**Objetivo: analizar una base de datos de un supermercado**

Datos:

Los datos en están en el archivo de texto enviado (datos\_arts.csv ) separado por el carácter coma (“ , “)

Los datos tienen el siguiente formato:

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| grupo | depto | item | articulo | fecha | precio | unidades | importe |
| 15 LIMPIEZA DEL HOGAR | CUIDADO DE LA ROPA | 1000025719 | 1000025719 JABON EN BARRA BULL DOG GLICERINA 200.00 G | 19/08/2020 | 54 | 7 | 378 |
| 15 LIMPIEZA DEL HOGAR | CUIDADO DE LA ROPA | 1000025724 | 1000025724 JABON LIQUIDO PAR D.P. 800 CC NEVEX 1.00 U | 19/08/2020 | 136 | 1 | 136 |
| 15 LIMPIEZA DEL HOGAR | CUIDADO DE LA ROPA | 1000025752 | 1000025752 JABON LIQ NEVEX MATIC MULTIAC DIF DOY REMUEVE 3.00 L | 19/08/2020 | 410 | 1 | 410 |

Ejercicios:

1. Cargar el archivo **datos\_arts.csv** en Python
2. Obtener un dataframe que tenga las siguientes variables (columnas):  
   * Fecha (un registro por día)
   * Cantidad\_articulos : cantidad de artículos distintos para cada fecha
   * Total\_importe: suma de los importes para cada fecha

El dataframe debería quedar más o menos así:  
  


1. En el archivo original, crear una nueva variable (SECTOR) que tenga los siguientes valores:

* Si la variable grupo es igual a **15 LIMPIEZA DEL HOGAR** o **18 PERFUMERIA**,   
  entonces SECTOR = “**NO COMESTIBLES**”
* Si la variable grupo es igual a **4 LACTEOS** o **9 FIAMBRERIA**,   
  entonces SECTOR = “**PERECEDEROS**”
* El resto de los casos SECTOR = grupo

1. Obtener un dataframe que tenga el siguiente formato:
   * Debe haber dos columnas para cada uno de los sectores creados en el ejercicio 3, dichas columna serán:  
     - promedio de precio  
     - suma de importe

Los nombres de las columnas son discrecionales y a decisión de Uds.

* + Las filas deben ser los meses obtenidos de la columna fecha

El dataframe debería quedar más o menos así:

