**Diccionario de datos**

* **Google Maps**

**metadata\_sitios**

La carpeta tiene 11 archivos .json donde se dispone la metadata contiene información del comercio, incluyendo localización, atributos y categorías.

// 'name': 'Walgreens Pharmacy',

// 'address': 'Walgreens Pharmacy, 124 E North St, Kendallville, IN 46755',

// 'gmap\_id': '0x881614ce7c13acbb:0x5c7b18bbf6ec4f7e',

// 'description': 'Department of the Walgreens chain providing prescription medications & other health-related items.',

// 'latitude': 41.451859999999996,

// 'longitude': -85.2666757,

// 'category': ['Pharmacy'],

// 'avg\_rating': 4.2,

// 'num\_of\_reviews': 5,

// 'price': '$$',

// 'hours': [['Thursday', '8AM–1:30PM'], ['Friday', '8AM–1:30PM'], ['Saturday', '9AM–1:30PM'], ['Sunday', '10AM–1:30PM'], ['Monday', '8AM–1:30PM'], ['Tuesday', '8AM–1:30PM'], ['Wednesday', '8AM–1:30PM']],

// 'MISC': {

'Service options': ['Curbside pickup', 'Drive-through', 'In-store pickup', 'In-store shopping'],

'Health & safety': ['Mask required', 'Staff wear masks', 'Staff get temperature checks'],

'Accessibility': ['Wheelchair accessible entrance', 'Wheelchair accessible parking lot'],

'Planning': ['Quick visit'],

'Payments': ['Checks', 'Debit cards']

},

// 'state': 'Closes soon ⋅ 1:30PM ⋅ Reopens 2PM',

// 'relative\_results': ['0x881614cd49e4fa33:0x2d507c24ff4f1c74', '0x8816145bf5141c89:0x535c1d605109f94b', '0x881614cda24cc591:0xca426e3a9b826432', '0x88162894d98b91ef:0xd139b34de70d3e03', '0x881615400b5e57f9:0xc56d17dbe420a67f'],

// 'url': 'https://www.google.com/maps/place//data=!4m2!3m1!1s0x881614ce7c13acb

b:0x5c7b18bbf6ec4f7e?authuser=-1&hl=en&gl=us'

**review-estados**

Los archivos donde se disponibiliza las reviews de los usuarios (51 carpetas, 1 por cada estado de USA, con varios archivos.json cada uno) se conforman de la siguiente manera

// 'user\_id': '101463350189962023774',

// 'name': 'Jordan Adams',

// 'time': 1627750414677,

// 'rating': 5,

// 'text': 'Cool place, great people, awesome dentist!',

// 'pics': [

{

'url': ['https://lh5.googleusercontent.com/p/AF1QipNq2nZC5TH4\_M7h5xRAd

61hoTgvY1o9lozABguI=w150-h150-k-no-p']

}

],

// 'resp': {

'time': 1628455067818,

'text': 'Thank you for your five-star review! -Dr. Blake'

},

// 'gmap\_id': '0x87ec2394c2cd9d2d:0xd1119cfbee0da6f3'

* **Dataset de Yelp!**

**business.pkl**

Contiene información del comercio, incluyendo localización, atributos y categorías.

// string, 22 caracteres id del negocio, refiere al negocio en business. json

"business\_id": "tnhfDv5Il8EaGSXZGiuQGg",

// string, nombre del negocio

"name": "Garaje",

// string, direccion completa del negocio

"address": "475 3rd St",

// string, ciudad

"city": "San Francisco",

// string, codigo de 2 letras del Estado donde se ubica el negocio

"state": "CA",

// string, el codigo postal

"postal code": "94107",

// float, latitud

"latitude": 37.7817529521,

// float, longitud

"longitude": -122.39612197,

// float, rating en estrellas, redondeado a 0 o 0.5

"stars": 4.5,

// entero, numero de reseñas

"review\_count": 1198,

// entero, 0 si esta cerrado, 1 si está abierto

"is\_open": 1,

// objeto, atributos del negocio como valores. Algunos valores de atributos también pueden ser objetos.

"attributes": {

"RestaurantsTakeOut": true,

"BusinessParking": {

"garage": false,

"street": true,

"validated": false,

"lot": false,

"valet": false

},

},

// lista de categorias de los negocios

"categories": [

"Mexican",

"Burgers",

"Gastropubs"

],

// objeto, de dia a hora, las horas son en 24hr

"hours": {

"Monday": "10:00-21:00",

"Tuesday": "10:00-21:00",

"Friday": "10:00-21:00",

"Wednesday": "10:00-21:00",

"Thursday": "10:00-21:00",

"Sunday": "11:00-18:00",

"Saturday": "10:00-21:00"

**review.json**

Contiene las reseñas completas, incluyendo el user\_id que escribió el review y el business\_id por el cual se escribe la reseña

// string, 22 caracteres id de reseña

"review\_id": "zdSx\_SD6obEhz9VrW9uAWA",

// string, 22 caracteres id único de usuario, refiere al usuario en user.json

"user\_id": "Ha3iJu77CxlrFm-vQRs\_8g",

// string, 22 caracteres id del negocio, refiere al negocio en business.json

"business\_id": "tnhfDv5Il8EaGSXZGiuQGg",

// entero, puntaje en estrellas de 1 al 5

"stars": 4,

// string, fecha formato YYYY-MM-DD

"date": "2016-03-09",

// string, la reseña en inglés

"text": "Great place to hang out after work: the prices are decent, and the ambience is fun. It's a bit loud, but very lively. The staff is friendly, and the food is good. They have a good selection of drinks.",

// entero, números de votos como reseña útil

"useful": 0,

// entero, número de votos como reseña graciosa

"funny": 0,

// entero, número de votos como reseña cool.

"cool": 0

**user.parquet**

Data del usuario incluyendo referencias a otros usuarios amigos y a toda la metadata asociada al usuario.

// string, 22 caracteres, id de usuario que refiere al usuario en user.json

"user\_id": "Ha3iJu77CxlrFm-vQRs\_8g",

// string, nombre del usuario

"name": "Sebastien",

// entero, numero de reseñas escritas

"review\_count": 56,

// string, fecha de creacion del usuario en Yelp en formato YYYY-MM-DD

"yelping\_since": "2011-01-01",

// lista con los id de usuarios que son amigos de ese usuario

"friends": [

"wqoXYLWmpkEH0YvTmHBsJQ",

"KUXLLiJGrjtSsapmxmpvTA",

"6e9rJKQC3n0RSKyHLViL-Q"

],

// entero, número de votos marcados como útiles por el usuario

"useful": 21,

// entero, número de votos marcados como graciosos por el usuario

"funny": 88,

// entero, número de votos marcados como cool por el usuario

"cool": 15,

// entero, número de fans que tiene el usuario

"fans": 1032,

// lista de enteros, años en los que el usuario fue miembro elite

"elite": [

2012,

2013

],

// float, promedio del valor de las reseñas

"average\_stars": 4.31,

// entero, total de cumplidos 'hot' recibidos por el usuario

"compliment\_hot": 339,

// entero, total de cumplidos varios recibidos por el usuario

"compliment\_more": 668,

// entero, total de cumplidos por el perfil recibidos por el usuario

"compliment\_profile": 42,

// entero, total de cumplidos 'cute' recibidos por el usuario

"compliment\_cute": 62,

// entero, total de listas de cumplidos recibidos por el usuario

"compliment\_list": 37,

// entero, total de cumplidos como notas recibidos por el usuario

"compliment\_note": 356,

// entero, total de cumplidos planos recibidos por el usuario

"compliment\_plain": 68,

// entero, total de cumplidos 'cool' recibidos por el usuario

"compliment\_cool": 91,

// entero, total de cumplidos graciosos recibidos por el usuario

"compliment\_funny": 99,

// entero, número de complidos escritos recibidos por el usuario

"compliment\_writer": 95,

// entero, número de cumplidos en foto recibidos por el usuario

"compliment\_photos": 50

**checkin.json**

Registros en el negocio.

// string, 22 caracteres id del negocio que se refiere al negocio en business.json

"business\_id": "tnhfDv5Il8EaGSXZGiuQGg"

// string que es una lista de fechas separados por coma, en formato YYYY-MM-DD HH:MM:SS

"date": "2016-04-26 19:49:16, 2016-08-30 18:36:57, 2016-10-15 02:45:18, 2016-11-18 01:54:50, 2017-04-20 18:39:06, 2017-05-03 17:58:02"

**tip.json**

Tips (consejos) escritos por el usuario. Los tips son más cortas que las reseñas y tienden a dar sugerencias rápidas.

// string, texto del tip

"text": "Secret menu - fried chicken sando is da bombbbbbb Their zapatos are good too.",

// string, fecha cuando se escribio el tip YYYY-MM-DD

"date": "2013-09-20",

// entero, cuantos cumplidos totales tiene

"compliment\_count": 172,

// string, 22 caracteres, id del negocio que se refiere al negocio en business.json

"business\_id": "tnhfDv5Il8EaGSXZGiuQGg",

// string, 22 caracteres de id de usuario, que se refieren al usuario en user.json

"user\_id": "49JhAJh8vSQ-vM4Aourl0g"