## Installation

curl -s -H "Content-Type: application/json" -XPOST "http://localhost:9200/default/\_bulk?pretty" -u 'elastic:secret' --data-binary @big-movies-elastic.json

Les documents ont la structure suivante :

{

**"fields"**: {

**"directors"**: [

"Joseph Gordon-Levitt"

],

**"release\_date"**: "2013-01-18T00:00:00Z",

**"rating"**: 7.4,

**"genres"**: [

"Comedy",

"Drama"

],

**"image\_url"**: "http://ia.media-imdb.com/images/M/MVNTQ3OQ@@.\_V1\_SX400\_.jpg",

**"plot"**: "A New Jersey guy dedicated to his family, friends, and church,

develops unrealistic expectations from watching porn and works to find

happ iness and intimacy with his potential true love.",

**"title"**: "Don Jon",

**"rank"**: 1,

**"running\_time\_secs"**: 5400,

**"actors"**: [

"Joseph Gordon-Levitt",

"Scarlett Johansson",

"Julianne Moore"

],

**"year"**: 2013

},

**"id"**: "tt2229499",

**"type"**: "add"

}

Interrogation

Exercice : Trouver les films « star Wars »

movies/movie/\_search

{

**"query"**: {

**"match"**: {

**"fields.title"**: "Star Wars"

}

}

}

Autre méthode

\_search?q=fields.title:Star+Wars

Résultat : Bride Wars et Dark Wars

Pour ne récupérer que les données pertinentes :

Variante :

{

**"query"**: {

**"match\_phrase"**: {

**"fields.title"**: "Star Wars"

}

}

}

2e variante :

{

**"query"**: {

**"match"**: {

**"fields.title"**: {

**"query"**: "Star Wars",

**"operator"**: "and"

}

}

}

}

Vous pouvez limiter la quantité d’informations qui se trouvent dans le champ \_source

de chacun des résultats comme ceci :

{

**"\_source"**: {

**"includes"**: [

"\*.title",

"\*.directors"

],

**"excludes"**: [

"\*.actors\*",

"\*.genres"

]

},

**"query"**: {

**"match"**: {

**"fields.title"**: "Star Wars"

}

}

}

Exercice : Trouver les Films “Star Wars” dont le réalisateur (directors) est “George Lucas”

{

**"query"**: {

**"bool"**: {

**"should"**: [

{

**"match"**: {

**"fields.title"**: "Star Wars"

}

},

{

**"match"**: {

**"fields.directors"**: "George Lucas"

}

}

]

}

}

}

Pour une recherche exacte :

{

**"query"**: {

**"bool"**: {

**"should"**: [

{

**"match\_phrase"**: {

**"fields.title"**: "Star Wars"

}

},

{

**"match"**: {

**"fields.directors"**: "George Lucas"

}

}

]

}

}

}

Variante :

{

**"\_source"**: {

**"includes"**: [ "\*.title" ],

**"excludes"**: [ "\*.actors\*", "\*.genres", "\*.directors" ]

},

**"query"**: {

**"bool"**: {

**"should"**: [

{

**"match"**: {

**"fields.title"**: {

**"query"**: "Star Wars",

**"operator"**: "and"

}

}

},

{

**"match"**: {

**"fields.directors"**: {

**"query"**: "Georges Lucas",

**"operator"**: "and"

}

}

}

]

}

}

}

ou

\_search?q=fields.title:Star+Wars directors:George+Lucas

Exercice : Films dans lesquels “Harrison Ford” a joué

{

**"query"**: {

**"match"**: {

**"fields.actors"**: "Harrison Ford"

}

}

}

ou \_search?q=fields.actors:Harrison+Ford

Exercice : Films dans lesquels “Harrison Ford” a joué dont le résumé (plot) contient “Jones”.

{**"query"**:{

**"bool"**: {

**"should"**: [

{ **"match"**: { **"fields.actors"**: "Harrison Ford" }},

{ **"match"**: { **"fields.plot"**: "Jones" }}

]

}}}

ou \_search?q=fields.actors=Harrison+Ford fields.plot:Jones

Exercice : Films dans lesquels “Harrison Ford” a joué dont le résumé (plot) contient “Jones” mais sans le mot “Nazis”

{**"query"**:{

**"bool"**: {

**"should"**: [

{ **"match"**: { **"fields.actors"**: "Harrison Ford" }},

{ **"match"**: { **"fields.plot"**: "Jones" }}

],

**"must\_not"** : { **"match"** : {**"fields.plot"**:"Nazis"}}

}}}

ou \_search?q=actors=Harrison+Ford plot:Jones -plot:Nazis

Exercice : Films de “James Cameron” dont le rang devrait être inférieur à 1000 (boolean + range query).

{**"query"**:{

**"bool"**: {

**"should"**: [

{ **"match"**: { **"fields.directors"**: "James Cameron" }},

{ **"range"**: { **"fields.rank"**: {**"lt"**:1000 }}}

]

}}}

Exercice : Films de “James Cameron” dont le rang **doit** être inférieur à 100

{**"query"**:{

**"bool"**: {

**"should"**: { **"match"**: { **"fields.directors"**: "James Cameron" }},

**"must"**: { **"range"**: { **"fields.rank"**: {**"lt"**:100 }}}

}}}

Exercice : Films de “Quentin Tarantino” dont la note (rating) doit être supérieure à 5, sans être un film d’action ni un drame.

{

**"\_source"**: {

**"includes"**: [

"\*.title"

],

**"excludes"**: [

"\*.actors\*"

]

},

**"query"**: {

**"bool"**: {

**"must"**: [

{

**"match\_phrase"**: {

**"fields.directors"**: "Quentin Tarantino"

}

},

{

**"range"**: {

**"fields.rating"**: {

**"gte"**: 5

}

}

}

],

**"must\_not"**: [

{

**"match"**: {

**"fields.genres"**: "Action"

}

},

{

**"match"**: {

**"fields.genres"**: "Drama"

}

}

]

}

}

}

Exercice : Films de “J.J. Abrams” sortis (released) entre 2010 et 2015

{

**"query"**: {

**"bool"**:{

**"must"**: {**"match"**: {**"fields.directors"**: "J.J. Abrams"}},

**"filter"**: {

**"range"**: {

**"fields.release\_date"**: { **"from"**: "2010-01-01", **"to"**: "2015-12-31"}

}

}

}

}

}

Aggregats

Un exemple, avec le nombre de films par année :

{

**"size"**: 0,

**"aggs"** : { **"nb\_par\_annee"** : {

**"terms"** : {**"field"** : "fields.year"}

}

}

}

Exercice : Donner la note (rating) moyenne des films.

{**"aggs"** : {

**"note\_moyenne"** : {

**"avg"** : {**"field"** : "fields.rating"}

}}}

Exercice : Donner la note (rating) moyenne, et le rang moyen des films de George Lucas.

{**"query"** :{

**"match"** : {**"fields.directors"** : "George Lucas"}

}

,**"aggs"** : {

**"note\_moyenne"** : {

**"avg"** : {**"field"** : "fields.rating"}

},

**"rang\_moyen"** : {

**"avg"** : {**"field"** : "fields.rank"}

}

}}

Exercice : Donnez la note (rating) moyenne des films par année. Attention, il y a ici une imbrication d’agrégats.

{**"aggs"** : {

**"group\_year"** : {

**"terms"** : {

**"field"** : "fields.year"

},

**"aggs"** : {

**"note\_moyenne"** : {

**"avg"** : {**"field"** : "fields.rating"}

}}

}}}

Exercice : Donner la note (rating) minimum, maximum et moyenne des films par année.

{**"aggs"** : {

**"group\_year"** : {

**"terms"** : {

**"field"** : "fields.year"

},

**"aggs"** : {

**"note\_moyenne"** : {**"avg"** : {**"field"** : "fields.rating"}},

**"note\_min"** : {**"min"** : {**"field"** : "fields.rating"}},

**"note\_max"** : {**"max"** : {**"field"** : "fields.rating"}}

}

}}}

Exercice : Donner le rang (rank) moyen des films par année et trier par ordre décroissant.

{**"aggs"** : {

**"group\_year"** : {

**"terms"** : {

**"field"** : "fields.year",

**"order"** : { **"rating\_moyen"** : "desc" }

},

**"aggs"** : {

**"rating\_moyen"** : {

**"avg"** : {**"field"** : "fields.rating"}

}}

}}}

Exercice : Compter le nombre de films par tranche de note (0-1.9, 2-3.9, 4-5.9…). Indice : group\_range.

{**"aggs"** : {

**"group\_range"** : {

**"range"** : {

**"field"** : "fields.rating",

**"ranges"** : [

{**"to"** : 1.9},

{**"from"** : 2, **"to"** : 3.9},

{**"from"** : 4, **"to"** : 5.9},

{**"from"** : 6, **"to"** : 7.9},

{**"from"** : 8}

]

}

}}}

Pertinence : Boosting

Quand on effectue des recherches sur plus d’un champ, il peut rapidement devenir pertinent de donner davantage de poids à l’un ou l’autre de ces champs, de façon à améliorer les résultats de recherche.

Saisissez la commande suivante et observez la position d’American Grafiti dans le classement, avec et sans l’option « boost ». Que se passe-t-il ?

{

**"\_source"**: {

**"includes"**: [

"\*.title"

],

**"excludes"**: [

"\*.actors\*",

"fields.genres",

"fields.directors"

]

},

**"query"**: {

**"bool"**: {

**"should"**: [

{

**"match"**: {

**"fields.title"**: {

**"query"**: "Star Wars",

**"boost"**: 4

}

}

},

{

**"match"**: {

**"fields.directors"**: {

**"query"**: "George Lucas"

}

}

}

]

}

}

}

Avec le boosting, American Grafiti est 9e, derrière Bride Wars, mieux classé car le boosting favorise la correspondance avec (au moins) un des mots du titre.

Exemple : avec la requête ci-dessous, on ne récupère que les films dont le titre contient Star Wars, mais l’on pondère négativement avec negative boost le réalisateur JJ Abrams, dont le film doit apparaître en queue de classement

{

**"\_source"**: {

**"includes"**: [

"\*.title"

],

**"excludes"**: [

"\*.actors\*"

]

},

**"query"**: {

**"boosting"**: {

**"positive"**: {

**"query"**: {

**"match\_phrase"**: {

"fields.title": {

"query": "Star Wars",

"boost": 2

}

}

}

},

"negative": {

"match": {

"fields.directors": "Abrams"

}

},

"negative\_boost": 0.5

}

}

}

}

Si les documents contiennent des valeurs numériques comme la popularité (les likes d’un statut de réseau social, le nombre d’achats d’un produit donné), ou une note, il est possible d’utiliser cet indicateur pour pondérer les documents.

On utilise pour cela field\_value\_factor. Avec nos documents, nous pouvons proposer une pondération avec la note (ce qui revient à ordonner par rating) :

{

**"\_source"**: {

**"includes"**: [

"\*.title",

"\*.rating"

],

**"excludes"**: [

"\*.actors\*"

]

},

**"query"**: {

**"function\_score"**: {

**"query"**: {

**"match\_phrase"**: {

**"fields.directors"**: {

**"query"**: "Sergio Leone"

}

}

},

**"functions"**: [

{

**"field\_value\_factor"**: {

**"field"**: "fields.rating"

}

}

]

}

}

}

Exercice : les films de James Cameron en pondérant négativement ceux qui durent plus de deux heures (choisissez par exemple la fonction exponentielle)

{

**"\_source"**: {

**"includes"**: [

"fields.title",

"fields\*.rating",

"fields.running\_time\_secs"

],

**"excludes"**: [

"fields\*.actors\*",

"fields\*.genres"

]

},

**"query"**: {

**"function\_score"**: {

**"query"**: {

**"match\_phrase"**: {

**"fields.directors"**: {

**"query"**: "James Cameron"

}

}

},

**"functions"**: [

{

**"exp"**: {

**"fields.running\_time\_secs"**: {

**"origin"**: "7200",

**"scale"**: "200",

**"decay"**: 0.5

}

}

}

]

}

}

}

Exercice : les meilleures comédies romantiques (il s’agit d’effectuer un simple tri par rating)

{

**"size"**: 25,

**"\_source"**: {

**"includes"**: [

"fields.title",

"fields\*.genres"

],

**"excludes"**: [

"fields\*.actors\*",

"fields\*.rating",

"fields.running\_time\_secs"

]

},

**"query"**: {

**"bool"**: {

**"must"**: [

{

**"match"**: {

**"fields.genres"**: "Romance"

}

}

]

}

},

**"sort"**: [

{

**"fields.rating"**: {

**"order"**: "desc"

}

}

]

}

Exercice : les films réalisés par Clint Eastwood, en affichant d’abord ceux dans lesquels il joue

{

**"size"**: 21,

**"\_source"**: {

**"includes"**: [

"\*.title",

"\*.actors",

"\*.directors"

],

**"excludes"**: [

"\*.genres"

]

},

**"query"**: {

**"bool"**: {

**"must"**: {

**"match\_phrase"**: {

**"fields.directors"**: {

**"query"**: "Clint Eastwood"

}

}

},

**"should"**: {

**"match\_phrase"**: {

**"fields.actors"**: {

**"query"**: "Clint Eastwood",

**"boost"**: 4

}

}

}

}

}

}

Exercice : les films de Sergio Leone, en les ordonnant du plus récent au plus ancien (deux requêtes possibles, avec boosting ou sans boosting mais avec un tri)

Sans boosting

{

**"size"**: 25,

**"\_source"**: {

**"includes"**: [

"fields.title",

"fields\*.directors",

"fields\*.release\_date"

],

**"excludes"**: [

"fields\*.actors\*",

"fields\*.rating",

"fields\*.genres",

"fields.running\_time\_secs"

]

},

**"query"**: {

**"bool"**: {

**"must"**: [

{

**"match\_phrase"**: {

**"fields.directors"**: "Sergio Leone"

}

}

]

}

},

**"sort"**: [

{

**"fields.release\_date"**: {

**"order"**: "desc"

}

}

]

}

Avec boosting

{

**"\_source"**: {

**"includes"**: [

"\*.title",

"\*.release\_date"

],

**"excludes"**: [

"\*.actors\*"

]

},

**"query"**: {

**"function\_score"**: {

**"query"**: {

**"match\_phrase"**: {

**"fields.directors"**: {

**"query"**: "Sergio Leone"

}

}

},

**"functions"**: [

{

**"field\_value\_factor"**: {

**"field"**: "fields.release\_date"

}

}

]

}

}

}

Exercice : les films du genre Western, en pondérant négativement ceux réalisés par Sergio Leone

{

**"size"**: 75,

**"\_source"**: {

**"includes"**: [

"\*.title"

],

**"excludes"**: [

"\*.actors\*"

]

},

**"query"**: {

**"boosting"**: {

**"positive"**: {

**"query"**: {

**"match\_phrase"**: {

**"fields.genres"**: {

**"query"**: "Western"

}

}

}

},

**"negative"**: {

**"match\_phrase"**: {

**"fields.directors"**: "Sergio Leone"

}

},

**"negative\_boost"**: 0.5

}

}

}

Exercice : les films de sport autour de la boxe, et assez courts. Indice : la durée du film se trouve dans fields.running\_time\_secs.

{

**"\_source"**: {

**"includes"**: [

"\*.title",

"fields.plot",

"\*.genres",

"fields.running\_time\_secs"

],

**"excludes"**: [

"\*.actors\*"

]

},

**"query"**: {

**"function\_score"**: {

**"query"**: {

**"bool"**: {

**"must"**: [

{

**"exists"**: {

**"field"**: "fields.running\_time\_secs"

}

},

{

**"exists"**: {

**"field"**: "fields.genres"

}

},

{

**"match"**: {

**"fields.genres"**: {

**"query"**: "Sport"

}

}

},

{

**"exists"**: {

**"field"**: "fields.plot"

}

}

],

**"should"**: [

{

**"wildcard"**: {

**"fields.plot"**: "box\*"

}

},

{

**"range"**: {

**"fields.running\_time\_secs"**: {

**"lte"**: 7200

}

}

}

]

}

}

}

}

}