Data Dictionary

Google_business_data

// string, 22 characters business id, refers to the business in google_business_data
"business_id": "tnhfDv5Il8EaGSXZGiuQGg",

// string, city

"city": "San Francisco",

// string, state where the business is located

"state": "Columbus",

// string, postal code

"postal code": "94107",

// float, latitude

"latitude": 37.7817529521,

// float, longitude

"longitude": -122.39612197,

// float, star rating rounded to 0 or 0.5

"stars": 4.5,

// integer, number of reviews

"review_count": 1198,

// string, 1 letter representing the reviews from Yelp (Y) or Google (G)

"source": "G",

Google_reviews_ulta_beauty

// string, 22 characters unique user id, refers to the user

 "user_id": "Ha3iJu77CxlrFm-vQRs_8g",

// string, 22 characters business id, refers to the business in google_business_data
"business_id": "tnhfDv5Il8EaGSXZGiuQGg",

// integer, star rating from 1 to 5

 "stars": 4,

// string, the review in English

"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, month in which the review was made

"month": 4,

// integer, year in which the review was made

"year": 2019,

// integer, hour at which the review was made

"hour": 19,

// string, 1 letter representing the reviews from Yelp (Y) or Google (G)

 "source": "G",

Ulta_beauty_sentiment_analysis

// string, 22 characters unique user id, refers to the user

"user_id": "Ha3iJu77CxlrFm-vQRs_8g",

// string, 22 characters business id, refers to the business in google_business_data
"business_id": "tnhfDv5Il8EaGSXZGiuQGg",

// float, star rating rounded to 0 or 0.5

"stars": 4.5,

// string, the review in English

"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, month in which the review was made

"month": 4,

// integer, year in which the review was made

"year": 2019,

// integer, hour at which the review was made

"hour": 19,

// string, 1 letter representing the reviews from Yelp (Y) or Google (G)

"source": "G",

// integer, 1 column representing the sentiment analysis with 1=positive, 0=negative
"sentiment": 1,

// string, label for the sentiment column, 1=positive, 0=negative

"sentiment_label": "negative",

yelp_business_data

// string, 22 characters business id, refers to the business in google_business_data
"business_id": "RMD5mNJgyQ1FGO_5YDCrzA",

// string, city

"city": "San Francisco",

// string, state where the business is located

"state": "Columbus",

// string, postal code

"postal code": "94107",

// float, latitude

"latitude": 37.7817529521,

// float, longitude

 "longitude": -122.39612197,

// float, star rating rounded to 0 or 0.5

 "stars": 4.5,

// integer, number of reviews

"review_count": 1198,

// string, 1 letter representing the reviews from Yelp (Y) or Google (G)

 "source": "G",

Yelp_reviews_ulta_beauty

// string, 22 characters unique user id, refers to the user

"user_id": "TsAJ4DjJFf2BL6RLveOYNw",

// string, 22 characters business id, refers to the business in google_business_data
"business_id": "tnhfDv5Il8EaGSXZGiuQGg",

// float, star rating rounded to 0 or 0.5

"stars": 4.5,

// string, the review in English

"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, month in which the review was made

"month": 4,

// integer, year in which the review was made

"year": 2019,

// integer, hour at which the review was made

"hour": 19,

// string, 1 letter representing the reviews from Yelp (Y) or Google (G)

"source": "G",