

Proyecto Grupal Final

de Data Science

HENRY

Integrantes Grupo N° 1:

Marta Inés Pedriel
Javier Báez Esqueda
Carlos Antonio Báez
Steven Yuval Barrantes Briceño
Hernán Pizarro

Data Analyst /Data Science
Data Engineer / Data Science
Data Engineer/Data Science
Data Engineer / Data Science
Data Analyst /Data Science

Henry Mentor:

Angela Aguirre



SMART
DATA

Sprint #3

Data analytics y machine learning



SMART
DATA

Índice

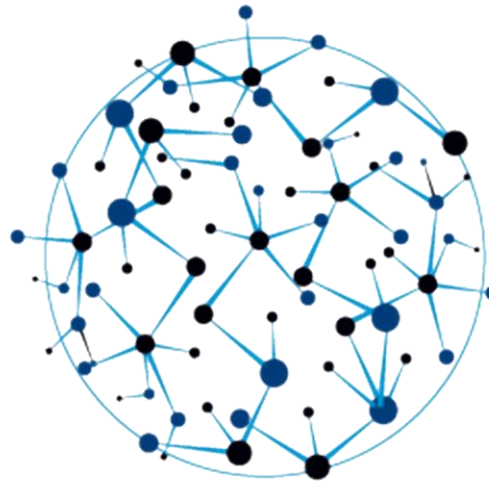
- ✓ Introducción
- ✓ Data Pipeline
- ✓ Visualización de datos
- ✓ Análisis de datos
- ✓ Insights
- ✓ Modelo de Machine Learning



Introducción

Rol del grupo

Somos la consultora **SMART DATA** y nos dedicamos al análisis de datos del mercado estadounidense.



SMART DATA

Contexto

Nos contrató la empresa **Gruma**, quiere lanzar al mercado nuevos productos y necesita saber si **el consumo de comida mexicana en restaurantes está en crecimiento en Estados Unidos** (desde año 2019 al 2022).



La población en EE.UU. de origen mexicana es de aproximadamente de 36 millones.



SMART
DATA

Plantas en América:

- Estados Unidos (26)
- México (25)
- Centroamérica (10)

Plantas en Europa:

- Reino Unido (2)
- Italia (1)
- Ucrania (1)
- Rusia (2)
- Tuvalu (1)
- España (2)

La empresa **Gruma**, es la mayor productora de harina de maíz y tortillas en el mundo.

- Con marcas líderes en la mayoría de sus mercados, opera principalmente en los Estados Unidos, México, Centroamérica, Europa, Asia y Oceanía.

Plantas en Oceanía:

- Australia (1)



SMART
DATA

Desarrollo del proyecto

1° Etapa

Puesta en marcha del proyecto
y Trabajo con Datos.

2° Etapa

Data Engineering.

3° Etapa

Data Analytics y Machine Learning.



SMART
DATA

Propuesta de trabajo

Objetivo general

- Realizar un análisis de mercado de los restaurantes mexicanos por estados de United States desde el año 2019 al 2022.



Objetivos específicos

- Crecimiento del 2 % en la cantidad de restaurantes de comidas mexicanas por trimestre por estado.
- Crecimiento del 3 % en las calificaciones iguales o mayores a tres estrellas por trimestre por estado.
- Reducción del 1 % en las calificaciones menores a tres estrellas por trimestre por estado.



Productos finales

Power Bi con KPI's de

- Incremento del 3% en cantidad de restaurantes por Trimestre
- 5% incremento de valoraciones de Estrellas Positivas
- 1% decremento de valoraciones de Estrellas Negativas
- Análisis geográfico por estado, año, análisis de sentimiento y cantidad de estrellas.

Modelo de Machine Learning

- Análisis de sentimiento de los restaurantes de comidas mexicanas en 2022.

Sistema de Recomendación.

- Top 5 de mayor calificación de los restaurantes mexicanos de los siete estados.
- Top 5 con mayor calificación de restaurantes mexicanos por estados.
- Top 5 con menor calificación de restaurantes mexicanos por estado.



SMART
DATA

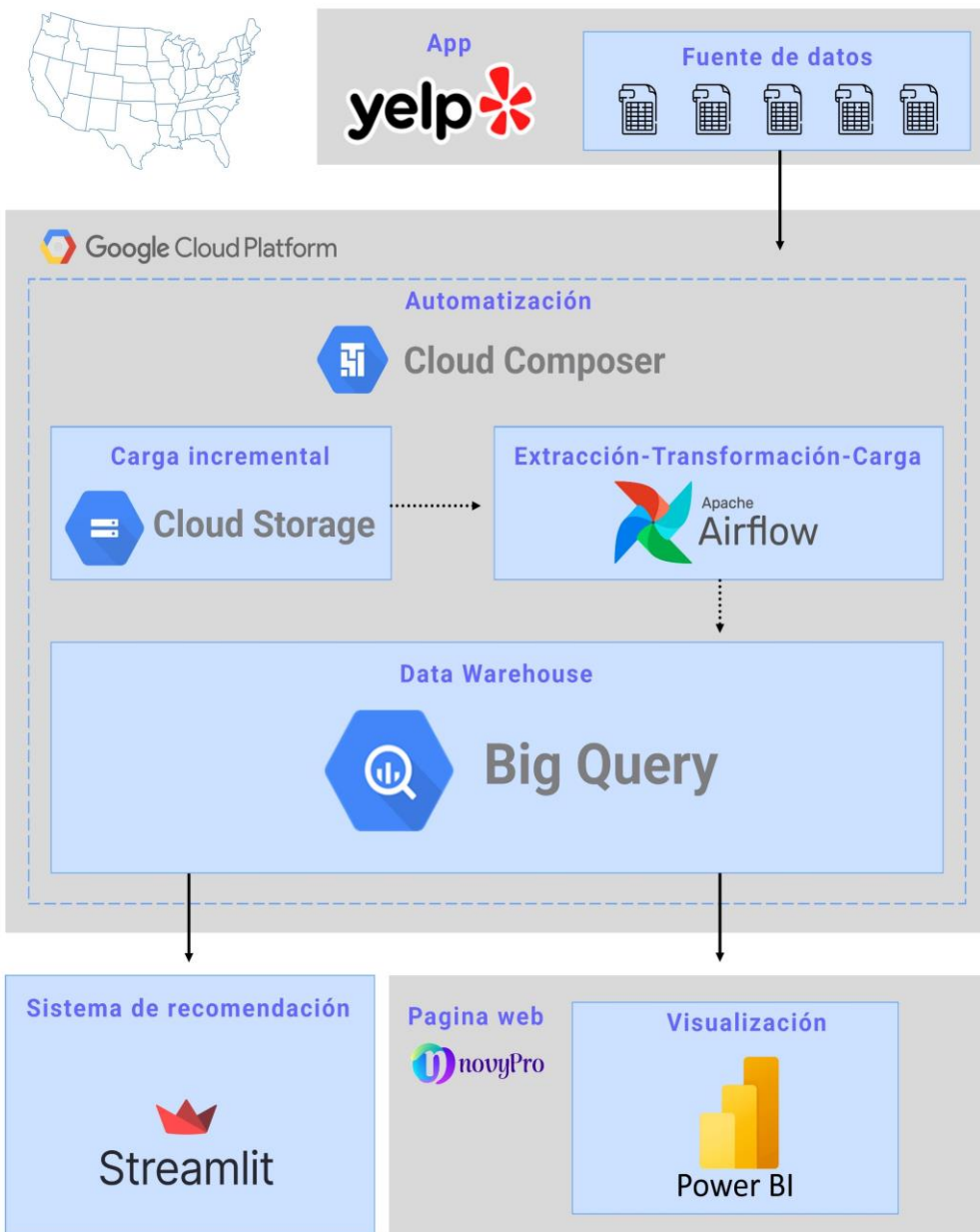
Data Pipeline



Google Cloud Platform



SMART
DATA



Data Pipeline - Stack tecnológico

- Se toman los datos de la plataforma de **Yelp**.
- Se los procesa con herramientas de **Google Cloud Platform**.
- Se almacenan en un **Data Warehouse** donde estarán disponibles.
- En **Power BI** se realizó su visualización.
- En **Streamlit** se realizó un sistema de recomendación.
- La plataforma **NovyPro** nos permite que el dashboard creado en Power BI esté disponible online.

Visualización de datos



Power BI



SMART
DATA



Power BI

[Link al video en YouTube](#)

[Link al dashboard en NovyPro](#)

Conclusiones

- Se evidencia una periodicidad de crecimiento y decrecimiento de la cantidad de Restaurantes
- Se observa una caída abrupta en el año 2020 relacionada con la Pandemia.
- Hay 4 cadenas de restaurantes que concentran la mayor cantidad de restaurantes.
- La comida mejicana tiene muy buena aceptación en este país, ya que los porcentajes de aceptación son muy altos.
- Las mayores ventas se concentran en las estaciones primavera-verano.

Análisis de datos



Visual Studio Code



SMART
DATA

Distribución de los restaurantes mexicanos en EE.UU.

Nº	ESTADOS	CANTIDAD DE RESTAURANTES
1	FLORIDA	369
2	PENSILVANIA	345
3	ARIZONA	281
4	INDIANA	266
5	TENNESSEE	263
6	MISSOURI	178
7	NEVADA	174
8	LUISIANA	106
9	CALIFORNIA	95
10	IDAHO	88
11	NUEVA JERSEY	84
12	ILLINOIS	42
13	DELAWARE	33

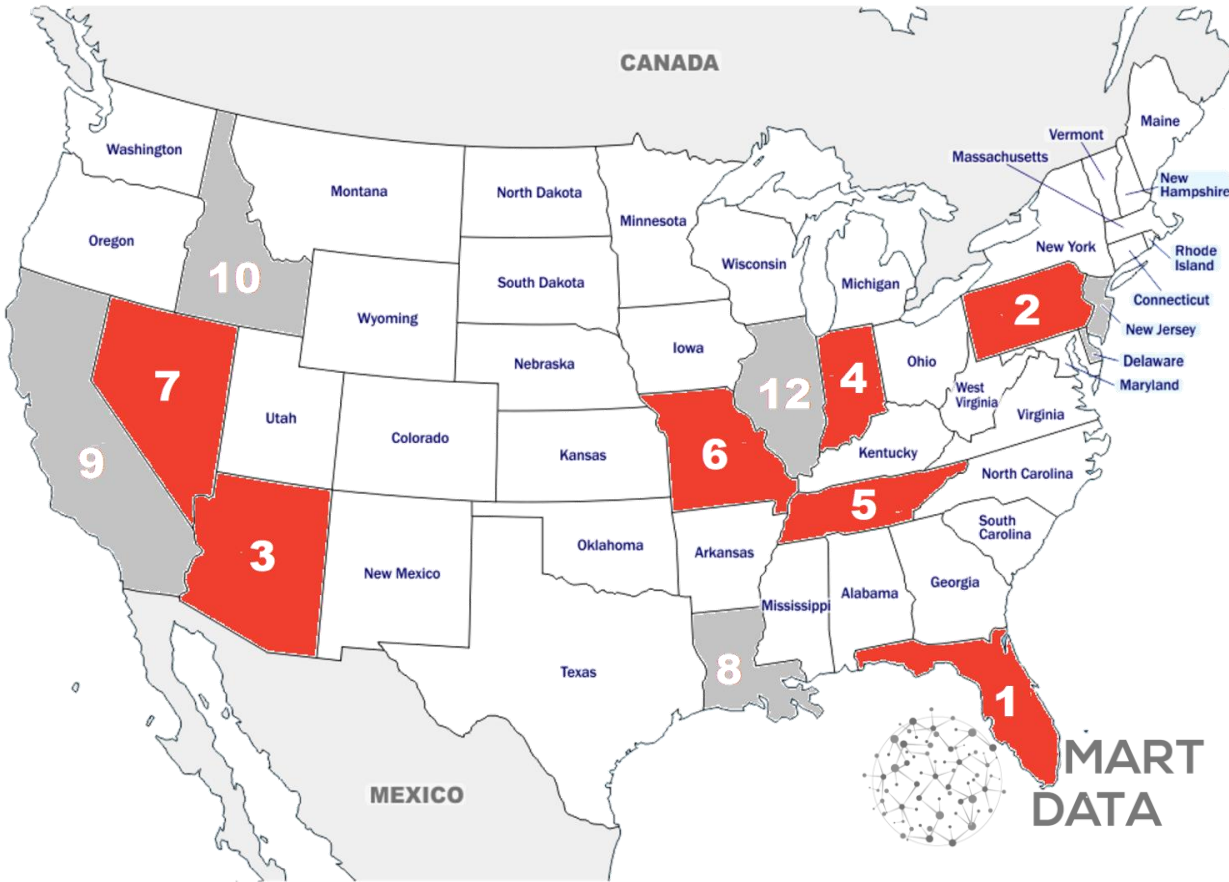
80 %

Restaurantes
mexicanos

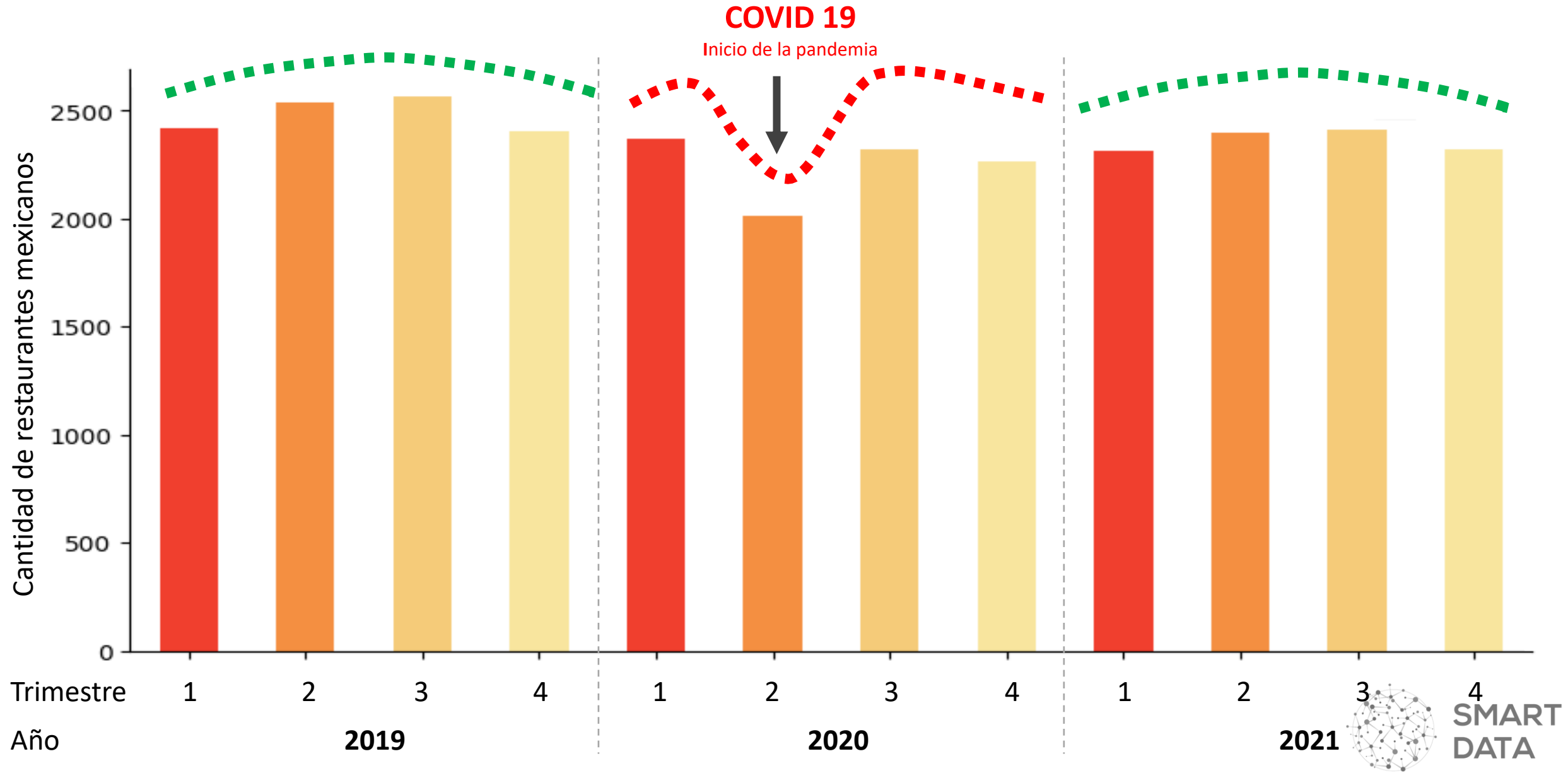
48 estados en EE.UU.

13 estados hay restaurantes mexicanos

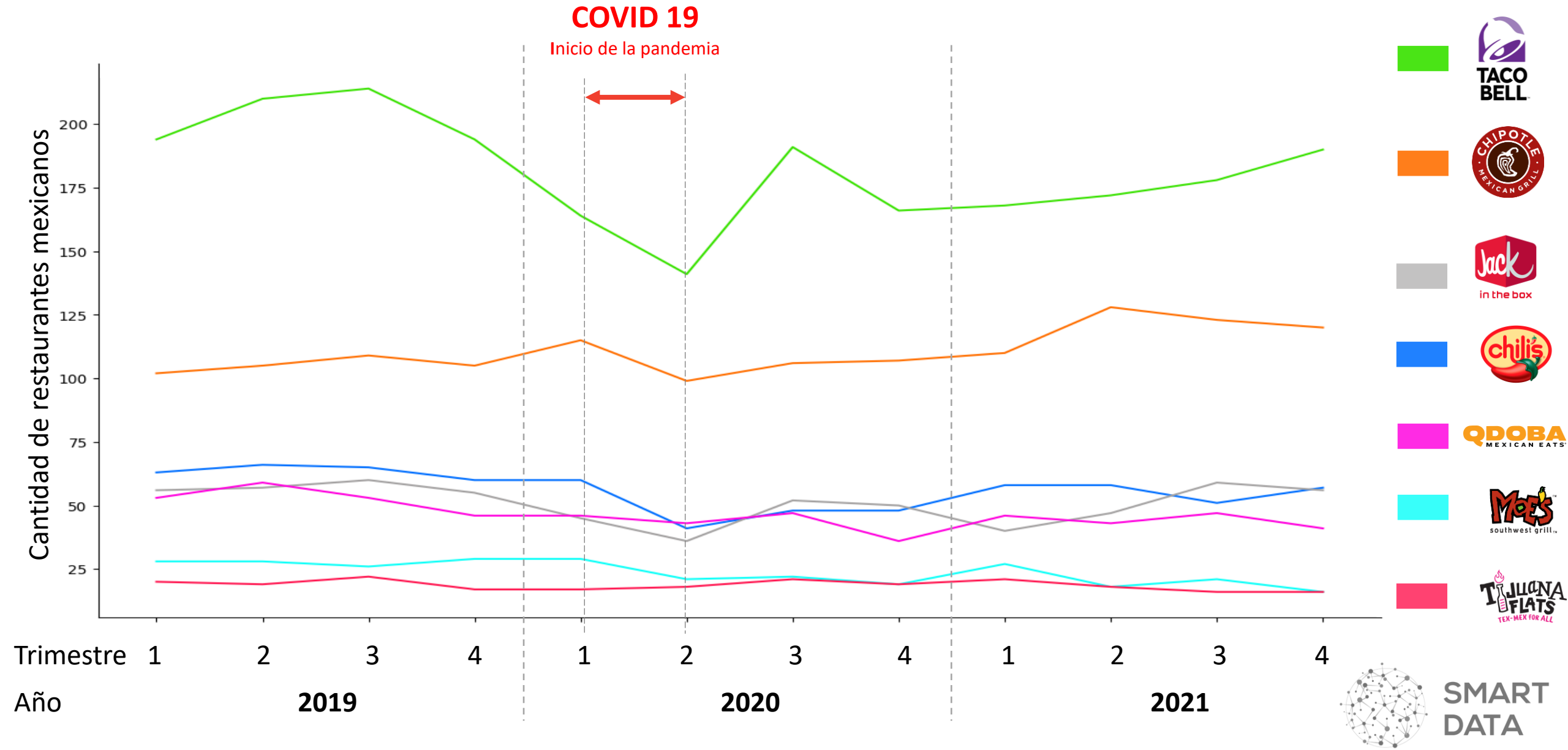
7 estados están el 80% de los restaurantes mexicanos



Cantidad de restaurantes mexicanos por trimestre



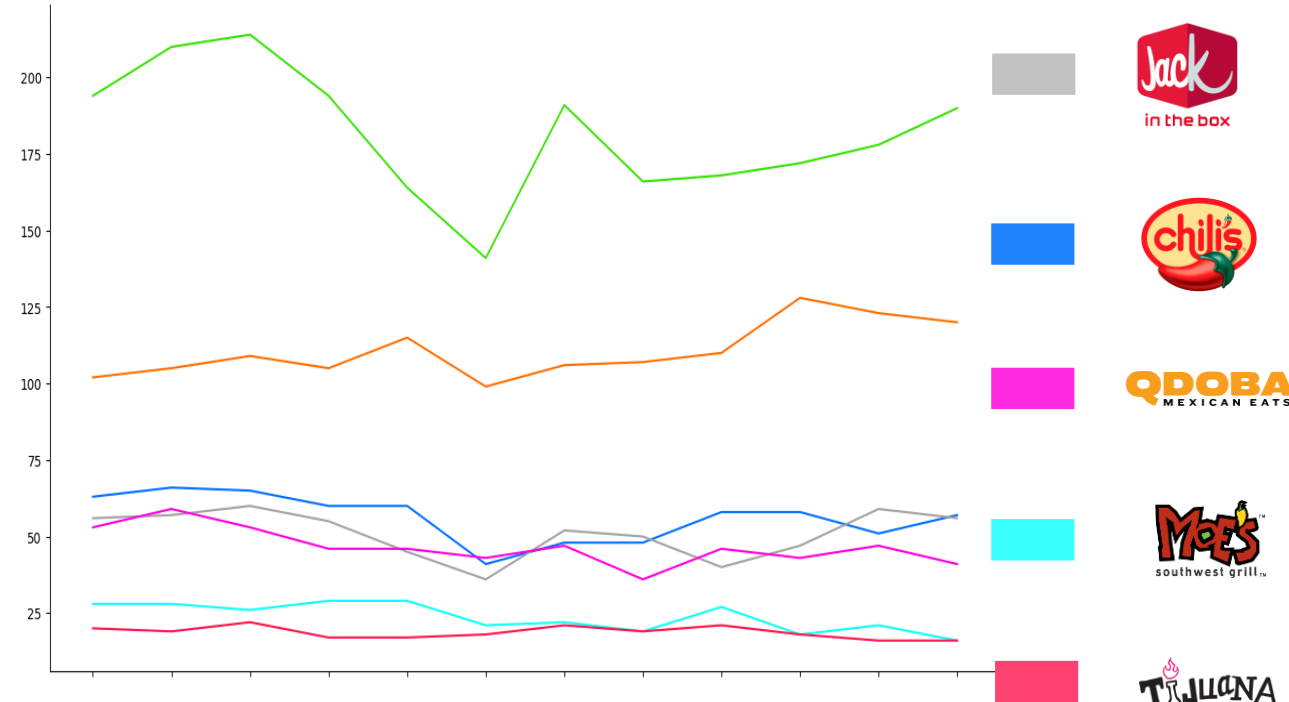
Cantidad de restaurantes mexicanos por trimestre por empresas



Parámetros estadísticos

Cantidad de restaurantes mexicanos por trimestre por empresas

EMPRESAS	MEDIA DE SUCURSALES / AÑO	DESVIO ESTANDAR
Taco Bell	181	20
Chipotle Mexican Grill	110	9
Chili's	56	7
Jack in the Box	51	7
QDOBA Mexican Eats	46	6
Moe's Southwest Grill	23	4
Tijuana Flats	18	2



Tijuana Flats
TEX-MEX FOR ALL
SMART DATA

Cotización de acciones de empresas de restaurantes mexicanos

El sector de la comida mexicana se mantiene estable y tiende a crecer.



Nube de palabras de las reseñas de los clientes



Procesamiento del lenguaje natural con Python

Las palabras más mencionadas en las reseñas son neutras a positivas

TOP 10 DE RESEÑAS DE LOS CLIENTES	
PLACE	LUGAR
FOOD	COMIDA
GOOD	BIEN
ORDER	ORDEN
TACO	TACO
DELICIOU	DELICIOSO
SERVICE	SERVICIO
ONE	UNO
ORDERED	ORDENADO
LOVE	AMOR

ChatGPT API



"I **love** this **place**, they have really **good food**."

"I **ordered** a **taco** and it was **delicious**."

"**Service one** is always prompt and friendly."

"I **love** the **taco** I **ordered**, it was **delicious**."

"This **place** has **good food**, I **ordered** a **taco** and it was **delicious**."

"I **love** the **service one** provides, they always take my **order** quickly."

"**Me encanta** este **lugar**, tienen muy **buena comida**."

"**Pedí un taco** y estaba **delicioso**".

"El **servicio** es siempre rápido y amigable."

"**Me encanta** el **taco** que **pedí**, estaba **delicioso**".

"Este **lugar** tiene **buena comida**, **pedí** un **taco** y estaba **delicioso**."

"**Me encanta** el **servicio** que brindan, siempre toman mi **pedido** rápidamente."



SMART
DATA

Insights

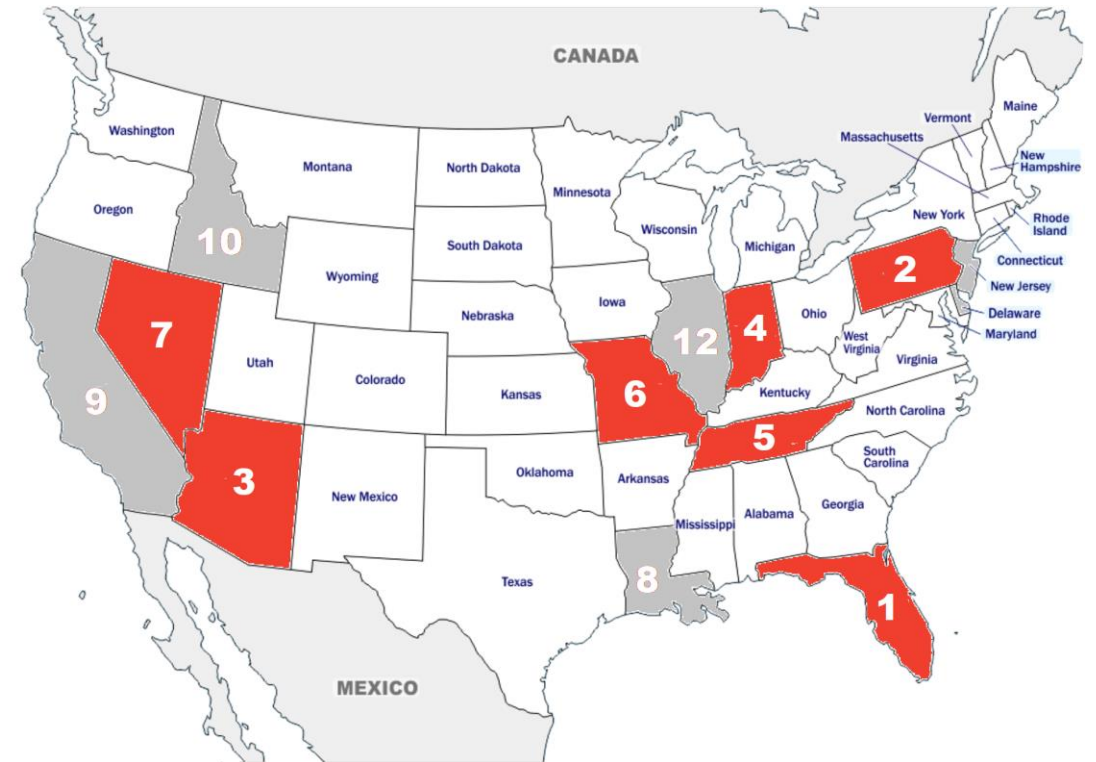


SMART
DATA



80 %

**de restaurantes mexicanos
lo tienen estas empresas**



80 %

**de los restaurantes mexicanos
se encuentran en estos estados**

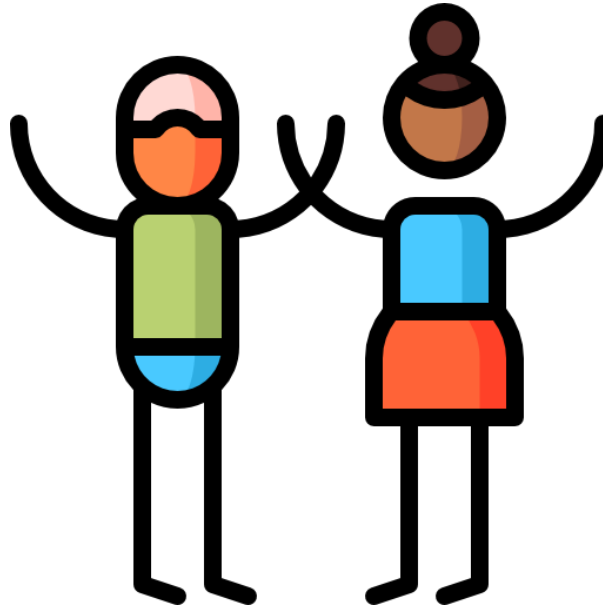


SMART DATA



Mas de ventas

en 2° y 3° trimestres



Cientes felices

Reseñas, calificaciones y análisis de sentimiento muy bueno



Crece el sector

de restaurantes mexicanos



SMART
DATA

Machine Learning



Streamlit



SMART
DATA



Streamlit

Link al sistema de
recomendación

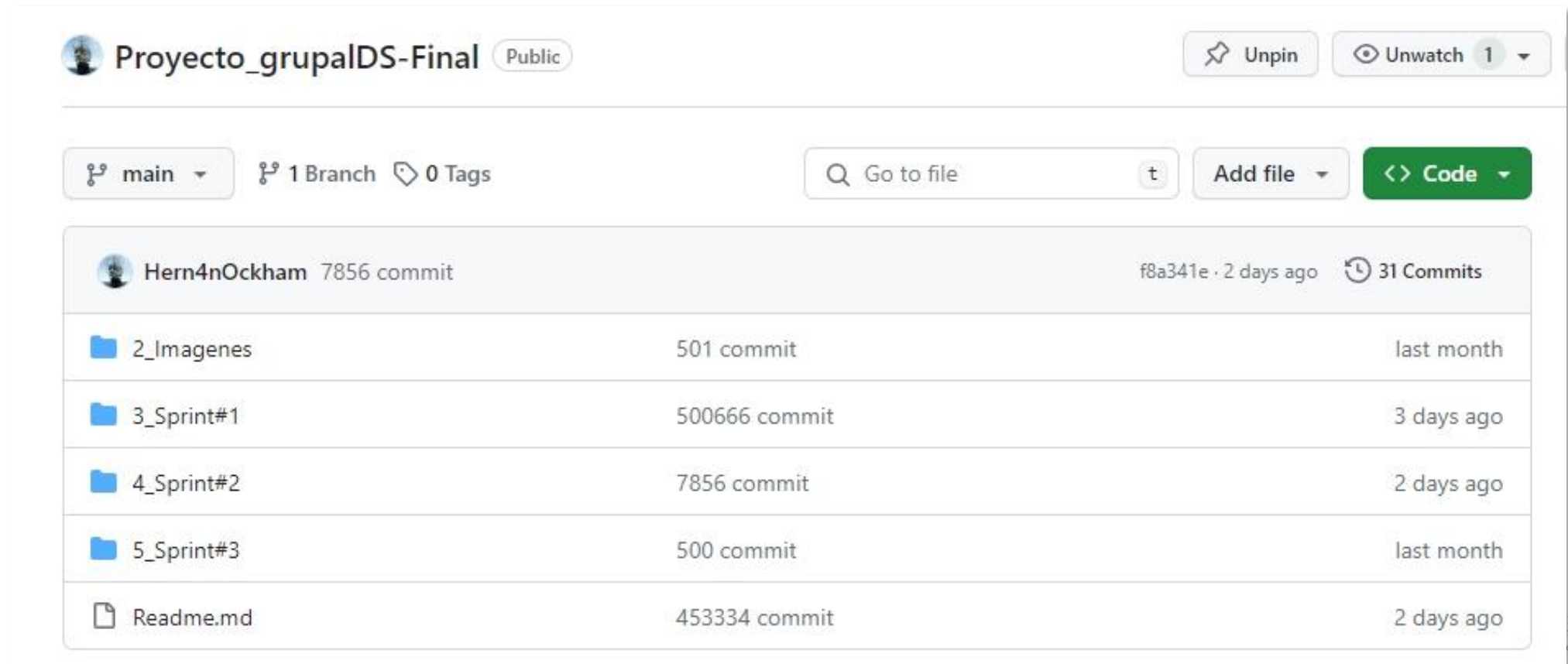
Link al sistema al modelo de
machine learning



SMART
DATA

Trabajo está disponible en GitHub

Repositorio 📌 Proyecto_grupalDS-Final



The screenshot shows the GitHub interface for the repository 'Proyecto_grupalDS-Final'. At the top, the repository name is displayed with a 'Public' badge. To the right are buttons for 'Unpin' and 'Unwatch' (with a dropdown showing '1'). Below this, the 'main' branch is selected, with '1 Branch' and '0 Tags' indicated. A search bar 'Go to file' and buttons for 'Add file' and 'Code' are present. The commit history table shows the following data:

Commit	Author	Commit Count	Time
f8a341e	Hern4nOckham	7856	2 days ago
501		501	last month
500666		500666	3 days ago
7856		7856	2 days ago
500		500	last month
453334		453334	2 days ago



SMART
DATA



GRACIAS



SMART
DATA