

CAとMCAの例

2025/09/04
計量分析セミナー

藤本一男

kazuo.fujimoto2007@gmail.com

CAとMCAの処理のさわり

- これからお話しするCA、MCAは、どんな処理なのか簡単にみておきます。
- CAの事例
 - データ：smoke
 - CAの結果
 - コードの詳細は、C01_1_CA_step1_smoke.Rmd resultはhtml
- MCAの事例
 - データ：taste_j
 - MCAの結果
 - コードの詳細は、C02_1_C02_1_CA_MCA_Q16.Rmd resultはhtml
- これらのresultの詳細は、追って説明していきます。

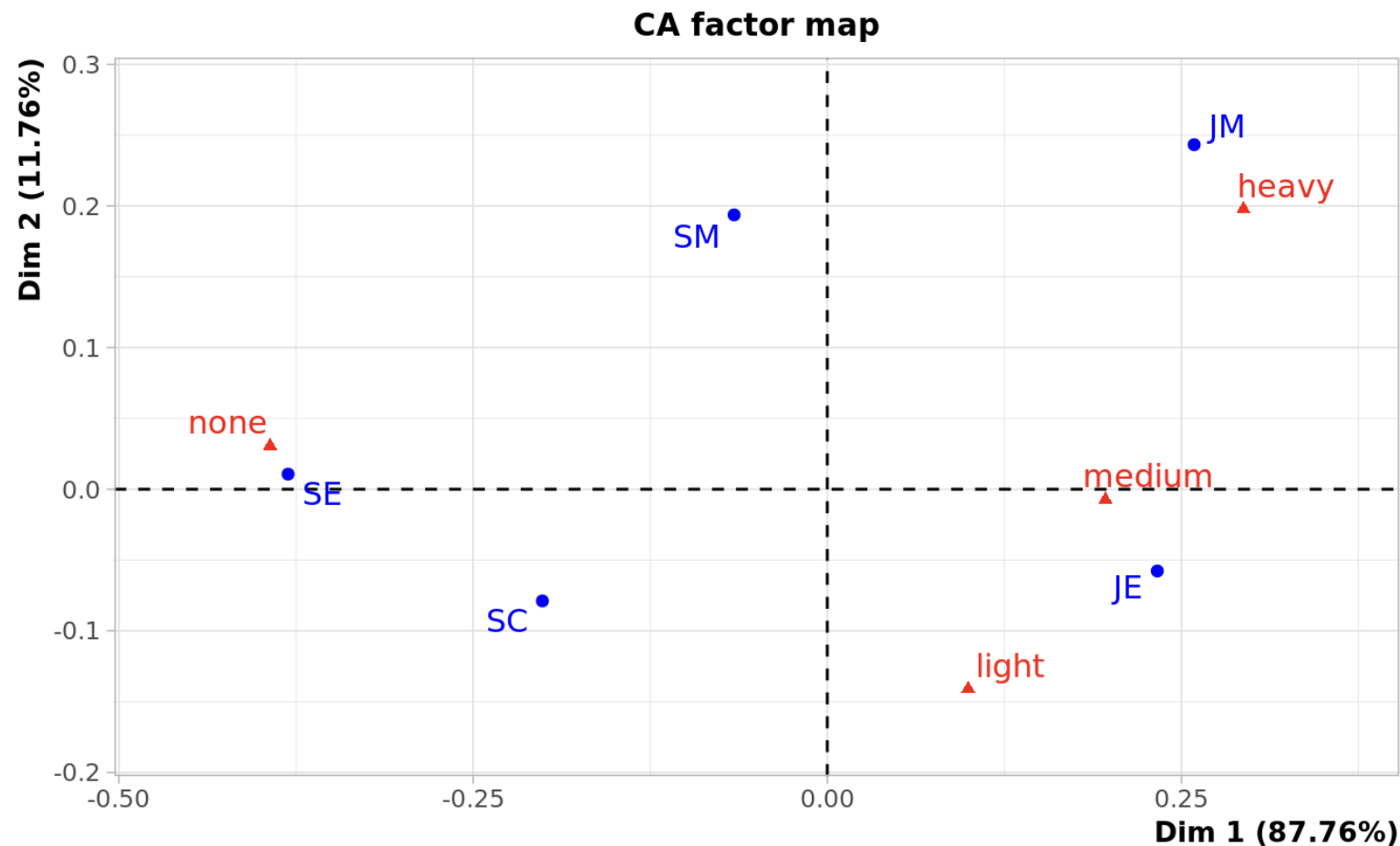
CAの事例 smoke

まずはさらっと！

データ smoke

• 入力するクロス表

	none	light	medium	heavy
SM	4	2	3	2
JM	4	3	7	4
SE	25	10	12	4
JE	18	24	33	13
SC	10	6	7	2



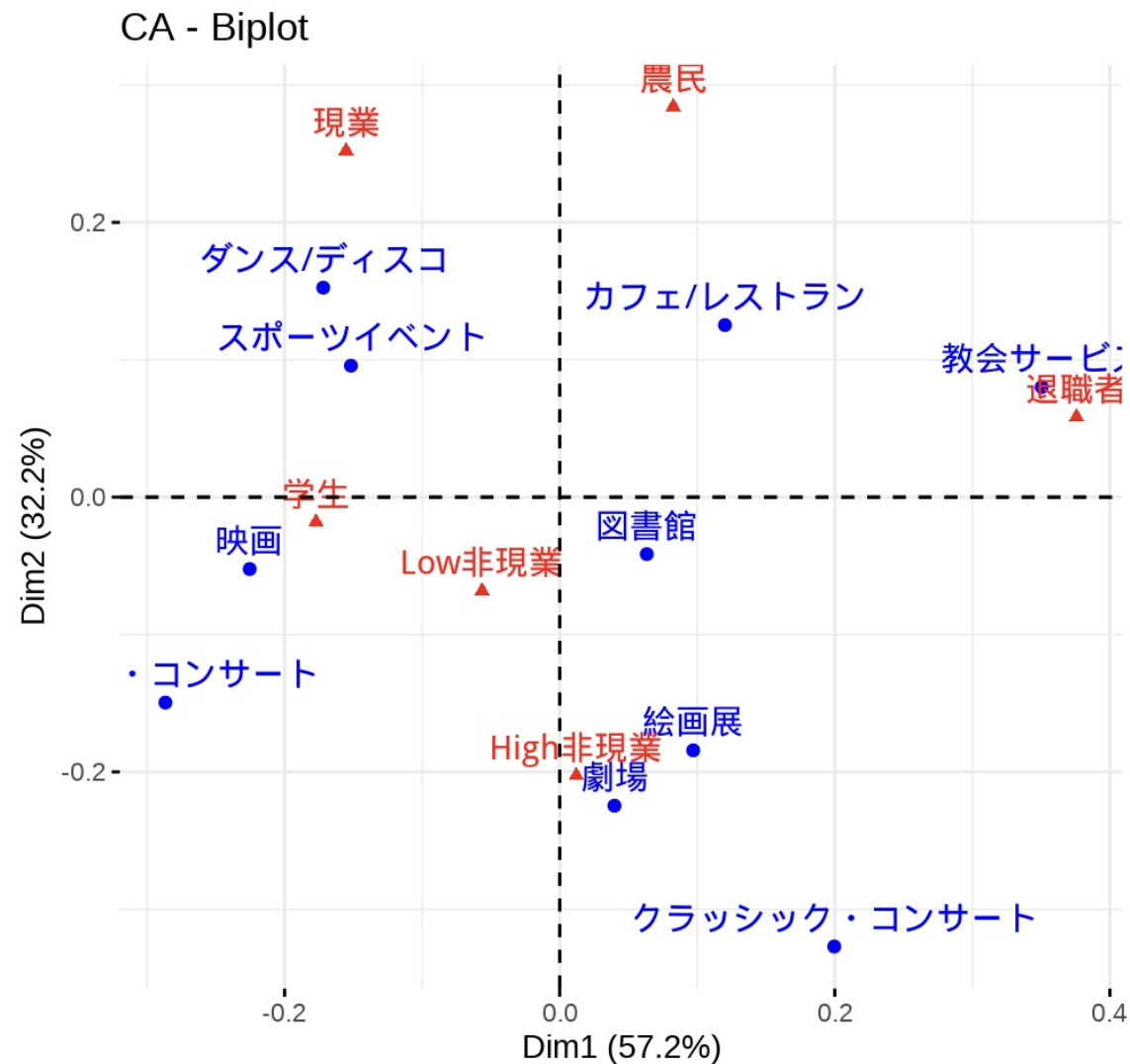
CAの事例 職種と余暇

Clausen 1987=2015の第一章のデータ

・データ

	現業	Low非現業	High非現業	農民	学生	退職者
スポーツイベント	301	497	208	50	254	187
映画	261	550	250	27	339	157
ダンス/ディスコ	361	534	204	59	324	216
カフェ/レストラン	463	766	334	72	350	601
劇場	89	350	195	12	143	167
クラシック・コンサート	23	182	124	10	60	110
ポップ・コンサート	117	298	145	11	184	56
絵画展	104	379	219	21	152	213
図書館	130	352	153	17	272	264
教会サービス	168	370	187	51	162	424

対称マップ（行変数と列変数の同時付置）



MCAの事例 「嗜好データ」

Le Roux & Rouanet 2010=2021で事例として使われているデータ。このデータは、ベネット他（訳：磯他）『文化・階級・卓越化』で分析されたデータのサブセット。

taste_J

- データの構成
 - 行 個体（回答者 1～1253）
 - 列 変数

	ID	Isup	TV	Film	Art	Eat	Gender	Age	Income
1	1	Active	TV-メロドラマ	映画-アクション	芸術-風景画	外食先-ステーキハウス	女性	55-64	£20-29
2	2	Active	TV-メロドラマ	映画-ホラー	芸術-静物画	外食先-インド料理店	女性	45-54	<£9
3	3	Active	TV-自然	映画-アクション	芸術-風景画	外食先-パブ	女性	55-64	<£9
4	4	Active	TV-メロドラマ	映画-時代劇	芸術-肖像画	外食先-イタリア料理店	女性	65+	£10-19
5	5	Active	TV-コメディ	映画-ホラー	芸術-静物画	外食先-インド料理店	女性	35-44	£10-19
6	6	Active	TV-コメディ	映画-ホラー	芸術-印象派	外食先-インド料理店	女性	18-24	<£9
7	7	Active	TV-ニュース	映画-アクション	芸術-風景画	外食先-インド料理店	女性	25-34	£10-19
8	8	Active	TV-ニュース	映画-ドキュメンタリー	芸術-パフォーマンス・アート	外食先-フィッシュ&チップス	男性	65+	£10-19
9	9	Active	TV-メロドラマ	映画-時代劇	芸術-風景画	外食先-ステーキハウス	女性	65+	<£9
10	10	Active	TV-ニュース	映画-アクション	芸術-風景画	外食先-フィッシュ&チップス	女性	65+	£10-19

Showing 1 to 10 of 1,253 entries

Previous

1

2

3

4

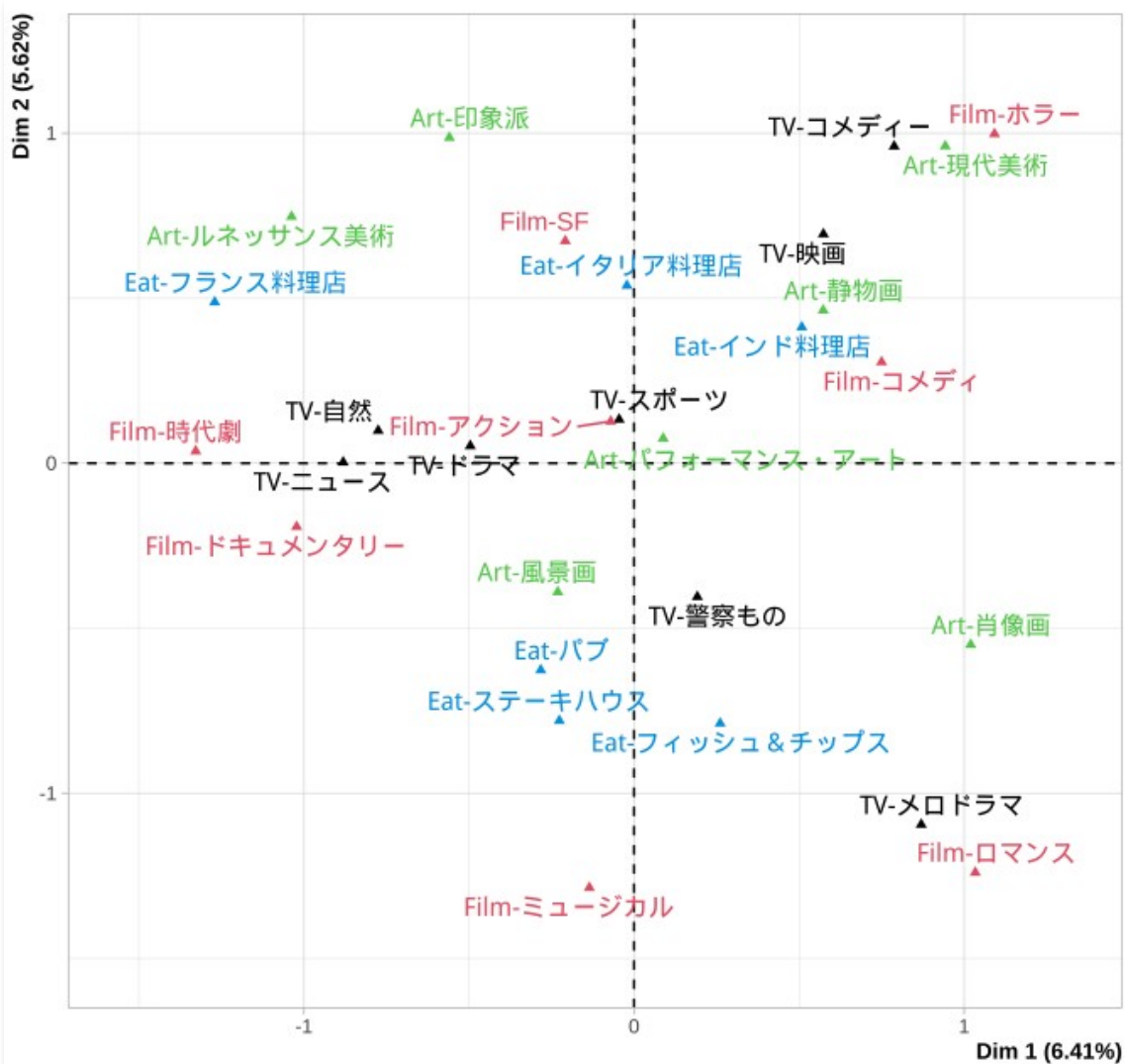
5

...

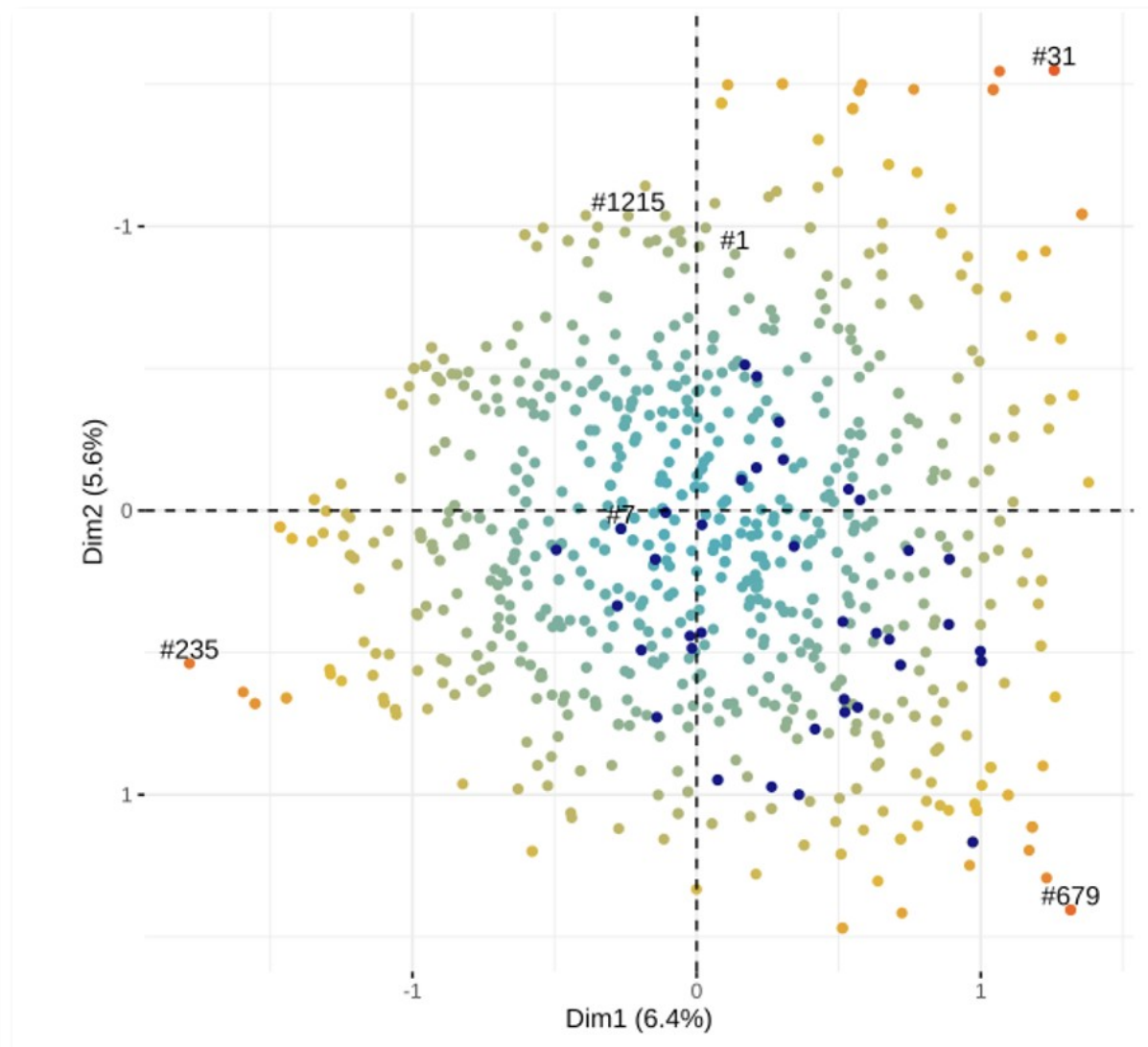
126

Next

変数カテゴリマップ



個体マップ



まとめ

- 二元表が入力 -> 二つの空間が生成される
 - 行空間/個体空間
 - 列空間/変数（カテゴリ）空間
- 生成された空間の座標軸は主軸と呼ばれる
 - 主軸技法
 - 分析対象のデータの分散をもっとも多く体現するところに第1軸（dim1）をひく。つぎに、この軸に直行する形で次にもっとも多く分散を体現する方向に第2軸を引く（dim2）。こうして、dim3、dim4、dim...と引いていく。
 - この軸は、各カテゴリポイントが影響をおよぼしている（寄与）。
 - その影響の大きさを方向を踏まえて、軸の解釈を行う。

数理的な基礎の確認

- 入力するデータは、多次元行列。 m 行、 n 列。
- この行列の持っている情報（分散）を100%表現するには、 m と n の小さい方から1を引いた次元の空間があれば可能。
 - somoke $5 \times 4 \rightarrow 4 - 1 = 3$ 次元で100%表現
 - 職種と余暇 $10 \times 6 \rightarrow 6 - 1 = 5$ 次元あれば100%表現
 - 嗜好データ $1215 \times 29 \rightarrow 29 - 1 = 28$ 次元あれば..
- こうした多次元行列を数学的な手法で表現する分散の大きさの順に次元縮減する。
- それぞれの軸ごとに何パーセントの情報（分散）を表現しているのか、で分析する次元を確定していきます。
- その仕組みを次にみていきましょう。