

COVID-19による購買行動の時空間変化の分析

山口 公平¹ 田村 直樹¹ 庄子 和之¹ 浦野 健太¹ 米澤 拓郎¹ 河口 信夫^{1,2}

概要：近年の社会の情報化に伴い、様々なデータが得られるようになってきた。その中でもクレジットカードの取引履歴は、個人情報がわからないように加工された状態での提供が行われており、データを所持する企業による分析サービスの提供もみられる。しかし、そのような分析はグラフでの可視化や統計的な分析にとどまり、取引履歴に付随する店舗情報の活用はされていない。例えば、消費が行われた場所と消費動向の変化の関係などを分析することで、COVID-19の影響がどのように各地域に現れたか分かり、行政が地域ごとに適切な支援を行えるなどのメリットが考えられる。そこで、本研究では時空間に紐づいた購買情報から地域の特徴づけ（エリアモデリング）をする方法の検討を行う。また、エリアモデリングを通じて、2019年から2020年にかけての名古屋市における消費動向の変化を分析する。実験ではクレジットカードの取引履歴を用いて、名古屋市各地の消費動向とその変化を可視化し、取引回数の落ち込んだ地域は平均的な購入金額も減少するなどの消費動向の変化を捉えられた。また、消費動向の変化が似ている地域における滞在の変化と購買行動との関係についても議論した。

Analysis of Spatio-Temporal Changes in Purchasing Behavior by COVID-19

KOHEI YAMAGUCHI¹ NAOKI TAMURA¹ KAZUYUKI SHOJI¹ KENTA URANO¹
TAKURO YONEZAWA¹ NOBUO KAWAGUCHI^{1,2}

1. はじめに

2020年から続くCOVID-19の影響により、人間の行動には大きな制限がかけられ、以前とは異なる生活様式を迫られた。人込みを避け、外出を自粛するといった行動の変容は経済活動にも影響を及ぼしており、2020年におけるGDP成長率はリーマンショックを下回った[1]。各地域におけるCOVID-19の打撃を把握するためにも、消費動向の変化の分析はより重要になっている。消費動向の分析のためのデータセットは住民の購買行動を代表する必要がある。そうしたデータセットの1つとして、クレジットカードの利用履歴が挙げられる。実際に、クレジットカードを運営する日本クレジットビューローや三井住友カード株式会社などは、個人が特定できないようにした状態で取引履歴を提供、分析するサービスを行っている[2][3]。また、社

会の情報化に伴い、クレジットカードが使用された店舗、すなわち位置情報を取引履歴に紐づけることが可能となった。しかし、現状は位置情報は使われずグラフでの可視化や統計的な分析に留まる。取引履歴に紐づいた位置情報を用いれば、消費動向の変化が地域ごとに分析でき、例えばCOVID-19の影響の把握や地域への経済的な支援を行う際の指標の提示が可能である。また、地域の消費動向の変化をモデリングすれば、出店・退店の目安といったマーケティングへの利用や、地価の推定などができると期待できる。そこで、本研究では時空間に紐づいた購買情報を用い、ある地域において人はどのような消費を行うか、という視点で地域の特徴付けを行う。

地域の特徴付け手法は、庄子らが提案したArea2Vec[4]から着想を得た。Area2Vecは、GPSによって取得された位置情報データから、ある時刻から何時間ユーザが滞在したかという滞在情報を抽出し、その滞在のされ方から地域を特徴付ける手法である。我々はこのArea2Vecの考え方をもとに、地域をその消費動向の変化で特徴付け、分散

¹ 名古屋大学大学院 工学研究科
Graduate School of Engineering, Nagoya University

² 名古屋大学 未来社会創造機構
Institutes of Innovation for Future Society, Nagoya University

表現を得る手法を提案する．分散表現を利用することにより，地域の分散表現同士のベクトル演算や，異なるデータ間の連携が容易になる．また，類似した消費動向を持つエリアの地理的な特徴を捉えるために，k-means 法を用いてクラスタリングを行う．本研究では，消費動向にはクレジットカードの取引履歴を用いる．複数の様式のクレジットカード取引履歴で特徴付けを行い，その結果をもとに名古屋市における COVID-19 が引き起こした購買行動の時間空間変化を分析する．

分析の結果，2019 年と比べ，て 2020 年では名古屋の中心市街や大規模商業施設・レジャー施設が含まれるクラスタにおいて取引回数が減少したものの，各地で取引回数が増加したクラスタも確認できた．さらに，異なる特徴量に着目して得られたクラスタや滞在情報をもとにした，エリアの具体的なモデリングについて議論した．2019 年と 2020 年を比べて，夜間の取引回数も平均的な消費金額も減少したという特徴や，エリア内部での消費が多く行われていたという特徴などが検出可能であり，スーパーや居酒屋などの業種を超えた具体的な分析・モデリングを達成できた．本研究の貢献は以下のとおりである．

- Word2Vec の構造を用いて，消費動向の変化を捉える手法を提案した
- 学習によって得られた分散表現をクラスタリングし，本手法が地域の購買行動の傾向を捉えられることを検証した
- POI 情報や店舗情報を用いず，エリアに含まれる業種を推定した
- 複数のエリアにおける実際の滞在情報を可視化し，消費動向との関係について検証した

2. 関連研究

種々の特徴量に基づいた，統計的な地域のモデリングやクラスタリングの例として，Grekousis らは国勢調査や独自のアンケートに基づいて地域やその住民をクラスタリングした [5]．また，Singleton らはアンケートやオープンソースのデータを用いて，イギリスにおけるデジタル不平等性を可視化し，各特長量から住民をクラスタリング・モデリングした [6]．

COVID-19 の経済的な影響を分析した研究として，Ascani らはイタリアの感染状況を，経済基盤指数や人口比率，失業率などの統計情報を用いて推定する回帰モデルを提案した [7]．また，Firano らはモロッコでの実際のデータを元に，政府の予算やエージェントの労働時間，消費財の生産量などのパラメータを用いて国家単位の感染者割合や経済状況を導く新たなモデルを提案した [8]．しかし，これらの手法は様々な統計情報が必要であり，また大規模な調査が必要であるために地理的単位を大きくせざるをえず，購買情報に紐づいた店舗情報の活用には向かない．比較的小規

表 1: クレジットカードの取引履歴の期間及び取引回数

期間	取引回数 [件]
2019/04/06 ~ 04/19	55226
2019/07/01 ~ 07/14	65846
2020/04/06 ~ 04/19	41416
2020/07/01 ~ 07/14	60644

模な範囲での消費動向分析として，消費が行われた時間帯や平均購入金額，平日か休日かなどの特徴量に基づいて店舗や人を分析・モデリングする研究が存在する [9][10][11]．

一方，庄子らによって，教師なし学習を用いて地域をモデリングする手法である，Area2Vec が提案されている [4]．Area2Vec は，曜日・滞在開始時刻・滞在時間を用いて，人々の滞在の特徴に基づく地域の使われ方をモデリングする手法である．この手法の利点として，地域をその特徴を加味したベクトル（分散表現）として表せる点がある．分散表現であるために表現力が高いほか，類似度の計算や分散表現同士の演算が可能であり，ほかのデータとの連携も容易である．また，庄子らの実験では独自に定めた 50m × 50m という比較的狭いエリアでモデル化を行っているが，Area2Vec はエリアに割り振られた番号をもとに学習するため，データの位置情報の粒度に応じてエリアを定められるという点も挙げられる．そこで，本研究では Area2Vec の考え方に基づき，購買情報によって地域をモデリングする手法を検討する．

3. データセット

データセットには三井住友カード株式会社から提供された，名古屋市におけるクレジットカードの利用履歴を使用する．表 1 にデータの期間と取引回数を示す．利用履歴は個人や店舗の特定を防ぐために，以下のように特徴量がまとめられている．

- 各時刻における消費回数・金額・ユーザー数
- 業種ごとの消費回数・金額・ユーザー数
- 性別，年代，居住地ごとの消費回数・金額・ユーザー数
- 日付ごとの消費回数・金額・ユーザー数

これらのデータは地域メッシュに基づく 500m × 500m のエリアごとに集計されている．また，時間的な粒度に関しては，日付ごとの取引データは 1 日ごと，他のものは 2 週間ごとにまとめられている．そのため，何時にこの業種に対して取引が行われたといった，各特徴量の種類を超えるようなデータの掛け合わせはできない．

各エリアの消費動向と地図上での位置を比べると，ショッピングモールが存在するエリアの中でも，日用品の購入先として利用されるエリアや，近隣住民以外の人によって休日に多くの消費がされるエリア，COVID-19 の影響による消費動向の変化という点では，夜間の消費が落ち込んだエリアや，巣籠り需要により平均的な購入金額が増加したエ

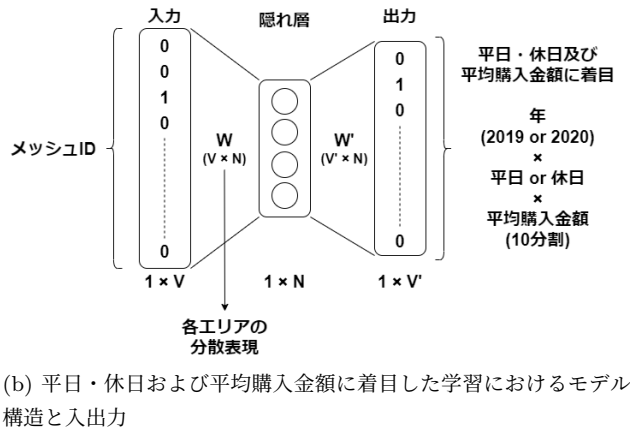
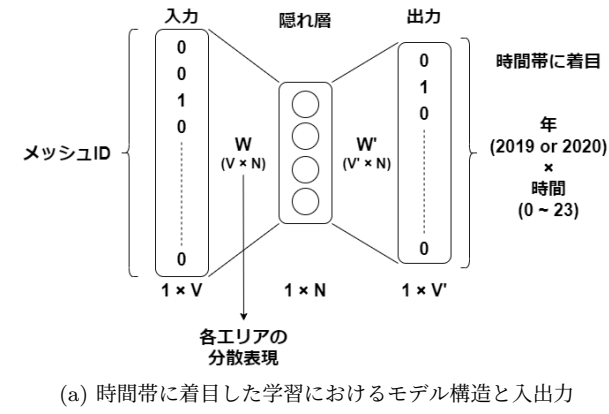


図 1: 提案手法のモデル構造と入出力

リアなどの違いがみられた。そこで、本研究では地域のモデリングのために、時間帯・平日と休日・平均購入金額という特徴量を用いて学習を行う。

4. 提案手法

4.1 購買情報による地域のモデリング

3章で述べたデータセットの特性により、以下の2つに分けて学習する。

- 時間帯に着目した学習
- 平日・休日および平均購入金額に着目した学習

また、5.2節では2つの学習により得られるクラスターを掛け合わせ、複数の特徴量を考慮した具体的なモデリングについて検討する。なお、学習はいずれの場合も学習回数を1000とし、隠れ層の次元数は4とした。図1に本手法のモデル構造と入出力を示す。学習ののち、入力層から隠れ層の重み W の各行を、エリアの分散表現として用いた。以下でそれぞれの学習方法の詳細を説明する。

時間帯に着目した学習

図2はすべての期間における、各時刻の取引回数である。このグラフは名古屋市における平均的な時間別の消費動向と捉えられる。しかし、名古屋市のそれぞれ単体のエリアでは図のとおりではなく、早朝に取引回数のピークを迎えるエリアや、深夜帯の取引が減少したエリアといった様々な消費動向とその変化が見られた。さらに、2019年と

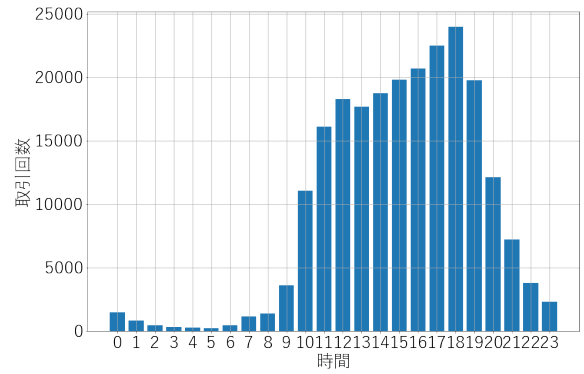


図 2: すべての期間における名古屋市の時間帯別の取引回数

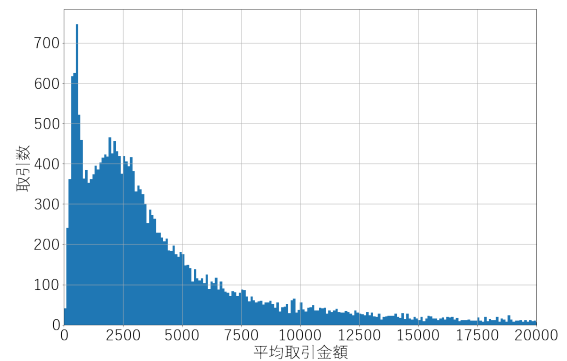


図 3: すべての期間における名古屋市の平均購入金額の分布

表 2: 平均購入金額の範囲及び各クラスの総取引回数

平均購入金額の範囲 [円]	取引回数 [件]
～ 600	2216
600 ～ 1200	2809
1200 ～ 1800	2420
1800 ～ 2400	2625
2400 ～ 3000	2385
3000 ～ 3800	2467
3800 ～ 5200	2770
5200 ～ 7700	2527
7700 ～ 14000	2448
14000 ～	2524

2020年の消費動向の変化という点では、変化があまり見られないエリアや深夜帯の取引回数が減少したエリア、昼の取引回数が増加したエリアなどが確認できた。そこで、2019年から2020年にかけての時間帯ごとの消費動向の変化が似ているエリアを抽出することを目的として、図1aのようにデータを整形し、学習を行う。入力には分割された名古屋市の各エリア（メッシュ）に振られたメッシュIDに対応する one-hot ベクトル、出力は2019年と2020年及び24時間に対応する48次元の one-hot ベクトルとする。こ

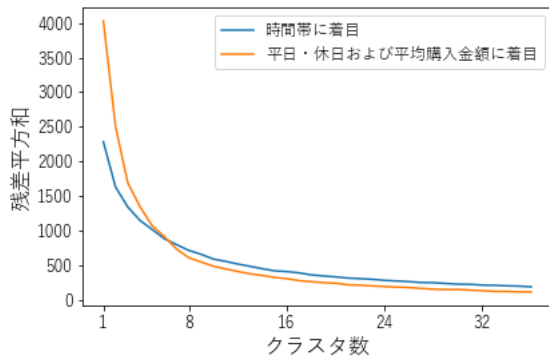


図 4: クラスタ数を変更した際の残差平方和

の入力と出力のセットは取引が 1 回行われたことを表すので、各時刻に行われた取引回数分訓練データとして追加する。この学習によって、エリアに応じた時間帯ごとの消費動向の COVID-19 による変化が反映された分散表現が得られると期待される。

平日・休日と平均購入金額に着目した学習

こちら、平日・休日に関係なく 1000 円程度の消費が多いエリアや、2020 年において休日の取引回数が特に減少したエリアなどの消費動向とその変化が見られた。平日・休日と平均購入金額の消費動向が似ているエリアを抽出することを目的として、同様に図 1b のようにデータを形成し、学習を行う。入力にメッシュ ID に対応する one-hot ベクトル、出力は 2019 年と 2020 年、平日と休日、表 2 のように離散化された平均購入金額に対応する 40 次元の one-hot ベクトルとする。離散化は各クラスの取引回数が均等になるように 10 分割した。図 3 に平均購入金額のヒストグラム示す。この学習も同様に、エリアに応じた平日・休日および平均購入金額の COVID-19 による変化が反映された分散表現が得られると期待される。

5. 実験と評価

本章では、学習で得られた分散表現をクラスタリングした結果を示し、その分析を行う。また、複数の学習方法で得られたクラスタリング結果を掛け合わせることで、より詳細な地域のモデリングが可能であるか検討する。更に、購買情報でクラスタリングされたあるクラスタに属するエリアにおける滞在情報の変化を可視化し、エリア内の滞在及び消費動向と滞在の関係性を分析する。

5.1 名古屋市内における消費動向の変化

学習によって得られた分散表現を k-means 法によりクラスタリングし、各クラスタの消費動向を分析する。クラスタ数はエルボー法により定めた。図 4 に示すとおり、クラスタ数 8 付近で残差平方和 (SSE) の減少が緩やかになることを確認したので、クラスタ数は 8 とした。

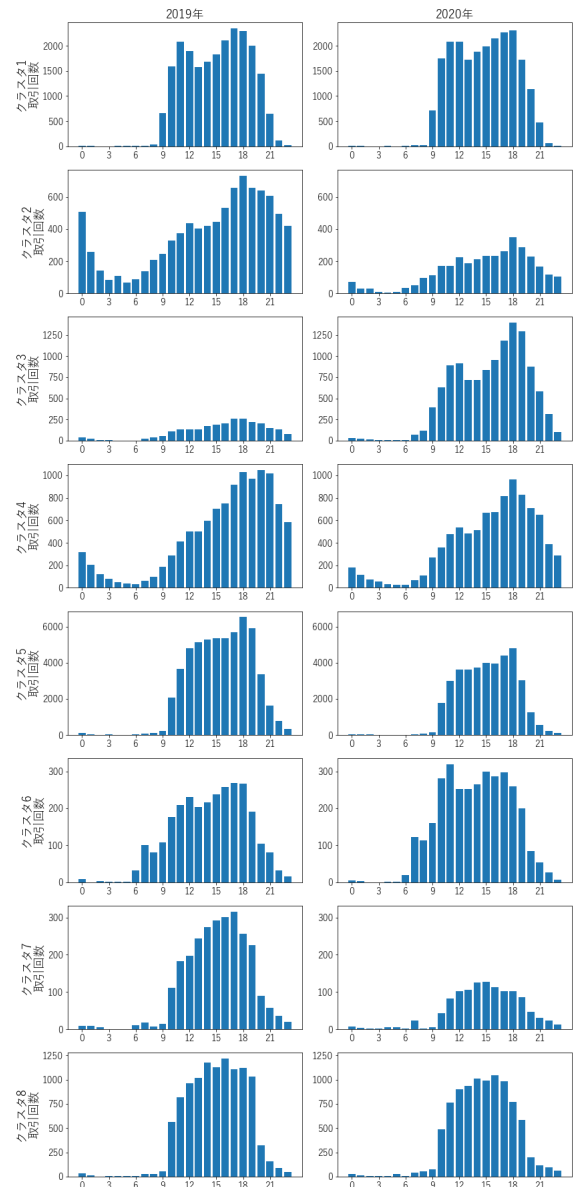


図 5: 各時間帯の取引回数の分布

時間帯に着目した学習

図 5 と図 6 は、取引が行われた時間帯の変化に着目した学習によって得られた分散表現のクラスタリング結果およびクラスタごとに地図上に色分けしたものである。図 5 の左列は 2019 年の、右列は 2020 年の時間帯別の取引回数のグラフであり、縦軸が取引回数を、横軸が時間帯を表す。

クラスタ 1, 4, 6 は、2019 年から 2020 年にかけて取引回数の時間帯別の分布の変化が小さかったクラスタであり、御器所・堀田・八事などの中心市街ではないが地方の消費の中心地であるエリアに見られた。クラスタ 1 は深夜帯の消費があまり見られず、12 時や 19 時付近に消費のピークが見られることからスーパーやレストランが多く存在するクラスタ、クラスタ 4 は夜間ほど取引回数が多くなるため、居酒屋などが多く存在するクラスタであると考えられる。クラスタ 3 は 2019 年と 2020 年を比べて、取引回数が増加

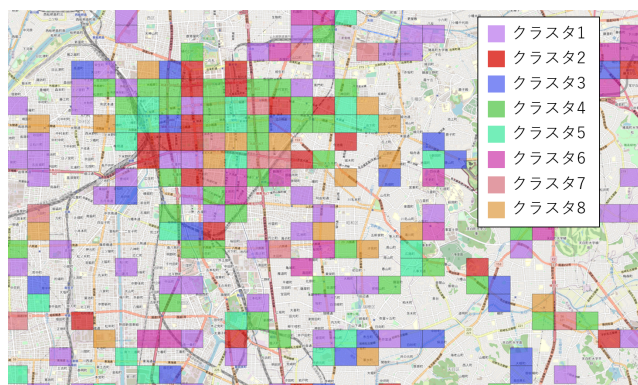


図 6: 時間帯に着目した学習によるクラスタ

したクラスタであり、名古屋市内の各地で見られた。クラスタ 2, 5, 7, 8 は取引回数が減少したクラスタである。これらのクラスタは、名古屋駅近辺や上前津、栄などの中心市街や、大規模商業施設・レジャー施設があるエリアに見られた。クラスタ 5, 7, 8 は取引回数の減少の程度に違いはあるものの、10 時から 20 時の時間帯に多くの消費がされていることから主にスーパーやショッピングモールが多く存在するクラスタ、クラスタ 2 は深夜帯における取引回数が特に顕著に減少しているため、居酒屋やバー、クラブなどがそのエリアの中心的な消費対象であるクラスタであると考えられる。

平日・休日および平均消費金額に着目した学習

図 7 と図 8 は、平日・休日および平均消費金額に着目した学習によって得られた分散表現のクラスタリング結果およびそれを地図上で色分けしたものである。左側の 2 列が 2019 年の、右側の 2 列が 2020 年であり、縦軸が取引回数を、横軸が離散化された平均購入金額を表す。

クラスタ 1 は 2019 年と 2020 年であり消費動向が変わらなかったクラスタである。名古屋市内にまばらに分散しており、このクラスタに属するエリアの多くにスーパーマーケットなどの商業施設が見られた。スーパーマーケットにおける 1 回あたりの平均購入金額は約 2000 円という調査があるため [12]、このクラスタは日用品の購入先だけではなく、近隣地域の外出先としても使用されていると推測する。クラスタ 2, 3, 7 は 2020 年において取引回数が増加したクラスタである。いずれのクラスタも 3000 円までの消費が増えており、名古屋市の全域にかけて分布することから、周辺地域の日常的な消費の先として使用されている。クラスタ 3 は、2019 年と 2020 年ともに平日と休日の消費動向が異なる。平日に低価格の消費が行われていることから、近隣の職場による需要であると考えられる。また、クラスタ 7 は特に消費の増加が顕著であるため、COVID-19 によって新たにこのクラスタに属するエリアで消費を行うようになったクラスタであると推測される。

クラスタ 4, 5, 6, 8 は消費が落ち込んだクラスタである。全体の傾向として、2019 年よりも 2020 年の方がより

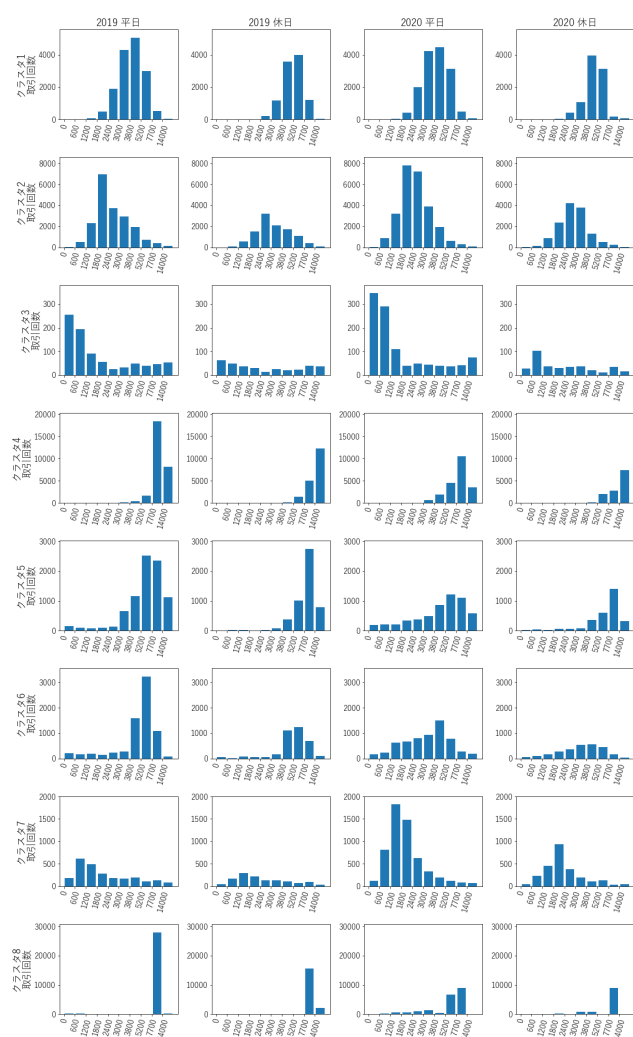


図 7: 各価格帯の取引回数の分布

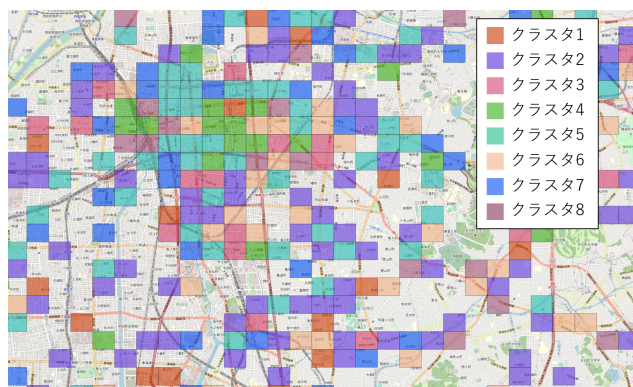


図 8: 平日・休日および平均購入金額に着目した学習によるクラスタ

低い価格帯の消費を行うようになった。クラスタ 5, 6 は 2020 年における平均購入金額の低下が特徴的なクラスタであり、名古屋市の全域に集中していた。クラスタ 4, 8 は約 8000 円以上の消費が顕著に多いエリアであり、名古屋市の中心市街である名古屋駅や栄駅周辺に集中的に分布していたほか、近隣地域の外出先として利用される星が丘や

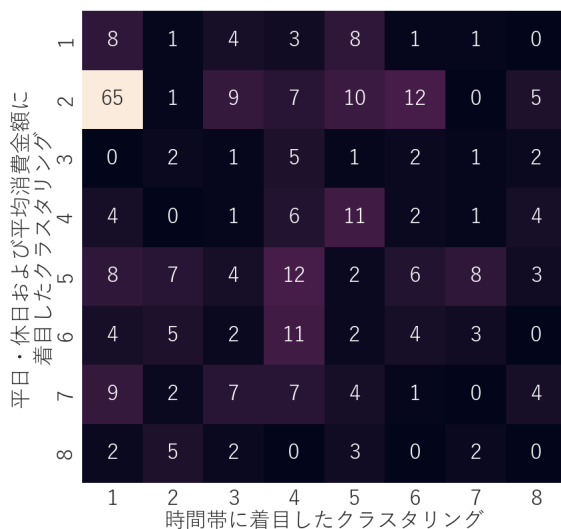


図 9: 各エリアが属するクラスタの分布

千種の駅前などでも見られた。

5.2 具体的な地域のモデリング

本節では、上記の2つのクラスタ間の相関について検討する。図9は各エリアが属するクラスタの分布であり、各マスの値はそれぞれのクラスタに対応するエリアの数である。2つのデータで有効な取引履歴が異なるため、どちらにも含まれるエリアのみを対象とした。

図10は時間帯に着目した学習におけるクラスタ1かつ、平日・休日および平均購入金額に着目した学習におけるクラスタ2であるエリアの分布である。これらの地域は12時と18時に多く消費が行われ、平均的に2000円付近の消費が多く、2019年と2020年を比べて時間帯も平均購入金額もあまり変化がなかったエリアと推測できる。また、図11は、時間帯に着目した学習におけるクラスタ4かつ、平日・休日および平均購入金額に着目した学習におけるクラスタ5であるエリアの分布である。これらは20時前後で消費のピークを迎え、2019年と2020年で時間帯別の取引回数では18時以降の取引回数が減少しており、平均購入金額は2019年が7000円付近の消費が多かったのに対して2020年では7000円付近の消費が落ち込んだエリアと推測できる。このように、本手法によって得られる複数の視点からのクラスタを用いて、地域の消費動向の具体的なモデリング・分析が可能である。

購買行動と滞在の関係

本手法によるクラスタリングの評価のため、株式会社プログウォッチャーより提供頂いたGPS滞在情報データセットを用いて、時間帯に着目したクラスタリング結果の各クラスタに属するエリア内の滞在の変化及び消費動向との関係について検討する。クレジットカードの取引履歴と

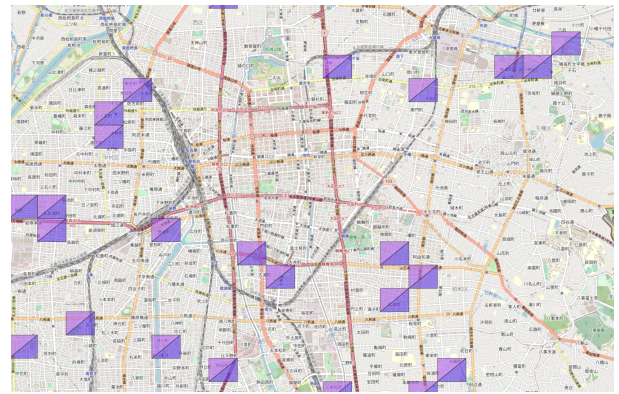


図 10: 平均購入金額に着目した学習におけるクラスタが2、時間帯に着目した学習におけるクラスタが1であるエリア

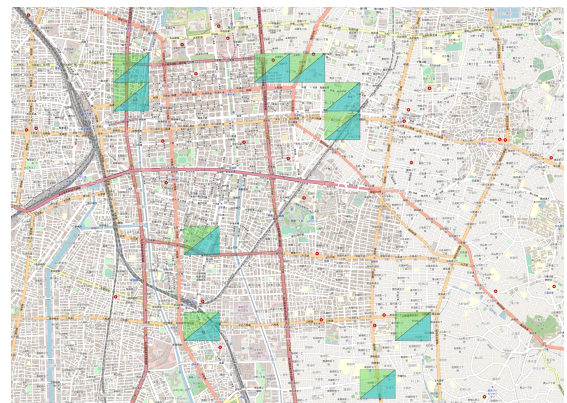


図 11: 平均購入金額に着目した学習におけるクラスタが5、時間帯に着目した学習におけるクラスタが4であるエリア

同じ期間の滞在データを使用し、滞在が取られた地点を取引履歴のエリアに分配した後、各クラスタごとにグラフに可視化した。図12は図5におけるクラスタ2とクラスタ4に属するメッシュ内の、2019年と2020年の滞在である。縦軸は確認された滞在件数の累計、横軸は30分ごとの時刻であり、グラフの各色は滞在時間の長さに対応する。クラスタ2と4の消費動向としては、どちらも2019年は夜間の取引が特に多かったが、2020年においてクラスタ2は大幅に取引回数が減少し、クラスタ4は大きな変化はなかった特徴が挙げられる。滞在においても、どちらも6時間以上の長期滞在の割合が増加しており、2時間以内の短期滞在は半分程度に減少している。よって、クラスタ2は消費者としてほかのエリアから来た短期滞在が主だったのに対し、クラスタ4はエリア内の住人による内需的な購買行動だったと推測できる。このように、本手法によって得られる複数の視点からのクラスタや滞在情報によって、地域の時空間的な消費動向のモデリングや分析が可能であることを確認できた。

6. まとめと展望

本研究では、エリアで行われた購買情報を用いて地域の

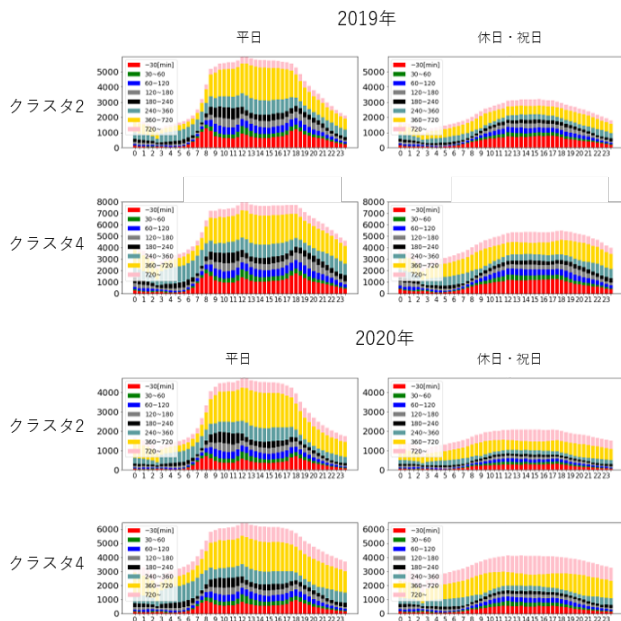


図 12: クラスタ 2 とクラスタ 4 の 2019 年と 2020 年における滞在開始時刻と滞在時間

特徴づけをする手法を提案した。本手法により、着目した特徴量に対してそれぞれのエリアの消費動向の特色が現れるような学習が行われたことを確認した。さらに、複数のクラスタリング結果や滞在情報を取り入れることで、地域の店舗の使われ方のより具体的な分析が可能であることを確認した。2019 年と 2020 年を比べて、夜間の取引回数も平均的な消費金額も減少したという特徴や、エリア内部での消費が多く行われていたという特徴などが検出可能であり、スーパーや居酒屋などの業種を超えた具体的な分析・モデリングを達成できた。

展望として、本研究で行ったエリアモデリングの分析のほかに、POI(Point Of Interest) 情報や土地利用動向調査、2021 年のクレジットカードの消費履歴を用いて、本手法のさらなる検証・改良を行いたい。また、Area2Vec を利用する利点の一つとして、得られた分散表現同士の演算ができる可能性が挙げられる。本研究ではエルボー法に基づいてクラスタ数を 8 としたが、クラスタ数を増やすことにより各クラスタが持つ分散表現上の意味を分離でき、深夜帯の消費が落ち込んだことを表す分散表現などの意味の基底的な分散表現を抽出できる可能性がある。こうしたベクトル演算を用いたエリアの分散表現自体の検証も今後の課題である。

謝辞

本研究の一部は、JST CREST (JPMJCR21F2), NICT 委託研究 (222C0101) 科学研究費補助金 (22H03696) に支援されています。また、データ提供にご協力いただきました三井住友カード株式会社、株式会社プログウォッチャーに感謝いたします。

参考文献

- [1] 内閣府. 2020 年度国民経済計算 (2015 年基準・2008sna) : 経済社会総合研究所 - 内閣府. https://www.esri.cao.go.jp/jp/sna/data/data.list/kakuhou/files/2020/2020_kaku-top.html, 3 2022. (Accessed on 05/06/2022).
- [2] 株式会社ナウキャスト, 株式会社ジェーシービー. Jcb 消費 now. <https://www.jcbconsumptionnow.com/>. (Accessed on 05/13/2022).
- [3] 三井住友カード株式会社. キャッシュレスデータでマーケティング課題を解決「custella」. <https://www.smbc-card.com/camp/custella/index.html>. (Accessed on 05/13/2022).
- [4] 庄子和之, 青木俊介, 米澤拓郎, 河口信夫. 大規模ユーザの滞在情報に基づくエリアの特徴付けと covid-19 による影響分析. 情報処理学会論文誌, Vol. 62, No. 10, pp. 1644–1657, oct 2021.
- [5] George Gekousis and Thomas Hatzichristos. Fuzzy clustering analysis in geomarketing research. *Environment and Planning B: Planning and Design*, Vol. 40, pp. 95 – 116, 2013.
- [6] Alex Singleton, Alexandros Alexiou, and Rahul Savani. Mapping the geodemographics of digital inequality in great britain: An integration of machine learning into small area estimation. *Computers, Environment and Urban Systems*, Vol. 82, p. 101486, 2020.
- [7] Andrea Ascani, Alessandra Faggian, and Sandro Montresor. The geography of covid-19 and the structure of local economies: The case of italy. *Journal of Regional Science*, Vol. 61, No. 2, pp. 407–441, 2021.
- [8] Zakaria Firano and Filali Adib Fatine. The covid-19: macroeconomics scenarii and role of containment in morocco. *One Health*, Vol. 10, p. 100152, 2020.
- [9] 田口東, 高松瑞代. 通勤時間帯の駅構内における購買行動の数理モデル. 日本応用数理学会論文誌, Vol. 27, No. 2, pp. 147–161, 2017.
- [10] 土井千章, 小西哲平, 中川智尋, 片桐雅二, 稲村浩, 太田賢. 店舗クラスタを用いた購買行動に関するライフスタイルの推定. In *DEIM Forum*, pp. F3–5, 2014.
- [11] 山田浩喜. 百貨店における顧客の来店行動メカニズムについて: 階層ベイズボアソン回帰モデルによる来店行動分析. *Review of economics and information studies*, Vol. 20, pp. 1–20, 2020.
- [12] 吉原慶. 消費増税とスーパーの利用状況に関する調査 - 市場調査・マーケティングリサーチ会社のアスマーク. <https://www.asmarq.co.jp/data/ex201912tax-increase/>, 12 2019. (Accessed on 05/17/2022).