

SNS と口コミデータによる魅力的なカフェの抽出とその空間分布

鈴木英之*

Extraction of Attractive Cafes and Spatial Distribution Using on Word-of-Mouth Data

Hideyuki SUZUKI*

Abstract: Continuing from the previous report, this report is an attempt to visualize and show the spatial location of the city's attractions using information obtained from word-of-mouth and review data written by visitors. Specifically, we clarified the spatial distribution of attractive cafes as a third place that is thought to influence the attractiveness of the city. Using word-of-mouth and review data, we classified cafes into subcategories and clarified the characteristics of each. As a result, we found that areas where cafes with low ratings are concentrated are also areas with few characteristic stores.

Keywords: 中心市街地活性化 (city center activation), クリギング (Kriging), 分散表現 (Distributed representation), t分布型確率的近傍埋め込み法 (t-SNE), プレイス・ブランディング (Place branding) 口コミデータ (word-of-mouth data)

1. 背景と研究目的

本報は前報に引き続き、街の魅力を計量化して可視化するための試行を目的とする。具体的には「魅力」的なカフェの空間分布を手掛かりに街の魅力について考察する。

前報では全国に所在する 1,142 の商店街に対する口コミを自然言語処理にかかる機械学習的な方法を用いて、ベクトル化し、多様な商店街イメージの分類や特徴の抽出を試みた。しかし、商店街に対する口コミは様々なレベルにおける種々雑多な体験の寄せ集めで構成されており、街の魅力を一枚の知覚地図 (パーセプションマップ) として作成するには必ずしも適切ではない。今回は「憩う」、「集う」、「語らう」という共通の機能を有した業種店舗 (カフェ・喫茶店) に限定した分析とする。

対象としてカフェを選択した理由として、口コミによるデータベース構築上、存在地点の地理的な偏在性やデータ収集の容易さのほかに、下記の3点が挙げられる。

1. 近年盛んになっているエリアマネジメント (リノベーションまちづくり) 等、遊休不動産の再活用事業では地区への来街集客やコミュニティ育成の面でカフェの新規出店が大きな役割を果たしており、魅力的なカフェの有無が街の魅力づくりに大きく影響していると考えられた。
2. ガイド誌等のカフェ紹介は、付近の他業種店や散策街路等と共にエリア内の回遊を前提に記述されることが多い。カフェ利用の大半は、目的来店ではなく付近での所要ついでや回遊の中でなされるものであり、街の魅力はカフェへの来店や利用のされ方と関連しているものと考えられた。
3. 上記に関連して、カフェ創業希望者の中で既存同業種店が多い特定の評判エリアで出店を希望する者は少なくない。街の魅力が利用者の集客につながるのみならず、意欲

的な新規事業者の誘致につながっている側面が推察される。カフェ店舗の魅力度の高低は、エリアの固有性や街の魅力から大きく影響を受け、またその逆に、店舗の存在が街全体の魅力に影響を及ぼしていると想定された。

以上より、本研究では、カフェ店舗の魅力計量化を以って街の魅力の可視化を試行することとした。

2. 先行研究

Adelfio et al. (2020) は街の魅力形成に大きな意味を持つカフェ等を含むサードプレイスを口コミ (GoogleAPI) のみを用いて、そこから抽出することが可能であることを確認した。

Reuschke et al. (2021) は、街の魅力形成に大きな意味を持つカフェ等を含むサードプレイスの主要な利用活動者であるクリエイティブクラス (コンサルタントやデザイナー等) による位置情報付きツイートの空間分布によって、創造性に関わる街の魅力の空間分布の可視化を試行した。

Van der Zee et al. (2020) は観光地域におけるカフェやレストランの口コミデータを用いたエリア評価を行った。評価の高い店舗が集中的に存在するホットスポットエリアやその反対に評価が低いレストランが集中的に存在するコールドスポットエリアの存在を明らかにし、今後のエリア開発における投資判断に対して示唆を提供した。

本報では、個別のカフェの属性 (評判や特徴) に空間的な偏在性が存在し、そのことがエリアの魅力価値に影響しているとする Van der Zee et al. (2020) の説を踏襲する。つまり魅力的なカフェの所在地が街の魅力エリアであるとみなして分析と考察を進める。

* 正会員 合同会社ファイン・アナリシス (Fine Analysis, LLC.)

〒545-0021 大阪府大阪市阿倍野区阪南町 2-24-12 Tel : 06-7503-9426 E-mail : suzuki@finea.co.jp

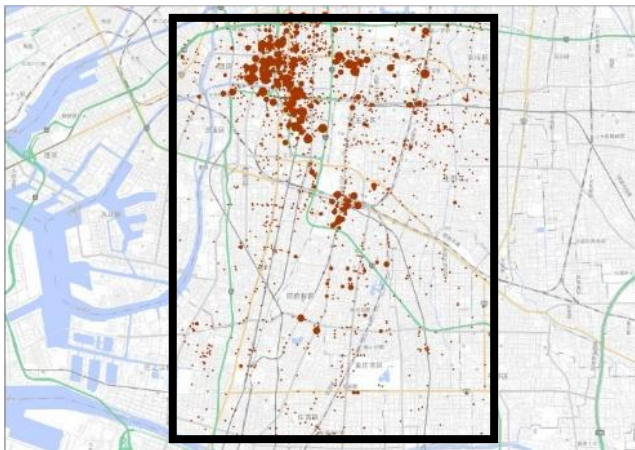


図1 分析対象エリア



図2 左：店舗密度（赤が高密度）、右：平均評価スコア（赤が高得点）

3. 方法

本報が収集したデータは国内レビューサイト「食べログ」(https://tabelog.com/)に掲載されたカフェ及び喫茶店に対する口コミである。図1の矩形エリア内に存在する4,277店のカフェ及び喫茶店に対して書き込まれた104,651件の口コミを利用した。個別の口コミは店舗ごとに連結した文書とし、Doc2vec:MIKOLOV(2014)を利用してそれぞれ200次元の分散表現を獲得した。口コミテキストの他に口コミ数や評価スコア(5点満点)等も利用する。

表1 クラスタに含まれる店舗名と特徴的な単語群(一部抜粋)

CL1 に含まれる主な店舗 旧ヤム邸 空堀店 サッチェズカリー ガネーシュー m 北浜店 ノムソン カリー モリ商店 ル コロン wocca rocca curry and tenroku cafe 黒岩珈琲 バザール スパイスカレーニューヤンティカ	CL12 に含まれる主な店舗 ドッブルベ・ボレロ 大阪本町店 バティスリー アクイユ マルナカ菓子店 バティスリー ショコラトリー オーディネール PHILLO&CO. ケントハウス ルビノー 玉造店 ボアール 帝塚山本店 バティスリーセボン お菓子のアトリエ ハンブルグ 十三本店
CL1 店舗群と類似するベクトルを持つ単語 キーマ カリー NONVEG マドン ラム・オレオパフェ VEG レモンパルシュガー キーマカレー ビスクグラタン スパイシーキーマカレー	CL12 店舗群と類似するベクトルを持つ単語 アチューゲ エクレア ムースリース フランポワーズ POIRE グリエールチーズ ブリュレタルト カフェノワゼット クリームソースメイン タルトフロヴァンサル
CL2 に含まれる主な店舗 クレープリー・スタンド シャンデレール 梅田店 CREPERIE Le Beurre Noisette クレープリー・アルション バーム 福島本店 サイドトリップ E S T 店 ロカンダ PALM Cafe&Creperie 茶屋町店 クレープバル&カフェ天神橋jiji 天満天六店 ブラッセリー マセナ Chabo	CL16 に含まれる主な店舗 バンデュース 石窯パンの店 ゴッホ プランジュリー&カフェグウ サク キッチン マルルー プランジュリー レ プティ アンジュ フードスケープ ハツタツ 天六店 プランジュリー フルティガラ 大阪店 PARTENAIRE
CL2 店舗群と類似するベクトルを持つ単語 ジャンデレール ガレット Peach バルフェ・プリン・ア・ラ・モード ナージュ シナモン・エクストラバージンオイル ブルターニュ レット アチューゲ バルフェ・イチゴ	CL16 店舗群と類似するベクトルを持つ単語 バンオショコラ ウィンナークロワッサン スバゲティボンゴレビアンコ チャンクシヨコ ライ麦 ノア・オランジュ ジューシージャーマンソーセージ フォウテ プレツェル フロマーズクッキーブラウニー

200次元でベクトル化された口コミは、t-SNE法:VAN DER MAATEN(2008)によって2次元に縮約した後、クラスタリングによって分類し、そのそれぞれの特徴(イメージや評価)を確認するとともに、小地域(駅勢圏)別に下位カテゴリーの構成等を確認することで、魅力的なカフェの属性的な特徴等を考察する。

4. 結果と考察

カフェ・喫茶店の多くは駅前型の立地であり、それらの所在地はターミナル3駅(なんば駅、心斎橋駅及び天王寺駅)周辺において著しい集中が確認される(図2左)。それに対して、店舗に対する評価スコアの地域平均の分布はエリア内において多峰的である(図2右)。交通の要衝とは言えない阿倍野区、住吉区および東住吉区内においても高評価店舗の集中地が存在するように見える。

店舗群をクラスタリングした結果、16のクラスター(以下CL)に分類することができた(店舗数が最小のCL:CL番号16;n=143,最大のCL:CL7;n=446)。

それぞれのCLは口コミに書き込まれた特徴的な単語やフレーズにおいて類似する店舗群に分類されていることが確認された。それぞれのCL店舗群のベクトルと類似するベクトルを持つ単語群を表1に掲げる。それぞれのCL店舗群に代表的な提供メニューやサービスの特徴を表す語群が並ぶ。例えばCL1の店舗群と類似度が高いベクトルを

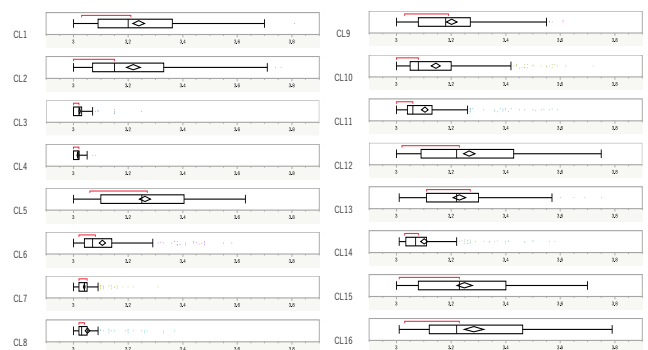


図3 クラスタ別店舗スコア分布

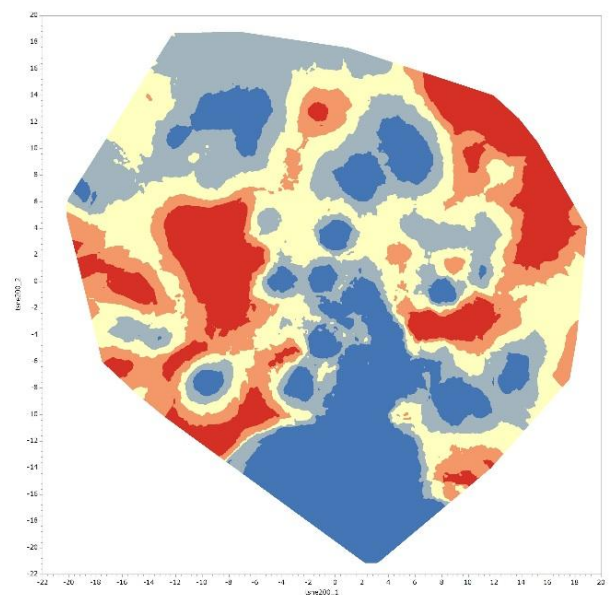
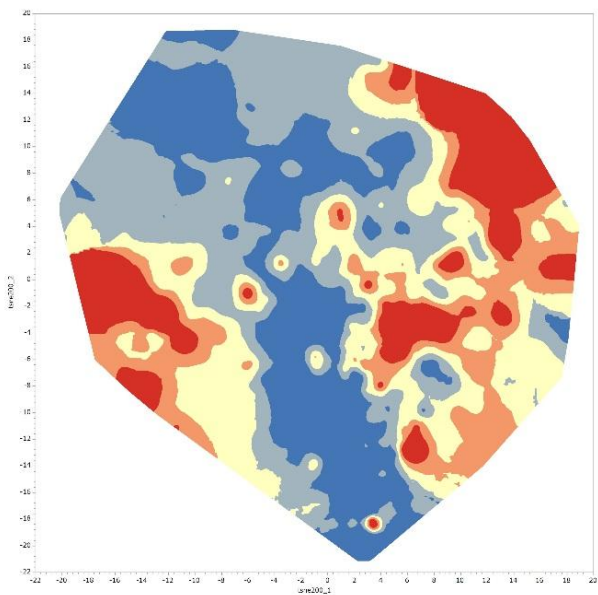
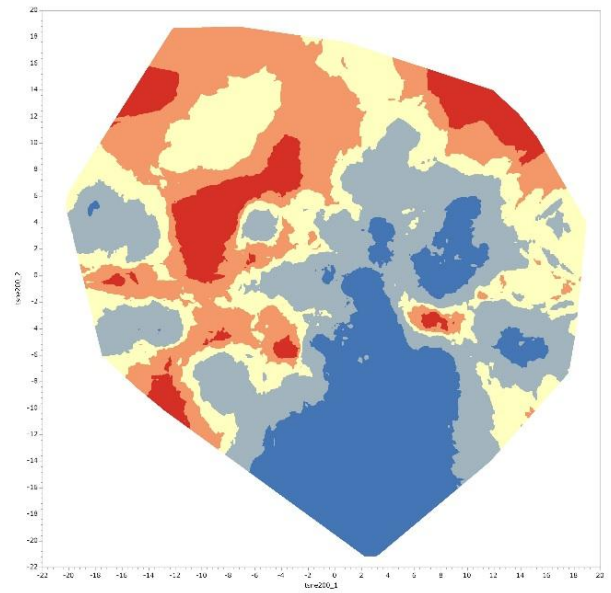
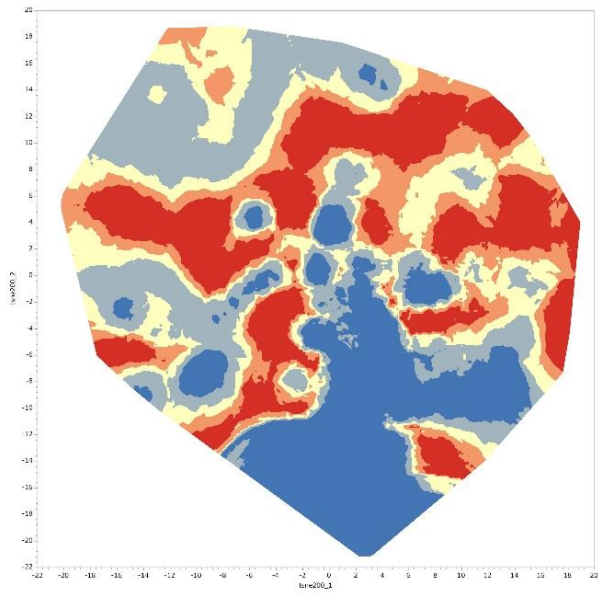
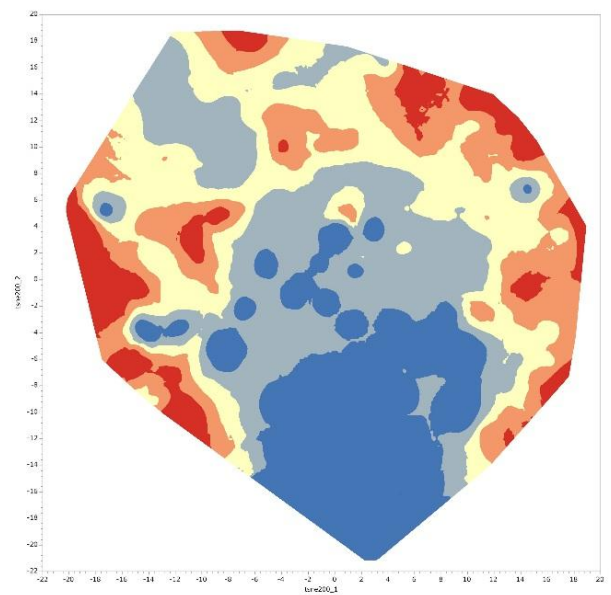
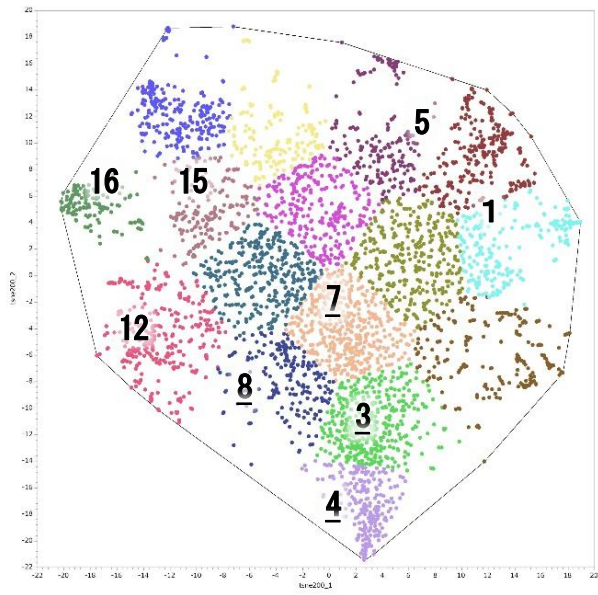


図4 t-SNE 法による店舗分散表現ベクトルの2次元散布図

上段左：各CLの配置（高評価CL：1, 5, 12, 15, 16，低評価CL：3, 4, 7, 8）

中段左：コストパフォーマンス（値頃感）スコア（赤が高評価）

下段左：予算（赤が高予算）

上段右：総合評価スコア（赤が高評価）

中段右：ドリンク評価（赤が高評価）

下段右：雰囲気スコア（赤が高評価）

持つ単語群はキーマ、ラム、NONVEG（野菜なし）等カレー専門店のメニューが頻出する。ここから CL1 はカレーメニューの提供をその特徴としてロコミが記述されている店舗群であるということが読み取れる。他に CL12 は本格的なスイーツ、CL2 はクレープやガレット、CL16 はリテールベーカリーの商品群を特徴として評判を得ていることが推測された。

消費者による一般的な店舗評価のプロセスとして、まず店舗での体験は、業種分類（カフェや喫茶店）やより細分化された下位カテゴリー（カレーカフェ、スイーツカフェ、クレープカフェ等）に適切に分類された後に、それぞれのカテゴリー店舗における過去の体験と詳細な比較評価がなされるというプロセスが仮定される。この場合に既存カテゴリーとの関係が不明瞭な場合には、積極的な評価がなされにくい。CL 別の店舗評価分布を図3に示す。多くの CL でスコアは広範囲に分布（高い店舗と低い店舗が混在）していることが確認されるが、CL3, 4, 7, 8 の4クラスターに限っては、店舗の多くが低評価に集中する。

これら低評価 CL は、特徴的な単語群として、モーニング、ダーズリンおよびスモーギーセットなど、一般的なカフェ用語が並ぶ。カレーカフェやスイーツカフェのような明確な下位カテゴリーとして認識されない、特筆性に乏しくブランド訴求が弱い単機能型店舗や駅前好立地で回転率重視の低付加価値型を中心とする店舗群によって構成され

た店舗群が多く含まれているものとみられる。

こうした店舗はビジネスマンの立ち寄りや待ち合わせ等のニーズも高く、地区においてその存在価値が他類型のカフェ店舗と比較し、決して劣るものではない。しかし、評判店となり店舗に滞在する経験価値それ自体が魅力となり、場合によっては名物店舗となって地区全体の印象を底上げし、街の魅力作りに寄与する価値は低いものとみられる。

図4に各店舗の分散表現ベクトルのt-SNE法による2次元散布図を掲げる。上段左に各 CL の配置を、上段右、中段及び下段には、各店舗に対する評価スコアをクリギング法で内挿推定することで、知覚地図（パーセプションマップ）として表現している。

上段右は、店舗の総合評価スコアを示す。中央部から真下に集中して低評価店舗が大半を占める CL3, 4, 7, 8 が配置されている。一方評価が高い店舗は下部を除く散布図外縁部に集中している。とりわけ評価が高い CL1, 5, 12, 15, 16 は、それぞれカレーカフェやスイーツカフェなどメニューやサービスに対して明瞭な特徴を持っている。中心から下部に集中する単機能・低付加価値型喫茶店からスタートして、遠方からの目的来店客が集客出来るほどのブランド・カフェに業態転換を意図する際には、散布図外縁部に向かって多様な方向性での検討が可能であるという実務的な示唆としても読み取れる。

中段以下は項目別評価である。本格派の老舗コーヒー店で構成される CL15 はドリンクとムードの評価が高い。最もコストパフォーマンスに優れるのは CL3 だが、その多く

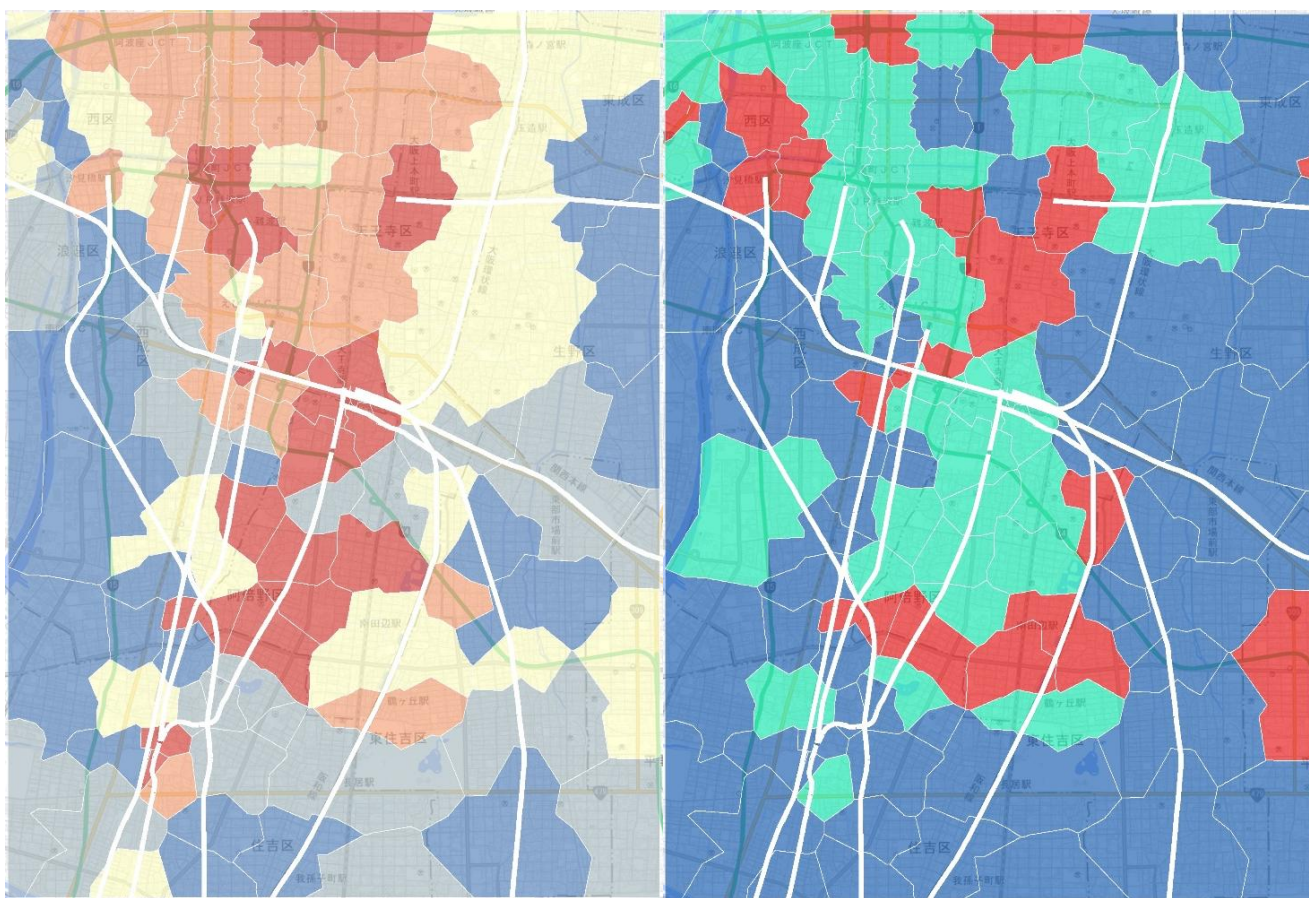


図5 左：駅勢圏別店舗総合スコア中央値（赤が高評価） 右：特化係数による優勢なCL（低評価CL地区は青、高評価CL地区は赤）

の店舗は残念ながら、総合スコアでは高い評価が得られてはいない。これら項目別の知覚地図からは、カフェの下位カテゴリのそれぞれが抱える事情を読み取ることが可能である。

最後に実空間（現実世界の地図）上での店舗の CL 別構成比について確認する。図 5 左に駅勢圏別に地区内におけるカフェの総合評価スコアの中央値を示す。先に記した通り高評価のカフェは必ずしもターミナル駅を中心とする業務地区・商業地区内のカフェ集中地区のみに分布しているのではない。住宅地区や混合地区のような立地においても高評価カフェの集積がみられる地区が存在する。今回の分析対象エリア内では、天王寺駅以南の住吉区（住吉大社）から阿倍野区（帝塚山，昭和町），東住吉区（山坂，南田辺）にかけて高評価の店舗が集まるエリアが広がっている。

それと比較して図 5 右は，それぞれの駅勢圏における優勢な CL を示したものである。先ずエリア別に CL 構成比による特化係数を算出し，それぞれのエリアにおいてもっとも特化係数の高い CL を優勢な CL とした。優勢な CL が低評価 CL（CL3，4，7，8）に含まれれば青色，高評価 CL（CL1，5，12，15，16）であれば赤色，それ以外を緑色として表現している。図 5 左で示した総合評価スコアが高いエリア（大阪環状線内の商業地区・繁華街および上記住吉区等）を除けば，低評価 CL が優勢なエリアが多くを占めていることが確認された。

これらを併せて検討すると，エリア別のカフェ評価の高低とカフェ下位カテゴリの構成効果とを関連付けて考えることが可能となる。まず，エリアにおける高評価店（低評価店）の多寡は，下位カテゴリの業態構成比率によってある程度説明可能だとみられることだ。低評価 CL に含まれる店舗の総合評価は低得点に集中しているため，低評価 CL が優勢なエリアでは，当然ながらエリアの平均スコアは下がる。

しかし，それだけでは留まらない。他のエリアでは，一定の割合で出現する高評価なカフェの存在が，後続店舗の出店促進など地区活性化の糸口につながる可能性があるが，低評価 CL が優勢なエリアでは，そのように地区が変化するための最初のきっかけを得ることは，より困難だとみられる。

5. おわりに

本報は，ロコミデータを用いてカフェをその下位カテゴリである CL に分類した上で，エリア別評価の高低が，CL 構成比と関係している可能性を示唆した。

小地域における CL 構成比からは，他にもより詳細な示唆（新規出店時の際の下位カテゴリの選択やそれら下位カテゴリの集積を活かした地域企画等）を得ることができると思われる。

高評価店舗の集積（低評価エリアの固定化）メカニズムについては仮説の域を出ない。今回，小地域における CL 構成比の経時的な動態変化については確認していない。これらは今後の課題としたい。

謝辞：

本研究は，東京大学空間情報科学研究センター 共同研究（No. 256）による成果である（利用データ：商業集積統計（2016 年），【空間配分版】2010 年近畿都市圏 人の流れデータセット）。研究遂行にあたり関係諸氏には日頃より，有益なご助言を戴いた。ここに感謝の意を表する。

文献：

- 鈴木英之（2021）ロコミ分散表現ベクトルを用いた街の魅力の類型化とその空間分布 地理情報システム学会講演論文集 vol30
- LE, Quoc; MIKOLOV, Tomas. (2014) Distributed representations of sentences and documents. In: International conference on machine learning. 2014. p. 1188-1196.
- VAN DER MAATEN, Laurens; HINTON, Geoffrey. (2008) Visualizing data using t-SNE. Journal of machine learning research, 2008, 9.11.
- Reuschke, Darja, Long, Jed and Bennett, Nicholas C (2021) Locating creativity in the city using Twitter data. Environment and Planning B: Urban Analytics and City Science, 48 (9), 2607-2622.
- Adelfio, Marco, et al. (2020) "Social activity in Gothenburg's intermediate city: Mapping third places through social media data." Applied Spatial Analysis and Policy 13.4: 985-1017.
- Van der Zee, Egbert, Dario Bertocchi, and Dominique Vanneste. (2020) "Distribution of tourists within urban heritage destinations: a hot spot/cold spot analysis of TripAdvisor data as support for destination management." Current Issues in Tourism 23.2: 175-1