



Elasticsearch 집계(aggregation) 조회

size=0

: 집계된 document들의 데이터는 불필요하므로 결과는 반환하지 않는다.

- 합산 집계(sum)를 통해 해당 서버로 유입된 데이터 집계

```
1 GET /apache-web-log/_search?size=0
2 {
3   "aggs": {
4     "total_bytes": {
5       "sum": {
6         "field": "bytes"
7       }
8     }
9   }
10 }
11 }
12
```

```
1 {
2   "took" : 13,
3   "timed_out" : false,
4   "_shards" : {
5     "total" : 5,
6     "successful" : 5,
7     "skipped" : 0,
8     "failed" : 0
9   },
10  "hits" : {
11    "total" : {
12      "value" : 10000,
13      "relation" : "gte"
14    },
15    "max_score" : null,
16    "hits" : [ ]
17  },
18  "aggregations" : {
19    "total_bytes" : {
20      "value" : 2.747282505E9
21    }
22  }
23 }
```

200 - OK 27 ms



Elasticsearch 집계(aggregation) 조회

- 특정 지역(Paris)에서 서버로 유입된 데이터 합산 집계

Console

Search Profiler

Grok Debugger

```
1 GET /apache-web-log/_search?size=0
2 {
3   "query" : {
4     "constant_score" : {
5       "filter" : {
6         "match" : { "geoip.city_name" : "Paris" }
7       }
8     }
9   },
10  "aggs" : {
11    "total_bytes" : {
12      "sum" : {
13        "field" : "bytes"
14      }
15    }
16  }
17 }
```

```
1 {
2   "took" : 3,
3   "timed_out" : false,
4   "_shards" : {
5     "total" : 5,
6     "successful" : 5,
7     "skipped" : 0,
8     "failed" : 0
9   },
10  "hits" : {
11    "total" : 21,
12    "max_score" : 0.0,
13    "hits" : [ ]
14  },
15  "aggregations" : {
16    "total_bytes" : {
17      "value" : 428964.0
18    }
19  }
20 }
```



Elasticsearch 집계(aggregation) 조회

- 평균 집계(avg)를 통해 해당 서버로 유입된 데이터의 평균 집계

Console

Search Profiler

Grok Debugger

```
1 GET /apache-web-log/_search?size=0
2 {
3   "aggs": {
4     "avg_bytes": {
5       "avg": {
6         "field": "bytes"
7       }
8     }
9   }
10 }
```



```
1 {
2   "took" : 17,
3   "timed_out" : false,
4   "_shards" : {
5     "total" : 5,
6     "successful" : 5,
7     "skipped" : 0,
8     "failed" : 0
9   },
10  "hits" : {
11    "total" : 10001,
12    "max_score" : 0.0,
13    "hits" : [ ]
14  },
15  "aggregations" : {
16    "avg_bytes" : {
17      "value" : 294456.8601286174
18    }
19  }
20 }
```



Elasticsearch 집계(aggregation) 조회

- 특정 지역(Paris)에서 서버로 유입된 데이터 평균 집계

Console

Search Profiler

Grok Debugger

```
1 GET _search?size=0
2 {
3   "query" : {
4     "constant_score" : {
5       "filter" : {
6         "match" : { "geoip.city_name" : "Paris" }
7       }
8     }
9   },
10  "aggs": {
11    "avg_bytes": {
12      "avg": {
13        "field": "bytes"
14      }
15    }
16  }
17 }
```

최대(max), 최소(min), 개수(value_count)
도 동일하게 처리.

```
1 {
2   "took" : 38,
3   "timed_out" : false,
4   "_shards" : {
5     "total" : 80,
6     "successful" : 80,
7     "skipped" : 0,
8     "failed" : 0
9   },
10  "hits" : {
11    "total" : 42,
12    "max_score" : 0.0,
13    "hits" : [ ]
14  },
15  "aggregations" : {
16    "avg_bytes" : {
17      "value" : 20426.85714285714
18    }
19  }
20 }
```



Elasticsearch 집계(aggregation) 조회

- 통계 집계(stats aggregation)는 모든 집계 결과값을 출력

Console Search Profiler Grok Debugger

```
1 GET /apache-web-log/_search?size=0
2 {
3   "aggs": {
4     "bytes_stats": {
5       "stats": {
6         "field": "bytes"
7       }
8     }
9   }
10 }
```

```
1 {
2   "took" : 27,
3   "timed_out" : false,
4   "_shards" : {
5     "total" : 5,
6     "successful" : 5,
7     "skipped" : 0,
8     "failed" : 0
9   },
10  "hits" : {
11    "total" : 10001,
12    "max_score" : 0.0,
13    "hits" : [ ]
14  },
15  "aggregations" : {
16    "bytes_stats" : {
17      "count" : 9330,
18      "min" : 35.0,
19      "max" : 6.9192717E7,
20      "avg" : 294456.8601286174,
21      "sum" : 2.747282505E9
22    }
23  }
24 }
```



Elasticsearch 집계(aggregation) 조회

- 특정 지역 통계 집계(stats aggregation)

Console

Search Profiler

Grok Debugger

```
1 GET /apache-web-log/_search?size=0
2 {
3   "query" : {
4     "constant_score" : {
5       "filter" : {
6         "match" : { "geoip.city_name" : "Paris" }
7       }
8     }
9   },
10  "aggs": {
11    "bytes_stats": {
12      "stats": {
13        "field": "bytes"
14      }
15    }
16  }
17 }
```

```
1 {
2   "took" : 11,
3   "timed_out" : false,
4   "_shards" : {
5     "total" : 5,
6     "successful" : 5,
7     "skipped" : 0,
8     "failed" : 0
9   },
10  "hits" : {
11    "total" : 21,
12    "max_score" : 0.0,
13    "hits" : [ ]
14  },
15  "aggregations" : {
16    "bytes_stats" : {
17      "count" : 21,
18      "min" : 1015.0,
19      "max" : 53270.0,
20      "avg" : 20426.85714285714,
21      "sum" : 428964.0
22    }
23  }
24 }
```



Elasticsearch 집계(aggregation) 조회

- 확장 통계 집계

```
1 GET /apache-web-log/_search?size=0
2 {
3   "aggs": {
4     "bytes_extended_stats": {
5       "extended_stats": {
6         "field": "bytes"
7       }
8     }
9   }
10 }
```

```
1 {
2   "took" : 28,
3   "timed_out" : false,
4   "_shards" : {
5     "total" : 5,
6     "successful" : 5,
7     "skipped" : 0,
8     "failed" : 0
9   },
10  "hits" : {
11    "total" : 10001,
12    "max_score" : 0.0,
13    "hits" : [ ]
14  },
15  "aggregations" : {
16    "bytes_extended_stats" : {
17      "count" : 9330,
18      "min" : 35.0,
19      "max" : 6.9192717E7,
20      "avg" : 294456.8601286174,
21      "sum" : 2.747282505E9,
22      "sum_of_squares" : 1.18280
23      "variance" : 1.25907136178
```



Elasticsearch 집계(aggregation) 조회

- 특정 지역 확장 통계 집계

```
1 GET /apache-web-log/_search?size=0
2 {
3   "query" : {
4     "constant_score" : {
5       "filter" : {
6         "match" : { "geoip.city_name" : "Paris"
7       }
8     }
9   },
10  "aggs": {
11    "bytes_extended_stats": {
12      "extended_stats": {
13        "field": "bytes"
14      }
15    }
16  }
17 }
```

```
1 {
2   "took" : 4,
3   "timed_out" : false,
4   "_shards" : {
5     "total" : 5,
6     "successful" : 5,
7     "skipped" : 0,
8     "failed" : 0
9   },
10  "hits" : {
11    "total" : 21,
12    "max_score" : 0.0,
13    "hits" : [ ]
14  },
15  "aggregations" : {
16    "bytes_extended_stats" : {
17      "count" : 21,
18      "min" : 1015.0,
19      "max" : 53270.0,
20      "avg" : 20426.85714285714,
21      "sum" : 428964.0,
22      "sum_of_squares" : 1
23        .8371748404E10,
24      "variance" : 4.575886693605442E8
25    },
26    "std_deviation" : 21391
27      .32229107271,
28    "std_deviation_bounds" : {
29      "upper" : 63209.501725002556,
30      "lower" : -22355.787439288277
31    }
32  }
33 }
```




Elasticsearch 집계(aggregation) 조회

- 카디널리티 집계(cardinality aggregation)는 중복 값을 제외한 고유 값 집계
 - 웹로그에서 미국의 몇 개 도시에서 데이터 유입이 있었는지 횟수 집계 (**terms** 이용)
 - 결과는 미국 내에서 요청 수가 가장 많은 도시 순으로 출력

```
1 GET /apache-web-log/_search?size=0
2 {
3   "query" : {
4     "constant_score" : {
5       "filter" : {
6         "match" : { "geoip.country_name" :
7           "United States" }
8       }
9     },
10    "aggs": {
11      "us_city_names": {
12        "terms": {
13          "field": "geoip.city_name.keyword",
14          "size": 20
15        }
16      }
17    }
18 }
```

```
15 "aggregations" : {
16   "us_city_names" : {
17     "doc_count_error_upper_bound" : 15
18     ,
19     "sum_other_doc_count" : 1295,
20     "buckets" : [
21       {
22         "key" : "Leander",
23         "doc_count" : 539
24       },
25       {
26         "key" : "Lititz",
27         "doc_count" : 273
28       },
29       {
30         "key" : "San Francisco",
31         "doc_count" : 230
32       }
33     ]
34   }
35 }
```



Elasticsearch 집계(aggregation) 조회

- 카디널리티 집계(stats aggregation)는 중복 값을 제외한 고유 값 집계
 - 미국 내 몇 개의 도시에서 유입되었는지 확인

```
1 GET /apache-web-log/_search?size=0
2 {
3   "query" : {
4     "constant_score" : {
5       "filter" : {
6         "match" : { "geoip.country_name" :
7                     "United States" }
8       }
9     },
10    "aggs": {
11      "us_cardinality": {
12        "cardinality": {
13          "field": "geoip.city_name.keyword"
14        }
15      }
16    }
17  }
18
19
```

```
1 {
2   "took" : 73,
3   "timed_out" : false,
4   "_shards" : {
5     "total" : 5,
6     "successful" : 5,
7     "skipped" : 0,
8     "failed" : 0
9   },
10  "hits" : {
11    "total" : 4255,
12    "max_score" : 0.0,
13    "hits" : [ ]
14  },
15  "aggregations" : {
16    "us_cardinality" : {
17      "value" : 249
18    }
19  }
20 }
```



Elasticsearch 집계(aggregation) 조회

- 백분위 수 집계를 통해 백분위에 대한 구간 분포 비율 확인

```
1 GET /apache-web-log/_search?size=0
2 {
3   "aggs": {
4     "bytes_percentiles": {
5       "percentiles": {
6         "field": "bytes"
7       }
8     }
9   }
10 }
11
```

```
1 {
2   "took" : 154,
3   "timed_out" : false,
4   "_shards" : {
5     "total" : 5,
6     "successful" : 5,
7     "skipped" : 0,
8     "failed" : 0
9   },
10  "hits" : {
11    "total" : 10001,
12    "max_score" : 0.0,
13    "hits" : [ ]
14  },
15  "aggregations" : {
16    "bytes_percentiles" : {
17      "values" : {
18        "1.0" : 229.0,
19        "5.0" : 358.0,
20        "25.0" : 3644.5,
21        "50.0" : 12238.009098899343,
22        "75.0" : 37348.580528846156,
23        "95.0" : 171303.6,
24        "99.0" : 1204031.8000000163

```



Elasticsearch 집계(aggregation) 조회

- 백분위에 대한 구간 분포 비율 지정을 통해 확인
 - [실무 팁] 이것을 활용해 서버 사양보다 너무 큰 데이터가 유입되는 경우 데이터 크기를 조절해 서비스 품질을 개선할 수 있다.

```
1 GET /apache-web-log/_search?size=0
2 {
3   "aggs": {
4     "bytes_percentiles": {
5       "percentiles": {
6         "field": "bytes",
7         "percents": [10, 20, 30, 40, 50, 60, 70, 80, 90]
8       }
9     }
10  }
11 }
```

```
1 {
2   "took" : 27,
3   "timed_out" : false,
4   "_shards" : {
5     "total" : 5,
6     "successful" : 5,
7     "skipped" : 0,
8     "failed" : 0
9   },
10  "hits" : {
11    "total" : 10001,
12    "max_score" : 0.0,
13    "hits" : [ ]
14  },
15  "aggregations" : {
16    "bytes_percentiles" : {
17      "values" : {
18        "10.0" : 1015.0,
19        "20.0" : 3638.0,
20        "30.0" : 4877.0,
21        "40.0" : 8725.630150831908,
22        "50.0" : 12266.037777777778,
23        "60.0" : 15897.161435791895,
24        "70.0" : 29941.0,
25        "80.0" : 50112.0,
26        "90.0" : 65748.0
```



Elasticsearch 집계(aggregation) 조회

- 백분위 수 Rank 집계
 - 백분위 수 집계와 반대로 백분위 수 집계는 백분위를 지정해서 백분위 수 확인하고, 이것은 특정 필드 수치를 통해 백분위 수 구간을 확인

```
1 GET /apache-web-log/_search?size=0
2 {
3   "aggs": {
4     "bytes_percentile_ranks": {
5       "percentile_ranks": {
6         "field": "bytes",
7         "values": [5000, 10000]
8       }
9     }
10  }
11 }
```

```
1 {
2   "took" : 33,
3   "timed_out" : false,
4   "_shards" : {
5     "total" : 5,
6     "successful" : 5,
7     "skipped" : 0,
8     "failed" : 0
9   },
10  "hits" : {
11    "total" : 10001,
12    "max_score" : 0.0,
13    "hits" : [ ]
14  },
15  "aggregations" : {
16    "bytes_percentile_ranks" : {
17      "values" : {
18        "5000.0" : 32.362841215178584,
19        "10000.0" : 45.02619946859613
20      }
21    }
22  }
```



Elasticsearch 집계(aggregation) 조회

- 지형 경계 집계는 지형 좌표 필드를 통해 해당 지역 경계를 계산 (geo_point)

```
1 GET /apache-web-log-applied-mapping/_search?size=0
2 {
3   "aggs" : {
4     "viewport" : {
5       "geo_bounds" : {
6         "field" : "geoip.location",
7         "wrap_longitude" : true
8       }
9     }
10  }
11 }
12
```

visualize > search 옆에 [+] Create Visualization > Region map > Name: apache-web-log 선택

```
1 {
2   "took" : 1,
3   "timed_out" : false,
4   "_shards" : {
5     "total" : 5,
6     "successful" : 5,
7     "skipped" : 0,
8     "failed" : 0
9   },
10  "hits" : {
11    "total" : 10001,
12    "max_score" : 0.0,
13    "hits" : [ ]
14  },
15  "aggregations" : {
16    "viewport" : {
17      "bounds" : {
18        "top_left" : {
19          "lat" : 69.34059997089207,
20          "lon" : -159.76670005358756
21        },
22        "bottom_right" : {
23          "lat" : -45.88390002027154,
24          "lon" : 176.91669998690486
25        }
26      }
27    }
28  }
29 }
```



Elasticsearch 집계(aggregation) 조회

- 유럽 지역 지형 경계 집계 (geo_point)

```
1 GET /apache-web-log-applied-mapping/_search?size=0
2 {
3   "query" : {
4     "constant_score" : {
5       "filter" : {
6         "match" : { "geoip.continent_code" : "EU"
7       }
8     }
9   },
10  "aggs" : {
11    "viewport" : {
12      "geo_bounds" : {
13        "field" : "geoip.location",
14        "wrap_longitude" : true
15      }
16    }
17  }
18 }
19
20 :
```

```
1 {
2   "took" : 32,
3   "timed_out" : false,
4   "_shards" : {
5     "total" : 5,
6     "successful" : 5,
7     "skipped" : 0,
8     "failed" : 0
9   },
10  "hits" : {
11    "total" : 3939,
12    "max_score" : 0.0,
13    "hits" : [ ]
14  },
15  "aggregations" : {
16    "viewport" : {
17      "bounds" : {
18        "top_left" : {
19          "lat" : 69.34059997089207,
20          "lon" : -16.358700077980757
21        },
22        "bottom_right" : {
23          "lat" : 28.534799963235855,
24          "lon" : 88.22059999220073
25        }
26      }
27    }
28  }
29 }
```