

# Ejercicio 9: Análisis global de los datos

En el análisis de correlación utilizo las tres características más significativas (alcohol, malic\_acid y flavanoids), al ejecutar mi código obtengo estos resultados:

```
35 # Calculo correlaciones entre las tres características seleccionadas
36 print("\n Calculo correlaciones entre las tres características seleccionadas: \n")
37 correlaciones = utilities.calcular_correlacion(df, caracteristicas)
38 print(correlaciones)
39
```

```
33 coef_pearson, p_valor = stats.pearsonr(df[caracteristicas[i]], df[caracteristicas[j]])
34 resultados[f'{caracteristicas[i]} vs {caracteristicas[j]}'] = {'Coeficiente de correlación': coef_pearson
35 # pasamos resultado(diccionario) a dataframe
36 return pd.DataFrame(resultados).T #transponemos para que las claves del diccionario pasen a columnas y los valores
37
38
39
```

PROBLEMS OUTPUT DEBUG CONSOLE GITLENS GITLENS INSPECT TERMINAL

```
(base) sandyrodriuezaponte@MacBook-Pro-M3-Pro-de-Sandy Aprendizaje_Automático % python AnalisisGlobalDatos.py
Nº de muestras: 178
Características que se analizan: ['alcohol', 'malic_acid', 'ash', 'alcalinity_of_ash', 'magnesium', 'total_phenols', 'flavanoids', 'nonflavanoid_phenols', 'proanthocyanins', 'color_intensity', 'hue', 'od280/od315_of_diluted_wines', 'proline']

Medidas de estadística descriptiva:
Range      Media Mediana  Moda Desviación estándar Asimetría Curtosis
alcohol      3.6    13.000618  13.05  12.37      0.811827 -0.051047 -0.86226
malic_acid   5.06    2.336348  1.865   1.73      1.117146  1.038869  0.257348
ash          1.87    2.366517  2.36   2.28      0.274344 -0.175207  1.078576
alcalinity_of_ash 19.4    19.494944  19.5   20.0      3.339564  0.211247  0.448823
magnesium    92.0    99.741573  98.0   88.0      14.282484  1.088915  2.012806
total_phenols 2.9     2.293112  2.355  2.2      0.626951  0.089907 -0.845637
flavanoids   4.74    2.02927  2.135  2.65      0.998859  0.025129 -0.889365
nonflavanoid_phenols 0.53    0.361854  0.34   0.26      0.124453  0.446349 -0.652952
proanthocyanins 3.17    1.590899  1.555  1.35      0.572359  0.512769  0.505671
color_intensity 11.72    5.05809   4.69   2.6       2.318286  0.861248  0.33737
hue          1.23    0.957449  0.865  1.04      0.228572  0.026913 -0.368825
od280/od315_of_diluted_wines 2.75    2.611885  2.78   2.87      0.789959 -0.30469 -1.089675
proline     1402.0   746.893258  673.5  520.0     314.907474  0.761336 -0.275

Calculo correlaciones entre las tres características seleccionadas:
Coeficiente de correlación  P-valor
alcohol vs malic_acid      0.094397  2.100820e-01
alcohol vs flavanoids      0.236815  1.459382e-03
malic_acid vs flavanoids   -0.411007  1.207022e-08
(base) sandyrodriuezaponte@MacBook-Pro-M3-Pro-de-Sandy Aprendizaje_Automático %
```

## Conclusiones:

### 1. Alcohol vs Malic Acid:

- **Coeficiente de correlación:** 0.09 (correlación muy débil).
- **P-valor:** 0.21 (no significativo,  $p > 0.05$ ).
- **Interpretación:** No hay una correlación significativa entre alcohol y malic\_acid. La relación es casi inexistente.
- 

### 2. Alcohol vs Flavanoids:

- **Coeficiente de correlación:** 0.24 (correlación positiva débil).
- **P-valor:** 0.00146 (significativo,  $p < 0.05$ ).
- **Interpretación:** Hay una correlación positiva débil, pero significativa, entre alcohol y flavanoids. Un ligero aumento en uno está relacionado con un aumento en el otro.

### 3. Malic Acid vs Flavanoids:

- **Coeficiente de correlación:** -0.41 (correlación negativa moderada).
- **P-valor:** 1.21e-08 (muy significativo).
- **Interpretación:** Existe una correlación negativa moderada y significativa entre malic\_acid y flavanoids, lo que indica que cuando una aumenta, la otra tiende a disminuir.

Observo que las correlaciones son débiles o moderadas, las relaciones entre estas variables son estadísticamente significativas, excepto en el caso de alcohol vs malic\_acid.