Package 'psmineR'

May 9, 2026

Type Package

Title Performance Spectrum Miner for Event Data

Version 0.1.1

Description Compute detailed and aggregated performance spectrum for event data. The detailed performance spectrum describes the event data in terms of segments, where the performance of each segment is measured and plotted for any occurrences of this segment over time and can be classified, e.g., regarding the overall population. The aggregated performance spectrum visualises the amount of cases of particular performance over time. Denisov, V., Fahland, D., & van der Aalst, W. M. P. (2018) <doi:10.1007/978-3-319-98648-7_9>.

Depends R (>= 3.5.0)

License MIT + file LICENSE

Encoding UTF-8

Imports bupaR (>= 0.5.1), dplyr, data.table, forcats, ggplot2, tidyr, rlang (>= 1.0.0), cli (>= 3.2.0), glue, stringi

RoxygenNote 7.3.3

URL https://bupar.net/, https://github.com/bupaverse/psmineR/, https://bupaverse.github.io/psmineR/

Suggests knitr, eventdataR, rmarkdown, covr, testthat (>= 3.1.3), edeaR

BugReports https://github.com/bupaverse/psmineR/issues

Config/testthat/edition 3

VignetteBuilder knitr

NeedsCompilation no

Author Greg Van Houdt [aut], Gert Janssenswillen [ctb, cre], Gerard van Hulzen [ctb]

Maintainer Gert Janssenswillen <gert.janssenswillen@uhasselt.be>

Repository CRAN

Date/Publication 2025-09-08 12:40:02 UTC

Contents

`plot` *Plot Methods*

Description

Visualize performance spectrum.

Usage

```
## S3 method for class 'ps_aggregated'
plot(x, ...)

## S3 method for class 'ps_detailed'
plot(x, ...)
```

Arguments

`x`	Object of class `ps_aggregated()` or `ps_detailed()`.
`...`	Additional variables to pass further.

Value

A `ggplot2` object, which can be customised further.

`psmineR` *psmineR*

Description

Performance Spectrum Miner For Event Data

ps_aggregated *Aggregated Performance Spectrum*

Description

Plots the aggregated performance spectrum. The performance spectrum describes the event data in terms of segments, i.e., pairs of related process steps. The performance of each segment is measured and plotted for any occurrences of this segment over time and can be classified, e.g., regarding the overall population. The aggregated performance spectrum visualises the amount of cases of particular performance over time (Denisov *et al.*, 2018). See **References** for more details.

Usage

```
ps_aggregated(
  log,
  segment_coverage,
  n_segments,
  classification = NULL,
  grouping = c("start", "complete"),
  scale = NULL,
  bins = 30
)

## S3 method for class 'log'
ps_aggregated(
  log,
  segment_coverage,
  n_segments,
  classification = NULL,
  grouping = c("start", "complete"),
  scale = NULL,
  bins = 30
)

## S3 method for class 'grouped_log'
ps_aggregated(
  log,
  segment_coverage,
  n_segments,
  classification = NULL,
  grouping = c("start", "complete"),
  scale = NULL,
  bins = NULL
)
```

Arguments

log	log: Object of class log or derivatives (grouped_log, eventlog, activitylog, etc.).
segment_coverage, n_segments	numeric: Provide either segment_coverage or n_segments. If neither is provided, segment_coverage = 0.2 will be used. segment_coverage: The percentage of cases (default 0.2) in which each segment must be present to be visualised in the spectrum. Ignored if n_segments is specified. n_segments: Visualise only the top n segments based on frequency.
classification	character (default NULL): The variable defining the colour legend. This variable should be present in log. If NULL (default) when log is a grouped_log, the first grouping variable will be used as classification. If NULL (default) or "quartile" when log is an eventlog or activitylog, a quartile variable dividing the durations of the segments in quartiles is calculated.
grouping	character (default "start"): The timestamps, "start" or "complete", which are binned in the histogram.
scale	ggplot2 scale function (default scale_fill_discrete_bupaR): Set color scale. Defaults to scale_fill_discrete_bupaR.
bins	numeric (default 30): The number of bins in the aggregated performance spectrum.

Value

A ggplot2 object describing the aggregated performance spectrum.

Methods (by class)

- ps_aggregated(log): Plot aggregated performance spectrum for a log.
- ps_aggregated(grouped_log): Plot aggregated performance spectrum for a grouped_log.

References

Denisov, V., Fahland, D., & van der Aalst, W. M. P. (2018). Unbiased, Fine-Grained Description of Processes Performance from Event Data. In M. Weske, M. Montali, I. Weber, & J. vom Brocke (Eds.), Proceedings of the 16th International Conference on Business Process Management (Vol. 11080, pp. 139–157). Springer International Publishing. doi:10.1007/9783319986487_9

See Also

ps_detailed()

Examples

```
library(psmineR)
library(eventdataR)
```

```
sepsis %>%
 ps_aggregated(segment_coverage = 0.2,
               classification = "quartile",
               grouping = "start",
               bins = 15)
```

---

ps_detailed *Detailed Performance Spectrum*

Description

Plots the detailed performance spectrum. The performance spectrum describes the event data in terms of segments, i.e., pairs of related process steps. The performance of each segment is measured and plotted for any occurrences of this segment over time and can be classified, e.g., regarding the overall population. The detailed performance spectrum visualises variability of durations in a segment across cases and time (Denisov *et al.*, 2018). See **References** for more details.

Usage

```
ps_detailed(
  log,
  segment_coverage,
  n_segments,
  classification = NULL,
  scale = NULL
)

## S3 method for class 'log'
ps_detailed(
  log,
  segment_coverage,
  n_segments,
  classification = NULL,
  scale = NULL
)

## S3 method for class 'grouped_log'
ps_detailed(
  log,
  segment_coverage,
  n_segments,
  classification = NULL,
  scale = NULL
)
```

Arguments

log	log: Object of class log or derivatives (grouped_log, eventlog, activitylog, etc.).
segment_coverage, n_segments	numeric: Provide either segment_coverage or n_segments. If neither is provided, segment_coverage = 0.2 will be used. segment_coverage: The percentage of cases (default 0.2) in which each segment must be present to be visualised in the spectrum. Ignored if n_segments is specified. n_segments: Visualise only the top n segments based on frequency.
classification	character (default NULL): The variable defining the colour legend. This variable should be present in log. If NULL (default) when log is a grouped_log, the first grouping variable will be used as classification. If NULL (default) or "quartile" when log is an eventlog or activitylog, a quartile variable dividing the durations of the segments in quartiles is calculated.
scale	ggplot2 scale function (default scale_color_discrete_bupaR): Set color scale. Defaults to scale_color_discrete_bupaR.

Value

A ggplot2 object describing the detailed performance spectrum.

Methods (by class)

- ps_detailed(log): Plot detailed performance spectrum for a log.
- ps_detailed(grouped_log): Plot detailed performance spectrum for a grouped_log.

References

Denisov, V., Fahland, D., & van der Aalst, W. M. P. (2018). Unbiased, Fine-Grained Description of Processes Performance from Event Data. In M. Weske, M. Montali, I. Weber, & J. vom Brocke (Eds.), Proceedings of the 16th International Conference on Business Process Management (Vol. 11080, pp. 139–157). Springer International Publishing. doi:10.1007/9783319986487_9

See Also

ps_aggregated()

Examples

```
library(psmineR)
library(eventdataR)

sepsis %>%
 ps_detailed(segment_coverage = 0.2,
             classification = "quartile")
```

Index