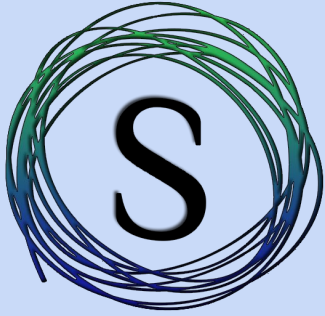


# Proyecto Final

Sinergia Global & CulinaryCrossRoads



Edith Cuellar | Leandro Ibarra | Matias Baez | Nicolás Ibarra | Tinmar Andrade

**HENRY**

# Consultora Sinergia Global

Roles y funciones:



Edith Cuellar  
*Data Scientist*



Leandro Ibarra  
*Data analyst*



Matias Baez  
*Machine Learning Engineer*



Nicolás Ibarra  
*Analista y coordinador*



Tinmar Andrade  
*Data Engineer*



# Culinary CrossRoads



**HENRY**



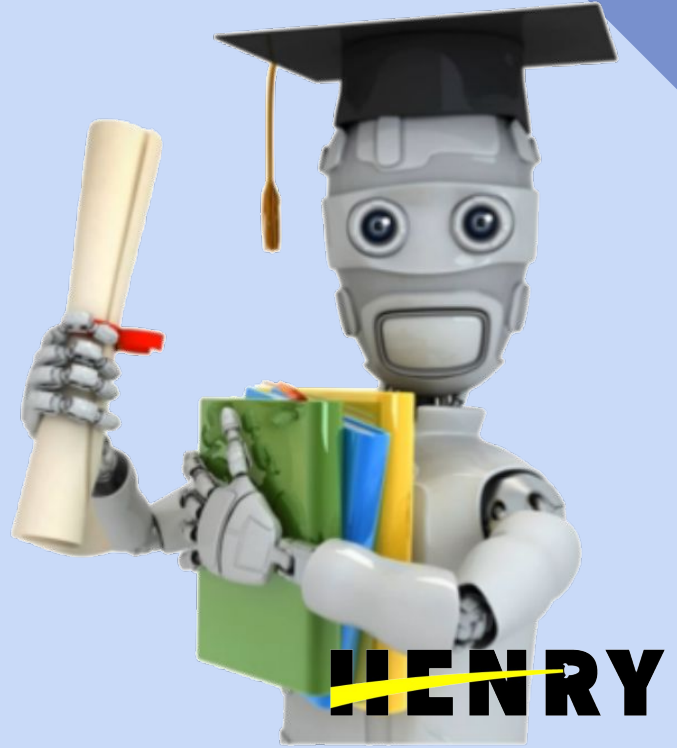
# Alcance

- ❖ Temporal
- ❖ Geográfico
- ❖ De los datos



# Objetivos

- ❖ Identificar oportunidades para nuevos negocios
- ❖ Analizar datos de usuarios para sistemas de recomendación
- ❖ Implementar modelos de aprendizaje automático:



**HENRY**



# KPI's

KPI	Fórmula	Medida	Objetivo
Variación trimestral del rating	$(\text{promedio rating actual} - \text{promedio rating anterior}) / \text{promedio rating anterior} ) * 100$	Trimestral	5%
Variación trimestral de las reseñas	$(\text{reseñas totales actual} - \text{reseñas totales anteriores} / \text{reseñas totales anteriores}) * 100$	Trimestral	5%
Reseñas positivas	$[(\text{reseñas positivas año actual} - \text{reseñas positivas año anterior}) / (\text{reseñas positivas año anterior})] * 100$	Anual	10%
Reseñas negativas	$[(\text{reseñas negativas año actual} - \text{reseñas negativas año anterior}) / (\text{reseñas negativas año anterior})] * 100$	Anual	-10%
Presencia	$[(\text{número de locales de nuestro rubro actual} / \text{número de locales totales actuales}) / (\text{número de locales de nuestro rubro año anterior} / \text{número de locales totales año anterior})] * 100$	Anual	2%
Calificación vs competencia	$[(\text{calificación promedio nuestros locales actual} / \text{calificación promedio de la competencia actual}) / (\text{calificación promedio nuestros locales anterior} / \text{calificación promedio de la competencia anterior})] * 100$	Anual	2%



# Medidas

Medida	Descripción
<b>Rubro que predomina por estado</b>	<ul style="list-style-type: none"><li>• Identificación del sector o industria con mayor presencia de locales por estado</li><li>• Comprender la influencia del rubro</li></ul>
<b>Cadenas con mejores reseñas</b>	<ul style="list-style-type: none"><li>• Identificación de las cadenas de restaurantes que cuentan con las calificaciones más altas</li><li>• visión general de la calidad percibida por los clientes.</li><li>• Patrones</li></ul>
<b>Palabras clave que identifican a las cadenas de restaurantes</b>	<ul style="list-style-type: none"><li>• Frases recurrentes</li><li>• Características distintivas</li><li>• Marketing</li></ul>
<b>Variación porcentual anual de la cantidad de reseñas a nivel general</b>	<ul style="list-style-type: none"><li>• Variación porcentual en el número total de reseñas.</li><li>• Permite entender el cambio en la retroalimentación de los clientes.</li></ul>



**Colaboración y comunicación**  
Google Meet, Slack y Discord,.



**Lenguaje de programación y manipulación de datos**  
Python, SQL y Big Query.



**Gestión de proyecto**  
Trello



**Análisis y procesamiento de imágenes**  
Pandas y NumPy



**Aprendizaje automático**

Scikit-learn para aprendizaje automático, Keras y TensorFlow para el desarrollo y entrenamiento de modelos de aprendizaje profundo.



**Gestión de flujos de trabajo**  
Apache Airflow



**Almacenamiento y análisis a gran escala**

Google Cloud Storage y Google BigQuery



**Visualización**  
Power BI



# Stack tecnológico:

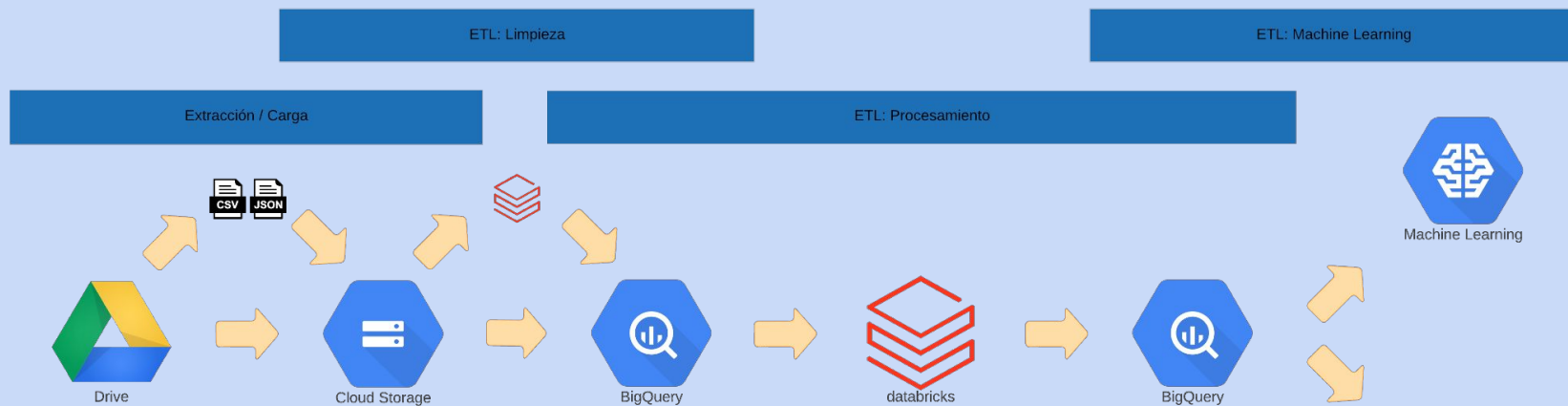
- ❖ Colaboracion y comunicacion
- ❖ Lenguajes
- ❖ Gestión del proyecto
- ❖ Análisis y procesamiento de datos
- ❖ Aprendizaje automático
- ❖ Gestión de flujos de trabajo
- ❖ Almacenamiento y análisis a gran escala
- ❖ Visualización de datos



# Arquitectura



Apache  
Airflow



Machine Learning



Power BI

**HENRY**





# Tareas realizadas

Diagrama de Gantt	✓
EDA Preliminar de los Datos	✓
Creación del Repositorio Git	✓
Readme	✓
Definición de Roles y KPIs	✓
Definición del Stack Tecnológico	✓



# Tareas por realizar

Creación del DW
Automatizar el DW
Carga de datos
Análisis de datos de muestra
MVP Dashboard
MVP Producto ML



# Modelo Predictivo ML

¿Qué parte?



Saturación del mercado



Sucursales  
por  
habitantes

Gracias  
¿Preguntas?

