

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方が異なる可能性がある。そこで、教職員と学生のそれぞれについて、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。

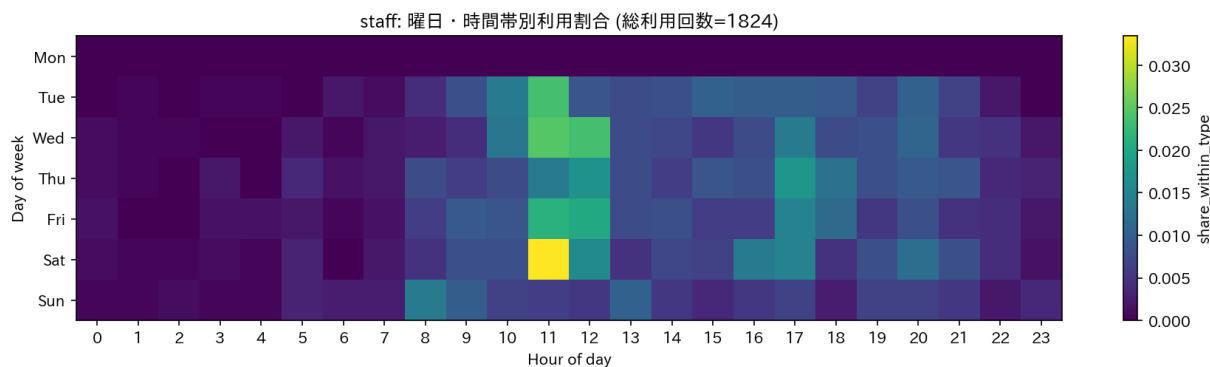


図 1: 教職員の曜日・時間帯別利用割合 (総利用回数=1824 回)

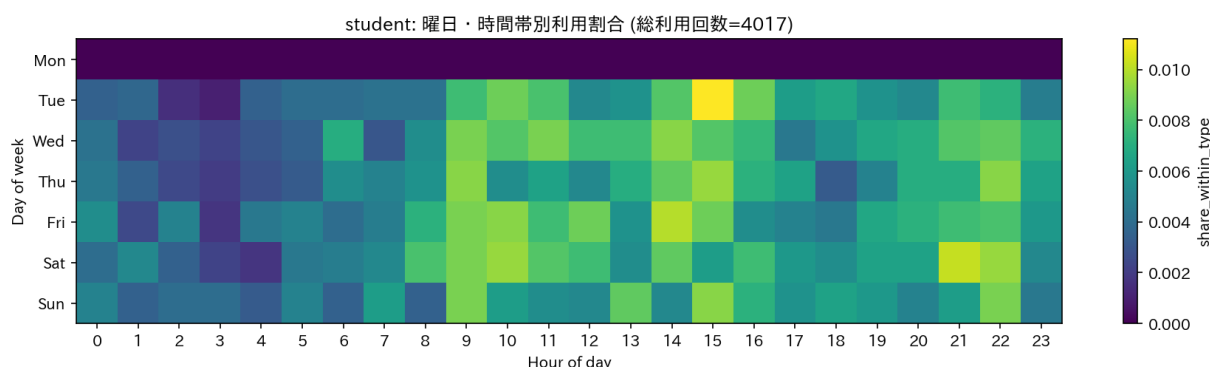


図 2: 学生の曜日・時間帯別利用割合 (総利用回数=4017 回)

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。

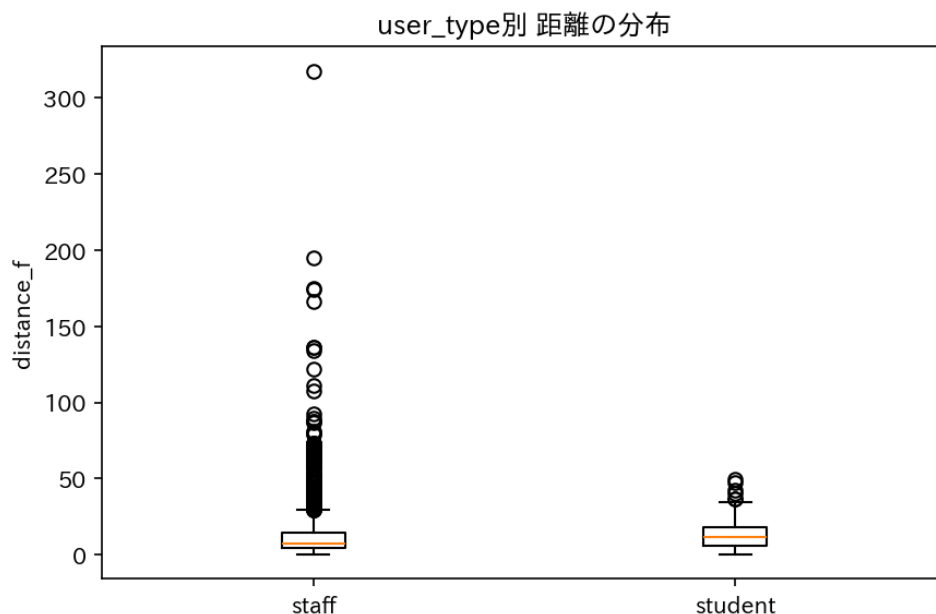


図 3: ユーザータイプ別の移動距離の分布

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 程度でほぼ同じである。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 程度に留まっている。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。

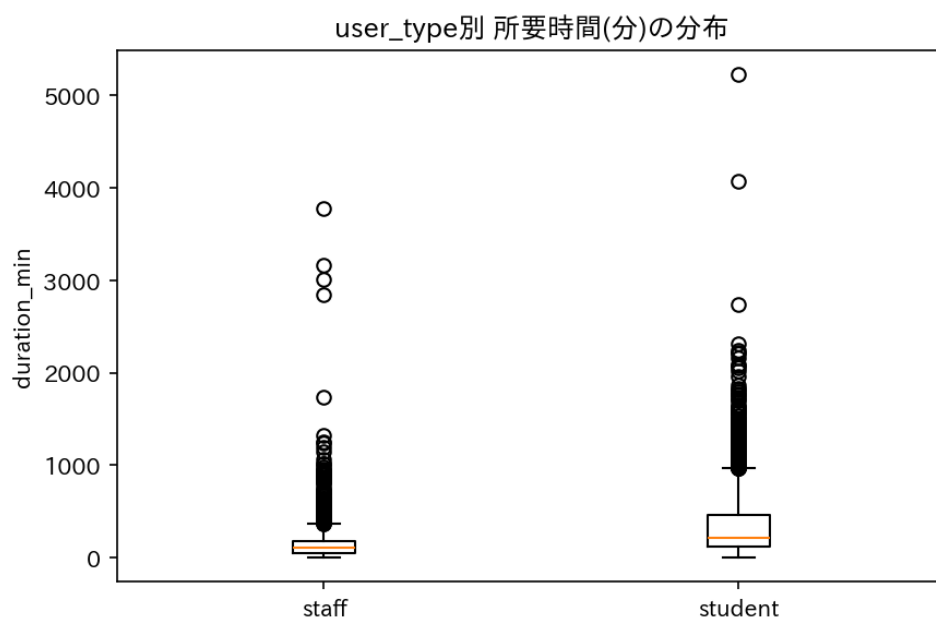


図 4: ユーザータイプ別の所要時間（分）の分布

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。

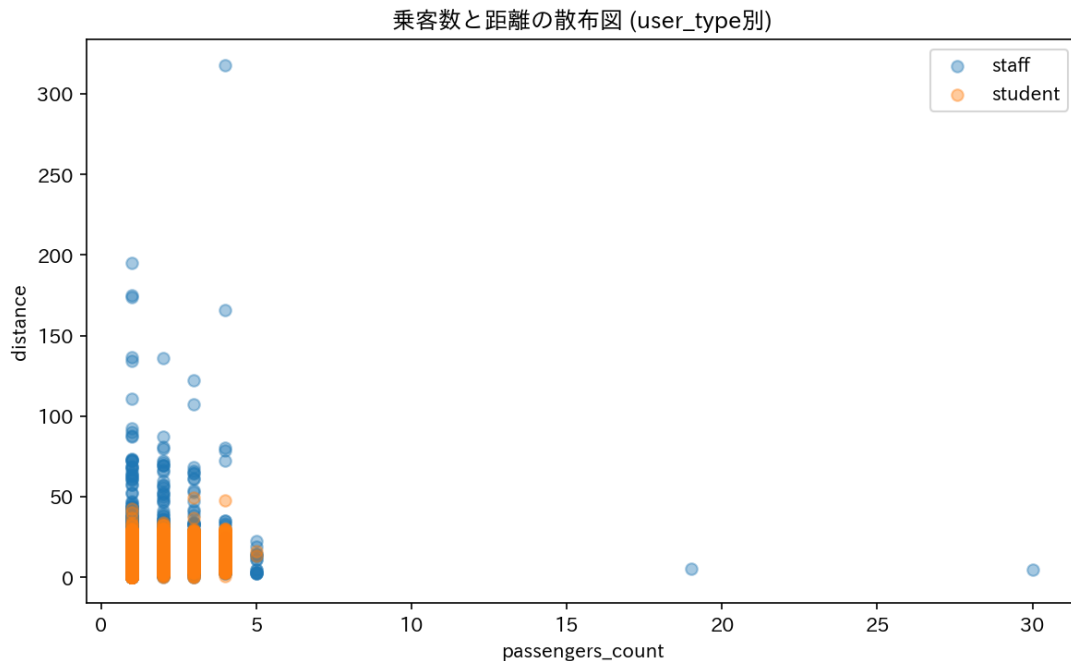


図 5: 乗車人数と移動距離の散布図（ユーザータイプ別）

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 人以上での利用は極めて少ない。乗車人数が増えるにつれてデータ点の数が急激に減少していることから、利用回数が減少していることが確認できる。

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）よりも距離の大きい範囲にデータ点が分布しており、長距離移動をする事例が相対的に多いことが分かる。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。

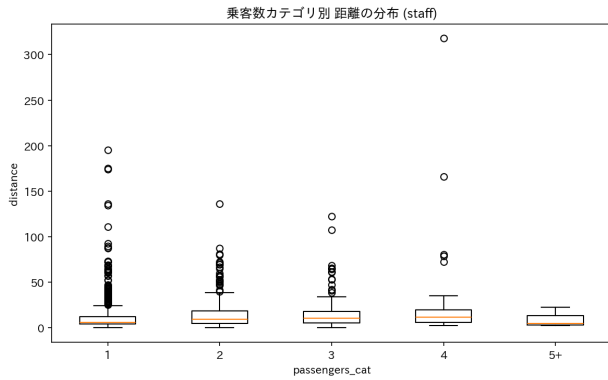


図 6: 教職員の乗車人数別移動距離

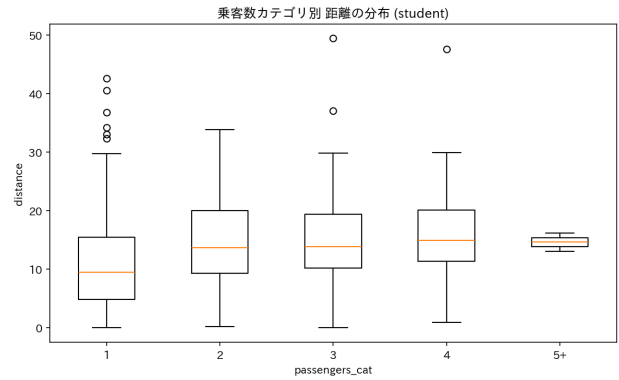


図 7: 学生の乗車人数別移動距離

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。目的地の数が増えるほど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詳細に分析した。

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 ユーザータイプ別の目的地数

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。

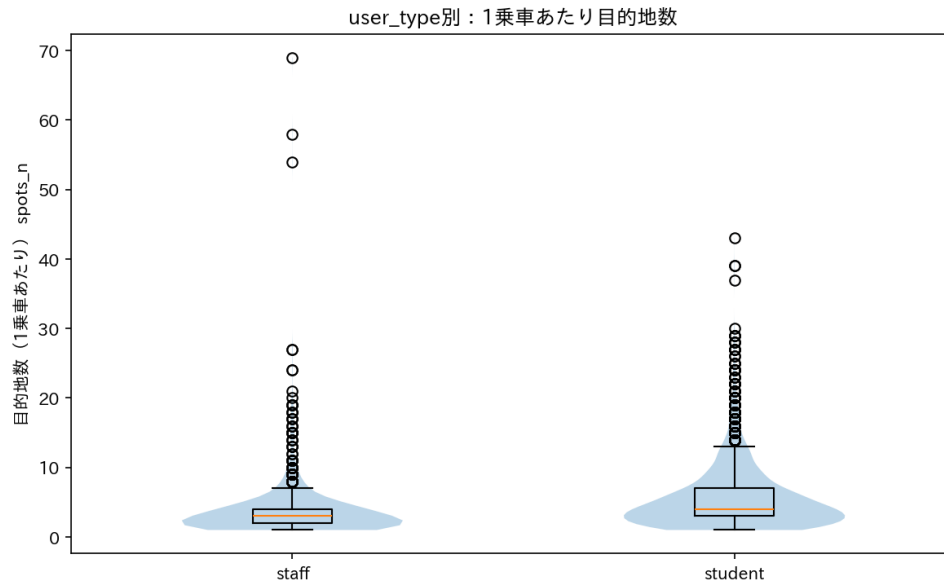


図 8: ユーザータイプ別の 1 乗車あたり目的地数の分布

バイオリンプロットとその中に重ねられた箱ひげ図から、教職員と学生で目的地数の分布に違いがあることが分かる。教職員の中央値は約 3 箇所であるのに対し、学生の中央値は約 5 箇所と、学生の方が 1 回の乗車で訪れる目的地数が多い傾向がある。

箱の広がりを見ると、学生の方が四分位範囲が広く、目的地数のばらつきが大きいことが確認できる。また、両ユーザータイプとも多数の外れ値が存在し、教職員で最大約 70 箇所、学生で最大約 43 箇所という極端なケースも見られる。これらは、長時間にわたって車を借りて多数の場所を巡った特殊な利用例と考えられる。

バイオリンの形状から、教職員は 2~4 箇所程度の利用が最も多く、比較的集中した分布を示している。一方、学生はより広い範囲に分散しており、少数の目的地から多数の目的地まで多様な利用パターンがあることが窺える。この違いは、学生の方が柔軟な行動パターンを持ち、1 回の乗車で複数の用事を済ませる傾向があることを示唆している。

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。

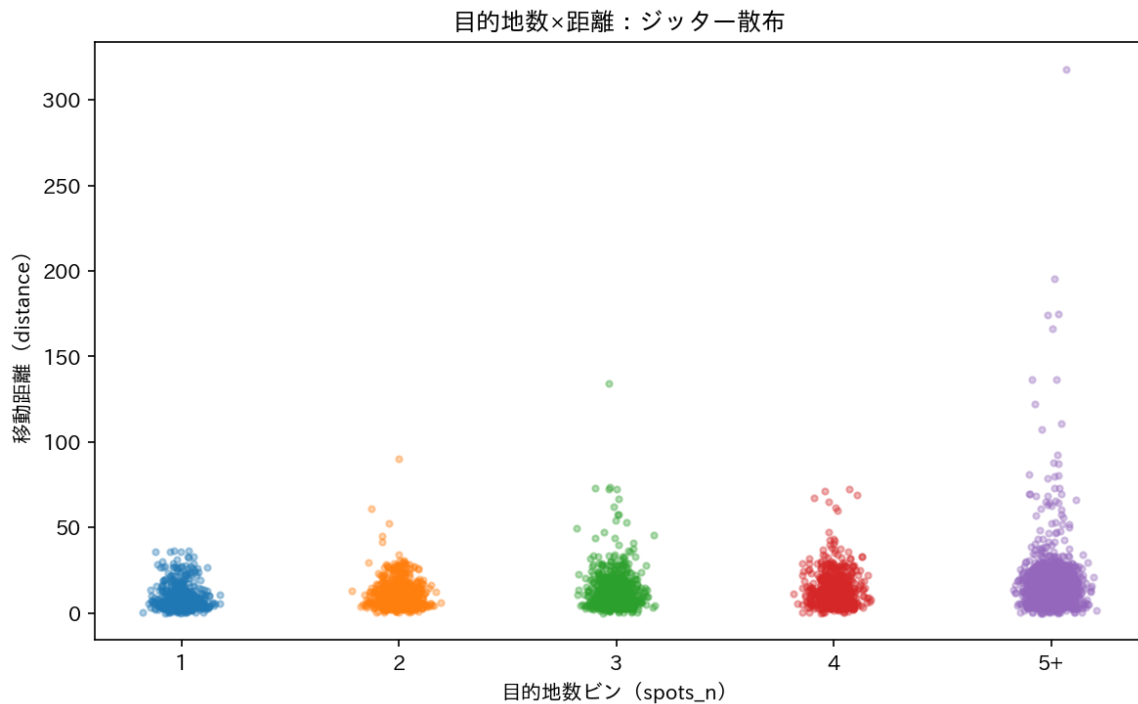


図 9: 目的地数別の移動距離分布（ジッター散布図）

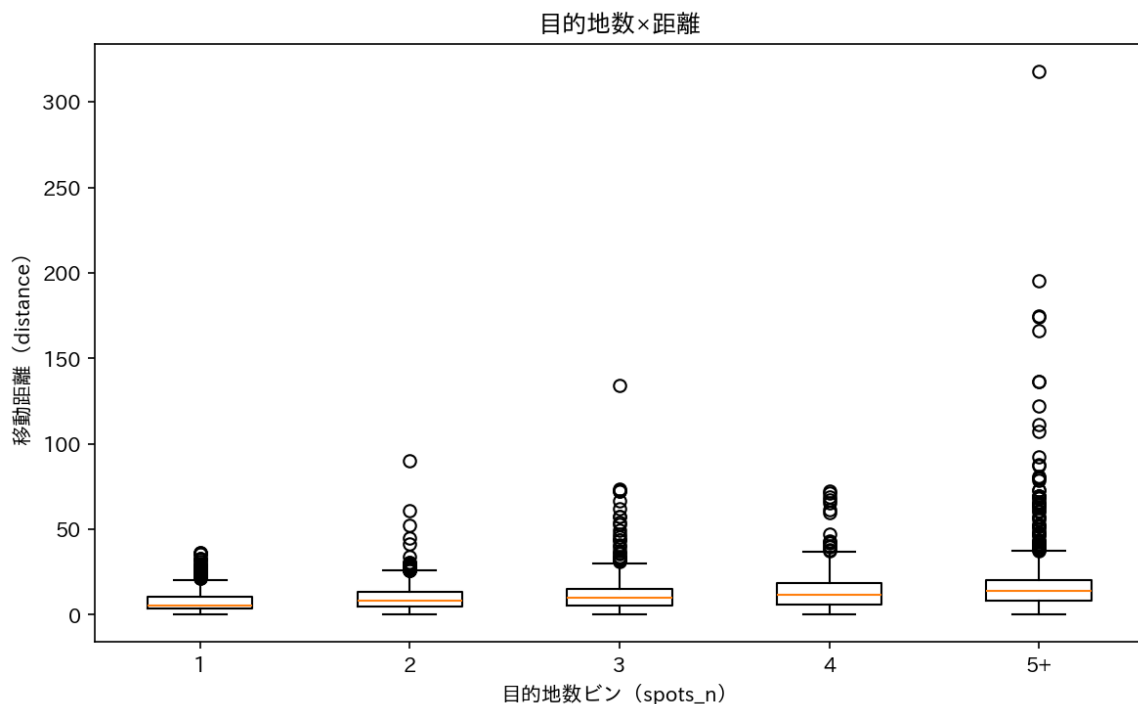


図 10: 目的地数別の移動距離分布（箱ひげ図）

ジッター散布図は、各目的地数カテゴリ内でのデータ分布を色分けして可視化したものである。この図から、目的地数が増えるにつれてデータ点全体が上方に移動しており、移動距離が増加する傾向が視覚的に明確である。

箱ひげ図から、より定量的な傾向を読み取ることができる。1 箇所のみを訪れる場合、中央値は約 10 程度と比較的短い。目的地が 2 箇所になると中央値は約 10~12、3 箇所です約 13、4 箇所です約 15 と段階的に上昇している。5 箇所以上になると中央値は約 18 まで増加し、目的地数が増えるほど移動距離も長くなることが明確に示されている。

各カテゴリの箱の大きさ（四分位範囲）を見ると、目的地数が増えるにつれてやや広がる傾向がある。これは、同じ目的地数でも、訪れる場所の位置関係や移動ルートによって移動距離に大きなばらつきが生じることを意味している。例えば、5 箇所以上の目的地を訪れる場合、効率的なルートを選べば比較的短い距離で済むが、非効率的なルー

トでは大幅に距離が伸びる可能性がある。

外れ値に関しては、全ての目的地数カテゴリで存在するが、特に5箇所以上のカテゴリで顕著である。最大で300を超える極端な長距離移動も記録されており、これらは遠方への出張や特殊な用途での利用と考えられる。

この結果は直感的にも理解しやすい。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自然である。ただし、同じ目的地数でも移動距離に大きなばらつきがあることから、目的地の選択や移動ルートが総移動距離に大きく影響することが分かる。

4.2.3 乗車人数と目的地数の組み合わせパターン

図11に乗車人数と目的地数の組み合わせパターンを、全体に対する割合として示す。

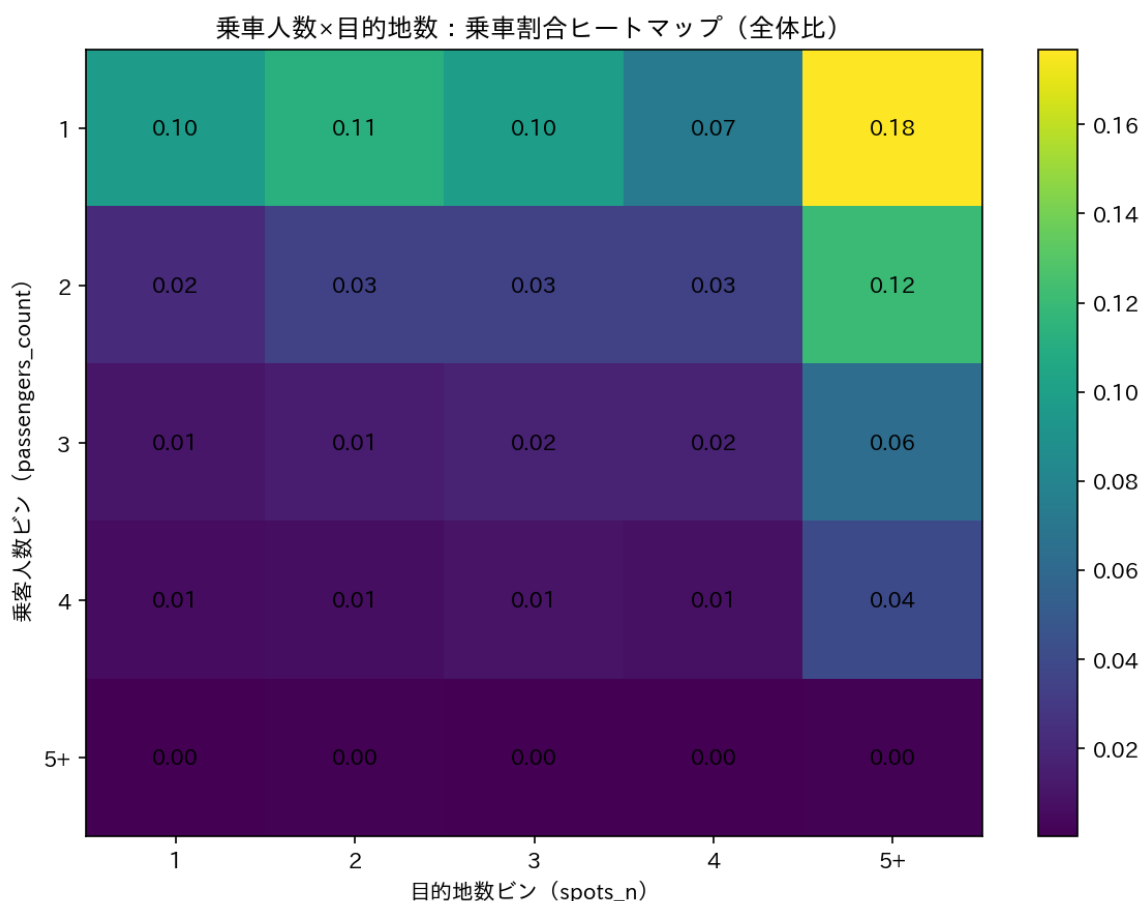


図 11: 乗車人数と目的地数の組み合わせパターン（全体比）

ヒートマップの各セルは、全体の乗車回数に対する割合を示している。最も高い割合を示しているのは「1人で5箇所以上の目的地を訪れる」パターンで、全体の18%を占めている。これは、単独でカーシェアリングを利用し、複数の用事を一度に済ませる利用形態が非常に多いことを示している。

次いで割合が高いのは「2人で5箇所以上」の12%である。1人での利用を見ると、「1箇所」「2箇所」「3箇所」「4箇所」がそれぞれ7～11%程度で比較的均等に分散している。これは、単独利用者の行動パターンが多様であることを示している。

乗車人数が2人の場合、1～4箇所の目的地を訪れるパターンはそれぞれ2～3%程度と低いが、5箇所以上になると12%に跳ね上がる。これは、2人での利用時には多くの目的地を巡る傾向があることを示唆している。

乗車人数が3人以上になると、全体的に割合が低下する。3人での利用は1～4箇所それぞれ1～2%、5箇所以上で6%である。4人での利用はさらに少なく、5箇所以上でも4%に留まる。5人以上での利用はほぼ0%であり、極めて稀であることが分かる。

この結果から、カーシェアリングは主に単独または2人での利用が中心であり、特に単独で複数の目的地を巡る利用形態が最も一般的であることが明らかになった。団体での利用は少なく、また団体に利用する場合でも目的地数は限定される傾向がある。

4.2.4 考察

分析3の結果から、以下の重要な知見が得られた。

まず、ユーザータイプによる違いとして、学生の方が教職員よりも1回の乗車で訪れる目的地数が多いことが明らかになった。学生は中央値で約5箇所、教職員は約3箇所であり、学生の方がカーシェアリングを効率的に活用して複数の用事を済ませる傾向がある。これは、分析1で見られた学生の利用時間の長さとも整合している。

目的地数と移動距離の関係については、明確な正の相関が確認された。目的地数が1箇所から5箇所以上に増えるにつれて、中央値が約10から約18へと段階的に上昇している。ただし、同じ目的地数でも移動距離には大きなばらつきがあり、目的地の配置や移動ルートを選択が重要であることが示唆された。

乗車人数と目的地数の組み合わせパターンからは、単独での多目的地利用が最も一般的な形態であることが分かった。全体の18%が「1人で5箇所以上」のパターンであり、これはカーシェアリングが個人の多様な用事を一度に済ませる手段として活用されていることを示している。

5 まとめ・今後の展望

本レポートでは、奈良先端科学技術大学院大学で実施さ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、ユーザータイプ別の利用パターン、移動距離・所要時間・乗車人数の関係、目的地数と移動距離の関係という3つの観点から分析を実施した。

本分析を通じて、カーシェアリングの利用パターンがユーザータイプによって大きく異なることが明らかになった。教職員は勤務時間内の短時間・業務目的の利用が中心であるのに対し、学生は長時間かけて多数の目的地を巡る柔軟な利用が主流である。利用回数も学生の方が2倍以上多く、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が示された。

移動距離に関しては、乗車人数と目的地数の両方が影響を与えることが定量的に確認された。

これらの知見は、カーシェアリングサービスの運用改善に活用できる。例えば、教職員と学生の利用時間帯が異なることを考慮した車両配置、学生の長時間利用に対応した料金体系の設計、多目的地利用に適した効率的なルート提案機能の実装などが考えられる。

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（レストラン、商業施設、公共施設など）や移動ルートの効率性、季節変動、天候の影響などを分析をしてみたいと思った。

なお、本レポートで使用した分析コードは以下のGitHubリポジトリで公開している。

<https://github.com/Kota-35/25-database>