

Rapport de projet MongoDB

Gaspard DANNET

Arthur HOUDANT

Modélisation

Compte tenu des besoins exprimés en section 1.3. , nous avons décidé de ne garder que les champs :

Pour les restaurants :

- business_id
- name
- city
- address
- categories
- attributes {
 Ambience
}

Pour les users :

- user_id
- name
- friends

Pour les reviews :

- review_id
- user_id
- stars
- business_id
- text

Pour simplifier les requêtes exprimées en section 1.3, nous avons ajoutés quelques champs.

Aux restaurants, nous avons ajouté le champ « nbAmbience » qui correspond au nombre d'ambiance dans le restaurant (= nombre d'ambiance de valeur true dans le dictionnaire restaurant[« attributes »][« Ambience »]). Ceci étant fait, le champ « attributes » ne sert plus par la suite, nous l'avons donc retiré.

Aux users, nous avons ajouté le champ « nbFriends » qui correspond au nombre d'ami de l'utilisateur.

L'extraction des données utiles et les modifications apportées ont été effectués dans le script `extraire_donnees.py` (à noter que dans ce même fichier `.py` a été implémenté l'ajout d'un champ « criteres » aux restaurants qui nous sera utile pour la section recommandation et plus particulièrement le calcul du facteur « préférence ». Ce champ représente Γ_r).

Ensuite, nous avons apportés d'autres modifications à cette base de données afin de faciliter le programme de recommandation.

A la collection de restaurants, nous avons donc rajoutés les champs « nbReview » et « avgStars » correspondant respectivement au nombre de review qui ont été rédigé par un utilisateur sur le restaurant et à la note moyenne du restaurant. Ces deux champs nous seront utiles pour calculer le facteur « appréciation » de la recommandation.

A la collection de users, nous avons ajoutés le champ « criteres » qui correspond aux critères préférés de l'utilisateur. Un critère sera dans « criteres » si l'utilisateur a donné une note supérieure à 2.5 à un restaurant qui possède ce critère dans son champ « criteres ». Ce champ représente Γ_u .

Au final, la base de donnée ressemble à :

Collection restaurants :

- business_id
- name
- city
- address
- categories
- nbAmbiance
- criteres
- nb_review
- avgStars

Collection users :

- user_id
- name
- friends
- nbFriends
- criteres

Collection reviews :

- review_id
- user_id
- business_id
- stars
- text

Interrogation des données

Dans cette partie nous allons définir les requêtes MongoDB répondant aux besoins exprimés et les résultats de ces requêtes en utilisant la modélisation détaillée précédemment.

1] Obtenir la liste des restaurants de la ville de Greenville

```
db.yelp_restaurants.find({ "city": "Greenville" })
```

```
[
  {
    _id: ObjectId('65803ef37cf2cb46f0887c66'),
    business_id: 'Vo4NZBtF3mr6EJitNnq-Wg',
    name: 'Pizzeria Maki',
    city: 'Greenville',
    address: '4007 Kennett Pike, Ste A',
    categories: 'Restaurants, Pizza, Salad, Sushi Bars',
    nbAmbience: 0,
    criteres: [ 'Restaurants', 'Pizza', 'Salad', 'Sushi Bars' ],
    nb_review: 13,
    avg_stars: 4.769230769230769
  },
  {
    _id: ObjectId('65803ef37cf2cb46f0887ccf'),
    business_id: 'sTvFxzQU_4bm-owSJXx0MQ',
    name: 'Cromwell's American Tavern and Taqueria',
    city: 'Greenville',
    address: '3858 Kennett Pike',
    categories: 'Nightlife, Bars, Mexican, American (New), Restaurants',
    nbAmbience: 1,
    criteres: [
      'Nightlife',
      'Bars',
      'Mexican',
      'American (New)',
      'Restaurants',
      'casual'
    ],
    nb_review: 101,
    avg_stars: 3.237623762376238
  },
  {
    _id: ObjectId('65803ef37cf2cb46f0887ce0'),
    business_id: '_ju1BXAF5g5QtAfZZKO1g',
    name: 'Brew Ha Ha',

```

```

city: 'Greenville',
address: '3838 Kennett Pike',
categories: 'Food, Breakfast & Brunch, American (New), Restaurants, Coffee & Tea',
nbAmbience: 1,
criteres: [
  'Food',
  'Breakfast & Brunch',
  'American (New)',
  'Restaurants',
  'Coffee & Tea',
  'hipster'
],
nb_review: 83,
avg_stars: 4.096385542168675
},
{
  _id: ObjectId('65803ef37cf2cb46f0887d22'),
  business_id: 'gRHfRxTqaf_oUapIB9MMUA',
  name: 'Cosi',
  city: 'Greenville',
  address: '3828 Kennett Pike Ste A',
  categories: 'Restaurants',
  nbAmbience: 0,
  criteres: [ 'Restaurants' ],
  nb_review: 8,
  avg_stars: 3
},
{
  _id: ObjectId('65803ef37cf2cb46f0887dd4'),
  business_id: 'hJ0QkeOdQ5y7MTmdR9NO1Q',
  name: 'The Copper Dram',
  city: 'Greenville',
  address: '3826 Kennett Pike',
  categories: 'Beer, Wine & Spirits, Cocktail Bars, American (New), Bars, Food, Nightlife, American (Traditional), Restaurants',
  nbAmbience: 3,
  criteres: [
    'Beer',
    'Wine & Spirits',
    'Cocktail Bars',
    'American (New)',
    'Bars',
    'Food',
    'Nightlife',
    'American (Traditional)',
    'Restaurants',
    'casual',
    'romantic',
    'upscale'
  ],
  nb_review: 59,
  avg_stars: 4.52542372881356
},
{
  _id: ObjectId('65803ef37cf2cb46f0887e64'),
  business_id: 'K3ZZua4ab6YLhPqZ_DEDIw',
  name: 'BBC Tavern and Grill',
  city: 'Greenville',
  address: '4019 Kennett Pike',
  categories: 'Nightlife, Restaurants, American (Traditional), Bars, Beer, Wine & Spirits, Food, Sports Bars',
  nbAmbience: 1,
  criteres: [
    'Nightlife',
    'Restaurants',
    'American (Traditional)',

```

```

    'Bars',
    'Beer',
    'Wine & Spirits',
    'Food',
    'Sports Bars',
    'casual'
  ],
  nb_review: 112,
  avg_stars: 3
},
{
  _id: ObjectId('65803ef37cf2cb46f0887eb3'),
  business_id: 'GsTHcKsO4c9ZnK0AdwfCpw',
  name: 'Starbucks',
  city: 'Greenville',
  address: '3828 Kennett Pike, Ste A',
  categories: 'Cafes, Coffee & Tea, Restaurants, Food, American (New)',
  nbAmbience: 0,
  criteres: [
    'Cafes',
    'Coffee & Tea',
    'Restaurants',
    'Food',
    'American (New)'
  ],
  nb_review: 9,
  avg_stars: 3.4444444444444446
}
]

```

2] Obtenir les utilisateurs qui ont rédigé des reviews avec une note inférieure ou égale à 2.0

db.yelp_review.find({ "stars": { \$lte: 2.0 } }, { "user_id": 1 })

```

[
  {
    _id: ObjectId('65803f0d24fc4ae889e31002'),
    user_id: 'TRFTDMJegE6DituEKBPkEQ'
  },
  {
    _id: ObjectId('65803f0d24fc4ae889e3100c'),
    user_id: '9kMI5dY91gmSWOqYM7SXkQ'
  },
  {
    _id: ObjectId('65803f0d24fc4ae889e3102c'),
    user_id: '5py7HurGiQkg5hKwelk3HA'
  },
  {
    _id: ObjectId('65803f0d24fc4ae889e31035'),
    user_id: '0Ler18OI57Ef6C4OzwACAg'
  },
  {
    _id: ObjectId('65803f0d24fc4ae889e3103a'),
    user_id: '6Rd-Oz0Iqns4_9NIhG9blw'
  },
  {
    _id: ObjectId('65803f0d24fc4ae889e3103b'),
    user_id: 'XjLvFaNKzzprEDRR-3kxsQ'
  },
  {
    _id: ObjectId('65803f0d24fc4ae889e3103c'),

```

```

    user_id: 'kBFb2l1OmfQ-zLOR1d112Q'
  },
  {
    _id: ObjectId('65803f0d24fc4ae889e31047'),
    user_id: 'wig0uC_qyPXfm3SmsjFCVQ'
  },
  {
    _id: ObjectId('65803f0d24fc4ae889e31051'),
    user_id: 'bCmOH5WRgd8MMLsz_jlfQA'
  }
]

```

3] Obtenir les noms et adresses des restaurants de la ville de New Castle qui entrent dans la catégorie Mexican

```

db.yelp_restaurants.find(
  { "city": "New Castle", "criteres": "Mexican" }, { "name": 1, "address": 1 })

```

```

[
  {
    _id: ObjectId('65803ef37cf2cb46f0887cac'),
    name: 'Three Countries Pizzeria',
    address: '202 E 6th St'
  },
  {
    _id: ObjectId('65803ef37cf2cb46f0887d2a'),
    name: 'Rivera Taco Express',
    address: '243 Christiana Rd'
  },
  {
    _id: ObjectId('65803ef37cf2cb46f0887d33'),
    name: 'Taqueria Atexquita',
    address: '220 New Castle Ave'
  },
  {
    _id: ObjectId('65803ef37cf2cb46f0887d63'),
    name: 'Chipotle Mexican Grill',
    address: '115A S Dupont Hwy'
  },
  {
    _id: ObjectId('65803ef37cf2cb46f0887d7d'),
    name: 'El Cantaro',
    address: '181 S Dupont Hwy'
  },
  {
    _id: ObjectId('65803ef37cf2cb46f0887d8a'),
    name: 'Los Taquitos De Puebla',
    address: '241 N Dupont Hwy'
  },
  {
    _id: ObjectId('65803ef37cf2cb46f0887d95'),
    name: 'Casa Mariachi',
    address: '519 E Basin Rd'
  },
  {
    _id: ObjectId('65803ef37cf2cb46f0887de0'),
    name: 'La Tlaxquenita',
    address: '218 New Castle Ave'
  },
  {
    _id: ObjectId('65803ef37cf2cb46f0887e7b'),

```

```

    name: 'Taco Bell',
    address: '116 North Dupont Highway'
  },
  {
    _id: ObjectId('65803ef37cf2cb46f0887f0d'),
    name: 'Los Jarritos',
    address: '110 N Dupont Hwy'
  },
  {
    _id: ObjectId('65803ef37cf2cb46f0887f64'),
    name: 'Mi Cosina',
    address: '165 S Dupont Hwy'
  },
  {
    _id: ObjectId('65803ef37cf2cb46f0887f77'),
    name: 'Mr. Taco',
    address: '101 East Hazeldell Ave'
  }
]

```

4] Obtenir le nombre d'utilisateurs qui n'ont pas d'ami

```
db.yelp_user.find({ "nbFriends": 0 }).count()
```

16047

5] Obtenir le nombre de restaurants pour chaque ville

```
db.yelp_restaurants.aggregate([ { $group: { _id: "$city", count: { $sum: 1 } } } ])
```

```

[
  { _id: 'Trolley Square', count: 1 },
  { _id: 'Talleyville', count: 3 },
  { _id: 'Claymont', count: 23 },
  { _id: 'Pike Creek', count: 2 },
  { _id: 'Stanton', count: 1 },
  { _id: 'wilmington', count: 1 },
  { _id: 'Newport', count: 3 },
  { _id: 'Wilmington ', count: 1 },
  { _id: 'Montchanin', count: 2 },
  { _id: 'Aston', count: 1 },
  { _id: 'Christiana', count: 6 },
  { _id: 'Greenville', count: 7 },
  { _id: 'Elsmere', count: 2 },
  { _id: 'Newark', count: 146 },
  { _id: 'New Castle', count: 102 },
  { _id: 'Marshallton', count: 1 },
  { _id: 'Hockessin', count: 37 },
  { _id: 'Wilmington Manor', count: 2 },
  { _id: 'Glen Mills', count: 1 },
  { _id: 'Wilmington', count: 619 }
]

```

6] Obtenir la note globale le plus élevée obtenue par un restaurant

```
db.yelp_restaurants.find({}).sort({ "avg_stars": -1 }).limit(1)
```

```
[
  {
    _id: ObjectId('65803ef37cf2cb46f0887c90'),
    business_id: 'wWcETNWBnFEi1J_W_9hXJA',
    name: 'El Fogon De Tato',
    city: 'Wilmington',
    address: '2304 Kirkwood Hwy',
    categories: 'Puerto Rican, Restaurants, Latin American, Caribbean',
    nbAmbience: 0,
    criteres: [ 'Puerto Rican', 'Restaurants', 'Latin American', 'Caribbean' ],
    nb_review: 7,
    avg_stars: 5
  }
]
```

7] Obtenir les restaurants qui proposent au moins 3 ambiances différentes

db.yelp_restaurants.find({ "nbAmbience": { \$gte: 3 } }, { _id:0, name: 1 })

```
[
  { name: '8th & Union Kitchen' },
  { name: 'Farmer & the Cow' },
  { name: 'Twisted Soul Restaurant & Bar' },
  { name: 'Pizza by Elizabeths' },
  { name: 'Rice Restaurant' },
  { name: 'First Watch' },
  { name: 'The Copper Dram' },
  { name: 'Trolley Square Oyster House' },
  { name: 'Satsuma Asian Kitchen & Bar' },
  { name: 'Turning Point of Christiana' },
  { name: 'Pizza by Elizabeths' },
  { name: 'Daily Veg - Pike Creek' },
  { name: 'Cafe New Castle' },
  { name: 'Constitution Yards' }
]
```

8] Obtenir les identifiants des restaurants qui ont été évalués par l'un des utilisateurs correspondants aux identifiants Yelp n2UpKhkU2N-66a1QQzrjYw et ucD25otZ0uqWPSJnl4muQQ

*db.yelp_review.find(
 { "user_id": { \$in:
 ["n2UpKhkU2N-66a1QQzrjYw", "ucD25otZ0uqWPSJnl4muQQ"] } },
 { "business_id": 1, "_id": 0 })*

```
[
  { business_id: 'Dy91wdWkwtl_qgjAlZ0Niw' },
  { business_id: 'Dy91wdWkwtl_qgjAlZ0Niw' },
  { business_id: '7O7YHTUB3uqXHsQSLKR0SA' },
  { business_id: 'citzbW-2MVEZRNbqVO-Jw' },
  { business_id: 'c4D851qxOujE5KTH8upXTA' },
  { business_id: '3zWNxtRhyR4rLPfXyQgSyw' },
  { business_id: 'jVJH1KfitQi87MYMh6vyMg' },
  { business_id: 'EROSyhoWNCUH53NrDrsHgQ' },
]
```



```

{ business_id: '3zWNxtRhyR4rLPfXyQgSyw' },
{ business_id: 'Dy91wdWkwTl_qgjAlZ0Niw' },
{ business_id: 'OIOVRlcEstxKK_6hMjy63A' },
{ business_id: 'uoMWvXch6sWnI9FMSE_-Vg' },
{ business_id: 'N-_ilHVIMQqzkgA7nHfA8A' },
{ business_id: 'EROSyhoWNCUH53NrDrsHgQ' },
{ business_id: '5RtAJPXFLTzwALie1mtNuA' },
{ business_id: 'Oksx3ALgKLRVxeAR8l_qA' },
{ business_id: '52RFRa6Ulblxy8EnWJQpJg' },
{ business_id: '8vvMGaj5biveUVE6-yTipQ' },
{ business_id: 'qFsh80AAL90tkOc0n98bqg' },
{ business_id: 'fd_fNaeZxcZfUza3K6Be9w' },

.
.
.

{ business_id: 'VbzoVWEPSfzaIGBq4KIVwQ' },
{ business_id: 'JvFDMS5yryf3IDwUPAb6OQ' },
{ business_id: 'riwJWxj4850tGBhRahDt7g' },
{ business_id: 'aAsIC0I5mZB8ojNbld0byA' },
{ business_id: 'v6AlogsSBf4sxrVHcSnMHQ' },
{ business_id: 'UREmwQBCMimpOuFyKPW3xw' },
{ business_id: 'Z53I9xwy0aMRy577GuM1KQ' },
{ business_id: 'v6AlogsSBf4sxrVHcSnMHQ' },
{ business_id: 'ELgR35KllpuTMeWi6tFP7g' }
]

```

9] Obtenir les identifiants Yelp et noms des utilisateurs qui ont évalué les deux restaurants d'identifiants Yelp UBX73ZWgCdgom4nv0UeGvg et 51Fo_uY52A5oal8YTHU5yg

```

db.yelp_review.aggregate([
  { $match: { "business_id":
    { $in: ["UBX73ZWgCdgom4nv0UeGvg","51Fo_uY52A5oal8YTHU5yg"] } } },
  { $group: { _id: "$user_id", count: { $sum: 1 } } },
  { $match: { count: { $gte: 2 } } },
  { $lookup: { from: "yelp_user", localField: "user_id", foreignField: "user_id",
    as: "user_info" } },
  { $project: { "_id": 0, "user_id": "$_id", "user_name": { $arrayElemAt:
    ["$user_info.name", 0] } } }
])

```

```

[
{ user_id: 'B9iPTih0dmpZn6fRyZnLLg' },
{ user_id: '0ojv9GMM4ljRiULkjyXYmQ' },
{ user_id: 'I7Fmd8WQA0FozWoeZQNdkw' },
{ user_id: 'gF9BKX9fuTPQL-EFvSok2Q' },
{ user_id: 'G1cT_L6toouKClehhgCzWQ' },
{ user_id: 'grk4WsJa4evlds1pFbfzg' },
{ user_id: 'B6uZIAMkd9EbPiX3k4fxsA' },
{ user_id: 'v4KYRLFGmeqwIntnDbj4Uw' },
{ user_id: 'LR5KLo9TkFbuLTUnQXz0xA' },
{ user_id: 'p-swY1Yb2qTOpnwJy65W3A' },
{ user_id: 'Q3MDOyTjrDmE72ZX3ZQPKA' },
{ user_id: 'qKOWuKXqa_16hmLnbntS1A' },

```

```

{ user_id: 'dXSUR6L9B10g-YMQMlnkag' },
{ user_id: 'DDBM1aIU5t9JfwjwQRk1Sg' },
{ user_id: 's9wPpM3DIceER-AZllqmmQ' },
{ user_id: 's8CI5iAMRD6b5yLn0EG2Tw' },
{ user_id: 'ehR-kzNZUBUm_RJmnTZ94Q' },
{ user_id: 'S6K4I-tzNekMg3YOKVHJ_A' },
{ user_id: 'uvEIDVQx3gctkUTBarAEng' },
{ user_id: 'YwEuCgKpQZ9MLvqBDJBAog' },
.
.
.
{ user_id: 'ksg167hKZcWBiPP1F8DtRw' },
{ user_id: 'KPCsL2cHkTnRbSPvYyaCTQ' },
{ user_id: 'RIEQ8qmFmHoRd_iQq7CUqg' },
{ user_id: 'kyVJXFslf_JaHMy51rc67w' }
]

```

10] Obtenir le nombre de restaurants qui ont reçu des notes au-dessus de 4.0 et également en-dessous de 2.0

```

db.yelp_restaurants.find(
  { $or: [
    { "avg_stars": { $gt: 4.0 } },
    { "avg_stars": { $lt: 2.0 } } ]
  }).count()

```

316

11] Obtenir les différentes catégories associées aux restaurants évalués par l'utilisateur d'identifiant Yelp FIXBpK_YzxLo27jcMdII1w

```

db.yelp_review.aggregate([
  { $match: { "user_id": "FIXBpK_YzxLo27jcMdII1w" } },
  { $lookup: { from: "yelp_restaurants", localField: "business_id",
    foreignField: "business_id", as: "restaurant" } },
  { $unwind: "$restaurant" },
  { $project: { "restaurant.categories": 1, "_id": 0 } }
])

[
  {
    restaurant: {
      categories: 'Cajun/Creole, Vegetarian, Restaurants, Tex-Mex, Cocktail Bars, Bars, Nightlife'
    }
  },
  {
    restaurant: {
      categories: 'Nightlife, Sports Bars, Seafood, Breakfast & Brunch, Bars, Sandwiches, Restaurants'
    }
  }
]

```

10

```
}
]
```

12] Obtenir le nombre d'amis de chaque utilisateur (y compris pour ceux qui n'en ont pas)

```
db.yelp_user.find({}, { _id:0, user_id: 1, nbFriends:1 })
```

```
[
  { user_id: 'XLs_PhrJ7Qwn_RfgMM7Djw', nbFriends: 0 },
  { user_id: '0juMoWXC7z4c7LgQP2s3sA', nbFriends: 4 },
  { user_id: '1-gnHNzZNbn8IFwYRF_CtA', nbFriends: 3 },
  { user_id: 'PadmV2GEoA6mWpQUpPh7Ig', nbFriends: 2 },
  { user_id: 'FlhZYUEhfpfm0Ynxacftwg', nbFriends: 5 },
  { user_id: 'FH14lvolAKhva9KVgvzKJw', nbFriends: 0 },
  { user_id: '7LmJ87EijrHcoYRu9WgoHg', nbFriends: 14 },
  { user_id: 'VFXKRVzEMbnsDcYFdwiiuA', nbFriends: 2 },
  { user_id: 'lpLRJY4CP3fXtlEd8Y4GFQ', nbFriends: 23 },
  { user_id: 'UJEJgi1xpmLR4b3J8fn2Q', nbFriends: 75 },
  { user_id: 'kLm3B6LeboEZWX_4otplLA', nbFriends: 2 },
  { user_id: 'mZHK_DPbeUTQBZOB0yoqlA', nbFriends: 6 },
  { user_id: 'nPJw7YpKgHE6fowxgl3l1A', nbFriends: 8 },
  { user_id: '-tnodf38uCXjogMhZtcWxw', nbFriends: 0 },
  { user_id: 'k_muz40jjaiY1PVga6NDXA', nbFriends: 17 },
  { user_id: 'GG0mFsEXb-02_dzFPqRV1Q', nbFriends: 39 },
  { user_id: 'y5qEpAfjQaOjjG8ZJWrp6w', nbFriends: 29 },
  { user_id: 'SBpz4xJ8MnrzXOpWiETCyw', nbFriends: 3 },
  { user_id: 'cwChkLd3jcvmqrk5UfUGJw', nbFriends: 3 },
  { user_id: 'NNyr4wknApjc5nF9QOO2Mg', nbFriends: 2 }
]
.
.
.
```

14] Obtenir les commentaires d'utilisateurs qui mentionnent le mot amazing

```
db.yelp_review.find(
  { "text": { $regex: "amazing", $options: "i" } }, { "text": 1, "_id": 0 })
```

```
[
  {
    text: "This is my favorite in Delaware. The menu is simple and perfectly executed. Rolls are reasonably priced, artfully made and are PACKED with fish! I do believe they use flash-frozen sashimi grade fish, but the color, texture, taste and smell are perfect. Easily the best spicy tuna I have ever had. I think the secret is the kewpie mayo along w/ the right amount of sriracha and sichimi powder. Frankly, I can't get enough of this place and have at times, eaten it every day for lunch for weeks. I've gone as far as to pick it up for dinner AFTER having had it for lunch. Their introduction of short-grain brown sushi rice has been great!\n" +
      '\n' +
      "The Sushiman, YC Park is amazing to watch. I've seen him skin a cucumber with a sword-like sushi knife in no time flat, in a single ribbon and then continue to shave it into a nori thin sheet for his tuna naruto roll roll. The temaki are as big as a carnival iced cream cone. You can barely finish one of them, there is so much fish in there. The nigiri are beautiful. Uniform pieces of fresh fish, perfectly perched on the best sushi rice I've ever had (just the right amount of vinegar and sugar).\n" +
      '\n' +
      'Skip the newly opened Kooma. Save Mikimotos for a night out, but go to Tokyo for lunch! The owners are lovely, so support them in this tough economy. If you must eat sushi, eat it here!\n' +
      '\n' +
      'Why not 5 stars? I have to save it for places that serve the local fresh fish, plucked out of the water that morning. Let me know if you find a place that does!'
```

```

},
{
  text: "It's good diner food. period.\n" +
    '\n' +
    'Nothing fancy, just simple food made well. Breakfast is a must. Homemade Belgium waffle covered in TONS
of fruit....amazing.\n' +
    '\n' +
    "Simple Said...it's good diner food at a good price with friendly waitresses. On a Saturday morning...what
more could you ask for?!"
},
{
  text: 'For a small local place, the crab cakes are very delicious. The restaurant is set up like a pizza shop. You
order at the counter to eat in or take out. They bring the food to you if you eat in. If you order to go, they call
the number on your receipt when your order is ready. \n' +
    '\n' +
    'My fiance and I tried the place out the first week it opened. We each ordered a crab cake platter since they
are known for their crab cakes. We also ordered a cup of soup each since it was cold out. The portions were
small but the crab cake was very good. It was full of crab and not fillers. \n' +
    '\n' +
    'The soups were amazing! I had a crab chowder and it was full of crab, not too creamy and great seasoning.
\n' +
    '\n' +
    'The second time we came, I ordered the crab cake sandwich. The brioche bun was a little too much for the
6 oz crab cake. I had to gut out the bun. That allowed me to enjoy the crab cake more. \n' +
    '\n' +
    'The next day my sister came to visit and wanted to try it. This time we ordered the crab balls. Definitely not
as good as the crab cakes. It was full of fillers and not a great way to introduce my sister to it. \n' +
    '\n' +
    'My fiance and i will definitely be back to give the rest of the menu a try.'
},
.
.
.
]

```

15] Obtenir les 3 prénoms d'utilisateurs les plus présents dans la base

```

db.yelp_user.aggregate([
  { $group: { _id: "$name", count: { $sum: 1 } } },
  { $sort: { count: -1 } }, { $limit: 3 }
])

[
  { _id: 'John', count: 314 },
  { _id: 'Michael', count: 236 },
  { _id: 'David', count: 218 }
]

```

Application : moteur de recommandation de restaurants

Nous allons utiliser dans cette partie le script python "recommandation.py" qui a pour but d'associer un score de pertinence entre un utilisateur et un restaurant. Cela en calculant différents facteurs (facteurs d'appréciation, facteur de préférence, facteur social et facteur commentaire) pour lesquels la modélisation a été pensée.

Nous allons utiliser la fonction "recommandation" qui a pour but de renvoyer les restaurants (au maximum 20) recommandés à l'utilisateur donné situé dans une ville donnée.

Cette fonction sera testée sur cinq exemples différents.

1] Utilisateur : kLm3B6LeboEZXW_4otpILA - Izzat, Ville : Greenville

recommandation("kLm3B6LeboEZXW_4otpILA","Greenville")

Recommandations pour la ville: Greenville

```
rang 1  Brew Ha Ha  score = 0.0642346685572492
rang 2  Cromwell's American Tavern and Taqueria  score = 0.05913965844402277
rang 3  BBC Tavern and Grill  score = 0.058621598877980366
rang 4  The Copper Dram  score = 0.05504596774193548
rang 5  Pizzeria Maki  score = 0.032322935779816514
rang 6  Cosi  score = 0.02493498759305211
rang 7  Starbucks  score = 0.02437452229299363
```

Execution time: 4.046916246414185 seconds

2] Utilisateur : kLm3B6LeboEZXW_4otpILA - Izzat, Ville : Newark

recommandation("kLm3B6LeboEZXW_4otpILA","Newark")

Recommandations pour la ville: Newark

```
rang 1  Border Cafe  score = 0.26772266443234183
rang 2  Firebirds Wood Fired Grill  score = 0.173152688172043
rang 3  Metro Diner  score = 0.1649719165085389
rang 4  Two Stones Pub  score = 0.1448476084538376
rang 5  Gyro Kabob House  score = 0.12604458525345622
rang 6  vPho  score = 0.11592580645161288
rang 7  The Cheesecake Factory - Newark  score = 0.10413514431239387
rang 8  Brio Italian Grille  score = 0.10408064516129031
rang 9  Eggspectation  score = 0.10284356435643563
rang 10  Sushi Sumo  score = 0.10051558228540186
rang 11  Bahama Breeze  score = 0.0951350806451613
rang 12  Potstickers Asian Grill  score = 0.09463117375982652
rang 13  Biden Welcome Center  score = 0.09215983643798273
rang 14  First Watch  score = 0.08511938809444629
rang 15  Turning Point of Christiana  score = 0.07973244781783681
rang 16  Horace Shellhammer's  score = 0.07691612903225806
```

rang 17 J.B. Dawson's Restaurant & Bar score = 0.07689983022071306
rang 18 Argilla Brewing score = 0.07676129032258064
rang 19 Pinang Asian Cuisine score = 0.07524346349745331
rang 20 Ted's Montana Grill - Christiana score = 0.07216363636363636
Execution time: 123.00794315338135 seconds

3] Utilisateur : IpLRJY4CP3fXtlEd8Y4GFQ – Robyn, Ville : Wilmington

recommendation("IpLRJY4CP3fXtlEd8Y4GFQ","Wilmington")

Recommandations pour la ville: Wilmington

rang 1 Iron Hill Brewery & Restaurant score = 0.3090322580645162
rang 2 Scrumptious score = 0.30335483870967744
rang 3 Goat Kitchen and Bar score = 0.2649548387096774
rang 4 El Diablo Burritos score = 0.25806451612903225
rang 5 Russell's Quality Food score = 0.23901935483870967
rang 6 Lucky's Coffee Shop score = 0.22967741935483874
rang 7 Big Fish Grill - Riverfront score = 0.22655483870967744
rang 8 Green Room score = 0.22492903225806452
rang 9 Piree's Piri Piri Grill score = 0.22361290322580646
rang 10 Soybean Asian Grille score = 0.21790967741935485
rang 11 La Fia score = 0.21775483870967743
rang 12 Pizza by Elizabeths score = 0.21527741935483874
rang 13 Ubun Thai Cuisine score = 0.2141161290322581
rang 14 Mikimotos Japanese Restaurant & Sushi Bar score = 0.21331612903225805
rang 15 LOMA Coffee score = 0.20978064516129036
rang 16 Maiale score = 0.20820645161290324
rang 17 DiMeo's Pizza score = 0.20583225806451616
rang 18 Washington Street Ale House score = 0.20541935483870966
rang 19 Johnnie's Dog House and Chicken Shack score = 0.2046709677419355
rang 20 Metro Diner score = 0.20420645161290324
Execution time: 301.26888823509216 seconds

4] Utilisateur : IpLRJY4CP3fXtlEd8Y4GFQ – Robyn, Ville : New Castle

recommendation("IpLRJY4CP3fXtlEd8Y4GFQ","New Castle")

Recommandations pour la ville: New Castle

rang 1 1 In Town score = 0.20348387096774195
rang 2 Pizza & Pasta score = 0.16658064516129034
rang 3 Wawa score = 0.16650322580645163
rang 4 Pat's Pizza & Pasta score = 0.16510967741935487
rang 5 Shake Shack score = 0.15561290322580645
rang 6 TGI Fridays score = 0.09277419354838712
rang 7 Jessop's Tavern score = 0.09236129032258066
rang 8 Ioannoni's Bullroasters score = 0.0672
rang 9 Nora Lee's score = 0.056283870967741935
rang 10 The Dog House score = 0.04738064516129032
rang 11 Lestardo's Crab House score = 0.03561290322580645
rang 12 Casablanca Moroccan Restaurant score = 0.03553548387096774
rang 13 Tasti Thai score = 0.03530322580645161
rang 14 Masala Kitchen score = 0.027638709677419354
rang 15 Golden Dove Diner Restaurant score = 0.02407741935483871
rang 16 Saap Lao Thai Cuisine score = 0.022761290322580643
rang 17 Cafe New Castle score = 0.022606451612903224

rang 18 Little Anthony's Pizza score = 0.01587096774193548
rang 19 Rivera Taco Express score = 0.014941935483870966
rang 20 Alex's Seafood Restaurant & Clam Bar score = 0.014941935483870966
Execution time: 51.202325105667114 seconds

5] Utilisateur : GG0mFsEXb-02_dzFPqRV1Q - Mandy, Ville : Wilmington

recommendation("GG0mFsEXb-02_dzFPqRV1Q","Wilmington")

Recommandations pour la ville: Wilmington

rang 1 Iron Hill Brewery & Restaurant score = 0.3523655913978495
rang 2 Big Fish Grill - Riverfront score = 0.33905483870967745
rang 3 Juliana's Kitchen score = 0.29777947214076245
rang 4 Sakura Japanese Sushi Restaurant score = 0.2966358351964229
rang 5 Scrumptious score = 0.2943330995792427
rang 6 El Pique - Wilmington score = 0.29239717741935484
rang 7 De La Coeur Café et Pâtisserie score = 0.29106912442396315
rang 8 Trolley Square Oyster House score = 0.28702005231037486
rang 9 Two Stones Pub score = 0.28643870967741936
rang 10 Goat Kitchen and Bar score = 0.2851321436742164
rang 11 Rasa Sayang score = 0.2799106382978723
rang 12 Bonefish Grill score = 0.27175806451612905
rang 13 Southeast Kitchen score = 0.2716394071490846
rang 14 Russell's Quality Food score = 0.27108832035595104
rang 15 Chelsea Tavern score = 0.26974227804738793
rang 16 Farmer & the Cow score = 0.26960540540540545
rang 17 Pochi score = 0.26749105020616054
rang 18 La Pizzeria Metro score = 0.2645211074422149
rang 19 Bellefonte Cafe score = 0.26425009487666035
rang 20 Cafe SiTALY score = 0.2619245218384242
Execution time: 962.0230007171631 seconds