



D A T A P R O J E C T 0 2

ÍNDICE DE CONTENIDO

Grupo WAKE

1

Presentación

2

Solución

3

Planificación
y costes

4

MVP

NUESTRO EQUIPO



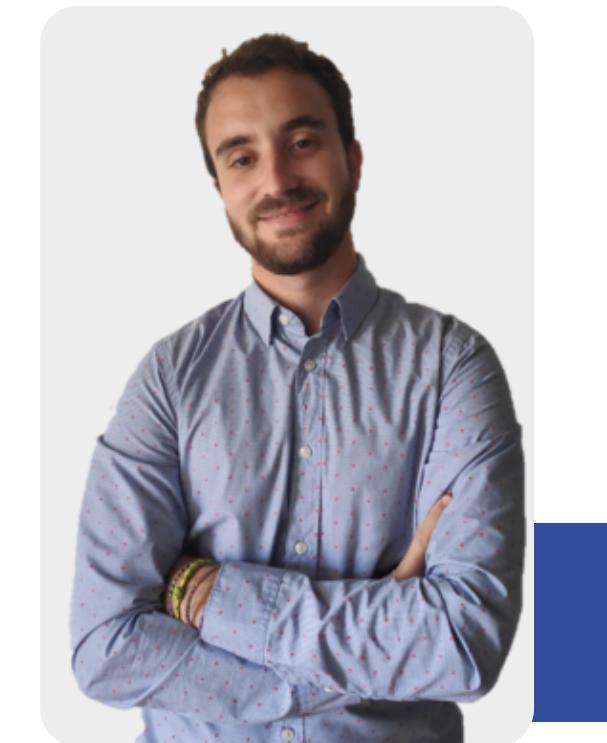
Vicent Asensio



Sofia Zander



Javier Briones



Jorge Camañez



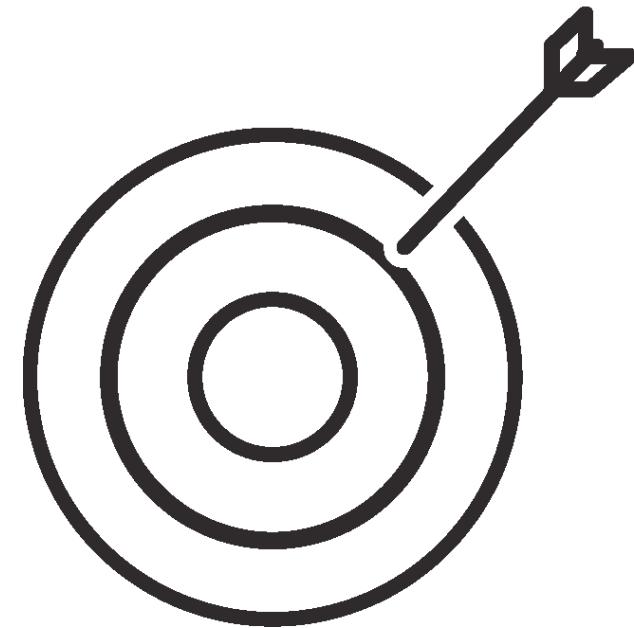
Vicente Gil

REQUERIMIENTOS



Nuestro cliente necesita responder en streaming a los users de Twitter que demandan viviendas y además ofrecer su cartera de pisos de la manera más personalizada y precisa según los gustos de los clientes.

OBJETIVO



Poner en **funcionamiento** una arquitectura de ingestión, procesamiento, almacenamiento y visualización en **streaming**, que cumpla con las **expectativas** del cliente y que además permita una **escalabilidad** a futuro.

¿POR QUÉ LA NUBE?



Coste

Optimización del gasto al establecer un pago por uso de los recursos.



Eficiencia

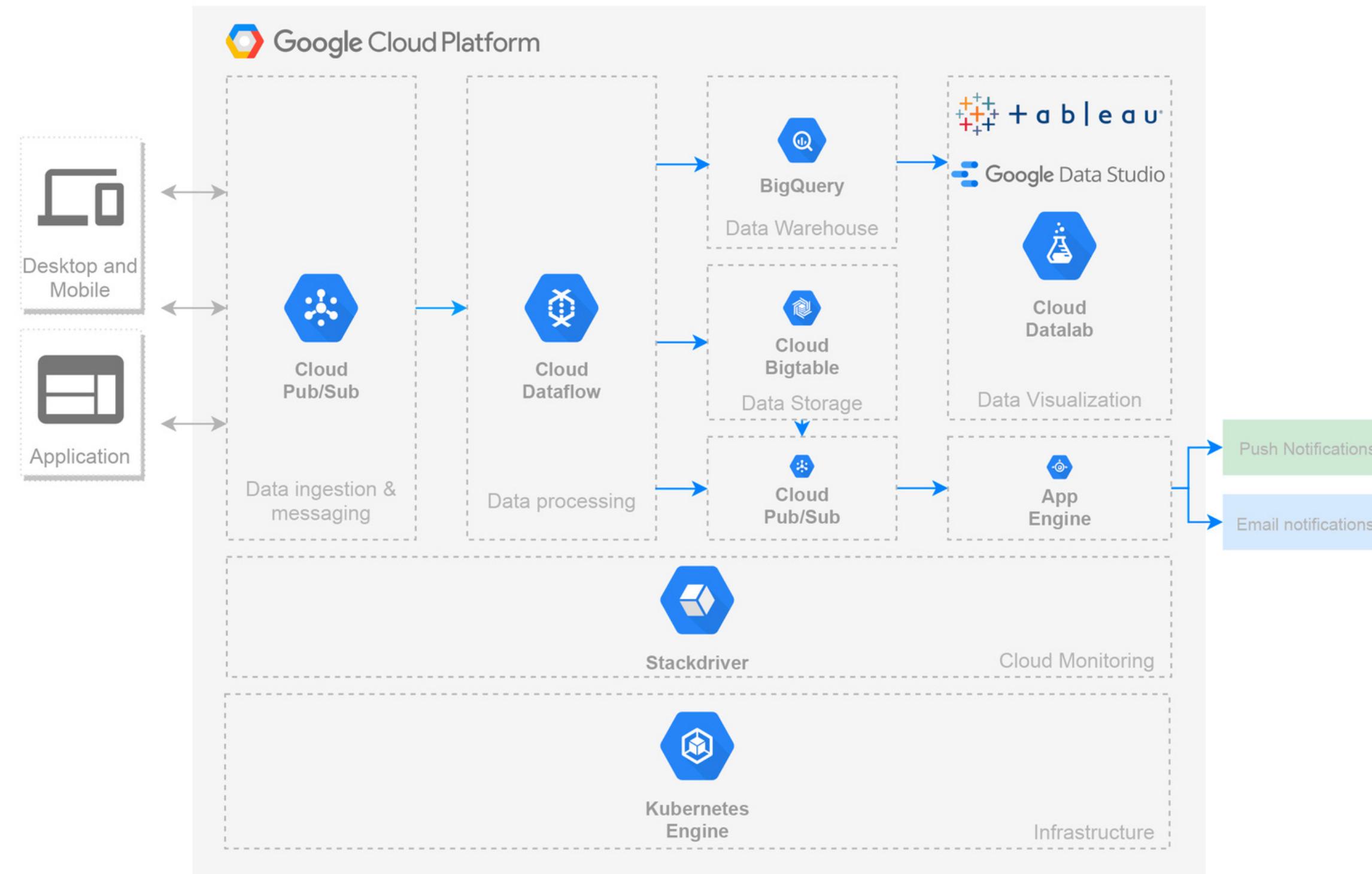
Permite focalizarse en tu aplicación sin importar la gestión de la infraestructura.



Infraestructura

Almacenamiento, rendimiento y autoescalado según las necesidades.

ARQUITECTURA



COMPONENTES



Cloud
Pub/Sub

Servicio gestionado de mensajería real-time.

Baja latencia de ingestión y envío.

Disponibilidad global.

Escalabilidad bajo demanda hasta 1M de mensajes por segundo.



Cloud
Dataflow

Servicio de procesado de dato gestionado.

Soporta trabajos Batch & Streaming.

Soporte de Event time & Processing time.

Tolerante a fallas.



Cloud
Bigtable

Base de datos NoSQL.

Altamente escalable con baja latencia.

Uso adecuado para apps que tienen una ingesta de datos frecuente.

Conexión fácil con ecosistema Apache



BigQuery

SQL contra PB de datos.

Proporciona ingesta de transmisión a conjuntos de datos ilimitados.

Almacenamiento duradero rentable para agregaciones históricas.

Almacenaje completamente administrado.



Kubernetes
Engine

Autoescalado de pods y clústeres.

Seguridad antes cargas de trabajo cambiantes.

Reducción de tiempos de respuesta.

Servicio completamente administrado.



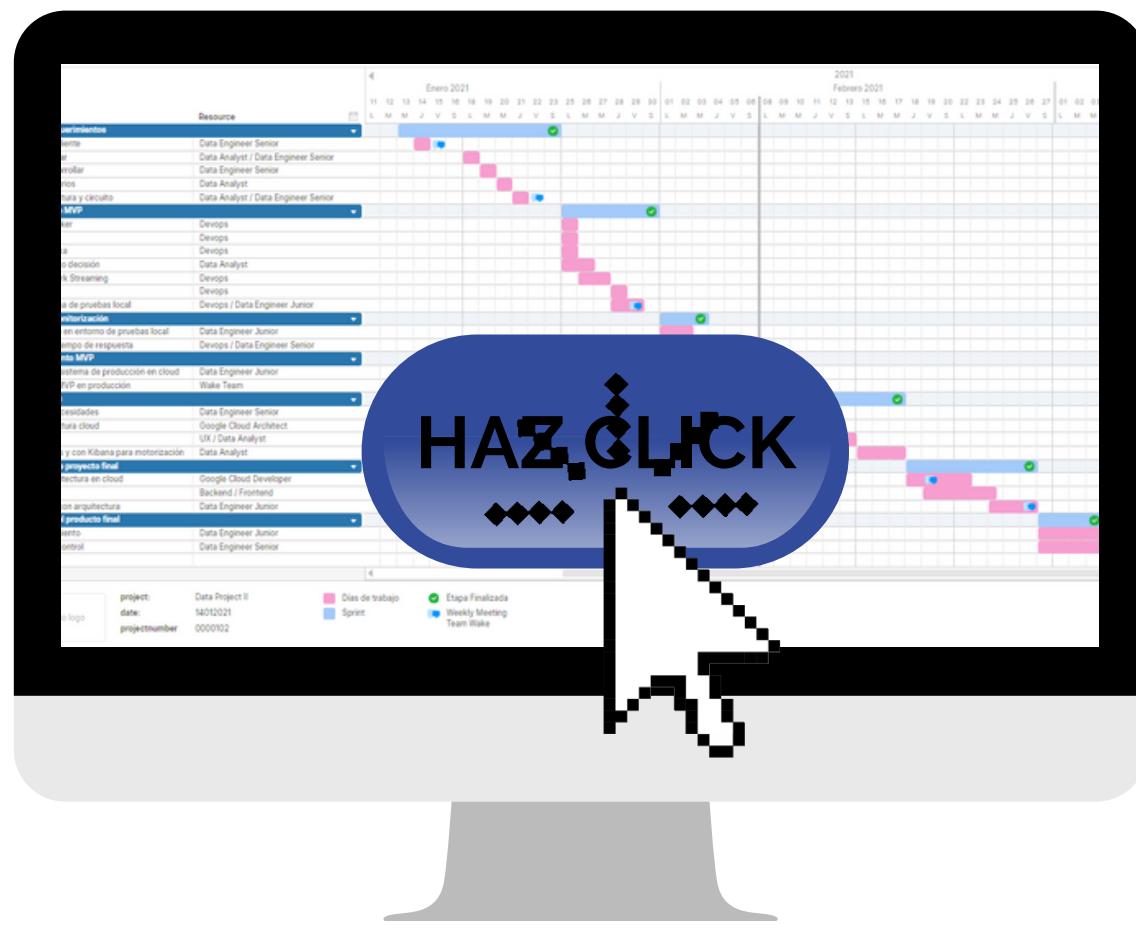
Cloud
Datalab

Exploración, visualización y transformación de datos de forma sencilla.

Utilización de lenguajes familiares como Python y/o SQL.

Notebooks de ML.

PLANIFICACIÓN DIAGRAMA DE GANTT



DATA TEAM



Senior Data Engineer & Cloud Architect

Oscar P.

8 años experiencia

- Developing Agile Projects as Product Owner and Scrum Master
- Big Data Migrations
- Architecting and Integrating Machine Learning Models
- Cloud Big data ETL migrations
- Cloud Architecting and Cloud based products insights based on analysis



Junior Data Engineer

Joaquin Montesinos

2 años experiencia

- Automating Business Intelligence Workflows
- ETL processes
- Scraping
- Database management and optimization
- Data analysis tools (Python, Bash, SQL)



Data Analyst

Ainoha Martin Mayordomo

7 años experiencia

- Creating the loans ETL in order to integrate the upstream data
- Data profiling tasks
- Data quality tasks
- Working on support teams analysing data in order to resolve production issues

DEVELOP TEAM



DevOps & Data Protection Officer

Carlos Sancho
+ 5 años de experiencia

- Kubernetes
- Cloud (Azure, AWS, Google Cloud,...)
- DevOps (Azure DevOps pipelines, Jenkins, Docker,...)



Full Stack Backend

Ávaro Ortega
8 años experiencia

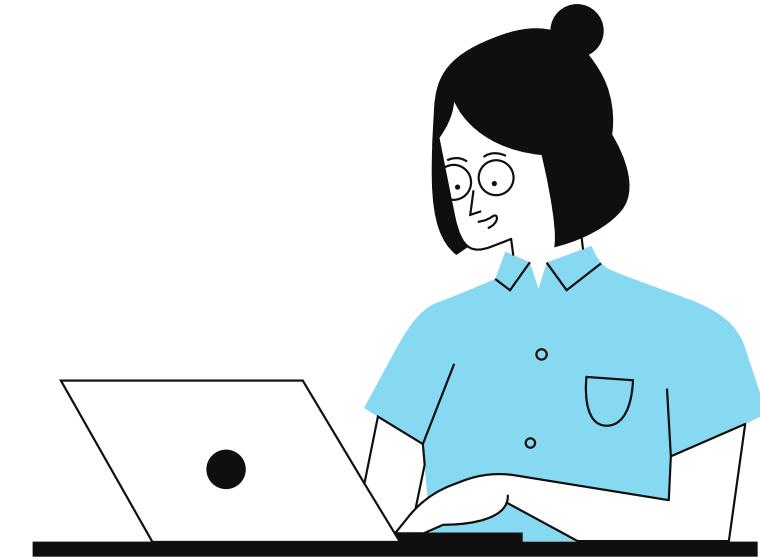
- Entornos desarrollo: Eclipse, Visual Studio, Xamarin, Mono, SQL Server Management Studio, Consola.
- Programación: C/C++, C#, Java, .Net...
- Tecnologías Web: ExtJS, JSF, Knockout, HTML, XML, Servlets, JSP, Json



Senior User Experience

Ana Ascorbe
8 años experiencia

COSTES DEL MVP

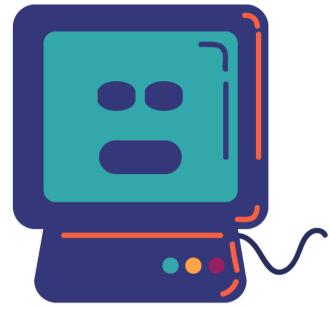


ETAPAS	D. engineer senior	D. engineer junior	Data analyst	Devops	Scrum master	
Etapa 1. Toma de requerimientos	16		24		2	
Etapa 2. Desarrollo MVP		16	10	28	2	
Etapa 3. Test y monitorización	16	8		16	2	
Etapa 4. Lanzamiento MVP	8	8			2	
TOTAL HORAS	40	32	34	44	8	158 Horas
TOTAL €	680	416	256	442	112	1906 Euros

COSTES DEL PROYECTO



ETAPAS	D.E senior	D.E. junior	Data analyst	Devops	C.A.	C.D.	Full stack dev.	UX	SC	h/etapa
Etapa 1. Toma de requerimientos	16		24						2	42
Etapa 2. Desarrollo MVP		16	10	28					2	56
Etapa 3. Test y monitorización	16	8		16					2	42
Etapa 4. Lanzamiento MVP	8	8							2	18
Etapa 5. Evolución	16		40		16				2	74
Etapa 6. Desarrollo proyecto final		16		20		40	160	80	2	318
Etapa 7. Soporte y mantenimiento	8	40							2	50
TOTAL HORAS	64	88	74	64	16	40	160	80	14	600 Horas
TOTAL €	1088	1144	704	962	272	560	1440	800	224	7194 Euros

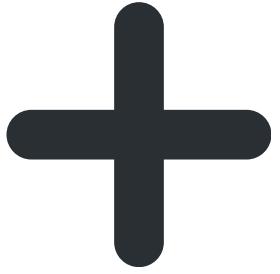
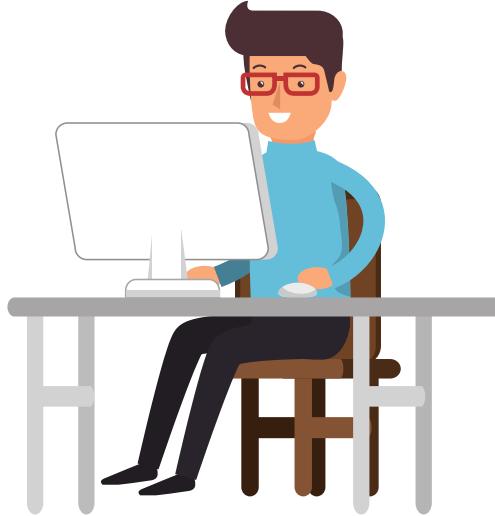


COSTES DEL PROYECTO

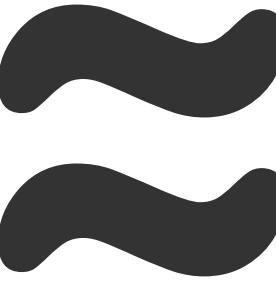
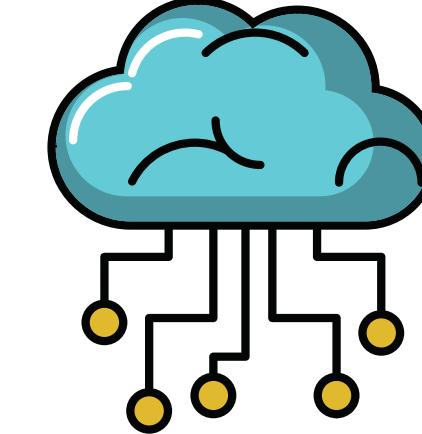
COMPONENTES	CARACTERISTICAS	REGIÓN	COSTE MENSUAL
Google BigQuery	1000 Gb Storage Pricing, Stream inserts 10 mb, query pricing 10 TB	Londres	80
Google PubSub	1000 Gb Volume Message	Londres	660
Google Dataflow	Streaming job, 720 houts/month, 2 workers nodes/job, n1-Standard8	Londres	1450
Google App Engine	4vCPUs, 32GB, 1000 persistent disc	Londres	442
Google Kubernetes Engine	2 nodes cluster, Machine class, general purpose, n-1 Standard8, 24/7 running	Londres	195
Google BigTable	1 node 1TB/node	Londres	470
TOTAL			3297

COSTES DEL PROYECTO

7.194€



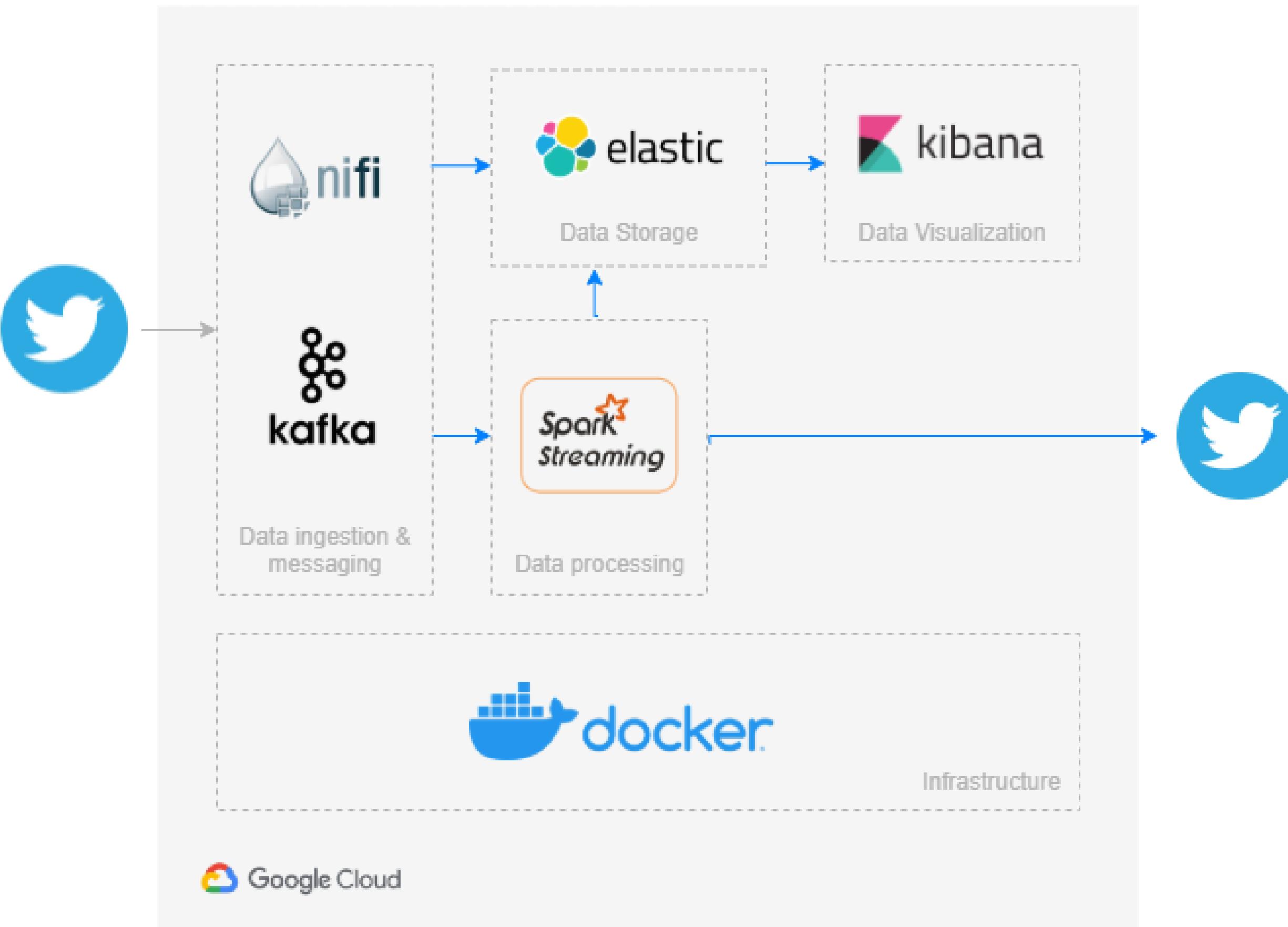
3.297€



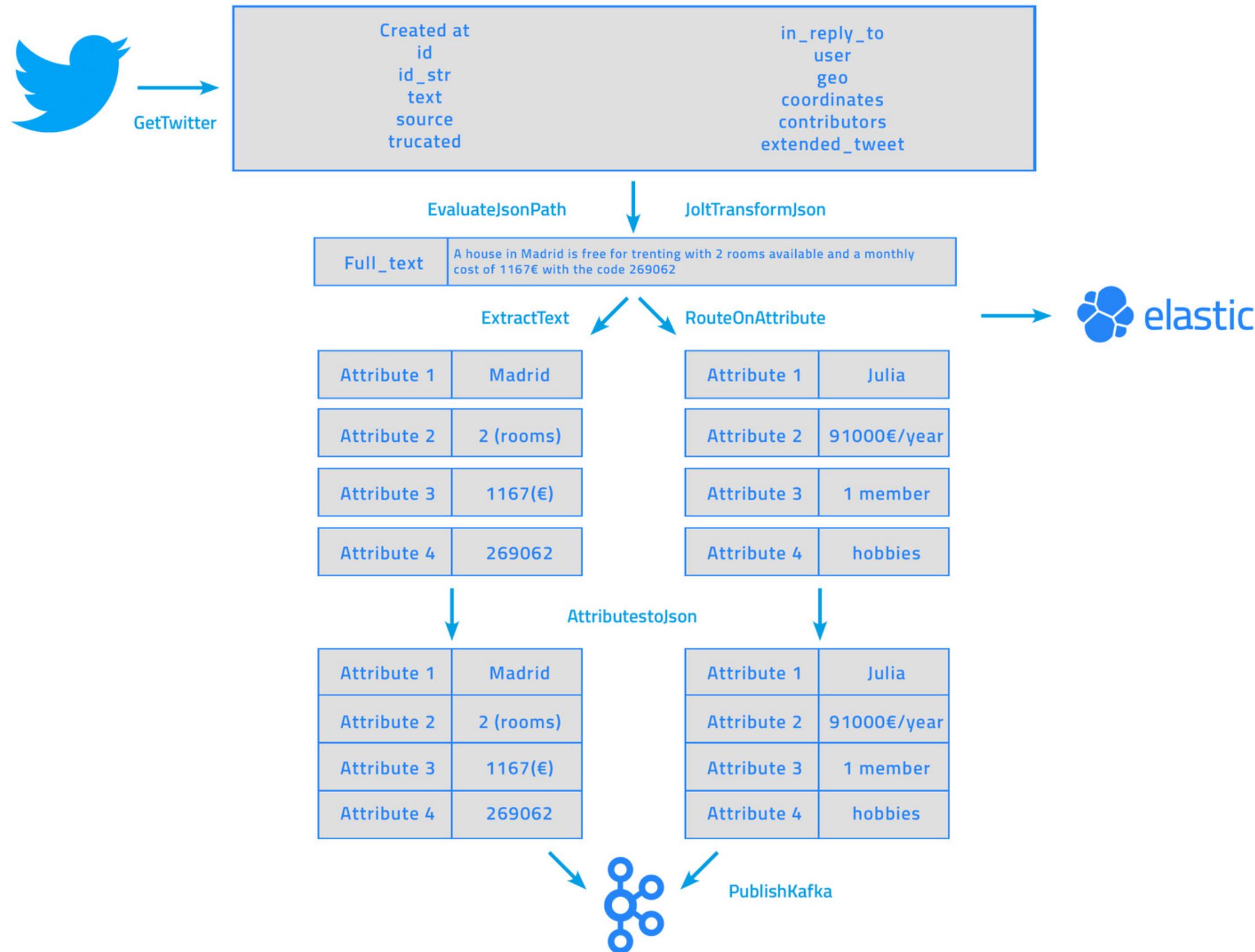
10.491€



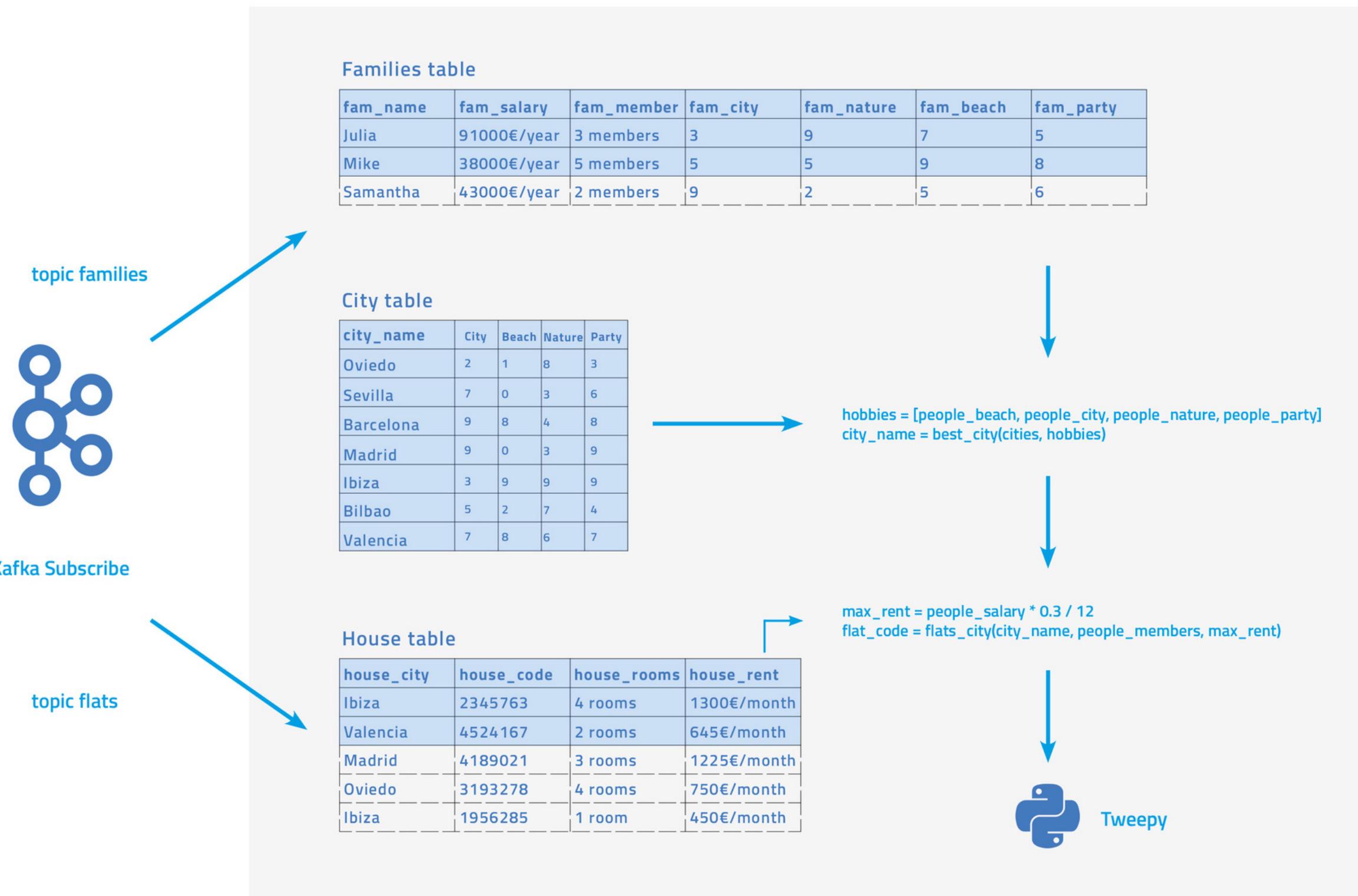
ARQUITECTURA MVP



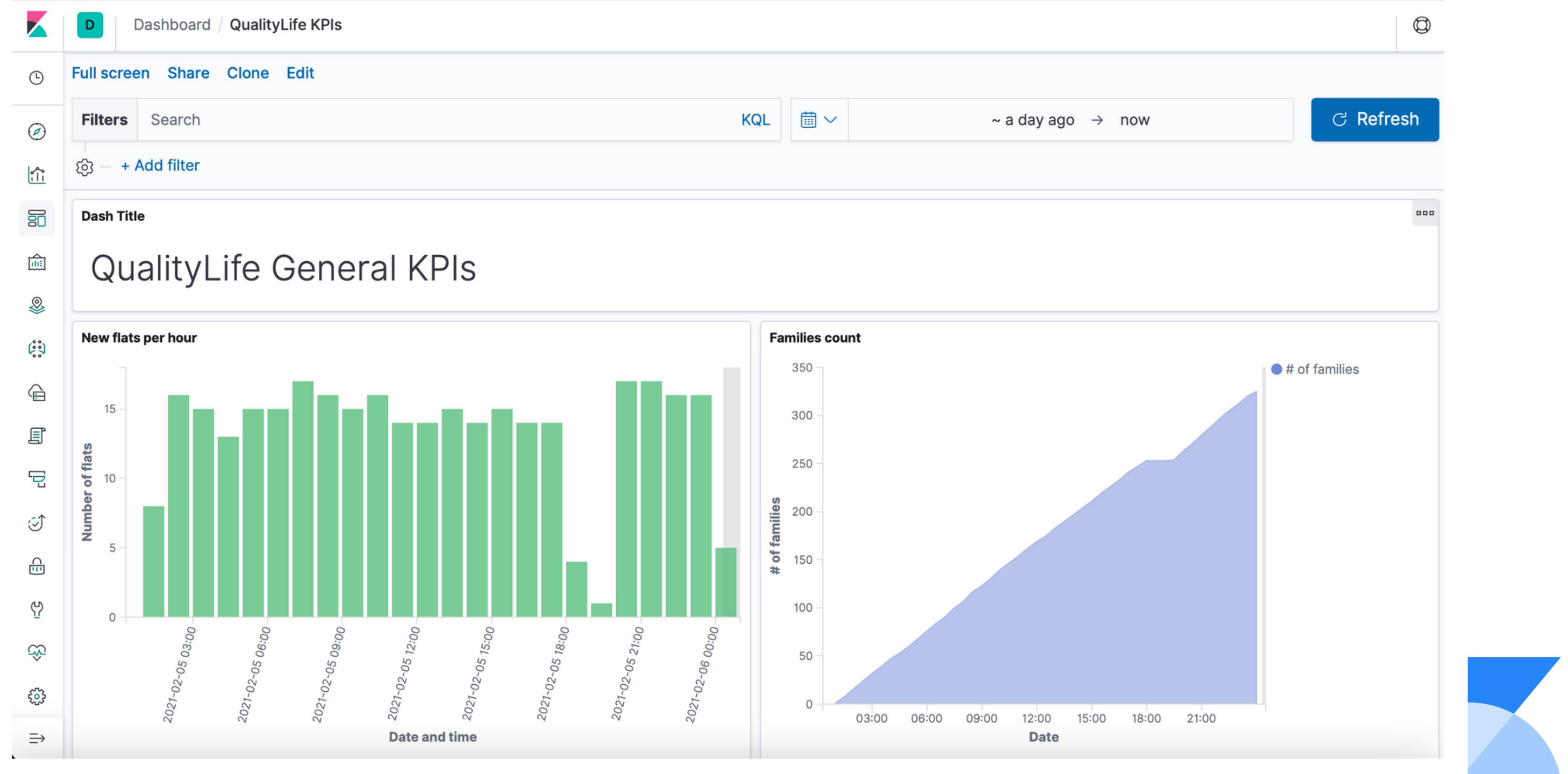
DATA INGESTION



DATA PROCESSING



DATA VISUALIZATION



RESULTADO

 **dlpexerciseprocess** @dlpexercisepro1 · 9m ...
My name is Courtney Rodriguez, my salary is 108925€ yearly and I am 28 years old. My family are 3 members. These are my hobbies: Beach(2), City(0), Nature(4), Party(0) #madaudem

3 ↗ ↘ ❤ ⬆

 **Wake Team** @javibbriones · 9m ...
Hi Courtney Rodriguez, your future house is waiting for you in Oviedo with code 845040. Please, contact [#QualityLife](#) for further details! [#WakeTeam](#) #madaudem

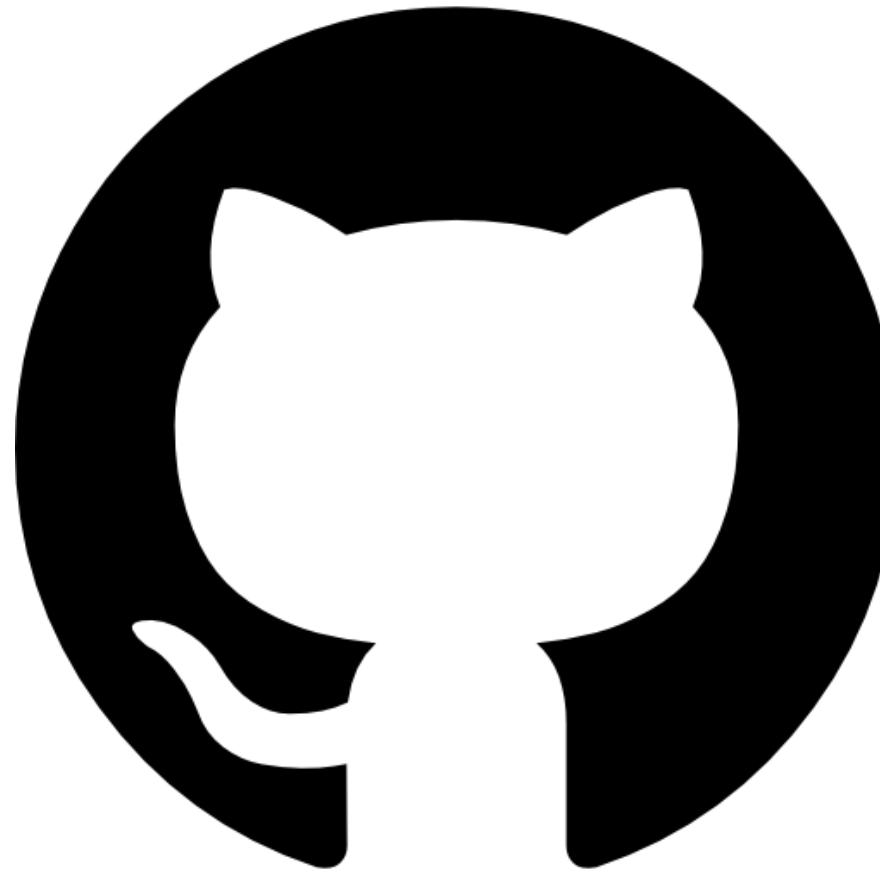


⋯ ↗ ↘ ❤ ⬆



**THANK
YOU**

ENLACES



HAZ CLICK



HAZ CLICK