

# DATACKET



"Análisis y Sistema de Recomendación de Restaurantes basado en opiniones de los usuarios de Yelp y Google Maps en Estados Unidos"



REPOSITORIO



# NUESTRO EQUIPO

Somos **Datacket**, una consultora con amplia experiencia e innovación en diferentes sectores de la industria.

Nos enfocamos en brindar soluciones de análisis de datos, inteligencia de negocios y modelado predictivo de datos innovadoras y personalizadas para nuestros clientes, y trabajamos estrechamente con ellos para identificar las habilidades y conocimientos clave que serán necesarios en el futuro.



**Ada Parhuana**  
Data Engineer



**Jesús Chávez**  
Data Analyst



**José Brito**  
Data Scientist



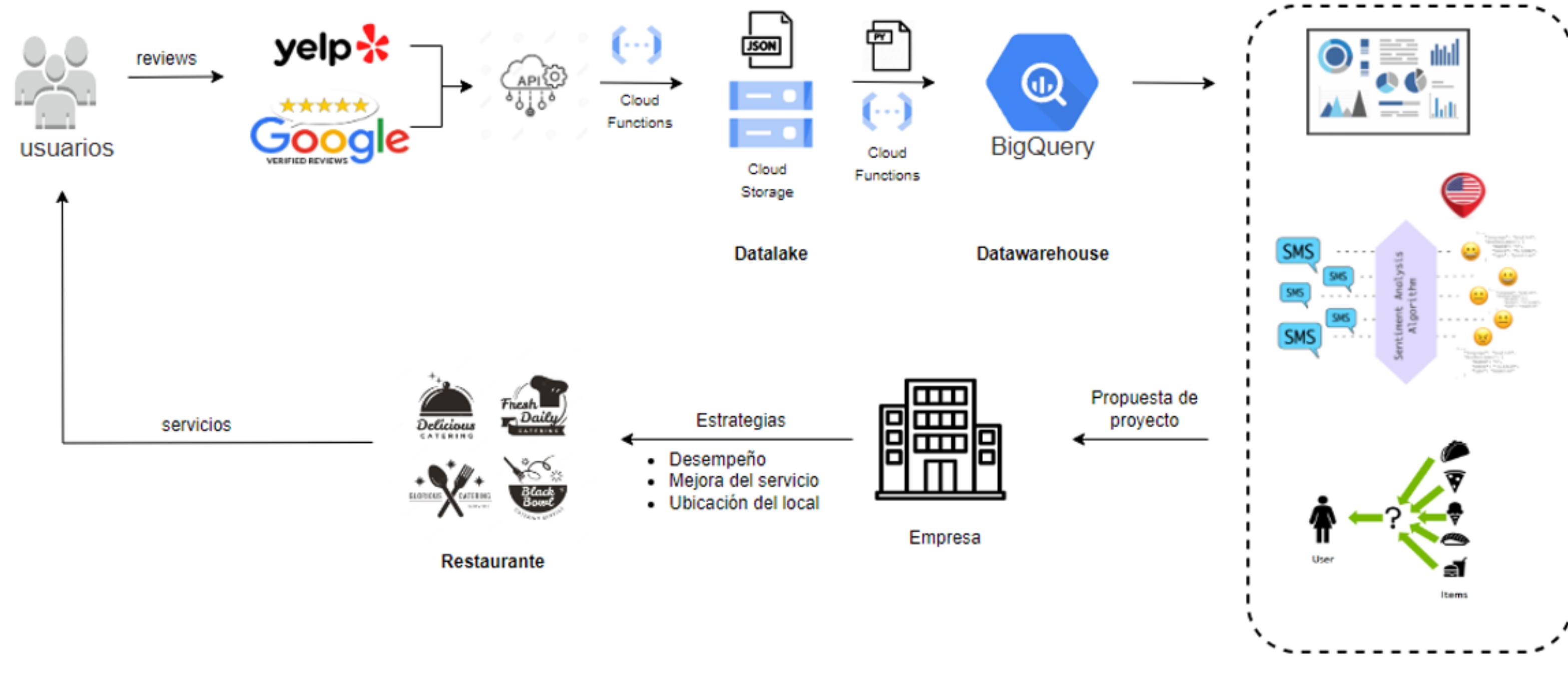
**Keily Lasso**  
Data Engineer



**Oriana Madriz**  
Data Scientist

# Planteamiento del problema

"Análisis y Sistema de recomendación de restaurantes basado en opiniones de los usuarios de Yelp y Google Maps en Estados Unidos"

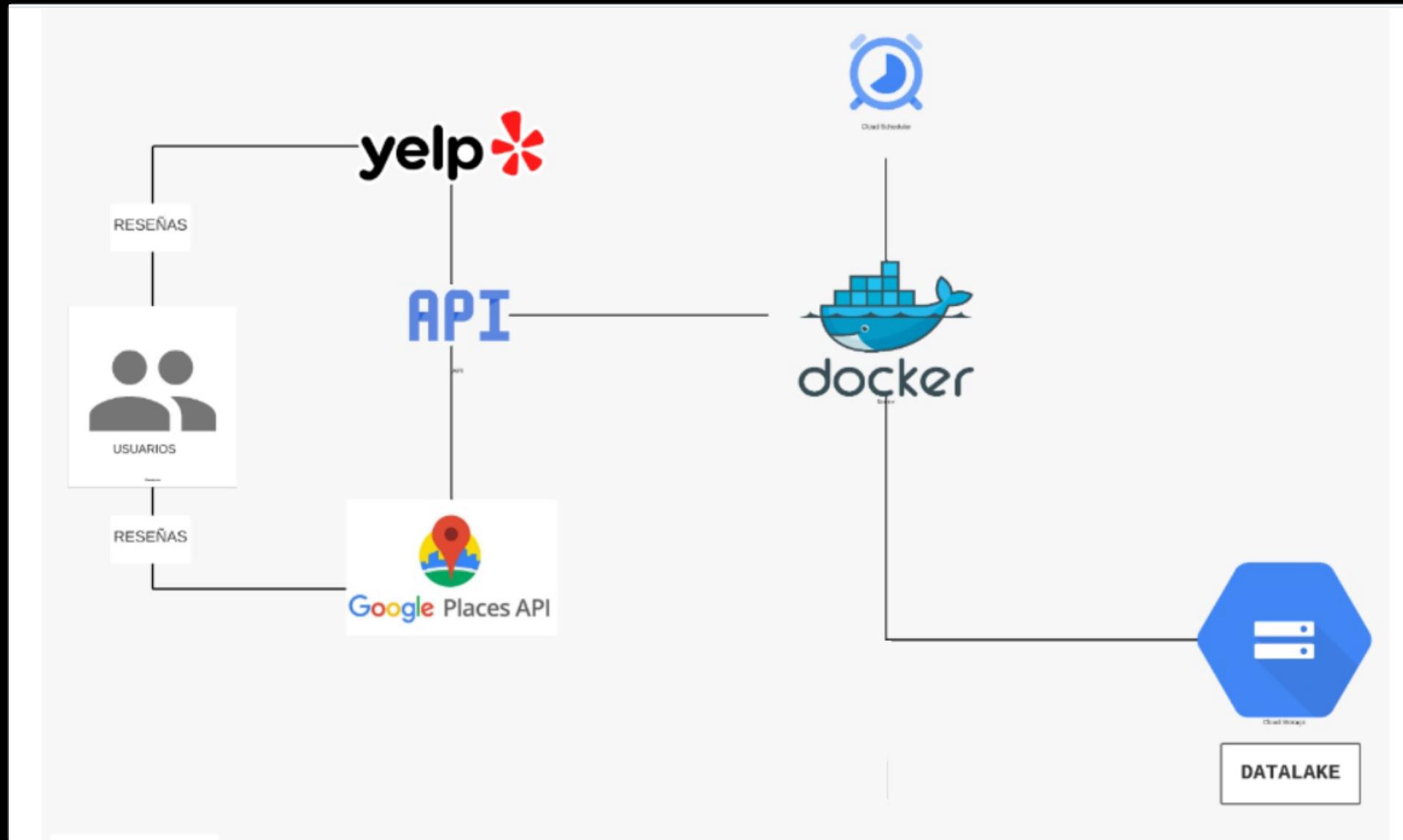


# Propuesta de Proyecto

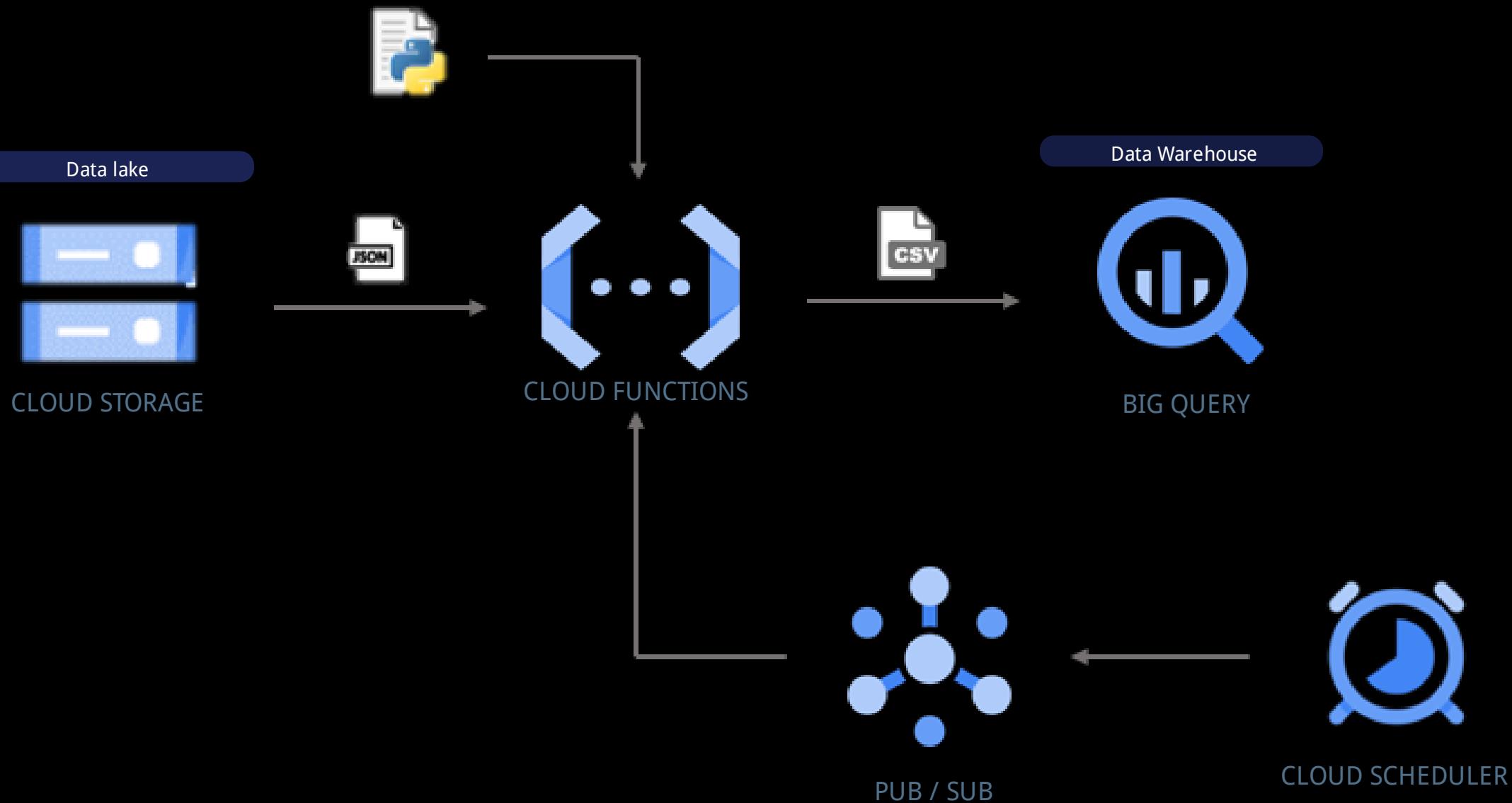
## Objetivos

- \*Hacer una recomendación basada en datos de los 5 estados en los que le conviene inaugurar sus primeros locales.
- \*Elaborar un dashboard mostrando diferentes indicadores que permitan conocer el estado del mercado gastronómico.
- \*Conocer los aspectos en los que resultaría más conveniente y efectivo focalizar sus esfuerzos de marketing.
- \*Evaluar el crecimiento del rubro de restaurantes y compararlo con el de otros rubros.
- \*Determinar cuáles son las variables que tienen en común los restaurantes o cadenas gastronómicas más exitosas en los estados a los que apunta el cliente
- \*Crear un sistema de recomendación

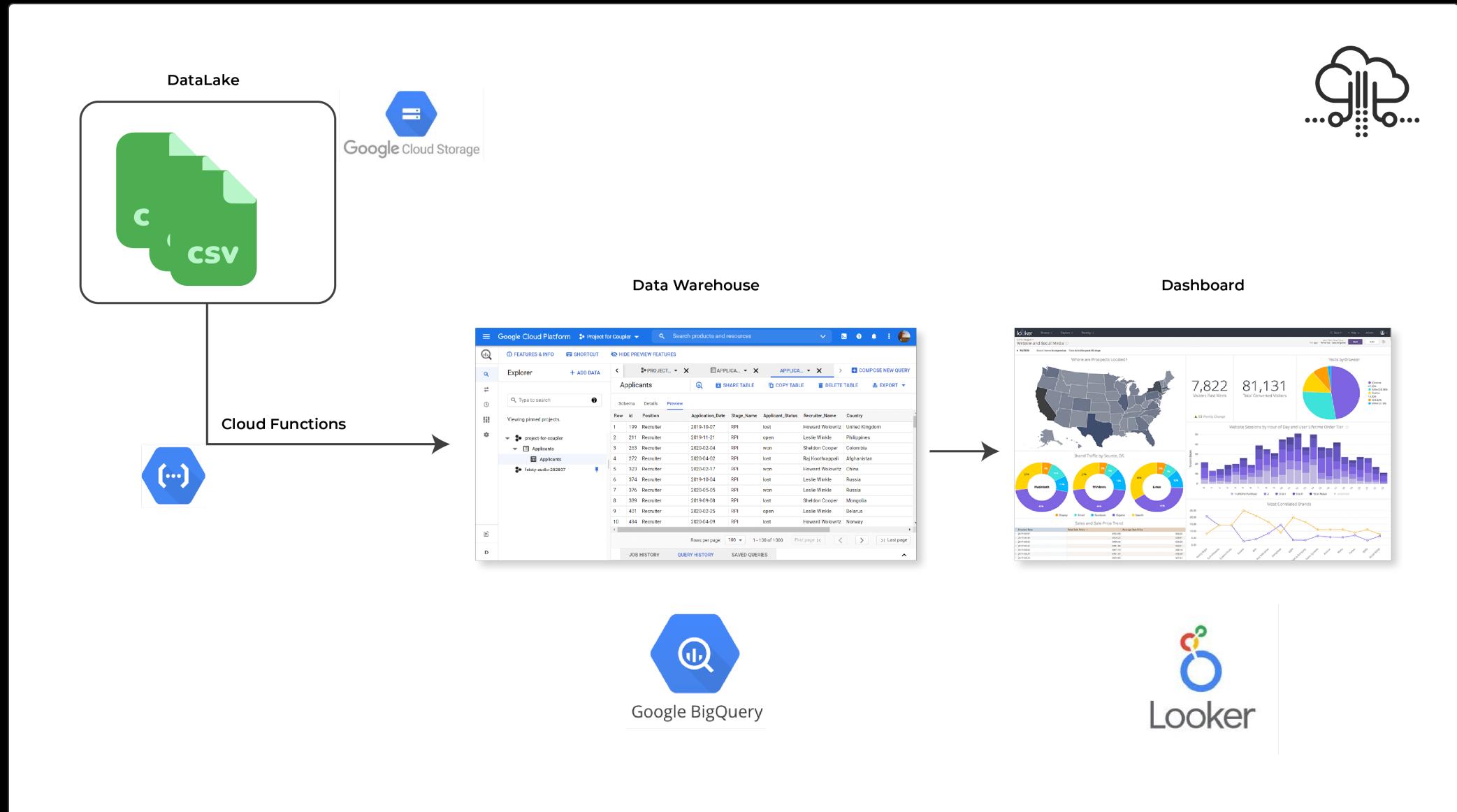
# Pipeline automatizado - Extracción



# Pipelines automatizado - ETL



# Carga incremental

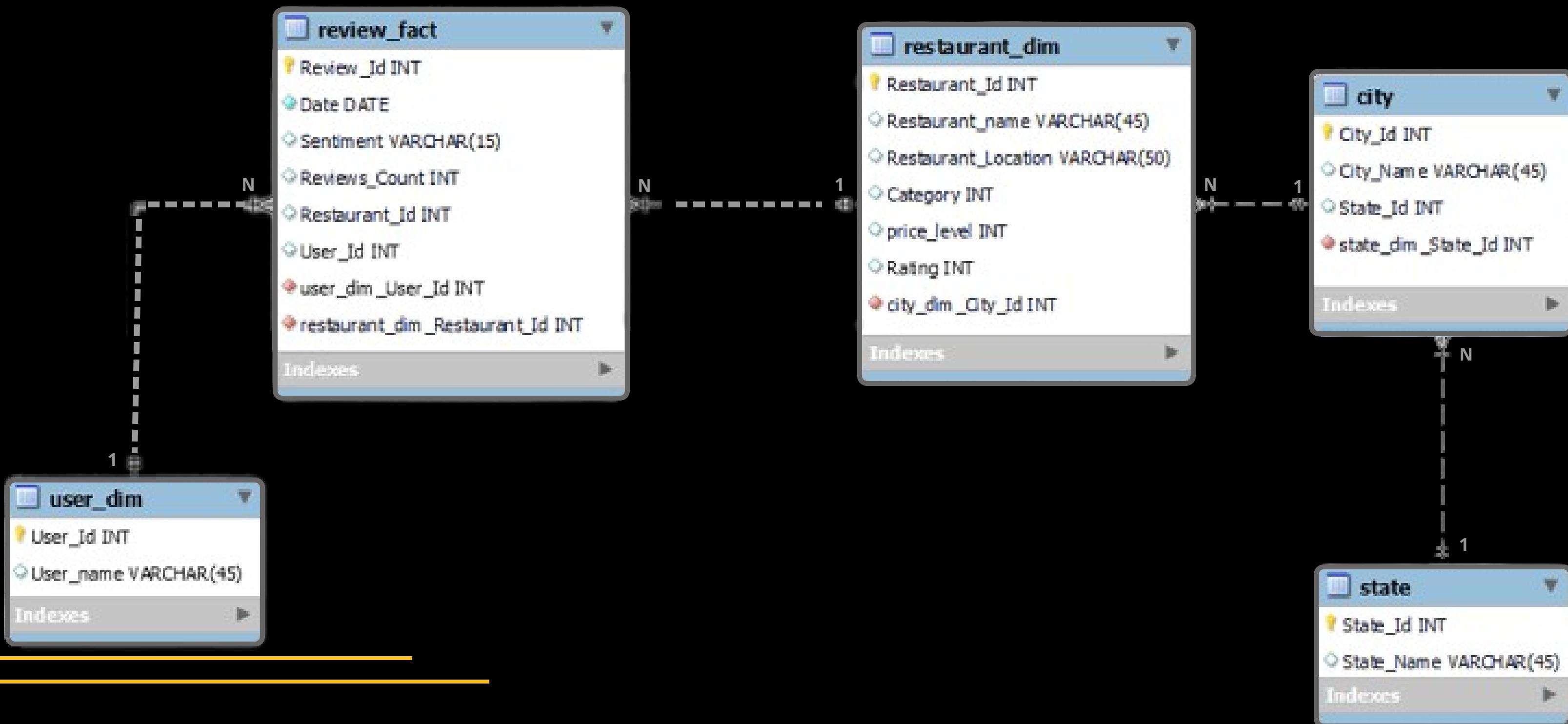


# Carga incremental



# Modelo ER

DATACKET 





state

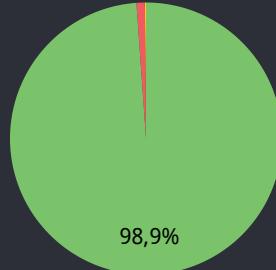
# of Restaurant  
264,0reviews  
1,6 mil# of User  
1,5 mil price\_level rating Escriba el término de búsqueda

<input checked="" type="checkbox"/> 2	5,1 mil
<input checked="" type="checkbox"/> 3	3,4 mil
<input checked="" type="checkbox"/> 4	1,7 mil
<input checked="" type="checkbox"/> 1	458

 sentiment rating Escriba el término de búsqueda

<input checked="" type="checkbox"/> neutral	5
<input checked="" type="checkbox"/> positive	5
<input checked="" type="checkbox"/> negative	2

- positive
- negative
- neutral

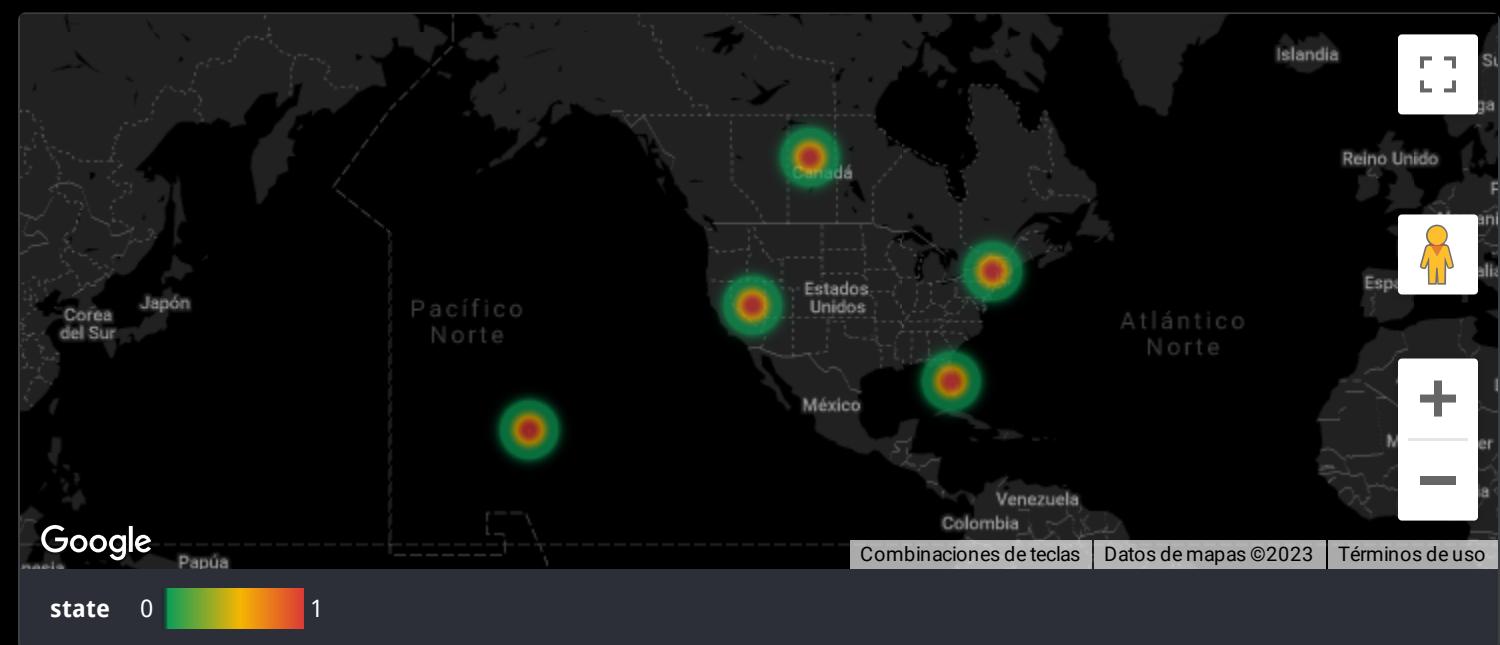
 restaurant\_name rating Escriba el término de búsqueda

<input checked="" type="checkbox"/> Tuome	5
<input checked="" type="checkbox"/> The Mission / East Village	5
<input checked="" type="checkbox"/> Red Carpet Italian Restaurant D...	5

 restaurant\_location reviews Escriba el término de búsqueda

<input checked="" type="checkbox"/> 129 E Fremont St, Las Vegas, NV...	19
<input checked="" type="checkbox"/> 1290 University Ave, Rochester, ...	13
<input checked="" type="checkbox"/> 274 N Goodman St, Rochester, ...	13

Los presentes KPIs nos brindan una visión integral de cómo su negocio está siendo percibido por los usuarios, así como oportunidades clave para mejorar y mantenerse competitivo en el mercado.



# KPI'S

## 1. Puntaje promedio por categoría:

Mide la calidad percibida por los usuarios de las diferentes categorías por restaurante. Ej.: Comida rápida, comida vegetariana, comida coreana, comida marina. Se calcula segmentando los restaurantes por categorías y promediando el puntaje (de 1 a 5 estrellas) por cada categoría. Es útil para medir las tendencias y preferencias del público. Objetivo: Mantener el puntaje de la categoría en la que se va a invertir siempre entre 4 y 5 estrellas mensualmente.

## 4.Sentimiento por estado:

Con un modelo de Machine Learning, se hace un análisis de las reviews que reciben los distintos tipos de restaurante en cada estado. Esto permite conocer la aceptación del público que tienen los diferentes tipos de restaurantes en cada estado. Esto puede resultar útil para segmentar de manera más precisa campañas de marketing o como un parámetro útil para decidir invertir en determinado tipo de restaurante en cada estado. Objetivo: Mantener el sentimiento positivo de la categoría del negocio o del estado a invertir por encima del 70% anualmente.

## 2.Cantidad de reseñas por restaurante:

Es útil para medir la interacción en el restaurante (o restaurantes, si tiene varias sucursales). Suma el total de reseñas que se obtuvieron en plataformas sociales (Gmaps y Yelp). Un aumento constante refleja el crecimiento de la relevancia y popularidad del negocio. Objetivo: Elevar la cantidad de reseñas por restaurante por lo menos un 15% anualmente. Objetivo: Mantenerse en el Top 5 de restaurantes con mayor número de reseñas

## 3.Impacto potencial promedio de los reviewers:

Promedia el número de fans que tienen las personas que escribieron reseñas del negocio en Yelp. Es útil para establecer objetivos de marketing y el desarrollo de campañas de comunicación. Objetivo: Mantener mensualmente el Impacto Potencial Promedio de los Reviewers por encima de 250.

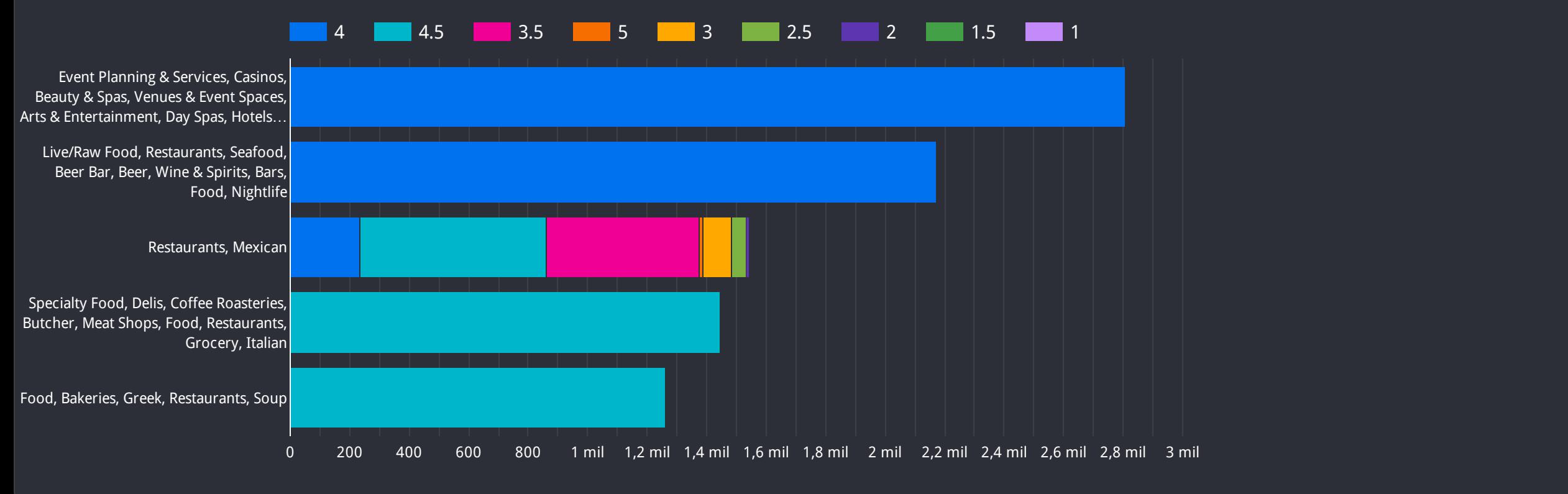
## 5.Tasa mensual de satisfacción de los reviewers:

Calcula el porcentaje de calificaciones positivas (4 o 5 estrellas) del total de puntuaciones que el restaurante recibió. Es útil para medir la reputación del negocio y establecer objetivos de mejora para el futuro .

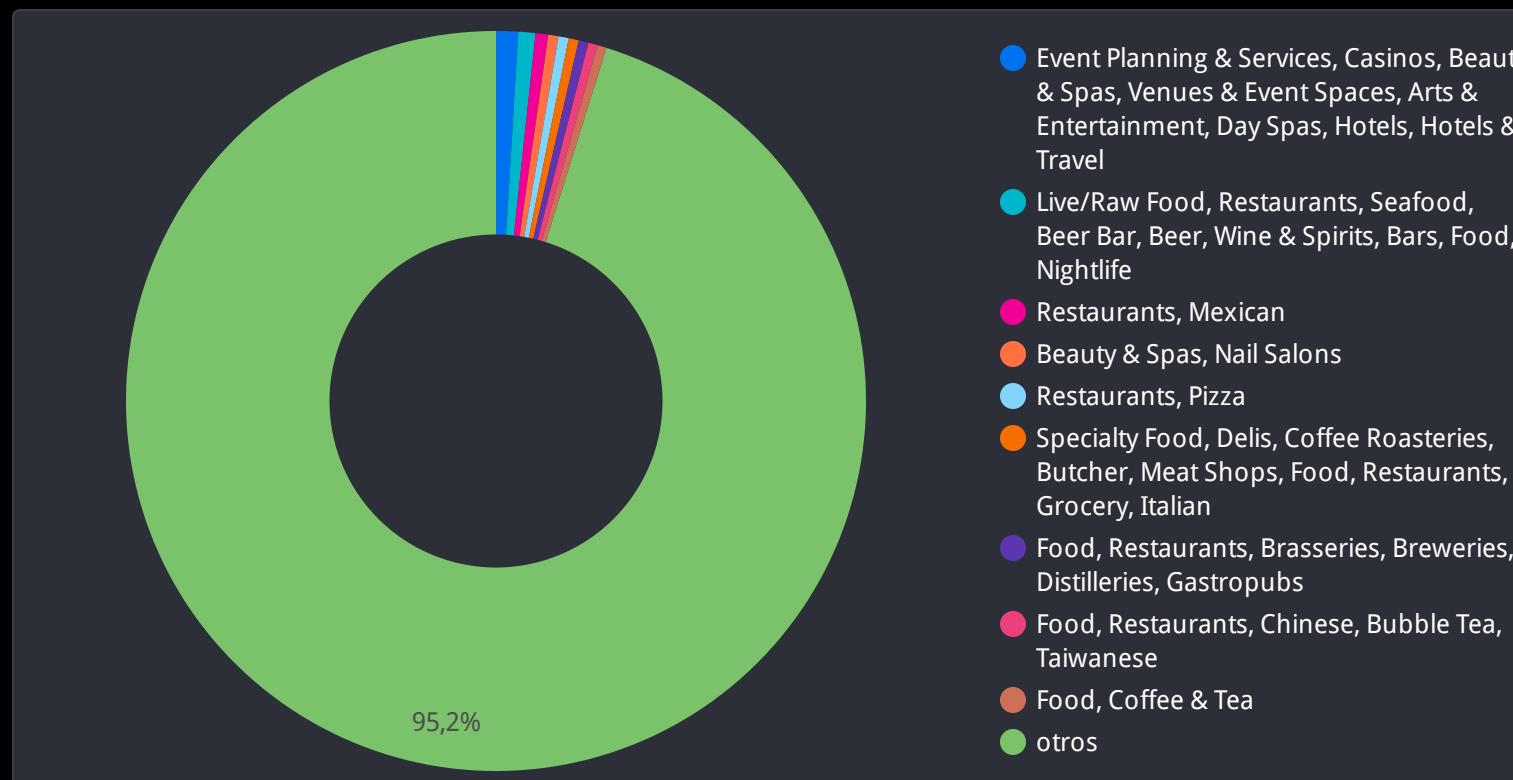
**Objetivo:** Mantener la Tasa Mensual de Satisfacción de los Reviewers siempre por sobre el 85%

$(\text{CantidadDeCalificaciones}/\text{CantidadDeCalificacionesPositivas}) * 100$

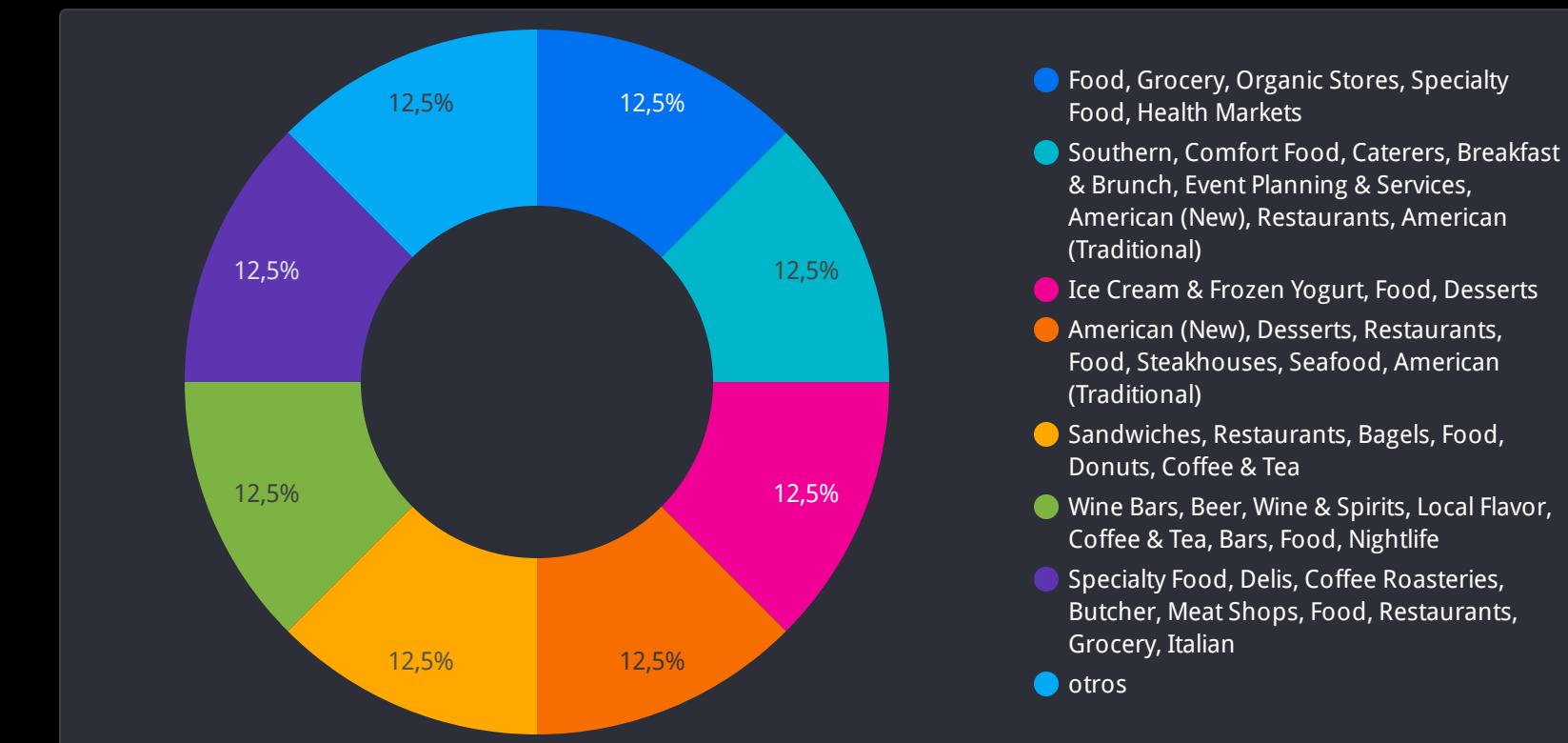
## Top 5 Categorías de Negocio



## Porcentaje por Categorías



## Sector de Restaurantes

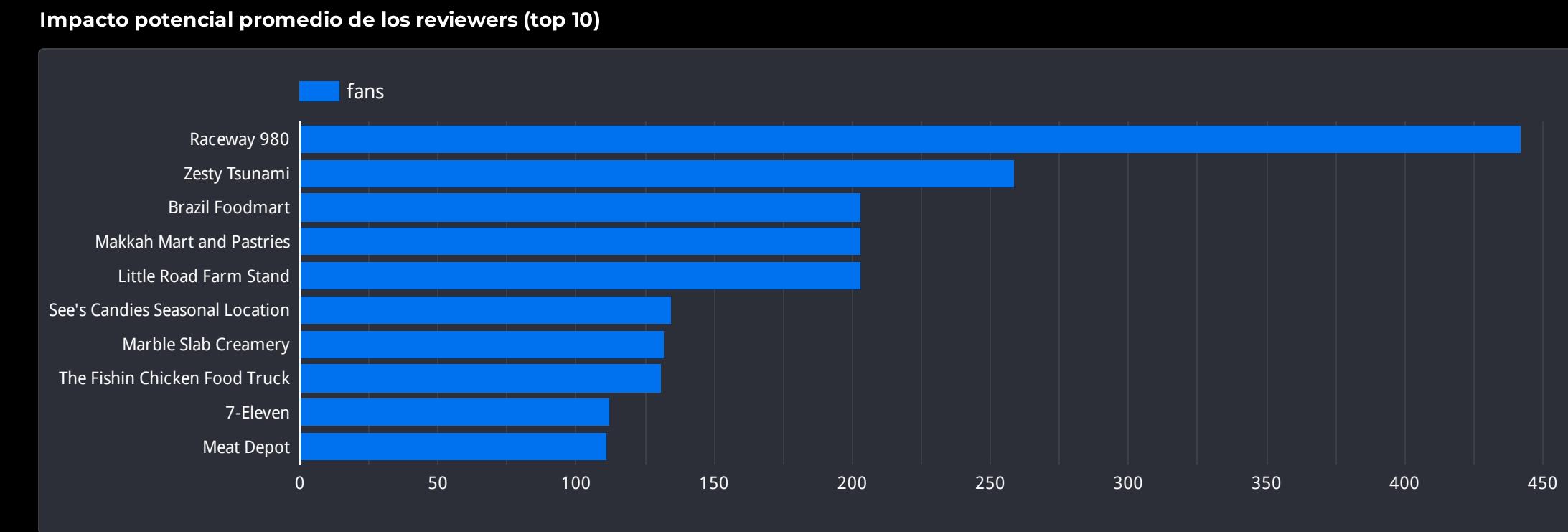
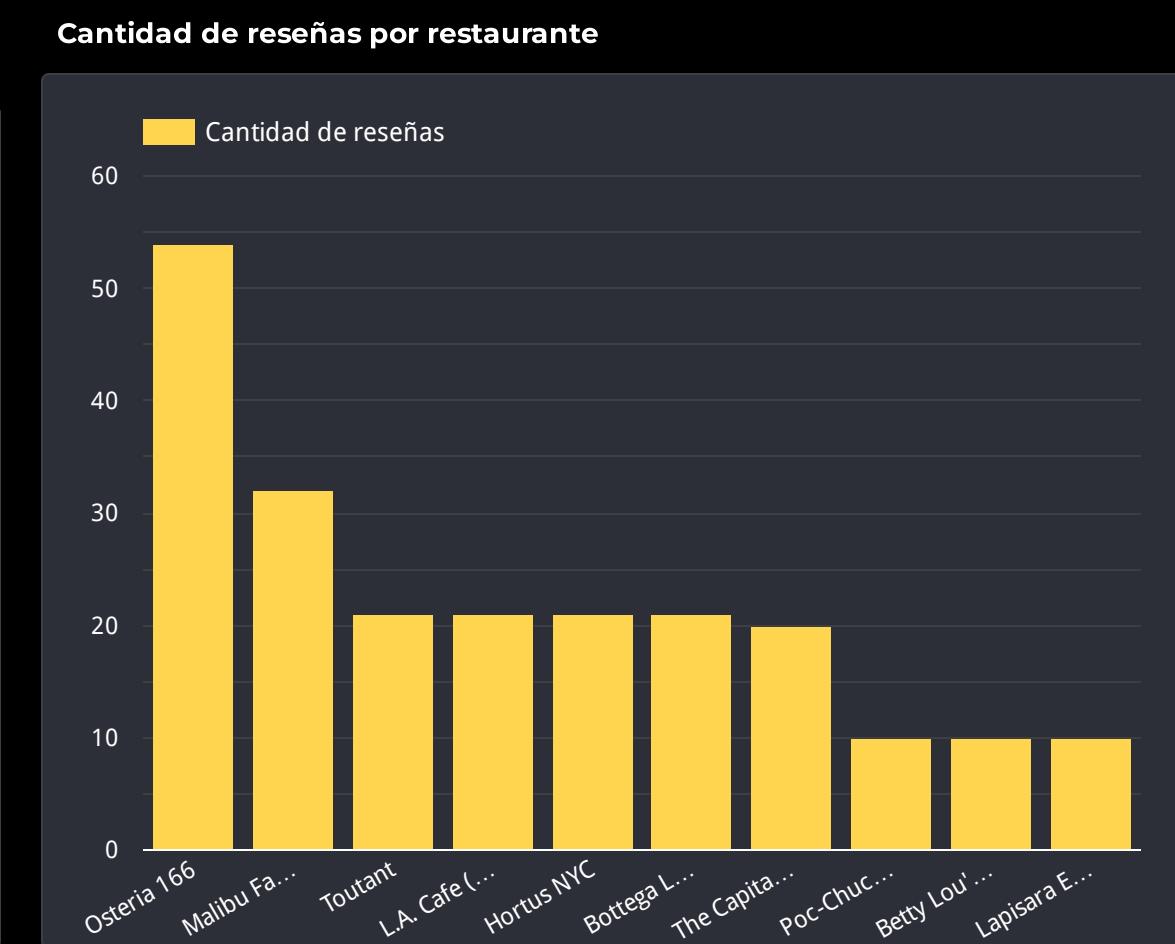
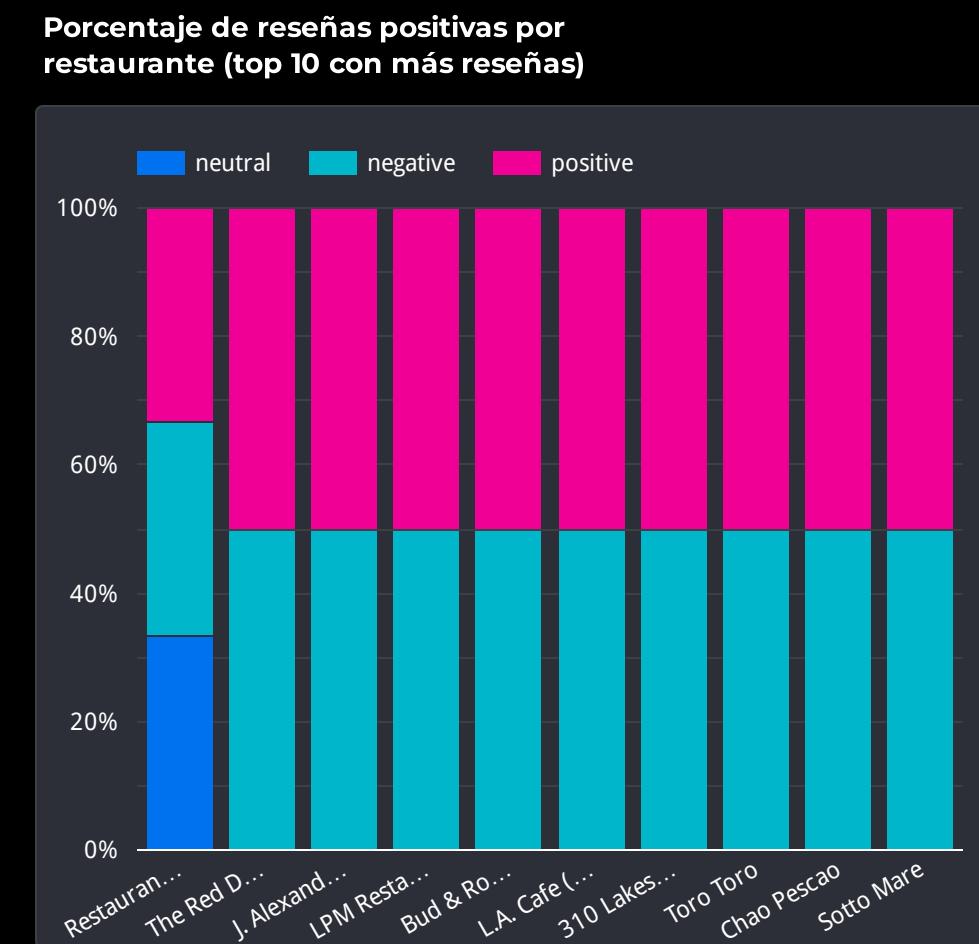
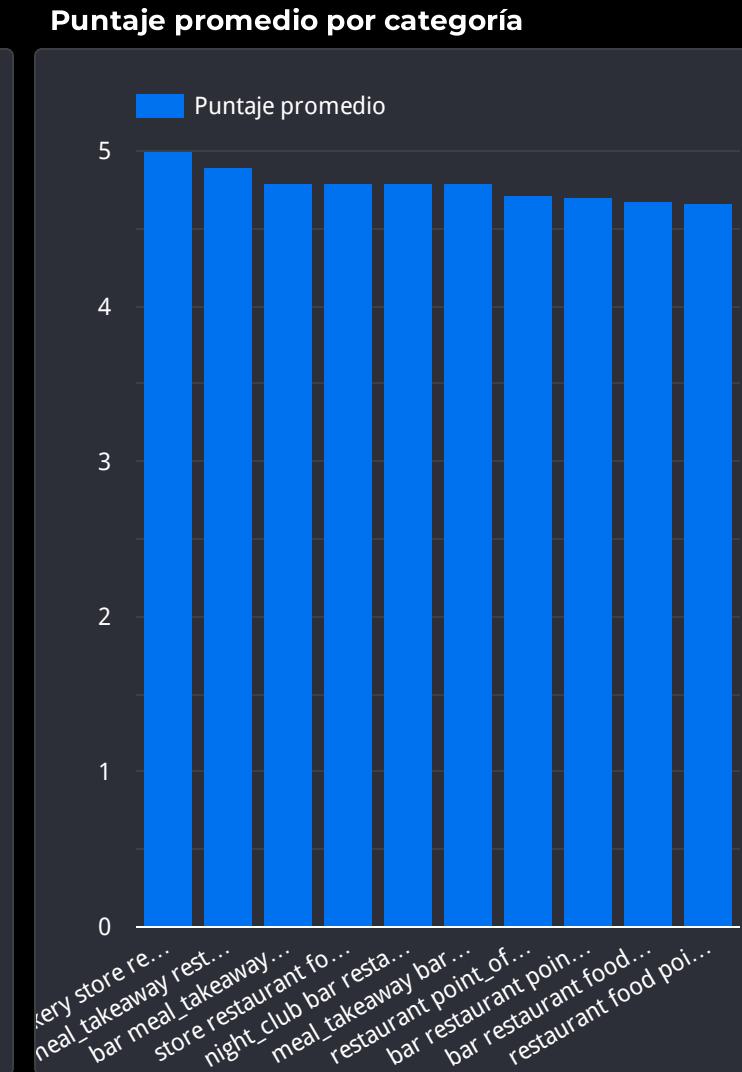


## Selecciona un periodo

# Número de restaurantes: 15.259

# Calificación promedio

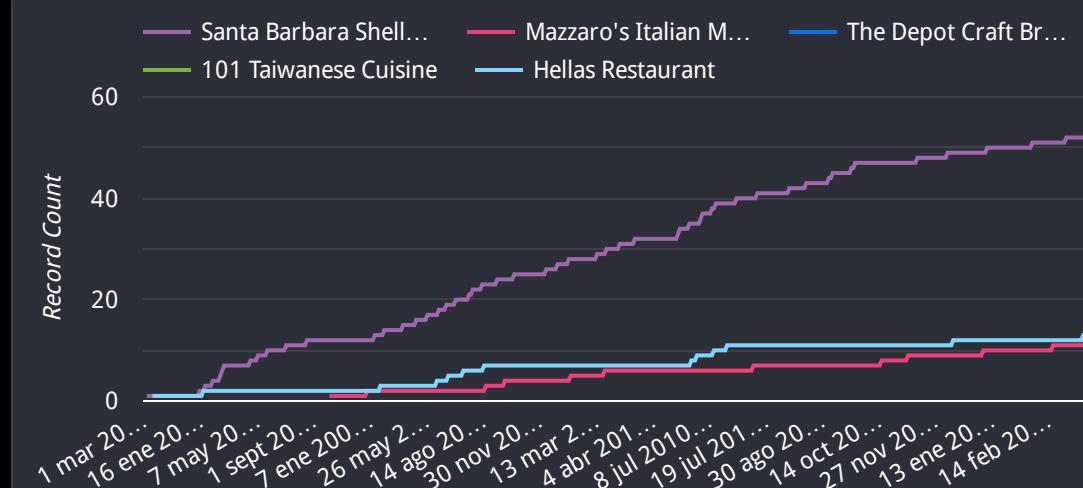
# 3,86



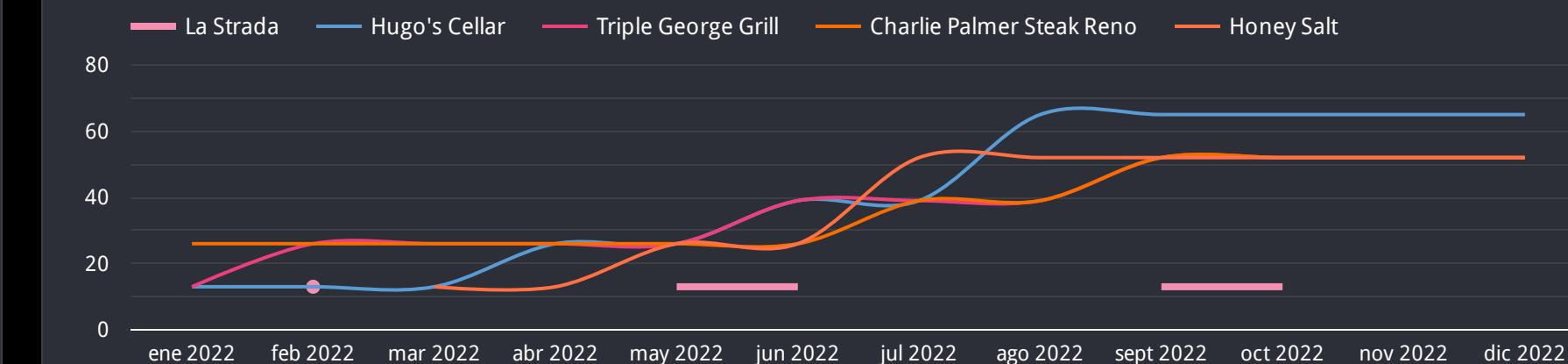


Total de reseñas  
15.259

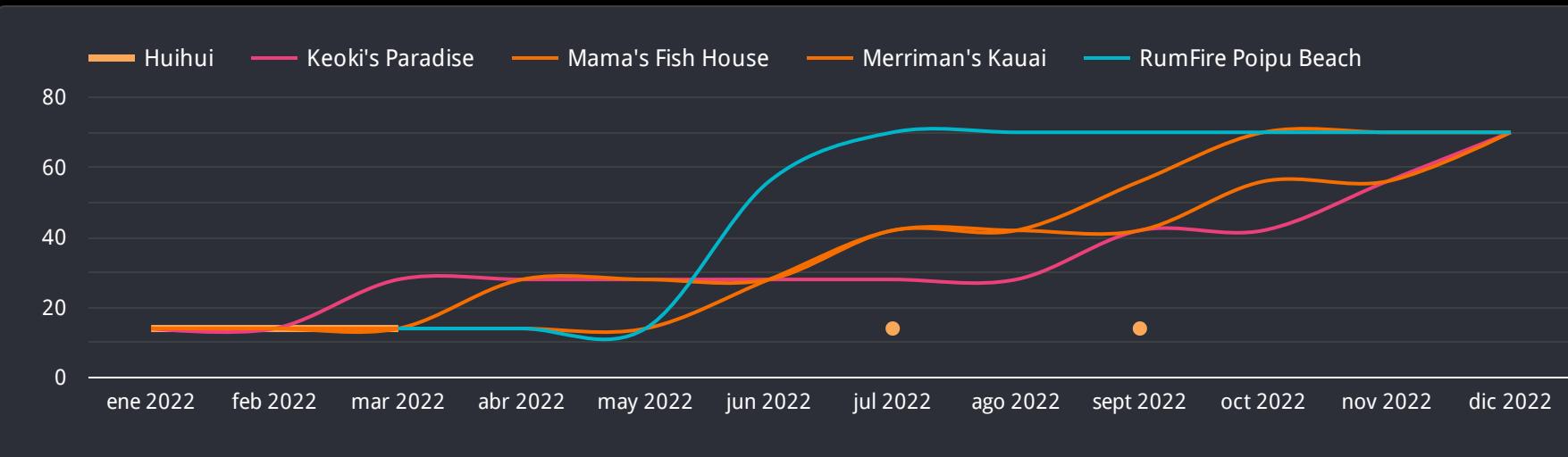
## TODO - Número de reseñas por restaurante



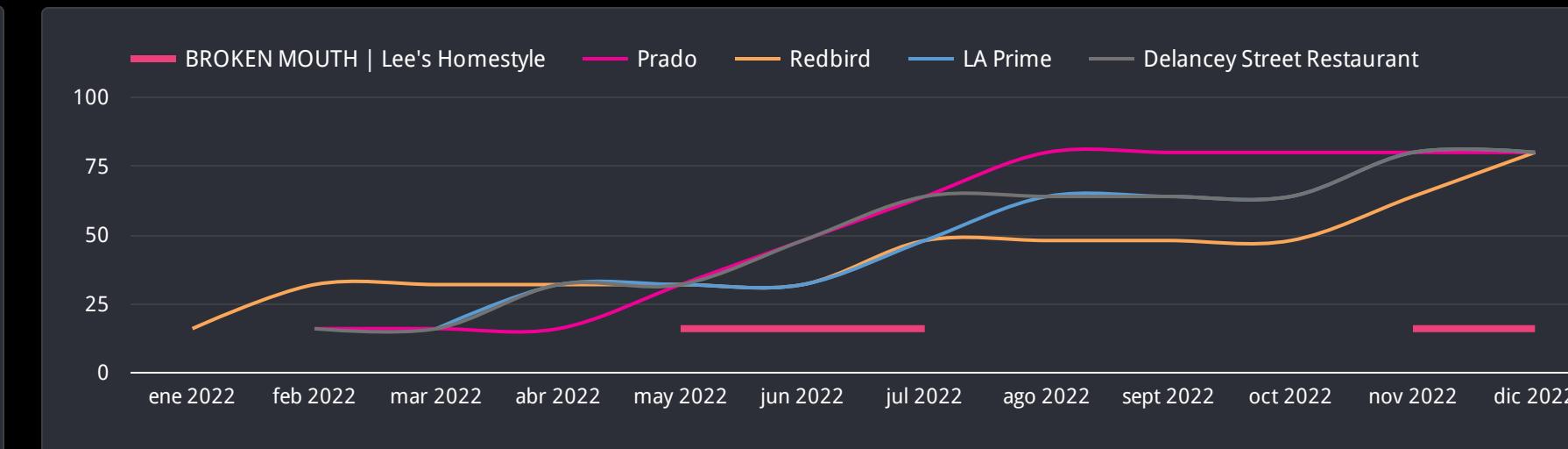
## NV - Número de reseñas por restaurante



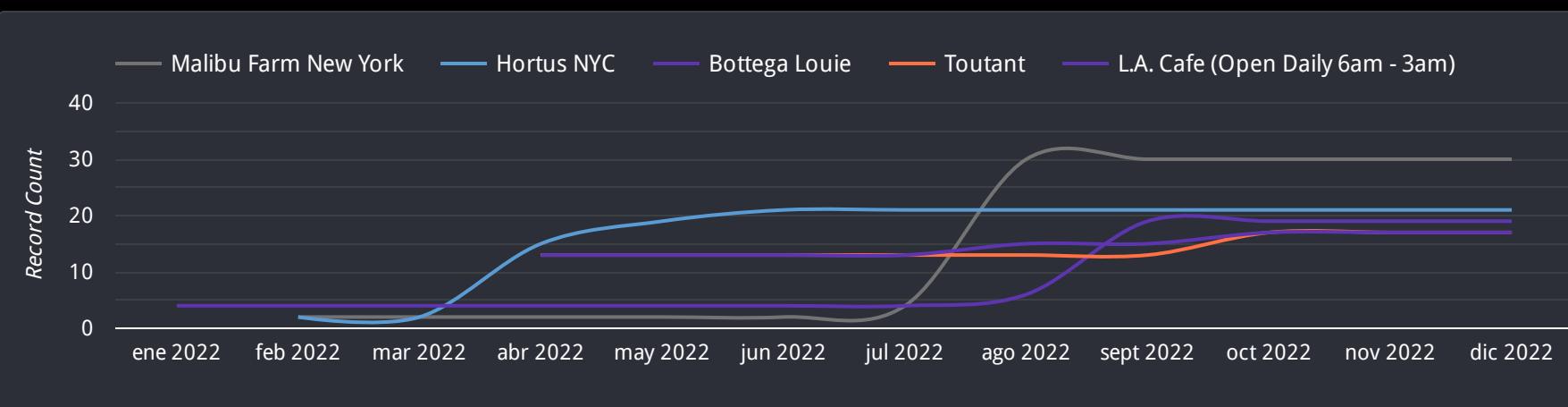
## HI - Número de reseñas por restaurante



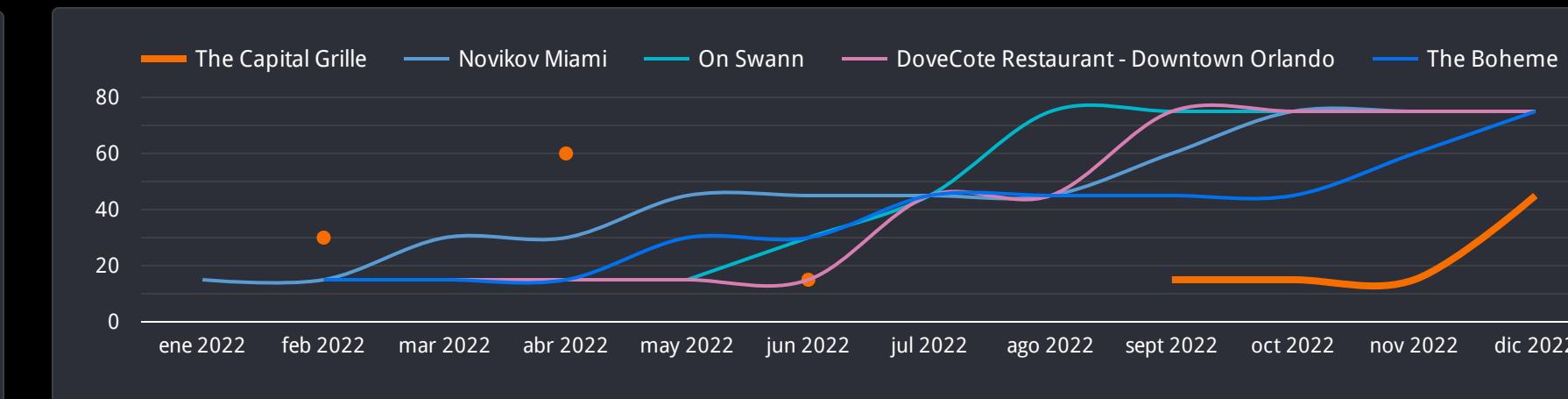
## CA - Número de reseñas por restaurante



## NY - Número de reseñas por restaurante



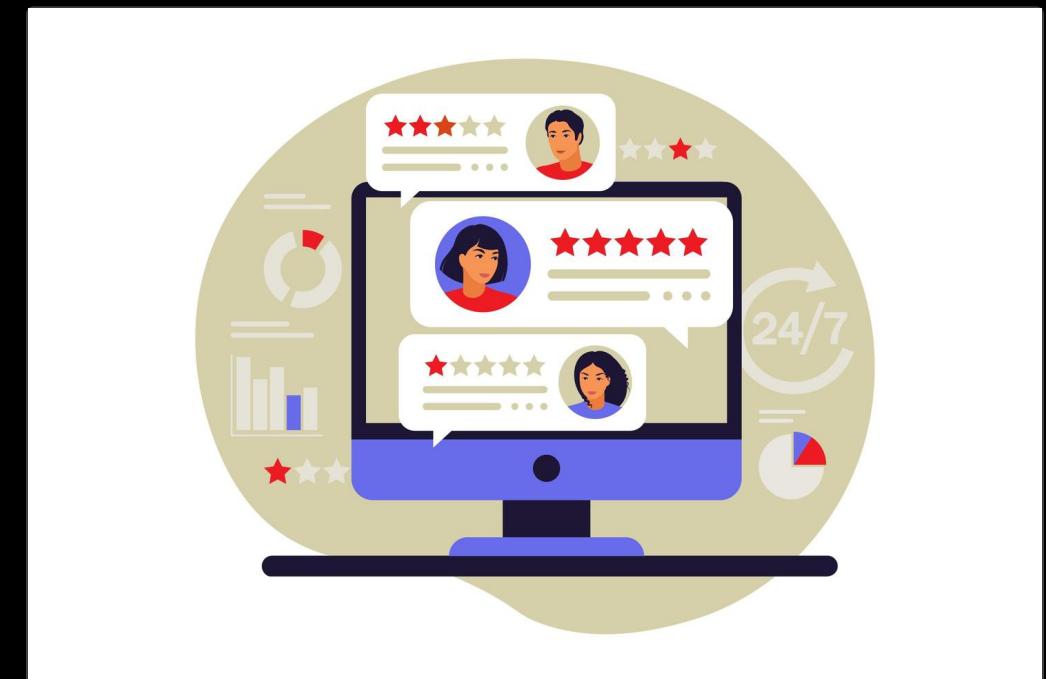
## FL - Número de reseñas por restaurante



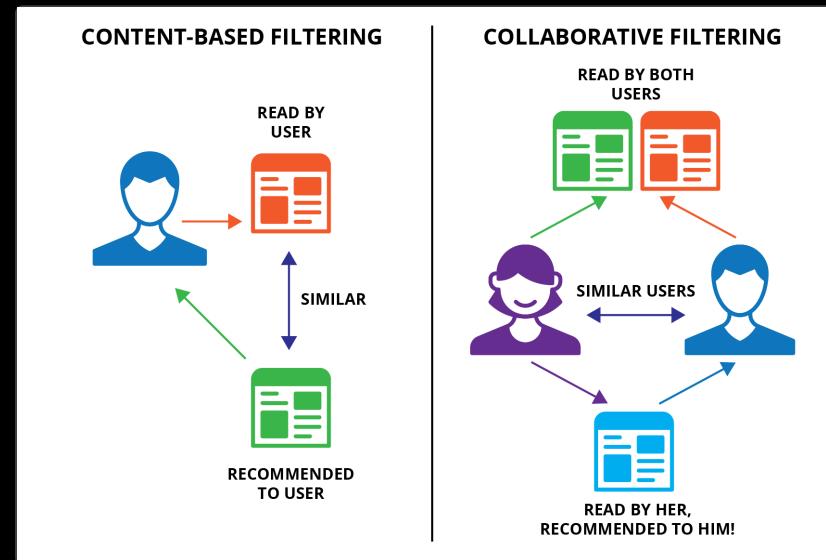
- \* Le recomendamos invertir en el sector de los restaurantes, la relación entre las 5 principales categorías y que 4 de ellas estén relacionadas con el sector de restaurantes; indica el interés de los clientes por la comida, los diversos platos y las experiencias gastronómicas. Además el constante aporte de los clientes le permite una retroalimentación constante para la mejora que ningún otro sector le ofrece.
- \* El cliente tiene un sentimiento positivo sobre el sector más allá de la categoría o estado, lo que es un buen indicador para la inversión dado que demuestra el interés por el desarrollo del mismo. El objetivo de su empresa debe centrarse en mantener este sentimiento positivo para que permanezca el interés por su negocio y le permita a futuro una expansión.
- \* Los estados considerados son grandes alternativas para establecer su primer local o una futura expansión dado el constante flujo de turistas nacionales como extranjero; lo que le permite un reconocimiento voz a voz que le facilite el crecimiento empresarial.

## Modelo de recomendación

Se propuso un modelo de recomendación para nuestro cliente, basado en los datos de Google, Yelp y Maps para encontrar la mejor opción para abrir una nueva cadena de restaurantes. Para esto, se utilizó un análisis de sentimientos en los comentarios de los usuarios para tener una visión más precisa de la perspectiva de los clientes potenciales para la nueva cadena de restaurantes. De esta manera, nuestro modelo toma como valor de ingreso el estado que se desee y arroja las ciudades y el tipo de restaurantes que podrían tener más éxito en estos lugares.



Este modelo utiliza técnicas avanzadas de análisis de datos para identificar patrones y tendencias en la información recopilada. Al combinar esta información con el análisis de sentimientos, podemos obtener una imagen más completa y precisa de las preferencias y necesidades de los clientes potenciales. Esto nos permite hacer recomendaciones más precisas y efectivas para la apertura de nuevos restaurantes.



stars_review	review	sentiment_review	city	state
3.0	Had a party of 6 here for hibachi. Our waitres...	neutral	Santa Barbara	California
1.0	Sat down over a half hour, only miso and drink...	negativo	Santa Barbara	California
3.0	Not bad. Sushi roll was good but options were ...	neutral	Santa Barbara	California
5.0	I'm very picky when it comes to Sushi. Quality...	positivo	Santa Barbara	California

## Conclusiones del Modelo de Recomendación

\*En resumen, el código que hemos presentado nos permite descubrir las mejores categorías en cada ciudad de un estado determinado, basándonos en las reseñas y el sentimiento promedio asociado a cada categoría. Estos resultados pueden ser de gran utilidad para diversos propósitos, como la planificación de negocios, la personalización de servicios o la toma de decisiones estratégicas. Al comprender las preferencias de las personas en cada ubicación, podemos adaptar nuestras ofertas y estrategias para satisfacer mejor las necesidades del mercado local.

\*Es importante destacar que este enfoque se basa en el análisis de datos y nos brinda información objetiva sobre las preferencias de los usuarios. Además, al utilizar Python y la biblioteca Pandas, podemos manejar grandes volúmenes de datos de manera eficiente y obtener resultados precisos.

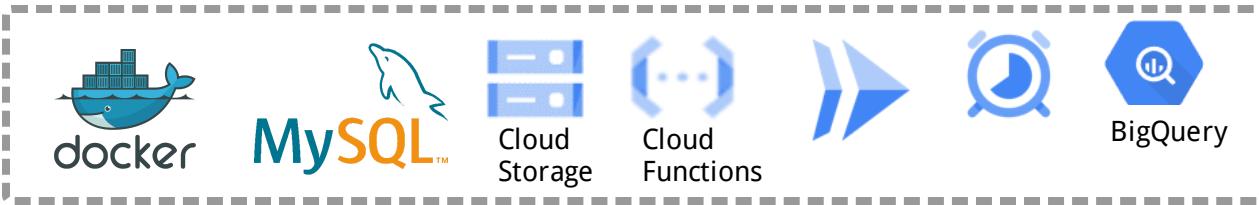
# STACK TECNOLÓGICO

Data Analytics



Looker Studio

Data Engineering



Machine Learning



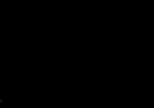
Collaborative Tools



Otros



DATACKET			Gantt Chart - Proyecto de Análisis y Recomendación de Restaurantes																	
Hitos	Tareas	Responsable	S1			S2			S3			S4								
			L	M	X	J	V	S	D	L	M	X	J	V	S	D	L	M	X	J
Propuesta del proyecto	Crear diagrama de Gannt	Jesús																		
	Distribución de tareas	Todos																		
	Planteamiento del problema	Ada																		
	Definir los objetivos y la propuesta	Todos																		
	Entendimiento de la situación	Keily/Oriana																		
	Definición de KPIs	Jesús/Oriana																		
	Diccionario de datos	Oriana																		
	Archivo presentación	José/Keily																		
	Commit en GitHub	Ada																		
	Forma de organización del proyecto	José																		
	Requerimientos semanales	Todos																		
	Análisis preliminar	José																		
	Stack tecnológico	Ada/Keily																		
	Diseño de presentación para propuesta	Ada/Jesús																		
	DEMO	Todos																		
Data Engineering	Modelo de ER	Ada / Jose																		
	Carga incremental	Oriana/Jesús/Keily																		
	ETL	Ada/Kelly																		
	Automatización Data Warehouse	Ada																		
	Validación de datos	Jesús/Oriana																		
	Documentación - Diccionario de datos	Jose / Oreana																		
	Diseño de presentación para propuesta	Ada/Keily																		
	Conexión con APIs	Ada/Keily/Oriana																		
	DEMO	Todos																		
Analytics y Machine Learning	Diseño de Dashboards	Oriana/Keily																		
	Visualización de KPIs	Jesús/Oriana																		
	Modelado de ML	Ada / José/Jesús																		
	Testeo y evaluación del modelo	José/Jesús																		
	Performance del modelo	Ada / José																		
	Integración de la propuesta de proyecto	Todos																		
	Documentación	Ada / Oriana																		
	DEMO	Todos																		

TÍTULO DEL PROYECTO:			Análisis y Sistema de Recomendación de Restaurantes basado en opiniones de los usuarios de Yelp y Google Maps en Estados Unidos												
RESPONSABLE DEL PROYECTO:			Grupo 6 Data Science												
NOMBRE DEL PO:			Gonzalo Posse												
<b>DATACKET</b> 															
Propuesta del proyecto	Hitos	Tareas	S1 S2 S3 S4												
	L M X J V S D L M X J V S D L M X J V S D L M X J V												S1 S2 S3 S4		
	L M X J V S D L M X J V S D L M X J V S D L M X J V												S1 S2 S3 S4		
	L M X J V S D L M X J V S D L M X J V S D L M X J V												S1 S2 S3 S4		
	L M X J V S D L M X J V S D L M X J V S D L M X J V												S1 S2 S3 S4		
	L M X J V S D L M X J V S D L M X J V S D L M X J V												S1 S2 S3 S4		
	L M X J V S D L M X J V S D L M X J V S D L M X J V												S1 S2 S3 S4		
	L M X J V S D L M X J V S D L M X J V S D L M X J V												S1 S2 S3 S4		
	L M X J V S D L M X J V S D L M X J V S D L M X J V												S1 S2 S3 S4		
	L M X J V S D L M X J V S D L M X J V S D L M X J V												S1 S2 S3 S4		
	L M X J V S D L M X J V S D L M X J V S D L M X J V												S1 S2 S3 S4		
	L M X J V S D L M X J V S D L M X J V S D L M X J V												S1 S2 S3 S4		
	L M X J V S D L M X J V S D L M X J V S D L M X J V												S1 S2 S3 S4		
	L M X J V S D L M X J V S D L M X J V S D L M X J V												S1 S2 S3 S4		
	L M X J V S D L M X J V S D L M X J V S D L M														