

楽天市場におけるレビューの総合的な統計・分析

経営学科 3年

山根拓真
高田真成

研究の動機

- 「レビュー」は現代の購買意思決定における重要な指標
- “サクラレビュー” に意味はあるのか？
- レビューの長所・短所を理解したい
- レビューを利用した新たなサービスなどの可能性を考えたい

進行中の内容

- 膨大なデータのため、コーヒーという単語を含むレビューデータだけを抽出
- “MeCab” という形態素解析エンジンにてレビューを品詞別に分解
- 用途別の出現頻度をソートしてグラフに起こし、相関性が無いかなどを調査

形態素解析とは

- 文法的な情報の注記の無い自然言語のテキストデータから、対象言語の文法や、辞書と呼ばれる単語の品詞等の情報にもとづき、形態素の列に分割し、それぞれの形態素の品詞等を判別する作業である。

すももももももものうち

すもも 名詞,一般,*,*,*,*,すもも,スモモ,スモモ

も 助詞,係助詞,*,*,*,*,も,モ,モ

もも 名詞,一般,*,*,*,*,もも,モモ,モモ

も 助詞,係助詞,*,*,*,*,も,モ,モ

もも 名詞,一般,*,*,*,*,もも,モモ,モモ

の 助詞,連体化,*,*,*,*,の,ノ,ノ

うち 名詞,非自立,副詞可能,*,*,*,うち,ウチ,ウチ

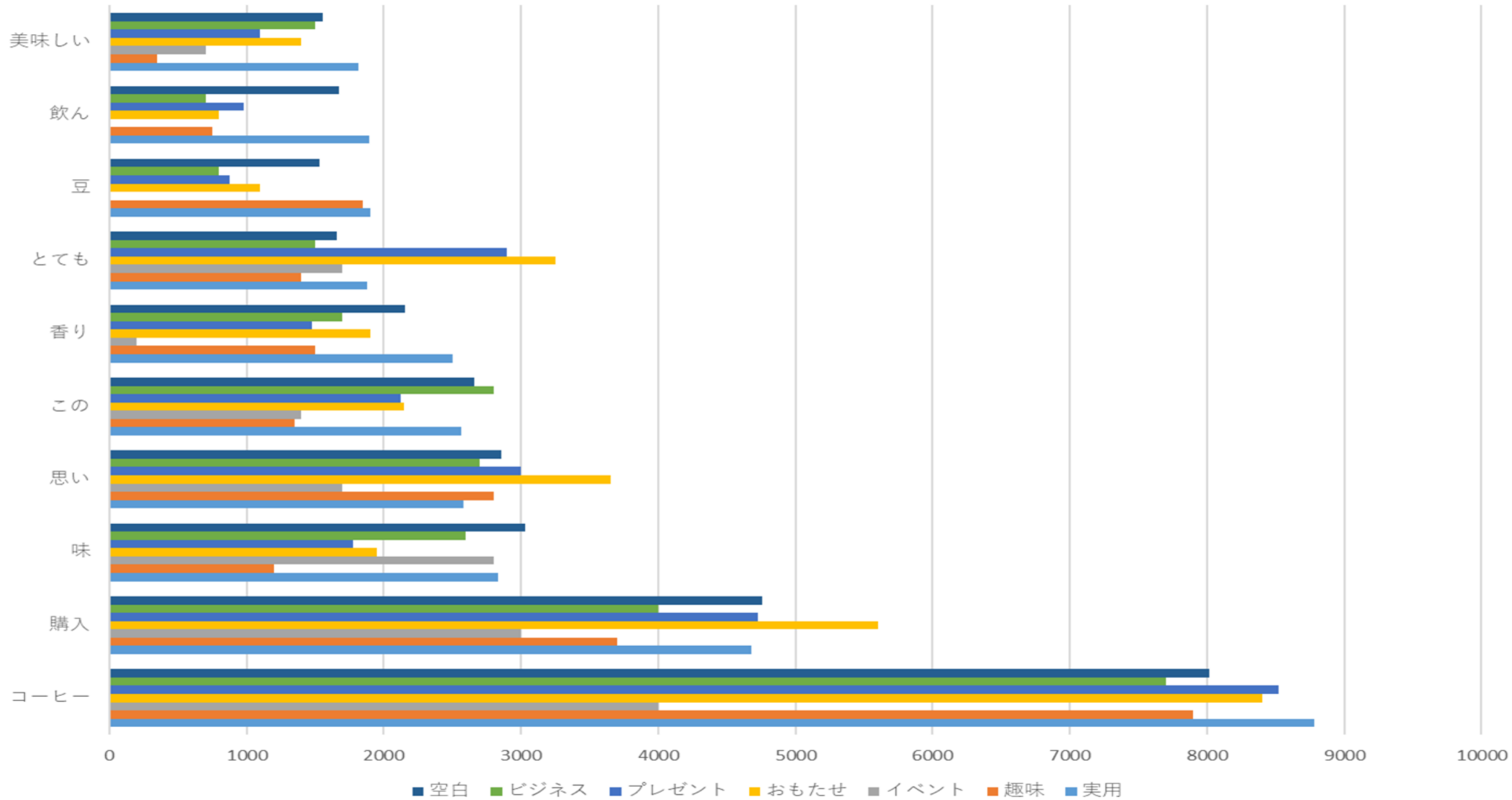
EOS

出現頻度の比率

7200/0.82 数字の重みを揃える

	82%	2%	1%	2%	4%	1%	7%								
全体	実用	趣味	イベント	おもたせ	プレゼント	ビジネス	空白		実用	趣味	イベント	おもたせ	プレゼント	ビジネス	空白
8738 コーヒー	7200	158	40	168	341	77	561	コーヒー	8780.488	7900	4000	8400	8525	7700	8014.2
4745 購入	3834	74	30	112	189	40	333	購入	4675.61	3700	3000	5600	4725	4000	4757.1
2786 味	2320	24	28	39	71	26	212	味	2829.268	1200	2800	1950	1775	2600	3028.5
2680 思い	2114	56	17	73	120	27	200	思い	2578.049	2800	1700	3650	3000	2700	2857.1
2533 この	2100	27	14	43	85	28	186	この	2560.976	1350	1400	2150	2125	2800	2657.1
2395 香り	2051	30	2	38	59	17	151	香り	2501.22	1500	200	1900	1475	1700	2157.1
1945 とても	1538	28	17	65	116	15	116	とても	1875.61	1400	1700	3250	2900	1500	1657.1
1794 豆	1559	37	#N/A	22	35	8	107	豆	1901.22	1850	#N/A	1100	875	800	1528.5
1794 飲ん	1554	15	#N/A	16	39	7	117	飲ん	1895.122	750	#N/A	800	975	700	1671.4
1741 美味しい	1491	7	7	28	44	15	109	美味しい	1818.293	350	700	1400	1100	1500	1557.1

グラフ タイトル



現在の課題

- レビュー研究という分野は、すでに開拓が進んでいる



独自性のある研究というのは難しい？

- コーヒーという有形の商品で研究しているが、サービスなどの無形商品に当てはめられるのか？

今後の方向性・展望

- 感情分析を含む形態素解析
- 「レビューによる購買意思決定」のモデル化
- レビューによる印象操作の有意性の調査



数字だけでなく、言葉、意思決定など別方面からのアプローチ