Numeric vector of expected patterns for schizophrenia spectrum disorder

MDD Numeric vector of expected patterns for major depressive disorder

AD_ADNI Numeric vector of expected patterns for Alzheimer's disease (ADNI cohort)

AD_ADNIOSYRIX Numeric vector of expected patterns for Alzheimer's disease (ADNI+OSYRIX cohort)

BD Numeric vector of expected patterns for bipolar disorder

PD Numeric vector of expected patterns for Parkinson's disease

Diabetes Numeric vector of expected patterns for diabetes

HighBP Numeric vector of expected patterns for high blood pressure

HighLipids Numeric vector of expected patterns for high lipids

MET Numeric vector of expected patterns for metabolic syndrome

DS_22q Numeric vector of expected patterns for 22q11.2 deletion syndrome

Suicide Numeric vector of expected patterns for suicide

OCD_pediatric Numeric vector of expected patterns for pediatric OCD

OCD_adult Numeric vector of expected patterns for adult OCD

AN Numeric vector of expected patterns for anorexia nervosa

Details

The dataset name has been kept as 'subcortical_patterns_tbl_df' to avoid confusion with other datasets in the R ecosystem. This naming convention helps distinguish this dataset as part of the NeuroDataSets package and assists users in identifying its specific characteristics. The suffix 'tbl_df' indicates that the dataset is a tibble. The original content has not been modified in any way.

Source

Data taken from the RVIpkg package version 0.3.2

TBI_age_tbl_df *Traumatic Brain Injury (TBI) Cases by Age Group and Injury Mechanism*

Description

This dataset, TBI_age_tbl_df, is a tibble containing information from the year 2014 on traumatic brain injury (TBI) cases across different age groups. The dataset provides details on the mechanisms that caused the injuries, the type of injury, the estimated number of observed cases, and the estimated rate of cases per 100,000 people.

Usage

```
data(TBI_age_tbl_df)
```

Format

A tibble with 231 observations and 5 variables:

age_group Age group category (character)

type Type of traumatic brain injury (character)

injury_mechanism Mechanism by which the injury occurred (character)

number_est Estimated number of observed cases in 2014 (numeric)

rate_est Estimated rate of cases per 100,000 population in 2014 (numeric)

Details

The dataset name has been kept as TBI_age_tbl_df to avoid confusion with other datasets in the R ecosystem. This naming convention helps distinguish this dataset as part of the NeuroDataSets package and assists users in identifying its specific characteristics. The suffix tbl_df indicates that the dataset is a tibble (a modern data frame). The original content has not been modified in any way. Variable names and values are provided exactly as they appear in the source.

Source

Data taken from Kaggle: https://www.kaggle.com/datasets/jessemostipak/traumatic-brain-injury-tbi

---

TBI_military_tbl_df *Traumatic Brain Injury (TBI) in U.S. Military Personnel*

Description

This dataset, TBI_military_tbl_df, is a tibble containing information on traumatic brain injuries (TBI) diagnosed among U.S. military personnel. The dataset includes the service branch, military component, severity of the injury, number of diagnosed cases, and the year of observation.

Usage

```
data(TBI_military_tbl_df)
```

Format

A tibble with 438 observations and 5 variables:

service Branch of military service (character)

component Status of the individual (active duty, reserve, or guard) (character)

severity Severity category of the traumatic brain injury (character)

diagnosed Number of diagnosed TBI cases (numeric)

year Year of recorded TBI diagnosis (numeric)

Details

The dataset name has been kept as TBI_military_tbl_df to avoid confusion with other datasets in the R ecosystem. This naming convention helps distinguish this dataset as part of the NeuroDataSets package and assists users in identifying its specific characteristics. The suffix tbl_df indicates that the dataset is a tibble (a modern data frame). The original content has not been modified in any way. Variable names and values are provided exactly as they appear in the source.

Source

Data taken from Kaggle: https://www.kaggle.com/datasets/jessemostipak/traumatic-brain-injury-tbi

---

TBI_steroids_df *Corticosteroids in Acute Traumatic Brain Injury*

Description

This dataset, TBI_steroids_df, is a data frame containing data from a systematic review evaluating the effects of corticosteroids on mortality in patients with acute traumatic brain injury. The dataset includes results from randomized controlled trials, including the influential MRC CRASH trial (Roberts et al. 2001). Variables include study identifiers, numbers of deaths in the corticosteroid and control groups, and corresponding sample sizes. These data are useful for meta-analytic investigations of corticosteroid efficacy in traumatic brain injury.

Usage

```
data(TBI_steroids_df)
```

Format

A data frame with 17 observations and 5 variables:

study Study identifier (character)

event.steroid Number of deaths in the corticosteroid group (numeric)

n.steroid Sample size of the corticosteroid group (numeric)

event.control Number of deaths in the control group (numeric)

n.control Sample size of the control group (numeric)

Details

The dataset name has been kept as TBI_steroids_df to avoid confusion with other datasets in the R ecosystem. This naming convention helps distinguish this dataset as part of the NeuroDataSets package and assists users in identifying its specific characteristics. The suffix df indicates that the dataset is a data frame. The original content has not been modified in any way.

Source

Data taken from the ratesci package version 1.0.0

tourette_ADHD_df *Attentional Dysfunction in Adults With Tourette's Syndrome*

Description

This dataset, tourette_ADHD_df, is a data frame containing accuracy scores from 51 adult participants grouped into three categories related to Tourette's Syndrome and attentional dysfunction. The data include performance accuracy and group membership, allowing comparison across diagnostic groups. Some observations may contain missing values. The dataset originates from research on attentional processes in adults with Tourette's Syndrome.

Usage

```
data(tourette_ADHD_df)
```

Format

A data frame with 51 observations and 2 variables:

accuracy Accuracy score (numeric)

group Participant group (factor with 3 levels)

Details

The dataset name has been kept as tourette_ADHD_df to avoid confusion with other datasets in the R ecosystem. This naming convention helps distinguish this dataset as part of the NeuroDataSets package and assists users in identifying its specific characteristics. The suffix df indicates that the dataset is a data frame. The original content has not been modified in any way.

Source

Data taken from the rcollectadhd package version 0.8

view_datasets_NeuroDataSets *View Available Datasets in NeuroDataSets*

Description

This function lists all datasets available in the 'NeuroDataSets' package. If the 'NeuroDataSets' package is not loaded, it stops and shows an error message. If no datasets are available, it returns a message and an empty vector.

Usage

```
view_datasets_NeuroDataSets()
```

Value

A character vector with the names of the available datasets. If no datasets are found, it returns an empty character vector.

Examples

```
if (requireNamespace("NeuroDataSets", quietly = TRUE)) {
  library(NeuroDataSets)
  view_datasets_NeuroDataSets()
}
```

WMpatterns_tbl_df *White Matter Patterns*

Description

This dataset, WMpatterns_tbl_df, is a tibble containing expected patterns of white matter in schizophrenia derived from large-scale meta-analyses by the ENIGMA consortium. It includes data from multiple neurological and psychiatric conditions for comparison.

Usage

```
data(WMpatterns_tbl_df)
```

Format

A tibble with 24 observations and 15 variables:

WM Character vector indicating white matter regions

SSD Numeric vector of expected patterns for schizophrenia spectrum disorder

MDD Numeric vector of expected patterns for major depressive disorder

AD_ADNI Numeric vector of expected patterns for Alzheimer's disease (ADNI cohort)

AD_ADNIOSYRIX Numeric vector of expected patterns for Alzheimer's disease (ADNI+OSYRIX cohort)

BD Numeric vector of expected patterns for bipolar disorder

Diabetes Numeric vector of expected patterns for diabetes

HighBP Numeric vector of expected patterns for high blood pressure

HighLipids Numeric vector of expected patterns for high lipids

MET Numeric vector of expected patterns for metabolic syndrome

DS_22q Numeric vector of expected patterns for 22q11.2 deletion syndrome

PTSD Numeric vector of expected patterns for post-traumatic stress disorder

TBI Numeric vector of expected patterns for traumatic brain injury

OCD_pediatric Numeric vector of expected patterns for pediatric OCD

OCD_adult Numeric vector of expected patterns for adult OCD

Details

The dataset name has been changed from 'white_matter_patterns_tbl_df' to 'WMpatterns_tbl_df' to follow the shorter naming convention adopted for the NeuroDataSets package while maintaining clarity. The suffix 'tbl_df' indicates that the dataset is a tibble. The original content has not been modified in any way.

Source

Data taken from the RVIpkg package version 0.3.2

Index