

## ワインレビューデータ タスクシナリオ

### 📌 プロジェクト概要

高級ワイン輸入商社より、新規取り扱いワインの選定支援を依頼されました。バイヤーは「高価なワインが良いのは当たり前だが、手頃な価格で高品質な『隠れた名品』を見つけたい」と考えています。

あなたのミッションは、**インタラクティブなダッシュボードを作成し、コストパフォーマンスの高いワイン産地や品種を発掘すること**です。

### 📊 データセット詳細: wineReview.csv

概要: 129971 行, 14 カラム

カラム名	データ型	統計量・詳細
csv_id	int64	最小: 0 最大: 129970 平均: 64985.00
country	object	ユニーク数: 43 上位 3 件: US (54504 件) France (22093 件) Italy (19540 件)
description	object	ユニーク数: 119955 上位 3 件: Cigar box, café au lait, and dried tobacco aromas are followed by coffee and blackberry flavors on the palate with a moderate finish. (3 件) Seductively tart in lemon pith, cranberry and pomegranate, this refreshing, light-bodied quaff is infinitely drool-worthy. (3 件) This communicates a sense of aromatic purity, with scents of white flower, stone fruit and citrus. The mouthfeel is creamy and rich, but supported by good acidity. (3 件)
designation	object	ユニーク数: 37979 上位 3 件: Reserve (2009 件) Estate (1322 件) Riserva (698 件)
points	int64	最小: 80 最大: 100 平均: 88.45

カラム名	データ型	統計量・詳細
price	float64	最小: 4.0 最大: 3300.0 平均: 35.36
province	object	ユニーク数: 425 上位 3 件: California (36247 件) Washington (8639 件) Bordeaux (5941 件)
region_1	object	ユニーク数: 1229 上位 3 件: Napa Valley (4480 件) Columbia Valley (WA) (4124 件) Russian River Valley (3091 件)
region_2	object	ユニーク数: 17 上位 3 件: Central Coast (11065 件) Sonoma (9028 件) Columbia Valley (8103 件)
taster_name	object	ユニーク数: 19 上位 3 件: Roger Voss (25514 件) Michael Schachner (15134 件) Kerin O'Keefe (10776 件)
taster_twitter_handle	object	ユニーク数: 15 上位 3 件: @voossroger (25514 件) @wineschach (15134 件) @kerinokeefe (10776 件)
title	object	ユニーク数: 118840 上位 3 件: Gloria Ferrer NV Sonoma Brut Sparkling (Sonoma County) (11 件) Korbel NV Brut Sparkling (California) (9 件) Segura Viudas NV Extra Dry Sparkling (Cava) (8 件)
variety	object	ユニーク数: 707 上位 3 件: Pinot Noir (13272 件) Chardonnay (11753 件) Cabernet Sauvignon (9472 件)

カラム名	データ型	統計量・詳細
winery	object	ユニーク数: 16757 上位 3 件: Wines & Winemakers (222 件) Testarossa (218 件) DFJ Vinhos (215 件)

## ■ バイヤーの仮説（検証対象）

バイヤーは以下の仮説を持っていますが、【】の部分が不明確です。ダッシュボードを作成し、この空白をファクトで埋めてください。

「弊社の基準では、Points が 80~85 点であれば『一般』、85~90 であれば『高品質』、90 以上であれば『最高品質』と定義しています。一般的に評価点 (Points) が高いワインは価格 (Price) も高い傾向にあります、この相関から外れた『コスパ最強』の生産国が存在するはずです。具体的には、【A】(国名) は平均価格が世界平均と比べて \$ [B] も安く、中でも 【C】(品種名) は圧倒的なコストパフォーマンスを誇っていると推測しています。この仮説を検証するためのダッシュボードを作成し、該当する国と詳細を特定してください。」

## ■ 納品物（作成するダッシュボードの要件）

この仮説を検証し、バイヤーに報告するために、以下の分析視点を網羅したダッシュボードを作成してください。

### 1. 価格と評価点の相関分析

- 目的: 価格と評価点の全体的な相関を確認し、全体の傾向を把握します。
- ヒント: 価格を横軸、評価点を縦軸にとり、個々のデータの分布を可視化できるグラフを作成してください。

### 2. 国別の平均価格比較

- 目的: 他の国と比較して、平均価格が低い「候補国」をいくつか特定します。
- ヒント: 国 (Country) ごとの平均価格を比較できるグラフを作成してください。価格が安い順に並べ替えると発見しやすくなります。

### 3. 候補国の品質検証

- 目的: 上記で特定した「低価格な国」の平均評価点が、バイヤーの求める「高品質 (85 点~90 点)」のレンジに収まっているかを確認し、最終的に国 [A] を特定します。
- ヒント: 特定した国の平均評価点を表示し、その数値が目標レンジ (85~90) のどこに位置するかを直感的に把握できるグラフを作成してください。

### 4. 【自由課題】国 [A] を担当するレビューの分析

- 目的: 国 [A] のワインを主に評価しているレビューは誰か特定し、その人物の評価傾向（他の国と比べて点数が甘いか厳しいか？どのような価格帯を好むか？）を深掘りしてください。
- ヒント: `taster_name` ごとのレコード数をカウントして主要なレビューを特定し、そのレビューに絞って分析を行います。

## ■ 実行手順

1. 分析アプリケーションにログインし、データソース **wine\_review** に接続します。
2. **Step 1:** 價格と評価点で散布図を作成し、全体の分布を確認します。
3. **Step 2:** 国ごとの平均価格を比較できるグラフを作成し、コストパフォーマンスが良さそうな（平均価格が低い）国の候補を見つけます。
4. **Step 3:** 候補国の平均評価点が **85 点以上 90 点未満（高品質）** の範囲に入っているかがひと目でわかるグラフを作成します。これにより、条件に合致する国【A】を確定させてください。
5. **Step 4: 自由課題** として、特定した国【A】を評価しているメインレビューを探し出し、その人物の評価傾向を分析するグラフを追加してください。
6. 完成したダッシュボードをもとに、報告内容をまとめてください。