

# 高 概念

似于 DSL 表 式, 聚合也有 可 合 的 法: 独立 元的功能可以被混合起来提供 需要的自定义 。 意味着只需要学 很少的基本概念, 就可以得到几乎无尽的 合。

要掌握聚合, 只需要明白 个主要的概念:

桶 (*Buckets*)

足特定条件的文 的集合

指 (*Metrics*)

桶内的文 行 算

就是全部了! 个聚合都是一个或者多个桶和零个或者多个指 的 合。翻 成粗略的SQL 句来解:

```
SELECT COUNT(color) ①
FROM table
GROUP BY color ②
```

① `COUNT(color)` 相当于指 。

② `GROUP BY color` 相当于桶。

桶在概念上 似于 SQL 的分 (GROUP BY), 而指 似于 `COUNT()`、`SUM()`、`MAX()` 等 方法。

我 深入 个概念 并且了解和 个概念相 的西。

## 桶

桶 来 就是 足特定条件的文 的集合:

- 一个雇 属于 男性 桶或者 女性 桶
- 奥 巴尼属于 桶
- 日期2014-10-28属于 十月 桶

当聚合 始被 行, 个文 里面的 通 算来决定符合 个桶的条件。如果匹配到, 文 将放入相 的桶并接着 行聚合操作。

桶也可以被嵌套在其他桶里面, 提供 次化的或者有条件的 分方案。例如, 辛辛那提会被放入俄亥俄州 个桶, 而 整个 俄亥俄州桶会被放入美国 个桶。

Elasticsearch 有很多 型的桶, 能 通 很多 方式来 分文 ( 、最受 迎的 、年 区 、地理位置等等)。其 根本上都是通 同 的原理 行操作: 基于条件来 分文 。

## 指

桶能 我 分文 到有意 的集合, 但是最 我 需要的是 些桶内的文 行一些指 的 算。分桶 是一 到目的的手段: 它提供了一 文 分 的方法来 我 可以 算感 趣的指 。

大多数 指 是 的数学 算（例如最小 、平均 、最大 ， 有 ）， 些是通 文 的 来 算。在 践中，指 能 算像平均薪 、最高出 格、95%的 延 的数据。

## 桶和指 的 合

聚合 是由桶和指 成的。 聚合可能只有一个桶，可能只有一个指 ，或者可能 个都有。也有可能有一些桶嵌套在其他桶里面。例如，我 可以通 所属国家来 分文 （桶），然后 算 个国家的平均薪酬（指 ）。

由于桶可以被嵌套，我 可以 非常多并且非常 的聚合：

- 1.通 国家 分文 （桶）
- 2.然后通 性 分 个国家（桶）
- 3.然后通 年 区 分 性 （桶）
- 4.最后， 个年 区 算平均薪酬（指 ）

最后将告 个 <国家， 性 ， 年 > 合的平均薪酬。所有的 些都在一个 求内完成并且只遍 一次数据！