**Notes on the Dataset**

Each file is composed of a single object type, one json-object per-line.  
Take a look at some examples to get you started: <https://github.com/Yelp/dataset-examples>.

**business**

{

'type': 'business',

'business\_id': (encrypted business id),

'name': (business name),

'neighborhoods': [(hood names)],

'full\_address': (localized address),

'city': (city),

'state': (state),

'latitude': latitude,

'longitude': longitude,

'stars': (star rating, rounded to half-stars),

'review\_count': review count,

'categories': [(localized category names)]

'open': True / False (corresponds to closed, not business hours),

'hours': {

(day\_of\_week): {

'open': (HH:MM),

'close': (HH:MM)

},

...

},

'attributes': {

(attribute\_name): (attribute\_value),

...

},

}

**review**

{

'type': 'review',

'business\_id': (encrypted business id),

'user\_id': (encrypted user id),

'stars': (star rating, rounded to half-stars),

'text': (review text),

'date': (date, formatted like '2012-03-14'),

'votes': {(vote type): (count)},

}

**user**

{

'type': 'user',

'user\_id': (encrypted user id),

'name': (first name),

'review\_count': (review count),

'average\_stars': (floating point average, like 4.31),

'votes': {(vote type): (count)},

'friends': [(friend user\_ids)],

'elite': [(years\_elite)],

'yelping\_since': (date, formatted like '2012-03'),

'compliments': {

(compliment\_type): (num\_compliments\_of\_this\_type),

...

},

'fans': (num\_fans),

}

**check-in**

{

'type': 'checkin',

'business\_id': (encrypted business id),

'checkin\_info': {

'0-0': (number of checkins from 00:00 to 01:00 on all Sundays),

'1-0': (number of checkins from 01:00 to 02:00 on all Sundays),

...

'14-4': (number of checkins from 14:00 to 15:00 on all Thursdays),

...

'23-6': (number of checkins from 23:00 to 00:00 on all Saturdays)

}, # if there was no checkin for a hour-day block it will not be in the dict

}

**tip**

{

'type': 'tip',

'text': (tip text),

'business\_id': (encrypted business id),

'user\_id': (encrypted user id),

'date': (date, formatted like '2012-03-14'),

'likes': (count),

}