

## **Diseño del DAaaS**

### *Estrategia del DAaaS*

Vamos a dar acceso de forma pública a descargarse un archivo desde mi Bucket de Google Cloud en el que diariamente pondré el ranking del precio medio de la cesta de la compra de los supermercados que más venden de España.

*-La idea surge por dos motivos :*

Por que conozco cuando cambian los precios al haber trabajado en Mercadona durante 7 años.

La importancia de ahorrar en la cesta de la compra con la gran subida de precios.

*-Objetivo:*

Sin ánimo de lucro. Ayudar a que las personas compren en los supermercados más económicos.

*-Inconveniente:*

La mayoría de supermercados se ponen de acuerdo en la subida de precios.

*-Datos a incorporar:*

Nombre y precio de los 30 productos de la cesta básica de la compra de los supermercados que más venden en España. En este caso serán : Mercadona, Lidl, Carrefour y Día.

Lista de productos:

Aceite, arroz, lentejas, garbanzos, macarrones, fideos, café, galletas, leche, huevos, margarina, fregasuelos, lavavajillas, sal, gel de baño, jabón de manos, pasta de dientes, champú, tomate canario, yogures, pan de molde, pan, chopped, caldo de verduras, caballa, caldo de pollo, tomate frito, zanahorias, lechuga y bananas.

Nota : al existir diversas marcas, escogeremos el más barato de cada producto.

## *Arquitectura DAaaS*

-Listado de componentes:

Windows Task Scheduler de Windows

Archivo con script de python

Crawler con Scrapy de las páginas web

Cloud Storage(Bucket) para almacenar los datos

Dataprocc para procesamiento y limpieza de datos

Cloud SQL para hacer la query.

Cloud Function : para el levantamiento de máquinas y el pase de datos  
Utilizaremos un total de ocho

## *DAaaS Operating Model Design and Rollout*

1. Cron job (Windows Task Scheduler ) : Mi ordenador, con sistema operativo Windows, ejecutará su encendido a través del programador de tareas a las 00:00 de la noche todos los días

2. Cron job (Windows Task Scheduler ) : Ejecutamos el programador de tareas de Windows para ejecutar un script de python que contendrá el código necesario para ejecutar el crawler

3. Script de python : contendrá un crawler con scrapy con el cuál accederemos a las páginas web de supermercados obteniendo todos sus datos

4. Cloud Function : levantamos un Bucket en Cloud Storage

5. Cloud Function : Crawler ---datos---> Bucket

6. Cloud Function : levantamos un Dataproc

7. Cloud Function : Bucket ---datos---> Dataproc

8. Dataproc : limpieza y procesamiento de datos

9. Cloud Function : levantamos un Cloud SQL

10. Cloud Function : Dataproc ---datos---> Cloud SQL

11. Cloud SQL : Ejecutamos la query que nos establezca la media de esos productos(AVG) y nos haga un ranking de forma descendente (order by desc)

12. Cloud Function : Cloud SQL ---datos---> Bucket

13. Cron job (Windows Task Scheduler ) : Ejecutamos el programador de tareas de Windows para apagar el ordenador todos los días a las 03:00 AM

*Nota :*

-El nombre del archivo del crawler y del Bucket siempre será el mismo, de esta forma quedarán sobrescritos los datos, sin guardar histórico.

## Desarrollo de la plataforma DAaaS

-Noticia que me inspiró:

<https://www.20minutos.es/noticia/5105590/0/la-cesta-de-la-compra-basica-sube-por-primera-vez-tras-la-bajada-del-iva-asi-varia-su-precio-segun-los-supermercados>

-Supermercados que más venden en España:

[https://www.consumidorglobal.com/noticias/noticias-servicios/ranking-actualizado-supermercados-mas-venden-espana\\_4770\\_102.html](https://www.consumidorglobal.com/noticias/noticias-servicios/ranking-actualizado-supermercados-mas-venden-espana_4770_102.html)

-Páginas web que vamos a crawlear:

<https://www.lidl.es>

<https://www.carrefour.es>

<https://www.dia.es>

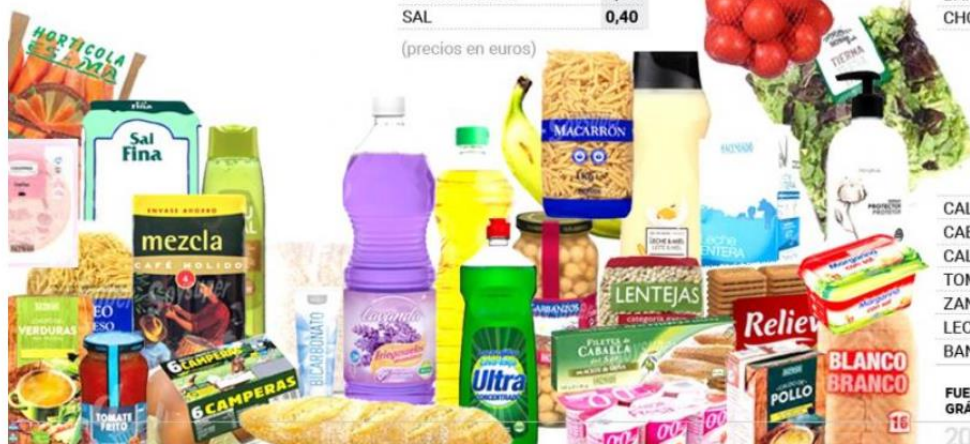
<https://www.mercadona.es>

### La cesta de la compra por 30 euros

ACEITE GIRASOL	2,55	CAFÉ	1,35
ARROZ	0,95	GALLETAS	1,35
LENTEJAS	1,75	LECHE	0,78
GARBANZO COCIDO	0,75	HUEVOS (6)	1,20
MACARRONES	0,77	MARGARINA	1,35
FIDEOS	0,78	FRIEGASUELOS	0,95
		LAVAJA JILLAS	1,10
		SAL	0,40

(precios en euros)

GEL DE BAÑO	1,00
JABÓN MANOS	0,80
PASTA DE DIENTES	1,05
CHAMPÚ	0,75
TOMATE CANARIO (5)	1,20
YOGURES (6)	0,85
PAN DE MOLDE	0,85
BARRA DE PAN	0,45
CHOPPED	1,25



CALDO VERDURAS	0,95
CABALLA (2 LATAS)	1,65
CALDO DE POLLO	0,75
TOMATE FRITO	0,85
ZANAHORIAS	0,59
LECHUGA	0,79
BANANAS (4)	0,84

FUENTE: Mercadona  
GRÁFICO: Carlos G. Kindelán

20minutos