

# スコッチウイスキーのフレーバーデータの可視化による 蒸留所の分類とその評価

## Distillery Classification and Evaluation by Visualization of Scotch Whisky Flavor Data

(2022年3月31日受理)

藤原 美佳 梶西 将司 福森 護  
Mika Fujiwara Shoji Kajinishi Mamoru Fukumori

**Key words** : 漢字グラフ, 拡張型漢字グラフ, クラスタ分析, スコッチウイスキー

本研究の目的は、拡張型漢字グラフをはじめとするグラフ手法を用いてスコッチウイスキーのフレーバーデータを可視化し、その評価を行うことである。分析には、スコットランドのストラスクライド大学が提供するスコッチウイスキー86銘柄のフレーバーに関するオープンデータを用いた。拡張型漢字グラフを適用した結果、スコッチウイスキーのフレーバーの違いや特徴が明確に示された。

## 1. はじめに

### 1-1. スコッチウイスキー

ウイスキーの生産地としては、スコットランド、アイルランド、日本、アメリカ、カナダが世界の5大ウイスキーとしてあげられる。スコッチウイスキーは、約100カ所の蒸留所を持ち、ウイスキーの生産量の6割を占める。スコッチウイスキーは、一般的に地形や土地の条件によって6つの地域に分類され、地域や蒸留所ごとの多種多様な自然環境により、さまざまな個性を持つウイスキーが醸造されている(土屋, 2018)(肥土, 2017)。ウイスキーの先行研究は、醸造に関する化学的な研究や官能評価と成分の分析(増田・小林, 1993)(戸塚, 1989)、ストレスとの関連(好田, 2015)など数多いが、地域特性とフレーバーとの関連性に関する統計的な検討は、ほとんど行われていない。

一般的にスコッチウイスキーは、スペイサイド、ハイランド、ローランド、キャンベルタウン、アイラ、アイランズの6つの地域に分類される。

スペイサイド スコッチのモルトウイスキーの半数近くにあたる50あまりの蒸留所が集中する、スコットランド最大のウイスキー生産地。エレガントで花のように華

やかなバランスに優れた銘酒が揃う。

ハイランド スペイサイドを除く40あまりの蒸留所がハイランドモルトに分類される。個性豊かな蒸留所が点在しており、地域全体としての共通の特徴を見いだすことは難しいエリアとされる。

ローランド 穏やかな気候風土の影響からか、ハイランドやスペイサイドに比べて飲みやすく繊細なライトタイプで、麦芽の香るやや辛口の風味が特徴。都会的な味わいである。

キャンベルタウン 「ブルニー」と表現される、麦芽風味のなかに感じる独特の塩辛さが特徴。港町の影響をストレートに受けている印象の味わいである。

アイランズ 現在、主な蒸留所は約7カ所で、それぞれの蒸留所が異なる個性をもっており、「島育ちのウイスキー」という地理的なくくり方と考えられる。

アイラ 全ての蒸留所が海沿いに建てられており、海岸沿いという立地や、アイラ島のピートが海藻を含んでいることが、ウイスキーに海の香りをもたらしている。独特のスモーキーフレーバーで有名。「潮の香り」をもつ個性的な味わいである。

## 1-2. 多変量グラフ表現

多変量グラフ表現は、データのもつ情報の要約や細かい特徴を直観的に把握するための可視化の方法であり、数量的に得られた結果を探索的に解釈するための手段として有効である（脇本 他, 1979）。

データの記述を主体とした代表的なグラフ表現として、円グラフや棒グラフ、散布図、レーダーチャートなどがある。さらに応用的なグラフ表現として、変数の値を目の傾き、口の位置、眉の長さなどの顔の部分に割り当てて、変数全体の特徴を顔の構造で表現する顔型グラフ（Chernoff, H., 1973）、三角多項式により曲線で表現する Andrews 曲線（Andrews, D. F., 1972）、変数を木の枝の傾きや葉の数、根の長さなどに対応させ、相関の度合を視覚的に判定できるようにした木型グラフ（Wakimoto, K., 1977）、変数を半円内にベクトルで表しそれらのベクトルを連結してその最終点に星を描く星座グラフ（Wakimoto, K., 1978）、漢字の縦と横のサイズに変数の値を割り当てて表現する漢字グラフ（平井 他, 1988）など、様々なグラフ手法が提案されている。多変量解析の結果を可視化したグラフ表現としては、クラスター分析におけるデンドログラム、主成分分析や因子分析において二次元の座標軸に変数やサンプルをプロットするグラフ表現、重回帰分析の統計量をグラフ表現したボンサイグラフ（石村, 2013）などがある。

これらのグラフ表現は、記憶や印象、直観に訴えかける方法であり、その導出と計算が複雑ではなく、客観性があり、結果の解釈が容易にできるといった利点がある。また、複数のグラフ表現を組み合わせることで、単一のグラフとは異なった観点からの解釈が可能になる（脇本 他, 1979）。

## 1-3. 研究目的

本研究では、フレーバーデータの可視化を行うことにより、スコッチウイスキーの地域の特徴を明らかにし、フレーバーによる蒸留所の分類と評価を行うことを目的とする。

### 分析1. 蒸留所の位置情報を用いたフレーバーの地域特性の評価

蒸留所の位置情報を用いて、スコットランドの地図上に、蒸留所をプロットし、フレーバーの要素である

Body, Sweetness, 香りの視点から地域特性を可視化する。  
分析2. フレーバーデータによる蒸留所の分類とその評価

フレーバーデータによりクラスター分析を行い、クラスターごとの特徴を拡張型漢字グラフで可視化する。また、正準判別分析により、地域による分類とクラスターによる分類を比較検討する。

## 2. 方法

### 2-1. 漢字グラフ

漢字グラフは、平井 他（1988）により提案されたグラフで、変数を一文字の漢字で表現し、漢字の大きさに変数の値を割り当てて可視化する手法である。

漢字グラフは以下のような手順で作成する。

- ①  $p$  個の変数を持つデータ  $x_{ij}$  ( $i=1, 2, \dots, n$ ,  $j=1, 2, \dots, p$ ) に対して次のような変換を行う。

$$x'_{ij} = \frac{x_{ij} - \bar{x}_j}{S_j} + b$$

ここで、 $b$  は正の定数で常に、 $x'_{ij} \geq 0$  である。なお、 $\bar{x}_j$  は平均、 $S_j$  は標準偏差である。

- ②  $p$  種類の一文字の漢字を用意する。
- ③ 漢字の縦と横の長さに対して、 $x'_{ij}$  に任意の倍率  $\theta$  を掛けた  $x'_{ij}$  の値を割り当てた大きさの漢字を描く。
- ④  $p$  番目の漢字に至るまで同様の手順で漢字を左から右に一行に配置して漢字の文字列を作る。
- ⑤ この手順を全個体について行う。

漢字グラフの利点として、次の点が挙げられる。

1. 漢字という親しみやすいデータの性質を表す文字で変数の大きさを表示できる。
2. グラフの形状によりデータの分類や特徴の分析が容易に出来る。
3. 複数の変数を同時に表現でき、データの特徴が読み取りやすい。
4. 変数の全体の長さにより総合評価が可能になる。

図1に漢字グラフの表示例として、平井 他（1988）で用いられた漢字グラフを示す。

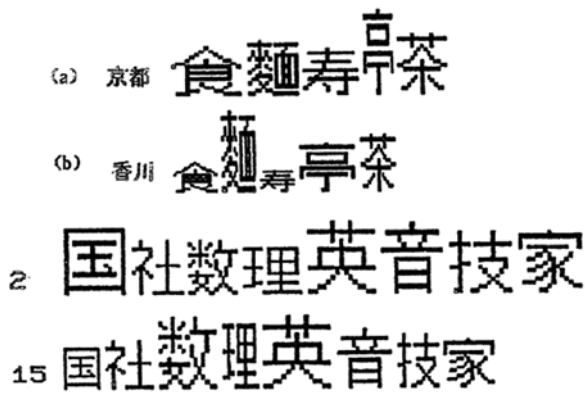


図1. (a) (b) 飲食関係のデータで京都府と香川県の漢字グラフを表示したもの. (2) (15) テスト成績のデータで2番と15番の生徒を漢字グラフで表示したもの

図1のように、漢字を左から右に一列に配置することにより、全体の文字列の長さから変数の総合評価の比較が可能になる。

しかし、群の平均値を用いる場合には、変数間で同じ平均であった場合でも散布度によって解釈が異なってくるため、本章では、漢字グラフを拡張して散布度（ここでは標準偏差）のレーダーチャートと漢字グラフを組み合わせた拡張型漢字グラフを紹介する。

## 2-2. 拡張型漢字グラフ

拡張型漢字グラフは、Fujiwara et al., (2021) により提案されたグラフで、各変数の標準偏差  $S_j$  によりレーダーチャートを作成し、レーダーチャート上に主成分分析の負荷量に基づいて漢字を配置する。このグラフにより、各群の平均と散布度を可視化し、その群を代表する特徴について分析することが可能になる。また、項目を漢字一文字で表現することにより項目の性質を直観的に可視化することができるため、群の特徴に関する解釈が容易に行えることが期待される。

拡張型漢字グラフの利点として、次の点が挙げられる。

1. 漢字の大きさから、変数の大きさを直観的に把握できる。
2. レーダーチャートにより、クラスタ内の各変数の散布度を把握できる。
3. 変数の性質を示す漢字を割り当てることにより、直観的に変数の意味を把握することができる。

グラフ作成の手順は次の通りである。

- ① レーダーチャートの変数の並び順を決定するために、主成分分析を行う。主成分分析の第1主成分と第2主成分の座標軸上に主成分負荷量をプロットし、その布置を参考にレーダーチャート上に変数を配置する。
- ②  $S_j$  を用いてレーダーチャートを作成する。
- ③ 各変数の性質を表現できる  $p$  個の漢字を決める。
- ④ 通常の漢字グラフと同様の手順により、漢字の縦と横の長さを決める。本研究のように、全ての変数において同一の評価基準である場合は、変数の値に任意の倍率  $l$  を掛けることにより文字のサイズを決める。
- ⑤ 縦と横に長さを割り当てた漢字をレーダーチャート上に配置する。

## 2-3. 使用データ

スコットランドのストラスクライド大学が提供するスコッチウイスキー86銘柄のフレーバーに関するオープンデータを用いた。フレーバーは、12種類 (Medicinal, Smoky, Body, Spicy, Winey, Nutty, Malty, Honey, Fruity, Sweetness, Floral, Tobacco) で構成され、それぞれ0点～4点までの5段階評価となっている。蒸留所の緯度・経度の位置情報も含まれている。

フレーバーデータについては、ノーディングサークル及びウイスキー辞典を参考に、Body, Sweetness, 香りの3つの指標に分類した。Bodyは、ウイスキーの深みやまろやかさ、味わい全体の印象を示す指標で、フルボディやライトボディといった言葉で表現される。Sweetnessは、甘さという味覚の指標で、軽快で澄んだ甘い香りやはちみつのような甘い香りがあり、味覚的な甘さを感じる指標である。また、香りは、スモーキーやフローラルなど10項目によって構成される指標となっている。

## 3. 結 果

### 分析1. 蒸留所の位置情報を用いたフレーバーの地域特性の評価

まず、6地域の蒸留所の位置を色分けしてプロットした。図2で示されるように、「region:1」で示されているスペイサイドは40カ所の蒸留所が狭い地域に密集

しており、「region:2」のハイランドは広い地域に蒸留所が点在している。「region:3」のローランドは3ヶ所の蒸留所のみで、それぞれ離れた場所に点在している。「region:4」はキャンベルタウンで2つの蒸留所で1ヶ所に集中している。「region:5」のアイランズは各島に蒸留所があり、「region:6」のアイラは、淡路島ほどの小さい島の中に7つの蒸留所がある。

次に、図3に示すように、Bodyの値を地図上に色分けしてプロットした。フルボディの「Body:3」「Body:4」の蒸留所は、スペイサイド、ハイランド、アイランズ、アイラなど複数の地域に点在している。また、「Body:2」のミディアムボディはローランド以外の全ての地域に幅広く点在しており、「Body:0」と「Body:1」緑と黄緑のライトボディもキャンベルタウン以外の全ての地域に点在していることがわかる。これらの結果から、Bodyの指標については、地域特性ではなく、個々の蒸留所の個性によるものと判断することができる。

次に、Sweetnessについて、図4に示すように、地図上に色分けしてプロットした。この図を見ると、最も甘さが強い「Sweetness:4」の蒸留所はスペイサイドに集中していることがわかる。またその周辺に「Sweetness:3」が点在しており、さらにその周辺に「Sweetness:2」、そして離れたエリアに「Sweetness:0」が点在していることがわかる。つまり、甘さが強い蒸留所はスペイサイドに集中しており、離れるにつれて甘さが弱くなる傾向が読み取れる。これらの結果から、甘さの指標は、スペイサイドの地域特性と判断することができる。

変数	Smoky	Medicinal	Tobacco	Honey	Spicy	Winey	Nutty	Malty	Fruity	Floral
主成分1	-0.757	-0.8691	-0.579	0.564	-0.02	0.239	0.218	0.417	0.526	0.605

主成分	固有値	寄与率	累積寄与率
1	2.904	29.04%	29.04%
2	1.53	15.03%	44.07%
3	1.144	11.44%	55.51%

表1. 主成分負荷量および寄与率

香りに関しては、10項目で構成されているため、まず主成分分析によって主成分を求めた。その結果、第1主成分は、ヨード香・スモーキーといった項目がマイナス方向、フローラル、ハニー、フルーティといった項目がプラス方向で高い値となっているため、「華やかな香

りとスモーキーでヨード香」の主成分と解釈した。この第1主成分の主成分得点を用いて、地図上にプロットしたものを図5に示す。この図を見ると、華やかな香りの「PCA1:1.335-3.025」「PCA1:0.303-1.335」の蒸留所は、スペイサイド、ローランド、ハイランドの比較的内陸部に点在していることが読み取れる。また、スモーキーでヨード香の強い「PCA1:-6.238--3.605」「PCA1:-3.605--0.920」の蒸留所はアイラ地方を中心にアイランズ、キャンベルタウン、ハイランドなどの沿岸部に点在していることがわかる。この結果から、華やかな香りのウイスキーは内陸部の蒸留所で生産され、スモーキーなウイスキーは沿岸部の蒸留所で生産されるという地域的な傾向があることがわかる。

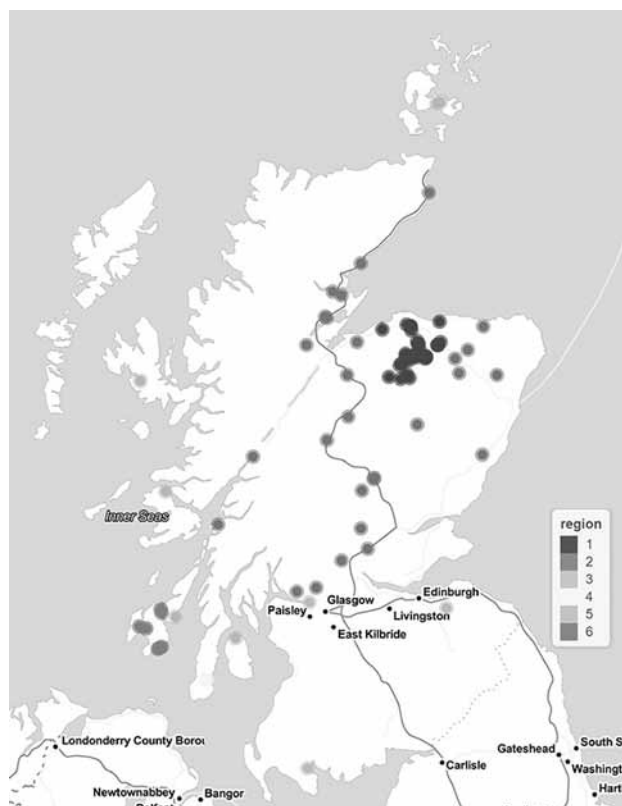


図2

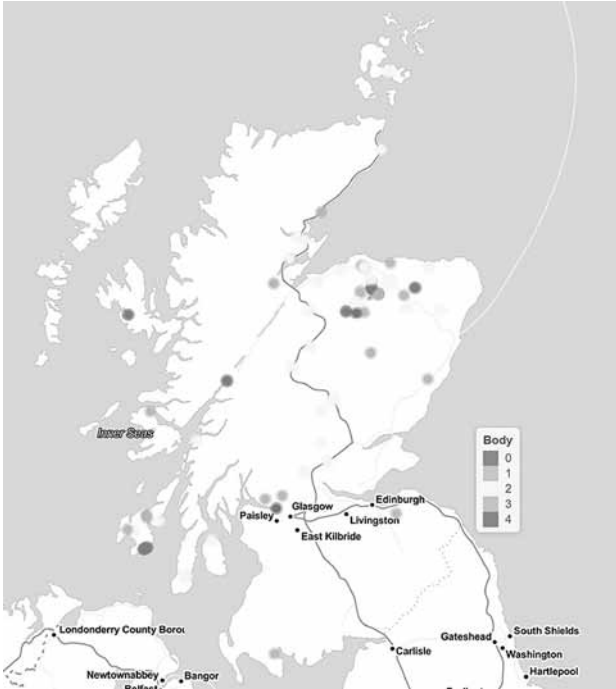


図 3

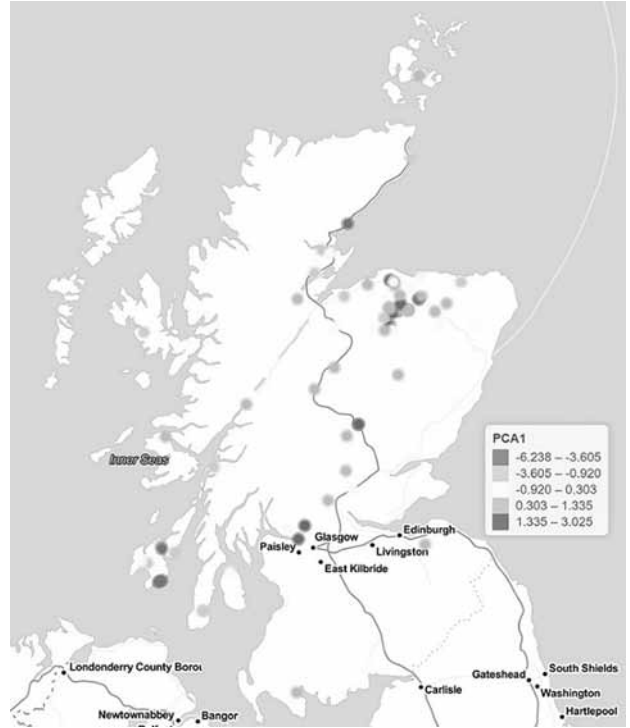


図 5

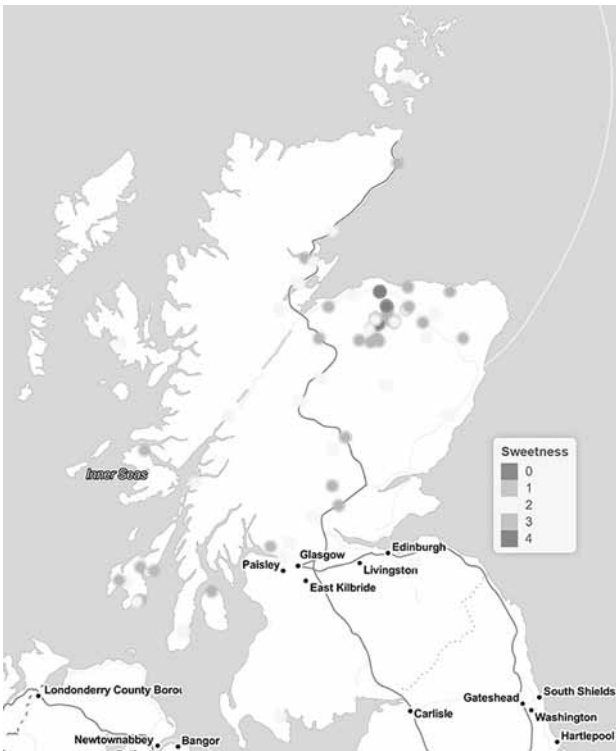


図 4

分析 2. フレーバーデータによる蒸留所の分類とその評価

フレーバーの12項目を用いてWard法によるクラスタ分析を適用し、86蒸留所を分類した。得られた結果から、6つのクラスタに分類できると解釈して、クラスタごとに平均値を求め、拡張型漢字グラフにより可視化した。その結果を図6～図11に示す。

図6より、クラスタ1は、全体的にフレーバーのばらつきが小さく、ミディアムボディでさまざまなフレーバーがバランスよく含まれた、バランス型のクラスタと解釈される。次に、図7より、クラスタ2は、フルボディで深さの中にも甘さとワイニーでフルーティなフレーバーを持つが、スモーキーさも融合されたリッチ&ワイニーのクラスタであると解釈できる。また、図8より、クラスタ3は、ライトボディで、モルトとフローラルでフルーティなフレーバーが特徴的でローランド地域と似た傾向、フローラル&モルティ型のクラスタであると解釈される。図9より、クラスタ4は、ミディアムボディで、甘さが際立ち、モルトの香りが特徴的なスイート&モルティ型のクラスタであると解釈される。図10より、クラスタ5は、ミディアムボディで、スモーキーな中にも甘さがあり、スパイシーさやフルーティでフローラル

な香りが融合されたスイート&スモーキー型のクラスタであると解釈される。最後に、図11より、クラスタ6は、フルボディで、煙、薬といったスモーキーでヨード香が漂う個性的なフレーバーを持ちアイラ地域と似た傾向、リッチ&ヨードスモーキー型クラスタであると解釈される。

このように、拡張型漢字グラフでは、変数の特徴を意味する漢字一文字を用いることで、個々のクラスタの詳細な特徴を直観的に把握できることが示された。

次に、クラスタごとに蒸留所が位置する地域を表2に示した。

表2. 地域別の蒸留所の度数

	スペイサイド	ハイランド	ローランド	アイランズ	キャンベルタウン	アイラ	計
クラスタ1	15	7	0	0	0	0	22
クラスタ2	4	3	0	0	0	0	7
クラスタ3	5	5	3	2	0	1	16
クラスタ4	12	5	0	1	0	0	18
クラスタ5	4	7	0	2	2	2	17
クラスタ6	0	1	0	1	0	4	6
計	40	28	3	6	2	7	86

クラスタ1は、スペイサイドとハイランドに点在しており、スペイサイド15ヶ所、ハイランド7ヶ所の蒸留所である。クラスタ2もスペイサイドとハイランドに点在しており、スペイサイド4ヶ所、ハイランド3ヶ所となっている。クラスタ3は、スペイサイド5ヶ所、ハイランド5ヶ所、ローランド3ヶ所、アイランズ2ヶ所、アイラ1ヶ所となっている。クラスタ4は、スペイサイド12ヶ所、ハイランド5ヶ所、アイランズ1ヶ所となっている。クラスタ5は、スペイサイド4ヶ所、ハイランド7ヶ所、アイランズ2ヶ所、キャンベルタウン2ヶ所、アイラ2ヶ所となっている。最後に、クラスタ6は、ハイランド1ヶ所、アイランズ1ヶ所、アイラ4ヶ所となっている。この表2から、クラスタ5及びクラスタ6のスモーキーさが特徴のクラスタは、比較的沿岸沿いに多く位置していることが分かる。また、フローラルが特徴のクラスタ3は、比較的内陸部に位置していることが示され、甘さが特徴のクラスタ4はスペイサイドに多く位置している。これらは、分析1を支持する結果と言える。

次に、フレーバーの評価において、地域による分類とクラスタによる分類のどちらが適しているかを確認するために、判別分析を行った。表3は、6つのクラスタを

目的変数に、12項目のフレーバーデータを説明変数にした結果である。表に示されている通り、判別率的中率は91.86%と高い値になっている。

表3. 判別分析の結果

判別関数	固有値	寄与率	累積寄与率	相関比 $\eta^2$
1	8.276	64.00%	64.00%	0.892
2	3.051	23.60%	87.60%	0.753
3	1.036	8.00%	95.60%	0.509

判別結果	予測値						判別率的中率	
	1	2	3	4	5	6		
観測値	1	21	0	0	0	1	0	95.45%
	2	0	7	0	0	0	0	100.00%
	3	1	0	15	0	0	0	93.75%
	4	0	0	3	15	0	0	83.33%
	5	1	0	0	1	15	0	88.24%
	6	0	0	0	0	0	6	100.00%
						全体	91.86%	

次に、6つの地域を目的変数に、12項目のフレーバーデータを説明変数にして判別分析を行った。

判別率的中率は61.63%となり、クラスタによる判別分析の結果と比較すると低い値になった。これらの結果から、フレーバーの評価においては、地域による分類よりもクラスタによる分類の方がよりよく判別できていることが示された。

## 4. ま と め

まず分析1では、Bodyの指標は地域よりも蒸留所の個性を反映しており、甘さの指標はスペイサイドが中心、香りの指標では、スモーキーな香りは沿岸部の蒸留所、華やかな香りは内陸部の蒸留所ということが示された。次に分析2では、フレーバーデータ12項目にクラスタ分析を適用して、蒸留所を分類したところ、6つのクラスタに分類された。クラスタの特徴については、拡張型漢字グラフにより、明確に解釈された。また、判別分析の結果、フレーバーデータにより分類されたクラスタは、6つの地域による分類よりも有効性が高いことが示された。

今後の課題として、本研究で得られた結果をもとに、

空間的な分析をさらに進めて、地域とフレーバーとの関連性についてさらに詳細な分析を行いたいと考えている。

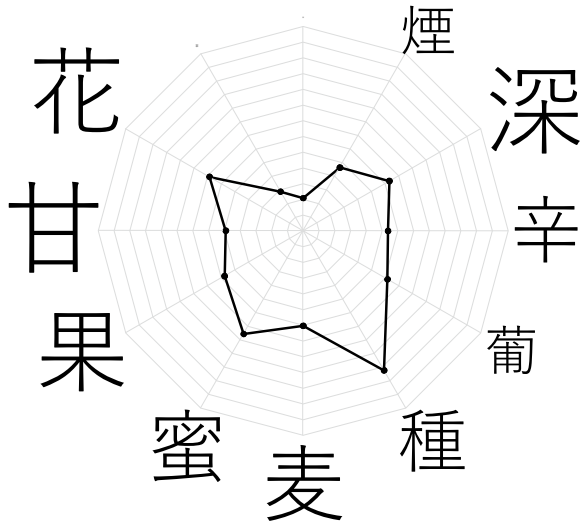


図6

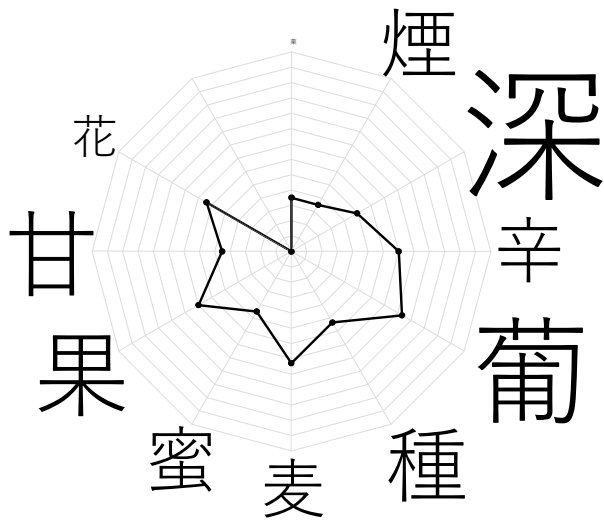


図7

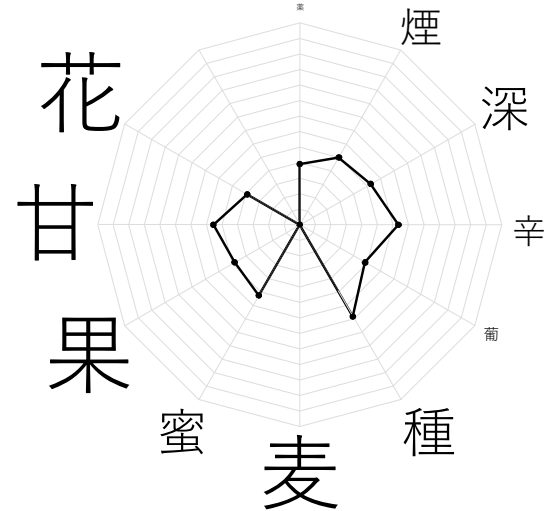


図8

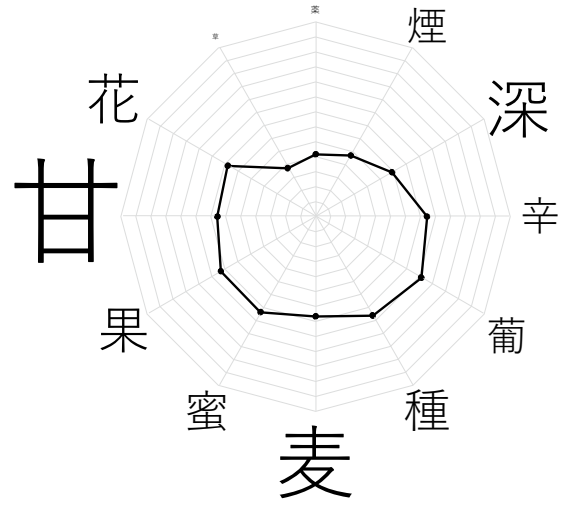


図9

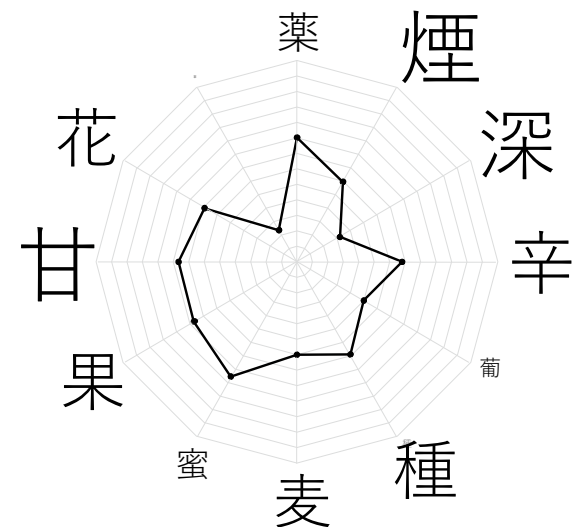


図10

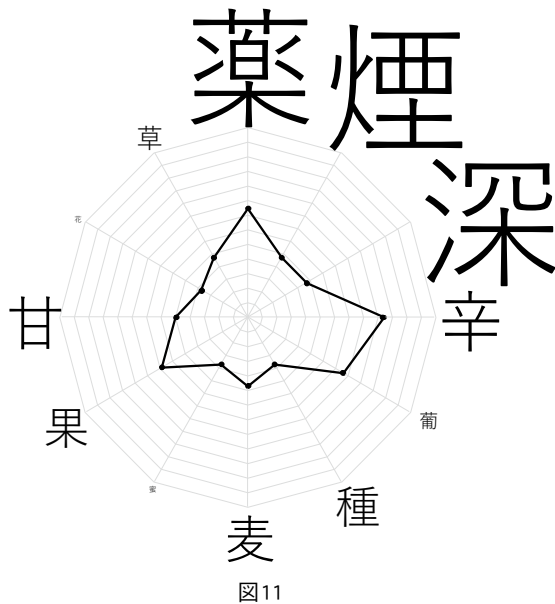


図11

## 参 考 文 献

- Andrews, D. F. (1972). Plots of high-dimensional data. *Biometrics*, 28, 125-136.
- Chernoff, H. (1973). The use of faces to represent points in k-dimensional space graphically, *Journal of the American Statistical Association*. 68, 361-368.
- Fujiwara, M., Kajinishi, S. & Kurihara, K. (2021). Visualization of multivariate data using expanded constellation and expanded kanji graphs and their application to clustering. *Journal of Environmental Science for Sustainable Society*, 10(1), 1-8.
- 肥土伊知郎 (2017). シングルモルト&ウイスキー完全バイブル, ナツメ社.
- 平井安久・福森護・脇本和昌 (1988). 多変量データの漢字グラフ表現とクラスタリングへの応用. *計算機統計学*, 1(1), 11-21.
- 石村友二郎 (2013). 重回帰分析のグラフ表現法. *計算機統計学*, 26(2), 93-103.
- 好田裕史 (2015). ウイスキー香気成分のストレス緩和作用, *日本醸造協会誌*, 110(2), 81-85.
- 増田正裕・小林啓 (1993). ウイスキーの味, 香り(その3) 官能評価と成分の寄与, *日本醸造協会誌*, 88(3), 201-207.
- サントリー ウイスキー用語辞典  
<https://www.suntory.co.jp/whisky/Ballantine/lex-a.html>
- 戸塚昭 (1989). ウイスキーとブランデーの官能評価, *日本醸造協会誌*, 84(11), 732-738.
- 土屋守 (2018). ウイスキー完全バイブル, ナツメ社.
- Wakimoto, K. (1977). Tree graph method for visual representation of multi-dimensional data, *Journal of the Japan Statistical Society*, 7, 27-34.
- 脇本和昌・後藤昌司・松原義弘 (1979). 多変量グラフ解析法. 朝倉書店.
- Wakimoto, K. & Taguri, M. (1978). Constellation graphical method for representing multi dimensional data, *Annals of the Institute of Statistical Mathematics*, 30, Part A, 77-84.
- Wishart, D. (2009). The Flavour of Whisky. *Significance*, 6(1).
- Wishart, D. (2018). Whisky Classified. *PAVILION*, 36-41.