

第8章：さまざまなデータマイニング手法

1元配置分散分析

下表は、各都市のワカメの消費金額データです。

出典：[【XLS】「家計調査\(二人以上の世帯\)：品目別都道府県庁所在市及び政令指定都市ランキング」](#)

- 東日本 (E)、中日本 (C)、西日本 (W)の3地域で、「ワカメの消費金額は地域による差がない」を帰無仮説とする仮説検定を行い、結果を判定してみましょう。

cons_wakame	city	region
2,236	盛岡市	E
1,455	仙台市	E
1,419	長野市	C
1,399	青森市	E
1,369	秋田市	E
1,269	福島市	E
1,220	山形市	E
1,208	新潟市	E
1,130	静岡市	C
1,112	宇都宮市	E
1,081	さいたま市	E
1,076	相模原市	E
1,060	水戸市	E
1,002	横浜市	E
993	甲府市	C
990	徳島市	W
971	千葉市	E
969	長崎市	E
945	福井市	C
939	浜松市	C
937	前橋市	E
883	富山市	C
882	奈良市	C
872	鳥取市	W
866	北九州市	W
866	東京都区部	E
866	高松市	W
806	佐賀市	W

cons_wakame	city	region
796	川崎市	E
794	熊本市	W
794	岐阜市	C
779	松江市	W
778	岡山市	W
767	広島市	W
765	神戸市	C
724	金沢市	C
712	高知市	W
710	堺市	C
708	名古屋市	C
692	津市	C
681	大津市	C
655	福岡市	W
633	和歌山市	C
587	大阪市	C
585	大分市	W
583	京都市	C
564	松山市	W
555	札幌市	E
545	山口市	W
488	鹿児島市	W
451	那覇市	W
423	宮崎市	W

因子分析

下表は1996年8月31日の日経流通新聞に載っていた、首都圏在住の15歳以上の女性モニター454人に聞いた「女性が選ぶ海外ブランド」に関するアンケート調査データです。

このデータを用いて、2軸の因子分析を行うことで、それぞれのブランドがどのような立ち位置（ポジション）にあるか分析してみましょう。

popularity: 人気度, awareness: 認知度, possession: 所有率, luxury: 高級感, proudness: 誇らしさ, reliability: 品質の信頼性, fashionability: センスのよさ, friendliness: 親しみやすさ, advertisement: 広告の魅力

brands	pop	awa	pos	lux	pro	rel	fas	fri	adv
CHANEL	159	377	209	318	136	150	123	36	86
HERMES	145	327	136	245	104	154	127	27	41
TIFFANY	145	327	136	182	86	136	136	77	59
LOUIS_VUITTON	136	359	186	177	77	186	82	109	18
GUCCI	123	350	154	163	73	141	114	68	32
POLO_RALPH_LAUREN	114	295	200	54	27	114	91	154	36
CARTIER	109	291	109	232	95	150	95	14	23
FERRAGAMO	109	286	68	159	64	109	77	32	18
PRADA	104	245	45	104	50	77	82	59	18
CALVIN_KLEIN	100	263	123	32	23	64	118	132	54
BENETTON	86	327	241	18	5	54	59	227	95

- 分析結果の2次元プロットについて、あなたなら第1軸と第2軸の成分にどのように名前をつけるか考えてみましょう。