

# Yelp Dataset Documentation & Terms

Use of the Yelp Dataset is governed by the terms and conditions outlined below. By using the Yelp Dataset, you expressly agree to these terms and conditions.

## Contents:

- **Yelp Dataset JSON Documentation**
- **Yelp Dataset Photos Documentation**
- **Terms and Conditions of Use**

## Yelp Dataset JSON

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>.

*Note: the follow examples contain inline comments, which are technically not valid JSON. This is done here to simplify the documentation and explaining the structure, the JSON files you download will not contain any comments and will be fully valid JSON.*

### **business.json**

Contains business data including location data, attributes, and categories.

```
{
  // string, 22 character unique string business id
  "business_id": "tnhfDv5Il8EaGSXZGiuQGg",

  // string, the business's name
  "name": "Garaje",

  // string, the full address of the business
  "address": "475 3rd St",

  // string, the city
  "city": "San Francisco",

  // string, 2 character state code, if applicable
  "state": "CA",

  // string, the postal code
  "postal code": "94107",

  // float, latitude
  "latitude": 37.7817529521,
```

```

    // float, longitude
    "longitude": -122.39612197,

    // float, star rating, rounded to half-stars
    "stars": 4.5,

    // integer, number of reviews
    "review_count": 1198,

    // integer, 0 or 1 for closed or open, respectively
    "is_open": 1,

    // object, business attributes to values. note: some attribute values
might be objects
    "attributes": {
      "RestaurantsTakeOut": true,
      "BusinessParking": {
        "garage": false,
        "street": true,
        "validated": false,
        "lot": false,
        "valet": false
      },
    },

    // an array of strings of business categories
    "categories": [
      "Mexican",
      "Burgers",
      "Gastropubs"
    ],

    // an object of key day to value hours, hours are using a 24hr clock
    "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

Contains full review text data including the user\_id that wrote the review and the business\_id the review is written for.

```

{
  // string, 22 character unique review id
  "review_id": "zdSx_SD6obEhz9VrW9uAWA",

```

```

// string, 22 character unique user id, maps to the user in user.json
"user_id": "Ha3iJu77CxlRfm-vQRs_8g",

// string, 22 character business id, maps to business in business.json
"business_id": "tnhfDv5Il8EaGSXZGiuQGg",

// integer, star rating
"stars": 4,

// string, date formatted YYYY-MM-DD
"date": "2016-03-09",

// string, the review itself
"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.",

// integer, number of useful votes received
"useful": 0,

// integer, number of funny votes received
"funny": 0,

// integer, number of cool votes received
"cool": 0
}

```

### user.json

User data including the user's friend mapping and all the metadata associated with the user.

```

{
  // string, 22 character unique user id, maps to the user in user.json
  "user_id": "Ha3iJu77CxlRfm-vQRs_8g",

  // string, the user's first name
  "name": "Sebastien",

  // integer, the number of reviews they've written
  "review_count": 56,

  // string, when the user joined Yelp, formatted like YYYY-MM-DD
  "yelping_since": "2011-01-01",

  // array of strings, an array of the user's friend as user_ids
  "friends": [
    "wqoXYLWmpkEH0YvTmHBsJQ",
    "KUXLLiJGrjtSsapmxmpvTA",
    "6e9rJKQC3n0RSKyHLViL-Q"
  ],

  // integer, number of useful votes sent by the user

```

```
"useful": 21,  
  
// integer, number of funny votes sent by the user  
"funny": 88,  
  
// integer, number of cool votes sent by the user  
"cool": 15,  
  
// integer, number of fans the user has  
"fans": 1032,  
  
// array of integers, the years the user was elite  
"elite": [  
2012,  
2013  
],  
  
// float, average rating of all reviews  
"average_stars": 4.31,  
  
// integer, number of hot compliments received by the user  
"compliment_hot": 339,  
  
// integer, number of more compliments received by the user  
"compliment_more": 668,  
  
// integer, number of profile compliments received by the user  
"compliment_profile": 42,  
  
// integer, number of cute compliments received by the user  
"compliment_cute": 62,  
  
// integer, number of list compliments received by the user  
"compliment_list": 37,  
  
// integer, number of note compliments received by the user  
"compliment_note": 356,  
  
// integer, number of plain compliments received by the user  
"compliment_plain": 68,  
  
// integer, number of cool compliments received by the user  
"compliment_cool": 91,  
  
// integer, number of funny compliments received by the user  
"compliment_funny": 99,  
  
// integer, number of writer compliments received by the user  
"compliment_writer": 95,
```

```
// integer, number of photo compliments received by the user
"compliment_photos": 50
}
```

### checkin.json

Checkins on a business.

```
{
  // string, 22 character business id, maps to business in business.json
  "business_id": "tnhfDv5Il8EaGSXZGiuQGg"

  // string which is a comma-separated list of timestamps for each checkin,
  // each with format 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 written by a user on a business. Tips are shorter than reviews and tend to convey quick suggestions.

```
{
  // string, text of the tip
  "text": "Secret menu - fried chicken sando is da bombbbbbbb Their zapatos are good too.",

  // string, when the tip was written, formatted like YYYY-MM-DD
  "date": "2013-09-20",

  // integer, how many compliments it has
  "compliment_count": 172,

  // string, 22 character business id, maps to business in business.json
  "business_id": "tnhfDv5Il8EaGSXZGiuQGg",

  // string, 22 character unique user id, maps to the user in user.json
  "user_id": "49JhAJh8vSQ-vM4Aourl0g"
}
```

### photo.json

Contains photo data including the caption and classification (one of "food", "drink", "menu", "inside" or "outside").

```
{
  // string, 22 character unique photo id
  "photo_id": "_nN_DhLXkfwEkwPNxne9hw",
  // string, 22 character business id, maps to business in business.json
  "business_id": "tnhfDv5Il8EaGSXZGiuQGg",
  // string, the photo caption, if any
  "caption": "carne asada fries",
  // string, the category the photo belongs to, if any
  "label": "food"
}
```

# Yelp Dataset Photos

## photo.json

Duplicate of photo.json from the [main dataset](#).

```
{
  // string, 22 character unique photo id
  "photo_id": "_nN_DhLXkfwEkwPNxne9hw",

  // string, 22 character business id, maps to business in business.json
  "business_id": "tnhfDv5Il8EaGSXZGiuQGg",

  // string, the photo caption, if any
  "caption": "carne asada fries",

  // string, the category the photo belongs to, if any
  "label": "food"
}
```

## photos/

Directory containing the photos themselves as "{photo\_id}.jpg".

## YELP DATASET TERMS OF USE

Last Updated: July 7, 2023 This document (“Data Agreement”) governs the terms under which you may access and use the data that Yelp makes available for download through this website (or made available by other means) solely for academic purposes (the “Data”).

### Terms:

By accessing or using the Data, you agree to be bound by the Data Agreement and represent that the contact information you provide to Yelp is correct. If you access or use the Data on behalf of a university, school, or other entity, you represent that you have authority to bind such entity and its affiliates to the Data Agreement and that it is fully binding upon them. In such a case, the term “you” and “your” will refer to such an entity and its affiliates. If you do not have authority, or if you do not agree with the terms of the Data Agreement, you may not access or use the Data. You should read and keep a copy of each component of the Data Agreement for your records. In the event of a conflict among them, the terms of this document will control.

1. Purpose The Data is made available by Yelp Inc. (“Yelp”) to enable you to access valuable local information to develop an academic project as part of an ongoing course of study. Yelp reserves the right to continually review and evaluate all uses of the Data provided under the Data Agreement. Notwithstanding the foregoing, for European Union Data, the Data is made available and licensed to you solely for academic research purposes.

2. Changes Yelp reserves the right to modify or revise the Data Agreement at any time. If the change is deemed to be material and it is foreseeable that such change could be adverse to your interests, Yelp will provide you notice of the change to this Data Agreement by sending you an email to the email you provided to Yelp. Your continued use of the Data after the notice of material change will constitute your acceptance of and agreement to such changes. IF YOU DO NOT WISH TO BE BOUND TO ANY NEW TERMS, YOU MUST TERMINATE THE DATA AGREEMENT BY IMMEDIATELY CEASING USE OF THE DATA AND DELETING IT FROM ANY SYSTEMS OR MEDIA.

3. License Subject to the terms set forth in the Data Agreement (specifically the restrictions set forth in Section 4 below), Yelp grants you a royalty-free, non-exclusive, revocable, non-sublicensable, non-transferable, fully paid-up right and license during the Term to use, access, and create derivative works of the Data in electronic form for solely for academic use. Academic use means use of the Data by registered nonprofits, government, educational institutions, and think tanks which (a) is not undertaken for profit, or (b) is not intended to produce works, services, or data for commercial use. You may not use the Data for any other purpose without Yelp’s prior written consent. You acknowledge and agree that Yelp may request information about, review, audit, and/or monitor your use of the Data at any time in order to confirm compliance with the Data Agreement. Nothing herein shall be construed as a license to use Yelp’s registered trademarks or service marks, or any other Yelp branding. Prior to any public presentation or publication of the academic results or conclusions that involve the Data and/or the Yelp brand name, you must submit your findings to Yelp for review and approval, and Yelp will approve of the public release within five (5) business days of its submission to Yelp.

4. Restrictions You agree that you will not, and will not encourage, assist, or enable others to: A. display, perform, or distribute any of the Data, or use the Data to update or create your own business listing information for commercial purposes (i.e. you may not publicly display any of the Data to any third party, especially reviews and other user generated content, as this is a private data set challenge and not a license to compete with or disparage with Yelp); B. use the Data in connection with any commercial purpose; C. use the Data in any manner or for any purpose that may violate any law or regulation, or any right of any person including, but not limited to, intellectual property rights, rights of privacy and/or rights of personality, or which otherwise may be harmful (in Yelp's sole discretion) to Yelp, its providers, its suppliers, end users of this website, or your end users; D. share or make available the Data to any third party; E. create, redistribute or disclose any summary of, or metrics related to, the Data (e.g., the number of reviewed business included in the Data and other statistical analysis) to any third party or on any website or other electronic media not expressly covered by this Agreement or without Yelp's prior consent, this provision however, excludes any disclosures necessary for academic purposes, including without limitation the publication of academic articles concerning your use of the Data; F. use the Data in a manner that is competitive in nature with Yelp; G. display Data in a manner that could reasonably imply an endorsement, relationship or affiliation with or sponsorship between you or a third party and Yelp, other than your permitted use of the Data under the terms of the Data Agreement; H. rent, lease, sell, transfer, assign, or sublicense, any part of the Data; I. modify, rate, rank, review, vote or comment on, or otherwise respond to the content contained in the Data; J. display the Data or publicly communicate in any way, or on any site, in a manner that disparages Yelp or its products or services, or infringes any Yelp intellectual property or other rights; K. use the Data in a manner that could reasonably be interpreted to suggest that Yelp is the author or entity that is responsible, in whole or in part, for the creation or development of any Data or that such Data represents the views of Yelp; or L. use the Data for any purpose prohibited by law.

5. Ownership As between you and Yelp, the Data and any derivative works you create from the Data, and all intellectual property rights contained in the foregoing, are and will at all times remain the sole and exclusive property of Yelp and are protected by applicable intellectual property laws and treaties (whether those rights happen to be registered or not, and wherever in the world those rights may exist), or as otherwise set forth in the contest rules where the various submitted solutions must be made available under a specified open source license, such as the MIT License.

6. Indemnity You agree that your use of the Data is at your own risk and you agree to hold harmless, defend (subject to Yelp's right to participate with counsel it selects) and indemnify Yelp and its subsidiaries, affiliates, officers, agents, employees and suppliers from and against any and all claims, damages, liabilities, costs and fees (including reasonable attorneys' fee) arising from, or in any way related to your or your end users' use or implementation of the Data. You will not agree to any settlement that imposes any obligation on Yelp without Yelp's prior consent.

7. No Warranties by Yelp; No Entitlement to Support from Yelp. THE DATA IS PROVIDED "AS IS", "WITH ALL FAULTS" AND "AS AVAILABLE" WITHOUT WARRANTY, OF ANY KIND AND AT YOUR SOLE RISK. EXCEPT TO THE MAXIMUM EXTENT REQUIRED BY APPLICABLE LAW, YELP DISCLAIMS ALL WARRANTIES, REPRESENTATIONS, CONDITIONS, AND DUTIES, WHETHER EXPRESS, IMPLIED OR STATUTORY, REGARDING THE DATA, INCLUDING, WITHOUT



LIMITATION, ANY AND ALL IMPLIED WARRANTIES OF MERCHANTABILITY, ACCURACY, RESULTS OF USE, RELIABILITY, FITNESS FOR A PARTICULAR PURPOSE, TITLE, INTERFERENCE WITH QUIET ENJOYMENT AND NON-INFRINGEMENT OF THIRD-PARTY RIGHTS. FURTHER, YELP DISCLAIMS ANY WARRANTY THAT YOUR USE OF THE DATA WILL BE UNINTERRUPTED, SECURE, TIMELY OR ERROR FREE. FOR THE AVOIDANCE OF DOUBT, YOU ACKNOWLEDGE AND AGREE THAT THE DATA AGREEMENT DOES NOT ENTITLE YOU TO ANY SUPPORT FOR THE DATA. NO ADVICE OR INFORMATION, WHETHER ORAL OR IN WRITING, OBTAINED BY YOU FROM YELP WILL CREATE ANY WARRANTY NOT EXPRESSLY STATED IN THE DATA AGREEMENT.

8. Limitation of Liability THE DATA IS BEING PROVIDED FREE OF CHARGE. ACCORDINGLY, YOU AGREE THAT YELP SHALL HAVE NO LIABILITY ARISING FROM OR BASED ON YOUR USE OF THE DATA. REGARDLESS OF WHETHER ANY REMEDY SET FORTH HEREIN FAILS OF ITS ESSENTIAL PURPOSE OR OTHERWISE, AND EXCEPT FOR BODILY INJURY, IN NO EVENT SHALL YELP OR ITS SUBSIDIARIES, AFFILIATES, OFFICERS, AGENTS, EMPLOYEES AND SUPPLIERS BE LIABLE TO YOU OR TO ANY THIRD PARTY UNDER ANY TORT, CONTRACT, NEGLIGENCE, STRICT LIABILITY OR OTHER LEGAL OR EQUITABLE THEORY FOR ANY LOST PROFITS, LOST OR CORRUPTED DATA, COMPUTER FAILURE OR MALFUNCTION, INTERRUPTION OF BUSINESS, OR OTHER SPECIAL, INDIRECT, INCIDENTAL OR CONSEQUENTIAL DAMAGES OF ANY KIND ARISING OUT OF THE USE OR INABILITY TO USE THE DATA, EVEN IF YELP HAS BEEN ADVISED OF THE POSSIBILITY OF SUCH LOSS OR DAMAGES AND WHETHER OR NOT SUCH LOSS OR DAMAGES ARE FORESEEABLE. ANY CLAIM ARISING OUT OF OR RELATING TO THE DATA AGREEMENT MUST BE BROUGHT WITHIN (1) YEAR AFTER THE OCCURRENCE OF THE EVENT GIVING RISE TO SUCH CLAIM. IF SUCH CLAIM IS NOT FILED, THEN THAT CLAIM IS PERMANENTLY BARRED. THIS APPLIES TO YOU AND YOUR SUCCESSORS, AND TO YELP AND ITS SUCCESSORS. NOTWITHSTANDING THE FOREGOING, SINCE THIS LICENSE IS PROVIDED TO YOU AT NO CHARGE, YELP'S MAXIMUM LIABILITY UNDER THIS DATA AGREEMENT SHALL NOT, IN ANY EVENT, EXCEED US\$50.00.

9. Limited Relationship Yelp and You are, and will remain, independent contractors, and nothing in the Data Agreement will be construed as creating an employer-employee relationship, partnership or joint venture. Although you are permitted to publicize your use of the Data, you agree not to make any other statements, without the prior written consent of Yelp, implying a different kind of relationship between you and Yelp, including any implied endorsement by Yelp. You do not have any authority of any kind to bind Yelp in any respect whatsoever.

10. Term and Termination This Data Agreement is effective as of the date you download or otherwise access the Data ("Effective Date") and shall continue in full force and effect for a term of twelve (12) months from the Effective Date, unless earlier terminated by the parties or expires in accordance with this Section 10 (the "Term"). Either party may immediately terminate this Data Agreement, for any reason or for no reason, by providing written notice to the other party. Yelp will provide notice of termination to the email account you provided to Yelp during registration and termination will be effective upon delivery of the email notice. Yelp reserves the right, in its sole discretion (for any reason or for no reason) and at any time without notice to you, to change, suspend or discontinue the Data and/or suspend or terminate your further access to the Data. Any termination of the Data Agreement will also immediately terminate the licenses granted to you

hereunder. Upon any termination of the Data Agreement, you will promptly: (i) delete and remove all Data from any location, including any web pages, scripts, widgets, applications and any other software in your possession or under your control; (ii) destroy and remove from all computers, hard drives, networks and other storage media in your possession or under your control all copies of any Data; and (iii) upon Yelp's request, certify in writing to Yelp that such actions have been taken.

11. Miscellaneous The Data Agreement encompasses the entire agreement between you and Yelp regarding the subject matter discussed therein. The Data Agreement, and any disputes arising from or relating to the interpretation thereof, will be governed by and construed under the laws of the State of California without regard to its conflict of law provisions. You agree to personal jurisdiction by and venue in the state and federal courts of the State of California, City of San Francisco. The failure of Yelp to exercise or enforce any right or provision of the Data Agreement will not constitute a waiver of such right or provision. The failure of either party to exercise in any respect any right provided for herein will not be deemed a waiver of any further rights hereunder. If any provision of the Data Agreement is found to be unenforceable or invalid, that provision will be replaced with terms that most closely match the intent of the provision that is not enforceable to the minimum extent necessary so that the remaining Data Agreement will otherwise remain in full force and effect and enforceable. The Data Agreement is not assignable, transferable or sublicensable, in whole or in part, by you except with Yelp's prior written consent. Any attempt to do so is void. Yelp may assign the Data Agreement, in whole or in part, at any time with or without notice to you. The section titles in the Data Agreement are for convenience only and have no legal or contractual effect.

12. Survival Sections 4 through 13 will survive any expiration or termination of this Data Agreement for any reason. 13. Contact and Violations Please contact Yelp with any questions regarding the Data Agreement. Please report any violations of the Data Agreement [dataset@yelp.com](mailto:dataset@yelp.com)