

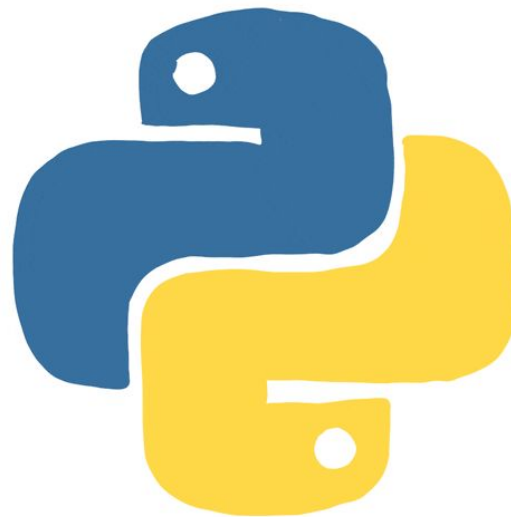
A faded, light-colored background image of a person's lower back and hands. The person's hands are placed on their lower back, suggesting discomfort or pain. The image is semi-transparent, allowing the text to be clearly visible over it.

SILLAS Y BIENESTAR

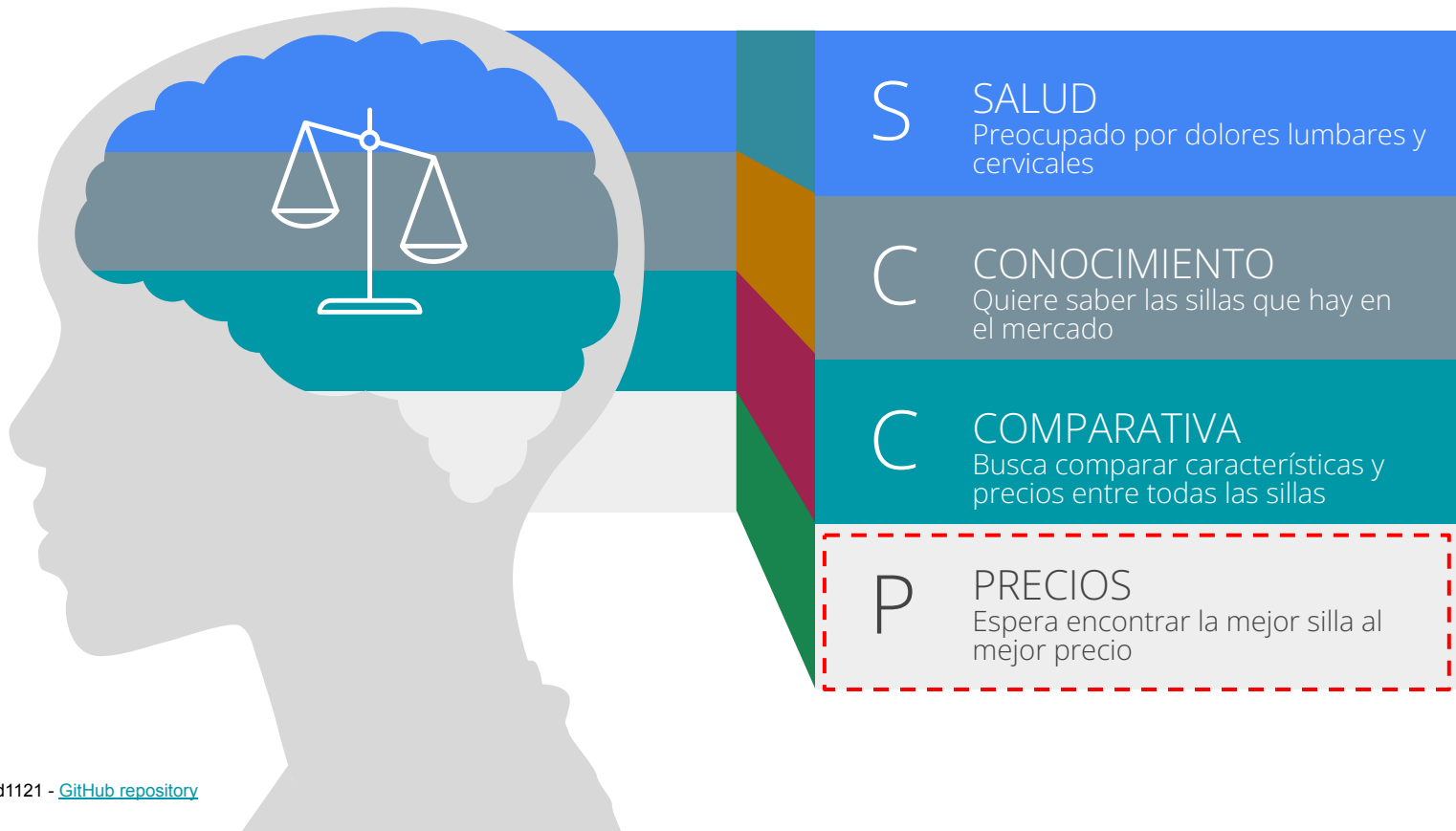
Encuentra tu silla perfecta

Evolución temporal de precios para mejorar la experiencia del usuario en la web

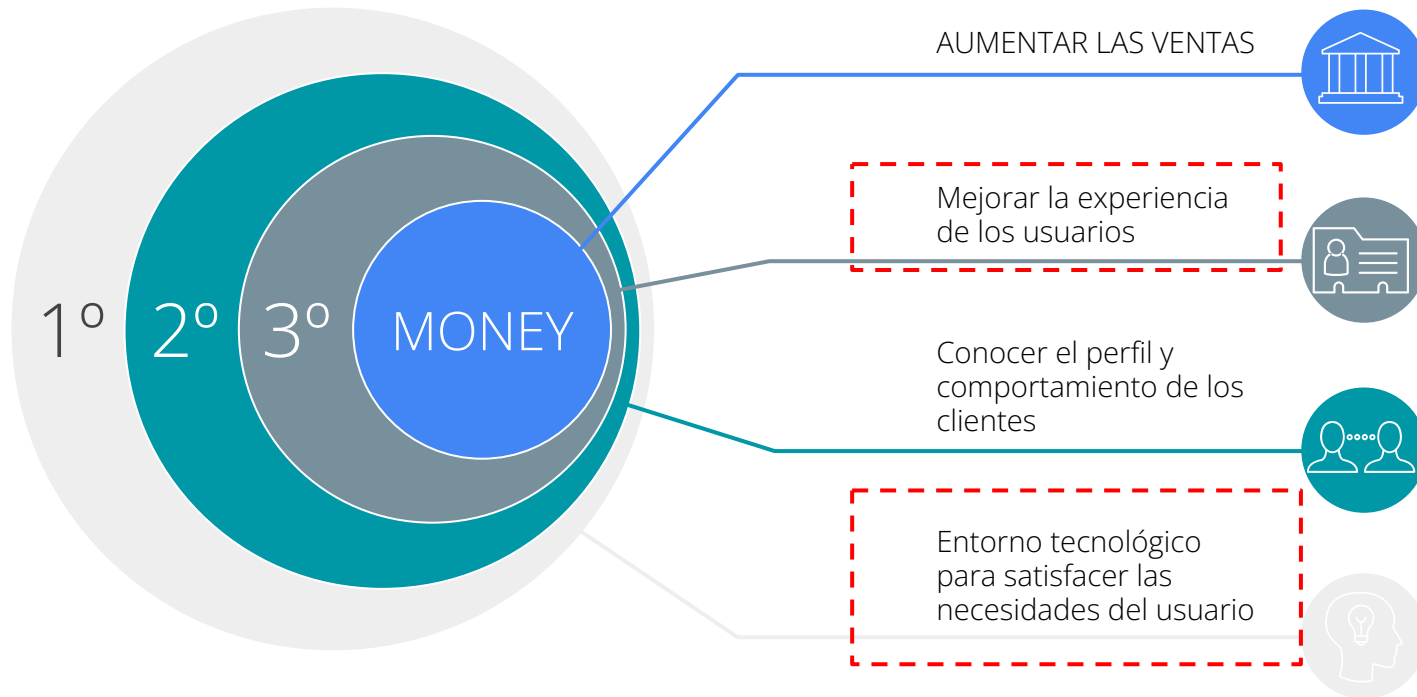
- Stakeholders y Objetivos
- Tecnologías utilizadas
- Librerías de Python
- Proceso y Outputs
 - Python Notebook | Exploratory Data Analysis
 - Gmail
 - Spreadsheets
 - Datastudio
 - Dashboarding
 - Publicación en producción en Wordpress
 - Subida del script y automatización
- Puesta común

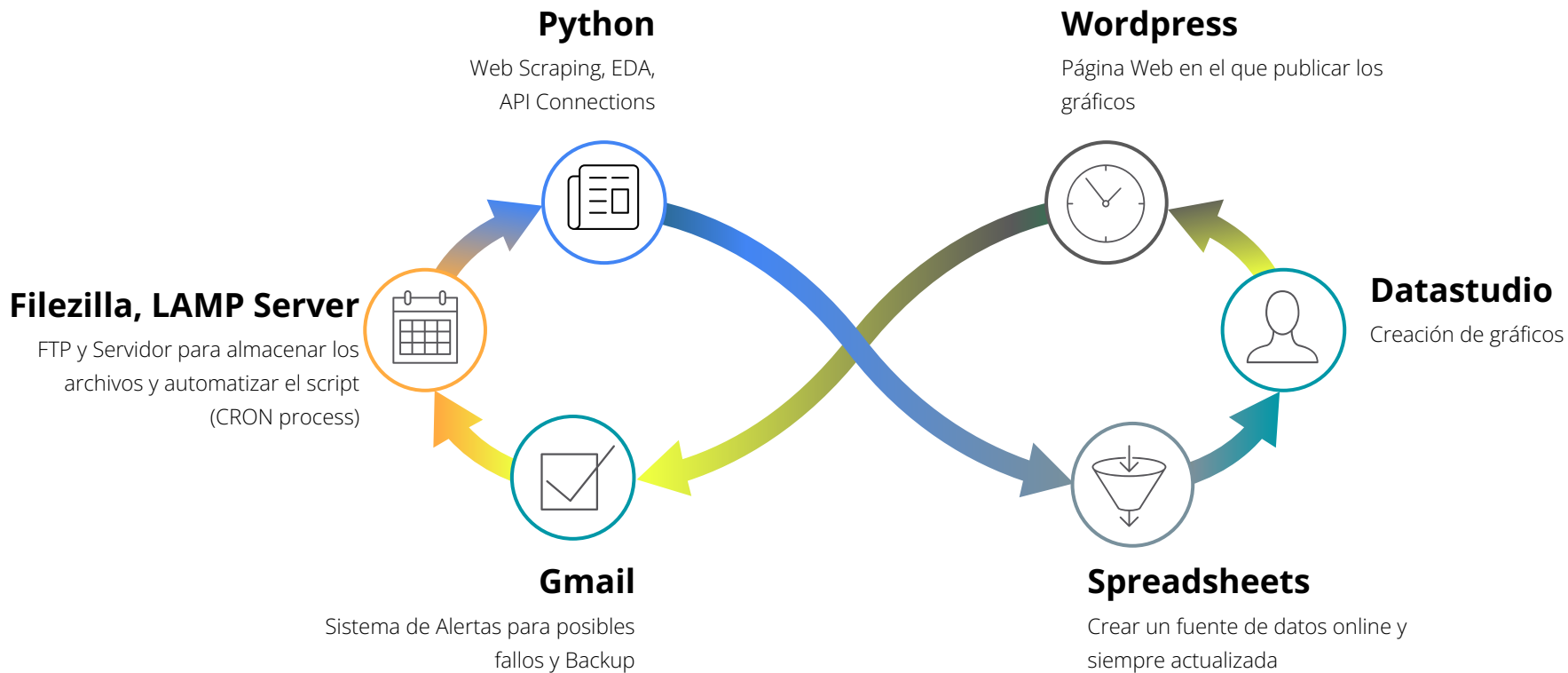


A. Usuarios



B. Administradores de la web







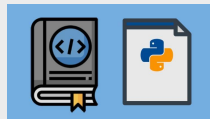
Creación
dataFrames

PANDAS



Modificación
de Arrays

NUMPY



Librería HTTP y
Web Scraping

REQUESTS, BS4

(.*)

Expresiones
Regulares

RE (regex)



Fechas en
diferentes
formatos

DATETIME



Google
Sheets

Tablas online
como BBDD

GSPREAD



Alertas y
Backup vía
Email

EMAIL, SMTPLIB, SSL



Gestión de
credenciales

DOTENV



p_acquisition

01

- Definición de URLs a scrapear
- Establecer las credenciales
- Web Scraping
- Creación de fechas
- Conexión con Spreadsheets



p_wrangling

02

- dataframe con datos totales
- dataframe con datos diarios
- Concatenación dataFrames
- Limpieza y Preparación
- Exportación a CSV



p_analysis

03

- Estudio de valores anómalos para crear alertas
- Identificación de productos sin stock o descatalogados
- Creación de backup
- Creación de alertas vía Email

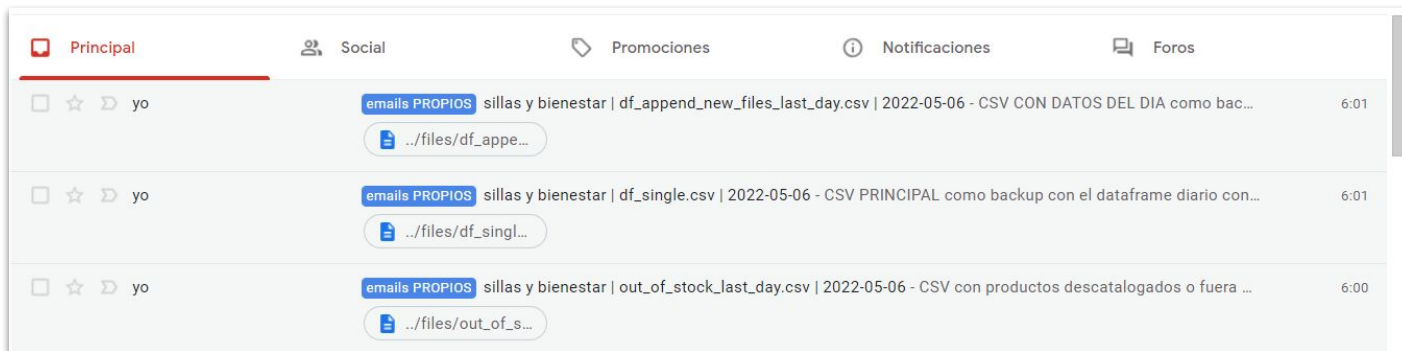


p_reporting

04

- Conexión con Spreadsheets:
 - a) MAIN dataframe
 - b) daily scrapped data dataframe
 - c) out of stock or discontinued dataframe
 - d) none values dataframe

PROCESO & OUTPUTS: 2º Conexión Python y Gmail



Cada día llegan 3 emails a las 6:00:

- alerta de productos “descatalogados” o “sin stock”
- backup del dataframe principal
- backup de los datos diarios del web scraping

PROCESO & OUTPUTS: 3º Conexión Python - Spreadsheets - Datastudio - WP



Data Studio



WORDPRESS

PROCESO & OUTPUTS: 3º Conexión Python - Spreadsheets - Datastudio - WP



business_afi_scraping_df_single

Archivo Editar Ver Insertar Formato Datos Herramientas Extensiones Ayuda Última modificación hace 7 horas por

100% € % .0_ .00 123 Predetermi... 10 B I G A

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O
	date_hyphen	date_slash	date_number	product_name	product_id	product_brand	price	status	url	product_mean	product_mean3	product_category	product_count	product_mean3	product_mean30
1	2022-04-14	2022/04/14	20220414	TLV TLV-MYX-801-1	B08T1TSMQO	TLV	89	correcto	https://sillas-ergonomicas.com/ergonomia/sillas-ergonomicas/review-individual/tlv-myx-801-1/	89	89	ergonomia	1	89	89
2	2022-04-14	2022/04/14	20220414	Songmics OBN55BK	B07ZCJVFDJ	Songmics	0	sin_stock	https://sillas-ergonomicas.com/ergonomia/sillas-ergonomicas/review-individual/songmics-obn55bk/	131	131	descatalogado	0	131	131

business_afi_scraping_last_day_files

Archivo Editar Ver Insertar Formato Datos Herramientas Extensiones Ayuda Última modificación hace 7 horas por ironhack-alvaro-scraping-price@ironhack-alvaro-scraping-price.iam.gserviceaccount.com

100% € % .0_ .00 123 Predetermi... 10 B I G A

	A	B	C	D	E	F	G	H	I	J	K	L	M	N
	date_hyphen	date_slash	date_number	product_name	product_id	product_brand	price	status	url					
1	2022-05-06	2022/05/06	20220506	TLV TLV-MYX-801-1	B08T1TSMQO	TLV	89	correcto	https://sillasbienestar.com/ergonomia/sillas-ergonomicas/review-individual/tlv-myx-801-1/					
2	2022-05-06	2022/05/06	20220506	Songmics OBN55BK	B07ZCJVFDJ	Songmics	131	correcto	https://sillasbienestar.com/ergonomia/sillas-ergonomicas/review-individual/songmics-obn55bk/					

business_afi_scraping_last_day_files

Archivo Editar Ver Insertar Formato Datos Herramientas Extensiones Ayuda Última modificación hace 7 horas por ironhack-alvaro-scraping-price@ironhack-alvaro-scraping-price.iam.gserviceaccount.com

100% € % .0_ .00 123 Predetermi... 10 B I G A

	A	B	C	D	E	F	G	H	I	J	K
	date_hyphen	date_slash	date_number	product_name	product_id	product_brand	price	status	url		
1	2022-05-06	2022/05/06	20220506	femor B08G8DR	B08G8DR75M	femor	0	descatalogado	https://sillasbienestar.com/ergonomia/sillas-ergonomicas/review-individual/femor-b08g8dr/		
2	2022-05-06	2022/05/06	20220506	Hbada HDNY10	B07QWG9LD4	Hbada	0	sin_stock	https://sillasbienestar.com/oficina-y-escritorio/sillas-ergonomicas/review-individual/hbada-hdny10/		
3	2022-05-06	2022/05/06	20220506	gtplayer bluetooth	B082F69ZKX	gtplayer	0	sin_stock	https://sillasbienestar.com/gaming/sillas-gaming/review-individual/gtplayer-bluetooth/		
4	2022-05-06	2022/05/06	20220506	bgeu-a136	B08D8BVHMW	racing gaming	0	sin_stock	https://sillasbienestar.com/gaming/sillas-gaming/review-individual/bgeu-a136/		
5	2022-05-06	2022/05/06	20220506	femor B08JD1T	B08JD1T3JK	femor	0	descatalogado	https://sillasbienestar.com/gaming/sillas-gaming/review-individual/femor-b08jd1t/		
6	2022-05-06	2022/05/06	20220506	dxracer king k06	B00ZRVDEJ6	dxracer	0	sin_stock	https://sillasbienestar.com/gaming/sillas-gaming/review-individual/dxracer-king-k06/		
7	2022-05-06	2022/05/06	20220506	DXRacer Formu	B073GF2L1H	dxracer	0	sin_stock	https://sillasbienestar.com/gaming/sillas-gaming/review-individual/dxracer-formu/		

PROCESO & OUTPUTS: 3º Conexión Python - Spreadsheets - Datastudio - WP



Data Studio

Nombre del campo	ID de campo
stock7	calc_loginvhuc
Fórmula ?	
<pre>1 CASE 2 WHEN price = 0 THEN "FUERA DE STOCK" 3 WHEN price < product_mean7 THEN "¡CHOLLO! HOY PRECIO POR DEBAJO DE LA MEDIA" 4 WHEN price > product_mean7 THEN "¡ATENCIÓN! HOY PRECIO POR ENCIMA DE LA MEDIA" 5 WHEN price = product_mean7 THEN "HOY PRECIO IGUAL A LA MEDIA" 6 END</pre>	



Data Studio



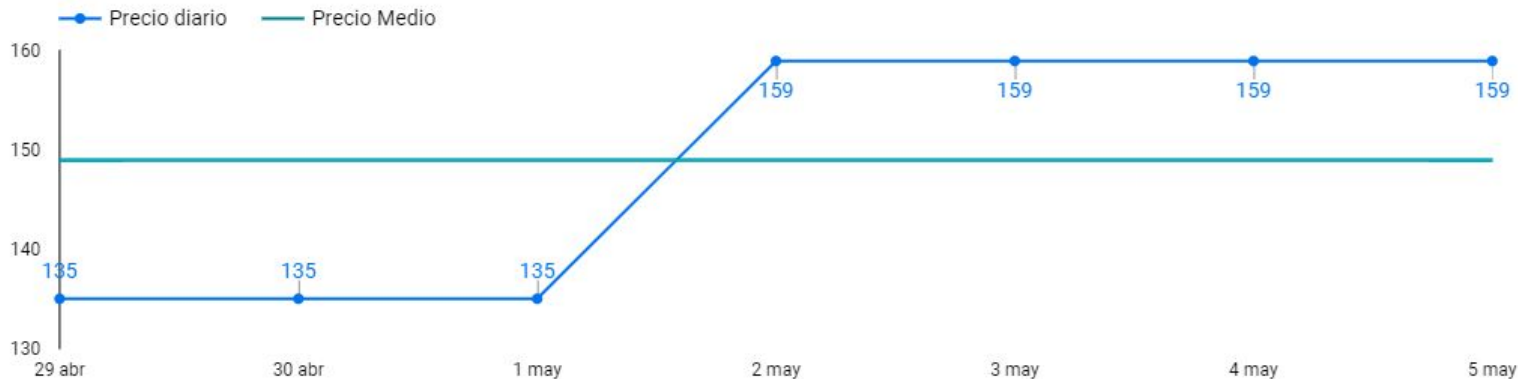
WORDPRESS

EVOLUCIÓN DEL PRECIO EN LOS ÚLTIMOS 7 DÍAS

¡ATENCIÓN! HOY PRECIO POR ENCIMA DE LA MEDIA

1

2





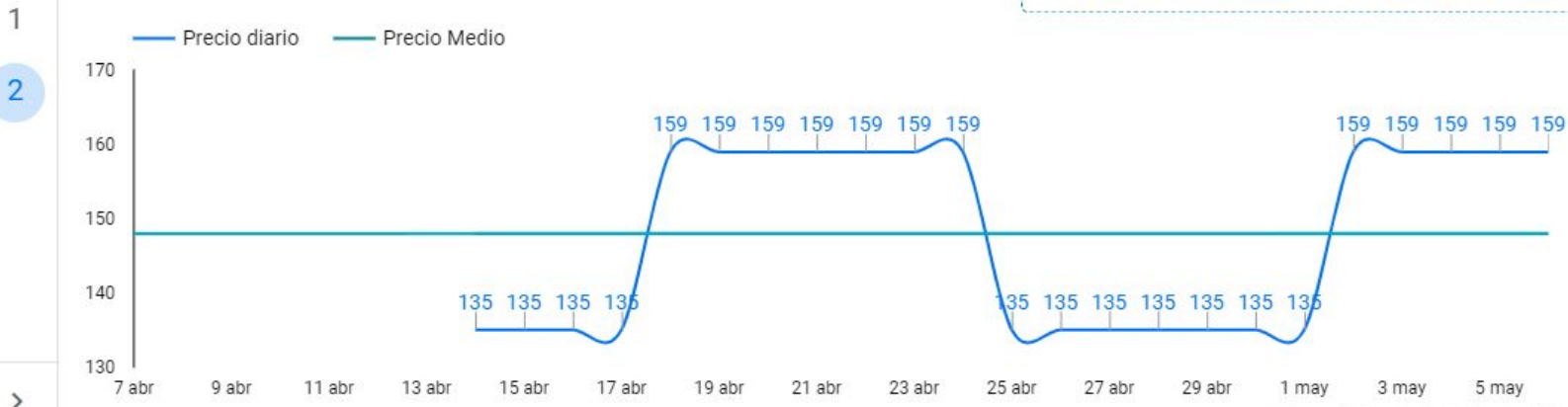
Data Studio



WORDPRESS

EVOLUCIÓN DEL PRECIO EN LOS ÚLTIMOS 30 DÍAS

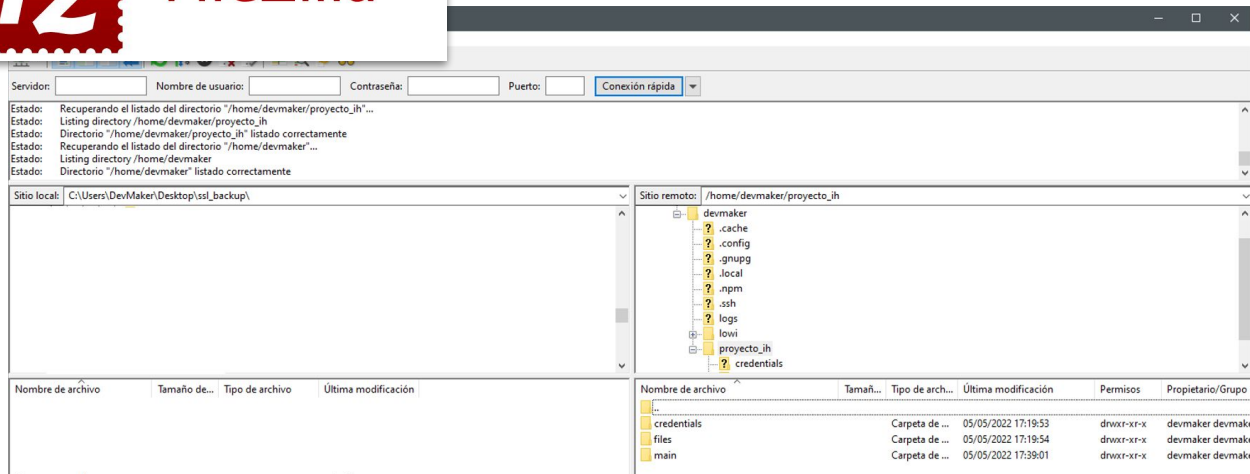
¡ATENCIÓN! HOY PRECIO POR ENCIMA DE LA MEDIA



PROCESO & OUTPUTS: 4º Subida de archivos al Servidor y Automatización de la ejecución



FileZilla



#PROYECTO IH

```
0 4 * * * cd /home/devmaker/proyecto_ih/main && python3 main_business_af.py >> main_business_af.txt
```

¡MUCHAS GRACIAS!

