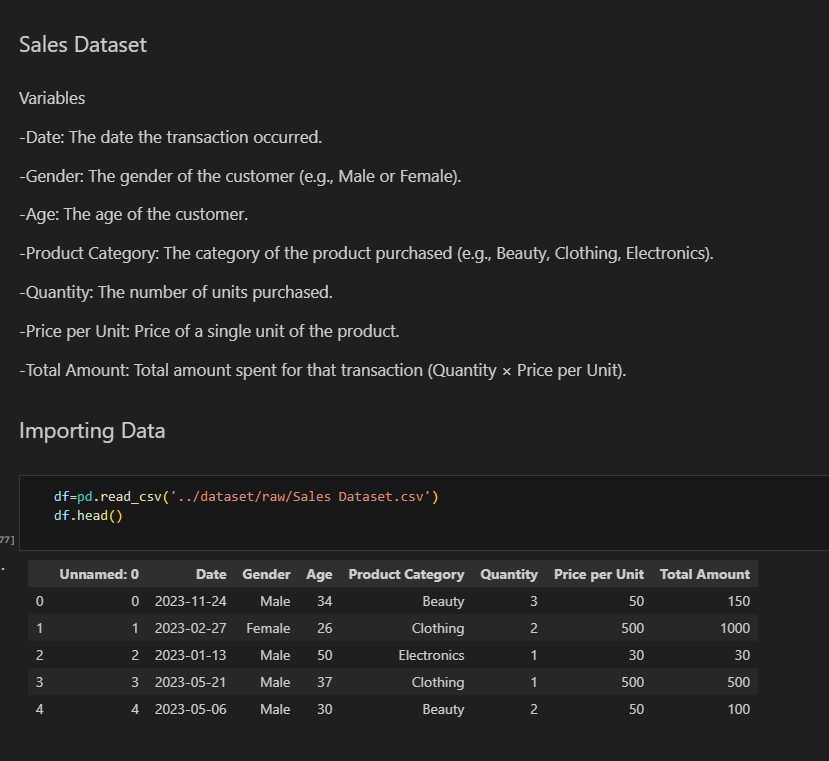
Para mi 1er proyecto para mi portafolio de data science escogi este dataset:



Primero en mi notebook visualize las primeras filas de dataframe, el nombre de las columnas, tipos de datos.

Para limpiar los datos

Elimine la columna “ Unnamed: 0” ya que solo ofrece uso de memoria innecesaria  
Con la funcion info(memory\_usage=”deep”), pude ver que no hay valores null y la cantidad de memoria que usa el dataframe(163.6 kb)

Optimizando datos

Converti algunas variables a valores que sean mas pequeños, como int8 y float32.

Separe la columna Gender a dos columnas Male y Female con valores True y False (Dummy variables)

Coloque la columna Date como el index ya que muchas de las graficas que hare sera usando ese columna.

Al final de la optimizacion volvi a chequear la memoria usada por el dataframe ahora con los cambios realizados y bajo hasta 83kb, siendo esto un 49% menos que antes.