

Diccionario de datos

1. Data Fuente

1.1 business.json

- **'business_id'**: String, hace referencia al negocio.
- **'name'**: String, nombre del negocio.
- **'address'**: String, dirección completa del negocio.
- **'city'**: String, ciudad.
- **'state'**: String, código del estado donde se encuentra la ciudad.
- **'postal_code'**: String, código postal.
- **'latitude'**: Float, Latitud.
- **'longitude'**: Float, Longitud.
- **'stars'**: Float, calificación en estrellas.
- **'review_count'**: Integer, cantidad de reseñas.
- **'attributes'**: Object, Atributos del negocio. Algunos atributos tambien pueden ser objetos.
- **'categories'**: List, Categorías del negocio.

1.2 review.json

- **'review_id'**: String, referencia a la reseña.
- **'user_id'**: String, id unico con referencia al usuario.
- **'business_id'**: String, refiere al negocio.
- **'stars'**: Integer, puntaje en estrellas del 1 al 5.
- **'text'**: String, reseña.
- **'date'**: Date, hace referencia a la fecha de la reseña.

1.3 user.json

- **'user_id'**: String, referencia al usuario.
- **'name'**: String, nombre del usuario.
- **'review_count'**: Integer, cantidad de reseñas
- **'yelping_since'**: Date, fecha de creación del usuario en Yelp.
- **'useful'**: Integer, cantidad de votos marcados por el usuario como utiles.
- **'funny'**: Integer, cantidad de votos marcados por el usuario como graciosos.
- **'cool'**: Integer, cantidad de votos marcados por el usuario como agradables.
- **'elite'**: List, Lista de años que el usuario fue catalogado como elite.
- **'friends'**: String, otros 'user_id' relacionados al usuario principal.
- **'fans'**: Integer, cantidad de reacciones de otros usuarios ante las reviews publicadas.
- **'average_stars'**: float, promedio del valor de las reseñas.

1.4 checkin.json

- **'business_id'**: String, refiere al negocio.
- **'date_id'**: String, fechas separadas por coma sobre las interacciones de usuarios del negocio con Yelp.

1.5 tip.json

- **'user_id'**: String, refiere al usuario.
- **'business_id'**: String, refiere al negocio.
- **'text'**: String, frases claves respecto al negocio.
- **'date'**: Date, fecha de creacion de la frase.

2. Extra Data

2.1 USA_total_population.csv

- **'City'**: String, nombre de la ciudad.
- **State**: String, estados del país.
- **Total_Population**: Integer, cantidad de personas por ciudad.

2.2 U.S.A_household_media_income.csv

- **'State'**: String, estados del país.
- **'2021', '2020', '2019', '2018', '2017', '2016', '2015', '2014', '2013', '2012', '2011', '2010', '2009', '2008', '2007', '2006', '2005'**: float, registran el ingreso medio de las familias por estado.