



Splunk Fundamentals 2

Document Usage Guidelines

- Should be used only for enrolled students
- Not meant to be a self-paced document, an instructor is needed
- Do not distribute

Generated for Anh Hoang (hoang_anhkim@bah.com) (C) Splunk Inc, not for distribution

Course Prerequisites

- To be successful in this course, you should have completed:
 - Splunk Fundamentals Part 1

Note



In order to receive credit for this course, you must complete all lab exercises.

Course Guidelines

- Hands-on lab exercises reinforce information presented in the lecture modules
- To receive a certificate of completion for the course, you must complete the lab exercises
- The lab exercises must be completed sequentially
 - Later lab exercises often depend on steps completed in previous lab exercises

Course Goals

- Use transforming commands and visualizations
- Filter and format the results of a search
- Correlate events into transactions
- Create and manage Knowledge Objects
- Create & manage extracted fields, field aliases, calculated fields
- Create tags and event types
- Create and use macros and workflow objects
- Create and manage data models
- Use the Splunk Common Information Model (CIM)

Generated for Anh Hoang (hoang_anhkim@bah.com) (C) Splunk Inc, not for distribution

Course Outline

Module 1: Beyond Search Fundamentals

Module 2: Using Transforming Commands for Visualization

Module 3: Using Trendlines, Mapping, and Single Value Commands

Module 4: Filtering Results and Manipulating Data

Module 5: Correlating Events

Module 6: Introduction to Knowledge Objects

Module 7: Creating and Managing Fields

Module 8: Creating Field Aliases and Calculated Fields

Module 9: Creating Tags and Event Types

Module 10: Creating and Using Macros

Module 11: Creating and Using Workflow Actions

Module 12: Creating Data Models

Module 13: Using the Common Information Model (CIM) Add-on

Generated for Anh Hoang (hoang_anhkim@bah.com) (C) Splunk Inc, not for distribution

Callouts

- Scenarios

- Many of the examples in this course relate to a specific scenario
 - For each example, a question is posed from a colleague or manager at Buttercup Games

- Notes & Tips

References for more information on a topic and tips for best practices

Scenario



The online sales manager wants to see the action, productId, and status of customer interactions in the online store.

Note



Lookups are discussed in the *Splunk Fundamentals Part 1* course.

Course Scenario

- Use cases in this course are based on Buttercup Games, a gaming company
- Searches and reports are based on:
 - Business analytics from the web access logs and lookups
 - Internal operations information from mail and internal network data
 - Security operations information from internal network and badge reader data

Buttercup Games, Inc.

- Multinational company with HQ in San Francisco and offices in Boston and London
- Sells products through its online store and 3rd party stores worldwide



Generated for Anh Hoang (hoang_anhkim@bah.com) (C) Splunk Inc, not for distribution

Your Role at Buttercup Games

- You are a Splunk power user with a great understanding of all your company's data
- Your responsibilities are to provide information to users throughout the company and to create and manage Splunk knowledge objects for your stakeholders
- You implement best practices for naming conventions of all knowledge objects
- You gather data and statistics, and report on Security, IT Operations, Operational Intelligence, etc.

Buttercup Games Network

Index	Description	Sourcetype	Host
web	Online transactions	access_combined	www1
			www2
			www3
security	Badge reader data	history_access	badgesv1
	AD/DNS data	winauthentication_security	adldapsv1
	Web login data	linux_secure	www1
	www2		
	www3		
sales	Retail sales data	vendor_sales	vendorUS1
	BI data	sales_entries	ecommsv1
network	Firewall data	cisco_firewall	cisco_router1
	Email data	cisco_esa	
	Web appliance data	cisco_wsa_squid	
games	Game logs	SimCubeBeta	sim_cube_server

Generated for Anh Hoang (hoang_anhkim@bah.com) (C) Splunk Inc, not for distribution

Module 1: Beyond Search Fundamentals

Generated for Anh Hoang (hoang_anhkim@bah.com) (C) Splunk Inc, not for distribution

Module Objectives

- Review basic search commands
- Use case correctly in searches
- Describe Splunk's search process

Basic Search Review

- **Keywords**

For example, search for a single word (e.g., error) or group of words (e.g., error password)

- **Booleans**

NOT, OR, AND; AND is implied; MUST be uppercase; can use ()'s to force precedence

sourcetype=vendor_sales OR (sourcetype=access_combined action=purchase)

- **Phrases**

"web error" (different than web AND error)

- **Field searches**

status=404, user=admin

- **Wildcard (*)**

- status=40* matches 40, 40a, 404, etc.

- Starting keywords with a wildcard is very inefficient, e.g. *dmin

- **Comparisons**

=, !=, <, <=, >, >=, > status>399, user!=admin

Generated for Anh Hoang (hoang_anhkim@bah.com) (C) Splunk Inc, not for distribution

Basic Search Review (cont.)

- **table**: returns table containing only specified fields in result set
- **rename**: renames a field in results
- **fields**: includes or excludes specified fields
- **dedup**: removes duplicates from results
- **sort**: sorts results by specified field
- **lookup**: adds field values from external source (e.g., csv files)

Case Sensitivity – Sensitive

Case sensitive	Examples
Boolean operators (uppercase)	AND, OR, NOT (Boolean operators) and, or, not (literal keywords)
Field names	productId vs. productid <code>eval cs_username = "Total Access"</code>
Field values from lookup (default, but configurable)	product_name="Tulip Bouquet" vs. product_name="tulip bouquet"
Regular expressions	\d\d\d vs. \D\D\D
Field values used with eval and where commands	<code>eval action;if(action=="view",...)</code> <code>where action="Purchase"</code> <code>stats count(eval(action="view")) as...</code>
Tags	tag=DMZ vs. tag=dmz

Case Sensitivity – Insensitive

Case insensitive	Examples
Command names	STATS, stats, sTaTs
Command clauses	AS used by stats, rename, ...; BY used by stats, chart, top, ...; WITH used by replace
Search terms	failed, FAILED, Failed
Statistical functions	avg, AVG, Avg used by stats, chart, ...
Field values	host=www1, host=WwW1 (unless coming from a lookup)

Note

You can perform a case-sensitive search by including the CASE directive in your search string; e.g., action=CASE(purchase).

Generated for Anh Hoang (hoang_anhkim@bah.com) (C) Splunk Inc, not for distribution

How Splunk Searches – Buckets

- As events come in, Splunk places them into an index's hot bucket (only writable bucket)
- As buckets age, they roll from the hot to warm to cold
- Each bucket has its own raw data, metadata, and index files
- Metadata files track source, sourcetype, and host
- Admins can configure settings and add more buckets



Generated for Anh Hoang (hoang_anhkim@bah.com) (C) Splunk Inc, not for distribution

How Splunk Searches – Searching

- When you search, Splunk uses the time range to choose which buckets to search and then uses the bucket indexes to find qualifying events
- When you search for `index=web password fail*` during the last 24 hours:
 - Splunk identifies the buckets for the last 24 hours
 - And searches those buckets for the search terms

Hot: Now to -3h	index	raw events
Hot: -3 to -6h	index	raw events
Hot: -5 to -8h	index	raw events
Warm: -9 to -12h	index	raw events
Warm: -12 to -15h	index	raw events
Warm: -14 to -17h	index	raw events
.....	index	raw events
Warm: -42 to -45h	index	raw events
Warm: -45 to -48h	index	raw events
Cold: -48 to -51h	index	raw events
Cold: -51 to -54h	index	raw events
Cold: -54 to -57h...	index	raw events

Generated for Anh Hoang (hoang_anhkim@bah.com) (C) Splunk Inc, not for distribution

General Search Practices

- As events are stored by time, `_time` is the most efficient filter
- After time, most powerful fields to filter on: `index`, `host`, `source`, and `sourcetype`
- To make searches more efficient, include as many terms as possible
 - Example: searching for `sourcetype=x failure` is better than `failure`
- Use the `fields` command to extract (discover) only fields you need
 - Example: Searching last 365 days, scans 566,720 events

`index=web sourcetype=access_combined`

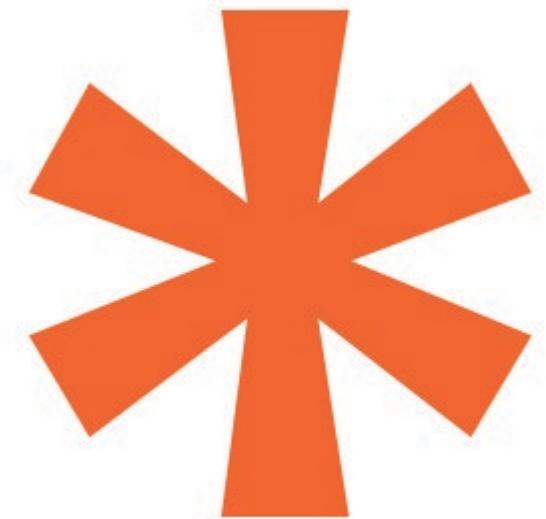
15.16 secs

`index=web sourcetype=access_combined | fields clientip, bytes, referrer`

4.49 secs

General Search Practices – Wildcards

- Splunk only searches for whole words, but wildcards allowed
- Only *trailing* wildcards make efficient use of index
 - Wildcards at *beginning* of string scan all events within time frame
 - Wildcards in *middle* of string may return inconsistent results
 - So use fail* (not *fail or *fail* or f*il)
- Wildcards tested after all other terms



General Search Practices

- Inclusion is generally better than exclusion
 - Searching for "access denied" is faster than NOT "access granted"
- Filter as early in your search as possible
 - Removing duplicates then sorting is faster than sorting then removing duplicates
- Use the appropriate search mode
 - Fast - performance over completeness
 - Smart [default]
 - Verbose - completeness over performance

Transforming Search Commands

- A transforming command:
 - Massages raw data into a data table
 - 'Transforms' specified cell values for each event into numerical values that you can use for statistical purposes
 - Is required to 'transform' search results into visualizations
- Transforming commands include:
 - top
 - rare
 - chart
 - timechart
 - stats
 - geostats

Generated for Anh Hoang (hoang_anhkim@bah.com) (C) Splunk Inc, not for distribution

Reviewing Search Mode – Fast Mode

- Emphasizes performance, returning only essential and required data
- For non-transforming searches:
 - ✓ Events – fields sidebar displays only those fields required for the search
 - ✓ Patterns
 - ✗ Statistics or visualizations
- Contents of interesting fields sidebar are lost

index=web sourcetype=access_combined

The screenshot shows a Splunk search interface with the search bar containing "index=web sourcetype=access_combined". Below the search bar, there are tabs for "Events (147)", "Patterns", "Statistics", and "Visualization". The "Events (147)" tab is selected. At the top right, there are buttons for "Format Timeline", "Zoom Out", "Zoom to Selection", and "Deselect". Below these are buttons for "List", "Format", and "20 Per Page". The main area displays a timeline with green bars representing event intervals. To the right is a table of event details:

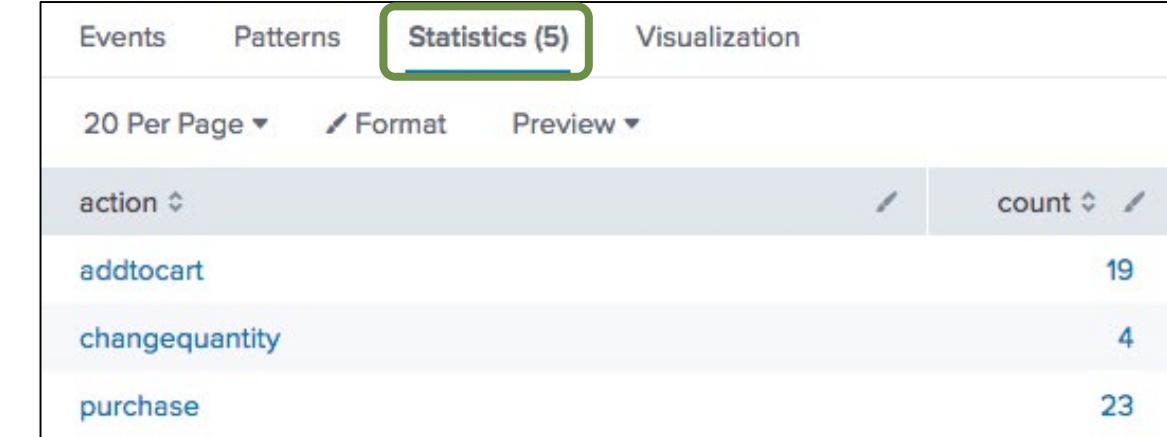
i	Time	Event
>	4/6/18 7:04:04.000 PM	188.138.40.16 SESSIONID=SD4S la/5.0 (Macin .38 Safari/53 host = www1
>	4/6/18 7:03:50.000 PM	188.138.40.16 DFF4961 HTTP ; U; Intel Ma 304 host = www1

On the left side, there are sections for "SELECTED FIELDS" (host 3, source 3, sourcetype 1) and "INTERESTING FIELDS" (index 1, linecount 1, splunk_server 4). A button "+ Extract New Fields" is also present.

Reviewing Search Mode – Fast Mode (cont.)

- For transforming searches:
 - ✗ Events
 - ✗ Patterns
 - ✓ Statistics or visualizations

```
index=web sourcetype=access_combined  
| stats count by action
```

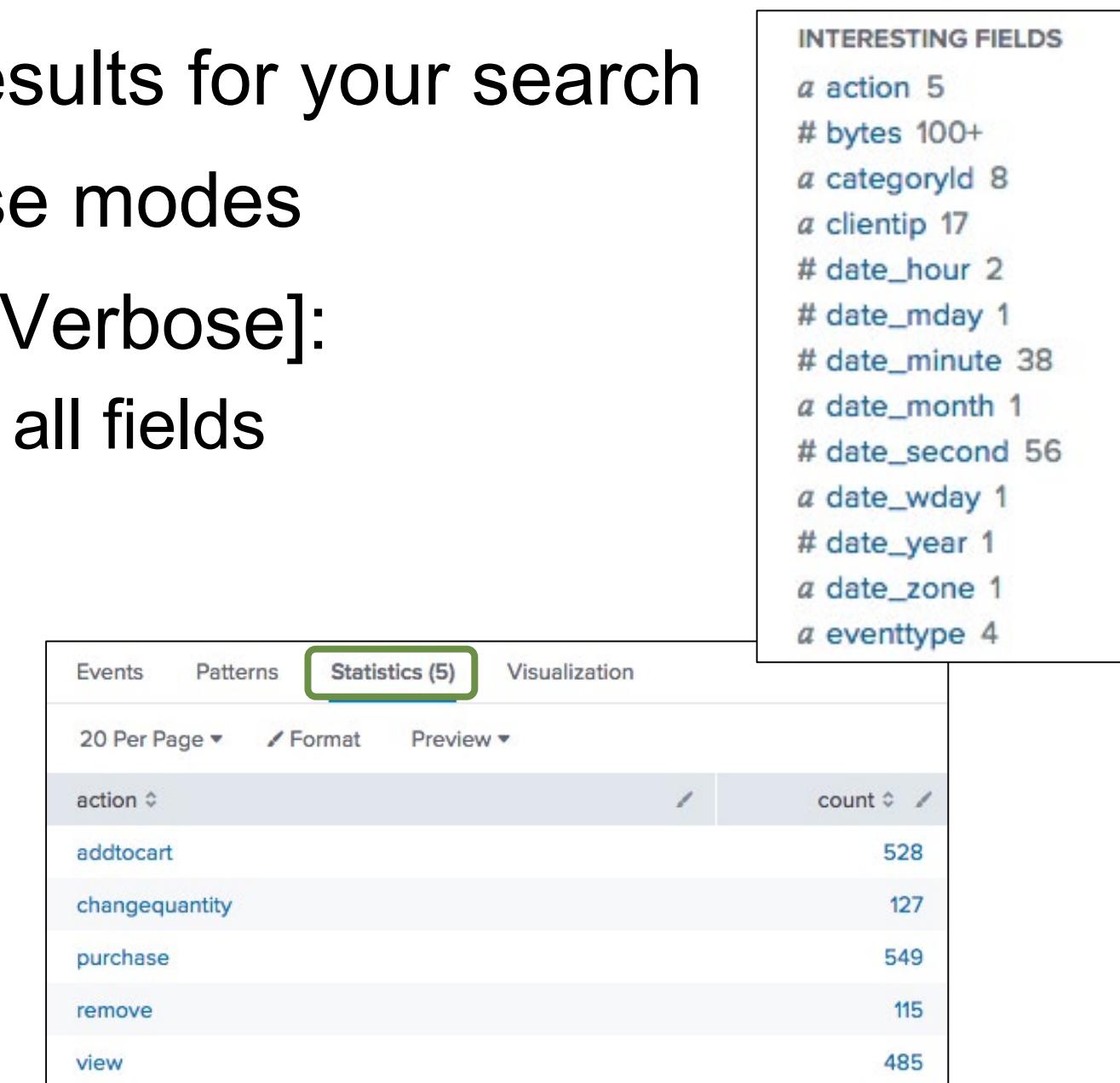


The screenshot shows a search interface with a green box highlighting the search command. Below it is a table titled 'Statistics (5)' with columns for 'action' and 'count'. The table lists four actions: 'addtocart' (19), 'changequantity' (4), and 'purchase' (23), along with a header row for 'action'.

action	count
addtocart	19
changequantity	4
purchase	23

Reviewing Search Mode – Smart Mode (Default)

- Designed to give you the best results for your search
- Combination of Fast and Verbose modes
- For non-transforming searches [Verbose]:
 - ✓ Events – fields sidebar displays all fields
 - ✓ Patterns
 - ✗ Statistics or visualizations
- For transforming searches:
 - ✗ Events
 - ✗ Patterns
 - ✓ Statistics or visualizations



Generated for Anh Hoang (hoang_anhkim@bah.com) (C) Splunk Inc, not for distribution

Reviewing Search Mode – Verbose Mode

- Emphasizes completeness by returning all possible field and event data
- For non-transforming searches:
 - ✓ Events – fields sidebar displays all fields
 - ✓ Patterns
 - ✗ Statistics or visualizations
- For transforming searches:
 - ✓ Events
 - ✓ Patterns
 - ✓ Statistics or visualizations

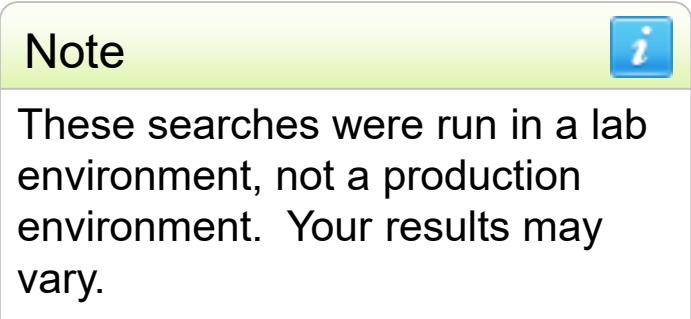
Events (3,672)	Patterns	Statistics (5)	Visualization
Format Timeline ▾	List ▾	/ Format	
◀ Hide Fields	☰ All Fields	i Time	
SELECTED FIELDS		> 1/12/18 1:23:32.000 PM	
a host 3		> 1/12/18 1:23:18.000 PM	
a source 3		> 1/12/18 1:23:03.000 PM	
a sourcetype 1		> 1/12/18 1:22:56.000 PM	
INTERESTING FIELDS		> 1/12/18 1:22:46.000 PM	
a action 5		> 1/12/18 1:22:46.000 PM	
# bytes 100+		> 1/12/18 1:22:46.000 PM	
a categoryId 8		> 1/12/18 1:22:46.000 PM	
a clientip 100+		> 1/12/18 1:22:46.000 PM	
# date_hour 24		> 1/12/18 1:22:46.000 PM	
# date_mday 2		> 1/12/18 1:22:46.000 PM	
# date_minute 60		> 1/12/18 1:22:46.000 PM	
a date_month 1		> 1/12/18 1:22:46.000 PM	
# date_second 60		> 1/12/18 1:22:46.000 PM	
a date_wday 2		> 1/12/18 1:22:46.000 PM	
# date_year 1		> 1/12/18 1:22:46.000 PM	
a date_zone 1		> 1/12/18 1:22:46.000 PM	
a eventtype 4		> 1/12/18 1:22:46.000 PM	
a file 14		> 1/12/18 1:22:46.000 PM	
a ident 1		> 1/12/18 1:22:46.000 PM	
a index 1		> 1/12/18 1:22:46.000 PM	
a itemId 14		> 1/12/18 1:22:46.000 PM	
a JSESSIONID 100+		> 1/12/18 1:22:46.000 PM	
# linecount 1		> 1/12/18 1:22:46.000 PM	
a method 2		> 1/12/18 1:22:46.000 PM	
# other 100+		> 1/12/18 1:22:46.000 PM	
# price 7		> 1/12/18 1:22:46.000 PM	
a product_name 14		> 1/12/18 1:22:46.000 PM	
a productid 15		> 1/12/18 1:22:46.000 PM	
a punct 84		> 1/12/18 1:22:46.000 PM	
a referer 100+		> 1/12/18 1:22:46.000 PM	
a referer_domain 4		> 1/12/18 1:22:46.000 PM	
a req_time 100+		> 1/12/18 1:22:46.000 PM	
# sale_price 6		> 1/12/18 1:22:46.000 PM	
a splunk_server 3		> 1/12/18 1:22:46.000 PM	
# status 9		> 1/12/18 1:22:46.000 PM	
a tag 2		> 1/12/18 1:22:46.000 PM	
a tag:eventtype 2		> 1/12/18 1:22:46.000 PM	
# timeendpos 6		> 1/12/18 1:22:46.000 PM	
# timestamppos 6		> 1/12/18 1:22:46.000 PM	
a url 100+		> 1/12/18 1:22:46.000 PM	
a url_path 14		> 1/12/18 1:22:46.000 PM	
a url_query 100+		> 1/12/18 1:22:46.000 PM	
a user 1		> 1/12/18 1:22:46.000 PM	

Generated for Anh Hoang (hoang_anhkim@bah.com) (C) Splunk Inc, not for distribution

Search Performance – Modes

- Use the most appropriate search mode:

```
index=web sourcetype=access_combined  
| chart count by product_name
```



- Time range: last 365 days

<u>Mode</u>	<u>Returned Results</u>	<u>Events Scanned</u>	<u>Time</u>
Fast	14	566,731	1.82
Smart	14	566,731	1.91
Verbose	14	566,731	15.21

Generated for Anh Hoang (hoang_anhkim@bah.com) (C) Splunk Inc, not for distribution

Search Job Inspector

- Tool allows you to examine:
 - Overall stats of search (e.g., records processed and returned, processing time)
 - How search was processed
 - Where Splunk spent its time
- Use to troubleshoot search's performance and understand impact of knowledge objects on processing (e.g., event types, tags, lookups)
- Any existing (i.e., not expired) search job can be inspected

Note

For more information, see:
docs.splunk.com/Documentation/Splunk/latest/Search/ViewsearchjobpropertieswiththeJobInspector

Search Job Inspector – 3 Components

index=web sourcetype=access_combined
| stats count by action

✓ 3,673 events (1/11/18 3:00:00.000 PM to 1/12/18 3:15:47.000 PM) No Event Sampling ▾

Events (3,673) Patterns Statistics (5) Visualization

Format Timeline ▾ List ▾ Format

Job ▾

- Edit Job Settings...
- Send Job to Background
- Inspect Job**
- Delete Job

Search job inspector

This search has completed and has returned 5 results by scanning 3,673 events in 0.778 seconds
(SID: 1515798947.29) [search.log](#)

> Execution costs

> Search job properties

Server Info: Splunk 7.1.0, 34.215.236.159, Fri Jan 12 15:18:14 2018 User: student1

- Header
- Execution costs
- Search job properties

Generated for Anh Hoang (hoang_anhkim@bah.com) (C) Splunk Inc, not for distribution

Search Job Inspector – Header

The screenshot shows the header section of a search job inspector. It has a dark header bar with the text "Search job inspector" in white. Below it is a light gray content area containing the following text:
This search has completed and has returned 5 results by scanning 3,673 events in 0.778 seconds
(SID: 1515798947.29) [search.log](#)

Top of Search job inspector provides basic information, including time to run and # of events scanned

Search Job Inspector – Execution Costs

- Provides details on cost to retrieve results, such as:
 - command.search.index**
Time to search the index for the location to read in rawdata files
 - command.search.filter**
Time to filter out events that do not match
 - command.search.rawdata**
Time to read events from the rawdata files

Duration (seconds)	Component	Invocations	Input count	Output count
0.02	command.addinfo	20	3,738	3,738
0.01	command.fields	20	3,738	3,738
0.02	command.prestats	20	3,738	80
0.34	command.search	20	-	3,738
0.29	command.search.expand_search	6	-	-
0.02	command.search.calcfIELDS	16	3,738	3,738
0.02	command.search.fieldallas	16	3,738	3,738
0.02	command.search.filter	16	-	-
0.00	command.search.index	4	-	-
0.00	command.search.index.usec_1_8	287	-	-
0.00	command.search.index.usec_8_64	4	-	-
0.28	command.search.rawdata	4	-	-
0.03	command.search.kv	16	-	-

Generated for Anh Hoang (hoang_anhkim@bah.com) (C) Splunk Inc, not for distribution

Lab Exercise 1

Time: 10 – 15 minutes

Tasks:

- Update your user profile: timezone, full name
- Use search job inspector to troubleshoot problems
- Check for issues with customer purchases in the online store
- Use search job inspector to view performance

Module 2: Using Transforming Commands for Visualizations

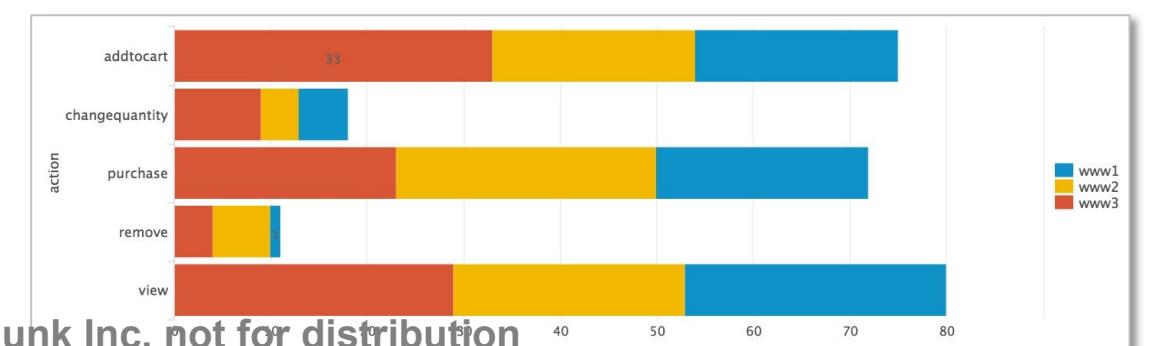
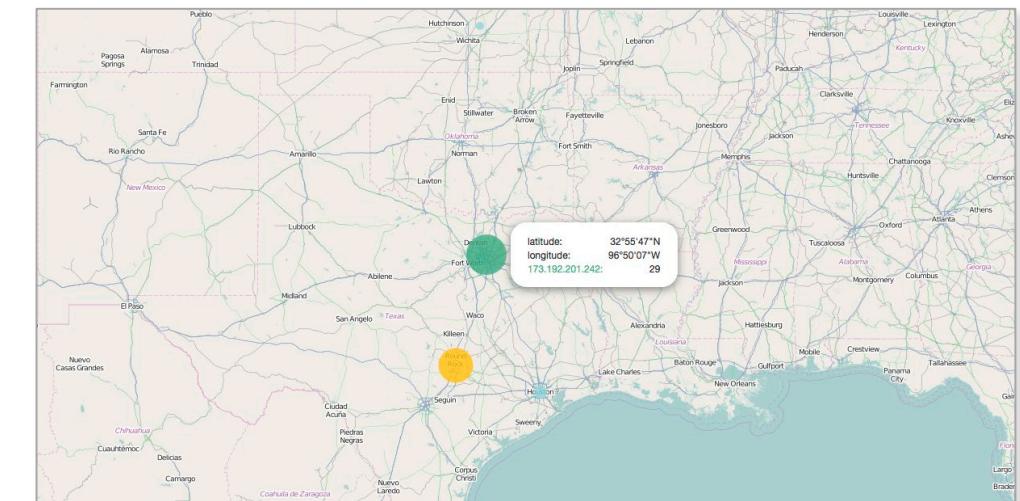
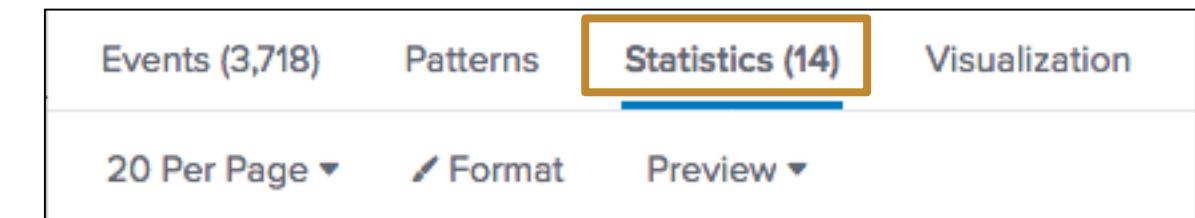
Generated for Anh Hoang (hoang_anhkim@bah.com) (C) Splunk Inc, not for distribution

Module Objectives

- Explore data structure requirements
- Explore visualization types
- Create and format charts
- Create and format timecharts
- Explain when to use each type of reporting command

Visualization Types

- When a search returns statistical values, results can be viewed with a wide variety of visualization types
 - Statistics table
 - Charts: Line, column, pie, etc.
 - Single value, gauges
 - Maps
 - Many more



Generated for Anh Hoang (hoang_anhkim@bah.com) (C) Splunk Inc. not for distribution

Viewing Results as a Visualization

- Not all searches can be visually represented
- A data series is a sequence of related data points that are plotted in a visualization
- Data series can generate various statistical or visualization results

New Search

index=web sourcetype=access_combined |(404 OR 500 OR 503) OR (error OR fail*) Last 60 minutes

Events (10) Patterns Statistics Visualization

Your search isn't generating any statistic or visualization results. Here are some possible ways to get results.

Pivot
Build tables and visualizations using multiple fields and metrics without writing searches.

Quick Reports
Click on any field in the events tab for a list of quick reports like 'Top Referrers' and 'Top Referrers by time'.

Search Commands
Use a transforming search command, like timechart or stats, to summarize the data.

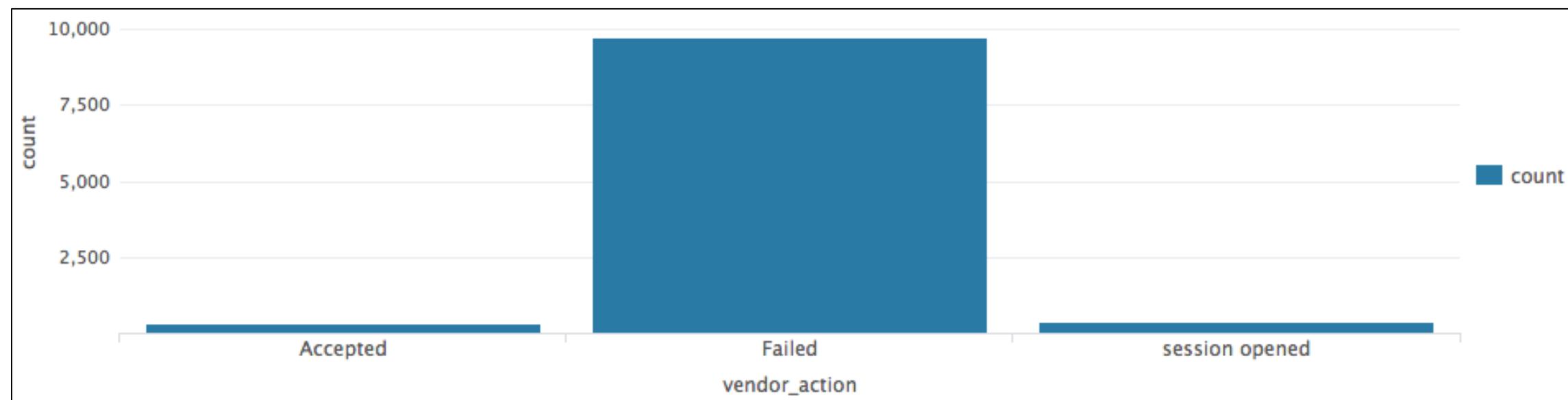
Generated for Anh Hoang (hoang_anhkim@bah.com) (C) Splunk Inc, not for distribution

Data Structure Requirements – Single Series

Most visualizations require a **single series** table (i.e., search results structured as a table with at least two columns)

- Leftmost column provides x-axis values
- Subsequent columns provide numeric y-axis values for each series in the chart

vendor_action	count
Accepted	301
Failed	9740
session opened	400

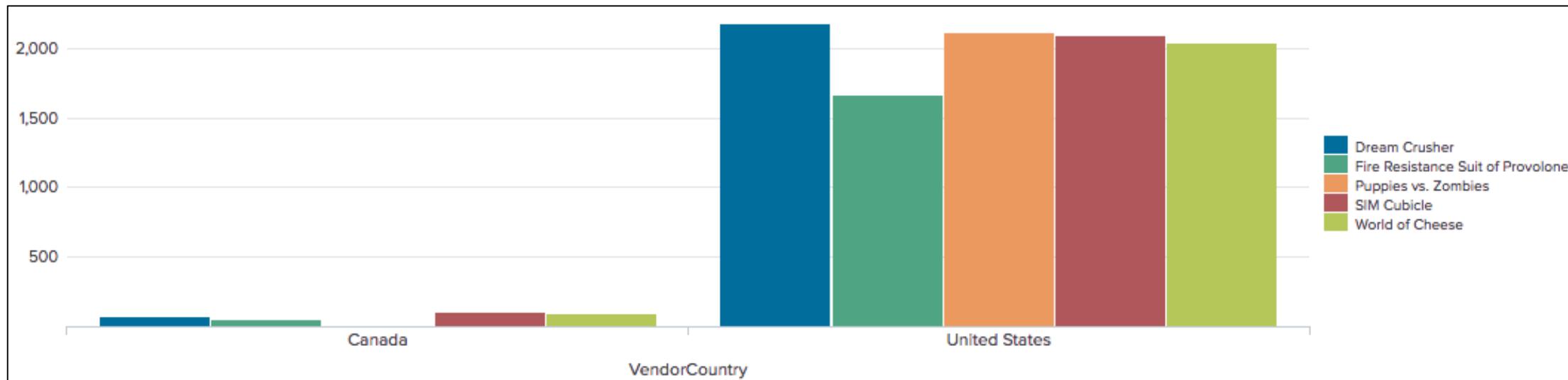


Generated for Anh Hoang (hoang_anhkim@bah.com) (C) Splunk Inc, not for distribution

Data Structure Requirements – Multi-Series

To get **multi-series** tables, you need to set up the underlying search with reporting search commands such as **chart** and **timechart**

VendorCountry	Dream Crusher	Fire Resistance Suit of Provolone	Puppies vs. Zombies	SIM Cubicle	World of Cheese
Canada	77	52	16	105	94
United States	2184	1668	2118	2098	2042



```
index=sales sourcetype=vendor_sales VendorID<4000  
| chart count over VendorCountry by product_name limit=5 useother=f
```

Last 30 days

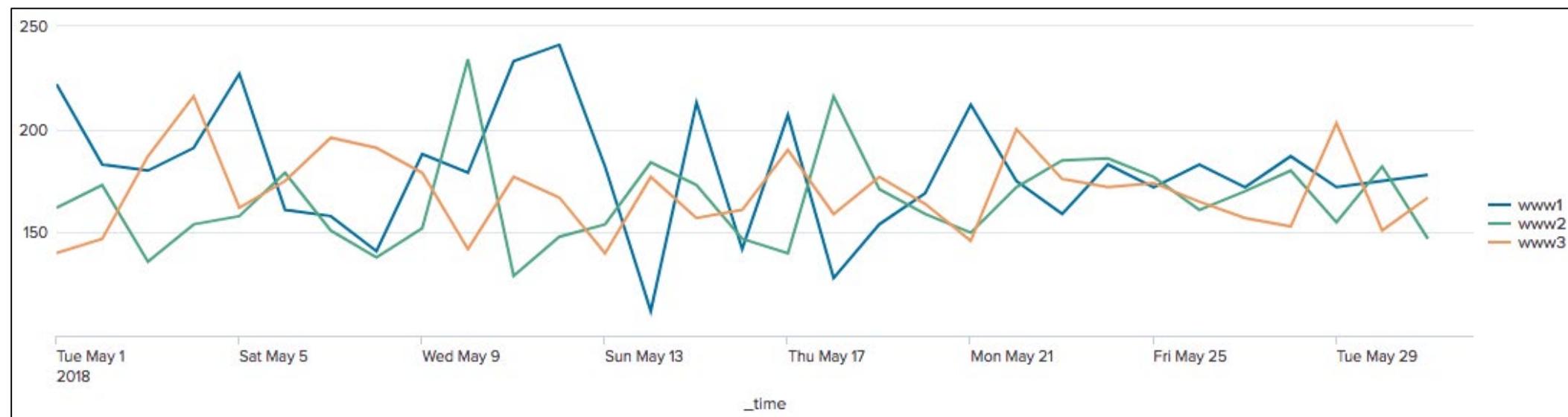
Note

The syntax `count by VendorCountry, product_name` would produce the same result as `count over VendorCountry by product_name`.

Data Structure Requirements – Time Series

- **Time series** displays statistical trends over time
- Can be single-series or multi-series

_time	www1	www2	www3
2018-05-01	222	162	140
2018-05-02	183	173	147
2018-05-03	180	136	187
2018-05-04	191	154	216
2018-05-05	227	158	162

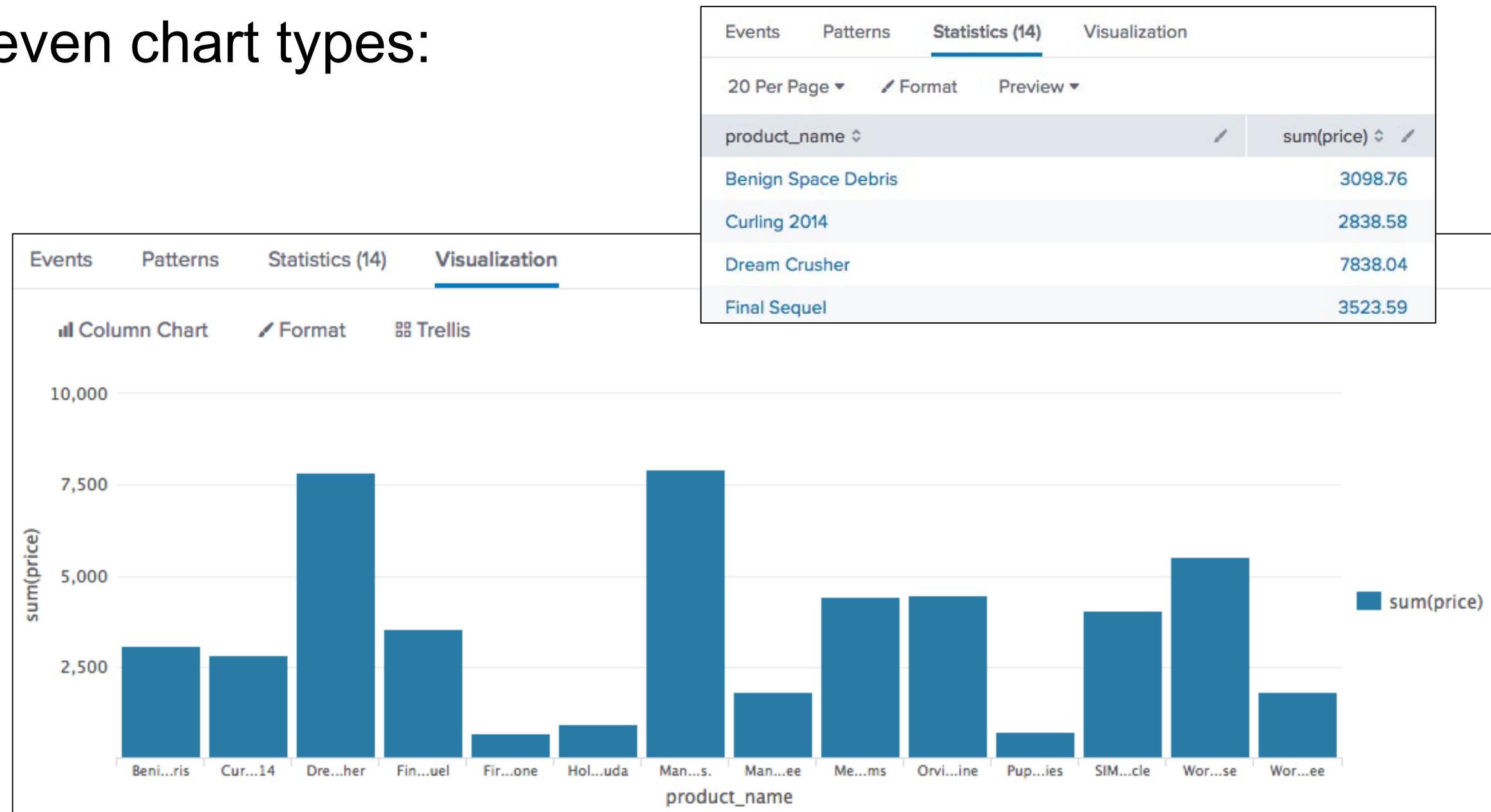


```
index=web sourcetype=access_combined action=purchase status=200  
| timechart count by host
```

Generated for Anh Hoang (hoang_anhkim@bah.com) (C) Splunk Inc, not for distribution

Viewing Results as a Chart

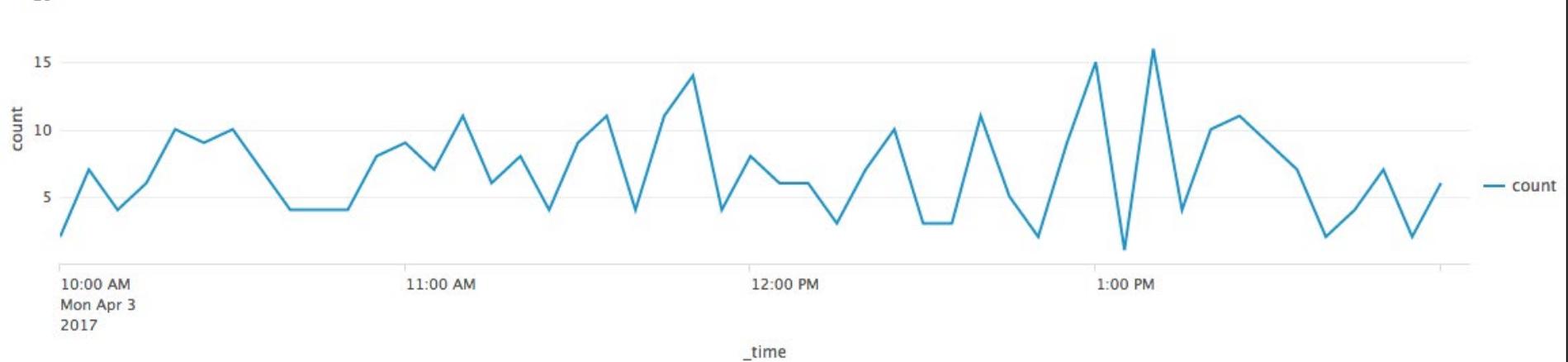
- There are seven chart types:
 - Line
 - Area
 - Column
 - Bar
 - Bubble
 - Scatter
 - Pie



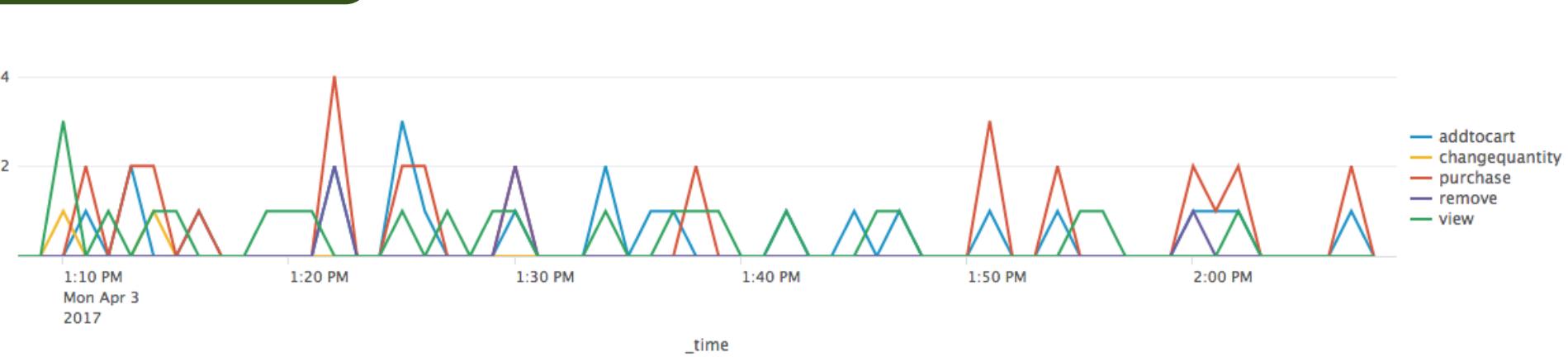
Generated for Anh Hoang (hoang_anhkim@bah.com) (C) Splunk Inc, not for distribution

Charts – Line

```
index=web sourcetype=access_combined action=*
| timechart count
```



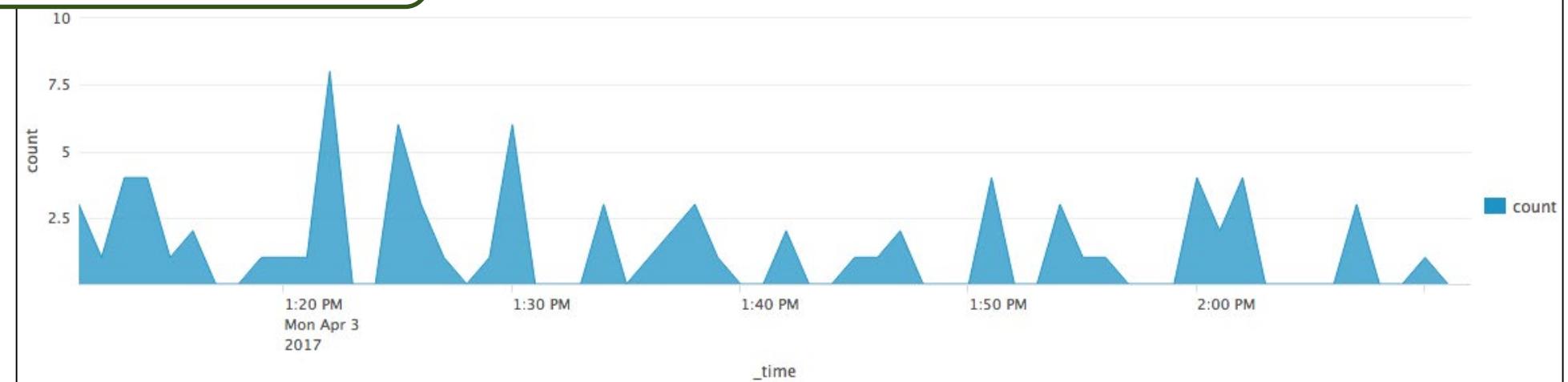
```
index=web sourcetype=access_combined action=*
| timechart count by action
```



Generated for Anh Hoang (hoang_anhkim@bah.com) (C) Splunk Inc. not for distribution

Charts – Area

```
index=web sourcetype=access_combined action=*
| timechart count
```



```
index=web sourcetype=access_combined action=*
| timechart count by action
```



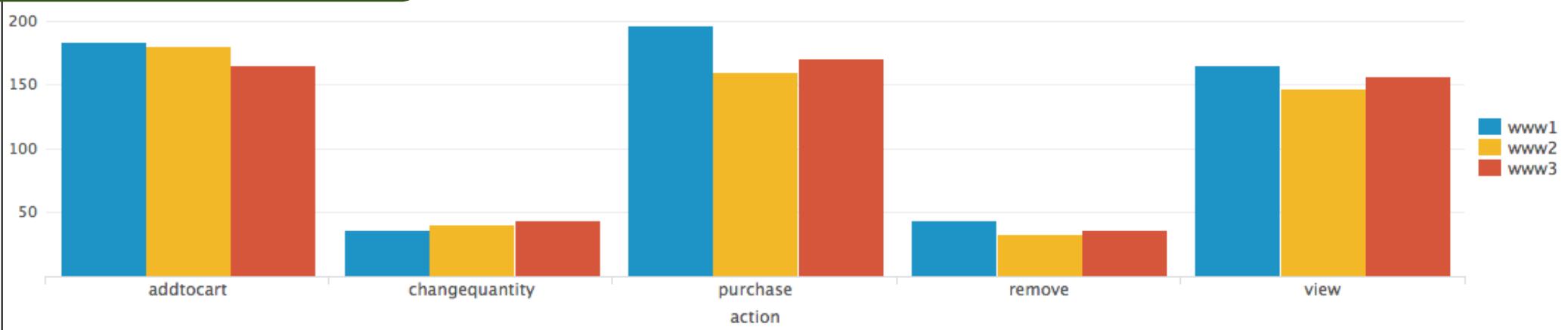
Generated for Anh Hoang (hoang_anhkim@bah.com) (C) Splunk Inc. not for distribution

Charts – Column

```
index=web sourcetype=access_combined action=*
| chart count over action
```



```
index=web sourcetype=access_combined action=*
| chart count over action by host
```



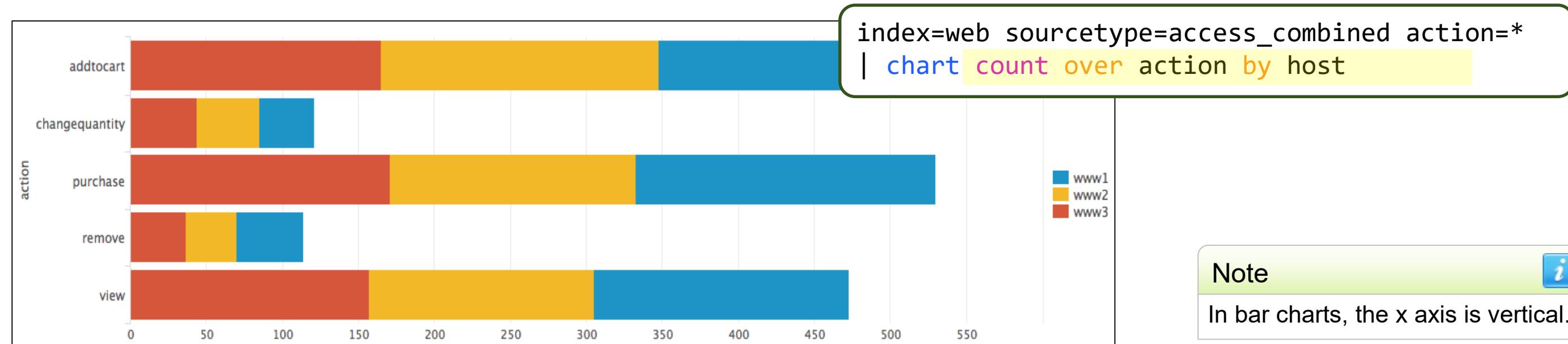
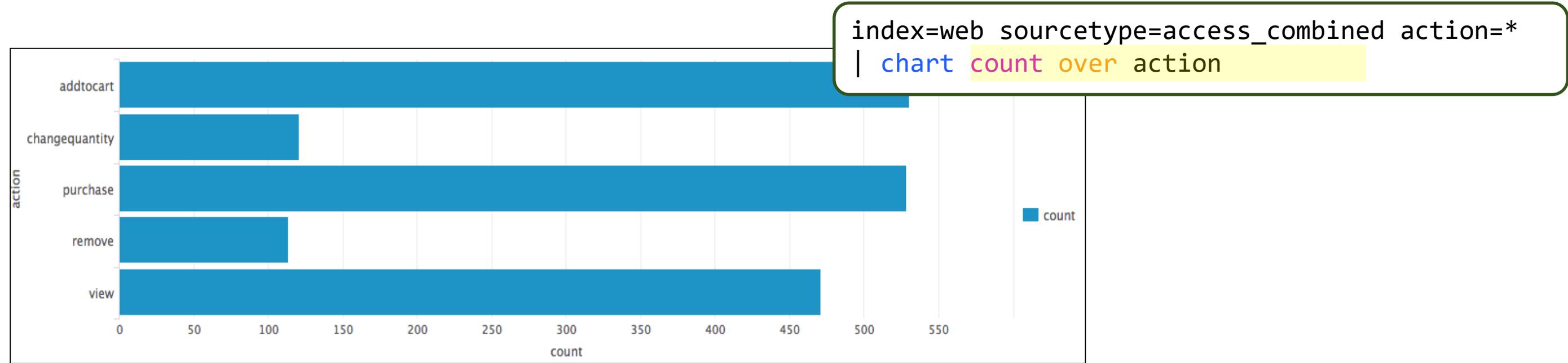
Generated for Anh Hoang (hoang_anhkim@bah.com) (C) Splunk Inc. not for distribution

Charts – Column (Formatted as Stacked)



Generated for Anh Hoang (hoang_anhkim@bah.com) (C) Splunk Inc, not for distribution

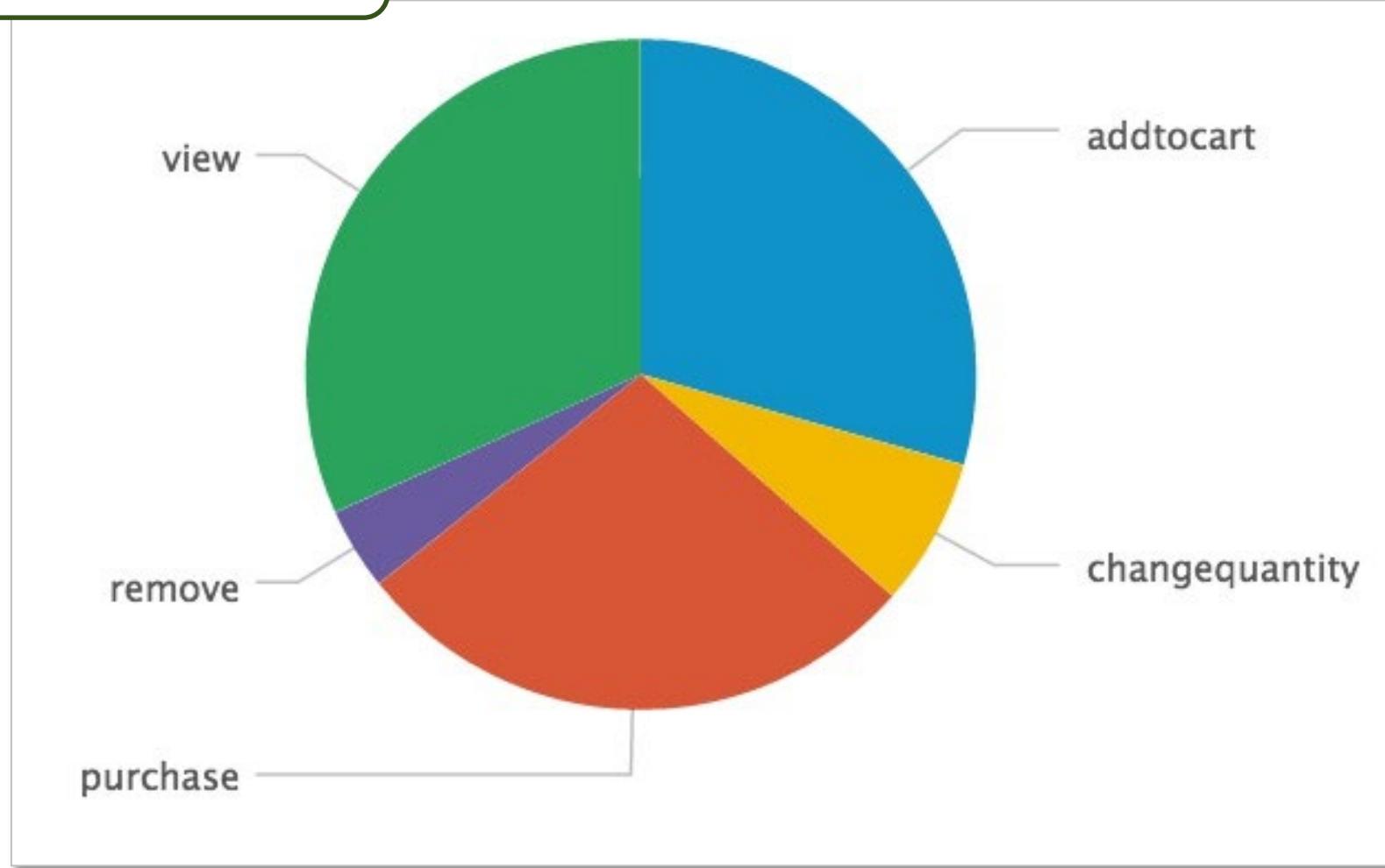
Charts – Bar



Generated for Anh Hoang (hoang_anhkim@bah.com) (C) Splunk Inc, not for distribution

Charts – Pie

```
index=web sourcetype=access_combined action=*  
| chart count over action
```

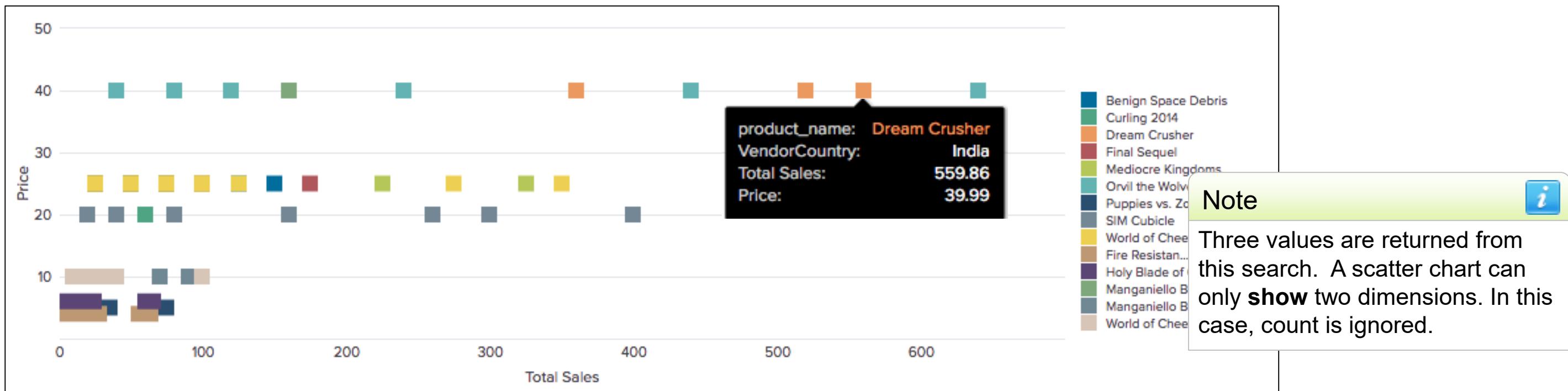


Generated for Anh Hoang (hoang_anhkim@bah.com) (C) Splunk Inc. not for distribution

Charts – Scatter

- Scatter chart shows trends in the relationships between discrete data values
- Generally, it shows discrete values that do not occur at regular intervals or belong to a series

```
index=sales sourcetype=vendor_sales  
VendorID >=7000 AND VendorID <=8999  
| stats sum(price) as "Total Sales",  
values(price) as "Price",  
count by VendorCountry, product_name
```

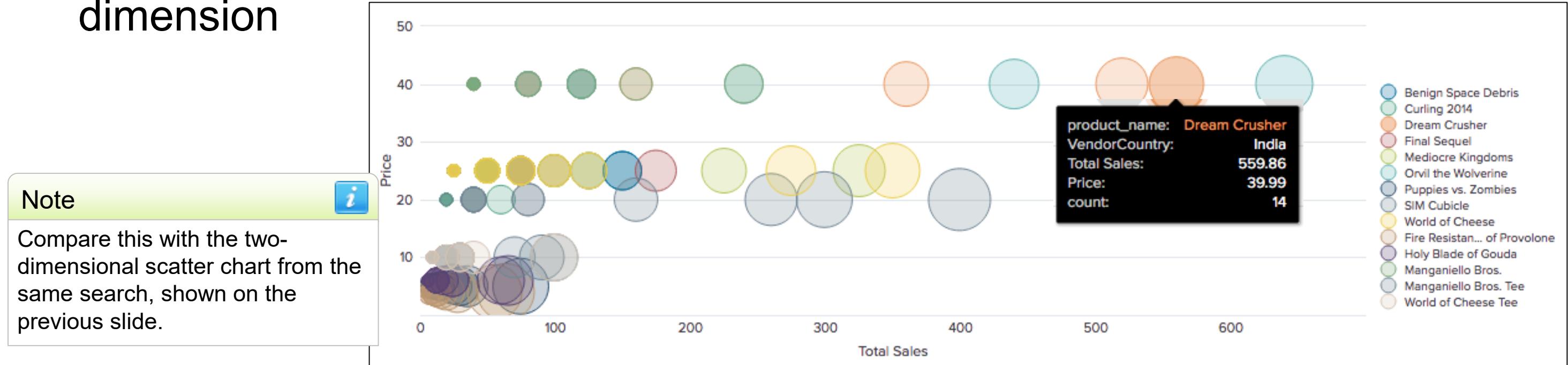


Generated for Anh Hoang (hoang_anhkim@bah.com) (C) Splunk Inc, not for distribution

Charts – Bubble

- Bubble chart provides a visual way to view a three dimensional series
- Each bubble plots against two dimensions on the X and Y axes
- The **size** of the bubble represents the value for the **third** dimension

```
index=sales sourcetype=vendor_sales  
VendorID >=7000 AND VendorID <=8999  
| stats sum(price) as "Total Sales",  
values(price) as "Price",  
count by VendorCountry, product_name
```



Generated for Anh Hoang (hoang_anhkim@bah.com) (C) Splunk Inc, not for distribution

chart Command

- chart command can display any data series plotted across one or two dimensions
- You decide which field to plot on the x-axis
 - The function defines the value of the y-axis, therefore it should be numeric
 - The first field after the over clause is the x-axis
 - Using the over and by clauses divides the data into sub-groupings
 - The values from the by clause display in the legend

chart avg(bytes) over host

The host values display over the x-axis

chart avg(bytes) over host by product_name

The host field is the x-axis and the series is further split by product_name

Generated for Anh Hoang (hoang_anhkim@bah.com) (C) Splunk Inc, not for distribution

chart Command – over field

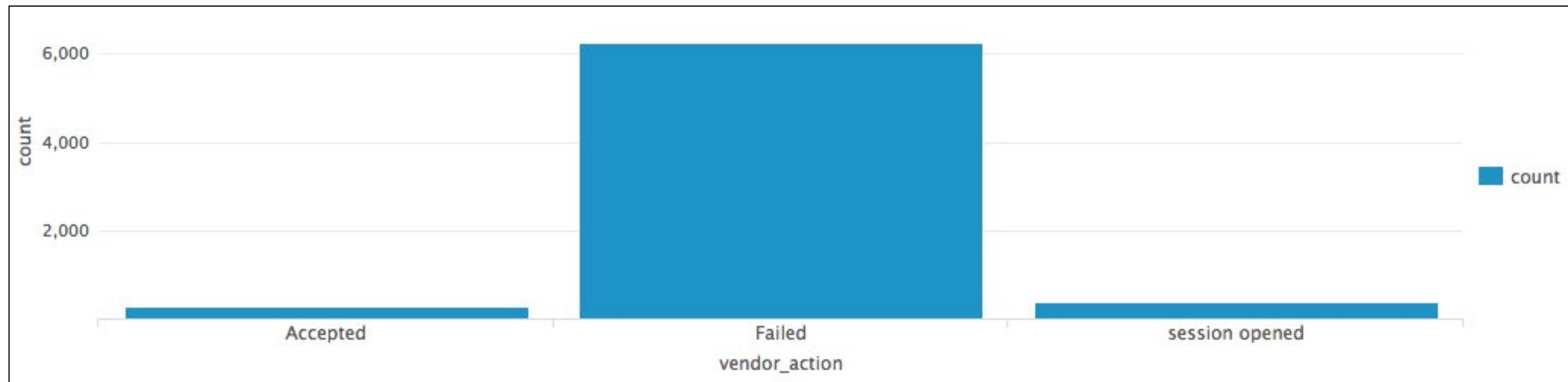
count function tallies the number of events for each value in the result set

Scenario



Display a count of vendor actions over the last 60 minutes.

```
index=security sourcetype=linux_secure  
| chart count over vendor_action
```



Generated for Anh Hoang (hoang_anhkim@bah.com) (C) Splunk Inc, not for distribution

chart Command – over *field* by *field*

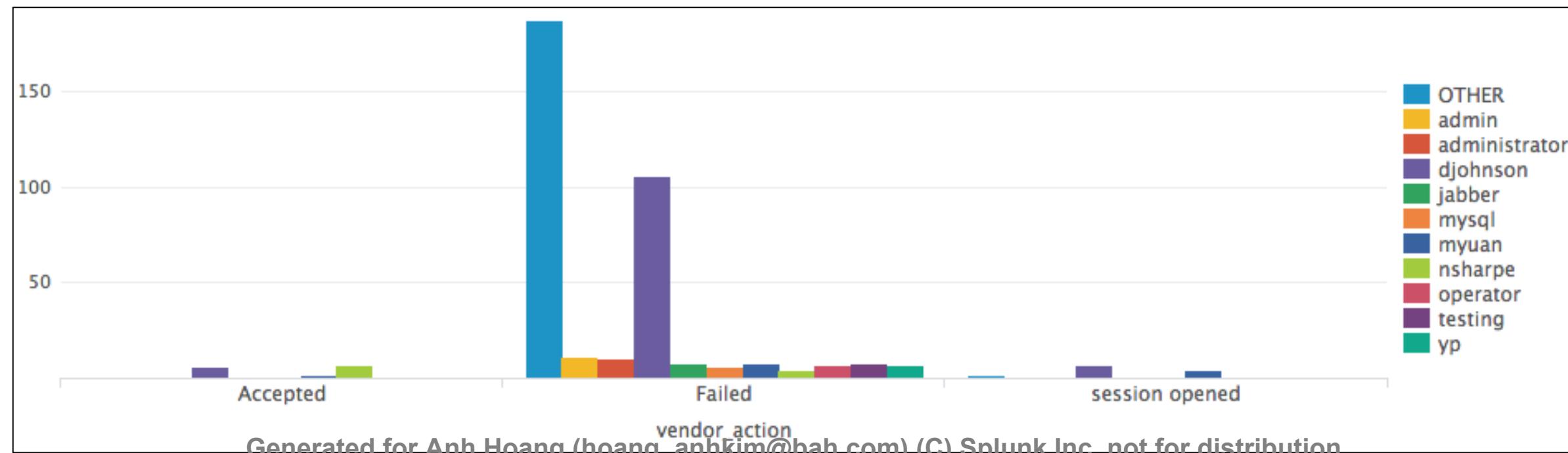
- You can use the **by** clause with the **over** clause to split results (over vendor_action by user)
- Alternatively, you can just use two **by** clauses (by vendor_action, user)
- You can only split chart results over TWO dimensions (unlike stats results)

Scenario



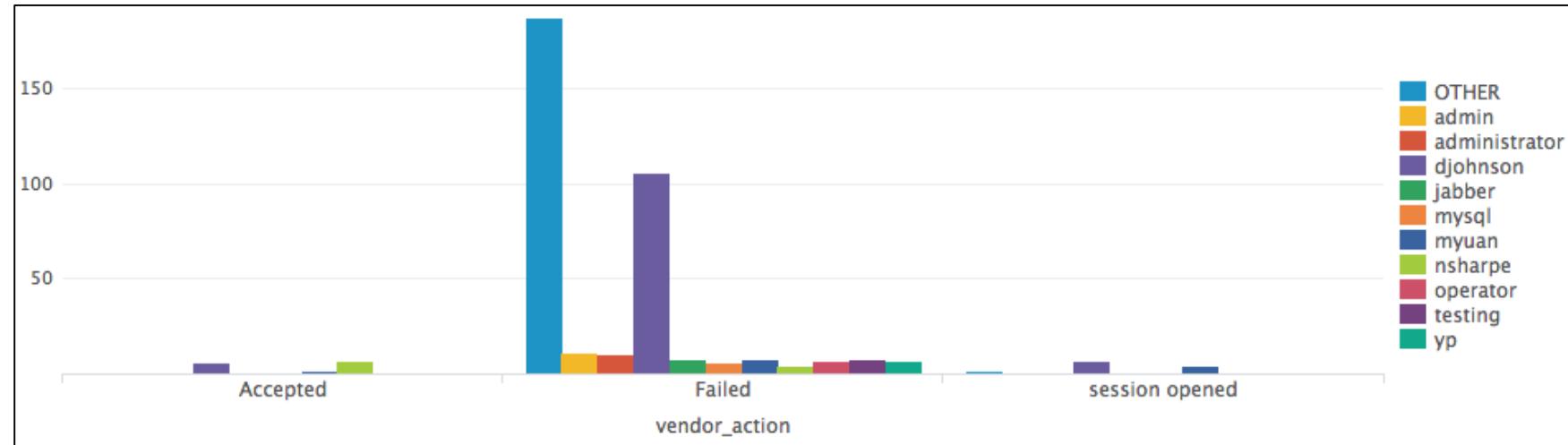
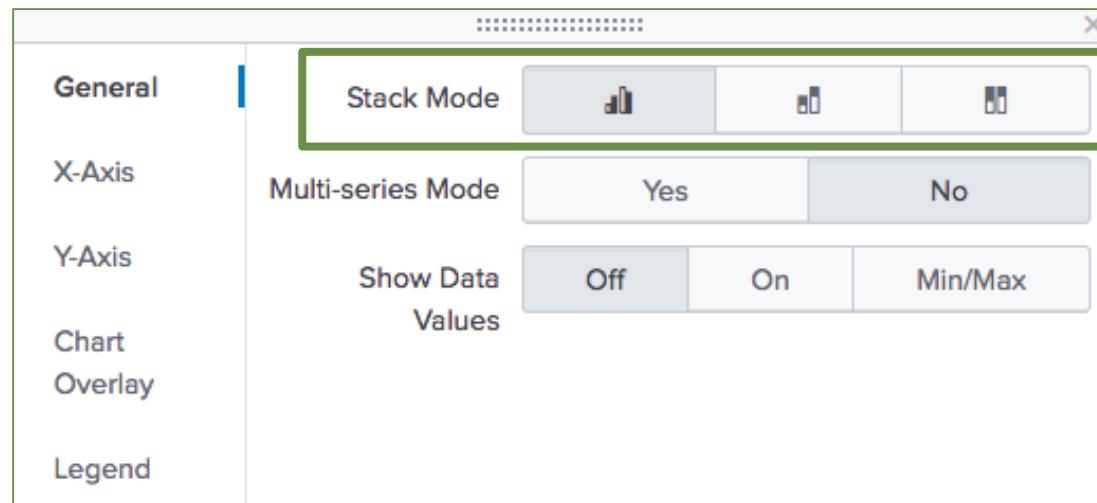
Display a count of vendor actions **by user** over the last 60 minutes.

```
index=security sourcetype=linux_secure  
| chart count over vendor_action by user
```

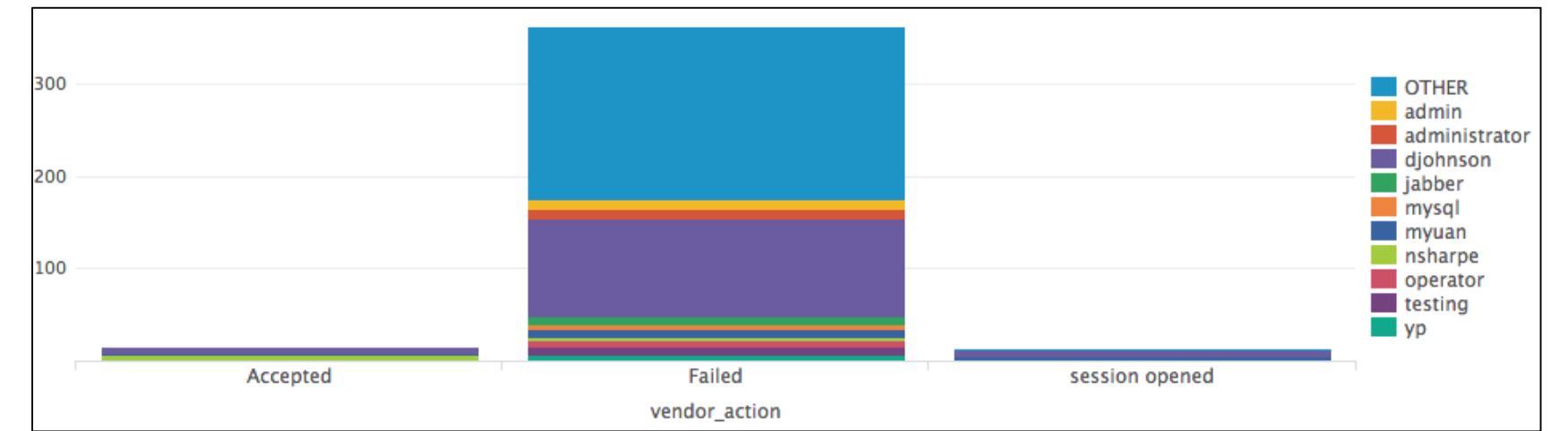


Stack Mode

Stack Mode OFF



Stack Mode ON



Generated for Anh Hoang (hoang_anhkim@bah.com) (C) Splunk Inc, not for distribution

Including NULL and OTHER Values

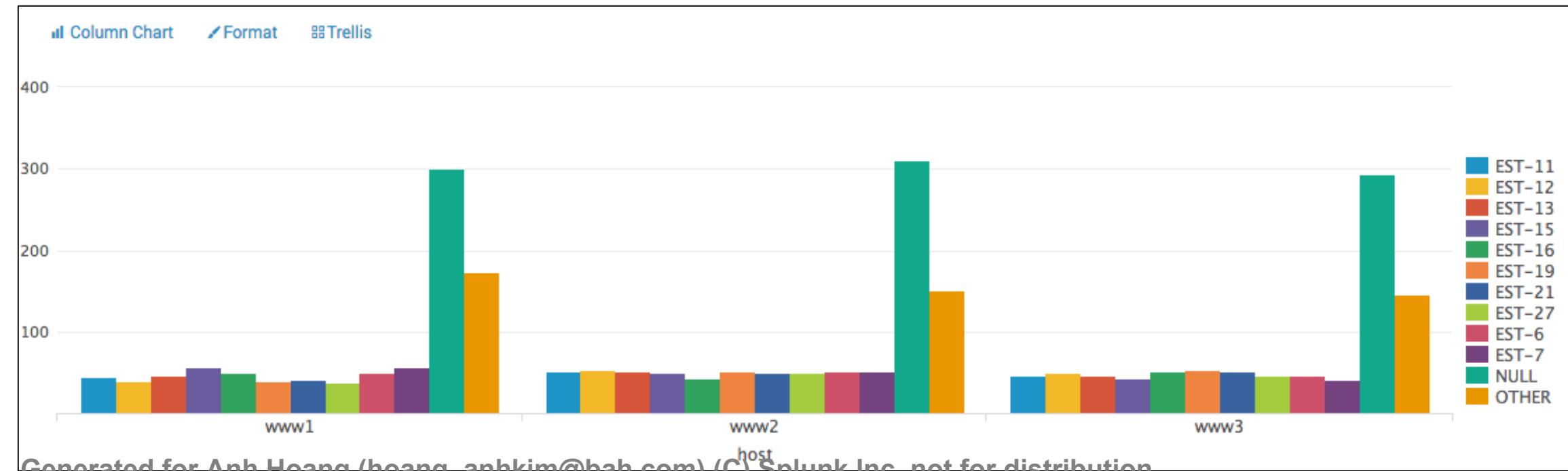
- chart and timechart commands automatically filter results to include the ten highest values
 - Surplus values are grouped into OTHER
- In this example, the results are skewed by NULL and OTHER
 - These values are shown by default

Scenario



Display a count of unsuccessful web transactions by host for each item over the last 7 days.

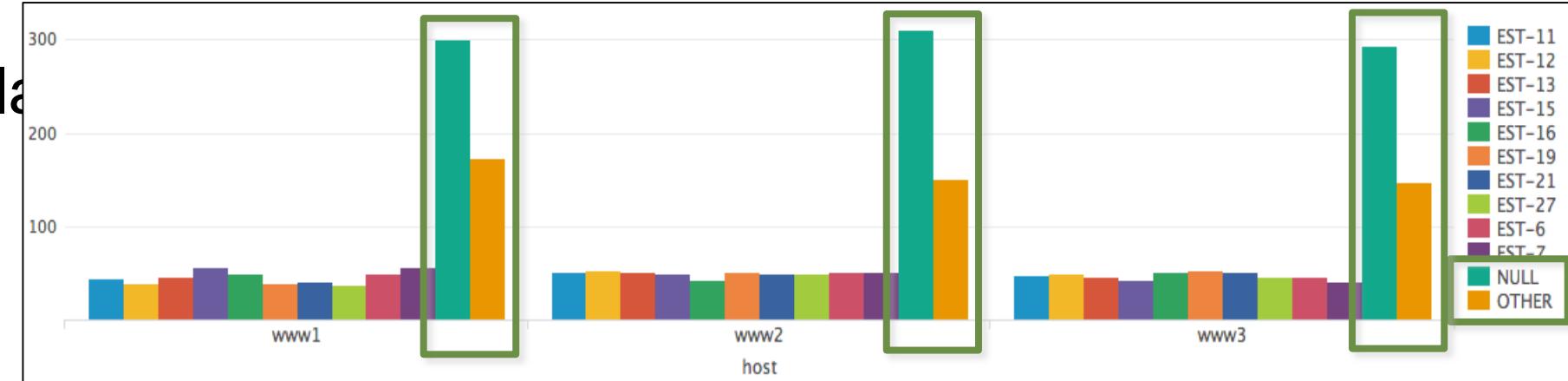
```
index=web sourcetype=access_combined  
status>399  
| chart count over host by itemId
```



Generated for Anh Hoang (hoang_anhkim@bah.com) (C) Splunk Inc, not for distribution

Omitting NULL and OTHER Values

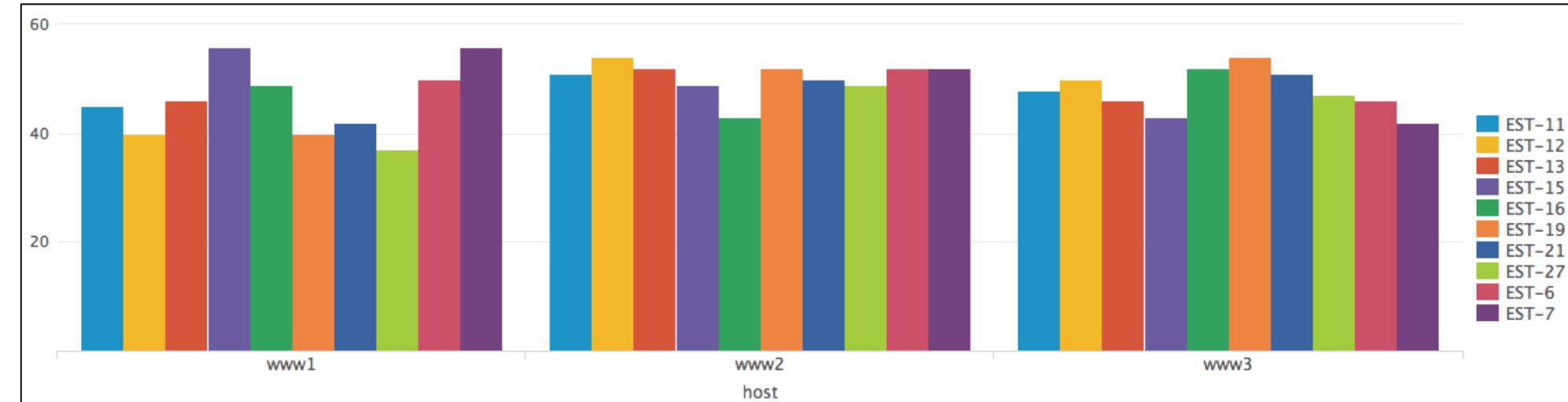
- To remove empty (NULL) and OTHER field values from the display, use these options:
 - useother=f
 - usenull=f



```
index=web sourcetype=access_combined status>399  
| chart count over host by itemId  
useother=f usenull=f
```

Note

To remove null values, add itemId=* to the base search.



Generated for Anh Hoang (hoang_anhkim@bah.com) (C) Splunk Inc, not for distribution

Limiting the Number of Values

- To adjust the number of plotted series, use the `limit` argument
- For unlimited values, use `limit=0`

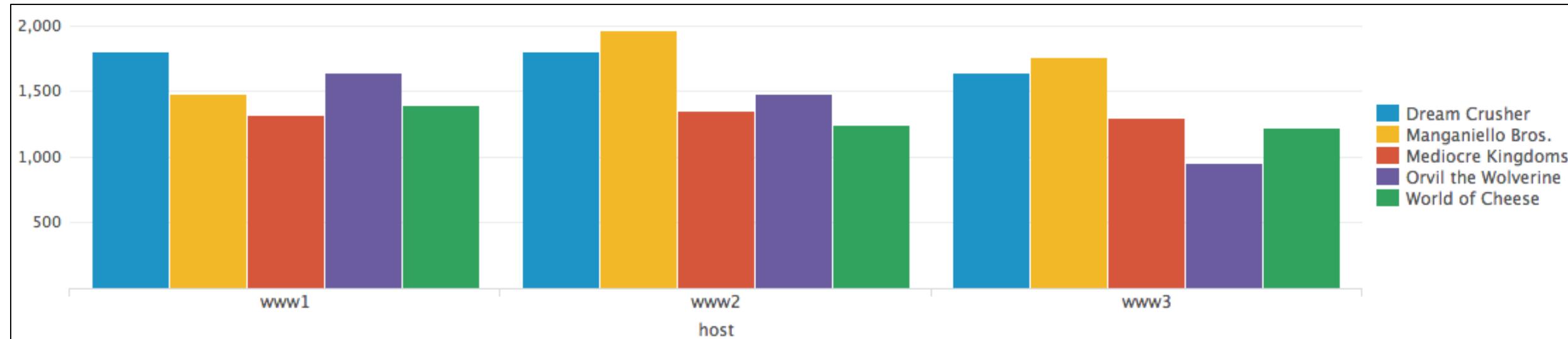
Note i

When splitting on two dimensions, the limit option applies to the second split (so in this case, `product_name`).

Scenario ?

Display sales per host for the top 5 best selling products over the last 7 days.

```
index=web sourcetype=access_combined  
action=purchase status=200  
| chart sum(price) over host  
by product_name limit=5 useother=f
```



Generated for Anh Hoang (hoang_anhkim@bah.com) (C) Splunk Inc, not for distribution

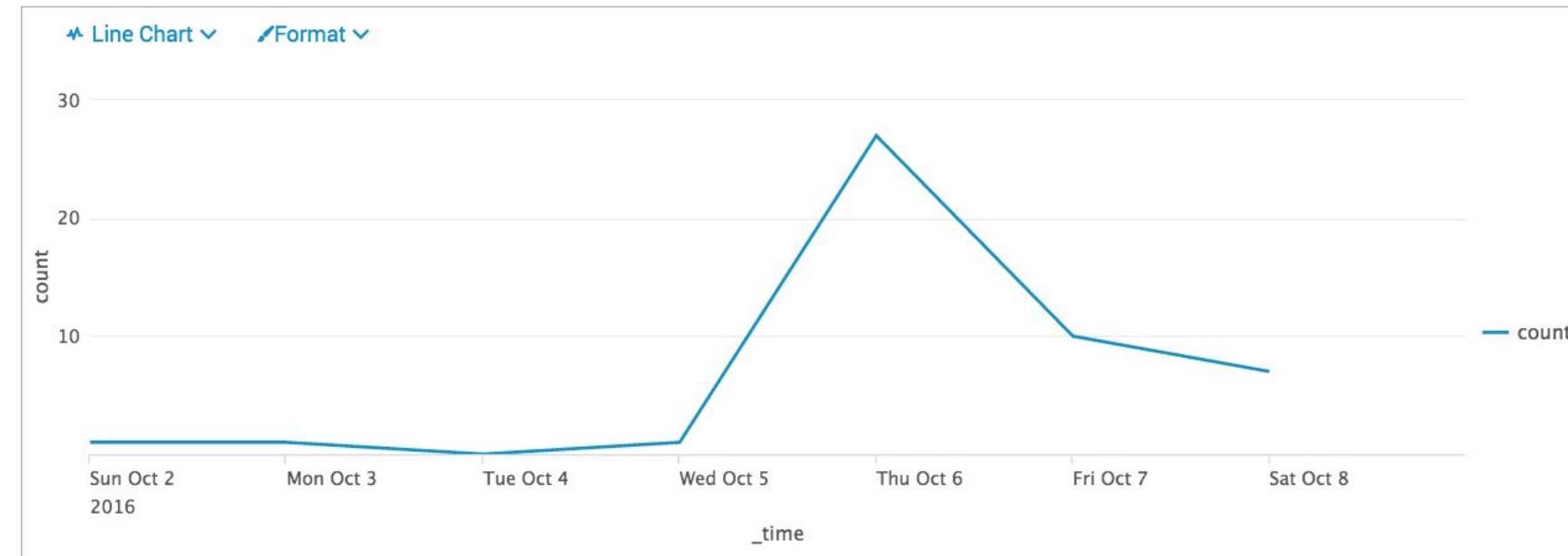
timechart Command – Overview

- timechart command performs statistical aggregations against time
- Plots and trends data over time
 - `_time` is always the x-axis
 - You can optionally split data using the `by` clause for one other field
 - Each distinct value of the split by field is a separate series in the chart
 - Timecharts are best represented as line or area charts

timechart Command – Example

Scenario ?
How many usage violations have occurred during the last 7 days?

```
index=network sourcetype=cisco_wsa_squid usage=Violation  
| timechart count
```



Note i
Functions and arguments used with stats and chart can also be used with timechart.

Generated for Anh Hoang (hoang_anhkim@bah.com) (C) Splunk Inc, not for distribution

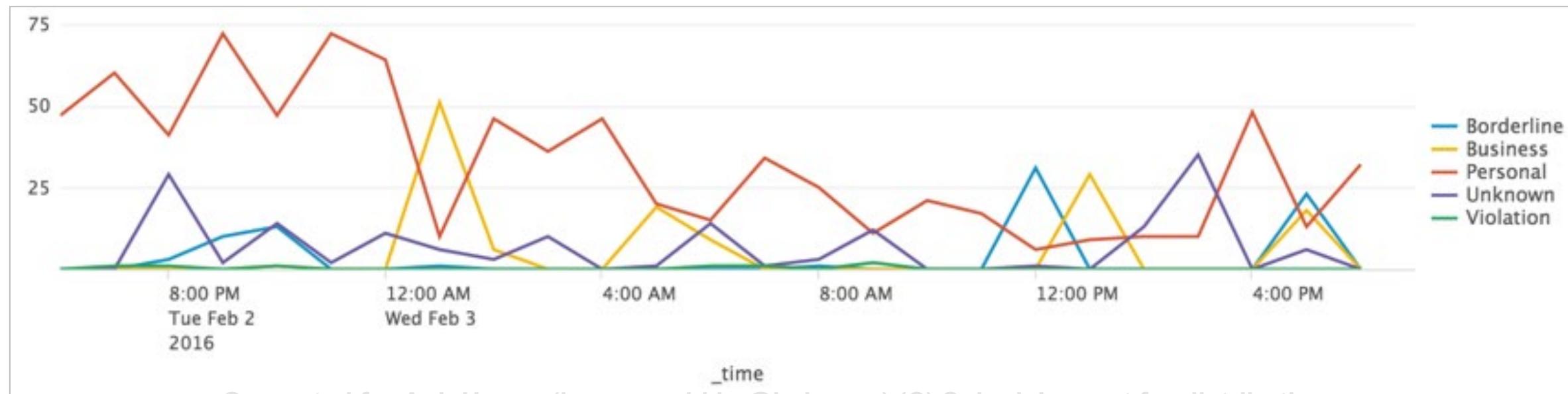
timechart Command – Multiple Values

- Splitting by the usage field, each line represents a unique field value
 - Unlike stats, only ONE field can be specified after by
- y-axis represents the count for each field value

Scenario ?
What is the overall usage trend
for the last 24 hours?

index=network sourcetype=cisco_wsa_squid
| timechart count by usage

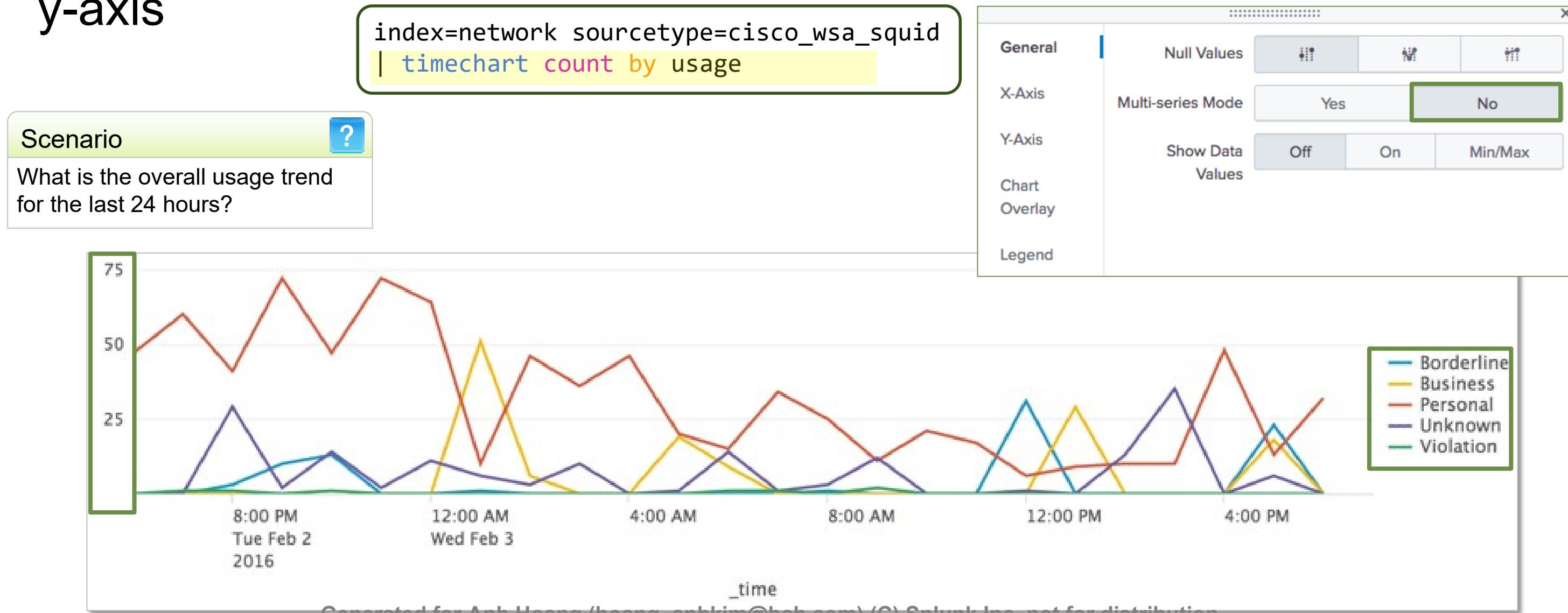
Note i
Using timechart, you can split by
a maximum of one field because
`_time` is the implied first by field.



Generated for Anh Hoang (hoang_anhkim@bali.com) (C) Splunk Inc. not for distribution

timechart Command – Multi-series: No

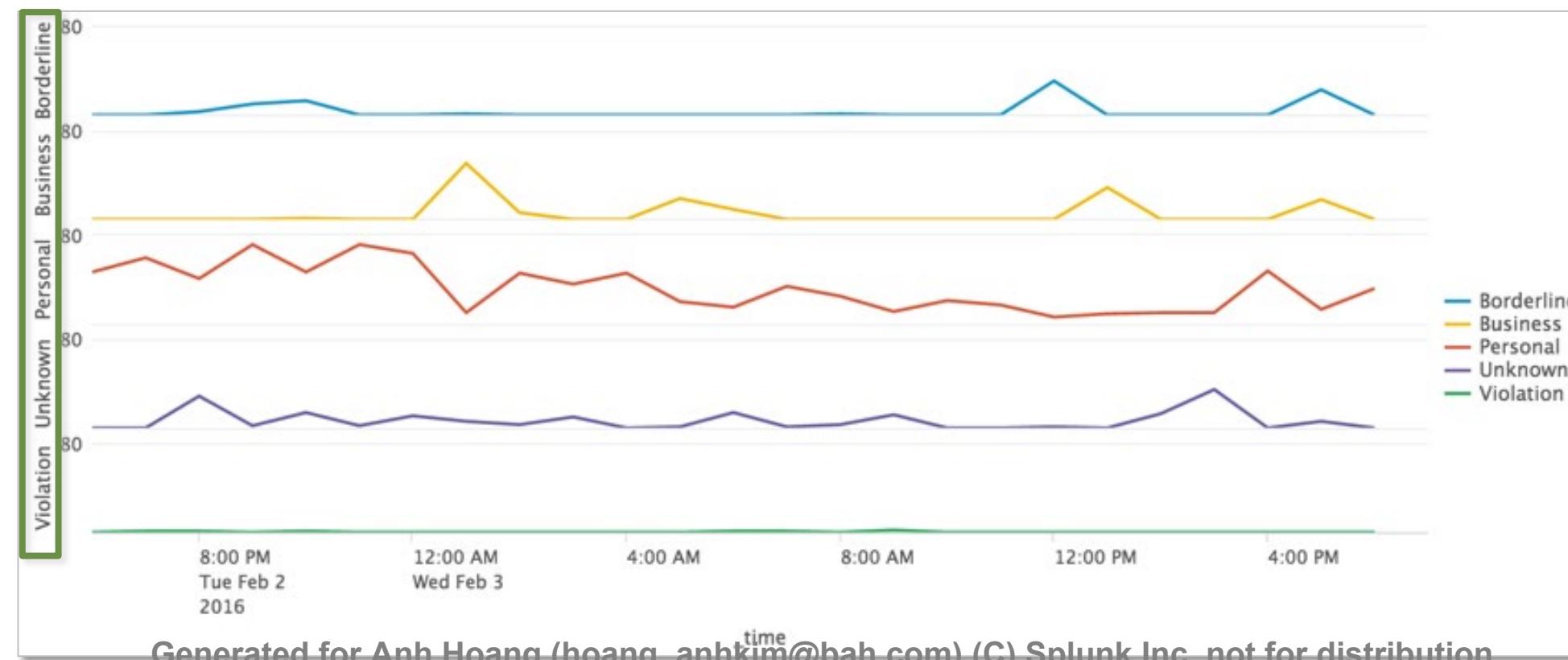
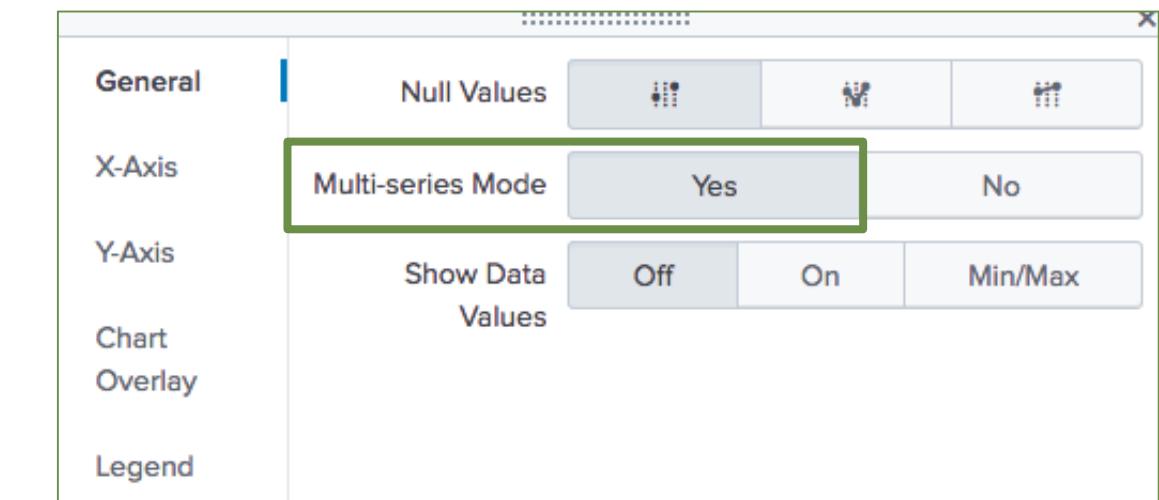
When the multi-series mode is set to **No**, all fields share the y-axis



Generated for Anh Hoang (hoang_anhkim@bah.com) (C) Splunk Inc, not for distribution

timechart Command – Multi-series: Yes

- Setting multi-series mode to **Yes** causes the y-axis to split for each field value
- y-axis is divided into sections, each spanning the same max and min count

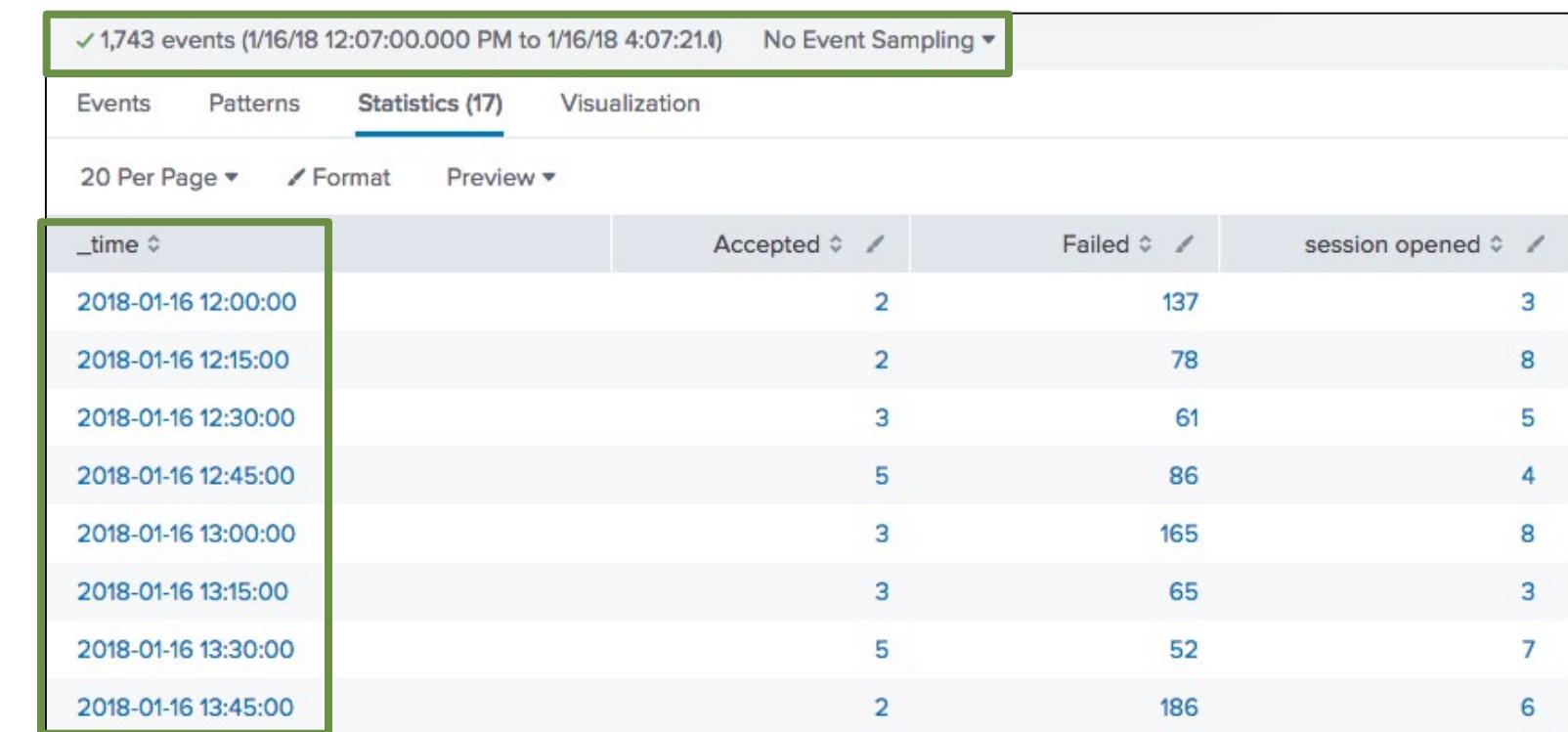


Generated for Anh Hoang (hoang_anhkim@bah.com) (C) Splunk Inc, not for distribution

timechart Command – Adjusting the Sampling Interval

- The timechart command "buckets" the values of the _time field
 - This provides dynamic sampling intervals, based upon the time range of the search
- Example defaults:
 - Last 60 minutes uses span=1m
 - Last 24 hours uses span=30m
- Adjust the interval using the span argument, e.g. span=15m

```
index=security sourcetype=linux_secure  
vendor_action=*  
| timechart span=15m count by vendor_action
```



The screenshot shows the Splunk Statistics view for 1,743 events from January 16, 2018, between 12:07:00.000 PM and 4:07:21.0. The view includes tabs for Events, Patterns, Statistics (17), and Visualization, with Statistics being the active tab. It also shows options for 20 Per Page, Format, and Preview. The main table displays the following data:

_time	Accepted	Failed	session opened
2018-01-16 12:00:00	2	137	3
2018-01-16 12:15:00	2	78	8
2018-01-16 12:30:00	3	61	5
2018-01-16 12:45:00	5	86	4
2018-01-16 13:00:00	3	165	8
2018-01-16 13:15:00	3	65	3
2018-01-16 13:30:00	5	52	7
2018-01-16 13:45:00	2	186	6

timechart Command – Statistical Functions

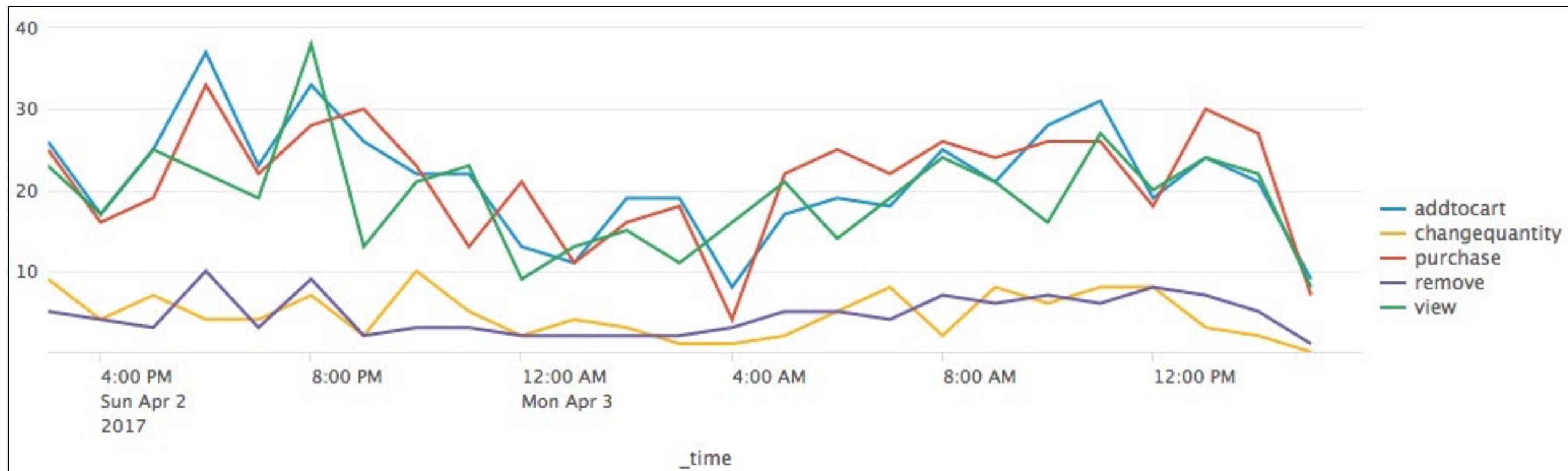
As with the stats and chart commands, you can apply statistical functions to the timechart command

Scenario



How much web activity of each type took place during the last 24 hours?

```
index=web sourcetype=access_combined action=*
| timechart span=1h count by action
```



Generated for Anh Hoang (hoang_anhkim@bah.com) (C) Splunk Inc, not for distribution

span Function with the chart Command

If x axis is numeric, span can be used with chart to group events on the x axis into buckets

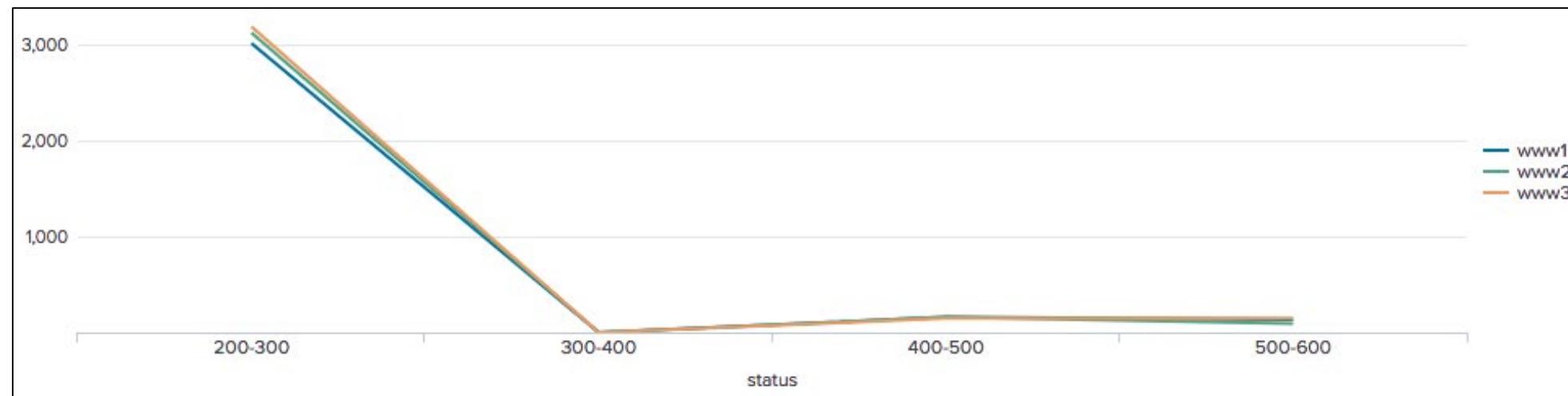
```
index=web sourcetype=access_combined  
| chart span=100 count over status by host
```

Scenario



The IT manager wants to see the volume of web activity over the past 24 hours by ranges of status codes.

status	www1	www2	www3
200-300	3014	3124	3188
300-400	0	0	0
400-500	161	161	149
500-600	136	93	148



Generated for Anh Hoang (hoang_anhkim@bah.com) (C) Splunk Inc, not for distribution

Formatting – Visualization

General Null Values

X-Axis Multi-series Mode Yes

Y-Axis Show Data Values

Chart Overlay

Legend

General Title Where Sold

X-Axis Label Rotation abc

Y-Axis Label Truncation Yes

Chart Overlay

Legend

```
index=web sourcetype=access_combined  
| chart count by host
```

General Title Products Sold

X-Axis Scale

Y-Axis Interval optional

Chart Overlay Min Value optional

Legend Max Value optional

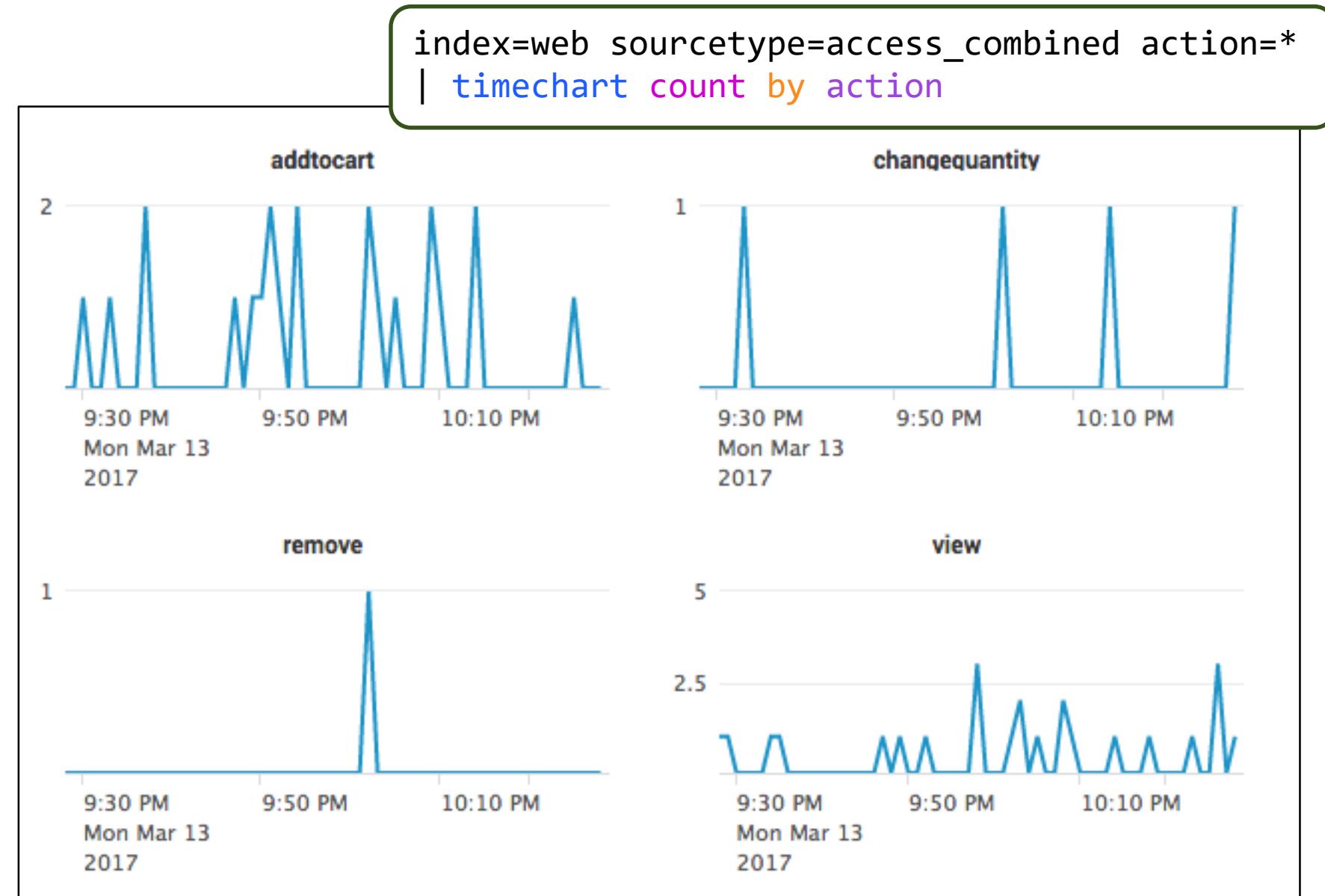
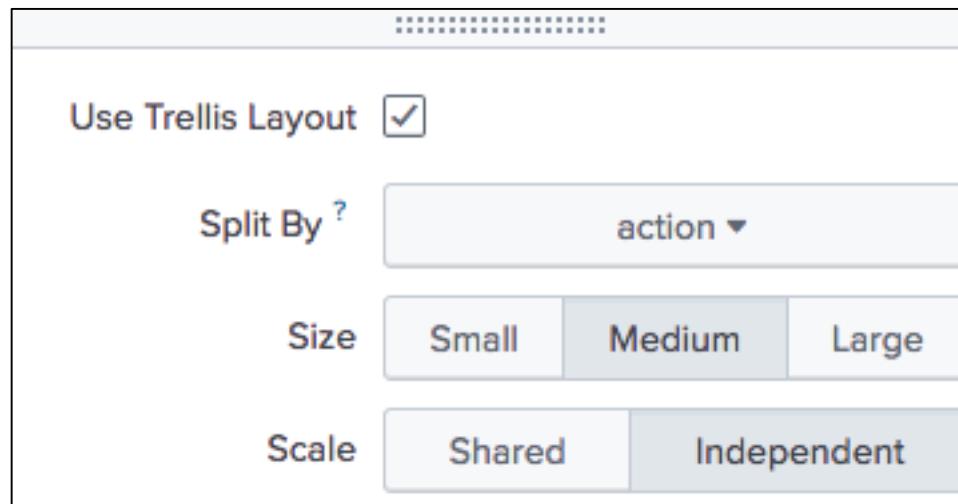
Number Abbreviations



Generated for Anh Hoang (hoang_anhkim@bah.com) (C) Splunk Inc, not for distribution

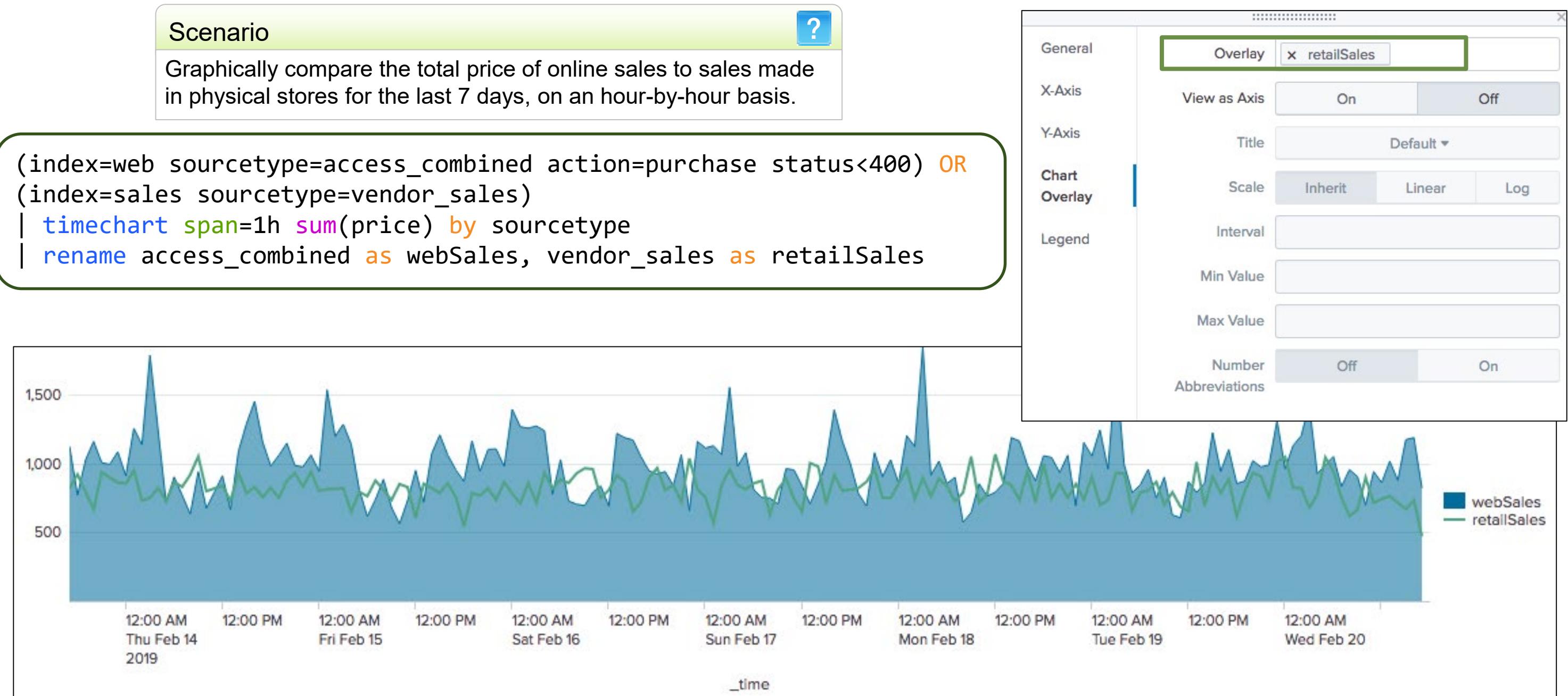
Formatting – Trellis Layout

- Display multiple charts based on one result set
- Allows visual comparison between different categories
- Data only fetched once



Generated for Anh Hoang (hoang_anhkim@bah.com) (C) Splunk Inc, not for distribution

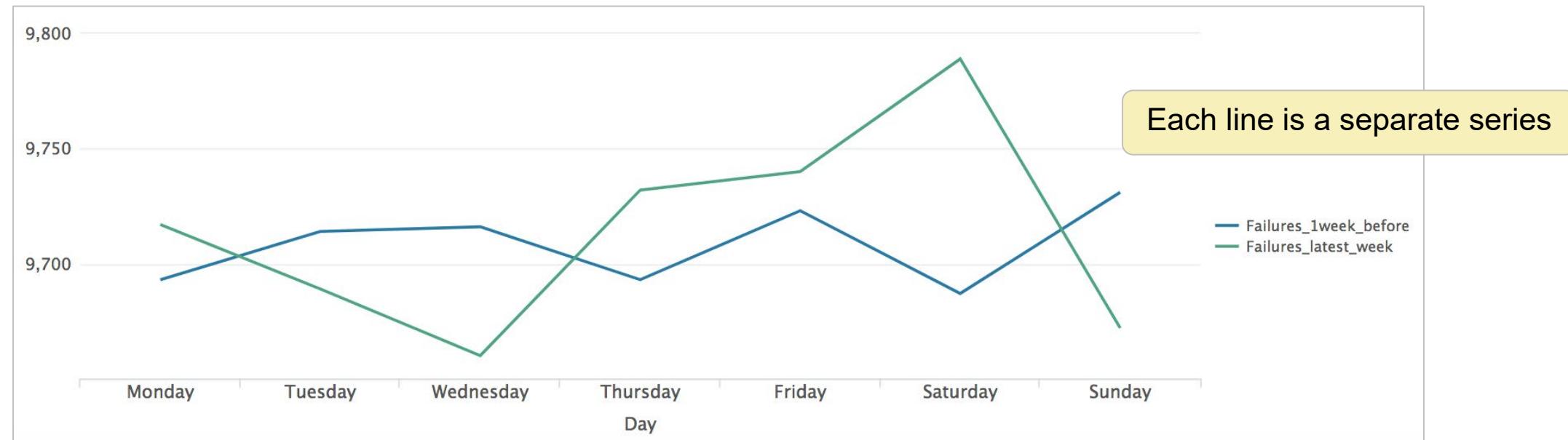
Formatting – Chart Overlay



Generated for Anh Hoang (hoang_anhkim@bah.com) (C) Splunk Inc, not for distribution

timewrap Command

- Displays the output of the timechart command, so that each time period is a separate series
- Can compare data over a specific time period, such as day-over-day or month-over-month



Generated for Anh Hoang (hoang_anhkim@bah.com) (C) Splunk Inc, not for distribution

timewrap Syntax and Example

- Syntax: `timewrap timewrap-span`
- *timewrap-span* can be second, minute, hour, day, week, month, quarter or year
- For example: `timewrap 1w`

Displaying Data Week Over Week

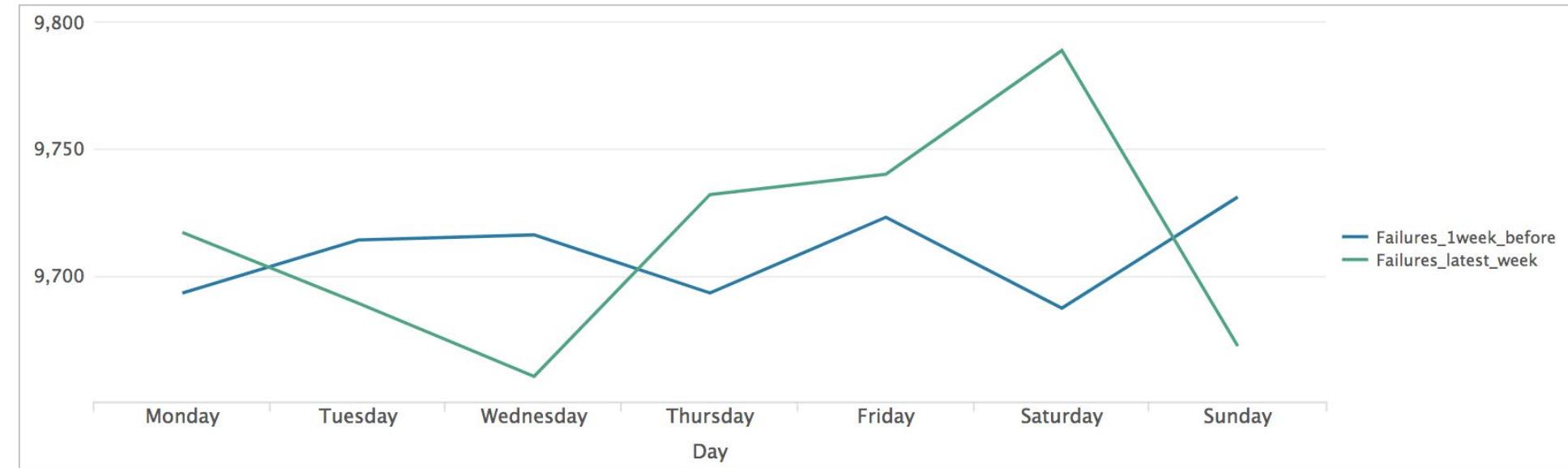
Scenario



Compare the number of password failures over the last week to password failures over the previous week.

- Earliest to latest spans 14 days – i.e., 2 weeks
- Specifying 1w with timewrap and using the line chart visualization causes two lines to be displayed

```
index=security "failed password"
earliest=-14d@d latest=@d
| timechart span=1d count as Failures
| timewrap 1w
| rename _time as Day
| eval Day = strftime(Day, "%A")
```



Note



The eval command is discussed in detail in Module 4.

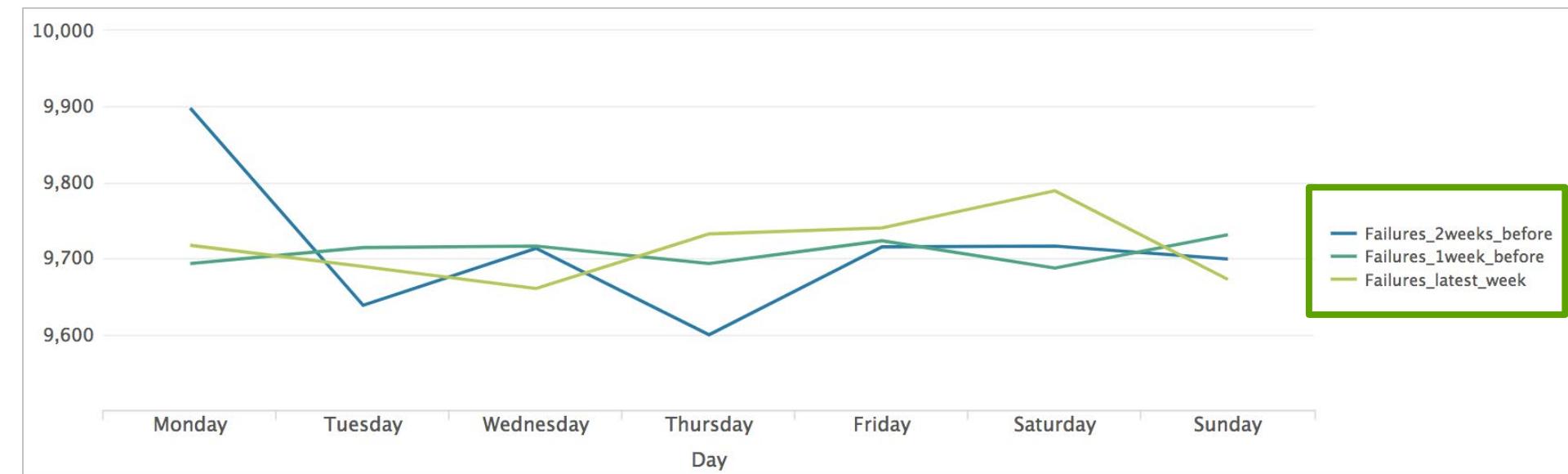
Displaying Data Week Over Week (cont.)

Scenario ?

Compare the number of password failures over the last *three* weeks.

```
index=security "failed password"  
earliest=-21d@d latest=@d  
| timechart span=1d count as Failures  
| timewrap 1w  
| rename _time as Day  
| eval Day = strftime(Day, "%A")
```

- Add additional lines to the chart by adding additional periods to the search
- For example, adding an extra week to the search that was shown earlier adds a line to the line chart



Transforming Command Summary

Feature	stats	chart	timechart
Multi-level breakdown [by clause]	Many	2	1
Limit # series shown	NA	<code>limit=n</code> <i>Default=10</i>	<code>limit=n</code> <i>Default=10</i>
Filter other series	NA	<code>useother=f</code>	<code>useother=f</code>
Filter null values	NA	<code>usenull=f</code>	<code>usenull=f</code>
Set value groups along the x-axis	NA	<code>span</code>	<code>span</code>

Generated for Anh Hoang (hoang_anhkim@bah.com) (C) Splunk Inc, not for distribution

Transforming Command Summary (cont.)

To count the frequency of a field(s), use top/rare

```
index=security sourcetype=linux_secure  
| top src_ip, user, vendor_action, app
```

src_ip	user	vendor_action	app	count	percent
235.166.61.39	nsharpe	Failed	sshd	122	7.850708
44.248.239.252	myuan	Failed	sshd	113	7.271557
16.201.248.137	nsharpe	Failed	sshd	113	7.271557
43.189.188.26	myuan	Failed	sshd	107	6.885457

```
index=security sourcetype=linux_secure  
| rare src_ip, user, vendor_action, app
```

src_ip	user	vendor_action	app	count	percent
107.3.146.207	admin	Failed	sshd	1	0.064185
107.3.146.207	desktop	Failed	sshd	1	0.064185
107.3.146.207	helpdesk	Failed	sshd	1	0.064185
107.3.146.207	inet	Failed	sshd	1	0.064185

Generated for Anh Hoang (hoang_anhkim@bah.com) (C) Splunk Inc, not for distribution

Transforming Command Summary (cont.)

Use stats to calculate statistics for two or more by fields (non time-based)

```
index=security sourcetype=linux_secure  
| stats count by src_ip, user, vendor_action, app
```

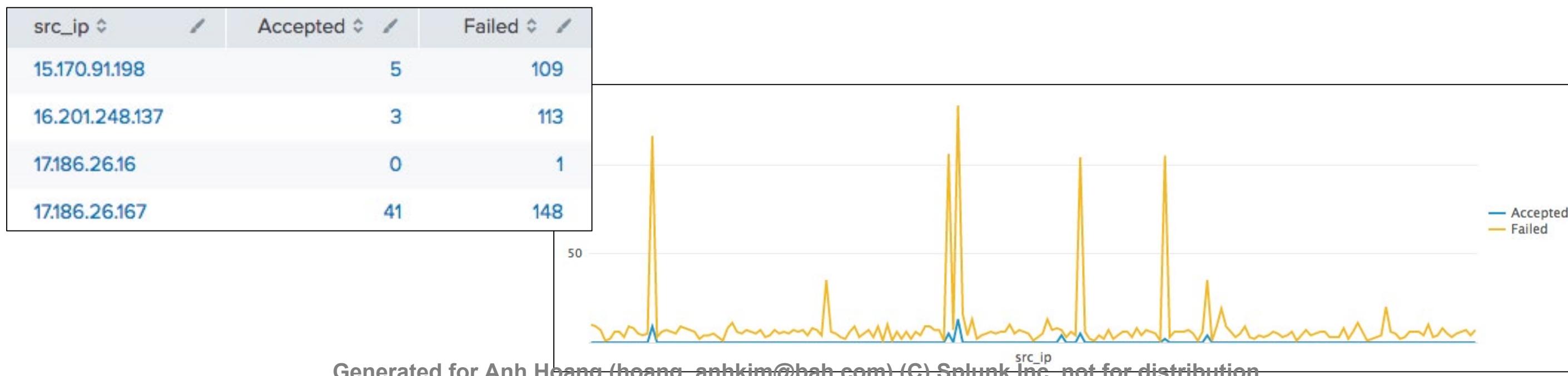
src_ip	user	vendor_action	app	count
105.131.160.148	djohnson	Failed	sshd	100
107.3.146.207	admin	Failed	sshd	2
107.3.146.207	administrator	Failed	sshd	3
107.3.146.207	angel	Failed	sshd	1
107.3.146.207	bin	Failed	sshd	1
107.3.146.207	couchdb	Failed	sshd	1
107.3.146.207	customer	Failed	sshd	1
107.3.146.207	db	Failed	sshd	1
107.3.146.207	desktop	Failed	sshd	4
107.3.146.207	divine	Failed	sshd	1

Generated for Anh Hoang (hoang_anhkim@bah.com) (C) Splunk Inc, not for distribution

Transforming Command Summary (cont.)

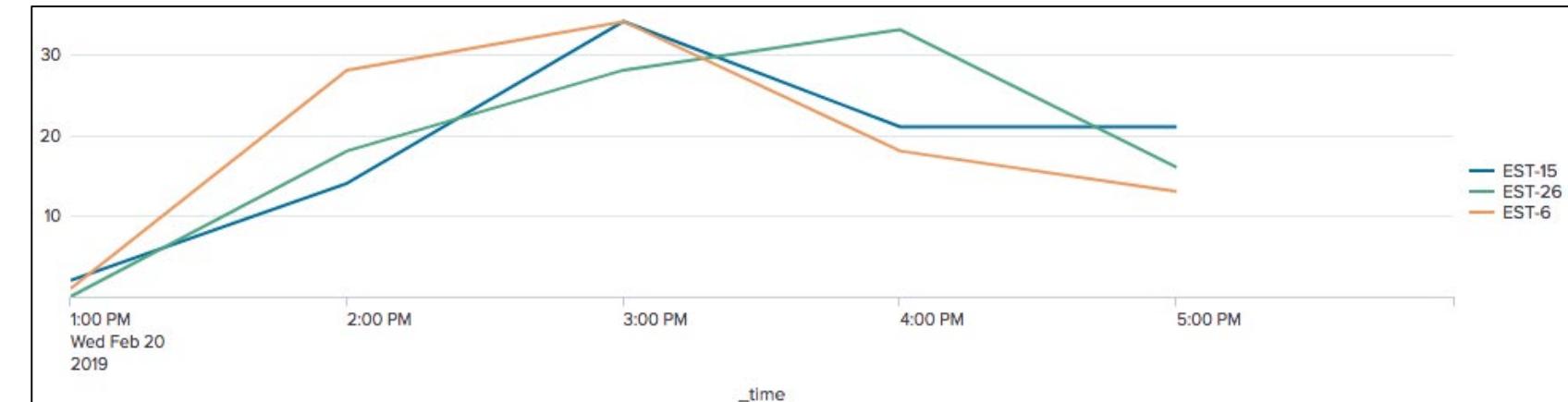
- To calculate statistics with an arbitrary field as the x-axis (not `_time`), use `chart`
 - When you use a `by` field, the output is a table
 - Each column represents a distinct value of the split-by field

```
index=security sourcetype=linux_secure  
| chart count over src_ip  
by vendor_action
```



Transforming Command Summary (cont.)

- Use timechart to calculate statistics with `_time` as the x-axis
- If a by field is used, the output is a table
- Each column represents a distinct value of the split-by field
 - | timechart span=1h count by itemId limit=3 useother=f usenull=f



_time	EST-15	EST-26	EST-6
2019-02-20 13:00	2	0	1
2019-02-20 14:00	14	18	28
2019-02-20 15:00	34	28	34
2019-02-20 16:00	21	33	18
2019-02-20 17:00	21	16	13

Generated for Anh Hoang (hoang_anhkim@bah.com) (C) Splunk Inc, not for distribution

Lab Exercise 2

Time: 45 – 50 minutes

Tasks:

- Report the failures on the network during the last 24 hours and add it to a new security dashboard as a column chart
- Chart the five best selling products in North America by country during the last 7 days
- Display internet usage in a timechart during the last 24 hours

****Challenge Exercise:**

- Display and compare online and vendor sales during the last 24 hours

Module 3: Using Trendlines, Mapping, and Single Value Commands

Generated for Anh Hoang (hoang_anhkim@bah.com) (C) Splunk Inc, not for distribution

Module Objectives

- Create a trendline
- Create maps
 - iplocation
 - geostats
 - geom
- Create and format single values
- Use the addtotals command

trendline Command

- Allows you to overlay a computed moving average on a chart

- trendline computes the moving averages of a field

`trendline <trendtype><period>(field) [AS newfield]`

- *trendtype*:

- sma - simple moving average

- ema - exponential moving average

- wma - weighted moving average

trendline Command (cont.)

- Must define the *period* over which to compute the trend
- *period* must be an integer between 2 and 10000
 - For example, `sma2(sales)` is valid
 - But `sma(sales)` would *fail* as it is missing an integer, the defining period

Scenario ?

Display total sales and sales trends over the past 24 hours.

New Search Save As ▾ Close

```
index=web sourcetype=access_combined action=purchase status=200
| timechart span=2h sum(price) as sales
| trendline sma(sales) as trend
```

Last 24 hours ?

! Error in 'trendline' command: command="trendline", Invalid trend period for argument 'sma(sales)'

Note ?

Autocomplete displays functions in purple. Here however, since it does not recognize sma as a function, it is shown in black.

Generated for Anh Hoang (hoang_anhkim@bah.com) (C) Splunk Inc, not for distribution

trendline Command – Example

Scenario

Display total sales and sales trends over the past 24 hours.



```
index=web sourcetype=access_combined action=purchase status=200  
| timechart span=2h sum(price) as sales  
| trendline sma2(sales) as trend
```

General Overlay trend

X-Axis View as Axis On Off

Y-Axis Title Default ▾

Chart Overlay Scale Inherit Linear Log

Legend Interval

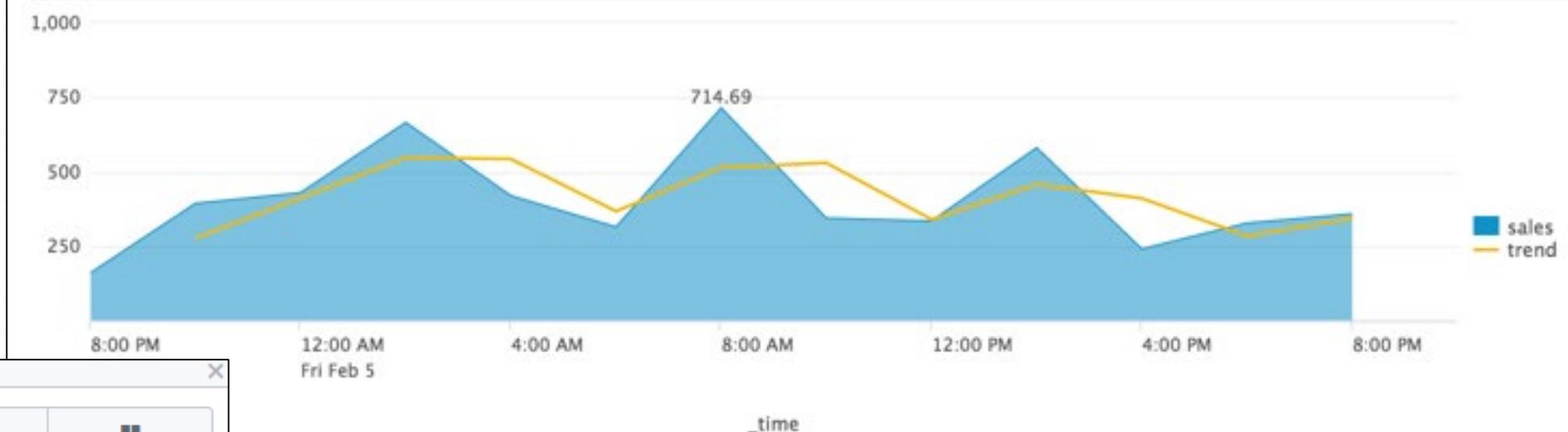
General Stack Mode

X-Axis Null Values

Y-Axis Multi-series Mode Yes No

Chart Overlay Show Data Values Off On Min/Max

Legend



Generated for Anh Hoang (hoang_anhkim@bah.com) (C) Splunk Inc, not for distribution

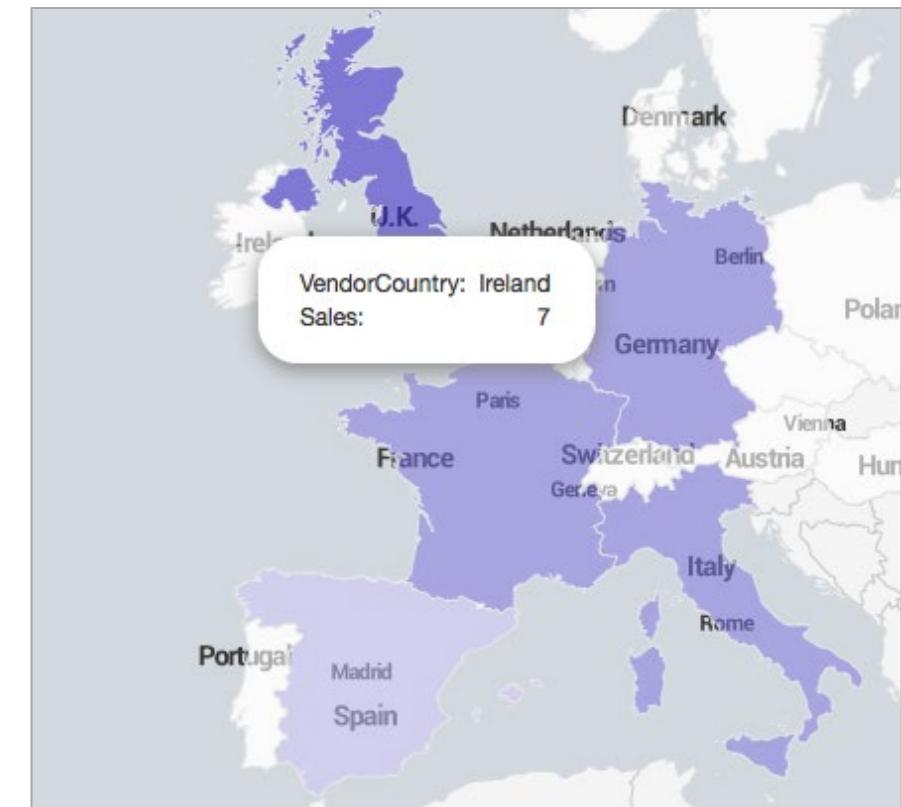
Viewing Results as a Map

There are two map types

Cluster Map



Choropleth Map



Generated for Anh Hoang (hoang_anhkim@bah.com) (C) Splunk Inc, not for distribution

iplocation Command

Scenario

Discover longitude and latitude data for src_ip for the last 60 minutes.



```
index=security sourcetype=linux_secure (fail* OR invalid)
| iplocation src_ip
```

- Use iplocation to look up and add location information to an event
 - This information includes city, country, region, latitude and longitude
- Not all of the information is available for all ip address ranges
- Automatically defines the default lat and lon fields required by geostats

INTERESTING FIELDS

a action 1

a app 1

a City 74

a Country 37

date_hour 2

lat 97

linecount 1

lon 98

pid 100+

a process 1

a punct 5

a Region 62

Note



This works with both IPv4 and IPv6.

geostats Command

- Use geostats to compute statistical functions and render a cluster map

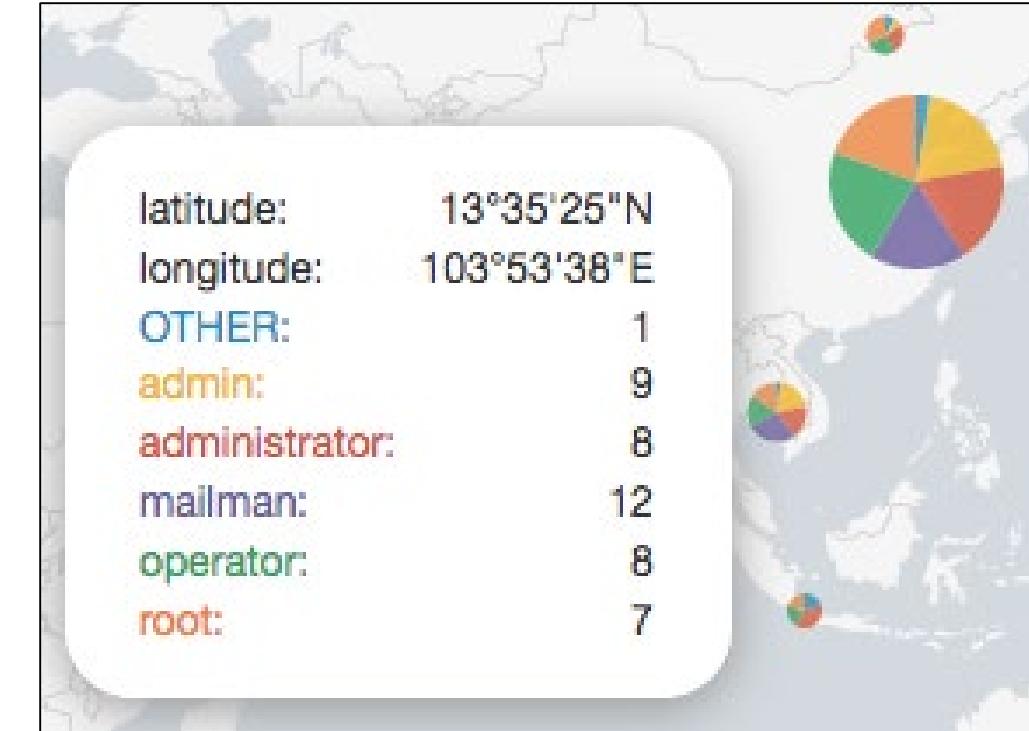
```
geostats [latfield=string] [longfield=string] [stats-agg-term]* [by-clause]
```

- Data must include latitude and longitude values
- Define the *Latfield* and *Longfield* only if they differ from the default lat and lon fields
- To control the column count, use the *globallimit* argument

geostats Command – Example

Scenario ?
Map the users of failed actions on the network worldwide during the last 24 hours.

```
index=security sourcetype=linux_secure  
(fail* OR invalid)  
| iplocation src_ip  
| geostats globallimit=5 count by user
```

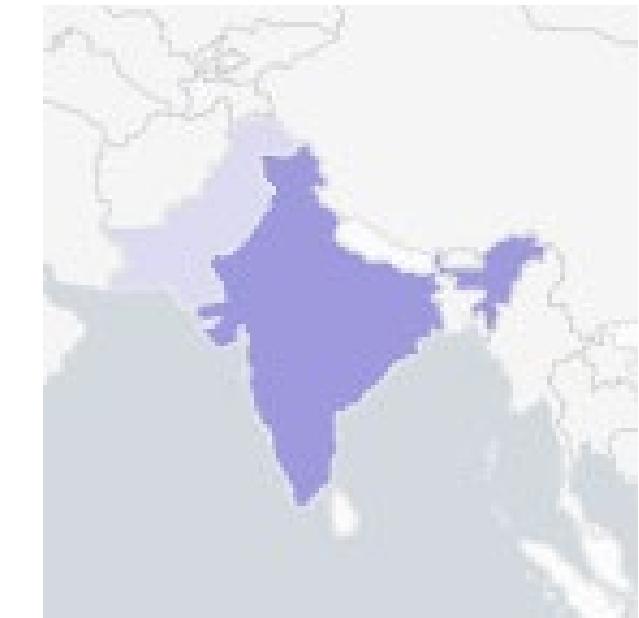


geobin	latitude	longitude	OTHER	admin	administrator	mailman	operator	root
bin_id_zl_0_y_2_x_2	-25.96763	-53.70309	3	4	5	4	2	
bin_id_zl_0_y_2_x_4	-29.00000	24.00000	1		1	3	3	1
bin_id_zl_0_y_2_x_7	-33.22925	151.62210	1	3	2		1	7
bin_id_zl_0_y_3_x_3	-8.05000	-34.90000	110					

Generated for Anh Hoang (hoang_anhkim@bah.com) (C) Splunk Inc, not for distribution

Choropleth Map

- Uses shading to show relative metrics, such as sales, network intruders, etc. for predefined geographic regions
- To define regional boundaries, you must have either a:
 - KML (Keyhole Markup Language) file
 - KMZ (compressed Keyhole Markup Language) file
- Splunk ships with:
 - geo_us_states, United States
 - geo_countries, countries of the world



... | geom [*featureCollection*] [*featureIdField*=*string*]

Note

For more information, see Appendix B: Creating New Choropleth Maps

Generated for Anh Hoang (hoang_anhkim@bah.com) (C) Splunk Inc, not for distribution

geom Command



Scenario



Display the previous week's retail sales in EMEA.

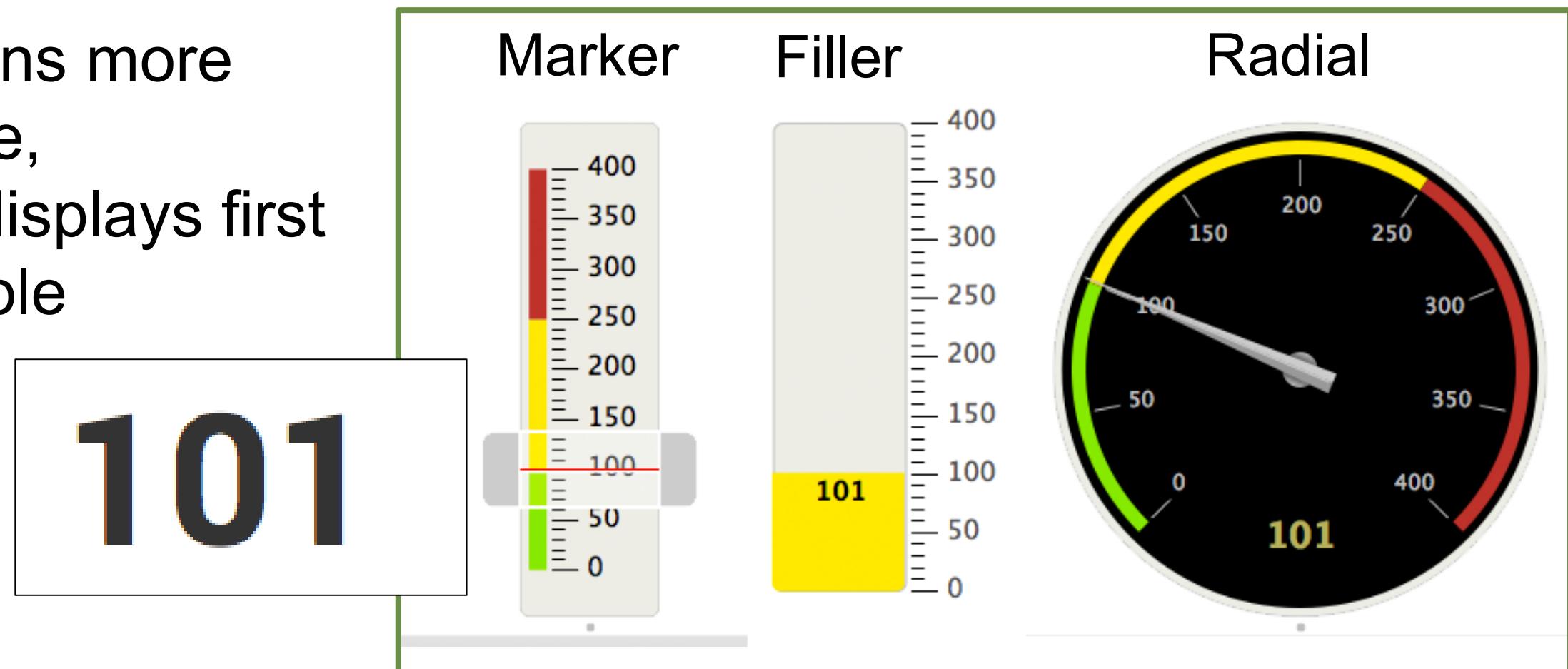
```
index=sales sourcetype=vendor_sales  
VendorID > 4999 AND VendorID < 6000  
| stats count as Sales by VendorCountry  
| geom geo_countries featureIdField=VendorCountry
```

0 - 25
25 - 50
50 - 75
75 - 100
100 - 125

Viewing Results as a Single Value

- Single value visualizations provide various formatting options
- If search returns more than one value, visualization displays first cell in data table

```
index=security sourcetype=linux_secure vendor_action=failed  
| stats count
```



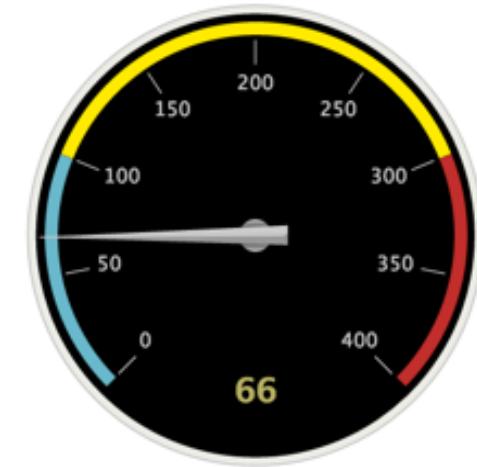
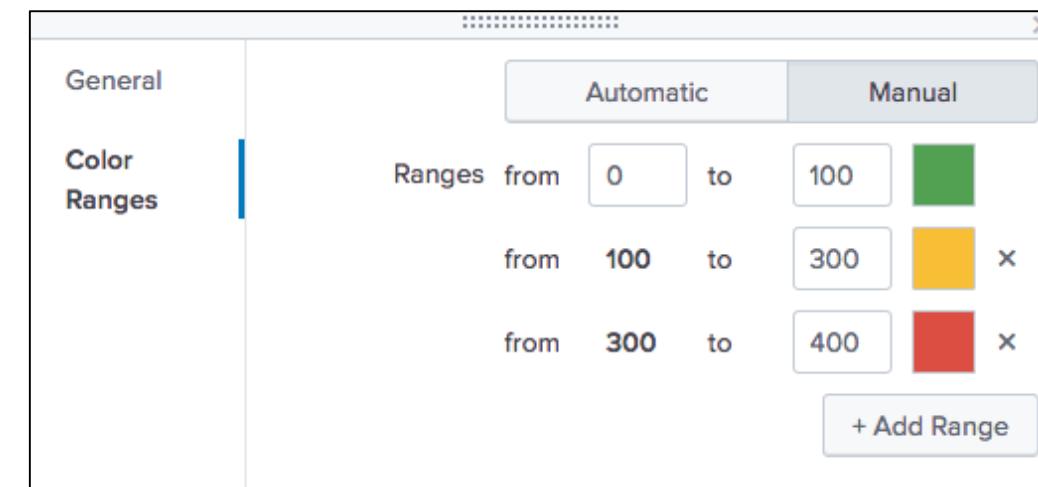
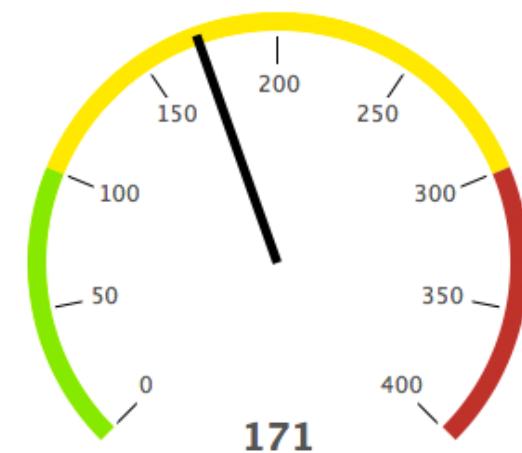
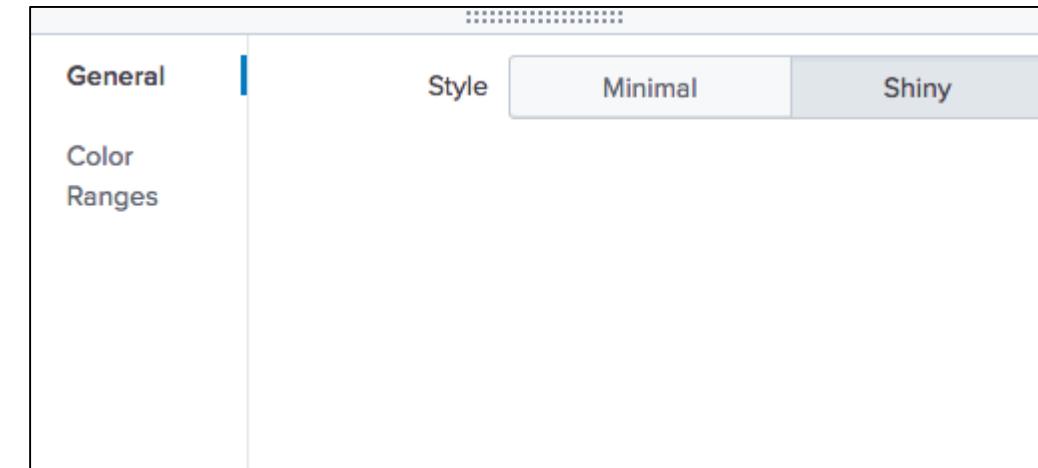
Generated for Anh Hoang (hoang_anh@pali.com) (C) Splunk Inc. not for distribution

Single Value Visualizations: Formatting

Set color using UI or with the gauge command



```
sourcetype=access_combined action=purchase  
| stats sum(price) as count  
| gauge count 0 5000 10000 15000
```



Generated for Anh Hoang (hoang_anhkim@bah.com) (C) Splunk Inc, not for distribution

Single Value Visualizations: Formatting (cont.)



Trellis layout displays multiple gauges when using a by clause in the stats command

```
sourcetype=access_combined  
| stats count by action
```

Single Value Visualizations: Formatting (cont.)

```
index=security sourcetype=linux_secure  
(fail* OR invalid)  
| stats count(vendor_action)
```

264

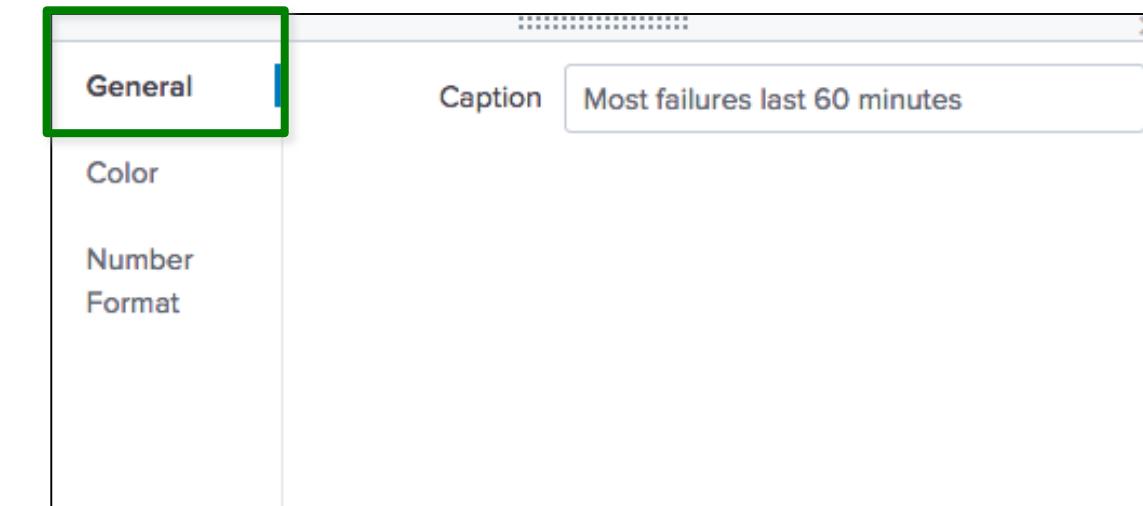
Failed/Invalid last 60 minutes



```
index=security sourcetype=linux_secure  
(fail* OR invalid)  
| chart count by src_ip  
| sort -count
```

10.1.10.172

Most failures last 60 minutes



Generated for Anh Hoang (hoang_anhkim@bah.com) (C) Splunk Inc, not for distribution

Single Value Visualizations: Formatting (cont.)

The screenshot shows the Splunk interface for a Single Value visualization. On the left, a configuration pane is open with tabs for General, Color (which is selected and highlighted with a green border), and Number Format. The Color tab contains settings for 'Use Colors' (Yes), 'Color by' (Value), and a color scale with ranges from 0 to 100. The preview pane on the right displays a large yellow box with the number '56' and the text 'Failed/Invalid last 15 minutes'. A green arrow points from the 'Color' tab in the configuration pane to the preview pane, indicating how the color mapping applies to the displayed value.

index=security sourcetype=linux_secure
(fail* OR invalid)
stats count

To resize the font,
resize the pane

Generated for Anh Hoang (hoang_anhkim@bah.com) (C) Splunk Inc, not for distribution

Single Value Visualizations: Formatting (cont.)

Optionally, specify number format information for the single value on the **Number Format** tab

```
index=security sourcetype=linux_secure  
(fail* OR invalid)  
| stats count
```

The screenshot shows the Splunk interface for configuring a single value visualization. On the left, there's a navigation bar with tabs: '42 Single Value' (selected), 'Format' (highlighted in blue), and 'Trellis'. A modal window titled 'Format' is open, showing settings for the single value: 'Precision' set to 0, 'Use Thousand Separators' set to 'Yes', 'Unit' set to 'Events', and 'Unit Position' set to 'Before'. To the right of the modal is the visualization itself, which displays the value '62Events' in large white text on a yellow background, with the subtitle 'Failed/invalid last 15 minutes' in smaller text below it.

Generated for Anh Hoang (hoang_anhkim@bah.com) (C) Splunk Inc, not for distribution

Single Value Visualizations: timechart

- With the **timechart** command, you can add a sparkline and a trend
- A **sparkline** is an inline chart
 - It is designed to display time-based trends associated with the primary key
- The **trend** shows the direction in which values are moving
 - It appears to the right of the single value

```
index=security sourcetype=linux_secure  
fail* OR invalid  
| timechart span=15m count(vendor_action)
```



Adding Totals Using Format Options

- Automatically total every column using the Format options
- When using this approach, you:
 - Cannot indicate which column to total; all columns are always totaled
 - Cannot add labels

Scenario ?

For the last 60 minutes, display the total number of events, with the total and average size (in bytes) by web server. Also, calculate the total bytes.

```
index=web sourcetype=access_combined  
| stats sum(bytes) as Bytes,  
avg(bytes) as avgBytes,  
count as totalEvents by host
```

host	Bytes	avgBytes	totalEvents
www1	173626	2066.9761904761904	84
www2	69789	2052.6176470588234	34
www3	92530	2372.5641025641025	39

General **3** Totals Summary

Percentages

host	Bytes	avgBytes	totalEvents
www1	173626	2066.9761904761904	84
www2	69789	2052.6176470588234	34
www3	92530	2372.5641025641025	39
	335945	6492.157940099117	157

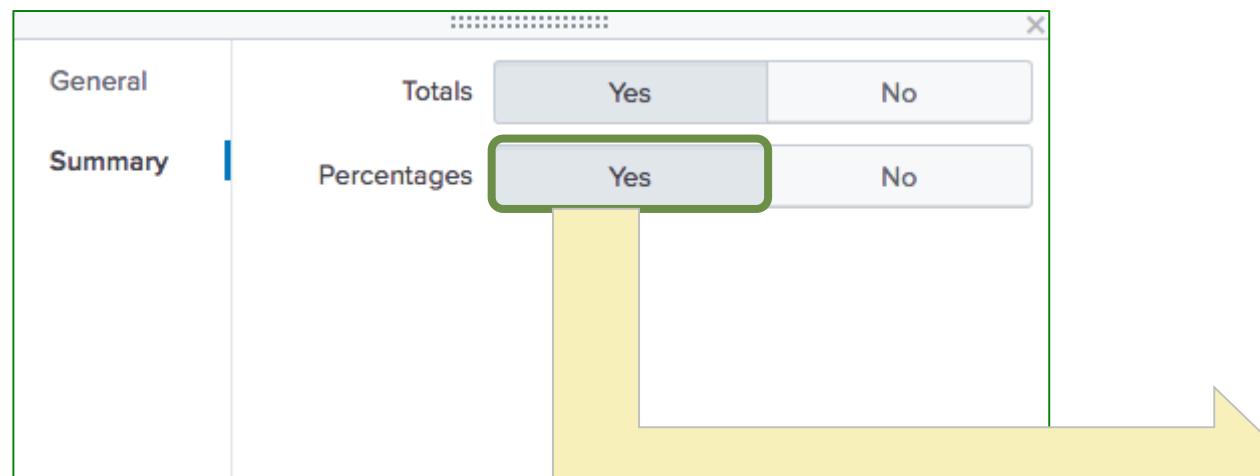
Generated for Anh Hoang (hoang_anhkim@bah.com) (C) Splunk Inc, not for distribution

Adding Totals Using Format Options (cont.)

- Using the **Summary** tab, you can add a % row to the end of the statistics table
- All columns are used to compute the percentage – i.e., all percentages from all columns combined will equal approximately 100%

Scenario ?

Display the retail products sold by country with totals by product and by country during the last 4 hours.



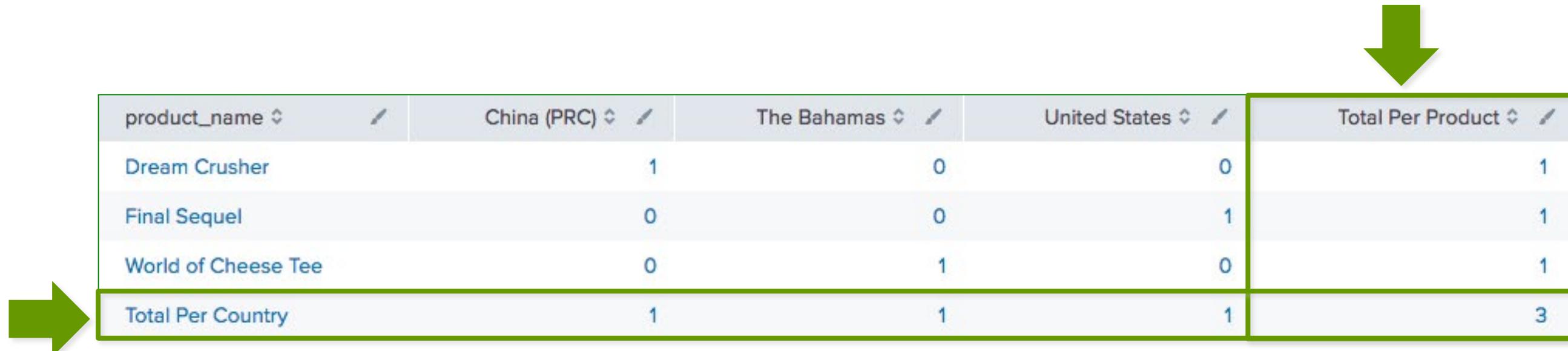
```
index=sales sourcetype=vendor_sales  
| chart count over product_name by VendorCountry
```

product_name	Haiti	The Bahamas	United States
Dream Crusher	0	0	3
Final Sequel	0	0	3
Holy Blade of Gouda	0	0	1
Manganiello Bros. Tee	1	0	1
Mediocre Kingdoms	0	0	2
Puppies vs. Zombies	0	0	1
World of Cheese Tee	0	1	1
	1	1	12
	7.1%	7.1%	85.7%

Generated for Anh Hoang (hoang_anhkim@bah.com) (C) Splunk Inc, not for distribution

Adding Totals Using addtotals Command

- Alternatively, use the `addtotals` command to:
 - Compute the sum of all **or selected** numeric fields for each column and place the total in the last row
 - Compute the sum of all **or selected** numeric fields for each **row** and place the total in the last column



product_name	China (PRC)	The Bahamas	United States	Total Per Product
Dream Crusher	1	0	0	1
Final Sequel	0	0	1	1
World of Cheese Tee	0	1	0	1
Total Per Country	1	1	1	3

Generated for Anh Hoang (hoang_anhkim@bah.com) (C) Splunk Inc, not for distribution

addtotals Command: Syntax

`addtotals [row=bool] [fieldname=field]
[col=bool][labelfield=field] [label=string] field-list`

Row Options		Column Options	
<code>row=true/false (Default= true)</code>	A column is created that contains numeric totals for each row.	<code>col=true/false (Default= false)</code>	A row is created that contains numeric totals for each column.
<code>fieldname=<i>field</i> (Default=Total)</code>	Defines a string used to create a field name for the totals column.	<code>label=<i>string</i> (Default=Total)</code>	Defines a string used to name the totals row.
		<code>labelfield=<i>fieldname</i></code>	Defines where the label string is placed. (Generally, you should make this the first column.)

General Options

`field-list=one or more numeric fields.
(Default: all numeric fields)`

Defines the numeric fields to be totaled.

addtotals Command – Example 1

- `row=t` (default) counts the fields in each row under a column named "Total Per Product"

```
index=sales sourcetype=vendor_sales  
| chart count over product_name by VendorCountry  
| addtotals  
  fieldname="Total Per Product" A  
  col=t B  
  label="Total Per Country" labelfield=product_name C
```

product_name	China (PRC)	The Bahamas	United States	Total Per Product
Dream Crusher	1	0	0	1
Final Sequel	0	0	1	A 1
World of Cheese Tee	0	1	0	1
Total Per Country	1	1	B 1	3

- `col=t` counts the fields in each row in a row named "Total Per Country"

Scenario

Calculate the total retail products sold by country totaled by product and by country during the last 60 minutes.

addtotals Command – Example 2

Scenario ?

For the last 60 minutes, display the total number of events with the total and average size (in bytes) by web server, and then total the bytes.

```
index=web sourcetype=access_combined  
| stats sum(bytes) as Bytes,  
avg(bytes) as avgBytes,  
count as totalEvents by host  
| addtotals row=f A col=t B label=totalBytes C  
labelfield=host D Bytes E
```

- A Do not total rows
- B Total columns
- C Add the label totalBytes
- D Place the label under the host column
- E Only total the Bytes column

D host	B Bytes	avgBytes	totalEvents
www1	141882	2149.7272727272725	66
www2	93781	2084.0222222222224	45
www3	122834	2233.3454545454547	55
C totalBytes	E 358497		

Lab Exercise 3

Time: 45 minutes

Tasks:

- Show the failures and the trend on the web server during the last 7 days
- Display the sales count of strategy games per day at Buttercup Games physical sales locations (i.e., not online) during the previous week
- Display a choropleth map of American retail sales during the previous week
- Display a map of online sales by country during the previous week
- Count the retail sales units sold by country and include a grand total row

Module 4: Filtering Results and Manipulating Data

Generated for Anh Hoang (hoang_anhkim@bah.com) (C) Splunk Inc, not for distribution

Module Objectives

- Use the eval command to:
 - Perform calculations
 - Convert values
 - Round values
 - Format values
 - Use conditional statements
- Use the search and where commands to filter calculated results
- Use fillnull command

eval Command – Overview

- Use eval to calculate and manipulate field values in your report
evalfieldname1 = expression1 [,fieldname2 = expression2...]
- Supports a variety of functions
- Results of eval written to either new or existing field you specify
 - If the destination field exists, the values of the field are replaced by the results of eval
 - Indexed data is not modified, and no new data is written into the index
 - Field values are treated in a case-sensitive manner
- Double quoted strings treated as field values
- Unquoted and single quoted fields treated as field names

Generated for Anh Hoang (hoang_anhkim@bah.com) (C) Splunk Inc, not for distribution

eval Command – Overview (cont.)

- Multiple expressions can be combined into one eval command
- Each subsequent expression references the results of previous expressions
- Expressions must be separated by commas

```
eval fieldname1 = expression1,  
      fieldname2 = expression2,  
      fieldname3 = expression3...
```

eval Command – Operators

- The eval command allows you to:
 - Calculate expressions
 - Place the results in a field
 - Use that field in searches or other expressions

- Type Operators

Arithmetic + - * / %

Concatenation + .

Boolean AND OR NOT XOR

Comparison < > <= >= != = LIKE

eval

[Learn More ↗](#)

Calculates an expression and puts the resulting value into a field. You can specify to calculate more than one expression.

Example:

... | eval velocity=distance/time

Note

For more info on common eval functions, see
<http://docs.splunk.com/Documentation/Splunk/latest/SearchReference/CommonEvalFunctions>

eval Command – Convert Values

- This example report displays the sum of bytes used for each usage category
- It is difficult to determine how much bandwidth is being used by looking at bytes
- First, use eval to convert the bytes value into megabytes

Scenario



What types of websites used the most bandwidth in bytes during the previous month?

```
index=network sourcetype=cisco_wsa_squid  
| stats sum(sc_bytes) as Bytes by usage
```

usage	Bytes
Borderline	54328755
Business	65534755
Personal	301103659
Unknown	77328971
Violation	3199635

eval Command – Convert Values (cont.)

- Results of eval must be set to a new or existing field
- In this example:
 - Calculate the number of bytes for each usage type
 - Create a new field named bandwidth
 - Convert the values of the Bytes field into MB by dividing Bytes field values by $(1024*1024)$

Scenario



What types of websites used the most bandwidth in megabytes during the previous month?

```
index=network sourcetype=cisco_wsa_squid  
| stats sum(sc_bytes) as Bytes A by usage  
| eval bandwidth B = Bytes/(1024*1024) C
```

usage	A Bytes	B bandwidth
Borderline	54011022	51.50892448425293
Business	66844576	63.747955322265625
Personal	299584913	285.7064371109009
Unknown	77187989	73.61220264434814
Violation	3203231	3.0548391342163086

eval Command – Round Values

- The results of Bandwidth are hard to read with so many decimal points
- `round(field/number, decimals)` function sets the value of a field to the number of decimals you specify
- In this example:
 - Divide the value of the Bytes field by $(1024*1024)$
 - Round the result to two decimal points
- If the number of decimals is unspecified, the result is a whole number

Scenario

What types of websites used the most bandwidth in megabytes, rounded to 2 decimal places, during the previous month? Sort by bandwidth.

```
index=network sourcetype=cisco_wsa_squid
| stats sum(sc_bytes) as Bytes by usage
| eval bandwidth = round(Bytes/(1024*1024), 2) A
| sort -bandwidth
| rename bandwidth as "Bandwidth (MB)"
```

usage	Bytes	Bandwidth (MB)
Personal	299584913	285.71 A
Unknown	77187989	73.61
Business	66844576	63.75
Borderline	54011022	51.51
Violation	3203231	3.05

Removing Fields

- The "Bandwidth (MB)" field has the data in the desired format
- The Bytes field is no longer needed
 - The Bytes field can be removed

Scenario

What types of websites used the most bandwidth in megabytes, rounded to 2 decimal places, during the previous month? Sort by bandwidth.

```
index=network sourcetype=cisco_wsa_squid  
| stats sum(sc_bytes) as Bytes by usage  
| eval bandwidth = round(Bytes/(1024*1024), 2)  
| sort -bandwidth  
| rename bandwidth as "Bandwidth (MB)"  
| fields - Bytes A
```

usage	Bandwidth (MB)
Personal	285.71
Unknown	73.61
Business	63.75
Borderline	51.51
Violation	3.05

Generated for Anh Hoang (hoang_anhkim@bah.com) (C) Splunk Inc, not for distribution

eval Command – Calculating Values

You can perform mathematical functions against fields with numeric field values

- A In this example, stats calculates the total list price and total sale price by product_name
- B eval calculates the discount percentage and formats the discount field

Scenario



Calculate total online sales for last week; include price, sales price, and discount percentage. Sort by descending discount value.

```
index=web sourcetype=access_combined product_name=*  
action=purchase  
A | stats sum(price) as tp, sum(sale_price) as tsp by product_name  
B | eval Discount = round(((tp - tsp)/ tp)*100)  
| sort -Discount  
| eval Discount = Discount.%"  
| rename tp as "Total List Price", tsp as "Total Sale Price",  
product_name as Product
```

Product	Total List Price	Total Sale Price	Discount
Puppies vs. Zombies	578.84	230.84	60%
Fire Resistance Suit of Provolone	594.51	296.51	50%
Holy Blade of Gouda	742.76	370.76	50%
Dream Crusher	6438.39	4023.39	38%
Manganiello Bros.	5758.56	3598.56	38%

Generated for Anh Hoang (hoang_anhkim@bah.com) (C) Splunk Inc, not for distribution

eval Command – Calculating Values (cont.)

- C sort lists the highest discounted items first
- D eval converts Discount to a string and concatenates the % character
- E rename provides user friendly headings

Scenario



Calculate total online sales for last week; include price, sales price, and discount percentage. Sort by descending discount value.

```
index=web sourcetype=access_combined product_name=*  
action=purchase  
stats sum(price) as tp, sum(sale_price) as tsp by product_name  
eval Discount = round(((tp - tsp)/ tp)*100)  
sort -Discount  
eval Discount = Discount.%"  
rename tp as "Total List Price", tsp as "Total Sale Price",  
product_name as Product
```

C
D
E

Product	Total List Price	Total Sale Price	Discount
Puppies vs. Zombies	578.84	230.84	60%
Fire Resistance Suit of Provolone	594.51	296.51	50%
Holy Blade of Gouda	742.76	370.76	50%
Dream Crusher	6438.39	4023.39	38%
Manganiello Bros.	5758.56	3598.56	38%

Generated for Anh Hoang (hoang_anhkim@bah.com) (C) Splunk Inc, not for distribution

Formatting and Sorting Values

- eval with added characters converts numeric field values to strings
- To order numerically, first sort, then use eval

```
index=web sourcetype=access_combined price=*
| stats values(price) as price by product_name
| eval price = "$".price
| sort -price
```

```
index=web sourcetype=access_combined price=*
| stats values(price) as price by product_name
| sort -price
| eval price = "$".price
```

product_name	price
Manganiello Bros. Tee	\$9.99
World of Cheese Tee	\$9.99
Holy Blade of Gouda	\$5.99
Puppies vs. Zombies	\$4.99
Dream Crusher	\$39.99
Manganiello Bros.	\$39.99
Orvil the Wolverine	\$39.99
Fire Resistance Suit of Provolone	\$3.99

Alpha

product_name	price
Dream Crusher	\$39.99
Manganiello Bros.	\$39.99
Orvil the Wolverine	\$39.99
Benign Space Debris	\$24.99
Final Sequel	\$24.99
Mediocre Kingdoms	\$24.99
World of Cheese	\$24.99
Curling 2014	\$19.99

Numeric

Generated for Anh Hoang (hoang_anhkim@bah.com) (C) Splunk Inc, not for distribution

eval Command – tostring Function

- **tostring** converts a numeric field value to a string

tostring(field, "option")

- Options:
 - "commas": applies commas
 - › If the number includes decimals, it rounds to two decimal places
 - "duration": formats the number as "hh:mm:ss"
 - "hex": formats the number in hexadecimal

Scenario

How much potential online sales revenue was lost during the previous week due to 503 server errors?

```
index=web sourcetype=access_combined  
action=purchase status=503  
| stats count(price) as NumberOfLostSales, A  
    avg(price) as AverageLostSales,  
    sum(price) as TotalLostRevenue  
| eval AverageLostSales =  
    "$" + tostring(AverageLostSales, "commas"), B  
TotalLostRevenue =  
    "$" + tostring(TotalLostRevenue, "commas") C
```

NumberOfLostSales	AverageLostSales	TotalLostRevenue
A 113	\$21.58 B	\$2,438.87 C

tostring Function – duration Option

This example shows "duration" option of `tostring` function

- A stats calculates `sessionTime` for each session (`JSESSIONID`)
 - Use the `range` function to return the difference between the max and min values of `_time`
- B `sort 5` displays the top 5 most frequent values
- C The `duration` option formats the time as "hh:mm:ss"

Scenario



Identify the five longest client sessions over the last 4 hours in HH:MM:SS format.

```
index=web sourcetype=access_combined  
| stats range(_time) as sessionTime by JSESSIONID A  
| sort 5 -sessionTime B  
| eval duration = tostring(sessionTime,"duration") C
```

JSESSIONID A	sessionTime A	duration C
SD6SL8FF3ADFF4951	175	00:02:55
SD6SL1FF10ADFF4966	158	00:02:38
SD6SL3FF1ADFF4959	141	00:02:21
SD0SL1FF2ADFF4963	135	00:02:15
SD4SL10FF2ADFF4961	132	00:02:12

eval Commands with Multiple Expressions

- A Based on the `values` of `list_price` and `current_sale_price`, calculate the `current_discount` percentage
- B Calculate the `new_discount` value by subtracting 5 from `current_discount`
- C Calculate the `new_sale_price` by applying the `new_discount` percentage

Scenario



Calculate a new sale price that is 5% less than the current discount percentage for online sales data over the last hour.

```
index=web sourcetype=access_combined price=*
| stats values(price) as list_price, values(sale_price)
as current_sale_price by product_name
| eval current_discount = round((list_price - current_sale_price)/list_price*100,2),
new_discount = (current_discount - 5),
new_sale_price = list_price - (list_price * (new_discount/100))
```

A
B
C

product_name	list_price	current_sale_price	current_discount	new_discount	new_sale_price
Benign Space Debris	24.99	19.99	20.01	15.01	21.24
Curling 2014	19.99	16.99	15.01	10.01	17.99
Dream Crusher	39.99	24.99	37.51	32.51	26.99
Final Sequel	24.99	16.99	32.01	27.01	18.24

Generated for Anh Hoang (hoang_anhkim@bah.com) (C) Splunk Inc, not for distribution

eval Command – if Function Syntax

`if(X,Y,Z)`

- The `if` function takes three arguments
- The first argument, `X`, is a boolean expression
 - If it evaluates to TRUE, the result evaluates to the second argument, `Y`
 - If it evaluates to FALSE, the result evaluates to the third argument, `Z`
- Non-numeric values must be enclosed in "double quotes"
- Field values are treated in a case-sensitive manner

eval Command – if Function Example

Scenario

Display retail sales for the previous week, broken down by Asia and the Rest of the World.

```
index=sales sourcetype=vendor_sales  
| eval SalesTerritory =  
| if((VendorID >= 7000 AND VendorID < 8000), "Asia", "Rest of the World")  
| stats sum(price) as TotalRevenue by SalesTerritory  
| eval TotalRevenue = "$" + tostring(TotalRevenue, "commas")
```

- Create a new field, **SalesTerritory**
- Evaluate **VendorID**
 - If ≥ 7000 AND < 8000 is TRUE, set result to "Asia"
 - Remember, arguments must be enclosed in quotes
 - If it evaluates to FALSE, set result to "Rest of the World"

SalesTerritory	TotalRevenue
Asia	\$1,371.26
Rest of the World	\$14,495.25

Generated for Anh Hoang (hoang_anhkim@bah.com) (C) Splunk Inc, not for distribution

eval Command – case Function Syntax

`case(X1,Y1,X2,Y2...)`

- The first argument, $X1$, is a boolean expression
- If it evaluates to TRUE, the result evaluates to $Y1$
- If it evaluates to FALSE, the next boolean expression, $X2$, is evaluated, etc.
- If none of the boolean expressions are true, the result evaluates to NULL

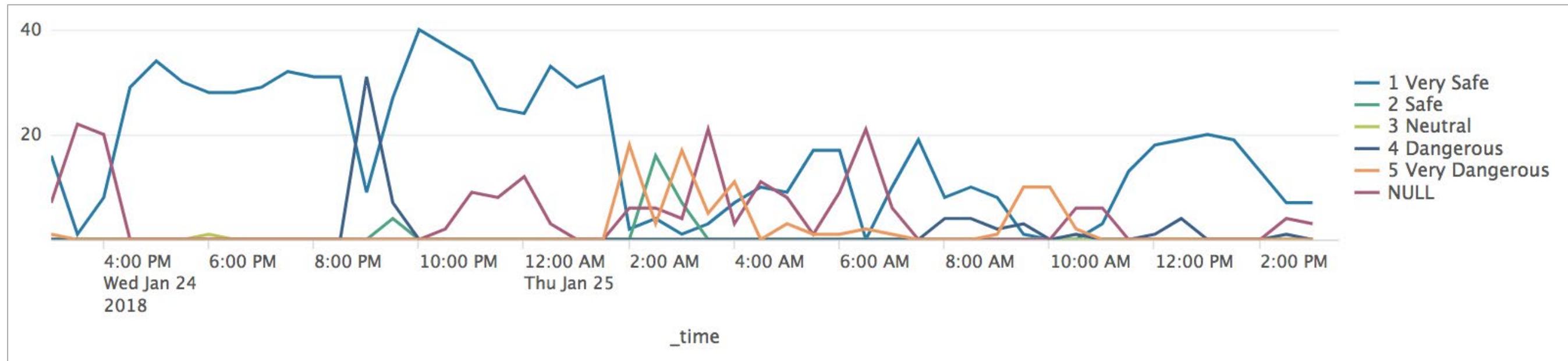
eval Command – case Function Example

Scenario

SecOps found a potential virus on a user's machine. Find and classify the number of internet visits by risk during the past 24 hours.



```
index=network sourcetype=cisco_wsa_squid
| eval Risk = case(x_wbrs_score >= 5,"1 Very Safe",
x_wbrs_score >= 3,"2 Safe",
x_wbrs_score >= 0,"3 Neutral",
x_wbrs_score >= -5,"4 Dangerous",
x_wbrs_score < -5, "5 Very Dangerous")
| timechart count by Risk
```

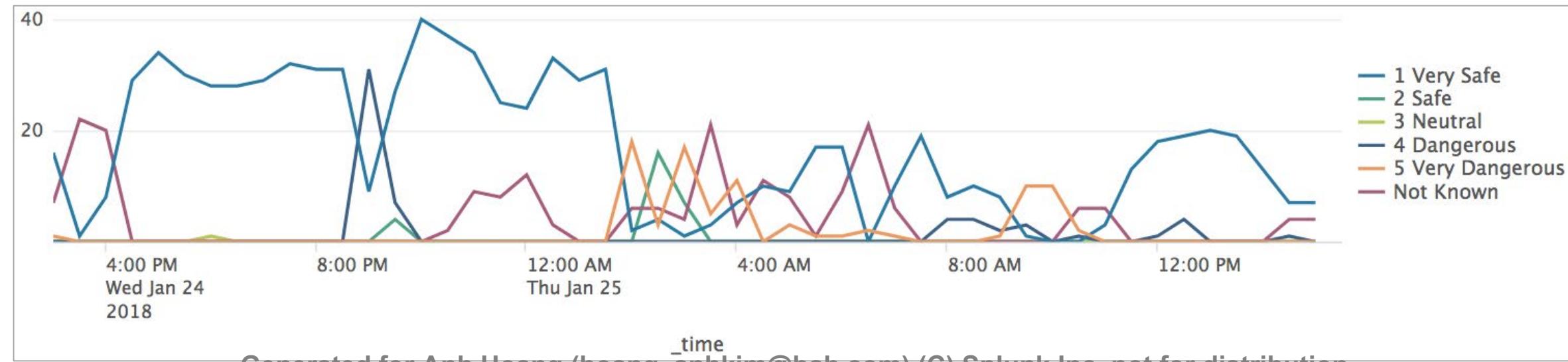


Generated for Anh Hoang (hoang_anhkim@bah.com) (C) Splunk Inc, not for distribution

case Function – Setting the Default

- If you want an “otherwise” clause, just test for a condition you know is true at the end (e.g., $0=0$)
- You can also use the `true()` function, which always evaluates to TRUE

```
index=network sourcetype=cisco_wsa_squid
| eval Risk = case(x_wbrs_score >= 5, "1 Very Safe",
x_wbrs_score >= 3, "2 Safe",
x_wbrs_score >= 0, "3 Neutral",
x_wbrs_score >= -5, "4 Dangerous",
x_wbrs_score < -5, "5 Very Dangerous",
true(), "Not Known")
| timechart count by Risk
```



Generated for Anh Hoang (hoang_anhkim@bah.com) (C) Splunk Inc, not for distribution

eval function

- To count the number of events that contain a specific field value, use the count and eval functions
 - Used within a transforming command, such as stats
 - Requires an as clause
 - Double quotes are required for character field values
 - Field values are case-sensitive

Scenario

Count the number of events that occurred yesterday where the vendor action was Accepted, Failed, or session opened.

```
index=security sourcetype=linux_secure vendor_action=*  
| stats  
count(eval(vendor_action="Accepted")) as Accepted, A  
count(eval(vendor_action="Failed")) as Failed, B  
count(eval(vendor_action="session opened")) as SessionOpened C
```

Accepted	Failed	SessionOpened
A 322	B 9676	C 407

Filtering Results – search and where

- The search and where commands both filter results
 - search
 - May be easier if you're familiar with basic search syntax
 - Treats field values in a case-insensitive manner
 - Allows searching on keyword
 - Can be used at any point in the search pipeline
 - where
 - Can compare values from two different fields
 - Functions are available, such as `isnotnull()`
 - Treats field values in a case-sensitive manner
 - Can't appear before first pipe in search pipeline

search Command

- To filter results, use search at any point in the search pipeline
- Behaves exactly like search strings before the first pipe
 - Uses the "*" wildcard
 - Treats field values in a case-insensitive manner

Scenario



Report which products during the last 24 hours have sold more than \$500 online.

```
index=web sourcetype=access_combined  
action=purchase status=200  
| stats sum(price) as sales by product_name  
| search sales>500 A  
| sort -sales  
| eval sales="$"+sales  
| rename sales as "Popular Products",  
product_name as "Product Name"
```

Product Name	Popular Products
Manganiello Bros.	\$839.79 A
Dream Crusher	\$799.80
SIM Cubicle	\$679.66

where Command

where eval-expression

- Uses same expression syntax and functions as eval command
- Uses boolean expressions to filter search results and only keeps results that are true
- Unquoted or single-quoted strings are treated as fields
- Double-quoted strings are interpreted as field values
- Treats field values in a case-sensitive manner

Note

To view all of the functions for where, see:
<http://docs.splunk.com/Documentation/Splunk/latest/SearchReference/Where?r=searchtip>

where Command – Example

- Filters search results using eval expressions
- Used to compare two different fields, which you can't do with search

Scenario



SalesOps wants to know which days over the previous week have seen more “remove” actions than “change quantity” actions.

```
index=web sourcetype=access_combined  
| timechart count(eval(action="changequantity"))  
as changes, count(eval(action="remove")) as removals  
| where removals > changes A
```

_time	changes	removals
2018-01-07	108	130 A
2018-01-09	126	135

where Command – Case Sensitivity

- Use where to perform case sensitive value searches

```
sourcetype=access_combined | where action="purchase"
```

- Remember, search is not case sensitive

```
sourcetype=access_combined action="purchase"
```

returns all variations of purchase, Purchase, PURCHASE, pUrChAsE

- To use the search command to perform a case-sensitive search on keywords, use the CASE directive

```
sourcetype=access_combined CASE(purchase)
```

Note

Don't confuse the **CASE** directive with the previously shown **case** function.

where Command With like Operator

- Can do wildcard searches with where command
- Use (_) for one character and (%) for multiple characters
- Must use the like operator with wildcards

src_ip	count
102.2.83.137	102
107.3.146.207	55
108.65.113.83	34
109.169.32.135	72

Generated for Anh Hoang (hoang_anhkim@bah.com) (C) Splunk Inc, not for distribution

Scenario



Report the number of events over the past 24 hours by IP address for a specific range of addresses.

```
index=security sourcetype=linux_secure  
| stats count by src_ip  
| where src_ip like "10_%"
```

Note



The **like** operator can also be used with the **case** function.

where Command With like Operator (cont.)

Scenario



IT wants a list of user accounts that are like admin (adm%).

- %dm% and _dm% both return an obvious user name: edmond

```
index=security sourcetype=linux_secure  
| where user like "_dm%"  
| dedup user  
| table user
```

user
administrator
edmond
adm
admin

- _dm misses sapadmin and sysadmin

```
index=security sourcetype=linux_secure  
| where user like "adm%"  
| dedup user  
| table user
```

user
administrator
adm
admin

where Command – isnull and isnotnull

Scenario



A sales campaign manager wants to know which 15 minute periods contained no sales during the previous year.

- Use `isnull` to find events with an empty value for a particular field
- Use `isnotnull` to find events that contain a non-empty value for a particular field

```
index=sales sourcetype=vendor_sales  
| timechart span=15m sum(price) as sum  
| where isnull(sum)
```

_time	sum
2018-03-20 19:15:00	
2018-03-20 19:30:00	
2018-03-20 19:45:00	

Generated for Anh Hoang (hoang_anhkim@bah.com) (C) Splunk Inc. not for distribution

fillnull Command

- Use `fillnull` to replace null values in fields
- Use `value=string` to specify a string you want displayed instead
 - Example: `fillnull value=NULL`
- If no `value=` clause, default replacement value is 0
- Optionally, restrict which fields `fillnull` apply to by listing them at end of command
 - Example: `fillnull VALUE="N/A" discount refund`

fillnull

Replaces null values with a specified value.

Example:

`sourcetype="web" | timechart count by host | fillnull value=NULL`

[Learn More](#)

fillnull Command – Examples

Scenario ?

Evaluate vendor sales by country for the last hour.

```
index=sales sourcetype=vendor_sales  
| chart sum(price) over product_name by VendorCountry  
| fillnull
```

product_name	Morocco	United States
Dream Crusher	39.99	39.99
Puppies vs. Zombies	0	4.99

```
index=sales sourcetype=vendor_sales  
| chart sum(price) over product_name by VendorCountry  
| fillnull value="No Value"
```

product_name	Morocco	United States
Dream Crusher	39.99	39.99
Puppies vs. Zombies	No Value	4.99

Lab Exercise 4

Time: 45 – 55 minutes

Tasks:

- Chart the total daily volume (in MB) of the web servers during the previous week
- Calculate the ratio of GET requests to POST requests for each web server
- Identify users with more than 3 failed logins during the last 60 minutes, sort in descending order
- Evaluate and classify the size of events on the web servers during the last 24 hours as a pie chart

****Challenge Exercises:**

- Classify and report employee web traffic by content type during the previous business week
- Report which products sold twice as much in the Buttercup Games online store than in the retail store during the previous week

Module 5: Correlating Events

Generated for Anh Hoang (hoang_anhkim@bah.com) (C) Splunk Inc, not for distribution

Module Objectives

- Identify transactions
- Group events using fields
- Group events using fields and time
- Search with transactions
- Report on transactions
- Determine when to use transaction vs. stats

What is a Transaction?

- A transaction is any group of related events that span time
- Events can come from multiple applications or hosts
 - Events related to a single purchase from an online store can span across an application server, database, and e-commerce engine
 - One email message can create multiple events as it travels through various queues
 - Each event in the network traffic logs represents a single user generating a single http request
 - Visiting a single website normally generates multiple http requests
 - HTML, JavaScript, CSS files
 - Flash, images, etc.

transaction Command

- **transaction *field-list***
 - *field-list* can be one field name or a list of field names
 - Events are grouped into transactions based on the values of these fields
 - If multiple fields are specified and a relationship exists between those fields, events with related field values are grouped into a single transaction

- Common constraints:

maxspan maxpause startswith endswith

transaction

Groups events into transactions.

Example:

... | transaction host cookie maxspan=30s maxpause=5s

[Learn More ↗](#)

Events That Have the Same JSESSIONID

- The log shows a number of events that share the same JSESSIONID value (SD0SL10FF3ADFF4950)
- However, it is difficult to:
 - View the events as a group
 - Gain insight as to what is happening with these events
 - Know if there are other events scattered in the results set

Scenario
Display customer transactions in the online store during the last 60 minutes.

index=web sourcetype=access_combined

> 1/17/18 4:01:08.000 PM	175.44.1.122 - - [18/Jan/2018:00:01:08] "GET /oldlink?itemId=EST-6&JSESSIONID=SD0SL10FF3ADFF4950 HTTP 1.1" 200 3516 "http://www.buttercupgames.com/oldlink?itemId=EST-6" "Mozilla/4.0 (compatible; MSIE 7.0; Windows NT 5.2; .NET CLR 1.1.4322; InfoPath.1; MS-RTC LM 8)" 614	JSESSIONID = SD0SL10FF3ADFF4950 host = www3 source = /opt/log/www3/access.log sourcetype = access_combined
> 1/17/18 4:01:05.000 PM	175.44.1.122 - - [18/Jan/2018:00:01:05] "GET /category.screen?categoryId=STRATEGY&JSESSIONID=SD0SL10FF3ADFF4950 HTTP 1.1" 200 2453 "http://www.buttercupgames.com/oldlink?itemId=EST-18" "Mozilla/4.0 (compatible; MSIE 7.0; Windows NT 5.2; .NET CLR 1.1.4322; InfoPath.1; MS-RTC LM 8)" 989	JSESSIONID = SD0SL10FF3ADFF4950 host = www3 source = /opt/log/www3/access.log sourcetype = access_combined
> 1/17/18 4:00:53.000 PM	175.44.1.122 - - [18/Jan/2018:00:00:53] "POST /category.screen?categoryId=SIMULATION&JSESSIONID=SD0SL10FF3ADFF4950 HTTP 1.1" 200 3526 "http://www.buttercupgames.com" "Mozilla/4.0 (compatible; MSIE 7.0; Windows NT 5.2; .NET CLR 1.1.4322; InfoPath.1; MS-RTC LM 8)" 222	JSESSIONID = SD0SL10FF3ADFF4950 host = www3 source = /opt/log/www3/access.log sourcetype = access_combined

transaction Command – Example 1

- The transaction command creates a single event from a group of events
 - The events must share the same value in a specified field
- Transactions can cross multiple tiers such as web servers or application servers
- For example, you can easily view the events for JSESSIONID SD0SL10FF3ADFF4950

Scenario



Group together Buttercup Games online store events based on the JSESSIONID value for the last 15 minutes.

```
index=web sourcetype=access_combined  
| transaction JSESSIONID
```

```
> 1/17/18 175.44.1.122 -- [18/Jan/2018:00:00:53] "POST /category.screen?categoryId=SIMULATION&JSESSIONID=SD0SL10FF3ADFF4950  
4:00:53.000 PM HTTP 1.1" 200 3526 "http://www.buttercupgames.com" "Mozilla/4.0 (compatible; MSIE 7.0; Windows NT 5.2; .NET CLR 1.1  
.4322; InfoPath.1; MS-RTC LM 8)" 222  
175.44.1.122 -- [18/Jan/2018:00:01:05] "GET /category.screen?categoryId=STRATEGY&JSESSIONID=SD0SL10FF3ADFF4950 HTTP  
1.1" 200 2453 "http://www.buttercupgames.com/oldlink?itemId=EST-18" "Mozilla/4.0 (compatible; MSIE 7.0; Windows NT  
5.2; .NET CLR 1.1.4322; InfoPath.1; MS-RTC LM 8)" 989  
175.44.1.122 -- [18/Jan/2018:00:01:08] "GET /oldlink?itemId=EST-6&JSESSIONID=SD0SL10FF3ADFF4950 HTTP 1.1" 200 3516  
"http://www.buttercupgames.com/oldlink?itemId=EST-6" "Mozilla/4.0 (compatible; MSIE 7.0; Windows NT 5.2; .NET CLR 1  
.1.4322; InfoPath.1; MS-RTC LM 8)" 614  
175.44.1.122 -- [18/Jan/2018:00:01:17] "GET /category.screen?categoryId=ACCESSORIES&JSESSIONID=SD0SL10FF3ADFF4950  
HTTP 1.1" 200 652 "http://www.buttercupgames.com/oldlink?itemId=EST-16" "Mozilla/4.0 (compatible; MSIE 7.0; Windows  
NT 5.2; .NET CLR 1.1.4322; InfoPath.1; MS-RTC LM 8)" 118
```

transaction Command – Example 2

- Use the search command at any point in the search pipeline to filter results
- Behaves exactly like search strings before the first pipe
 - search uses the "*" wildcard and treats field values in a case-insensitive manner
 - status=404 finds the errors
 - highlight highlights the terms you specify

Scenario



Display transactions that included a 404 error during the last 60 minutes.

```
index=web sourcetype=access_combined  
| transaction JSESSIONID A  
| search status=404  
| highlight JSESSIONID, 404 B
```

```
1/17/18      ... 2 lines omitted ...  
4:00:53.000 PM 175.44.1.122 -- [18/Jan/2018:00:01:08] "GET /oldlink?itemId=-6&JSESSIONID=SD0SL10FF3ADFF4950  
HTTP 1.1" 200 3516 "http://www.buttercupgames.com/oldlink?itemId=EST-6" "Mozilla/4.0 (compatible  
; MSIE 7.0; Windows NT 5.2; .NET CLR 1.1.4322; InfoPath.1; MS-RTC LM 8)" 614  
... 6 lines omitted ...  
175.44.1.122 -- [18/Jan/2018:00:01:55] "POST /product.screen?productId=BS-1&JSESSIONID=SD0  
SL10FF3ADFF4950 HTTP 1.1" 200 1537 "http://www.buttercupgames.com/category.screen?categoryId=ARC  
ADE" "Mozilla/4.0 (compatible; MSIE 7.0; Windows NT 5.2; .NET CLR 1.1.4322; InfoPath.1; MS-RTC L  
M 8)" 782  
175.44.1.122 -- [18/Jan/2018:00:02:00] "GET /oldlink?itemId=E1&JSESSIONID=SD0SL10FF3ADFF495  
0 HTTP 1.1" 408 2190 "http://www.buttercupgames.com/cart.do?action=view&itemId=EST-18" "Mozilla/  
4.0 (compatible; MSIE 7.0; Windows NT 5.2; .NET CLR 1.1.4322; InfoPath.1; MS-RTC LM 8)" 512  
175.44.1.122 -- [18/Jan/2018:00:02:15] "GET /search.do?item A 2&JSESSIONID=SD0SL10FF3ADFF4950  
HTTP B 404 3969 "http://www.buttercupgames.com/cart.do?action=view&itemId=EST-6" "Mozilla/4.0  
(compatible; MSIE 7.0; Windows NT 5.2; .NET CLR 1.1.4322; InfoPath.1; MS-RTC LM 8)" 859
```

transaction Command – Example 3

Scenario

For failed network logins, display different users from the same IP during the last 60 minutes.



```
index=security sourcetype=linux_secure failed  
| transaction src_ip
```

```
> 1/17/18      Wed Jan 17 2018 23:49:43 www1 sshd[2796]: Failed password for backup from 92.46.53.223 port 3138 ssh2  
3:49:43.000 PM Thu Jan 18 2018 00:23:12 www2 sshd[2462]: Failed password for invalid user operator from 92.46.53.223 port 1075 ssh2  
host = www1 host = www2 | source = /opt/log/www1/secure.log source = /opt/log/www2/secure.log | sourcetype = linux_secure  
  
> 1/17/18      Wed Jan 17 2018 23:49:43 mailsv1 sshd[5594]: Failed password for invalid user db2 from 118.142.68.222 port 4345 ssh2  
3:49:43.000 PM Thu Jan 18 2018 00:19:15 www3 sshd[1373]: Failed password for invalid user admin from 118.142.68.222 port 1568 ssh2  
Thu Jan 18 2018 00:23:12 www3 sshd[2574]: Failed password for backup from 118.142.68.222 port 1700 ssh2  
host = mailsv1 host = www3 | source = /opt/log-mailsv1/secure.log source = /opt/log/www3/secure.log | sourcetype = linux_secure
```

transaction Command – Specific Fields

- The transaction command produces additional fields, such as:
 - duration – the difference between the timestamps for the first and last event in the transaction
 - eventcount – the number of events in the transaction

transaction Command – maxspan/maxpause

- You can also define a max overall time span and max gap between events

- **maxspan=10m**

- ▶ Maximum total time between the *earliest* and *latest* events
 - ▶ If not specified, default is -1 (or no limit)

- **maxpause=1m**

- ▶ Maximum total time *between* events
 - ▶ If not specified, default is -1 (or no limit)

Note

Assumptions: Transactions spanning more than 10 minutes with the same client IP are considered unrelated. Also, there can be no more than one minute between any two related events.

Scenario

Display customer actions on the website during the last 4 hours.

```
index=web sourcetype=access_combined  
| transaction clientip maxspan=10m maxpause=1m  
| eval duration = tostring(duration,"duration")  
| sort -duration  
| table clientip duration action  
| rename clientip as "Client IP",  
| action as "Client Actions"
```

Client IP	duration	Client Actions
91.199.80.24	00:01:27	addtocart changequantity purchase view
12.130.60.5	00:01:27	addtocart purchase view
60.220.218.88	00:01:25	view

transaction Command – startswith/endswith

- To form transactions based on terms, field values, or evaluations, use `startswith` and `endswith` options
- In this example:

- The first event in the transaction includes `addtocart`
- The last event includes `purchase`

Scenario



Determine the length of time spent to complete a purchase by customers in the online store over the last 24 hours.

```
index=web sourcetype=access_combined  
| transaction clientip JSESSIONID  
  startswith=eval(action="addtocart")  
  endswith=eval(action="purchase")  
| table clientip, JSESSIONID, duration, eventcount
```

clientip	JSESSIONID	duration	eventcount
223.205.219.198	SD1SL9FF5ADFF4953	3	2
201.3.120.132	SD10SL5FF2ADFF4956	4	2
201.3.120.132	SD10SL5FF2ADFF4956	1	2
203.45.206.135	SD10SL7FF4ADFF4964	2	2

Investigating with Transactions

- Transactions can be useful when a single event does not provide enough information
- This example searches email logs for the term “REJECT”
- Events that include the term do not provide much information about the rejection

Scenario

Find emails that were rejected during the last 24 hours.

```
index=network sourcetype=cisco_esa REJECT
```

i	Time	Event
>	1/18/18 12:24:58.000 AM	Thu Jan 18 08:24:58 2018 Info: ICID 744783 REJECT SG BLACKLIST match sbrs[10.0: 3.0] SBRS 10.0 host = cisco_router1 source = /opt/log/cisco_router1/cisco_ironport_mail.log sourcetype = cisco_esa
>	1/17/18 10:41:48.000 PM	Thu Jan 18 06:41:48 2018 Info: ICID 744777 REJECT SG BLACKLIST match sbrs[10.0: 3.0] SBRS 6.8 host = cisco_router1 source = /opt/log/cisco_router1/cisco_ironport_mail.log sourcetype = cisco_esa
>	1/17/18 9:47:37.000 PM	Thu Jan 18 05:47:37 2018 Info: ICID 744774 REJECT SG BLACKLIST match sbrs[10.0: 3.0] SBRS 10.0 host = cisco_router1 source = /opt/log/cisco_router1/cisco_ironport_mail.log sourcetype = cisco_esa
>	1/17/18 9:27:59.000 PM	Thu Jan 18 05:27:59 2018 Info: ICID 744773 REJECT SG BLACKLIST match sbrs[10.0: 3.0] SBRS 4.0 host = cisco_router1 source = /opt/log/cisco_router1/cisco_ironport_mail.log sourcetype = cisco_esa

Investigating with Transactions (cont.)

- By creating a transaction, you can then search and see additional events related to the rejection, such as:
 - IP address of sender
 - Reverse DNS lookup results
 - Action taken by the mail system following the rejection
- **mid** – Message ID
- **dcid** – Delivery Connection ID
- **icid** – Incoming Connection ID

Scenario

Find emails that were rejected in the last 24 hours.

```
index=network sourcetype=cisco_esa  
| transaction mid dcid icid  
| search REJECT
```

i	Time	Event
>	1/17/18 8:02:16.000 AM	Wed Jan 17 17:27:43 2018 Info: New SMTP ICID 744722 interface Management (192.16 8.3.120) address 82.90.51.177 reverse dns host host177 51 static.90 82 b.busines s.telecomitalia.it verified yes Wed Jan 17 17:27:55 2018 Info: ICID 744722 REJECT SG BLACKLIST match sbrs[10.0: 3.0] SBRS 10.0 Wed Jan 17 17:28:02 2018 Info: ICID 744722 close host = cisco_router1 source = /opt/log/cisco_router1/cisco_ironport_mail.log sourcetype = cisco_esa
>	1/17/18 8:02:16.000 AM	Wed Jan 17 18:01:11 2018 Info: New SMTP ICID 744726 interface Management (192.16 8.3.120) address 86.43.70.62 reverse dns host unknown verified no Wed Jan 17 18:01:16 2018 Info: ICID 744726 REJECT SG BLACKLIST match sbrs[10.0: 3.0] SBRS 10.0 Wed Jan 17 18:01:27 2018 Info: ICID 744726 close host = cisco_router1 source = /opt/log/cisco_router1/cisco_ironport_mail.log sourcetype = cisco_esa

Generated for Anh Hoang (hoang_anhkim@bah.com) (C) Splunk Inc, not for distribution

Reporting on Transactions

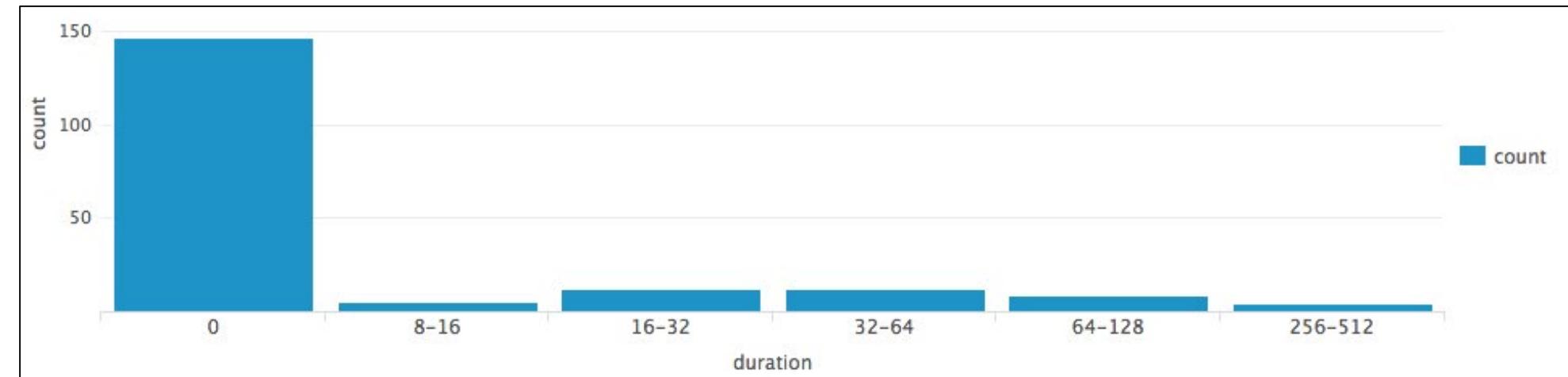
- You can use statistics and reporting commands with transactions
- In this example, a transaction is defined by events that share clientip and fit within a 10 minute span
- count() function counts number of transactions and separates the count by the duration of each

Scenario



Create a chart to show the number of purchase transactions based on their duration.

```
index=web sourcetype=access_combined  
status=200 action=purchase  
| transaction clientip maxspan=10m  
| chart count BY duration span=log2
```



Generated for Anh Hoang (hoang_anhkim@bah.com) (C) Splunk Inc, not for distribution

transaction vs. stats

- When you have a choice, use stats—it's faster and more efficient, especially in large Splunk environments
- Only use transaction when you:
 - Need to see events correlated together
 - Must define event grouping based on start/end values or segment on time
- Use stats when you:
 - Want to see the results of a calculation
 - Can group events based on a field value (e.g., by `src_ip`)
- By default, there's a limit of 1,000 events per transaction
 - No such limit applies to stats
 - Admins can change limit by configuring `max_events_per_bucket` in `limits.conf`

transaction vs. stats: Example 1

```
index=web  
sourcetype=access_combined  
earliest=-1y@y latest=@y  
| transaction JSESSIONID  
| table JSESSIONID,  
    action, product_name  
| sort JSESSIONID
```

A

```
index=web  
sourcetype=access_combined  
earliest=-1y@y latest=@y  
| stats values(action)  
    as "action",  
    values(product_name)  
    as "product_name"  
    by JSESSIONID  
| sort JSESSIONID
```

B

Scenario

Find online purchase transactions over the past year.

JSESSIONID	action	product_name
SD0SL10FF3ADFF4950	addtocart purchase remove view	Benign Space Debris Mediocre Kingdoms SIM Cubicle World of Cheese
SD0SL10FF6ADFF4954	remove	Benign Space Debris SIM Cubicle
SD0SL1FF1ADFF4954	addtocart purchase	Fire Resistance Suit of Provolone Manganiello Bros. Tee
SD0SL1FF1ADFF4955	addtocart purchase view	Benign Space Debris World of Cheese Tee

- Searches produce same result
- A took 23.381 seconds
- B took 2.077 seconds
- stats faster than transaction

transaction vs. stats: Example 2

A

```
index=security  
sourcetype=linux_secure failed  
| transaction src_ip  
| table src_ip, eventcount  
| sort - eventcount
```

Note

- 1. **transaction** has a limit of 1,000
- 2. Count of transactions vs. count of IPs

B

```
index=security  
sourcetype=linux_secure failed  
| stats count as eventcount  
| by src_ip  
| sort - eventcount
```

src_ip	eventcount
87.194.216.51	1000
211.166.11.101	840
128.241.220.82	662
194.215.205.19	617
87.194.216.51	134
236.45.28.248	133
229.156.63.239	133

- A took 6.163 seconds
- B took 4.643 seconds

src_ip	eventcount
87.194.216.51	1134
211.166.11.101	840
128.241.220.82	662
194.215.205.19	617
109.169.32.135	443
216.221.226.11	432
188.138.40.166	429

Generated for Anh Hoang (hoang_anhkim@bah.com) (C) Splunk Inc, not for distribution

Lab Exercise 5

- **Time:**

30 minutes

- **Tasks:**

- Analyze transactions in the online store during the last 60 minutes
- Display the online store purchase transactions lasting more than one minute and include the number of events in each transaction
- Search for online store transactions that begin with an addtocart action and end with a purchase action

****CHALLENGE Exercise:**

- Report common HTTP status errors that occurred during the last 30 days on the online sales web servers and the internal web appliance within a proximity of 30 seconds or less

Note



You may need to expand time on sourcetype=cisco_esa to the Last 24 hours.

Module 6: Introduction to Knowledge Objects

Generated for Anh Hoang (hoang_anhkim@bah.com) (C) Splunk Inc, not for distribution

Module Objectives

- Identify the categories of knowledge objects
- Define the role of a knowledge manager
- Identify naming conventions
- Review permissions
- Manage knowledge objects
- Describe the Splunk Common Information Model (CIM)

What are Knowledge Objects?

- Knowledge objects are tools you use to discover and analyze various aspects of your data
 - Data interpretation – Fields and field extractions
 - Data classification – Event types
 - Data enrichment – Lookups and workflow actions
 - Normalization – Tags and field aliases
 - Datasets – Data models



Generated for Anh Hoang (hoang_anhkim@bah.com) (C) Splunk Inc, not for distribution

What are Knowledge Objects? (cont.)

- Shareable
 - Can be shared between users
- Reusable
 - Persistent objects that can be used by multiple people or apps, such as macros and reports
- Usable in searches
 - Since the objects are persistent, they can be used in a search

Note



For more information go to:
<http://docs.splunk.com/Documentation/Splunk/latest/Knowledge/WhatIsSplunkknowledge>

What is a Knowledge Manager?

- Oversees knowledge object creation and usage for a group or deployment
- Normalizes event data
- Creates data models for Pivot users

Defining Naming Conventions

- This course uses simple names for lab exercises, but using a naming convention in your production environment is recommended. For example:
 - **Group:** Corresponds to the working group(s) of the user saving the object (examples: SEG. NEG. OPS. NOC)
 - **Object Type:** Indicates the type of object (alert, report, summary-index-populating) (examples: Alert, Report, Summary)
 - **Description:** A meaningful description of the context and intent of the search, limited to one or two words if possible; ensures the search name is unique
- Example: `SEG_Alert_WinEventlogFailures`

Note 

For more information go to:
<http://docs.splunk.com/Documentation/Splunk/latest/Knowledge/Developmentnamingconventionsforknowledgedobjecttitles>

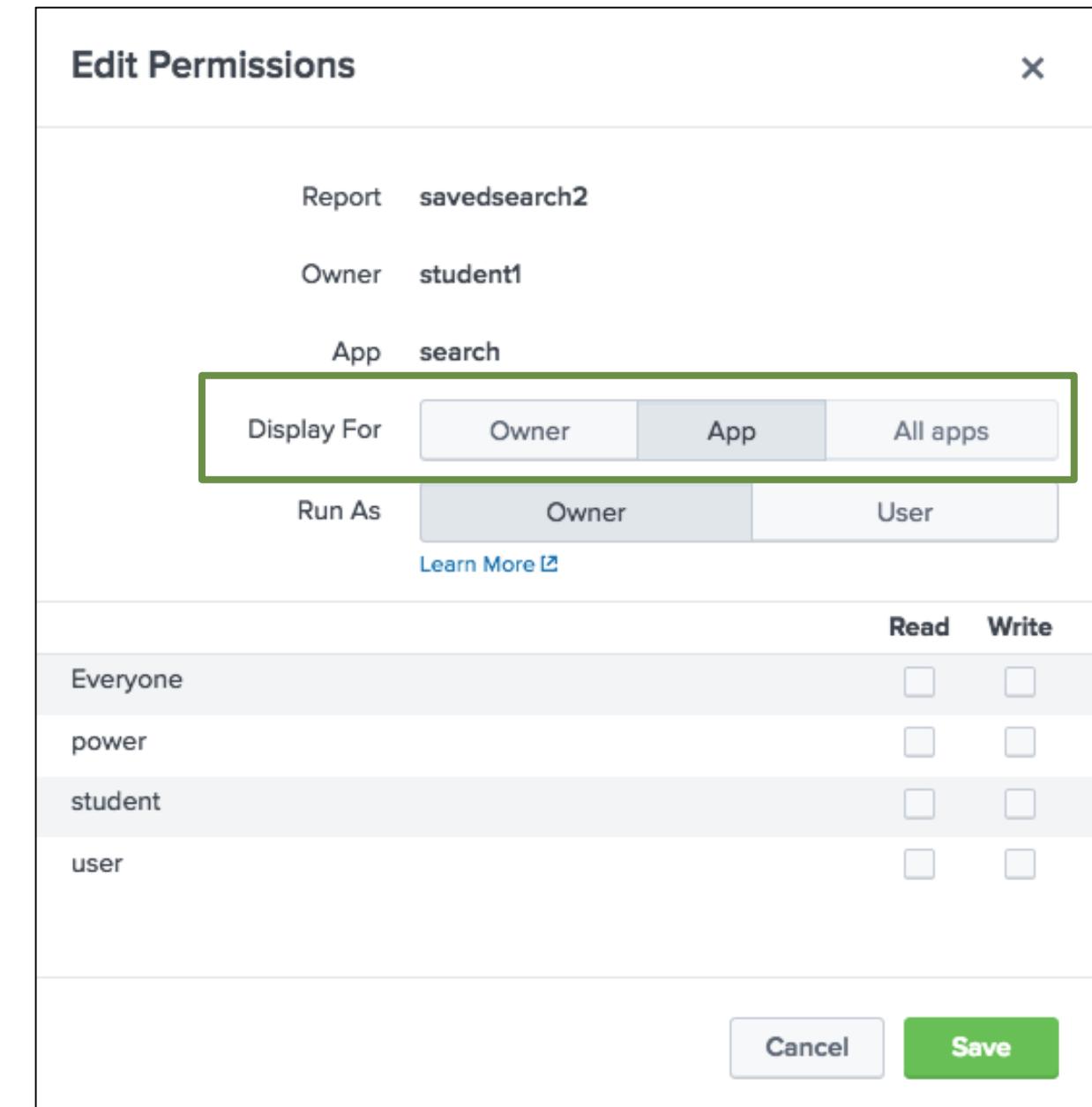
Reviewing Permissions

	Description	Create	Read	Edit (write)
Private	Only the person who created the object can use it and edit it	User Power Admin	Person who created it Admin*	Person who created it Admin
This app only	Object persists in the context of a specific app	Power Admin	User* Power* Admin	User* Power* Admin
All apps	Object persists globally across all apps	Admin	User* Power* Admin	User* Power* Admin

* Permission to read and/or write if creator gives permission to that role

Reviewing Permissions (cont.)

- When an object is created, Display For is set to **Owner** by default
- When object's permissions are set to **App** or **All apps**, any/all roles listed beneath can be given read permission
 - Write permission is reserved for admin and the object creator unless the creator edits permissions
- Only the admin role can promote an object to **All apps**
 - Other roles have **All Apps** button grayed out

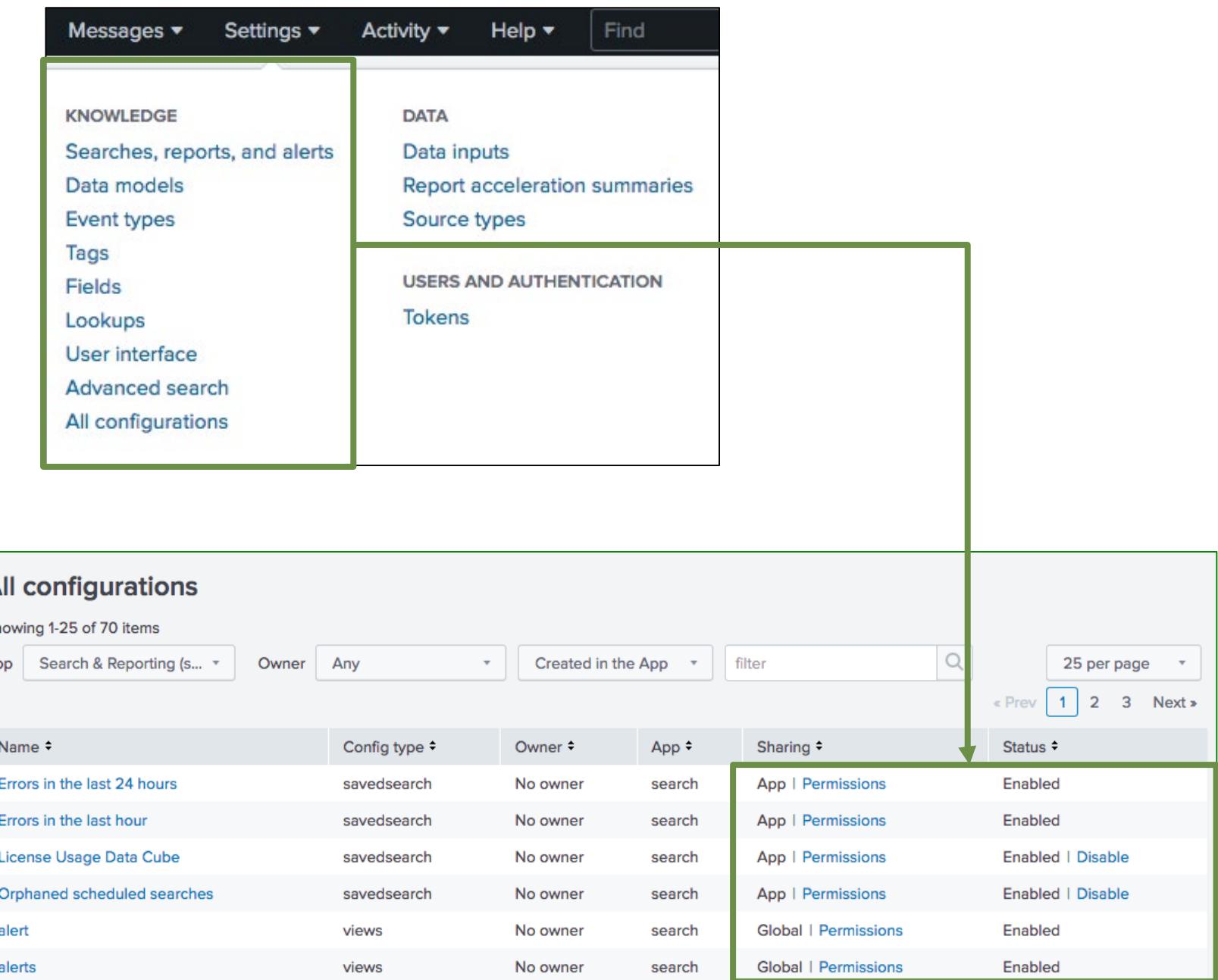


Generated for Anh Hoang (hoang_anhkim@bah.com) (C) Splunk Inc, not for distribution

Managing Knowledge Objects

- Knowledge objects are centrally managed from **Settings > Knowledge**
- Your role and permissions determine your ability to modify an object's settings

Note 
By default, objects for all owners are listed.



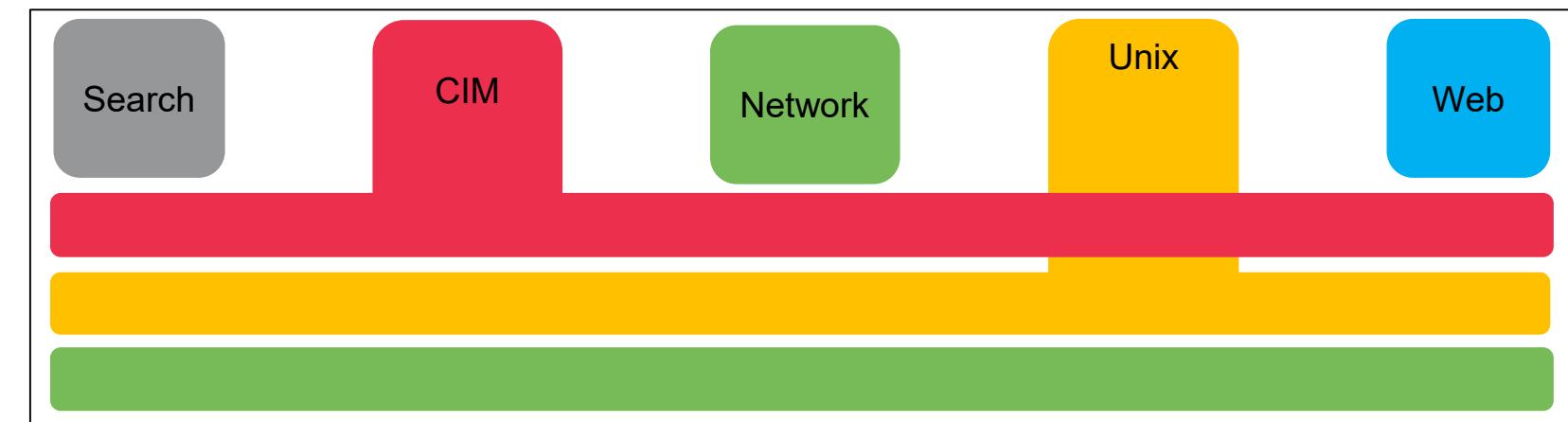
The screenshot shows the Splunk Settings > Knowledge interface. The left sidebar is titled 'KNOWLEDGE' and lists: Searches, reports, and alerts; Data models; Event types; Tags; Fields; Lookups; User interface; Advanced search; and All configurations. The main area is titled 'All configurations' and shows a table of 70 items. The table columns are: Name, Config type, Owner, App, Sharing, and Status. The table data includes:

Name	Config type	Owner	App	Sharing	Status
Errors in the last 24 hours	savedsearch	No owner	search	App Permissions	Enabled
Errors in the last hour	savedsearch	No owner	search	App Permissions	Enabled
License Usage Data Cube	savedsearch	No owner	search	App Permissions	Enabled Disable
Orphaned scheduled searches	savedsearch	No owner	search	App Permissions	Enabled Disable
alert	views	No owner	search	Global Permissions	Enabled
alerts	views	No owner	search	Global Permissions	Enabled

Generated for Anh Hoang (hoang_anhkim@bah.com) (C) Splunk Inc, not for distribution

Using the Splunk Common Information Model (CIM)

- Methodology for normalizing data
- Easily correlate data from different sources and source types
- Leverage to create various objects discussed in this course—field extractions, field aliases, event types, tags
- More details discussed in Module 13



Generated for Anh Hoang (hoang_anhkim@bah.com) (C) Splunk Inc, not for distribution

Module 7: Creating and Managing Fields

Generated for Anh Hoang (hoang_anhkim@bah.com) (C) Splunk Inc, not for distribution

Module Objectives

- Review the Field Extractor (FX) methods
 - Regex
 - Delimiter
- Identify the different options to get to the Field Extractor
 - Settings menu
 - Fields sidebar
 - Event Actions menu
- Review the process of extracting fields manually using regular expressions
- Use the Field Extraction Manager to modify extracted fields

Generated for Anh Hoang (hoang_anhkim@bah.com) (C) Splunk Inc, not for distribution

Field Auto-Extraction

- Splunk automatically discovers many fields based on source type and key/value pairs found in the data
- Prior to search time, some fields are already stored with the event in the index
 - Meta fields, such as `host`, `source`, and `sourcetype`
 - Internal fields such as `_time` and `_raw`
- At search time, *field discovery* extracts fields from raw event data, including those directly related to the search's results

Performing Field Extractions

- In addition to the many fields Splunk auto-extracts, you can also extract your own fields with the Field Extractor (FX)
- Use FX to extract fields that are static and that you use often in searches
 - Graphical UI
 - Extract fields from events using regex or delimiter
 - Extracted fields persist as knowledge objects
 - Can be shared and re-used in multiple searches
- Access FX via Settings, Fields Sidebar, or Event Actions menu

Note

For more information, see:
<http://docs.splunk.com/Documentation/Splunk/latest/Knowledge/ExtractfieldsinteractivelywithIFX>

Field Extraction Methods

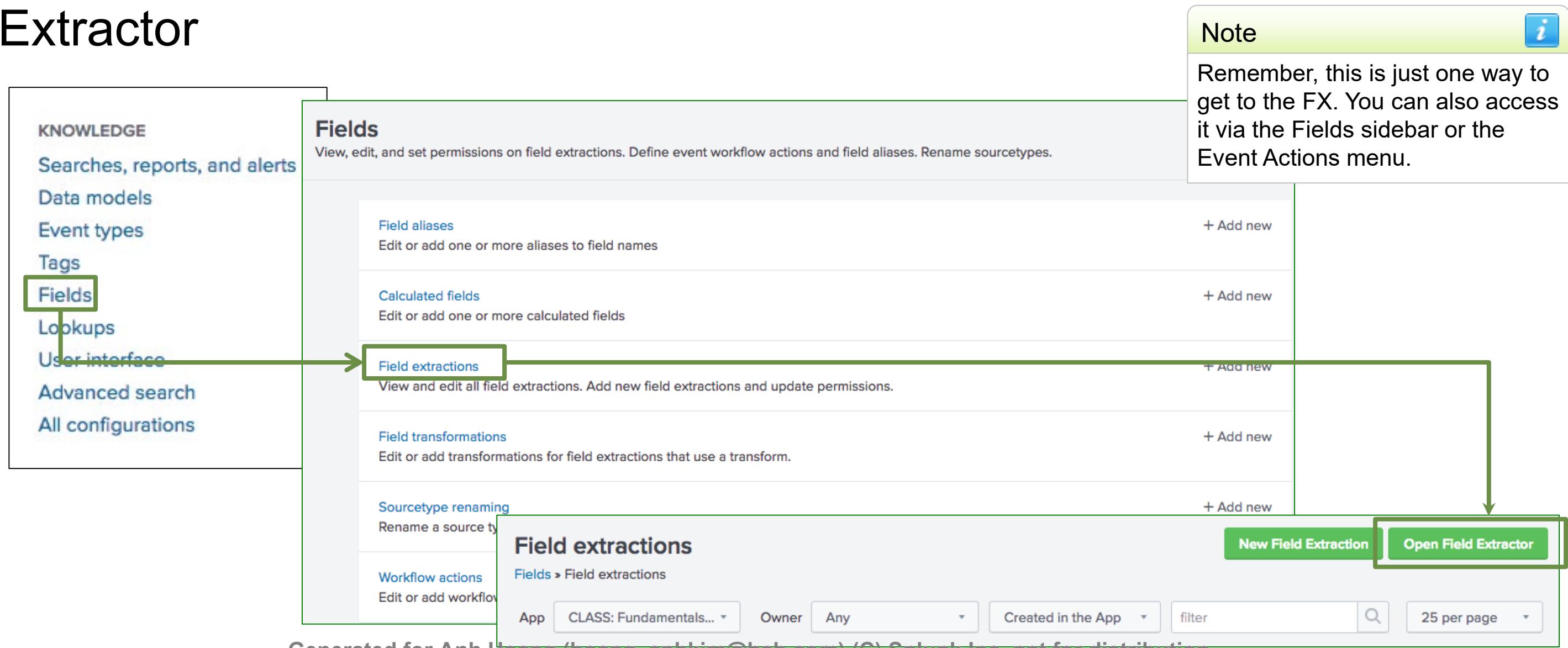
- **Regex**
 - Use this option when your event contains unstructured data like a system log file
 - FX attempts to extract fields using a Regular Expression that matches similar events
- **Delimiter**
 - Use this option when your event contains structured data like a .csv file
 - The data doesn't have headers and the fields must be separated by delimiters (spaces, commas, pipes, tabs, or other characters)

Field Extraction Workflows

- For either method (Regex or Delimiter), there are 3 ways to get to the Field Extractor:
 - Settings menu
 - Fields sidebar
 - Event Actions menu
- Now we'll walk through two example workflows from beginning to end
 - Example #1: Regex field extractions from Settings menu
 - Example #2: Delimiters field extractions from Event Actions menu

Workflow Example #1: Regex Field Extractions from Settings

Start by going to Settings > Fields > Field extractions > Open Field Extractor



Generated for Anh Hoang (hoang_anhkim@bah.com) (C) Splunk Inc, not for distribution

Regex Field Extractions from Settings (cont.)

1. Select the Data Type

- sourcetype
- source

2. Select the Source Type

Extract Fields Select Sample Select Method Select Fields Save Next >

Select Sample Event

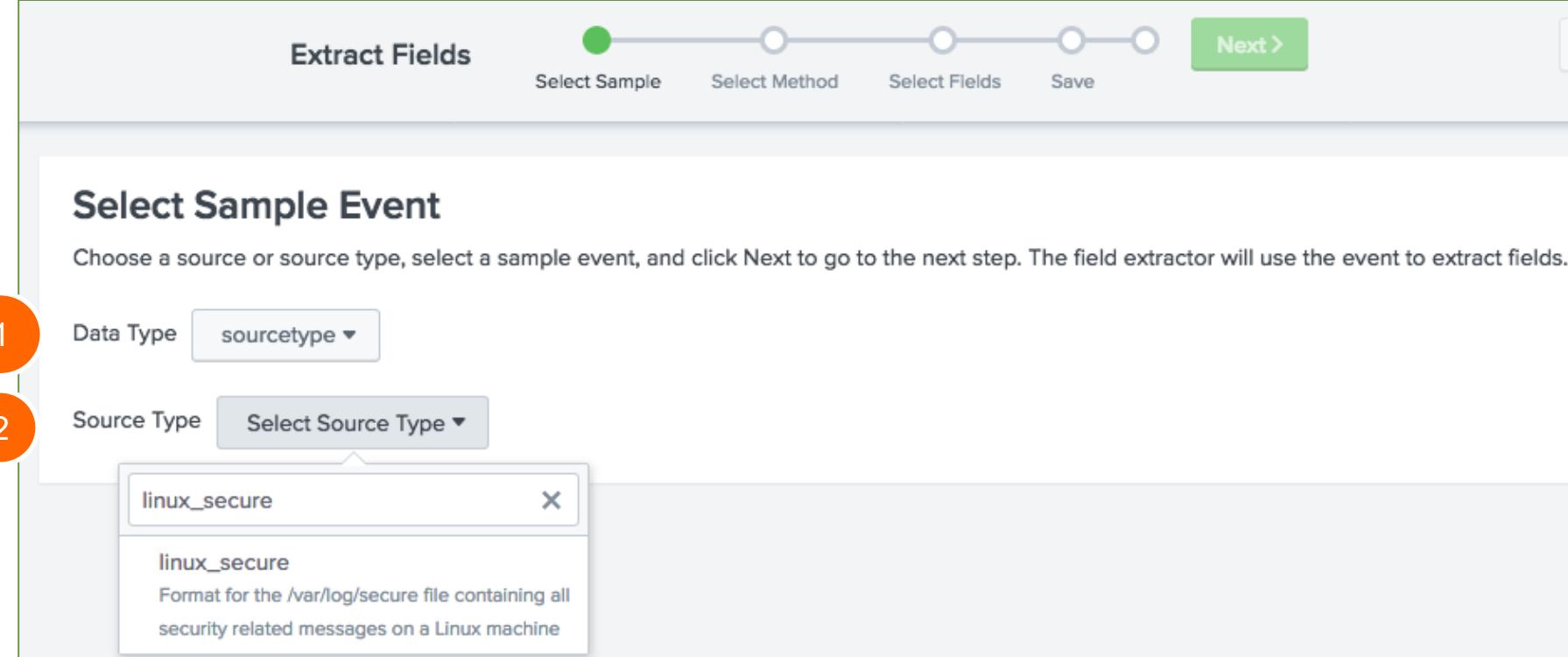
Choose a source or source type, select a sample event, and click Next to go to the next step. The field extractor will use the event to extract fields.

1 Data Type sourcetype ▾

2 Source Type Select Source Type ▾

linux_secure x

linux_secure
Format for the /var/log/secure file containing all security related messages on a Linux machine



Generated for Anh Hoang (hoang_anhkim@bah.com) (C) Splunk Inc, not for distribution

Regex Field Extractions from Settings – Select Sample

3. Select a sample event by clicking on it

4. Click Next >

The screenshot shows the 'Extract Fields' wizard in Splunk. The current step is 'Select Sample'. The interface includes:

- A top navigation bar with tabs: Extract Fields, Select Sample (highlighted with a green dot), Select Method, Select Fields, Save, and a 'Next >' button.
- An 'Existing fields >' link on the right.
- A main section titled 'Select Sample Event' with instructions: 'Choose a source or source type, select a sample event, and click Next to go to the next step. The field extractor will use the event to extract fields.' It also includes a link 'I prefer to write the regular expression myself >'.
- Configuration dropdowns for 'Data Type' (set to 'sourcetype') and 'Source Type' (set to 'linux_secure').
- A list of events under the heading 'Events'. The first event is highlighted with a blue background:

```
Thu Jan 18 2018 19:30:28 www3 sshd[3014]: Failed password for myuan from 187.60.191.199 port 2961 ssh2
```
- Below the event list are filters: 'filter' and 'Apply', and pagination controls: '20 per page', page numbers 1 through 9, and 'Next >'. There are also buttons for 'Sample: 1,000 events' and 'All events'.
- A preview section at the bottom shows raw log entries:

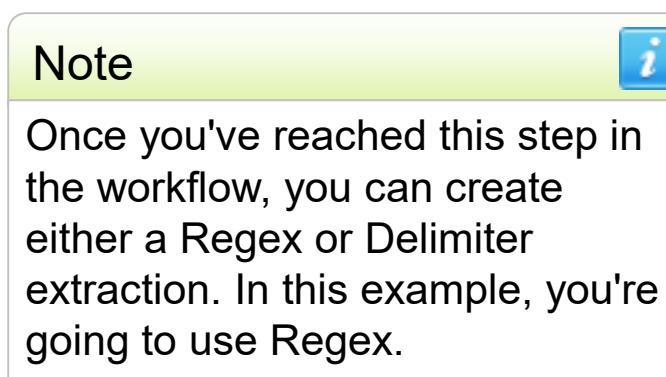
```
_raw ◀  
Thu Jan 18 2018 19:35:01 www3 sshd[3273]: Failed password for invalid user tavi from 128.241.220.82 port 4820 ssh2  
Thu Jan 18 2018 19:34:02 www3 sudo: nsharpe ; TTY=pts/0 ; PWD=/home/nsharpe ; USER=root ; COMMAND=/bin/su  
Thu Jan 18 2018 19:32:56 www3 sshd[2934]: Server listening on :: port 22.
```

Generated for Anh Hoang (hoang_anhkim@bah.com) (C) Splunk Inc, not for distribution

Regex Field Extractions from Settings – Select Method

5. Select Regular Expression

6. Click Next >



Extract Fields

Select Sample Select Method Select Fields Validate Save < Back Next > Existing fields > 6

Select Method

Indicate the method you want to use to extract your field(s). [Learn more](#) [I prefer to write the regular expression myself](#)

Source type `linux_secure`

Thu Jan 18 2018 19:30:28 www3 sshd[3014]: Failed password for myuan from 187.60.191.199 port 2961 ssh2

5 (.*)?
Regular Expression

Splunk Enterprise will extract fields using a Regular Expression.

6 x|y|z
Delimiters

Splunk Enterprise will extract fields using a delimiter (such as commas, spaces, or characters). Use this method for delimited data like comma separated values (CSV files).

Generated for Anh Hoang (hoang_anhkim@bah.com) (C) Splunk Inc, not for distribution

Regex Field Extractions from Settings – Select Values

7. Select the value(s) you want to extract. In this example, two fields are being extracted
8. Provide a field name
9. Click Add Extraction

The screenshot shows the 'Extract Fields' wizard at the 'Select Fields' step. A sample event is shown with several values highlighted in blue. A configuration panel on the right allows setting a field name and sample value, with a green 'Add Extraction' button. Three orange circles numbered 7, 8, and 9 point to the highlighted values, the configuration panel, and the 'Add Extraction' button respectively.

Note i

Require option – only events with the highlighted string are included in the extraction.

Regex Field Extractions from Settings – Preview

10. Preview the sample events

11. Click Next >

The screenshot shows the 'Extract Fields' wizard at step 11, 'Select Fields'. The interface includes a progress bar with five steps: 'Select Sample' (green), 'Select Method' (green), 'Select Fields' (orange), 'Validate' (light gray), and 'Save' (light gray). A red circle with the number '11' is overlaid on the 'Next >' button. The main area is titled 'Select Fields' with a sub-instruction: 'Highlight one or more values in the sample event to create fields. You can indicate one value is required, meaning it must exist in an event for the regular expression to match. Click on highlighted values in the sample event to modify them. To highlight text that is already part of an existing extraction, first turn off the existing extractions.' Below this is a sample event: 'Thu Jan 18 2018 20:00:39 www1 sshd[1085]: Failed password for invalid user oracle from 211.140.3.183 port 2081 ssh2'. Buttons for 'Show Regular Expression >' and 'View in Search' are present. A green box highlights the 'Preview' section, which contains a table of 1,000 events. The table has columns for '_raw', 'src', and 'port'. The first three rows of the table are shown:

_raw	src	port
✓ Thu Jan 18 2018 20:00:39 www3 sshd[1558]: Failed password for myuan from 187.60.191.199 port 2142 ssh2	187.60.191.199	2142
✓ Thu Jan 18 2018 20:00:39 www2 sshd[4904]: Failed password for invalid user carrie from 182.236.164.11 port 3838 ssh2	182.236.164.11	3838
✓ Thu Jan 18 2018 20:00:39 www2 sshd[2336]: Failed password for bin from 95.163.78.227 port 3555 ssh2	95.163.78.227	3555

Generated for Anh Hoang (hoang_anhkim@bah.com) (C) Splunk Inc, not for distribution

Regex Field Extractions from Settings – Validate

12. Validate the proper field values are extracted

13. Click Next >

The screenshot shows the 'Extract Fields' wizard with the 'Validate' step selected. The 'Events' tab is active, showing a list of 1,000 events. A specific event is highlighted with orange circles around the numbers 12 and 13, corresponding to the steps in the text. The event details are as follows:

_raw	src	port
Thu Jan 18 2018 20:00:39 www3 sshd[1558]: Failed password for myuan from 187.60.191.199 port 2142 ssh2	187.60.191.199	2142
Thu Jan 18 2018 20:00:39 www2 sshd[4904]: Failed password for invalid user carrie from 182.236.164.11 port 3838 ssh2	182.236.164.11	3838
Thu Jan 18 2018 20:00:39 www2 sshd[2336]: Failed password for bin from 95.163.78.227 port 3555 ssh2	95.163.78.227	3555

Generated for Anh Hoang (hoang_anhkim@bah.com) (C) Splunk Inc, not for distribution

Regex Field Extractions from Settings – Save

14. Review the name for the extraction knowledge object and set permissions

15. Click Finish

Extract Fields  

Save

Name the extraction and set permissions.

Extractions Name	EXTRACT- 	14	
Owner	student1		
App	class_Fund2		
Permissions	Owner	App	All apps

Source type linux_secure

Sample event Thu Jan 18 2018 20:00:39 www1 sshd[1085]: Failed password for invalid user oracle from 211.140.3.183
port 2081 ssh2

Fields src,port

Regular Expression `^(\\w+\\s+)+(\\d+\\s+)+\\d+:\\d+(\\d+\\s+)+\\w+(\\d+\\s+)+\\w+\\d+:\\s+(\\w+\\s+)+(?P<src>[^]+) port (?P<port>\\d+)`

Note 
An extractions name is provided by default. However, this name can be changed.

Generated for Anh Hoang (hoang_anhkim@bah.com) (C) Splunk Inc, not for distribution

Using the Extracted Fields

New Search

index=security sourcetype=linux_secure
| table src, port

Last 24 hours 

✓ 11,138 events (1/17/18 12:00:00.000 PM to 1/18/18 12:08:46.000 PM) No Event Sampling Job II ⚡ Smart Mode

Events Patterns Statistics (11,138) Visualization

20 Per Page Format Preview < Prev 1 2 3 4 5 6 7 8 9 ... Next >

src	port
91.205.40.22	1217
182.236.164.11	2101
27.101.11.11	1329
29.11.106.88	8137
60.18.93.11	3259
221.207.229.6	2169
142.162.221.28	4998
91.205.40.22	1636
188.143.232.202	3320

Generated for Anh Hoang (hoang_anhkim@bah.com) (C) Splunk Inc, not for distribution

Editing Regex for Field Extractions

1. From Select Method, click Regular Expression

2. Click Next >

The screenshot shows the 'Extract Fields' wizard at the 'Select Method' step. The steps are: Select Sample (green), Select Method (green), Select Fields (light gray), Validate (light gray), Save (light gray). The 'Select Method' step is active. The 'Source type' is set to 'linux_secure'. A log entry is shown: 'Fri Mar 16 2018 16:16:36 www2 sshd[4004]: Failed password for invalid user sys from [223.205.219.198] port [3148] ssh2'. Two extraction methods are displayed: 'Regular Expression' (highlighted with a blue box and circled with a red number 1) and 'Delimiters' (circled with a red number 2). The 'Regular Expression' method uses the pattern '(.*?)'.

Generated for Anh Hoang (hoang_anhkim@bah.com) (C) Splunk Inc, not for distribution

Editing Regex for Field Extractions – Select Field

3. Select the field to extract
4. Provide a Field Name
5. Click Add Extraction

Note

For more information about Splunk Regular Expressions, see:
<http://docs.splunk.com/Documentation/Splunk/latest/Knowledge/AboutSplunkregularexpressions>

Extract Fields ● Select Sample ● Select Method ● Select Fields ○ Validate ○ Save < Back Next > Existing fields >

Select Fields

Highlight one or more values in the sample event to create fields. You can indicate one value is required, meaning it must exist in an event for the regular expression to match. Click on highlighted values in the sample event to modify them. To highlight text that is already part of an existing extraction, first turn off the existing extractions. [Learn more](#)

Fri Mar 16 2018 16:16:36 www2 sshd[4004]: Failed password for invalid user sys from 223.205.219.198 port 3148 ssh2

Extract Require

Field Name user 3

Sample sys
Value

5 Add Extraction

Generated for Anh Hoang (hoang_anhkim@bah.com) (C) Splunk Inc, not for distribution

Editing Regex for Field Extractions – Show Regex

6. Click Show Regular Expression >

7. Click Edit the Regular Expression

Extract Fields

Select Sample Select Method Select Fields Validate Save < Back Next > Existing fields >

Select Fields

Highlight one or more values in the sample event
Click on highlighted values in the sample event

```
Thu Jan 18 2018 23:04:34 www3 sshd[2888]
```

Show Regular Expression >

6

Extract Fields

Select Sample Select Method Select Fields Validate Save < Back Next > Existing fields >

Select Fields

Highlight one or more values in the sample event to create fields. You can indicate one value is required, meaning it must exist in an event for the regular expression to match. Click on highlighted values in the sample event to modify them. To highlight text that is already part of an existing extraction, first turn off the existing extractions. [Learn more](#)

```
Fri Mar 16 2018 16:16:36 www2 sshd[4004]: Failed password for invalid user sys from 223.205.219.198 port 3148 ssh2
```

Hide Regular Expression > View in Search

```
^\\w+\\s+\\w+\\s+\\d+\\s+\\d+:\\d+:\\d+:\\d+\\s+\\w+\\d+\\s+\\w+\\|\\d+\\|:\\s+\\w+\\s+\\w+\\s+\\w+\\s+\\w+\\s+\\w+\\s+(?P<user>[^ ]+)
```

7 Edit the Regular Expression

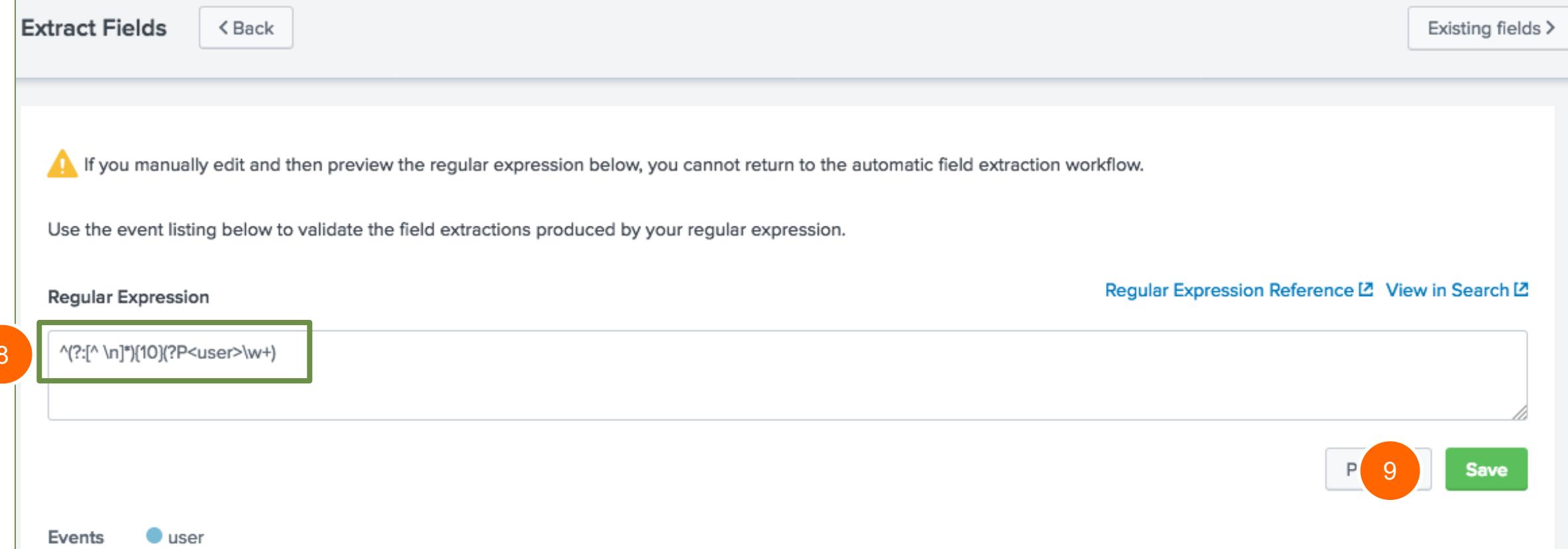
Generated for Anh Hoang (hoang_anhkim@bah.com) (C) Splunk Inc, not for distribution

Editing Regex for Field Extractions – Modify Regex

8. Update the regular expression

9. Click Save

Warning 
After you edit the regular expression, you cannot go back to the Field Extractor UI.



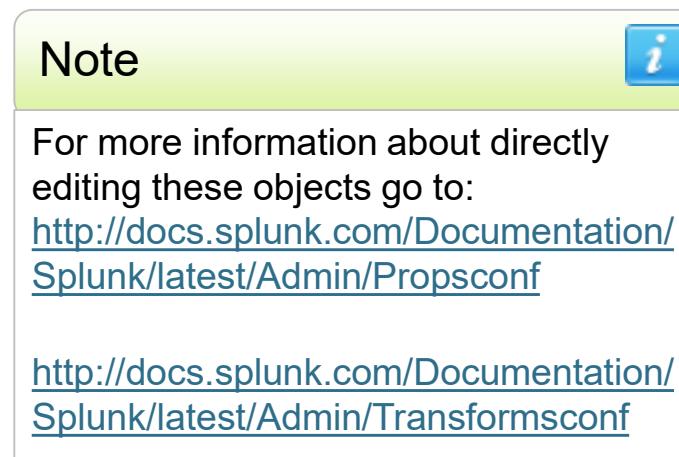
The screenshot shows the 'Extract Fields' interface. At the top left is the title 'Extract Fields' and a 'Back' button. On the right are 'Existing fields >' and a 'Save' button. A warning message at the top says: '⚠ If you manually edit and then preview the regular expression below, you cannot return to the automatic field extraction workflow.' Below it, a note says: 'Use the event listing below to validate the field extractions produced by your regular expression.' A 'Regular Expression' input field contains the value '^(?:[^ \n]*(10)(?P<user>\w+))'. To the right of this field are links to 'Regular Expression Reference' and 'View in Search'. A green box highlights the regular expression input field. A red circle with the number '8' is positioned to the left of the input field. At the bottom right is a 'Save' button with a red circle containing the number '9'. At the bottom left are buttons for 'Events' and 'user'.

Generated for Anh Hoang (hoang_anhkim@bah.com) (C) Splunk Inc, not for distribution

Editing Regex for Field Extractions - Save

10. Review the Extractions Name and set permissions

11. Click >Finish



Extract Fields < Back > Finish 11

Save

Name the extraction and set permissions.

10 Extractions Name EXTRACT-user

Owner student2

App search

Permissions Owner App All apps

Source type linux_secure

Fields user

Regular Expression `^(?:[^ \n]*)(10)(?P<user>\w+)`

Generated for Anh Hoang (hoang_anhkim@bah.com) (C) Splunk Inc, not for distribution

Delimited Field Extractions

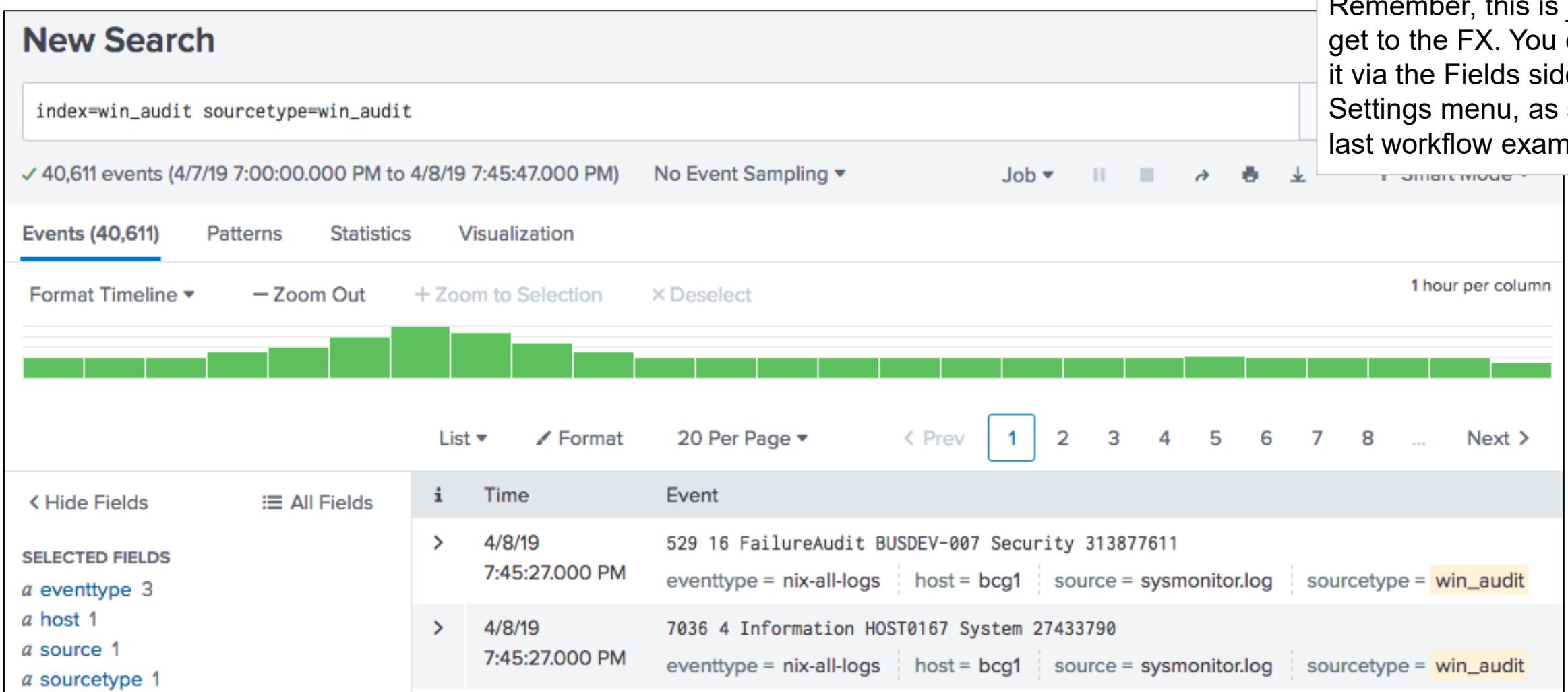
- Use delimited field extractions when the event log does not have a header and fields are separated by spaces, commas, or characters
- In this example, the fields are separated by commas

i	Time	Event
>	1/30/18 10:43:39.000 PM	"2018-01-30T22:43:39000-0400",29,1>Error,HOST0167,System,772103058 host = adldapsv1 source = /opt/log/adldapsv1/sysmonitor.log sourcetype = win_audit
>	1/30/18 10:43:37.000 PM	"2018-01-30T22:43:37000-0400",35,4,Information,HOST0201,System,507701378 host = adldapsv1 source = /opt/log/adldapsv1/sysmonitor.log sourcetype = win_audit
>	1/30/18 10:43:36.000 PM	"2018-01-30T22:43:36000-0400",35,4,Information,HOST0201,System,753380719 host = adldapsv1 source = /opt/log/adldapsv1/sysmonitor.log sourcetype = win_audit

Generated for Anh Hoang (hoang_anhkim@bah.com) (C) Splunk Inc, not for distribution

Workflow Example #2: Delimited Field Extractions from Event Actions

Start by searching against the sourcetype from which you want to extract fields



The screenshot shows a Splunk search interface with the following details:

- Search Bar:** index=win_audit sourcetype=win_audit
- Results Summary:** 40,611 events (4/7/19 7:00:00.000 PM to 4/8/19 7:45:47.000 PM) | No Event Sampling
- Event View:** Events (40,611) | Patterns | Statistics | Visualization
- Timeline:** Format Timeline | Zoom Out | Zoom to Selection | Deselect | 1 hour per column
- Pagination:** List | Format | 20 Per Page | < Prev | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | ... | Next >
- Selected Fields:** eventtype 3, host 1, source 1, sourcetype 1
- Event List:** Two events are visible:
 - 4/8/19 7:45:27.000 PM | 529 16 FailureAudit BUSDEV-007 Security 313877611 | eventtype = nix-all-logs | host = bcg1 | source = sysmonitor.log | sourcetype = win_audit
 - 4/8/19 7:45:27.000 PM | 7036 4 Information HOST0167 System 27433790 | eventtype = nix-all-logs | host = bcg1 | source = sysmonitor.log | sourcetype = win_audit

Note



Remember, this is just one way to get to the FX. You can also access it via the Fields sidebar or the Settings menu, as shown in the last workflow example.

Generated for Anh Hoang (hoang_anhkim@bah.com) (C) Splunk Inc, not for distribution

Delimited Field Extractions from Event Actions (cont.)

1. Expand an event, and from the event details select **Event Actions > Extract Fields**

The screenshot shows the Splunk search interface with a search bar containing "index=win_audit sourcetype=win_audit" and a time range of "Last 24 hours". The results section shows 39,617 events. The "Events (39,617)" tab is selected. A specific event is expanded, showing fields like Time, Event, and a detailed view of the event content. An "Event Actions" dropdown menu is open over the event details, with the "Extract Fields" option highlighted. A red circle with the number "1" is drawn around this menu item.

Generated for Anh Hoang (hoang_anhkim@bah.com) (C) Splunk Inc, not for distribution

Delimited Field Extractions (Event Actions) – Select Method

2. Select Delimiters

3. Click Next >

Note

Once you've reached this step in the workflow, you can create either a Regex or Delimiter extraction. In this example, you're going to use Delimiter.

The screenshot shows the 'Extract Fields' workflow interface. The top navigation bar has steps: 'Extract Fields' (selected), 'Select Sample' (green dot), 'Select Method' (green dot), 'Rename Fields' (grey dot), and 'Save' (grey dot). A 'Next >' button is on the right, with a red circle containing the number '3' above it, indicating the current step.

Select Method

Indicate the method you want to use to extract your field(s). [Learn more](#) [I prefer to write the regular expression myself](#)

Source type `win_audit`

"2018-01-30T23:38:40000-0400",7036,4,Information,HOST0167,System,296651410

Regular Expression
Splunk Enterprise will extract fields using a Regular Expression.
(.*)?

Delimiters
Splunk Enterprise will extract fields using a delimiter (such as commas, spaces, or characters). Use this method for delimited data like comma separated values (CSV files).
2 x|y|z

Generated for Anh Hoang (hoang_anhkim@bah.com) (C) Splunk Inc, not for distribution

Delimited Field Extractions (Event Actions) – Select Delimiter

4. Select the Delimiter used in your event

4

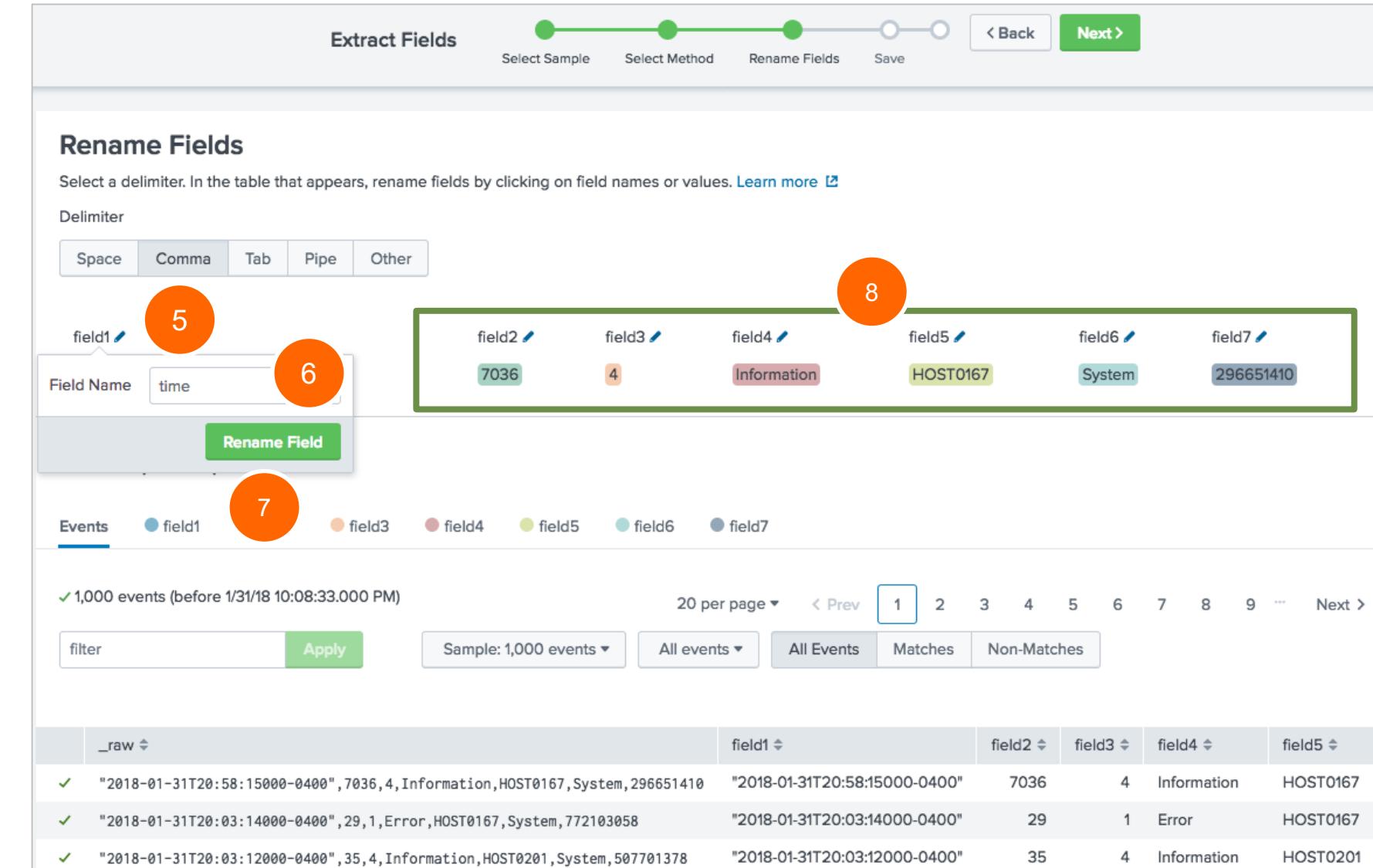
The screenshot shows the 'Extract Fields' process in Splunk. The current step is 'Rename Fields'. The interface includes a navigation bar with 'Select Sample', 'Select Method', 'Rename Fields', and 'Save' steps. Below the navigation is a 'Rename Fields' section with a subtitle: 'Select a delimiter. In the table that appears, rename fields by clicking on field names or values.' A 'Learn more' link is provided. A 'Delimiter' tab is selected, showing tabs for 'Space', 'Comma', 'Tab', 'Pipe', and 'Other'. The main area displays a table with seven columns: field1, field2, field3, field4, field5, field6, and field7. The first column contains the value "2018-01-31T20:58:15000-0400". The second column contains the value 7036. The third column contains the value 4. The fourth column contains the value Information. The fifth column contains the value HOST0167. The sixth column contains the value System. The seventh column contains the value 296651410. Below the table is a 'Preview (7 fields)' section. It shows a legend where 'Events' is highlighted in blue, followed by field1 (light blue), field2 (light green), field3 (light orange), field4 (light pink), field5 (light yellow), field6 (light teal), and field7 (light purple). It also shows a message: '✓ 1,000 events (before 1/31/18 10:08:33.000 PM)'. There are buttons for 'filter' and 'Apply', and dropdowns for 'Sample: 1,000 events', 'All events', 'All Events', 'Matches', and 'Non-Matches'. A page navigation bar shows pages 1 through 9.

_raw	field1	field2	field3	field4	field5	field6	field7
"2018-01-31T20:58:15000-0400",7036,4,Information,HOST0167,System,296651410	"2018-01-31T20:58:15000-0400"	7036	4	Information	HOST0167	System	296651410
"2018-01-31T20:03:14000-0400",29,1>Error,HOST0167,System,772103058	"2018-01-31T20:03:14000-0400"	29	1	Error	HOST0167		
"2018-01-31T20:03:12000-0400",35,4,Information,HOST0201,System,507701378	"2018-01-31T20:03:12000-0400"	35	4	Information	HOST0201		
"2018-01-31T20:03:11000-0400",35,4,Information,HOST0201,System,753380719	"2018-01-31T20:03:11000-0400"	35	4	Information	HOST0201		

Generated for Anh Hoang (hoang_anhkim@bah.com) (C) Splunk Inc, not for distribution

Delimited Field Extractions (Event Actions) – Rename Field

5. Click the  icon next to the default field name
6. Enter a new field name
7. Click Rename Field
8. Repeat these steps for all fields



The screenshot shows the 'Extract Fields' interface in Splunk, specifically the 'Rename Fields' step. At the top, a progress bar indicates the current step is 'Rename Fields'. Below the progress bar, there's a section titled 'Rename Fields' with instructions to select a delimiter and rename fields by clicking on field names or values. A 'Delimiter' dropdown menu is open, showing options: Space, Comma, Tab, Pipe, and Other. The 'Other' option is selected. A table below lists field names and their current values. The first row, 'field1', has its value 'time' highlighted with a red box and a red circle containing the number 6. The second row, 'field2', has its value '7036' highlighted with a red box and a red circle containing the number 5. The third row, 'field3', has its value '4' highlighted with a red box and a red circle containing the number 7. The fourth row, 'field4', has its value 'Information' highlighted with a red box and a red circle containing the number 8. The fifth row, 'field5', has its value 'HOST0167' highlighted with a red box and a red circle containing the number 8. The sixth row, 'field6', has its value 'System' highlighted with a red box and a red circle containing the number 8. The seventh row, 'field7', has its value '296651410' highlighted with a red box and a red circle containing the number 8. Below the table, there's a legend for 'Events' showing colored dots corresponding to the field names: field1 (blue), field3 (orange), field4 (pink), field5 (yellow-green), field6 (teal), and field7 (dark blue). At the bottom, there's a summary of 1,000 events before 1/31/18 10:08:33.000 PM, a filter input, and a table preview.

	_raw	field1	field2	field3	field4	field5	field6	field7
✓	"2018-01-31T20:58:15000-0400",7036,4,Information,HOST0167,System,296651410	"2018-01-31T20:58:15000-0400"	7036	4	Information	HOST0167		
✓	"2018-01-31T20:03:14000-0400",29,1>Error,HOST0167,System,772103058	"2018-01-31T20:03:14000-0400"	29	1	Error	HOST0167		
✓	"2018-01-31T20:03:12000-0400",35,4,Information,HOST0201,System,507701378	"2018-01-31T20:03:12000-0400"	35	4	Information	HOST0201		

Generated for Anh Hoang (hoang_anhkim@bah.com) (C) Splunk Inc, not for distribution

Delimited Field Extractions (Event Actions) – Rename Field (cont.)

9. After all the fields are renamed, click Next >

The screenshot shows the 'Extract Fields' wizard in Splunk. The current step is 'Rename Fields'. At the top right, there is a green 'Next >' button with a large orange circle around it, indicating the next action. The 'Rename Fields' section contains a table with columns: time, eventcode, eventtype, type, computername, logname, and recordnumber. Each column has its current value displayed below it. The 'Preview (7 fields)' section shows a table of 1,000 events with the same seven columns and their corresponding values.

_raw	time	eventcode	eventtype	type	computername	logname	recordnumber
"2018-01-31T20:58:15000-0400"	7036	4	Information	HOST0167	System	296651410	
"2018-01-31T20:03:14000-0400"	29	1	Error	HOST0167	System	772103058	
"2018-01-31T20:03:12000-0400"	35	4	Information	HOST0201	System	507701378	

Generated for Anh Hoang (hoang_anhkim@bah.com) (C) Splunk Inc, not for distribution

Delimited Field Extractions (Event Actions) – Save

10. Name your extraction

11. Click Finish>

Extract Fields

Select Sample Select Method Rename Fields Save

11

Finish >

Save

Name the extraction and set permissions.

Extractions Name	REPORT- sysmon	10	
Owner	alifeson		
App	search		
Permissions	Owner	App	All apps

Source type win_audit

Sample event "2018-01-30T23:38:40000-0400",7036,4,Information,HOST0167,System,296651410

Fields time,eventcode,eventtype,type,computername,logname,recordnumber

Delimiter comma

Using a Delimited Field Extraction

New Search

index=_* OR index=* sourcetype=win_audit type=failureaudit

Last 24 hours

✓ 16 events (1/30/18 6:00:00.000 PM to 1/31/18 6:12:00.000 PM) No Event Sampling

Job

Events (16) Patterns Statistics Visualization

Format Timeline

1/30/18 10:40:39.000 PM "2018-01-30T22:40:39000-0400",529,16,FailureAudit,"BUSDEV-007",Security,319037447

Event Actions

Type	Field	Value	Actions
Selected	host	adldapsv1	
	source	/opt/log/adldapsv1/sysmonitor.log	
	sourcetype	win_audit	
Event	computername	BUSDEV-007	
	eventcode	529	
	eventtype	nix-all-logs	
	logname	Security	
	recordnumber	319037447	
	time	2018-01-30T22:40:39000-0400	
	type	FailureAudit	

Time _time 2018-01-30T22:40:39.000+00:00

List Format 20 Per P

< Hide Fields All Fields

SELECTED FIELDS

- a host 1
- a source 1
- a sourcetype 1

INTERESTING FIELDS

- a computername 7
- # date_hour 5
- # date_mday 1
- # date_minute 6
- a date_month 1
- # date_second 6
- a date_wday 1
- # date_year 1
- a date_zone 1
- # eventcode 3

Time _time 2018-01-30T22:40:39.000+00:00

Generated for Anh Hoang (hoang_anhkim@bah.com) (C) Splunk Inc, not for distribution

Module 7 Lab Exercise

Time: 25 minutes

Tasks:

- Using the regex field extraction method, extract the IP address and port fields from the `linux_secure` logs
- Using the delimiter field extraction method, extract and name fields from the `SimCubeBeta` sourcetype

Module 8: Creating Field Aliases and Calculated Fields

Generated for Anh Hoang (hoang_anhkim@bah.com) (C) Splunk Inc, not for distribution

Module Objectives

- Create and use field aliases
- Create calculated fields

Field Aliases

- A way to normalize data over any default field (per host, source or sourcetype)
- Multiple aliases can be applied to one field
- Applied after field extractions, before lookups
- Can be referenced by a lookup

Field Alias Example

- Several source types contain some type of a username field
- To make data correlation and searching easier, normalize the username field

The screenshot shows a Splunk search interface with three main sections corresponding to different source types:

- sourcetype=cisco_firewall**: Shows a single event with a 'Username' field value of 'dhale'. A green box highlights this field.
- sourcetype=cisco_wsa_squid**: Shows a single event with a 'username' field value of 'dhale'. A green box highlights this field.
- sourcetype=winauthentication_security**: Shows a single event with a 'User' field value of 'dhale'. A green box highlights this field.

Each section has a dropdown menu showing other fields and their values. Arrows point from the highlighted 'username' fields in the first two sections up to the corresponding 'User' field in the third section, illustrating how a single alias ('User') is used to represent the same data across different source types.

Selected	host	cisco_router1
Event	source	/opt/log/cisco_router1/cisco_firewall.log
	sourcetype	cisco_firewall
	Duration	9h:21m:4s
	Group	buttercupgames
	IP	10.3.10.241
	Username	dhale
	bcg_ip	10.3.10.241
	bcg_workstation	BG03-dhale

severity	5-None
splunk_role	power
src	123.196.113.11
src_ip	123.196.113.11
status	200
url	http://www.uncw.edu /www/screenNewMast.css
usage	Personal
username	dhale
x_acltag	DEFAULT_CASE-DefaultGroup-Demo_Clients-NONE-NONE-DefaultRouting

ComputerName	BG03-dhale
EventCode	4634
EventType	8
LogName	Security
Message	Successful
RecordNumber	9787
Sid	S-1-5-21-57989841-920026266-725345543-6444
SidType	1
SourceName	Security
Type	Success
User	dhale

Generated for Anh Hoang (hoang_anhkim@bah.com) (C) Splunk Inc, not for distribution

Creating a Field Alias

Settings > Fields > Field Aliases > New Field Alias

1. Select the app associated with the field alias
2. Enter a Name for the field alias
3. Apply the field alias to a default field (either host, source, or sourcetype)
4. Enter the name for the existing field and the new alias

Add new
Fields > Field aliases > Add new

1 Destination app class_Fund2

2 Name * cisco_firewall_aliases

3 Apply to sourcetype named * cisco_firewall

4 Field aliases Username = user

5 + Add another field Overwrite field values

existing field name new field alias Save

The screenshot shows the 'Add new' field alias configuration screen. It includes fields for 'Destination app' (set to 'class_Fund2'), 'Name' (set to 'cisco_firewall_aliases'), 'Apply to' (set to 'sourcetype' with 'named' set to 'cisco_firewall'), and a 'Field aliases' section. In the 'Field aliases' section, there is a row for 'Username' with a value of 'user'. A callout with five numbered circles (1-5) points to various fields: 1 points to 'Destination app', 2 points to 'Name', 3 points to 'Apply to', 4 points to the 'Field aliases' section, and 5 points to the 'Overwrite field values' checkbox. Two green arrows point upwards from labels 'existing field name' and 'new field alias' to the 'Username' field and its value 'user' respectively. A 'Save' button is located at the bottom right.

5. Optionally, select Overwrite field values
 - If alias name already exists, replaces alias field name with original name
 - If original field doesn't exist or has no value, removes alias field name

Creating a Field Alias (cont.)

In this example, one field alias is used for new ‘user’ fields in multiple source types

New field alias required for each sourcetype

The screenshots illustrate the creation of three field aliases:

- cisco_firewall_aliases**: Applied to sourcetype `cisco_firewall`. Field alias: `Username`.
- cisco_wsa_squid_aliases**: Applied to sourcetype `cisco_wsa_squid`. Field alias: `username`.
- winauthentication_security_aliases**: Applied to sourcetype `winauthentication_security`. Field alias: `User`.

Each configuration includes a dropdown for 'Destination app' (set to `class_Fund2`), a required 'Name' field, an 'Apply to' dropdown set to 'sourcetype', and a 'named' field containing the sourcetype name. There is also a checkbox for 'Overwrite field values'.

Generated for Anh Hoang (hoang_anhkim@bah.com) (C) Splunk Inc, not for distribution

Testing the Field Alias

After the field alias is created, perform a search using the new field alias

New Search

user=dhale Last 7 days

✓ 96 events (1/18/18 11:00:00.000 AM to 1/25/18 11:27:08.000 AM) No Event Sampling

Events (96) Patterns Statistics Visualization

Format Timeline

sourcetype

3 Values, 100% of events Yes No

Reports

Top values Top values by time Rare values

Events with this field

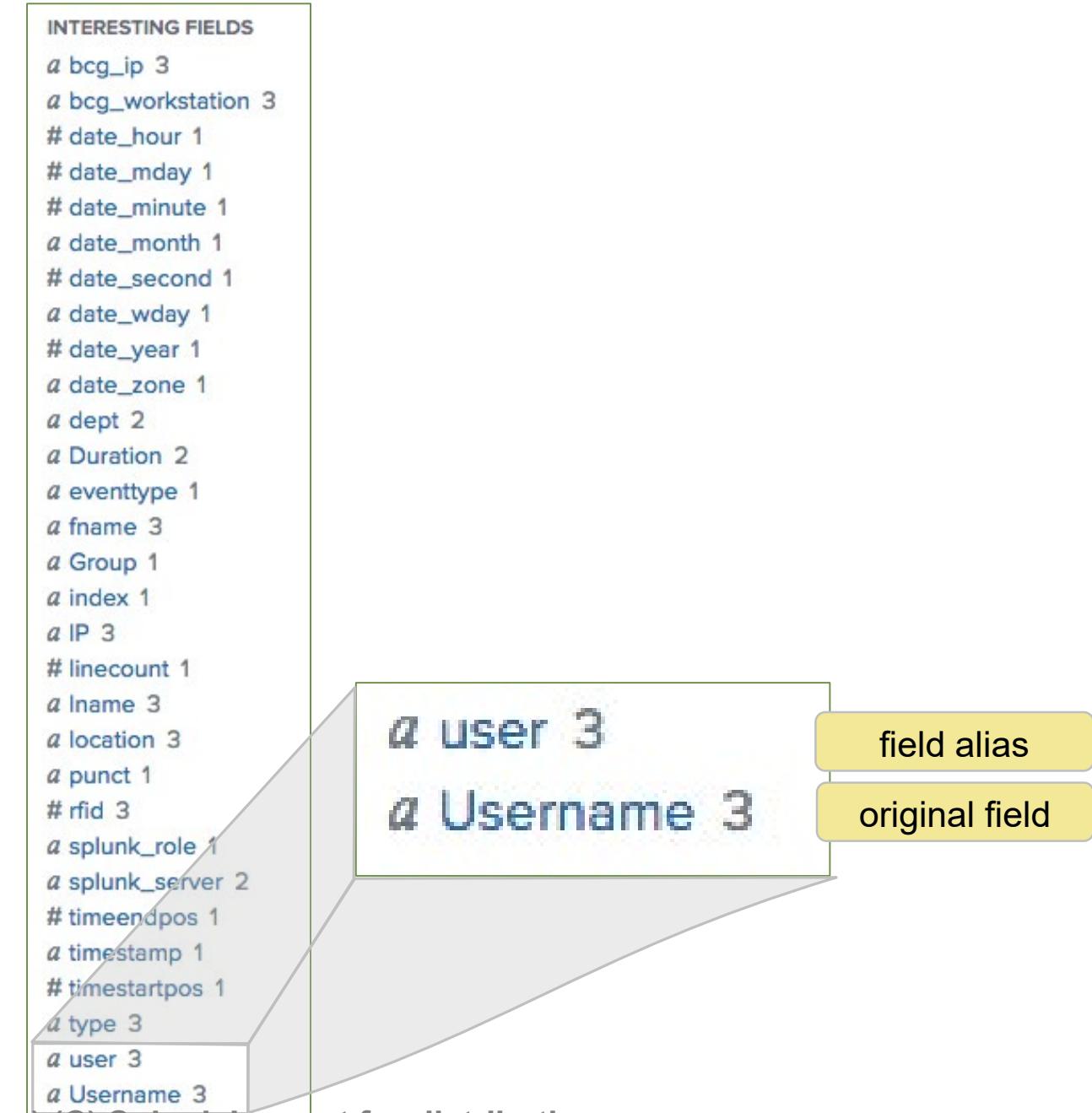
Values

	Count	%
cisco_wsa_squid	78	81.25%
winauthentication_security	16	16.667%
cisco_firewall	2	2.083%

Generated for Anh Hoang (hoang_anhkim@bah.com) (C) Splunk Inc, not for distribution

Field Alias and Original Fields

- When you create a field alias, the original field is not affected
- Both fields appear in the All Fields and Interesting Fields lists, if they appear in at least 20% of events



Generated for Anh Hoang (hoang_anhkim@bah.com) (C) Splunk Inc, not for distribution

Field Aliases and Lookups

After you have defined your field aliases, you can reference them in a lookup table

The screenshot shows the Splunk Field Alias configuration interface on the left and a CSV file named 'employees.csv' on the right.

Field aliases:

- Username = user
- + Add another field
- Overwrite field values

employees.csv:

rfid	fname	lname	user	email	dept	location	ip
108423575302	Allen	Pucci	apucci	apucci@buttercupgames.com	Sales	Boston	10.3.10.53
672903009231	Dwight	Hale	dhale	dhale@buttercupgames.com	Sales	Boston	10.3.10.241
398009643042	Phyllis	Bunch	pbunch	pbunch@buttercupgames.com	ITOps	Boston	10.3.10.227
374765319282	Enrique	Maxwell	emaxwell	emaxwell@buttercupgames.com	ITOps	Boston	10.3.10.46
227128834140	David	Johnson	djohnson	djohnson@buttercupgames.com	Engineering	Boston	10.3.10.180
371211812887	Galina	Zuyeva	gzuyeva	gzuyeva@buttercupgames.com	Engineering	Boston	10.3.10.67
249772079712	Louis	Sagers	lsagers	lsagers@buttercupgames.com	SecOps	Boston	10.3.10.21
417852300683	Amanda	Curry	acurry	acurry@buttercupgames.com	SecOps	San Francisco	10.1.10.252
542830538161	Alan	Dombrowski	adombrowski	adombrowski@buttercupgames.com	SecOps	San Francisco	10.1.10.129
768166372290	Cerys	Farrell	cfarrell	cfarrell@buttercupgames.com	Sales	San Francisco	10.1.10.107
153218951159	Placido	Toscani	ptoscani	ptoscani@buttercupgames.com	Sales	San Francisco	10.1.10.38
994499284304	Ian	King	iking	iwing@buttercupgames.com	Sales	San Francisco	10.1.10.201
531253083348	Gabriel	Voronoff	gvoronoff	gvoronoff@buttercupgames.com	Marketing	San Francisco	10.1.10.163
520156890727	Bao	Lu	blu	blu@buttercupgames.com	Marketing	San Francisco	10.1.10.100
727896988001	Lien	Teng	lteng	lteng@buttercupgames.com	ITOps	San Francisco	10.1.10.15
936901629743	Gabriel	Voronoff	gvoronoff	gvoronoff@buttercupgames.com	ITOps	San Francisco	10.1.10.163
230876363319	Meng	Yuan	myuan	myuan@buttercupgames.com	Engineering	San Francisco	10.1.10.172
271108583080	Patrick	Callahan	pcallahan	pcallahan@buttercupgames.com	Engineering	San Francisco	10.1.10.98
569361105570	Kathleen	Percy	kpercy	kpercy@buttercupgames.com	Compliance Officer	San Francisco	10.1.10.216
145297537706	Nigella	Pearce	npearce	npearce@buttercupgames.com	SecOps	London	10.2.10.70
632071692298	Yanto	Owen	yowen	yowen@buttercupgames.com	Sales	London	10.2.10.170
862417886973	Finlay	Bryan	fbryan	fbryan@buttercupgames.com	Sales	London	10.2.10.166
890313901800	Bradley	Hussain	bhussain	bhussain@buttercupgames.com	ITOps	London	10.2.10.22
425932411002	Naomi	Sharpe	nsharpe	nsharpe@buttercupgames.com	ITOps	London	10.2.10.163

Generated for Anh Hoang (hoang_anhkim@bah.com) (C) Splunk Inc, not for distribution

What is a Calculated Field?

- Shortcut for performing repetitive, long, or complex transformations using the eval command
- Must be based on an extracted field

New Search

```
index=network sourcetype=cisco_wsa_squid
| eval megabytes = sc_bytes/(1024*1024)
| stats sum(megabytes) as Megabytes by usage
| sort Megabytes
```

Last 24 hours 

✓ 1,211 events (1/24/18 12:00:00.000 PM to 1/25/18 12:21:26.000 PM) No Event Sampling ▾ Job ▾ II ▾ Smart Mode ▾

Events Patterns Statistics (5) Visualization

100 Per Page ▾ Format Preview ▾

usage	Megabytes
Violation	0.0172176361083984380
Business	0.29117774963378906000
Borderline	2.04081058502197270000
Unknown	2.59564876556396500000
Personal	7.97706794738769500000

Note 

Output fields from a lookup table or fields/columns generated from within a search string are not supported.

Generated for Anh Hoang (hoang_anhkim@bah.com) (C) Splunk Inc, not for distribution

Creating a Calculated Field

Settings > Fields > Calculated fields > New Calculated Field

1. Select the app that will use the calculated field
2. Select host, source, or sourcetype to apply to the calculated field and specify the related name
3. Name the calculated field
4. Define the eval expression

Add new

Fields > Calculated fields > Add new

1 Destination app class_Fund2

2 Apply to sourcetype named * cisco_wsa_squid

3 Name * megabytes
Name of the field whose value will be calculated

4 Eval expression * sc_bytes/(1024*1024)
A valid eval expression, e.g. x + 3

Cancel Save

Generated for Anh Hoang (hoang_anhkim@bah.com) (C) Splunk Inc, not for distribution

Using a Calculated Field

After you have created a calculated field, you can use it in a search like any other extracted field

New Search

```
index=network sourcetype=cisco_wsa_squid
| stats sum(megabytes) as Megabytes by usage
| sort Megabytes
```

Last 24 hours ▾ 

✓ 1,229 events (1/24/18 12:00:00.000 PM to 1/25/18 12:35:16.000 PM) No Event Sampling ▾ Job ▾ II ▾ Smart Mode ▾

Events Patterns Statistics (5) Visualization

100 Per Page ▾ Format Preview ▾

usage	Megabytes
Violation	0.0172176361083984380
Business	0.29117774963378906000
Borderline	2.04081058502197270000
Unknown	2.65176105499267600000
Personal	8.19471740722656200000

Generated for Anh Hoang (hoang_anhkim@bah.com) (C) Splunk Inc, not for distribution

Module 8 Lab Exercise

Time: 15 minutes

Tasks:

- Create a field alias for the user field
- Create a calculated field that converts bytes to megabytes

Module 9: Working with Tags and Event Types

Generated for Anh Hoang (hoang_anhkim@bah.com) (C) Splunk Inc, not for distribution

Module Objectives

- Create and use tags
- Describe event types and their uses
- Create an event type
- Understand the search-time operation sequence

Describing Tags

- A knowledge object that enables you to search for events that contain particular field values
- Tags are like labels that you create for related field/value pairs
- Tags make your data more understandable and less ambiguous
- You can create one or more tags for any field/value combination
- Tags are case sensitive

Creating Tags

To create a tag:

1. Click on the arrow for event details
2. Under Actions, click the down arrow
3. Select Edit Tags
4. Name the tags, separated by commas

1

> 1/25/18 Thu Jan 25 2018 21:09:26 www2 sshd[43407]: Failed password for user nsharpe
1:09:26.000 PM .77 port 3073 ssh2
host = www3 | source = /opt/log/www3/secure.log | sourcetype = linux_secure

The screenshot shows the Splunk interface for creating tags. At the top, there is a log entry: "Thu Jan 25 2018 21:09:26 www2 sshd[43407]: Failed password for user nsharpe from 26.171.161.77 port 3073 ssh2". Below it is a table of event details with checkboxes for selecting fields like host, source, and sourcetype. A dropdown menu labeled "Actions" is open, with the "Edit Tags" option highlighted and circled in orange. A green box surrounds the "Edit Tags" button. In the bottom right corner, a modal window titled "Create Tags" is open, showing a field value "user=root" and a separate input field containing "privileged", which is also circled in orange. A green box surrounds the input field. The modal has "Cancel" and "Save" buttons at the bottom.

Generated for Anh Hoang (hoang_anhkim@bah.com) (C) Splunk Inc, not for distribution

Viewing Tags

- When tagged field/value pairs are selected, the tags appear:
 - In the results as tags A
 - In parentheses next to the associated field/value pairs B

The screenshot shows a Splunk search interface. At the top, a sidebar titled "tag" is open, displaying a list of tags: authentication, error, failure, os, privileged, remote, and unix. Below this is a table of search results with columns for Event and _source. The first result is highlighted with a green box and labeled 'A'. The event details show a failed password attempt from 190.113.128.150 port 4655 ssh2. The source is /opt/log/mailsv1/secure.log and the sourcetype is linux_secure. The tags listed are authentication, error, failure, os, privileged, remote, and unix. The second result is labeled 'B' and shows a user named root with a privilege level of privileged.

Event	_source
1/25/18 Thu Jan 25 2018 21:36:43 mailsv1 sshd[2411]: Failed password for root from 190.113.128.150 port 1:36:43.000 PM 4655 ssh2 host = mailsv1 source = /opt/log/mailsv1/secure.log sourcetype = linux_secure tag = authentication tag = error tag = failure tag = os tag = privileged tag = remote tag = unix	ix) on(authentication remote) process sshd src_ip 190.113.128.150 src_port 4655 sshd_protocol ssh2 user root(privileged)

Generated for Anh Hoang (hoang_anhkim@bah.com) (C) Splunk Inc, not for distribution

Using Tags

To use tags in a search, use the syntax: **tag=<tag name>**

New Search

Save As ▾ Close

index=security sourcetype=linux_secure tag=priv* src_ip!=NULL
| stats count by src_ip, host

Last 24 hours ▾

✓ 228 events (1/24/18 2:00:00.000 PM to 1/25/18 2:02:52.000 PM) No Event Sampling ▾ Job ▾ II ■ ▶ 🔍 ⏪ ⏩ Smart Mode ▾

Events Patterns Statistics (192) Visualization

100 Per Page ▾ Format Preview ▾ < Prev 1 2 Next >

src_ip	host	count
107.3.146.207	mailsv1	1
108.65.113.83	www2	1
109.169.32.135	www2	2
110.159.208.78	www2	1
111.161.27.20	www1	1
111.161.27.20	www2	1
111.161.27.20	www3	1
112.111.162.4	www1	1
117.21.246.164	www2	2

Generated for Anh Hoang (hoang_anhkim@bah.com) (C) Splunk Inc, not for distribution

Searching for Tags

- To search for a tag associated with a value:

`tag=<tagnname>`

```
tag=privileged
```

Note



Tag names are case sensitive.

- To search for a tag using a partial field value:

Use (*) wildcard

```
tag=p*
```

- To search for a tag associated with a value on a specific field:

`tag::<field>=<tagnname>`

```
tag::user=privileged
```

Managing Tags – List by Field Value Pair

- **Settings > Tags > List by field value pair**
 - Edit permissions
 - Disable all tags for pair: disables the tag in searches and prevents it from being listed under List by Tag Name and All unique tag objects

List by field value pair

[Tags](#) » List by field value pair

Showing 1-2 of 2 items

App CLASS: Fundamentals... Owner Any Created in the App filter 25 per page

Field value pair	Tag name	App	Sharing	Status	Actions
user=administrator	privileged	class_Fund2	Private	Enabled Disable all tags for pair	Clone Move Delete
user=root	privileged	class_Fund2	Private	Enabled Disable all tags for pair	Clone Move Delete

Generated for Anh Hoang (hoang_anhkim@bah.com) (C) Splunk Inc, not for distribution

Adding/Changing the Tag Name

Click **List by field value pair** to add another tag or change the name of the tag

The screenshot shows the Splunk UI for managing tags. On the left, there's a main panel titled "List by field value pair" with a "New Tag" button. It displays a table of items with columns: "Field value pair", "Tag name", and "App". Two rows are shown: one for "user=administrator" and one for "user=root". A green arrow points from the "user=administrator" row to a modal dialog on the right. The modal has a title "user=administrator" and a "Tag name" input field containing "privileged". It also has "Delete" buttons for each row and a "+ Add another field" button. At the bottom are "Cancel" and "Save" buttons.

Field value pair	Tag name	App
user=administrator	privileged	class_Fund2
user=root	privileged	class_Fund2

user=administrator

Tags > List by field value pair > user=administrator

Tag name Enter one tag per textfield

privileged

+ Add another field

Cancel Save

Generated for Anh Hoang (hoang_anhkim@bah.com) (C) Splunk Inc, not for distribution

Adding/Changing the Field Value Pair

Click **List by tag name** to add or edit the field value pair for the tag

The screenshot shows the Splunk UI for managing tags. On the left, the 'List by tag name' page is displayed. It includes a breadcrumb 'Tags > List by tag name', a search bar, and filters for 'App' (CLASS: Fundamentals...), 'Owner' (Any), 'Created in the App', and 'filter'. A green 'New Tag' button is at the top right. Below these are dropdowns for 'Tag name' (set to 'privileged') and 'Field value pair' (set to 'user=administrator, user=root'). A green arrow points from the 'privileged' tag name in the list to the 'privileged' tag name in the modal. The modal window is titled 'privileged' and shows the same tag name and field value pair. It also includes a 'Field value pair' input field ('example: host=splunk.com'), three delete buttons for each field value, and a '+ Add another field' button. At the bottom are 'Cancel' and 'Save' buttons.

Generated for Anh Hoang (hoang_anhkim@bah.com) (C) Splunk Inc, not for distribution

Describing Event Types

- A method of categorizing events based on a search
- A useful method for institutional knowledge capturing and sharing
- Can be tagged to group similar types of events

Creating an Event Type from the Search Page

1. Run a search and verify that all results meet your event type criteria
2. From the Save As menu, select **Event Type**
3. Provide a Name for your event type (name should not contain spaces)

The screenshot shows the Splunk search interface. On the left, a search bar contains the query `index==* status>499`. Below the search bar, it says "195 events (1/24/18 3:00:00.000 PM to 1/25/18 3:06:20.000 PM) No Event Sampling". The "Events (195)" tab is selected. A context menu is open at the top right, with "Event Type" highlighted. An arrow points from this menu to a "Save As Event Type" dialog box on the right.

Save As Event Type

- Name: web_error
- Tags: Optional
- Color: yellow
- Priority: 1 (Highest)

Determines which style wins, when an event has more than one event type.

Note

Must be a basic search (cannot contain pipes or subsearches).

Generated for Anh Hoang (hoang_anhkim@bah.com) (C) Splunk Inc, not for distribution

Using the Event Type Builder

1. From the event details, select **Event Actions > Build Event Type**

The screenshot shows the Splunk search interface with the following details:

- Search Bar:** index=*& status>499, Last 24 hours.
- Event Summary:** 195 events (1/24/18 3:00:00.000 PM to 1/25/18 3:06:20.000 PM), No Event Sampling.
- Event Timeline:** A timeline visualization showing event times over a 1-hour period.
- Event Details View:** Shows a single event from 1/25/18 3:06:04.000 PM. The event data includes:
 - Time: 1/25/18 3:06:04.000 PM
 - Event ID: 202.91.242.117 - - [25/Jan/2018:23:06:04] "POST /cart.do?action=view&itemId=EST-19&JSESSIONID=SD0SL3FF2ADFF4962 HT TP 1.1" 500 2017 "http://www.buttercupgames.com" "Mozilla/4.0 (compatible; MSIE 7.0; Windows NT 5.1; .NET CLR 2.0.50727; .NET CLR 3.0.4506.2152; .NET CLR 3.5.30729; InfoPath.1; .NET4.0C; .NET4.0E; MS-RTC LM 8)" 609
- Event Actions:** A dropdown menu is open, with the "Build Event Type" option highlighted.
- Interesting Fields:** A list of fields including host, source, sourcetype, and tag.
- Table:** A table showing the selected field "host" with value "www2".

Generated for Anh Hoang (hoang_anhkim@bah.com) (C) Splunk Inc, not for distribution

Using the Event Type Builder (cont.)

2. Refine the criteria for your event type such as:

- Search string
- Field values
- Tags

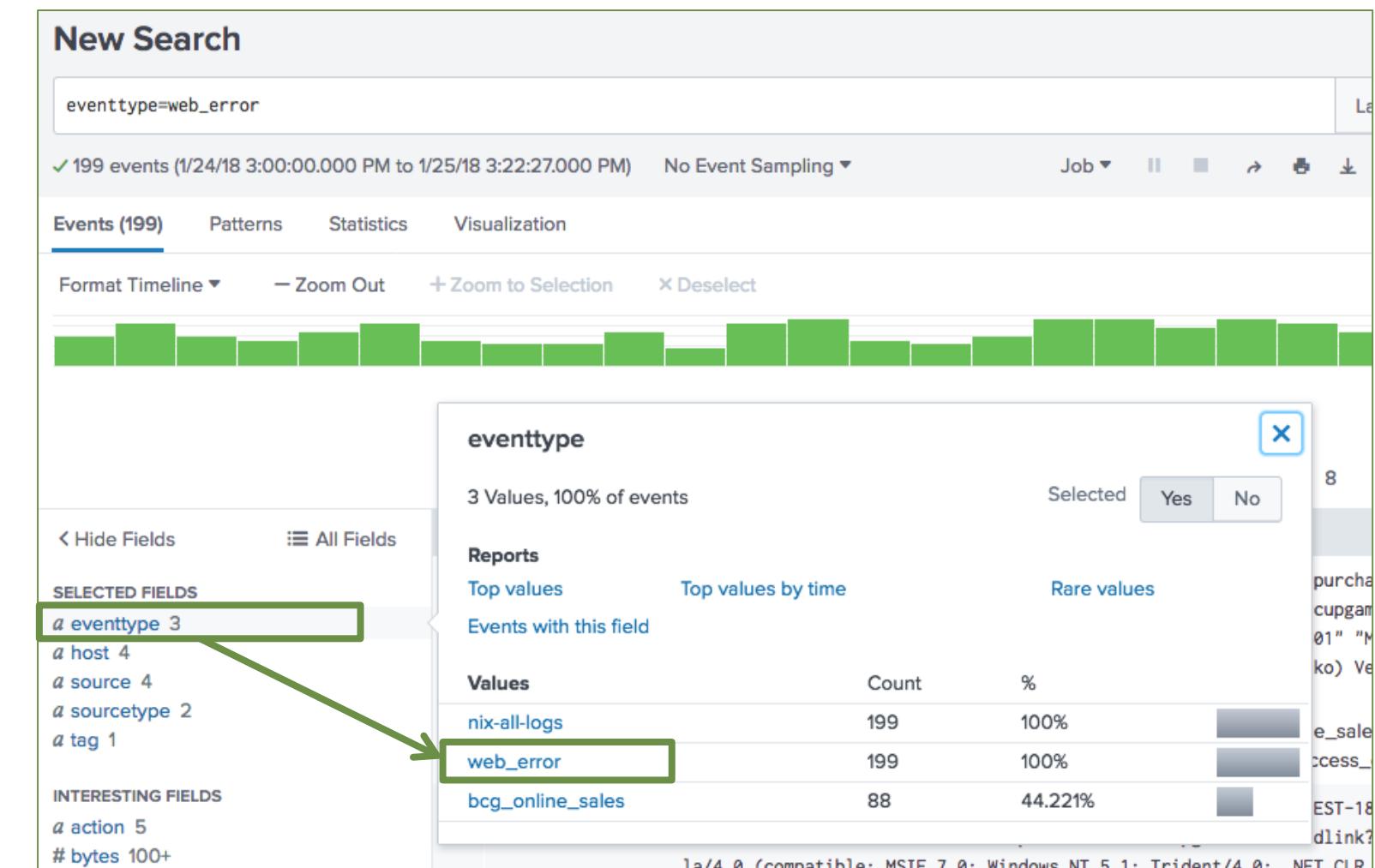
3. Verify your selections and click Save

The screenshot shows the Splunk Event Type Builder interface. At the top right, there is a panel titled "Generated event type" containing the search query: "(index=!* status>499)". Below it are three buttons: "Edit", "Test", and "Save". To the right of this panel is a vertical sidebar titled "Suggested field values" which lists various event fields like ident, linecount, tag, tag:eventtype, version, eventtype, index, method, and sourcetype, each with a list of possible values. The main area of the interface is titled "Generated event type" and contains two sections: "Suggested field values" and "Sample events". The "Suggested field values" section contains the same list of fields as the sidebar, with the selected values highlighted in blue. The "Sample events" section displays a list of log entries matching the search query. A large gray diagonal shadow is cast over the entire interface from the top right corner.

Generated for Anh Hoang (hoang_anhkim@bah.com) (C) Splunk Inc. not for distribution

Using Event Types

- To verify the event type, search for `eventtype=web_error`
- ‘`eventtype`’ displays in the Fields sidebar and can be added as a selected field
- Splunk evaluates the events and applies the appropriate event types at search time
- Using the Fields sidebar, you can easily view the individual event types, the number of events, and percentage



Generated for Anh Hoang (hoang_anhkim@bah.com) (C) Splunk Inc, not for distribution

Tagging Event Types

You can tag event types two ways:

1. **Settings > Event types**
2. **Event details > Actions**

Event Actions ▾

Type	Field	Value	Actions
Selected	<input checked="" type="checkbox"/>	eventtype	▼
		nix-all-logs	▼
		web_error(error web)	▼
		bog_online_sales(wet)	▼
	<input checked="" type="checkbox"/>	host	www1

Edit Tags

Create Tags

Field Value	Tag(s)
eventtype=web_error	error, web

Comma or space separated list of tags.

Save

web_error

Event types > web_error

Search string * index== status>499

Tag(s) **error web**
Enter a comma-separated list of tags.

Color **yellow**

Priority **1 (Highest)**
Highest priority shows up first in a result.

Cancel **Save**

Priority controls which event type color displays for an event

Generated for Anh Hoang (hoang_anhkim@bah.com) (C) Splunk Inc, not for distribution

Event Types vs. Saved Reports

- Event Types

- Categorize events based on a search string
- Tag event types to organize data into categories
- The eventtype field can be included in a search string
- Does not include a time range

- Saved Reports

- Search criteria will not change
- Includes a time range and formatting of the results
- Can be shared with Splunk users and added to dashboards

Search-time Operation Sequence

- Search-time operations are always applied in the same order
- Each operation can reference fields derived from operations that **precede** them
- No operation can reference fields that are derived by operations that **follow** them

Search-time Operation Sequence (cont.)

1. Inline field extractions
2. Field extractions that use a field transform
3. Automatic key-value field extractions
4. Field aliases
5. Calculated fields
6. Lookups
7. Event types
8. Tags

Note

Details are available at
<https://docs.splunk.com/Documentation/Splunk/latest/Knowledge/Searchtimeoperationssequence#>

Search-time Operation Sequence (cont.)

- You can perform search operations such as:
 - Creating a calculated field that references a field alias, or
 - Creating an event type that references a calculated field
- But you **can't**:
 - Create a field alias that references a calculated field, or
 - Create a calculated field that references a field added through a lookup operation
- Splunk has to generate the object before it can be used

Module 9 Lab Exercise

Time: 15 minutes

Tasks:

- Create a **privileged** tag for all user values containing the word *admin*
- Create a **web_error** event type for all web servers/devices that receive an error of 500 or greater

Module 10:

Creating and Using

Macros

Generated for Anh Hoang (hoang_anhkim@bah.com) (C) Splunk Inc, not for distribution

Module Objectives

- Describe macros
- Manage macros
- Create a basic macro
- Use a basic macro
- Define arguments / variables for a macro
- Add and use arguments with a macro

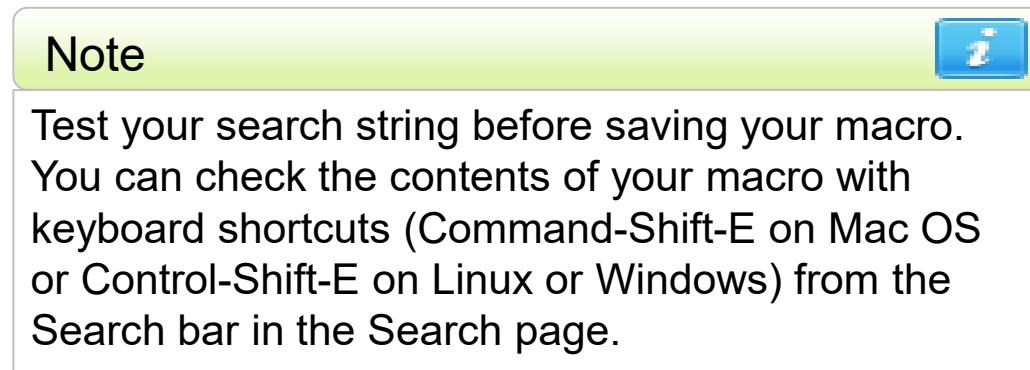
Macros Overview

- Useful when you frequently run searches or reports with similar search syntax
- The time range is selected at search time
- Macros can be a full search string or a portion of a search that can be reused in multiple places
- Allows you to define one or more arguments within search segment
 - Pass parameter values to macro at execution time
 - Macro uses values to resolve search string

Creating a Basic Macro

Settings > Advanced search > Search macros

1. Click **New Search Macro**
2. Select the destination app
3. Enter a name
4. Type the search string
5. Save



Add new
Advanced search > Search macros > Add new

Destination app class_Fund2

Destination app class_Fund2

Name * Enter the name of the macro. If the search macro takes an argument, indicate of arguments to the name. For example: mymacro(2)
US_sales

Definition * Enter the string the search macro expands to when it is referenced in another included, enclose them in dollar signs. For example: \$arg1\$
index=sales sourcetype=vendor_sales VendorCountry="United States"
| stats sum(price) as USD by product_name
| eval USD = "\$" + tostring('USD', "commas")

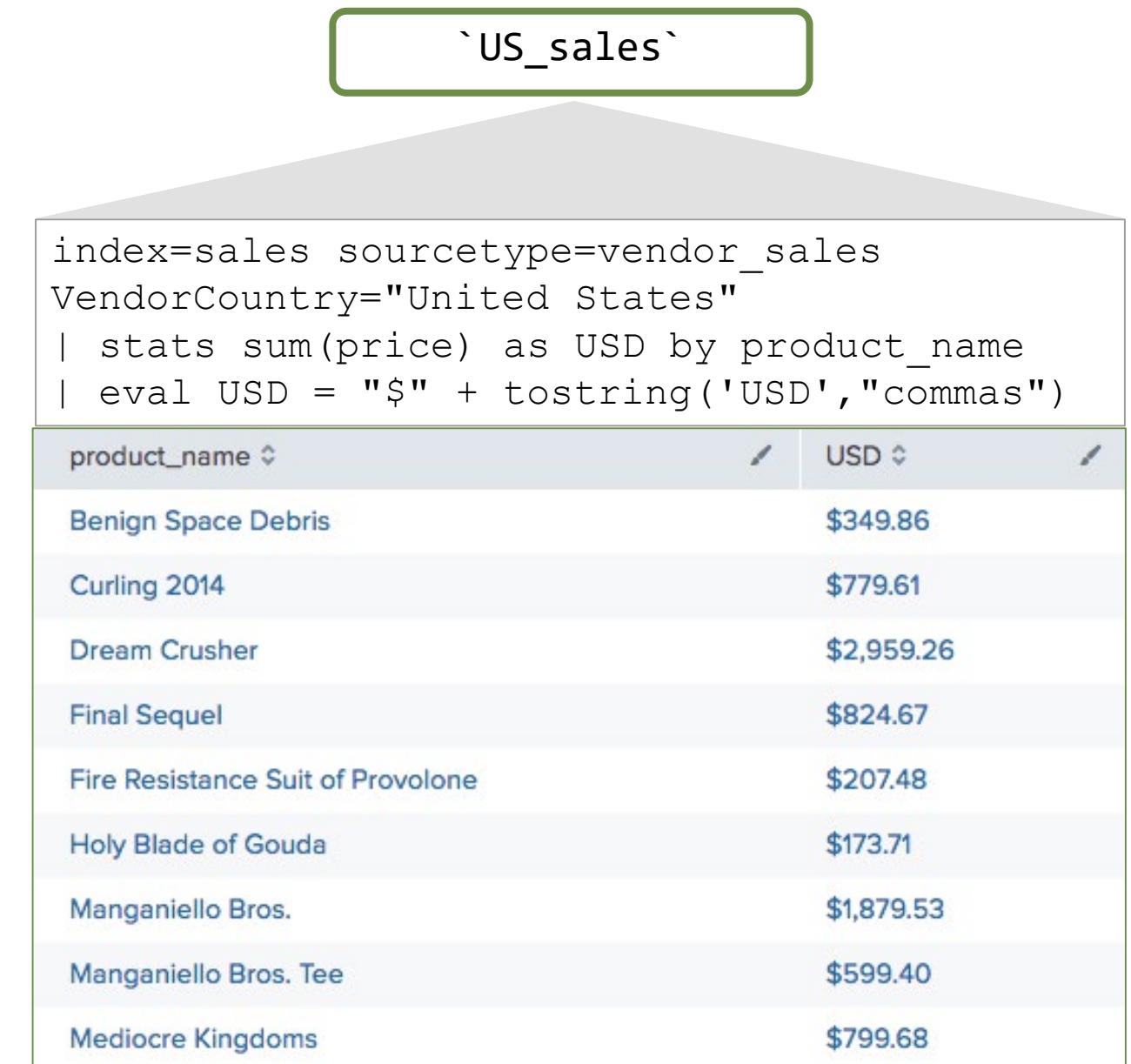
Validation Error Message Enter a message to display when the validation expression returns 'false'.

Cancel Save

Generated for Anh Hoang (hoang_anhkim@bah.com) (C) Splunk Inc, not for distribution

Using a Basic Macro

- Type the macro name into the search bar
- Surround the macro name with the **backtick** (or grave accent) character
 - **`macroname`** != 'macroname'
 - Do not confuse with single-quote character (')
- Pipe to more commands, or precede with search string



The screenshot shows a search results table with two columns: **product_name** and **USD**. The table lists various items with their prices in USD. Above the table, a search bar contains the macro name ``US_sales``.

product_name	USD
Benign Space Debris	\$349.86
Curling 2014	\$779.61
Dream Crusher	\$2,959.26
Final Sequel	\$824.67
Fire Resistance Suit of Provolone	\$207.48
Holy Blade of Gouda	\$173.71
Manganiello Bros.	\$1,879.53
Manganiello Bros. Tee	\$599.40
Mediocre Kingdoms	\$799.68

Adding Arguments

- Include the number of arguments in parentheses after the macro name
 - `monthly_sales(3)`
- Within the search definition, use `arg`
 - `currency=$currency$`
 - `symbol=$symbol$`
 - `rate=$rate$`
- In the **Arguments** field, enter the name of the argument(s)
- Provide one or more variables of the macro at search time

Add new
Advanced search > Search macros > Add new

Destination app class_Fund2

Name * Enter the name of the macro. If the search macro takes an argument, indicate this by appending the number of arguments to the name. For example: mymacro(2)

monthly_sales(3)

Definition * Enter the string the search macro expands to when it is referenced in another search. If the definition includes variables, enclose them in dollar signs. For example: \$arg1\$

stats sum(price) as USD by product_name
| eval \$currency\$ = "\$symbol\$" + tostring(USD*\$rate\$, "commas"),
USD = "\$" + tostring(USD, "commas")

Use eval-based definition?

Arguments Enter a comma-delimited string of argument names. Argument names may only contain letters, numbers, and underscores.

currency,symbol,rate

Using Arguments

- When using a macro with arguments, include the argument(s) in parentheses following the macro name
- Be sure to pass in the arguments in the same order as you defined them

```
index=sales sourcetype=vendor_sales  
VendorCountry=Germany OR VendorCountry=France  
OR VendorCountry=Italy |  
`monthly_sales(euro,€,0.79)`
```

```
index=sales sourcetype=vendor_sales  
VendorCountry=Germany OR VendorCountry=France OR  
VendorCountry=Italy |  
stats sum(price) as USD by product_name  
| eval euro = "€" + tostring(USD*0.79,  
"commas"), USD = "$" + tostring(USD, "commas")
```

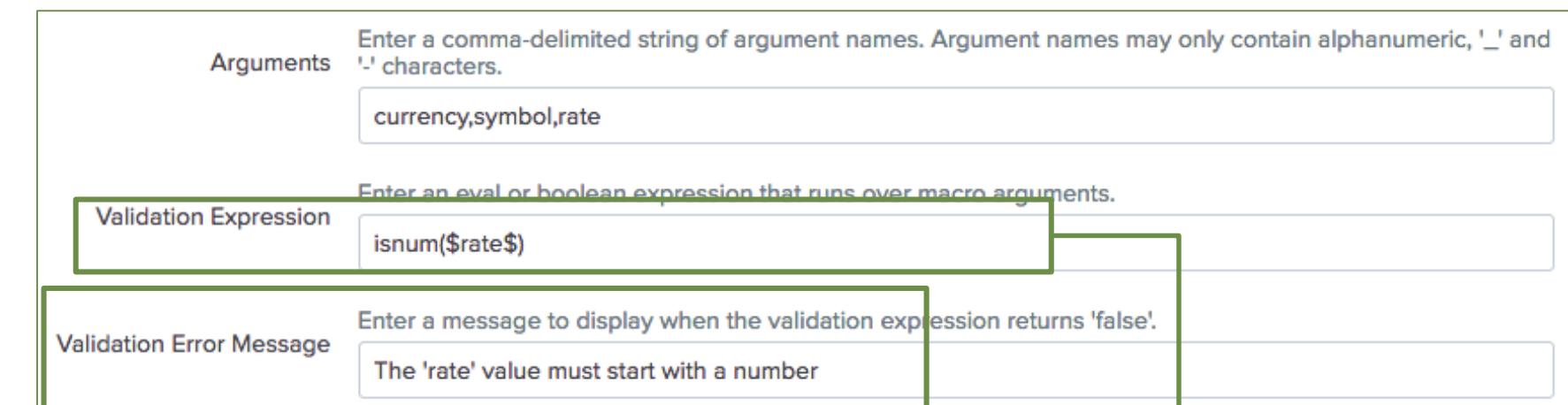
product_name	USD	euro
Benign Space Debris	\$174.93	€138
Curling 2014	\$179.91	€142
Dream Crusher	\$119.97	€95
Final Sequel	\$24.99	€20
Fire Resistance Suit of Provolone	\$31.92	€25
Holy Blade of Gouda	\$23.96	€19
Manganiello Bros.	\$719.82	€569
Manganiello Bros. Tee	\$69.93	€55
Mediocre Kingdoms	\$124.95	€99
Orvil the Wolverine	\$199.95	€158
Puppies vs. Zombies	\$9.98	€7.9
SIM Cubicle	\$219.89	€174
World of Cheese	\$99.96	€79
World of Cheese Tee	\$99.90	€79

Generated for Anh Hoang (hoang_anhkim@bah.com) (C) Splunk Inc, not for distribution

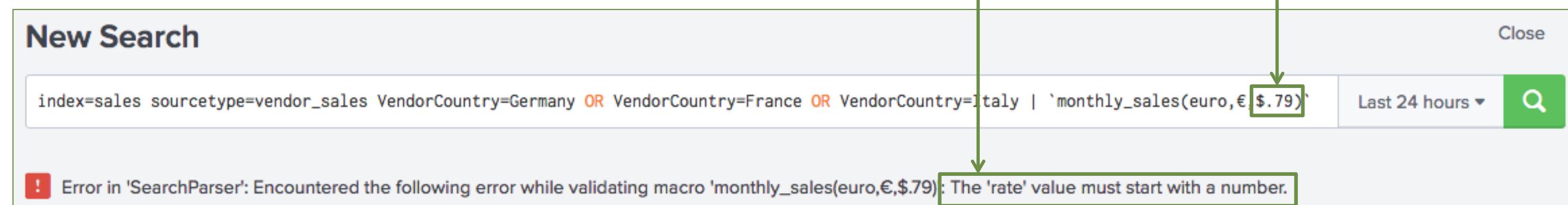
Validating Macros

- You can validate argument values in your macro
- Validation Expression: you can enter an expression for each argument
 - Argument must be enclosed in dollar signs
- Validation Error Message: message that appears when you run the macro with an incorrect argument

Note  Don't create macros with leading pipes—someone may put a pipe in front of the macro when using it in the actual search string.



The screenshot shows the configuration of a Splunk macro. It includes fields for Arguments (currency,symbol,rate), Validation Expression (isnum(\$rate\$)), and Validation Error Message (The 'rate' value must start with a number). The Validation Expression and Validation Error Message fields are highlighted with green boxes and arrows pointing to them from the corresponding sections in the list above.



Generated for Anh Hoang (hoang_anhkim@bah.com) (C) Splunk Inc, not for distribution

Module 10 Lab Exercise

Time: 25 minutes

Task:

- Use the **VendorCountry** field to create a macro that provides the sales totals for the month and converts US dollars to foreign currency
- Use the **isnum** expression to validate a numeric argument

Module 11: Creating and Using Workflow Actions

Generated for Anh Hoang (hoang_anhkim@bah.com) (C) Splunk Inc, not for distribution

Module Objectives

- Create a GET workflow action
- Create a POST workflow action
- Create a Search workflow action

What are Workflow Actions?

- Execute workflow actions from an event or field in your search results to interact with external resources or run another search
 - **GET** - retrieve information from an external resource
 - **POST** - send field values to an external resource
 - **Search** - use field values to perform a secondary search

The screenshot shows a Splunk search results page. An event is selected, displaying its timestamp (4/6/19 10:28:56.817 PM), source, and type. Below the event details, there is a "Event Actions" button. A dropdown menu titled "Build Event Type" is open, showing three options: "Get info for IP:95.130.170.231" (which is highlighted with a green border), "Extract Fields", and "Show Source". To the right of the dropdown, a table lists fields with their values and actions. The table has columns for "Value" and "Actions". The fields listed are "nix-all-logs", "cisco_router1", and "/opt/log/cisco_router1/cisco_ironport_web.log".

Value	Actions
nix-all-logs	▼
cisco_router1	▼
/opt/log/cisco_router1/cisco_ironport_web.log	▼

Generated for Anh Hoàng (hoang_anhkim@bah.com) (C) Splunk Inc, not for distribution

Creating a GET Workflow Action

Settings > Fields > Workflow actions > New Workflow Action

1. Select the app
2. Name the workflow action with no spaces or special characters
3. Define the label, which will appear in the Event Action and/or Fields menu
4. Determine if your workflow action applies to a field or event type

Add new

Fields > Workflow actions > Add new

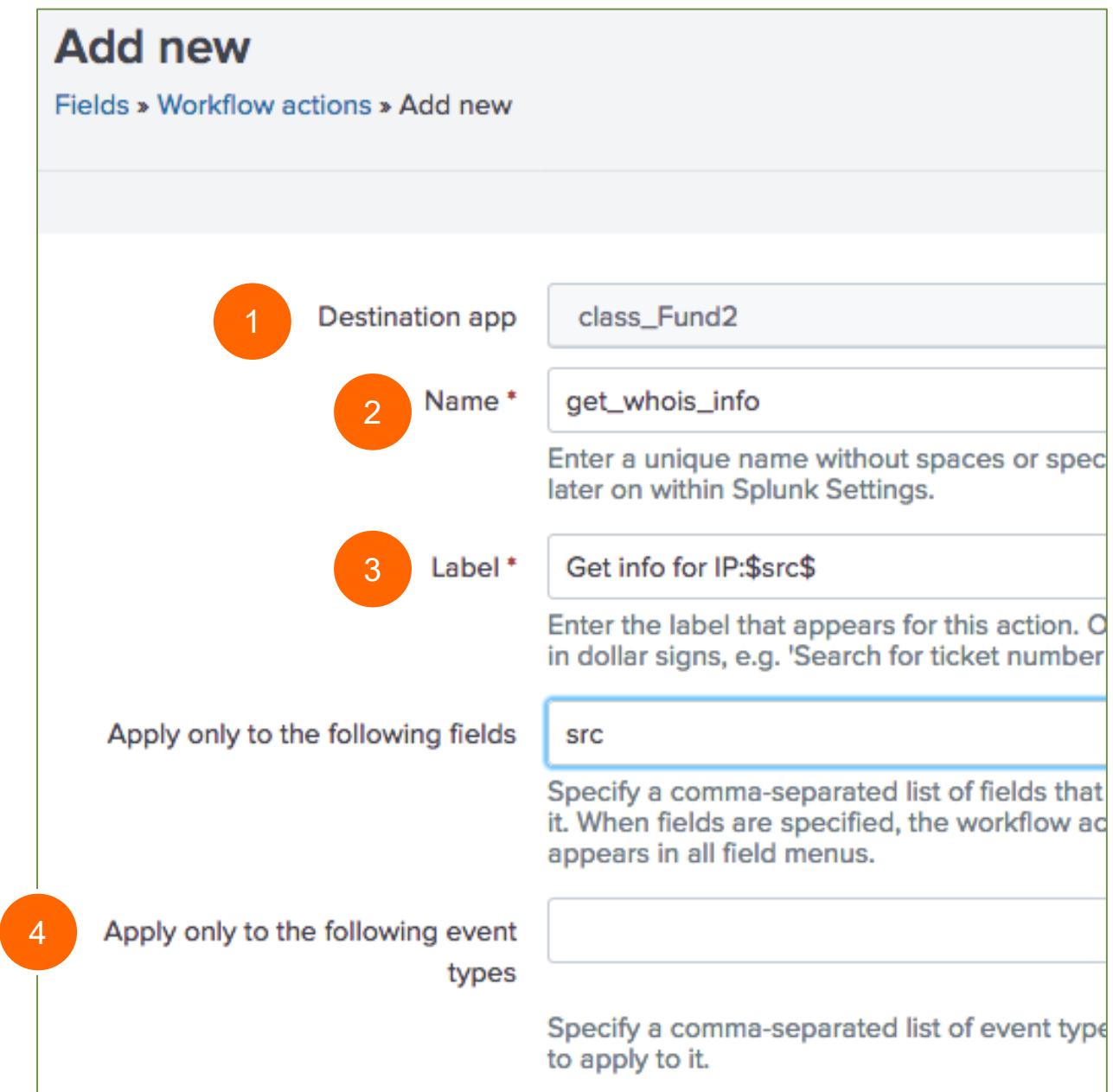
1 Destination app class_Fund2

2 Name * get_whois_info
Enter a unique name without spaces or spec later on within Splunk Settings.

3 Label * Get info for IP:\$src\$
Enter the label that appears for this action. O in dollar signs, e.g. 'Search for ticket number

4 Apply only to the following fields src
Specify a comma-separated list of fields that it. When fields are specified, the workflow ac appears in all field menus.

Apply only to the following event types
Specify a comma-separated list of event type to apply to it.



Creating a GET Workflow Action (cont.)

5. From the **Show action in** dropdown list, select **Event menu**, **Fields menu**, or **Both**
6. From Action type dropdown list, select **link**
7. Enter the URI of where the user will be directed
8. Specify if the link should open in a **New window** or **Current window**
9. Select the Link method of **get**
10. Save

The screenshot shows the configuration of a GET Workflow Action in the Splunk UI. The form fields are as follows:

- Show action in: Event menu (numbered 5)
- Action type *: link (numbered 6)
- URI *: http://who.is/whois-ip/ip-address/\$src\$ (numbered 7)
- Open link in: New window (numbered 8)
- Link method: get (numbered 9)
- Buttons: Cancel (gray) and Save (green) (numbered 10)

Testing the GET Workflow Action

The screenshot illustrates the workflow for testing a GET action. It shows three main sections: a top panel with event details, a middle panel with a search bar and dropdown menu, and a right panel displaying IP Whois information.

Event Details:

4/6/19 10:28:56.817 PM 1554589736.817 105 95.130.170.231 TCP_REFRESH_HIT/200 1270 GET http://www.fftimes.com/themes/ftimes2009/icons/rrr.png tzielinski@buttercupgames.com DIRECT/www.fftimes.com image/png DEFAULT T_CASE-DefaultGroup-Demo_Clients-NONE-NONE-DefaultRouting <IW_news> -> - http://www.fftimes.com/

Event Actions:

- Build Event Type
- Get info for IP:95.130.170.231
- Extract Fields
- Show Source

Value

nix all logo
cisco_router1
/opt/log/cisco_router1/cisco_i

95.130.170.231 address profile

Whois Diagnostics

IP Whois

NetRange:	95.0.0.0 – 95.255.255.255
CIDR:	95.0.0.0/8
NetName:	95-RIPE
NetHandle:	NET-95-0-0-0-1
Parent:	()
NetType:	Allocated to RIPE NCC
OriginAS:	
Organization:	RIPE Network Coordination Centre (RIPE)
RegDate:	2007-07-30
Updated:	2009-05-18
Comment:	These addresses have been further assigned to users in
Comment:	the RIPE NCC region. Contact information can be found in
Comment:	the RIPE database at http://www.ripe.net/whois
Ref:	https://rdap.arin.net/registry/ip/95.0.0.0

Generated for Anh Hoang (hoang_anhkim@bah.com) (C) Splunk Inc, not for distribution

Creating a POST Workflow Action

**Settings > Fields >
Workflow actions > New
Workflow Action**

Complete steps 1 – 6 as described in the previous example, Creating a GET Workflow Action

The screenshot shows the 'Add new' workflow action configuration page in Splunk. The page has a header 'Add new' under 'Fields > Workflow actions > Add new'. There are several input fields with numbered callouts:

- Destination app:** class_Fund2 (Step 1)
- Name ***: multiple_attempts_to_open_ports (Step 2)
- Label ***: Create ticket - multiple attempts to port:\$port\$ (Step 3)
- Apply only to the following fields:** port (Step 4)
- Apply only to the following event types:** (Step 5)
- Show action in:** Event menu (Step 6)
- Action type ***: link

Generated for Anh Hoang (hoang_anhkim@bah.com) (C) Splunk Inc, not for distribution

Creating a POST Workflow Action (cont.)

7. Enter the URI of where the user will be directed
8. Open the link in a **New window** or Current window
9. Select the Link method of **post**
10. Provide post argument parameters
11. Save

Link configuration

7 URI *
Enter the location to link to. Optionally, specify fields by enclosing the field name in dollar signs, e.g.
http://www.google.com/search?q=\$host\$.

8 Open link in

9 Link method

Post arguments

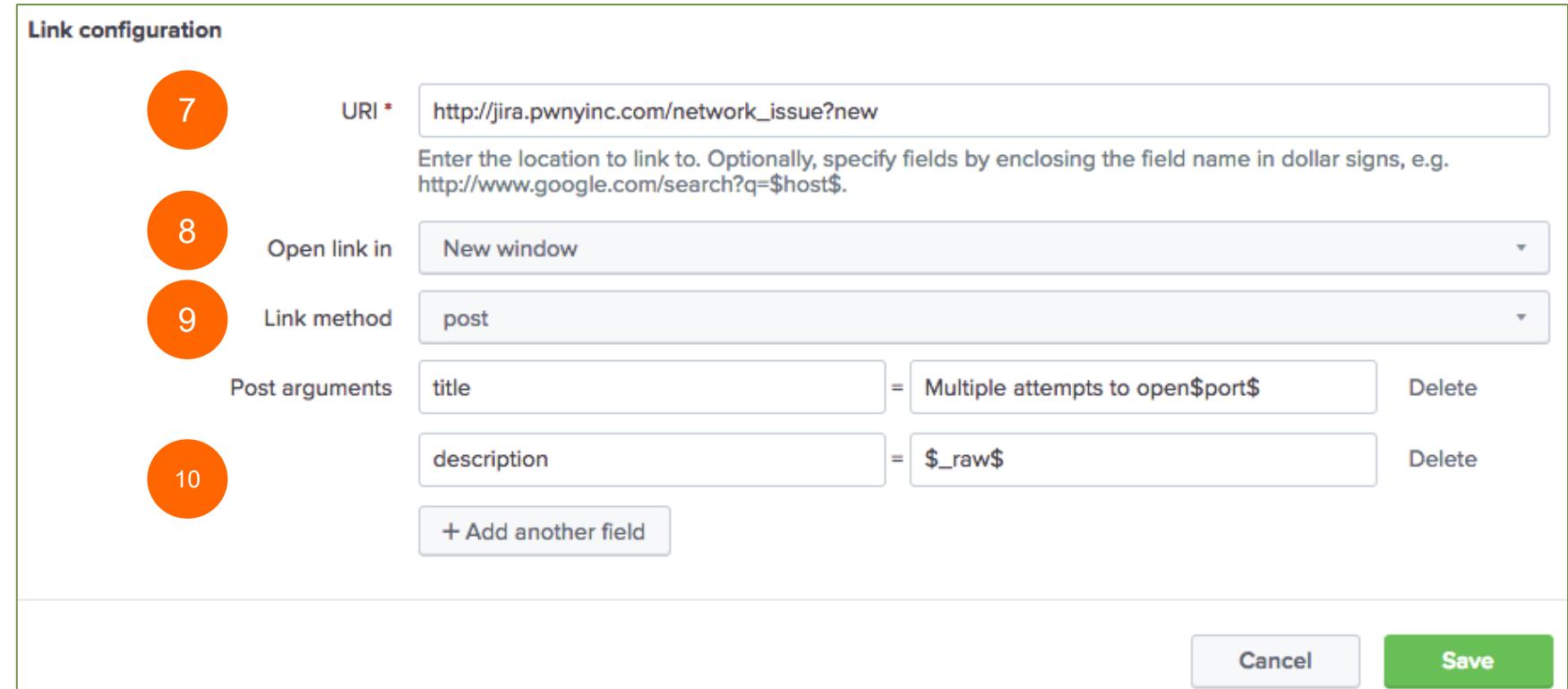
10 title = Delete

 description = Delete

+ Add another field

Cancel Save

11



Creating a Search Workflow Action

Settings > Fields > Workflow actions > New Workflow Action

Complete steps 1 – 5 as described in the previous example, Creating a GET Workflow Action

6. From the Action type dropdown list, select **search**

Add new

Fields > Workflow actions > Add new

1 Destination app class_Fund2

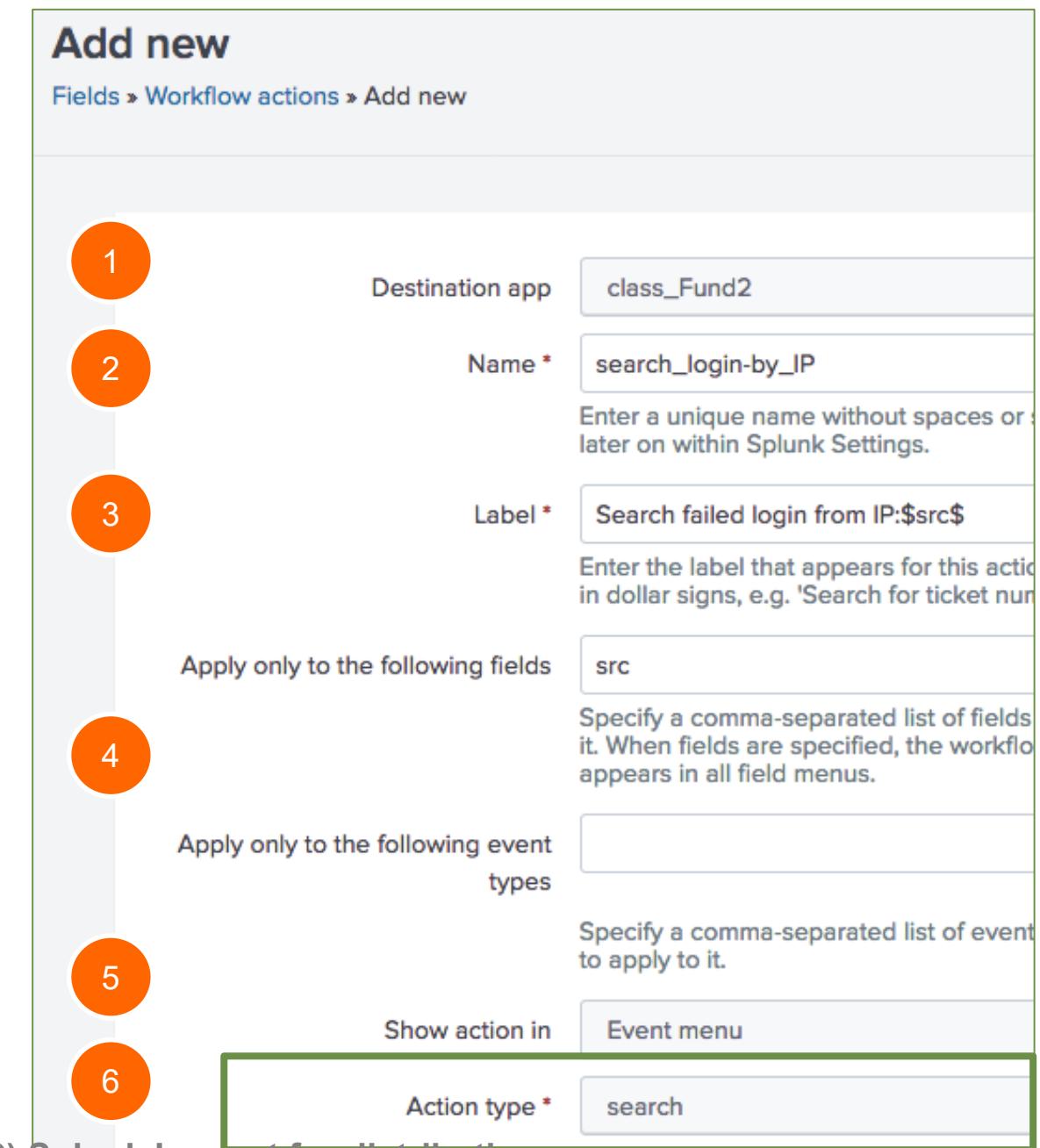
2 Name * search_login-by_IP
Enter a unique name without spaces or later on within Splunk Settings.

3 Label * Search failed login from IP:\$src\$
Enter the label that appears for this action in dollar signs, e.g. 'Search for ticket number \$src\$'

4 Apply only to the following fields src
Specify a comma-separated list of fields it. When fields are specified, the workflow appears in all field menus.

5 Apply only to the following event types
Specify a comma-separated list of event types to apply to it.

6 Show action in Event menu
Action type * search



Creating a Search Workflow Action (cont.)

7. Enter the Search string
8. Select the app if it is different from the current app
9. Enter the view name where the search will execute
10. Indicate if the search should run in a New window or the Current window
11. Either
 - Enter Earliest time and Latest time, or
 - Just click the checkbox to use the same time range as the original search

Search configuration

Search string * 7

Enter the search for this action. Optionally, specify fields as \$fieldname\$, e.g. \$sourcetype=rails\$ controller=\$controller\$ error=*.

Run in app 8

Choose an app for the search to run in. Defaults to the current app.

Open in view 9

Enter the name of a view for the search to open in. Defaults to the current view.

Run search in 10

New window

Time range

Earliest time

Latest time 11

Use the same time range as the search that created the field listing

12

12. Save

Testing the Search Workflow Action

The screenshot shows the Splunk interface with a search results page and a sidebar.

Search Results:

- Event Actions:** A dropdown menu containing:
 - Build Event Type
 - Extract Fields
 - Search failed login from IP:188.173.152.100** (highlighted with a green box)
 - Show Source
- New Search:** A search bar containing the query `index=security sourcetype=linux_secure failed src_ip=188.173.152.100`. Below it is a timeline visualization showing event counts over time.
- Events (22):** A table listing 22 events. The first two events are shown in detail:

Time	Event
Fri Jan 26 2018 22:09:38	Failed password for invalid user administrator from 188.173.152.100 port 3454 ssh2 eventtype = errOr error eventtype = failed_login eventtype = nix-all-logs eventtype = nix_errors error eventtype = nix_security os unix eventtype = sshd_authentication authentication remote host = www2 source = /opt/log/www2/secure.log sourcetype = linux_secure tag = authentication tag = error tag = failure tag = os tag = privileged tag = remote tag = unix
Fri Jan 26 2018 21:59:13	Failed password for invalid user rdb from 188.173.152.100 port 1623 ssh2 eventtype = errOr error eventtype = failed_login eventtype = nix-all-logs eventtype = nix_errors error eventtype = nix_security os unix eventtype = sshd_authentication authentication remote host = www2 source = /opt/log/www2/secure.log sourcetype = linux_secure tag = authentication tag = error tag = failure tag = os tag = remote tag = unix

Generated for Anh Hoang (hoang_anhkim@bah.com) (C) Splunk Inc, not for distribution

Module 11 Lab Exercise

Time: 20 minutes

Tasks:

- Use the `src_ip` field to create a GET workflow action that opens a new browser window with information about a source IP address
- Create a POST workflow action that uses fields from events with errors to create a ticket in the IT ticket tracking system
- Create a Search workflow action that performs a secondary search for all failed password events associated with a specific IP address

Module 12:

Creating Data Models

Generated for Anh Hoang (hoang_anhkim@bah.com) (C) Splunk Inc, not for distribution

Objectives

- Describe the relationship between data models and Pivot
- Identify data model datasets
- Identify dataset fields
- Create a data model
- Use a data model in Pivot

Reviewing Pivot

- Used for creating reports and dashboards
- Pivot reports are based on datasets
- As a knowledge manager, you're responsible for building the **data model** that provides those datasets

Note



Using pivot is discussed in the live, instructor-led version of *Splunk Fundamentals 1*.

The screenshot illustrates the process of creating a pivot report. On the left, the 'Datasets' page shows a list of 9 datasets, with one dataset selected ('Buttercup Games Online Sales > Web Requests') highlighted. A green arrow points from the 'Visualize with Pivot' button in the dataset's Actions menu to the 'New Pivot' interface on the right. The 'New Pivot' interface displays a summary of 74,789 events from December 27, 2017, to January 26, 2018. It includes sections for Filters (Last 30 days), Split Columns (product name), Split Rows (action), Column Values (Count of Web...), and a main data grid. The data grid contains columns for various actions and their counts, such as addtocart, changequantity, purchase, remove, and view.

action	Benign Space Debris	Curling 2014	Dream Crusher	Final Sequel	Fire Resistance Suit of Provolone	Holy Blade of Gouda	Manganiello Bros.	Manganiello Bros. Tee	Mediocre Kingdoms	NULL	Orvil the Wolverine	Puppies vs. Zombies	SIM Cubicle	World of Cheese
addtocart	807	767	1183	1041	1123	959	1131	1038	1272	281	835	895	1291	1317
changequantity	179	165	305	239	250	225	277	223	292	129	209	227	293	278
purchase	460	398	661	569	616	577	611	575	685	118	482	512	707	730
remove	170	171	274	231	267	188	247	267	301	120	188	190	300	277
view	689	697	1136	939	1066	912	975	996	1141	503	753	823	1136	1227

Generated for Anh Hoang (hoang_anhkim@bah.com) (C) Splunk Inc, not for distribution

Overview of Data Models

- Hierarchically structured datasets containing searches and fields
- Each event, search, or transaction is saved as a separate dataset
- Data models can be accelerated for faster performance

The screenshot shows the Splunk Enterprise interface for managing data models. The top navigation bar includes 'splunk>enterprise', 'App: Search & Reporting', 'student1', 'Messages', 'Settings', 'Activity', 'Help', and a search bar. The main title is 'Buttercup Games Site Activity' under 'Buttercup_Games_Site_Activity'. Below the title are buttons for 'Edit', 'Download', 'Pivot', and 'Documentation'. A sidebar on the left lists 'Datasets' and 'EVENTS'. Under 'EVENTS', the 'Web Requests' dataset is selected, showing its structure. It contains two main sections: 'Successful Requests' (with sub-categories like 'purchases', 'addtocart', 'remove') and 'Failed Requests' (with sub-categories like 'failed purchases', 'failed addtocart', 'failed remove'). To the right of the sidebar, the 'Web Requests' dataset is detailed. It includes a 'CONSTRAINTS' section with the query 'index=web sourcetype=access_combined', a 'Constraint' button, and an 'Edit' button. Below this are sections for 'INHERITED' fields (_time, host, source, sourcetype) and 'EXTRACTED' fields (action, bytes, categoryId, change_type, clientip, cookie, status). Each field has an 'Override' or 'Edit' link.

Data Model Dataset Types

A data model can consist of 3 types of datasets

- Events
- Searches
- Transactions

Events

Searches

Transactions

Buttercup Games Site Activity

Buttercup_Games_Site_Activity

< All Data Models

Datasets Add Dataset ▾

EVENTS

Web Requests

Web_Requests

CONSTRAINTS

index=web sourcetype=access_combined

Constraint Edit

Bulk Edit ▾

INHERITED

_time Time

host String

source String

sourcetype String

Override

Override

Override

EXTRACTED

action String

bytes Number

categoryid String

change_type String

clientip IPv4

cookie String

Edit

Edit

Edit

Edit

Edit

Add Field ▾

Generated for Anh Hoang (hoang_anhkim@bah.com) (C) Splunk Inc, not for distribution

Data Model Events

- **Event datasets** contain constraints and fields
- **Constraints** are essentially the search broken down into a hierarchy
- **Fields** are properties associated with the events

The screenshot shows the Splunk Data Model Editor interface. On the left, there's a tree view under the 'EVENTS' tab labeled 'Web Requests'. It has two main branches: 'Successful Requests' containing 'purchases', 'addtocart', and 'remove'; and 'Failed Requests' containing 'failed purchases', 'failed addtocart', and 'failed remove'. To the right of the tree view, the 'Web Requests' dataset details are shown. At the top, it says 'Web Requests' and 'Web_Requests'. Below that is the 'CONSTRAINTS' section with the search string 'index=web sourcetype=access_combined'. A yellow box highlights this section with the label 'base search'. Underneath is the 'INHERITED' section, which includes the '_time' field. Another yellow box highlights this section with the label 'fields'. The 'fields' section lists several fields with their types and edit options: '_time' (Time), 'host' (String, Override), 'source' (String, Override), 'sourcetype' (String, Override), 'action' (String, Edit), 'bytes' (Number, Edit), 'categoryid' (String, Edit), 'change_type' (String, Edit), 'clientip' (IPv4, Edit), 'cookie' (String, Edit), and 'date_hour' (Number, Edit). There are also 'Edit' buttons for each row.

Event Object Hierarchy and Constraints

The screenshot illustrates the Splunk search interface, specifically focusing on event object hierarchy and constraints. The interface is organized into three main sections:

- Left Panel (Datasets):** Shows a tree structure of event types under "Web Requests".
 - Successful Requests:** Contains "purchases", "addtocart", "remove", and "Failed Requests" (with "failed purchase", "failed addtocart", "failed remove").
 - Failed Requests:** Contains "failed purchases", "failed addtocart", and "failed remove".
- Middle Panel (Search Results):** Displays search results for "base search – all access_combined events".
 - Constraint 1:** index=web sourcetype=access_combined (highlighted with a green box).
 - Constraint 2:** index=web sourcetype=access_combined status<400 (highlighted with a green box). This constraint is applied to the "Successful Requests" object.
 - Constraint 3:** index=web sourcetype=access_combined status<400 action=purchase productId=* (highlighted with a green box). This constraint is applied to the "purchases" object under "Successful Requests".
- Right Panel (Constraints Summary):** A summary box stating "Each constraint inherits the parent search string".

Annotations and callouts:

- A yellow callout box labeled "base search – all access_combined events" points to the search results.
- A yellow callout box labeled "successful requests" points to the second search result.
- A yellow callout box labeled "successful requests for purchases" points to the third search result.
- An annotation box on the right states "Each constraint inherits the parent search string".

Generated for Anh Hoang (hoang_anhkim@bah.com) (C) Splunk Inc, not for distribution

Dataset Fields

- Select the fields you want to include in the dataset
- Like constraints, fields are inherited from parent objects

The screenshot shows the 'Web Requests' dataset configuration page. At the top right are 'Rename' and 'Delete' buttons. Below that is a 'CONSTRAINTS' section with the constraint 'index=web sourcetype=access_combined'. To the right of the constraint are 'Constraint' and 'Edit' buttons. A 'Bulk Edit' dropdown is located below the constraint section. On the far right is an 'Add Field' button. The main area is divided into 'INHERITED' and 'EXTRACTED' sections. The 'INHERITED' section lists fields: '_time' (Time), 'host' (String), 'source' (String), and 'sourcetype' (String). The 'EXTRACTED' section lists fields: 'action' (String), 'bytes' (Number), 'categoryId' (String), 'change_type' (String), 'clientip' (IPv4), 'cookie' (String), 'date_hour' (Number), and 'status' (Number). Each field entry includes a checkbox, a data type, and an 'Edit' button.

Field	Type	Action
_time	Time	
host	String	Override
source	String	Override
sourcetype	String	Override
action	String	Edit
bytes	Number	Edit
categoryId	String	Edit
change_type	String	Edit
clientip	IPv4	Edit
cookie	String	Edit
date_hour	Number	Edit
status	Number	Edit

Generated for Anh Hoang (hoang_anhkim@bah.com) (C) Splunk Inc, not for distribution

Creating a Data Model

Settings > Data models

The screenshot shows the Splunk Enterprise interface with the title "Data Models". A green callout box points from the "New Data Model" button in the top right to the "New Data Model" dialog box. Inside the dialog, a yellow callout box points from the "Title" field to the text "ID is automatically populated from Title, but can be overridden". Another yellow callout box points from the "App" dropdown to the text "Choose app context". The dialog also includes fields for "ID", "App", and "Description".

Data Models

Data models enable users to easily create reports in the Pivot tool. [Learn More](#)

24 Data Models App: CLASS: Fundamentals 2 (class_Fund2) Created in the App Owner: Any filter 20 per page

New Data Model

ID is automatically populated from Title, but can be overridden

Choose app context

Title: Buttercup Games Site Activity
ID: Buttercup_Games_Site_Activity
App: Search & Reporting
Description: optional

Cancel Create

Generated for Anh Hoang (hoang_anhkim@bah.com) (C) Splunk Inc, not for distribution

Adding a Root Event

Note

Starting in Splunk 7.3, dataset constraints must specify at least one index.

The diagram illustrates the process of adding a root event dataset. It begins with a screenshot of the 'Buttercup Games Site Activity' dashboard, where the 'Root Event' button is highlighted. An arrow points from this button to the 'Add Event Dataset' configuration page. On this page, the 'Dataset Name' is set to 'Web Requests' and the 'Constraints' field contains the search term 'index=web sourcetype=access_combined'. A callout box explains that constraints are essentially search terms and suggests adding child events to narrow the search. The 'Save' button is also highlighted. Below the configuration, a preview window shows sample events, with a callout instructing to click 'Preview' to view the events that the constraint returns.

Add Event Dataset

Data Model: Buttercup Games Site Activity

Dataset Name: Web Requests

Dataset ID: Web_Requests

Constraints: index=web sourcetype=access_combined

Examples:

uri="*.php*" OR uri="*.py*"
NOT (referer=null OR referer="-")

Cancel Preview Save

Click Preview to view the events that the constraint returns

Event

90.205.111.169 - - [29/Jan/2018:18:56:34] "GET /oldlink?itemId=EST-15&JSESSIONID=SD3SL8FF8ADFF4965 HTTP 1.1" 200 2732 "http://www.buttercupgames.com/oldlink?itemId=EST-15" "Mozilla/4.0 (compatible; MSIE 7.0; Windows NT 5.1; Trident/4.0; .NET CLR 2.0.50727; MS-RTC LM 8; InfoPath.2)" 987

90.205.111.169 - - [29/Jan/2018:18:56:19] "POST /cart/success.do?JSESSIONID=SD3SL8FF8ADFF4965 HTTP 1.1" 200 867 "http://www.buttercupgames.com/cart.do?action=purchase&itemId=EST-6" "Mozilla/4.0 (compatible; MSIE 7.0; Windows NT 5.1; Trident/4.0; .NET CLR 2.0.50727; MS-RTC LM 8; InfoPath.2)" 735

90.205.111.169 - - [29/Jan/2018:18:56:19] "POST /cart.do?action=purchase&itemId=EST-6&JSESSIONID=SD3SL8FF8ADFF4965 HTTP 1.1" 200 3202 "http://www.buttercupgames.com/cart.do?action=addtocart&itemId=EST-6&categoryId=ARCADE&productId=BS-AG-C09" "Mozilla/4.0 (compatible; MSIE 7.0; Windows NT 5.1; Trident/4.0; .NET CLR 2.0.50727; MS-RTC LM 8; InfoPath.2)" 815

Generated for Anh Hoang (hoang_anhkim@bah.com) (C) Splunk Inc, not for distribution

Adding a Root Event (cont.)

- In this example, the root event of this data model represents all web requests
- The Inherited attributes are default fields
- Use **Add Field > Auto-Extracted** to add more fields

Buttercup Games Site Activity
Buttercup_Games_Site_Activity

[Edit](#) [Download](#) [Pivot](#) [Documentation](#)

[All Data Models](#)

Datasets [Add Dataset](#)

EVENTS

Web Requests

Web_Requests

[Rename](#) [Delete](#)

CONSTRAINTS

index=web sourcetype=access_combined

Constraint [Edit](#)

Bulk Edit [Add Field](#)

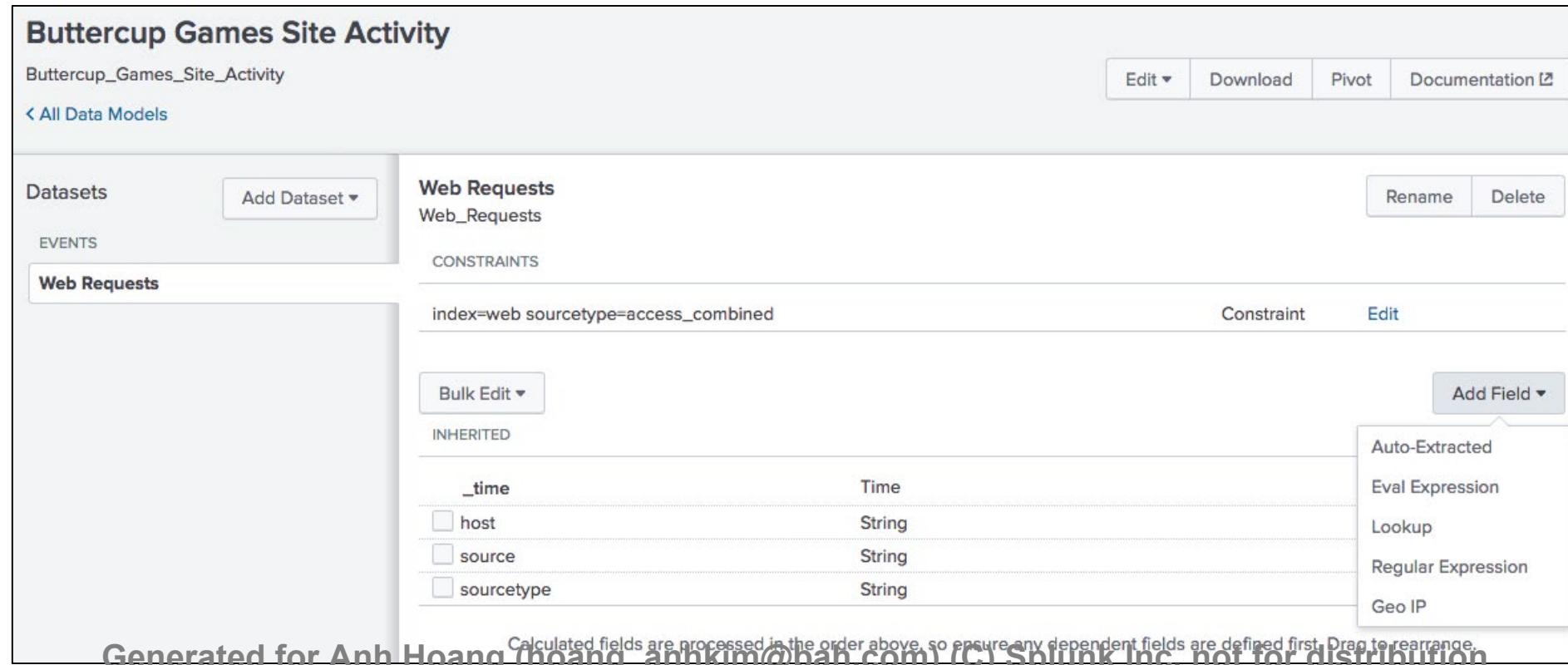
INHERITED

<input type="checkbox"/>	_time	Time
<input type="checkbox"/>	host	String
<input type="checkbox"/>	source	String
<input type="checkbox"/>	sourcetype	String

Calculated fields are processed in the order above, so ensure any dependent fields are defined first. Drag to rearrange.

Generated for Anh Hoang (hoang_anhkim@bah.com) (C) Splunk Inc, not for distribution

Auto-Extracted
Eval Expression
Lookup
Regular Expression
Geo IP



Adding Fields

- **Auto-Extracted** – can be default fields or manually extracted fields
- **Eval Expression** – a new field based on an expression that you define
- **Lookup** – leverage an existing lookup table
- **Regular Expression** – extract a new field based on regex
- **Geo IP** – add geographical fields such as latitude/longitude, country, etc.

The screenshot shows the 'Web Requests' field configuration page in Splunk. At the top, there are 'Rename' and 'Delete' buttons. Below that, the 'CONSTRAINTS' section contains the constraint 'index=web sourcetype=access_combined'. To the right of this are 'Constraint' and 'Edit' buttons. In the center, there's a 'Bulk Edit' dropdown and an 'INHERITED' section listing '_time' (Time), 'host' (String), 'source' (String), and 'sourcetype' (String). Below this is an 'EXTRACTED' section listing various fields: 'action' (String), 'bytes' (Number), 'categoryid' (String), 'change_type' (String), 'clientip' (IPv4), 'cookie' (String), 'date_hour' (Number), and 'status' (Number). Each extracted field has an 'Edit' button to its right. On the far right, a vertical dropdown menu is open, showing options: 'Auto-Extracted', 'Eval Expression', 'Lookup', 'Regular Expression', and 'Geo IP', all enclosed in a green rectangular box.

Adding Fields – Auto-Extracted

Fields that already exist for the constraint can be added as attributes to the data model

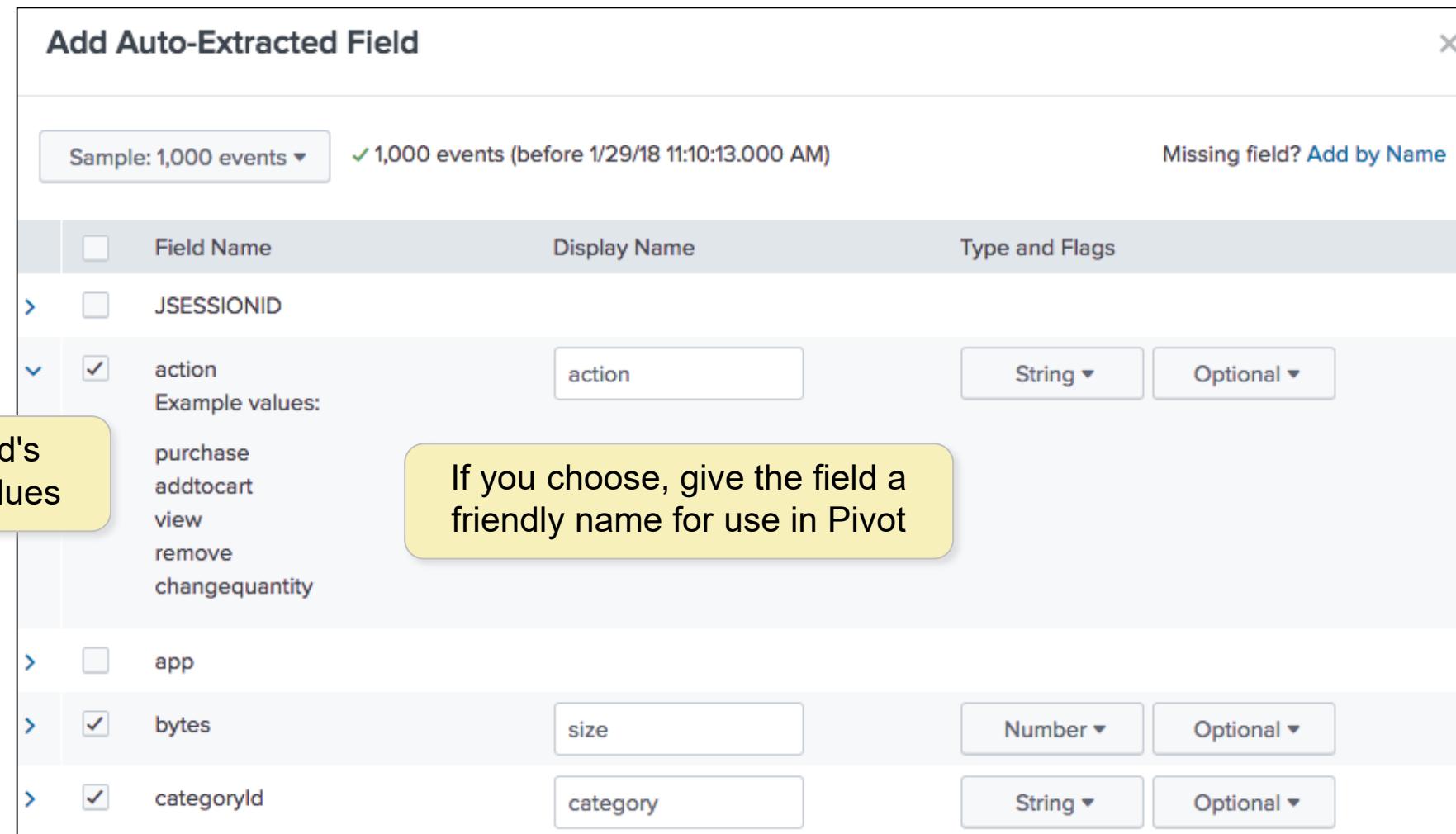
Add Auto-Extracted Field

Sample: 1,000 events ✓ 1,000 events (before 1/29/18 11:10:13.000 AM) Missing field? Add by Name

	Field Name	Display Name	Type and Flags
>	<input type="checkbox"/> JSESSIONID		
▼	<input checked="" type="checkbox"/> action	action	String ▾ Optional ▾
	Example values:	purchase addtocart view remove changequantity	
>	<input type="checkbox"/> app		
>	<input checked="" type="checkbox"/> bytes	size	Number ▾ Optional ▾
>	<input checked="" type="checkbox"/> categoryId	category	String ▾ Optional ▾

View a field's example values

If you choose, give the field a friendly name for use in Pivot

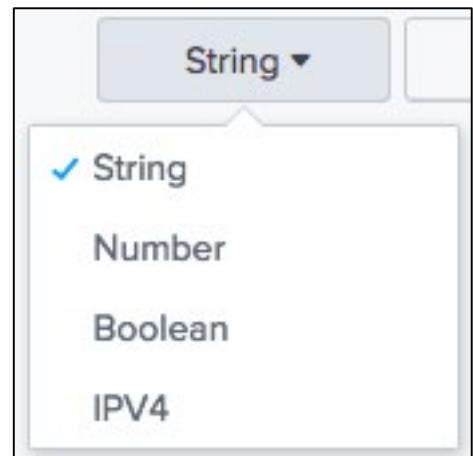


Note

Providing a friendly name for display in Pivot does not change the name of the field as it appears in searches. So remember that there will now be two different names referencing the same field, depending on whether or not you're in Pivot.

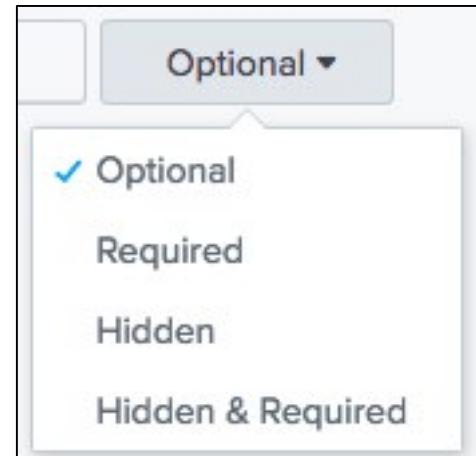
Field Types

- String: Field values are recognized as alphanumeric
- Number: Field values are recognized as numeric
- Boolean: Field values are recognized as true/false or 1/0
- IPV4: Field values are recognized as IP addresses
 - This is an important field type, as at least one IPV4 attribute type must be present in the data model in order to add a Geo IP attribute



Field Flags

- Optional: This field doesn't have to appear in every event
- Required: Only events that contain this field are returned in Pivot
- Hidden: This field is not displayed to Pivot users when they select the dataset in Pivot
 - Use for fields that are only being used to define another field, such as an eval expression
- Hidden & Required: Only events that contain this field are returned, and the fields are hidden from use in Pivot



Adding Fields – Eval Expressions

- You can define a new field using an eval expression
 - In this example, you create a field named Error Reason that evaluates the value of the status field

The screenshot shows the Splunk interface for creating a new field. On the left, a sidebar menu has 'Eval Expression' selected. A green arrow points from this selection to the main configuration window. The main window is titled 'Add Fields with an Eval Expression' and specifies 'Data Model: Buttercup Games Site Activity' and 'Dataset: Web Requests'. The 'Eval Expression' field contains the code: `if(status>399,"Web error","OK")`. To the right, a 'Field' section is configured with 'Field Name: errorReason', 'Display Name: Error Reason', 'Type: String', and 'Flags: Optional'. A yellow callout bubble says 'Click Preview to verify your eval expression returns events'. Below the preview button, there's a sample event table showing four log entries. The table includes columns for _time, errorReason, host, source, sourcetype, JSESSIONID, action, app, bytes, categoryId, and change_type.

_time	errorReason	host	source	sourcetype	JSESSIONID	action	app	bytes	categoryId	change_type
2018-01-29 11:14:59	OK	www3	/opt/log/www3/access.log	access_combined	SD2SL2FF3ADFF4952			1327	STRATEGY	
2018-01-29 11:14:51	OK	www3	/opt/log/www3/access.log	access_combined	SD2SL2FF3ADFF4952	changequantity		1639		
2018-01-29 11:14:37	OK	www3	/opt/log/www3/access.log	access_combined	SD2SL2FF3ADFF4952	view		800		
2018-01-29 11:14:22	OK	www3	/opt/log/www3/access.log	access_combined	SD2SL2FF3ADFF4952			1978	ARCADE	

Generated for Anh Hoang (hoang_anhkim@bah.com) (C) Splunk Inc, not for distribution

Adding Fields – Lookups

- Leverage an existing lookup definition to add fields to your event object
- Configure the lookup attribute in the same way as an automatic lookup

Add Fields with a Lookup

Data Model: Buttercup Games Site Activity Dataset: Web Requests

Lookup Table: http_status_lookup

Input:

Field in Lookup: Field in Dataset:
code = status Remove

Add New

Output:

Field in Lookup: Field in Dataset: code Display Name: Type: Flags:
 code code String Optional
 description description String Optional

Auto-Extracted
Eval Expression
Lookup
Regular Expression
Geo IP

Cancel Preview Save

Generated for Anh Hoang (hoang_anhkim@bah.com) (C) Splunk Inc, not for distribution

Adding Fields – Lookups (cont.)

- Use Preview to test your lookup settings
- Use the Events and Values tab to verify your results

Add Fields with a Lookup

Data Model: Buttercup Games Site Activity Dataset: Web Requests

Documentation

Lookup Table

http_status_lookup

Events Values

✓ 1,000 events (before 1/29/18 11:59:20.000 AM)

Sample: 1,000 events

20 per page

Input

Field in Lookup: Field in Dataset:
code = status Remove

Add New Output

Field in Lookup: Field in Dataset:
 code code
 description description

Events Values

✓ 1,000 events (before 1/29/18 11:59:20.000 AM)

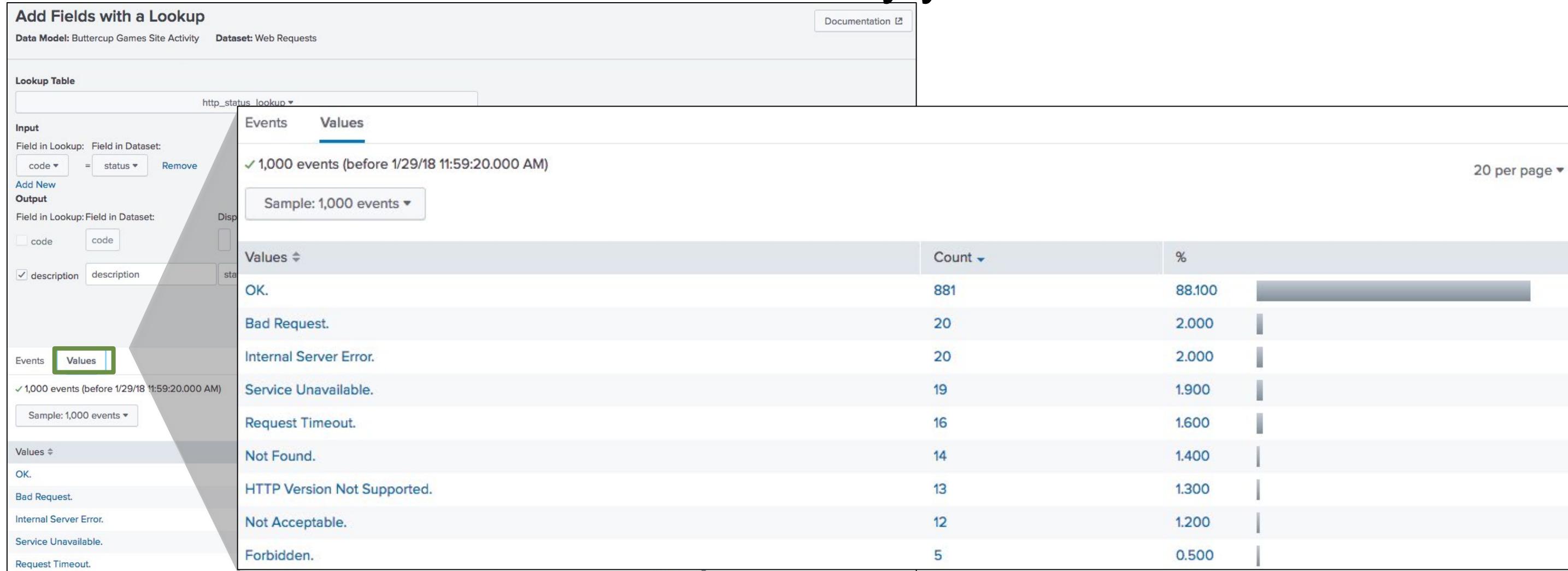
Sample: 1,000 events

Values

	Count	%
OK.	881	88.100
Bad Request.	20	2.000
Internal Server Error.	20	2.000
Service Unavailable.	19	1.900
Request Timeout.	16	1.600
Not Found.	14	1.400
HTTP Version Not Supported.	13	1.300
Not Acceptable.	12	1.200
Forbidden.	5	0.500

Values

OK.
Bad Request.
Internal Server Error.
Service Unavailable.
Request Timeout.
Not Found.
HTTP Version Not Supported.
Not Acceptable.
Forbidden.



Generated for Anh Hoang (hoang_anhkim@bah.com) (C) Splunk Inc, not for distribution

Adding Fields – Regular Expression

You can define a new field using a regular expression

Add Fields with a Regular Expression

Data Model: Buttercup Games Site Activity Dataset: Web Requests Documentation

Extract From: _raw Regular Expression: userAgent=(?<browser>[()]+)

Field(s): Field Name: browser Display Name: browser Type: String Flags: Optional

Example: From: (?<from>.* To: (?<to>.*))

Learn More

Events: browser

✓ 1,000 events (before 1/29/18 12:12:34.000 PM) 20 per page < Prev **1** 2 3 4 5 6 7 8 9 ... Next >

filter Apply Sample: 1,000 events All events All Events Matches Non-Matches

_raw

```
173.192.201.242 - - [29/Jan/2018:20:12:30] "GET /category.screen?categoryId=SHOOTER&JSESSIONID=SD10SL5FF8ADFF4952 HTTP/1.1" 200 1814 "http://www.buttercupgames.com/product.screen?"  
173.192.201.242 - - [29/Jan/2018:20:12:17] "GET /category.screen?categoryId=STRATEGY&JSESSIONID=SD10SL5FF8ADFF4952 HTTP/1.1" 200 1032 "http://www.buttercupgames.com/cart.do?action=...  
173.192.201.242 - - [29/Jan/2018:20:12:11] "GET /cart.do?action=addtocart&itemId=EST-13&productId=FS-SG-G03&JSESSIONID=SD10SL5FF8ADFF4952 HTTP/1.1" 200 3700 "http://www.buttercupg...  
173.192.201.242 - - [29/Jan/2018:20:11:55] "GET /oldlink?itemId=EST-6&JSESSIONID=SD10SL5FF8ADFF4952 HTTP/1.1" 200 2096 "http://www.buttercupgames.com/product.screen?productId=MB-AG...  
173.192.201.242 - - [29/Jan/2018:20:11:49] "POST /cart.do?action=remove&itemId=EST-15&productId=DC-SG-G02&JSESSIONID=SD10SL5FF8ADFF4952 HTTP/1.1" 200 2259 "http://www.yahoo.com"  
233.77.49.94 - - [29/Jan/2018:20:09:13] "GET /product.screen?productId=BS-AG-G09&JSESSIONID=SD2SL5FF5ADFF4952 HTTP/1.1" 200 3473 "http://www.buttercupgames.com/oldlink?itemId=EST-..."
```

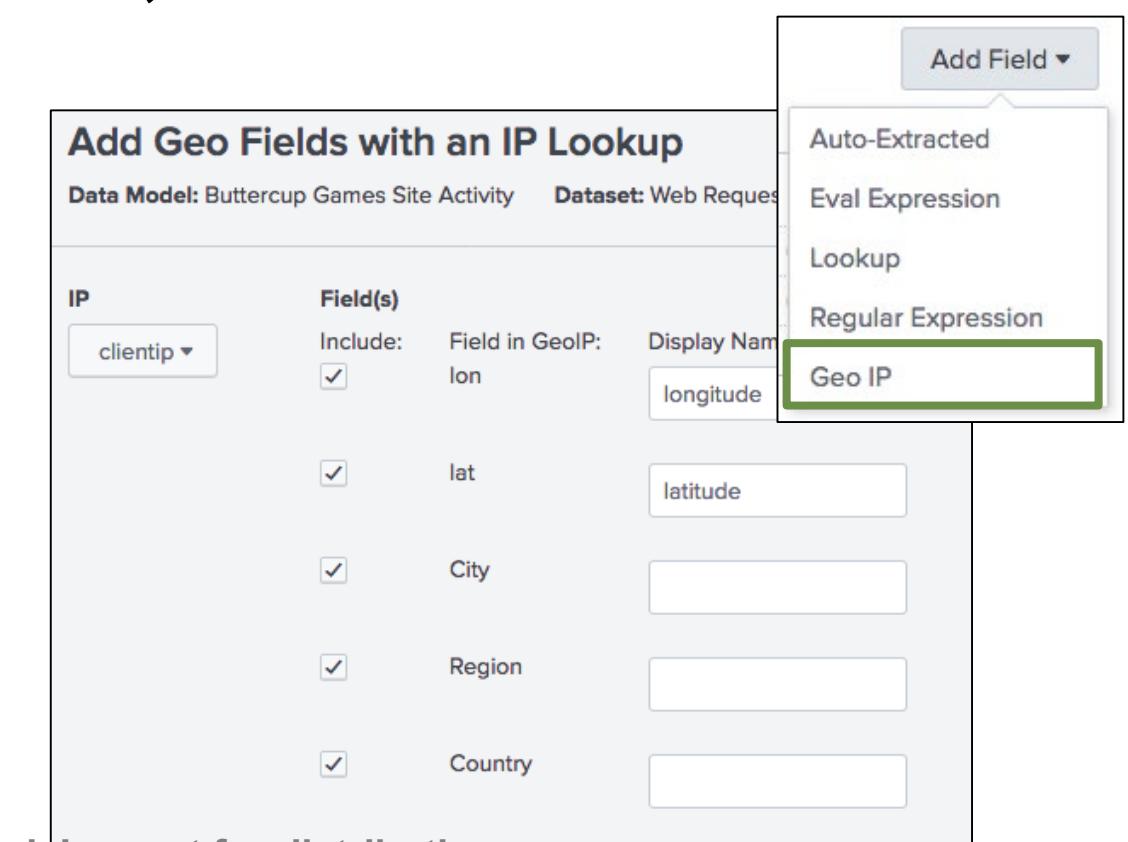
Generated for Anh Hoang (hoang_anhkim@bah.com) (C) Splunk Inc, not for distribution

Add Field ▾

- Auto-Extracted
- Eval Expression
- Lookup
- Regular Expression**
- Geo IP

Adding Fields - GeoIP

- Map visualizations require latitude/longitude fields
- To use Geo IP Lookup, at least one IP field must be configured as an IPv4 type
- While the map function isn't available in Pivot, the data model can be called using the `| pivot` command and `<map>` element in a dashboard population search
 - Select the field that contains the mapping to lat/lon
 - Identify the lat/lon and geo fields in the data



Generated for Anh Hoang (hoang_anhkim@bah.com) (C) Splunk Inc, not for distribution

Adding Child Datasets

When you create a new child dataset, you give it one or more additional constraints

Buttercup Games Site Activity

Buttercup_Games_Site_Activity

[Edit](#) [Download](#) [Pivot](#) [Documentation](#)

[All Data Models](#)

Datasets

EVENTS

Web Requests

Root Event
Root Transaction
Root Search
Child

Add Dataset ▾

Web Requests
Web_Requests

CONSTRAINTS

index=web sour

Add Child Dataset

Data Model: Buttercup Games Site Activity

Dataset Name: Successful Requests

Additional Constraints: status<400

All events that have a status less than 400 (successful http request)

Dataset ID: Successful_Requests

Inherit From: Web Requests

Examples:
uri="*.php*" OR uri="*.py*"
NOT (referer=null OR referer="-")

Field ▾

Generated for Anh Hoang (hoang_anhkim@bah.com) (C) Splunk Inc, not for distribution

Adding Child Datasets (cont.)

- Child datasets inherit all fields from the parent events
- You can add more fields to child datasets

The screenshot shows the Splunk Data Model Editor interface. At the top, there's a title bar with 'Successful Requests' and 'Successful_Requests'. On the right, there are 'Rename' and 'Delete' buttons. Below the title, there's a 'CONSTRAINTS' section with two entries: 'index=web sourcetype=access_combined' and 'status<400'. To the right of these constraints are 'Inherited' and 'Constraint' buttons, with 'Edit' next to the constraint button. Underneath the constraints, there's a 'Bulk Edit ▾' button. The main area is divided into 'INHERITED' and 'ADDED' sections. The 'INHERITED' section contains fields: '_time' (Time), 'clientip' (IPv4), 'host' (String), 'source' (String), 'sourcetype' (String), and 'status' (Number). The 'ADDED' section contains fields: '_index' (Auto-Extracted, highlighted in blue), '_score' (Eval Expression, also highlighted in blue), and 'geoip' (Geo IP). A context menu is open over the 'Eval Expression' option in the 'ADDED' section, listing other options: 'Auto-Extracted', 'Eval Expression' (selected and highlighted in blue), 'Lookup', 'Regular Expression', and 'Geo IP'. There's also an 'Override' button at the bottom right of the 'ADDED' section.

Generated for Anh Hoang (hoang_anhkim@bah.com) (C) Splunk Inc, not for distribution

Testing the Data Model

- Click Pivot to access the Select a Dataset window
- Choose an object from the selected data model to begin building the report

Buttercup Games Site Activity

Buttercup_Games_Site_Activity

< All Data Models

Datasets Add Dataset ▾

EVENTS

Web Requests

Web_Requests

CONSTRAINTS

index=web sourcetype=access_combined

Bulk Edit ▾

INHERITED

_time Time

host String

source String

sourcetype String

EXTRACTED

Constraint

Edit

Failed Requests

11 Objects in Buttercup Games Site Activity

Web Requests

Successful Requests

purchases

addtocart

remove

Failed Requests

failed purchases

failed addtocart

failed remove

visit duration

User

Generated for Anh Hoang (hoang_anhkim@bah.com) (C) Splunk Inc, not for distribution

Using the Data Model in Pivot

The New Pivot window automatically populates with a count of events for the selected dataset

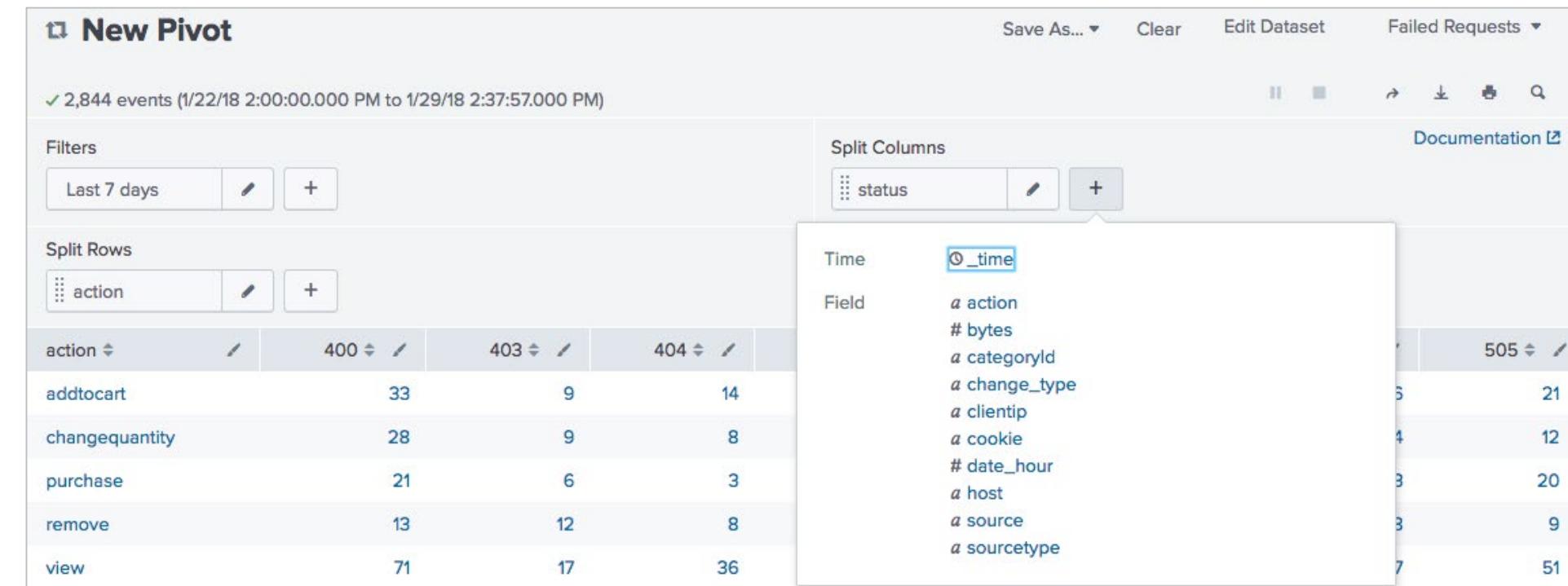
The screenshot illustrates the Splunk Pivot interface. On the right, a 'Select a Dataset' dropdown menu is open, listing various datasets under '11 Objects in Buttercup Games Site Activity'. One item, 'Failed Requests', is highlighted with a green border and a green arrow points from it to the corresponding value in the Pivot window below. The Pivot window itself shows a summary of 2,844 events from January 22 to 29, 2018. It includes sections for Filters (Last 7 days), Split Rows, and a main table. The table has a single row with 'Count of Failed Requests' set to '2844'. The 'Failed Requests' dataset is also listed in the Pivot's 'Column Values' section.

Dataset
11 Objects in Buttercup Games Site Activity
Web Requests
Successful Requests
purchases
addtocart
remove
Failed Requests
failed purchases
failed addtocart
failed remove
visit duration
User

Generated for Anh Hoang (hoang_anhkim@bah.com) (C) Splunk Inc, not for distribution

Pivot – Using Fields

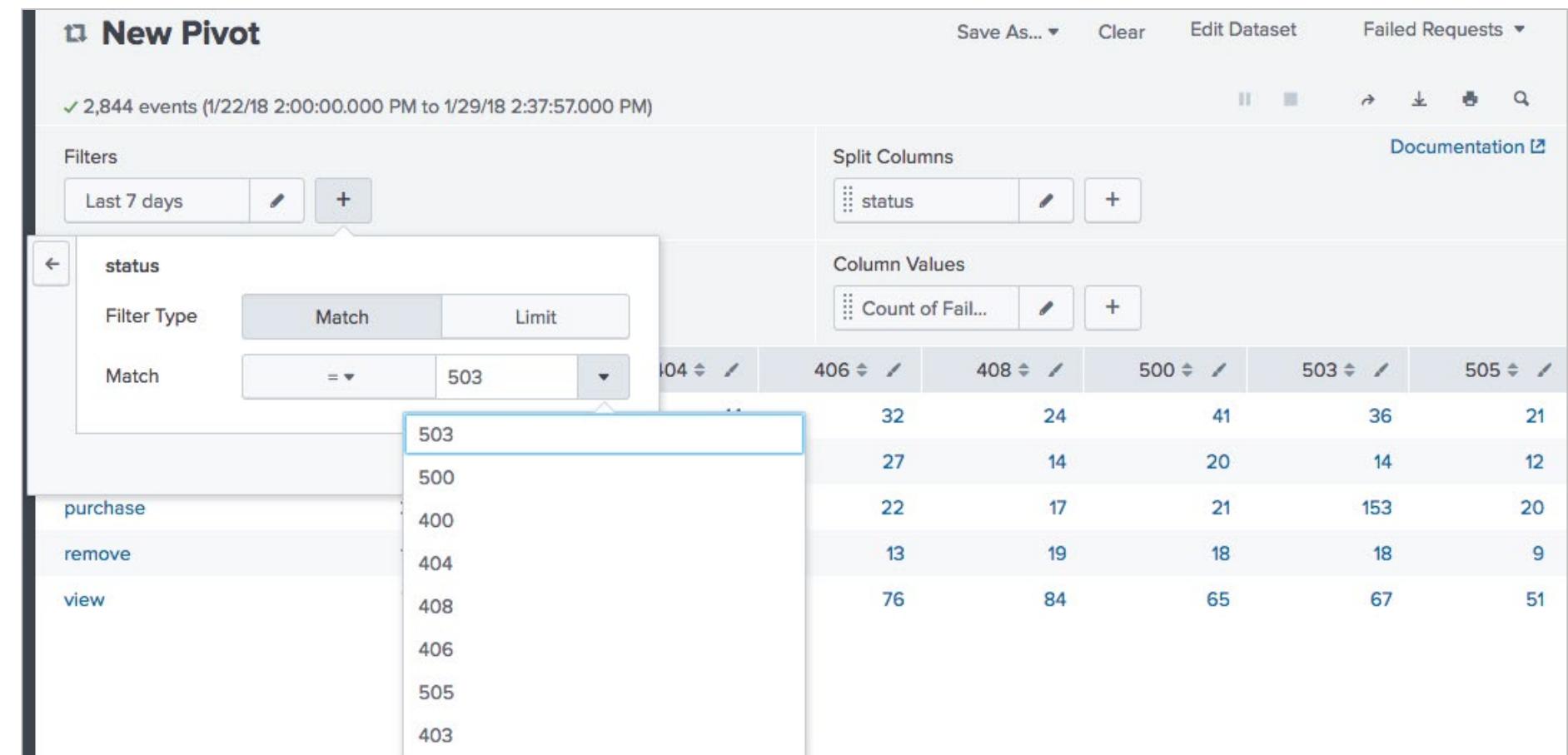
- The fields associated with each dataset are available as splits for rows or columns
- In this example, the Pivot report will show a count of failed request actions by status



Generated for Anh Hoang (hoang_anhkim@bah.com) (C) Splunk Inc, not for distribution

Pivot – Using Fields (cont.)

- Fields can also be used to filter events in the Pivot interface
- In this example, the Pivot report is filtered to only return results where status=503



Generated for Anh Hoang (hoang_anhkim@bah.com) (C) Splunk Inc, not for distribution

Underlying Search

The screenshot shows two main sections of a Splunk search interface. The top section is a "New Pivot" window with a search bar containing the command:

```
| pivot Buttercup_Games_Site_Activity Failed_Requests count(Failed_Requests) AS "Count of Failed Requests" SPLITROW action AS action  
SPLITCOL status FILTER status = 503 SORT 100 action ROWSUMMARY 0 COLSUMMARY 0 NUMCOLUMNS 1000 SHOWOTHER 0
```

The bottom section is a "New Search" window titled "New Search". It displays the same search command in its search bar. The search results table shows the following data:

Action	Count
addtocart	36
changequantity	14
purchase	153
remove	18
view	67

Both windows have a green box highlighting the search bar area.

Generated for Anh Hoang (hoang_anhkim@bah.com) (C) Splunk Inc, not for distribution

Underlying Search (cont.)

Data model name

Object name

```
| pivot Buttercup_Games_Site_Activity Failed_Requests  
count(Failed_Requests) AS "Count of Failed requests"
```

Split row field (or attribute)

```
SPLITROW action AS action TOP 100  
count(failed_request)
```

Split column field (or attribute) and filter field/value pair

```
SPLITCOL status  
FILTER status = 503
```

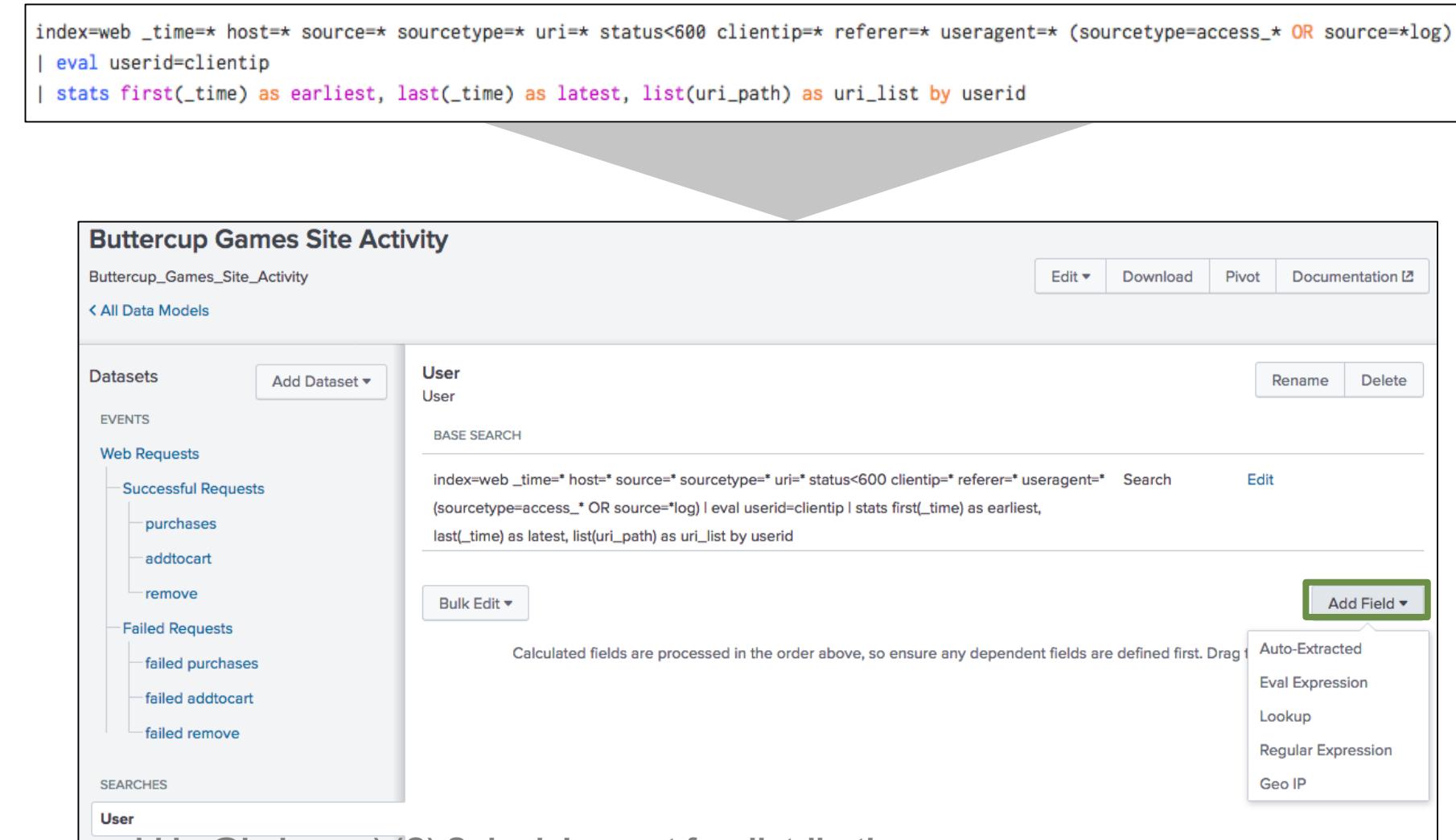
Generated for Anh Hoang (hoang_anhkim@bah.com) (C) Splunk Inc, not for distribution

Data Model Search Datasets

- As mentioned, a data model can consist of 3 types of datasets: **Events** (the kind of dataset we've been discussing so far), **Searches**, and **Transactions**
- A **search dataset** defines a dataset based on a search that includes transforming commands
- Search datasets can also have fields, which you can add via the **Add Field** button

Note

Starting in Splunk 7.3, dataset constraints must specify at least one index.



```
index=web _time=* host=* source=* sourcetype=* uri=* status<600 clientip=* referer=* useragent=* (sourcetype=access_* OR source=log)
| eval userid=clientip
| stats first(_time) as earliest, last(_time) as latest, list(uri_path) as uri_list by userid
```

The screenshot shows the Splunk Data Model Editor interface. A search dataset named "Buttercup Games Site Activity" is selected. The search command is displayed in a code editor window. The interface includes a sidebar with "Datasets" and "Events" sections, and a main panel with "User" and "BASE SEARCH" sections. A note at the bottom states: "Calculated fields are processed in the order above, so ensure any dependent fields are defined first. Drag and drop fields to change their order." A dropdown menu on the right labeled "Add Field" offers options: Auto-Extracted, Eval Expression, Lookup, Regular Expression, and Geo IP.

Generated for Anh Hoang (hoang_anhkim@bah.com) (C) Splunk Inc, not for distribution

Data Model Transaction Datasets

- A **transaction dataset** defines a dataset based on a transaction
- Uses fields that have already been added to the data model via either event or search datasets

Buttercup Games Site Activity

Buttercup_Games_Site_Activity

[Edit](#) [Download](#) [Pivot](#) [Documentation](#)

[Rename](#) [Delete](#)

Datasets [Add Dataset](#)

EVENTS

Web Requests

- Successful Requests
 - purchases
 - addtocart
 - remove
- Failed Requests
 - failed purchases
 - failed addtocart
 - failed remove

SEARCHES

User

TRANSACTIONS

visit duration

visit duration
visit_duration

CONSTRAINTS

Group Datasets	Web_Requests	Transaction
Group By	clientip	Edit
Max Pause	10s	
Max Span		

Bulk Edit [Add Field](#)

INHERITED

_time	Time	Required
duration	Number	Required
eventcount	Number	Required
host	String	Override
source	String	Override
sourcetype	String	Override

EXTRACTED

action	String	Edit
bytes	Number	Edit
categoryid	String	Edit
change_type	String	Edit

Generated for Anh Hoang (hoang_anhkim@bah.com) (C) Splunk Inc, not for distribution

Adding a Transaction Dataset

The transaction dataset below would equate to the search:
sourcetype=access_* | transaction clientip maxpause=10s

Add Transaction Dataset

Data Model: Buttercup Games Site Activity

You must specify at least one of the optional fields.

Dataset Name: visit duration

Dataset ID ?: visit_duration

Can only contain letters, numbers and underscores.

Select a dataset from the data model to base the transaction on

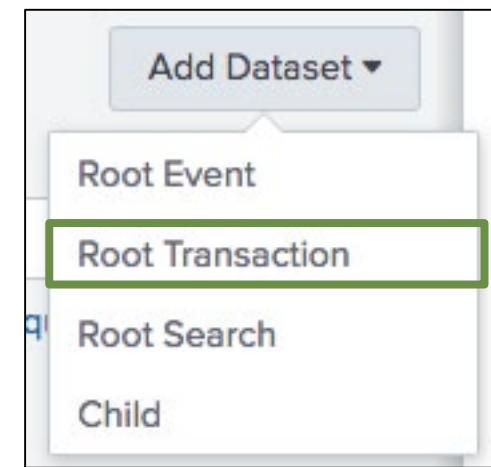
Group Datasets: Web Requests

Group by: clientip

Duration: Max Pause: 10 Seconds

Max Span: Seconds

Select a group by field



Optionally select max pause and max span

Adding a Transaction Dataset (cont.)

- You can add an eval expression or any other field to your transaction to further define the results
- This example shows dividing the duration field value by 60 to convert the duration field to minutes

visit duration
visit_duration

CONSTRAINTS

Group Datasets	Web_Requests	Transaction
Group By	clientip	Edit
Max Pause	10s	
Max Span		

Bulk Edit ▾

INHERITED

time	Time	Required
duration	Number	Required
eventcount	Number	Required
host	String	
source	String	

Add Field ▾

- Auto-Extracted
- Eval Expression
- Lookup
- Regular Expression
- Geo IP

Add Fields with an Eval Expression

Data Model: Buttercup Games Site Activity Dataset: visit duration

Documentation ↗

Eval Expression

```
duration/60
```

Field

Field Name:	Display Name:	Type:	Flags:
visitDuration	visitDuration	Number ▾	Optional ▾

Generated for Anh Hoang (hoang_anhkim@bah.com) (C) Splunk Inc, not for distribution

Search and Transaction Dataset Considerations

- There must be at least one event or search dataset before adding a transaction dataset
- Search and transaction datasets cannot benefit from persistent data model acceleration
- As you learn to create data models, consider the types of reports your users will run
 - Can the same report be achieved with event datasets?
 - Will users need raw events or transactional data?

Note



Data model acceleration is discussed briefly later in this module, and in-depth in *Splunk Fundamentals 3*.

Set Permissions

- When a data model is created, the owner can determine access based on the following permissions:
 - Who can see the data models
 - Owner, App, or All Apps
 - Which users can perform which actions (Read/Write)
 - Everyone
 - Power
 - User
 - Admin-defined roles, if applicable

Edit Permissions

Data Model	Buttercup Games Site Activity
Owner	student1
App	search
Display For	<input type="radio"/> Owner <input checked="" type="radio"/> App <input type="radio"/> All apps
Everyone	<input checked="" type="checkbox"/> Read <input type="checkbox"/> Write
power	<input type="checkbox"/> Read <input type="checkbox"/> Write
student	<input type="checkbox"/> Read <input type="checkbox"/> Write
user	<input type="checkbox"/> Read <input type="checkbox"/> Write

Cancel Save

Download and Upload Data Models

- Use the Splunk Web interface to download or upload data models:
 - Back up important data models
 - Collaborate with other Splunk users to create/modify/test data models
 - Move data models from a test environment to production instance

Downloading a Data Model

Buttercup Games Site Activity

Buttercup_Games_Site_Activity

< All Data Models

Datasets Add Dataset ▾

EVENTS

Web Requests

Web_Requests

CONSTRAINTS

index=web sourcetype=access_combined

Bulk Edit ▾

INHERITED

_time Time

host String

Note

An HTML 5 supported browser must be used to download data models.

Constraint Edit

Opening Buttercup_Games_Site_Activity.json

You have chosen to open:

Buttercup_Games_Site_Activity.json
which is: JSON file (1.3 KB)
from: http://54.201.235.11

What should Firefox do with this file?

Open with Sublime Text (default)

Save File

Do this automatically for files like this from now on.

Cancel OK

Generated for Anh Hoang (hoang_anhkim@bah.com) (C) Splunk Inc, not for distribution

Uploading a Data Model

The screenshot illustrates the process of uploading a data model in Splunk Enterprise. It consists of three main panels:

- Top Panel:** The Splunk Enterprise navigation bar with links for student1, Messages, Settings, Activity, Help, Find, and a search bar. Below it is the "Data Models" page, which lists 27 data models. A green box highlights the "Upload Data Model" button, and a green arrow points from this button to the "Upload New Data Model" dialog.
- Middle Panel:** The "Upload New Data Model" dialog. It shows a file input field containing "Buttercup_Games_Site_Activity.json", an ID field with "Buttercup_Games_Site_Activity_AC", and an "App" dropdown set to "Search & Reporting".
- Bottom Panel:** The "Buttercup Games Site Activity" data model configuration page. It includes sections for "Datasets" (with "Web Requests" selected), "Web Requests" (listing "Web_Requests" with constraints like "index=web sourcetype=access_combined"), and "INHERITED" fields. Buttons for "Upload" (highlighted with a green box) and "Edit", "Download", "Pivot", and "Documentation" are also visible.

Generated for Anh Hoang (hoang_anhkim@bah.com) (C) Splunk Inc, not for distribution

Data Model Acceleration

- Uses automatically created summaries to speed completion times for pivots
- Takes the form of inverted time-series index files (`tsidx`) that have been optimized for speed
- Discussed in more detail in *Splunk Fundamentals 3*

Note



Reports can also be accelerated, as discussed in *Splunk Fundamentals 3*.

Accelerating a Data Model

- With persistent data model acceleration, all fields in the model become "indexed" fields
- You must have administrative permissions or the **accelerate_datamodel** capability to accelerate a data model
- Private data models cannot be accelerated
- Accelerated data models cannot be edited

The screenshot shows the Splunk Data Models interface. At the top, there are buttons for 'Upload Data Model' and 'New Data Model'. Below that, a search bar contains the text 'buttercup'. A green arrow points from the 'Edit Acceleration' option in the context menu of a data model named 'Buttercup Games Online Sales' to the 'Edit Acceleration' dialog window.

The 'Edit Acceleration' dialog window is open. It shows the 'Data Model' set to 'Buttercup Games Online Sales'. The 'Accelerate' checkbox is checked. A note below it states: 'Acceleration may increase storage and processing costs.' The 'Summary Range' dropdown is set to '1 Day', with other options including '7 Days', '1 Month', '3 Months', '1 Year', 'All Time', and 'Custom'. A green 'Save' button is at the bottom right. A yellow 'Note' box in the top right corner contains the text: 'Only root events can be accelerated. If there are multiple root events, only the first root event is accelerated.'

Generated for Anh Hoang (hoang_anhkim@bah.com) (C) Splunk Inc, not for distribution

Module 12 Lab Exercise

Time: 30 minutes

Tasks:

- Create a data model using the **access_combined** source type
- Create a dashboard and add a Pivot report using your data model

NOTE: You will save the data model you create to the Search & Reporting app, **NOT** the CLASS: Fundamentals 2 app

Module 13: Using the Common Information Model (CIM) Add-On

Generated for Anh Hoang (hoang_anhkim@bah.com) (C) Splunk Inc, not for distribution

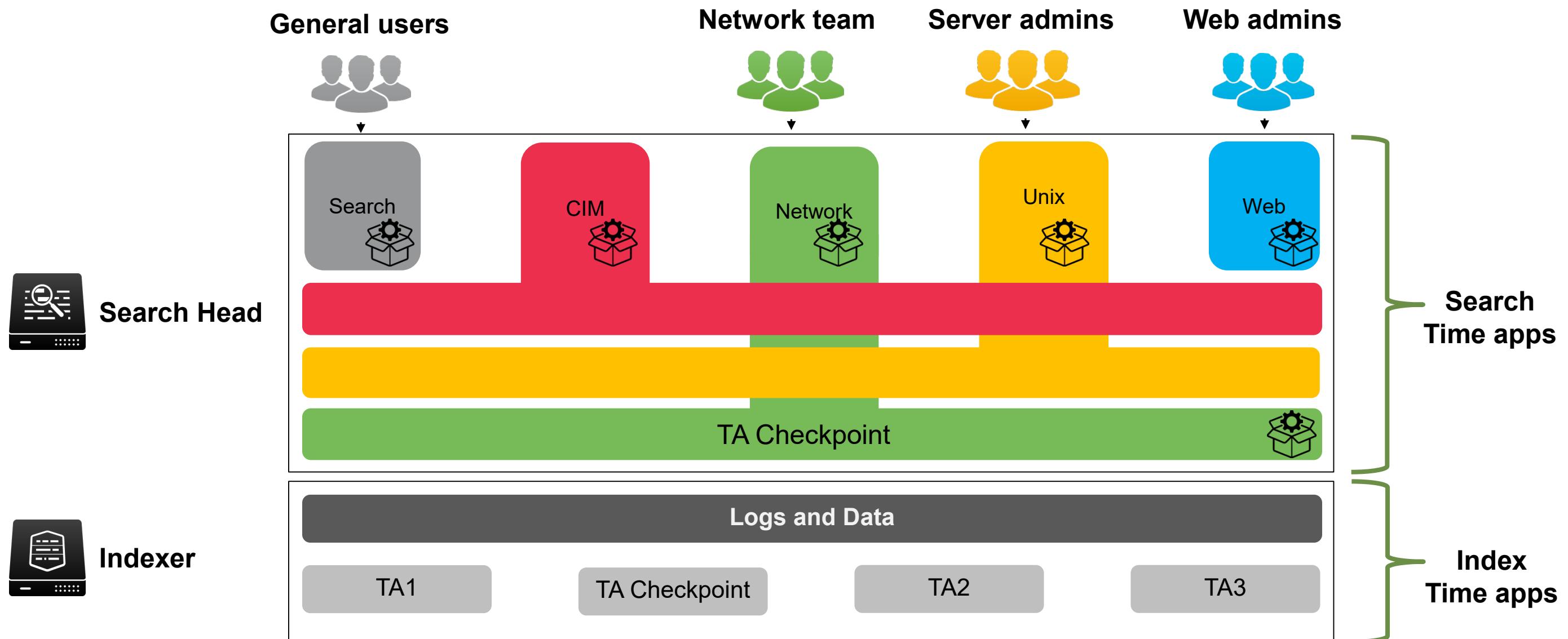
Objectives

- Describe the Splunk Common Information Model
- List the knowledge objects included with the Splunk CIM Add-on
- Use the CIM Add-on to normalize data

What is the Common Information Model (CIM)?

- The Splunk Common Information Model provides a methodology to normalize data
- Leverage the CIM when creating field extractions, field aliases, event types, and tags to ensure:
 - Multiple apps can co-exist on a single Splunk deployment
 - Object permissions can be set to global for the use of multiple apps
 - Easier and more efficient correlation of data from different sources and source types

How the Splunk CIM Works



Generated for Anh Hoang (hoang_anhkim@bah.com) (C) Splunk Inc, not for distribution

Normalized Field Names – Email Data

Field name	Data type	Description	Possible values
action	string	Action taken by the reporting device.	delivered, blocked, quarantined, deleted, unknown
duration	number	The amount of time for the completion of the messaging event, in seconds.	Email
src	string	The system that sent the message. May be aliased from more specific fields such as src_host, src_ip, or src_name.	

Normalized Field Names – Network Traffic

Field name	Data type	Description	Possible values
action	string	The action taken by the network device.	allowed, blocked, dropped, unknown
bytes	number	Total count of bytes handled by this device/interface (bytes_in + bytes_out).	
bytes_in	number	How many bytes this device/interface received.	
bytes_out	number	How many bytes this device/interface transmitted.	
src	string	The source of the network traffic (the client requesting the connection). May be aliased from more specific fields such as src_host, src_ip, or src_name.	

Generated for Anh Hoang (hoang_anhkim@bah.com) (C) Splunk Inc, not for distribution

Normalized Field Names – Web Data

Field name	Data type	Description	Possible values
action	string	The action taken by the server or proxy.	
duration	number	The time taken by the proxy event, in milliseconds.	
http_method	string	The HTTP method used in the request.	GET, PUT, POST, DELETE, etc.
src	string	The source of the network traffic (the client requesting the connection).	
status	string	The HTTP response code indicating the status of the proxy request.	404, 302, 500, and so on.

Generated for Anh Hoang (hoang_anhkim@bah.com) (C) Splunk Inc, not for distribution

Splunk CIM Add-on

- Set of 22 pre-configured data models
 - Fields and event category tags
 - Least common denominator of a domain of interest
- Leverage the CIM so that knowledge objects in multiple apps can co-exist on a single Splunk deployment
- Available on splunkbase:
 - <https://splunkbase.splunk.com/app/1621/>
- Use the CIM Reference Tables
 - <https://docs.splunk.com/Documentation/CIM/4.9.0/User/Howtousethesreferencetables>

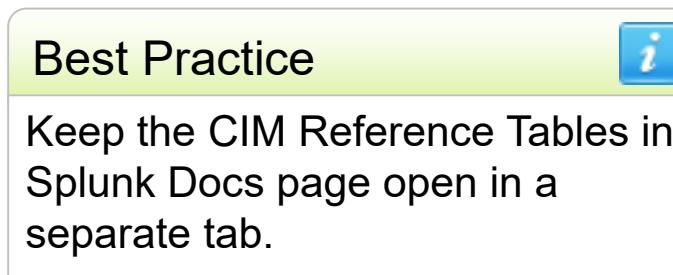
Splunk CIM Add-On Data Models	
Alerts	Java Virtual Machines (JVM)
Application State	Malware
Authentication	Network Resolution (DNS)
Certificates	Network Sessions
Change Analysis	Network Traffic
CIM Validation (S.o.S)	Performance
Databases	Splunk Audit Logs
Email	Ticket Management
Interprocess Messaging	Updates
Intrusion Detection	Vulnerabilities
Inventory	Web

Note

The data models included in the CIM add-on are configured with data model acceleration turned off.

Using the CIM Add-on

1. Examine your data
 - Go to **Settings > Data models**
 - Identify a data model relevant to your dataset



Data Models

Data models enable users to easily create reports in the Pivot tool. [Learn More](#)

24 Data Models App: CLASS: Fundamentals 2 (class_Fund2) ▾ Visible in the App ▾ Owner: Any ▾ filter 20 per page ▾

i	Title ▾	Type	Actions	App	Owner	Sharing
>	Alerts	data model	Clone Pivot	Splunk_SA_CIM	nobody	Global
>	Application State	data model	Clone Pivot	Splunk_SA_CIM	nobody	Global
>	Authentication	data model	Clone Pivot	Splunk_SA_CIM	nobody	Global
>	Buttercup Games Online Sales	data model	Clone Pivot	ao-bcg	nobody	Global
>	Certificates	data model	Clone Pivot	Splunk_SA_CIM	nobody	Global
>	Change Analysis	data model	Clone Pivot	Splunk_SA_CIM	nobody	Global
>	CIM Validation (S.o.S.)	data model	Clone Pivot	Splunk_SA_CIM	nobody	Global
>	Data Loss Prevention	data model	Clone Pivot	Splunk_SA_CIM	nobody	Global
>	Databases	data model	Clone Pivot	Splunk_SA_CIM	nobody	Global
>	Email	data model	Clone Pivot	Splunk_SA_CIM	nobody	Global
>	Interprocess Messaging	data model	Clone Pivot	Splunk_SA_CIM	nobody	Global
>	Intrusion Detection	data model	Clone Pivot	Splunk_SA_CIM	nobody	Global
>	Inventory	data model	Clone Pivot	Splunk_SA_CIM	nobody	Global
>	JVM	data model	Clone Pivot	Splunk_SA_CIM	nobody	Global
>	Malware	data model	Clone Pivot	Splunk_SA_CIM	nobody	Global
>	Network Resolution (DNS)	data model	Clone Pivot	Splunk_SA_CIM	nobody	Global
>	Network Sessions	data model	Clone Pivot	Splunk_SA_CIM	nobody	Global
>	Network Traffic	data model	Clone Pivot	Splunk_SA_CIM	nobody	Global
>	Performance	data model	Clone Pivot	Splunk_SA_CIM	nobody	Global
>	Splunk Audit Logs	data model	Clone Pivot	Splunk_SA_CIM	nobody	Global

Generated for Anh Hoang (hoang_anhkim@bah.com) (C) Splunk Inc, not for distribution

Using the CIM Add-on (cont.)

2. Create event types & tags

- Identify the CIM datasets relevant to your events
- Observe which tags are required for that dataset or any parent datasets
- Apply those tags to your events using event types

Add new
Event types > Add new

Destination App: class_Fund2

Name *: bcg_online_sales

Search string *: index=web sourcetype=access_combined action="*"

Tag(s): web
Enter a comma-separated list of tags.

Color: blue

Priority: High

Save As Event Type

Name: bcg_online_sales

Tags: web

Color: blue

Priority: High

Determines which style wins, when an event has more than one event type.

Cancel Save

Using the CIM Add-on (cont.)

3. Create field aliases

- Determine whether any existing fields in your data have different names than the names expected by the data models
- Define field aliases to capture the field with a different name in your original data and map it to the field name that the CIM expects

Add new
Fields > Field aliases > Add new

Destination app	class_Fund2	Field name in CIM object		
Name *	access_combined_aliases			
Apply to	sourcetype	named *	access_combined	
Field aliases	clientip	=	src	Delete
	host	=	dest	Delete
	useragent	=	http_user_agent	Delete
	+ Add another field			
	Field name in your data			
			Cancel	Save

Using the CIM Add-on (cont.)

4. Add missing fields

- Create field extractions
- Write lookups to add fields and normalize field values

5. Validate against data model

- Use the datamodel command
- Use Pivot in Splunk Web

Note

For more information, see the *Common Information Model Add-on Manual*:
<http://docs.splunk.com/Documentation/CIM/latest/User/Overview>

Add new
Lookups > Automatic lookups > Add new

Destination app: class_Fund2
Name*: Action
Lookup table*: cim_access_action_lookup
Apply to: sourcetype
Lookup input fields: LI_ACTION = action_taken
Lookup output fields: action_primary = action

Extract Fields
Select sample → Select method → Select fields → Save
Cancel Save

1	2
1 LI_ACTION1	action_primary
2	1 Quarantine
3	2 Rename
4	3 Delete
5	4 Leave alone
6	5 Clean
7	6 Clean or delete macros

Generated for Anh Hoang (hoang_anhkim@bah.com) (C) Splunk Inc, not for distribution

from and datamodel Commands

- Both can be used to retrieve data from a specified data model dataset
- Both are generating commands—must be the first command in the pipeline

from

[Learn More ↗](#)

Retrieves data from a named dataset, saved search, report, or lookup file. Must be the first command in a search.

Example:

```
I from datamodel:"internal_server.splunkdaccess"
```

datamodel

[Learn More ↗](#)

Allows users to examine data models and search data model datasets.

Example:

```
I datamodel
```

from Command

- Retrieves data from a data model

```
| from datamodel:Web
```

- Can also retrieve data from other named datasets (saved searches, reports, or lookup files)

```
| from savedsearch:mysecurityquery
```

INTERESTING FIELDS

a action 5
bytes 100+
a dest 3
a http_content_type 1
a http_method 1
a http_referrer 1
a http_user_agent 26
http_user_agent_length 19
is_not_Proxy 1
is_Proxy 1
a src 100+
status 9
a uri_path 13
a uri_query 100+
a url 1
url_length 1
a user 1
a vendor_product 1

datamodel Command

- Retrieves data from a data model or displays its structure
- Without search command, returns a description of all or a specified data model and its objects

```
| datamodel Web Web
```

A B C

- A Command
- B Data model name
- C Data model dataset name

```
i Time Event
> { [-]
  autoextractSearch: | search (index=* OR index=_*) ((`cim_Web_indexes`)
tag=web)
  calculations: [ [+]
  ]
  comment: { [+]
  }
  constraints: [ [+]
  ]
  displayName: Web
  fields: [ [+]
  ]
  indexScopeWarning: true
  lineage: Web
  objectAccelerationSearch: | search (index=* OR index=_*) ((`cim_Web_indexes`)
tag=web) | eval nodename = "Web" | eval action=if(isnull(action) OR
```

datamodel Command (cont.)

With search command, retrieves data from data model—like from, but with dataset name prepended to field names in data

```
| datamodel Web Web search
```

A B C D

- A Command
- B Data model name
- C Data model dataset name
- D Command

Important

The object name and search keyword aren't valid unless preceded by the data model name. The command search cannot be substituted with a search string or name. When using the datamodel command, the data model name and dataset name are case-sensitive.

INTERESTING FIELDS

a Web.action 7
Web.bytes 100+
a Web.dest 10
a Web.http_content_type 1
a Web.http_method 2
a Web.http_referrer 1
a Web.http_user_agent 1
Web.http_user_agent_length 1
Web.is_not_Proxy 1
Web.is_Proxy 1
a Web.src 10
Web.status 10
a Web.uri_path 11
a Web.uri_query 100+
a Web.url 10
Web.url_length 7
a Web.user 9
a Web.vendor_product 2

Additional CIM Resources

- Understand and use the CIM Add-on

<http://docs.splunk.com/Documentation/Splunk/latest/Knowledge/UnderstandandusetheCommonInformationModel>

- Overview of the Splunk CIM

<http://docs.splunk.com/Documentation/CIM/latest/User/Overview>

- Use the CIM to normalize data at search time

<http://docs.splunk.com/Documentation/CIM/latest/User/UsetheCIMtonormalizedataatsearchtime>

Module 13 Lab Exercise

Time: 20 minutes

Tasks:

- Normalize your data from the `access_combined` sourcetype to the CIM Web data model
 - Create the appropriate tags, field aliases and event types
- Validate your data against the Web data model

What's in the Appendices?

- Appendix A: Splunk Premium Solutions and Apps
 - An overview of ITSI, ES, UBA, and the most popular Splunk apps
- Appendix B: Creating New Choropleth Maps
 - Details on creating your own choropleth maps
- Appendix C: Buttercup Games
 - An overview of the data in the lab environment

What's Next?

- Splunk Certification program
 - https://www.splunk.com/en_us/training/faq-training.html
 - Splunk Core Certified User
 - Splunk Core Certified Power User
 - Splunk Enterprise Certified Admin
 - Splunk Enterprise Certified Architect
 - Splunk Certified Developer
- Program information
 - <https://www.splunk.com/pdfs/training/Splunk-Certification-Handbook-v.8.31.2018.pdf>
- Exam registration
 - <https://www.splunk.com/pdfs/training/Exam-Registration-Tutorial.pdf>
- If you have questions, send an email to: certification@splunk.com

Note



You must complete the Splunk User certification before you can take the exam to achieve Splunk Power User certification.

What's Next?: *Splunk Fundamentals* 3

- Use a wide variety of comparison, conversion, mathematical, and statistical functions for the eval command
- Create advanced lookups and alerts
- Advanced field extraction using regex
- Work with self-describing file types such as JSON and XML
- Create more complex and nested search macros
- Accelerate reports and data models, and explore the tsidx files created
- And much more

Splunk Fundamentals 3 Preview: replace Function – Example

Scenario ?

Show sales information for the 3 best-selling products of the last 24 hours. Mask the middle octets of the customer IP addresses.

```
index=web sourcetype=access_combined  
| chart sum(price) as totalSales over clientip by product_name  
| limit=3 useother=f  
| eval clientip =  
| replace(clientip, "(\d+\.)\d+\.\d+(\.\d+)", "\1xxx.xxx\2")  
| fillnull
```

clientip	Dream Crusher	Manganiello Bros.	Orvil the Wolverine
87.xxx.xxx.51	39.99	559.86	79.98
202.xxx.xxx.117	239.94	79.98	159.96
173.xxx.xxx.226	119.97	39.99	279.93
188.xxx.xxx.166	39.99	159.96	239.94

Note

The `limit` argument of the `chart` command limits the number of products displayed to the top `x` values (in this case, the top 3).



splunk® > .conf19

.conf19

October 21-24, 2019

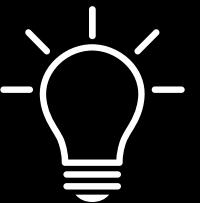
Splunk University

October 19-21, 2019

Las Vegas, NV

The Venetian Sands Expo

4 Days of Innovation



350 Education Sessions



20 Hours of Networking



"Hands down the most beneficial and attendee focused conference I have attended!"

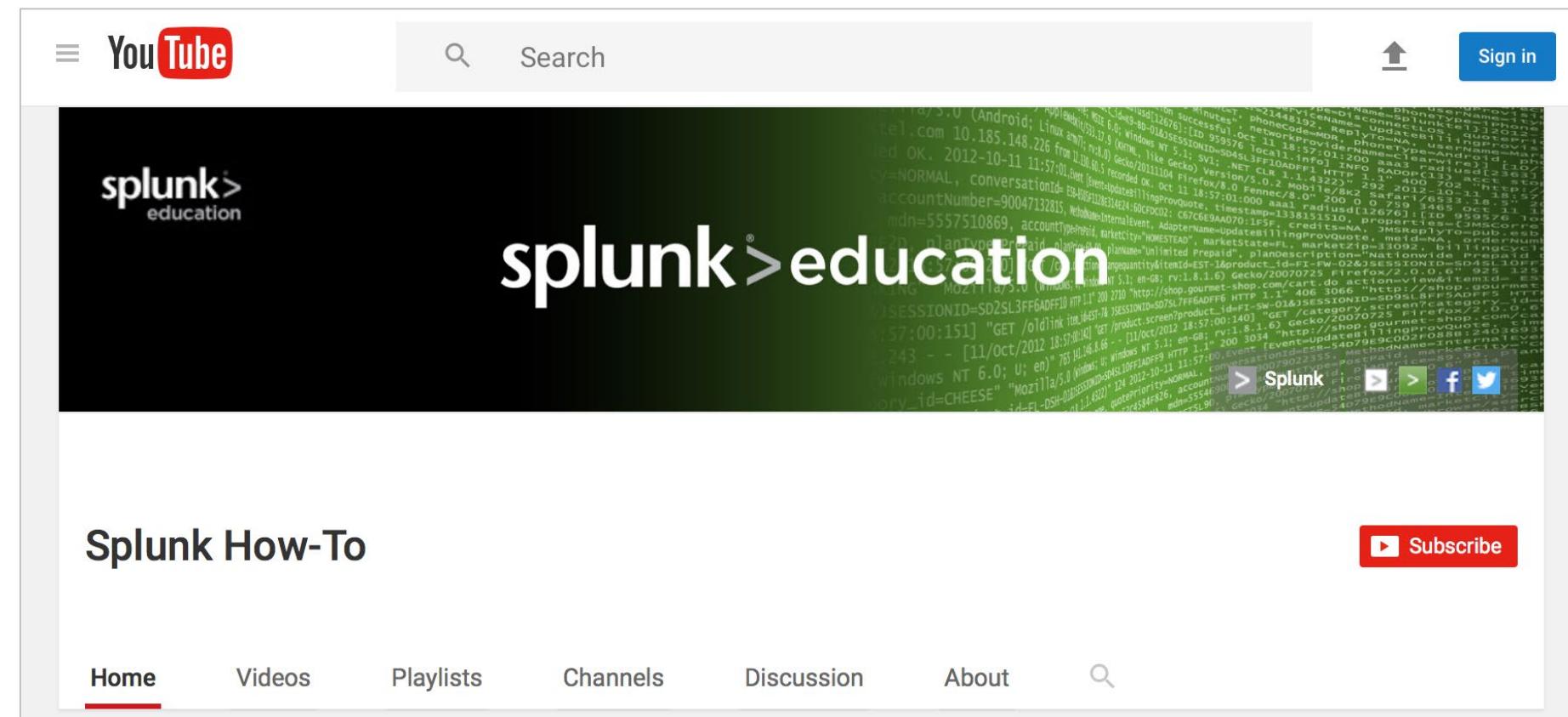
– Michael Mills, Senior Consultant, Booz Allen Hamilton

Generated for Anh Hoang (hoang_anhkim@bah.com) (C) Splunk Inc, not for distribution

sign up for notifications @ conf.splunk.com

YouTube: The Splunk How-To Channel

- In addition to our roster of training courses, check out the Splunk Education How-To channel: <http://www.youtube.com/c/SplunkHowTo>
- This site provides useful, short videos on a variety of Splunk topics



Generated for Anh Hoang (hoang_anhkim@bah.com) (C) Splunk Inc, not for distribution

Support Programs and Resources

- Global Support: dedicated resource for critical issues to manage your account, 24 x 7 x 365
 - Phone: (855) SPLUNK-S or (855) 775-8657
 - Web: http://www.splunk.com/index.php/submit_issue
- Enterprise Support: customer support team available to manage your cases, 24 x 7 (depending on support contract)
- Splunk Docs (<http://docs.splunk.com>)
Constantly updated, so be sure to select the version of Splunk you're using

Splunk Community

- Splunk App Repository
splunkbase.splunk.com
- Splunk Answers
answers.splunk.com
- Splunk Blogs
splunk.com/blog
- Splunk User Groups
usergroups.splunk.com
- Splunk .conf
conf.splunk.com
- Splunk Community Portal
splunk.com/en_us/community.html
- Splunk Live
splunklive.splunk.com
- Splunk Docs on Twitter
twitter.com/splunkdocs
- Over 100 channels in splunk-usergroups on Slack
<http://splk.it/slack>

Appendix A: Splunk Premium Solutions and Apps

Generated for Anh Hoang (hoang_anhkim@bah.com) (C) Splunk Inc, not for distribution

Topics

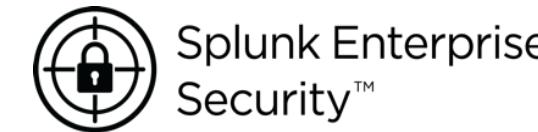
- Splunk IT Service Intelligence (ITSI)
- Splunk Enterprise Security (ES)
- Splunk User Behavior Analytics (UBA)
- Other Splunk Apps on Splunkbase

Splunk Premium Solutions and Apps

Splunk Premium Solutions



Splunk IT Service
Intelligence™



Splunk Enterprise
Security™



Splunk User Behavior
Analytics™

Rich Ecosystem of Apps & Add-Ons



amazon
web services



splunk®>enterprise

splunk®>cloud™

splunk®> Platform for Operational Intelligence



Forwarders



Syslog/
TCP



Mobile



IoT
Devices



Network
Wire Data



Hadoop



Relational
Databases

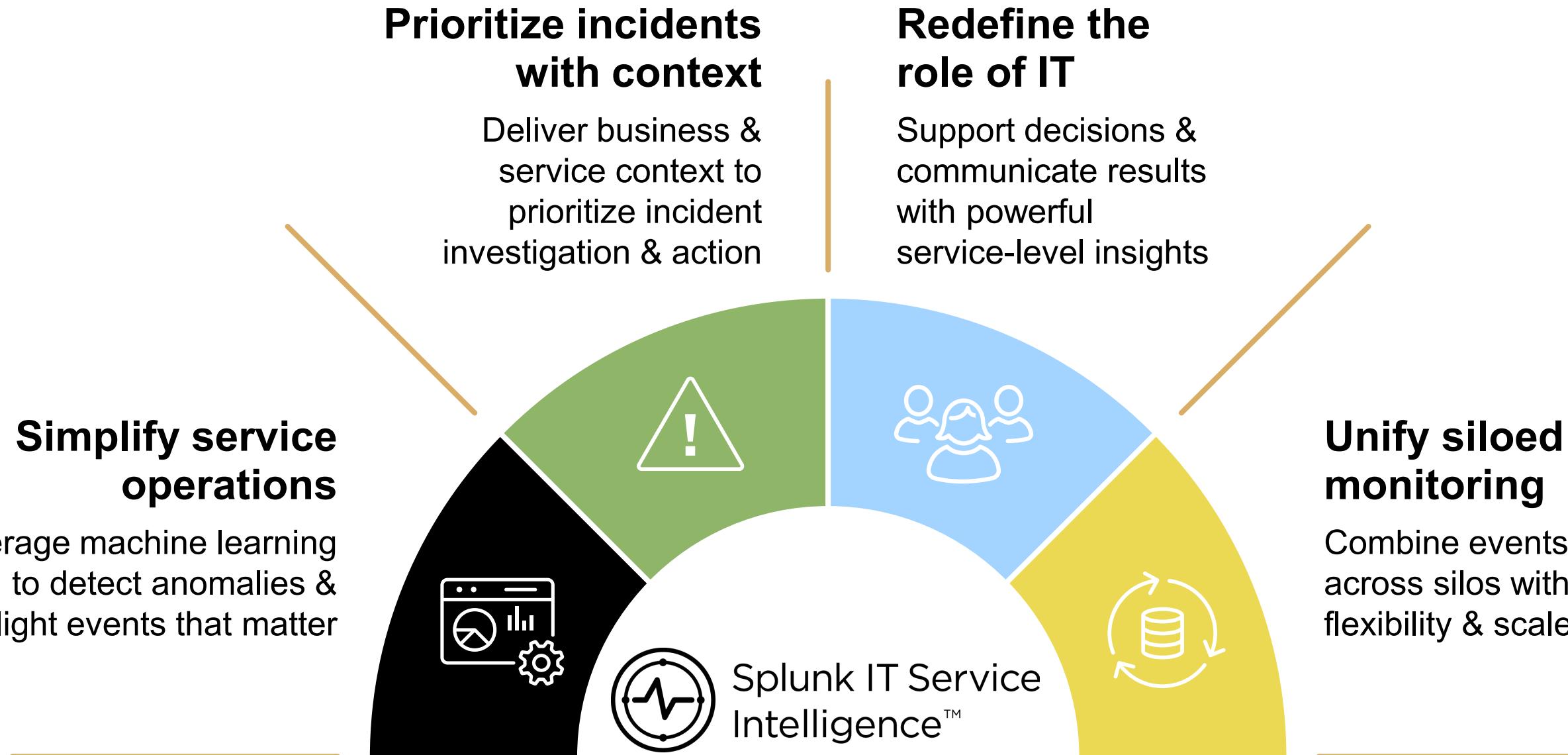


Mainframe
Data

Generated for Anh Hoang (hoang_anhkim@bah.com) (C) Splunk Inc, not for distribution

Splunk IT Service Intelligence (ITSI)

Artificial Intelligence Powering an Analytics-Driven IT



How IT Usually Operates: Stack Focus

- The way many in IT think of their world
- Each layer is a silo
- Dedicated experts with domain tools focus just on that layer's health
- Layer's health based on aggregated health of each component in layer
- If 2 out of 100 DBs are struggling, you're still having a good day!

Applications, business/mission services

Web Server (Apache, TomCat)

App Server (WebLogic, JBoss EAP, WebSphere)

Database (Oracle, SQL Server, MySQL)

Guest OS (Windows/Linux/*Nix)

Hypervisor (ESX, HyperV, Citrix)

Physical Server (Dell, HP, CISCO blades or servers)

SAN/NAS Storage (EMC, NetApp)

Network

What IT Really Needs: Service Focus

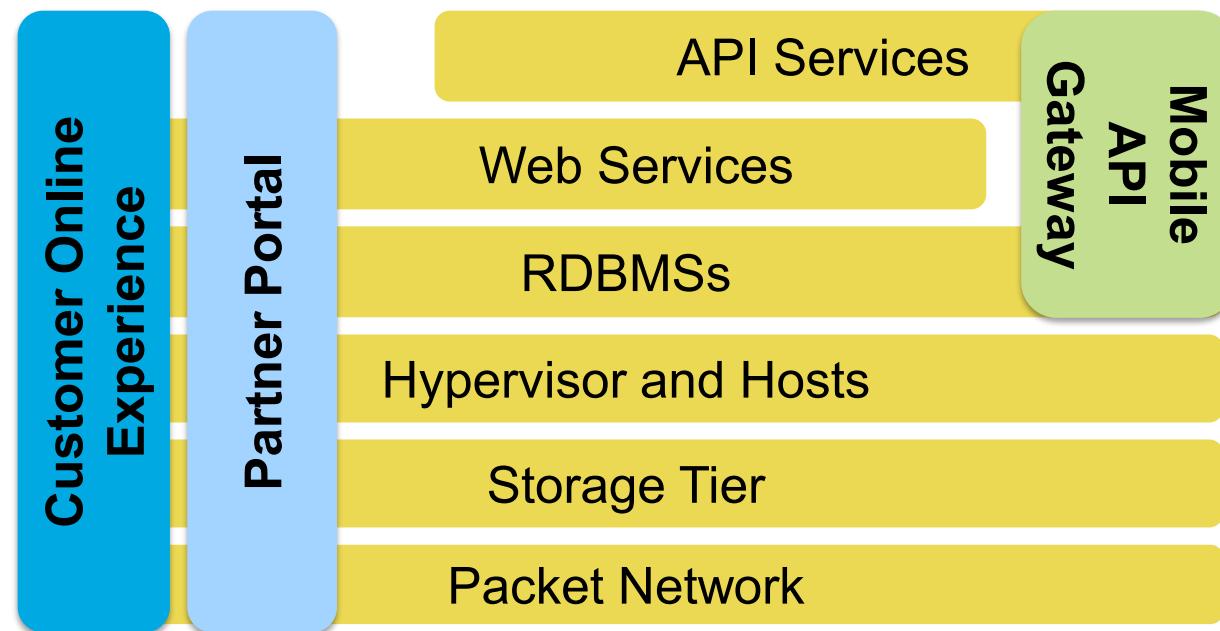
- The health of the layer isn't what's most relevant—**dependencies** matter
- The health of an app or service depends on the health of **each** component of **each** layer that app depends on
- If your app depends on 1 or more of those 2 struggling DB servers, you're having a **bad** day!

Service/App	Claims	Outage!	
		Status	
Web Server	(1,2,3,4,5,6,7,8,9,10...N)	100%	Green
App Server	(1,2,3,4,5,6,7,8,9,10...N)	100%	Green
Database	(1, 2,3,4 ,5,6,7,8,9,10...100)	98%	Yellow
Guest OS	(1,2,3,4,5,6,7,8,9,10...N)	100%	Green
VM/Hypervisor	(1,2,3,4,5, 6,7,8,9,10 ...N)	95%	Red
Physical Server	(1,2,3,4,5,6,7,8,9,10...N)	100%	Green
SAN/NAS Storage	(1,2,3,4,5,6,7,8,9,10...N)	100%	Green
Network		100%	Green

Generated for Anh Hoang (hoang_anhkim@bah.com) (C) Splunk Inc, not for distribution

What's a Service?

- A collection of IT objects that:
 - Relate to your business goals
 - Need to be monitored together

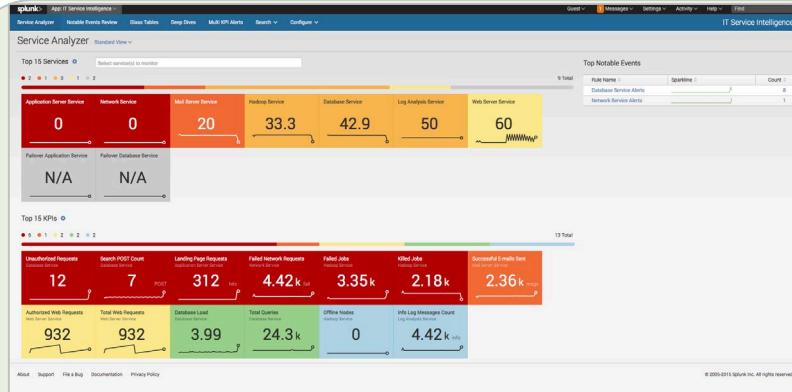


- Services can be
 - High-level, business-oriented
 - Low-level, technically-oriented
 - Tangible, like a storage tier
 - Abstract, multi-tier concepts, like a partner portal
 - Groups of people or objects
 - Dynamic or static
 - Wide or narrow in scope—strategic vs. tactical, global vs. local, corporate vs. team

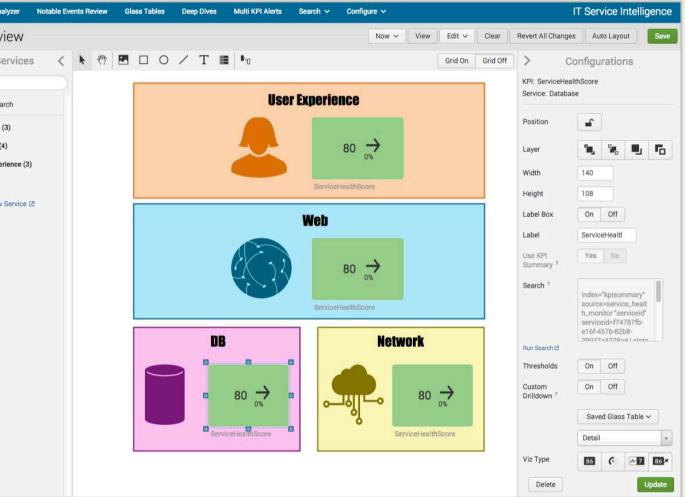
So What Does Service Intelligence Do?

- Enables a business-aware IT
 - Measures and reports on KPI's (Key Performance Indicators) such as Average Response Time or Errors Per Day—indicators that really matter
- Unlocks operational efficiencies
 - Collects information across all layers, removing “silos” to improve service operations
- Enables data-based decision making
 - Solves problems and anticipates pitfalls with sophisticated analytics and powerful insights

ITSI User Interface



Service Analyzer: quick view / filter for only the Services and KPIs you want

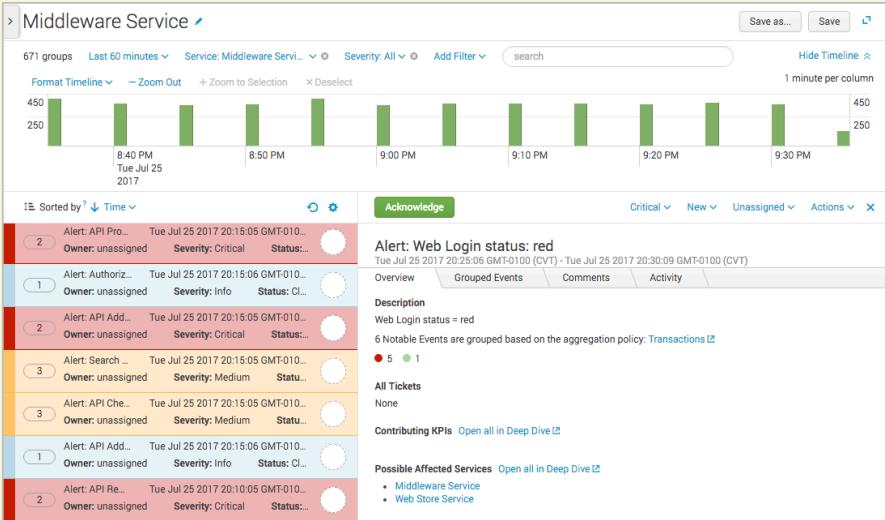


Glass table: Draw your own custom workflow and context with scores



Deep Dives: Investigate when things go wrong

Multi KPI Alerts: Alerts & Notable Events dashboard to review incidents



Generated for Anh Hoang (hoang_anhkim@bah.com) (C) Splunk Inc, not for distribution

ITSI Use Cases

- Transaction troubleshooting: alert and troubleshoot, prioritize based on KPIs important to you
- IT infrastructure: monitor entire system, reduce connectivity issues and systems downtime
- Policy service: detect changes in customer behavior before sales are significantly impacted
- Call Center monitoring: improve customer experience, reduce wait time and dropped calls

Note 

To learn more, register for one of the ITSI classes: *Using Splunk IT Service Intelligence* or *Administering Splunk IT Service Intelligence*.

Splunk Enterprise Security (ES)

Complement, replace, and go beyond traditional Security Information and Event Management (SIEM)



Incident
Investigations
& Forensics



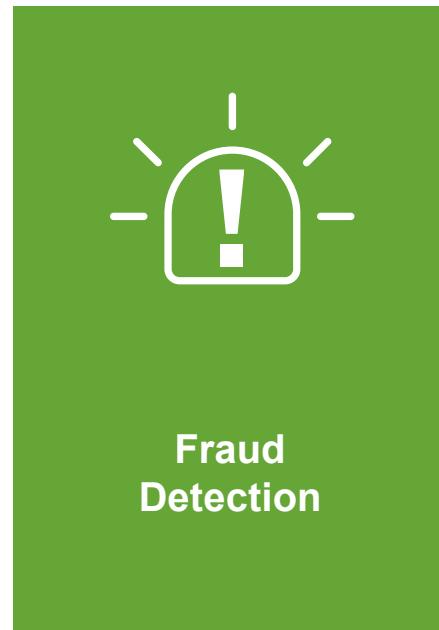
Security &
Compliance
Reporting



Real-time
Monitoring of
Known Threats



Detecting
Unknown
Threats



Fraud
Detection



Insider
Threat

splunk®>

Generated for Anh Hoang (hoang_anhkim@bah.com) (C) Splunk Inc, not for distribution

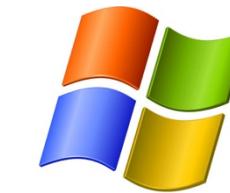
What is Splunk ES?

- Makes **all** your data—not just your “security data”—relevant to your security effort
- Detect, prevent, and respond to security threats and incidents
- Efficiently manage, analyze, and mitigate security breaches
- Highly customizable for your specific enterprise requirements
- Real-time, scalable, context-aware, focused on content



How Does ES Work?

1. Data acquired by add-ons in your enterprise from servers, routers, etc.
2. Data forwarded to Splunk indexers and stored as events
3. ES runs real-time searches, looking for indicators of threats, vulnerabilities, or attacks
4. If a search discovers something that needs attention, ES displays it on one or more of its dashboards
5. You can then investigate the issue, track it, analyze it, and take the appropriate action



Analytics Driven SIEM

- Risk-Based Analytics align security operations with business
 - Risk scoring framework enhances decision making by applying risk scores to any data
 - Quickly and easily assign any KSI or KPI to any event to align with your current priorities
 - Expose the contributing factors of a risk score for deeper insights
- Visualize and discover relationships for faster detection and investigation
 - Visually fuse data, context, and threat intel across the stack and time to discern relationships

Analytics Driven SIEM (cont.)

- Pre-built correlations, alerts, and dashboards for detection, investigation, and compliance
 - Workflow actions and automated lookups enhance context building
-
- Enrich security analysis with threat intelligence
 - Automatically apply threat intelligence from any number of providers
 - Apply threat intelligence to event data as well as wire data
 - Conduct historical analysis using new threat intelligence across all data

ES Functional Areas

Perimeter Defense

- Known threats
- Vulnerability alerts
- Unexpected/prohibited processes or traffic
- Threat activity
- Risk framework

Preventative Analysis

- Anomaly detection
- Pattern matching
- Traffic analysis
- Statistical analysis

Breach Response

- Investigation journaling
- Incident tracking
- Forensics tools
- Asset and identity management
- Audit

- Security posture

- Correlation searches

- Threat intelligence

- Network capture

- Predictive analytics

- Insider threat analysis

- Advanced threat analysis

- Risk analysis

- Protocol intelligence

- Adaptive response

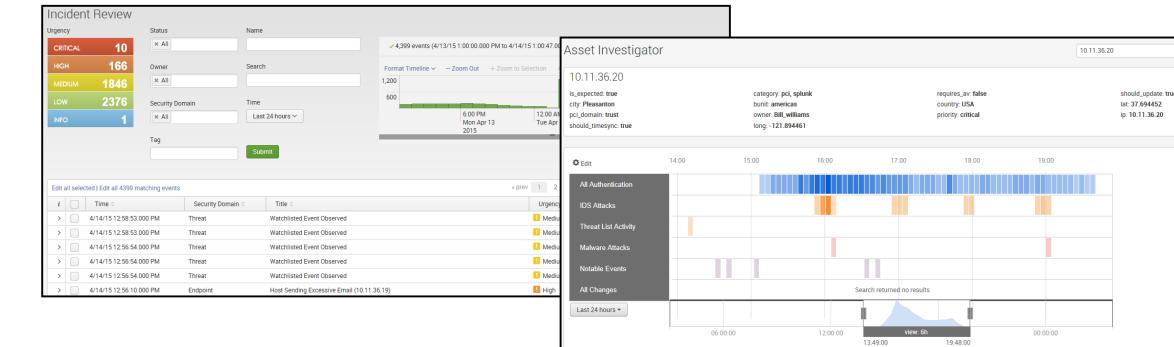
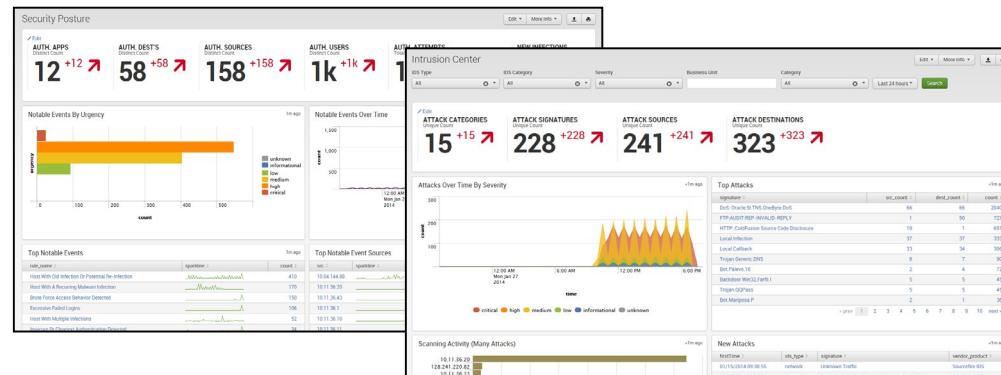
- Investigation timeline

- Asset and identity investigators

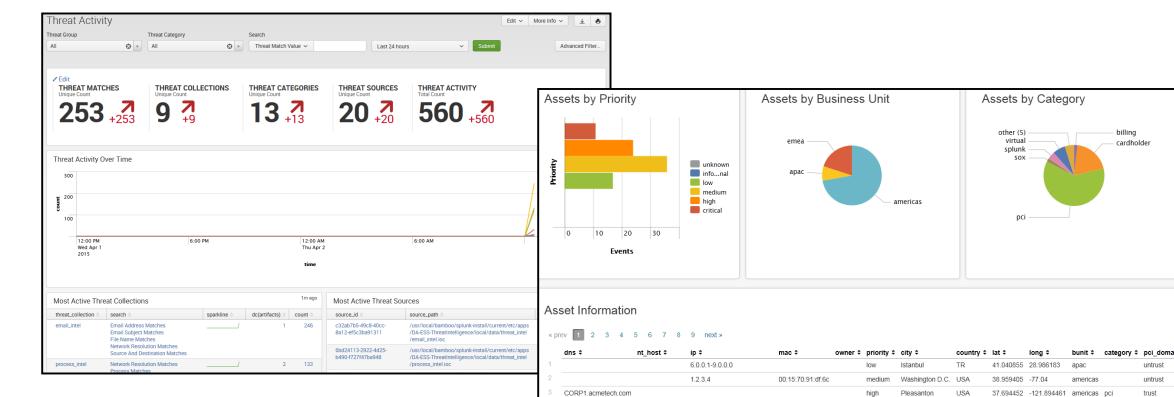
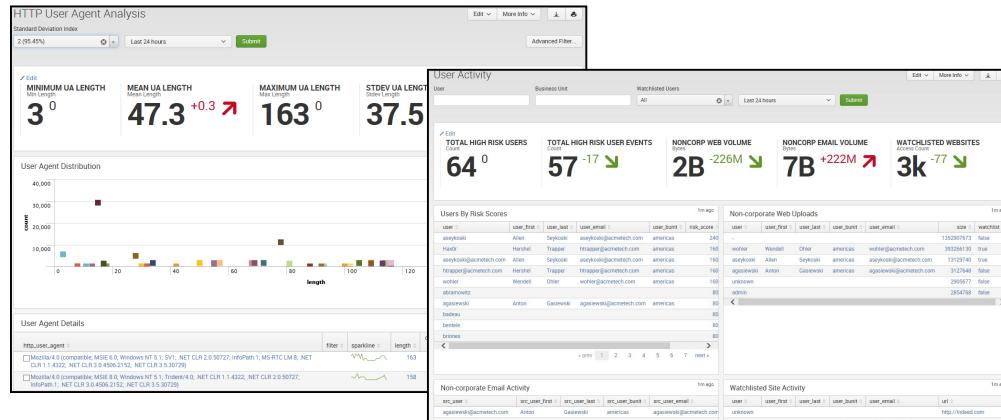
- Dashboards, searches, and reports for forensics and auditing

ES User Interface

Pre-built searches, alerts, reports, dashboards, incident workflow, and threat intelligence feeds



Alerts, dashboards, and reports



Statistical outliers, risk scoring, and user activity
Generated for Anh Hoang (hoang_anhkim@bah.com) (C) Splunk Inc, not for distribution

Threat intel and asset & identity integration

ES Use Cases

- Malware protection
 - Detection
 - Patient zero identification
 - Zero-day investigations
- Insider threat
 - Data exfiltration
 - Suspicious privileged account activity
 - Monitor threat activity with a glass table
- Detailed use case descriptions at:

docs.splunk.com/Documentation/ES/latest/Usecases

Note 

To learn more, register for one of the ES classes: *Using Splunk Enterprise Security* or *Administering Splunk Enterprise Security*.

Splunk User Behavior Analytics (UBA)

Out-of-the-box solution that helps organizations find unknown threats and anomalous behavior via **machine learning**



Real-Time & Big
Data Architecture



Behavior Baseline
& Modelling



Unsupervised
Machine Learning



Anomaly
Detection



Threat
Detection

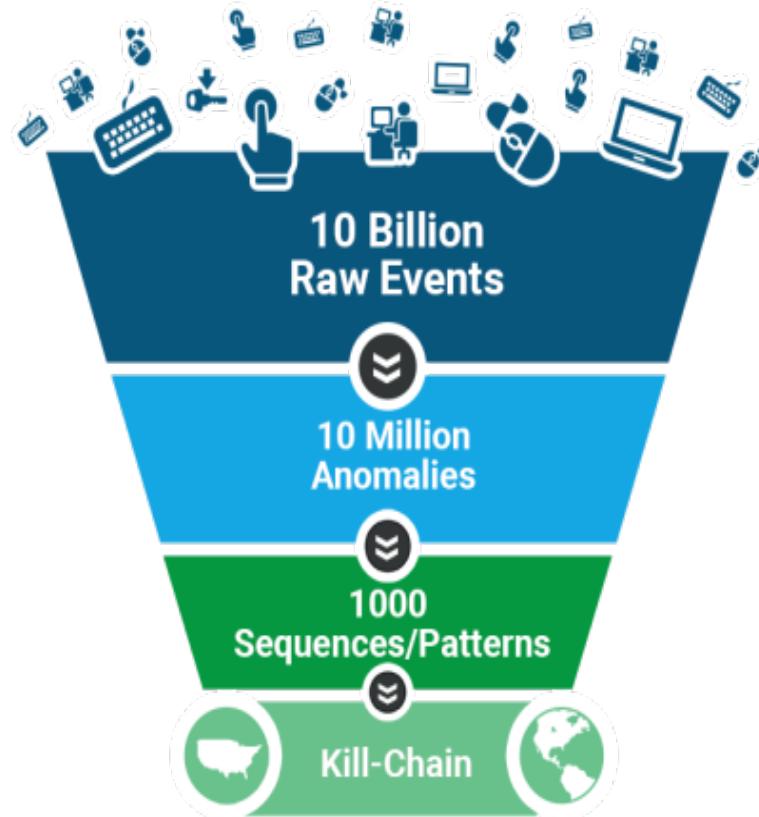


Splunk User Behavior
Analytics™

splunk> Platform for Machine Data

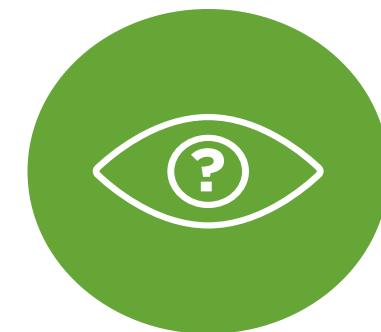
Generated for Anh Hoang (hoang_anhkim@bah.com) (C) Splunk Inc, not for distribution

What Does UBA Do?



⚠️ Top-10
critical and actionable
unknown threats

- Finds known, unknown, and hidden threats by analyzing user behavior and flagging unusual activity
- The more you use it, the “smarter” it becomes about what is and isn’t suspicious behavior in your environment
- Integrates with ES



Anomalous
Behavior



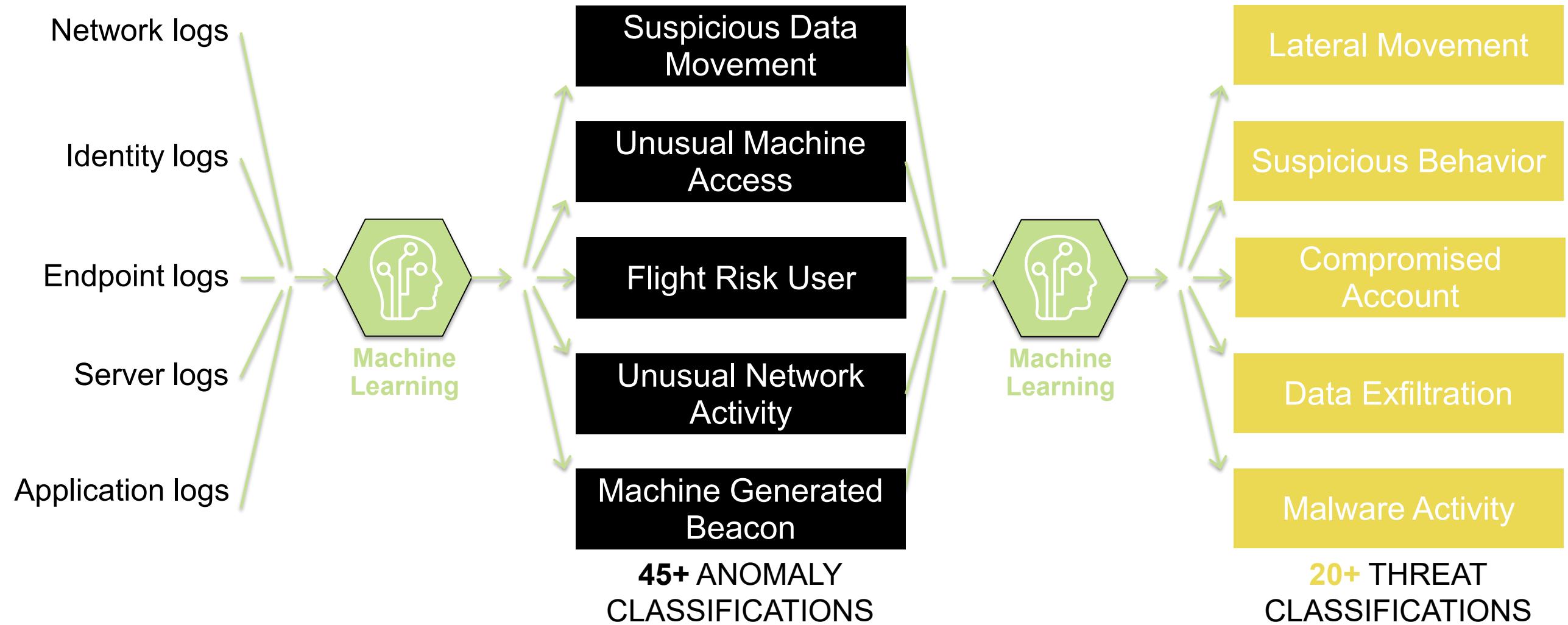
Risky User
Monitoring



Unknown
Threats

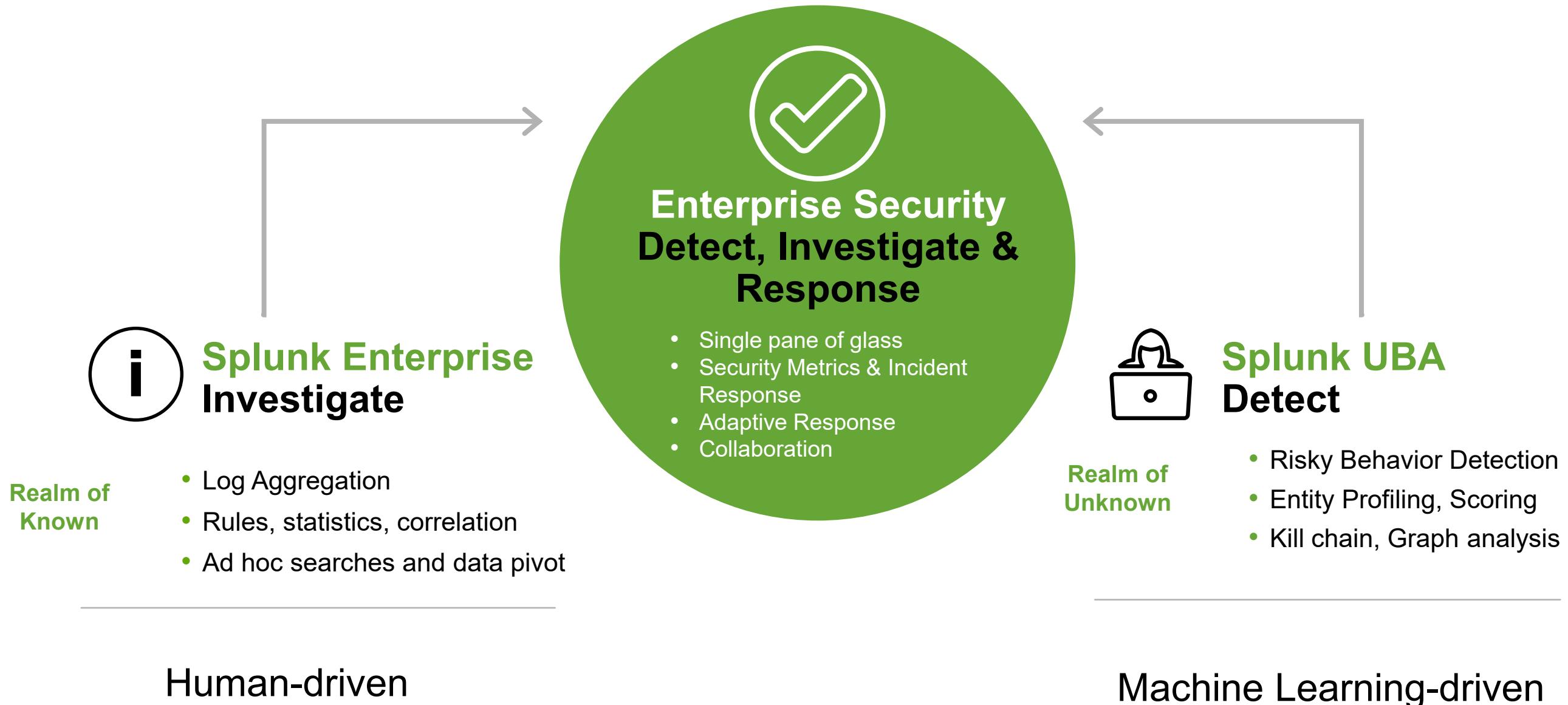
Generated for Anh Hoang (hoang_anhkim@bah.com) (C) Splunk Inc, not for distribution

How Does UBA Work?



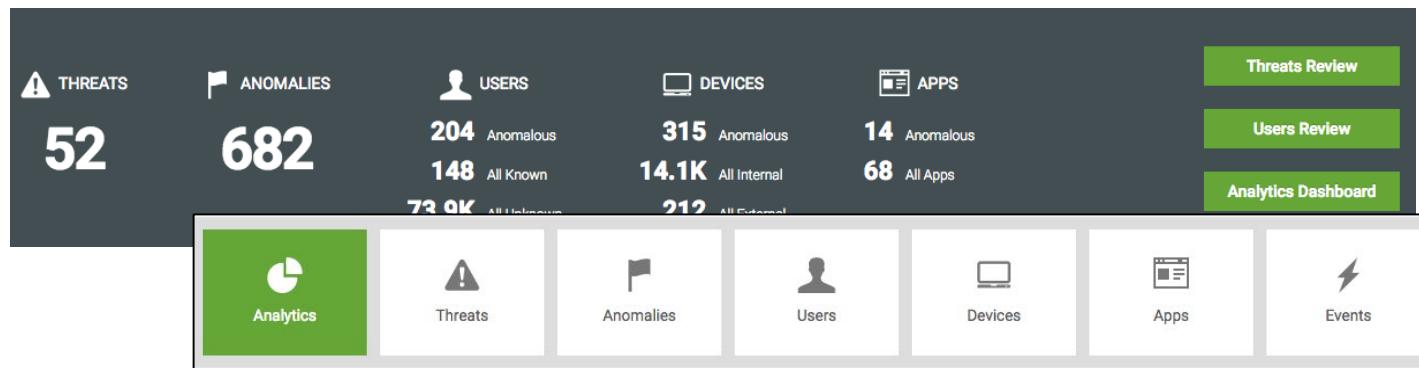
Generated for Anh Hoang (hoang_anhkim@bah.com) (C) Splunk Inc, not for distribution

How Does UBA Complement ES?

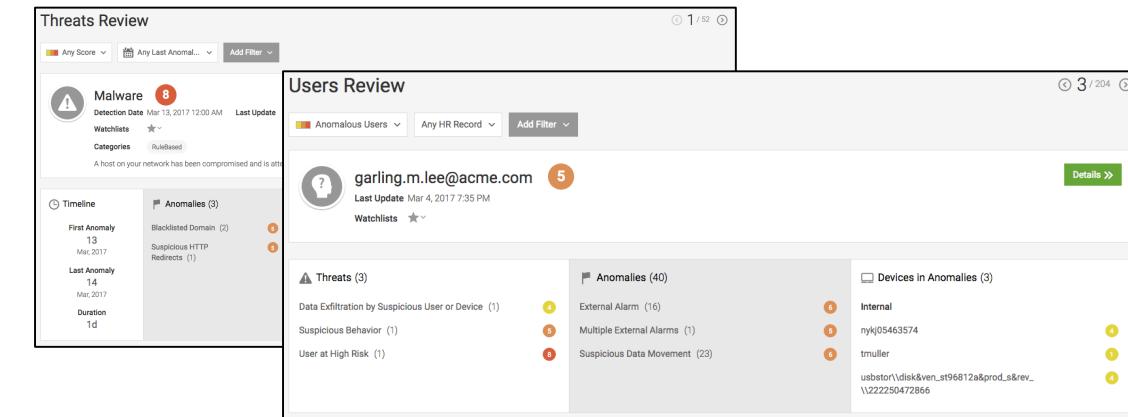


Generated for Anh Hoang (hoang_anhkim@bah.com) (C) Splunk Inc, not for distribution

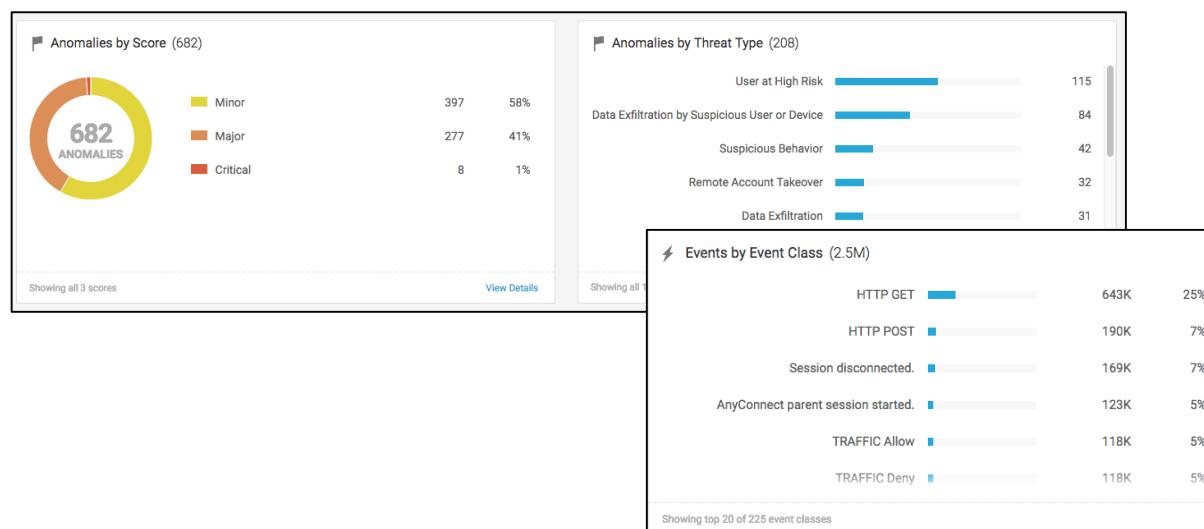
UBA User Interface



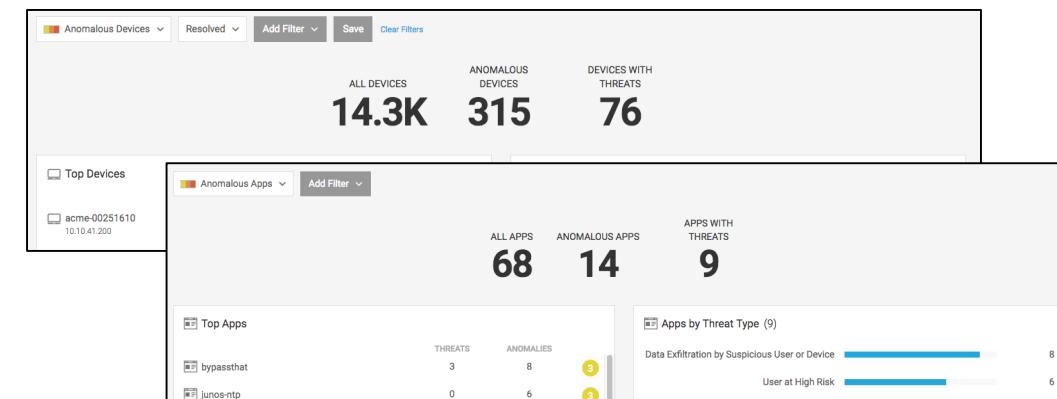
Summary and detailed dashboards



Review by specific threats and users



Zoom in on specific events and anomalies



Break down by specific devices and apps

Generated for Anh Hoang (hoang_anhkim@bah.com) (C) Splunk Inc, not for distribution

UBA Use Cases

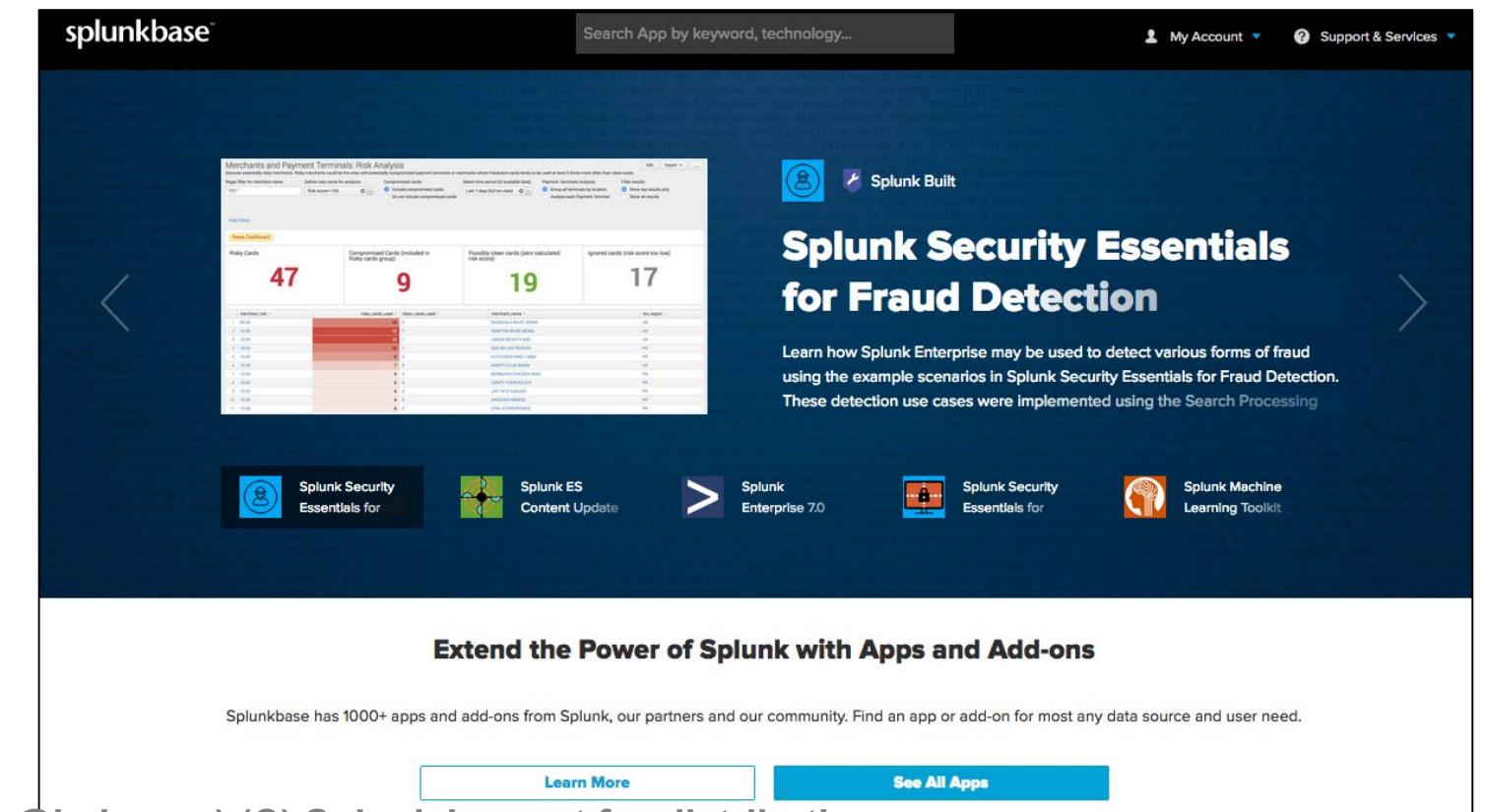
- Detect compromised user accounts
- Detect data exfiltration
- Detect compromised endpoints
- Detect insider access abuse, including privilege abuse
- Provide information for investigations

Note 

To learn more, register for the eLearning class *Splunk User Behavior Analytics*.

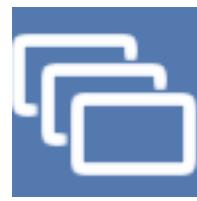
Splunkbase

- 1000+ other ready-made apps and add-ons addressing a wide variety of use cases available on Splunkbase (splunkbase.splunk.com)
- Some of the more popular apps include:
 - Dashboard Examples
 - Machine Learning Toolkit
 - Splunk App for Windows
 - Splunk App for Unix and Linux
 - Splunk App for AWS
 - Splunk Security Essentials



Generated for Anh Hoang (hoang_anhkim@bah.com) (C) Splunk Inc, not for distribution

Popular Splunk Apps



- Splunk Dashboard Examples
 - Teaches basic concepts and tools to rapidly create rich dashboards using XML
 - Includes extensions to XML for further customization of layout, interactivity, and visualizations
- Splunk Machine Learning Toolkit
 - Delivers new SPL commands, custom visualizations, assistants, and examples to explore a variety of ML concepts
 - Includes examples with datasets, plus the ability to apply the visualizations and SPL commands to your own data



Popular Splunk Apps (cont.)

INF

- Splunk App for Windows Infrastructure
 - Provides examples of pre-built data inputs, searches, reports, and dashboards for Windows server and desktop management
 - Monitor, manage, and troubleshoot Win OS's, including AD elements, all from one place

*nix

- Splunk App for Unix and Linux
 - Enables admins to quickly identify performance and capacity bottlenecks and outliers in large-scale *nix environments
 - Pre-packaged alerting capability, flexible service-based hosts grouping, and easy management of many data sources

Popular Splunk Apps (cont.)



- Splunk App for AWS
 - Provides critical insights into your Amazon Web Services account
 - Contains pre-built dashboards, reports, and alerts

- Splunk Security Essentials
 - Free app that detects insiders and advanced attackers in your environment
 - Showcases 55+ working examples of anomaly detection related to entity behavior analysis (UEBA) that can immediately be put to use in your environment
 - Can send results to either Splunk ES or Splunk UBA



Appendix B: Creating New Choropleth Maps

Generated for Anh Hoang (hoang_anhkim@bah.com) (C) Splunk Inc, not for distribution

Choropleth Maps

- Uses shading to show relative metrics for predefined geographic regions
- Splunk ships with two:
 - geo_us_states, United States
 - geo_countries, countries of the world
- You can import other choropleth maps or create your own



Generated for Anh Hoang (hoang_anhkim@bah.com) (C) Splunk Inc, not for distribution

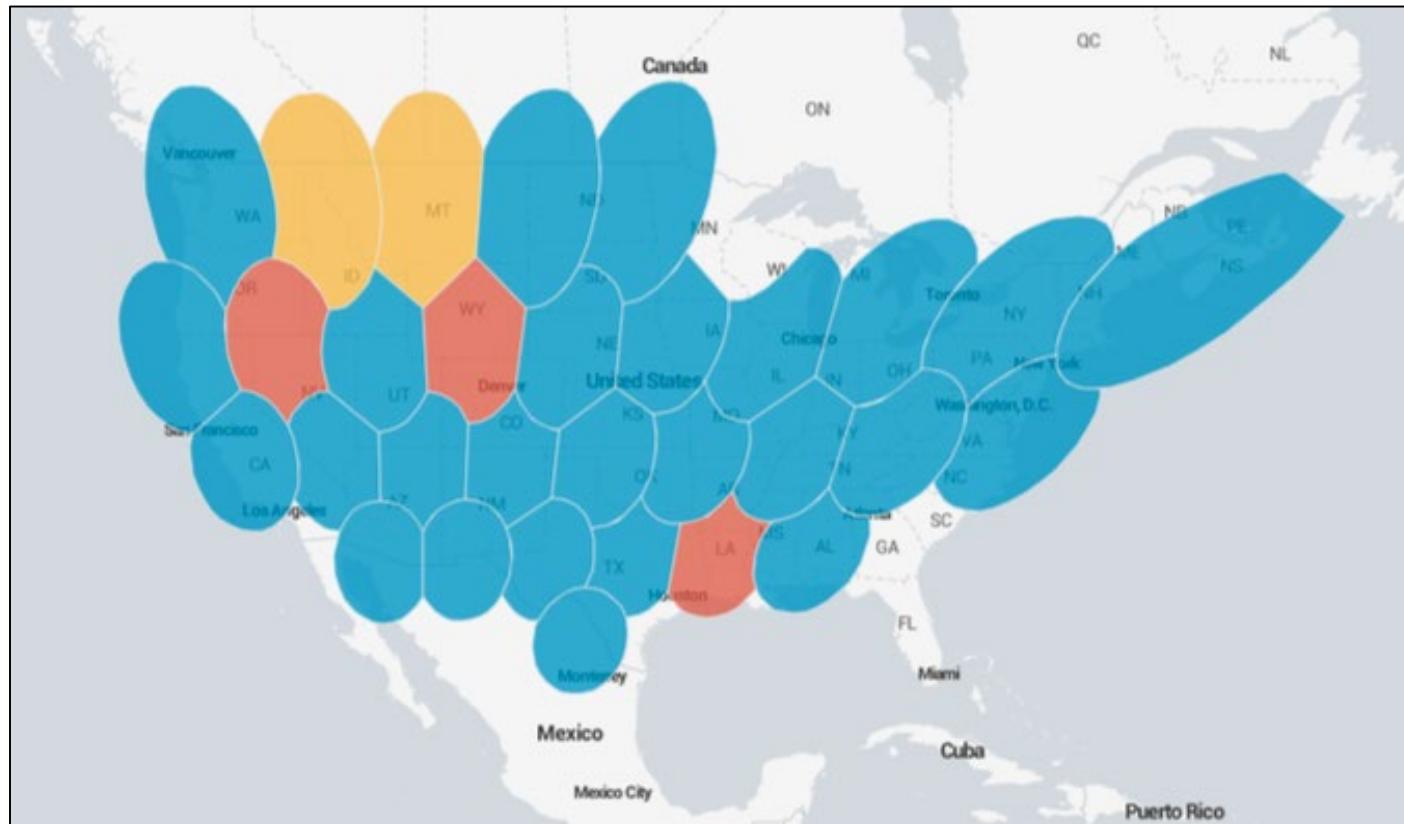
Choropleth Terminology

- KML (Keyhole Markup Language): type of XML developed by Google and others
- KMZ: a zipped KML file
- Polygon: the specific KML tag that Splunk uses to define its choropleth map data

```
<?xml version="1.0" encoding="UTF-8"?>
<kml xmlns="http://www.opengis.net/kml/2.2">
<Document><name>My document</name>
<description>Content</description>
<Style id="Lump">
<LineStyle><color>CD0000FF</color><width>2</width>
<PolyStyle><color>9AFF0000</color></PolyStyle>
</Style>
<Style id="Path">
<LineStyle><color>FF0000FF</color><width>3</width>
</Style>
<Style id="markerstyle">
<IconStyle><Icon><href>
http://maps.google.com/intl/en_us/mapfiles/ms/
</href></Icon></IconStyle>
</Style>
<Placemark><name>C</name>
<description></description>
<styleUrl>#Lump</styleUrl>
<Polygon>
<tessellate>1</tessellate>
<altitudeMode>clampToGround</altitudeMode>
<outerBoundaryIs><LinearRing><coordinates>
-86.264648,28.091366,0.0 -86.704102,28.188244,
-87.561035,27.722436,0.0 -87.604980,27.137368,
-87.209473,26.293415,0.0 -86.726074,26.194877,
-86.088867,26.431228,0.0 -86.022949,26.588527,
-86.418457,26.549223,0.0 -86.682129,26.470573,
-87.209473,26.725987,0.0 -87.253418,27.117813,
-87.011719,27.839076,0.0 -86.748047,27.858504,
-86.352539,27.761330,0.0 -86.220703,27.702984,
-86.264648,28.091366,0.0 </coordinates></LinearRing>
</Polygon>
```

Generated for Anh Hoang (hoang_anhkim@bah.com) (C) Splunk Inc, not for distribution

How Choropleth Maps Work

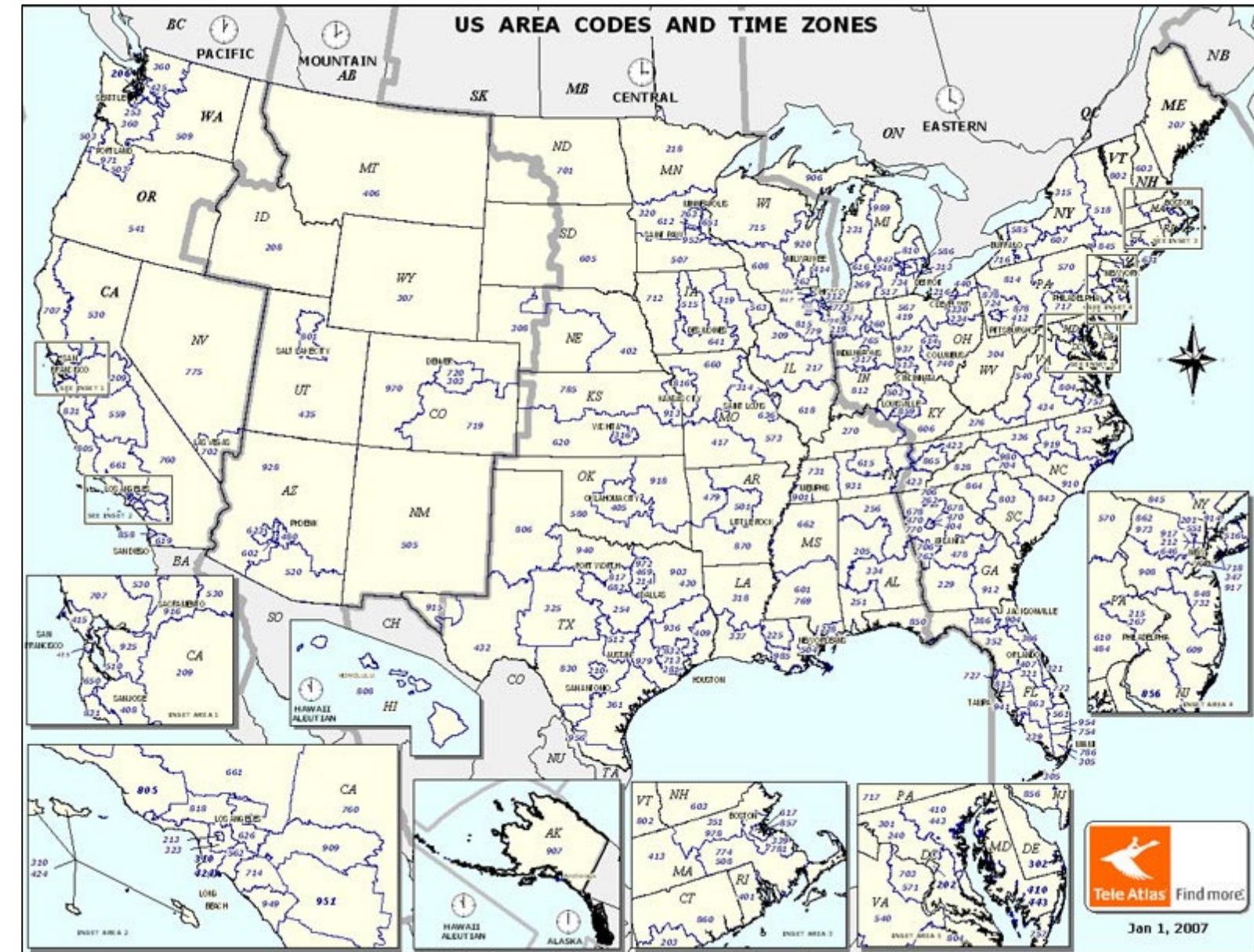


- Choropleth map data serves two purposes:
 1. Defines polygons to produce the colored map
 2. Provides method to determine within which polygon a given latitude/longitude is located
- Splunk can use a choropleth KML file as a lookup

Finding Other KML Choropleth Data Files

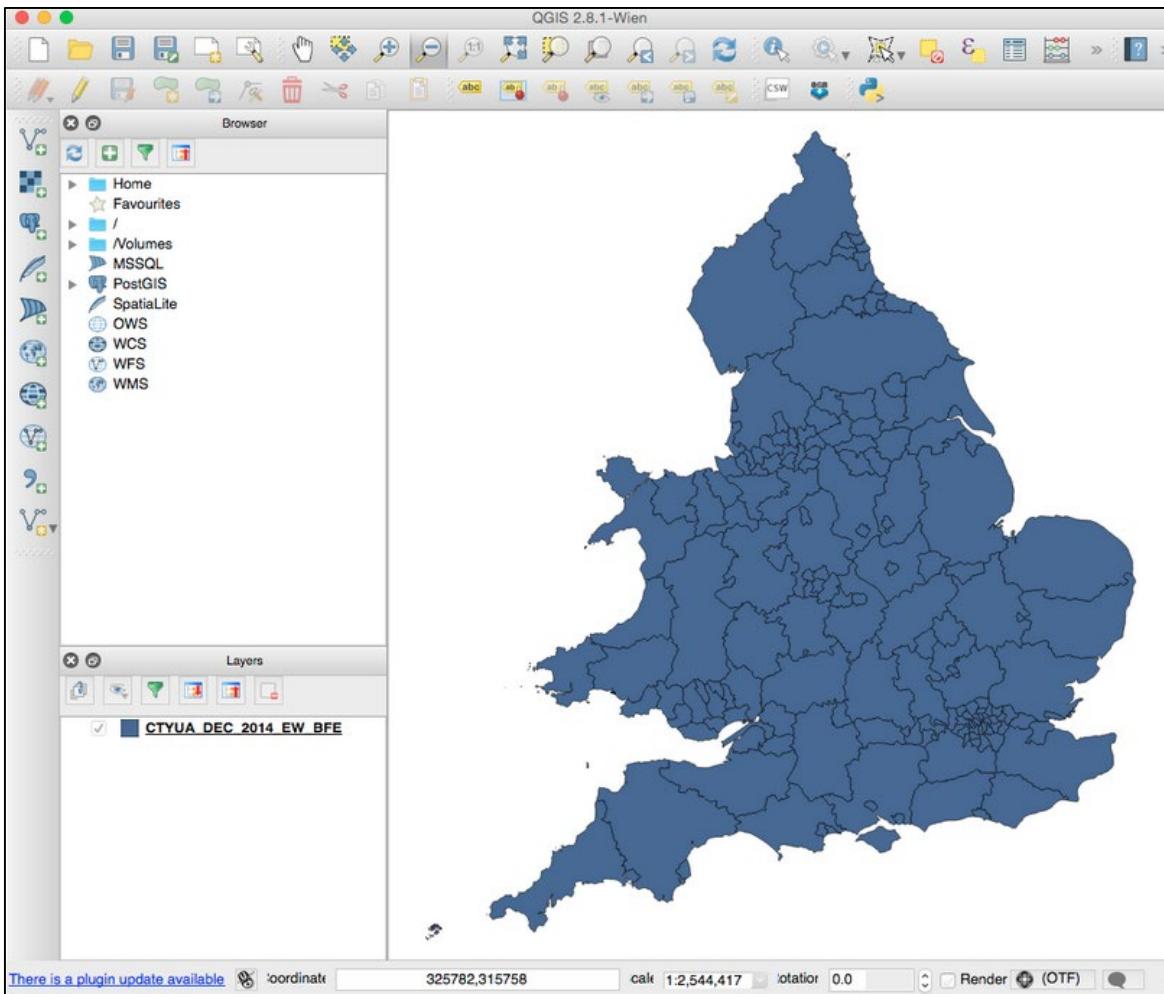
- Census bureau sites for US, UK, Australia
- Lots of other free KML/KMZ files available online
- For example, Google published a KMZ for phone area codes

<https://productforums.google.com/forum/#topic/gec-tools/6y6PrVFdPIY>



Generated for Anh Hoang (hoang_anhkim@bah.com) (C) Splunk Inc, not for distribution

Converting Other File Types to KML



- You can also convert choropleth files to KML from other formats, such as Shapefile
- Mapping systems have been around for over 20 years—some formats not so easy to work with

Note

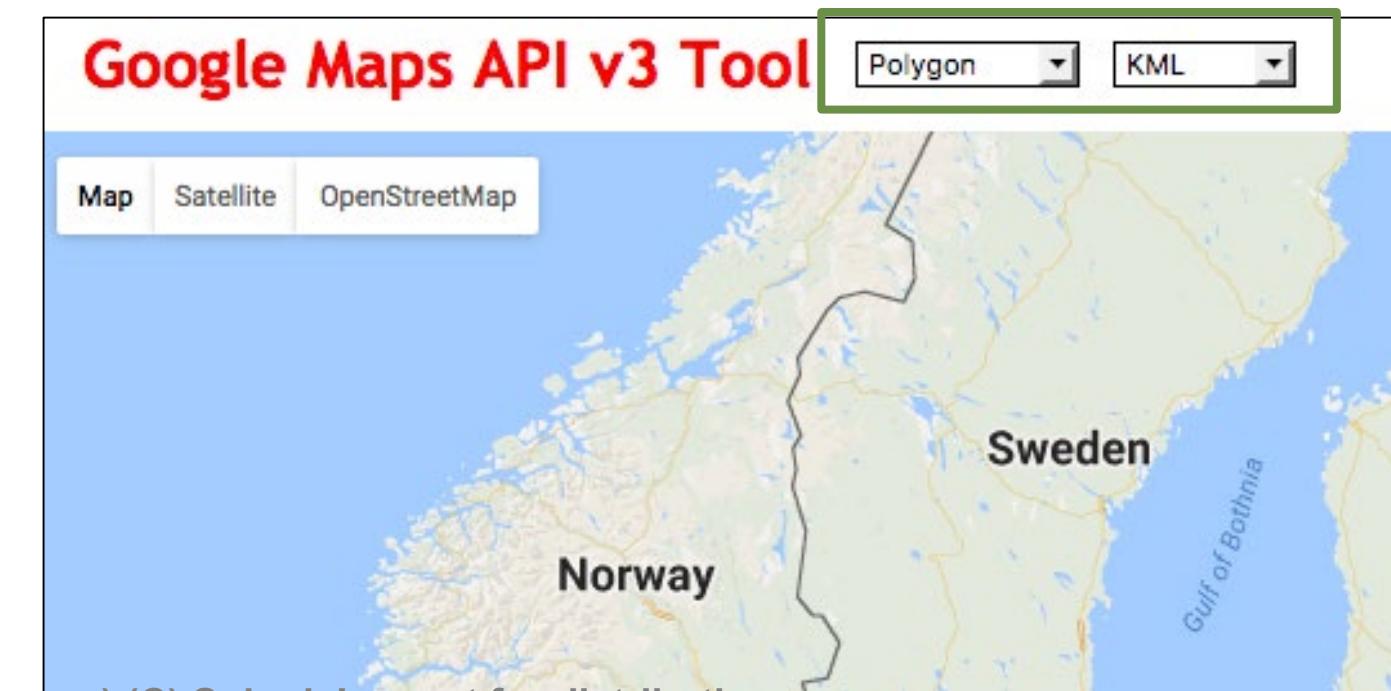


For complete details, see <http://blogs.splunk.com/2015/10/01/use-custom-polygons-in-your-choropleth-maps/>

Generated for Anh Hoang (hoang_anhkim@bah.com) (C) Splunk Inc, not for distribution

Build Your Own Choropleth Files using Online Tools

- Google Earth (<http://earth.google.com>)
- Sketchup (<http://www.sketchup.com>)
- Other online point-and-click tools (for example, <http://www.birdtheme.org/useful/v3tool.html>)
- Make sure:
 - Shapes being created are Polygon, not Polyline
 - Polygons are closed (start and end at the same coordinate)
 - No carriage returns in coordinates list (Splunk won't accept them)



Generated for Anh Hoang (hoang_anhkim@bah.com) (C) Splunk Inc, not for distribution

Using KML Files in Splunk

1. Upload the KML/KMZ file into Splunk as a lookup file
2. In the search, indicate an events data source that contains either featureID (location name) or latitude and longitude
If file contains only lat/long, you can use lookup to find location name (e.g.,
`|lookup my_geo_map latitude longitude`)
3. Use transforming command to aggregate data by location name
For example, `stats count by featureId |`
4. Optionally, select and configure a visualization
5. Create the choropleth map using the geom command

For example, `geom my_geo_map`

Note

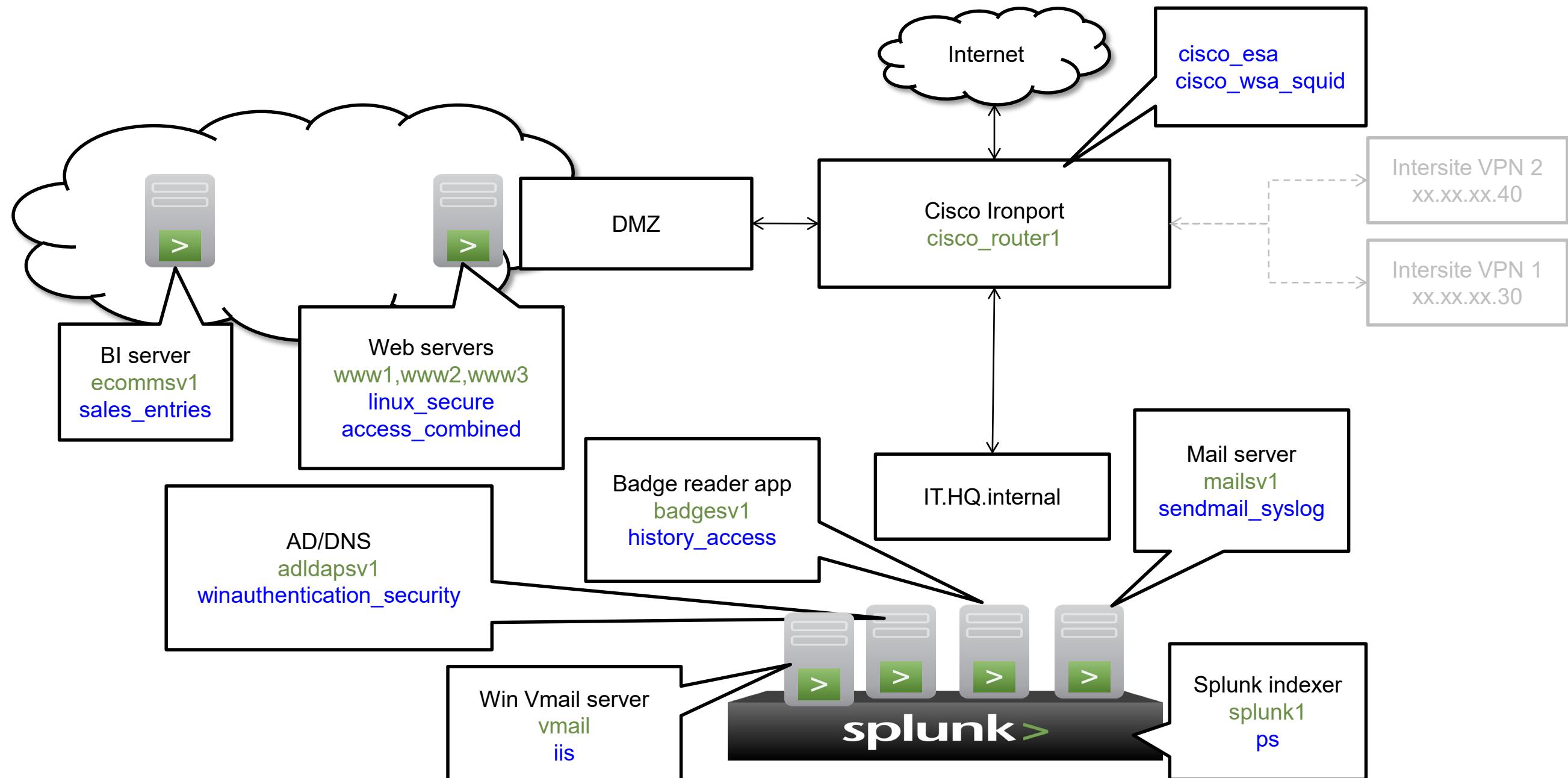
For complete details, see

<https://docs.splunk.com/Documentation/SplunkCloud/latest/Viz/ChoroplethGenerate>

Appendix C: Buttercup Games

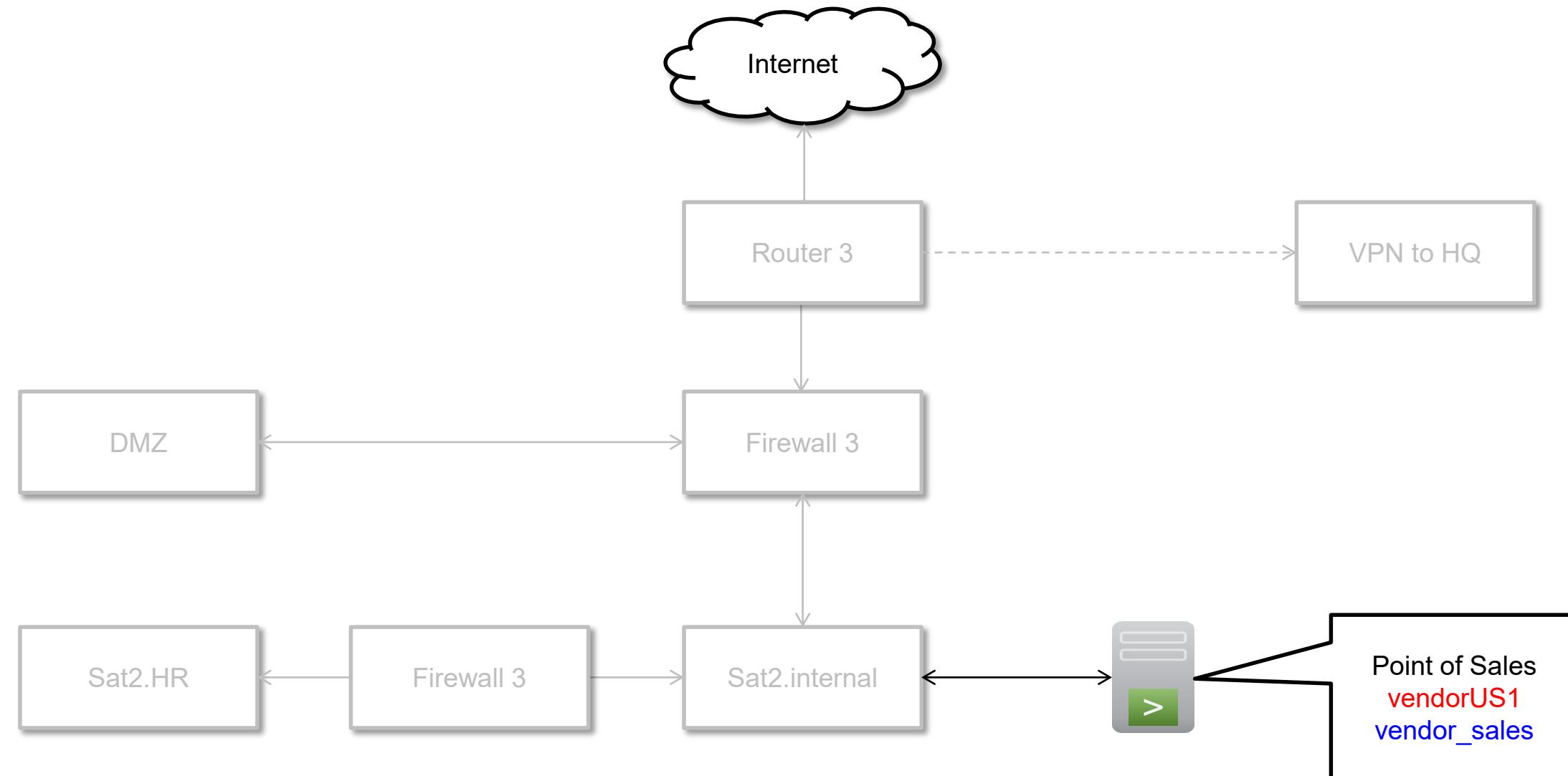
Generated for Anh Hoang (hoang_anhkim@bah.com) (C) Splunk Inc, not for distribution

San Francisco – Headquarters



Generated for Anh Hoang (hoang_anhkim@bah.com) (C) Splunk Inc, not for distribution

Boston – Satellite Office 2



Generated for Anh Hoang (hoang_anhkim@bah.com) (C) Splunk Inc, not for distribution

Buttercup Games Environment

Data	host	sourcetype
AD/DNS data	adldapsv1	WinEventLog:Security
Badge reader data	badgesv1	history_access
BI server data	ecommsv1	sales_entries
Email data	cisco_router1	cisco_esa
Online transactions & Web server	www1	access_combined
	www2	linux_secure
	www3	
Retail sales data	vendorUS1	vendor_sales
Splunk indexer data	splunk1	ps
Web appliance data	cisco_router1	cisco_wsa_squid

Generated for Anh Hoang (hoang_anhkim@bah.com) (C) Splunk Inc, not for distribution

Buttercup Games – HQ Employees

emp_no	birth_date	first_name	last_name	gender	hire_date	location
1	1963-12-25	Suzanne	Flaemmchen	F	2008-05-03	San Francisco
2	1960-04-20	Huang	Sham	M	2008-05-03	San Francisco
3	1950-06-09	Stefano	Pahkthecah	M	2008-05-03	San Francisco
4	1962-01-01	Shawn	Scallion	M	2008-05-03	San Francisco
5	1992-02-29	Shane	Youngin	M	2008-05-03	San Francisco
11	1969-08-19	Placido	Toscani	M	2009-06-09	San Francisco
12	1988-12-06	Meng	Yuan	F	2009-06-09	San Francisco
13	1963-09-29	Amanda	Curry	F	2009-06-09	San Francisco
14	1978-10-31	Bao	Lu	M	2009-06-09	San Francisco
			:			
			:			
68	1978-09-19	Pat	Leuchs	NR	2011-02-04	San Francisco
70	1964-05-19	Patricia	dAbbeville	F	2009-03-14	San Francisco
72	1978-07-10	Saran	Wrappe	F	2011-04-16	San Francisco
73	1988-12-01	Thomasina	Cugina	F	2012-05-19	San Francisco
75	1963-06-28	Frazer	Ullian	M	2013-12-13	San Francisco
76	1964-05-19	Mitsuko	Oh	F	2008-07-04	San Francisco
77	1962-04-01	Yurij	Schonegge	M	2010-01-11	San Francisco
81	1970-01-01	Buttercup	Pony	P	2008-05-03	San Francisco

Generated for Anh Hoang (hoang_anhkim@bah.com) (C) Splunk Inc, not for distribution

Buttercup Games – Satellite Employees

emp_no	birth_date	first_name	last_name	gender	hire_date	location
7	1963-06-07	Daniil	Piazza	M	2009-06-09	Boston
8	1961-05-02	Enrique	Dutra	M	2009-06-09	Boston
9	1974-06-19	Louis	Sagers	M	2009-06-09	Boston
23	1978-02-19	Saniya	Kalloffi	M	2009-09-15	Boston
			:			
69	1962-09-18	Kish	Perna	F	2008-10-21	Boston
79	1973-10-18	Debatosh	Khasidashvili	M	2009-01-30	Boston
emp_no	birth_date	first_name	last_name	gender	hire_date	location
10	1986-02-12	Cosima	Quinn	F	2009-06-09	London
32	1977-05-23	Tzvetan	Zielinski	F	2010-02-10	London
34	1986-02-12	Berni	Genin	M	2010-03-11	London
35	1966-11-14	Cedric	Munson	M	2010-03-18	London
37	1983-09-02	Gianpaolo	Facello	M	2010-06-26	London
			:			
71	1977-01-27	Gioia	Bottazzi	F	2013-05-12	London
74	1963-06-07	Moses	Adeyemi	M	2013-05-11	London
78	1984-05-27	Santino	Sbarro	M	2009-11-06	London
80	1975-07-22	Giancarlo	Rao	M	2008-10-21	London

Generated for Anh Hoang (hoang_anhkim@bah.com) (C) Splunk Inc, not for distribution

Buttercup Games – Employee Information

emp no	RFID	IP	user	host	email	splunk role
1	417852300683	10.1.10.201	sflaemmchen	BG01-sflaemmchen	sflaemmchen@buttercupgames.com	user
2	542830538161	10.1.10.231	hsham	BG01-hsham	hsham@buttercupgames.com	user
3	520156890727	10.1.10.230	spahkthecah	BG01-spahkthecah	spahkthecah@buttercupgames.com	user
4	564931543224	10.1.10.216	sscallion	BG01-sscallion	sscallion@buttercupgames.com	user
5	534931200268	10.1.10.241	syoungin	BG01-syoungin	syoungin@buttercupgames.com	power
6	768166372290	10.1.10.290	lhaddadi	BG01-lhaddadi	lhaddadi@buttercupgames.com	power
7	659636929855	10.2.10.38	dpiazza	BG02-dpiazza	dpiazza@buttercupgames.com	user
8	559129672655	10.2.10.77	edutra	BG02-edutra	edutra@buttercupgames.com	power
9	960318676000	10.2.10.45	lsagers	BG02-lsagers	lsagers@buttercupgames.com	power
10	513908343176	10.3.10.28	cquinn	BG03-cquinn	cquinn@buttercupgames.com	admin
11	125179529264	10.1.10.234	ptoscani	BG01-ptoscani	ptoscani@buttercupgames.com	power
12	382839148784	10.1.10.238	myuan	BG01-myuan	myuan@buttercupgames.com	power
13	713929421175	10.1.10.246	acurry	BG01-acurry	acurry@buttercupgames.com	power
14	900191452102	10.1.10.252	blu	BG01-blu	blu@buttercupgames.com	user
				:		
				:		
81	999999999999	10.1.10.1	bpony	BG01-bpony	bpony@buttercupgames.com	user

Generated for Anh Hoang (hoang_anhkim@bah.com) (C) Splunk Inc, not for distribution

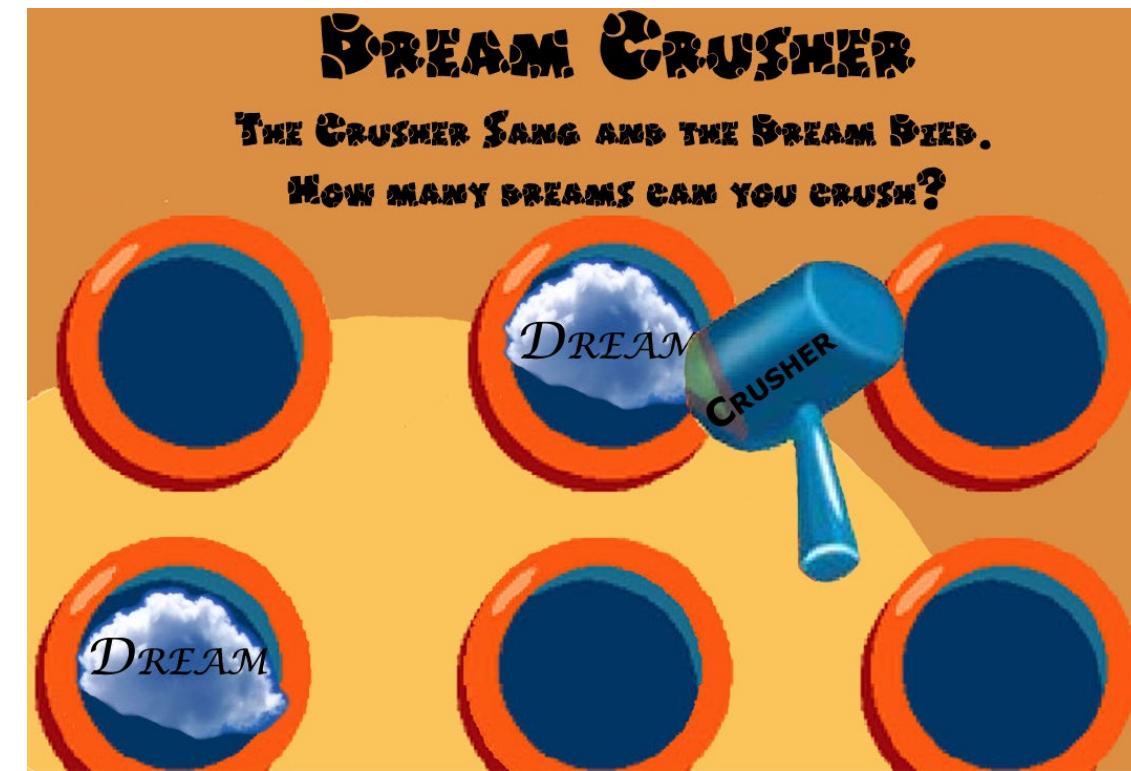
Vendor Sales – Sample

<u>Vendor</u>	<u>VendorCity</u>	<u>VendorCountry</u>	<u>VendorID</u>	<u>VendorLatitude</u>	<u>VendorLongitude</u>	<u>VendorStateProvince</u>
Frozen Fun General Store	Amundsen-Scott Station	Antarctica	9999	90.0000	139.2667	Antarctica
Jeremy's House of Hobbies	Fort-Lamy	Chad	9116	12.134846	15.055742	Chari-Baguirmi
Pan-African RC and Toys	Ouagadougou	Burkina Faso	9115	12.364637	-1.533864	Nord Region
Passe-Temps	Yamoussoukro	Cote d'Ivoire	9114	6.816667	-5.283333	Lacs
Kahled's Amusements	Tripoli	Libya	9113	5.560735	-0.193087	Tripoli
Mburo Games	Kampala	Uganda	9112	0.313611	32.581111	Kampala
Pan-African RC and Toys	Yaounde	Cameroon	9111	3.866667	11.516667	Centre Region
Comics and Games	Dar es Salaam	Tanzania	9110	-6.822921	39.269661	Dar es Salaam
Pan-African RC and Toys	Mombasa	Kenya	9109	-4.043477	39.668207	Mombasa
Pan-African RC and Toys	Accra	Ghana	9108	5.555717	-0.196306	Greater Accra
Seminna-Werq Games Warehouse	Addis Ababa	Ethiopia	9107	9.022736	38.746799	Oromia
RTL Boutique de Train Miniature	Tunis	Tunisia	9106	36.81881	10.16596	Tunis
Rick's Toy Shop and Cafe	Casablanca	Morocco	9105	33.533333	-7.583333	Grand Casablanca
Laval's Joke and Toy Store	Oran	Algeria	9104	35.696944	-0.633056	Oran
Sweepstake Games	Lagos	Nigeria	9103	6.441158	3.417977	Lagos
Lightening Games of Johannesburg	Johannesburg	South Africa	9102	-26.204103	28.047305	Gauteng
Natal Games of Pietermaritzburg	Pietermaritzburg	South Africa	9101	-29.600607	30.379412	KwaZulu-Natal
Peers Games of Cape Town	Cape Town	South Africa	9100	-33.924868	18.424055	Western Cape
Kiwi Game Warehouse	Auckland	New Zealand	7045	-36.84846	174.763332	Auckland
Kiwi Game Warehouse	Christchurch	New Zealand	7044	-43.529854	172.637888	Canterbury

Generated for Anh Hoang (hoang_anhkim@bah.com) (C) Splunk Inc, not for distribution

Buttercup Games – Products

productId	product_name	categoryId
WC-SH-T02	World of Cheese Tee	TEE
WC-SH-G04	World of Cheese	SHOOTER
WC-SH-A02	Fire Resistance Suit of Provolone	ACCESSORIES
WC-SH-A01	Holy Blade of Gouda	ACCESSORIES
SC-MG-G10	SIM Cubicle	SIMULATION
PZ-SG-G05	Puppies vs. Zombies	STRATEGY
MB-AG-T01	Manganiello Bros. Tee	TEE
MB-AG-G07	Manganiello Bros.	ARCADE
FS-SG-G03	Final Sequel	STRATEGY
FI-AG-G08	Orvil the Wolverine	ARCADE
DC-SG-G02	Dream Crusher	STRATEGY
DB-SG-G01	Mediocre Kingdoms	STRATEGY
CU-PG-G06	Curling 2014	SPORTS
BS-AG-G09	Benign Space Debris	ARCADE



Generated for Anh Hoang (hoang_anhkim@bah.com) (C) Splunk Inc, not for distribution