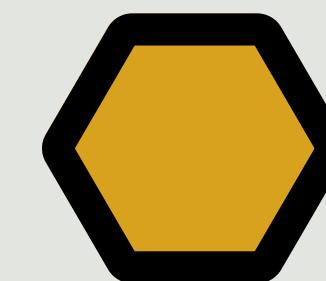


PROYECTO GRUPAL

Yelp & Google Maps
Reseñas y Recomendaciones

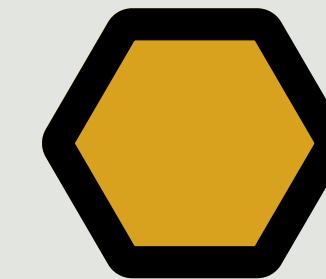


CONTEXTO



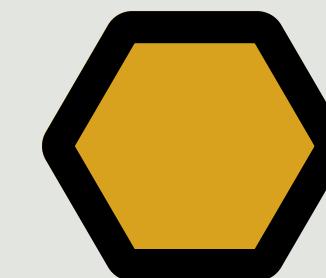
TAREA

- Análisis detallado del mercado estadounidense



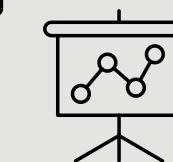
FUENTE DE DATOS

- Yelp! 
- Google Maps 

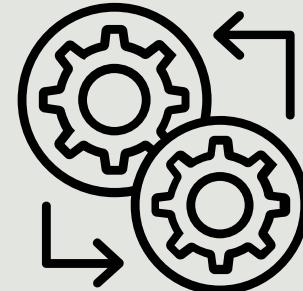
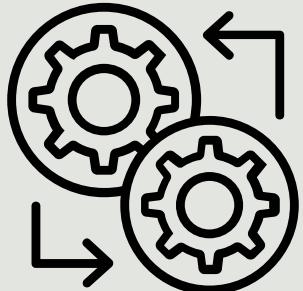


NUESTROS SERVICIOS

- Tecnologías de Big Data
- Herramientas de visualización
- Algoritmos de Machine Learning



NUESTRO EQUIPO



Dany Cruzado

Data Engineer



Luz Oblea

Data Engineer



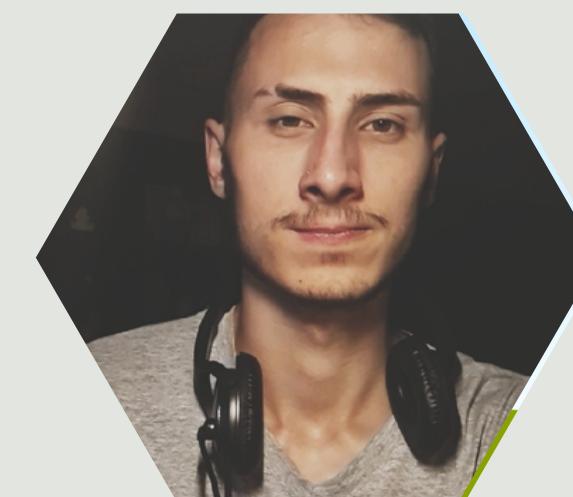
Ricardo Zegarra

Data Analyst



Tobías Oliva

Data Analyst

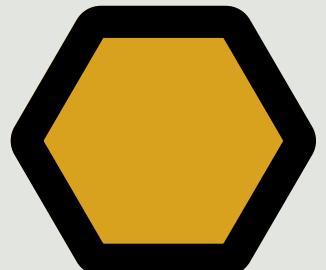


Joel Camatta

Data Scientist

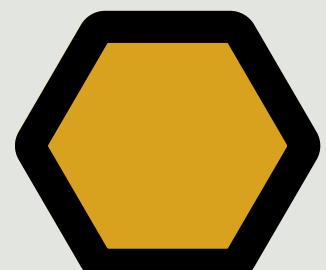
Click en la imagen para redireccionar a cada Linkedin

ENFOQUE



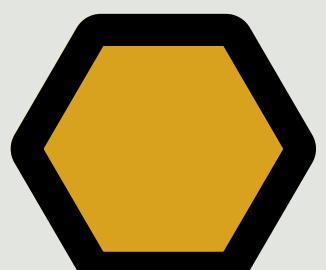
TIPOS DE SERVICIOS

Predicción del comportamiento de los distintos rubros en el mercado.



LOCACIONES

Determinar las áreas convenientes para ubicar nuevas unidades de negocio.



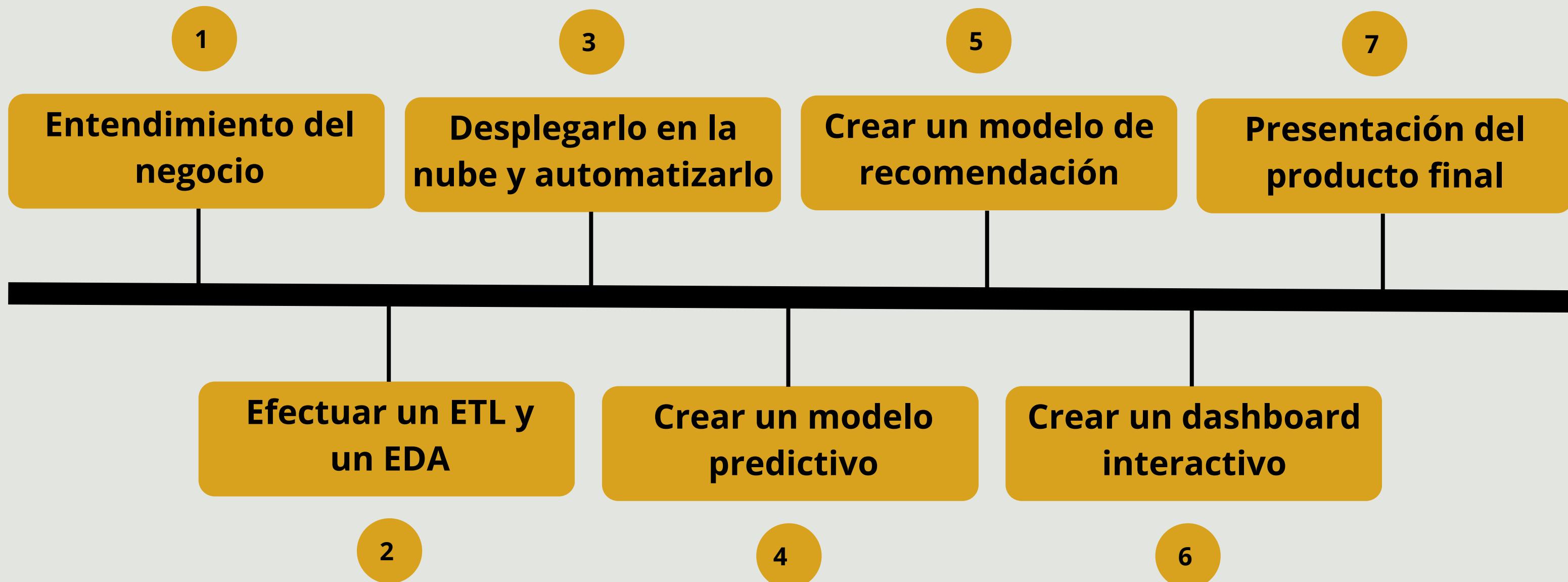
EXPERIENCIA DEL USUARIO

Crear un sistema de recomendación para los usuarios.

-
-
-
-
-
-
-
-
-
-



OBJETIVOS



CRONOGRAMA

- 1 PROPUESTA DE PROYECTO**
- 2 DATA ENGINEERING**
- 3 DATA ANALYTICS -
MACHINE LEARNING**
- 4 PRESENTACIÓN**



DIAGRAMA DE GANT

PRIMERA SEMANA

Tareas

Entendimiento de la situación actual

Planteamiento de objetivos

Definición del alcance del proyecto

Planteamiento de KPI's

Creación de repositorio de Github

Planteamiento de solución propuesta

Lunes

Martes

Miércoles

Jueves

Viernes



Critico

No Critico

Entrega

DIAGRAMA DE GANT

SEGUNDA SEMANA

Tareas

Diseño del modelo ER

Creación de Pipelines y
Data Warehouse

Automatización

Validación de datos

Documentación

Lunes

Martes

Miércoles

Jueves

Viernes

Critico

No Critico

Entrega



DIAGRAMA DE GANT

TERCERA SEMANA



REPOSITORIO GITHUB

tobiasolivapinto / **ProyectoGrupal-SoyHenry** Public

Pin Unwatch 1 Fork 0 Star 0

Code Issues Pull requests Actions Projects Wiki Security Insights Settings

main 1 branch 0 tags Go to file Add file Code

jcamatta Update ReadME.md 31c987d 42 minutes ago 16 commits

Data Analytics	DA	3 hours ago
Data-Engineering	load_data	1 hour ago
General	DA	11 hours ago
MLOps	README.md for MLOps	1 hour ago
.DS_Store	DA	11 hours ago
ReadME.md	Update ReadME.md	42 minutes ago

ReadME.md

Proyecto Grupal N°1 - Henry



Yelp & Google Maps - Reseñas y Recomendaciones



About

Proyecto grupal de la carrera Data Science en Soy Henry - Camatta, Oblea, Cruzado, Zegarra, Oliva

Readme Activity 0 stars 1 watching 0 forks

Releases

No releases published Create a new release

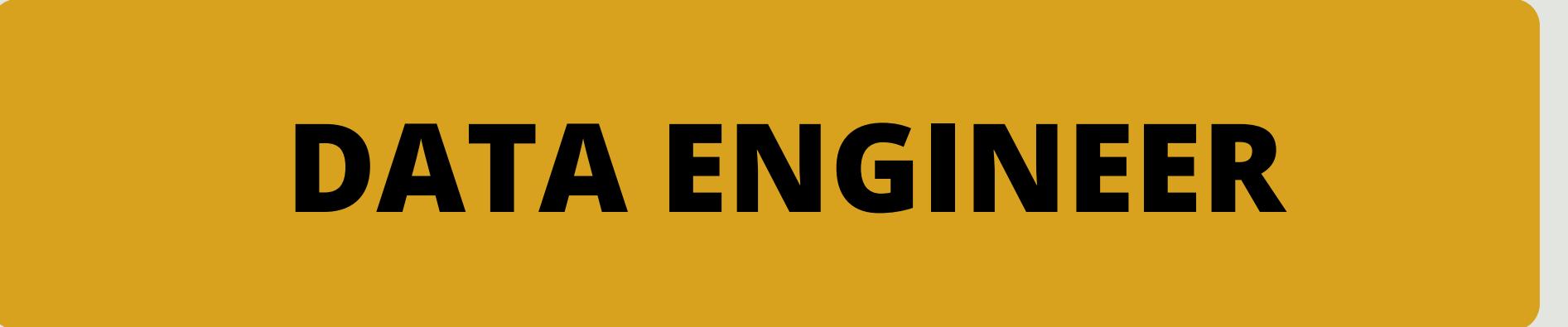
Packages

No packages published Publish your first package

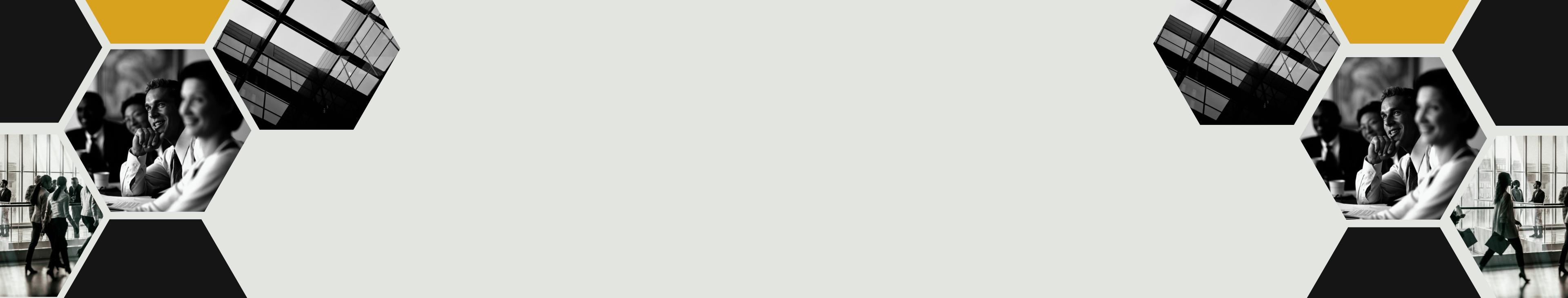
Contributors 3

 tobiasolivapinto Tobias Oliva Pinto

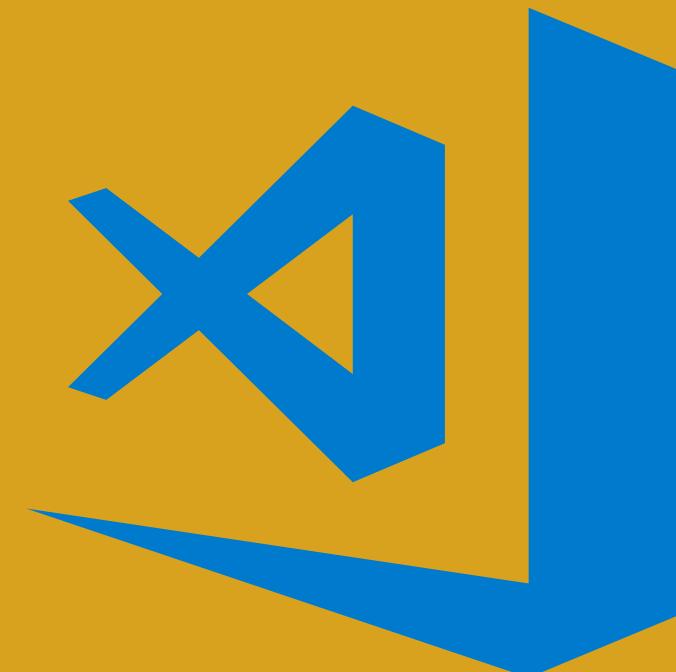
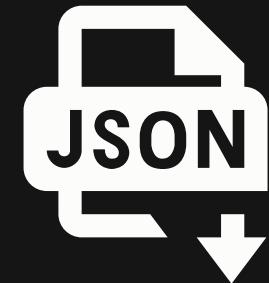
 jcamatta realcamatta



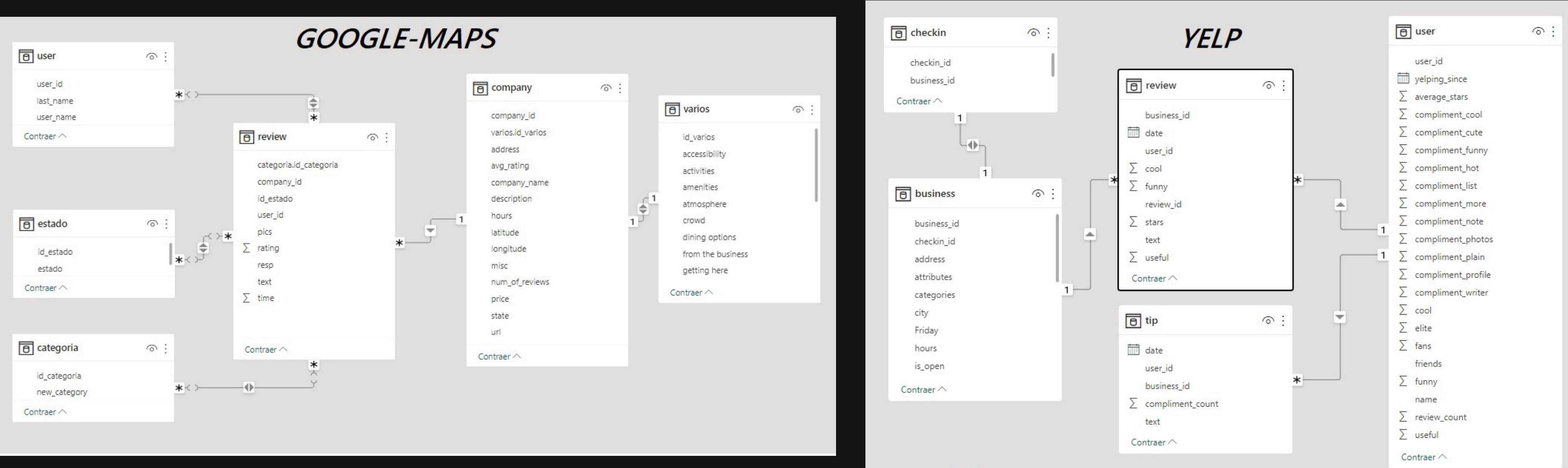
DATA ENGINEER



ETL



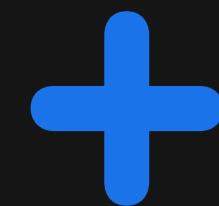
MODELO ER - COPO DE NIEVE



DATA LAKE



Google Cloud Storage

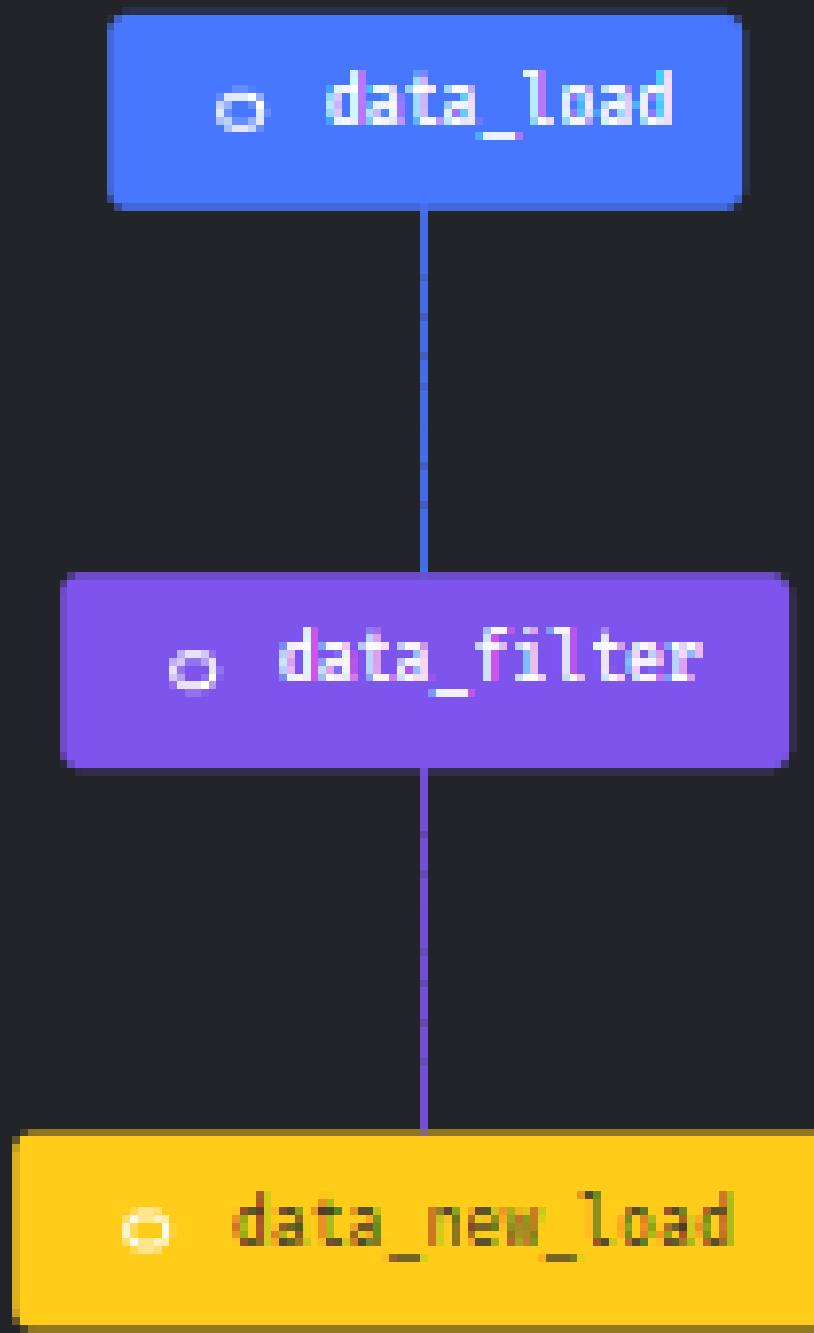


Google
Big Query

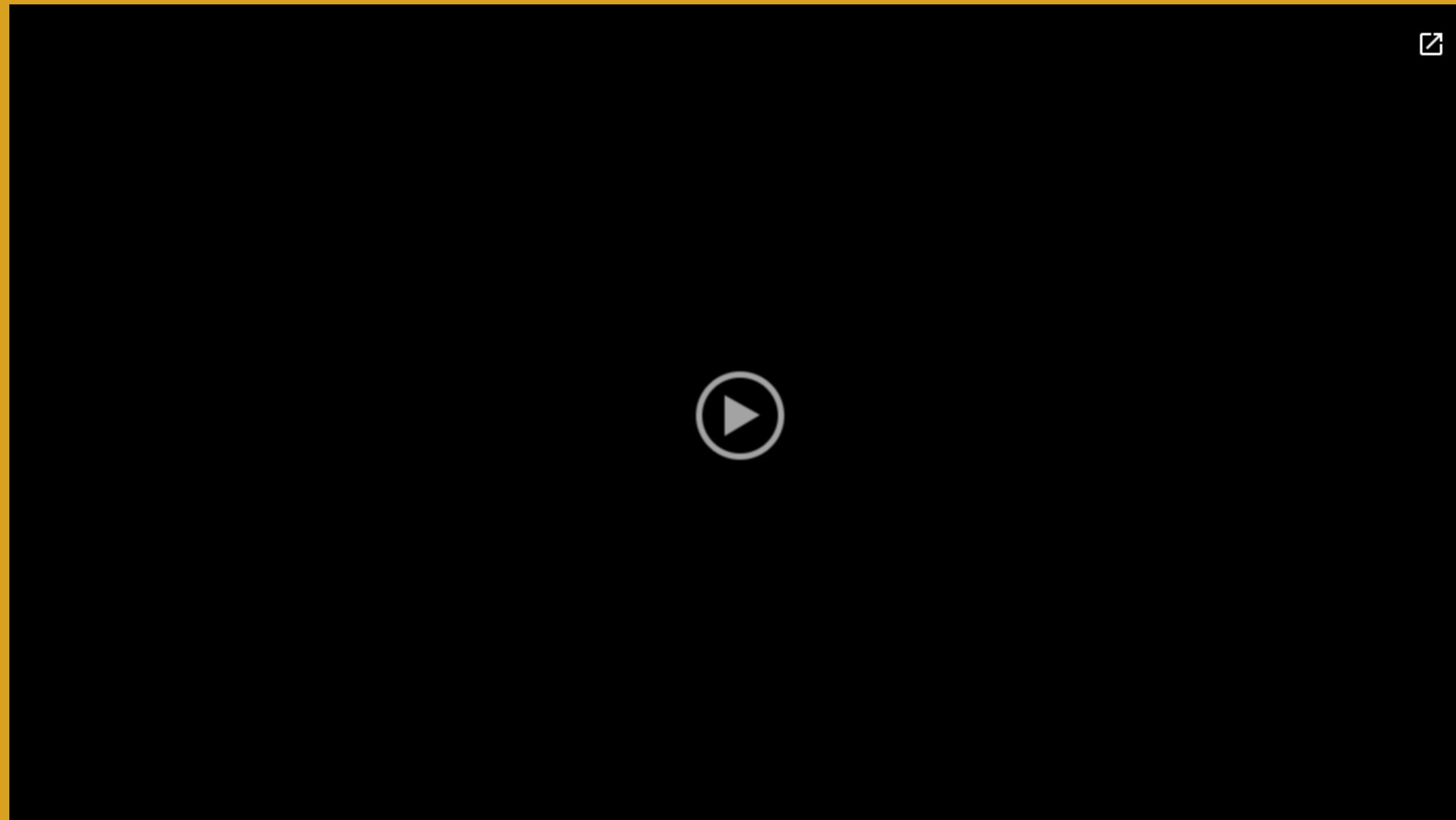
PIPELINE



Mage.ai

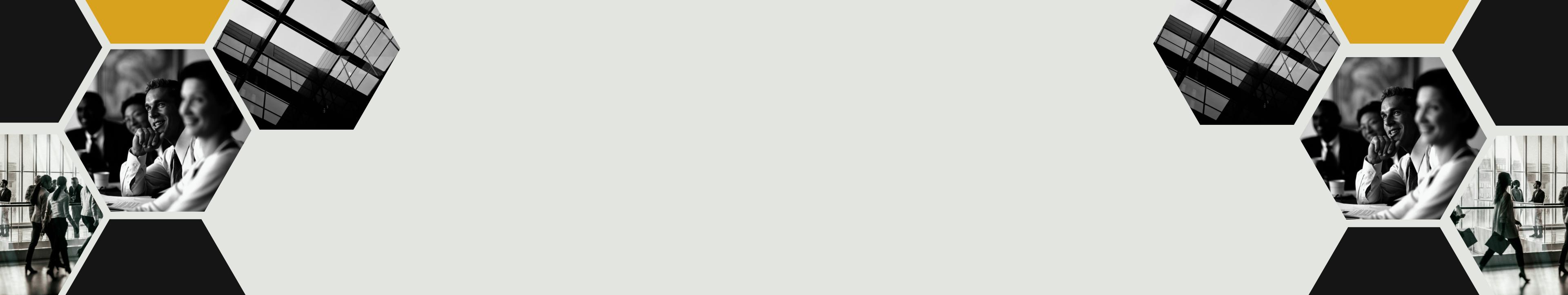


DEMO PIPELINE





DATA ANALYTICS

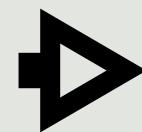


DATA ANALYTICS

TAREAS

1

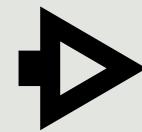
ANALISIS EXPLORATORIO DE DATOS (EDA)



BIG QUERY

2

ELABORACIÓN DEL DASHBOARD

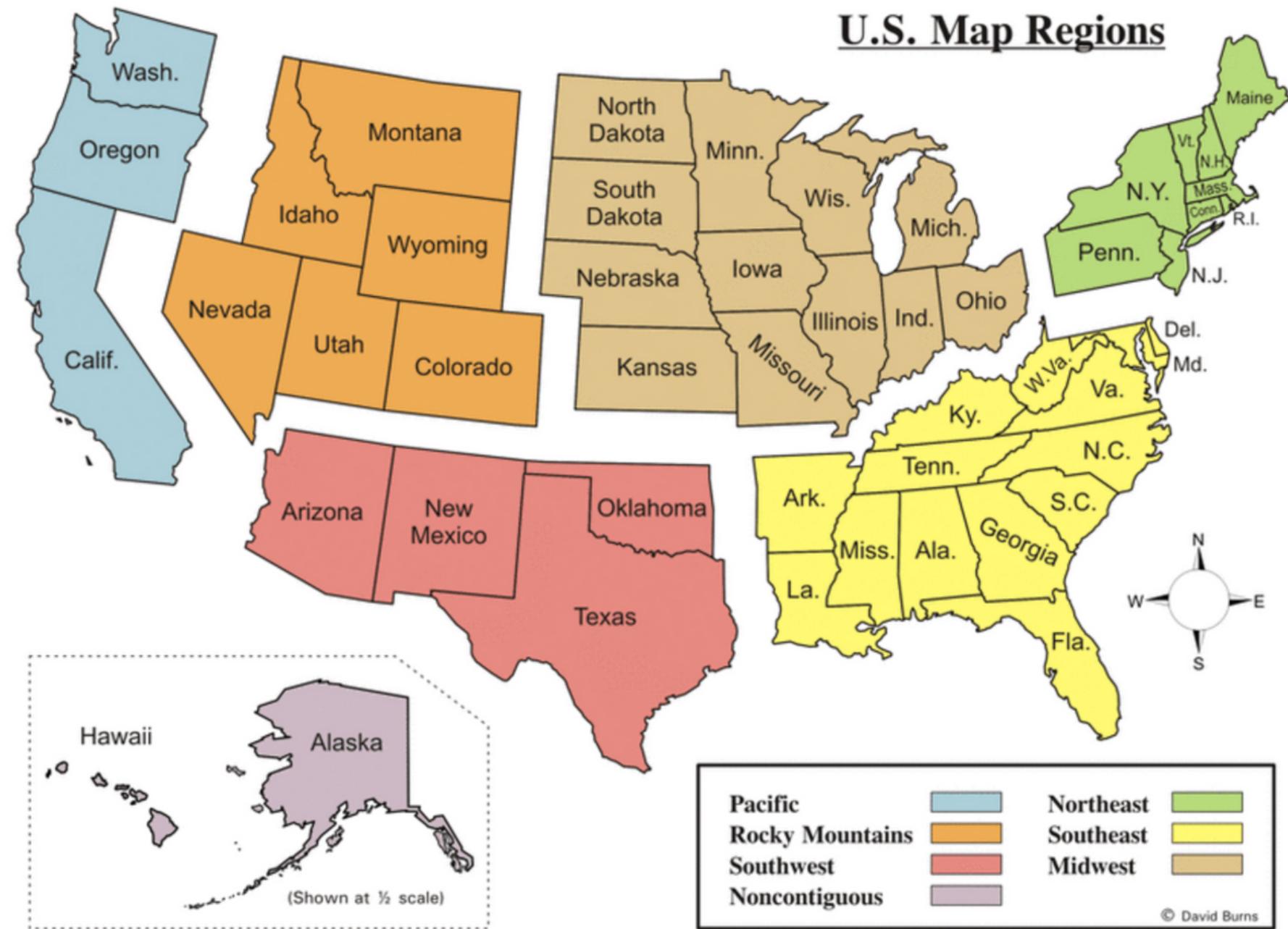


LOOKER STUDIO



CONTEXTO

MAPA DE E.E.U.U - REGIONES



Pacifico

Rocosas

Suroeste

No Contiguo

Noreste

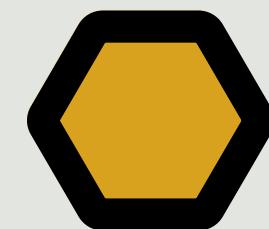
Sureste

Medioeste



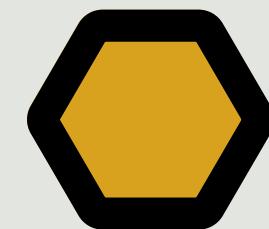
CONTEXTO

TIPOS DE ESTADOS



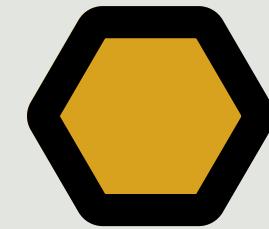
E.cosmopolistas

- Reciben un gran número de inmigrantes.



E. turísticos

- Reciben un gran número de turistas al año.

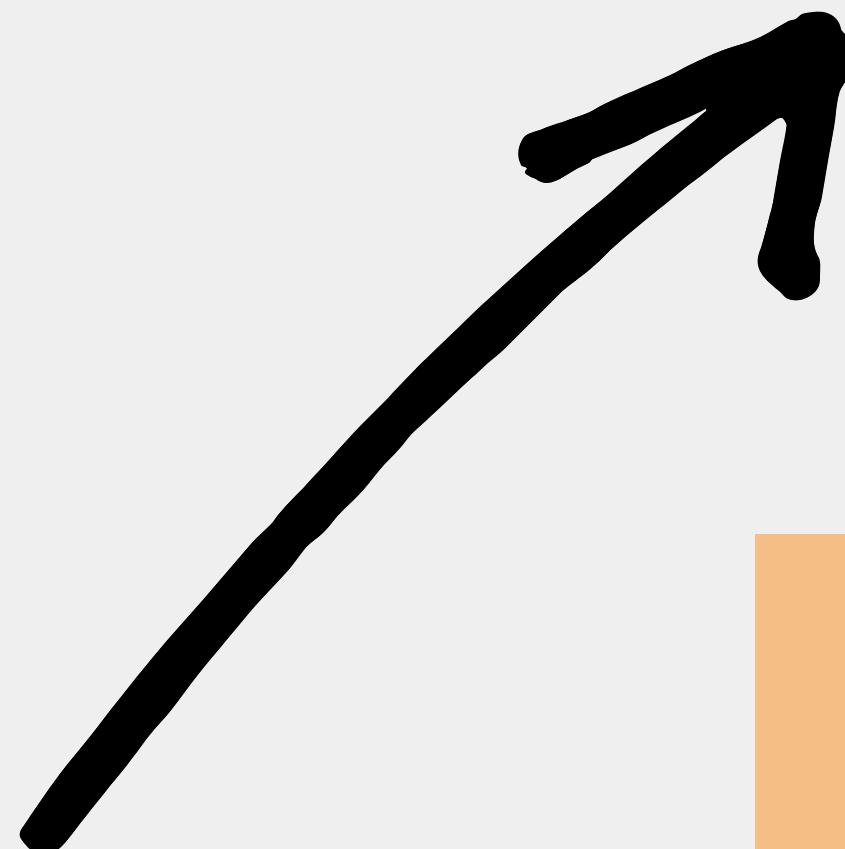


E. tradicionales

- Hogar de los "country people"

**MACHINE
LEARNING**

ESQUEMA DE LA PRESENTACION



MUESTRA DEL
PRODUCTO

DESCRIPCION DEL
PRODUCTO FINAL

RECAPITULACION
CAMINO DEL DATO

RECORDANDO LAS PROPUESTAS

**SISTEMA DE
RECOMENDACION**

**RESUMEN
INFORMATIVO**

FILTRADO COLABORATIVO

SIMILITUD ENTRE
USUARIOS

BASADO EN LAS
RESEÑAS

FILTRO POR CONTENIDO

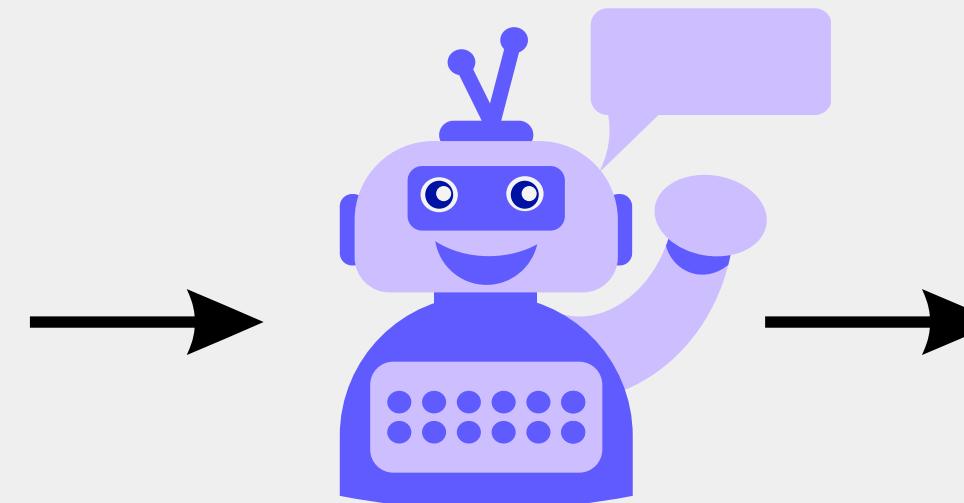
SIMILITUD ENTRE
PRODUCTO

BASADO EN LA
METADATA

EMPRESA



**MODELO DE
LENGUAJE**



RESUMEN

The ground or stays
iverse is vast, and you
s also beautiful. You a
nthing bigger than yo
t of something that ma
most of your time. Tal
e a blog post. Make a
...

STACK TECNOLOGICO FINAL



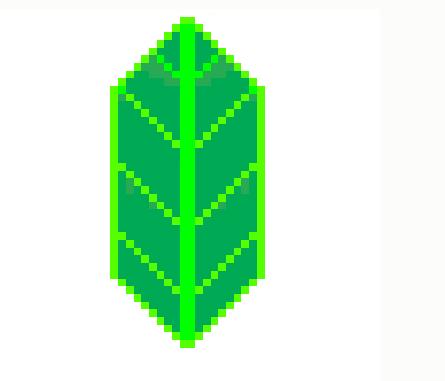
spaCy



SQLAlchemy

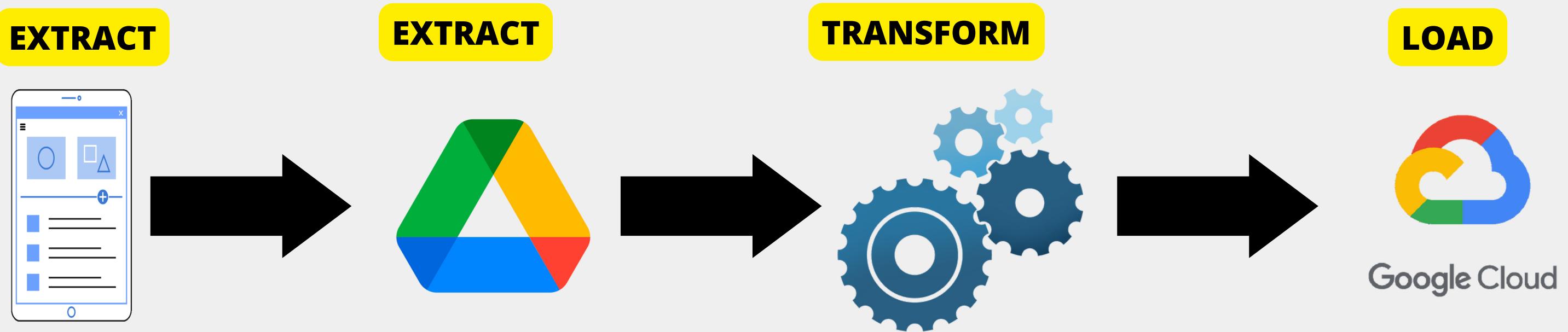
FastAPI

FOLIUM

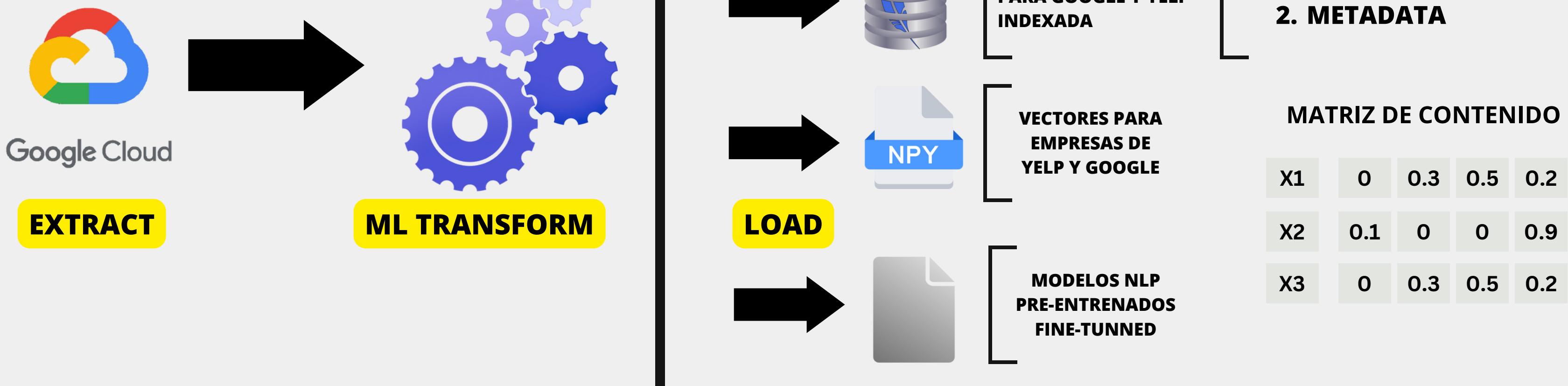


ROADMAP DE LA DATA

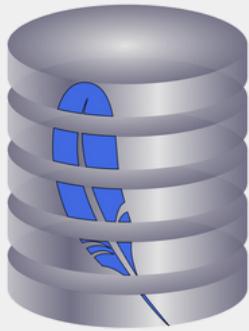
PRE-CLOUD



POST-CLOUD



ARQUITECTURA



**SQLITE DATABASE
PARA GOOGLE Y YELP
INDEXADA**



**VECTORES PARA
EMPRESAS DE
YELP Y GOOGLE**



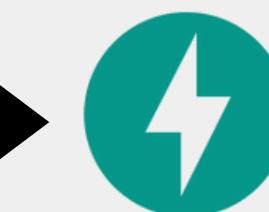
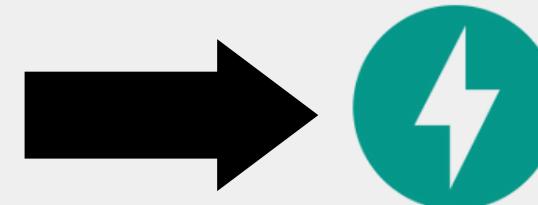
**MODELOS NLP
PRE-ENTRENADOS
FINE-TUNED**

SISTEMA DE RECOMENDACION

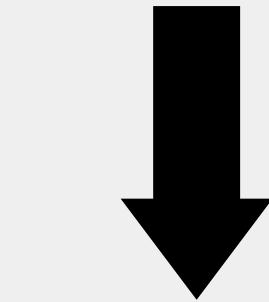
RESPUESTA A REVIEWS

TOPIC CLASSIFICATION

SUMMARIZATION



FastAPI



Streamlit

API ENDPOINTS



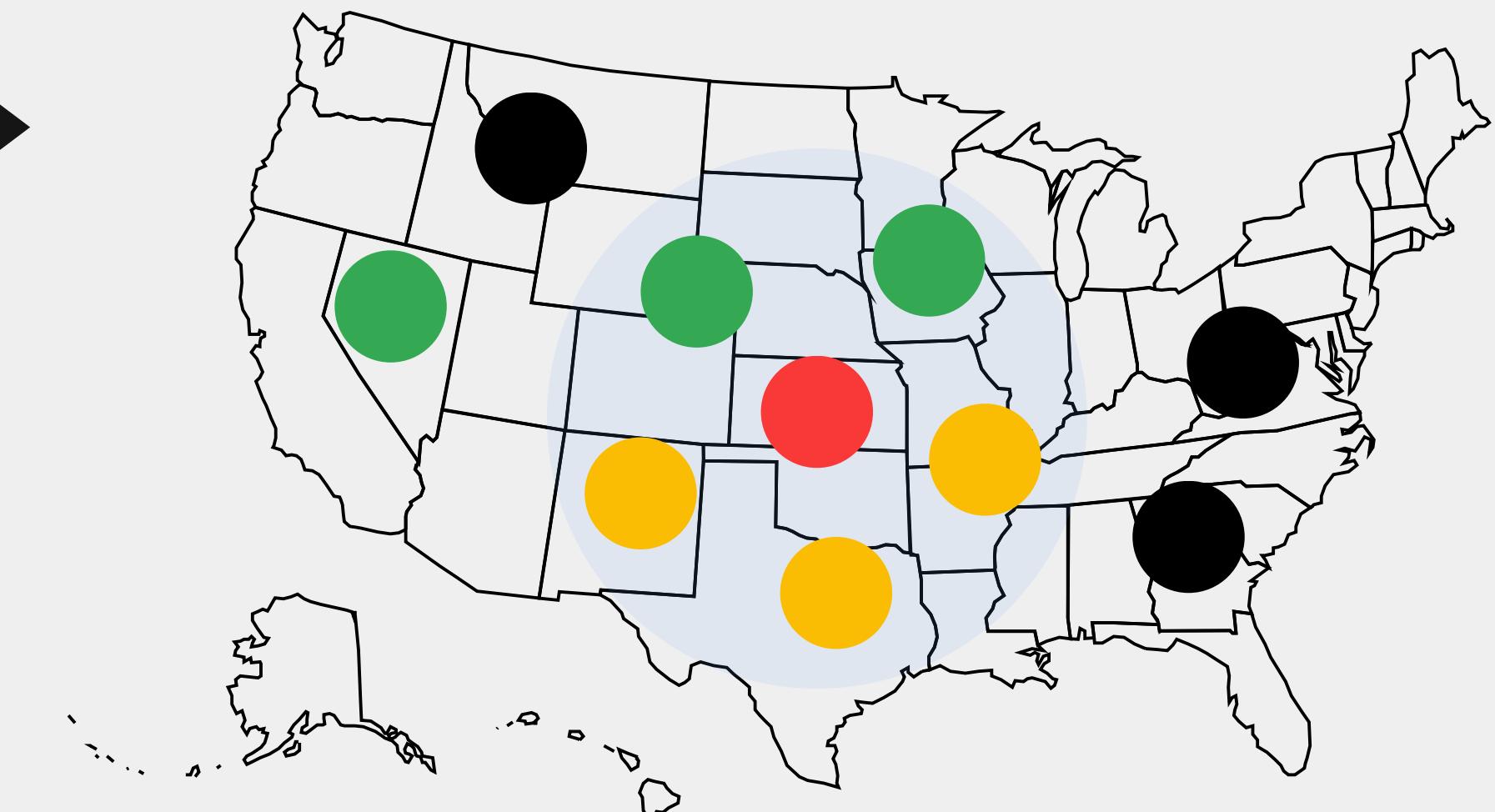
SISTEMA DE RECOMENDACION

NLP MODELOS

EXTRAS: DEVUELVE LAS REVIEWS

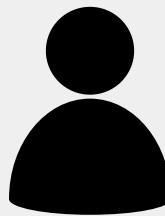
GET	/user/google/{user_id}/{user_coord}/{category}/{threshold}	Google Recomendation	▼
GET	/user/yelp/{user_id}/{user_coord}/{category}/{threshold}	Yelp Recomendation	▼
GET	/business/topic_classification/{review}	Topic Classification	▼
GET	/business/review_response/{review}	Review Response	▼
GET	/business/summarize/{review}	Summarize	▼
GET	/business/reviews/google/{company_index}	Get Company Reviews	▼
GET	/business/reviews/yelp/{company_index}	Get Company Reviews	▼

SISTEMA DE RECOMENDACION



- USUARIO
- INSTITUCIONES VISITADAS
- INSTITUCIONES NO VISITADAS
- INSTITUCIONES NO CONSIDERADAS

NOTA EXTRA



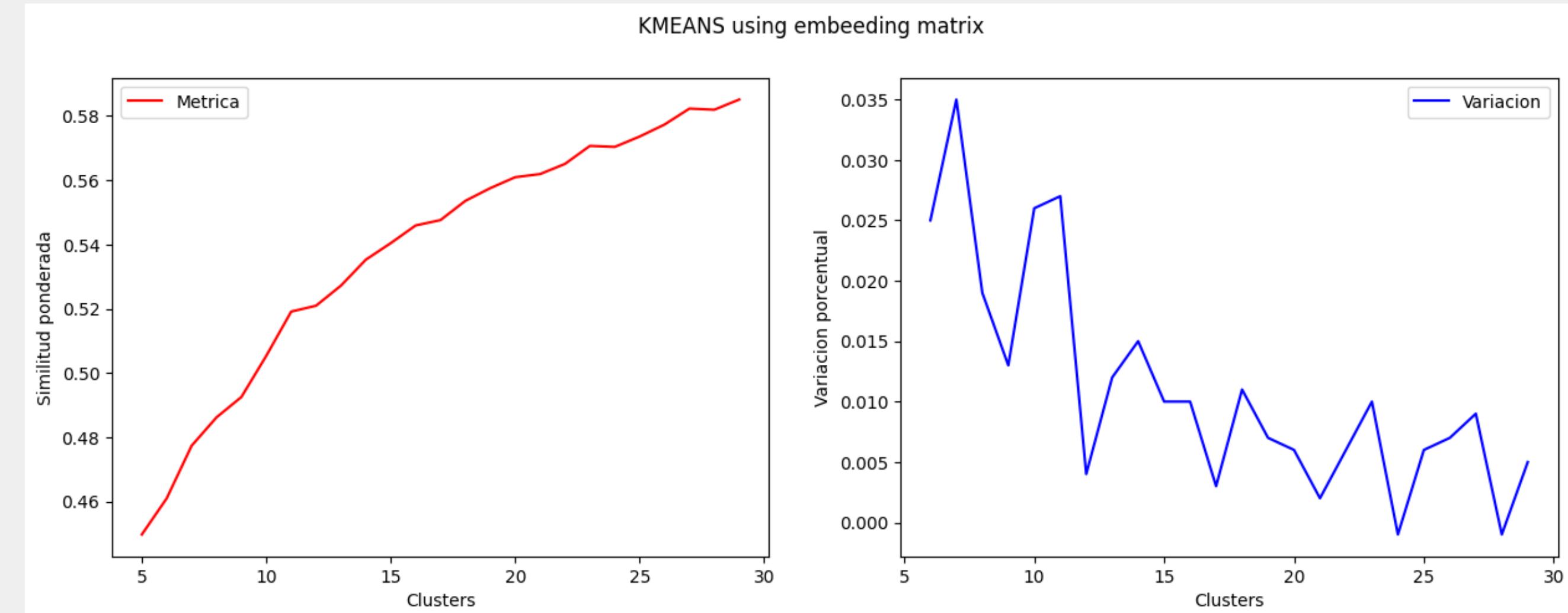
CATEGORIA

13 CATEGORIAS PARA YELP

11 CATEGORIAS PARA GOOGLE

EJEMPLO PARA GOOGLE

- healthcare services
- food and beverage establishments
- personal care services
- retail stores
- wholesale suppliers
- service establishments
- places of interest and worship
- business services
- clubs and sports
- creative professionals
- educational institutions



VAMOS A LA
APLICACION





**GRACIAS POR SU
ATENCIÓN**

