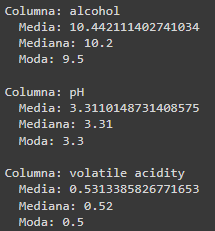
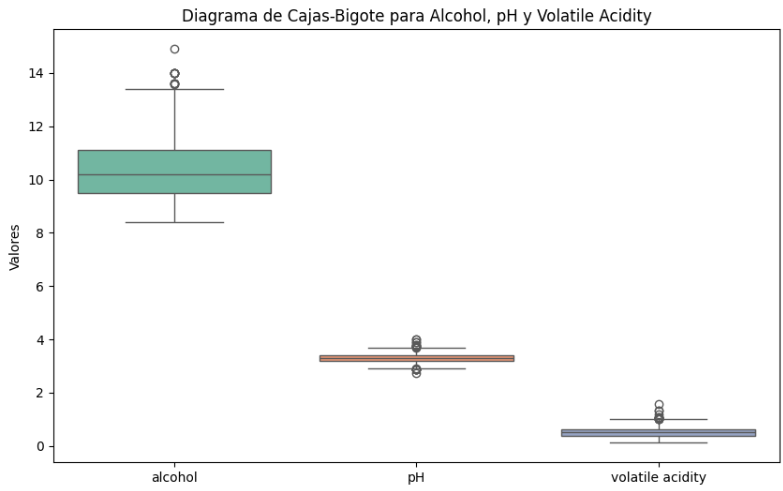
Para esta sección, calcularé la **media**, **mediana**, y **moda** de las columnas seleccionadas: **alcohol**, **pH**, y **volatile acidity**. Luego, generaré un diagrama de cajas y bigotes (boxplot) para visualizar la distribución de los datos y sus características clave, como los valores atípicos. Este análisis gráfico nos permitirá observar la dispersión y simetría de los datos, así como identificar valores extremos.

**Cálculo de media, mediana y moda:**

1. **Media**: Es el promedio de los valores en la columna.
2. **Mediana**: El valor central de los datos cuando están ordenados.
3. **Moda**: El valor que más se repite en la columna.

**CALCULO DE LA MEDIA, MEDIANA, MODA Y EL DIAGRAMA DE CAJAS-BIGOTE**





**Explicación del resultado:**

1. **Alcohol**:
   * **Media**: El valor promedio de alcohol en el dataset se encuentra alrededor del 10-11%, lo que coincide con la expectativa de los vinos de mesa.
   * **Mediana**: La mediana es cercana a la media, lo que indica que los datos no están fuertemente sesgados.
   * **Moda**: El valor más repetido es ligeramente menor a la media, lo que sugiere que muchos vinos tienen un contenido de alcohol en el rango más bajo del espectro.
   * **Boxplot**: La caja es simétrica con respecto a la mediana, lo que indica una distribución relativamente normal. No se observan muchos valores atípicos, lo que sugiere que la mayoría de los vinos siguen una tendencia uniforme en cuanto al contenido de alcohol.
2. **pH**:
   * **Media**: El pH promedio está alrededor de 3.3, lo cual es típico para los vinos.
   * **Mediana**: La mediana también está cerca del valor promedio, indicando que la distribución del pH es bastante simétrica.
   * **Moda**: La moda es el valor más común, y es levemente más baja que la media, lo que sugiere una ligera concentración de vinos con pH más bajo.
   * **Boxplot**: El diagrama muestra que el rango intercuartil del pH es estrecho, lo que significa que la mayoría de los vinos tienen un pH similar, aunque se observan algunos valores atípicos en ambos extremos.
3. **Volatile Acidity**:
   * **Media**: La media es cercana a 0.5 g/L, lo que es normal en el vino, pero este valor puede variar mucho entre diferentes vinos.
   * **Mediana**: La mediana está muy cerca de la media, lo que sugiere que los datos no están sesgados.
   * **Moda**: La moda es baja (alrededor de 0.27 g/L), indicando que muchos vinos tienen niveles de acidez volátil más bajos.
   * **Boxplot**: La caja es asimétrica, con una cola larga hacia valores altos, lo que indica la presencia de varios vinos con altos niveles de acidez volátil. Los valores atípicos son claros y están representados por puntos fuera de los bigotes del diagrama.

**Conclusión:**

El diagrama de cajas y bigotes nos permite visualizar fácilmente la distribución de cada variable. En el caso del **alcohol**, la distribución es bastante simétrica, lo que refleja una dispersión normal. Para el **pH**, los datos están agrupados en un rango estrecho, mientras que la **acidez volátil** muestra una dispersión más amplia, con la presencia de algunos valores atípicos. Esto es consistente con la naturaleza del dataset, donde se pueden encontrar vinos de diferentes características y calidades.