

CARGA DE DATOS

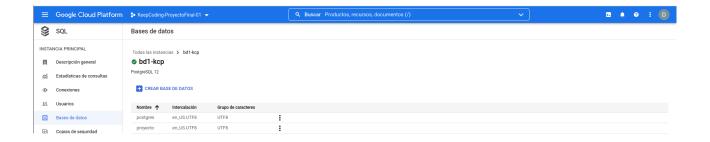
Una vez que las funciones de recolección de precios (web scrapping) y de detección se realizan, dejarán los resultados dentro de un bucket generado en Google Cloud Storage con ficheros de formatos CSV o XLSX. Es entonces cuando se ejecutará la función para cargar estos datos en una Base de Datos generada dentro un servicio Google Cloud SQL.

La estrategia de contar con una Base de Datos para esta información es la de tener de manera versátil el acceso a la información. También se necesitará contar con un historial de dicha información para su utilización en diferentes puntos de la aplicación.

La base de datos se ha generado dentro de un servicio Google Cloud SQL para contar con las ventajas que ofrece un servicio de estas características, tales como Accesibilidad, Escalabilidad y Seguridad.



La elección de PostgreSQL se debe, en primer lugar, a que se trata de un sistema de gestión de Base de Datos con licencia libre. Este sistema también cuenta con funciones para trabajar con información JSON, aspecto que en futuras evoluciones de la aplicación será muy conveniente.





Actualmente se cuentan con 4 tablas para la aplicación:

 Tabla Store. En esta tabla se almacena la información de precio obtenida a través de la función de scraping.

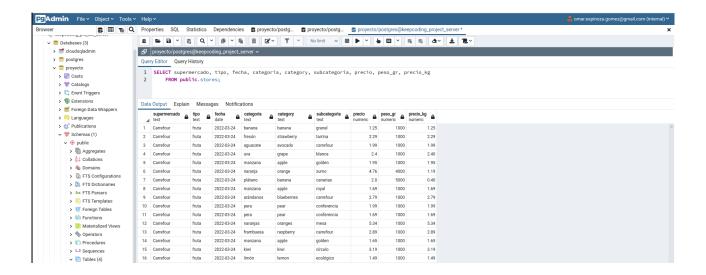


 Tabla Detección. Para cada detección realizada por la función correspondiente se genera un fichero que será almacenado en esta tabla. Se indica un ID por cada detección, y por cada clase detectada en una imagen se genera un registro con el número de elementos detectados de dicha clase.

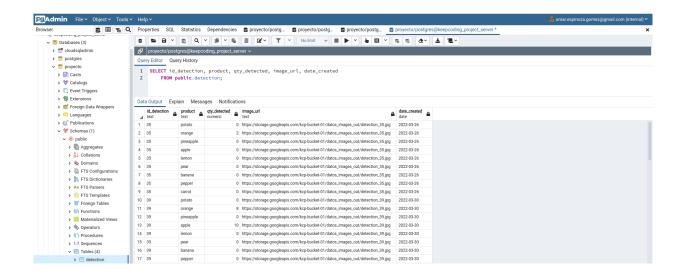




 Tabla Shoplist. Esta tabla almacena la lista de compra de los productos que el usuario requiere para una determinada fecha.

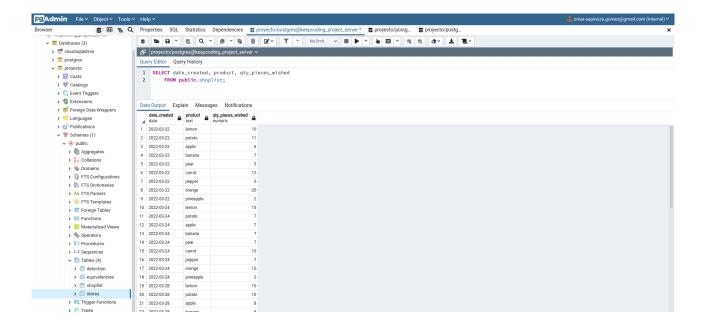
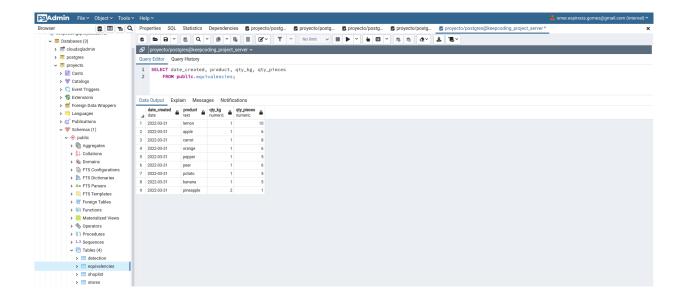


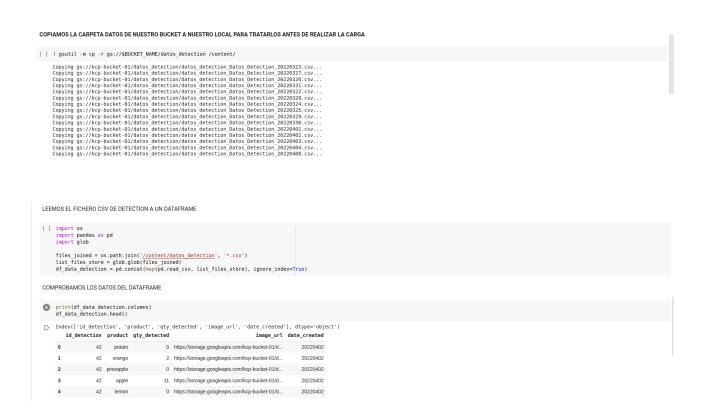
 Tabla Equivalencies. Dado que los precios de la tabla Store se determinan por kilogramos, esta tabla se genera para determinar la equivalencia en piezas de cada producto por kilo.



ALGORITMO DE DETECCIÓN



La función de Carga de Datos obtendrá los datos desde un bucket generado en Cloud Storage. Para cada tabla se tiene una carpeta correspondiente dentro del bucket donde los datos serán descargados, se procesaran y se realizará la inserción de estos mismos en la tabla.



Una vez realizada la carga de los datos en la tabla, la información de la carpeta correspondiente se moverá a la carpeta de datos_processed para dejar así libre las carpetas de datos para el siguiente envió de información por parte las funciones de recolección de precios y detección.

En futuras actualizaciones, se generaría una función para la carga de información para las tablas shoplist y equivalencies.