

DATASET WINEQT

Nicole Susan Loza Ticona

December 9, 2024

1. Descripción detallada del dataset

El dataset contiene **1,143 observaciones** y **13 variables**, las cuales describen características físico-químicas y la calidad del vino. A continuación, se presenta una descripción de las columnas:

Columna	Descripción
fixed acidity	Cantidad de ácidos no volátiles presentes en el vino (g/dm ³).
volatile acidity	Acidez volátil que puede afectar el sabor (g/dm ³).
citric acid	Contenido de ácido cítrico (g/dm ³).
residual sugar	Azúcar residual presente después de la fermentación (g/dm ³).
chlorides	Contenido salino (g/dm ³).
free sulfur dioxide	Cantidad de dióxido de azufre libre presente (mg/L).
total sulfur dioxide	Cantidad total de dióxido de azufre presente (mg/L).
density	Densidad relativa del vino (g/cm ³).
pH	Medida de la acidez o alcalinidad.
sulphates	Concentración de sulfatos (g/dm ³).
alcohol	Porcentaje de alcohol por volumen (%).
quality	Evaluación de calidad del vino (escala de 3 a 8).
Id	Identificador único para cada registro.

A continuación, presentamos un resumen estadístico de las variables numéricas del dataset:

Variable	Mínimo	1er Cuartil	Mediana	Media	3er Cuartil	Máximo
Fixed acidity	4.6	7.1	7.9	8.31	9.1	15.9
Volatile acidity	0.12	0.39	0.52	0.53	0.64	1.58
Residual sugar	0.9	1.9	2.2	2.53	2.6	15.5
Chlorides	0.012	0.07	0.079	0.0869	0.09	0.611
Alcohol	8.4	9.5	10.2	10.44	11.1	14.9
Quality	3	5	6	5.66	6	8

2. Objetivo de investigación

El objetivo principal de este análisis es **determinar cómo las características físico-químicas del vino influyen en su calidad**, representada por la variable **quality**. Se busca identificar:

- Las características más relevantes que afectan la calidad.
- Relación entre propiedades químicas y la percepción sensorial del vino.
- Posibilidad de predecir la calidad del vino a partir de las variables químicas y físicas.