

オケージョナル・ペーパー No.123

飲食業ウェブデータにおける「業態」の表記揺れを
統一的な表記に補正するデータクレンジング手法の開発

2022年12月

法政大学

日本統計研究所

オケージョナル・ペーパー No.123

飲食業ウェブデータにおける「業態」の表記揺れを
統一的な表記に補正するデータクレンジング手法の開発

2022年12月

法政大学

日本統計研究所

飲食業ウェブデータにおける「業態」の表記揺れを 統一的な表記に補正するデータクレンジング手法の開発

城 宏樹

1 はじめに

現在日本において飲食店に関するデータは公的統計・民間企業の提供するデータ共に複数存在する。公的統計では「経済センサスー活動調査」や「サービス産業動向調査」、民間企業の提供するデータでは、VRESAS に提供している Retty 株式会社「Food Data Platform」などがある。日本統計研究所では民間企業が毎月収集している全国飲食店舗名簿（ミクロのパネルデータ）を数か月に 1 回、定期的に購入している（以下、「飲食業ウェブデータ」と呼ぶ。）。これらは新型コロナウイルス流行の飲食業への影響を分析する上で有効なデータであると考えられる。

しかし同データはビッグデータの種類であり、それゆえにビッグデータ一般に存在する様々な問題が同データにもある。中田・溝口（2020）はビッグデータ一般の利用における課題として「代表性・信頼性」「継続性」「比較可能性」があるとしている。ちなみに、ビッグデータ一般の「比較可能性」の問題について内閣府（2015）は「重要な課題としては、ビッグデータが価値ある情報だけでなく、誤った情報、ノイズ的信息、不規則な情報もある玉石混合であることから、『ノイズの除去や活用しやすい形式への変換』が挙げられる。データクレンジング技術の高度化、企業・業界横断的にデータ形式の標準化、蓄積の方法の統一化などが必要である。」（3 頁）としている。

本稿で取り上げる飲食業ウェブデータについては、「業態」について表記揺れが多々あることから「比較可能性」に問題があり、それが同データを利用する上での障害になっている。そこで飲食業ウェブデータの「業態」の表記揺れを統一的な表記に補正するデータクレンジング手法の開発を行った。ここでは統一的な表記として日本標準産業分類を参照した。すなわち、飲食業ウェブデータをテキストマイニングツールである KH Coder と Excel を用いてデータクレンジングを行い、同データの「業態」を日本標準産業分類（本稿では「大分類 M 宿泊業，飲食サービス業」および「中分類 76 飲食店」の「小分類 760 管理，補助的経済活動を行う事業所（76 飲食業）」を除く）へ産業格付けする手法を開発した。なお、本来「業態」と「業種」は異なる概念であるが、同データにおける「業態」は日本標準産業分類に類似した分類である。

本稿の構成は以下の通りである。まず飲食業ウェブデータにどのようなデータが含まれているのかを説明する。次に、分析に用いる KH Coder についての説明とテキストマイニングの変遷を述べる。次に分類先である日本標準産業分類を説明する。次に、KH Coder を用

いた飲食業ウェブデータの日本標準産業分類への産業格付けを説明する。最後にテキストマイニングを用いた飲食業ウェブデータの産業格付けについての今後の課題を述べることとする。

2 飲食業ウェブデータについて

本稿で用いる飲食業ウェブデータは、各飲食店情報を年月別・都道府県別にまとめたExcel データである。データとしては「店名」「ヨミガナ」「URL」「料理ジャンル」「ジャンル」「エリア」「最寄駅」「住所」「緯度/経度」「TEL」「FAX」「E-mail」「WEB」「業態」「アクセス」「営業時間」「定休日」「平均予算」「その他料金」「クレジットカード」「総席数」「宴会最大人数」「貸切可能人数」「駐車場」「個室」「設備・サービス」「お子様連れ」「携帯・Wi-Fi・電源」「メニューのサービス」「外国語対応」「ペット同伴」「禁煙・喫煙」の情報が記載されている（図表1～3）。

図表1 飲食業ウェブデータの内容例（「店名」から「住所」）

店名	ヨミガナ	URL	料理ジャンル
食肉センター 高崎商店 昭島	ショクニクセンターカサキシヨウテン アキシマ	https://r.gnavi.co.jp/hpadv2sp0000/	焼肉・ホルモン・鉄板焼き
昭和の森ガーデンCAFE&BBQ	ショウワノモリガーデンカフェアンドバーベキュー	https://r.gnavi.co.jp/amsmkvgr0000/	焼肉・ホルモン・鉄板焼き
すずみやき家 串陣 中神店	スズミヤキヤクシジン ナカガミテン	https://r.gnavi.co.jp/9vuxcppk0000/	串陣
昭和の森 車屋	ショウワノモリクルマヤ	https://r.gnavi.co.jp/g008010/	和食
魚民 中神北口駅前店	ウオタミ ナカガミキタグチエキマエテン	https://r.gnavi.co.jp/gajn166/	魚民
山内農場 昭島南口駅前店	ヤマウチノウジョウ アキシマミナミグチエキマエテン	https://r.gnavi.co.jp/gcku362/	山内農場
魚民 拝島南口駅前店	ウオタミ ハイジマミナミグチエキマエテン	https://r.gnavi.co.jp/gau4671/	魚民
目利きの銀次 昭島南口駅前店	メキキノギンジ アキシマミナミグチエキマエテン	https://r.gnavi.co.jp/gajn165/	目利きの銀次
和食レストランとんでん 昭島店	ワシヨクレストランテン アキシマテン	https://r.gnavi.co.jp/p9cvgg710000/	とんでん
魚民 昭島南口駅前店	ウオタミ アキシマミナミグチエキマエテン	https://r.gnavi.co.jp/gajn164/	魚民
千年の宴 東中神南口駅前店	センネンノウタゲ ヒガシナカガミナミグチエキマエテン	https://r.gnavi.co.jp/gcku363/	鍋
La Faro	ラファエロ	https://r.gnavi.co.jp/k3367hdf0000/	イタリアン・フレンチ
暖家 昭島店	ダンケ アキシマテン	https://r.gnavi.co.jp/a797701/	焼肉・ホルモン・鉄板焼き
ハイカラ酒場	ハイカラサカバ	https://r.gnavi.co.jp/4a76mnhz0000/	居酒屋
遊食Dining 月宵	ユウシヨクダイニングツキヨイ	https://r.gnavi.co.jp/ntu1z4wh0000/	バー・パル・ダイニングバー
中国料理 花林 フォレスト・イン昭和館	チュウゴクリョウワカリン フォレストインシヨウワカン	https://r.gnavi.co.jp/g857201/	中華
うまもん酒場 かさ 拝島駅前店	ウマモンサカバカサ ハイジマエキマエテン	https://r.gnavi.co.jp/e023700/	居酒屋
海鮮居酒屋 はなの舞 昭島南口店	カイセンザカヤハナノマイ アキシマミナミグチテン	https://r.gnavi.co.jp/g863240/	はなの舞
寿司居酒屋 やがずし 中神町	スシザカヤヤダイズシ ナカガミチヨウ	https://r.gnavi.co.jp/e919320/	やがずし

ジャンル	エリア	最寄駅	住所
焼肉・ホルモン・鉄板焼き	東京/立川・八王子・青	昭島駅/中神駅/拝島駅	196-0015 東京都昭島市昭和町5-11-17
焼肉・ホルモン・鉄板焼き	東京/立川・八王子・青	昭島駅/西武立川駅/中	196-0014 東京都昭島市田中町576-1
居酒屋/和食/焼き鳥	東京/立川・八王子・青	中神駅/東中神駅/昭島	196-0022 東京都昭島市中神町1176-32 HIMビル1F
日本料理・郷土料理	昭 東京/立川・八王子・青	昭島駅/西武立川駅/中	196-8601 東京都昭島市昭和の森 フォレスト・イン昭和館1F
鍋	昭島市×鍋×個室/中	東京/立川・八王子・青	中神駅/東中神駅/昭島 196-0022 東京都昭島市中神町1180-4 中神駅北口T Kビル2・3F
鍋	昭島市×鍋×個室/新	東京/立川・八王子・青	昭島駅/中神駅/拝島駅 196-0015 東京都昭島市昭和町5-14-17 メディックスビル2F
鍋	昭島市×鍋×個室/中	東京/立川・八王子・青	拝島駅/龍川駅/昭島駅 196-0003 東京都昭島市松原町4-12-6 増田ビル1F
海鮮・魚介料理	昭島市	東京/立川・八王子・青	昭島駅/中神駅/西武立 196-0015 東京都昭島市昭和町2-1-14 松葉家ビルB1
海鮮・魚介料理/海鮮	東京/立川・八王子・青	中神駅/昭島駅/東中神	196-0015 東京都昭島市昭和町3-11-12
鍋	昭島市×鍋×個室/中	東京/立川・八王子・青	昭島駅/中神駅/西武立 196-0015 東京都昭島市昭和町2-1-14 松葉家ビル1F・2F
鍋	昭島市×鍋×個室/新	東京/立川・八王子・青	東中神駅/西武立川駅/中 196-0034 東京都昭島市玉川町1-1-9 匠ビル1F
イタリアン・フレンチ	昭島市	東京/立川・八王子・青	昭島駅/西武立川駅/東 196-0021 東京都昭島市武蔵野3-2-51
焼肉・ホルモン・鉄板焼き	東京/立川・八王子・青	東中神駅/中神駅/小宮	196-0031 東京都昭島市福島町2-35-2
居酒屋/和食/焼き鳥	東京/立川・八王子・青	拝島駅/龍川駅/牛浜駅	196-0003 東京都昭島市松原町4-13-20 停車場ビル1F
バー・パル・ダイニングバー	東京/立川・八王子・青	拝島駅/龍川駅/牛浜駅	196-0003 東京都昭島市松原町5-1-20 第二野村ビル2F
中華	昭島市×中華×個室	東京/立川・八王子・青	昭島駅/西武立川駅/中 196-8601 東京都昭島市拝島町4017-3 フォレスト・イン昭和館
居酒屋/日本料理・郷土	東京/立川・八王子・青	拝島駅/龍川駅/牛浜駅	196-0003 東京都昭島市松原町4-13-20 停車場ビル1F・101
居酒屋	昭島市×居酒屋	東京/立川・八王子・青	昭島駅/中神駅/西武立 196-0015 東京都昭島市昭和町2-5-12 昭島駅前ビル2F
居酒屋/和食	東京/立川・八王子・青	中神駅/東中神駅/昭島	196-0025 東京都昭島市朝日町1-4-1 平田ビル1F

出典:筆者作成

図表 2 飲食業ウェブデータの内容例(「緯度/経度」から「個室」)

緯度/経度	TEL	FAX	E-mail	WEB	業態
35.7127046/139.3600503	050-3373-9540				食肉センター
35.71624696529713/139.050-3313-8574					レストラン型B&B
35.7095027/139.3765654	050-3374-1033				居酒屋 個室
35.71866802/139.364163	050-3469-3763	042-542-3393		http://www.kuruma-ya.co.jp	日本料理
35.709244/139.375622	042-545-9288	042-545-9291		http://www.monterozo.co.jp/	和風居酒屋
35.712921/139.35932	042-546-7588			http://www.monterozo.co.jp/	地鶏 居酒屋
35.719962/139.343382	042-544-8988	042-544-8999		http://www.monterozo.co.jp/	のみくい処
35.713299/139.361226	042-544-6388	042-544-6392		http://www.monterozo.co.jp/	漁師料理とうま
35.706635/139.366913	050-3476-1593	042-545-7966		http://www.tonden.co.jp/	和食レストラン
35.713299/139.361226	042-544-3288	042-544-5414		http://www.monterozo.co.jp/	和風居酒屋
35.7058533/139.3845455	042-543-2188			http://www.monterozo.co.jp/	和風居酒屋
35.718813/139.379158	050-3469-1409			http://lafaro2000.amebaownd.jp	(イタリアン)
35.69696788/139.376439	050-3467-6586	042-543-9777			焼肉
35.72117521784622/139.050-3313-2418					居酒屋 個室
35.7212419/139.3422458	050-3374-5430				ダイニングバー
35.71864516/139.364155	050-3466-5178	042-542-7039	service.fis@showa-aircraft.co.jp	http://www.showakan.co.jp/	広東料理
35.7211821/139.3426707	042-545-1929				から揚げ☆丸
35.71301242/139.360019	042-500-5868 / 042-500-5868			http://www.chimney.co.jp/	海鮮居酒屋 (
35.70839/139.375573	050-3490-8026	042-549-1808		http://www.yataiya-group.com/	本格寿司職人

アクセス	営業時間	定休日	平均予算	その他料金	クレジットカード	総席数	宴会最大人数	買切可能人数	駐車場	個室
J R 青梅線	月～土17:00	(不定休日あり)	2,500円	(通常平均)		80席掘りごた	80名様 (着席時)		無	
J R 青梅線	11:30～16:00	(不定休日あり)	5,000円	(通常平均)	5,000円 (ランチ平均)	200席				
J R 青梅線	17:00～24:00	無※年末年始	3,000円	(通常平均)	3,500円 (宴会平均)	86席座敷席あり	カウンター席あり			
J R 青梅線	月～金ランチ	無	8,000円	(通常平均)	1 VISA UC	ダイナースクラ	260席座敷席	80名様 (着席時)	～80名様	有: 共有400台 (共用)
J R 青梅線	月～木・日・夜	無	2,900円	(通常平均)	2.5 VISA MasterCard UC	124席掘りごた	2つ席あり			有 (詳細は店舗までお問い合わせ)
J R 青梅線	日・祝日17:00	無	3,500円	(通常平均)	2.5 VISA MasterCard UC	105席				有 (詳細は店舗までお問い合わせ)
J R 拝島駅	月～木・日・夜	無	2,700円	(通常平均)	2.5 VISA MasterCard UC	77席				有 (詳細は店舗までお問い合わせ)
J R 青梅線	月～土・祝前	無	2,800円	(通常平均)	2.5 VISA MasterCard UC	78席				有 (詳細は店舗までお問い合わせ)
JR 昭島駅	11:00～24:00	無	1,400円	(通常平均)	2.5 VISA MasterCard UC	138席				有: 専用無料55台
J R 青梅線	月～木・日・夜	無	2,600円	(通常平均)	2.5 VISA MasterCard UC	104席掘りごた	2つ席あり			有 (詳細は店舗までお問い合わせ)
J R 青梅線	月～木・日・夜	無	2,500円	(通常平均)	2.5 VISA MasterCard UC	97席				無
J R 青梅線	月・水～日・夜	火曜日	2,500円	(通常平均)	3,600円 (宴会平均)	850	24席	24名様 (着席時)	15名様～24名様	有: 無料3台
J R 青梅線	17:00～23:00	水曜日	4,500円	(通常平均)	4.5 VISA MasterCard UC	60席	60名様 (着席時)			有: 専用25台 (専用)
J R 拝島駅	17:00～24:00	(不定休日あり)	2,500円	(通常平均)	3,000円 (宴会平均)	56席	カウンター席あり			
J R 拝島駅	火～木・日・夜	月曜日	2,800円	(通常平均)		80席				
J R 青梅線	月～金ランチ	無	4,800円	(通常平均)	1 VISA UC	ダイナースクラ	68席ソファー	席30名様 (着席時)		有: 共有無料400台 (※)
J R 拝島駅	月～木・日	17年末年始	3,000円	(通常平均)	3 VISA MasterCard UC	90席掘りごた	50名様 (着席時)			
J R 青梅線	月～木16:00	無	2,700円	(通常平均)	3.3 VISA UC DC UFJ	ダイ	131席座敷席	40名様 (着席時)		
J R 青梅線	月～木・日・夜	年中無休	2,500円	(通常平均)	230円	60席	40名様 (着席時)			

出典:筆者作成

図表 3 飲食業ウェブデータの内容例(「設備・サービス」から「禁煙・喫煙」)

設備・サービス	お子様連れ	携帯・Wi-Fi・電源	メニューのサービス	外国語対応	ペット同伴	禁煙・喫煙
日曜営業あり					同伴不可	喫煙可 (全席喫煙可)
日曜営業あり						喫煙可 (全席喫煙可)
マイク利用可				外国語メニューあり: 英語メニューあり		喫煙可 (全席喫煙可)
23時以降も食事ができる 朝までオープン	日曜営業あり	カラオケあり				喫煙可 (全席喫煙可)
23時以降も食事ができる 朝までオープン	日曜営業あり	カラオケあり		外国語メニューあり: 英語メニューあり		喫煙可 (一部禁煙席)
23時以降も食事ができる 朝までオープン	日曜営業あり	カラオケあり				喫煙可 (一部禁煙席)
23時以降も食事ができる 朝までオープン	日曜営業あり	カラオケあり				店内全面禁煙 (店外)
23時以降も食事ができる 朝までオープン	日曜営業あり	カラオケあり				喫煙可 (一部禁煙席)
23時以降も食事ができる 日曜営業あり		シェフがいる			同伴不可	店内全面禁煙 (店外)
日曜営業あり						喫煙可
日曜営業あり	ソムリエがいる店	シェフがいる	ホテル内・ホテル直営	宿泊施設あり	外国語メニューあり: 英語メニューあり	店内全面禁煙
23時以降も食事ができる					同伴可?小型犬:	喫煙可
23時以降も食事ができる						喫煙可 (一部禁煙席)

出典:筆者作成

図表 1～3 から、各飲食店によって記載されている情報と記載されていない情報があることがわかる。そのため、本稿ではすべての飲食店で記載をされている、または記載されていなくても補完することのできる「店名」「URL」「住所」「緯度/経度」「業態」「アクセス」を用いる。また、「店名」「住所」「緯度/経度」「業態」「アクセス」には表記揺れがあるため、同一飲食店かどうかを見分ける際には菅（2017）同様、「URL」を用いた¹。

3 テキストマイニングツール「KH Coder」について

「KH Coder」とは、樋口（2004）の中で提案されたテキストマイニングツールであり、テキストマイニングにおける従来の Correlational アプローチと Dictionary-based アプローチを統合させ、第一段階に Correlational アプローチ、第二段階に Dictionary-based アプローチを行うテキストマイニングツールである²。本稿では飲食業ウェブデータの各店飲食店の日本標準産業分類への格付けを行うが、飲食業ウェブデータに含まれる形態素（テキストデータの最小単位）の出現頻度や共起関係を確認し、分類基準を作成する際に KH Coder を用いた。

KH Coder の第一段階で使われている Correlational アプローチは、同一文書内に頻繁に出現する言葉のグループや共通する言葉を含む文書のグループを多変量解析によって自動的に発見分類するコンピュータを用いるアプローチであり³、対象となる文書の中からどのような言葉がどのような基準（コーディング規則）によって取り上げられたのかを第三者が把握することができる客観性を持っている⁴。一方で Correlational アプローチは、多変量解析に依存しているために、理論や仮説に基づいて分類を自動化することができないというデメリットが存在する⁵。Correlational アプローチによる英語でのテキストマイニングシステムとしては WORDS (Iker & Harway 1969) があり、日本語では WordMiner (大隅・Lebart 2000) が存在する⁶。

KH Coder の第二段階で使われている Dictionary-based アプローチは、分析者が作成した基準（コーディング規則 (dictionary)）に従って文書内の言葉を分類するためにコンピュータを用いるアプローチであり、基準（コーディング規則）を作成することで分析者の理論や問題意識を自由に操作することができるという利点がある⁷。一方で Dictionary-based アプローチは、意図的・無意識的に理論や仮説に都合の良い基準（コーディング規則）が作成されやすいという危険性がある⁸。Dictionary-based アプローチによる英語でのテキス

¹ 菅（2017）、p. 51 参照。

² 樋口耕一（2004）、p. 104 参照。

³ 前掲書、p. 102 参照。

⁴ 前掲書、p. 104 参照。

⁵ 前掲書、p. 104 参照。

⁶ 前掲書、p. 104 参照。

⁷ 前掲書、pp. 102. 103 参照。

⁸ 前掲書、p. 103 参照。

トマイニングシステムとしては General Inquirer (Stone et al. 1966) があり、日本語では Autocode (佐藤 1992、田中・太郎丸 1996) が存在する⁹。

4 日本標準産業分類と飲食業ウェブデータ「業態」の分類

本稿の目的である飲食業ウェブデータの分類を行う際、最終的な目標としているのは「表記揺れのある各飲食店ウェブデータを日本標準産業分類に沿って分類すること」である。前述のとおり、本稿では飲食業ウェブデータのうち各飲食店に記載されている・補完することのできる「店名」「URL」「住所」「緯度/経度」「業態」「アクセス」を用いる。そのうち日本標準産業分類への格付けの際には、各飲食店の営む業態を示す「業態」を用いる。

本節では、最終的に飲食業ウェブデータを分類したい分類先である日本標準産業分類のうち「大分類 M 宿泊業, 飲食サービス業」「中分類 76 飲食店」に含まれる細分類（「小分類 760 管理, 補助的経済活動を行う事業所 (76 飲食業)」を除く）について述べる。具体的には、飲食業ウェブデータの「業態」にどのような形態素が含まれていたら、どの細分類へ格付けるべきか、例を挙げて説明する。

4.1 日本標準産業分類について

日本標準産業分類とは、「統計の正確性と客観性を保持し、統計の相互比較性と利用の向上を図ることを目的として設定された統計基準であり、全ての経済活動を産業別に分類している」¹⁰ものである。日本標準産業分類は 20 の大分類、99 の中分類、530 の小分類、1,460 の細分類からなっており¹¹、今回の分類では「大分類 M 宿泊業, 飲食サービス業」「中分類 76 飲食店」（「小分類 760 管理, 補助的経済活動を行う事業所 (76 飲食業)」を除く）を用いる。

4.2 本稿で使用する日本標準産業分類細分類について

本稿では日本標準産業分類のうち 15 の細分類を用いており、飲食業ウェブデータの「業態」にどのような形態素が含まれていた場合に各細分類に格付けていくか、総務省 HP を参照し、以下に簡潔にまとめる。

⁹ 前掲書、p. 104 参照。

¹⁰ 総務省「日本標準産業分類の変遷と第 13 回改訂の概要」

URL : https://www.soumu.go.jp/toukei_toukatsu/index/seido/sangyo/02toukatsu01_03000043.html (最終閲覧日 : 2022 年 5 月 5 日)

¹¹ 総務省「日本標準産業分類一般原則」

URL : https://www.soumu.go.jp/main_content/000286955.pdf (最終閲覧日 : 2022 年 5 月 5 日)

「細分類 7611 食堂，レストラン（専門料理店を除く）」

「細分類 7611 食堂，レストラン（専門料理店を除く）」は「主として主食となる各種の料理品をその場所で飲食させる事業所」¹²であり、食堂や大衆食堂などがそれにあたる。飲食業ウェブデータの「業態」に「食堂」「レストラン」などの言葉が含まれている場合、この細分類へ格付ける。

「細分類 7621 日本料理店」

「細分類 7621 日本料理店」は「主として特定の日本料理（そば，うどん，すしを除く）をその場所で飲食させる事業所」¹³であり、てんぷら料理店やうなぎ料理店などがそれにあたる。飲食業ウェブデータの「業態」に「てんぷら」「うなぎ」などの言葉が含まれている場合、この細分類へ格付ける。

「細分類 7622 料亭」

「細分類 7622 料亭」は「主として日本料理を提供し，客に遊興飲食させる事業所」¹⁴であり、料亭や待合がそれにあたる。飲食業ウェブデータの「業態」に「料亭」「待合」などの言葉が含まれている場合、この細分類へ格付ける。

「細分類 7623 中華料理店」

「細分類 7623 中華料理店」は「主として中華料理をその場所で飲食させる事業所」¹⁵であり、中華料理店や上海料理店などがそれにあたる。飲食業ウェブデータの「業態」に「中華」「上海」などの言葉が含まれている場合、この細分類へ格付ける。

「細分類 7624 ラーメン店」

「細分類 7624 ラーメン店」は「主としてラーメンをその場所で飲食させる事業所」¹⁶であり、ラーメン店や中華そば店がそれにあたる。飲食業ウェブデータの「業態」に「ラーメン」「中華そば」などの言葉が含まれている場合、この細分類へ格付ける。

¹² 総務省「大分類M－宿泊業，飲食サービス業」

URL： https://www.soumu.go.jp/main_content/000290732.pdf（最終閲覧日：2022年5月6日）

¹³ 同上

¹⁴ 同上

¹⁵ 同上

¹⁶ 同上

「細分類 7625 焼肉店」

「細分類 7625 焼肉店」は「主として焼肉（自ら網で焼くもの）をその場所で飲食させる事業所」¹⁷であり、焼肉店がそれにあたり、ステーキハウスやジンギスカン料理店はそれにあたらない。飲食業ウェブデータの「業態」に「焼肉」などの言葉が含まれている場合、この細分類へ格付ける。

「細分類 7629 その他の専門料理店」

「細分類 7629 その他の専門料理店」は「主として他に分類されない特定の料理をその場所で飲食させる事業所」¹⁸であり、西洋料理店やフランス料理店などがそれにあたる。飲食業ウェブデータの「業態」に「西洋料理」「フランス料理」などの言葉が含まれている場合、この細分類へ格付ける。

「細分類 7631 そば・うどん店」

「細分類 7631 そば・うどん店」は「主としてそばやうどんなどをその場所で飲食させる事業所」¹⁹であり、そば屋やうどん店などがそれにあたる。飲食業ウェブデータの「業態」に「そば」「うどん」などの言葉が含まれている場合、この細分類へ格付ける。

「細分類 7641 すし店」

「細分類 7641 すし店」は「主としてすしをその場所で飲食させる事業所」²⁰であり、すし屋がそれにあたる。飲食業ウェブデータの「業態」に「すし」「寿司」などの言葉が含まれている場合、この細分類へ格付ける。

「細分類 7651 酒場、ビヤホール」

「細分類 7651 酒場、ビヤホール」は「主として酒類及び料理をその場所で飲食させる事業所」²¹であり、大衆酒場や居酒屋などがそれにあたる。飲食業ウェブデータの「業態」に「酒場」「居酒屋」などの言葉が含まれている場合、この細分類へ格付ける。

¹⁷ 同上

¹⁸ 同上

¹⁹ 同上

²⁰ 同上

²¹ 同上

「細分類 7661 バー、キャバレー、ナイトクラブ」

「細分類 7661 バー、キャバレー、ナイトクラブ」は「主として洋酒や料理などを提供し、客に遊興飲食させる事業所」²²であり、バーやスナックバーなどがそれにあたる。飲食業ウェブデータの「業態」に「バー」「スナック」などの言葉が含まれている場合、この細分類へ格付ける。

「細分類 7671 喫茶店」

「細分類 7671 喫茶店」は「主としてコーヒー、紅茶、清涼飲料などの飲料や簡易な食事などをその場所で飲食させる事業所」²³であり、喫茶店やカフェなどがそれにあたる。飲食業ウェブデータの「業態」に「喫茶店」「カフェ」などの言葉が含まれている場合、この細分類へ格付ける。

「細分類 7691 ハンバーガー店」

「細分類 7691 ハンバーガー店」は「主としてハンバーガーをその場所で飲食させる事業所」²⁴であり、ハンバーガー店がそれにあたる。飲食業ウェブデータの「業態」に「ハンバーガー」などの言葉が含まれている場合、この細分類へ格付ける。

「細分類 7692 お好み焼・焼きそば・たこ焼き店」

「細分類 7692 お好み焼・焼きそば・たこ焼き店」は「主としてお好み焼・焼きそば・たこ焼きをその場所で飲食させる事業所」²⁵であり、お好み焼点や焼きそば店などがそれにあたる。飲食業ウェブデータの「業態」に「お好み焼（き）」「焼きそば」などの言葉が含まれている場合、この細分類へ格付ける。

「細分類 7699 他に分類されない飲食店」

「細分類 7699 他に分類されない飲食店」は「主として大福、今川焼、ところ天、汁粉、湯茶など他に分類されない飲食料品をその場所で飲食させる事業所」²⁶であり、大福屋や今川焼屋などがそれにあたる。飲食業ウェブデータの「業態」に「大福」「今川焼」などの言葉が含まれている場合、この細分類へ格付ける。

²² 同上

²³ 同上

²⁴ 同上

²⁵ 同上

²⁶ 同上

5 KH Coder での飲食業ウェブデータの分類基準作成

本章では飲食店ウェブデータを KH Coder を用いて日本標準産業分類への格付け基準を作成していく。本章の分類では西東京市の 2018 年 7 月、2019 年 7 月、2020 年 7 月の飲食店ウェブデータを使用する。

5.1 元データの貼り付け

Excel のフィルター機能を使い、「住所」情報から「西東京市」で検索を行い、西東京市のみ対象年月のデータを元ファイルからコピーし、各年を 1 つのシートとし、値として張り付ける。この作業を各年で行い、1 つの Excel ファイルにまとめる (図表 4)。このとき、元データはいじらないように留意する。

図表 4 対象年月ごとのシート作成 (西東京市)

	A	B	C	D	E	F	
1	店名	ヨミガナ	URL	料理ジャンル	ジャンル	エリア	最新
2	貸切&ハワカシキリ	ア	https://r.g	居酒屋	居酒屋	田無	東京/練馬
3	釜めし・ギカメシヤ	ア	https://r.g	居酒屋	居酒屋	田無	東京/練馬
4	焼鳥 チキヤキトリ	ア	https://r.g	居酒屋	居酒屋	田無	東京/練馬
5	個室居酒屋 コシツイヤ	ア	https://r.g	居酒屋	居酒屋	田無	東京/練馬
6	大衆酒場 タイシュウ	ア	https://r.g	居酒屋	居酒屋	田無	東京/練馬
7	肉×クラブ ニククラブ	ア	https://r.g	居酒屋	居酒屋	田無	東京/練馬
8	ピストロ	ア	https://r.g	居酒屋	居酒屋	田無	東京/練馬
9	ワイン酒場 ワインサナ	ア	https://r.g	居酒屋	居酒屋	田無	東京/練馬
10	居酒屋 よいザカヤ	ア	https://r.g	居酒屋	居酒屋	田無	東京/練馬
11	やきとり家 ヤキトリヤ	ア	https://r.g	居酒屋	居酒屋	田無	東京/練馬
12	旨唐揚げとウマカラ	ア	https://r.g	ミライザカ	居酒屋	田無	東京/練馬
13	木曾路 田	ア	https://r.g	木曾路	鍋田無	×	東京/練馬
14	鳥良商店	ア	https://r.g	磯丸水産	和食/鳥料	東京/練馬	田無
15	産直牡蠣・サンチョク	ア	https://r.g	さかなや道	居酒屋	田無	東京/練馬
16	沖縄健康長オキナワク	ア	https://r.g	和食	和食	田無	×
17	九州料理	ア	https://r.g	居酒屋	居酒屋	/和	東京/練馬
18	ワインレス	ア	https://r.g	イタリアン	イタリアン	東京/練馬	田無
		201807	201907	202007			+

出典:筆者作成

5.2 不要な情報の削除と並び替え

前述のとおり、本稿での用いる飲食業ウェブデータの情報は「店名」「URL」「住所」「緯度/経度」「業態」「アクセス」である。よってそれ以外の情報は列ごと削除する。このとき、一番左をA列とし、「URL」「店名」「住所」「業態」「緯度/経度」「アクセス」の順番で並び変えておく（図表5）。

図表5 不要な情報の削除と並び替え(西東京市)

	A	B	C	D	E	F	G
1	URL	店名	住所	業態	緯度/経度	アクセス	
2	https://r.g	貸切&ハワ	202-0002	ハワイにG	35.752912	西武池袋線	ひは
3	https://r.g	釜めし・ヤ	202-0001	釜めし・炭	35.751495	西武池袋線	ひは
4	https://r.g	焼鳥 チキ	188-0011	炭火烧き鳥	35.728782	西武新宿線	田無
5	https://r.g	個室居酒屋	188-0011	田無 居酒屋	35.727951	西武新宿線	田無
6	https://r.g	大衆酒場	188-0011	焼き鳥と	35.728740	西武新宿線	田無
7	https://r.g	肉×クラブ	202-0002	肉×クラブ	35.754079	西武池袋線	ひは
8	https://r.g	ビストロ	188-0012	居酒屋ビブ	35.726952	西武新宿線	田無
9	https://r.g	ワイン酒場	202-0001	生ハム好き	35.750904	西武池袋線	ひは
10	https://r.g	居酒屋よ	188-0011	炭火で焼い	35.729343	西武新宿線	田無
11	https://r.g	やきとり	202-0002	大山どり	35.752277	西武池袋線	ひは
12	https://r.g	旨唐揚げ	188-0011	田無居酒屋	35.728239	西武新宿線	田無
13	https://r.g	木曾路 田	188-0004	しゃぶしゃ	35.732485	西武新宿線	田無
14	https://r.g	鳥良商店	188-0011	鶏料理	35.72928/	西武新宿線	田無
15	https://r.g	産直牡蠣	188-0012	海鮮居酒屋	35.727041	西武新宿線	田無
16	https://r.g	沖縄健康	202-0002	沖縄海鮮	35.752431	西武池袋線	ひは
17	https://r.g	九州料理	188-0012	緑園・モッ	35.727021	西武新宿線	田無
18	https://r.g	ワインレ	202-0002	肉ビストロ	35.752565	西武池袋線	ひは

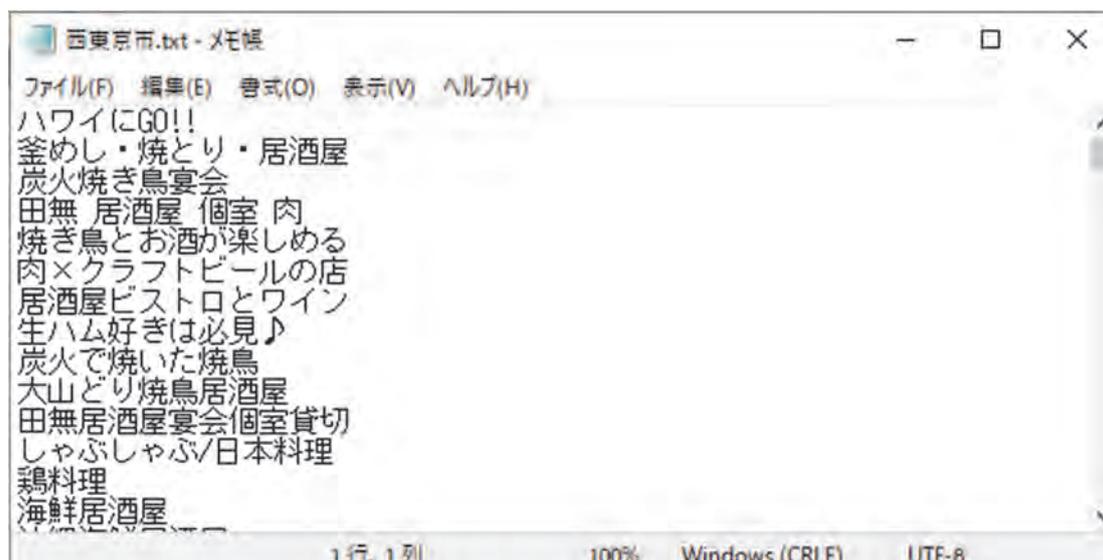
出典:筆者作成

5.3 KH Coder での前処理

まずKH Coder で分析を行うためにテキスト（拡張子「.txt」）ファイルを作成する。分類対象である飲食業ウェブデータの「業態」を各年7月からコピーし、テキストファイルを作成できるソフトに張り付けていき、2018～2020年7月の業態をひとつのテキストファイルにする。このとき、1つの業態名につき1行になるように張り付ける（図表6）。その際、年と年の間をあける必要はない（空の行を作成しないようにする）。

次にKH Coder での分類を行う。KH Coder を起動し、ホーム画面から「プロジェクト」「新規」を選択、新規プロジェクトのウィンドウが出たら、「参照」から先ほど作成したテキストファイルを選択し、「OK」を選択する。次にホーム画面から「前処理」「テキストのチェック」を実行する。次にホーム画面から「前処理」「前処理の実行」「OK」を選択、「処理が完了しました」のウィンドウが現れ、図表7のような画面に切り替われば分類準備は完了である。この作業により、テキストファイルに含まれる形態素を形態素解析し、抽出語としてKH Coder 内で分析することができるようになる。

図表6 「業態」のテキストファイル作成



出典：筆者作成

図表 7 「KH Coder」での前処理(西東京市)



出典:筆者作成

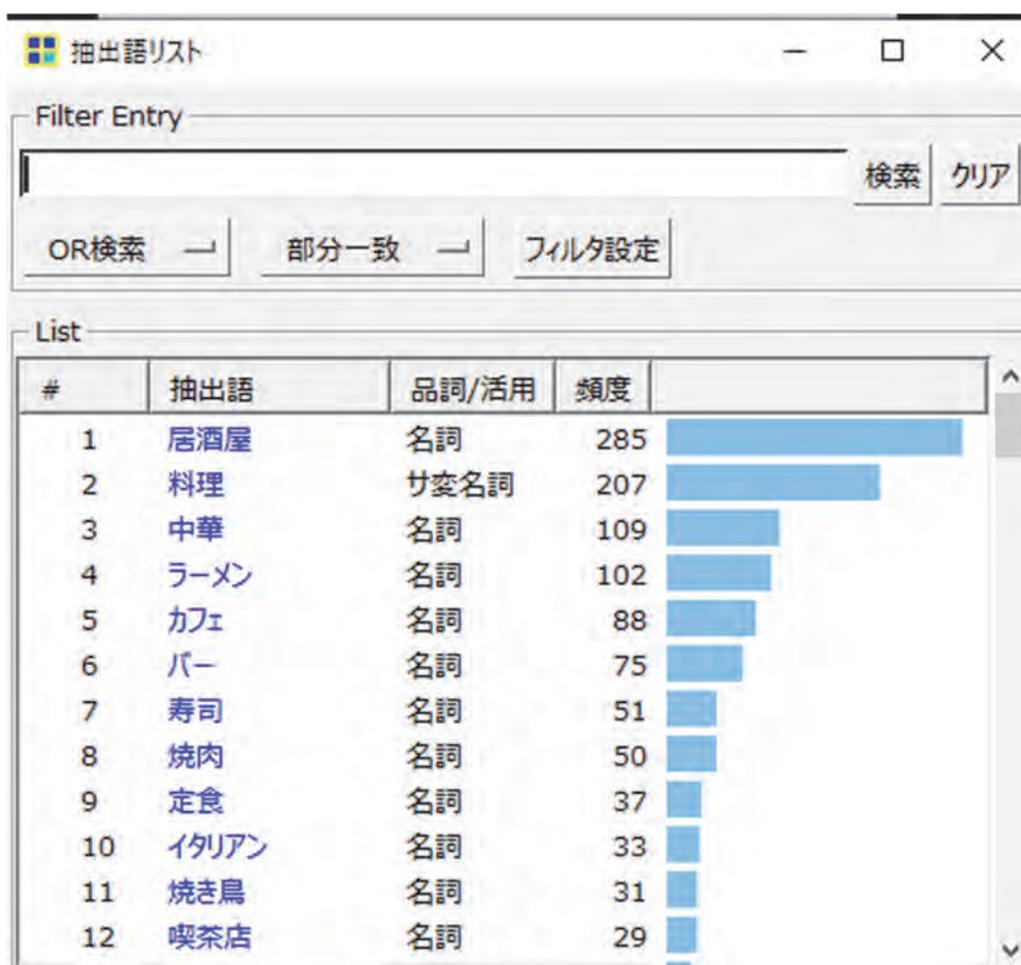
5.4 KH Coder での分析

次に KH Coder を用いて抽出語（形態素解析された形態素）の分析を行う。KH Coder では様々な分析をすることができるが、本稿では飲食業ウェブデータの「業態」に含まれる抽出語から日本標準産業分類の 15 の飲食店に関する細分類への格付け基準の作成するため、「抽出語リスト」と抽出語から作成される「共起ネットワーク」での分析を行う。

5.4.1 「抽出語リスト」での分析

KH Coder ホーム画面から「ツール」「抽出語」「抽出語リスト」を選択する。すると頻出語順に並び替えられた抽出語のヒストグラムが新しいウィンドウとして現れる（図表 8）。

図表 8 抽出語リスト(西東京市)



出典:筆者作成

図表 8 より、飲食業ウェブデータ「業態」にどのような抽出語が多く含まれているのかが確認できる。この結果から、西東京市にはどのような業種を営む飲食業が多いのかを大まかに把握することができる。またここで表記揺れを確認することもできる。例えば図表 8 から 5 番目に頻度の高い「カフェ」という抽出語と 12 番目に頻度の高い「喫茶店」という抽出語を確認することができる。前述の第 4 章第 2 節によれば、「カフェ」「喫茶店」の抽出語が含まれる場合は「細分類 7671 喫茶店」に分類されるべきである。しかし抽出語上では、これらは別の業態として判別されていることがわかる。

5.4.2 「共起ネットワーク」での分析

次に KH Coder ホーム画面から「ツール」「抽出語」「共起ネットワーク」を選択する。すると「抽出語・共起ネットワーク：オプション」というウィンドウが出てくるので、ウィンドウ右側の「共起ネットワークの設定」から「強い共起関係ほど濃い線に」「最小スパニング・ツリーだけを描画」のチェックボックスにチェックを入れ、「OK」を選択する（図表 9）。

図表 9 共起ネットワークのオプション選択画面

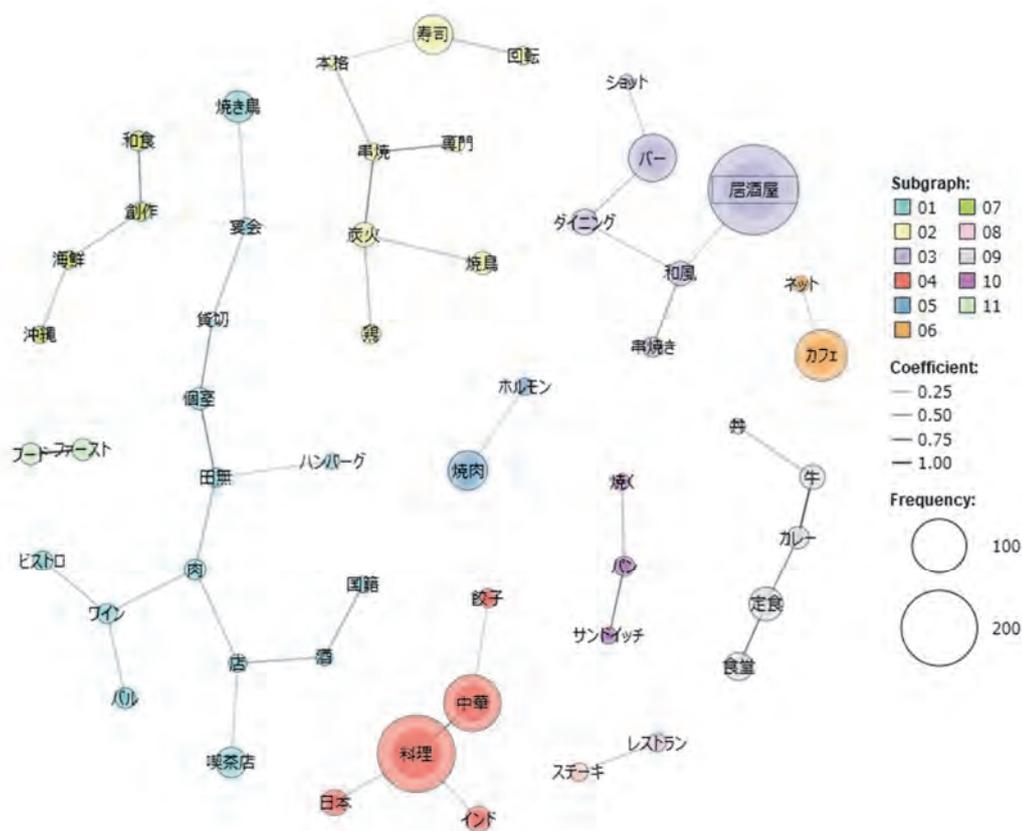
The screenshot shows the '抽出語・共起ネットワーク：オプション' dialog box. It contains the following settings:

- 集計単位と抽出語の選択**
 - 集計単位: 段落 (dropdown), サンプルング: 250000 (input)
 - 最小/最大 出現数による語の取捨選択: 最小出現数: 5, 最大出現数: (empty)
 - 最小/最大 文書数による語の取捨選択: 最小文書数: 1, 最大文書数: (empty)
 - 品詞による語の取捨選択:
 - 名詞
 - サ変名詞
 - 形容動詞
 - 固有名詞
 - 組織名
 - 人名
 - 地名
 - ナイ形容
 - 現在の設定で利用される語の数: チェック 75
- 共起ネットワークの設定**
 - 共起関係 (edge) の種類:
 - 語 - 語
 - 語 - 外部変数・見出し
 - 外部変数・見出し: 利用不可 (dropdown)
 - 描画する共起関係 (edge) の選択: Jaccard (dropdown)
 - 上位 60
 - 係数 0.2 以上
 - 係数を標準化する
 - 強い共起関係ほど濃い線に 係数を表示
 - バブルプロット: バブルの大きさ 100 %
 - 小さめの円ですべての語を描画
 - 最小スパニング・ツリーだけを描画
 - 共起パターンの変化を探る (相関)
 - 対象とする数値変数: 出現位置 (dropdown)
 - グレースケールで表現
 - フォントサイズ: 100 % 太字 プロットサイズ: 640
 - 実行時にこの画面を閉じない

出典：筆者作成

すると、どのような抽出語同士が同時に出現している（共起している）か、図として確認することができる（図表 10）。図表 10 より、「中華」と「餃子」、「居酒屋」と「和風」、「定食」と「食堂」が共起関係にあることがわかる。共起関係にある抽出語は、日本標準産業分類への格付け基準を作成する際の判断基材料となる。しかし、図表 10 から確認できるように、「料理」という抽出語が含まれていることで本来別であると考えられる「中華」と「日本」に共起関係があるように描画されてしまっている。

図表 10 共起ネットワーク(西東京市)



出典:筆者作成

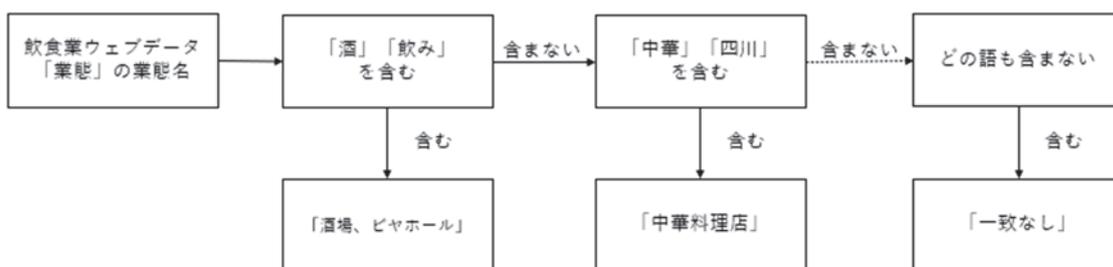
6 Excel での飲食業ウェブデータの分類

次に KH Coder での分析結果を基準に、飲食業ウェブデータ「業態」を日本標準産業分類の各細分類に分類していく。KH Coder の拡張機能を用いることで KH Coder 上でも細分類への格付けを行うことはできるが、分類したデータを其他分析ツール (R や Stata 等) でさらに分析していくことを考慮し、本稿では複数のソフトで互換性のある拡張子でデータを作成することができる Excel 上で細分類への格付けを行う。

Excel での分類を行う前に、前章まででどのような分類基準があったかを再度確認する。第 4 章では、本稿の分類に用いる 15 の細分類がどのような業務内容か、どのような語 (形態素、抽出語) が含まれているとその細分類当たるとのかについて述べた。第 5 章では、各図表で示した通り、KH Coder を用いて飲食業ウェブデータの「業態」について分析を行った。その結果、各抽出語の出現頻度や抽出語同士の共起関係を確認することができた。本章では、第 3 章・第 4 章で述べたことを結び合わせることで、飲食業ウェブデータの「業態」を日本標準産業分類の 15 の細分類へ格付けることを目的とする。

しかし、これらの産業格付け作業を手作業で行うには量が膨大である。そこで本稿では Excel の IFS 関数と COUNTIF 関数を組み合わせることで、飲食業ウェブデータの「業態」を前章までの分類基準を踏まえた上で自動的に日本標準産業分類の 15 の細分類へ格付けていく。この作業を行っていく際、図表 12 のようなフローチャートで分類を判断させる式を上述の関数を用いて作成していく。

図表 12 業態情報の日本標準産業分類への格付け判別フローチャート



出典:筆者作成

ただし産業格付けをしていく際、機械的にすべての飲食店を判別させるのは不可能である。理由は 2 つある。1 つ目は、飲食業ウェブデータの「業態」からでは日本標準産業分類の 15 の細分類への格付けが不可能な飲食店があるためである。これは「業態」にあたる部分が飲食店の従業者により手入力されているため、飲食店の業態とは全く関係のない情報が記入されている場合があるためである。これらを分類するためには各飲食店の URL から

元情報を確認する必要がある。しかし、この作業には膨大な時間がかかるためにウェブデータを含むビッグデータの機械的なデータクレンジング作業には適さない。2つ目は、本稿で用いる飲食業ウェブデータには飲食店以外にも飲食の提供を行う商業施設や精肉店、カラオケボックスなどの情報が記載されているためである。そのためこれらは図表 12 にある「一致なし」に分類し、区別する必要がある。

また作成した式を用いて年月ごとに別途産業格付け作業を実行すると、同一店舗であるのにもかかわらず時間の経過により飲食業ウェブデータ「業態」が飲食店側により書き換えられている影響で分類結果が表記揺れを起こす可能性がある。そのため、図表 13 のように、各年月のデータの先頭に「年」という時間を示す列を作成した後、比較したい年月の最も古いデータのみ（本稿の西東京市の例であれば 2018 年 7 月）を日本標準産業分類の細分類に分類し、それ以降の同飲食店の新しい年月のデータに関しては「URL」で昇順ないし降順に並び替えたうえで最も古いデータで分類した結果を張り付ける形で処理するのが良い。またさらに、この並び替えを行った際に「2018 年だけない」「2019 年のみある」といったような飲食店ごとのデータが存在する。これは飲食店の参入退出の影響によるものである。

図表 13 同飲食店の時系列での並び替え(西東京市)

A	B	C	D	E	F	G	H	I
年	URL	店名	有無	住所	業態	緯度/経度	アクセス	業態判別
2018	https://r.g	山城	0	188-0014	中華料理	35.733845	西武新宿線	中華料理店
2019	https://r.g	山城	1	188-0014	中華料理	35.733845	西武新宿線	中華料理店
2020	https://r.g	山城	1	188-0014	中華料理	35.733845	西武新宿線	中華料理店
2018	https://r.g	長寿庵	1	202-0005	そば	35.744125	西武池袋線	そば・うどん店
2019	https://r.g	長寿庵	0	202-0005	そば	35.744125	西武池袋線	そば・うどん店
2020	https://r.g	長寿庵	0	202-0005	そば	35.744125	西武池袋線	そば・うどん店
2018	https://r.g	ホルモンハウス	1	202-0022	居酒屋	35.728462	西武新宿線	酒場、ビヤホール
2019	https://r.g	ホルモンハウス	0	202-0022	居酒屋	35.728462	西武新宿線	酒場、ビヤホール
2020	https://r.g	ホルモンハウス	0	202-0022	居酒屋	35.728462	西武新宿線	酒場、ビヤホール
2018	https://r.g	食通天	1	202-0001	中華料理	35.751557	西武池袋線	中華料理店
2019	https://r.g	食通天	1	202-0001	中華料理	35.751557	西武池袋線	中華料理店
2020	https://r.g	食通天	1	202-0001	中華料理	35.751557	西武池袋線	中華料理店
2018	https://r.g	居酒屋 せん	1	202-0002	居酒屋	35.752968	西武池袋線	酒場、ビヤホール
2019	https://r.g	居酒屋 せん	1	202-0002	居酒屋	35.752968	西武池袋線	酒場、ビヤホール
2020	https://r.g	居酒屋 せん	0	202-0002	居酒屋	35.752968	西武池袋線	酒場、ビヤホール
2018	https://r.g	本格焼きたりとワイン 門福 ひばり	0	202-0002	居酒屋	35.753506	西武池袋線	酒場、ビヤホール
2019	https://r.g	本格焼きたりとワイン 門福 ひばり	1	202-0002	居酒屋	35.753506	西武池袋線	酒場、ビヤホール
2020	https://r.g	本格焼きたりとワイン 門福 ひばり	1	202-0002	居酒屋	35.753506	西武池袋線	酒場、ビヤホール

出典:筆者作成

飲食業ウェブデータの分類結果を用いてさらに分析を行う際は、必要に応じて変数（項目）を飲食業ウェブデータに追加する必要がある。例えば参入・退出を各飲食店で把握したい場合は、その年月に飲食店のウェブページが存在したか否かを「0」「1」で示す「有無」の列を追加し、飲食店情報が存在しない年月のデータを「0」とする。そして2時期並んだ

「有無」が古い年月から「0」「0」なら飲食店なし、「1」「0」なら退出、「0」「1」なら参入、「1」「1」なら存続と判断することで参入・退出に関する変数（項目）を作成することができる。

7 今後の課題

本稿では上述のようにテキストマイニングツール（KH Coder）と Excel を用いて、飲食業ウェブデータの「業態」の表記揺れを補完し、日本標準産業分類の 15 の細分類へ格付ける手法の開発を行った。しかしこの手法には課題がある。それは Excel の関数を用いた飲食業ウェブデータの「業態」の分類では Excel の仕様上文字数に限界がある点である。図表 12 ではフローチャートによる Excel の関数を用いた処理の一例を示しているが、日本標準産業分類への飲食業ウェブデータの「業態」の分類の精度を上げていくためには文字数が増加していく。しかし、Excel を提供する Microsoft のサポートページによると、数式の長さの上限は 8,192 文字までであり²⁷、これを超える数式の文字数をセルに入力するとエラーが表示され、関数を実行することができない。

この問題を解決するためには文字数に制限されない方法での産業格付け手法の開発が必要である。そこで VBA（Visual Basic for Applications）や Python などのプログラミング言語を用いた方法が考えられる。飲食業ウェブデータの「業態」をプログラミング言語で処理することで日本標準産業分類へ分類し、Excel 上に分類結果を表示させることができれば、課題である文字数の制限は解決する可能性が高い。そのため今後はプログラミング言語を用いた飲食業ウェブデータの「業態」の日本標準産業分類への産業格付けプログラムの開発が考えられる。

参考文献

- 大隅昇、L. Lebart. (2000) 「調査における自由回答データの解析 —infoMiner による探索的テキスト型データ解析—」『統計数理』統計数理研究所、第 48 巻第 2 号、pp. 339-376。
- 佐藤裕 (1992) 「自由回答のコンピュータコーディング」『第 14 回数理社会学会大会研究報告要旨集』数理社会学会。
- 菅幹雄 (2017) 「飲食業ウェブデータを用いたローリング・センサスのシミュレーション」『研究所報』法政大学日本統計研究所、第 49 号、pp. 45-57。
- 総務省「日本標準産業分類の変遷と第 13 回改訂の概要」
URL:https://www.soumu.go.jp/toukei_toukatsu/index/seido/sangyo/02toukatsu01_03

²⁷ Microsoft 「Excel の仕様と制限」

URL:<https://support.microsoft.com/ja-jp/office/excel-%E3%81%AE%E4%BB%95%E6%A7%98%E3%81%A8%E5%88%B6%E9%99%90-1672b34d-7043-467e-8e27-269d656771c3>（最終閲覧日：2022 年 10 月 2 日）

000043.html (最終閲覧日:2022年5月5日)

総務省「日本標準産業分類一般原則」

URL:https://www.soumu.go.jp/main_content/000286955.pdf (最終閲覧日:2022年5月5日)

総務省「大分類M-宿泊業,飲食サービス業」

URL:https://www.soumu.go.jp/main_content/000290732.pdf (最終閲覧日:2022年5月6日)

総務省「大分類N-生活関連サービス業,娯楽業」

URL:https://www.soumu.go.jp/main_content/000290733.pdf (最終閲覧日:2022年5月6日)

総務省統計局「サービス産業動向調査」

URL:<https://www.stat.go.jp/data/mssi/index.html> (最終閲覧日:2022年11月10日)

田中重人・太郎丸博(1996)「文章データのコンピュータコーディング—プログラムとその応用—」『第22回数理社会学会大会研究報告要旨集』数理社会学会、pp.20-23。

内閣府(2015)「公的統計におけるビッグデータの活用に関する調査研究」、平成26年度内閣府大臣官房統計委員会担当室請負調査、NTTデータ経営研究所。

中田理恵・溝端幹雄(2020)「ビッグデータの公的統計への貢献と課題」『大和総研調査季報』大和総研リサーチ本部、第40巻、66-83。

樋口耕一(2004)「テキスト型データの計量的分析—2つのアプローチの峻別と統合—」『理論と方法』数理社会学会、第19巻第1号、pp.101.115。

Iker, H. P. & N. I. Harway. 1969. "A Computer Systems Approach toward the Recognition and Analysis of Content." in G. A. Gerbner, O. R. Holsti, K. Krippendorff, W. J. Paisly & P. J. Stone (eds.) *The Analysis of Communication Content: Developments in Scientific Theories and Computer Techniques*. New York: Wiley & Sons.: 381.486.

Retty「Food Data Platform」

URL:<https://campaign.retty.me/fooddataplatfrom/> (最終閲覧日:2022年11月10日)

Stone, P. J., D. C. Dunphy & D. M. Ogilvie. 1966. *The General Inquirer: A Computer Approach to Content Analysis*. Cambridge: MIT Press.

オケージョナル・ペーパー(既刊一覧)

号	タイトル	刊行年月
100	第1回国勢調査が記録した社会移動 —生涯移動から見た転入移動圏の特徴を中心に—	2019.09
101	第1回国勢調査の出生地データによる県間生涯移動分析	2019.08
102	わが国の1980年代後半期以降の社会移動に関する一考察 —純移動選好度の人口加重平均値による地域の転入・ 転出超過状況の評価—	2019.09
103	QGISによる西武国分寺線沿線の産業構造分析Ⅱ	2020.02
104	明治2年駿河国人別調における静態把握と動態把握	2020.02
105	地域勘定における一般政府勘定について	2020.04
106	駿河国人別調と甲斐国現在人別調における人口の静態把握 —家別表の調査項目の比較を中心に—	2020.05
107	地租改正にともなう土地評価の改定: 東京府日本橋区・京橋区の事例	2020.05
108	駿河国人別調と甲斐国人員運動調における動態把握 —わが国における人口動態統計前史(1)—	2020.05
109	甲斐国人員運動調について —わが国における人口動態統計前史(2)—	2020.06
110	明治4年「一般戸籍の法」における人口の社会動態の把握	2020.06
111	明治前期の戸籍法制と社会移動の統計的把握 —明治4年「戸籍の法」による社会移動把握の制度改定を中心に—	2020.07
112	甲斐国現在人別調における人口概念	2020.07
113	明治初期における物産調査の展開 —明治16年農商務通信規則成立前史—	2020.09
114	明治16年農商務通信規則について	2020.09
115	明治16年農商務通信規則による工業通信事項と附録様式	2020.09
116	東京湾岸地域の人口増加と郵便局の考察	2021.01
117	明治16年農商務通信規則の史的系譜 —農事通信三規則との比較を中心に—	2021.02
118	高齢化状況による介護保険サービスの利用料量の地域差	2021.07
119	自動車保有台数からみるSS過疎地に関する考察	2021.08
120	高齢化率と経年変化からみる事業所数モデルの考察	2021.09
121	甲斐国現在人別調における調査過程	2021.10
122	明治16年農商務通信規則による工場票情報の情報特性	2022.03

オケージョナル・ペーパー No.123

2022年12月20日

発行所 法政大学日本統計研究所

〒194-0298 東京都町田市相原4342

Tel 042-783-2325、2326

Fax 042-783-2332

jsri@adm.hosei.ac.jp

発行人 菅 幹雄

再生紙使用

