

TheseusPlot: Visualizing Decomposition of Differences in Rate Metrics

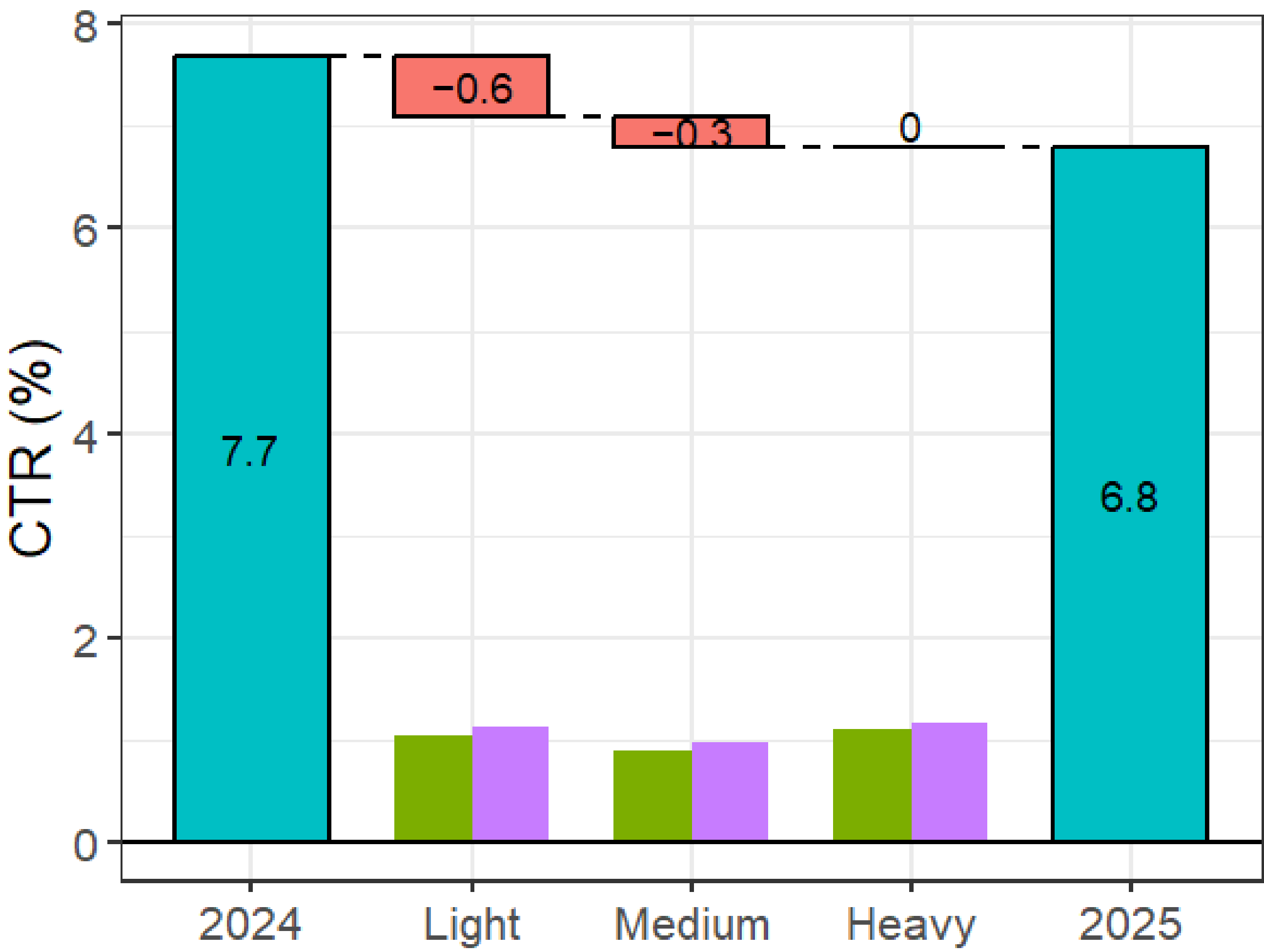
Koji Makiyama (HOXO-M Inc.), Kazuyuki Sano (LY Corp.), Daisuke Ichikawa (Kibaroku), Shinichi Takayanagi

1. Overview

- Rate metrics, such as click-through rate (CTR), conversion rate, and churn rate, often differ between two groups.
- For rate metrics, the overall difference cannot be decomposed by simply adding subgroup rates.
- We propose **TheseusPlot**, a visualization method and an R package for identifying which subgroups contribute to differences in rate metrics.

2. Toy Example

- Suppose users are categorized into Light, Medium, and Heavy user groups.
- CTR decreased from 2024 to 2025.
- **TheseusPlot** replaces one user category at a time and recalculates CTR after each replacement.
- The resulting Theseus Plot is shown below.



- Light users contribute most to the CTR decline.
- Heavy users make little to no contribution.

3. How to Read the Plot

- The left and right endpoints show the metric values for the two groups.
- Each step represents a subgroup's contribution to the overall difference in the rate metric.
- Downward steps indicate negative contributions; upward steps indicate positive contributions.
- The bottom bars show subgroup sample sizes, scaled for display, for each group.
- Large steps may partly reflect subgroup size, so contributions should be interpreted together with the sample-size bars.

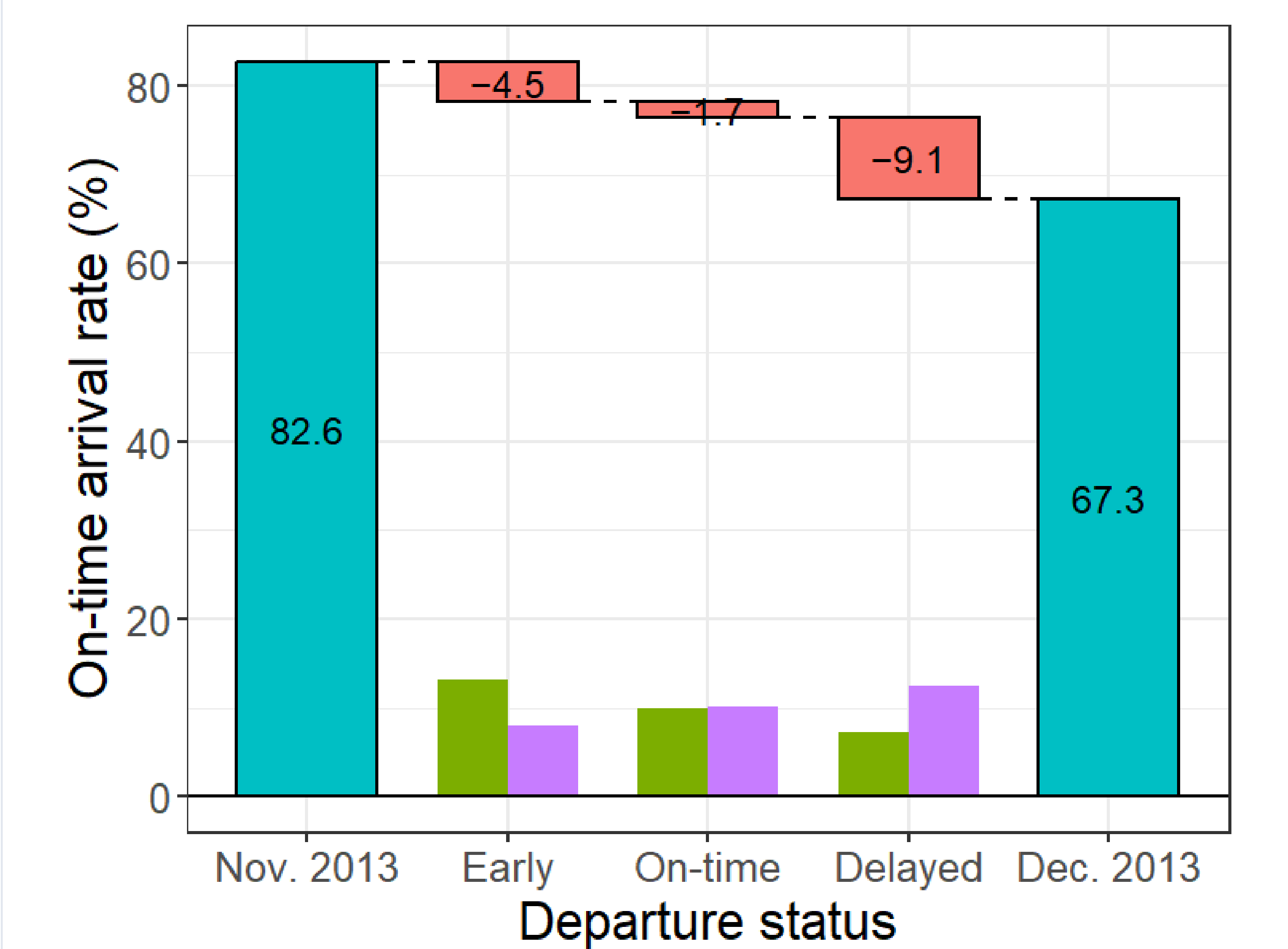
Main Takeaway

TheseusPlot turns rate-metric decomposition problems into interpretable visual summaries.

By showing both subgroup contributions and sample sizes, it helps identify which subgroups contribute most to observed differences in rate metrics.

4. Application: NYC Flights

- Using **nycflights13**, we analyze the drop in on-time arrival rate from November to December 2013.
- **TheseusPlot** decomposes the difference by departure status: Early, On-time, and Delayed.
- Fewer early departures and more delayed departures contributed substantially to the decline.



5. Package

- The **TheseusPlot** package is available on CRAN.
- ```
install.packages("TheseusPlot")
```
- Usage:
- ```
library(TheseusPlot)

ship <- create_ship(data1, data2, y)

ship$plot(subgroup_column)
```
- More features:
 - Contribution tables: `ship$table()`
 - Axis-flipped plots: `ship$plot_flip()`
 - Grouping low-contribution subgroups
 - Binning continuous variables
 - For more details:
<https://hoxo-m.github.io/TheseusPlot/>

