LogExplain Operator

The **LogExplain** operator allows you to compare sets of structured logs based on events you're interested in. Structured logs can be in JSON, CSV, key-value, or any structured format.

You'll need to specify an event of interest as a conditional statement – this is called the **Event Condition**. You can specify a condition to compare against the event-of-interest condition, this is called the **Against Condition**. If no Against Condition is provided, LogExplain will generate the comparison data set based on the fields in your Event Condition.

The syntax is:

```
| logexplain <event_condition> [against <against_condition>] on
<fieldname>
```

LogExplain will process your data against the specified conditions and create separate data sets to compare:

- A control data set from normal operations data.
- An event-of-interest data set.

LogExplain gathers frequent (at least 5% higher) joint-column entries, such as key-value pairs that occur more frequently compared to the control set. The results indicate what entities correlate with the event you're interested in.

To explore the LogExplain operator, copy and run the following query on CloudTrail:

```
_sourceCategory=*cloudtrail*
| json field=_raw "userIdentity.userName" as userName nodrop
| json field=_raw
"userIdentity.sessionContext.sessionIssuer.userName" as
userName_role nodrop
| if (isNull(userName), if(!isNull(userName_role),userName_role,
"Null_UserName"), userName) as userName
| json field=_raw "eventSource" as eventSource
| json field=_raw "eventName" as eventName
| json field=_raw "awsRegion" as awsRegion
| json field=_raw "errorCode" as errorCode nodrop
| json field=_raw "errorMessage" as errorMessage nodrop
| json field=_raw "sourceIPAddress" as sourceIp nodrop
| json field=_raw "requestParameters.bucketName" as bucketName
nodrop
| json field=_raw "recipientAccountId" as accountId
| where eventSource matches "s3.amazonaws.com"  and accountId
matches "*"
```

```
| logexplain (errorCode matches "*AccessDenied*") on sourceIp,
userName, awsRegion, eventName, bucketName
```

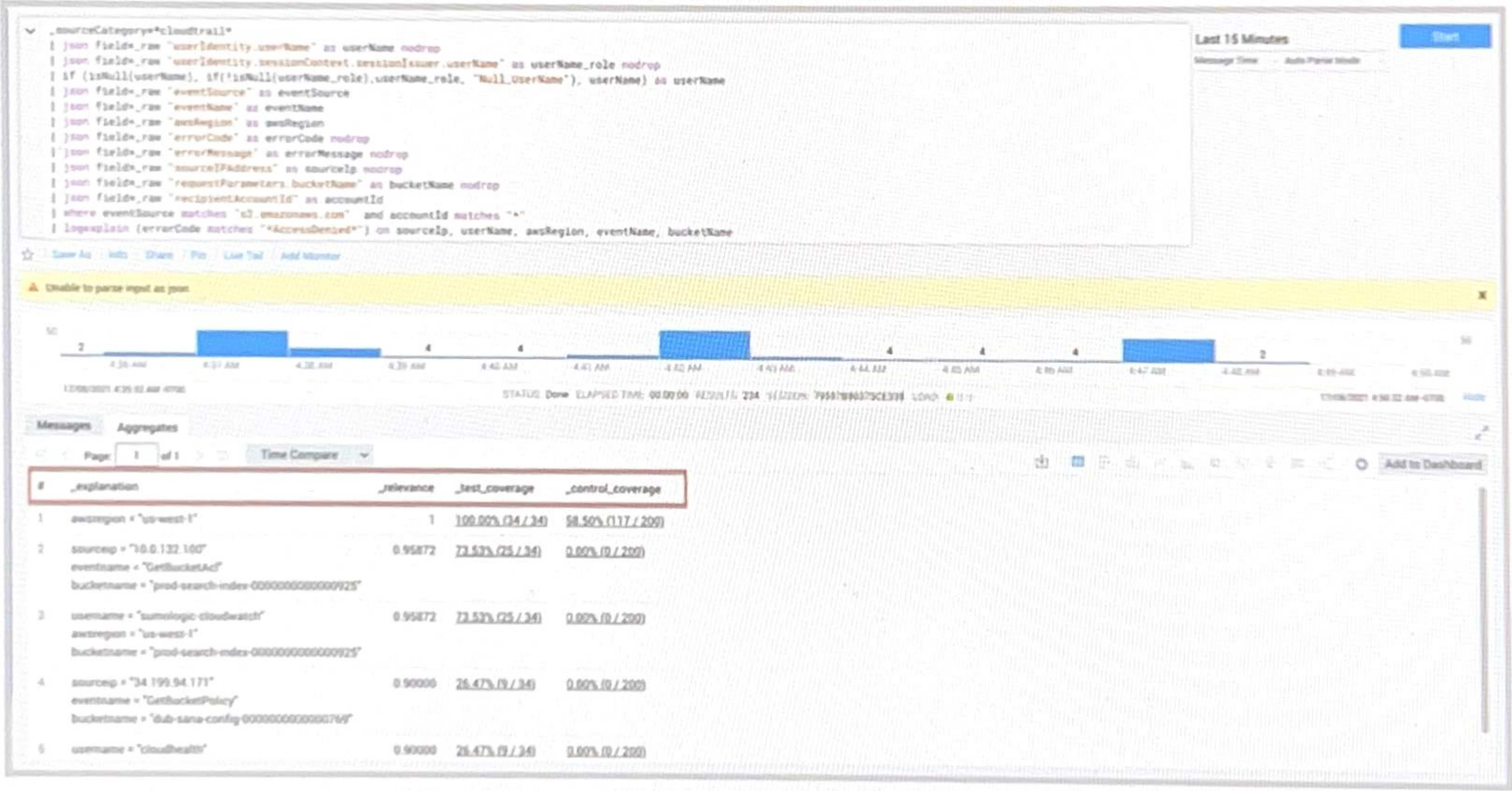
LogExplain returns following fields in results:

- _explanation
  The fields and respective values from the comparison.

- _relevance
  The probability that the explanation occurs in the event-of-interest data set. Values are 0 to 1.

  _test_coverage
  The percentage of data in the event-of-interest set that has the explanation.The link opens a new search that drills down to these logs based on the related explanation.

- _control_coverage
  The percentage of control data in the event-of-interest set that has the explanation.The link opens a new search that drills down to these logs based on the related explanation.



```
_sourceCategory=*cloudtrail*
| json field=_raw "userIdentity.userName" as userName nodrop
| json field=_raw "userIdentity.sessionContext.sessionIssuer.userName" as userName_role nodrop
| if (isNull(userName), if(!isNull(userName_role),userName_role, "Null_UserName"), userName) as userName
| json field=_raw "eventSource" as eventSource
| json field=_raw "eventName" as eventName
| json field=_raw "awsRegion" as awsRegion
| json field=_raw "errorCode" as errorCode nodrop
| json field=_raw "errorMessage" as errorMessage nodrop
| json field=_raw "sourceIPAddress" as sourceIp nodrop
| json field=_raw "requestParameters.bucketName" as bucketName nodrop
| json field=_raw "recipientAccountId" as accountId
| where eventSource matches "s3.amazonaws.com" and accountId matches "*"
| logexplain (errorCode matches "*AccessDenied*") on sourceIp, userName, awsRegion, eventName, bucketName
```

⚠ Unable to parse input as json

#	_explanation	_relevance	_test_coverage	_control_coverage
1	awsregion = "us-west-1"	1	100.00% (34 / 34)	58.50% (117 / 200)

2	sourceip = "10.0.132.100"
	eventname = "GetBucketAcl"
	bucketname = "prod-search-index-000000000000925"
		0.95872	73.53% (25 / 34)	0.00% (0 / 200)

3	username = "sumologic-cloudwatch"
	awsregion = "us-west-1"
	bucketname = "prod-search-index-000000000000925"
		0.95872	73.53% (25 / 34)	0.00% (0 / 200)

4	sourceip = "34.199.94.171"
	eventname = "GetBucketPolicy"
	bucketname = "dub-sana-config-000000000000769"
		0.90000	26.47% (9 / 34)	0.00% (0 / 200)

5	username = "cloudhealth"	0.90000	26.47% (9 / 34)	0.00% (0 / 200)

With the provided results you can:

- Click the provided links to drill down and further explore logs from each explanation.
- Run subsequent searches.

For example, if an IP address is an outlier you might search for logs referencing that IP address for further investigation.