select_n_taxa.R

June 19, 2025

Contents

select_n_taxa	*Select N Taxa*

Description

Selects the top N taxa based on their occurrence across datasets.

Usage

```
select_n_taxa(data = NULL, n_taxa = Inf, per = "dataset_name")
```

Arguments

data	A data frame containing the input data. Must include columns "taxon" and the column specified in the 'per' parameter.
n_taxa	A numeric value specifying the number of taxa to select. Default is Inf.
per	A character string specifying the column name to group by. Default is "dataset_name".

Details

The function identifies the most common taxa across datasets by counting their occurrences. It then filters the input data to include only the selected taxa. If no taxa are found, an error is raised.

Value

A data frame containing the filtered data with the top N taxa.

Index