



ACTIVIDAD 2

Tipo actividad: Carga de datos y la visualización de información empleando las herramientas instaladas.

Para inicial la exploración de datos, abriremos R studio y crearemos un nuevo script. Esto se logra haciendo clic en el botón que tiene una hoja blanca y un ícono verde con el símbolo de + en la parte superior izquierda de la ventana. Luego seleccionaremos la opción R script como se muestra en la figura 1.

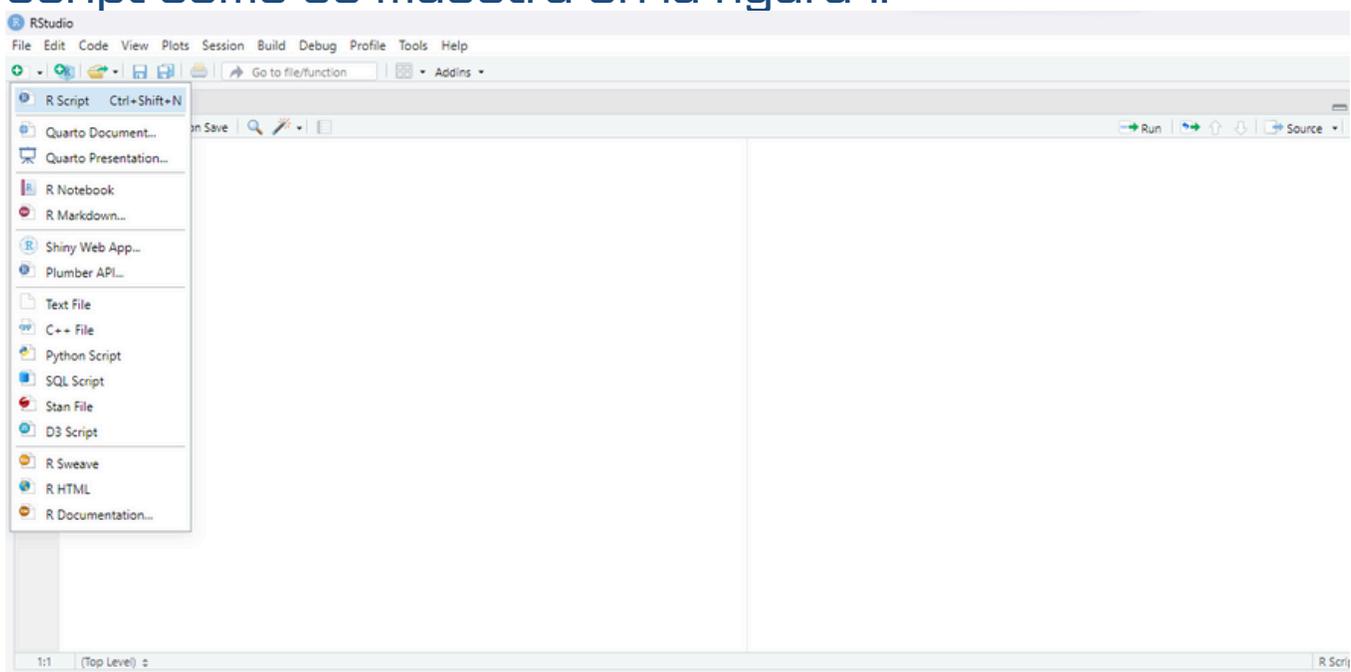


Figura 1: creación de un nuevo R script.

Luego, guardaremos el script dando clic en el botón del diskette (o con la combinación de teclas ctrl+s) y escogeremos una carpeta. Se recomienda crear una carpeta vacía en un lugar de fácil acceso para facilitar el trabajo con los datos.



Posteriormente descargaremos el dataset de vinos (wine) desde el enlace: <https://archive.ics.uci.edu/dataset/186/wine+quality>. Y lo descargaremos haciendo clic en el botón Download. Esto descargará un archivo .zip que debe descomprimirse en la misma carpeta en donde guardamos el script de R creado.

El dataset de vinos contiene datos resultantes del análisis químico de vinow cuyas uvas han crecido en la misma región de Italia pero de tres cultivos diferentes. El análisis ha terminado las cantidades de 13 componentes encontrados en los tres tipos de vinos. El dataset tiene 178 observaciones con 12 componentes analizados.

Cada una de las características del conjunto de datos contiene::

- Fixed Acidity/Acidez fija: Representa la cantidad de ácidos fijos en el vino, como ácido tartárico. Es importante para el equilibrio y la estructura del vino.
- Volatile Acidity/Acidez volátil: Indica la cantidad de ácidos volátiles presentes en el vino, como el ácido acético. Un exceso de acidez volátil puede dar lugar a sabores avinagrados no deseados.
- Citric Acid/Acido cítrico: Mide la cantidad de ácido cítrico en el vino, el cual puede contribuir a la frescura y la acidez del vino.
- Residual Sugar/Azúcar residual: Indica la cantidad de azúcar que queda en el vino después de la fermentación. Este factor influye en el sabor y la dulzura percibida del vino.



- Chlorides/Cloruros: Representa la concentración de cloruros en el vino, que pueden provenir de la uva, el suelo o los procesos de vinificación.
- Free Sulfur Dioxide/Dióxido de azufre libre: Indica la cantidad de dióxido de azufre presente en forma libre en el vino, el cual actúa como antioxidante y agente antimicrobiano.
- Total Sulfur Dioxide/Dióxido de azufre total: Representa la suma de dióxido de azufre libre y unido en el vino. Este parámetro también está relacionado con la estabilidad y la conservación del vino.
- Density/Densidad: La densidad del vino, que está influenciada por la cantidad de azúcar y alcohol presentes.
- pH/pH: Una medida de la acidez o alcalinidad del vino. Un pH más bajo indica mayor acidez, mientras que un pH más alto indica menor acidez.
- Sulphates/Sulfatos: La cantidad de sulfatos presentes en el vino, que también pueden actuar como antioxidantes y ayudar a proteger el vino contra la oxidación.
- Alcohol: El contenido de alcohol del vino, generalmente expresado como un porcentaje del volumen.
- Quality/Calidad: Una calificación de calidad asignada al vino, que fue determinada por expertos catadores.

Nota para el docente: continuar el ejercicio en el cuaderno de R adjunto, en donde se realiza y explica el paso a paso del ejercicio.