



# Yelp Dataset

Welcome to the Yelp treasure chest of rich research data!

Warning: links will expire in 30 seconds.

Warning: The photo dataset is a single 5.9GB tar file, so it might take some time to download.

If you are a student and meet our eligibility criteria, you can [submit your research](#) findings to us by **December 31, 2016** to be considered for the \$5,000 [Dataset Challenge Award](#).

## 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),
}
```

### photos (from the photos auxiliary file)

This file is formatted as a JSON list of objects.

```
[
  {
    "photo_id": (encrypted photo id),
    "business_id" : (encrypted business id),
    "caption" : (the photo caption, if any),
    "label" : (the category the photo belongs to, if any)
  },
  {...}
]
```

### About

[About Yelp](#)  
[Order Food on Eat24](#)  
[Careers](#)  
[Press](#)  
[Investor Relations](#)  
[Content Guidelines](#)  
[Terms of Service](#)

### Discover

[The Local Yelp](#)  
[Yelp Blog](#)  
[Support](#)  
[Yelp Mobile](#)  
[Developers](#)  
[RSS](#)

### Yelp for Business Owners

[Claim your Business Page](#)  
[Advertise on Yelp](#)  
[Online Ordering from Eat24](#)  
[Yelp Reservations](#)  
[Business Success Stories](#)  
[Business Support](#)  
[Yelp Blog for Business Owners](#)

### Languages

[English ▾](#)

### Countries

[United States ▾](#)

---

Site Map [Atlanta](#) | [Austin](#) | [Boston](#) | [Chicago](#) | [Dallas](#) | [Denver](#) | [Detroit](#) | [Honolulu](#) | [Houston](#) | [Los Angeles](#) | [Miami](#) | [Minneapolis](#) | [New York](#) | [Philadelphia](#)  
| [Portland](#) | [Sacramento](#) | [San Diego](#) | [San Francisco](#) | [San Jose](#) | [Seattle](#) | [Washington, DC](#) | [More Cities](#)

Copyright © 2004–2016 Yelp Inc. Yelp, **yelp** ,  and related marks are registered trademarks of Yelp.