

# Splunk: Exploring SPL

## INTRODUCTION

Splunk is a powerful SIEM solution that provides the ability to search and explore machine data. **Search Processing Language (SPL)** is used to make the search more effective. It comprises various functions and commands used together to form complex yet effective search queries to get optimized results.

This room will dives into some key fundamentals of searching capability, like chaining SPL queries to construct simple to complex queries.

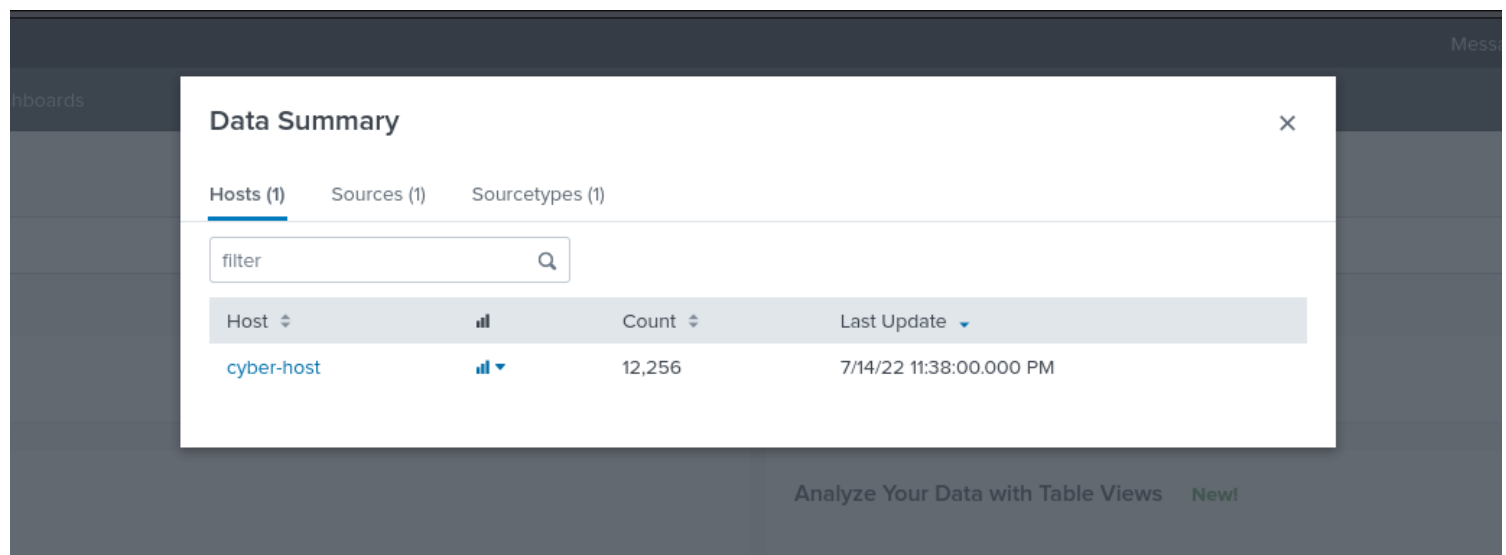
What is the name of the host in the Data Summary tab?

cyber-host

✓ Correct

The filter "index=windowslogs\*" is used to filter data from the "windowslogs" index, showing all events represented by the asterisk ().

Time filter set to all, to get all the existing logs. By this I was able to retrieve the host name from the data summary tab.



In the search History, what is the 7th search query in the list? (excluding your searches from today)

index=windowslogs | chart count(EventCode) by Image

✓ Correct

Setting no time filter to get all possible logs with the time duration set to week to date.

I also ignored all the logs for today and obtained the answer at the 7th place from the first search that did not occur today just as seen below.

**splunk>enterprise** Apps ▾ Messages ▾

Search Analytics Datasets Reports Alerts Dashboards

### Search

1 enter search here...

No Event Sampling ▾

▼ Search History ⓘ

filter 🔍 No Time Filter ▾ 20 Per Page ▾

Search ▾

- > index=windowslogs\*
- > Windowslogs
- > source="Event\_Logs.json" host="cyber-host" index="windowslogs" sourcetype="\*\_json"
- > index=\*
- > \*
- > index=\* | delete
- > index=\* | stats count by index
- > index=windowslogs | chart count by Image
- > **index=windowslogs | chart count(EventCode) by Image**
- > index=windowslogs | chart count(EventCode) by User
- > index=windowslogs | rare User

In the left field panel, which Source IP has recorded max events?

172.90.12.11

✓ Correct

I used the new search query to search for source ip. host="cyber-host" sourceip as shown in the image below.

```
a EventReceivedTime 14
a EventType 14
# ExecutionProcessID 3
a extracted_EventType 1
a extracted_host 1
a Hostname 3
a Image 10
a Index 1
a Initiated 2
# Keywords 1
# linecount 1
a Message 86
a Opcode 1
# OpcodeValue 1
# port 1
a ProcessGuid 20
# Processid 19
a Protocol 2
a ProviderGuid 1
a punct 1
# RecordNumber 86
a RuleName 1
a Severity 1
# SeverityValue 1
a SourceHostname 1
a Sourceip 7
a SourceipV6 2
a SourceModuleName 1
a SourceModuleType 1
a SourceName 1
# SourcePort 46
a SourcePortName 1
a splunk_server 1
a tags[] 1
# Task 1
# ThreadID 3
a timestamp 61
a UserID 1
a UtcTime 69
# Version 1
```

+ Extract New Fields

```
Protocol: udp
Initiated: false
SourceIsIpv6: true
SourceIp: 0:0:0:0:0:0:1
SourceHostname: -
SourcePort: 53
SourcePortName: -
DestinationIsIpv6: true
DestinationIp: 0:0:0:0:0:0:1
DestinationHostname: -
DestinationPort: 62132
DestinationPortName: -
Opcode: Info
OpcodeValue: 0
ProcessGuid: {ef3dcf22-42bc-5f5f-4c00-000000000300}
```

**Sourceip** X

7 Values, 100% of events Selected

**Reports**

Top values Top values by time Rare values

Events with this field

Values	Count	%
172.90.12.11	53	61.628%
172.18.38.5	15	17.442%
fe80:0:0:c86e:cb04:bc03:d64f	10	11.628%
0:0:0:0:0:0:1	4	4.651%
fe80:0:0:7976:d2f2:1752:21b5	2	2.326%
224.0.0.251	1	1.163%
ff02:0:0:0:0:0:fb	1	1.163%

```
version: 5
host: cybertees.net
port: 60427
tags: [ [+ ] ]
timestamp: 2022-04-15T12:06:43.672Z
}
```

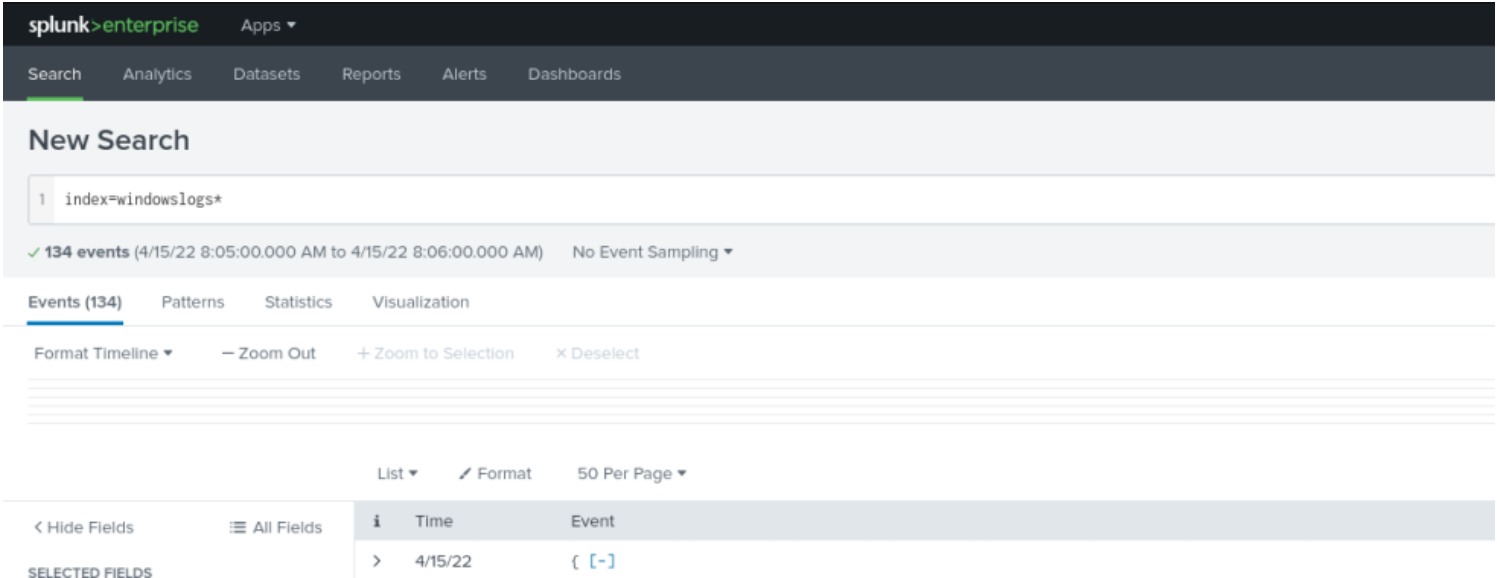
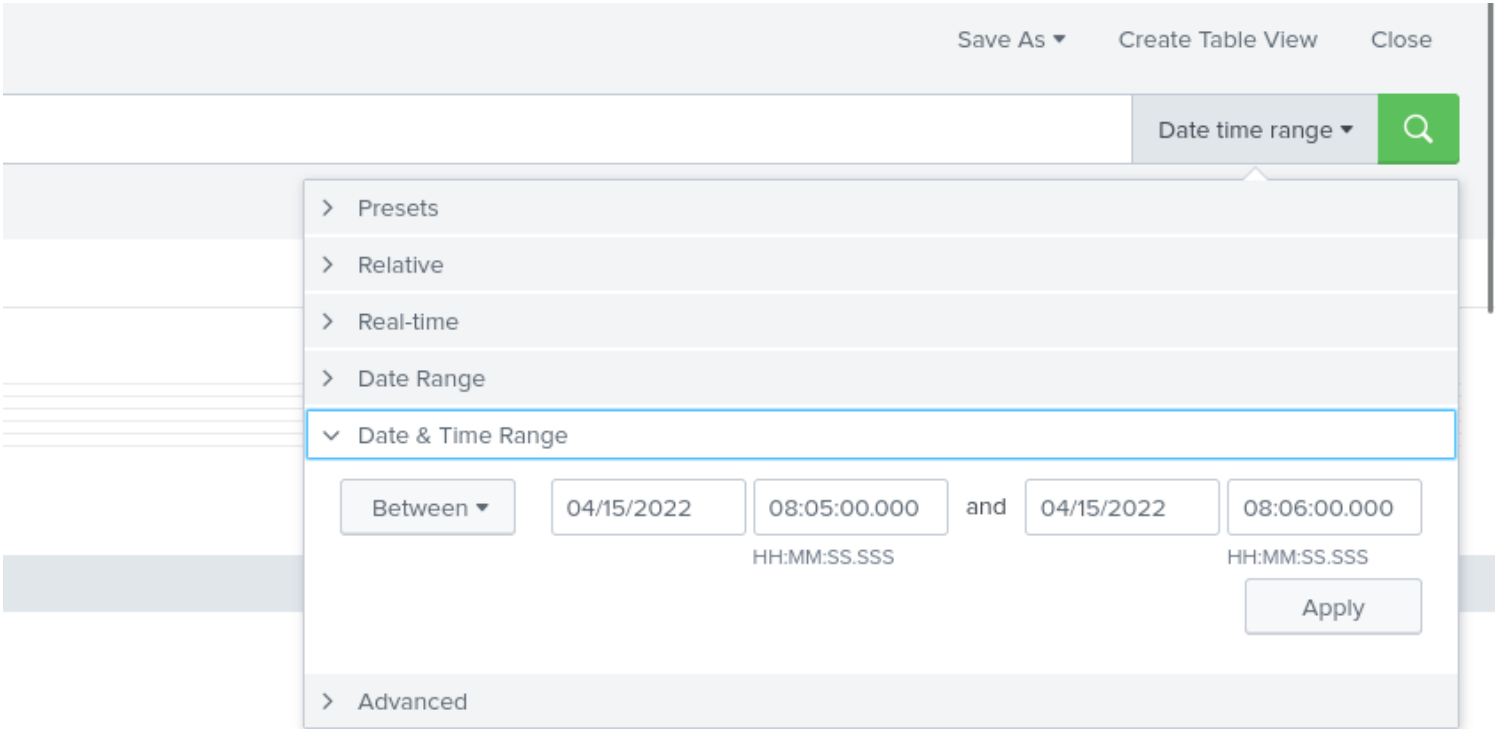
Show as raw text

How many events are returned when we apply the time filter to display events on 04/15/2022 and Time from 08:05 AM to 08:06 AM?

134

✓ Correct

I first applied the filters from the question in the splunk web interface as seen below, and then proceeded to search for the results as shown from the second image below.



How many Events are returned when searching for Event ID 1 **AND** User as \*James\*?

4

✓ Correct

Four events are returned as seen from the image below. This is after applying the boolean operator “AND” in my SPL.

splunk>enterprise Apps ▾

Search Analytics Datasets Reports Alerts Dashboards

## New Search

1 index=windowslogs\* EventID = 1 AND User = \*James\*

✓ 4 events (before 7/12/24 5:44:09.000 AM) No Event Sampling ▾

Events (4) Patterns Statistics Visualization

Format Timeline ▾ — Zoom Out + Zoom to Selection × Deselect

List ▾ ✎ Format 50 Per Page ▾

< Hide Fields	≡ All Fields	i	Time	Event
SELECTED FIELDS		>	4/15/22 8:06:02.000 AM	{ [-] @version: 1

host 1

How many events are observed with Destination IP 172.18.39.6 AND destination Port 135?

4 ✓ Correct

Four events are observed after filtering the results by setting dest ip and dest port as shown in the image below.

splunk>enterprise Apps ▾

Search Analytics Datasets Reports Alerts Dashboards

## New Search

1 index=windowslogs\* DestinationIp="172.18.39.6" DestinationPort="135"

✓ 4 events (before 7/12/24 5:54:13.000 AM) No Event Sampling ▾

Events (4) Patterns Statistics Visualization

Format Timeline ▾ — Zoom Out + Zoom to Selection × Deselect

List ▾ ✎ Format 50 Per Page ▾

< Hide Fields	≡ All Fields	i	Time	Event
SELECTED FIELDS		>	4/15/22 8:06:02.000 AM	{ [-] @version: 1 AccountName: SYSTEM AccountType: User Category: Network connection detected (rule: NetworkConnect) Channel: Microsoft-Windows-Sysmon/Operational

host 1  
source 1  
sourcetype 1  
User 1

What is the Source IP with highest count returned with this Search query?

Search Query: index=windowslogs Hostname="Salena.Adam" DestinationIp="172.18.38.5"

172.90.12.11 ✓ Correct

Given the **Search Query: index=windowslogs \* Hostname="Salena.Adam" DestinationIp="172.18.38.5"**, From the results returned, I checked the ip from the sourceip column as seen below.

```
# ProcessId 4
a Protocol 1
a ProviderGuid 1
a punct 1
# RecordNumber 19
a RuleName 1
a Severity 1
# SeverityValue 1
a SourceHostname 1
a SourceIp 2
a SourceIpV6 1
a SourceModuleName 1
a SourceModuleType 1
a SourceName 1
# SourcePort 18
a SourcePortName 1
a splunk_server 1
a tags[] 1
# Task 1
# ThreadID 1
a timestamp 16
a UserID 1
a UtcTime 18
```

PROCESSID: 100  
Protocol: tcp  
ProviderGuid: {5770385F-C22A-43E0-BF4C-06F5698FFBD9}

SourceIp

2 Values, 100% of events

Selected Yes No

Reports

Top values Top values by time Rare values

Events with this field

Values	Count	%
172.90.12.11	17	89.474%
172.18.38.5	2	10.526%

User: NT AUTHORITY\SYSTEM  
UserID: S-1-5-18  
UtcTime: 2022-04-15 12:06:27.548  
Version: 5  
host: cybertees.net  
port: 60427

In the index windowslogs, search for all the events that contain the term **cyber** how many events returned?

0

✓ Correct

Using this SPL "index=windowslogs\* cyber" events that contain the cyber term is zero.

splunk>enterprise Apps ▾

Search Analytics Datasets Reports Alerts Dashboards

## New Search

1 index=windowslogs\* cyber

✓ 0 events (before 7/12/24 6:16:23.000 AM) No Event Sampling ▾

Events (0) Patterns Statistics Visualization

Format Timeline ▾ — Zoom Out + Zoom to Selection × Deselect

Now search for the term **cyber\***, how many events are returned?

12256

✓ Correct

Though by adding \* as per the instructions from the quiz above, the result returns 12256 events recorded.

splunk>enterpriseApps

SearchAnalyticsDatasetsReportsAlertsDashboards

New Search

1index=windowslogs\*cyber\*

✓12,256 events (before 7/12/24 6:18:17.000 AM)No Event Sampling

Events (12,256)PatternsStatisticsVisualization

Format TimelineZoom OutZoom to SelectionDeselect

ListFormat50 Per Page

< Hide FieldsAll Fields

SELECTED FIELDS  
a host 1  
a source 1  
a sourcetitle 1

i

TimeEvent

What is the third EventID returned against this search query?

Search Query: `index=windowslogs | table _time EventID Hostname SourceName | reverse`

4103

✓ Correct

Using this **Search Query:**  
`index=windowslogs | table _time EventID Hostname SourceName | reverse`, I was able to retrieve the results as shown below.

splunk>enterpriseApps

SearchAnalyticsDatasetsReportsAlertsDashboards

Messages

New Search

1index=windowslogs | table \_time EventID dedup Hostname SourceName  
2 | reverse

✓12,256 events (before 7/12/24 6:26:33.000 AM)No Event Sampling

Events (12,256)PatternsStatistics (12,256)Visualization

100 Per PageFormatPreview

_time	EventID	dedup	Hostname	SourceName
2022-04-15 08:05:46	800		James.browne	PowerShell
2022-04-15 08:05:46	800		James.browne	PowerShell
2022-04-15 08:05:46	4103		James.browne	Microsoft-Windows-PowerShell
2022-04-15 08:05:46	800			PowerShell
2022-04-15 08:05:46	4103		James.browne	Microsoft-Windows-PowerShell

Use the dedup command against the Hostname field before the reverse command in the query mentioned in Question 1. What is the first username returned in the H

Salena.Adam

✓ Correct

dedup cmd in SPL is used to remove duplicate values. By applying it in the Hostname field, I was able to retrieve the first username as shown in the image below.

splunk>enterprise

Apps

Messages

Settings

Search

Analytics

Datasets

Reports

Alerts

Dashboards

## New Search

1 index=windowslogs | table \_time EventID dedup Hostname SourceName | dedup Hostname

2 | reverse

✓ 12,256 events (before 7/12/24 6:27:23.000 AM)

No Event Sampling

Job

Events (12,256)

Patterns

Statistics (3)

Visualization

100 Per Page

Format

Preview

_time	EventID	dedup	Hostname	SourceName
2022-04-15 08:06:38	3		Salena,Adam	Microsoft-Windows-Sysmon
2022-07-14 23:37:59	5156		James.browne	Microsoft-Windows-Security-Auditing
2022-07-14 23:37:59	10		Micheal.Beaven	Microsoft-Windows-Sysmon

Using the Reverse command with the search query `index=windowslogs | table _time EventID Hostname SourceName` - what is the HostName that comes on the

✓ Correct A

Structuring the results using the "tail" cmd alongside with the SPL query given in the question above, I was able to retrieve the Hostname that appeared at the top.

splunk>enterprise
Apps
Messages
Settings

Search
Analytics
Datasets
Reports
Alerts
Dashboards

## New Search

```
1 index=windowslogs | table _time EventID Hostname SourceName
2 | tail
```

✓ 12,256 events (before 7/12/24 6:32:08.000 AM)
No Event Sampling
Jobs

Events (12,256)
Patterns
Statistics (10)
Visualization

100 Per Page
Format
Preview

_time	EventID	Hostname	SourceName
2022-04-15 08:05:46	800	James.browne	PowerShell
2022-04-15 08:05:46	800	James.browne	PowerShell
2022-04-15 08:05:46	4103	James.browne	Microsoft-Windows-PowerShell
2022-04-15 08:05:46	800	James.browne	PowerShell
2022-04-15 08:05:46	4103	James.browne	Microsoft-Windows-PowerShell
2022-04-15 08:05:46	800	James.browne	PowerShell

What is the last EventID returned when the query in question 1 is updated with the **tail** command?

✓ Correct

Updating the query in the previous question, I Structuring the results using the "tail" cmd alongside with the SPL query given in the question above, I was able to retrieve the EventID that appeared at the bottom.

splunk>enterprise Apps

Search Analytics Datasets Reports Alerts Dashboards

## New Search

```
1 index=windowslogs | table _time EventID Hostname SourceName
2 | tail
```

✓ 12,256 events (before 7/12/24 6:32:08.000 AM) No Event Sampling

Events (12,256) Patterns **Statistics (10)** Visualization

100 Per Page Format Preview

_time	EventID	Hostname	SourceName
2022-04-15 08:05:46	800	James.browne	PowerShell
2022-04-15 08:05:46	800	James.browne	PowerShell
2022-04-15 08:05:46	4103	James.browne	Microsoft-Windows-PowerShell
2022-04-15 08:05:46	800	James.browne	PowerShell
2022-04-15 08:05:46	4103	James.browne	Microsoft-Windows-PowerShell
2022-04-15 08:05:46	800	James.browne	PowerShell
2022-04-15 08:05:46	4103	James.browne	Microsoft-Windows-PowerShell
2022-04-15 08:05:46	800	James.browne	PowerShell
2022-04-15 08:05:46	800	James.browne	PowerShell
2022-04-15 08:05:46	4103	James.browne	Microsoft-Windows-PowerShell

Sort the above query against the SourceName. What is the top SourceName returned?

Microsoft-Windows-Directory-Services-SAM

✓ Correct

Using the sort keyword, I added it to my SPL query against the SourceName, and successfully retrieved the top sourcename as in the image below.

splunk>enterprise Apps

Search Analytics Datasets Reports Alerts Dashboards

## New Search

```
1 index=windowslogs | table _time EventID Hostname SourceName
2 | sort SourceName
```

✓ 12,256 events (before 7/12/24 6:34:04.000 AM) No Event Sampling

Events (12,256) Patterns **Statistics (10,000)** Visualization

100 Per Page Format Preview

_time	EventID	Hostname	SourceName
2022-04-15 08:06:07	16977	James.browne	Microsoft-Windows-Directory-Services-SAM
2022-04-15 08:06:07	1582	James.browne	Microsoft-Windows-GroupPolicy
2022-04-15 08:06:48	4183	James.browne	Microsoft-Windows-PowerShell
2022-04-15 08:06:48	4183	James.browne	Microsoft-Windows-PowerShell
2022-04-15 08:06:48	4183	James.browne	Microsoft-Windows-PowerShell

List the top 8 Image processes using the top command - what is the total count of the 6th Image?

196

✓ Correct

I used the transformational SPL “**top limit=8 image**” to filter the results and checked out the results in the “image” field as shown from the image below.



UtcTime: 2022-04-15 12:05:44.355  
SourceProcessGUID: {83d0c8c3-43cb-5f5f-1000-00000000400}

**Image** ×

35 Values, 34.497% of events Selected Yes No

**Reports**

Top values Top values by time Rare values

Events with this field

**Top 10 Values**

	Count	%	
C:\windows\system32\svchost.exe	1,642	38.836%	<div></div>
C:\windows\system32\backgroundTaskHost.exe	547	12.938%	<div></div>
C:\Windows\System32\svchost.exe	426	10.076%	<div></div>
C:\windows\system32\taskhostw.exe	250	5.913%	<div></div>
C:\Windows\System32\BackgroundTransferHost.exe	210	4.967%	<div></div>
C:\Windows\System32\backgroundTaskHost.exe	196	4.636%	<div></div>
C:\Windows\System32\wbem\WmiPrvSE.exe	108	2.554%	<div></div>
C:\Windows\System32\usocoreworker.exe	95	2.247%	<div></div>
C:\Windows\System32\RuntimeBroker.exe	93	2.2%	<div></div>
C:\windows\system32\BackgroundTransferHost.exe	88	2.081%	<div></div>

TargetProcessGUID: {83d0c8c3-4479-5f5f-5602-00000000400}  
TargetProcessId: 1268  
Task: 10

James ✓ Correct

splunk>enterprise

Apps

Messages

Search

Analytics

Datasets

Reports

Alerts

Dashboards

## New Search

1

index=windowslogs | rare User , EventID

✓ 12,256 events (before 7/12/24 6:45:34.000 AM)

No Event Sampling

Events (12,256)

Patterns

Statistics (9)

Visualization

100 Per Page

Format

Preview

User	EventID	count
Cybertees\James	3	1
Cybertees\James	1	4
NT AUTHORITY\NETWORK SERVICE	1	5
Cybertees\Alberto	1	7
Cybertees\Alberto	23	8
Cybertees\Alberto	3	9
NT AUTHORITY\SYSTEM	1	

70 ✓ Correct

9/11

```

a EventReceivedTime 8
a EventTime 5
# ExecutionProcessID 3
a extracted_EventType 3
a extracted_host 1
a FileVersion 15
a Hashes 33
a Hostname 2
a Image 2
a ImageLoaded 33
a index 1
# Keywords 3
# linecount 1
a Message 100
a Opcode 1
# OpcodeValue 1
a OriginalFileName 32
# port 1

```

```

SourceImage: C:\windows\system32\svchost.exe
TargetProcessGUID: {2d351099-5bb3-5f5f-2104-00000000400}

```

## Image

2 Values, 72% of events

Selected

Yes

No

## Reports

Top values

Top values by time

Rare values

Events with this field

## Values

Count

%

C:\Windows\System32\conhost.exe

70

97.222%

C:\windows\system32\svchost.exe

2

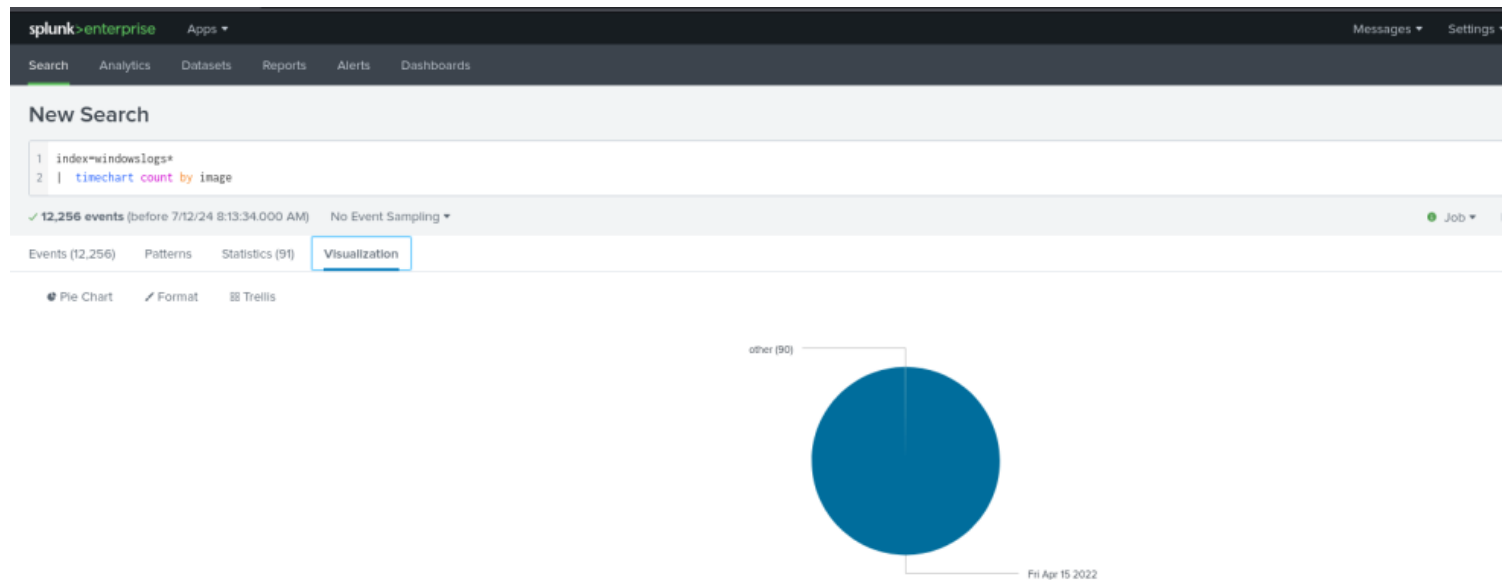
2.778%

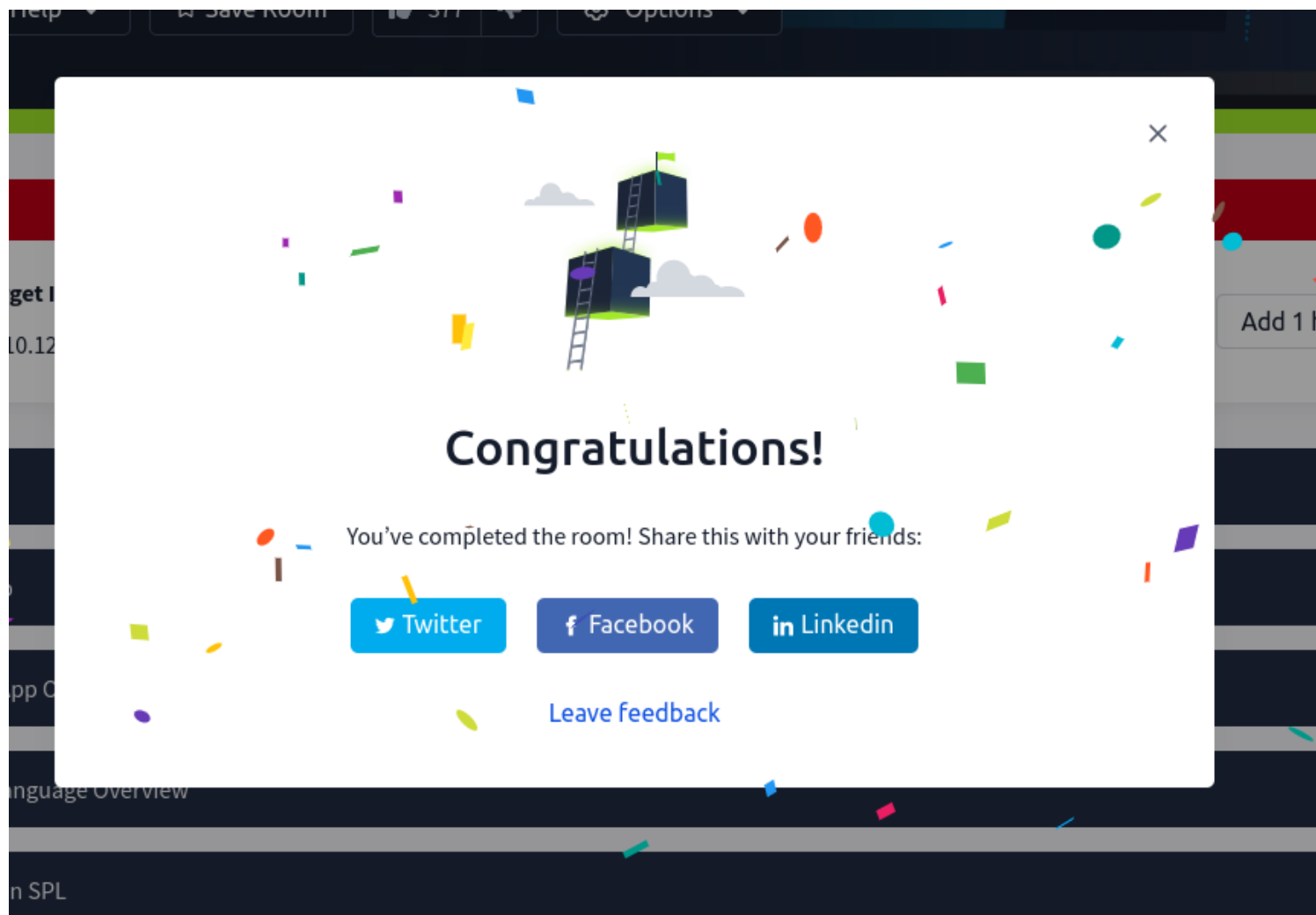
SourceImage: C:\windows\system32\svchost.exe

SourceModuleName: eventlog

SourceModuleType: im msvistalog

Using the `timechart count by Image` cmd, the query displayed the Image chart based on the time as seen below.





<https://tryhackme.com/r/room/splunkexploringspl>

## CONCLUSION

In this session, we explored the fundamental aspects of Splunk's powerful search and analysis features, particularly focusing on the Search Processing Language (SPL). As a newcomer to Splunk, it is evident that the tool offers immense potential for analyzing machine-generated data, making it a valuable asset for any organization aiming to leverage its data for actionable insights.

### Key Learnings:

- 1. Understanding Search Processing Language (SPL)**
- 2. Applying Filters to Narrow Down Results**
- 3. Using Transformational Commands**
- 4. Changing the Order of the Results**

As a new user exploring Splunk for the first time, I find it to be an incredibly powerful and intuitive tool. The ability to search and explore machine data using SPL opens up numerous possibilities for data analysis and operational efficiency. The structured approach to learning SPL, from understanding basic queries to applying filters and using transformational commands, provides a clear pathway to mastering this tool. Indeed, the first interaction with Splunk has been super great, and it is exciting to see the potential it holds for future data-driven endeavors. In conclusion, this introductory session on Splunk and SPL has been enlightening and promising. The robust functionality and flexibility of Splunk make it an indispensable tool for anyone looking to harness the power of their machine data.