

Bershka – Data Engineer

A continuación, se plantea el problema a resolver. Para ello debe emplearse el lenguaje de programación Python (última versión preferiblemente).

Microservicio con proceso ETL:

Se desea desarrollar un microservicio para el suministro de los datos de ticket medio de las diferentes tiendas disponibles. Este microservicio debe habilitar un endpoint al que solicitar dichos datos a través de una petición HTTP GET, y responder con un json que contenga la información.

Partiendo de la base de que la estructura del microservicio ya se encuentra implementada, y que al realizar la petición HTTP GET al endpoint este lanza una consulta a una fuente interna y obtiene un CSV, montar la lógica que permite transformar el CSV en un JSON.

Tanto el CSV del que partir como el JSON que se desea obtener se encuentran disponibles en la misma carpeta que el actual documento.

Observar que el CSV contiene diferentes tipos de columnas:

1. El nombre no contiene barra baja e indica que se abre anidación, por lo que en el CSV aparece sin contenido. Un ejemplo de este caso es la columna MarketsList
2. El nombre no contiene barra baja e indica que es una propiedad de primer nivel, por lo que en el CSV la columna aparece con contenido. Un ejemplo de este caso es la columna schemaVersion.
3. El nombre contiene barra baja e identifica a la izquierda de esta el origen y a la derecha el destino. Si en el CSV la columna aparece sin contenido indica una nueva anidación, por el contrario, si contiene información indica que es una propiedad del JSON. Un ejemplo de este caso son las columnas MarketsList_Market y Market_isoCode.

**** Se recomienda el uso de las librerías JSON y Pandas para el desarrollo de la funcionalidad.**