

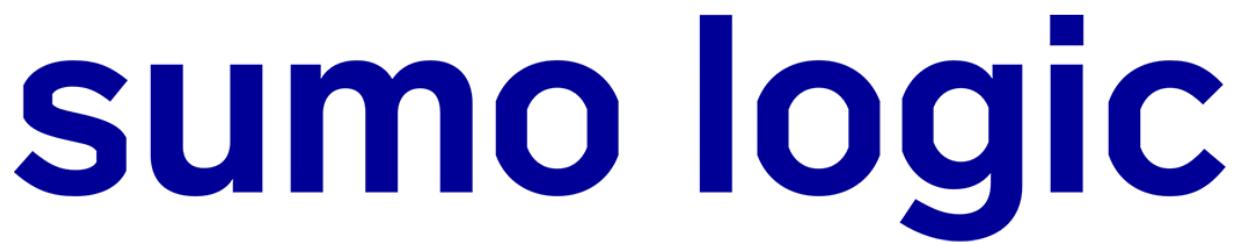
# Parse Operators

---

 [dev.classmethod.jp/articles/i-tried-hands-on-some-of-the-parse-operators-of-sumo-logic](https://dev.classmethod.jp/articles/i-tried-hands-on-some-of-the-parse-operators-of-sumo-logic)

HemanthKumar R

August 15, 2023



## 目次

---

## Introduction

---

Hemanth of Alliance Department here. The blog focuses on sumo logic parse operators. The different way to analyse logs in the sumo logic by the parse operators.

---

## Sumo Logic

---

Before going further let's understand what sumo logic is. A cloud-based log management and analytics software called Sumo Logic enables businesses to exploit their machine data for useful insights. Sumo Logic's flexible capabilities make log data analysis simple and offer real-time visibility into operational and security insights.

The ability to manually and ad-hoc extract fields from log messages within a query is provided via parse operators. One of technique to make the most of data is using this method. Let's explore some of the parse operators that Sumo Logic offers.

# Parse Variable pattern using Regix

The Parse Regex operator, also known as the extract operator, is designed for users who are familiar with the syntax of regular expressions. You may easily extract complex data from log lines with this operator.

# Parsing a simple IP address

## Parse multi

The screenshot shows the Sumo Logic interface with a search query in the top bar: `sourceCategory=Logs/Apache/Access | parse regex \"^<ip_address><id1><id2><id3>|<id1><id2><id3>\" multi`. The search results table displays 4399 results over a time range from 08/15/2023 12:50:00 PM +0000 to 08/15/2023 1:05:30 PM +0000. The results show various Apache access logs with multiple IDs extracted using the multi-line parser.

#	Time	ip_address	Message
2	08/15/2023 1:04:19.355 PM +0000	19.174.45.8	Host apache-prod = Name: http: Input = Category:Logs/Apache/Access 19.174.45.8 - - [2023-08-15 04:04:19.355 +0000] "GET /downloads/Datasheet.pdf HTTP/1.1" 200 1563 "http://www.mozilla.org/Firefox/33.6" Host apache-prod = Name: http: Input = Category:Logs/Apache/Access
3	08/15/2023 1:04:19.355 PM +0000	128.198.198.281	128.198.198.281 - - [2023-08-15 04:04:19.355 +0000] "GET /includes/follow/follow_us.php HTTP/1.1" 200 2584 "http://www.google.com/1share/drive/plog?#management&source=web&id=4" "Mozilla/5.0 (Macintosh; U; Intel Mac en-us) AppleWebKit/533.25.1 (KHTML, like Gecko) Chrome/19.0.1084.52 Safari/533.5" Host apache-prod = Name: http: Input = Category:Logs/Apache/Access
4	08/15/2023 1:04:19.250 PM +0000	17.233.159.68	17.233.159.68 - - [2023-08-15 04:04:19.255 +0000] "GET /_sns/index/follow/follow_us.php HTTP/1.1" 302 5596 "http://www.google.com/" "FacebookExternalIframe/9.1 (http://www.facebook.com/external_iframe.php)" Host apache-prod = Name: http: Input = Category:Logs/Apache/Access
5	08/15/2023 1:04:19.250 PM +0000	78.59.152.165	78.59.152.165 - - [2023-08-15 04:04:19.255 +0000] "GET /media/play_button_gray.png HTTP/1.1" 200 5121 "http://www.linkedin.com/" "GalaxyY7.0 [en] (Huawei Y7 (Huawei Y7)) AppleWebKit/536.28 (KHTML, like Gecko) Version/6.0 Mobile/16405560 Safari/536.2" Host apache-prod = Name: http: Input = Category:Logs/Apache/Access
6	08/15/2023 1:04:19.250 PM +0000	78.238.33.64	78.238.33.64 - - [2023-08-15 04:04:19.255 +0000] "GET /_sns/index/follow/follow_us.php HTTP/1.1" 302 7217 "http://www.greylock.com" (iPad; CPU OS 6.0 like Mac OS X) AppleWebKit/536.28 (KHTML, like Gecko) Version/6.0 Mobile/16405560 Safari/536.2 Host apache-prod = Name: http: Input = Category:Logs/Apache/Access
7	08/15/2023 1:04:19.250 PM +0000	169.187.142.237	169.187.142.237 - - [2023-08-15 04:04:19.355 +0000] "GET /media/bio_kid.jpg HTTP/1.1" 200 2474 "http://www.singaporesocialmediareview.com/SearchIndex/10100" "Mozilla/5.0 (iPhone; U; CPU iPhone OS 4.1 like Mac OS X; en-us) AppleWebKit/534.32 (KHTML, like Gecko) Version/4.0.5 Mobile/8039777 Safari/534.32" Host apache-prod = Name: http: Input = Category:Logs/Apache/Access
8	08/15/2023 1:04:19.250 PM +0000	78.238.33.64	78.238.33.64 - - [2023-08-15 04:04:19.255 +0000] "GET /_sns/index/follow/follow_us.php HTTP/1.1" 302 5581 "http://www.Mozilla/5.0 (Macintosh; Intel Mac OS X 10.5.8) AppleWebKit/537.12 (KHTML, like Gecko) Version/5.1.7 Safari/534.59" Host apache-prod = Name: http: Input = Category:Logs/Apache/Access

## Parse JSON Formatted Logs

JSON logs are full of structured data. The `JSON` operator, combined with strong JSONPath expressions, allows you to precisely extract certain values from these logs.

## Extracting multiple fields

```
1 _sourceCategory=Labs/AWS/CloudTrail
2 | json "eventTime", "awsRegion"
3 | fields eventTime, awsRegion
```

500 0:45m

1:01 PM 1:03 PM 1:05 PM 1:07 PM 1:09 PM 1:11 PM 1:13 PM 1:15 PM 06/15/2023 1:01:42 PM +0900 STATUS Done ELAPSED TIME 00:00:00 RESULTS 584 SESSION 33E0CC10HEC190C332 LOAD 0/0 06/15/2023 1:14:42 PM +0900 Hide Histogram Hide Log Levels Maximize View

Messages Page 1 of 24 LogReduce LogCompare

#	Time	eventTime	awsRegion	Message
1	06/15/2023 1:17:37.000 PM +0900	2023-06-15T04:12:07Z	us-east-1	<pre>eventVersion: "1.08", userIdentity: "...", eventTime: "2023-06-15T04:12:07Z", eventSource: "kafka.amazonaws.com", eventName: "DescribeClusterV2", awsRegion: "us-east-1", sourceIPAddress: "144.219.219.172", userAgent: "Mozilla/5.0 (Macintosh; Intel Mac OS X 10_15_7) AppleWebKit/537.36 (KHTML, like Gecko) Chrome/111.0.0.0 Safari/537.36", errorCode: "NotFoundException", requestParameters: {...},</pre> <p>Host: 54.219.55.58 ~ None:Http Input ~ Category:Labs/AWS/CloudTrail ~</p>
2	06/15/2023 1:17:34.000 PM +0900	2023-06-15T04:12:04Z	us-east-1	<pre>eventVersion: "1.08", userIdentity: "...", eventTime: "2023-06-15T04:12:04Z", eventSource: "kafka.amazonaws.com", eventName: "DescribeClusterV2", awsRegion: "us-east-1", sourceIPAddress: "144.219.219.172", userAgent: "Mozilla/5.0 (Macintosh; Intel Mac OS X 10_15_7) AppleWebKit/537.36 (KHTML, like Gecko) Chrome/111.0.0.0 Safari/537.36", errorCode: "NotFoundException", requestParameters: {...},</pre> <p>Host: 54.219.55.58 ~ None:Http Input ~ Category:Labs/AWS/CloudTrail ~</p>
3	06/15/2023 1:17:34.000 PM +0900	2023-06-15T04:12:04Z	us-east-1	<pre>eventVersion: "1.08", userIdentity: "...", eventTime: "2023-06-15T04:12:04Z", eventSource: "kafka.amazonaws.com", eventName: "DescribeClusterV2", awsRegion: "us-east-1", sourceIPAddress: "144.219.219.172", userAgent: "Mozilla/5.0 (Macintosh; Intel Mac OS X 10_15_7) AppleWebKit/537.36 (KHTML, like Gecko) Chrome/111.0.0.0 Safari/537.36", errorCode: "NotFoundException", requestParameters: {...},</pre> <p>Host: 54.219.55.58 ~ None:Http Input ~ Category:Labs/AWS/CloudTrail ~</p>

Displayed Fields

Time  awsRegion  eventTime  Message

Hidden Fields

Collector  Size  Source  Source Category  Source Host  Source Name  Aid  \_loglevel  \_siemforward  \_size  \_source  \_sourceCategory  \_sourceHost  \_sourceName  \_index  account  additionalEventData  awRegion  eventId  eventsName  eventsource  message

## Using Nested Array with wildcard

```
1 _sourceCategory=Labs/AWS/CloudTrail
2 | json "resources[*].accountId" as Aid
```

100 0:45m

1:08 PM 1:13 PM 1:12 PM 1:14 PM 1:15 PM 1:16 PM 1:17 PM 1:20 PM 1:22 PM 06/15/2023 1:08:04 PM +0900 STATUS Done ELAPSED TIME 00:00:00 RESULTS 67 SESSION 18F03A0D1300CD9 LOAD 0/0 06/15/2023 1:23:04 PM +0900 Hide Histogram Hide Log Levels Maximize View

Messages Page 1 of 3 LogReduce LogCompare

#	Time	Aid	Message
1	06/15/2023 1:17:46.000 PM +0900	<pre>[{"accountID": "123456789012", "resourceType": "Lambda"}, {"accountID": "123456789012", "resourceType": "Kinesis"}]</pre>	<pre>eventVersion: "1.06", userIdentity: "...", eventTime: "2023-06-15T04:17:46Z", eventSource: "sts.amazonaws.com", eventName: "Invoke", awsRegion: "us-east-2", sourceIPAddress: "20.168.149.149", userAgent: "apigateway.amazonaws.com", requestParameters: {...}, responseParameters: null, additionalEventDetails: [...], requestID: "0a1d9aa8-81a6-4f96-92dc-23c409a10881", eventID: "9f2371e3-04a5-4caa-adc9-0052ab4a651a", readOnly: false, resources: [...]   + {     accountId: "123456789012",     type: "AWS::Lambda::Function",     ARN: "arn:aws:lambda:us-east-1:123456789012:function:dump"   },   + {     accountId: "123456789012",     type: "AWS::Kinesis::Stream",     ARN: "arn:aws:kinesis:us-east-2:123456789012:stream/KinesisTest"   } ], eventType: "ApiApicall", managementEvent: false, recipientAccountId: "123456789012", sharedEventID: "41499ea-a022-4fa7-9cb4-00944a711190"</pre> <p>Host: 54.212.214.211 ~ None:Http Input ~ Category:Labs/AWS/CloudTrail ~</p>
2	06/15/2023 1:17:46.000 PM +0900	<pre>[{"accountID": "123456789012", "resourceType": "Lambda"}]</pre>	

Displayed Fields

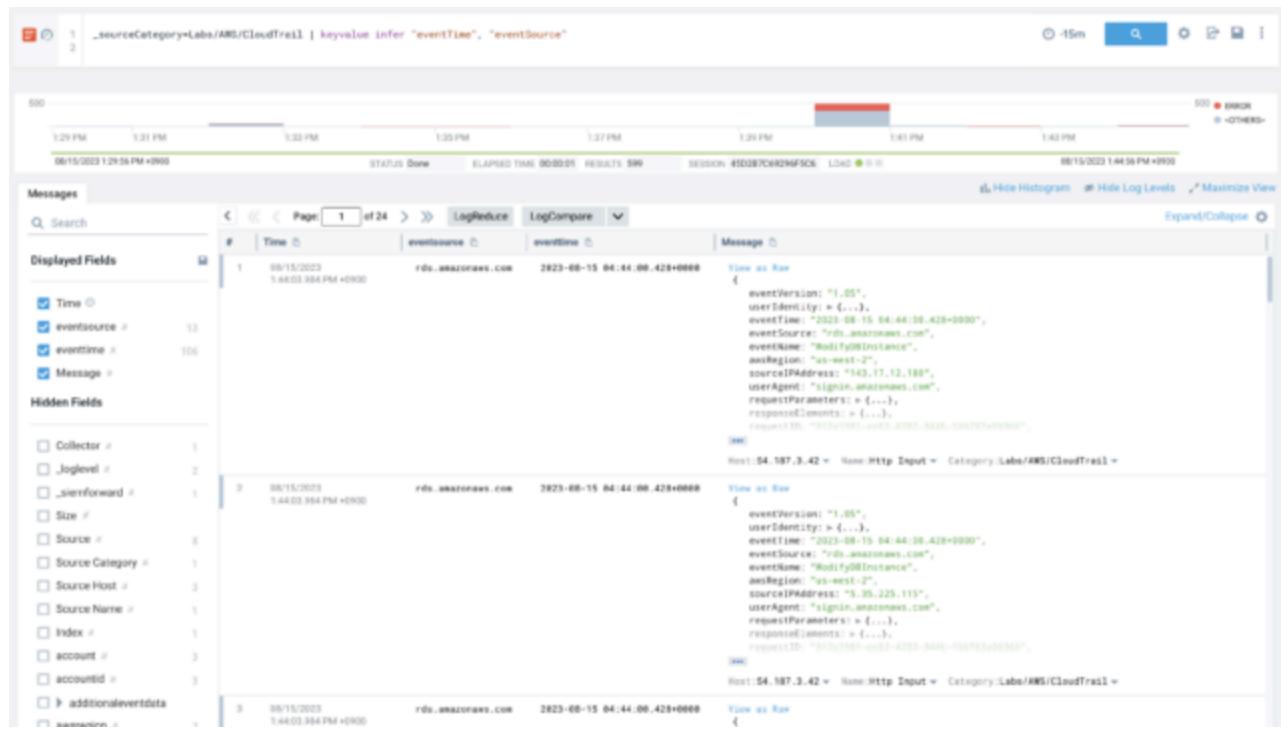
Time  Aid  Message

Hidden Fields

Collector  Size  Source  Source Category  Source Host  Source Name  Index  account  additionalEventData  awRegion  eventId  eventsName  eventsource  message

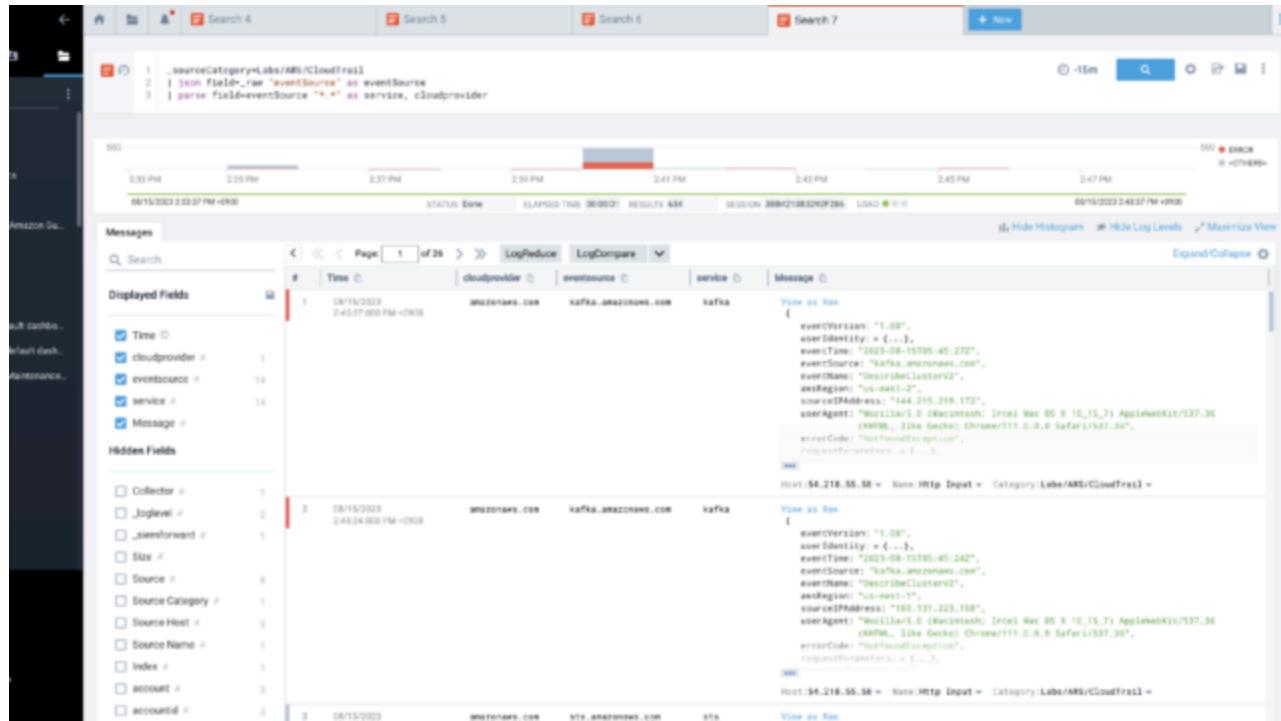
# Parse Keyvalue Formatted Logs

Key-value pairs are a common structure for log files. By defining the key associated with each value, the key-value operator enables you to extract values from a log message.



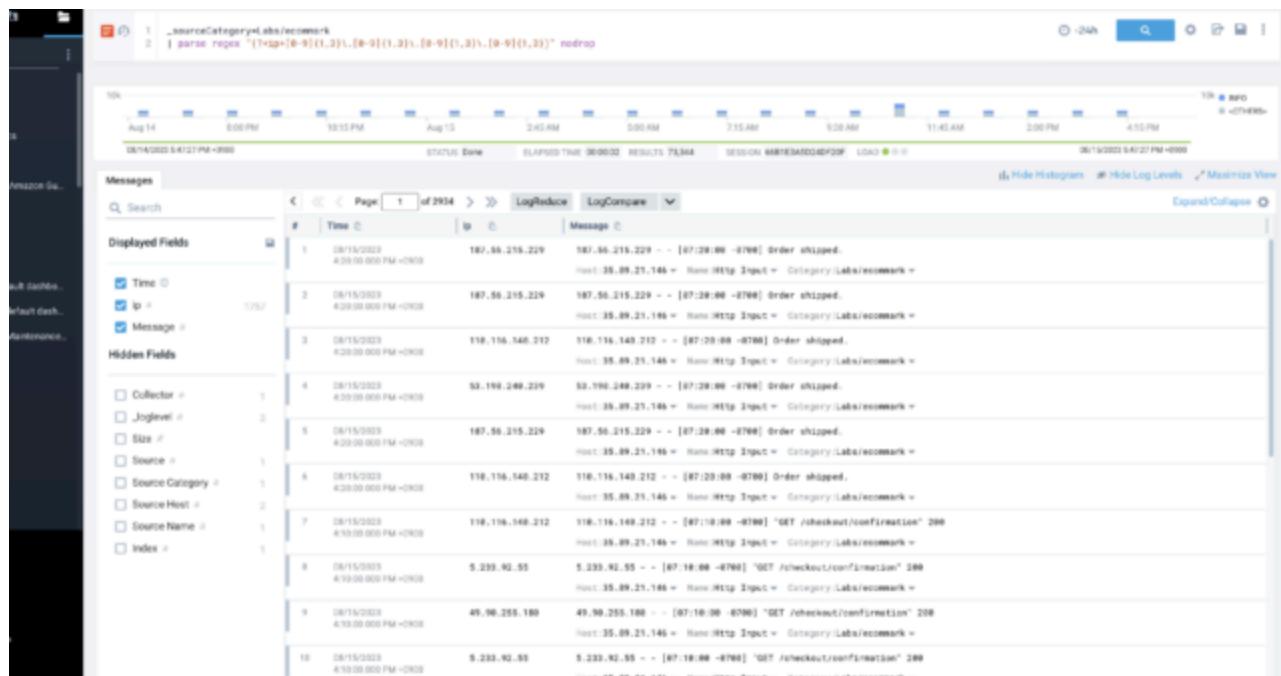
## Parse field option

This is the idea of using the "field" syntax to parse previously extracted fields or metadata fields..



## Parse Nodrop option

Even unmatched segments of parse expressions are included in the results with the Nodrop option, ensuring that no valuable data is lost.



## parseDate

The `parseDate` operator extracts date and time information from strings, delivering millisecond-accurate timestamps.

```

1 _sourceCategory=Labs/AWS/CloudTrail
2 | json Fields->@eventTime as eventTime
3 | parseDate(eventTime, 'yyyy-MM-dd'T'HH:mm:ss.SSSX') as milliseconds

```

Messages

#	Time	eventTime	milliseconds	Message
1	08/15/2023 5:49:49 PM +0000	2023-08-15T08:59:29.829Z	1,492,889,948,829	View as Raw
2	08/15/2023 5:50:23.862 PM +0000	2023-08-15T08:59:49.826Z	1,492,889,989,826	View as Raw
3	08/15/2023 5:50:49.825Z	2023-08-15T08:59:49.825Z	1,492,889,989,825	View as Raw

Displayed Fields

- Time
- eventTime
- milliseconds
- Message

Hidden Fields

- Collector
- LogLevel
- \_idforward
- Size
- Source
- Source Category
- Source Host
- Source Name
- Index
- account
- \_index
- \_id
- \_score
- \_version
- awsRegion

## parseHex

The `parseHex` operator makes it simple to convert hexadecimal strings into numerical values.

```

1 _sourceCategory=Labs/eecomark
2 | parseHex("12003F") as decimalValue

```

Messages

#	Time	decimalValue	Message
1	08/15/2023 5:29:00 PM +0000	1,233,456	118.116.148.212 - - [08:29:00 -0700] Order shipped. Host: 35.89.21.146 v Name: http: Input: Category:Labs/eecomark v
2	08/15/2023 5:29:00 PM +0000	1,233,456	32.198.248.239 - - [08:29:00 -0700] Order shipped. Host: 35.89.21.146 v Name: http: Input: Category:Labs/eecomark v
3	08/15/2023 5:29:00 PM +0000	1,233,456	86.95.185.127 - - [08:29:00 -0700] Order shipped. Host: 35.89.21.146 v Name: http: Input: Category:Labs/eecomark v
4	08/15/2023 5:29:00 PM +0000	1,233,456	148.53.33.29 - - [08:29:00 -0700] Order shipped. Host: 35.89.21.146 v Name: http: Input: Category:Labs/eecomark v
5	08/15/2023 5:29:00 PM +0000	1,233,456	149.13.33.29 - - [08:29:00 -0700] Order shipped. Host: 35.89.21.146 v Name: http: Input: Category:Labs/eecomark v
6	08/15/2023 5:29:00 PM +0000	1,233,456	118.116.148.212 - - [08:29:00 -0700] Order shipped. Host: 35.89.21.146 v Name: http: Input: Category:Labs/eecomark v
7	08/15/2023 5:29:00 PM +0000	1,233,456	137.249.174.157 - - [08:19:00 -0700] "GET /checkout/confirmation" 200 Host: 35.89.21.146 v Name: http: Input: Category:Labs/eecomark v
8	08/15/2023 5:29:00 PM +0000	1,233,456	118.116.148.212 - - [08:19:00 -0700] "GET /checkout/confirmation" 200 Host: 35.89.21.146 v Name: http: Input: Category:Labs/eecomark v
9	08/15/2023 5:29:00 PM +0000	1,233,456	5.233.92.55 - - [08:19:00 -0700] "GET /checkout/confirmation" 200 Host: 35.89.21.146 v Name: http: Input: Category:Labs/eecomark v
10	08/15/2023 5:29:00 PM +0000	1,233,456	86.95.183.127 - - [08:19:00 -0700] "GET /checkout/confirmation" 200 Host: 35.89.21.146 v Name: http: Input: Category:Labs/eecomark v

Displayed Fields

- Time
- decimalValue
- Message

Hidden Fields

- Collector
- LogLevel
- \_idforward
- Size
- Source
- Source Category
- Source Host
- Source Name
- Index
- account
- \_index
- \_score
- \_version
- awsRegion

## Conclusion

I hope you now have a better knowledge of parse operators. These are some of the operators available that can be used to transform raw data into meaningful insights.



## EVENTS

---



[【1/29（木）】クラスメソッドの会社説明会を開催します](#)

開催前



[【1/28 \(水\)】クラスメソッドの新卒向け会社説明会を開催します](#)

開催前



[【CMグループ/エンド直案件特集】ITフリーランス向け「CMパートナーズ」説明会 by クラスメソッド](#)

開催前



[【2/5 \(木\) 東京】オペレーターの生産性を50%アップ! 見て、聞いて、納得できるAIコールセンター実演セミナー](#)

開催前



[【2/25 \(水\)】AI駆動開発、実際どうなの?【実践編】～現場でぶつかる課題と乗り越え方～](#)

[開催前](#)



[【1/29 \(木\)】今日から始めるAWSセキュリティ対策 3ステップでわかる実践ガイド](#)

[開催前](#)

[セミナー一覧](#) [会社説明会一覧](#) [勉強会一覧](#)