

Codebook

Tammy Truong

2022-06-09

Variables

- **business_id**: unique business id (categorical)
- **date**: date of the review formatted in YYYY-MM-DD (categorical)
- **review_id**: unique review id (categorical)
- **stars**: star rating (categorical)
- **text**: the review itself
- **type**: type of entry (categorical)
- **user_id**: unique user id (categorical)
- **cool**: number of cool votes received (numerical)
- **useful**: number of useful votes received (numerical)
- **funny**: number of funny votes received (numerical)

Summary of the Dataset

The Yelp review dataset consists of the reviews that users post to business pages. Yelp has released this along with other open datasets for its dataset challenge. This utilizes the data from the year 2015 and in locations of the metropolitan areas centered on Montreal, Calgary, Toronto, Pittsburgh, Charlotte, Urbana-Champaign, Phoenix, Las Vegas, Madison, and Cleveland.