Package 'NeuroDataSets'

May 27, 2025

Type Package

Title A Comprehensive Collection of Neuroscience and Brain-Related Datasets

Version 0.1.0

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, epilepsy, schizophrenia, gliomas, and mental health.
Datasets cover structural and functional brain data, neurotransmission, gene expression, cognitive performance, and treatment outcomes.
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]

Repository CRAN

Date/Publication 2025-05-27 08:40:06 UTC

Contents

aba_phenotype_data_df . . . 3
ability_intelligence_list . . . 3
adolescent_mental_health_df . . . 4
alzheimers_biomarkers_tbl_df . . . 5
alzheimer_smoking_df . . . 6
bilingual_brains_df . . . 7
blood_brain_barrier_df . . . 7
brains_cognitive_matrix . . . 8
brain_litter_mammals_df . . . 9
brain_size_iq_df . . . 10
brain_string_players_df . . . 11
cocaine_dopamine_df . . . 12
dopamine_schizophrenia_tbl_df . . . 12
epilepsy_drug_qol_df . . . 13
epilepsy_drug_trial_df . . . 14
epilepsy_RCT_tbl_df . . . 15
gm_expected_patterns_tbl_df . . . 16
guineapig_neurotransmission_df . . . 17
hippocampus_lesions_df . . . 18
mammals_brain_body_df . . . 18
markers_brain_df . . . 19
markers_human_brain_df . . . 20
markers_mouse_brain_df . . . 21
migraine_treatment_df . . . 21
neanderthal_brains_df . . . 22
neurocognitive_psychiatric_df . . . 23
NeuroDataSets . . . 24
neuro_pointprocess_matrix . . . 24
oasis_dementia_mri_df . . . 25
parkinsons_dopamine_list . . . 26
pediatric_glioma_tbl_df . . . 27
sleep_performance_df . . . 28
subcortical_patterns_tbl_df . . . 29
view_datasets_neuro . . . 30
white_matter_patterns_tbl_df . . . 31

Index **32**

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

---

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).

`alzheimers_biomarkers_tbl_df`

Alzheimer's Disease Biomarkers Study

Description

This dataset, alzheimers_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(alzheimers_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 'alzheimers_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.

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.

`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.

`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.

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

dopamine_schizophrenia_tbl_df *Dopamine β-Hydroxylase Activity in Schizophrenia*

Description

This dataset, 'dopamine_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(dopamine_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 kept as `dopamine_schizophrenia_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 **BSDA** package version 1.2.2

`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_neurotransmission_df`
Neurotransmission in Guinea Pig Brains

Description

This dataset, guineapig_neurotransmission_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_neurotransmission_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 kept as 'guineapig_neurotransmission_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 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

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.

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.

`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

---

neurocognitive_psychiatric_df

Neurocognitive Performance in Psychosis

Description

This dataset, neurocognitive_psychiatric_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(neurocognitive_psychiatric_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 kept as 'neurocognitive_psychiatric_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 research: Hartman, L.I. (2016) Schizophrenia and Schizoaffective Disorder: One Condition or Two? Unpublished PhD dissertation, York University.

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

Description

This package provides a wide variety of datasets focused on the brain, nervous system, and related disorders including Parkinson's disease, Alzheimer's, epilepsy, schizophrenia, gliomas, 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

`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.

`oasis_dementia_mri_df` *OASIS Aging-Dementia Longitudinal MRI*

Description

This dataset, oasis_dementia_mri_df, is a data frame containing longitudinal neuroimaging and clinical data from 150 older adults (60-96 years) with repeated MRI scans over multiple visits. The study includes both nondemented and demented individuals, with 373 total imaging sessions featuring 3-4 T1-weighted scans per session.

Usage

```
data(oasis_dementia_mri_df)
```

Format

A data frame with 373 observations and 15 variables:

Subject.ID Character: Unique subject identifier

MRI.ID Character: Unique MRI session identifier

Group Factor with 3 levels: Diagnostic group classification

Visit Integer: Visit number

MR.Delay Integer: Days since first visit

Gender Character: Subject gender

Hand Character: Handedness

Age Integer: Subject age in years

EDUC Integer: Years of education

SES Integer: Socioeconomic status

MMSE Integer: Mini-Mental State Examination score (0-30)

CDR Numeric: Clinical Dementia Rating (0-3)

eTIV Integer: Estimated total intracranial volume (mm^3)

nWBV Numeric: Normalized whole brain volume

ASF Numeric: Atlas scaling factor

Details

The dataset name has been kept as 'oasis_dementia_mri_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 jointest package version 1.0. Original study: Marcus, D.S. et al. (2007) Open Access Series of Imaging Studies (OASIS): Cross-sectional MRI Data in Young, Middle Aged, Nondemented and Demented Older Adults. *Journal of Cognitive Neuroscience*, 19(9), 1498-1507.

`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

---

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

`view_datasets_neuro` *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_neuro()
```

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_neuro()
}
```

white_matter_patterns_tbl_df
Expected Patterns of White Matter

Description

This dataset, white_matter_patterns_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(white_matter_patterns_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 kept as 'white_matter_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

Index

aba_phenotype_data_df, 3
ability_intelligence_list, 3
adolescent_mental_health_df, 4
alzheimer_smoking_df, 6
alzheimers_biomarkers_tbl_df, 5

bilingual_brains_df, 7
blood_brain_barrier_df, 7
brain_litter_mammals_df, 9
brain_size_iq_df, 10
brain_string_players_df, 11
brains_cognitive_matrix, 8

cocaine_dopamine_df, 12

dopamine_schizophrenia_tbl_df, 12

epilepsy_drug_qol_df, 13
epilepsy_drug_trial_df, 14
epilepsy_RCT_tbl_df, 15

gm_expected_patterns_tbl_df, 16
guineapig_neurotransmission_df, 17

hippocampus_lesions_df, 18

mammals_brain_body_df, 18
markers_brain_df, 19
markers_human_brain_df, 20
markers_mouse_brain_df, 21
migraine_treatment_df, 21

neanderthal_brains_df, 22
neuro_pointprocess_matrix, 24
neurocognitive_psychiatric_df, 23
NeuroDataSets, 24
NeuroDataSets-package *(*NeuroDataSets*)*, 24

oasis_dementia_mri_df, 25

parkinsons_dopamine_list, 26
pediatric_glioma_tbl_df, 27

sleep_performance_df, 28
subcortical_patterns_tbl_df, 29

view_datasets_neuro, 30

white_matter_patterns_tbl_df, 31