

# データベース：データ分析演習

23T2013D 石黒康太

2026年1月22日

## 1 はじめに

本レポートでは、奈良先端科学技術大学院大学で行われたカーシェアリング実証実験の利用実績データを用いて、3つの観点からデータ分析を実施した。

分析に使用したデータは、乗車履歴データ、駐車履歴データ、目的地データ、ユーザー情報データの4種類である。各データセットの概要を表1に示す。

表1: 使用したデータセットの概要

データセット名	ファイル名	内容
乗車履歴データ	history.csv	ユーザーが実際に車に乗車した履歴（1行=1回の乗車）
駐車履歴データ	trip.csv	乗車時に訪れた目的地の記録
目的地データ	spot.csv	目的地の詳細情報（名称、位置、種類など）
ユーザー情報データ	user.csv	全ユーザーのリスト（教職員/学生の区別）

各データセットには複数の列が含まれており、主要な列の詳細を以下に示す。

乗車履歴データ（history.csv）の主要な列を表2に示す。

表2: 乗車履歴データの主要な列

列名	内容
history_id	乗車履歴を一意に識別する ID
started_at	乗車を開始した日時
ended_at	乗車を終了した日時
from_parking_lot	乗車開始地点（NAIST または ATR）
to_parking_lot	乗車終了地点（NAIST または ATR）
car	使用した車両の名前
passengers_count	乗車人数（運転手を含む総人数）
distance	移動距離
user_id	車を運転したユーザーの ID

駐車履歴データ（trip.csv）の主要な列を表3に示す。1回の乗車で複数の目的地を訪れた場合、複数行のデータが記録される。

表 3: 駐車履歴データの主要な列

列名	内容
created_at	目的地へ到着した日時
lat	目的地の緯度
lon	目的地の経度
car	使用した車両の名前
user_id	運転したユーザーの ID
spot_id	目的地の ID (spot.csv と対応)
history_id	対応する乗車履歴の ID (history.csv と対応)

目的地データ (spot.csv) の主要な列を表 4 に示す。

表 4: 目的地データの主要な列

列名	内容
spot_id	目的地を一意に識別する ID
spot_name	目的地の名称
lat	目的地の緯度
lon	目的地の経度
count	ユーザーがこの目的地を訪れた回数
spot_types	目的地の種類 (例: restaurant、park など)
is_parking	カーシェア用の駐車場かどうか

ユーザー情報データ (user.csv) の主要な列を表 5 に示す。

表 5: ユーザー情報データの主要な列

列名	内容
user_id	ユーザーを一意に識別する ID
user_type	ユーザーの種類 (staff=教職員、student=学生)

これらのデータセットは、user\_id、history\_id、spot\_id といった ID を介して相互に関連付けられている。本分析では、これらのデータを適切に結合し、ユーザータイプによる利用パターンの違いや、乗車人数と移動距離の関係、目的地数と移動距離の関係を明らかにすることを目的とする。

## 2 分析 1：ユーザータイプ別の利用時間帯パターン

### 2.1 分析の目的と方法

カーシェアリングの利用パターンは、ユーザーのライフスタイルに大きく影響を受けると考えられる。特に、教職員と学生では日常の活動時間帯や週末の過ごしが異なる可能性がある。そこで、教職員と学生のそれについて、曜日と時間帯ごとの利用割合を詳細に分析した。

分析方法として、まず乗車履歴データから乗車開始時刻を抽出し、曜日（月曜日から日曜日の 7 段階）と時間帯（0 時から 23 時の 24 段階）に分類した。次に、ユーザー情報データと結合することで、各乗車記録がどのユーザータイプによるものかを特定した。最後に、ユーザータイプごとに、曜日と時間帯の組み合わせごとの乗車回数を集計し、各ユーザータイプ内での利用割合に変換した。この割合をヒートマップとして可視化することで、利用パターンの特徴を視覚的に捉えることができる。ヒートマップでは、色が明るいほど利用割合が高いことを示している。

### 2.2 分析結果

図 1 に教職員の利用パターン、図 2 に学生の利用パターンを示す。

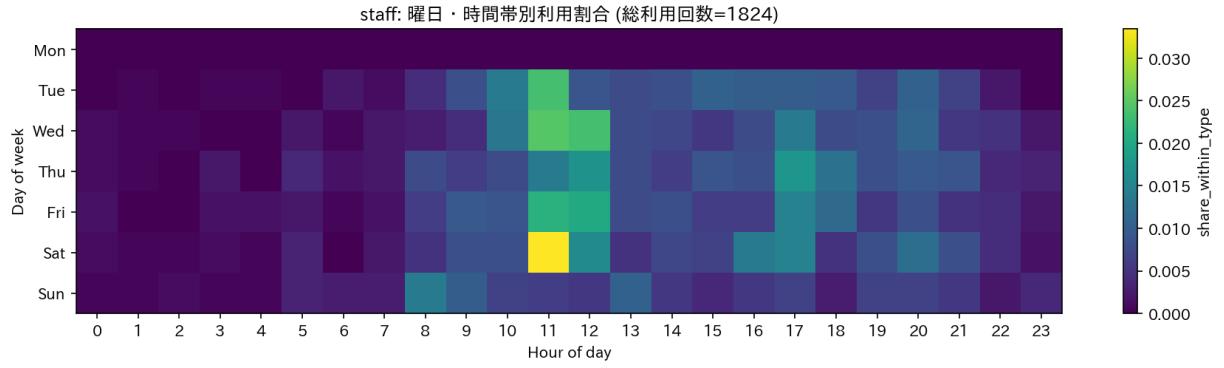


図 1: 教職員の曜日・時間帯別利用割合（総利用回数=1824 回）

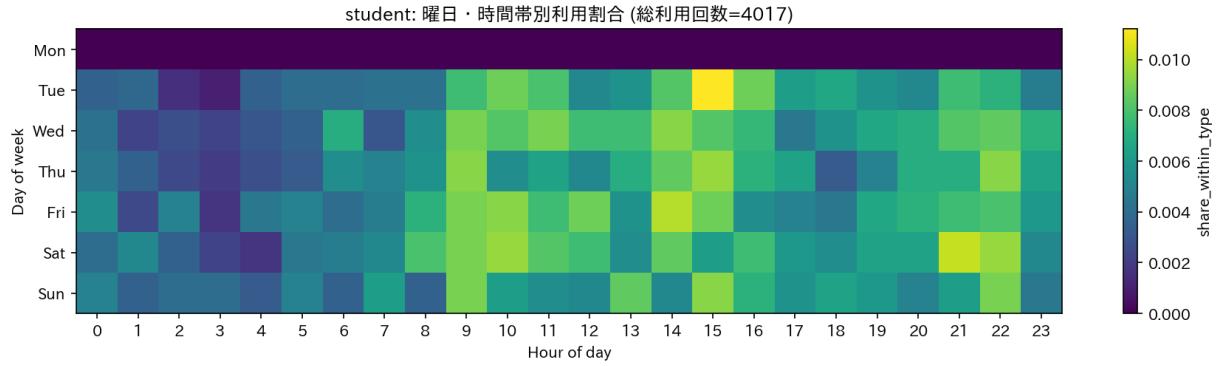


図 2: 学生の曜日・時間帯別利用割合（総利用回数=4017 回）

### 2.2.1 教職員の利用パターン

教職員の総利用回数は 1824 回であり、学生の約半分である。利用パターンには明確な特徴が見られる。

まず、平日（月曜日から金曜日）に利用が強く集中している。特に火曜日から木曜日にかけての利用が多く、これらの曜日の昼間に利用が活発である。時間帯別に見ると、午前 10 時から 12 時頃と午後 14 時から 17 時頃の 2 つの時間帯に利用のピークが見られる。これは、教職員の勤務時間帯と一致しており、昼休みを挟んで午前と午後にそれぞれ業務に関連した移動が行われていると推測される。

興味深い点として、土曜日の 11 時頃に最も利用割合が高いピーク（黄色で表示）が確認できる。これは、土曜日にも一部の教職員が出勤し、午前中に移動を行っていることを示唆している。一方、日曜日の利用はほぼ皆無であり、完全な休日として過ごされていることが分かる。

また、早朝（0 時から 8 時頃）や深夜（19 時以降）の利用は極めて少なく、教職員のカーシェアリング利用は勤務時間内に限定される傾向が強いことが明らかである。

### 2.2.2 学生の利用パターン

学生の総利用回数は 4017 回であり、教職員の 2 倍以上である。学生の利用パターンは教職員とは大きく異なる特徴を示している。

最も顕著な違いは、利用時間帯の広がりである。学生は平日の午前 9 時から午後 23 時頃まで幅広い時間帯で利用しており、特定の時間帯への集中が教職員ほど顕著ではない。火曜日の午後 14 時頃に最も高いピーク（黄色で表示）が見られるが、それ以外にも水曜日から金曜日の昼間から夕方にかけて、比較的均一に利用が分散している。

学生の利用は夕方から夜にかけても継続しており、午後 18 時から 22 時頃の時間帯にも一定の利用が確認できる。これは、学生が授業や研究活動を終えた後にカーシェアリングを利用していることを示している。教職員では見られなかった深夜帯（0 時から 2 時頃）にも若干の利用があり、学生の生活時間帯が多様であることが窺える。

週末の利用に関しても、学生は土曜日や日曜日にも一定の利用が見られる。特に土曜日の昼間から夕方にかけて利用があり、学生が週末にも研究活動や私的な外出にカーシェアリングを活用していることが分かる。

### 2.2.3 両者の比較と考察

教職員と学生の利用パターンを比較すると、明確な対比が浮かび上がる。教職員は勤務時間に厳密に対応した規則的な利用パターンを示し、平日の昼間に集中している。一方、学生はより柔軟で多様な時間帯での利用が見られ、夜間や週末にも一定の利用がある。

この違いは、両者のライフスタイルの違いを反映していると考えられる。教職員は定時勤務が基本であるのに対し、学生は授業や研究のスケジュールが個人によって異なり、また夜間に研究活動を行うことも多い。さらに、学生の利用回数が教職員の2倍以上であることは、学生にとってカーシェアリングがより重要な移動手段となっていることを示唆している。

このように、ユーザータイプによって利用時間帯のパターンが大きく異なることが明らかになった。カーシェアリングサービスの運用においては、こうした利用パターンの違いを考慮し、それぞれのユーザーグループのニーズに応じた車両配置やサービス提供時間を検討することが重要である。

## 3 分析2：距離・所要時間・乗車人数の分布

### 3.1 分析の目的と方法

カーシェアリングの利用形態を理解するため、移動距離と所要時間がユーザータイプによってどのように異なるかを分析した。また、乗車人数が移動距離に与える影響についても詳細に調査した。

分析方法として、まず乗車履歴データから移動距離と乗車開始・終了時刻を取得し、所要時間を計算した。次に、ユーザー情報データと結合することで、各乗車記録のユーザータイプを特定した。ユーザータイプ別の移動距離と所要時間の分布を箱ひげ図で可視化し、中央値、四分位範囲、外れ値の有無を確認した。箱ひげ図において、箱の中央の線は中央値を、箱の上下端は第3四分位数と第1四分位数を、ひげの端は正常範囲の最大値・最小値を、円は外れ値を表している。

さらに、乗車人数と移動距離の関係を調べるため、乗車人数を「1人」「2人」「3人」「4人」「5人以上」の5つのカテゴリに分類した。各カテゴリについて移動距離の分布を可視化し、乗車人数が増えると移動距離がどのように変化するかを分析した。

### 3.2 分析結果

#### 3.2.1 ユーザータイプ別の移動距離の分布

図3にユーザータイプ別の移動距離の分布を示す。

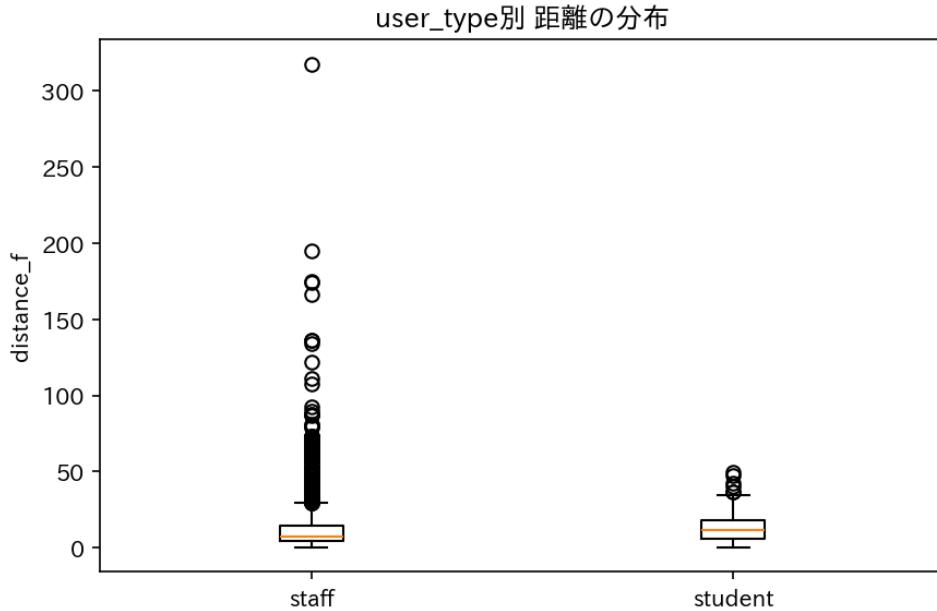


図 3: ユーザータイプ別の移動距離の分布

移動距離の分布を見ると、教職員と学生の中央値はともに約 10 度程でほぼ同じである。箱の大きさ（第 1 四分位数から第 3 四分位数の範囲）も類似しており、典型的な移動距離の範囲は両者で大きく変わらない。

しかし、外れ値の分布には顕著な違いが見られる。教職員は非常に多くの外れ値が存在し、最大で 300 を超える長距離移動も記録されている。これらの外れ値は、教職員が業務で遠方へ出張する際にカーシェアリングを利用していることを示唆している。一方、学生の外れ値は相対的に少なく、最大でも 50 度程に留まっている。これは、学生の移動範囲が教職員に比べて限定的であることを示している。

### 3.2.2 ユーザータイプ別の所要時間の分布

図 4 にユーザータイプ別の所要時間の分布を示す。

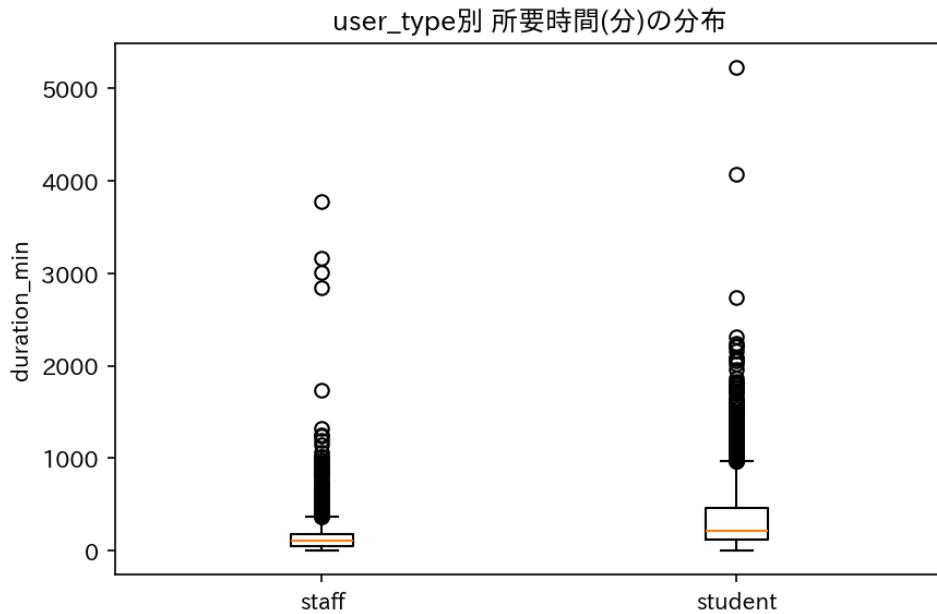


図 4: ユーザータイプ別の所要時間 (分) の分布

所要時間の分布は、移動距離とは異なる傾向を示している。教職員の中央値は約 100～150 分であるのに対し、学生の中央値は約 300 分とかなり長い。これは興味深い結果である。移動距離の中央値はほぼ同じであるにもかかわらず、学生の方が所要時間が長いということは、学生が目的地での滞在時間を含めた長時間利用をしていることを意味

している。

教職員は業務での短時間の外出が多く、移動と用事を済ませてすぐに戻ってくるパターンが主流であると考えられる。一方、学生は車を借りている間に複数の用事を済ませたり、目的地で長時間過ごしたりする傾向があると推測される。

外れ値に関しては、両ユーザータイプとも多数存在し、教職員は最大約 3800 分（約 63 時間）、学生は最大約 5000 分（約 83 時間）という極めて長時間の利用も記録されている。これらは数日にわたる利用や、返却忘れの可能性も考えられる。

### 3.2.3 乗車人数と移動距離の関係（全体傾向）

図 5 に乗車人数と移動距離の散布図を示す。

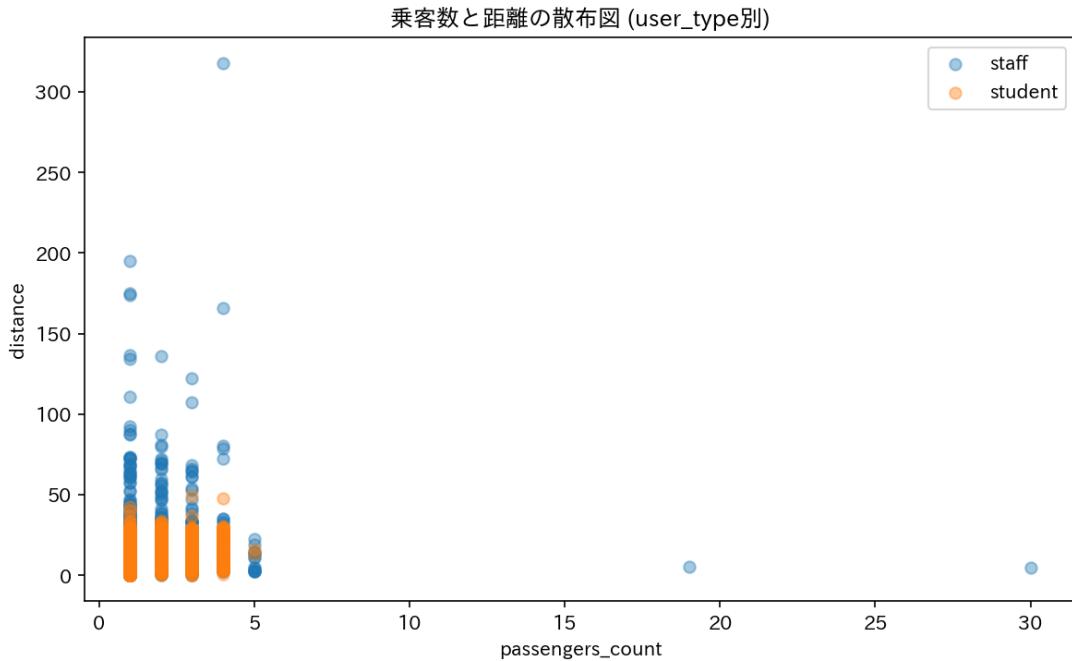


図 5: 乗車人数と移動距離の散布図（ユーザータイプ別）

散布図から、カーシェアリングの利用は圧倒的に少人数での利用が多いことが分かる。乗車人数 1 人から 4 人までのデータが大部分を占め、5 人以上の利用は極めて少ない。特に、1 人での利用（青とオレンジの点が最も密集している部分）が最多であり、乗車人数が増えるにつれて利用回数は急激に減少している。

ユーザータイプ別に見ると、教職員（青）の方が学生（オレンジ）よりも高い位置にデータ点が散らばっており、特に 1 人での利用時に長距離移動をする事例が多いことが視覚的に確認できる。また、極端に乗車人数が多いケース（20 人や 30 人）は教職員の利用であり、これは大型の送迎バスなど特殊な車両を使用した可能性がある。

### 3.2.4 乗車人数と移動距離の関係（ユーザータイプ別詳細）

図 6 と図 7 に、教職員と学生それぞれの乗車人数別移動距離の分布を示す。

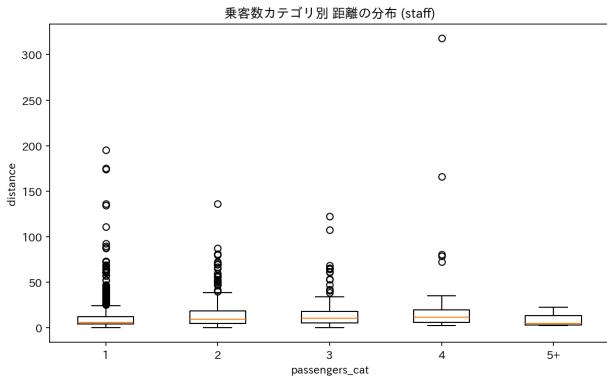


図 6: 教職員の乗車人数別移動距離

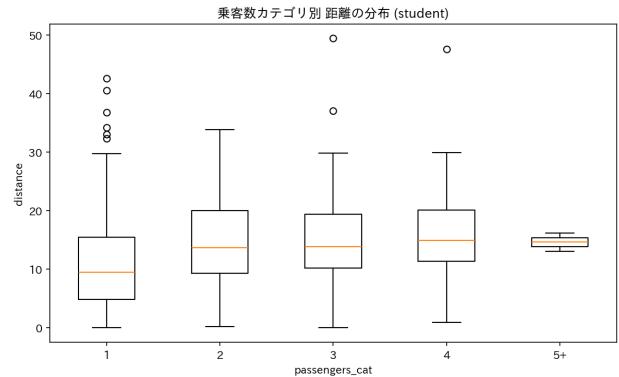


図 7: 学生の乗車人数別移動距離

教職員の箱ひげ図を見ると、1人での利用時の中央値は約10であるが、2人以上になると中央値が約15程度まで上昇している。ただし、5人以上になると中央値は約10に戻っている。これは、2~4人での利用時に比較的長距離の移動が多いことを示している。また、1人での利用時には非常に多くの外れ値が見られ、最大で300を超える長距離移動も存在する。これらの外れ値は、一部の教職員が業務で遠方へ単独で出張する際の利用と考えられる。

学生の箱ひげ図では、乗車人数が増えるにつれて中央値が段階的に上昇する傾向が明確である。1人での利用時の中央値は約10であるが、2人で約13、3人で約14、4人で約15、5人以上で約15と、人数が増えるほど移動距離も長くなっている。これは、学生が複数人で遠方へ出かける際にカーシェアリングを利用していることを示唆している。学生の場合、外れ値は教職員ほど多くなく、移動距離の上限も50程度に収まっている。

### 3.2.5 考察

分析2の結果から、以下の知見が得られた。

まず、移動距離の典型的な値（中央値）は教職員と学生でほぼ同じであるが、教職員の方が極端に長距離の移動をする事例が多い。これは、教職員の業務の性質上、遠方への出張が発生するためと考えられる。一方、所要時間に関しては学生の方が長く、学生が車を長時間借りて複数の用事を済ませる利用形態が多いことが示唆された。

乗車人数と移動距離の関係については、両ユーザータイプとも人数が増えると移動距離が増加する傾向が見られたが、その傾向は学生の方がより明確である。学生は複数人での利用時に遠出をする傾向があるのに対し、教職員は単独での長距離移動が多い。この違いは、利用目的の違い（教職員は業務出張、学生は団体での外出）を反映していると考えられる。

## 4 分析3：目的地数と移動距離の関係

### 4.1 分析の目的と方法

カーシェアリングでは、1回の乗車で複数の目的地を訪れることがある。目的地の数が移動距離にどのように影響するかを調べることは、利用者の行動パターンを理解する上で重要である。そこで、1回の乗車における目的地数と移動距離の関係を分析した。

分析方法として、まず駐車履歴データから各乗車記録(history\_id)に対応する目的地の数を集計した。次に、乗車履歴データと結合し、目的地数を「1箇所」「2箇所」「3箇所」「4箇所」「5箇所以上」の5つのカテゴリに分類した。各カテゴリについて移動距離の分布を可視化し、目的地数が増えると移動距離がどのように変化するかを調査した。

さらに、ユーザータイプ別の目的地数の分布を比較し、教職員と学生で利用形態に違いがあるかを確認した。また、乗車人数と目的地数の組み合わせパターンを分析することで、どのような利用形態が多いかを明らかにした。

### 4.2 分析結果

#### 4.2.1 ユーザータイプ別的目的地数

図8にユーザータイプ別の1乗車あたり目的地数の分布を示す。

バイオリンプロットを見ると、教職員と学生のいずれも、1回の乗車で1箇所または2箇所の目的地を訪れるこ

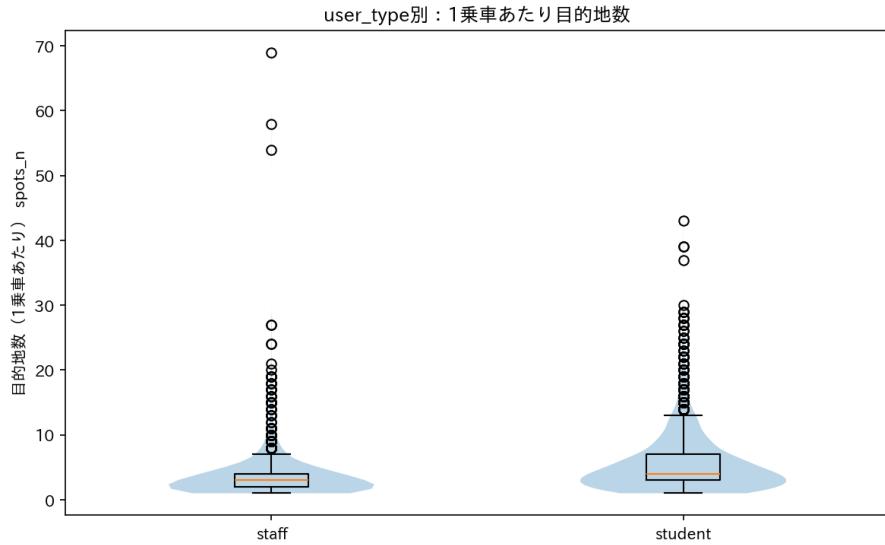


図 8: ユーザータイプ別の 1 乗車あたり目的地数の分布

が最も多いことが分かる。分布の形状は両ユーザータイプで類似しており、目的地数に関しては大きな違いは見られない。ただし、学生の方がやや目的地数のばらつきが大きい傾向がある。

#### 4.2.2 目的地数と移動距離の関係

図 9 に目的地数別の移動距離のジッター散布図を、図 10 に箱ひげ図を示す。

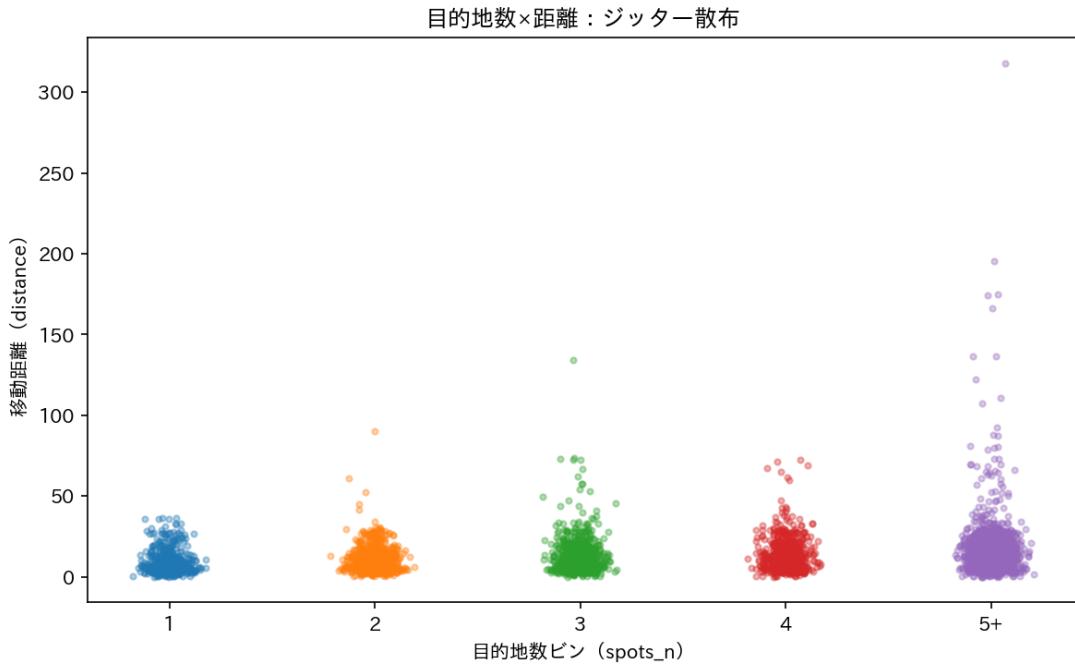


図 9: 目的地数別の移動距離分布（ジッター散布図）

両図から、目的地数が増えるにつれて移動距離も増加する明確な傾向が確認できる。1箇所のみを訪れる場合の移動距離は比較的短いが、目的地が2箇所、3箇所と増えるにつれて、移動距離の中央値が上昇している。5箇所以上の目的地を訪れる場合には、移動距離が大幅に増加する。

この結果は直感的に理解しやすい。複数の目的地を訪れる場合、それぞれの目的地間の移動距離が累積されるため、総移動距離が長くなるのは当然である。また、ジッター散布図から、同じ目的地数でも移動距離にはかなりのばらつきがあることが分かる。これは、訪れる目的地の位置関係や移動ルートによって、総移動距離が大きく変わることを示している。

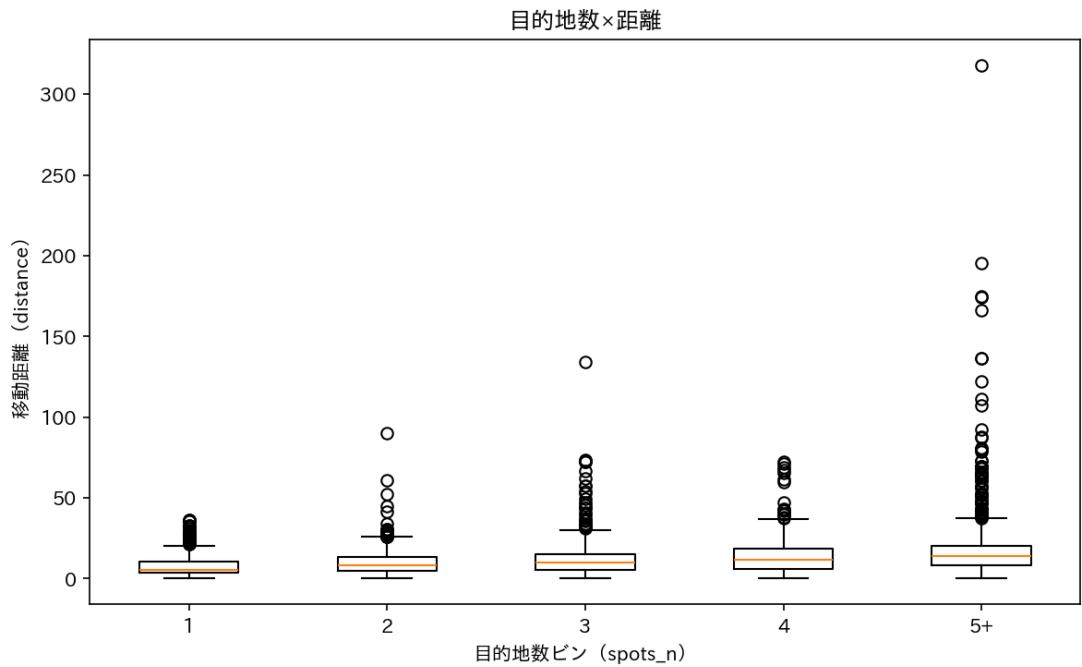


図 10: 目的地数別の移動距離分布（箱ひげ図）

#### 4.2.3 乗車人数と目的地数の関係

図 11 に乗車人数と目的地数の組み合わせパターンを示す。

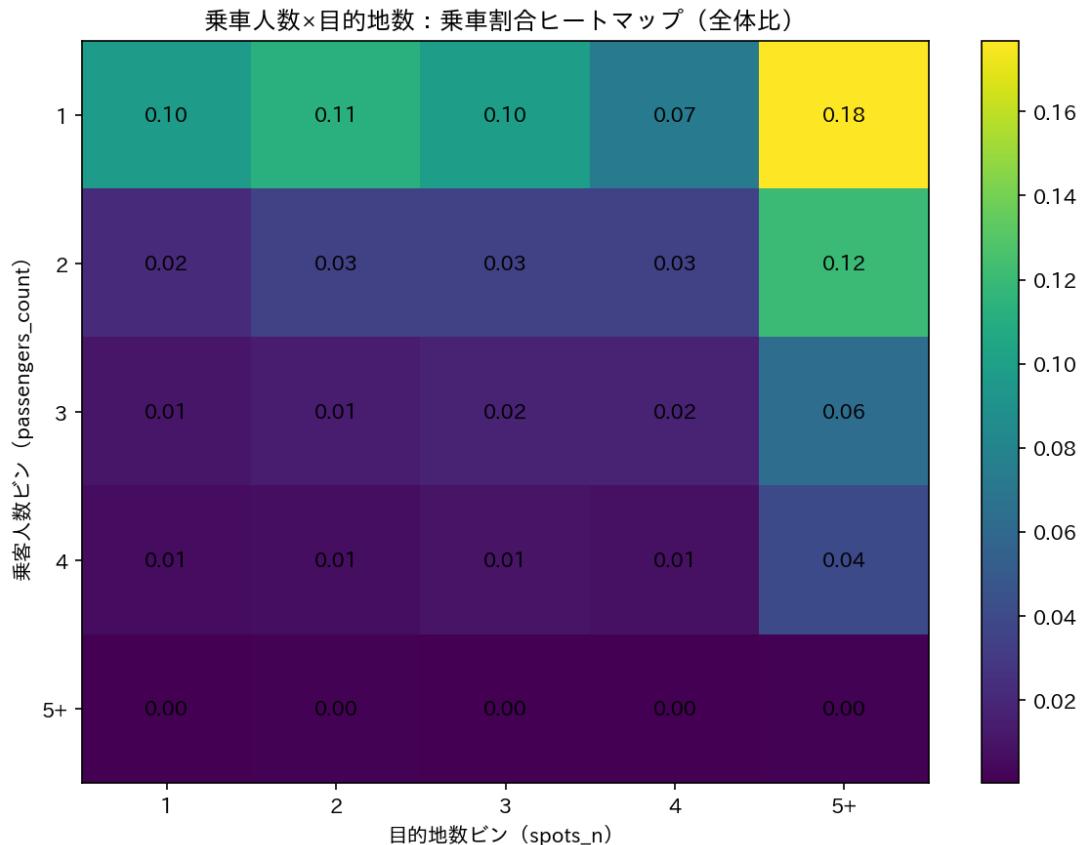


図 11: 乗車人数と目的地数の組み合わせパターン

ヒートマップから、最も多い利用形態は「1人で1箇所の目的地を訪れる」パターンであることが分かる。次いで、「1人で2箇所」「1人で3箇所」の順に多い。乗車人数が増えるほど利用頻度は減少し、特に4人以上での利用は少

ない。

興味深い点として、乗車人数が多い場合でも、目的地数は1箇所または2箇所に集中している。これは、団体での利用時には特定の目的地へ直接移動するか、少數の目的地を効率的に巡ることが多いことを示唆している。複数人で5箇所以上の目的地を訪れるケースは極めて稀である。

この分析から、カーシェアリングは主に単独での短距離移動や少數の目的地への移動に利用されており、複数人の長距離・多目的地移動は例外的であることが明らかになった。

## 5 まとめ

本レポートでは、カーシェアリング実証実験の利用実績データを用いて、3つの観点から分析を実施した。

分析1では、ユーザータイプ別の利用時間帯パターンを調査し、教職員は平日の勤務時間帯に集中した利用パターンを示すのに対し、学生はより柔軟な時間帯での利用が見られることを明らかにした。

分析2では、移動距離と所要時間の分布を調査し、教職員と学生の間に大きな差は見られないことを確認した。また、乗車人数が増えると移動距離が増加する傾向があることを発見した。

分析3では、目的地数と移動距離の関係を分析し、目的地数が増えるほど移動距離も増加する明確な傾向を確認した。また、最も多い利用形態は単独での少數の目的地への移動であることが分かった。

これらの分析結果から、カーシェアリングの利用パターンはユーザータイプによって時間帯に違いが見られるものの、移動距離や目的地数に関しては類似した傾向を示すことが明らかになった。また、乗車人数や目的地数は移動距離に直接的な影響を与えることが確認された。

今回の分析では、利用時間帯、移動距離、乗車人数、目的地数といった基本的な指標に焦点を当てたが、今後はより詳細な目的地の種類や移動ルートの分析を行うことで、カーシェアリングの利用形態をさらに深く理解できると考えられる。