

数据库结构说明

1.MySQL 数据说明

数据库由 business, tip 以及 user 三个表组成。

其中, business 一共有两列, 第一列是 business_id, 它是字符串类型; 第二列是 business_info, 它是 JSON 类型。

首先执行以下 SQL 语句:

```
select business_id, JSON_PRETTY(business_info) as business_info from business limit 1;
```

它的格式如下所示:

```
| ---kPU91CF4Lq2-WlRu9Lw |           //这是它的 business_id

//下面是它的 business_info
{
  "city": "New Port Richey",
  "name": "Frankie's Raw Bar",
  "hours": {
    "Friday": "11:0-22:0",
    "Monday": "0:0-0:0",
    "Tuesday": "11:0-21:0",
    "Saturday": "11:0-22:0",
    "Thursday": "11:0-21:0",
    "Wednesday": "11:0-21:0"
  },
  "stars": 4.5,
  "state": "FL",
  "address": "4903 State Rd 54",
  "is_open": 1,
  "latitude": 28.2172884,
  "longitude": -82.7333444,
  "attributes": {
    "BYOB": "True",
    "WiFi": "u'no'",
    "HasTV": "False",
    "Caters": "True",
    "Alcohol": "u'none'",
    "Corkage": "False",
    "Ambience": "{ 'touristy': False, 'hipster': False, 'romantic': False,
    'divey': False, 'intimate': False, 'trendy': False, 'upscale': False,
    'classy': True, 'casual': False}",
    "BikeParking": "True",
    "DogsAllowed": "True",
```

```

    "OutdoorSeating": "True",
    "BusinessParking": "{ 'garage': False, 'street': False, 'validated': False,
    'lot': True, 'valet': False }",
    "RestaurantsTakeOut": "True",
    "RestaurantsDelivery": "False",
    "WheelchairAccessible": "True",
    "BusinessAcceptsBitcoin": "False",
    "RestaurantsReservations": "False",
    "RestaurantsTableService": "True",
    "BusinessAcceptsCreditCards": "True"
  },
  "categories": "Seafood, Restaurants, Latin American, Food, Food Trucks",
  "postal_code": "34652",
  "review_count": 24
}
|

```

可以看到在 hours, attributes 中又包含若干键值对, hours 表示该商户的开店时间, attributes 表示该商户的一些属性, hours 以及 attributes 的内容以及键值对的数量不是固定的, 比如, 一个商户可能有属性 BikeParking, 另一个则没有; 同理, 一个商户可能会在周天歇业, 因此它的 hours 里边不会有 Sunday, 但是可能另一个商户在周天营业, 它的 hours 里就会有 Sunday。

user 与 business 类似, 它由 user_id 与 user_info 两列组成, user_id 是字符串类型, 而 user_info 则是 JSON 类型, 执行下面的语句:

```
select user_id, JSON_PRETTY(user_info) as user_info from user limit 1;
```

```

| ---1lKK3aK0uomHnwAkAow |          //user_id

//下面是 user_info
{
  "cool": 206,
  "fans": 17,
  "name": "Monera",
  "elite": "2010,2011,2012",
  "funny": 181,
  "useful": 510,
  "friends": " 这里面是该用户朋友的 user_id",
  "review_count": 266,
  "average_stars": 3.93,
  "yelping_since": "2007-06-04 01:37:45",
  "compliment_hot": 2,
  "compliment_cool": 9,
  "compliment_cute": 1,

```

```

"compliment_list": 0,
"compliment_more": 3,
"compliment_note": 5,
"compliment_funny": 9,
"compliment_plain": 9,
"compliment_photos": 0,
"compliment_writer": 9,
"compliment_profile": 2
}
|

```

对于 tip 表，执行如下语句：

```
select * from tip limit 1;
```

```

mysql> select * from tip limit 1;
+-----+-----+-----+
| user_id | business_id | tip_info |
+-----+-----+-----+
| AGHJyVwZUey3gcPCJ76iw | 3uLgwr0qeCNWjKenHJwPGQ | {"date": "2012-05-18 02:17:21", "text": "Avengers time with the ladies.", "compliment_count": 0} |
+-----+-----+-----+
1 row in set (0.01 sec)

```

Figure 1: 查询结果

它存储了用户对商户的评论数据，它包含三个字段，user_id, business_id 以及 tip_info，tip_info 是 JSON 类型，它包含三个键值对，分别是评论的日期 date，评论的内容 text，以及评论的点赞数 compliment

2.MonGoDB 数据说明

数据库由 business, user 以及 review 表组成，分别执行查询命令来查看它们的结构：

查看 business 结构

```
db.business.find().limit(1)
```

查询结果如下，它的结构与 MySQL 相差不大

```

[
  {
    _id: ObjectId('6016c6b4af81085b0f2183c1'),
    business_id: '1SWheh84yJXfytoVILX0AQ',
    name: 'Arizona Biltmore Golf Club',
    address: '2818 E Camino Acequia Drive',
    city: 'Phoenix',
    state: 'AZ',
    postal_code: '85016',

```

```

latitude: 33.5221425,
longitude: -112.0184807,
stars: 3,
review_count: 5,
is_open: 0,
attributes: { GoodForKids: 'False' },
categories: [ 'Golf', 'Active Life' ],
hours: null,
loc: { type: 'Point', coordinates: [ -112.0184807, 33.5221425 ] }
}
]

```

查看 user 结构

```
db.user.find().limit(1)
```

```

[
  {
    _id: ObjectId('600d35c72cb85d6290190e31'),
    user_id: 'l6BmjZMeQD3rDxWUbiAiw',
    name: 'Rashmi',
    review_count: 95,
    yelping_since: '2013-10-08 23:11:33',
    useful: 84,
    funny: 17,
    cool: 25,
    elite: '2015,2016,2017',
    friends: [
      'c78V-rj8NQcQjOI8KP3UEA',
      'alRMgPcngYSCJ5naFRBz5g',
      'ajcnq75Z5xxkvUSmmJ1bCg',
      'BSMAmp2-wMzCkhTfq9ToNg',
      'jka10dk9ygX76hJG0gfPZQ',
      'dut0e4xvme7QSlesOychQA',
      '14151BnK356zBua7B-UJ6Q',
      '0HicM00s-M_g12e0-zES4Q',
      '_uI57wL2fLyftrcSFpfSGQ',
      'T4_Qd0YWbC3co6WSMw4vxg',
      'iBRoLWPtWmsI1kdbE9ORSA',
      'xjrUcid6Ymq0DoTJELkYyw',
      'GqadWVzJ6At-vgLzK_SKgA',
      'DvB13VJBmSnbFXBVBsKmdA',
      'vRP9nQkYTeNioDjtxZlVhg',
      'gTOA1iN3eeQ8EMAJJhwQtw',
      '6yCWjFPtp_AD4x93WAwmnw',
    ],
  }
]

```

'1dKzpNnib-JlViKv8_Gt5g',
'3Bv4_JxHXq-gVLOxYMQX0Q',
'ikQyfuliViYh8TOus7wiFQ',
'f1GGltNaB7K5DR1jf3d0mg',
'tgeFUChlh7v8bZfV12-hjQ',
'-9-9oyXlqsMG2he5xIWdLQ',
'Adj9fBPVJad8vSs-mIP7gw',
'Ce49RY8CKXVsTifxRYFTsw',
'M1_7TLi8CbdaA89nFL1H4iw',
'wFsNv-hqbW_F5-IRqfBN6g',
'0Q1L7zXHocaUZ2gsG2XJeg',
'cBFgmOCBdhYa0xoFEAzp_g',
'VrD_AgiFvzqtlR15vir3SQ',
'cpE-7HK514Sr5vpSen9CEQ',
'F1UYelhPFB-zIKlt0ygIZg',
'CQAL1hvsLMCzuJf9AglSxw',
'1KnY1wr15WfEWIRLB9IS6g',
'QWFQ-kXBilbid-lm5Jr3dQ',
'nymT8liFugCrM16lTy0ZfQ',
'qj69bdd885heDvUPCyHd2Q',
'DySCZZcgbdrlHgEovk5y9w',
'lZMJIDuvhT9Dy4KyquLXyA',
'b_9Gn7wS93AoPZPR0dIJqQ',
'N07g1IaLh0_6sUjtiSRe4w',
'YdfPX_7DxSnKvvdCJ57i0w',
'8GYryZPD22W7WgQ8kvMkEQ',
'cpQmAgOWatghp14h1pn1dQ',
'EnchhymLYMqftCRjqvVmw',
'-JdfKhFktE7Zs9BMDFcPeQ',
'uWhC9eof98zPkvsalgaqJw',
'eyTlNDDaiPatfe6mheIZ0g',
'VfHq0o73aKsODvfAhwAQtg',
'kvD5tICngLAaQDujSFWupA',
'dXacwEhqi9-3_XT6JeH00g',
'NfU0zDaTMEQ4-X9dbQWd9A',
'cTHWBdjSKbctSUIvWsgFwx',
'3IEtCbSDF5t7RkZ20T6s9A',
'HJJXTrp6UybYyPdQ9DA0JA',
'JaXogQFVjzGRAeBvzamBHg',
'NUonfKkjS1iVqnNITtgXZg',
'D5vaJAYp0sOrGfsj9qvsMA',
'H27Ecwwu4FGAlLgICourw',
'S8HrLmMiE4u8FWYwkNEoTw',
'Io36Y3xWQcIX9rYvPcYfXQ',
'J5mcqh8KxYpqjaLBNlwcig',

```

'-nTB3_08g06fD0GT8AtDBQ',
'wMpFA46lihK8oFns_5p65A',
'RZGFJHeomGJCWp3xcL3ejA',
'ZoQSzzXoSP1RxOD4Amv9Bg',
'qzM0EBOSkuuGIFv0adjQAA',
'HuM6vvuveken-fPZ7d4o1A',
'H3oukHpGpn9n_mJwSDSQyQ',
'PkmsJsQ8FIZe8eh8c_u96g',
'wSBYVbwME4MzgkJaFyfvNg',
'YEVqknoDmrHAoUbHX0nPnA',
'li3vsK1XAPmeJYAUTYf1HQ',
'MKc8yXi0glbPYt0Qb4PECw',
'fQPH6W9fksi27gkuUPnFaA',
'amrCMrDsoRetYFg2kwwdFA',
'84dVQ6n6r2ezNaTuc7RkKA',
'yW9QjWY0o1v5-uRKv3t_Kw',
'5XJDj7c3eoidfQ3jW18Zgw',
'txSc6a6pIDctvwyBeu7Aqg',
'HFbbDCyyqP9xPkUlcxeIdg',
'hTUv5oh2do6Z3OppPuuiJA',
'gSqonG9J4fNM-fl_fE71AA',
'pd9mgTFpBTg5F9x-MsczNg',
'j3VE22V2GcHiH8UZxfFLfw',
'NYXlMW-T-3V4Jqr4r-iOWg',
'btXgAZedxX8IWhMifA7Xkg',
'-Hp5mPLiRJNFnyeX5Ygzag',
'P6-DwVg6-t2JuQwIUEk0iQ',
'OI2TvxYvZrAodBG_RF53Xw',
'bHxf_VPKmZur1Bier-6A2A',
'Et_Sb39cVm81_Xe9HDM8ZQ',
'5HwG12UyYbaRq8aD6YC-fA',
'ZK228WmCCKLo5thcjD7rdw',
'iTf8wojwfm0N0i7d0iz3Nw',
'btYRxQYNJjpecf1NHtFH0A',
'Kgo42FzpW_dXFgDKoewbtg',
'MNk_1Q_dq0Y3xxHZKe08VQ',
'AlwD504T9k0m51kg3k5g6Q'
],
fans: 5,
average_stars: 4.03,
compliment_hot: 2,
compliment_more: 0,
compliment_profile: 0,
compliment_cute: 0,
compliment_list: 0,

```

```

    compliment_note: 1,
    compliment_plain: 1,
    compliment_cool: 1,
    compliment_funny: 1,
    compliment_writer: 2,
    compliment_photos: 0
  }
]

```

查看 review 结构

```
db.review.find().limit(1)
```

```

[
  {
    _id: ObjectId('600d7ea4f5e9bd91d7c30018'),
    review_id: 'Q1sbwvVQXV2734tPgoKj4Q',
    user_id: 'hG7bOMtEbXx5QzbzE6C_VA',
    business_id: 'ujmEBvifdJM6h6RLv4wQIg',
    stars: 1,
    useful: 6,
    funny: 1,
    cool: 0,
    text: 'Total bill for this horrible service? Over $8Gs. These crooks
    actually had the nerve to charge us $69 for 3 pills. I checked
    online the pills can be had for 19 cents EACH! Avoid Hospital ERs
    at all costs.',
    date: '2013-05-07 04:34:36'
  }
]

```

与 MySQL 中的 tip 相比多出了一些信息，比如 useful, funny, cool 等

3.Neo4j 数据说明

由于 Neo4j 具有图形化的界面，可以点击左侧的节点，然后就会在右侧展示出它的数据结构，比较方便查看。

这里需要说明一下各个节点之间的关系：

```

(BusinessNode)-[IN_CITY]->(CityNode)
(UserNode)-[Review]->(ReviewNode)
(BusinessNode)<-[Reviewed]-(ReviewNode)
(UserNode)-[HasFreind]-(UserNode)
(BusinessNode)-[IN_CATEGORY]->(CategoryNode)

```