Find the Mario

Joint A TEAM

Takumi SATO 佐藤匠
Atsushi UECHI 上地安諄
Kazuki NAKAMURA 仲村一起
Kentaro BABA 馬場健太郎
Masahiro KATO 加藤真大
Naoto MINAKAWA 皆川直人
Wataru KOMATSUBARA 小松原航

Background・Objective・Data Summary 背景・目的・データ概要

【Background】

・Diversification of work styles
・Diversification of actions
・Changes in Transportation Demand

Clearify planned / unplanned actions

Choice of Transportation by Income

【Objective】

Understanding differences in transportation mode choice based on behavioral tendencies

【Data】

Toyosu PP Data(2021)

Basic Analysis

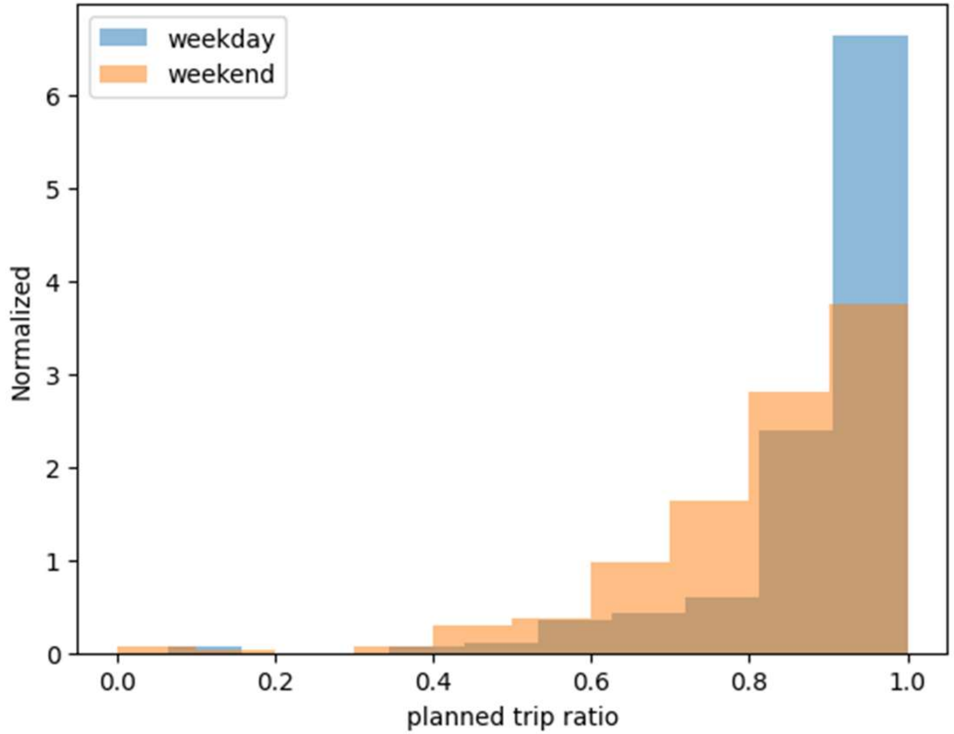
・Do people change their behavior between weekday and weekend ?
・Analyze the proportion of daily activities by categorizing trips into typical and atypical types for weekdays and weekends.
・It is expected that typical trips will be more frequent on weekdays, while atypical trips will be more common on weekends.

※Typical trips include commuting, returning home, and shopping (daily shopping is a major component).
※Atypical trips include leisure, strolling, and other activities.

Basic Analysis

・planned_trip_ratio is defined as typical trip divided by all trip

・The graph shows that the planned_tip_ratio is lower on weekends, which suggests that people tend to do more atypical trips in weekend than in weekday.



Analytical Framework 分析の枠組み

Hawkes Process

The Hawkes process is a *self-exciting point process*. Suitable for modeling burstiness and clustering in event sequences such as departure times.

Users: $i = 1, \ldots, N$, Day type: $d \in \{W, H\}$ (Weekday/ Holiday)
For user i and day type d, observed departure times in an observation window $[0, T_{i,d}]$:

$$\mathcal{T}_{i,d} = \{t_{i,d,1}, \ldots, t_{i,d,n_{i,d}}\}, \quad 0 < t_{i,d,1} < \cdots < t_{i,d,n_{i,d}} < T_{i,d}$$

For each user i, estimate Hawkes parameters separately for Weekday (W) and Holiday (H).

$$\lambda_{i,d}(t) = \mu_{i,d} + \sum_{t_{i,d,k}<t} \alpha_{i,d}\beta_{i,d} \exp(-\beta_{i,d}(t - t_{i,d,k}))$$

$\mu_{i,d}$: baseline intensity
$\alpha_{i,d}$: excitation strength (per-event jump size)
$\beta_{i,d}$: decay rate (how fast self-excitation fades)

Branching ratio $n_{i,d} = \alpha_{i,d}/\beta_{i,d}$ as the key indicator of event chain intensity.
Applying a two-group classification (low- and high-chain) to both weekdays and holidays produces four distinct user clusters.

Analytical Framework 分析の枠組み

Label 0: Mode-Focused Routine (Weekday: Low / Holiday: Low)

- Intermediate in trips and weekend share, no strong differences. Possibly reflects **specific mode dependency** rather than temporal behavior.

Label 1: Stable Majority (Weekday: High / Holiday: Low)

- Lowest trip counts, longest interarrival time. Represents the **baseline habitual travelers**: stable, weekday-centered, moderate in mode use.

Label 2: Weekend Bursty (Weekday: Low / Holiday: High):

- Highest trip counts, shortest interarrival, significantly higher weekend share. Clear evidence of **bursty, weekend-concentrated mobility**.

Label 3: High-Frequency Diverse (Weekday: High / Holiday: High):

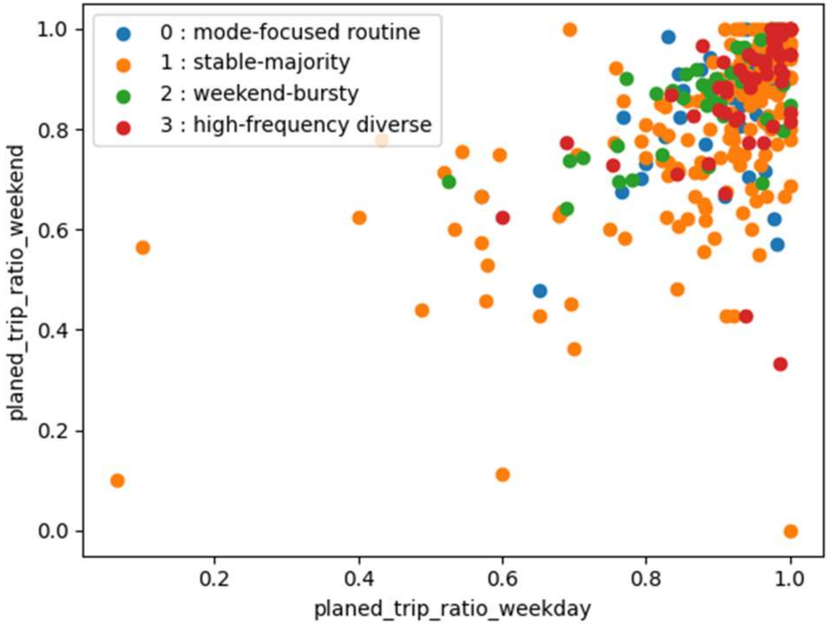
- High trip counts and short interarrival times (like Label 2) but not as weekend-oriented. Represents **active travelers** with more balanced weekday/ weekend spread.

Dimension	Significant Pairwise Differences (Tukey)	Cluster-Level Interpretation
Number of trips	Label 2 > Label 1 (**$p<0.001$**), Label 3 > Label 1 (**$p<0.001$**)	Label 2 & 3 are **high-frequency travelers**, Label 1 is **low-frequency / stable**
Weekend share	Label 2 > Label 0 (**$p\approx0.025$**), Label 2 > Label 1 (**$p<0.001$**), Label 2 > Label 3 (**$p\approx0.049$**)	Label 2 is **weekend-dominant / bursty**, others remain **weekday-centered**
Median interarrival hours	Label 1 > Label 2 (**$p<0.001$**), Label 1 > Label 3 (**$p\approx0.003$**)	Label 1 has **long gaps** → **stable routine**, Labels 2 & 3 are **short gaps** → **bursty or high-frequency**

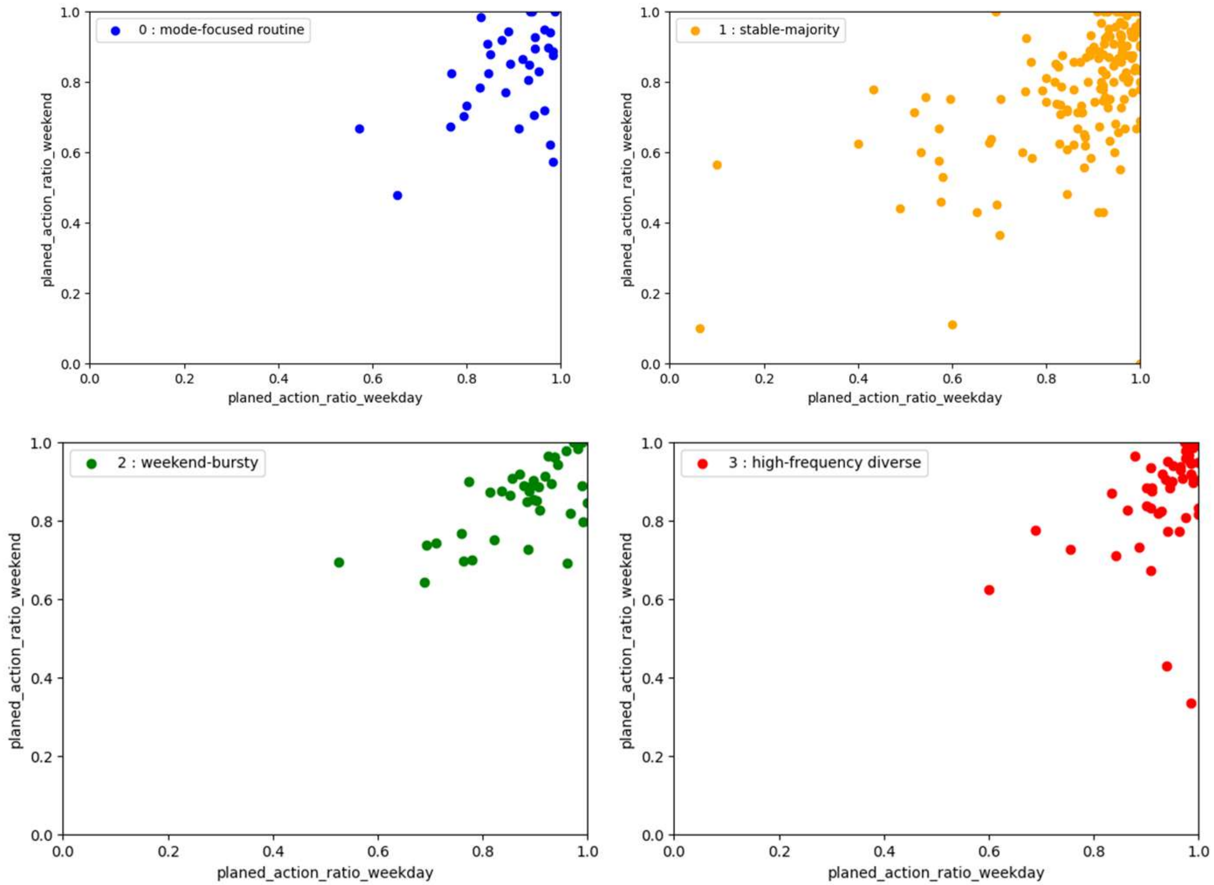
Result 結果

・Plot the weekday planned trip ratio on the horizontal axis and the weekend planned trip ratio on the vertical axis, coloring the four categories of people classified by the hawk model.

・the data points in the upper-right area—those who make routine trips on both weekdays and weekends—consist of a mixture of 0 : mode-focused routine individuals and 3 : high-frequency diverse individuals, contrary to the intuition.

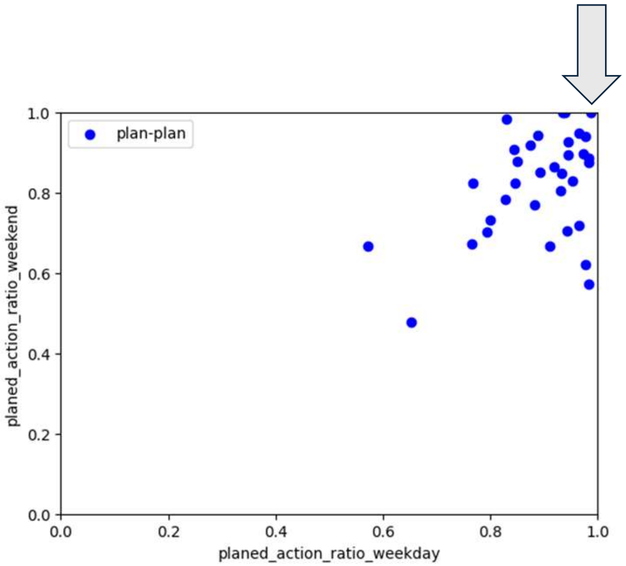


・Rather than interpreting “bursty” as acting without a plan, it may be better understood as people who pack their schedules. For that reason, even those classified as bursty tend to have a high proportion of typical trips.

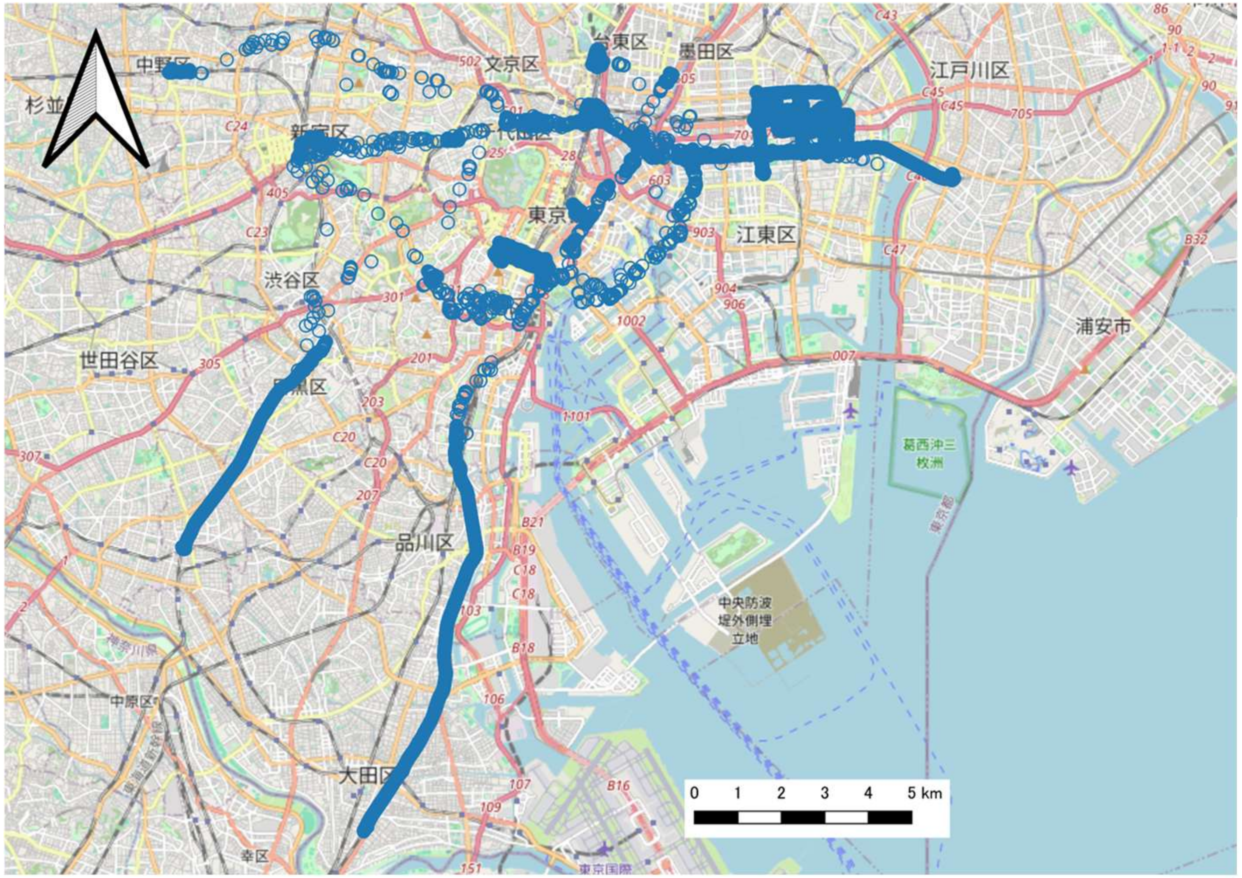


Result 結果

A person in mode-focused routine category seems to trip as planned

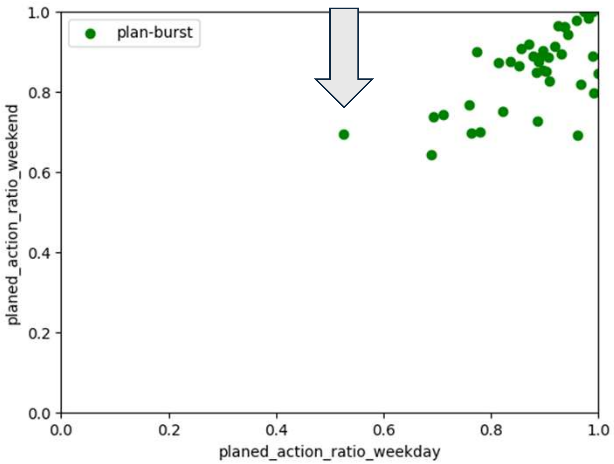


【0 : mode-focused routine】

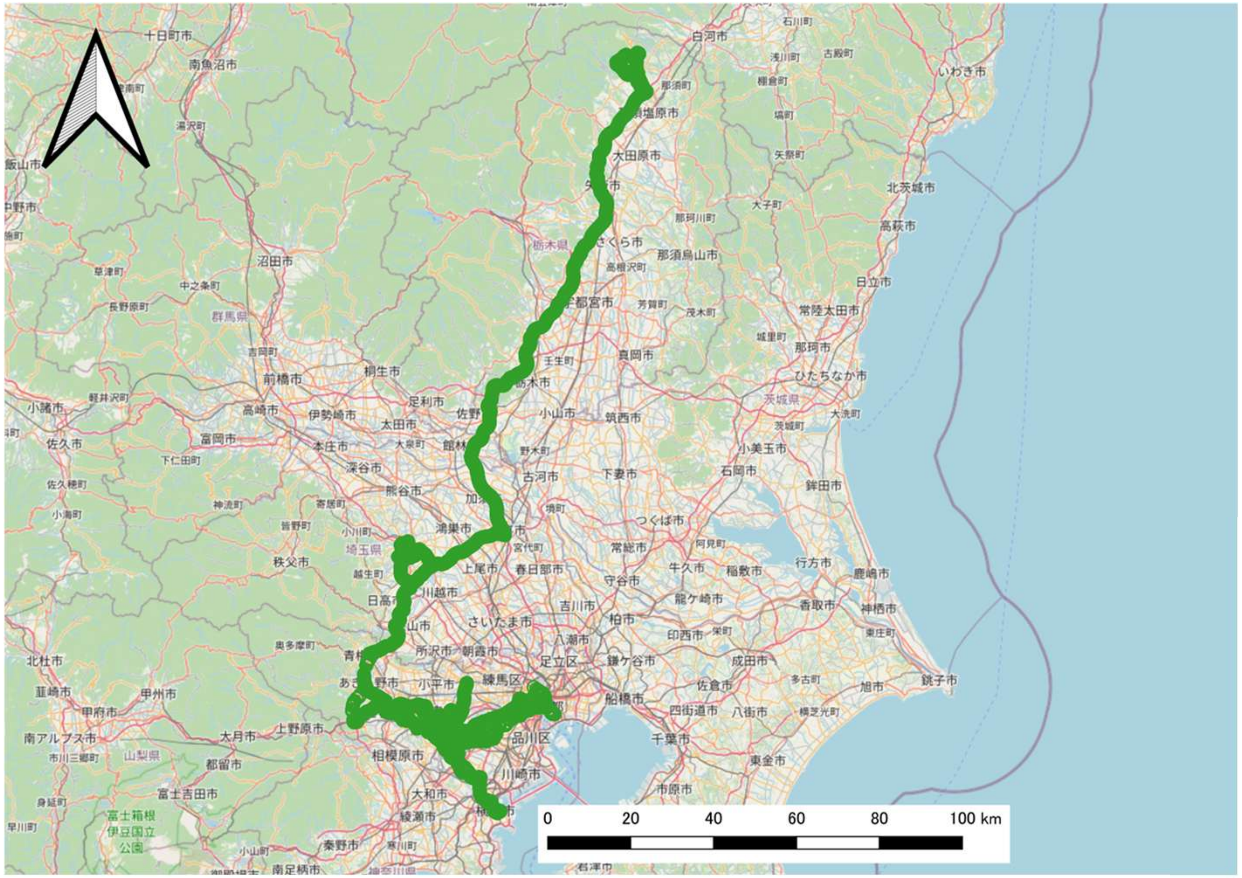


Result 結果

・A person in stable-majority category seems to trip without a plan

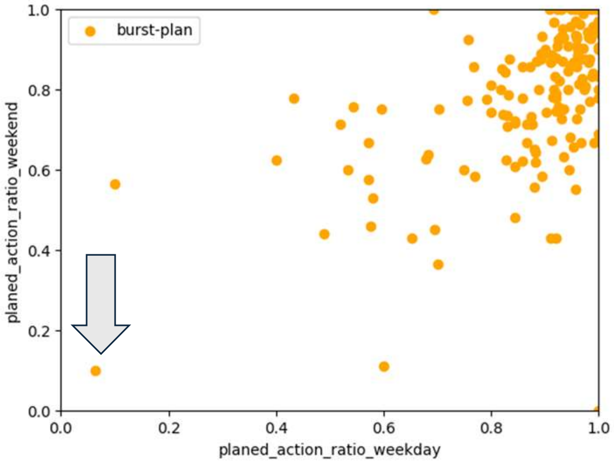


【1 : stable-majority】

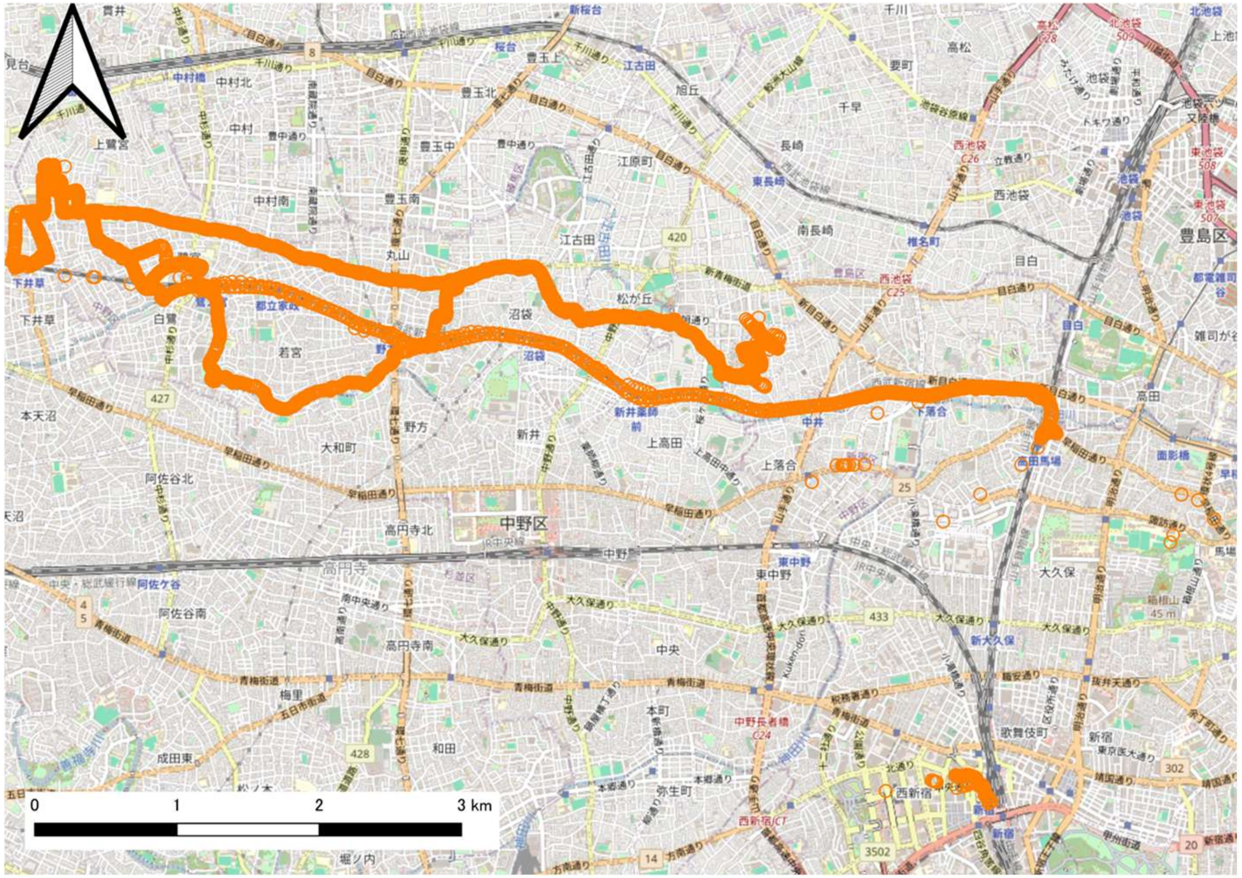


Result 結果

・A person in weekend-bursty category seems to trip without a plan

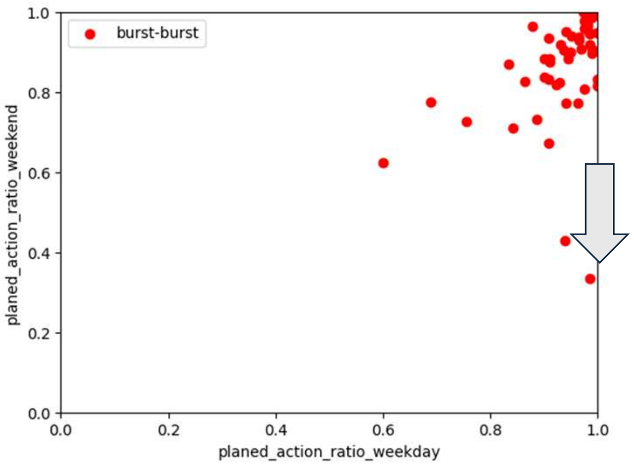


【2 weekend bursty】

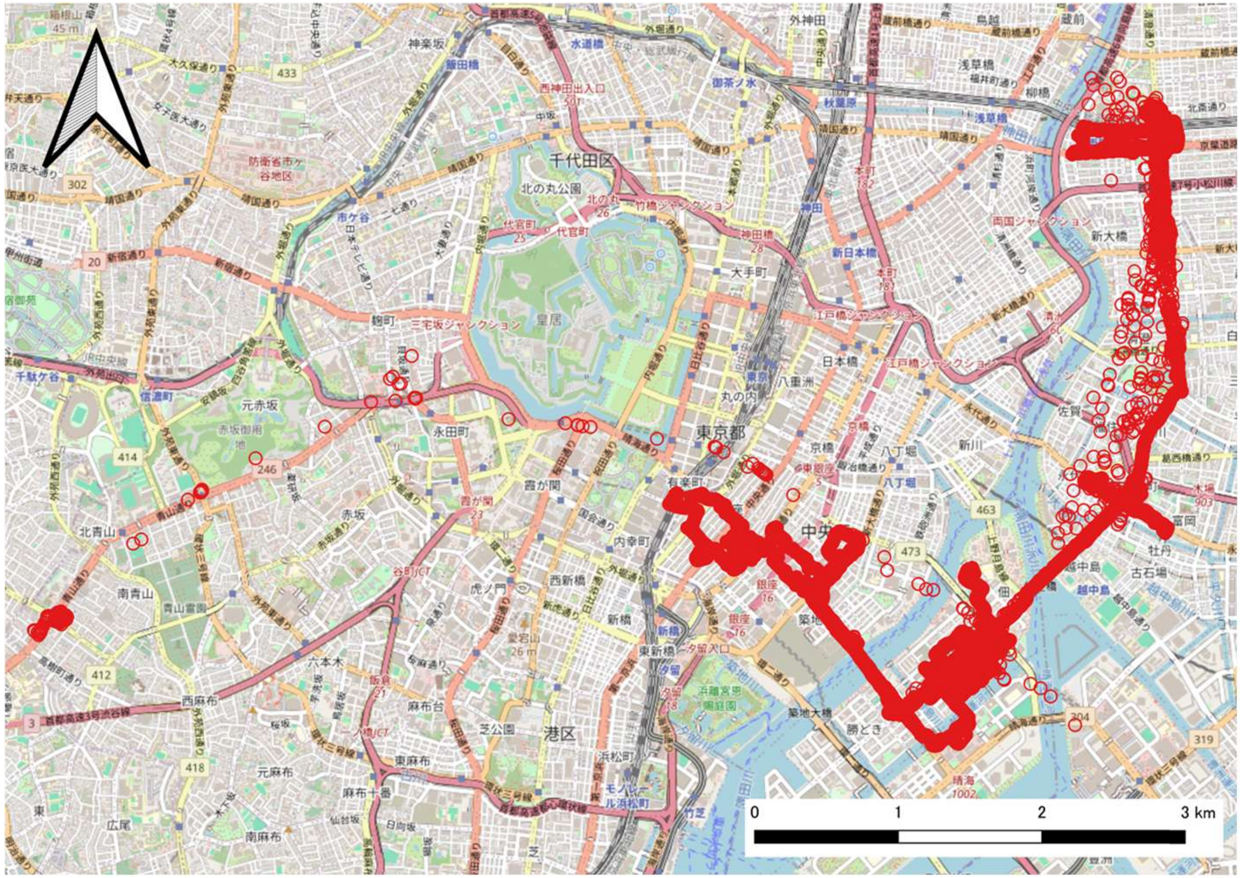


Result 結果

・A person in high frequency diverse category seems to trip without a plan



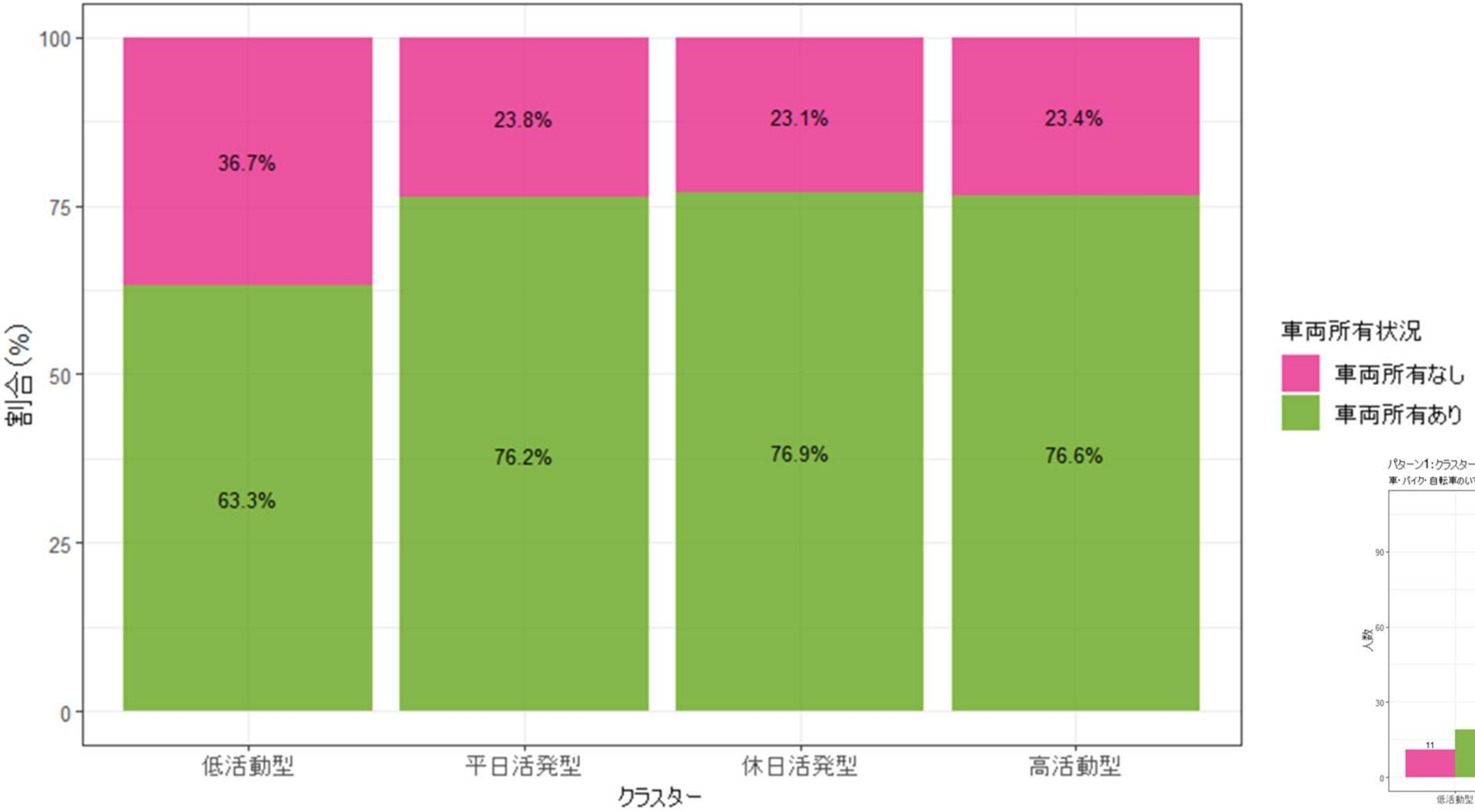
【high frequency diverse】

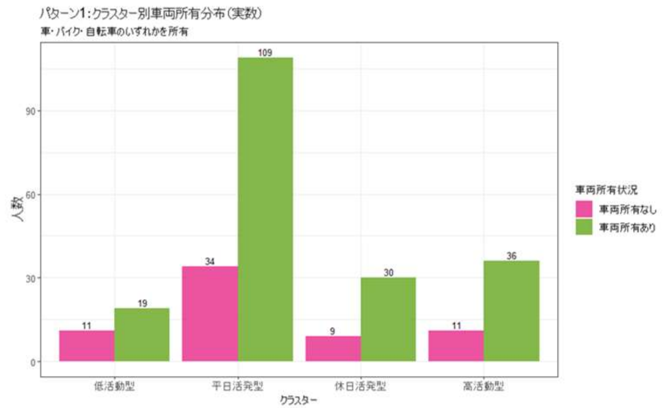


Result

Relationship between Activity Clusters and Vehicle Ownership

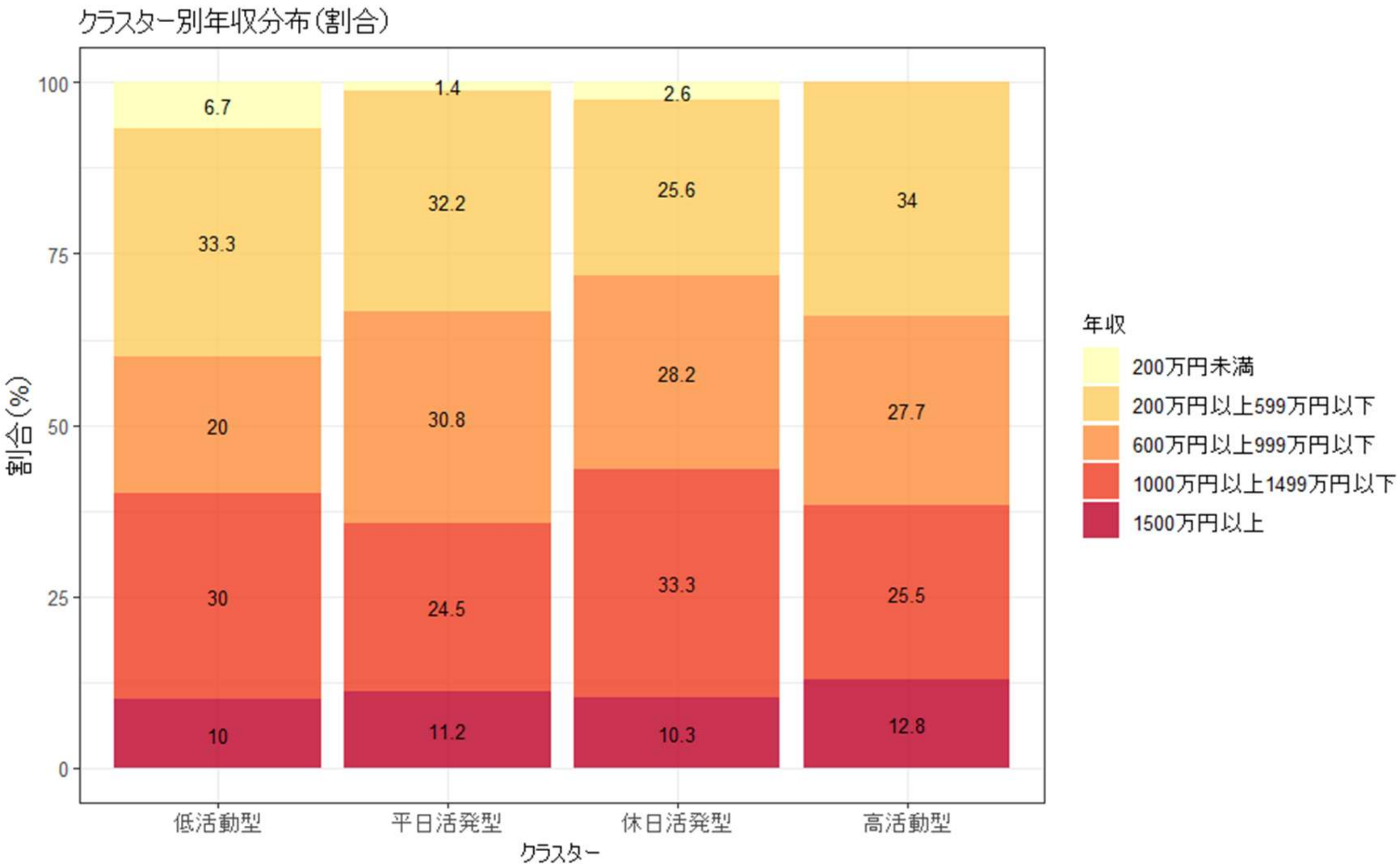
※Ownership of any vehicle type (car, motorcycle, or bicycle)





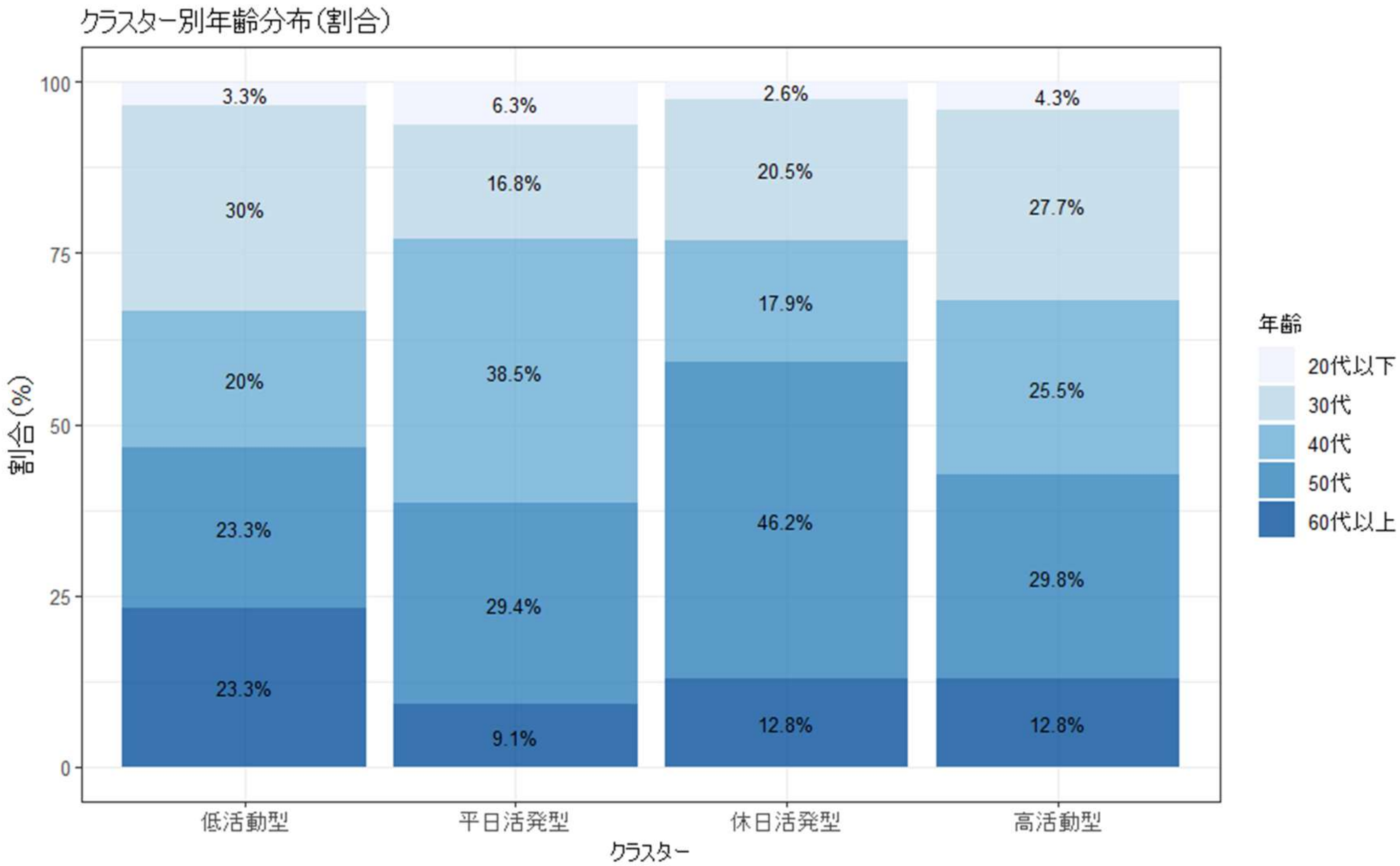
Result

Distribution of Income for people categorized by hawks model



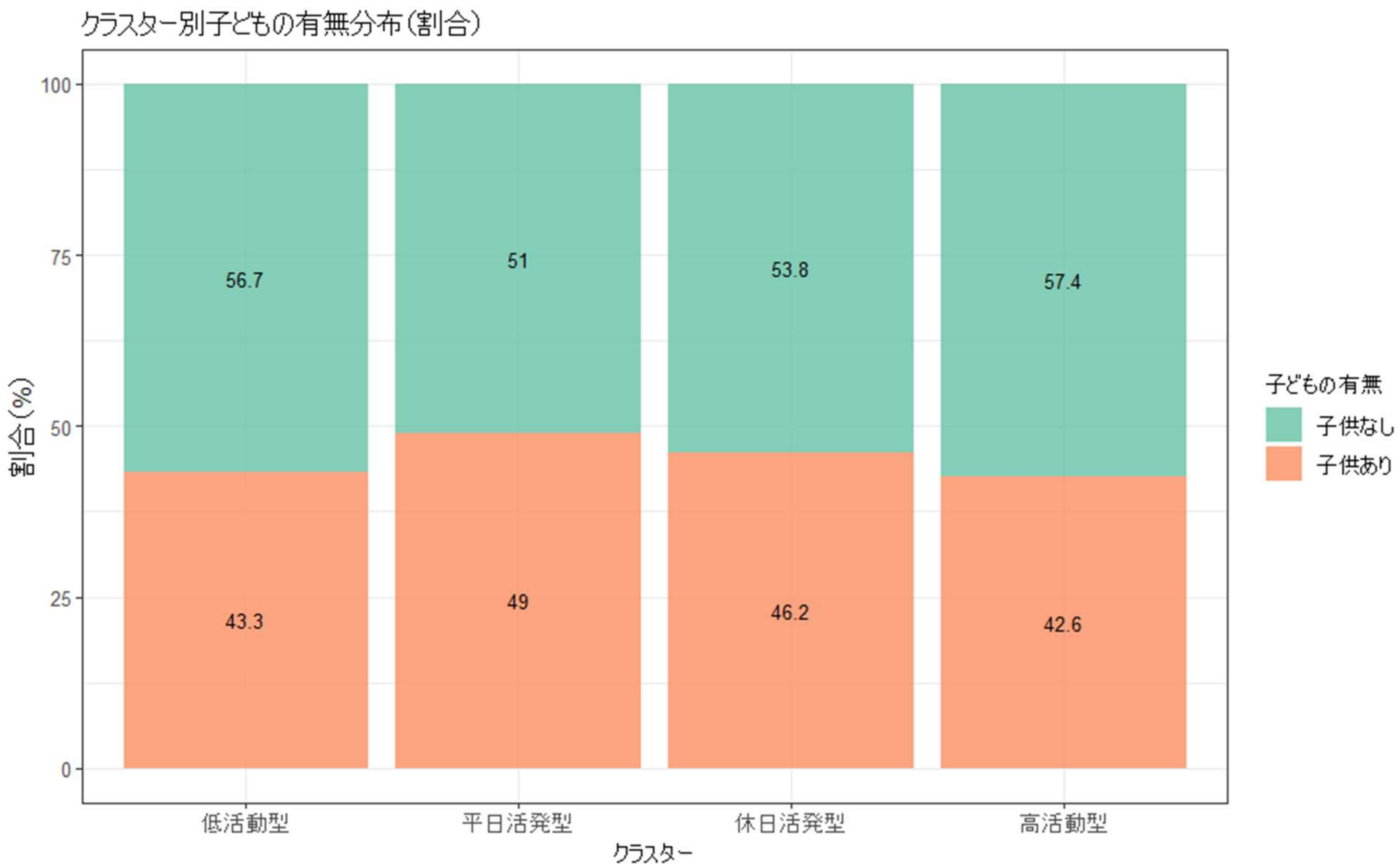
Result

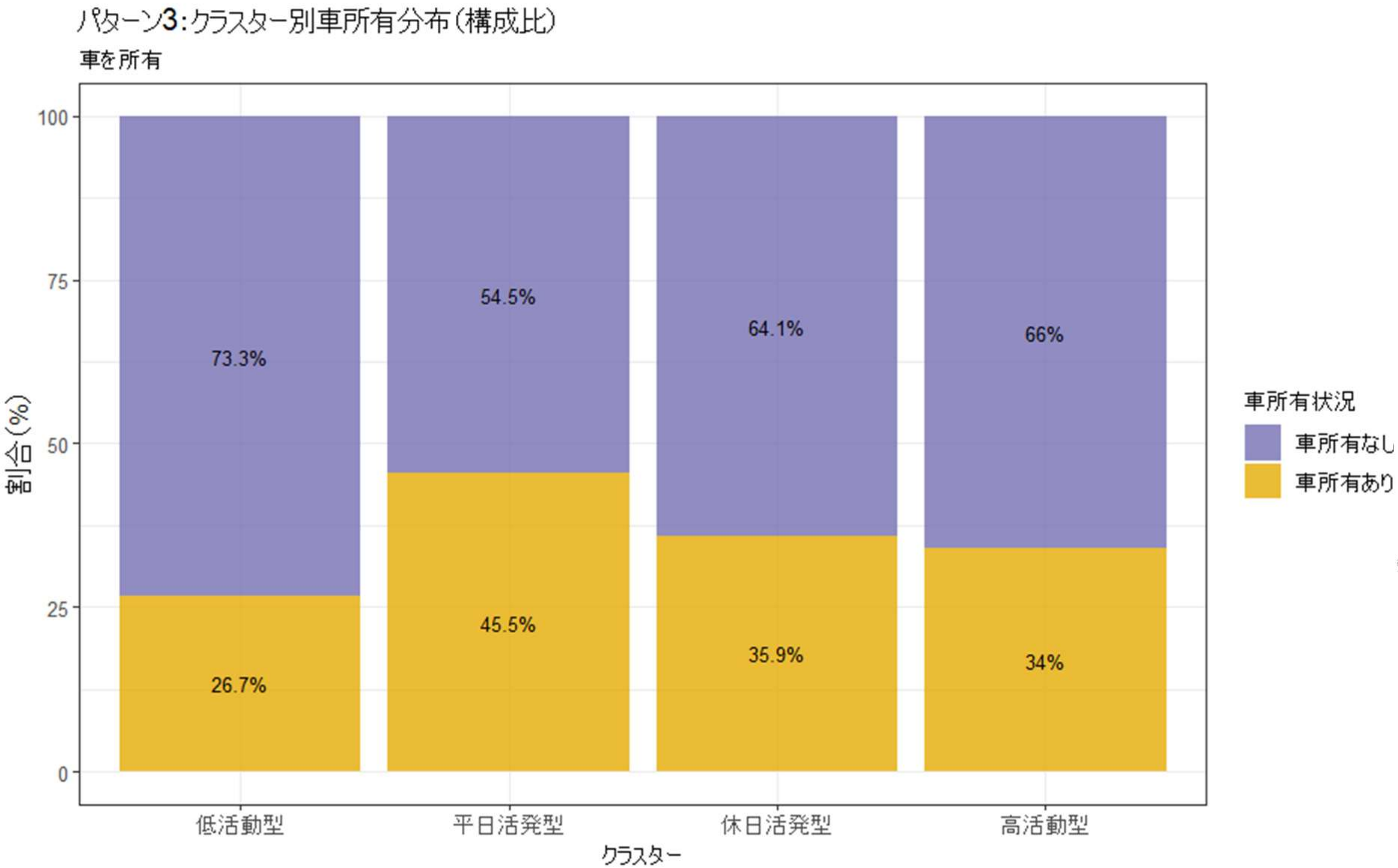
Distribution of Age for people categorized by hawks model

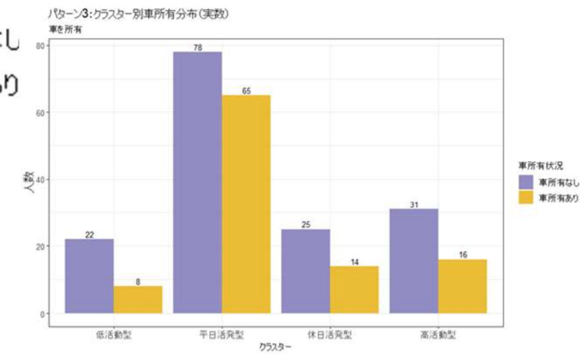


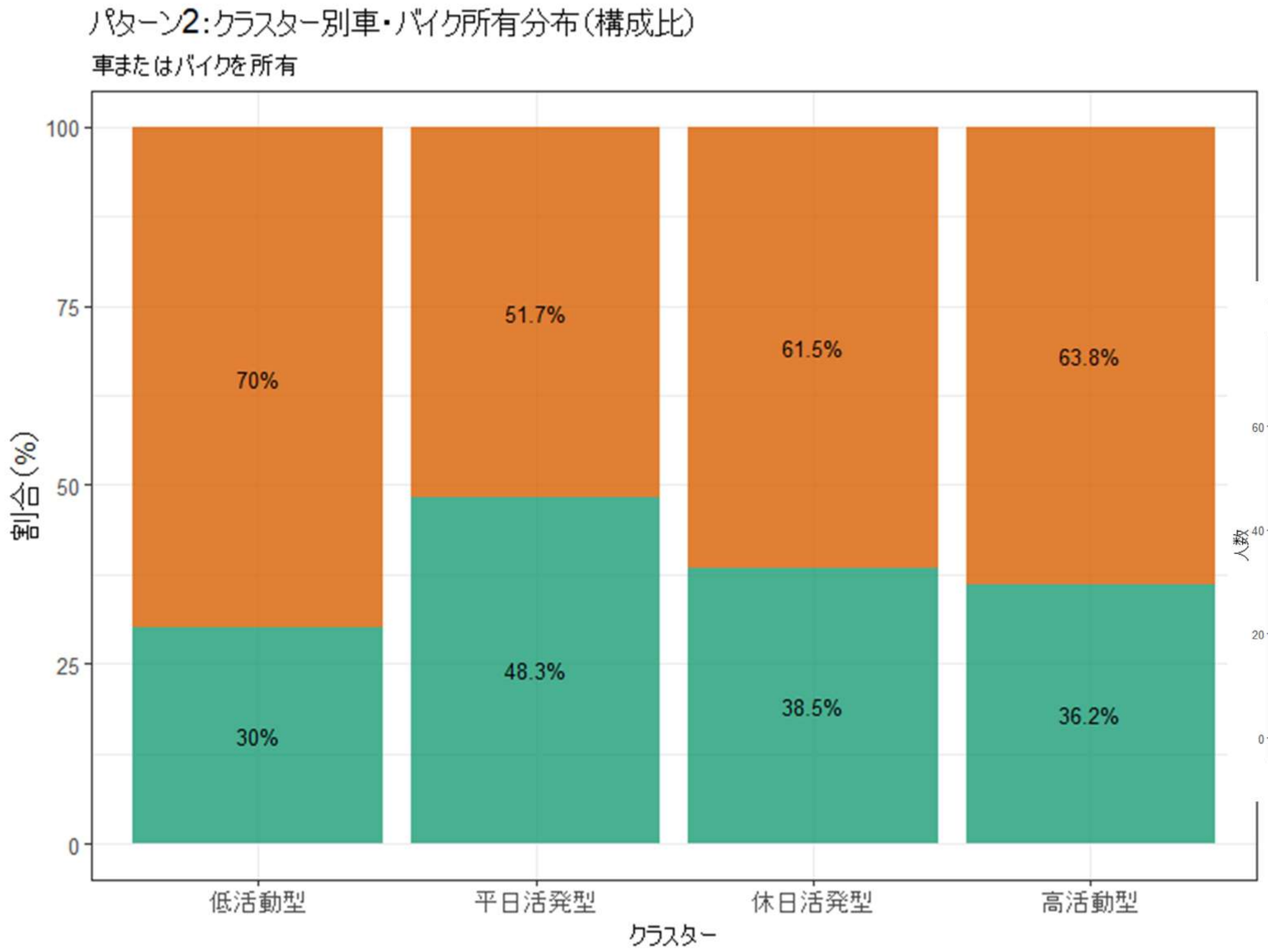
Result

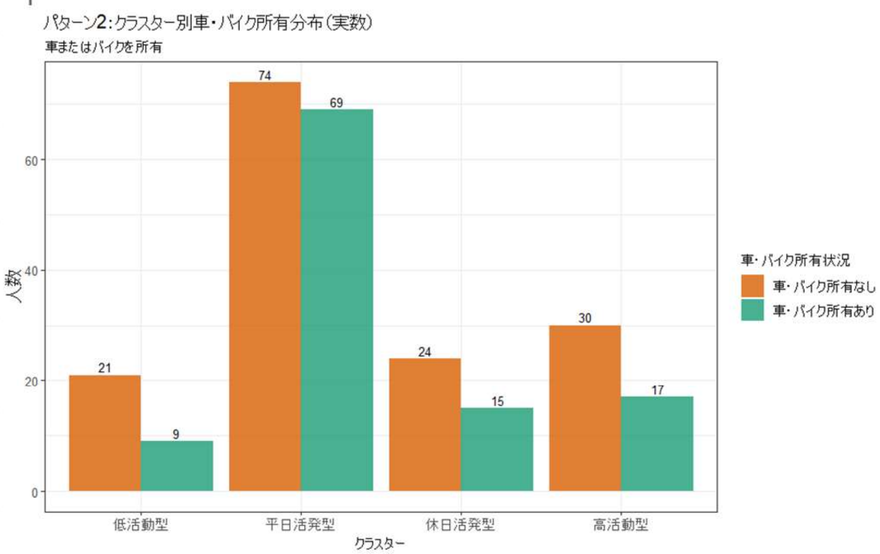
Distribution of Presence of Children for people categorized by hawks model




パターン3：クラスター別車所有分布（構成比）
車を所有
割合（%）
100
75
50
25
0
73.3%
26.7%
54.5%
45.5%
64.1%
35.9%
66%
34%
低活動型
平日活発型
休日活発型
高活動型
クラスター
車所有状況
車所有なし
車所有あり


パターン3：クラスター別車所有分布（実数）
車を所有
22
8
78
65
25
14
31
16
低活動型
平日活発型
休日活発型
高活動型
クラスター
車所有状況
車所有なし
車所有あり

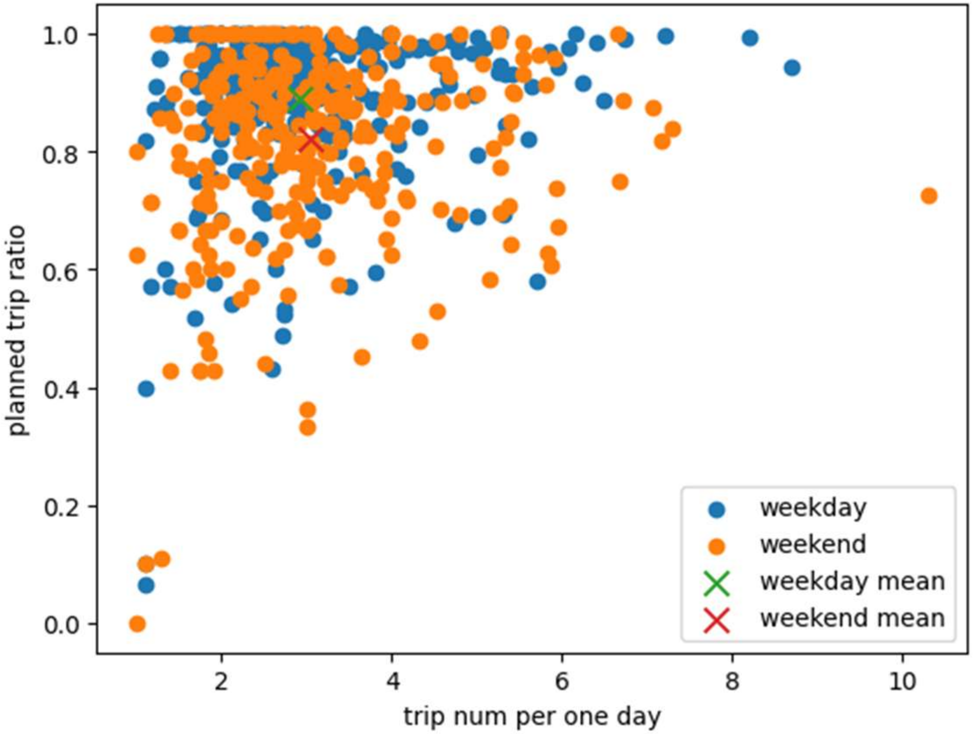

パターン2：クラスター別車・バイク所有分布（構成比）
車またはバイクを所有
割合(%)
100
75
50
25
0
70%
30%
51.7%
48.3%
61.5%
38.5%
63.8%
36.2%
低活動型
平日活発型
休日活発型
高活動型
クラスター


パターン2：クラスター別車・バイク所有分布（実数）
車またはバイクを所有
人数
60
40
20
0
21
9
74
69
24
15
30
17
低活動型
平日活発型
休日活発型
高活動型
クラスター
車・バイク所有状況
車・バイク所有なし
車・バイク所有あり

Result 結果

・Average number of trips per day on the horizontal axis and the planned_trip_ratio on the vertical axis.

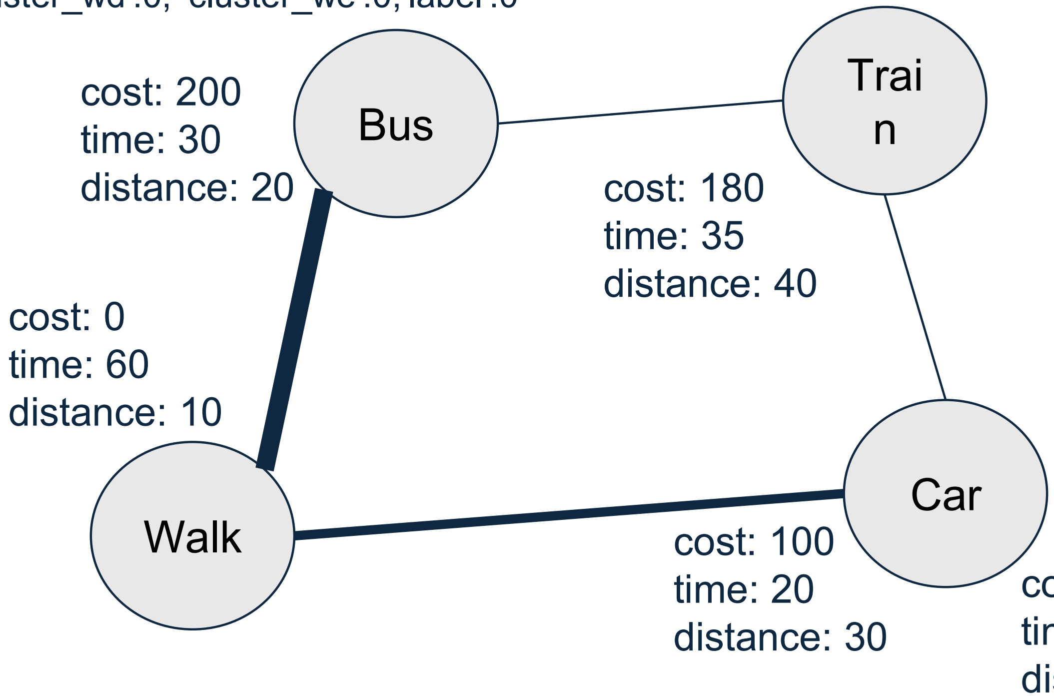
・Compared to weekdays, weekends show a slight overall increase in the average number of trips per day, but a slight decrease in the planned_trip_ratio, which also suggests that people tend to do more atypical behavior in weekend.



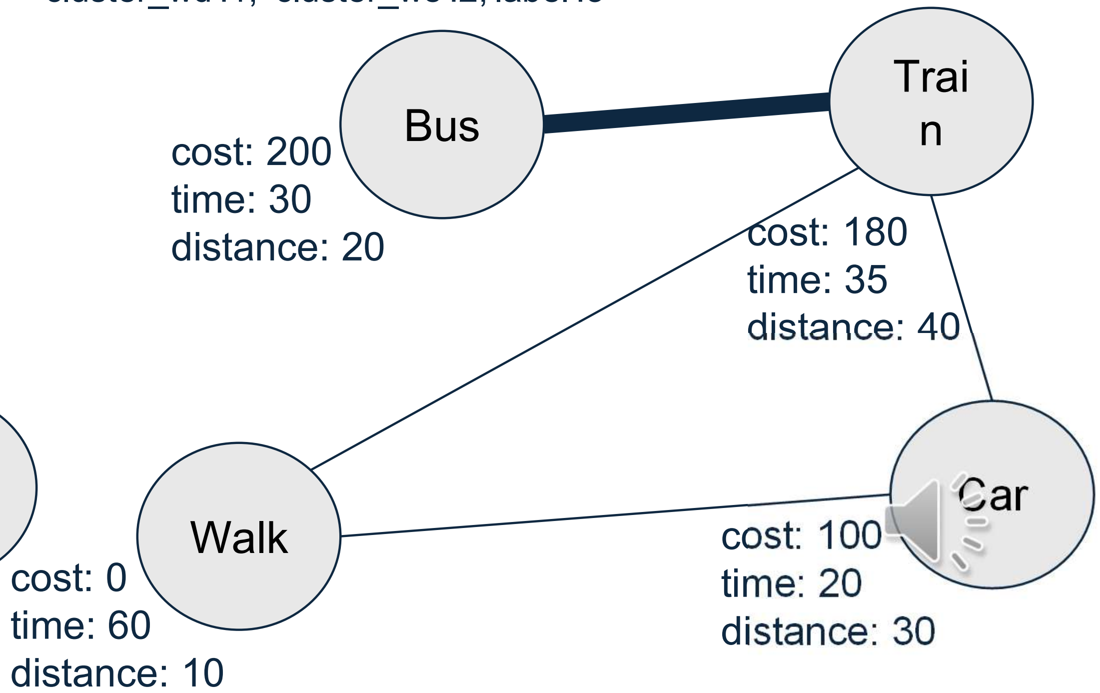
Understand travel mode preference by NestGNN

Travel mode preference can be different for persona, circumstances, etc.

Person A（working mom）：
'Gender': F, 'Age': 40, 'Occupation':, 'Income'; 7M, 'Type of Residence': rent,' Living Alone': 0, 'Spouse (Husband/Wife)': 1, 'Own Children': 1, 'Number of Children': 2, 'Car Ownership':1 , 'cluster_wd':0, 'cluster_we':0,'label':0

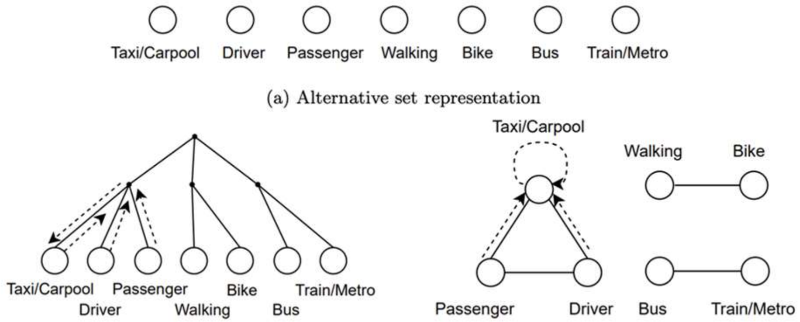


Person B（living alone）：
'Gender': M, 'Age': 75, 'Occupation': no occupation, 'Income'; 3M, 'Type of Residence': own,' Living Alone': 1, 'Spouse (Husband/Wife)': 0, 'Own Children': 0, 'Number of Children': 0, 'Car Ownership':0 , 'cluster_wd':1, 'cluster_we':2,'label':3



Understand travel mode preference by NestGNN

NestGNN is generalized version of nested logit, incorporating NL functional form



(a) Alternative set representation

(b) Nest representation and computation in nested logit models

(c) Alternative graph representation and message passing in NestGNN

Figure 1: Comparison of alternative representations with seven alternatives as an example. The arrow with the dotted line means the direction of message passing. In NestGNN's alternative graph, the solid line represents the edge between nodes.

$$V_{ni} = \phi(x_{ni}) + A(\left\{\mathbf{m}_{i,j}^{(1)}, \forall j \in \mathcal{N}(i)\right\}),$$

$$P_{ni} = \frac{\exp\left(\phi(x_{ni}) + A(\left\{\mathbf{m}_{i,j}^{(1)}, \forall j \in \mathcal{N}(i)\right\})\right)}{\sum_{m \in \mathcal{V}} \exp\left(\phi(x_{nm}) + A(\left\{\mathbf{m}_{i,m}^{(1)}, \forall j \in \mathcal{N}(i)\right\})\right)}.$$

NestGNN reduces to MNL and NL by choice of aggregation function, update function, node features and etc.

Example 1. *NestGNN reduces to a MNL model when its hyperparameter space $\mathcal{H}$ takes the following form:* $\{L = 0, M(x_i) = \emptyset, A = \emptyset, U = \emptyset, R(x_i) = w_i^T x_i^{(0)}\}$.

Here the NestGNN does not use message passing algorithm but only initializes the zero-layer node feature $x_i^{(0)}$. Without message passing algorithm, the utility function in the NestGNN enables the inputs from only alternative i's own attributes. The choice probability function of alternative i equals to $P_i = e^{w_i^T x_i^{(0)}} / \sum_{j \in \mathcal{V}} e^{w_j^T x_j^{(0)}}$, which is the same as the MNL model.

Example 2. *NestGNN reduces to a ASU-DNN model when its hyperparameter space $\mathcal{H}$ takes the following form:* $\{L = 0, M(x_i) = \emptyset, A = \emptyset, U = \emptyset, R(x_i^{(0)}) = MLP(x_i^{(0)})\}$.

Again, the NestGNN does not use message passing but only initializes the zero-layer node feature $x_i^{(0)}$. Different from the linear mapping in MNL, here the multilayer perception (a.k.a., feedforward neural networks) is used as the readout function. As a result, the choice probability function of alternative i equals to $P_i = e^{MLP(x_i^{(0)})} / \sum_{j \in \mathcal{V}} e^{MLP(x_j^{(0)})}$, which is the same as Wang et al. (2020a)'s work. In both examples, the MNL and ASU-DNN models follow the IIA constraint and thus exhibit the proportional substitution pattern among alternatives.

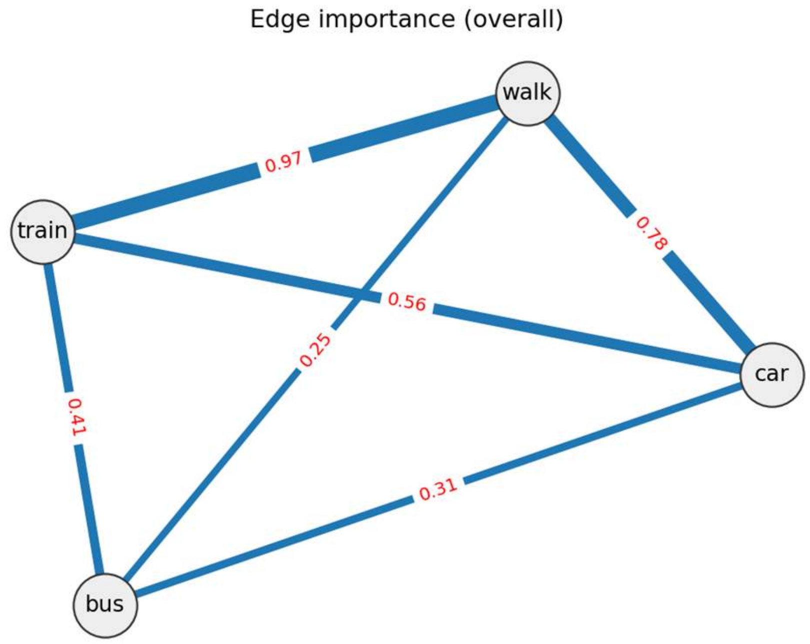
Example 3. *NestGNN reduces to a NL model when its hyperparameter space $\mathcal{H}$ takes the following form:* $L = 1, M(x_j^{(0)}) = w_j^\top x_j^{(0)} / \mu_k, A(m_i) = (\mu_k - 1) LSE_{j \in \mathcal{N}_*(i)}(M(x_j^{(0)}), U(x_i, a_i) = M(x_i^{(0)}) + A(m_i), R(x_i^{(1)}) = x_i^{(1)}\}$.

Zhou et al (2025) *NestGNN: A Graph Neural Network Framework Generalizing the Nested Logit Model for Travel Mode Choice*, https://arxiv.org/pdf/2509.07123

Understand travel mode preference by NestGNN

Travel mode preference can be visualized by GNNExplainer

Intuition: Larger edge values = “in the same nest”



Visualization of travel mode preference relation by NestGNN

- Differentiate log P(choice) with respect to elements of the adjacency matrix A (i.e., transportation mode relationship graph)
- Larger edge value indicates that similar comparisons and trade-offs are being made (i.e., in the same “nest”)
- Example: Suppose the importance value between car and taxi is high: Car and taxi utilities share similar structures
- This relation can be compared for different segments

Understand travel mode preference by NestGNN

For label 0, nest structure is more sparse (i.e., single transportation mode)

For label 1 and 2, nest structure is similar to overall

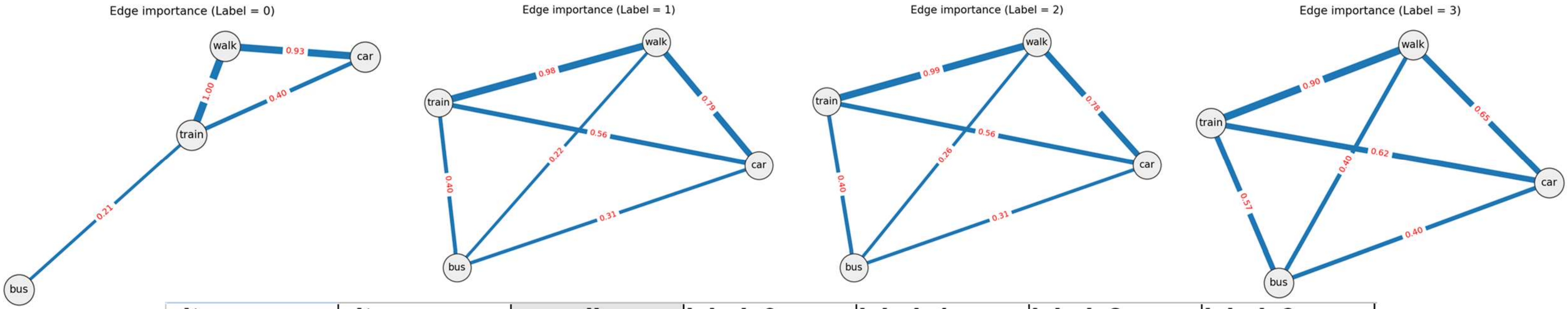
Label_0 (routine behavior on weekdays) and Label_3 (Outing on weekends) are contrary

0: routine behavior (Weekdays)

1: outing behavior (Weekdays)

2: routine behavior (Weekends)

3: outing behavior (Weekdays)



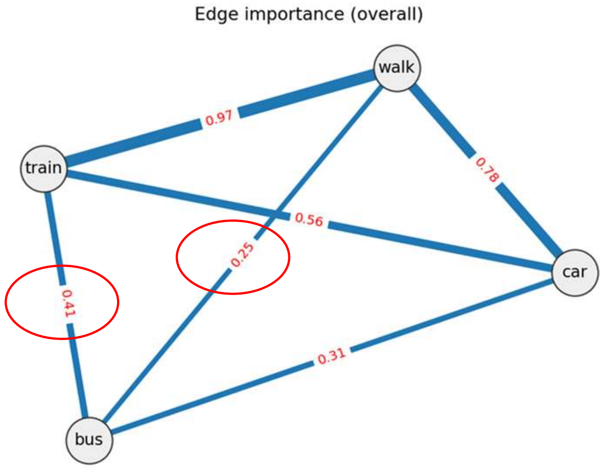
alt_u	alt_v	overall	label_0	label_1	label_2	label_3
walk	train	**0.97**	1.00	0.98	0.99	0.90
walk	car	**0.78**	0.93	0.79	0.78	0.65
car	train	**0.56**	0.40	0.56	0.56	0.62
bus	train	**0.41**	0.21	0.40	0.40	0.57
car	bus	**0.31**	0.19	0.31	0.31	0.40
walk	bus	**0.25**	0.07	0.22	0.26	0.40

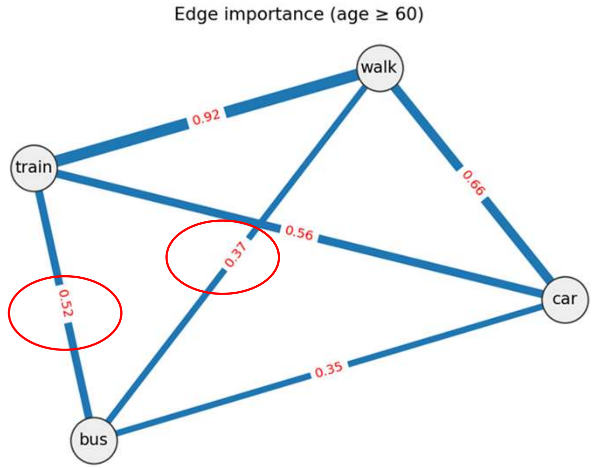
Understand travel mode preference by NestGNN

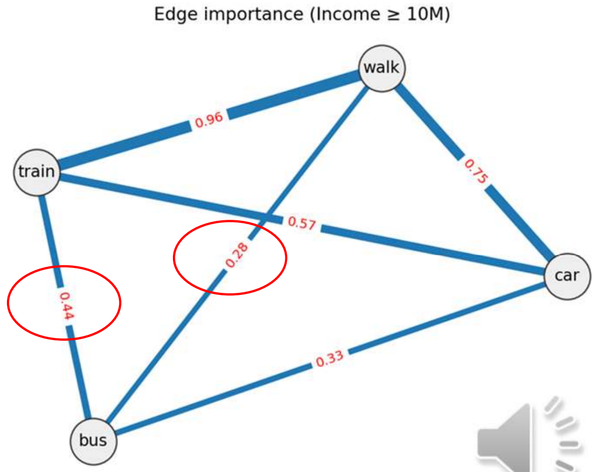
Compare overall, senior (over age 60), high income (over 10M yen)

For senior people, [1] train and bus and [2] bus and walk are tied stronger than for overall and people with high income earners

For high income earners, no particular relationship can be observed







Appendix: Trip count by personas (Age, Income, Label)

Overall, Age >= 60, Income >=10M, Label (0, 1, 2, 3)

```
df.Age.value_counts()
✓ 0.0s

Age
4    11003
3     9248
2     7563
5     4530
1     1117
Name: count, dtype: int64
```

```
df.Income.value_counts()
✓ 0.0s

Income
2    10307
3    10038
4     9195
5     3469
1      452
Name: count, dtype: int64
```

```
df.label.value_counts()
✓ 0.0s

label
1.0    14948
3.0     7063
2.0     6384
0.0     3869
Name: count, dtype: int64
```

Future Work ()

- Consider other alternative modes than Car, Bus, Train, Walk
    - Such as bicycle shareing
- Visualize feature importance for explaining edge
    - Sophisticate GNNExplainer method
- Visualize feature importance for explaining edge
    - Sophisticate GNNExplainer method