Package ‘siie’

December 1, 2022

Type Package
Title Superior Identification Index and Its Extensions
Version 0.3.0
Depends R (>= 4.0.0)
Description Calculate superior identification index and its extensions. Measure the performance of journals based on how well they could identify the top papers by any index (e.g. citation indices) according to Huang & Yang. (2022) <doi:10.1007/s11192-022-04372-z>. These methods could be extended to evaluate other entities such as institutes, countries, etc.
License MIT + file LICENSE
Encoding UTF-8
Imports data.table
Suggests ggrepplot2, knitr, rmarkdown, testthat (>= 3.0.0), tidyfst
RoxygenNote 7.2.2
VignetteBuilder knitr
Config/testthat/edition 3
NeedsCompilation no
Author Tian-Yuan Huang [aut, cre] (<https://orcid.org/0000-0002-3591-4203>)
Maintainer Tian-Yuan Huang <huang.tian-yuan@qq.com>
Repository CRAN
Date/Publication 2022-12-01 04:00:02 UTC

\textbf{R topics documented:}

<table>
<thead>
<tr>
<th>Topic</th>
<th>Pages</th>
</tr>
</thead>
<tbody>
<tr>
<td>prp</td>
<td>2</td>
</tr>
<tr>
<td>p_sie</td>
<td>3</td>
</tr>
<tr>
<td>siie</td>
<td>4</td>
</tr>
</tbody>
</table>

Index 6
Calculation of Paper Rank Percentile (PRP)

Description

Paper rank percentile refers to the journals’ average ranking of papers within the field. If Journal A has a PRP of 90, it means the papers in Journal A has an average ranking of 90 out of 100. This metric could be extended to measure other entities such as institutes and countries.

Usage

prp(df, group, index)

Arguments

df A data.frame containing at least two columns (namely the group and the index of each individual).
group The group variable. In the context of our paper, this could be the name or ISSN of a journal.
index The indicator of individuals. In the context of our paper, this could be citation index of papers.

Value

A data.table with 3 columns, with the group, total number in the group (total_no) and PRP.

References


Examples

```r
set.seed(19960822)
nr_of_rows = 1e4
data.frame(
  Id = 1:1e4,
  Journal = sample(LETTERS,nr_of_rows,replace = TRUE),
  CiteCount = sample(1:100,nr_of_rows,replace = TRUE)
) -> journal_table

prp(journal_table,group = "Journal",index = "CiteCount")
```
p_sie

Get data for p-SIE curve

Description

Calculate data for p-SIE curve.

Usage

```
p_sie(df, group, index, p = 1:100, to_compare)
```

Arguments

- `df`: A data.frame containing at least two columns (namely the group and the index of each individual).
- `group`: The group variable. In the context of our paper, this could be the name or ISSN of a journal.
- `index`: The indicator of individuals. In the context of our paper, this could be citation index of papers.
- `p`: A group of p (Cutoff of superior). Defaults to integers from 1 to 100.
- `to_compare`: Which groups to compare with.

Value

A data.table with 3 columns, with the group, p and the according SIE.

References


Examples

```
library(ggplot2)
library(tidyfst)

set.seed(19960822)
nr_of_rows = 1e4
data.frame(
  Id = 1:1e4,
  Journal = sample(LETTERS,nr_of_rows,replace = TRUE),
  CiteCount = sample(1:100,nr_of_rows,replace = TRUE)
) -> journal_table

p_sie(journal_table,group = "Journal",
      index = "CiteCount",to_compare = c("A","B","C")) -> p_sie_df
```
Calculation of SII and SIE

Description

Calculate SII (Superior Identification Index) and SIE (Superior Identification Efficiency) for each group with individual values. In the context of the paper, we have citation counts of papers from different journals. This function could calculate SII and SIE for each journal within the field.

Usage

siie(df, group, index, p = 10)

Arguments

df A data.frame containing at least two columns (namely the group and the index of each individual).
group The group variable. In the context of our paper, this could be the name or ISSN of a journal.
index The indicator of individuals. In the context of our paper, this could be citation index of papers.
p Cutoff of superior. Defaults to 10, meaning top 10 percent individuals are regarded as superior.

Details

In the context, SII indicates how well a journal could identify the top papers (superior research), whereas SIE quantifies the efficiency of a journal to identify the top papers.

Value

A data.table with 5 columns, with the group, superior number (superior_no), total number in the group (total_no), SII (sii) and SIE (sie).
siie

References

Examples

```r
set.seed(19960822)
nr_of_rows = 1e4
data.frame(
  Id = 1:1e4,
  Journal = sample(LETTERS,nr_of_rows,replace = TRUE),
  CiteCount = sample(1:100,nr_of_rows,replace = TRUE)
) -> journal_table
siie(journal_table,group = "Journal",index = "CiteCount")
```
Index

p_sie, 3
prp, 2
siie, 4