Práctico Especial

Fundamentos de la Ciencia de Datos

Calidad del vino

Hermida Matias: mhermida@alumnos.exa.unicen.edu.ar
Russo Dolores Magalí: drusso@alumnos.exa.unicen.edu.ar
Velis Ulises: uvelis@alumnos.exa.unicen.edu.ar

ÍNDICE

INTRODUCCIÓN

En el presente trabajo, se llevará a cabo un análisis detallado de un conjunto de datos con 3231 muestras de vino, obtenidas mediante pruebas fisicoquímicas en la bodega Del Sol, elaboradas a partir de dos tipos de uva: Moscatel y Syrah.
Primero se realizará un análisis exploratorio de los datos (EDA) para obtener una visión general de las características de cada variable, detectar valores atípicos, nulos o outliers y buscar relaciones entre dos o más variables. A partir de estos hallazgos, se tomarán decisiones sobre cómo proceder con estos valores inusuales aplicando las transformaciones necesarias.
Luego, una vez conocido el conjunto de datos se comenzarán a plantear hipótesis centradas en la calidad del vino, la variable objetivo del dataset, y en otros aspectos que van surgiendo en el análisis. Hipótesis a las cuales se les realizará un test para validarlas.

ANÁLISIS EXPLORATORIO DE LOS DATOS

El dataset cuenta con las siguientes variables:

1. type: tipo de uva con la que se elabora el vino.
2. fixed acidity: cantidad de ácidos no volátiles presentes en el vino, medida en gramos por litro.
3. volatile acidity: cantidad de ácidos volátiles presentes en el vino, medida en gramos por litro.
4. citric acid: contenido de ácido cítrico en el vino, medido en gramos por litro.
5. residual sugar: cantidad de azúcar que queda en el vino después de la fermentación, medida en gramos por litro.
6. chlorides: concentración de cloruros (sales) en el vino, medida en gramos por litro.
7. free sulfur dioxide: cantidad de dióxido de azufre que no está ligado químicamente en el vino, medida en miligramos por litro.
8. total sulfur dioxide: suma del dióxido de azufre libre y el combinado en el vino, medida en miligramos por litro.
9. density: medida de la masa por unidad de volumen del vino, utilizada para estimar la concentración de sólidos disueltos, medida en gramos por centímetro cúbico.
10. pH: medida de la acidez o alcalinidad del vino.
11. sulphates: concentración de sales de sulfato en el vino, medida en gramos por litro.
12. alcohol: contenido alcohólico del vino, medido en porcentaje de volumen (% vol).
13. quality: puntuación del vino, con una escala que va de 0 a 10.

Análisis univariado

Tipos incorrectos

El dataframe con el que se trabajó detectó que la variable “type” es de tipo *object,* pero es de tipo *string*, por lo que se transformó para que sea del tipo correspondiente.
La variable “alcohol” también fue afectada, ya que se encontraron valores que se supuso que son periódicos mal cargados. La columna no se identificaba como *float* sino como *object*. Se dispuso de estas filas y se transformó la columna al tipo correspondiente.

Outliers

La columna “free sulfur dioxide” tiene un valor extremo muy alto en comparación a los demás valores, con una diferencia de más de 160 mg/l con el valor anterior a él. Este valor afecta a la columna “total sulfur dioxide” generando también un outlier en su último valor ya que depende de ella. Según el INV de Argentina el valor 440 mg/l de dióxido de azúfre total (anhídrido sulfuroso total) es muy superior al límite máximo aceptado (210 mg/l).
Previamente se buscó la “Bodega del Sol” y se encontró una posible coincidencia en España, por lo que se indagó sobre el valor máximo aceptado del anhídrido sulfuroso total en la Unión Europea y aún así lo supera. Por estos motivos, se decidió eliminar la muestra.

Valores atípicos

Según el INV (Instituto Nacional de Vitivinicultura), los valores típicos de densidad en los vinos suelen variar entre 0.990 g/cm³ para vinos secos y hasta aproximadamente 1.020 g/cm³ para vinos dulces, ya que estos últimos contienen más azúcar residual. Sin embargo, en el análisis de nuestro conjunto de datos, se encontraron valores de densidad que superan significativamente el rango esperado, alcanzando cifras superiores a 100 g/cm³. Las 75 filas involucradas se eliminaron.

En la columna “alcohol” se tomaron los valores atípicos como números periódicos mal cargados ya que además del hecho de que son periódicos, el alcohol está medido en porcentaje (siendo 100% el máximo) y se observan valores mayores a 100 e incluso 900. En un principio se pensó en modificarlos a 10% y 9% respectivamente ya que los demás valores no superan el 15%, pero como las filas involucradas representan tan solo 0,01% del total de muestras (40), se decidió descartarlas.

Análisis bivariado

Para este análisis, se consideró el dataset completo y también se lo dividió en dos, según el tipo de uva. Para los tres conjuntos se graficó la matriz de correlación con el fin de analizar las relaciones entre las variables y comenzar a encontrar ciertos hallazgos o posibles hipótesis.

General

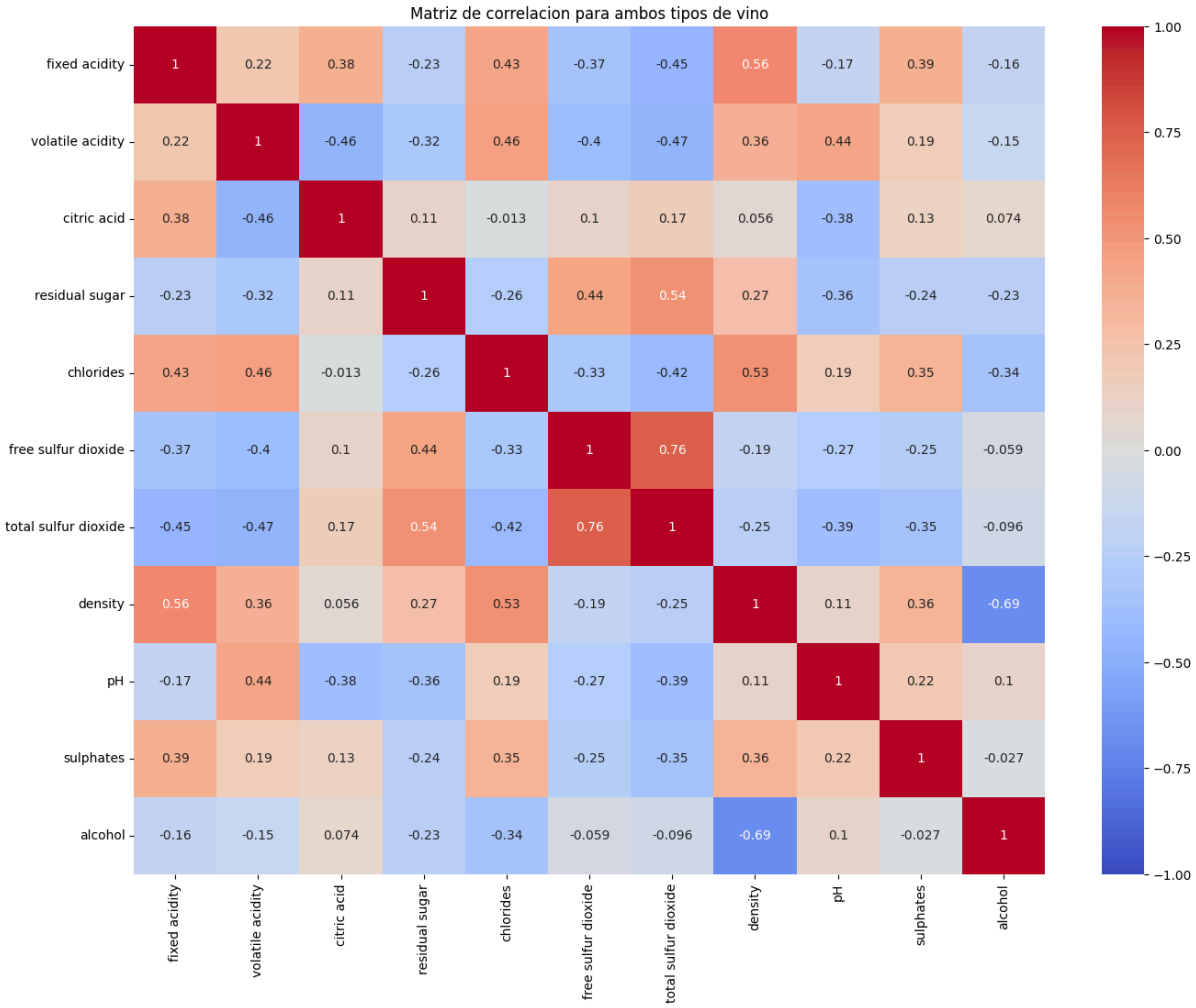

Gráfico 1

Considerando los vinos de ambos tipos de uva, se identificaron correlaciones significativas entre las variables “total sulfur dioxide” y “free sulfur dioxide” con un coeficiente de correlación de 0,79. Esto no sorprende, entendiendo que “total sulfur dioxide” es la suma del dióxido de azufre libre (“free sulfur dioxide”) y el combinado en el vino.
Otra correlación que vale la pena mencionar es la de las columnas “density” y “alcohol”, con un coeficiente de correlación de -0,69. Esto podría indicar que **los vinos con mayor densidad tienen menor porcentaje de alcohol** (o viceversa).

Como se podrá ver, en los dataset divididos por tipo de uva no existen las mismas correlaciones entre variables.

Moscatel

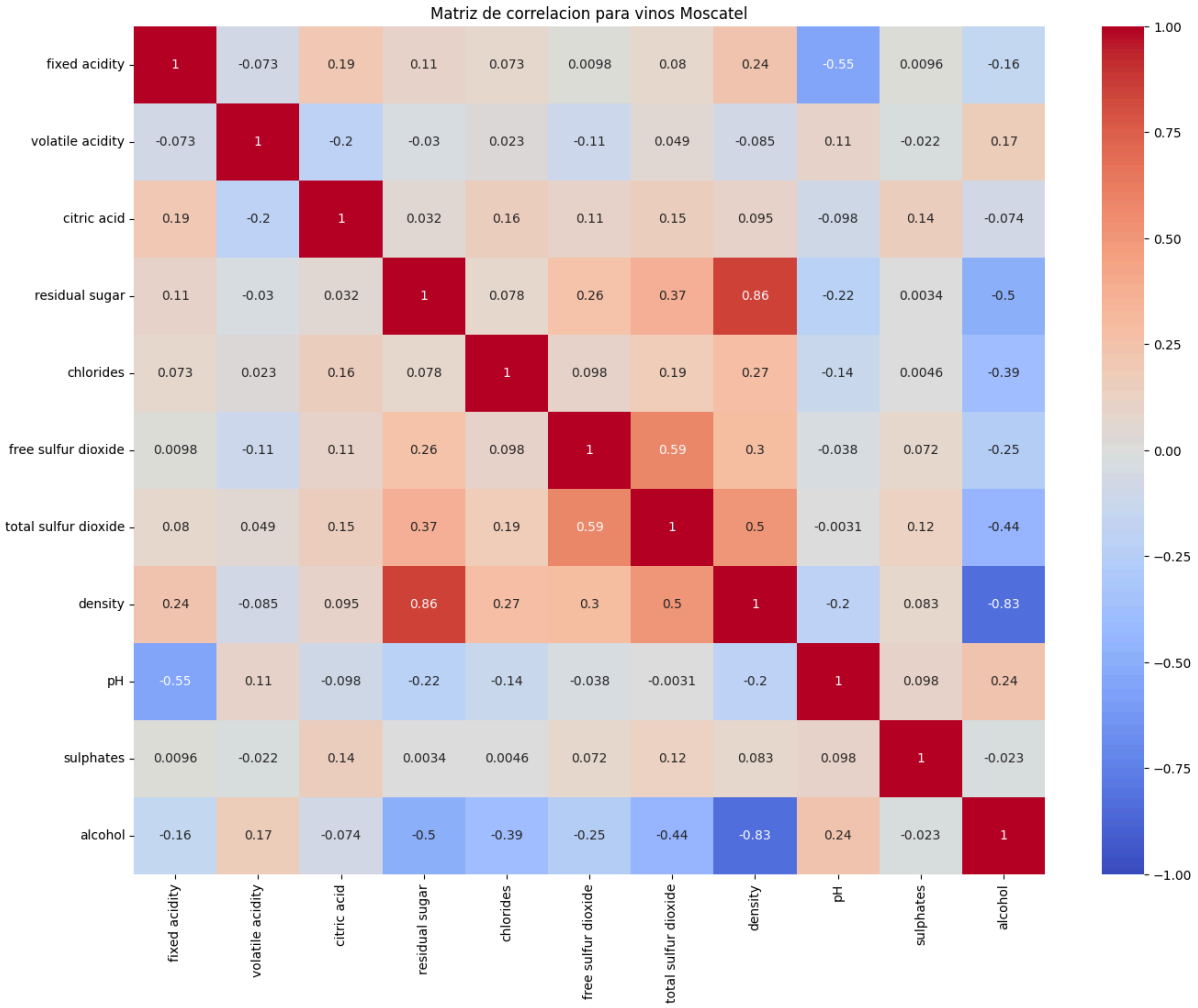

Gráfico 2

Para los vinos del tipo de uva Moscatel se encontró una correlación significativa entre las variables “alcohol” y “density” (en línea con lo descubierto previamente) con un coeficiente de correlación de -0,83. Con las variables “residual sugar” y “density” cabe señalar también una correlación significativa, con un coeficiente de correlación de 0,86.
Se realizó un gráfico de dispersión para complementar el análisis de esta última relación visualmente.

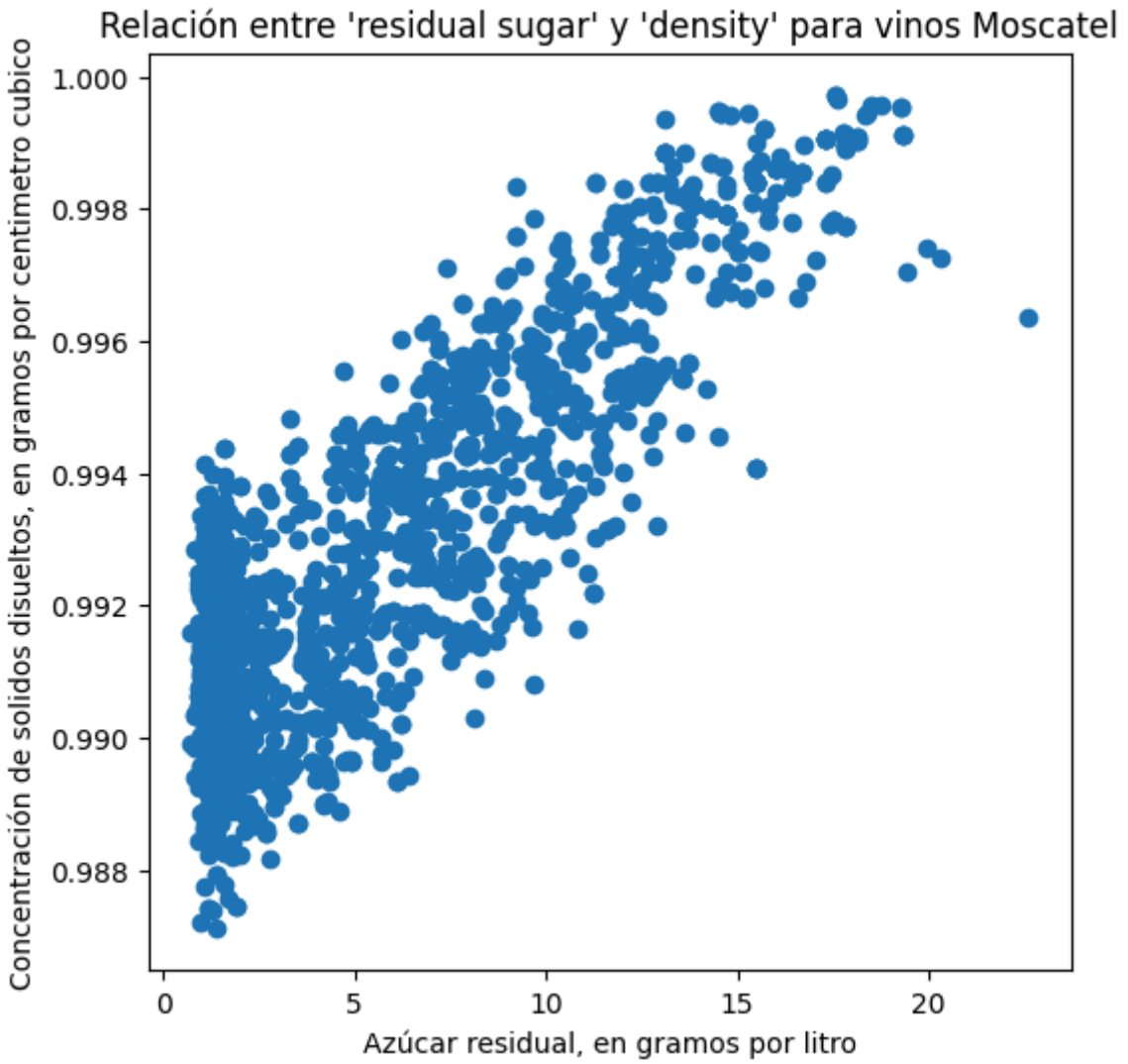

Gráfico 3

Para este caso “residual sugar” y “density” son directamente proporcionales. Se puede intuir entonces, que **en los vinos Moscatel, los que tienen mayor azúcar residual tienen mayor concentración de sólidos disueltos** (densidad).

Syrah

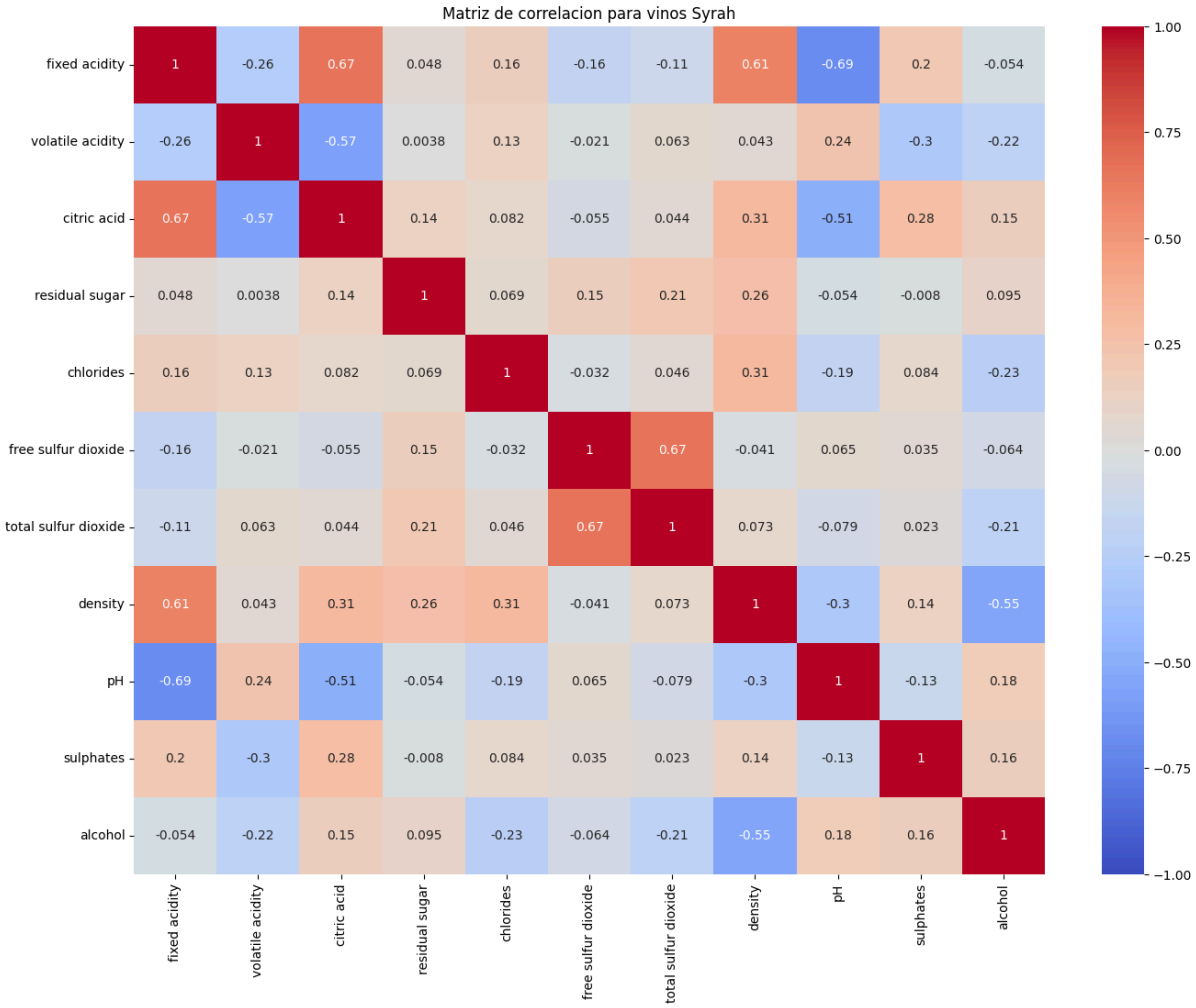

Gráfico 4

En el dataset del tipo de uva Syrah se encontraron correlaciones significativas entre las variables “free sulfur dioxide” y “total sulfur dioxide” (0,67) (correlación también presente en el dataset completo), y entre “pH” y “fixed acidity” (-0,69). Para esta última, como son inversamente proporcionales, se puede decir que **en los vinos Syrah, a mayor alcalinidad** (pH) **menor acidez no volátil**.

Análisis multivariado

Con el objetivo de seguir obteniendo información de los datos, se utilizó PCA como técnica de reducción de dimensionalidad y el gráfico biplot para profundizar el análisis (ahora incorporando múltiples variables).

En primer lugar, como PCA se basa en la varianza de los datos y las variables tienen escalas muy diferentes, se estandarizaron los datos. Luego de aplicar PCA para todas las variables (menos las cualitativas), se obtuvo una varianza explicada del 53%. Luego se utilizó el gráfico biplot, que ofrece una perspectiva comparativa, donde se puede ver la

relación de cada una de las variables con cada tipo de vino, además de observarse las relaciones entre las propias variables.
En el gráfico se observan las variables que más influyen sobre cada tipo de vino.

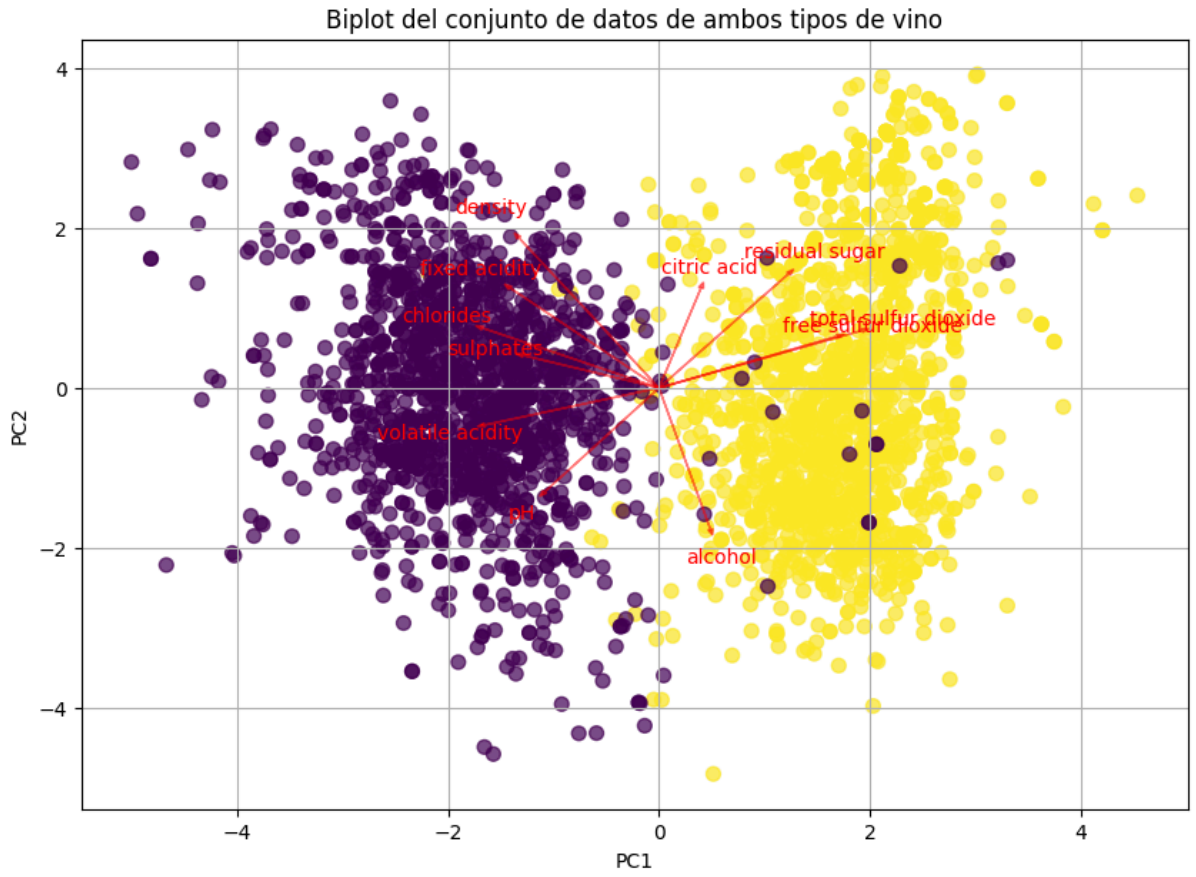

Gráfico 5

Los Moscatel (en amarillo) presentan valores elevados en “total sulfur dioxide”, “free sulfur dioxide”, “residual sugar”, “alcohol” y “citric acid”. Mientras que los Syrah (en violeta) muestran valores más altos en “pH”, “volatile acidity”, “sulphates”, “chlorides”, “fixed acidity” y “density”, y más bajos en el resto. En el punto (0,0) se encuentran los vinos con valores promedio.

Cómo son claramente diferenciables los dos tipos de vino, se puede decir que la primera componente (eje x), es la que permite diferenciarlos. Esta componente explica alrededor del 34% de la varianza. Las variables que más contribuyen a esta componente son las que más diferencian los vinos de distinto tipo de uva.

Las variables que explican mayormente la primera componente (“total sulfur dioxide”, “chlorides”, “free sulfur dioxide”, “density” y “volatile acidity”) son las variables en las que esos vinos son diferentes. Para la variable “total sulfur dioxide” que tiene gran peso en la componente horizontal, se puede decir que **los Moscatel tienen mayor dióxido de azúfre total que los vinos Syrah**. Luego para la variable “density” se puede decir que **los vinos Syrah tienen mayor densidad que los vinos Moscatel**. Para la variable “free sulfur dioxide” se puede decir que **los vinos Moscatel tienen mayor cantidad de dióxido de azufre libre que los vinos Syrah**. Para la variable “chlorides” se puede decir que **los vinos Syrah tienen una mayor concentración de cloruros que los vinos Moscatel**. Por último,

para la variable “volatile acidity” se puede decir que **los vinos Syrah tienen mayor acidez volátil que los vinos Moscatel**.

Para profundizar en cada tipo de vino, se aplicó PCA para poder proyectar a un espacio de dos dimensiones en los datasets separados por tipo de uva.

Luego de aplicar PCA en los Moscatel, se obtuvo una proyección que explica el 44% de la varianza. Al gráfico de esta proyección se le pintaron las muestras por calidad.

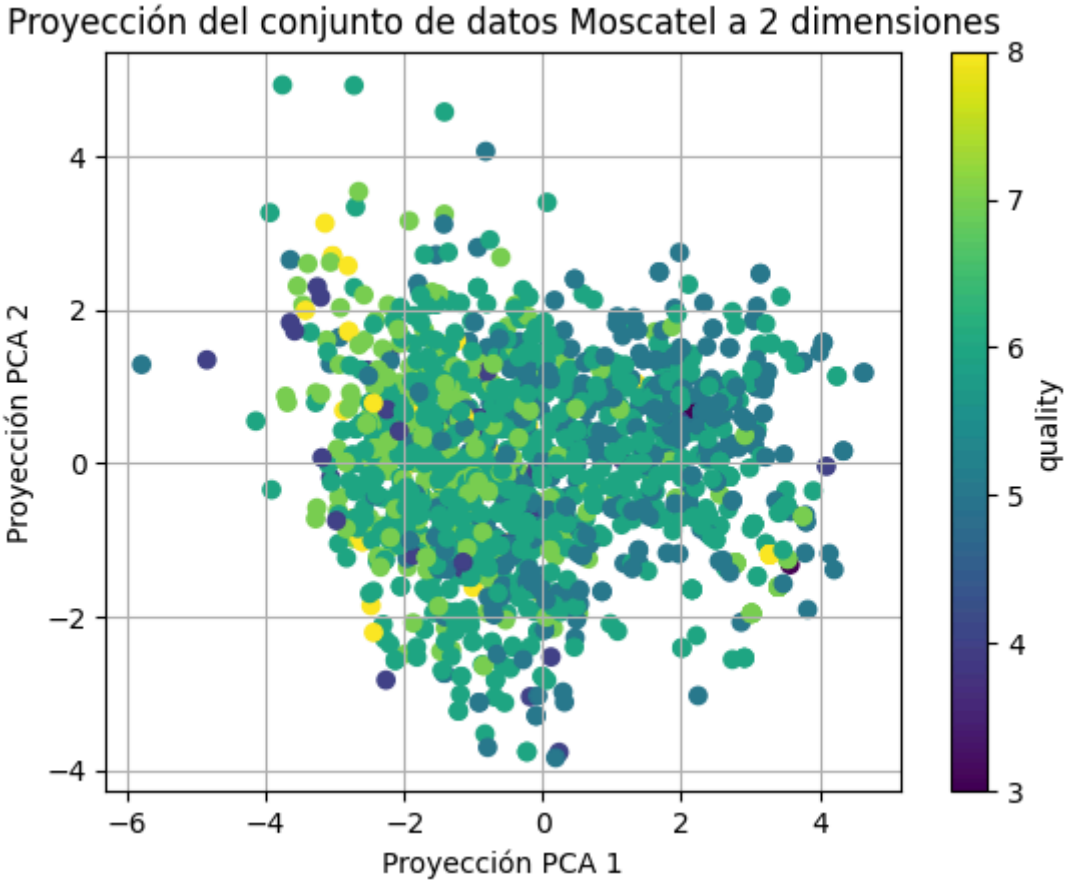

Gráfico 6

Como se puede observar, los vinos con mayor puntaje de calidad se encuentran en la parte izquierda. Luego, se observó el mismo gráfico pero se pintaron las muestras en base a otras variables. Por ejemplo “alcohol”:

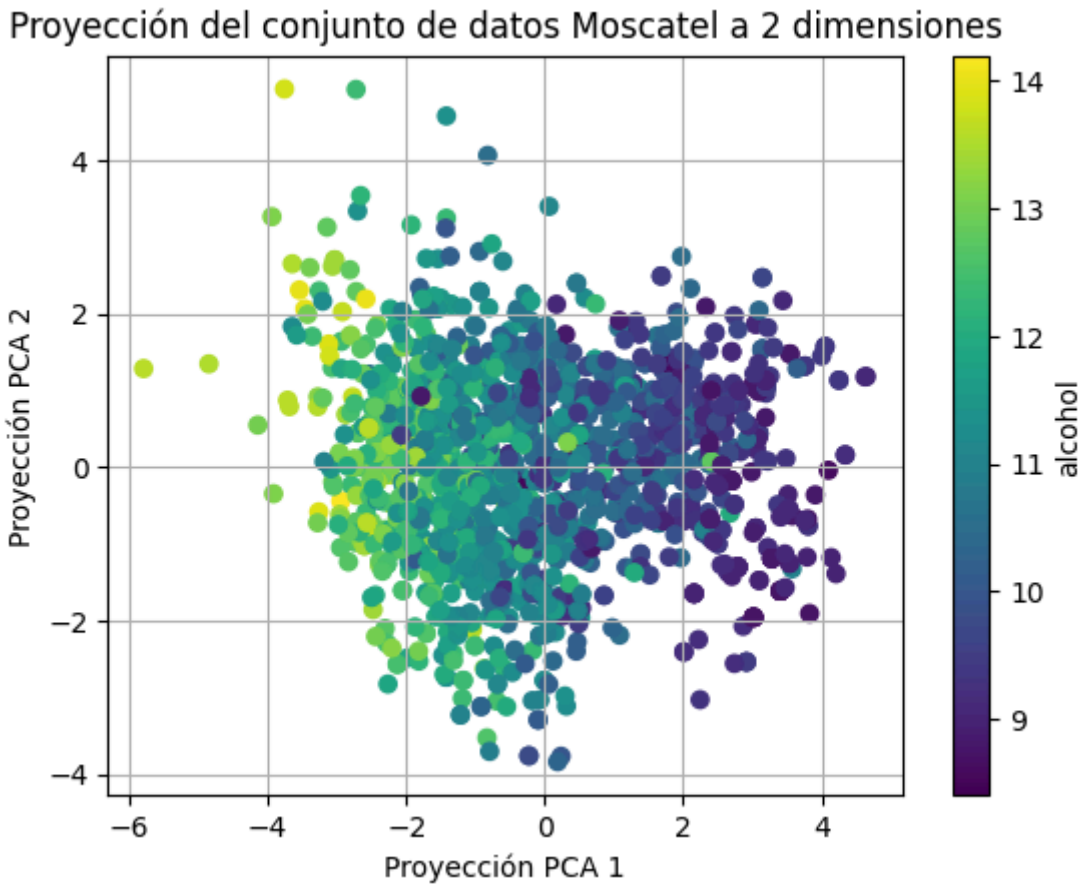

Gráfico 7

Para este caso, se puede ver que aproximadamente los mismos vinos que tienen alto porcentaje de alcohol coinciden con aquellos con puntuación alta. Se puede decir entonces que ***en los vinos Moscatel, los vinos con más contenido alcohólico tienen mayores puntuaciones***.
Una situación distinta es la del caso “density”:

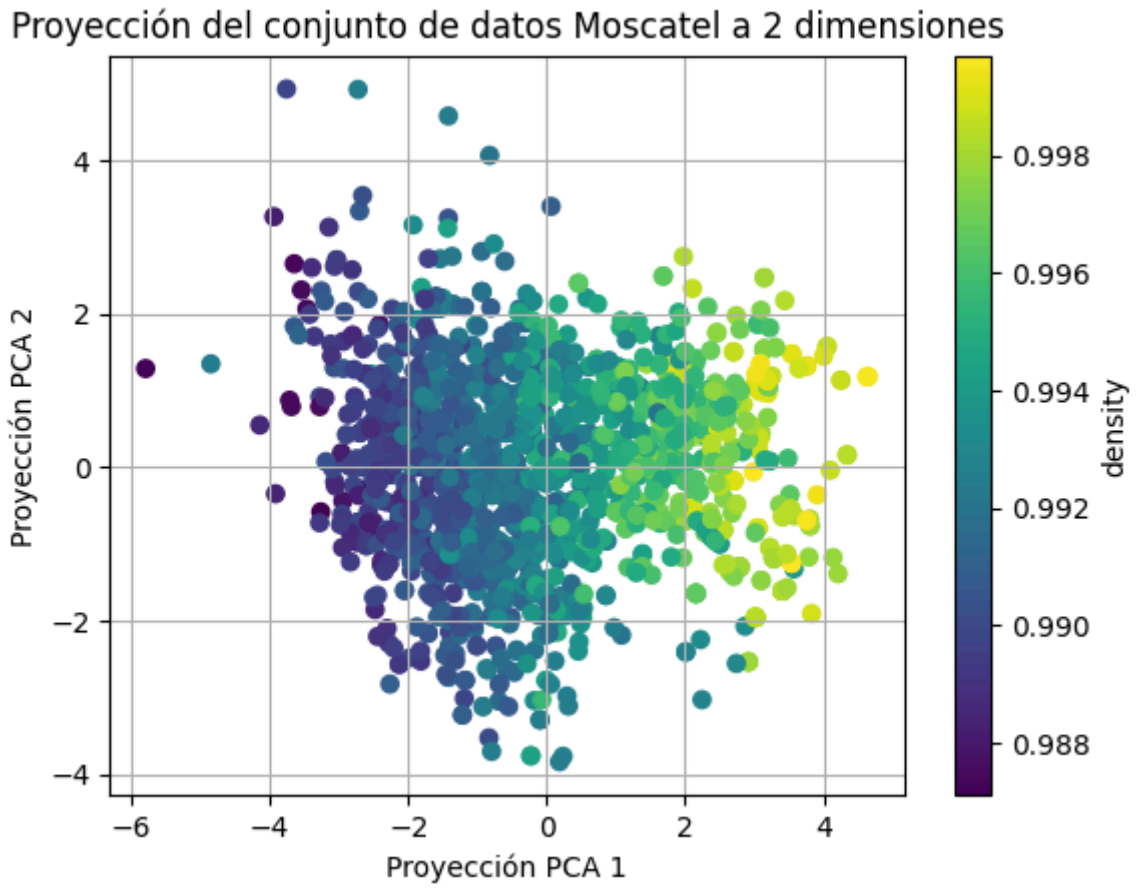

Gráfico 8

En este caso, se puede observar que **en los vinos Moscatel, aquellos con mayor densidad tienen menores puntuaciones de calidad**.

Luego de aplicar PCA en los Syrah, se obtuvo una proyección que explica el 44% de la varianza. Al igual que para Moscatel, al gráfico de esta proyección se le pintaron las muestras por calidad.

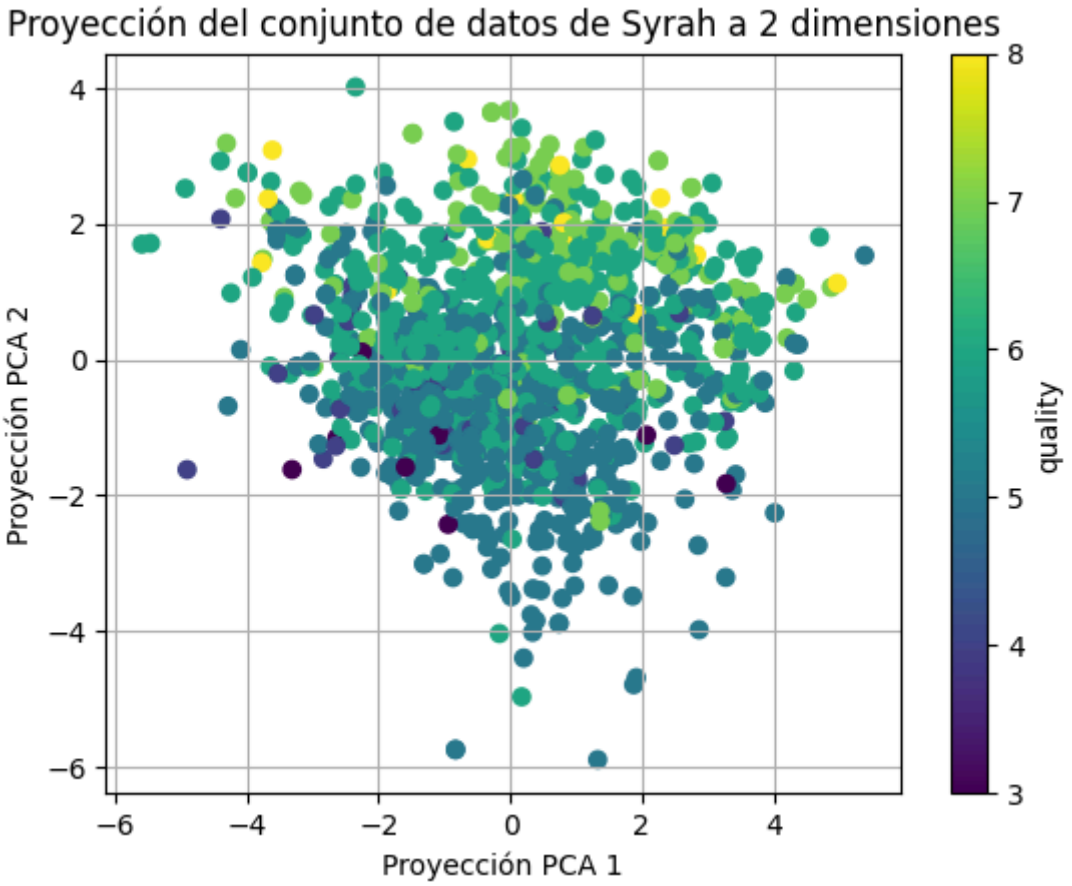

Gráfico 9

En este gráfico se puede observar que los vinos con mayor puntaje se encuentran en la parte superior.
Se realizaron algunos hallazgos con la variable “free sulfur dioxide”:

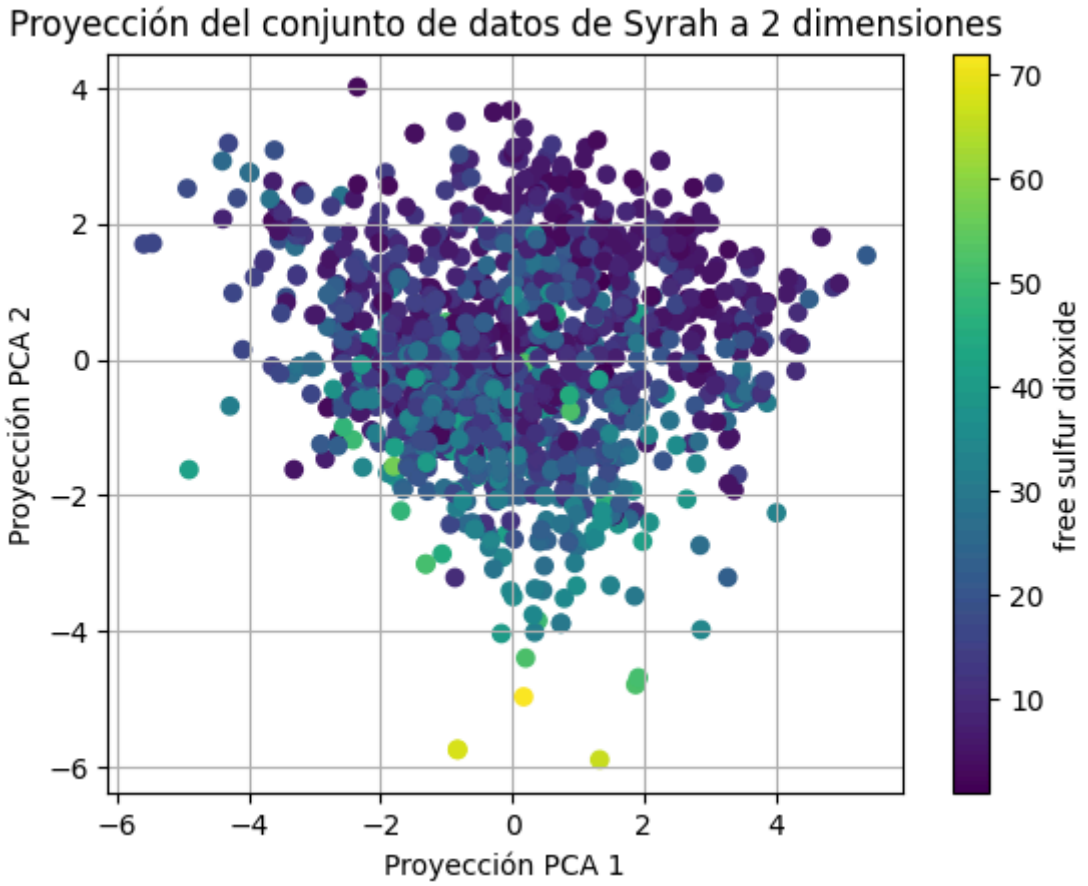

Gráfico 10

Por lo tanto, se puede decir que **en los vinos Syrah, a menor cantidad de dióxido de azufre libre mayor puntuación de calidad**.

Y de manera similar al caso de los Moscatel, se indagó con la variable “alcohol”:

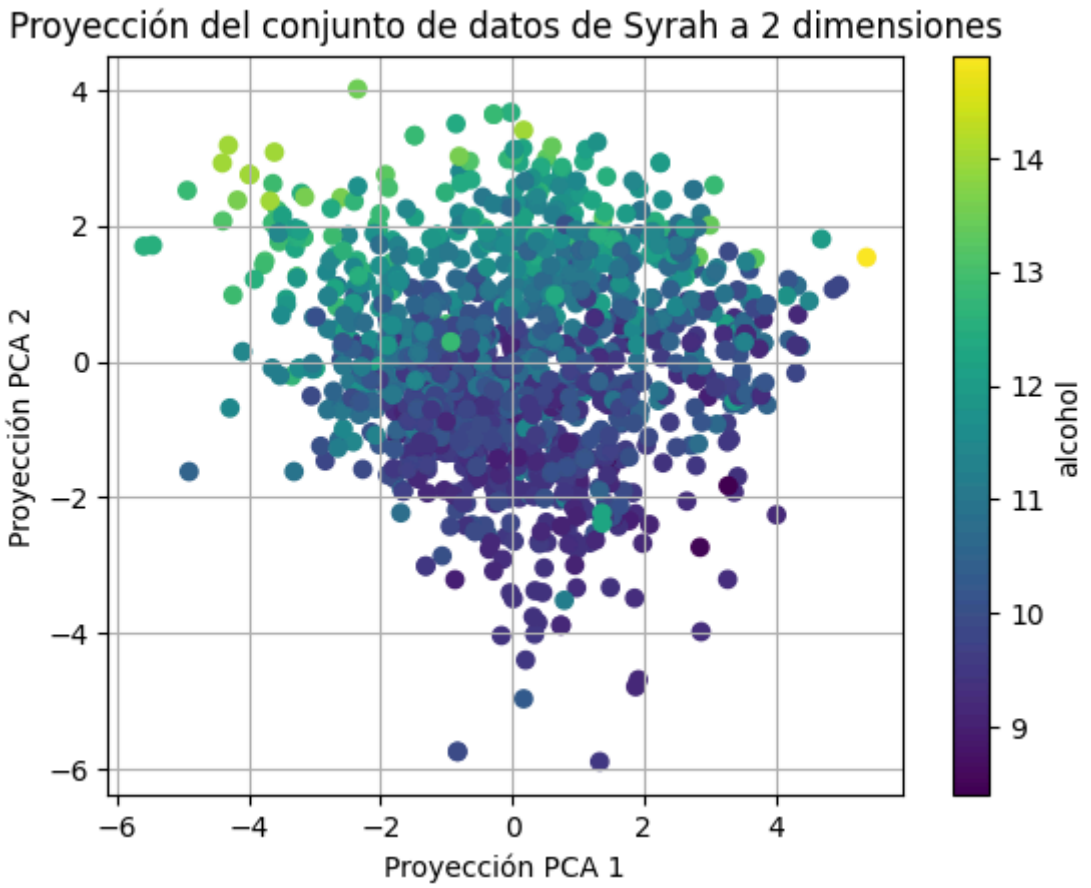

Gráfico 11

En este caso, también **en los vinos Syrah, a mayor contenido alcohólico mayor puntuación de calidad.**

Para concluir con el análisis exploratorio, cabe mencionar que utilizando PCA hay baja reconstrucción. Ya que la varianza explicada es del 53% para el dataset completo, lo que indica que hay otro 47% de los datos que no se está contemplando.
Por este motivo es que se usará regresión logística como una herramienta de confirmación o validación de las hipótesis planteadas, entrenando un modelo con la mejor precisión posible.

REGRESIÓN LOGÍSTICA

Por tipo

Con el objetivo de fortalecer las hipótesis que surgieron durante el análisis exploratorio, se utilizó la regresión logística para obtener un modelo que intente predecir el tipo de uva de un vino.
Para ello, primero se estratificó el dataset y se estandarizaron los conjuntos de entrenamiento y muestra.
Luego se buscó el hiperparámetro que devuelva los mejores resultados y se obtuvieron los siguientes pesos para cada variable:

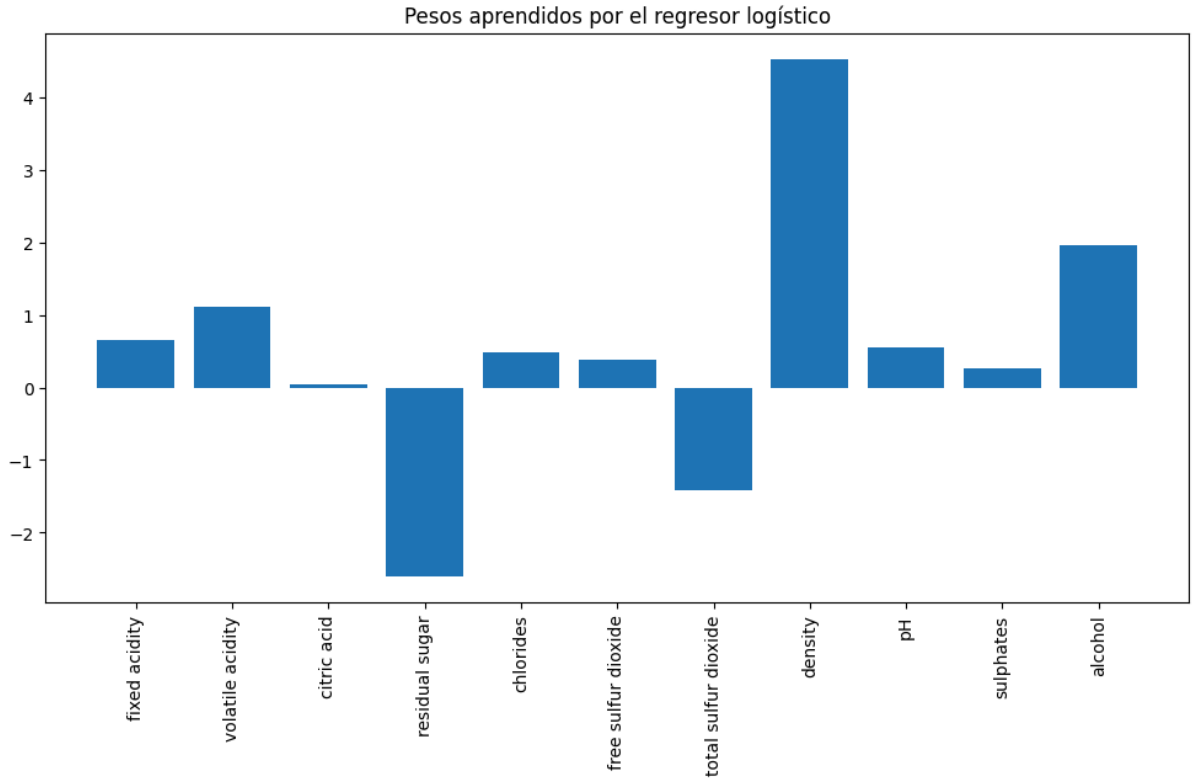

Gráfico 12

Como se puede ver, “density” tiene el mayor peso dentro de todas las variables en diferenciar el tipo de uva de los vinos. Se puede ver entonces, una coincidencia con lo planteado en el análisis multivariado: **los vinos Syrah tienen mayor densidad que los vinos Moscatel**.
Un caso similar es el de “total sulfur dioxide”, aunque esta variable no tenga tanto peso como otras. Previamente se había planteado: **los Moscatel tienen mayor dióxido de azúfre total que los vinos Syrah**.
Vale la pena destacar también aquellas variables que no fueron especialmente identificadas en el análisis multivariado pero que incrementaron en significancia con este nuevo gráfico.
Se anuncian entonces, estas posibles hipótesis:

- **Los vinos Moscatel tienen mayor azúcar residual que los vinos Syrah.**
- **Los vinos Moscatel tienen mayor alcohol que los vinos Syrah.**

El modelo logístico que se entrenó tiene una exactitud del 99%.

Por calidad

Para profundizar el análisis sobre la variable objetivo del dataset (“quality”), se utilizó regresión logística para obtener un modelo que intente predecir la calidad de un vino.
Para ello, primero se estratificó el dataset y se estandarizaron los conjuntos de entrenamiento y muestra.
Luego se buscó el hiperparámetro que devuelva los mejores resultados y se obtuvieron los siguientes pesos para cada variable:

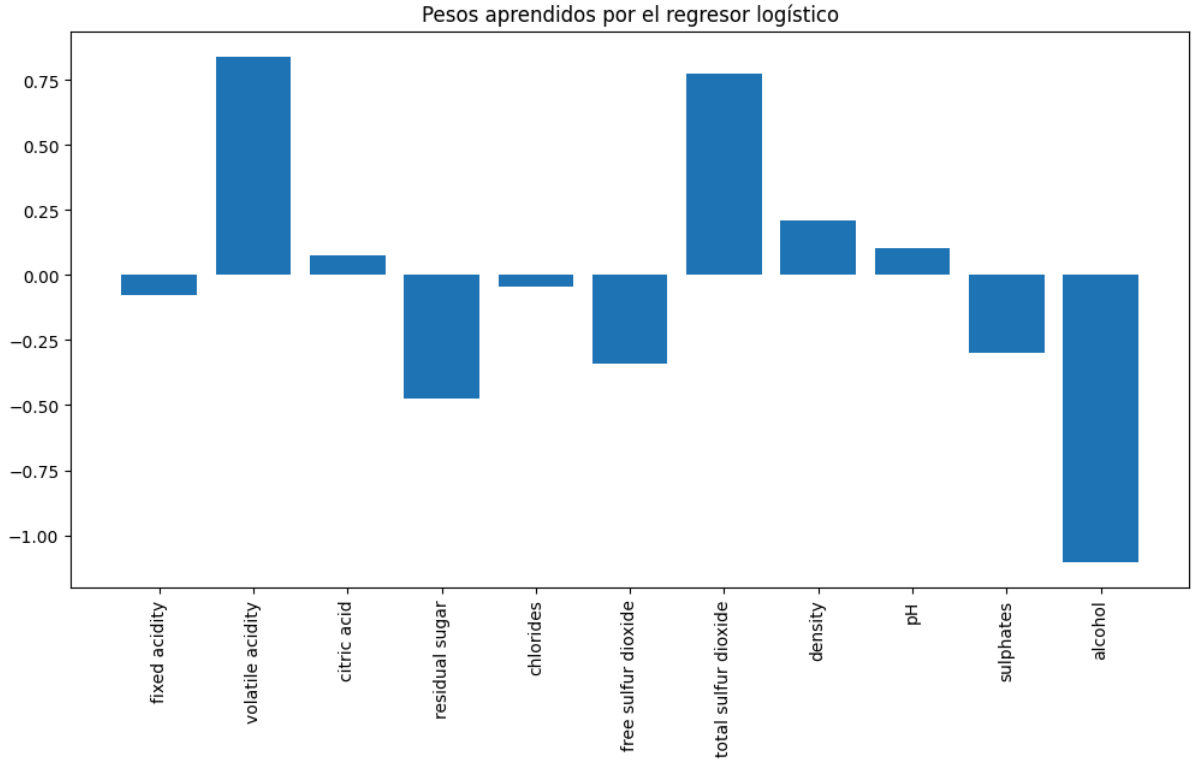

Gráfico 13

Como se puede observar, la variable “alcohol” es la que tiene mayor peso. Previamente se planteó una hipótesis que involucraba esta variable pero sólo se tenía en cuenta para los vinos del tipo de uva Moscatel. Se la amplió entonces a: **Los vinos con mayor contenido alcohólico tienen mayores puntuaciones de calidad.**
A partir del gráfico también se podría identificar a la variable “volatile acidity” como gran influyente en la calidad del vino. Se planteó entonces, la siguiente hipótesis:

- **Los vinos con mayor cantidad de ácidos volátiles tienen menor puntuación.**

El modelo logístico que se entrenó tiene una exactitud del 75%.

HIPÓTESIS Y RESULTADOS

Para las hipótesis que se fueron planteando en el análisis multivariado y lo visto con la regresión logística, se puede ver que estas hipótesis involucran variables que, en conjunto, tienen una significancia en qué tipo de vino es una muestra, ¿univariadamente es lo mismo? A continuación se desarrollaron algunas de estas con el objetivo de responder esta pregunta.

1. Los vinos con mayor densidad tienen menor porcentaje de alcohol.
2. En los vinos Moscatel, los que tienen mayor azúcar residual tienen mayor concentración de sólidos disueltos (densidad).
3. En los vinos Syrah, a mayor alcalinidad (pH) menor acidez no volátil.
4. **Los vinos Moscatel tienen mayor dióxido de azúfre total que los vinos Syrah.**
5. **Los vinos Syrah tienen mayor densidad que los vinos Moscatel.**
6. Los vinos Moscatel tienen mayor cantidad de dióxido de azufre libre que los vinos Syrah.
7. Los vinos Syrah tienen una mayor concentración de cloruros que los vinos Moscatel.
8. Los vinos Syrah tienen mayor acidez volátil que los vinos Moscatel.
9. En los vinos Moscatel, aquellos con mayor densidad tienen menores puntuaciones de calidad.
10. En los vinos Syrah, a menor cantidad de dióxido de azufre libre mayor puntuación de calidad.
11. En los vinos Syrah, a mayor contenido alcohólico mayor puntuación de calidad.
12. **Los vinos Moscatel tienen mayor azúcar residual que los vinos Syrah.**
13. **Los vinos Moscatel tienen mayor alcohol que los vinos Syrah.**
14. **Los vinos con mayor contenido alcohólico tienen mayores puntuaciones de calidad.**
15. **Los vinos con mayor cantidad de ácidos volátiles tienen menor puntuación.**

Los vinos Moscatel tienen mayor dióxido de azúfre total que los vinos Syrah.

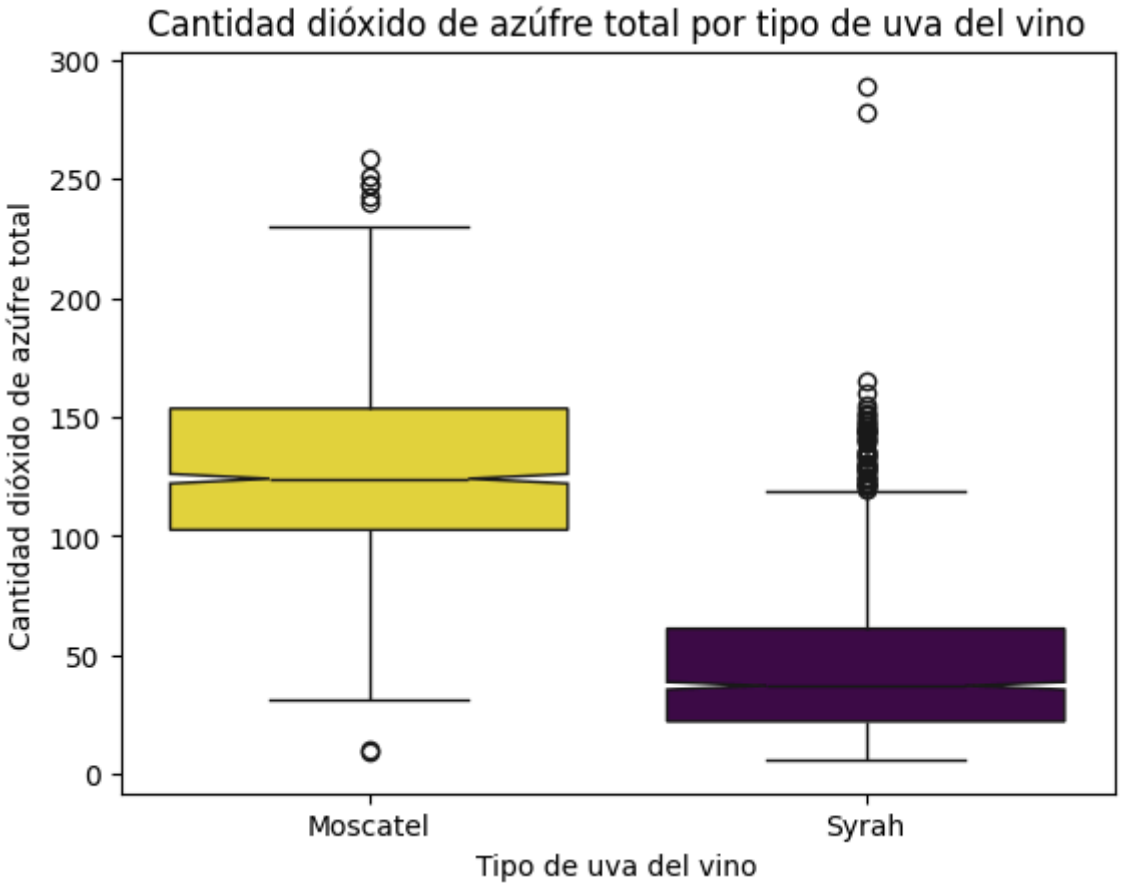

Gráfico 14

El gráfico Boxplot muestra que los notch no se superponen entre los vinos Moscatel y Syrah, lo que sugiere posibles diferencias en el contenido de dióxido de azufre total.

Pruebas Estadísticas

Se realizó el test de Shapiro-Wilk para verificar la normalidad de los datos, obteniendo p-valor=0.000 para “total sulfur dioxide” para ambos tipos de uva, lo que indica que los datos no siguen una distribución normal. Además, la prueba de Levene mostró que los datos no son homocedásticos (p-valor = 0.000), lo que indica que las varianzas entre los grupos no son iguales.

Dado que los datos no cumplen con los supuestos de normalidad ni homocedasticidad, se utilizó el test de Kruskal-Wallis. El p-valor resultante fue 0.000, lo que llevó a rechazar la hipótesis nula, ya que no se encontró evidencia suficiente para afirmar que los vinos Moscatel tienen un contenido significativamente mayor de dióxido de azufre total en comparación con los vinos Syrah.

Los vinos Syrah tienen mayor densidad que los vinos Moscatel.

Análisis Gráfico:

- Se usó un Boxplot para comparar la densidad entre los vinos Syrah y Moscatel.
- Los notch no se superponen, sugiriendo que podría haber una diferencia en las densidades de ambos tipos de vino.

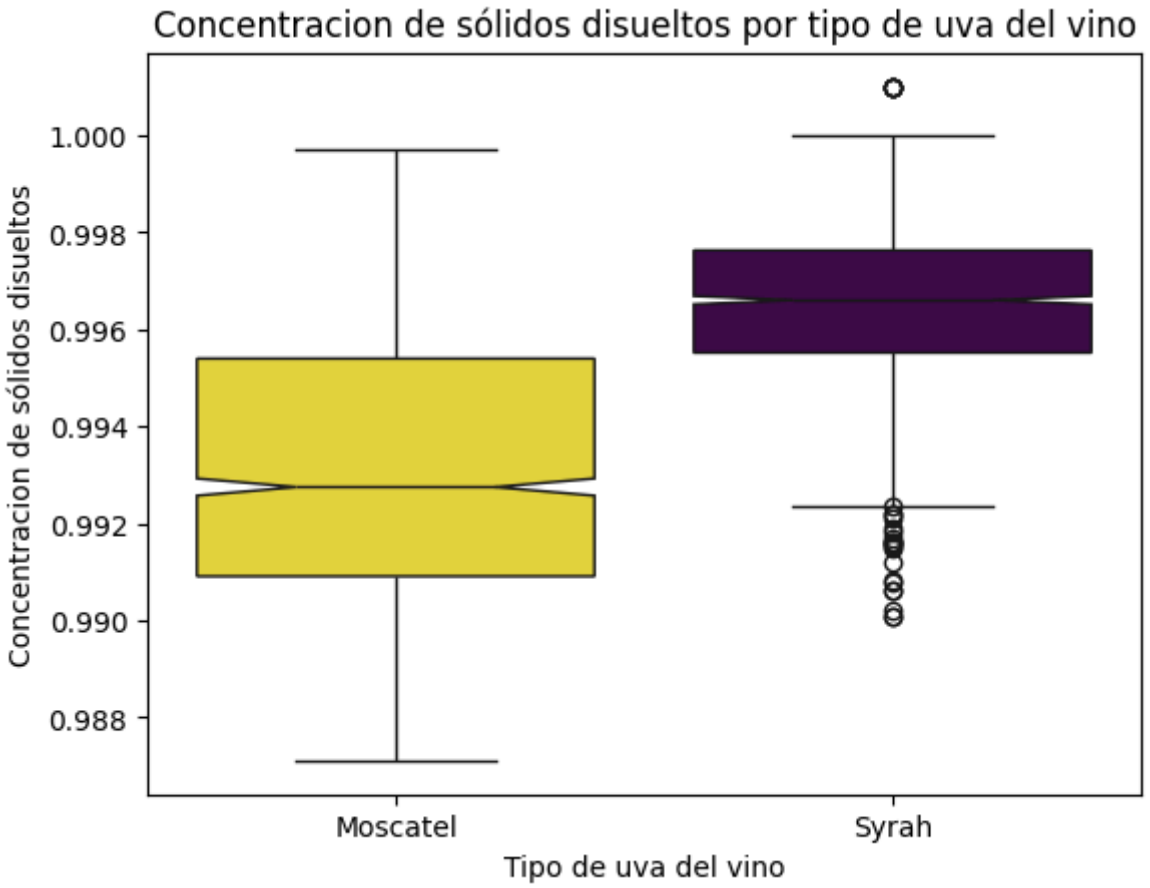

Gráfico 15

Pruebas Estadísticas

- **Shapiro-Wilk**:
  - Para Syrah: p-valor = 0.000 (no normalidad).
  - Para Moscatel: p-valor = 0.000 (no normalidad).
- **Levene**:
  - p-valor = 0.000 (no homocedasticidad).

Test de Kruskal-Wallis

- p-valor = 0.000.
- **Se rechaza la hipótesis nula.**
- No se encontró evidencia suficiente para concluir que los vinos Syrah tienen mayor densidad que los Moscatel.

Los vinos Moscatel tienen mayor azúcar residual que los vinos Syrah.

Análisis Gráfico

El gráfico boxplot mostró que los notch no se superponen, sugiriendo una diferencia en la cantidad de azúcar residual entre los vinos Syrah y Moscatel.

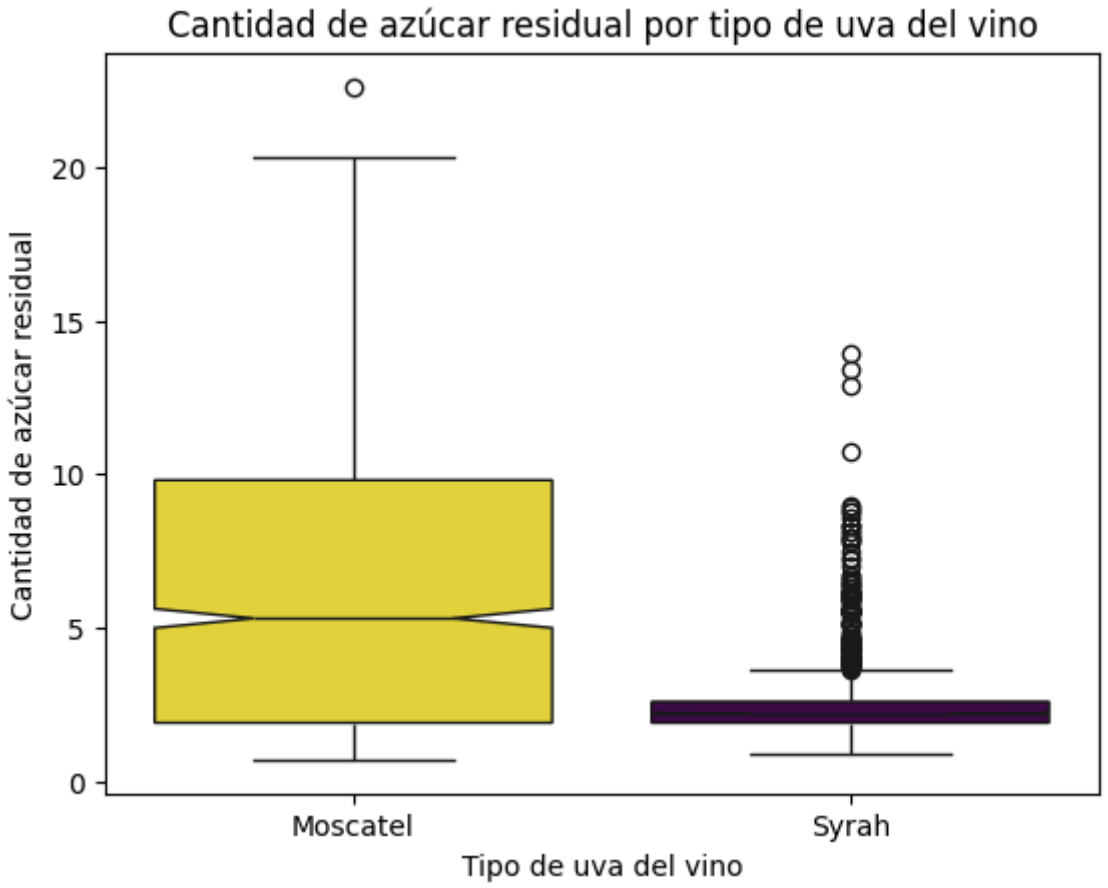

Gráfico 16

Pruebas Estadísticas

El test de Shapiro-Wilk reveló que los datos no siguen una distribución normal (p-valor = 0.000). La prueba de Levene también rechazó la homocedasticidad (p-valor = 0.000). Dado que los datos no son normales ni homocedásticos, se utilizó el test de Kruskal-Wallis, que arrojó un p-valor de 0.000, indicando que no hay una diferencia significativa en la cantidad de azúcar residual entre los vinos Syrah y Moscatel y **se rechaza la hipótesis nula**.

Los vinos Moscatel tienen mayor alcohol que los vinos Syrah

Análisis Gráfico

El gráfico Boxplot muestra que los notches no se superponen entre los vinos Moscatel y Syrah, lo que sugiere posibles diferencias en el contenido de alcohol.

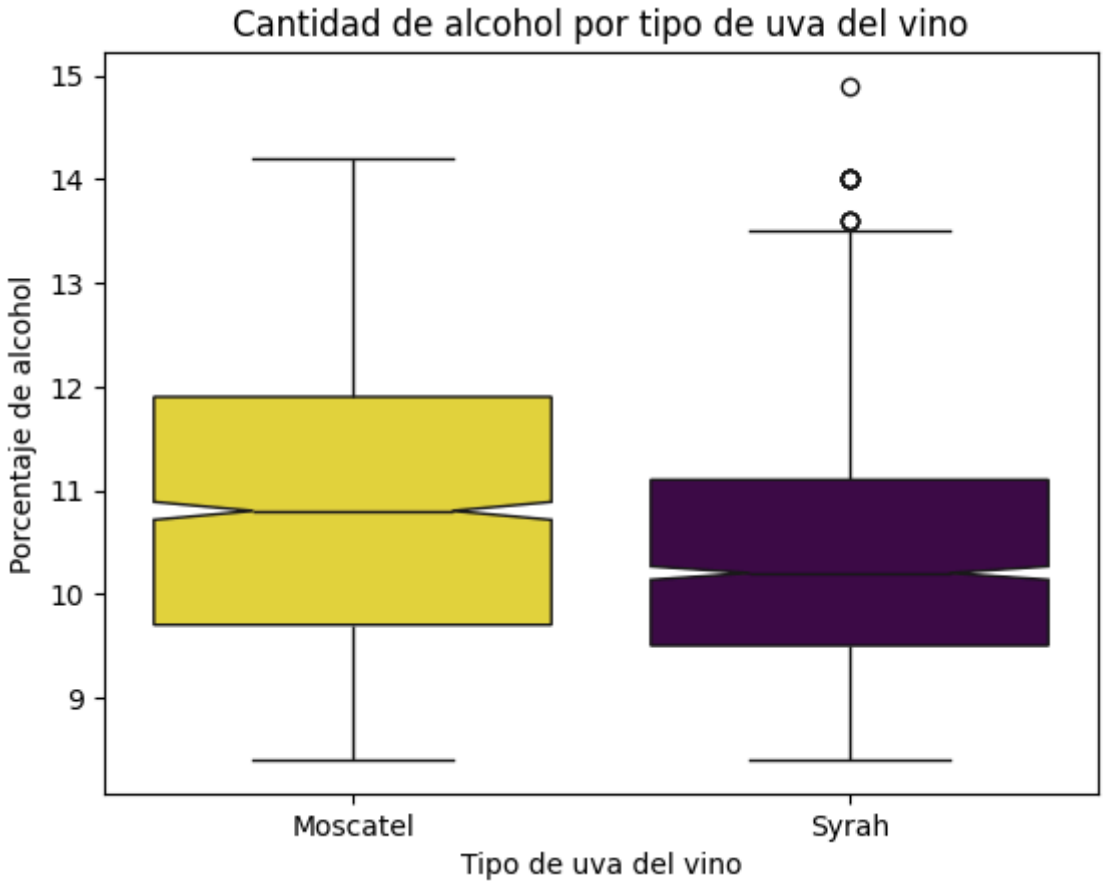

Gráfico 17

Pruebas Estadísticas

Se realizó el test de Shapiro-Wilk para verificar la normalidad de los datos, obteniendo p-valor=0.000, lo que indica que los datos no siguen una distribución normal. Además, la prueba de Levene mostró que los datos no son homocedásticos (p-valor = 0.000), lo que significa que las varianzas entre los grupos no son iguales.

Debido a la falta de normalidad y homocedasticidad, se utilizó el test de Kruskal-Wallis, el cual resultó en un p-valor de 0.000, lo que indica que **se rechaza la hipótesis nula**, es decir, no existe una diferencia significativa en el porcentaje de alcohol entre los vinos Moscatel y Syrah.

Los vinos con mayor contenido alcohólico tienen mayores puntuaciones de calidad.

Para llevar a cabo esta hipótesis en relación a la calidad, inicialmente se estableció una clasificación en dos grupos: "Malo" y "Bueno" (referido a la calidad del vino). La categoría de "Malo" incluyó las puntuaciones de 3, 4 y 5, mientras que la de "Bueno" abarcó las puntuaciones de 6, 7 y 8. Esta división en dos grupos fue preferida sobre una clasificación en tres grupos (3-4, 5-6 y 7-8), ya que esta última resultaba demasiado desequilibrada en cuanto a la distribución de las muestras. La primera categoría (3-4) solo representaba un (3%) de las 3093 muestras totales, la segunda (5-6) un 79% y la última (7-8) un 18%.

Análisis Gráfico:

- Se visualizó el contenido alcohólico según la calidad del vino usando el gráfico Boxplot, observando que los notch no se superponen. Esto inicialmente sugiere una posible diferencia en el contenido alcohólico entre vinos buenos y malos.

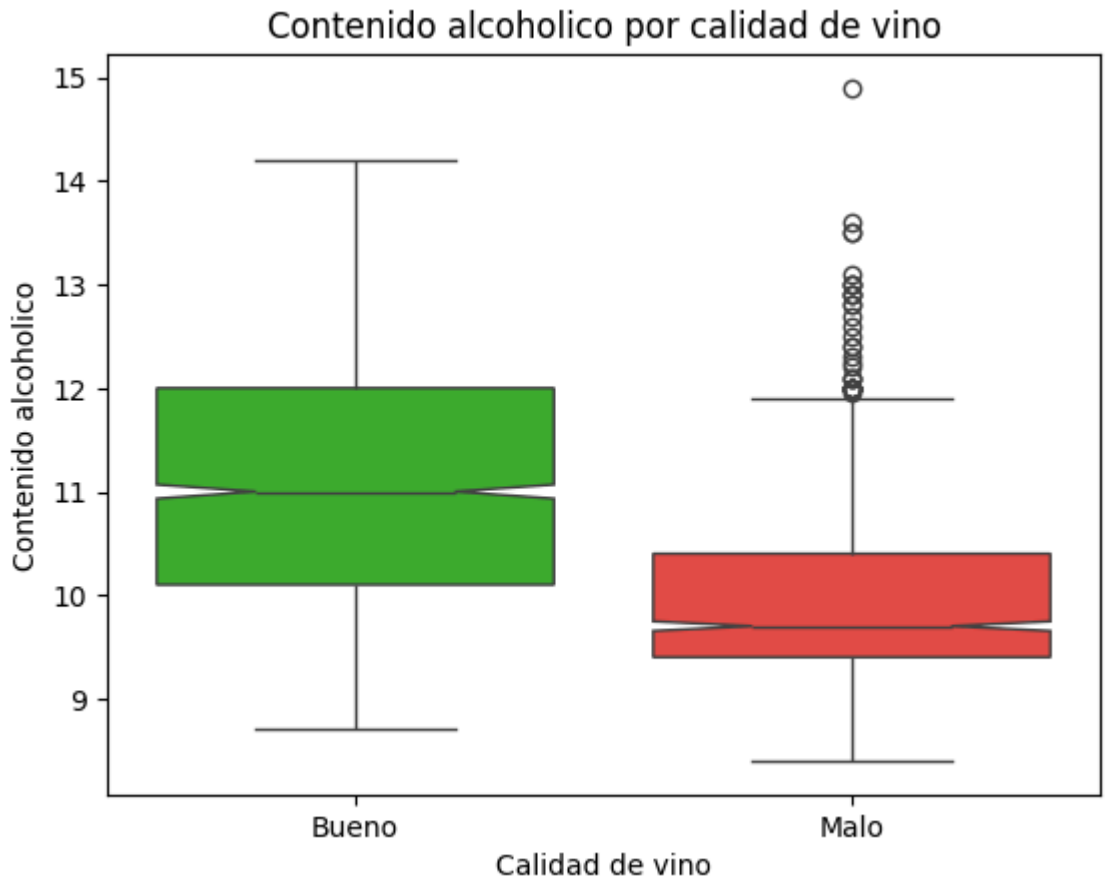

Gráfico 18

Pruebas Estadísticas:

- **Prueba de Normalidad**: Se aplicó el test de Shapiro-Wilk y se encontró que los datos no siguen una distribución normal en ambas categorías de calidad (p-valor=0.000 para ambas).
- **Prueba de Homocedasticidad**: Se usó la prueba de Levene para evaluar si las varianzas son similares entre ambas categorías. La prueba rechaza la homocedasticidad (p-valor = 0.000), indicando que las variabilidades en el contenido alcohólico entre vinos buenos y malos son diferentes.

Comparación de Grupos con Kruskal-Wallis:

- Dado que los datos no son normales y no presentan homocedasticidad, se utilizó el test de Kruskal-Wallis, que es una prueba no paramétrica para comparar los contenidos alcohólicos entre ambas categorías de calidad.
- El test de Kruskal-Wallis resultó en un p-valor de 0.000, permitiendo **rechazar la hipótesis nula** indicando que no hay diferencia significativa en el contenido alcohólico entre los vinos buenos y malos.

Los vinos con mayor cantidad de ácidos volátiles tienen menor puntuación.

Análisis Gráfico:
Se analizó la cantidad de ácidos volátiles según la calidad del vino utilizando un gráfico Boxplot, observando que los notches no se superponen, lo cual sugiere una posible diferencia en la cantidad de ácidos volátiles entre vinos de buena y mala calidad.

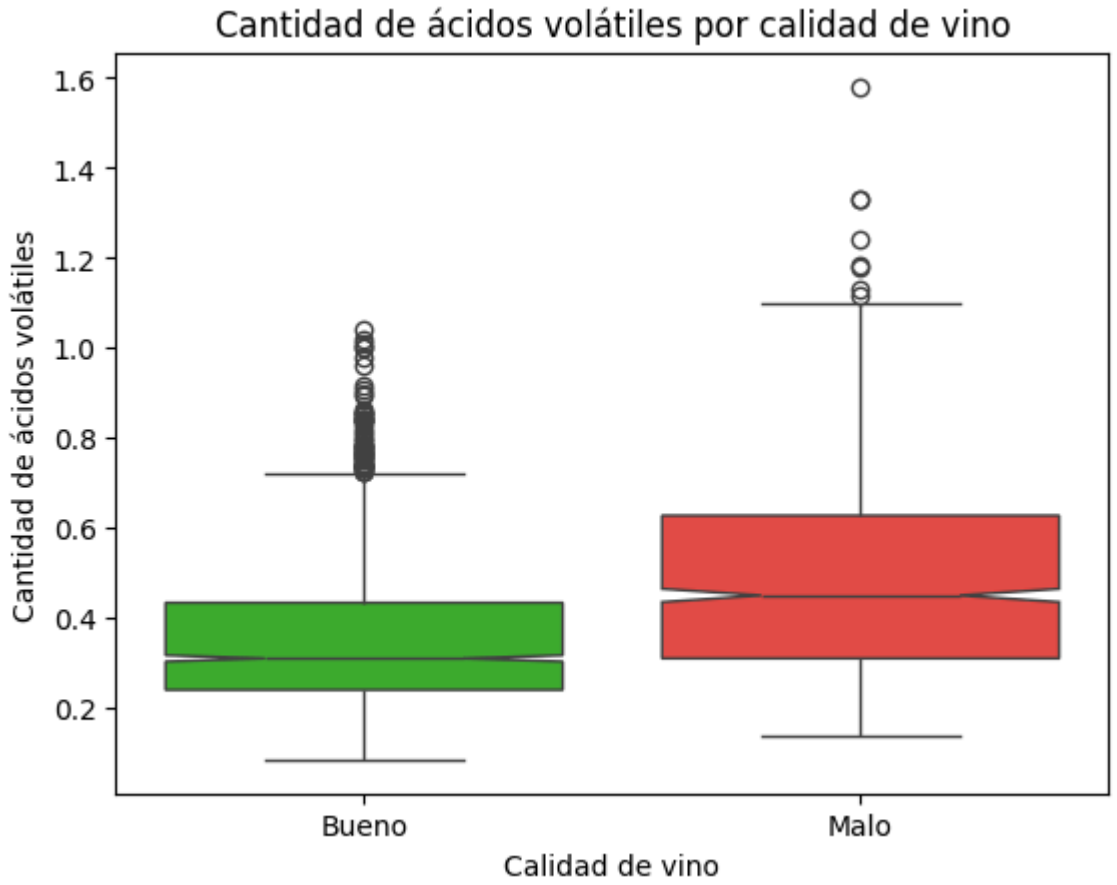

Gráfico 19

Pruebas Estadísticas:
Para evaluar la normalidad de los datos, se aplicó el test de Shapiro-Wilk a ambos grupos. El resultado (p-valor=0.000 para ambos grupos) indica que los datos no presentan distribución normal en ninguna de las categorías. Luego, se verificó la homocedasticidad con la prueba de Levene, que arrojó un p-valor de 0.000, rechazando la hipótesis de varianzas iguales, lo que sugiere diferencias en la variabilidad de la cantidad de ácidos volátiles entre vinos buenos y malos.

Comparación de Grupos con Kruskal-Wallis:
Dado que los datos no presentan normalidad ni homocedasticidad, se empleó el test no paramétrico de Kruskal-Wallis para comparar la cantidad de ácidos volátiles entre ambas categorías de calidad. El test devolvió un p-valor de 0.000, lo cual permite rechazar la

hipótesis nula. Esto sugiere que no existe una diferencia significativa en la cantidad de ácidos volátiles entre vinos buenos y malos.

CONCLUSIONES

La realización de este trabajo práctico especial invitó a indagar y conocer más sobre los vinos y sus características. Con el objetivo de entender mejor el conjunto de datos con el que se trabajó, se indagó en varias webs, tiendas y documentos legales. Nos llevó a comunicarnos con bodegas locales (con una comunicación exitosa) y con bodegas internacionales (sin haber obtenido respuesta).

En base a los análisis realizados, se plantearon hipótesis que se esperaba iban a probarse válidas pero no fue el caso con ninguna. Los p-valores obtenidos para la validación de las hipótesis en las pruebas de Kruskal-Wallis fueron todos inferiores al nivel de significancia (0.05), lo que llevó a rechazar las hipótesis nulas en cada uno de los casos. Por lo tanto, a la pregunta que se planteó, de si las variables individualmente tienen significancia en qué tipo de vino es una muestra, gracias a las hipótesis que se probaron podemos afirmar que la respuesta es no.

Parte de las decisiones sobre los datos fue el tratamiento de outliers. Al descartar outliers en todas las variables, se notó que para el análisis multivariado la varianza explicada en el PCA aumenta a un 56% pero con un costo de 600 muestras menos, por lo que se decidió no perder esa gran cantidad de muestras por obtener solamente un 3% más de representación.

REFERENCIAS

https://www.acenologia.com/compuestos_so2_fml_cienc1211/
https://www.argentina.gob.ar/normativa/nacional/resoluci%C3%B3n-76-1985-314374/texto
https://www.boletinoficial.gob.ar/detalleAviso/primera/191289/20180910
https://www.interempresas.net/Vitivinicola/Articulos/369765-Azucares-en-la-enologia-tipos-evolucion-y-metodos-de-medicion.html
https://www.argentina.gob.ar/normativa/nacional/resoluci%C3%B3n-275-2017-287281/texto
https://www.sherry.wine/es/vinos-de-jerez/vinos-dulces-naturales/moscatel
https://es.pinord.com/producto/moscatel/
https://www.usc.gal/caa/MetAnalisisStgo1/enologia.pdf
https://vinoslof.cl/wp-content/uploads/2020/09/Ficha-Tecnica-LOF-Syrah-2018.pdf
https://www.vinosdeayerbe.com.ar/producto/capitulo-i-syrah/