

Package ‘pTITAN2’

October 14, 2022

Title Permutations of Treatment Labels and TITAN2 Analysis

Version 1.0.2

Description Permute treatment labels for taxa and environmental gradients to generate an empirical distribution of change points. This is an extension for the 'TITAN2' package <<https://cran.r-project.org/package=TITAN2>>.

License MIT + file LICENSE

Encoding UTF-8

Language en-US

LazyData true

URL <https://github.com/USEPA/pTITAN2>

RoxygenNote 7.1.2

Depends R (>= 3.5.0)

Imports data.table

Suggests bookdown, DiagrammeR, dplyr (>= 1.0.0), knitr, magrittr, readr, rmarkdown, TITAN2, testthat (>= 3.0.0), tidy

VignetteBuilder knitr

Config/testthat/edition 3

NeedsCompilation no

Author Peter DeWitt [aut, cre] (<<https://orcid.org/0000-0002-6391-0795>>),
Stephanie Figary [aut] (<<https://orcid.org/0000-0002-4876-9850>>),
Naomi Detenbeck [aut] (<<https://orcid.org/0000-0002-9868-0475>>)

Maintainer Peter DeWitt <peter.dewitt@cuanschutz.edu>

Repository CRAN

Date/Publication 2022-03-07 17:00:02 UTC

R topics documented:

datasets	2
occurrences	3

permutation_example	4
permute	5
pTITAN2	6
Index	7

datasets	<i>Example Data Sets</i>
----------	--------------------------

Description

Four Example data sets for use the pTITAN2 package.

Usage

C_IC_D_06_wID

C_IC_N_06_wID

CD_06_Mall_wID

CN_06_Mall_wID

Format

An object of class data.frame with 251 rows and 2 columns.

An object of class data.frame with 124 rows and 2 columns.

An object of class data.frame with 251 rows and 501 columns.

An object of class data.frame with 124 rows and 501 columns.

Details

Example data sets are from publicly available macroinvertebrate survey data from California. The data sets are broken down between the environmental variable, in this case percent impervious cover, and macroinvertebrate data. Separate data files are provided for each 'treatment' that is explored. In this case, the treatments are data from either drought (dry) or normal precipitation years in the Chaparral region of California.

CN_06_Mall_wID (Chaparral Region, Treatment = Normal) file contains raw macroinvertebrate density data for 500 possible macroinvertebrate codes for each taxonomic level (class, order, family, genus).

The raw data files are provided for your use as well. See example below for accessing these files.

Examples

```

head(C_IC_D_06_wID) # Environemntal Gradient, Dry Treatment
head(C_IC_N_06_wID) # Environemntal Gradient, Normal Treatment
head(CD_06_Mall_wID) # Taxonomic, Dry Treatment
head(CN_06_Mall_wID) # Taxonomic, Normal Treatment

# Get the paths to the raw data files
list.files(system.file("extdata", package = "pTITAN2"))

```

occurrences	<i>Occurrence</i>
-------------	-------------------

Description

Generate a taxonomic data set with codes which have at least n occurrences.

Usage

```
occurrences(data, n = 5L)
```

Arguments

data	A data.frame wit
n	the minimum number of occurrences.

Details

Of the codes with at least n occurrences, the code with the most taxonomic detail needs to be selected for the `titan` run. This means if the macroinvertebrate count has at least n occurrences in a genus code, the family, order, and class codes associated with these counts should be removed. Or, for another example, if there are too few counts at the genus level, but at least n counts at the family level- the family code would be retained and the order and class codes would be removed.

Coding: All of the macroinvertebrates were coded so the first two letters indicate the class, the second two letters indicate the order, the third two letters indicate the family, and the last two numbers indicate the genus. "00" indicates that there is no information at that level. For example: A code that is 'Bi000000' is the Bivalvia class, while BiVe0000 is the Bivalvia class, Veneroida order. BiVeSh00 is the Bivalvia class, Veneroida order, Spheriridae family. BiVeSh01 is a genus within that family.

NOTE: The example script that inspired the development of this function required the data set to have the column names in alpha order. This function relaxes that requirement by using the `arrange` call.

Value

a data.frame with six columns: taxon, Class, Order, Family, Genus, and count.

See Also

```
vignette(topic = "pTITAN2", package = "pTITAN2")
```

Examples

```
# Report the tax with minimum of five (default) occurrences.  
occurrences(CN_06_Mall_wID[, -1])
```

```
# Report the tax with at least six occurrences  
occurrences(CN_06_Mall_wID[, -1], n = 6)
```

permutation_example *Permutation Example Result*

Description

Results for a permutation example.

Usage

```
permutation_example
```

Format

An object of class `data.frame` with 10 rows and 7 columns.

Details

The code to generate this example permutation set can be found via `system.file("example-scripts", "permutation_example.R", package = "pTITAN2")`

See Also

```
vignette("pTITIAN2")
```

permute	<i>Permute</i>
---------	----------------

Description

Permute treatment labels for a taxa and associated environmental gradients.

Usage

```
permute(taxa, envs, sid)

permute2(..., minTaxonFreq = 3L, trys = 100L)
```

Arguments

taxa	a list of data.frames with the taxa. See Details.
envs	a list of data.frames with the environmental gradients. See Details
sid	a character vector of length one with the name of the column identifying the station id.
...	passed to permute
minTaxonFreq	min number of occurrences for each taxon
trys	maximum number of attempts to generate a meaningful permutation

Details

The taxa and envs lists are expected to be of equal length and that the *i*th element of taxa list is associated with the *i*th element of the envs list. That is, the taxa and environmental gradient for treatment 1 are both the first elements of the respective lists, the taxa and environmental gradient for treatment 2 are the second elements for the respective lists, etc.

The environmental gradient data.frames are expected to have two columns, one with the station ID and one with the data defining the gradient.

The taxa data.frames are expected to have the station ID column as well. **Important** The station ID column name needs to be the same for all the taxa and environmental gradient data.frames.

Value

A list of lists of lists. At the top level the elements are the treatment groups. There are as many elements as the length of the lists taxa and envs.

The second level are the taxa and environmental gradient.

See Also

```
vignette(topic = "pTITAN2", package = "pTITAN2")
```

Examples

```
example_permutation <-  
  permute(taxa = list(CD_06_Mall_wID, CN_06_Mall_wID),  
         envs = list(C_IC_D_06_wID, C_IC_N_06_wID),  
         sid = "StationID")  
str(example_permutation, max.level = 2)
```

pTITAN2

Permutations of Treatment Labels and TITAN2 Analysis

Description

Permute treatment labels for taxa and environmental gradients to generate an empirical distribution of change points.

Index

* datasets

datasets, [2](#)

permutation_example, [4](#)

arrange, [3](#)

C_IC_D_06_wID (datasets), [2](#)

C_IC_N_06_wID (datasets), [2](#)

CD_06_Mall_wID (datasets), [2](#)

CN_06_Mall_wID (datasets), [2](#)

datasets, [2](#)

occurrences, [3](#)

permutation_example, [4](#)

permute, [5](#)

permute2 (permute), [5](#)

pTITAN2, [6](#)

titan, [3](#)