Tutoriel Kibana

Ce tutoriel a été conçu pour la version 7.5 d'Elasticsearch et de Kibana

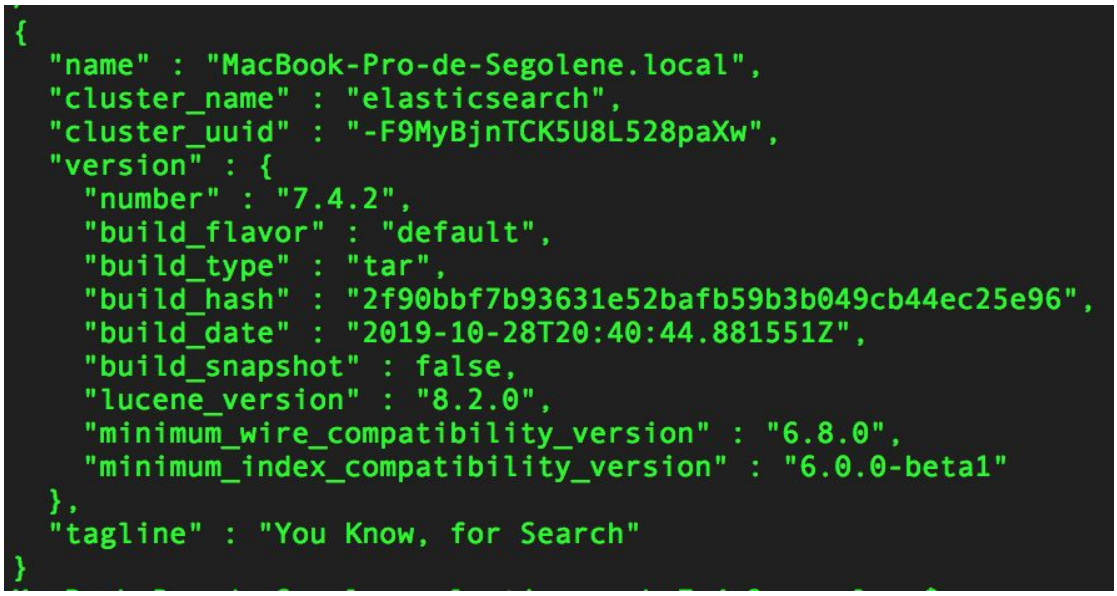
INSTALLATION

https://www.elastic.co/fr/start

Elasticsearch

Lien de téléchargement : https://www.elastic.co/fr/downloads/elasticsearch

- Télécharger la version d'ElasticSearch correspondant à son système d'exploitation
- Dézipper le fichier
- Ouvrir un terminal dans le dossier `elasticsearch-[version]`[1]
- Lancer la commande `bin/elasticsearch` (`bin\elasticsearch.bat` sur Windows) : veiller à avoir assez de place sur son ordinateur
- Pour vérifier qu'ElasticSearch s'est bien lancé, ouvrir un second terminal dans le dossier `elasticsearch-[version]` et exécuter la commande `curl http://localhost:9200/`



Une réponse similaire devrait apparaître sur le terminal

Kibana

Lien de téléchargement : https://www.elastic.co/fr/downloads/kibana

- Télécharger la version de Kibana correspondant à son système d'exploitation
- Dézipper le fichier
- Exécuter la commande `bin/kibana` (`bin\kibana.bat` sur Windows) dans un terminal ouvert dans le dossier `kibana-[version]`
- Ouvrir son navigateur à l'adresse `http://localhost:5601`
- Tada !

[1] Remplacer `[version]` par le numéro de la version effectivement téléchargée

RELANCER ELASTICSEARCH ET KIBANA

En fonction de la méthode d'installation d'Elasticsearch et Kibana, le processus de démarrage varie[2]. Dans le cas d'une installation avec des paquets d'archives (comme expliqué précédemment) :

Elasticsearch

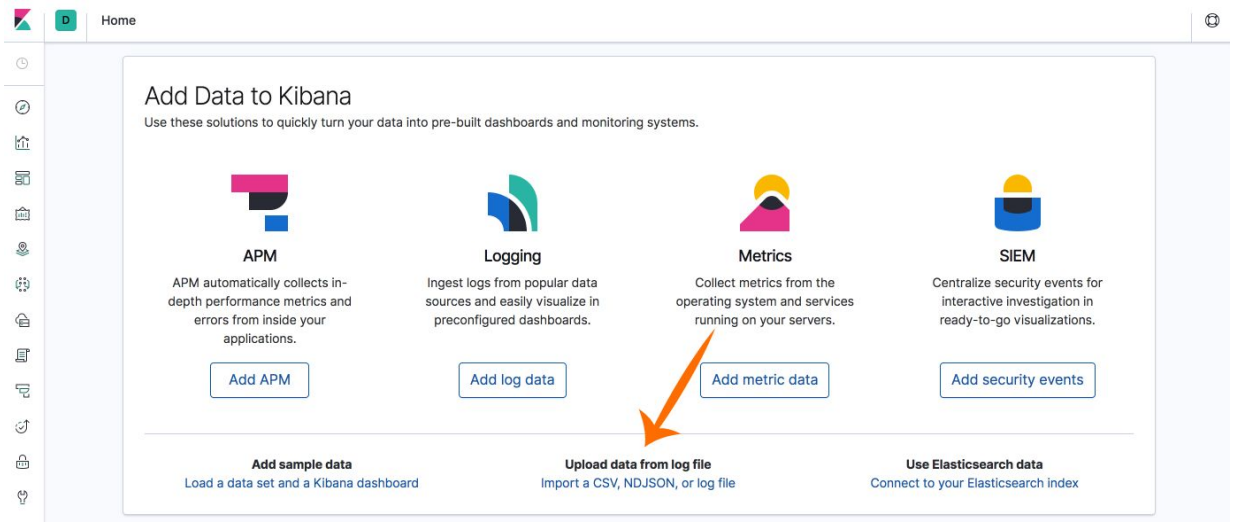
- Ouvrir un terminal dans le dossier `elasticsearch-[version]`
- Lancer la commande `bin/elasticsearch` (`bin\elasticsearch.bat` sur Windows)

Kibana

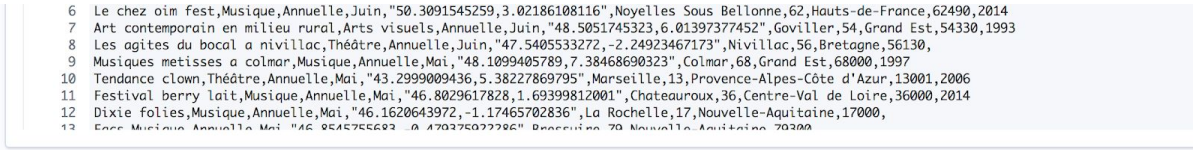
- Ouvrir un terminal dans le dossier `kibana-[version]`
- Lancer la commande `bin/kibana` (`bin\kibana.bat` sur Windows)

[2] Elasticsearch : https://www.elastic.co/guide/en/elasticsearch/current/starting-elasticsearch
Kibana : https://www.elastic.co/guide/en/kibana/current/start-stop

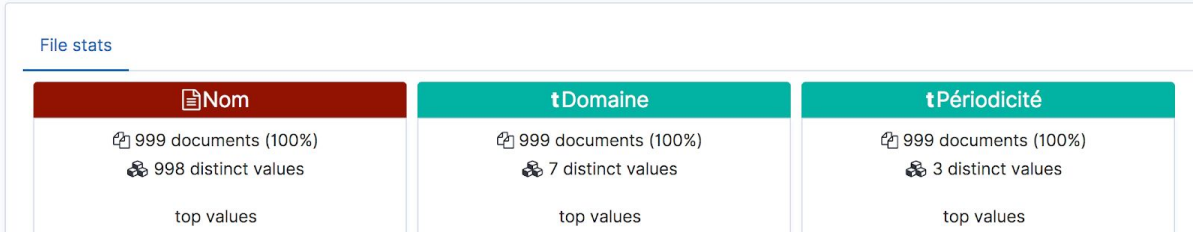
IMPORTATION DE DONNÉES CSV



- Sur la page d'accueil de Kibana, cliquer sur "*Upload data from log file*" (ou dans l'onglet "*Machine learning*")
- Importer le fichier `.csv` choisi : Kibana parse le fichier automatiquement

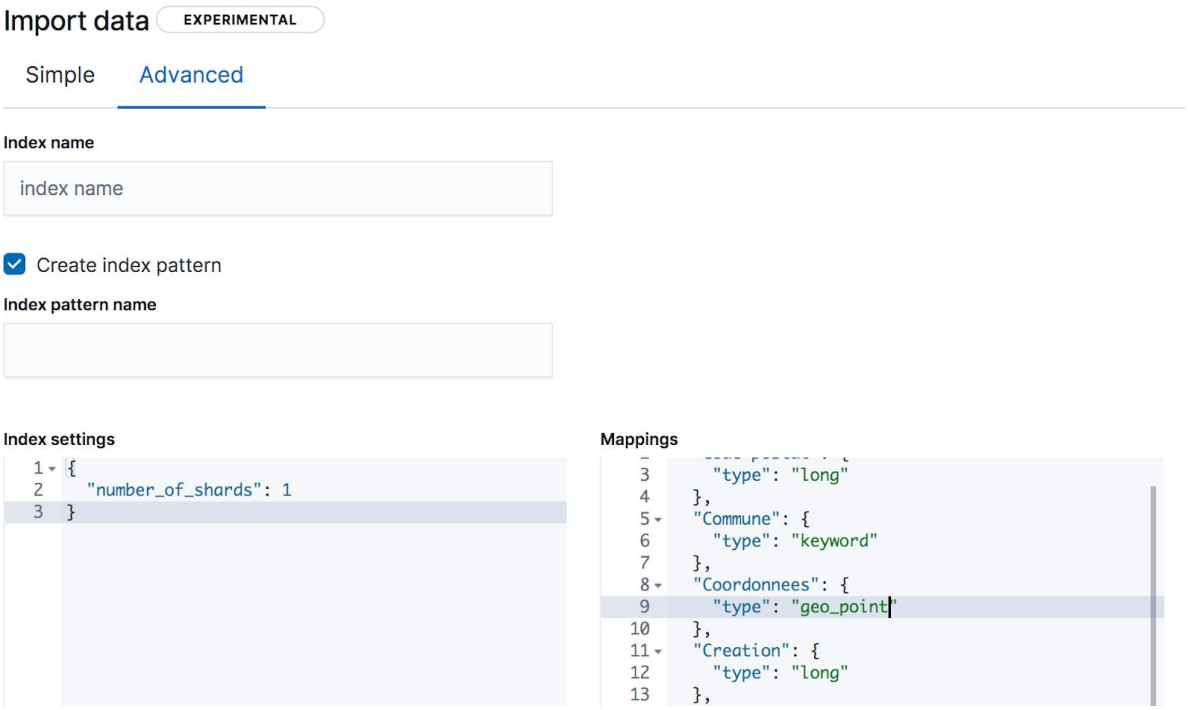


```
Le chez oim fest,Musique,Annuelle,Juin,"50.3091545259,3.02186108116",Noyelles Sous Bellonne,62,Hauts-de-France,62490,2014
Art contemporain en milieu rural,Arts visuels,Annuelle,Juin,"48.5051745323,6.01397377452",Goviller,54,Grand Est,54330,1993
Les agites du bocal a nivillac,Théâtre,Annuelle,Juin,"47.5405533272,-2.24923467173",Nivillac,56,Bretagne,56130,
Musiques metisses a colmar,Musique,Annuelle,Mai,"48.1099405789,7.38468690323",Colmar,68,Grand Est,68000,1997
Tendance clown,Théâtre,Annuelle,Mai,"43.2999009436,5.38227869795",Marseille,13,Provence-Alpes-Côte d'Azur,13001,2006
Festival berry lait,Musique,Annuelle,Mai,"46.8029617828,1.69399812001",Chateauroux,36,Centre-Val de Loire,36000,2014
Dixie folies,Musique,Annuelle,Mai,"46.1620643972,-1.17465702836",La Rochelle,17,Nouvelle-Aquitaine,17000,
```



- Pour modifier le titre des colonnes, le caractère délimiteur, etc., cliquer sur "*Override settings*" : cliquer ensuite sur "*Import*"

- Pour ajuster le type des différentes colonnes (notamment pour permettre des traitements ultérieurs), choisir le mode “*Advanced*” afin de modifier le mapping du fichier



Cela est nécessaire notamment pour forcer la reconnaissance de certains types de données (pour la liste de tous les types utilisés dans ElasticSearch : [c'est ici](#)). Par exemple :

- Géographiques (format “`lat,long`”) : changer le type pour “`geo_point`”
- Temporelles ([formats](#)) : changer le type pour “`date`” et préciser un format. Par exemple :

```
"Creation": {
  "type": "date",
  "format": "yyyy"
```

- Choisir un nom (sans majuscule ni accent) pour son ensemble de données dans le champ “*Index name*” puis cliquer sur le bouton “*Import*”
- Les données sont ensuite converties en NDJSON et chargée dans votre base de données ElasticSearch ![3]

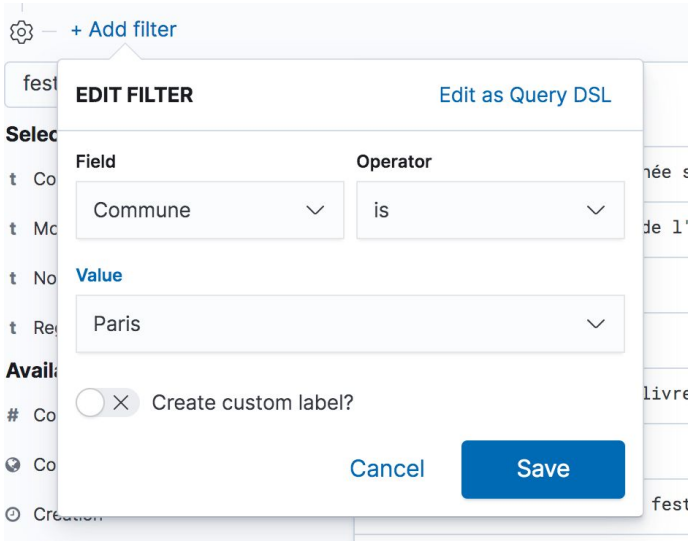
De plus amples informations à cette [adresse](#).

[3] Trois champs ont été créés automatiquement par elasticsearch : “`_id`”, “`_index`” et “`_type`”. Chaque enregistrement dispose ainsi d’un identifiant, et connaît à la fois l’index auquel il appartient et son type.

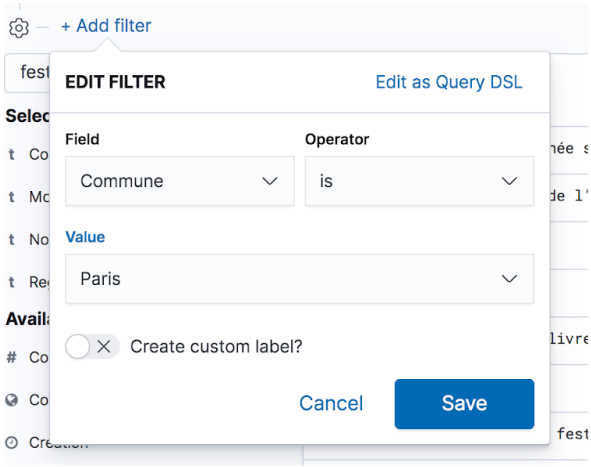
ONGLET “DISCOVER”

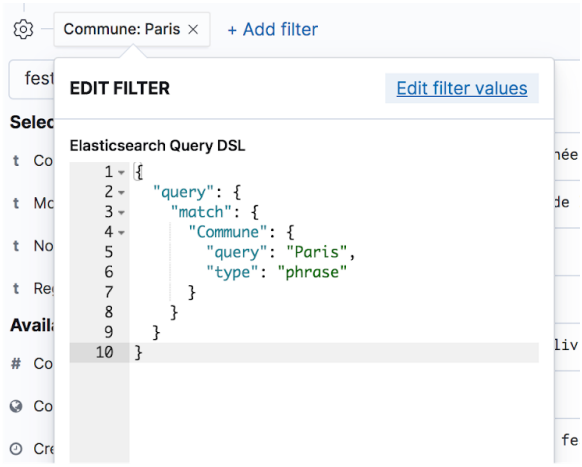
- Réorganiser les données en sélectionnant les champs à faire apparaître dans la barre latérale

- Ajouter/Modifier des filtres pour dégager des ensembles dans les données présentes : l’autocomplétion (au clic sur “*Add filter*”) permet de formuler des filtres très facilement. Le nombre de “*hits*” qui apparaît au dessus des résultats correspond au nombre d’enregistrement de l’index correspondant à la recherche.



Tip : Une fois appliqué, en cliquant sur le filtre puis dans “*Edit filter*”, lorsque l’on clique sur “*Edit as Query DSL*”[4], le filtre est formulé en tant que requête au format JSON : cela permet d’avoir des requêtes écrites automatiquement.



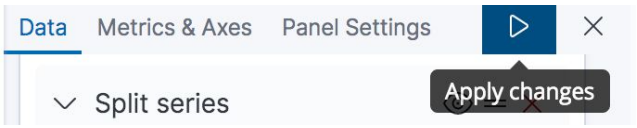


- La barre de recherche propose également de l’autocomplétion pour formuler des requêtes.
- Si l’on souhaite conserver un état “filtré” d’un index, il suffit de cliquer sur “Save” en haut de l’écran. Cette requête peut être ensuite utilisée pour réaliser des visualisations sur des sous-ensembles d’un index.

[4] “*Query DSL*” est le langage de requête utilisé par Elasticsearch (équivalent à SQL pour une base de données relationnelle)

ONGLET “VISUALIZE”

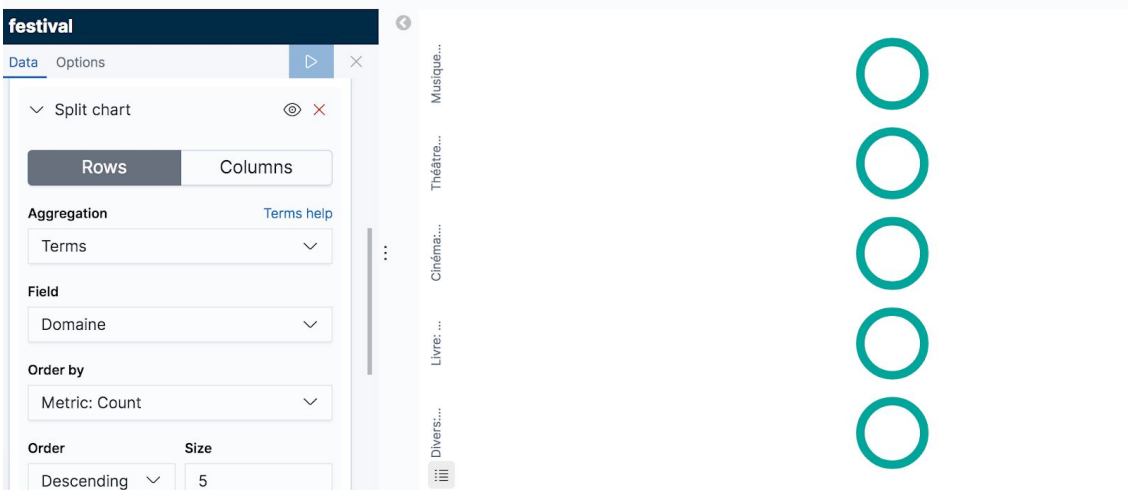
- Cliquer sur “*Create new visualization*” et sélectionner le type de visualisation à réaliser
- Sélectionner l’index ou le sous-ensemble d’index (requête enregistrée) que l’on souhaite visualiser
- Configurer son graphique puis cliquer sur “*Apply changes*” pour voir la visualisation se mettre à jour

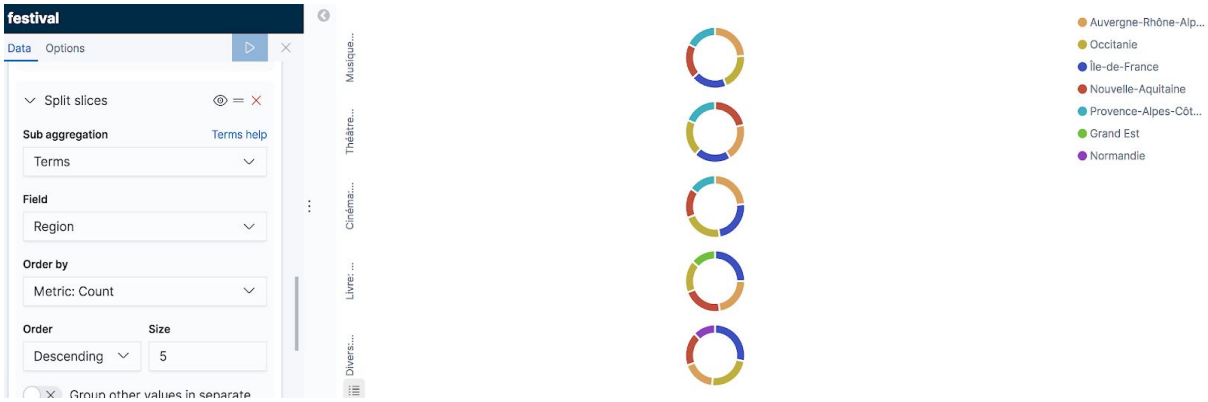


Tip : il est aussi possible d’avoir une vue d’ensemble sur un index dans l’onglet “*Machine learning*” avec l’outil “*Data visualizer*”

Pie chart

- Choisir la mesure (“*Metrics*”) en cliquant dessus
- Configurer les *buckets* et la manière de les visualiser : “*Split slices*” répartit les données d’un même graphique dans plusieurs champs tandis que “*Split chart*” opère la répartition des données dans des graphiques différents. “*Split chart*” doit être fait avant “*Split slices*”
- L'agrégation de “*Terms*” correspond à la création de sous-ensemble de données partageant la même valeur pour un champ
  “*Order by*” spécifie quels *buckets* doivent ressortir en premier : ceux premiers dans l’ordre alphabétique ? Premier en termes de nombre ? etc.
  Ne pas oublier d’augmenter la taille – de base limitée à 5 – pour obtenir des données plus complètes (ce que je n’ai pas fait sur les *screenshots*)

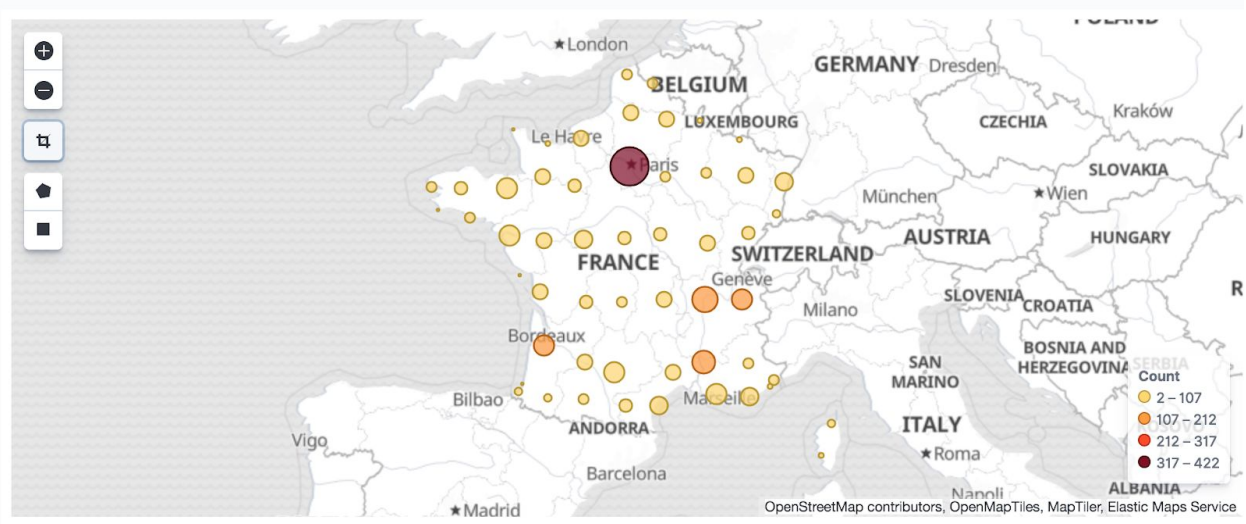




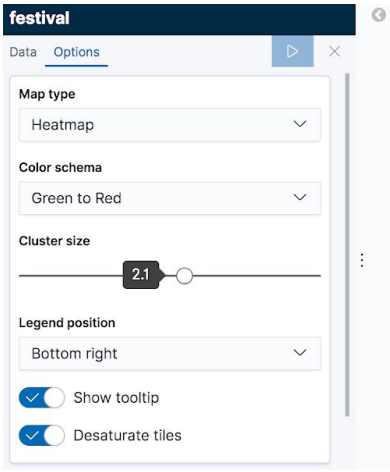
Coordinate map

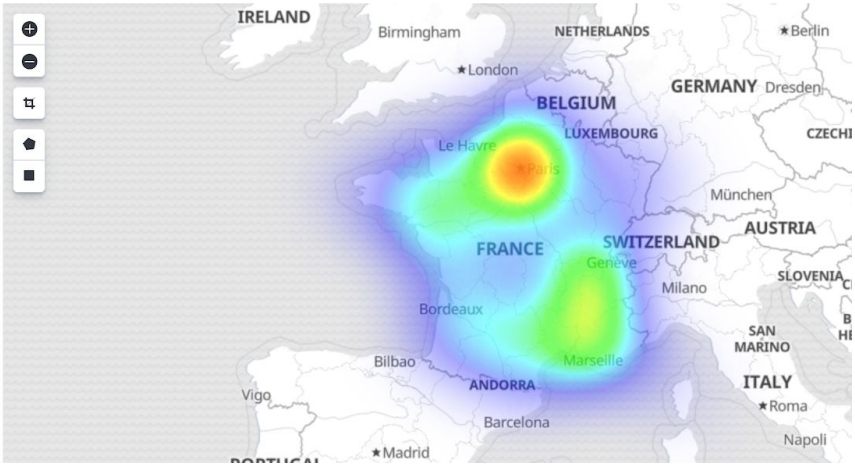
La carte ne peut être réalisée que si l'index à visualiser possède un de type "`geo_point`".

- Choisir la mesure ("*Metrics*")
- Définir un *bucket* à visualiser
- Cliquer sur l'icône "*Fit data bounds*" (pictogramme de recadrage) pour resserrer le zoom autour des points présents sur la carte. En zoomant, la carte se précise



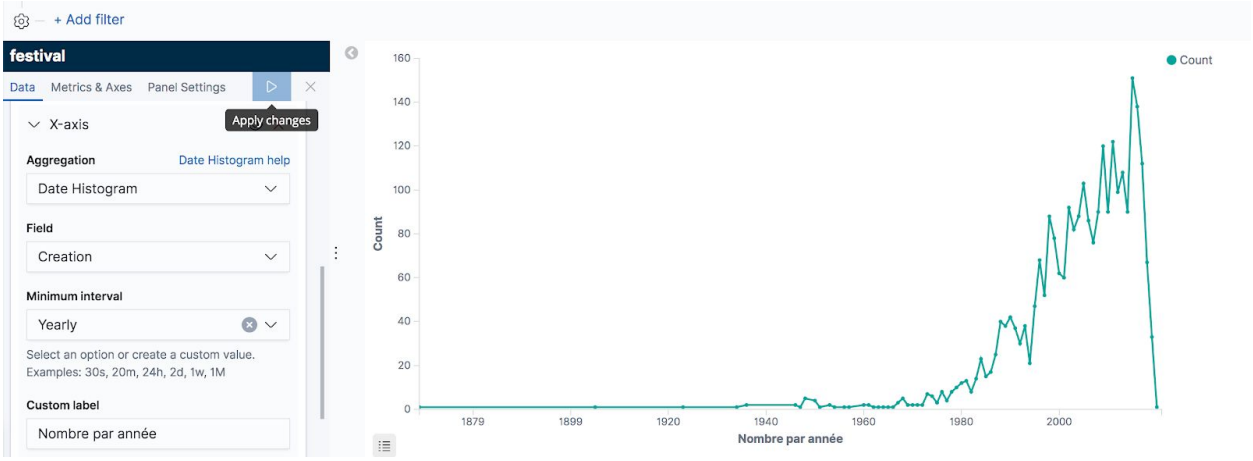
- Avec l'outil "*Draw a polygon*" (pictogramme en pentagone) ou "*Draw a rectangle*" (pictogramme en carré), il est possible de filtrer pour ne garder que les données présentes dans un certain périmètre
- Dans les options, il est possible de changer les paramètres d'affichage et de créer une carte de chaleur



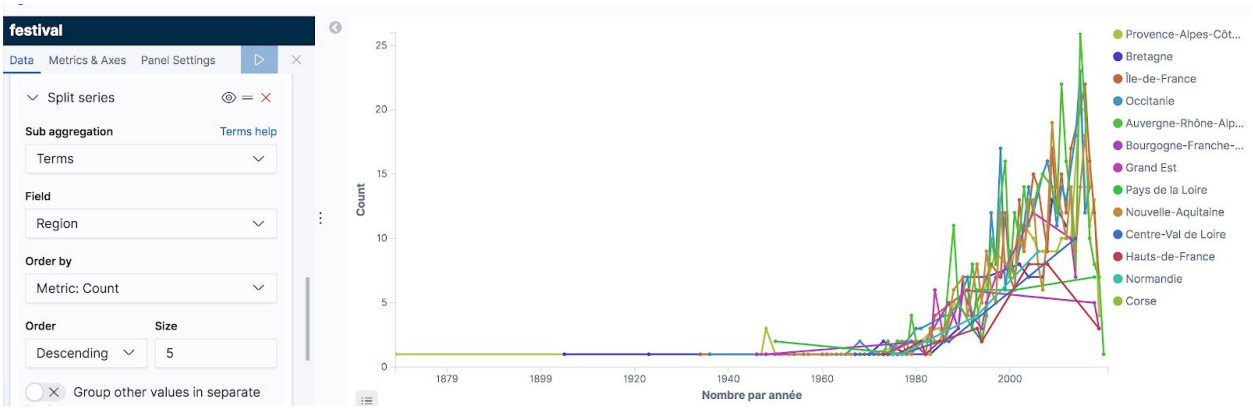


Line chart

- Choisir la mesure (“*Metrics*”) associée à l’axe vertical (“*Y-axis*”)
- Définir le *bucket* à visualiser sur l’axe horizontal
- Si l’on dispose d’un champ de type date, il est possible de choisir de réaliser un “*Date histogram*” : on va définir un intervalle de temps pour lequel la mesure sera effectué[5]



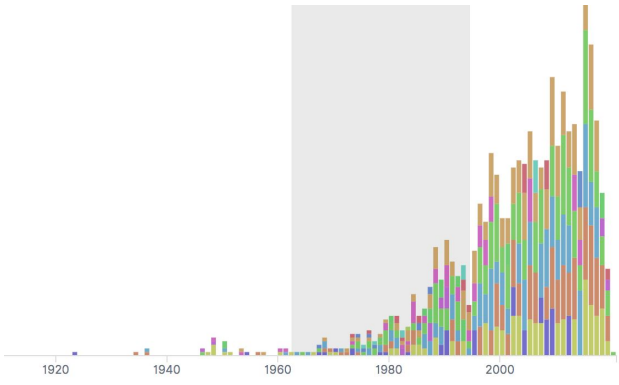
- Pour détailler le graphique en fonction d’un autre champ : ajouter un *bucket* de type “*Split series*”



- Si l’on souhaite changer le type de graphique : cliquer sur “Metrics & Axes” en haut du panneau de configuration et sélectionner le mode de visualisation

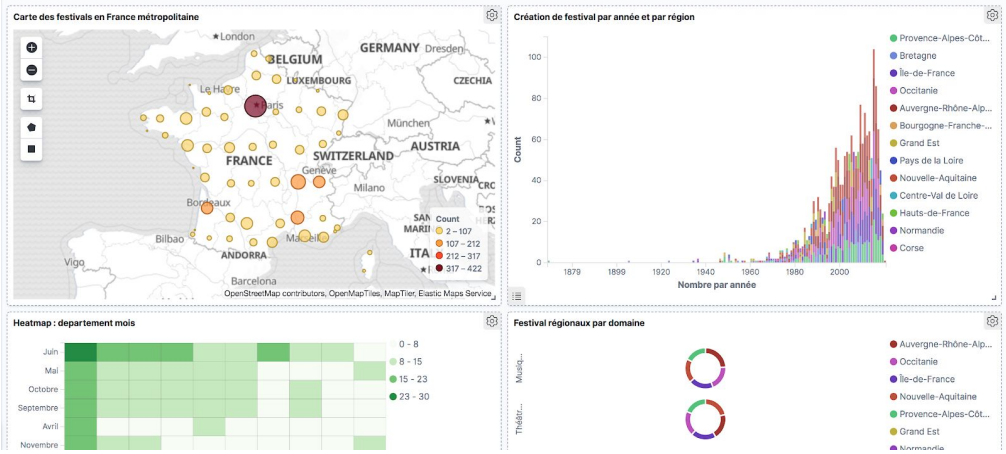
[5] Par exemple, pour un jeu de données comprenant des dates précises, la fonction histogramme permet de considérer ces dates par *cluster* d’année plutôt qu’individuellement

- Pour zoomer sur une zone précise du graphique, utiliser la sélection avec la souris

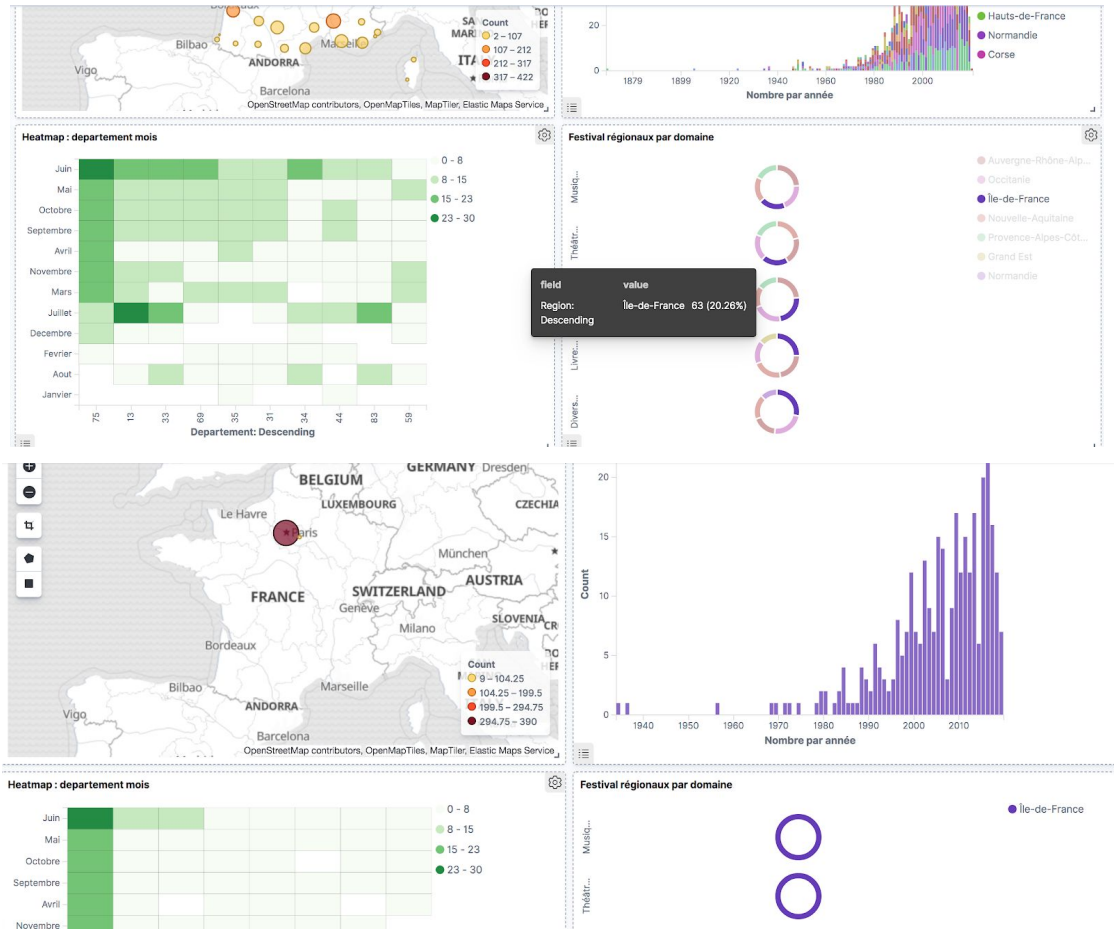


ONGLET “DASHBOARD”

- Créer un *dashboard* puis cliquer sur le bouton “*Add*” : sélectionner toutes les visualisations à disposer sur le *dashboard*
- Disposer les différentes visualisations

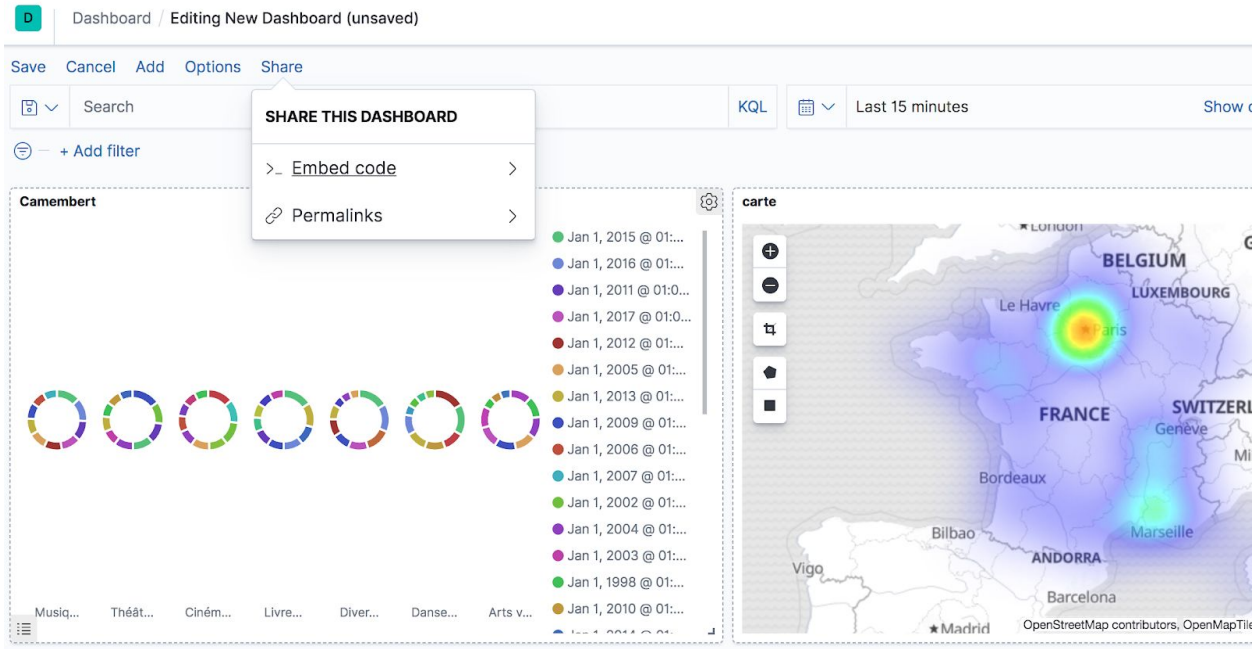


- Sélectionner une donnée pour opérer un filtrage sur l’ensemble du *dashboard* : lorsque la donnée cliquée est ambigüe, une boîte de dialogue s’affiche pour proposer les différents filtres possibles

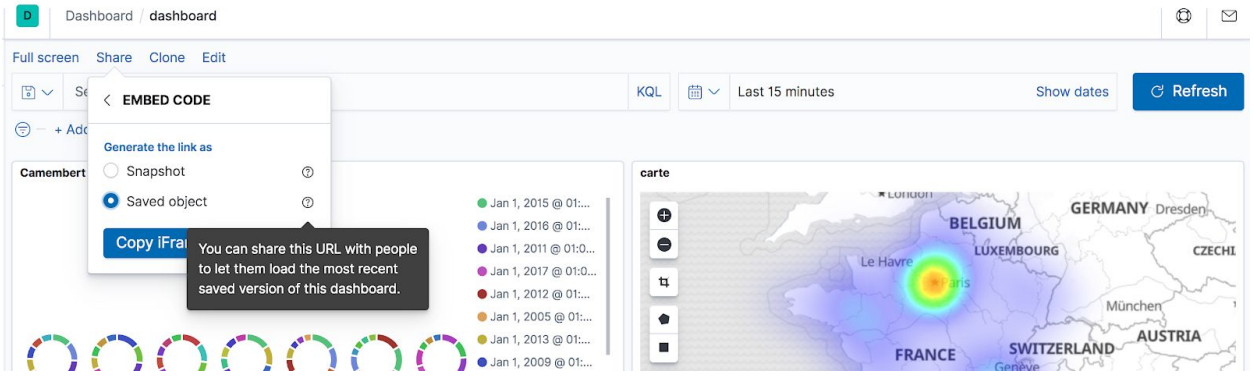


EXPORTER UNE VISUALISATION

- Pour exporter un *dashboard*, il est possible de cliquer sur le bouton “*Share*” en haut de l’écran pour avoir accès à différents modes d’export



- Le mode “*Embed code*” permet d’obtenir le code d’une *iframe* HTML pour l’insert du dashboard au sein d’une page HTML. Une iframe est une fenêtre sur l’application Kibana, ce qui signifie que si Elasticsearch et Kibana ne sont pas lancés sur le serveur où le dashboard est créé, le dashboard intégré à la page HTML ne s’affichera pas ; ainsi, un *dashboard* créé sur un *localhost* ne sera visible que depuis le même ordinateur



- Pour un affichage minimal du dashboard, copier le code dans un template de page HTML comme suit :

```
<!doctype html>
<html lang='en'>
	<head><meta charset='utf-8'></head>
	<body>
		<!-- copier le code généré par Kibana ici -->
	</body>
</html>
```

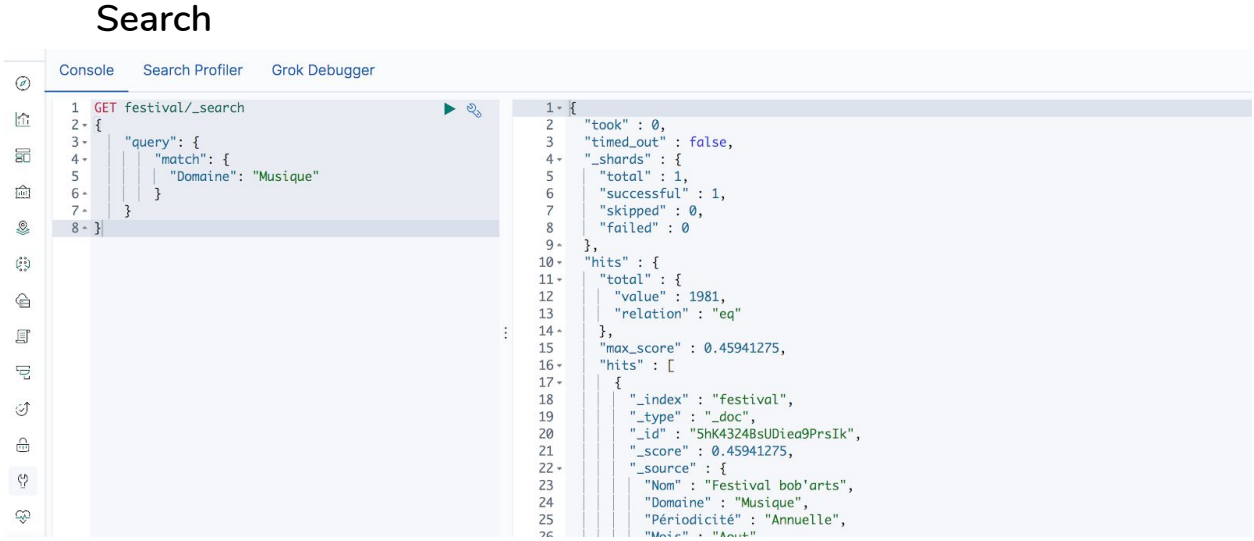
- Enregistrer ce code dans un fichier au format `.html` puis l'ouvrir dans un navigateur pour voir le résultat.

L'option "*Permalink*" permet d'obtenir directement un lien partageable. Les mêmes fonctionnalités existent également pour les visualisations seules.

L'export de visualisation au format `.pdf`, `.svg` et `.png` est possible avec l'extension *Reporting* contenu dans le X-Pack, c'est-à-dire la version payante de la suite Elastic.

ONGLET “DEV TOOLS”

C’est l’onglet pour écrire des requêtes dans le langage d’elasticsearch. Pour se familiariser avec ces requêtes : quelques exemples simples. Des nombreuses fonctionnalités d’autocomplétion sont proposées ce qui rend la formulation des requêtes particulièrement facile[6] : attention, le JSON doit être parfaitement valide, c’est-à-dire, interdiction d’utiliser des guillemets simples pour les chaînes de caractères et d’oublier une virgule à la fin d’une liste.



La réponse renvoyée par l’API d’elasticsearch est toujours structurée de la même manière : c’est une longue chaîne de caractères en JSON avec un certain nombre de clefs :

- “`took`” : temps en milliseconde pris par elasticsearch pour répondre à la requête
- “`time_out`” : si la requête a expiré (généralement “`false`”)
- “`_shards`” : informations vis-à-vis des *shards* utilisés (“éclats” de la base de données)
- “`hits`” : informations relatives aux réponses de la requête
    - “`total`” : nombre de résultats obtenus (“`value`”) et si le nombre donné est exact (“`eq`”) ou s’il est en dessous du véritable nombre de réponses (“`gte`”)
    - “`max_score`” : score maximal récolté parmi les réponses, c’est-à-dire le score de la réponse la plus pertinente par rapport à la requête
    - “`hits`” : liste (*array* `[...]`) des résultats obtenus

Chaque *hit* est contenu dans un “dictionnaire” (*object* `{...}`) structuré ainsi :

- “`_index`” : nom de l’index auquel appartient le *hit*
- “`_type`” : son type
- “`_id`” : son identifiant
- “`_score`” : son score vis-à-vis de la requête
- “`_source`” : ses champs de métadonnées

[6] Kibana garde un historique des requêtes

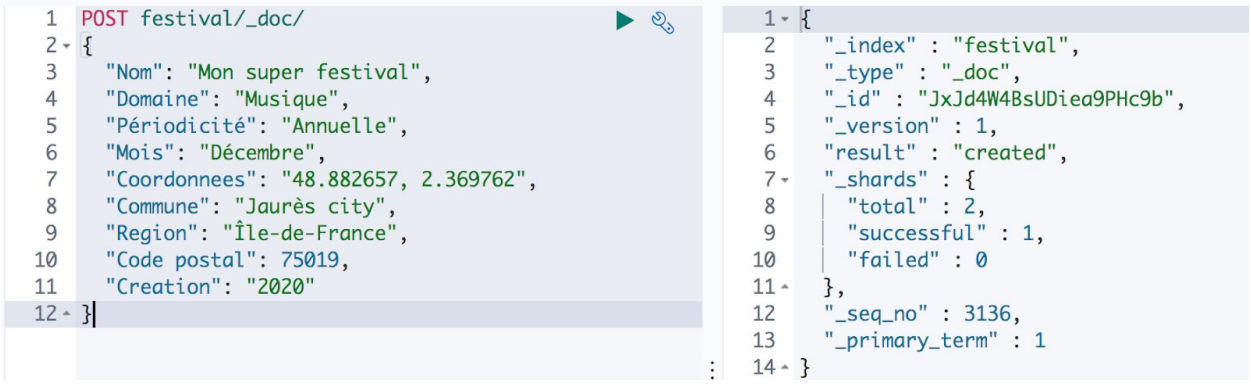
La requête de base pour retirer tous les enregistrements d'un index se formule ainsi :

```
GET festival/_search
{
    "query": {
        "match_all": {}
    }
}
```

Le nombre de résultats renvoyé est de base limité à 10 : pour augmenter cette taille, il suffit de rajouter une clef à la requête comme suit : `"size": ${nombre-de-resultats}`

Put

Il est aussi possible d'ajouter de nouveaux enregistrements à ses index via la console de l'onglet "Dev tools"



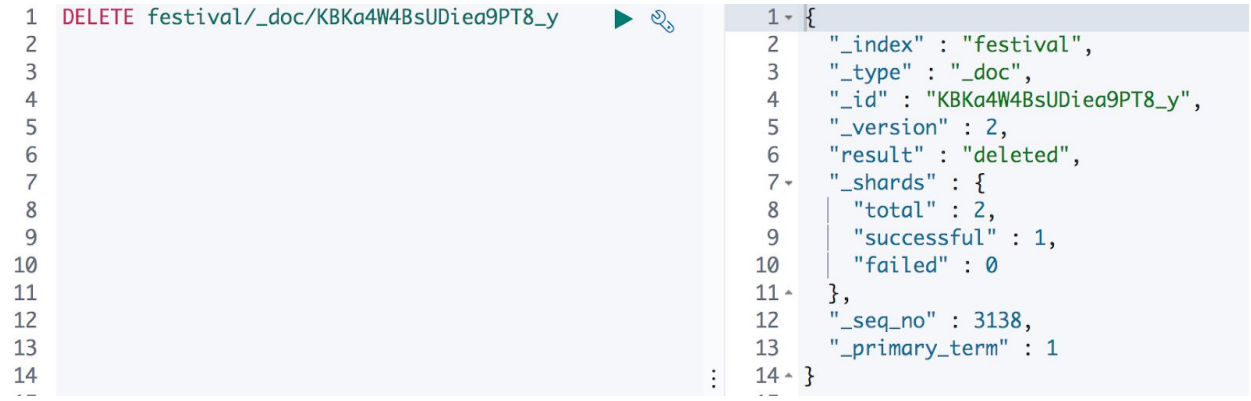
```
POST festival/_doc/
{
  "Nom": "Mon super festival",
  "Domaine": "Musique",
  "Périodicité": "Annuelle",
  "Mois": "Décembre",
  "Coordonnees": "48.882657, 2.369762",
  "Commune": "Jaurès city",
  "Region": "Île-de-France",
  "Code postal": 75019,
  "Creation": "2020"
}
```

```
{
  "_index" : "festival",
  "_type" : "_doc",
  "_id" : "JxJd4W4BsUDiea9PHc9b",
  "_version" : 1,
  "result" : "created",
  "_shards" : {
    "total" : 2,
    "successful" : 1,
    "failed" : 0
  },
  "_seq_no" : 3136,
  "_primary_term" : 1
}
```

Les visualisations et *dashboards* créés se mettront à jour en fonction des ajouts.

Delete

De la même manière, il est possible de retirer des données d'un index en donnant son identifiant



```
DELETE festival/_doc/KBKa4W4BsUDiea9PT8_y
```

```
{
  "_index" : "festival",
  "_type" : "_doc",
  "_id" : "KBKa4W4BsUDiea9PT8_y",
  "_version" : 2,
  "result" : "deleted",
  "_shards" : {
    "total" : 2,
    "successful" : 1,
    "failed" : 0
  },
  "_seq_no" : 3138,
  "_primary_term" : 1
}
```

Communiquer avec l'API hors Kibana

Pour formuler une requête SEARCH dans une URL (dans un navigateur) :

- `"http://${port-elasticsearch}/"`
  `+ "${nom-index}/_search?source_content_type=application/"`
  `+ "json&source=${requete}"`

Par exemple :

```
http://localhost:9200/festival/_search?source_content_type=application/json&source={"query":{"match_all": {}}}
```

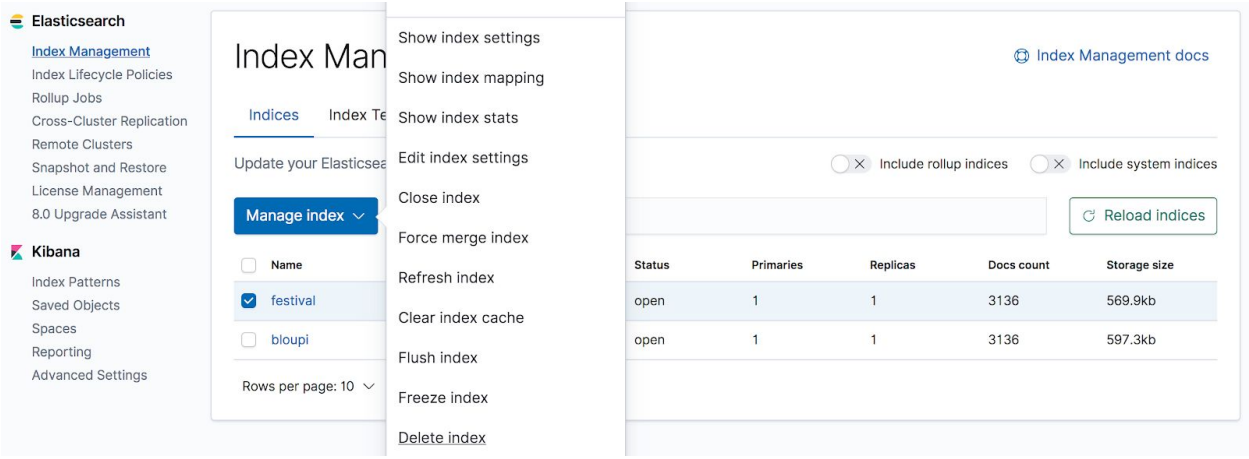
Pour formuler une requête SEARCH sous forme cURL (dans un terminal)[7] :

```
curl -XGET "http://localhost:9200/festival/_search" -H 'Content-Type: application/json'-d'{"query":{"match_all":{}}}'
```

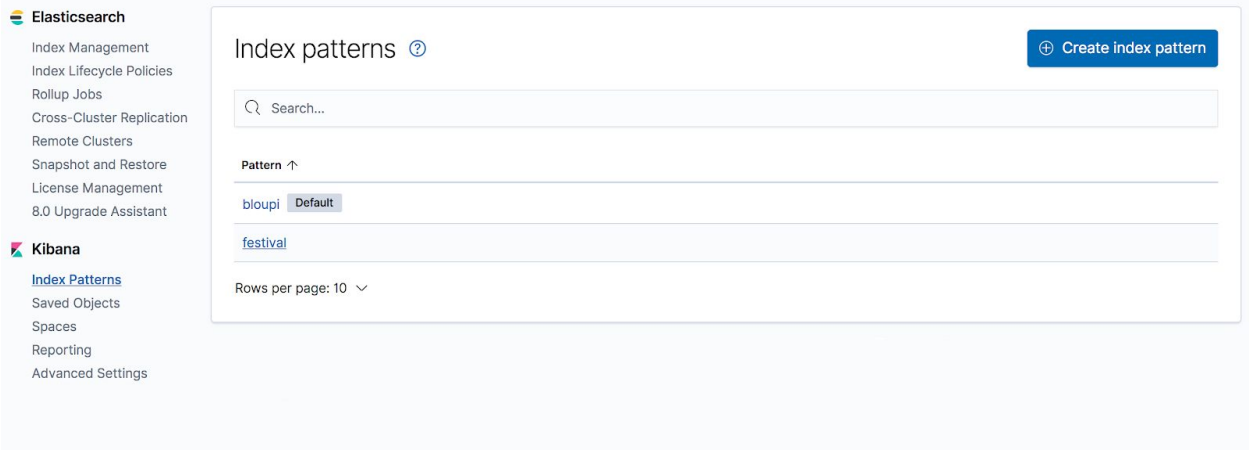
[7] En cliquant sur la clef à molette dans la console Kibana, il y a une fonctionnalité pour transposer la requête en cURL

SUPPRIMER UN INDEX

- Se rendre dans l'onglet "*Management*", tout en bas de la barre latérale
- Dans la partie "Elasticsearch", aller dans l'onglet "Index Management"
- Sélectionner l'index à supprimer et cliquer sur "*Manage index*" puis "*Delete index*"



- Dans la partie "Kibana", aller dans l'onglet "*Index Patterns*"
- Cliquer sur l'index à supprimer



- Cliquer sur l'icône de suppression