Splunk Fundamentals 2
Instructor on Demand

Generated for (v.ms.besetti@jpmchase.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.

Generated for (v.ms.besetti@jpmchase.com) (C) Splunk Inc, not for distribution

Copyright © 2018 Splunk, Inc. All rights reserved | 14 June 2018

Course Guidelines

- Hands-on lab exercises reinforce information presented in the lecture modules
- To receive a 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

Generated for (v.ms.besetti@jpmchase.com) (C) Splunk Inc, not for distribution

Copyright © 2018 Splunk, Inc. All rights reserved | 14 June 2018

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 (v.ms.besetti@jpmchase.com) (C) Splunk Inc, not for distribution

Copyright © 2018 Splunk, Inc. All rights reserved | 14 June 2018

Course Outline

Module 1: Course Introduction
Module 2: Beyond Search Fundamentals
Module 3: Commands for Visualizations
Module 4: Advanced Visualizations
Module 5: Filtering and Formatting Data
Module 6: Correlating Events
Module 7: Introduction to Knowledge Objects
Module 8: Creating and Managing Fields
Module 9: Creating Field Aliases and Calculated Fields
Module 10: Creating Tags and Event Types
Module 11: Creating and Using Macros
Module 12: Creating and Using Workflow Actions
Module 13: Creating Data Models
Module 14: Using the Common Information Model (CIM) Add-On

Generated for (v.ms.besetti@jpmchase.com) (C) Splunk Inc, not for distribution

Copyright © 2018 Splunk, Inc. All rights reserved | 14 June 2018

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

Generated for (v.ms.besetti@jpmchase.com) (C) Splunk Inc, not for distribution

Copyright © 2018 Splunk, Inc. All rights reserved | 14 June 2018

Buttercup Games, Inc.

- Multinational company with HQ in San Francisco and offices in Boston and London
- Sells product mainly through its worldwide chain of third party stores, but also sells through its online store

Generated for (v.ms.besetti@jpmchase.com) (C) Splunk Inc, not for distribution

Copyright © 2018 Splunk, Inc. All rights reserved

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.

Generated for (v.ms.besetti@jpmchase.com) (C) Splunk Inc, not for distribution

Copyright © 2018 Splunk, Inc. All rights reserved

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 (v.ms.besetti@jpmchase.com) (C) Splunk Inc, not for distribution

Copyright © 2018 Splunk, Inc. All rights reserved | 14 June 2018

Module 2: Beyond Search Fundamentals

Generated for (v.ms.besetti@jpmchase.com) (C) Splunk Inc, not for distribution

Copyright © 2018 Splunk, Inc. All rights reserved

Module Objectives

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

Generated for (v.ms.besetti@jpmchase.com) (C) Splunk Inc, not for distribution

Copyright © 2018 Splunk, Inc. All rights reserved | 14 June 2018

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 (v.ms.besetti@jpmchase.com) (C) Splunk Inc, not for distribution

Copyright © 2018 Splunk, Inc. All rights reserved | 14 June 2018

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)

Generated for (v.ms.besetti@jpmchase.com) (C) Splunk Inc, not for distribution

Copyright © 2018 Splunk, Inc. All rights reserved | 14 June 2018

Case Sensitivity – Sensitive

Case sensitive	Examples
Boolean operators (uppercase)	`AND`, `OR`, `NOT` (Boolean operators) and, or, not (literal keywords)
Field names	`productId` vs. `productid` `eval cs_username = "Total Access"`
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`

eval and where commands	`eval action=if(action=="view",...)`
	`where action="Purchase"`
	`stats count(eval(action="view") as…`

Tags	`tag=DMZ` vs. `tag=dmz`

Generated for (v.ms.besetti@jpmchase.com) (C) Splunk Inc, not for distribution

Copyright © 2018 Splunk, Inc. All rights reserved | 14 June 2018

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`

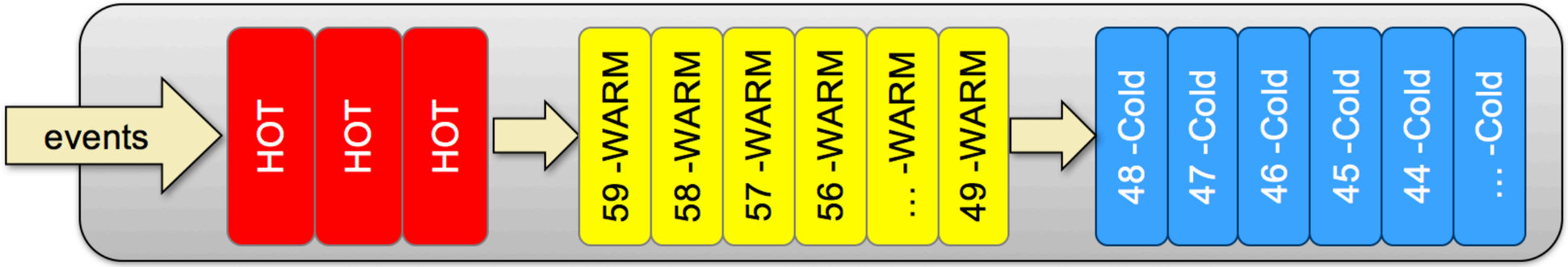
Search terms	`failed, FAILED, Failed`
Statistical functions	`avg, AVG, Avg used by stats, chart, …`
Field values	`host=www1, host=WWW1` (unless coming from a lookup)

Generated for (v.ms.besetti@jpmchase.com) (C) Splunk Inc, not for distribution

Copyright © 2018 Splunk, Inc. All rights reserved | 14 June 2018

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 add more



Generated for (v.ms.besetti@jpmchase.com) (C) Splunk Inc, not for distribution

Copyright © 2018 Splunk, Inc. All rights reserved | 14 June 2018

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 the indexes of 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 (v.ms.besetti@jpmchase.com) (C) Splunk Inc, not for distribution

Copyright © 2018 Splunk, Inc. All rights reserved | 14 June 2018

General Search Practices

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

`index=web sourcetype=access_combined`	15.16
`index=web sourcetype=access_combined | fields clientip bytes referrer`	4.49

Generated for (v.ms.besetti@jpmchase.com) (C) Splunk Inc, not for distribution

Copyright © 2018 Splunk, Inc. All rights reserved | 14 June 2018

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

Generated for (v.ms.besetti@jpmchase.com) (C) Splunk Inc, not for distribution

Copyright © 2018 Splunk, Inc. All rights reserved | 14 June 2018

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

Generated for (v.ms.besetti@jpmchase.com) (C) Splunk Inc, not for distribution

Copyright © 2018 Splunk, Inc. All rights reserved | 14 June 2018

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`

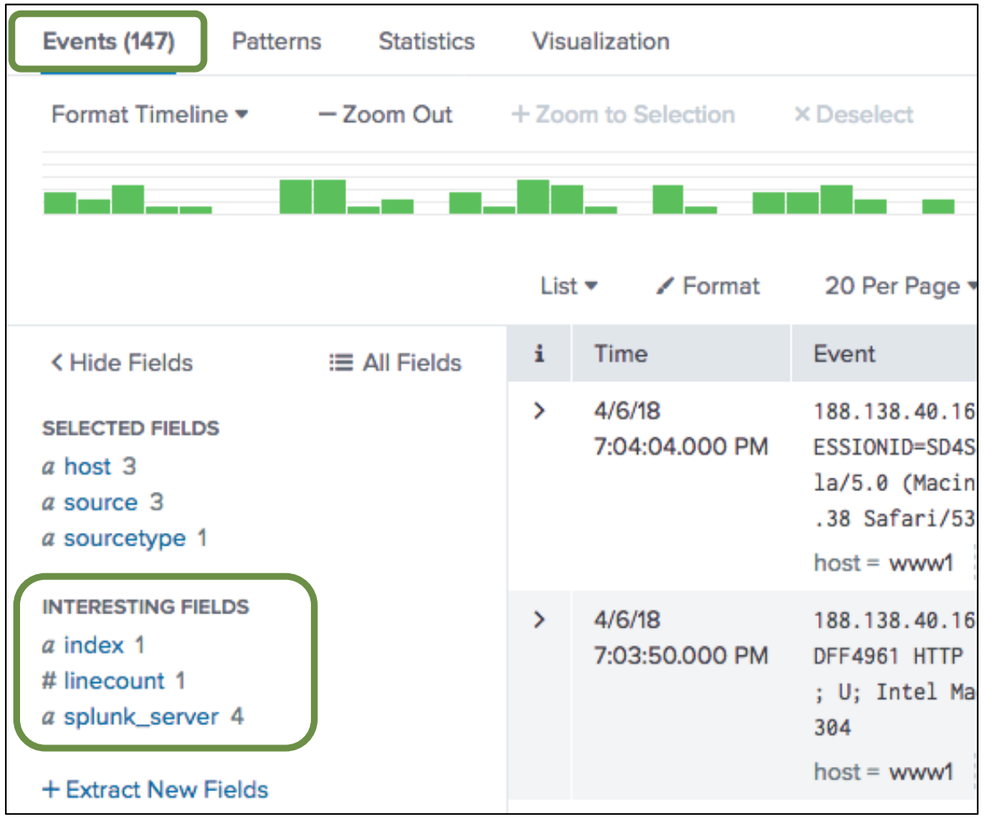
Generated for (v.ms.besetti@jpmchase.com) (C) Splunk Inc, not for distribution

Copyright © 2018 Splunk, Inc. All rights reserved | 14 June 2018

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
```



Generated for (v.ms.besetti@jpmchase.com) (C) Splunk Inc, not for distribution

Copyright © 2018 Splunk, Inc. All rights reserved | 14 June 2018

Reviewing Search Mode – Fast Mode (cont.)

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

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

Events | Patterns | Statistics (5) | Visualization

20 Per Page ▾ | Format | Preview ▾

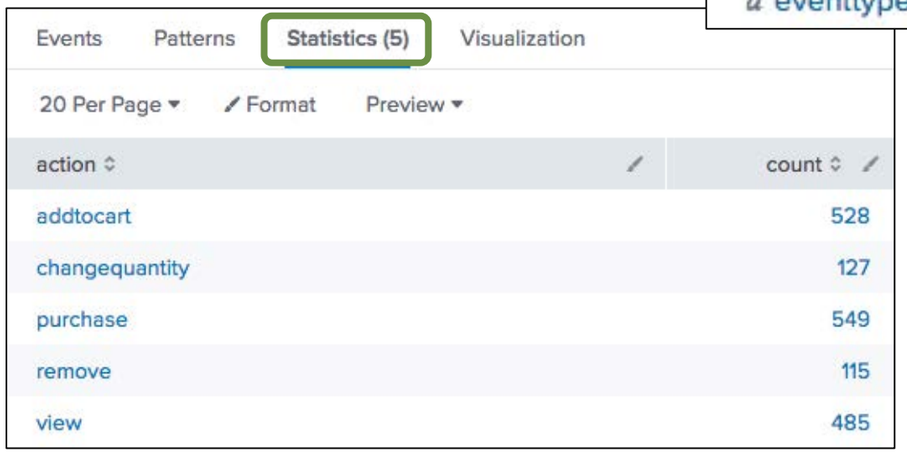
action	count
addtocart	19
changequantity	4
purchase	23

Generated for (v.ms.besetti@jpmchase.com) (C) Splunk Inc, not for distribution

Copyright © 2018 Splunk, Inc. All rights reserved | 14 June 2018

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

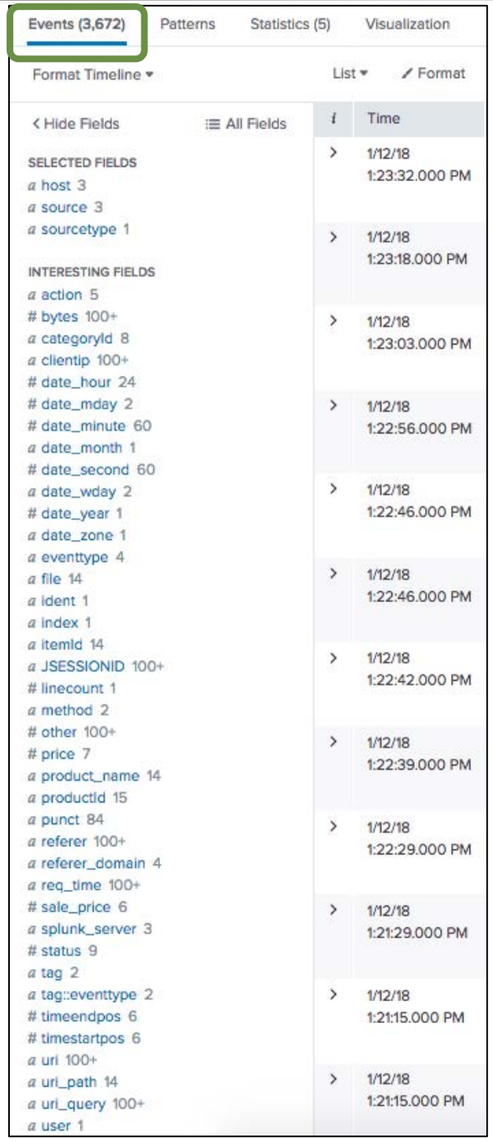


Generated for (v.ms.besetti@jpmchase.com) (C) Splunk Inc, not for distribution

Copyright © 2018 Splunk, Inc. All rights reserved | 14 June 2018

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



Generated for (v.ms.besetti@jpmchase.com) (C) Splunk Inc, not for distribution

Copyright © 2018 Splunk, Inc. All rights reserved | 14 June 2018

Search Performance – Modes

- Use the most appropriate search mode:

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

- Time range: `last 365 days`

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

Note

These searches were run in our lab environment, not a production environment - your results may vary.

Generated for (v.ms.besetti@jpmchase.com) (C) Splunk Inc, not for distribution

Copyright © 2018 Splunk, Inc. All rights reserved | 14 June 2018

Search Performance – Types of Searches

Search	Description	Indexer throughput
Dense	A large percentage of the data matches the search	Up to 50K matching EPS (Events per second) *CPU bound*
	Use Cases: computing stats, reporting	
	`index=web sourcetype=access_combined | timechart count`	
Sparse	A small percentage of data matches the search	Up to 5K matching EPS *CPU bound*
	Use Cases: troubleshooting, error analysis	
	`index=web sourcetype=access_combined status=404 | timechart count`	
Super Sparse	Returns a small number of results from each index bucket matching the search	Up to 2 seconds per index bucket *I/O bound*
	I/O intensive as the indexer looks through all of an index's buckets	
	With a lot of data, with a lot of buckets, it can take a long time to finish	
	`index=network sourcetype=cisco_wsa_squid action=denied src_ip=10.2.3.11`	
Rare	The indexer checks all buckets to find results, but bloom filters eliminate those buckets that don't include search results	Up to 10-50 index buckets/second *I/O bound*
	Use Cases: user behavior tracking	
	`index=web sourcetype=access_combined sessionID=1234`	

Generated for (v.ms.besetti@jpmchase.com) (C) Splunk Inc, not for distribution

Copyright © 2018 Splunk, Inc. All rights reserved

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

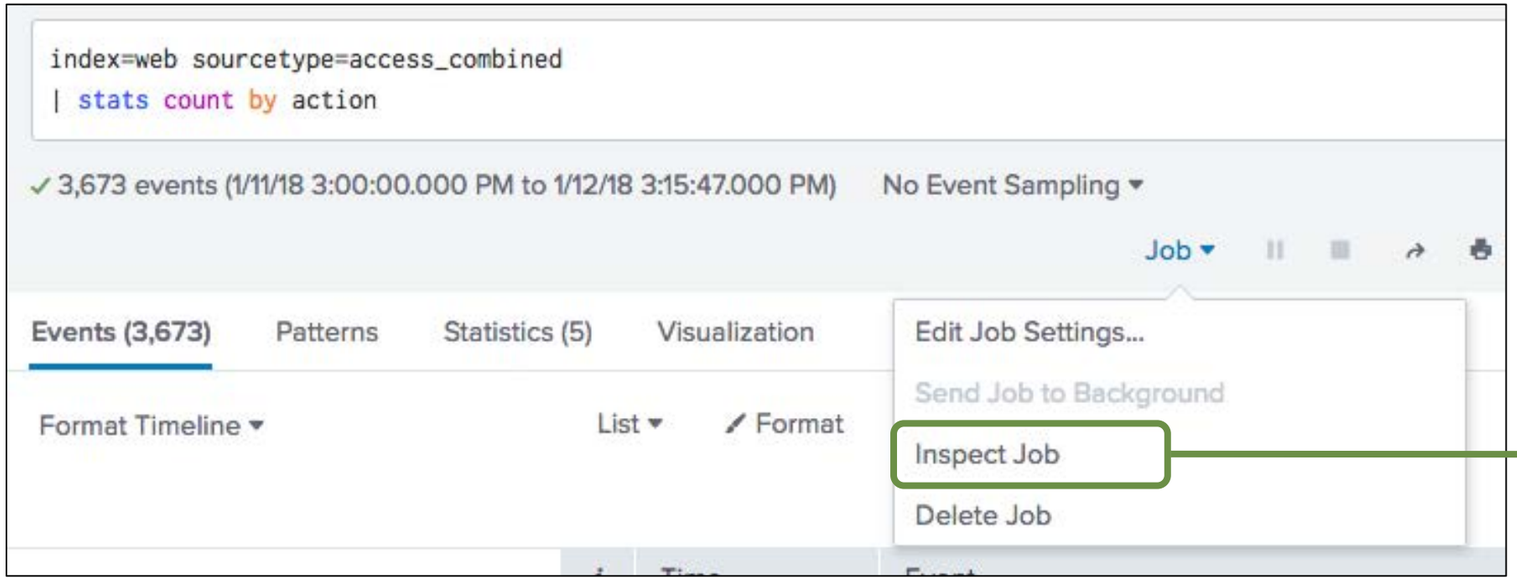
Note

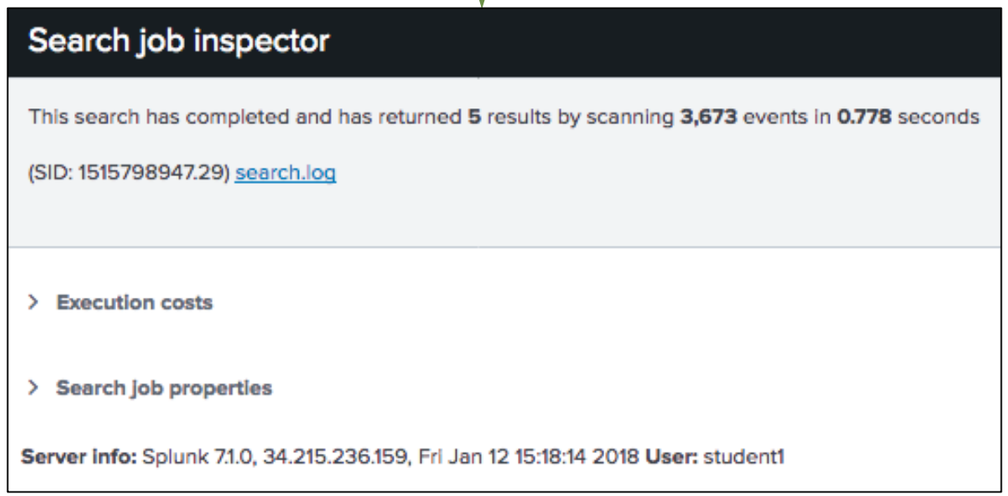
For more information, see:
docs.splunk.com/Documentation/Splunk/latest/Search/ViewsearchjobpropertieswiththeJobInspector

Generated for (v.ms.besetti@jpmchase.com) (C) Splunk Inc, not for distribution

Copyright © 2018 Splunk, Inc. All rights reserved | 14 June 2018

Search Job Inspector – 3 Components



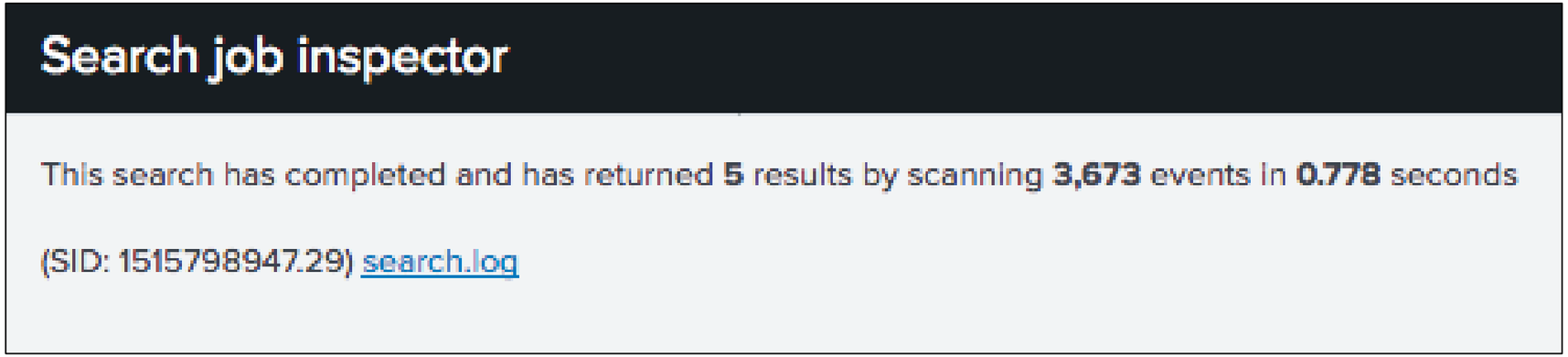


- Header
- Execution costs
- Search job properties

Generated for (v.ms.besetti@jpmchase.com) (C) Splunk Inc, not for distribution

Copyright © 2018 Splunk, Inc. All rights reserved

Search Job Inspector – Header



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

Generated for (v.ms.besetti@jpmchase.com) (C) Splunk Inc, not for distribution

Copyright © 2018 Splunk, Inc. All rights reserved | 14 June 2018

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.fieldalias	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 (v.ms.besetti@jpmchase.com) (C) Splunk Inc, not for distribution

Copyright © 2018 Splunk, Inc. All rights reserved | 14 June 2018

Search Job Inspector – Search Job Properties

Example:

```
index=sales sourcetype=vendor_sales VendorID>=7000 AND VendorID<9000
```

Last 365 days

- Produces `scanCount` of 127,201 events
- Returns `resultCount` of 2,144 in 3.01 seconds
- To calculate performance:
  - Do **not** use `resultCount/time` 2,144 / 3.01 = 712 EPS*
  - Rather, calculate `scanCount/time` 127,201/ 3.01 = 40,892 EPS

* EPS= events per second

Generated for (v.ms.besetti@jpmchase.com) (C) Splunk Inc, not for distribution

Copyright © 2018 Splunk, Inc. All rights reserved | 14 June 2018

Module 3:
Commands for Visualizations

Generated for (v.ms.besetti@jpmchase.com) (C) Splunk Inc, not for distribution

Copyright © 2018 Splunk, Inc. All rights reserved

Module Objectives

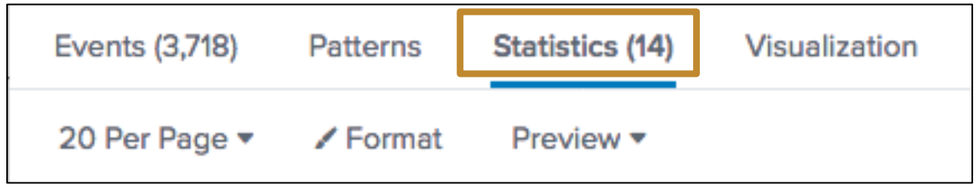
- Explore data structure requirements
- Explore visualization types
- Create and format charts
- Create and format `timecharts`
- Explain when to use each type of reporting command

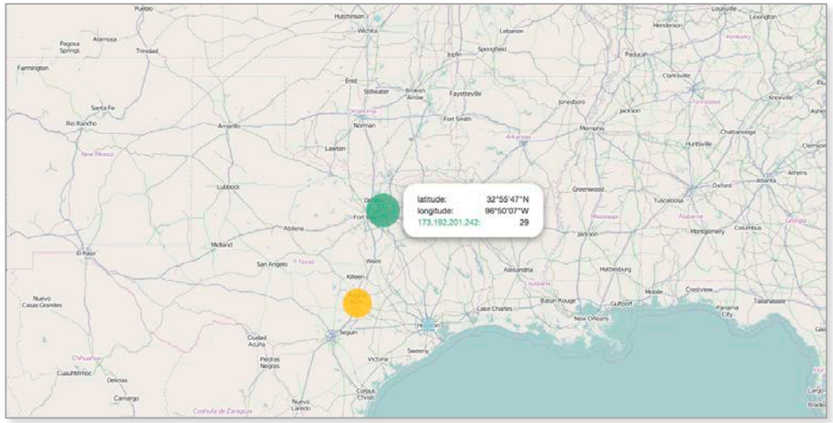
Generated for (v.ms.besetti@jpmchase.com) (C) Splunk Inc, not for distribution

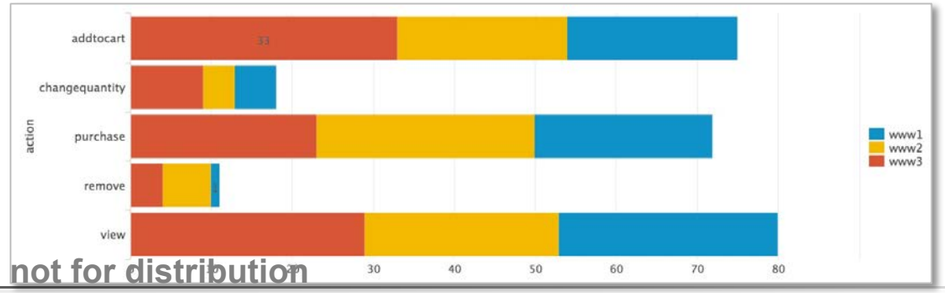
Copyright © 2018 Splunk, Inc. All rights reserved | 14 June 2018

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





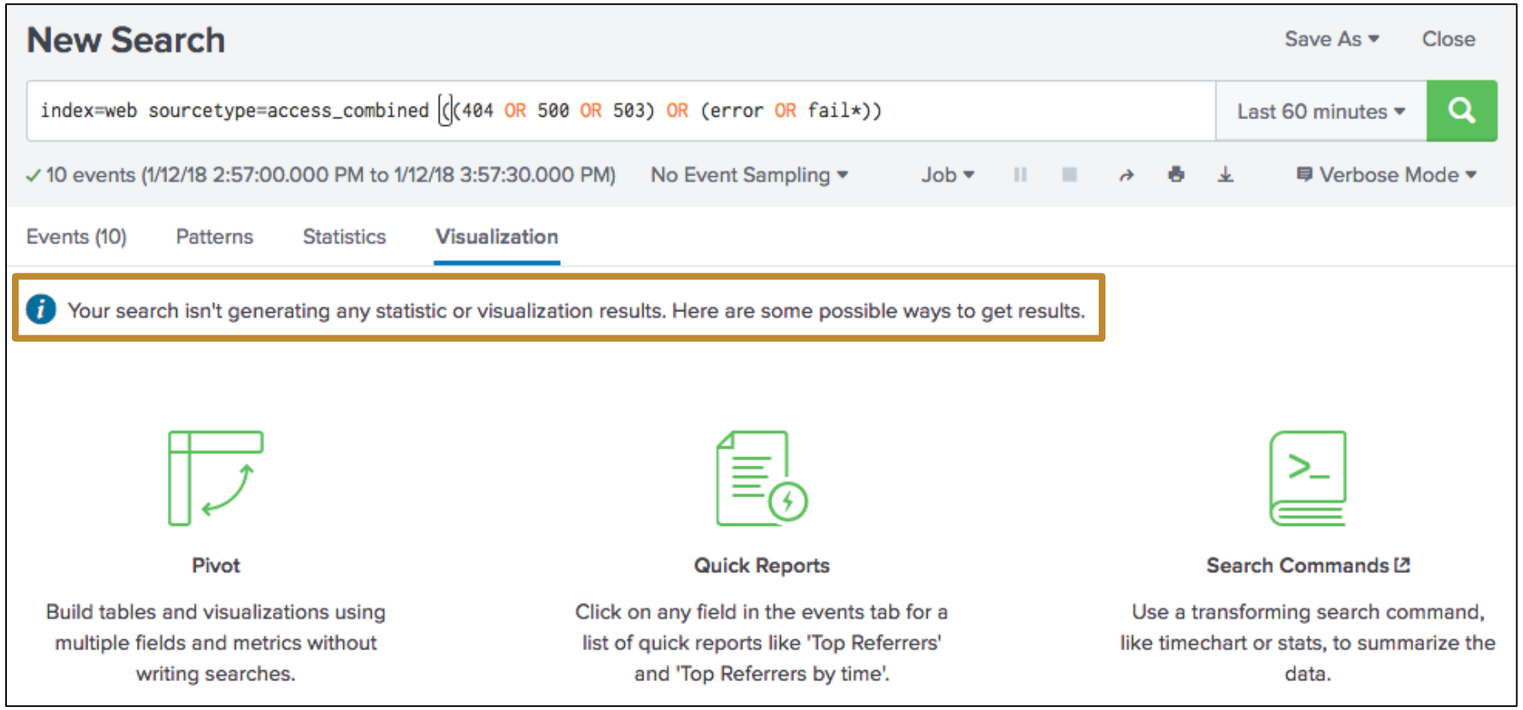


Generated for (v.ms.besetti@jpmchase.com) (C) Splunk Inc, not for distribution

Copyright © 2018 Splunk, Inc. All rights reserved

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 any statistical or visualization results



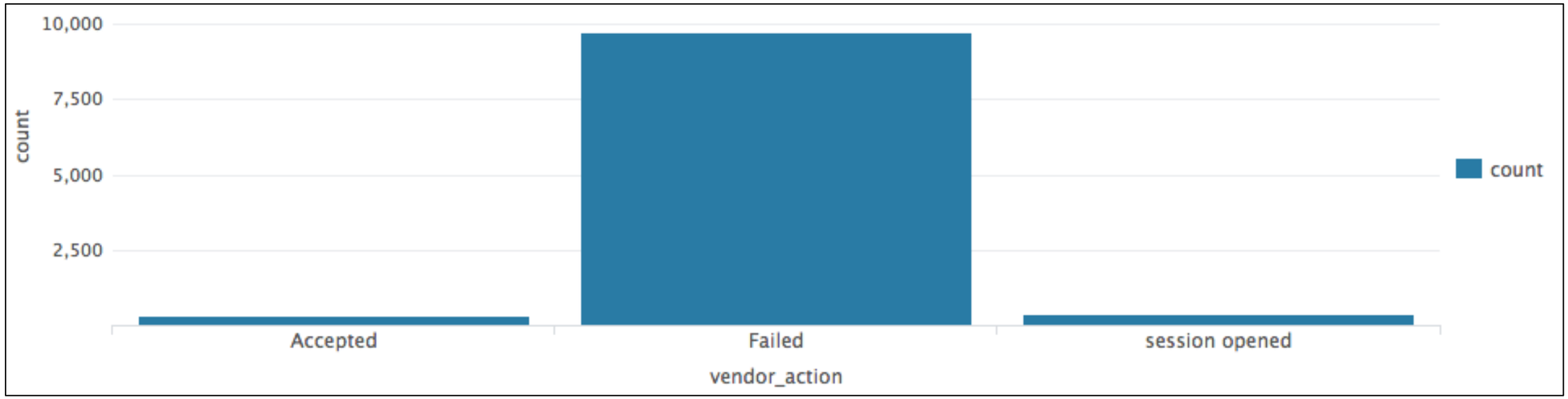
Generated for (v.ms.besetti@jpmchase.com) (C) Splunk Inc, not for distribution

Copyright © 2018 Splunk, Inc. All rights reserved | 14 June 2018

Data Structure Requirements – Single Series

- Most visualizations require search results structured as tables, with at least two columns, a **single series**
  - 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



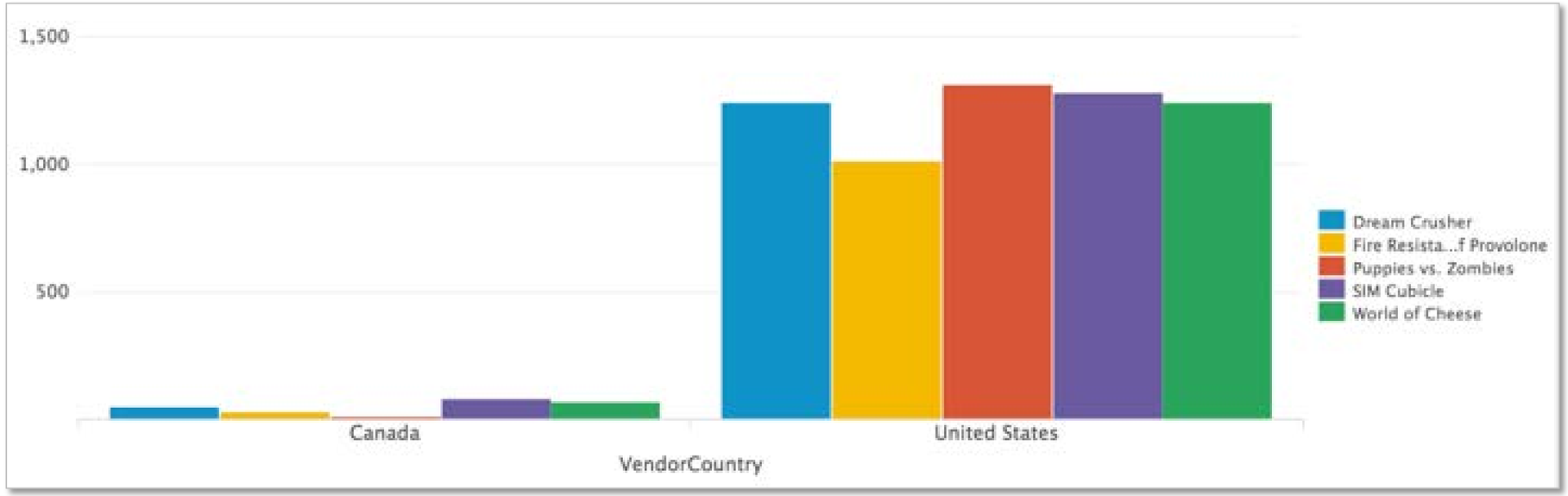
Generated for (v.ms.besetti@jpmchase.com) (C) Splunk Inc, not for distribution

Copyright © 2018 Splunk, Inc. All rights reserved | 14 June 2018

Data Structure Requirements – Multi-Series

To get **multi-series** tables, you need to set up the underlying search with reporting search commands like **`chart`** or **`timechart`**

VendorCountry	Dream Crusher	Fire Resistance Suit of Provolone	Puppies vs. Zombies	SIM Cubicle	World of Cheese
Canada	10	3	2	13	13
United States	230	202	230	228	223




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

Last 30 days

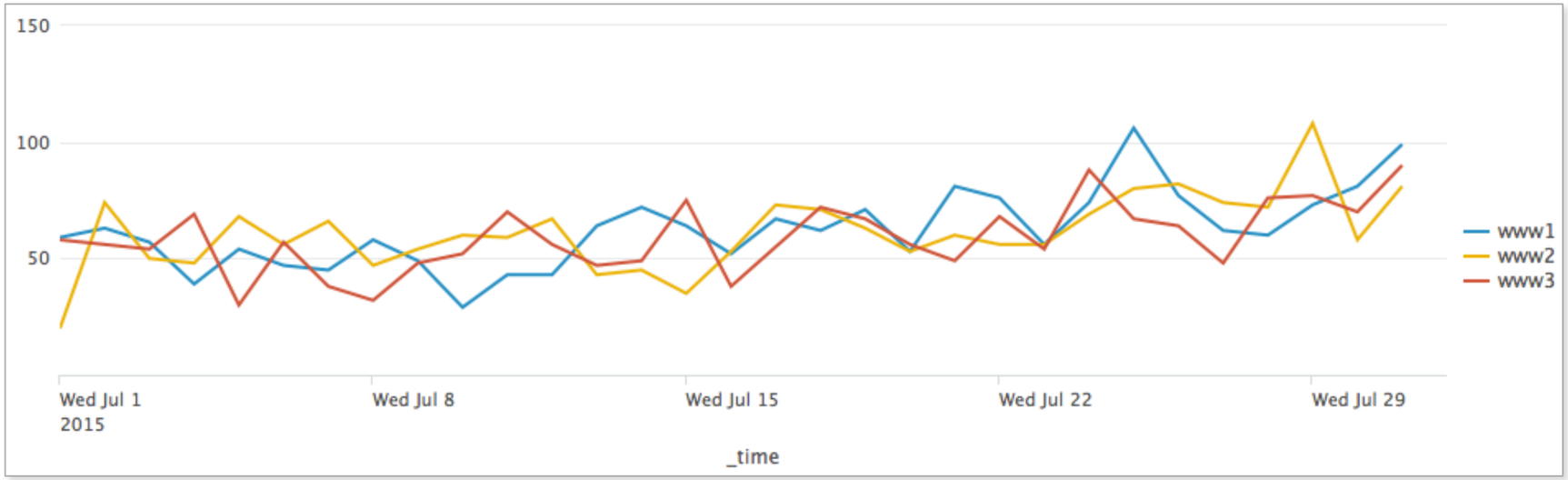
Generated for (v.ms.besetti@jpmchase.com) (C) Splunk Inc, not for distribution

Copyright © 2018 Splunk, Inc. All rights reserved | 14 June 2018

Data Structure Requirements – Time Series

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

_time	www1	www2	www3
2017-12-01	168	190	158
2017-12-02	156	171	168
2017-12-03	138	150	151
2017-12-04	169	162	195
2017-12-05	135	170	150




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

Previous month

Generated for (v.ms.besetti@jpmchase.com) (C) Splunk Inc, not for distribution

Copyright © 2018 Splunk, Inc. All rights reserved

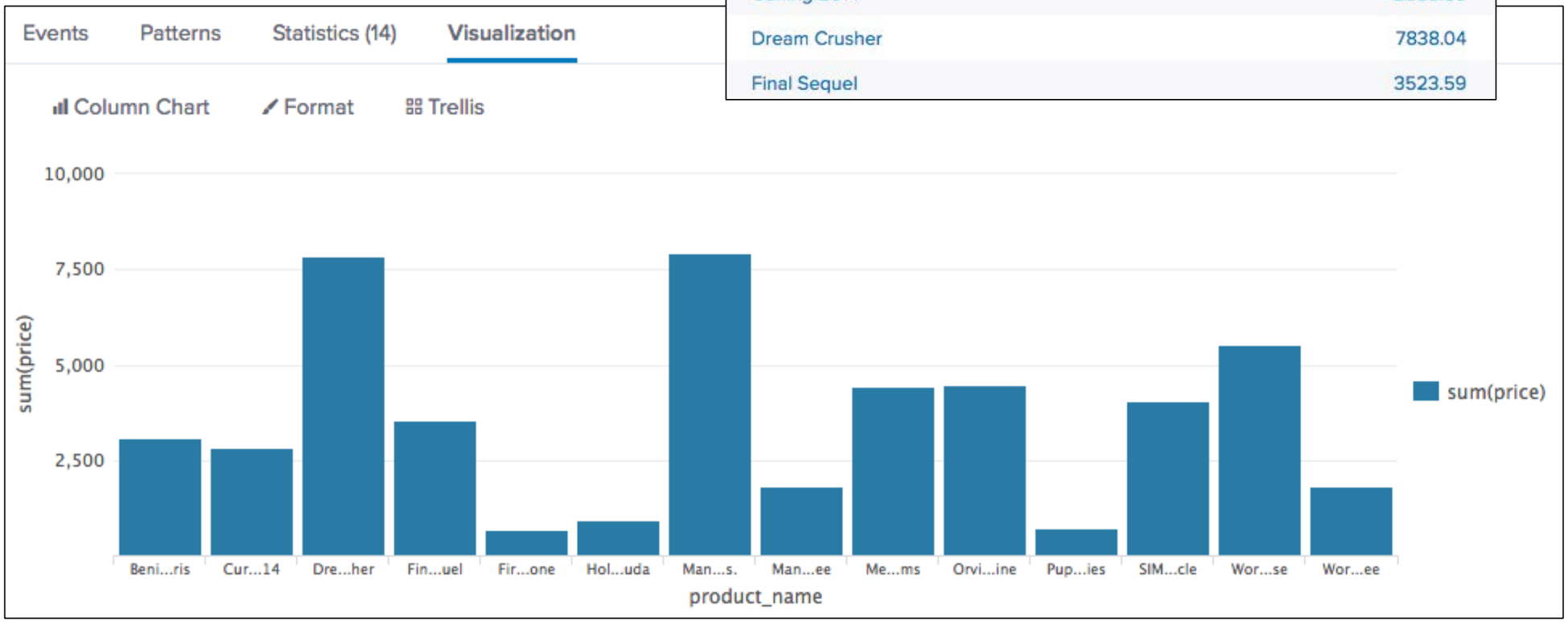
Viewing Results as a Chart

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

Events Patterns Statistics (14) Visualization

20 Per Page ▾ Format Preview ▾

product_name	sum(price)
Benign Space Debris	3098.76
Curling 2014	2838.58
Dream Crusher	7838.04
Final Sequel	3523.59

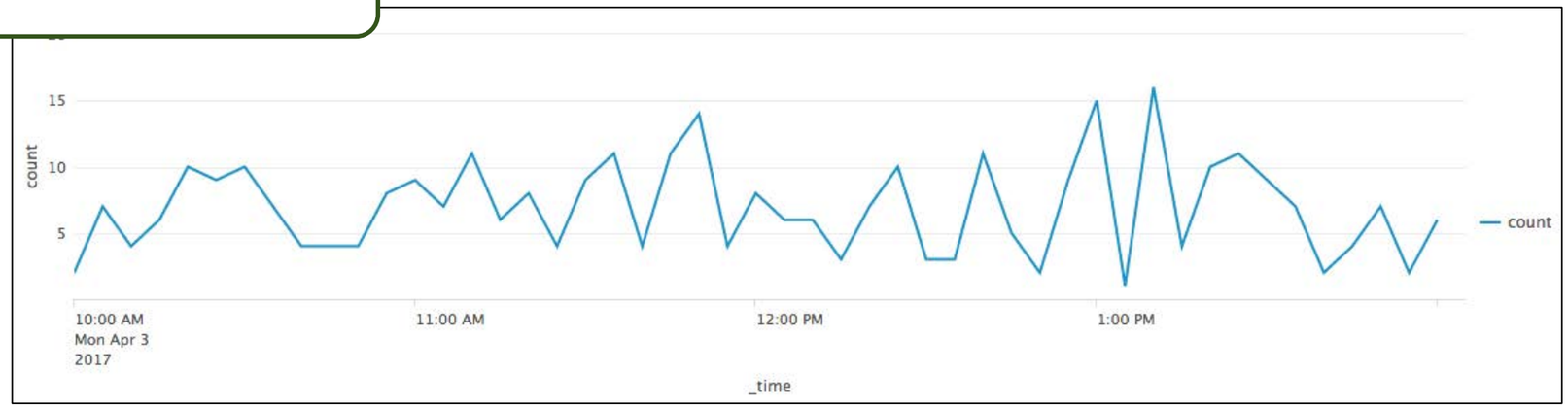


Generated for (v.ms.besetti@jpmchase.com) (C) Splunk Inc, not for distribution

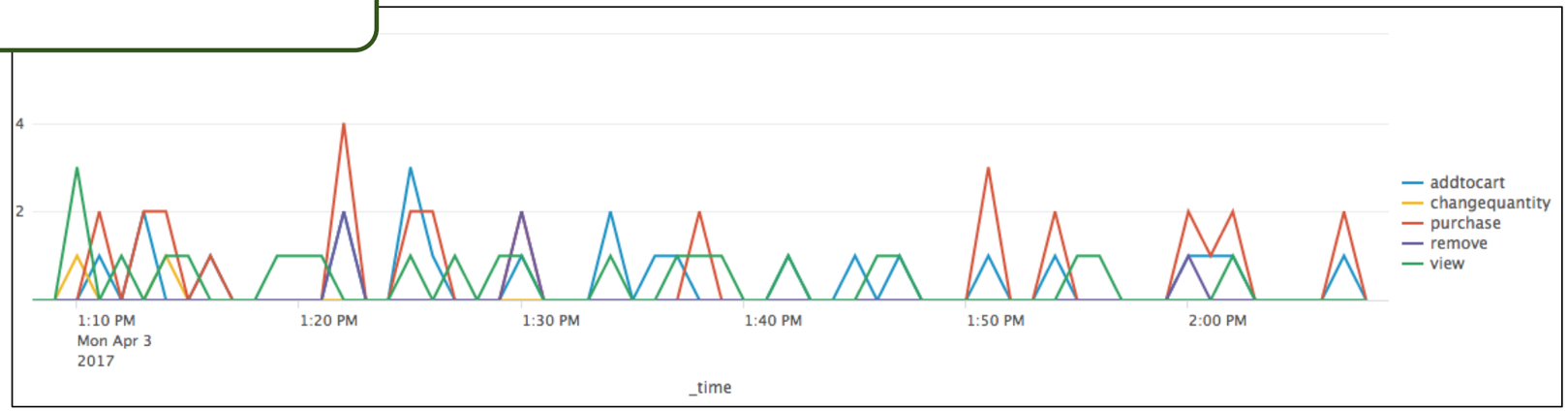
Copyright © 2018 Splunk, Inc. All rights reserved | 14 June 2018

Charts – Line

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




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

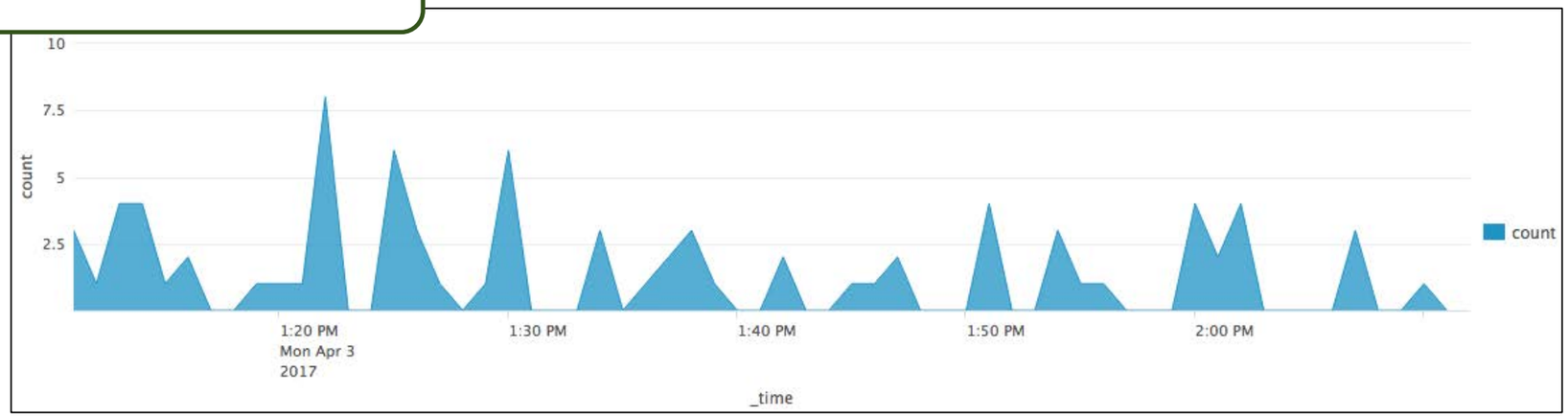


Generated for (v.ms.besetti@jpmchase.com) (C) Splunk Inc, not for distribution

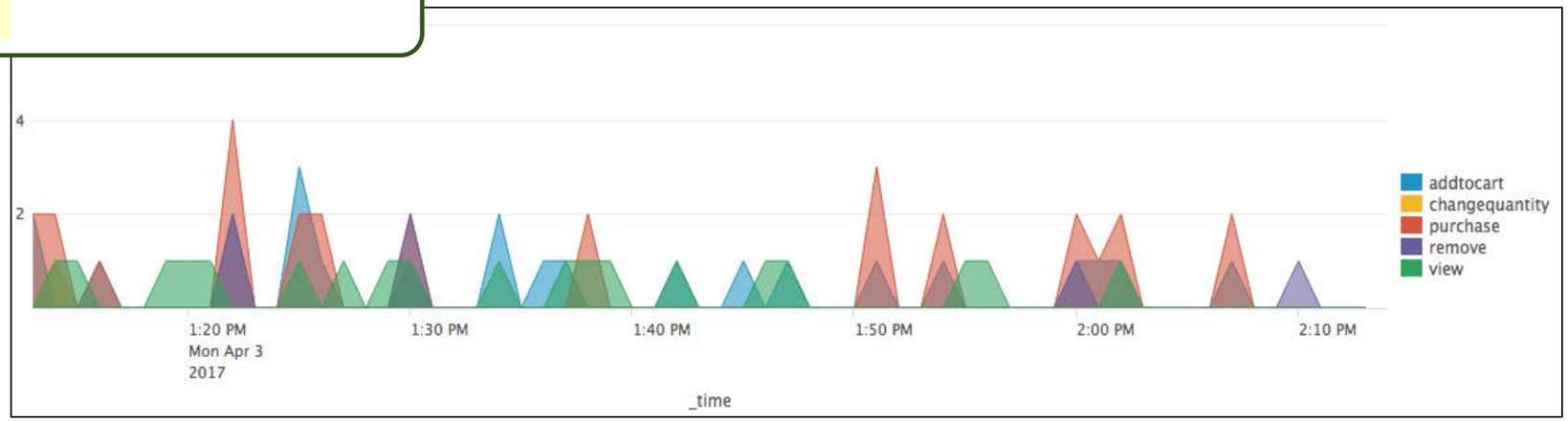
Copyright © 2018 Splunk, Inc. All rights reserved | 14 June 2018

Charts – Area

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




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



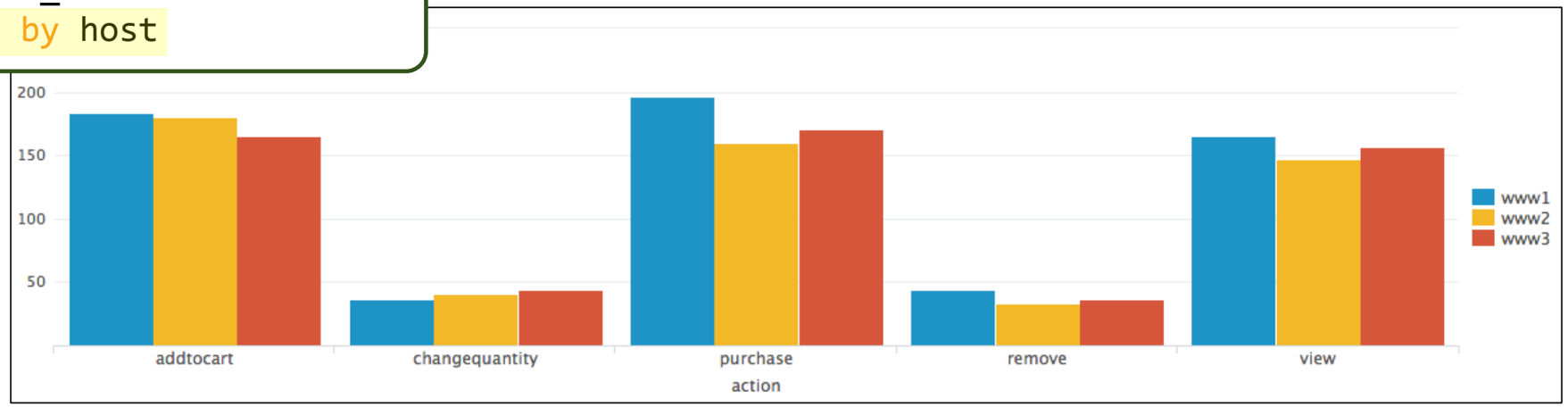
Generated for (v.ms.besetti@jpmchase.com) (C) Splunk Inc, not for distribution

Copyright © 2018 Splunk, Inc. All rights reserved | 14 June 2018

Charts – Column

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

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

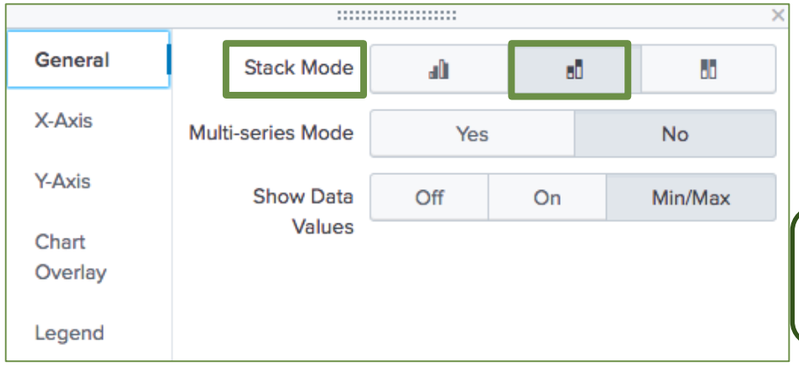


Generated for (v.ms.besetti@jpmchase.com) (C) Splunk Inc, not for distribution

Copyright © 2018 Splunk, Inc. All rights reserved | 14 June 2018

Charts – Column (Formatted as Stacked)

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



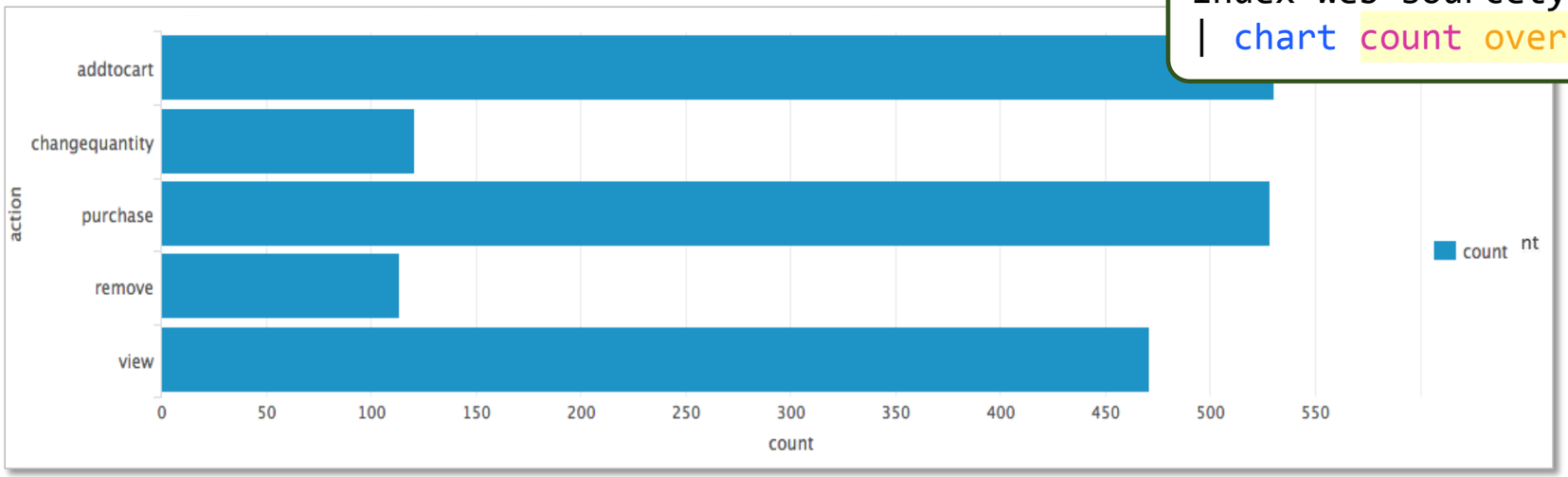
```
index=web sourcetype=access_combined action=*
| chart count over action by host
```

Generated for (v.ms.besetti@jpmchase.com) (C) Splunk Inc, not for distribution

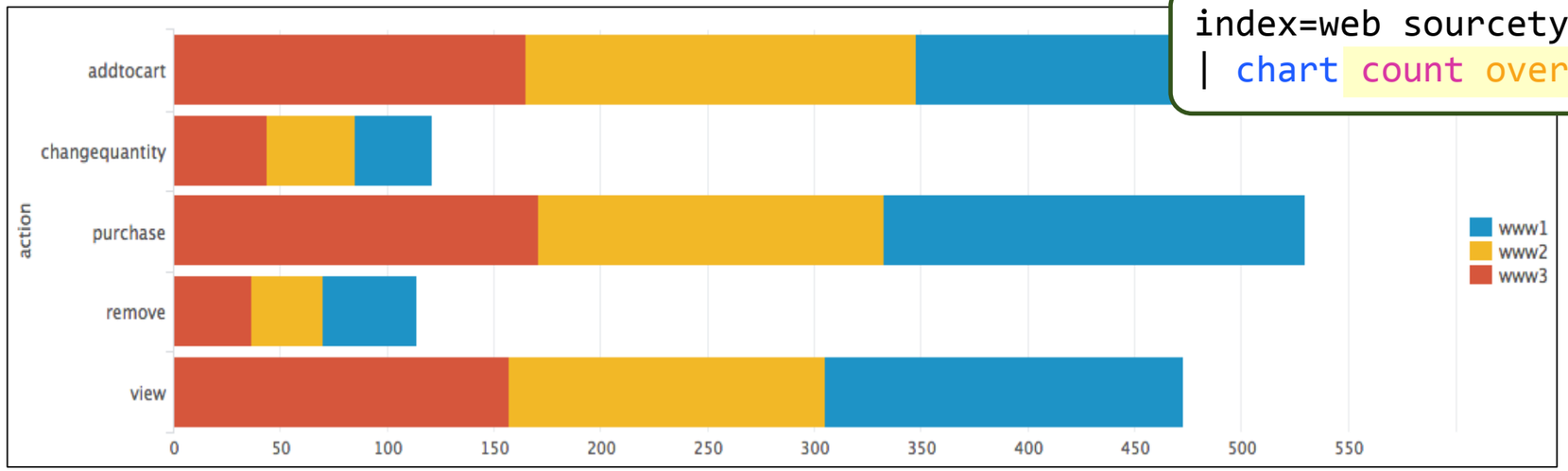
Copyright © 2018 Splunk, Inc. All rights reserved | 14 June 2018

Charts – Bar

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




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

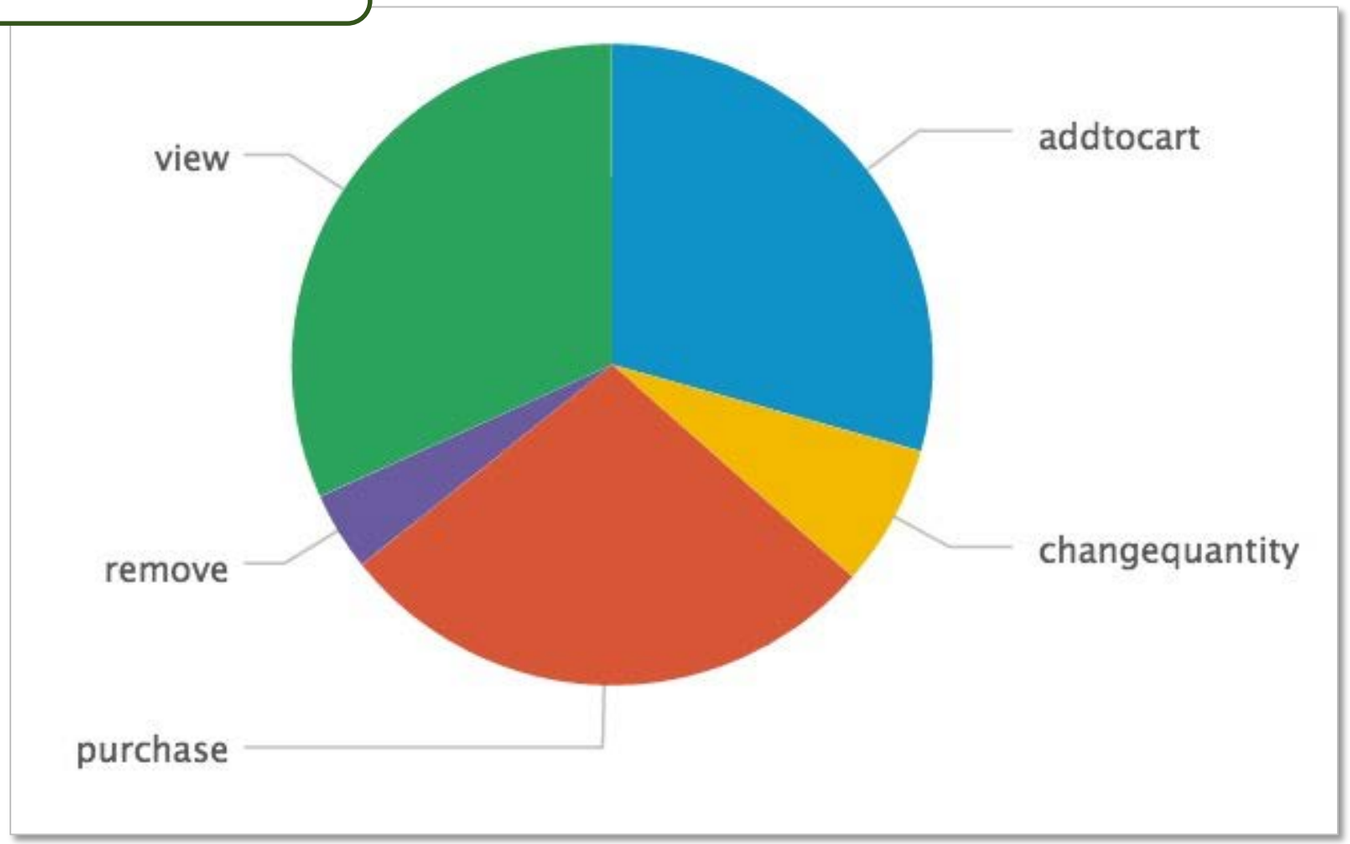


Generated for (v.ms.besetti@jpmchase.com) (C) Splunk Inc, not for distribution

Copyright © 2018 Splunk, Inc. All rights reserved | 14 June 2018

Charts – Pie

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



Generated for (v.ms.besetti@jpmchase.com) (C) Splunk Inc, not for distribution

Copyright © 2018 Splunk, Inc. All rights reserved | 14 June 2018

Charts – Scatter

- Scatter chart shows trends in the relationships between discrete data values

```
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
```

- Generally, it shows discrete values that do not occur at regular intervals or belong to a series

Note

Although three values are returned from this search, a scatter chart can only **show** two dimensions.

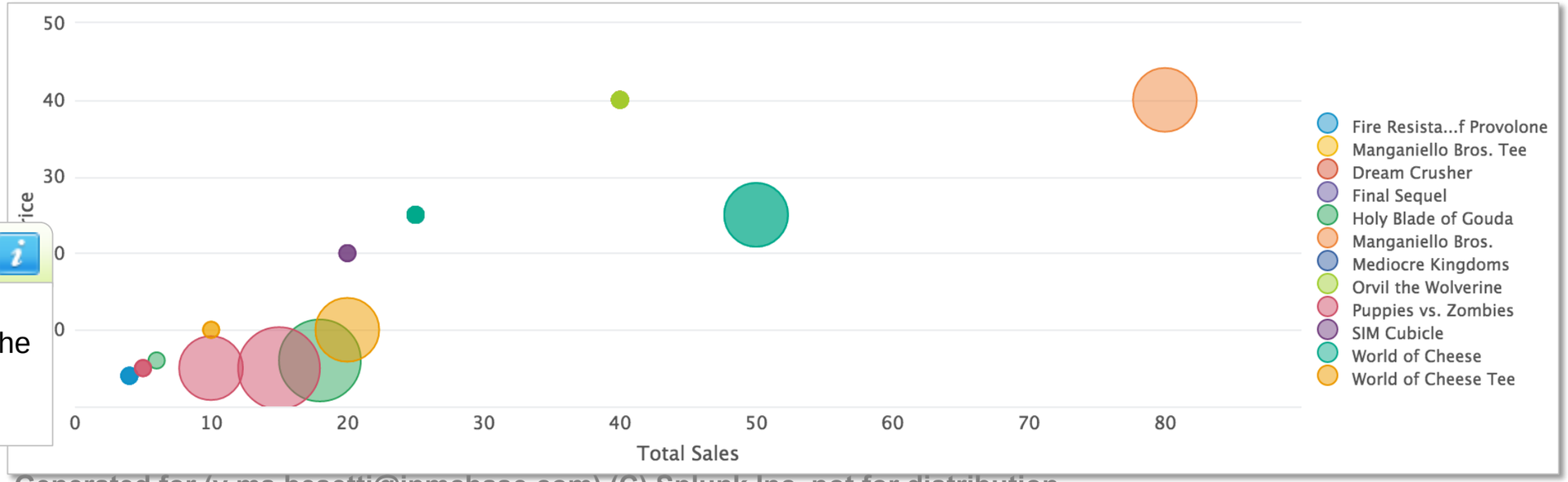
Generated for (v.ms.besetti@jpmchase.com) (C) Splunk Inc, not for distribution

Copyright © 2018 Splunk, Inc. All rights reserved | 14 June 2018

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
```



Note

Compare this with the two-dimensional scatter chart from the same search, shown on the previous slide.

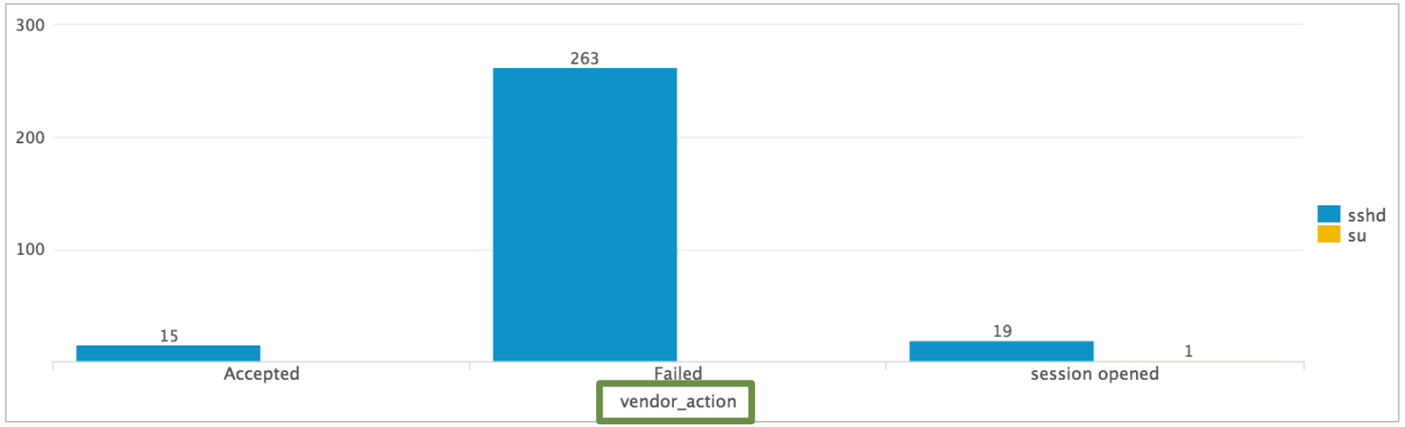
Generated for (v.ms.besetti@jpmchase.com) (C) Splunk Inc, not for distribution

Copyright © 2018 Splunk, Inc. All rights reserved | 14 June 2018

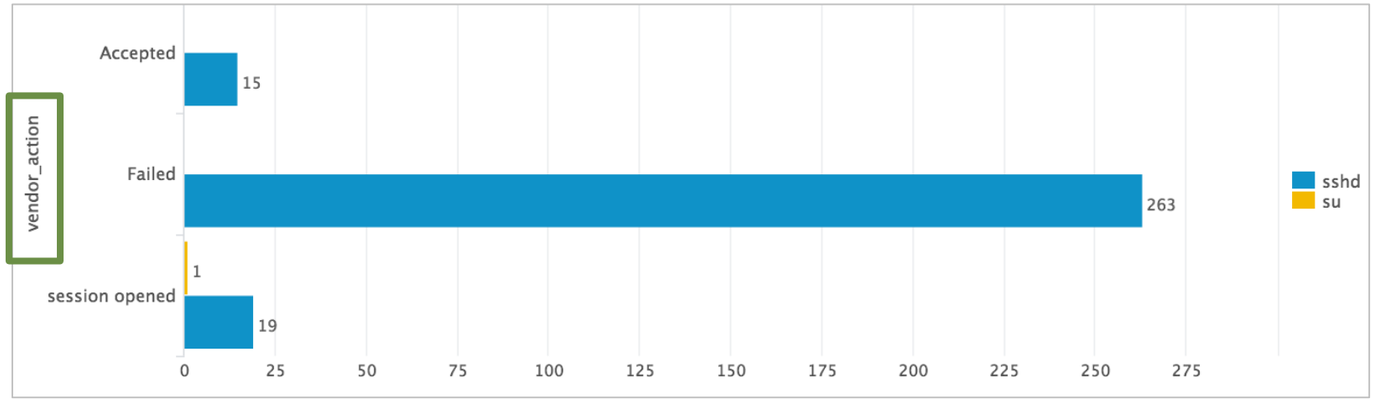
Visualizations – x and y Axes

- For line, area, and column charts, the x axis is horizontal

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



- For bar chart, the x axis is vertical



Generated for (v.ms.besetti@jpmchase.com) (C) Splunk Inc, not for distribution

Copyright © 2018 Splunk, Inc. All rights reserved | 14 June 2018

`chart` Command

- `chart` command can display any series of data that you want to plot
- 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 (v.ms.besetti@jpmchase.com) (C) Splunk Inc, not for distribution

Copyright © 2018 Splunk, Inc. All rights reserved | 14 June 2018

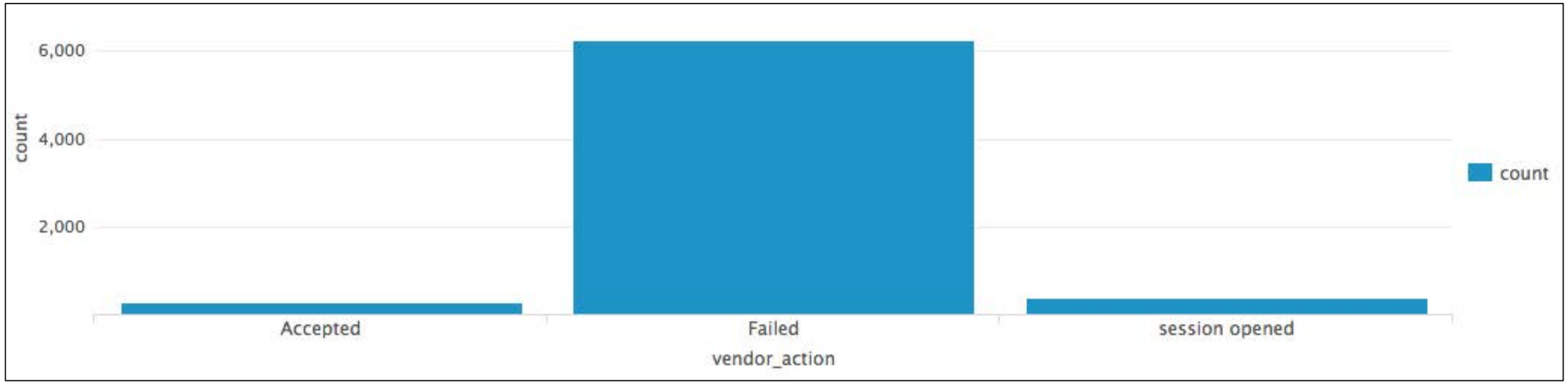
chart Command – over *field*

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 (v.ms.besetti@jpmchase.com) (C) Splunk Inc, not for distribution

Copyright © 2018 Splunk, Inc. All rights reserved | 14 June 2018

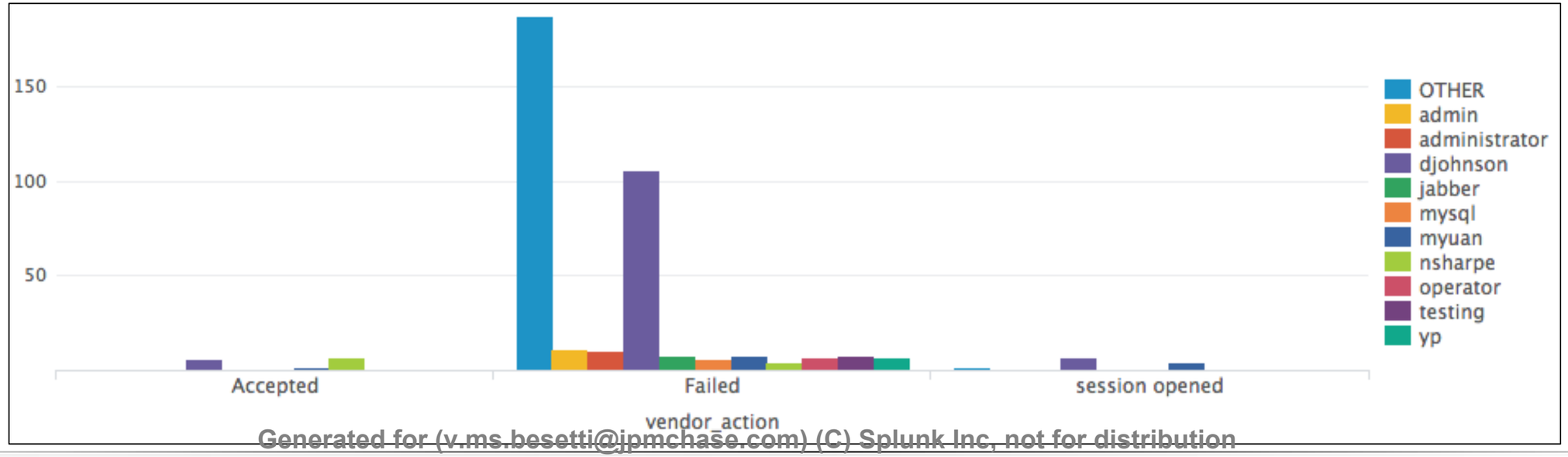
chart Command – over *field* by *field*

- 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
```

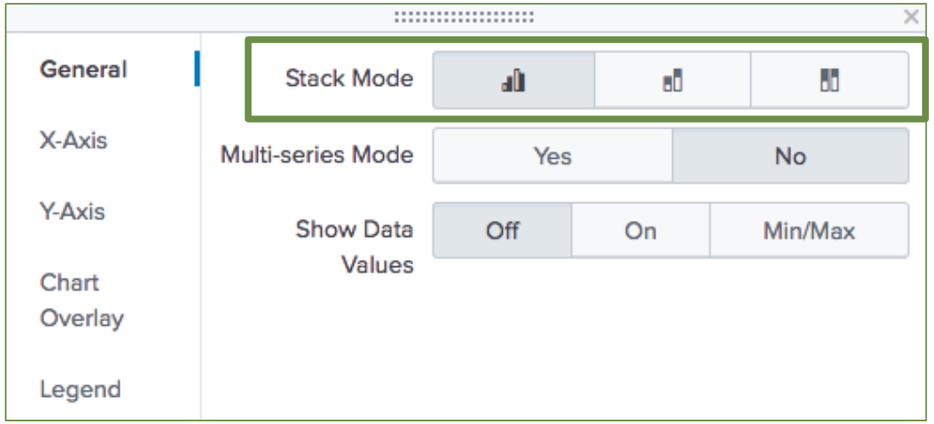


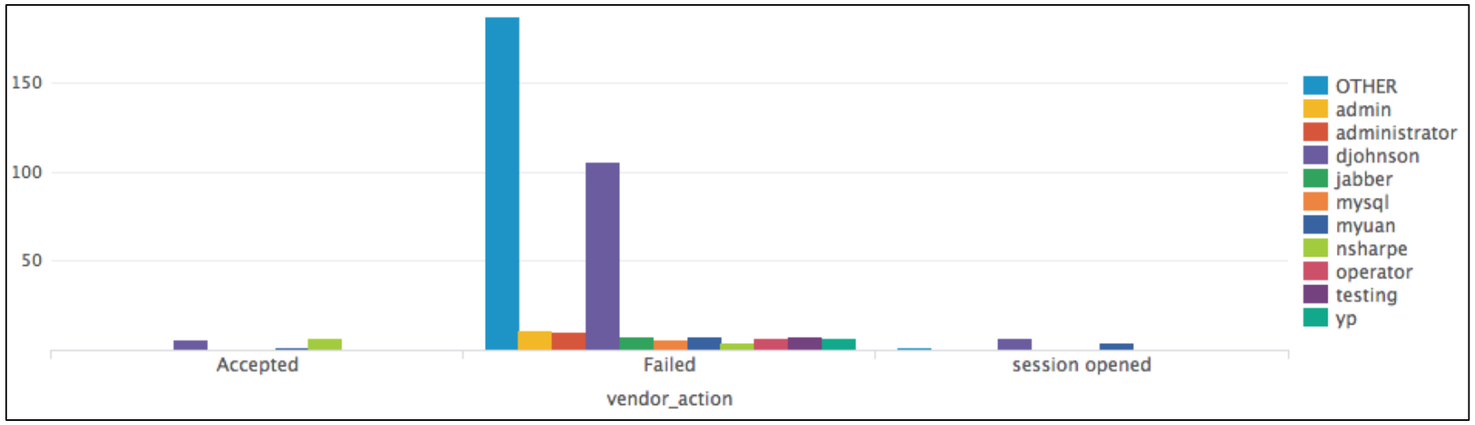
Generated for (v.ms.besetti@jpmchase.com) (C) Splunk Inc, not for distribution

Copyright © 2018 Splunk, Inc. All rights reserved | 14 June 2018

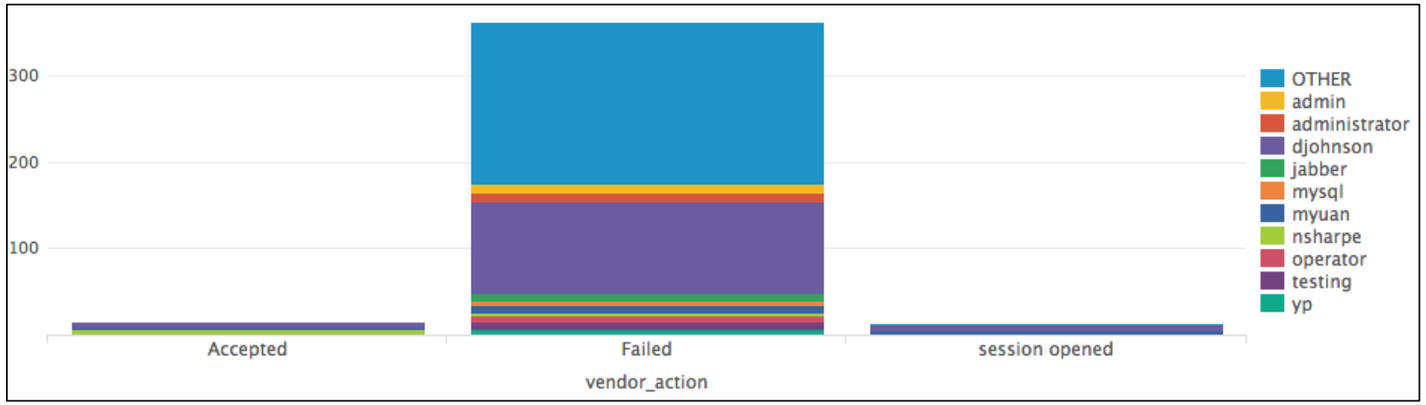
Stack Mode

Stack Mode **OFF**





Stack Mode **ON**



Generated for (v.ms.besetti@jpmchase.com) (C) Splunk Inc, not for distribution

Copyright © 2018 Splunk, Inc. All rights reserved | 14 June 2018

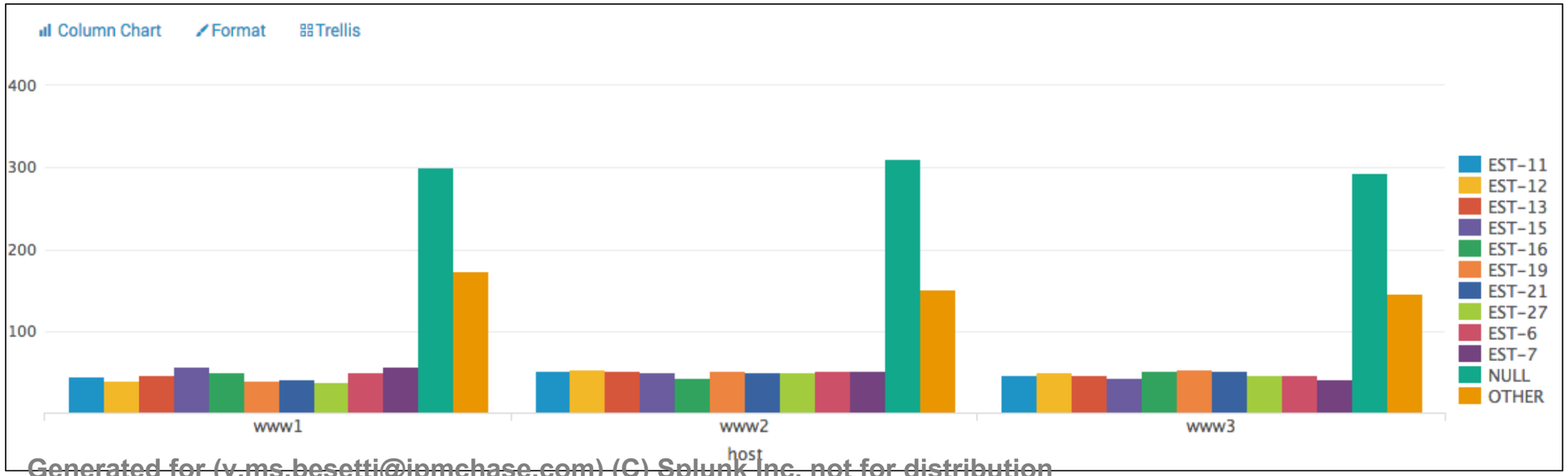
Including NULL and OTHER Values

- `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
```

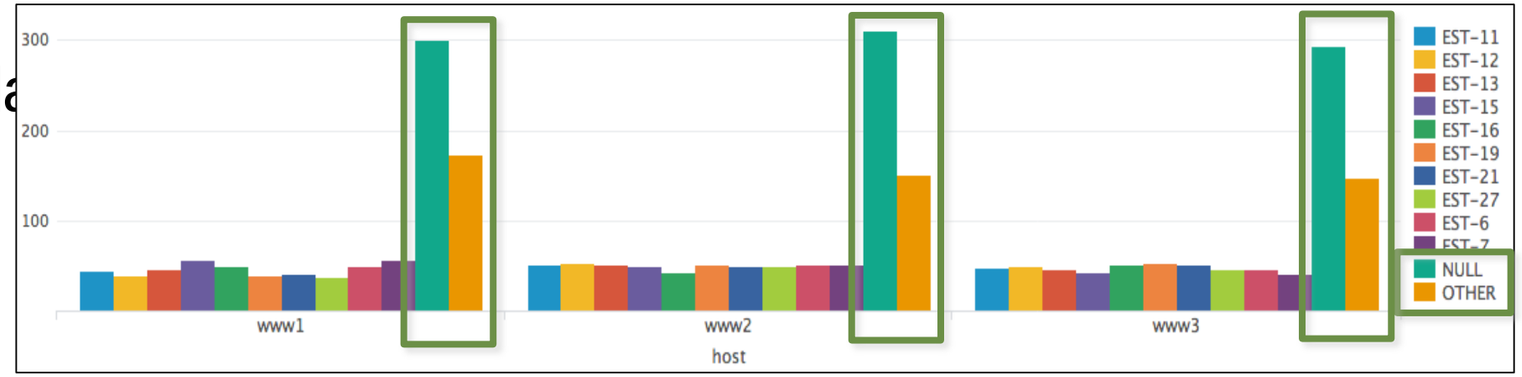


Generated for (v.ms.besetti@jpmchase.com) (C) Splunk Inc, not for distribution

Copyright © 2018 Splunk, Inc. All rights reserved | 14 June 2018

Omitting NULL and OTHER Values

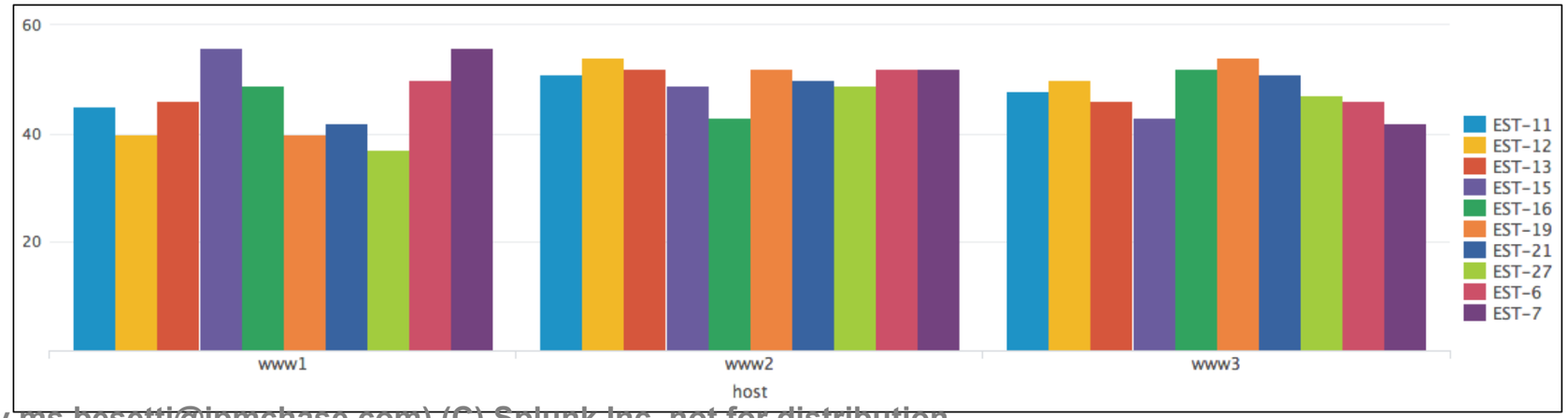
- To remove empty (NULL) and OTHER field values from the displa use these options:
  - useother=f
  - usenull=f




```
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 (v.ms.besetti@jpmchase.com) (C) Splunk Inc, not for distribution

Copyright © 2018 Splunk, Inc. All rights reserved | 14 June 2018

Limiting the Number of Values

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

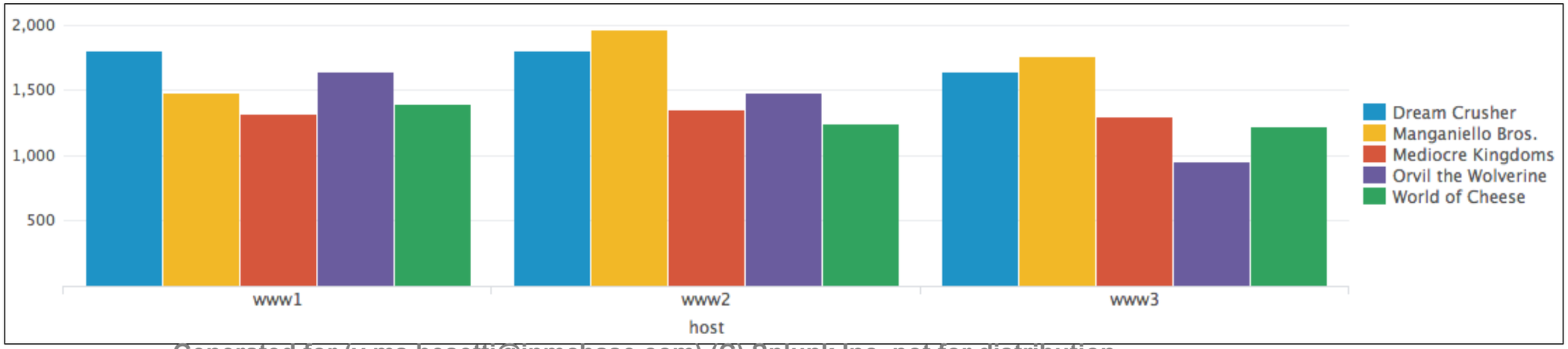
Scenario

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

Note

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

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



Generated for (v.ms.besetti@jpmchase.com) (C) Splunk Inc, not for distribution

Copyright © 2018 Splunk, Inc. All rights reserved | 14 June 2018

`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

Generated for (v.ms.besetti@jpmchase.com) (C) Splunk Inc, not for distribution

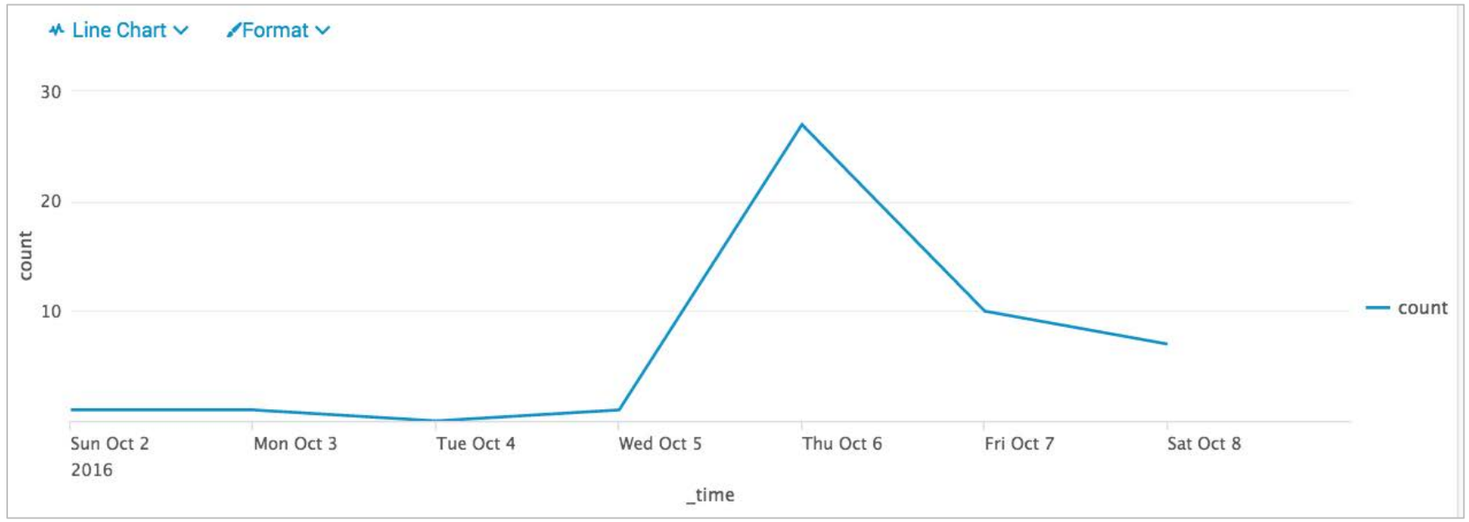
Copyright © 2018 Splunk, Inc. All rights reserved | 14 June 2018

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

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

Generated for (v.ms.besetti@jpmchase.com) (C) Splunk Inc, not for distribution

Copyright © 2018 Splunk, Inc. All rights reserved | 14 June 2018

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

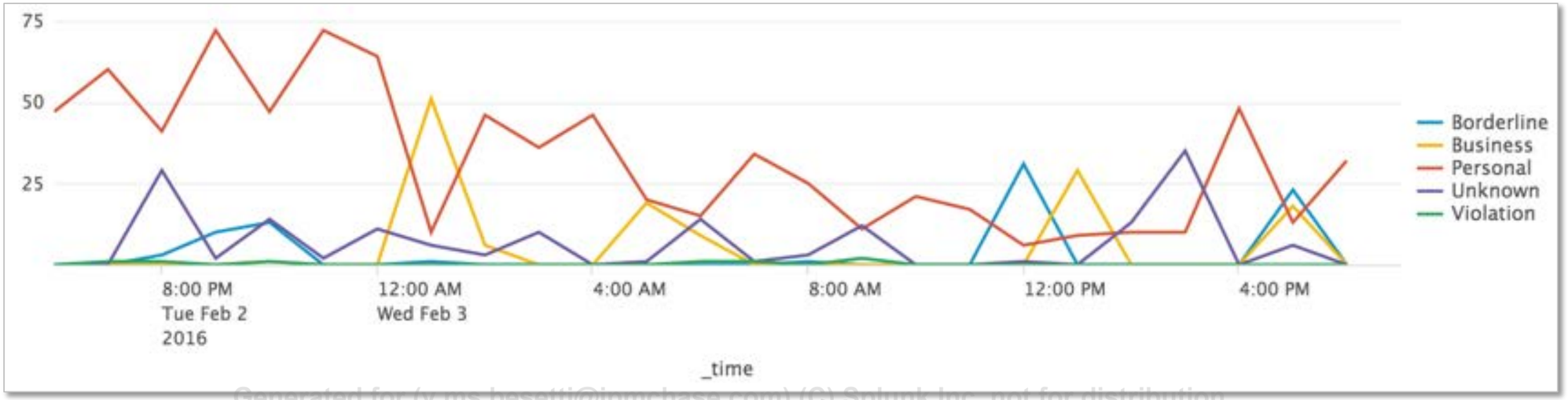
Scenario

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

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

Note

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



Generated for (v.ms.besetti@jpmchase.com) (C) Splunk Inc, not for distribution

Copyright © 2018 Splunk, Inc. All rights reserved | 14 June 2018

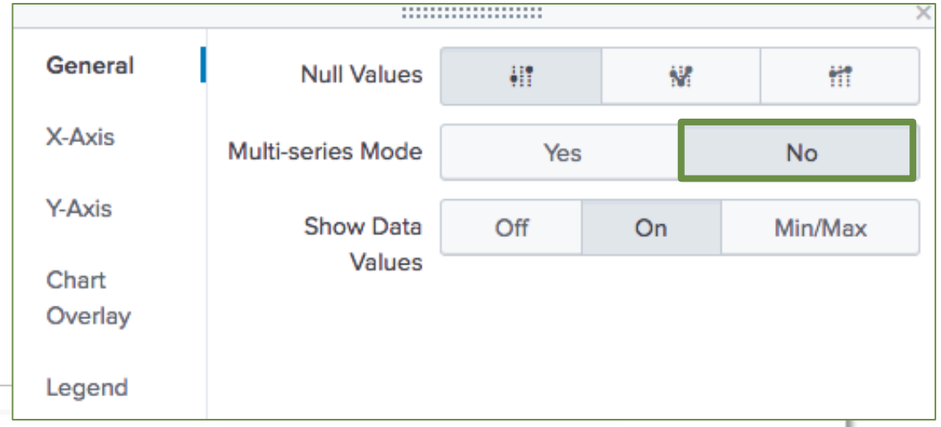
`timechart` Command – Multi-series: No

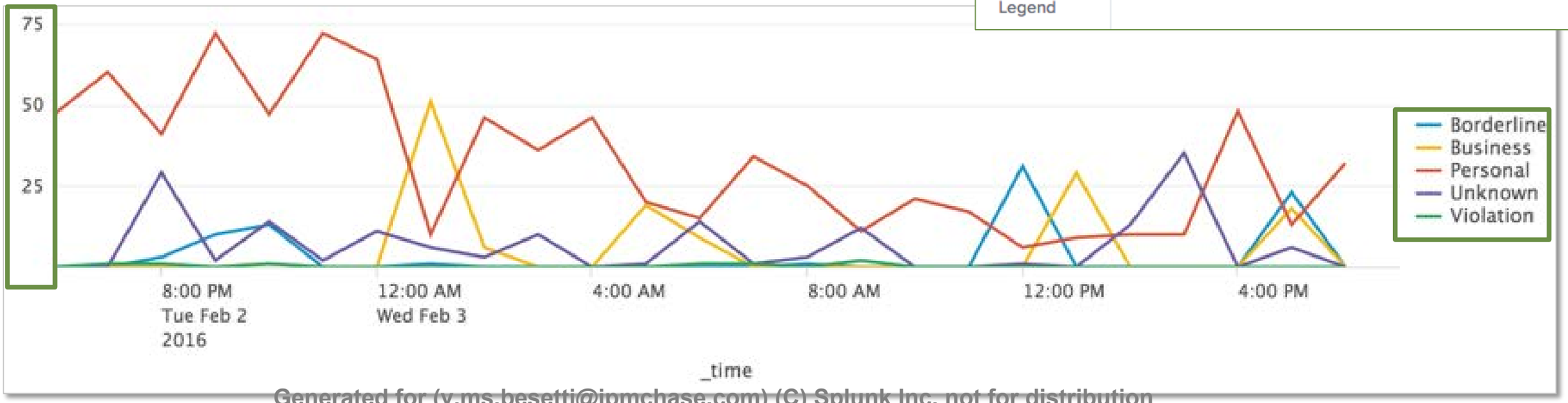
When the multi-series mode is set to **No**, all fields share the y-axis

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

Scenario

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



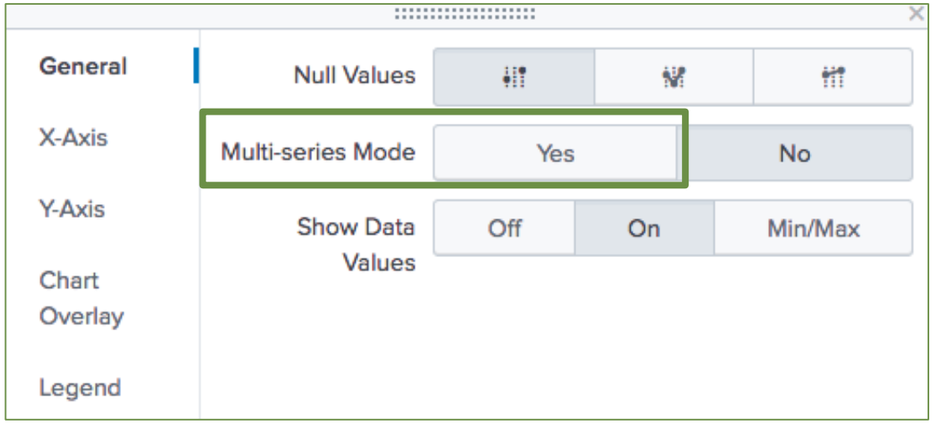


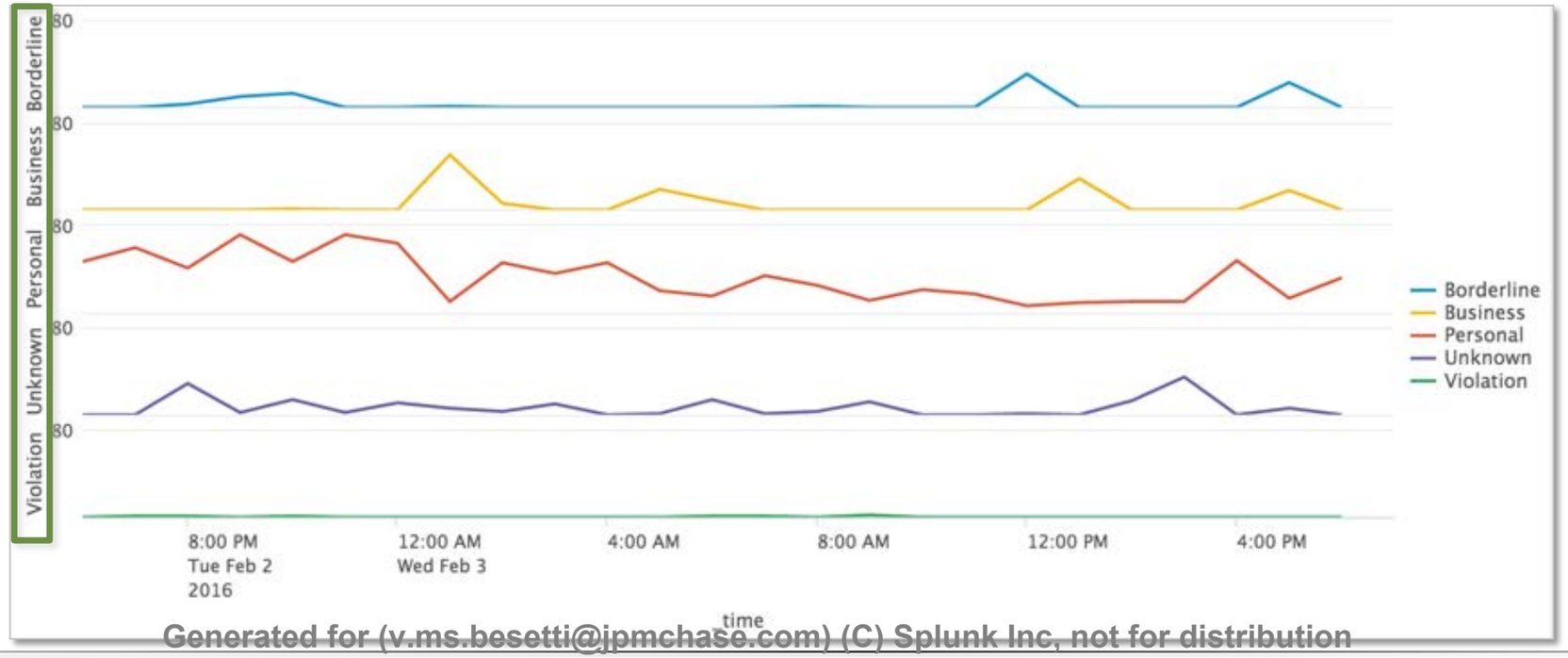
Generated for (v.ms.besetti@jpmchase.com) (C) Splunk Inc, not for distribution

Copyright © 2018 Splunk, Inc. All rights reserved | 14 June 2018

`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 (v.ms.besetti@jpmchase.com) (C) Splunk Inc, not for distribution

Copyright © 2018 Splunk, Inc. All rights reserved | 14 June 2018

`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
```

✓ 1,743 events (1/16/18 12:07:00.000 PM to 1/16/18 4:07:21.0) No Event Sampling

Events Patterns Statistics (17) Visualization

20 Per Page Format Preview

_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

Generated for (v.ms.besetti@jpmchase.com) (C) Splunk Inc, not for distribution

Copyright © 2018 Splunk, Inc. All rights reserved | 14 June 2018

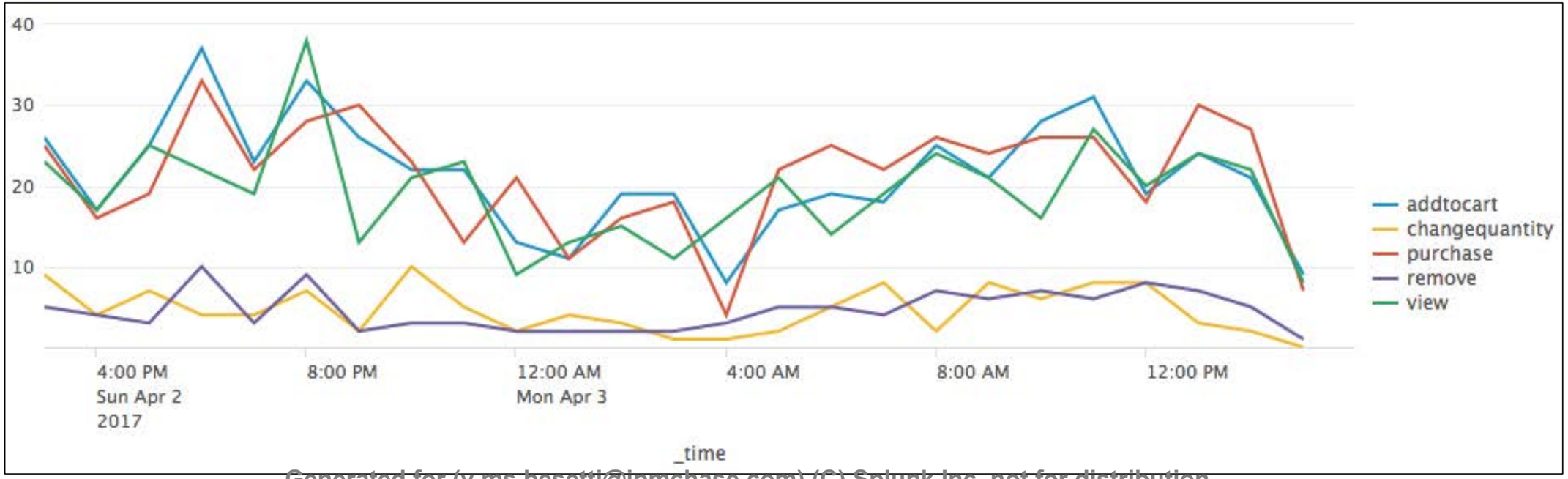
timechart Command – Statistical Functions

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
```



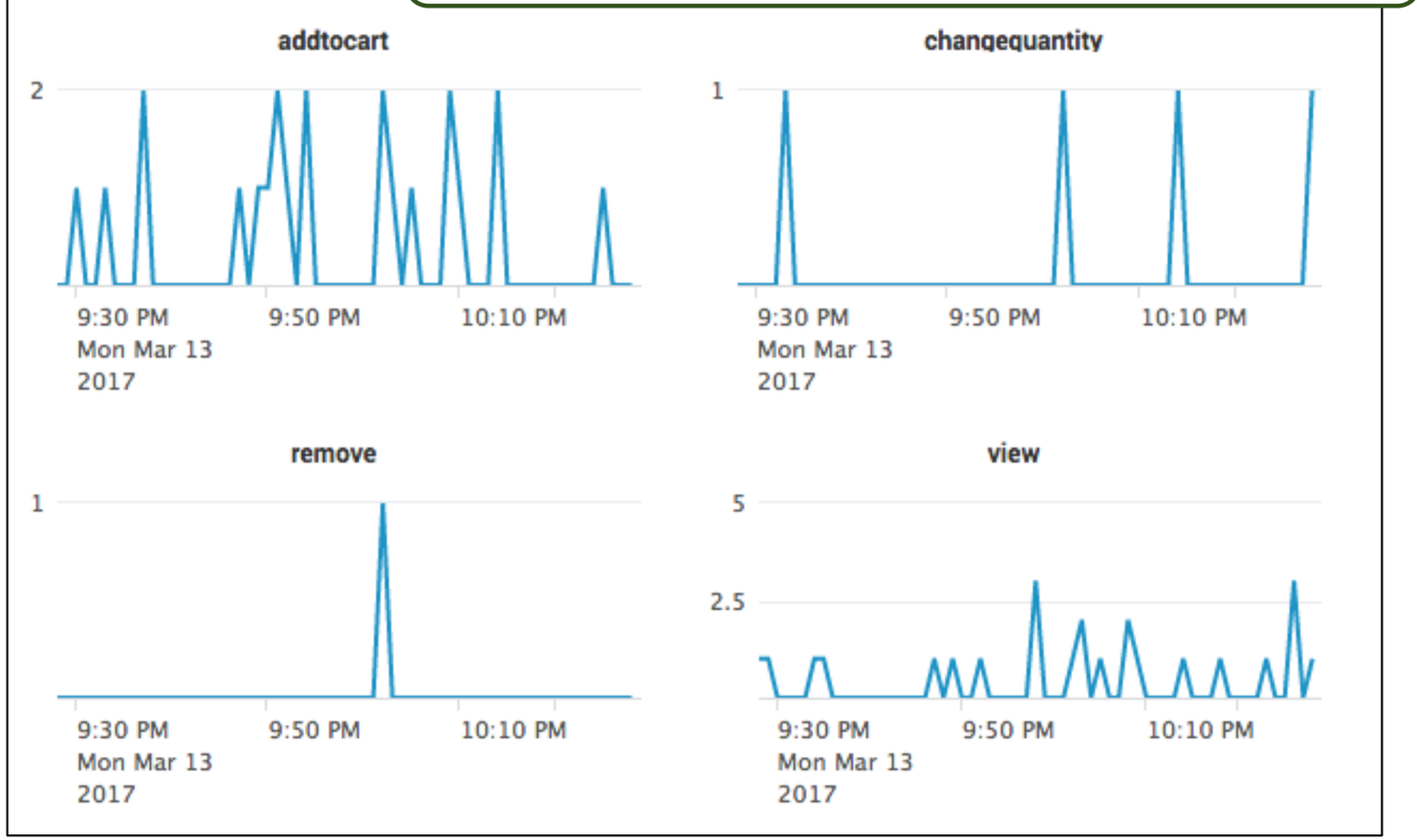
Generated for (v.ms.besetti@jpmchase.com) (C) Splunk Inc, not for distribution

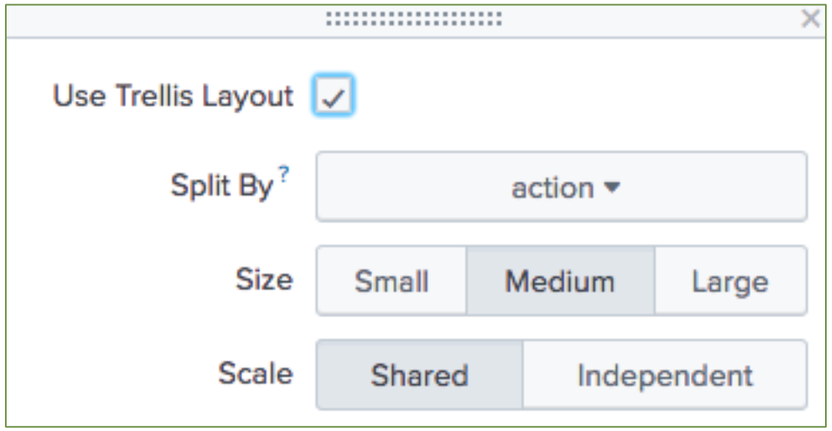
Copyright © 2018 Splunk, Inc. All rights reserved | 14 June 2018

Trellis Layout

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

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

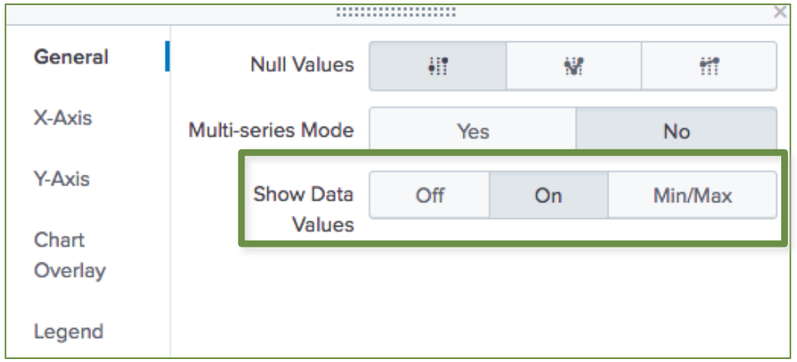




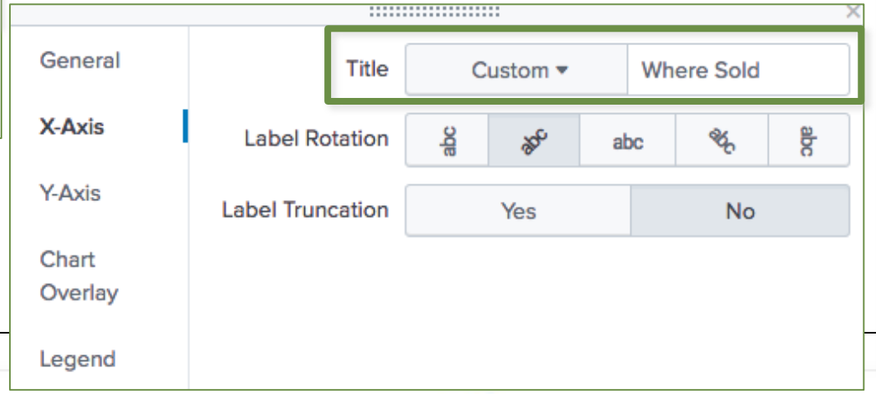
Generated for (v.ms.besetti@jpmchase.com) (C) Splunk Inc, not for distribution

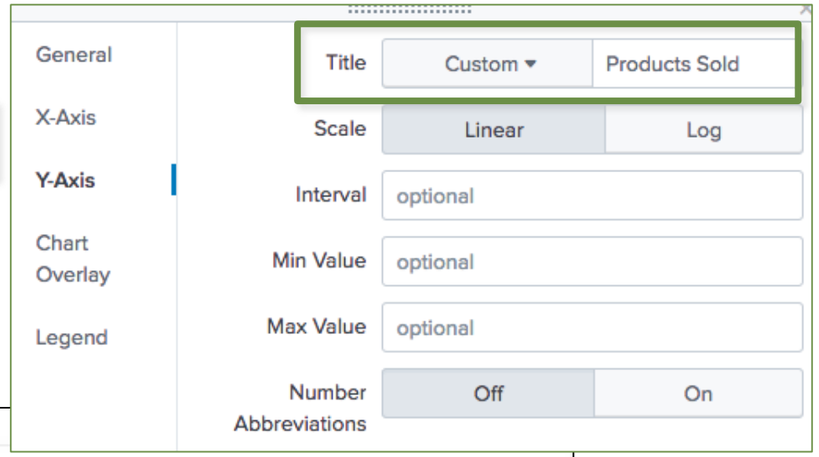
Copyright © 2018 Splunk, Inc. All rights reserved

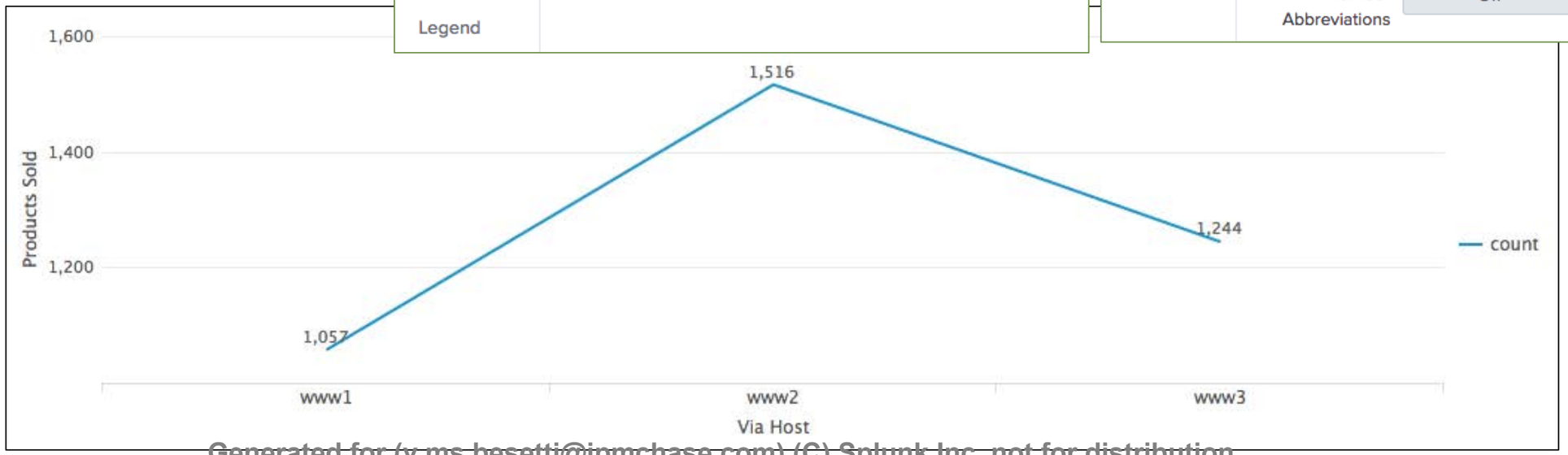
Visualization Formatting




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







Generated for (v.ms.besetti@jpmchase.com) (C) Splunk Inc, not for distribution

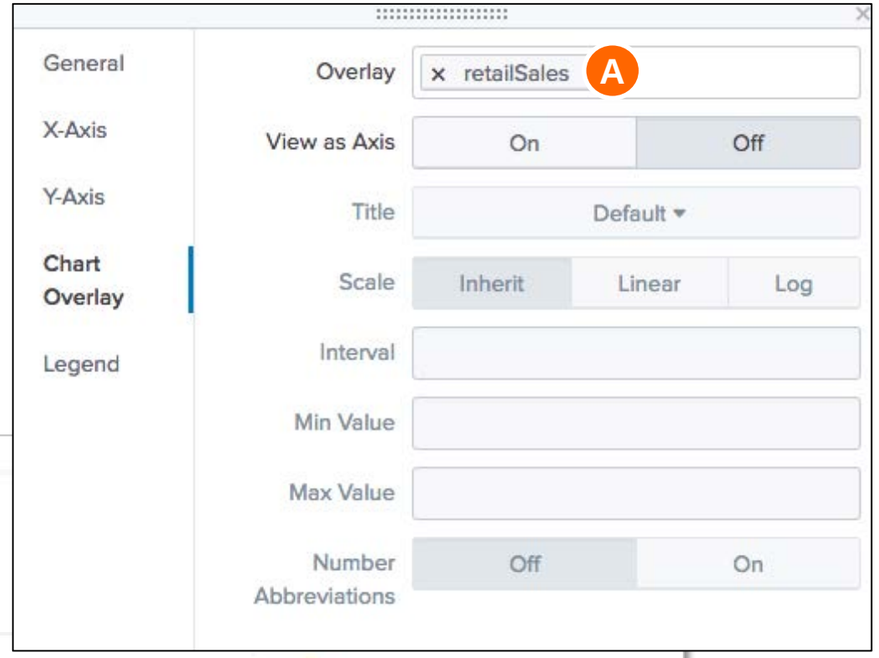
Copyright © 2018 Splunk, Inc. All rights reserved

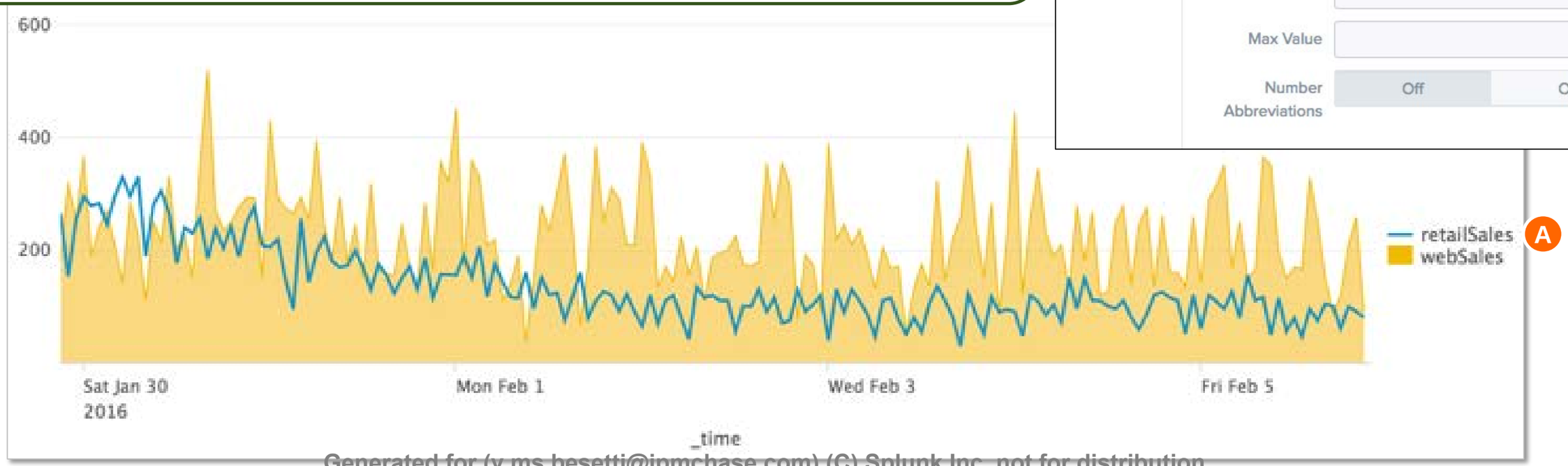
Formatting – Chart Overlay

Scenario

Graphically compare the total price of online sales to sales made in physical stores for the last 7 days, on an hour-by-hour basis.

```
(index=web sourcetype=access_combined action=purchase status<400) OR
(index=sales sourcetype=vendor_sales)
| timechart span=1h sum(price) by sourcetype
| rename access_combined as webSales, vendor_sales as retailSales
```





Generated for (v.ms.besetti@jpmchase.com) (C) Splunk Inc, not for distribution

Copyright © 2018 Splunk, Inc. All rights reserved | 14 June 2018

Transforming Command Summary

Feature	`stats`	`chart`	`timechart`
Multi-level breakdown [`by` clause]	Many	2	1
Limit # series shown	NA	`limit=`*n* *Default=10*	`limit=`*n* *Default=10*
Filter `other` series	NA	`useother=f`	`useother=f`
Filter null values	NA	`usenull=f`	`usenull=f`
Set time value on x axis	NA	NA	`span`

Generated for (v.ms.besetti@jpmchase.com) (C) Splunk Inc, not for distribution

Copyright © 2018 Splunk, Inc. All rights reserved | 14 June 2018

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 (v.ms.besetti@jpmchase.com) (C) Splunk Inc, not for distribution

Copyright © 2018 Splunk, Inc. All rights reserved | 14 June 2018

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 (v.ms.besetti@jpmchase.com) (C) Splunk Inc, not for distribution

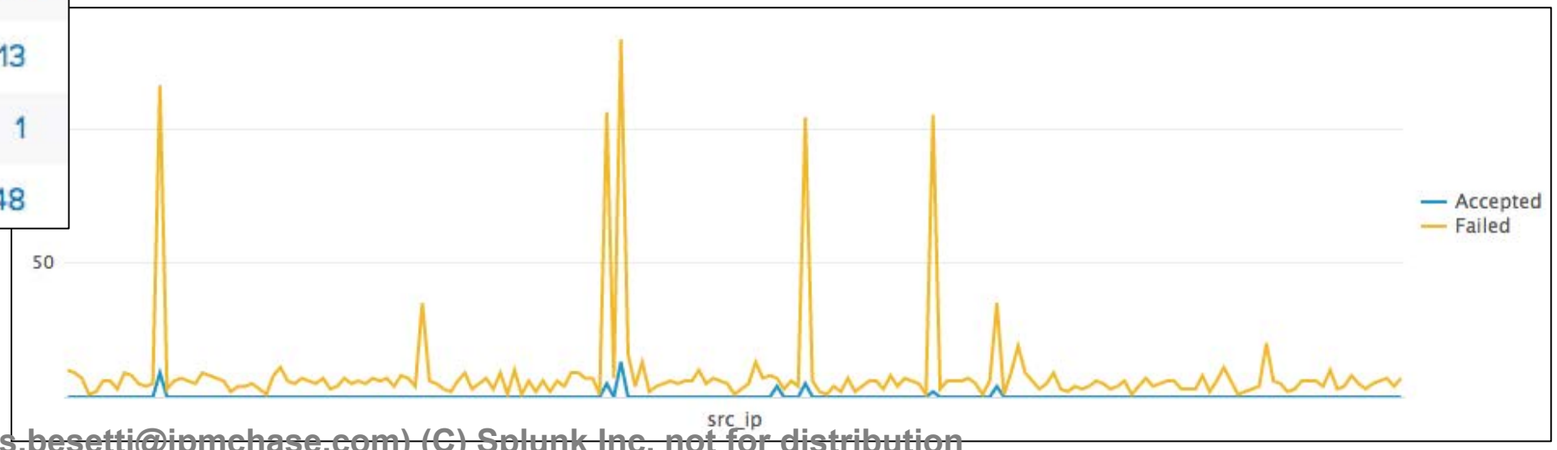
Copyright © 2018 Splunk, Inc. All rights reserved | 14 June 2018

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
```

src_ip	Accepted	Failed
15.170.91.198	5	109
16.201.248.137	3	113
17.186.26.16	0	1
17.186.26.167	41	148

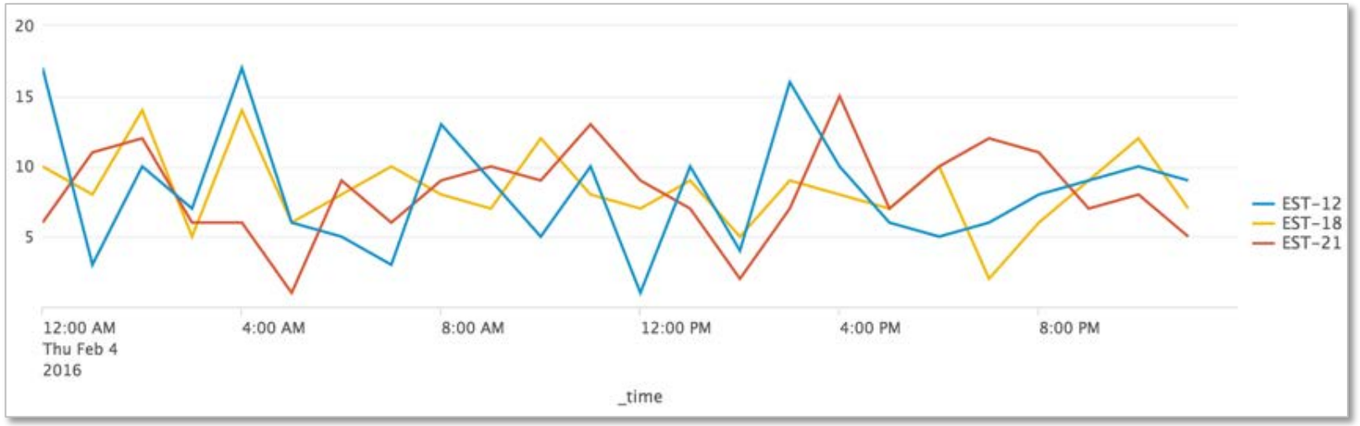


Generated for (v.ms.besetti@jpmchase.com) (C) Splunk Inc, not for distribution

Copyright © 2018 Splunk, Inc. All rights reserved | 14 June 2018

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
```

_time	EST-13	EST-15	EST-7	NULL
2018-01-16 12:00	7	5	4	28
2018-01-16 13:00	15	6	7	53
2018-01-16 14:00	8	20	15	42
2018-01-16 15:00	3	8	9	38
2018-01-16 16:00	5	8	4	34

Generated for (v.ms.besetti@jpmchase.com) (C) Splunk Inc, not for distribution

Copyright © 2018 Splunk, Inc. All rights reserved | 14 June 2018

Module 4: Advanced Visualizations

Generated for (v.ms.besetti@jpmchase.com) (C) Splunk Inc, not for distribution

Copyright © 2018 Splunk, Inc. All rights reserved

Module Objectives

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

Generated for (v.ms.besetti@jpmchase.com) (C) Splunk Inc, not for distribution

Copyright © 2018 Splunk, Inc. All rights reserved | 14 June 2018

`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

Generated for (v.ms.besetti@jpmchase.com) (C) Splunk Inc, not for distribution

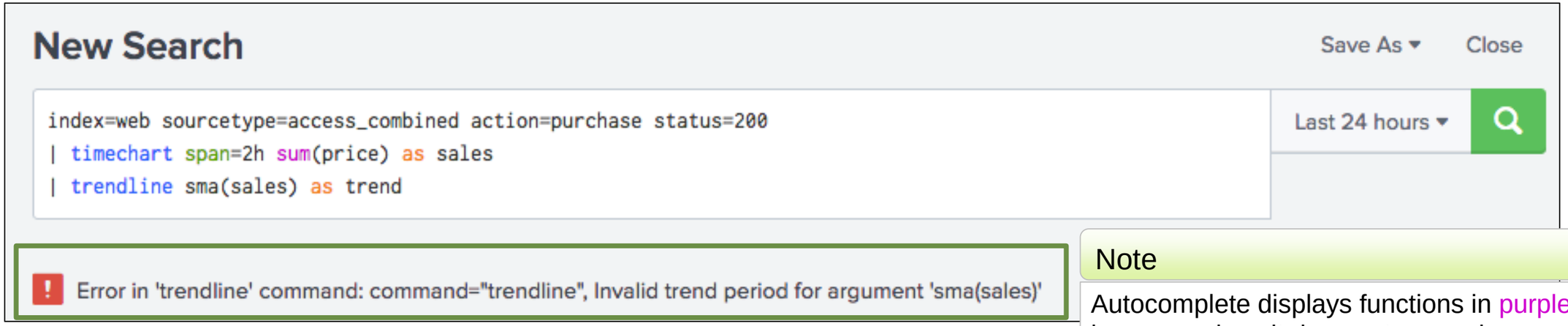
Copyright © 2018 Splunk, Inc. All rights reserved | 14 June 2018

`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.



Note

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

Generated for (v.ms.besetti@jpmchase.com) (C) Splunk Inc, not for distribution

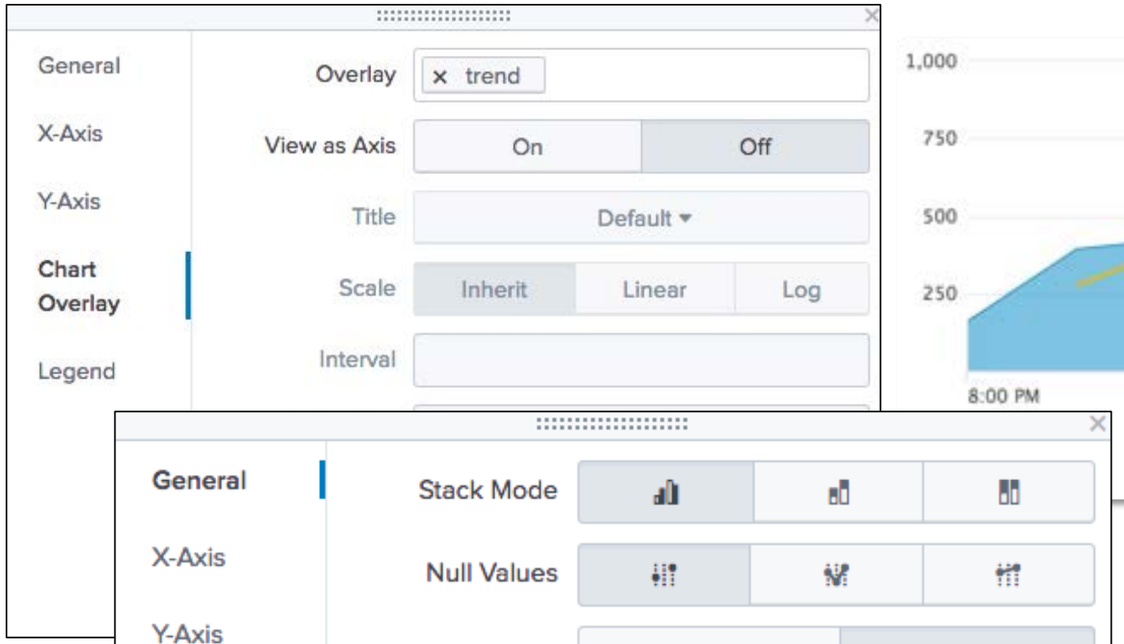
Copyright © 2018 Splunk, Inc. All rights reserved | 14 June 2018

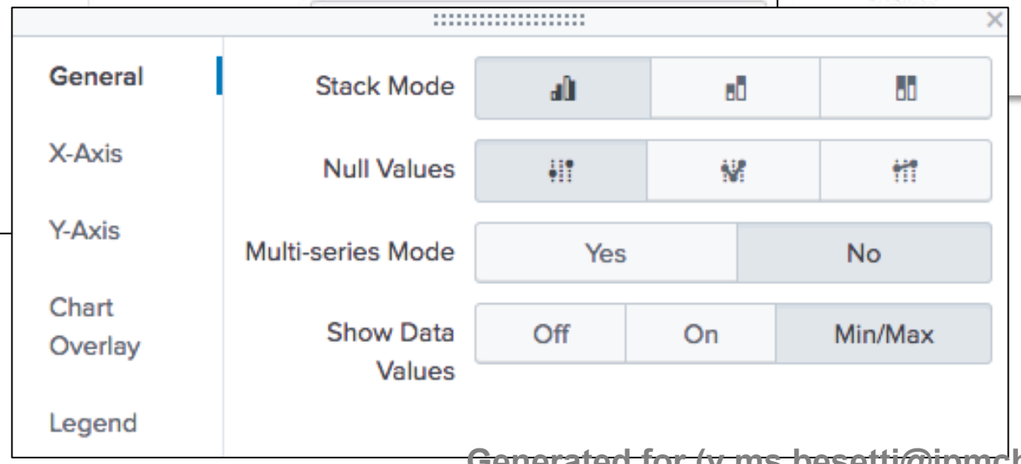
trendline Command – Example

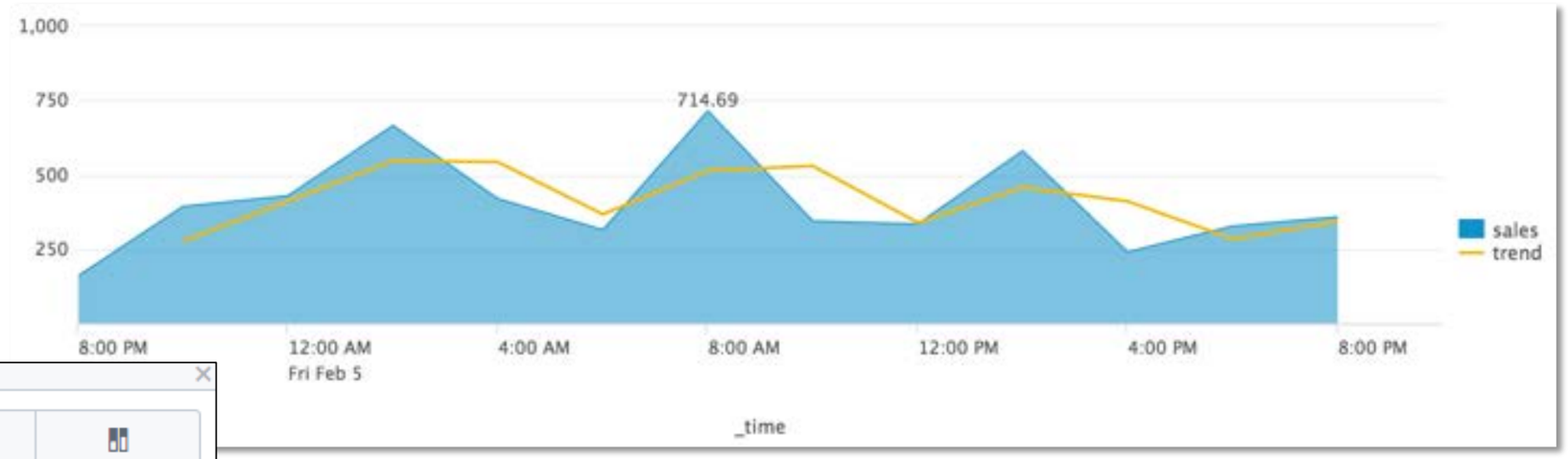
Scenario

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
```







Generated for (v.ms.besetti@jpmchase.com) (C) Splunk Inc, not for distribution

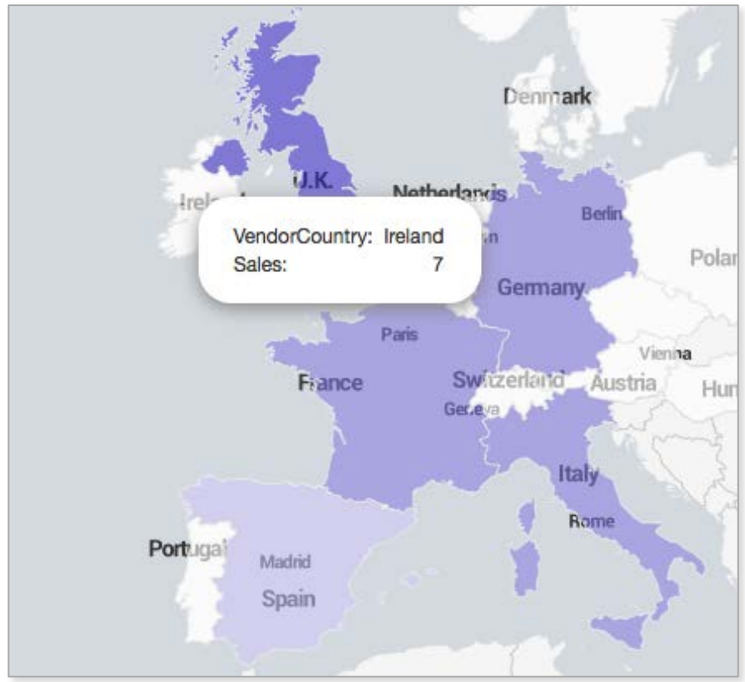
Copyright © 2018 Splunk, Inc. All rights reserved | 14 June 2018

Viewing Results as a Map

There are two map types

Cluster Map

Choropleth Map



Generated for (v.ms.besetti@jpmchase.com) (C) Splunk Inc, not for distribution

Copyright © 2018 Splunk, Inc. All rights reserved | 14 June 2018

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

Generated for (v.ms.besetti@jpmchase.com) (C) Splunk Inc, not for distribution

Copyright © 2018 Splunk, Inc. All rights reserved | 14 June 2018

`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:
  - On a global level, use the `globallimit` argument
  - On a local level, depending on where your focus is (i.e., where you've zoomed in), use the `locallimit` argument

Generated for (v.ms.besetti@jpmchase.com) (C) Splunk Inc, not for distribution

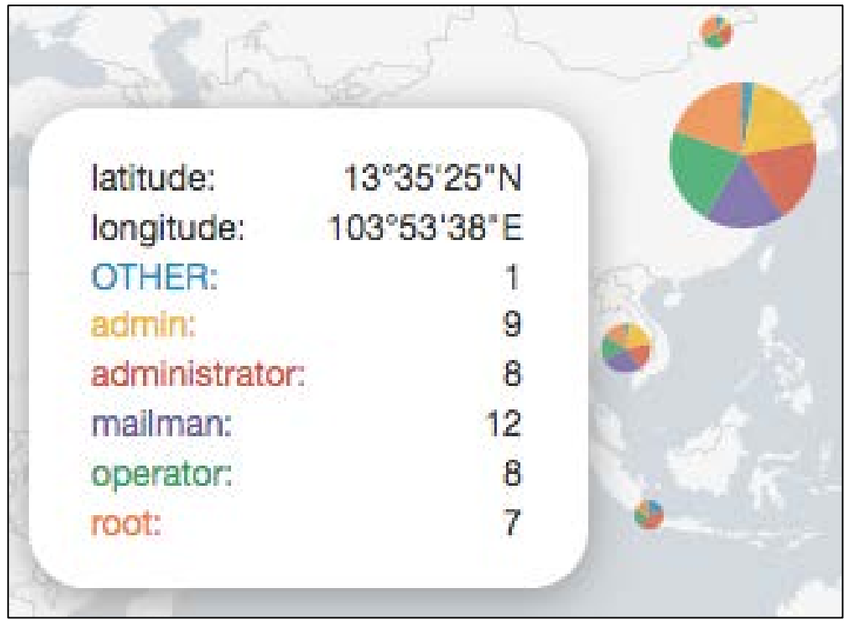
Copyright © 2018 Splunk, Inc. All rights reserved | 14 June 2018

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 (v.ms.besetti@jpmchase.com) (C) Splunk Inc, not for distribution

Copyright © 2018 Splunk, Inc. All rights reserved | 14 June 2018

Choropleth Map

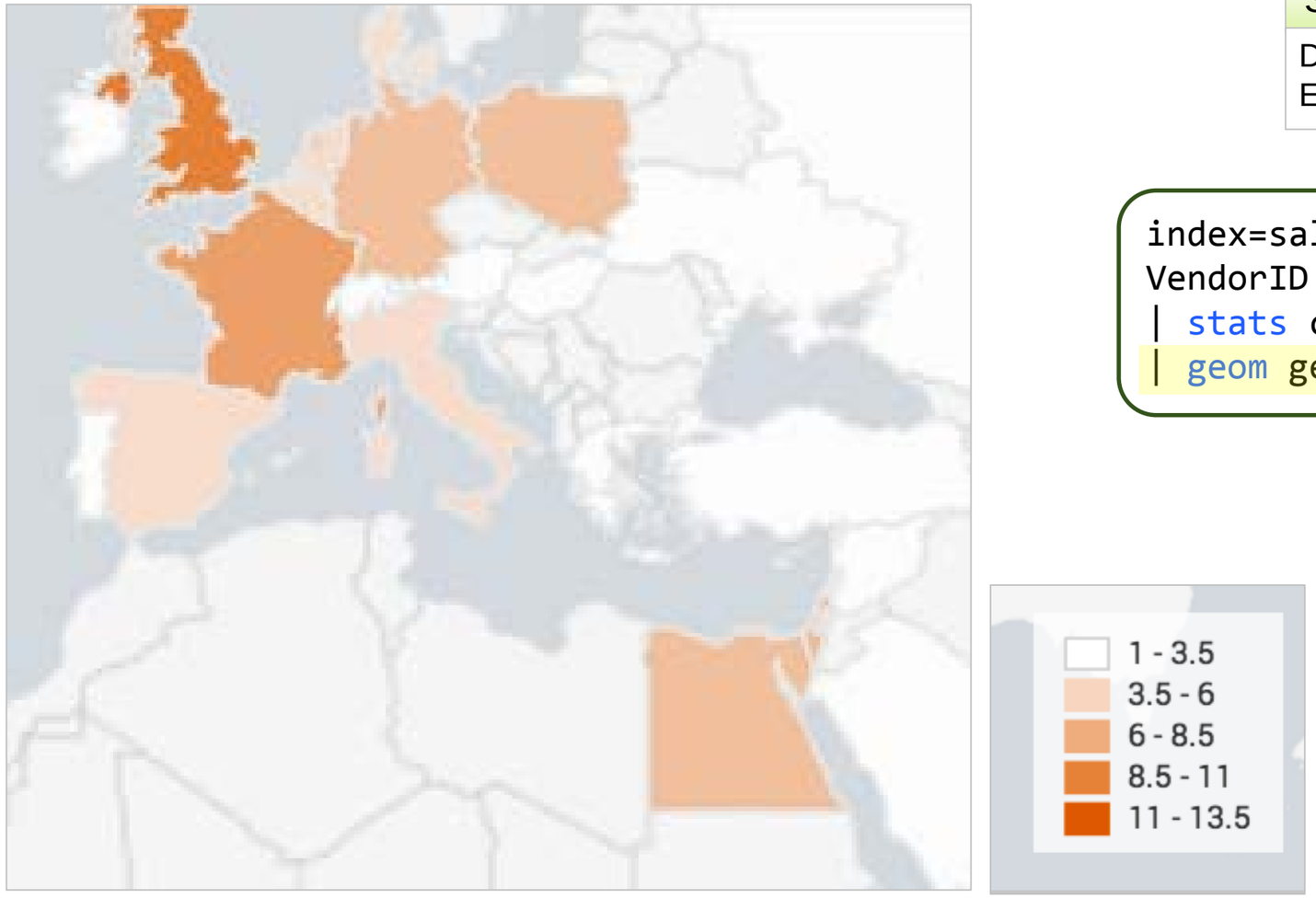
- Uses shading to show relative metrics, such as sales, network intruders, etc. for predefined geographic regions
- To define regional boundaries, you must have either a:
  - KML (Keyhole Markup Language) file
  - KMZ (compressed Keyhole Markup Language) file
- Splunk ships with:
  - `geo_us_states`,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

…| `geom [`*`featureCollection`*`] [featureIdField=`*`string`*`]`

Generated for (v.ms.besetti@jpmchase.com) (C) Splunk Inc, not for distribution

Copyright © 2018 Splunk, Inc. All rights reserved | 14 June 2018

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
```

Generated for (v.ms.besetti@jpmchase.com) (C) Splunk Inc, not for distribution

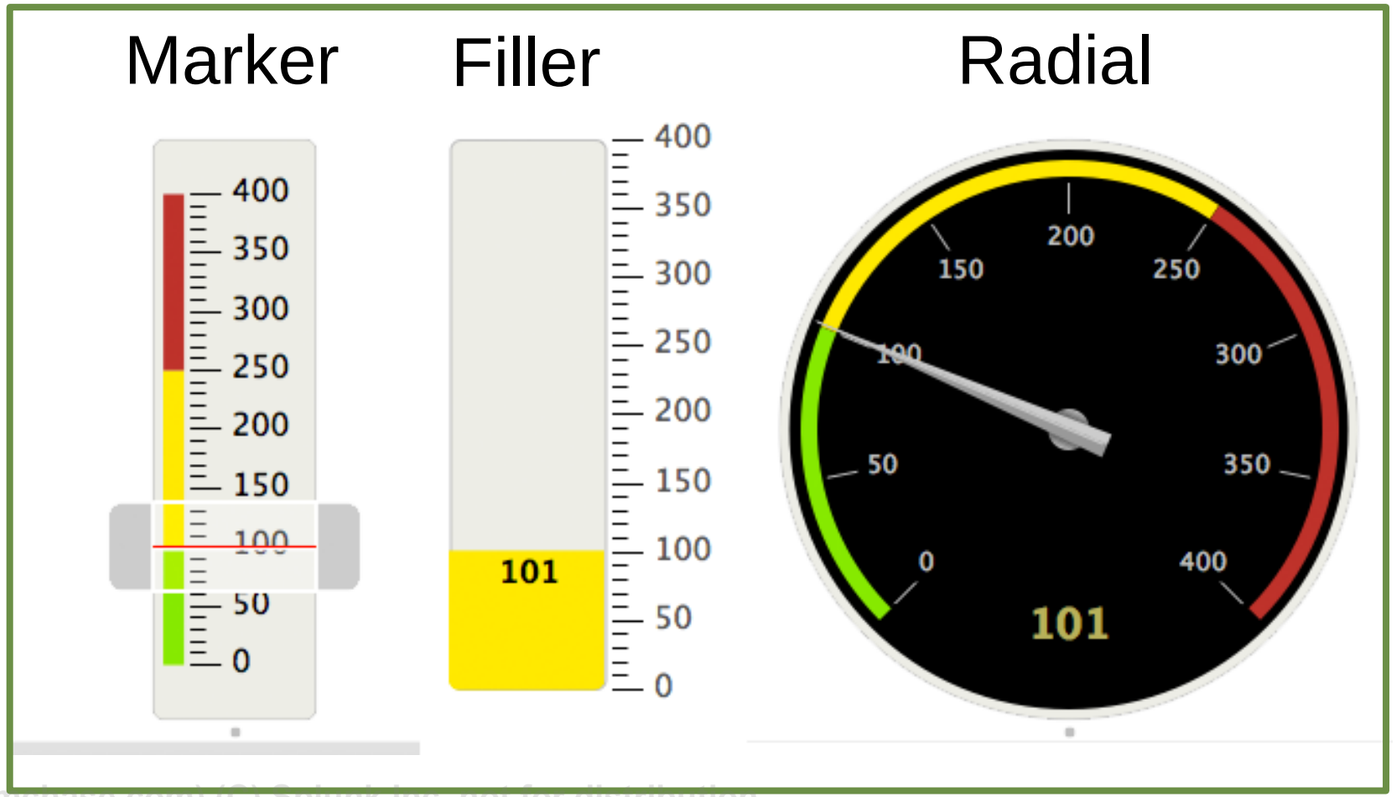
Copyright © 2018 Splunk, Inc. All rights reserved | 14 June 2018

Viewing Results as a Single Value

Single value visualizations provide various formatting options

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

101

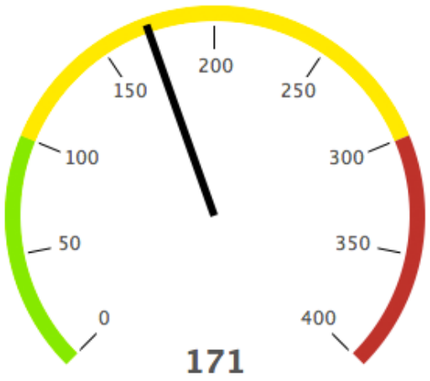


Generated for (v.ms.besetti@jpmchase.com) (C) Splunk Inc, not for distribution

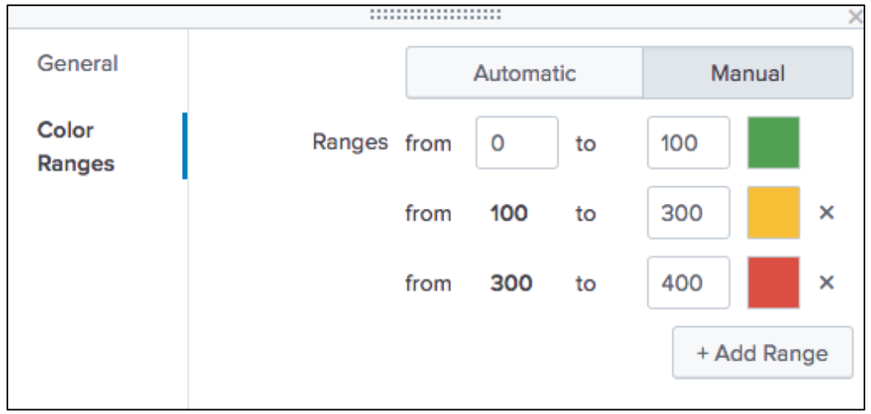
Copyright © 2018 Splunk, Inc. All rights reserved

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 (v.ms.besetti@jpmchase.com) (C) Splunk Inc, not for distribution

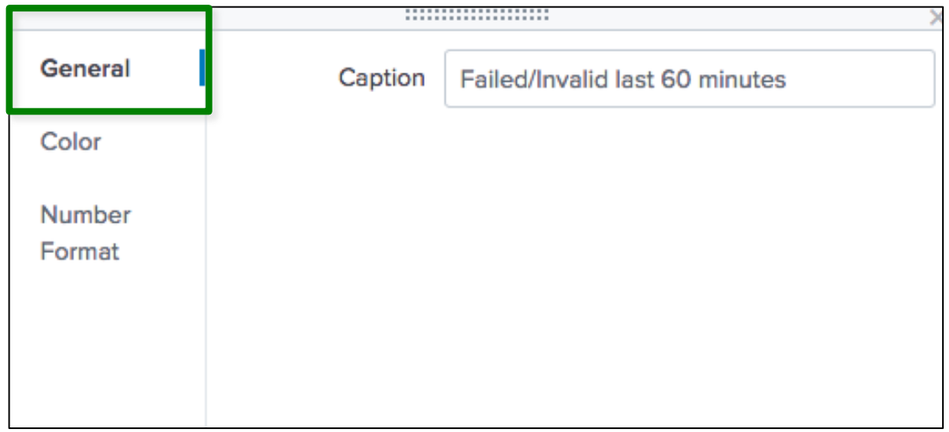
Copyright © 2018 Splunk, Inc. All rights reserved | 14 June 2018

Single Value Visualizations: Formatting (cont.)

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

264

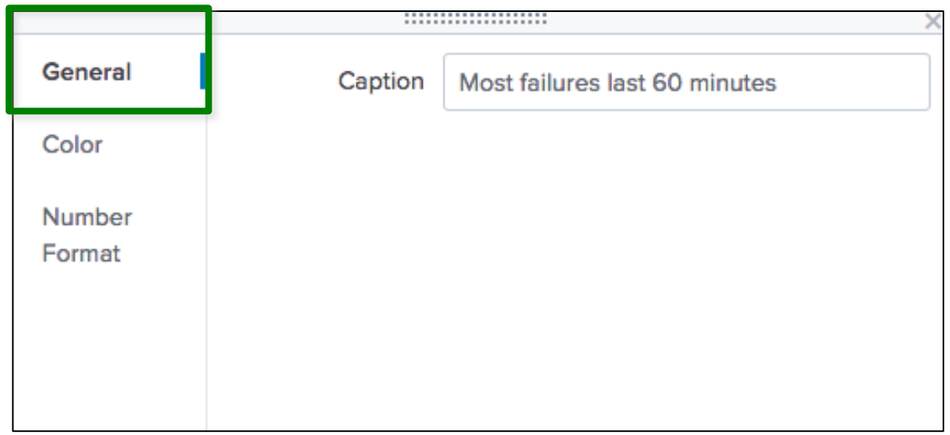
Failed/Invalid last 60 minutes




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

10.1.10.172

Most failures last 60 minutes

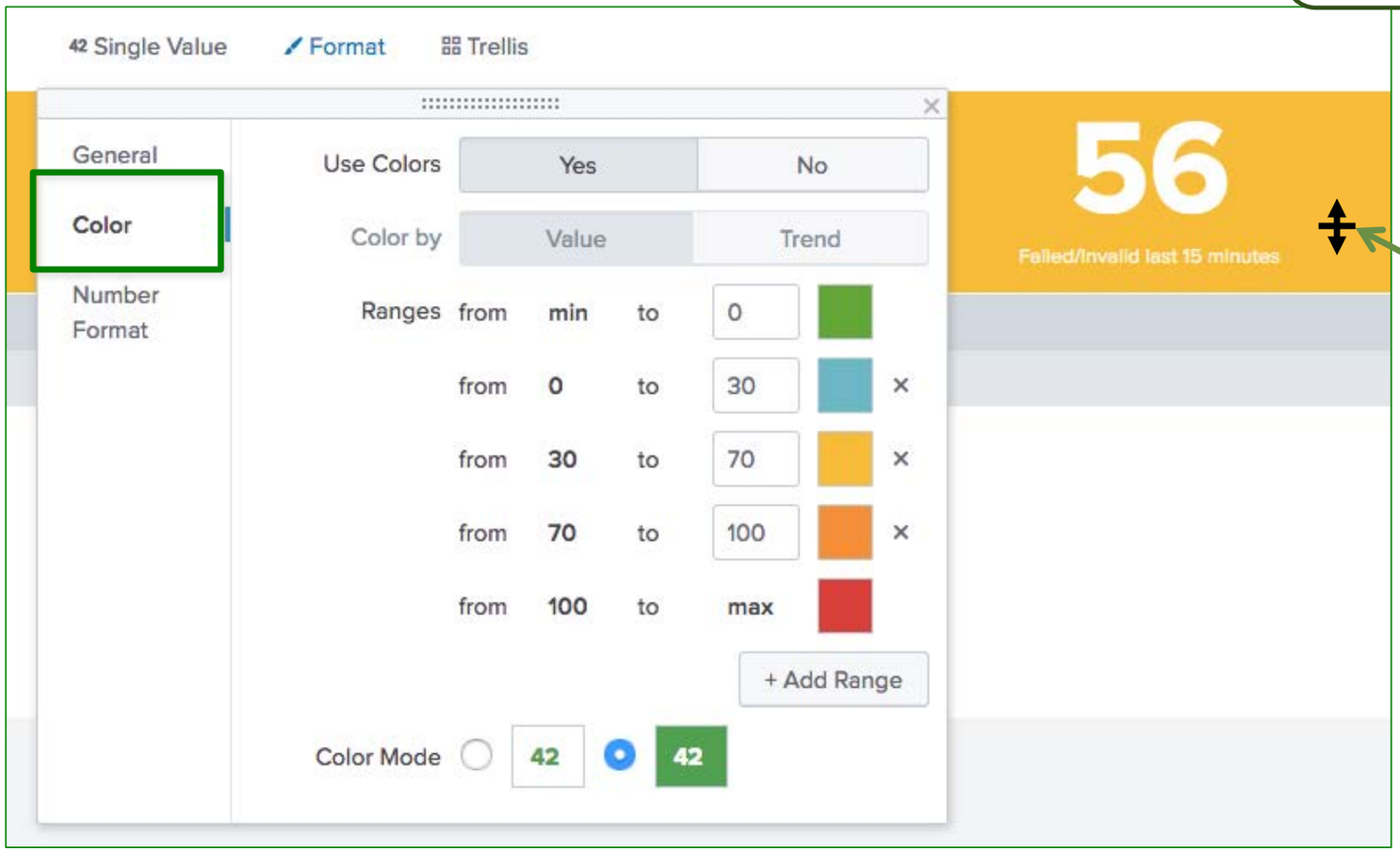


Generated for (v.ms.besetti@jpmchase.com) (C) Splunk Inc, not for distribution

Copyright © 2018 Splunk, Inc. All rights reserved | 14 June 2018

Single Value Visualizations: Formatting (cont.)

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



To resize the font, resize the pane

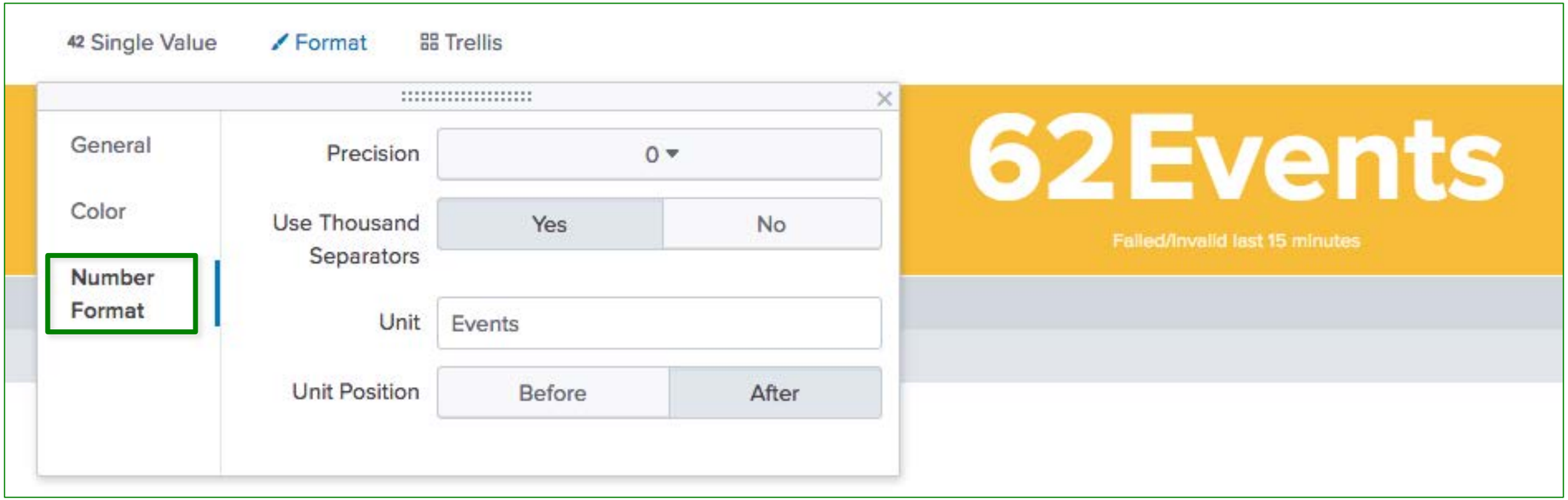
Generated for (v.ms.besetti@jpmchase.com) (C) Splunk Inc, not for distribution

Copyright © 2018 Splunk, Inc. All rights reserved | 14 June 2018

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
```



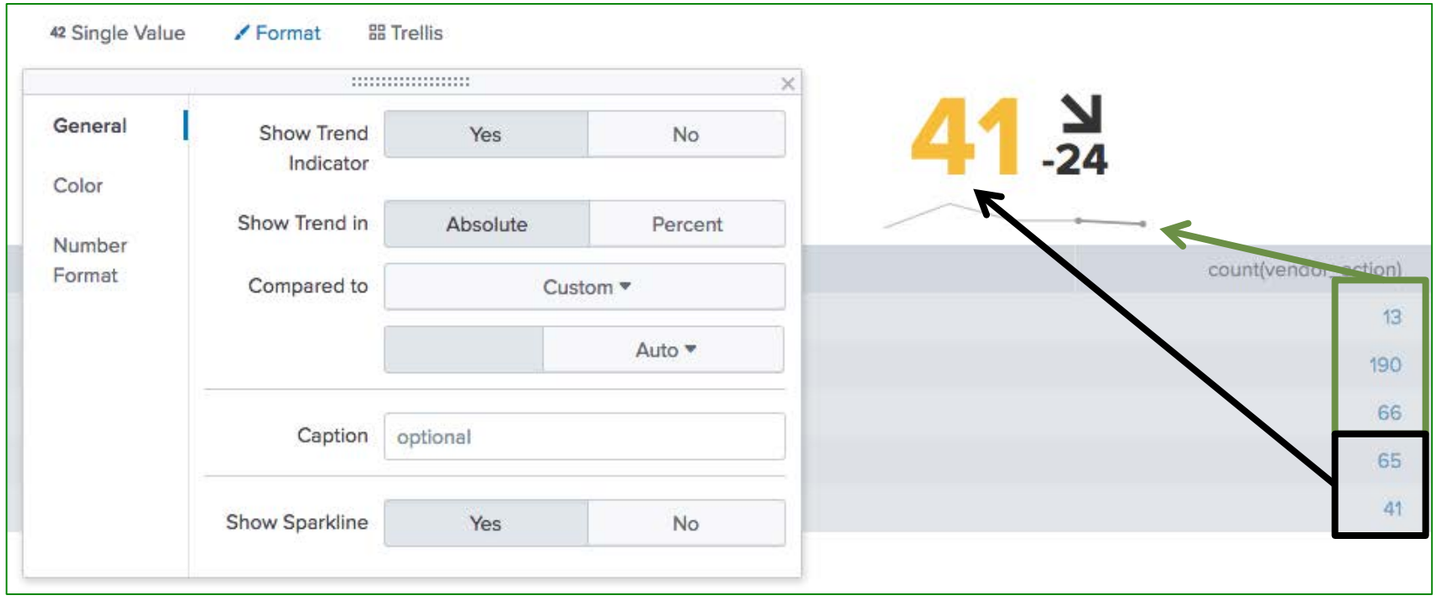
Generated for (v.ms.besetti@jpmchase.com) (C) Splunk Inc, not for distribution

Copyright © 2018 Splunk, Inc. All rights reserved | 14 June 2018

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)
```



Generated for (v.ms.besetti@jpmchase.com) (C) Splunk Inc, not for distribution

Copyright © 2018 Splunk, Inc. All rights reserved | 14 June 2018

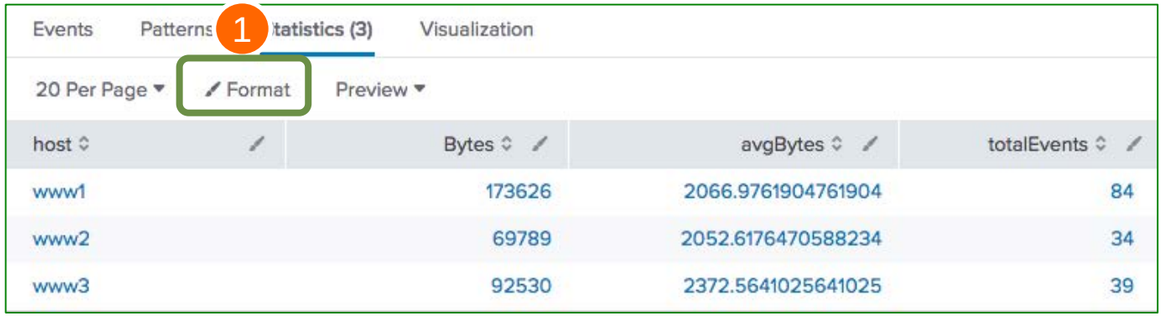
Adding Totals Using Format Options

- 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

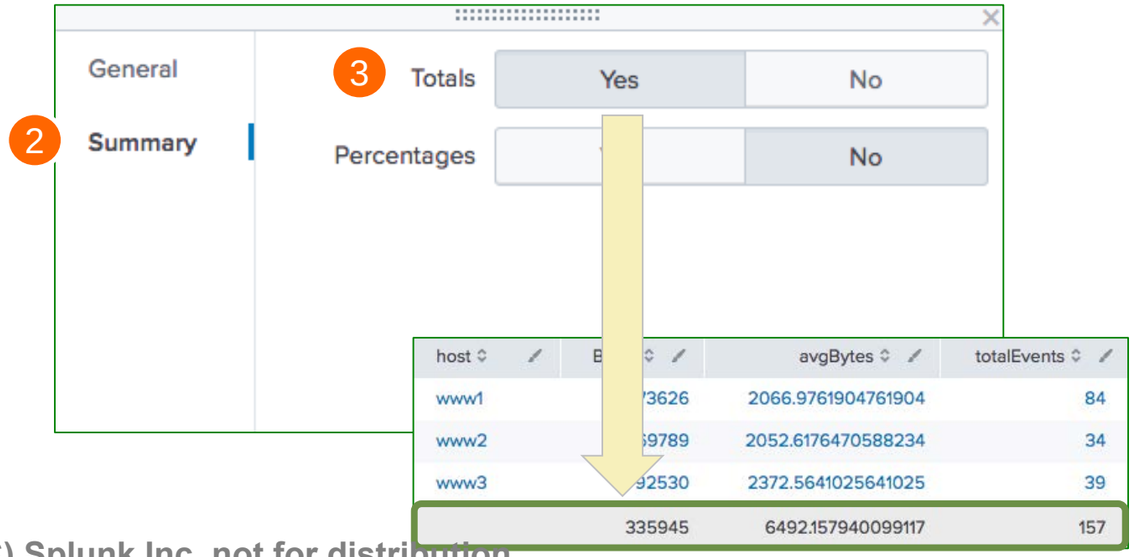
Scenario

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



host	B	avgBytes	totalEvents
www1	3626	2066.9761904761904	84
www2	9789	2052.6176470588234	34
www3	92530	2372.5641025641025	39
	335945	6492.157940099117	157

Generated for (v.ms.besetti@jpmchase.com) (C) Splunk Inc, not for distribution

Copyright © 2018 Splunk, Inc. All rights reserved | 14 June 2018

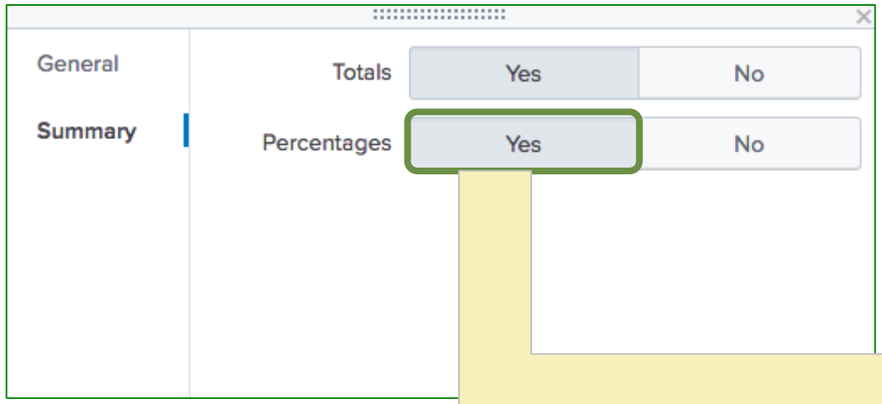
Adding Totals Using Format Options (cont.)

- 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 (v.ms.besetti@jpmchase.com) (C) Splunk Inc, not for distribution

Copyright © 2018 Splunk, Inc. All rights reserved | 14 June 2018

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 (v.ms.besetti@jpmchase.com) (C) Splunk Inc, not for distribution

Copyright © 2018 Splunk, Inc. All rights reserved | 14 June 2018

addtotals Command: Syntax

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

Row Options		Column Options	
`row=true/false` (Default= `true`)	A column is created that contains numeric totals for each row.	`col=true/false` (Default= `false`)	A row is created that contains numeric totals for each column.
`fieldname=`*field* (Default=`Total`)	Defines a string used to create a field name for the totals column.	`label=`*string* (Default=`Total`)	Defines a string used to name the totals row.
		`labelfield=` *fieldname*	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.

Generated for (v.ms.besetti@jpmchase.com) (C) Splunk Inc, not for distribution

Copyright © 2018 Splunk, Inc. All rights reserved | 14 June 2018

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 C	China (PRC)	The Bahamas	B United States	A 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 C	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.

Generated for (v.ms.besetti@jpmchase.com) (C) Splunk Inc, not for distribution

Copyright © 2018 Splunk, Inc. All rights reserved | 14 June 2018

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		

Generated for (v.ms.besetti@jpmchase.com) (C) Splunk Inc, not for distribution

Copyright © 2018 Splunk, Inc. All rights reserved | 14 June 2018

Module 5: Filtering and Formatting Data

Generated for (v.ms.besetti@jpmchase.com) (C) Splunk Inc, not for distribution

Copyright © 2018 Splunk, Inc. All rights reserved

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

Generated for (v.ms.besetti@jpmchase.com) (C) Splunk Inc, not for distribution

Copyright © 2018 Splunk, Inc. All rights reserved | 14 June 2018

`eval` Command – Overview

- `eval` allows you to calculate and manipulate field values in your report

  `eval` *`fieldname1 = 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

Generated for (v.ms.besetti@jpmchase.com) (C) Splunk Inc, not for distribution

Copyright © 2018 Splunk, Inc. All rights reserved | 14 June 2018

eval Command

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

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

-

Type	Operators
Arithmetic	`+ - * / %`
Concatenation	`+ .`
Boolean	`AND OR NOT XOR`
Comparison	`< > <= >= != = == LIKE`

Note

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

Generated for (v.ms.besetti@jpmchase.com) (C) Splunk Inc, not for distribution

Copyright © 2018 Splunk, Inc. All rights reserved | 14 June 2018

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

Generated for (v.ms.besetti@jpmchase.com) (C) Splunk Inc, not for distribution

Copyright © 2018 Splunk, Inc. All rights reserved | 14 June 2018

eval Command – Convert Values (cont.)

- Results of `eval` must be set to a new or existing field
- In this example:
  - A Calculate the number of bytes for each usage type
  - B Create a new field named `bandwidth`
  - C 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	C 51.50892448425293
Business	66844576	63.747955322265625
Personal	299584913	285.7064371109009
Unknown	77187989	73.61220264434814
Violation	3203231	3.0548391342163086

Generated for (v.ms.besetti@jpmchase.com) (C) Splunk Inc, not for distribution

Copyright © 2018 Splunk, Inc. All rights reserved | 14 June 2018

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)
| sort -bandwidth
| rename bandwidth as "Bandwidth (MB)"
```

A

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

A

Generated for (v.ms.besetti@jpmchase.com) (C) Splunk Inc, not for distribution

Copyright © 2018 Splunk, Inc. All rights reserved | 14 June 2018

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	A	73.61
Business		63.75
Borderline		51.51
Violation		3.05

Generated for (v.ms.besetti@jpmchase.com) (C) Splunk Inc, not for distribution

Copyright © 2018 Splunk, Inc. All rights reserved | 14 June 2018

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 (v.ms.besetti@jpmchase.com) (C) Splunk Inc, not for distribution

Copyright © 2018 Splunk, Inc. All rights reserved | 14 June 2018

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)
C | sort -Discount
D | eval Discount = Discount."%"
E | 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 (v.ms.besetti@jpmchase.com) (C) Splunk Inc, not for distribution

Copyright © 2018 Splunk, Inc. All rights reserved | 14 June 2018

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

Generated for (v.ms.besetti@jpmchase.com) (C) Splunk Inc, not for distribution

Copyright © 2018 Splunk, Inc. All rights reserved | 14 June 2018

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

Generated for (v.ms.besetti@jpmchase.com) (C) Splunk Inc, not for distribution

Copyright © 2018 Splunk, Inc. All rights reserved | 14 June 2018

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
```

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

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

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 (v.ms.besetti@jpmchase.com) (C) Splunk Inc, not for distribution

Copyright © 2018 Splunk, Inc. All rights reserved | 14 June 2018

`eval` Commands with Multiple Expressions

- 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...
```

Generated for (v.ms.besetti@jpmchase.com) (C) Splunk Inc, not for distribution

Copyright © 2018 Splunk, Inc. All rights reserved | 14 June 2018

eval Commands with Multiple Expressions – Example

- 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 -  A
  current_sale_price)/list_price*100,2),
  new_discount = (current_discount - 5),      B
  new_sale_price = list_price - (list_price *
  (new_discount/100))  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 (v.ms.besetti@jpmchase.com) (C) Splunk Inc, not for distribution

Copyright © 2018 Splunk, Inc. All rights reserved | 14 June 2018

`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

Generated for (v.ms.besetti@jpmchase.com) (C) Splunk Inc, not for distribution

Copyright © 2018 Splunk, Inc. All rights reserved | 14 June 2018

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")
```

X Y Z

```
| 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 (v.ms.besetti@jpmchase.com) (C) Splunk Inc, not for distribution

Copyright © 2018 Splunk, Inc. All rights reserved | 14 June 2018

eval Command – case Function

`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 you want an "otherwise" clause, just test for a condition you know is true at the end (e.g., `0=0`)

Note

The **`like`** operator is discussed later in this module.

```
index=web sourcetype=access_combined
| eval rating = case(productId LIKE "WC%", "Teen", productId LIKE "FS%", Mature",
0=0, "Unrated")
```

X1 Y1 X2 Y2 X3 Y3

Generated for (v.ms.besetti@jpmchase.com) (C) Splunk Inc, not for distribution

Copyright © 2018 Splunk, Inc. All rights reserved | 14 June 2018

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 <u>are</u> 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

Generated for (v.ms.besetti@jpmchase.com) (C) Splunk Inc, not for distribution

Copyright © 2018 Splunk, Inc. All rights reserved | 14 June 2018

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

Generated for (v.ms.besetti@jpmchase.com) (C) Splunk Inc, not for distribution

Copyright © 2018 Splunk, Inc. All rights reserved | 14 June 2018

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
| sort -sales
| eval sales="$"+sales
| rename sales as "Popular Products",
  product_name as "Product Name"
```

A

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

A

Generated for (v.ms.besetti@jpmchase.com) (C) Splunk Inc, not for distribution

Copyright © 2018 Splunk, Inc. All rights reserved | 14 June 2018

`where` Command

`where eval-expression`

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

Note

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

Generated for (v.ms.besetti@jpmchase.com) (C) Splunk Inc, not for distribution

Copyright © 2018 Splunk, Inc. All rights reserved | 14 June 2018

where Command – Example

- Filters search results using eval expressions
- Used to compare two different fields

Scenario

Report 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
2018-01-09	126	135

A

Generated for (v.ms.besetti@jpmchase.com) (C) Splunk Inc, not for distribution

Copyright © 2018 Splunk, Inc. All rights reserved | 14 June 2018

`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

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_.%"
```

src_ip	count
102.2.83.137	102
107.3.146.207	55
108.65.113.83	34
109.169.32.135	72

Note

As you've seen, the **`like`** operator can also be used with the **`case`** function.

Generated for (v.ms.besetti@jpmchase.com) (C) Splunk Inc, not for distribution

Copyright © 2018 Splunk, Inc. All rights reserved | 14 June 2018

`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 Learn More

Replaces null values with a specified value.

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

Generated for (v.ms.besetti@jpmchase.com) (C) Splunk Inc, not for distribution

Copyright © 2018 Splunk, Inc. All rights reserved | 14 June 2018

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

Generated for (v.ms.besetti@jpmchase.com) (C) Splunk Inc, not for distribution

Copyright © 2018 Splunk, Inc. All rights reserved | 14 June 2018

Module 6: Correlating Events

Generated for (v.ms.besetti@jpmchase.com) (C) Splunk Inc, not for distribution

Copyright © 2018 Splunk, Inc. All rights reserved

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`

Generated for (v.ms.besetti@jpmchase.com) (C) Splunk Inc, not for distribution

Copyright © 2018 Splunk, Inc. All rights reserved | 14 June 2018

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.

Generated for (v.ms.besetti@jpmchase.com) (C) Splunk Inc, not for distribution

Copyright © 2018 Splunk, Inc. All rights reserved | 14 June 2018

`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 Learn More

Groups events into transactions.

Example:

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

Generated for (v.ms.besetti@jpmchase.com) (C) Splunk Inc, not for distribution

Copyright © 2018 Splunk, Inc. All rights reserved | 14 June 2018

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

Generated for (v.ms.besetti@jpmchase.com) (C) Splunk Inc, not for distribution

Copyright © 2018 Splunk, Inc. All rights reserved | 14 June 2018

`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
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
175.44.1.122 - - [18/Jan/2018:00:01:05] "GET /category.screen?categoryId=STRATEGY&JSESSIONID=SD0SL10FF3ADFF4950 HTT
P 1.1" 200 2453 "http://www.buttercupgames.com/oldlink?itemId=EST-18" "Mozilla/4.0 (compatible; MSIE 7.0; Windows N
T 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
```

Generated for (v.ms.besetti@jpmchase.com) (C) Splunk Inc, not for distribution

Copyright © 2018 Splunk, Inc. All rights reserved | 14 June 2018

`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
4:00:53.000 PM

```
... 2 lines omitted ...
175.44.1.122 - - [18/Jan/2018:00:01:08] "GET /oldlink?itemId A -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- A 9&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=E A &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
```

Generated for (v.ms.besetti@jpmchase.com) (C) Splunk Inc, not for distribution

Copyright © 2018 Splunk, Inc. All rights reserved | 14 June 2018

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 3:49:43.000 PM	Wed Jan 17 2018 23:49:43 www1 sshd[2796]: Failed password for backup from 92.46.53.223 port 3138 ssh2 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 3:49:43.000 PM	Wed Jan 17 2018 23:49:43 mailsv1 sshd[5594]: Failed password for invalid user db2 from 118.142.68.222 port 4345 ssh2 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

Generated for (v.ms.besetti@jpmchase.com) (C) Splunk Inc, not for distribution

Copyright © 2018 Splunk, Inc. All rights reserved | 14 June 2018

`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

Generated for (v.ms.besetti@jpmchase.com) (C) Splunk Inc, not for distribution

Copyright © 2018 Splunk, Inc. All rights reserved | 14 June 2018

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)

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

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.

Generated for (v.ms.besetti@jpmchase.com) (C) Splunk Inc, not for distribution

Copyright © 2018 Splunk, Inc. All rights reserved | 14 June 2018

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

Generated for (v.ms.besetti@jpmchase.com) (C) Splunk Inc, not for distribution

Copyright © 2018 Splunk, Inc. All rights reserved | 14 June 2018

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

Generated for (v.ms.besetti@jpmchase.com) (C) Splunk Inc, not for distribution

Copyright © 2018 Splunk, Inc. All rights reserved | 14 June 2018

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 (192.168.3.120) address 82.90.51.177 reverse dns host host177 51 static.90 82 b.business.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 (192.168.3.120) 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 (v.ms.besetti@jpmchase.com) (C) Splunk Inc, not for distribution

Copyright © 2018 Splunk, Inc. All rights reserved | 14 June 2018

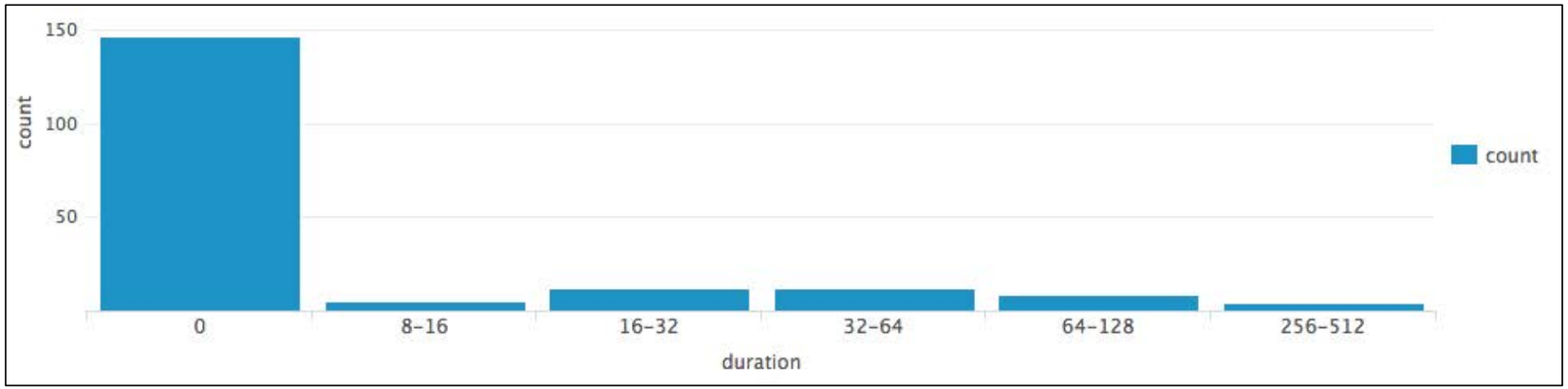
Reporting on Transactions

- 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 (v.ms.besetti@jpmchase.com) (C) Splunk Inc, not for distribution

Copyright © 2018 Splunk, Inc. All rights reserved | 14 June 2018

`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`

Generated for (v.ms.besetti@jpmchase.com) (C) Splunk Inc, not for distribution

Copyright © 2018 Splunk, Inc. All rights reserved | 14 June 2018

transaction vs. stats: Example 1

A

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

B

```
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
```

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`

Generated for (v.ms.besetti@jpmchase.com) (C) Splunk Inc, not for distribution

Copyright © 2018 Splunk, Inc. All rights reserved | 14 June 2018

transaction vs. stats: Example 2

A

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

Note

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

- A took 6.163 seconds
- B took 4.643 seconds

B

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

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 (v.ms.besetti@jpmchase.com) (C) Splunk Inc, not for distribution

Copyright © 2018 Splunk, Inc. All rights reserved | 14 June 2018

Module 7: Introduction to Knowledge Objects

Generated for (v.ms.besetti@jpmchase.com) (C) Splunk Inc, not for distribution

Copyright © 2018 Splunk, Inc. All rights reserved

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)

Generated for (v.ms.besetti@jpmchase.com) (C) Splunk Inc, not for distribution

Copyright © 2018 Splunk, Inc. All rights reserved | 14 June 2018

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 (v.ms.besetti@jpmchase.com) (C) Splunk Inc, not for distribution

Copyright © 2018 Splunk, Inc. All rights reserved | 14 June 2018

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
- Searchable
  - Since the objects are persistent, they can be used in a search

Note

For more information go to: http://docs.splunk.com/Documentation/Splunk/7.0.0/Knowledge/WhatisSplunkknowledge

Generated for (v.ms.besetti@jpmchase.com) (C) Splunk Inc, not for distribution

Copyright © 2018 Splunk, Inc. All rights reserved | 14 June 2018

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

Generated for (v.ms.besetti@jpmchase.com) (C) Splunk Inc, not for distribution

Copyright © 2018 Splunk, Inc. All rights reserved

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
- So, for example: `SEG_Alert_WinEventlogFailures`

Note

For more information go to: http://docs.splunk.com/Documentation/Splunk/7.0.0/Knowledge/Developnamingconventionsforknowledgeobjecttitles

Generated for (v.ms.besetti@jpmchase.com) (C) Splunk Inc, not for distribution

Copyright © 2018 Splunk, Inc. All rights reserved | 14 June 2018

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

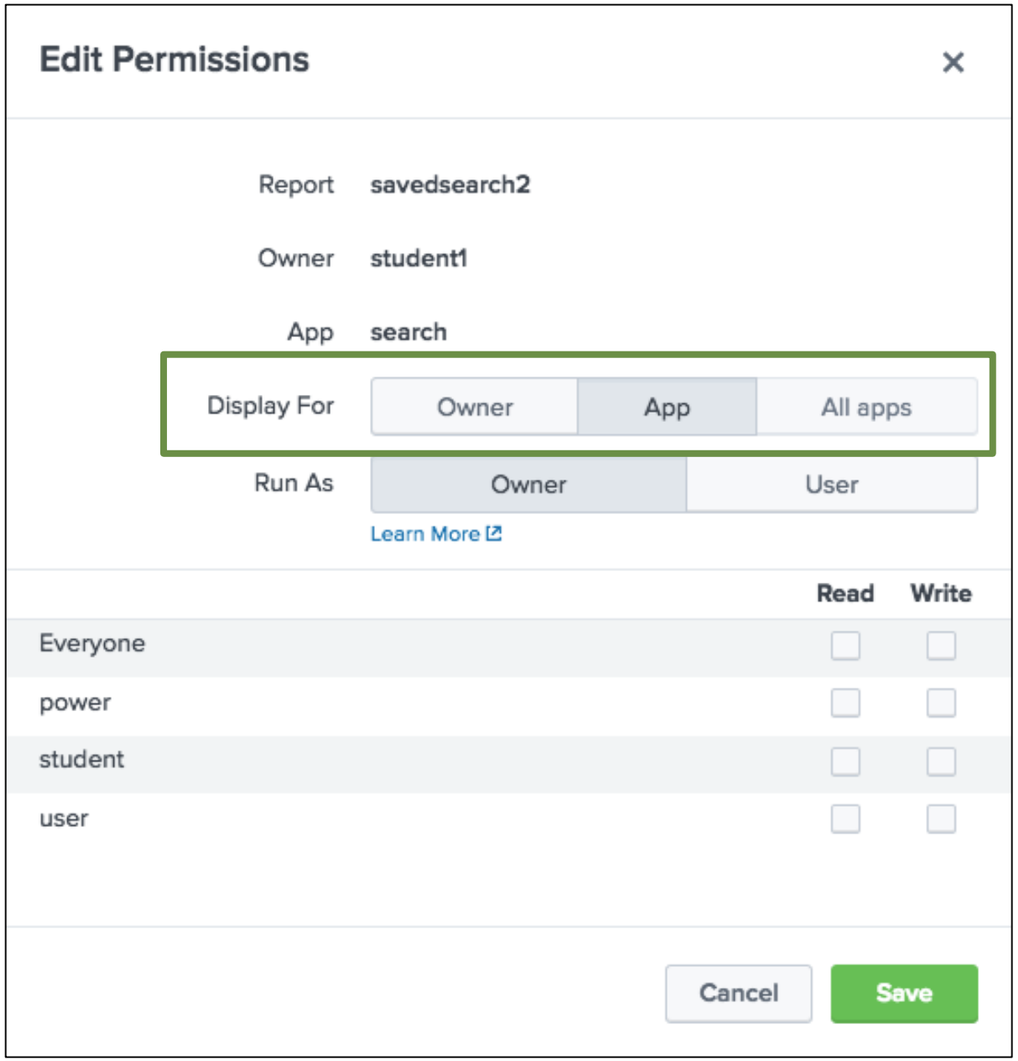
* Permission to read and/or write if creator gives permission to that role

Generated for (v.ms.besetti@jpmchase.com) (C) Splunk Inc, not for distribution

Copyright © 2018 Splunk, Inc. All rights reserved | 14 June 2018

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**, all roles are 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

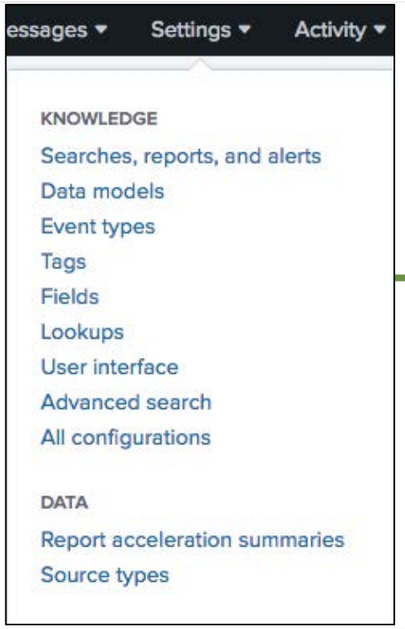


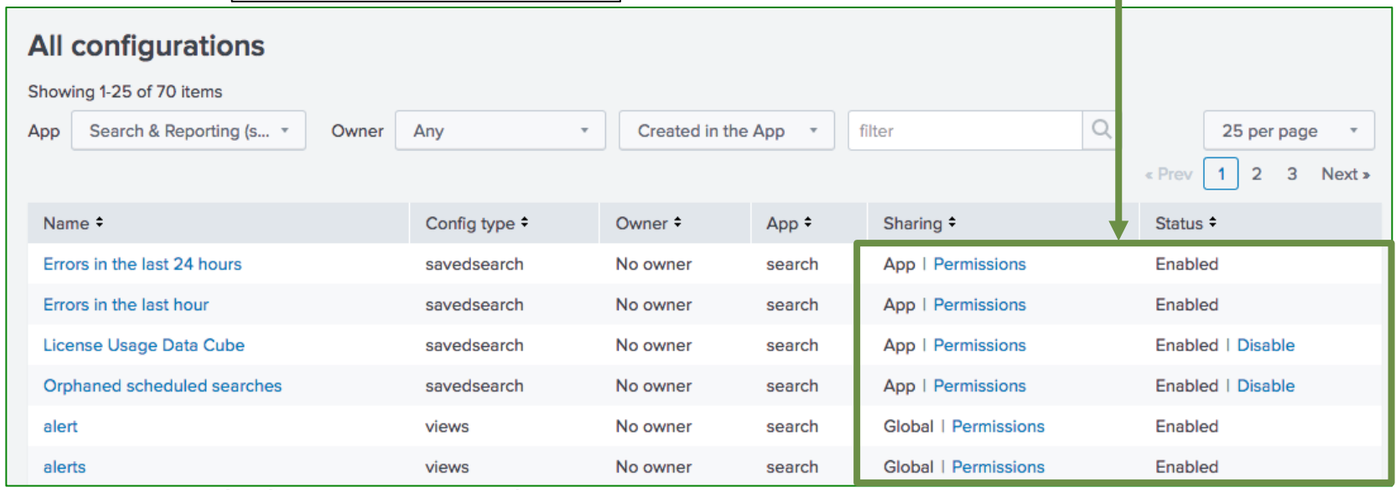
Generated for (v.ms.besetti@jpmchase.com) (C) Splunk Inc, not for distribution

Copyright © 2018 Splunk, Inc. All rights reserved | 14 June 2018

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

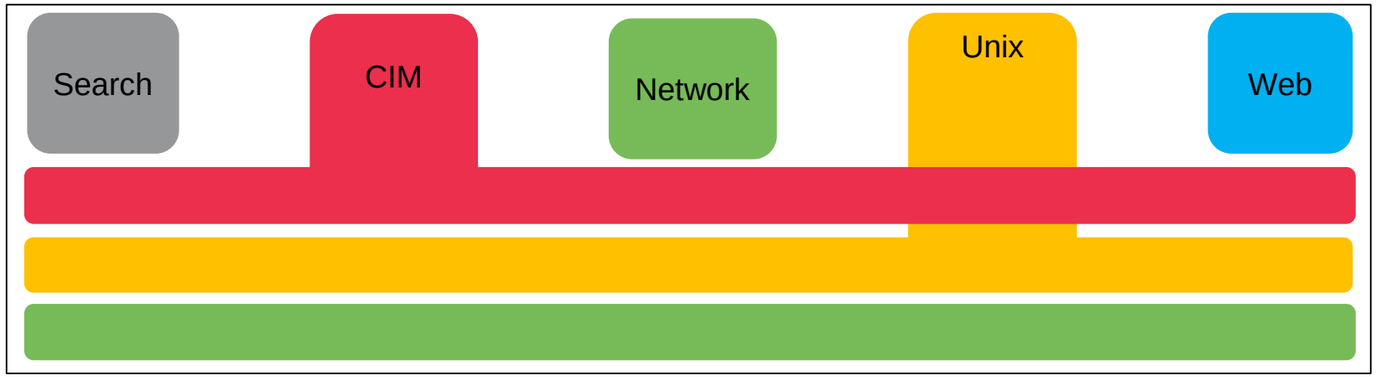
By default, objects for all owners are listed.

Generated for (v.ms.besetti@jpmchase.com) (C) Splunk Inc, not for distribution

Copyright © 2018 Splunk, Inc. All rights reserved | 14 June 2018

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 (v.ms.besetti@jpmchase.com) (C) Splunk Inc, not for distribution

Copyright © 2018 Splunk, Inc. All rights reserved | 14 June 2018

Module 8: Creating and Managing Fields

Generated for (v.ms.besetti@jpmchase.com) (C) Splunk Inc, not for distribution

Copyright © 2018 Splunk, Inc. All rights reserved

Module Objectives

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

Generated for (v.ms.besetti@jpmchase.com) (C) Splunk Inc, not for distribution

Copyright © 2018 Splunk, Inc. All rights reserved | 14 June 2018

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* discovers fields directly related to the search's results
- Splunk may also extract other fields from raw event data that aren't directly related to the search

Generated for (v.ms.besetti@jpmchase.com) (C) Splunk Inc, not for distribution

Copyright © 2018 Splunk, Inc. All rights reserved | 14 June 2018

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/7.0.0/Knowledge/ExtractfieldsinteractivelywithIFX

Generated for (v.ms.besetti@jpmchase.com) (C) Splunk Inc, not for distribution

Copyright © 2018 Splunk, Inc. All rights reserved | 14 June 2018

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)

Generated for (v.ms.besetti@jpmchase.com) (C) Splunk Inc, not for distribution

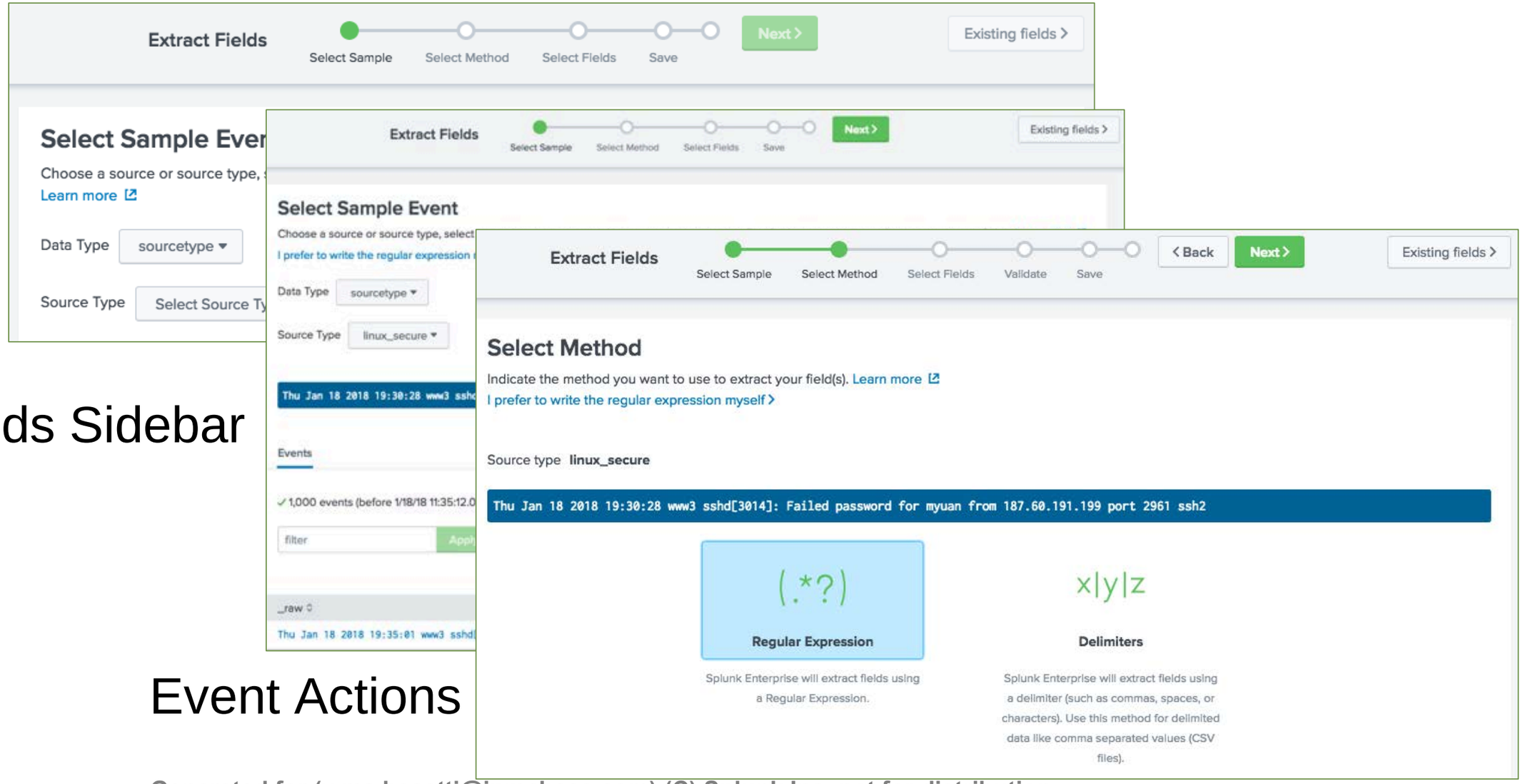
Copyright © 2018 Splunk, Inc. All rights reserved | 14 June 2018

Field Extraction Workflows – RegEx

Settings

Fields Sidebar

Event Actions

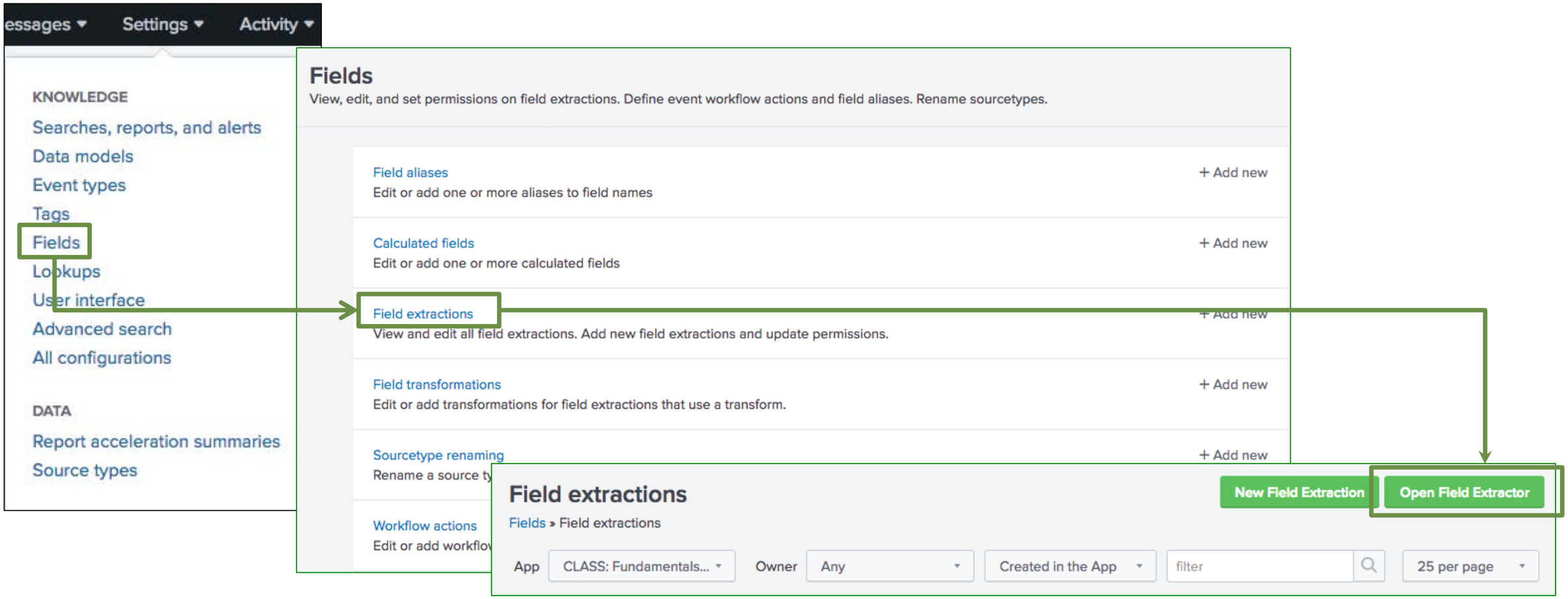


Generated for (v.ms.besetti@jpmchase.com) (C) Splunk Inc, not for distribution

Copyright © 2018 Splunk, Inc. All rights reserved

Regex Field Extractions from Settings

Settings > Fields > Field extractions > Open Field Extractor

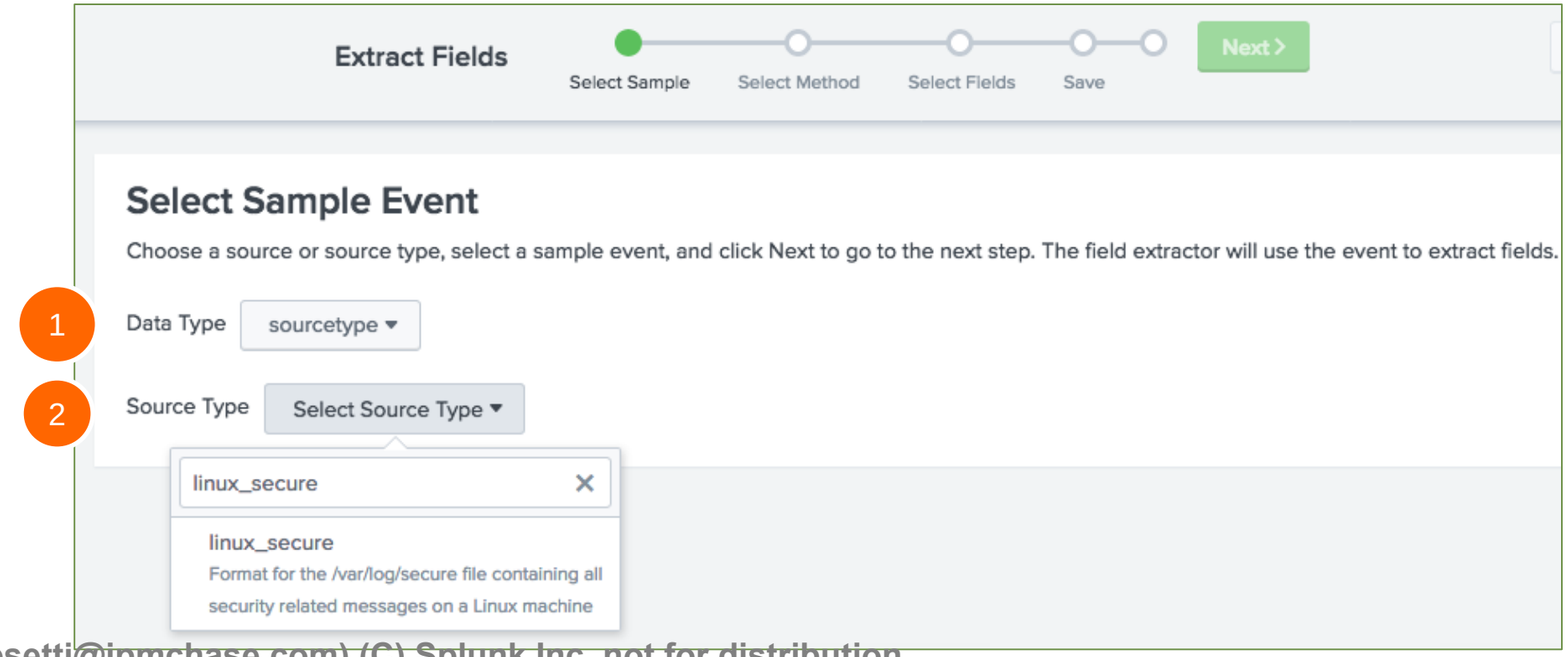


Generated for (v.ms.besetti@jpmchase.com) (C) Splunk Inc, not for distribution

Copyright © 2018 Splunk, Inc. All rights reserved | 14 June 2018

Regex Field Extractions from Settings (cont.)

1. Select the Data Type
   - sourcetype
   - source
2. Select the Source Type

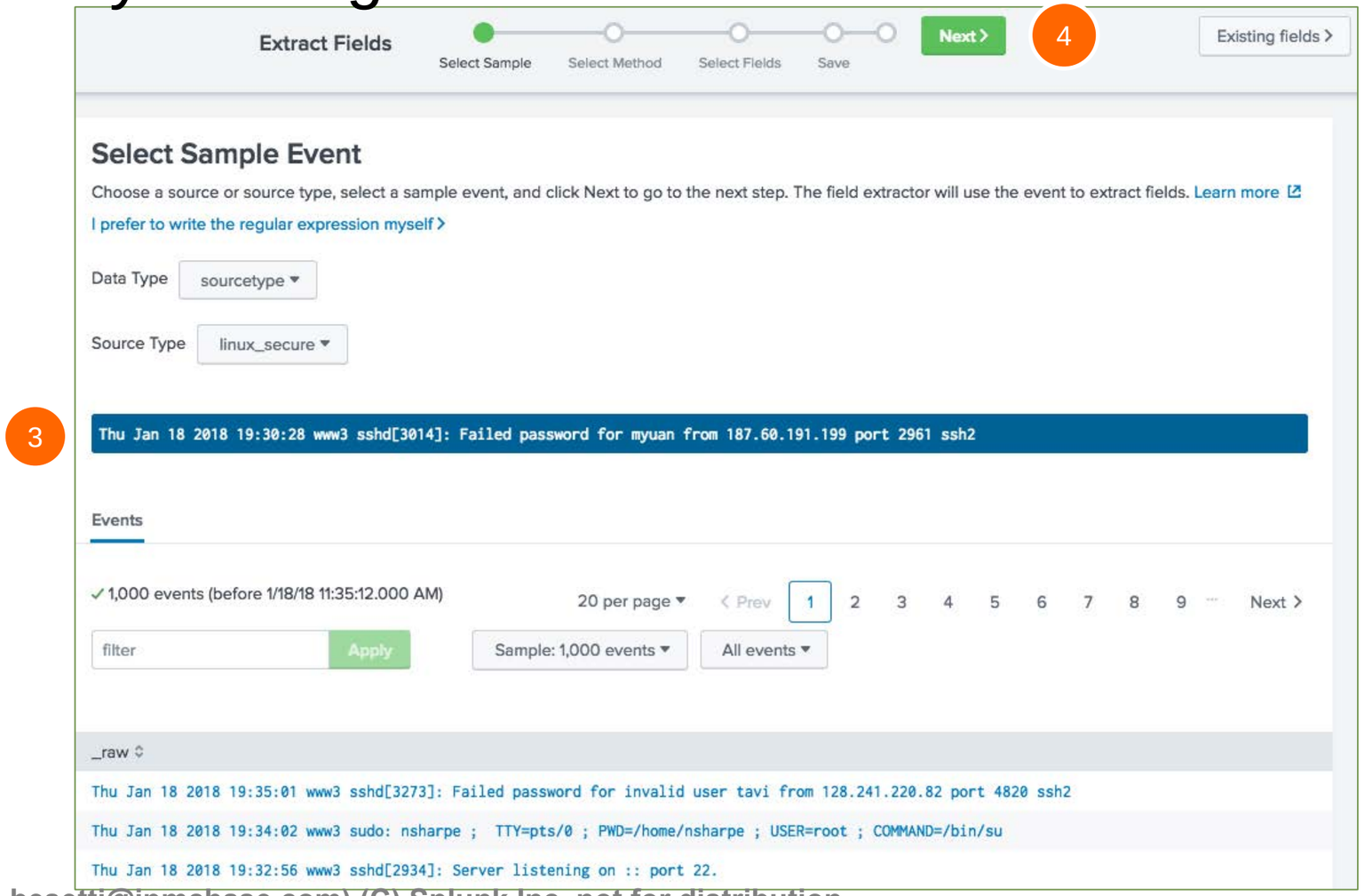


Generated for (v.ms.besetti@jpmchase.com) (C) Splunk Inc, not for distribution

Copyright © 2018 Splunk, Inc. All rights reserved | 14 June 2018

Regex Field Extractions from Settings – Select Sample

3. Select a sample event by clicking on it
4. Click Next >

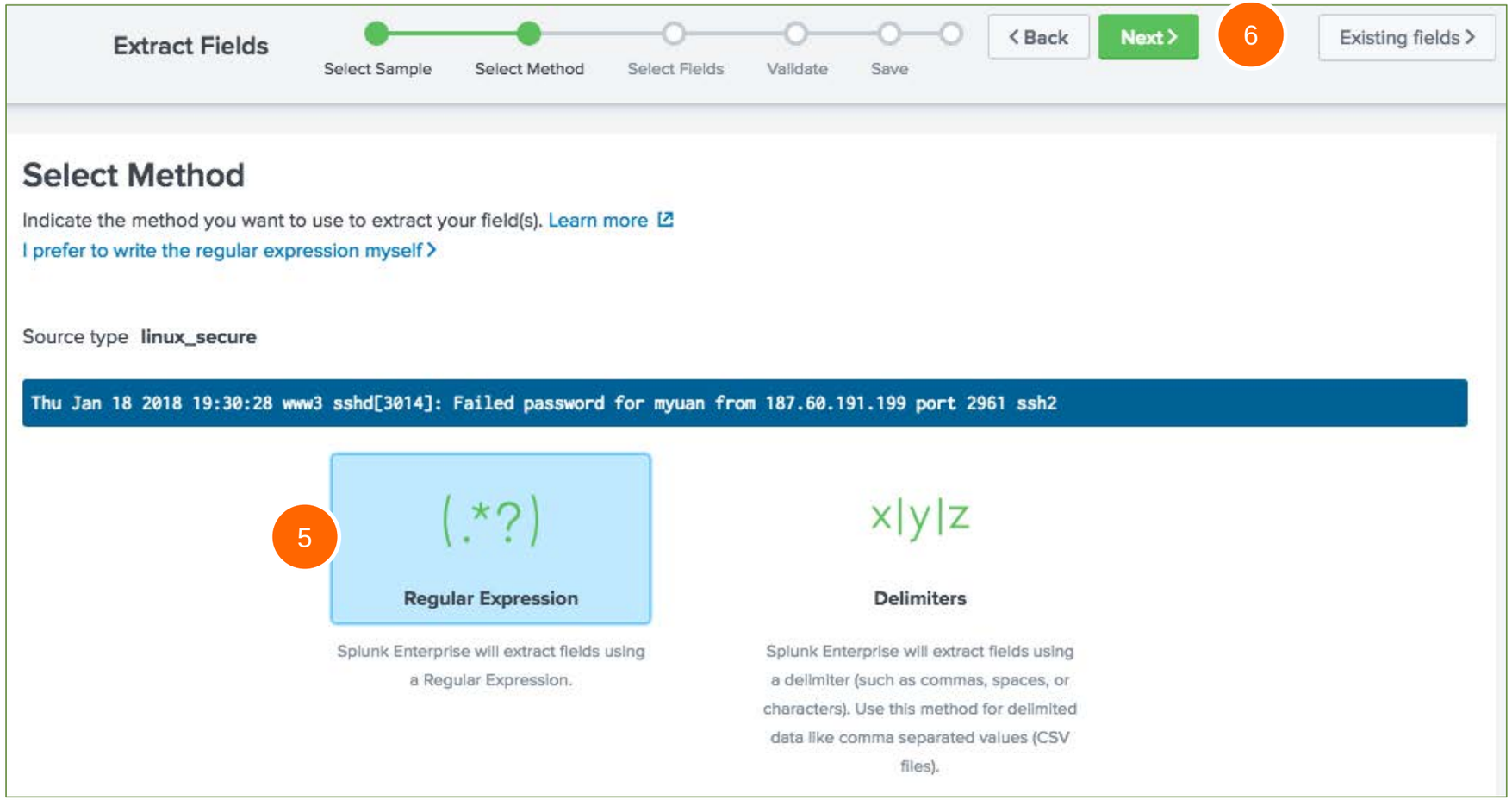


Generated for (v.ms.besetti@jpmchase.com) (C) Splunk Inc, not for distribution

Copyright © 2018 Splunk, Inc. All rights reserved | 14 June 2018

Regex Field Extractions from Settings – Select Method

5. Select Regular Expression
6. Click Next >

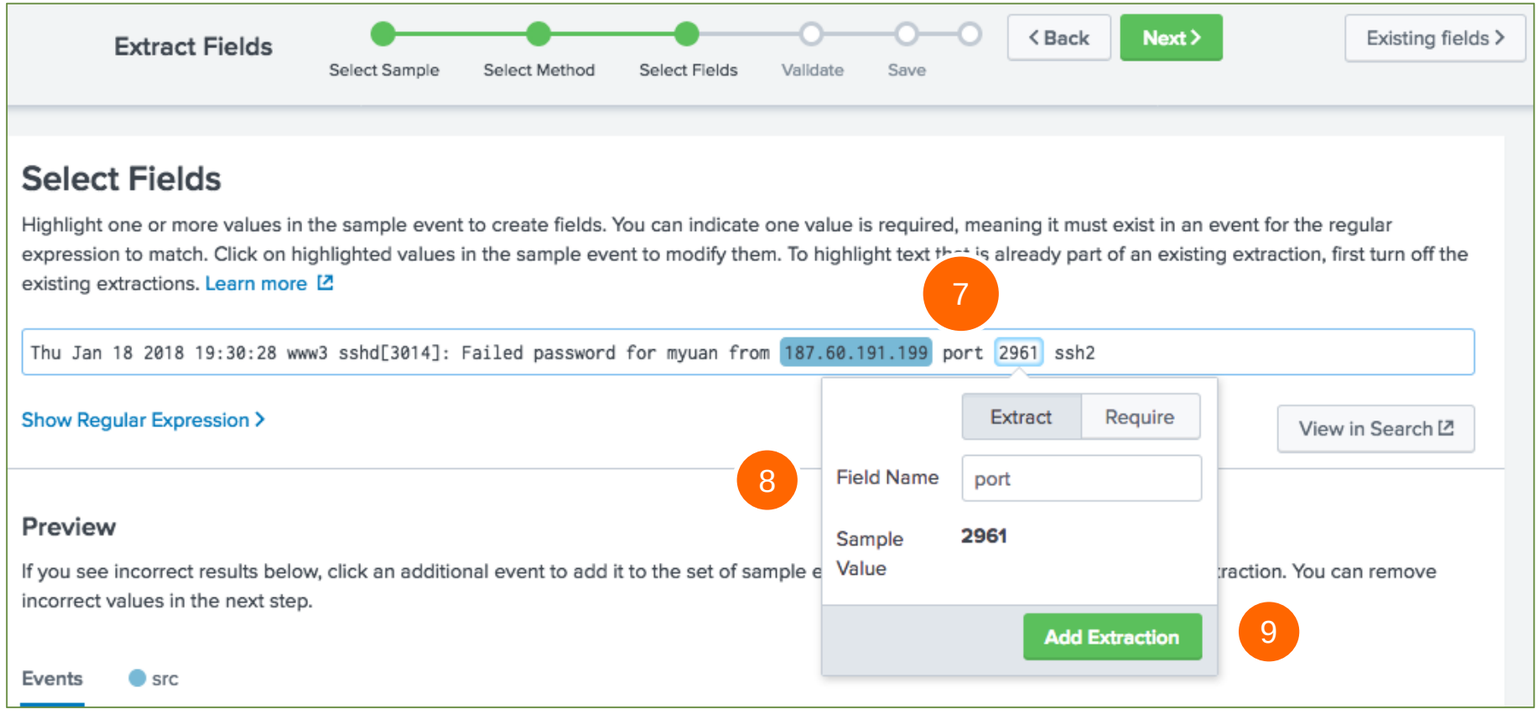


Generated for (v.ms.besetti@jpmchase.com) (C) Splunk Inc, not for distribution

Copyright © 2018 Splunk, Inc. All rights reserved | 14 June 2018

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



Note

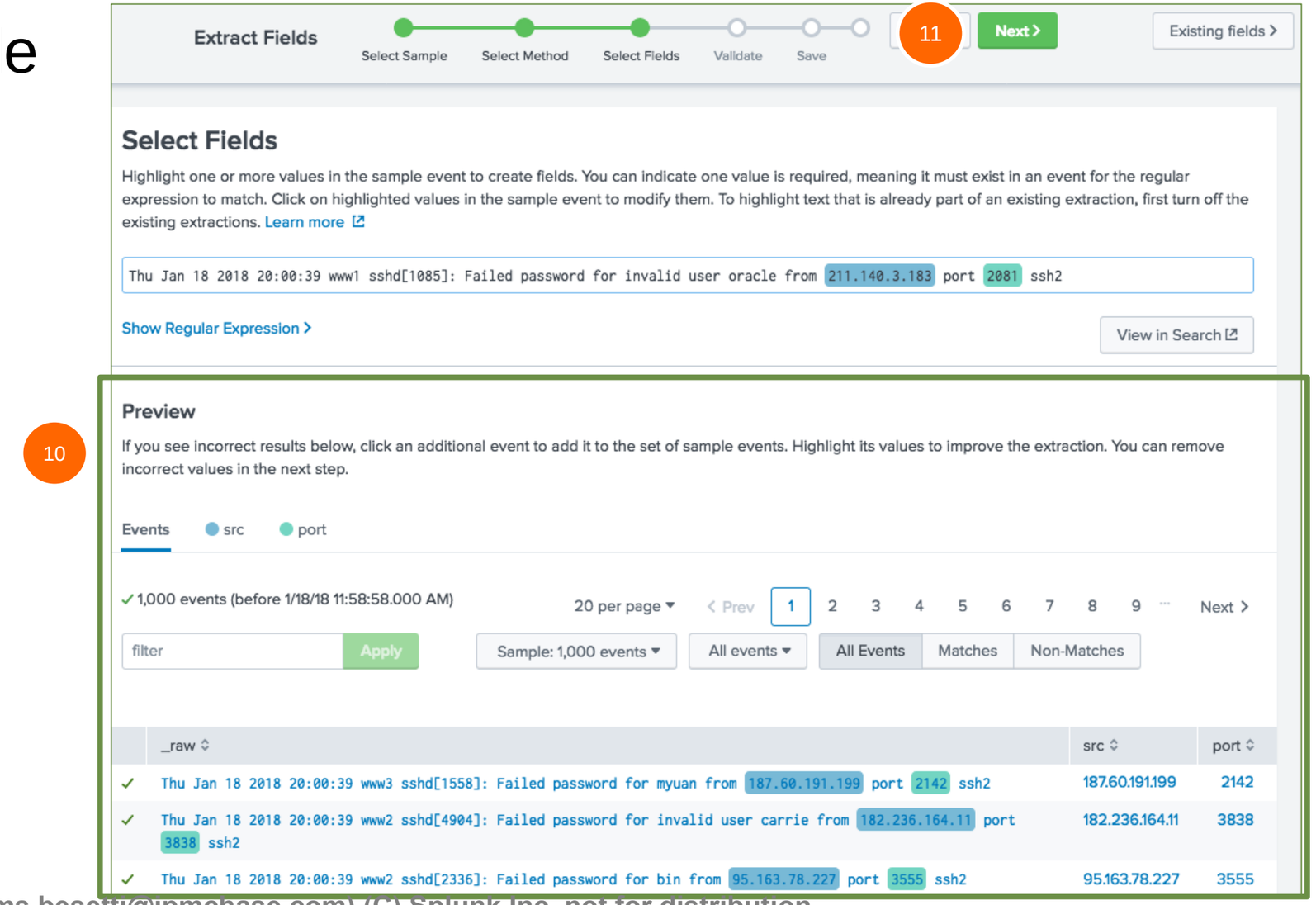
Require option – only events with the highlighted string are included in the extraction.

Generated for (v.ms.besetti@jpmchase.com) (C) Splunk Inc, not for distribution

Copyright © 2018 Splunk, Inc. All rights reserved | 14 June 2018

Regex Field Extractions from Settings – Preview

10. Preview the sample events
11. Click Next >

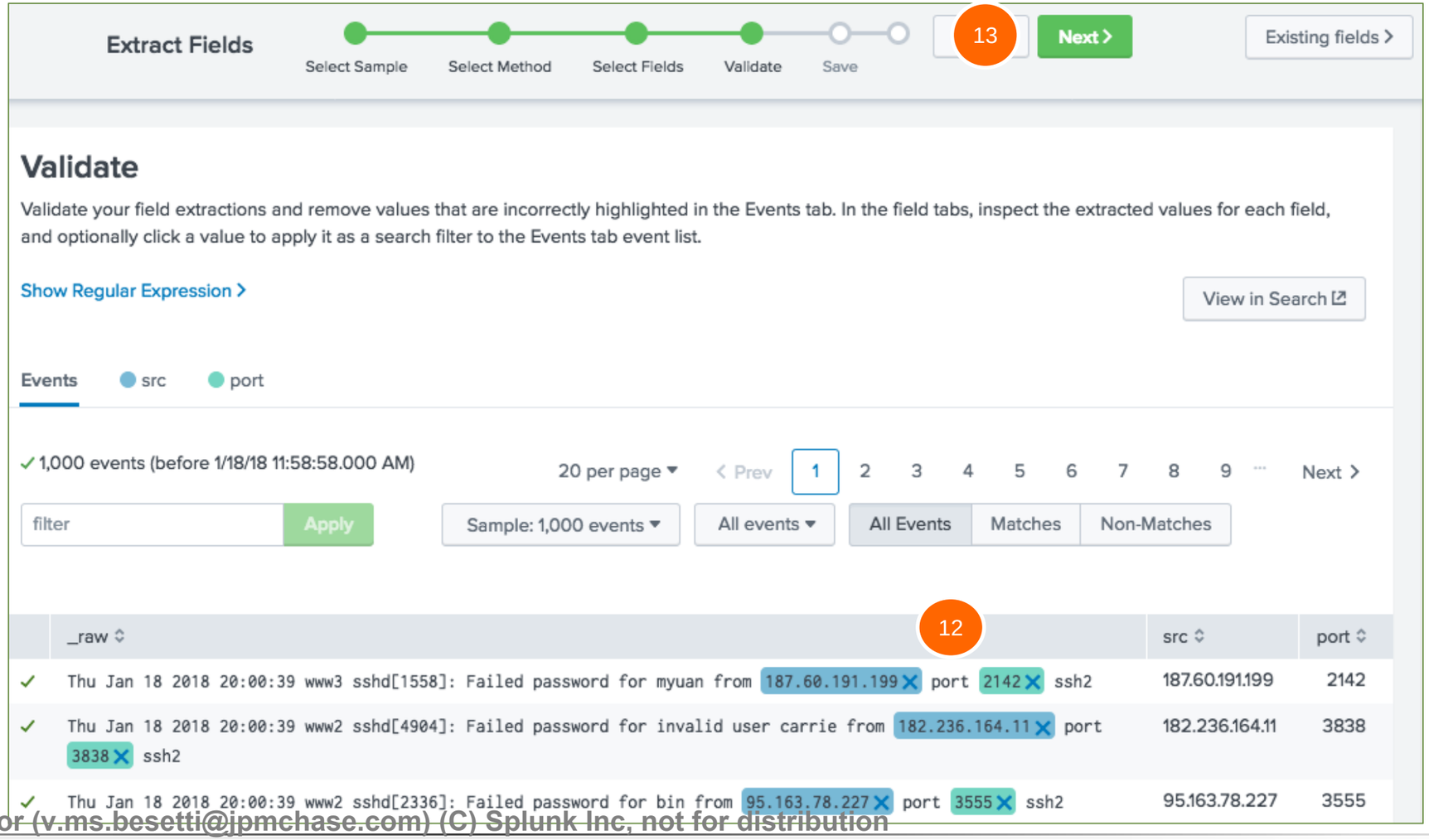


Generated for (v.ms.besetti@jpmchase.com) (C) Splunk Inc, not for distribution

Copyright © 2018 Splunk, Inc. All rights reserved | 14 June 2018

Regex Field Extractions from Settings – Validate

12. Validate the proper field values are extracted
13. Click Next >

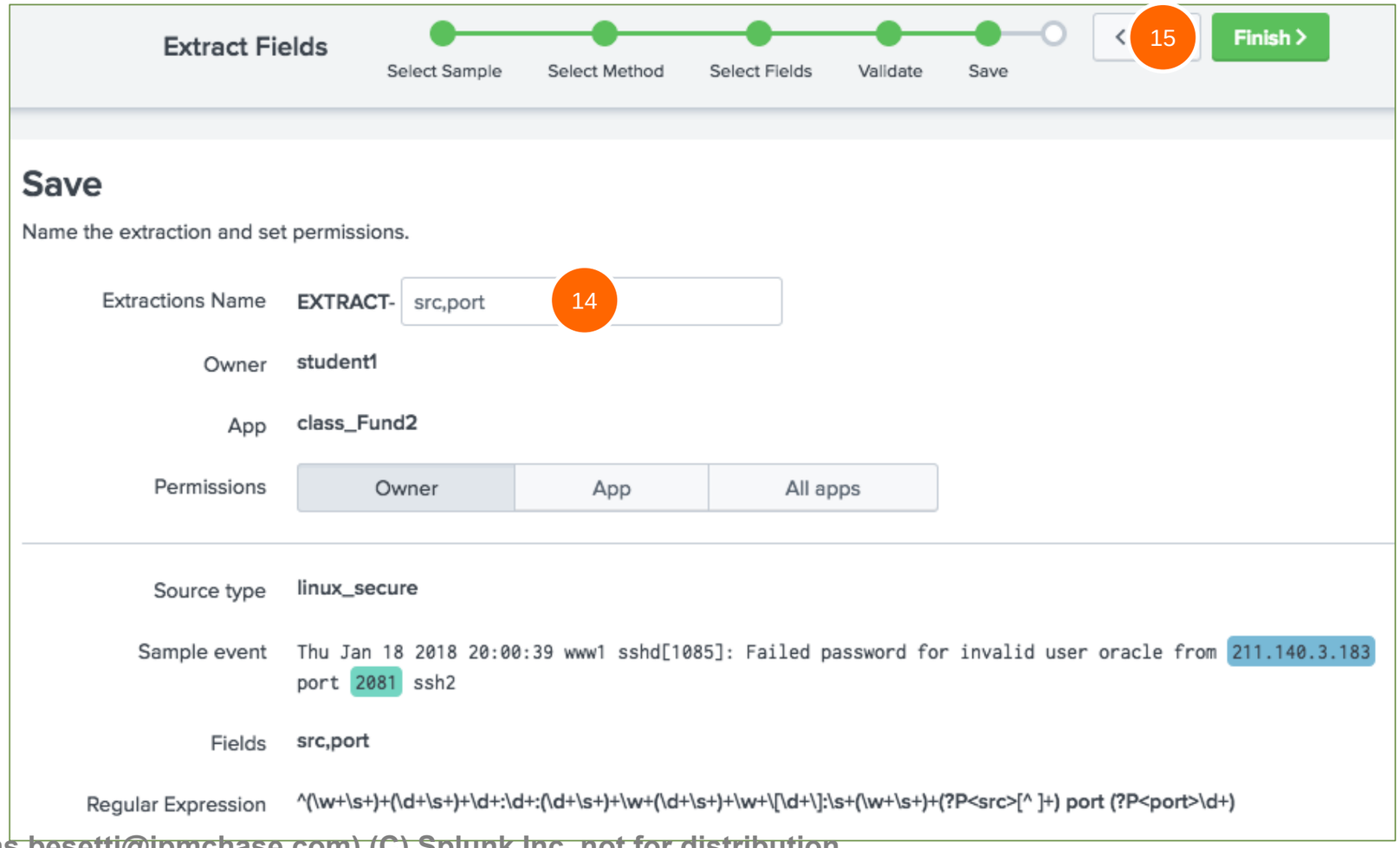


Generated for (v.ms.besetti@jpmchase.com) (C) Splunk Inc, not for distribution

Copyright © 2018 Splunk, Inc. All rights reserved | 14 June 2018

Regex Field Extractions from Settings – Save

14. Review the name for the newly extracted fields and set permissions
15. Click Finish



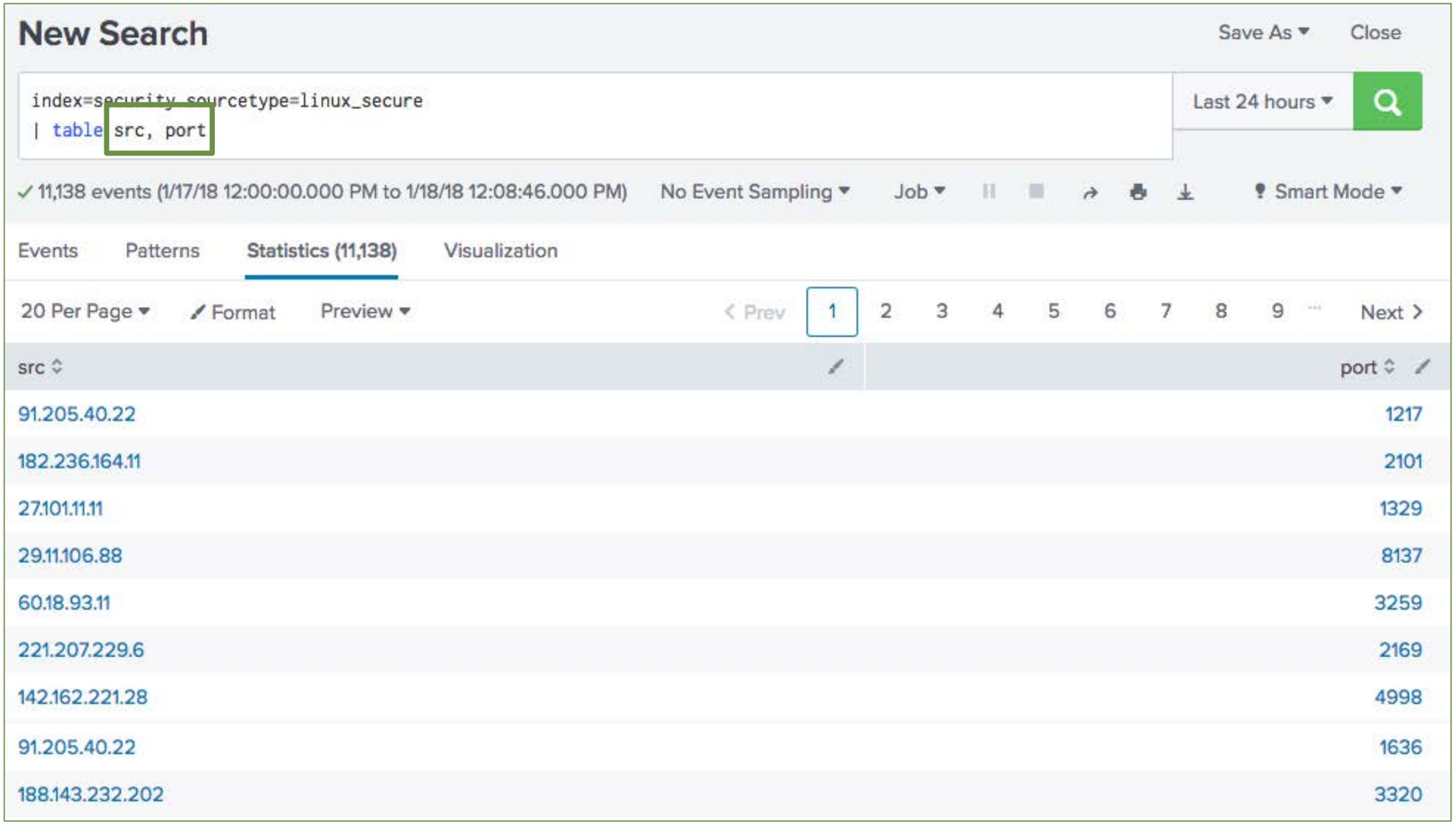
Note

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

Generated for (v.ms.besetti@jpmchase.com) (C) Splunk Inc, not for distribution

Copyright © 2018 Splunk, Inc. All rights reserved | 14 June 2018

Using the Extracted Fields

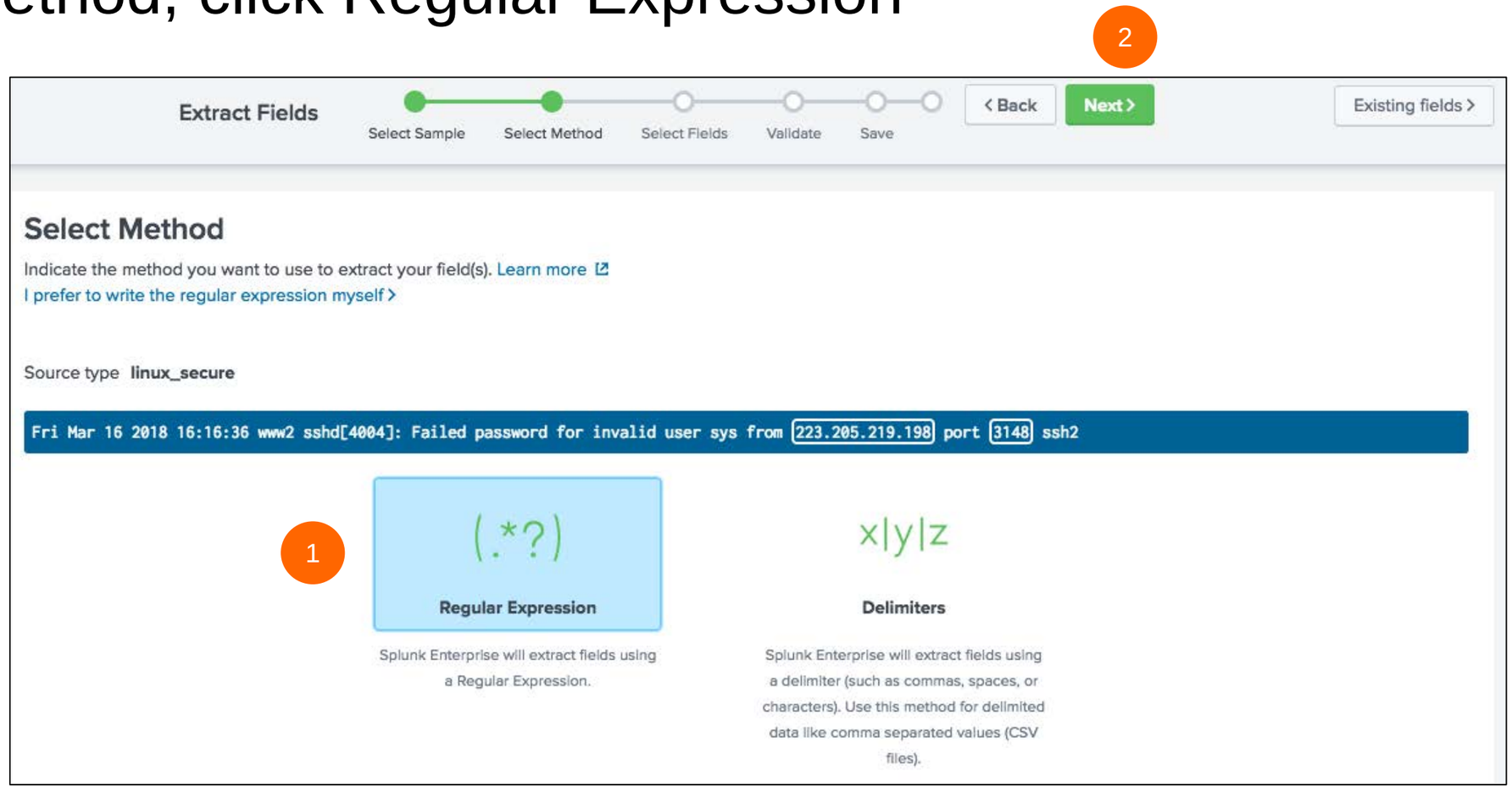


Generated for (v.ms.besetti@jpmchase.com) (C) Splunk Inc, not for distribution

Copyright © 2018 Splunk, Inc. All rights reserved | 14 June 2018

Editing Regex for Field Extractions

1. From Select Method, click Regular Expression
2. Click Next >



Generated for (v.ms.besetti@jpmchase.com) (C) Splunk Inc, not for distribution

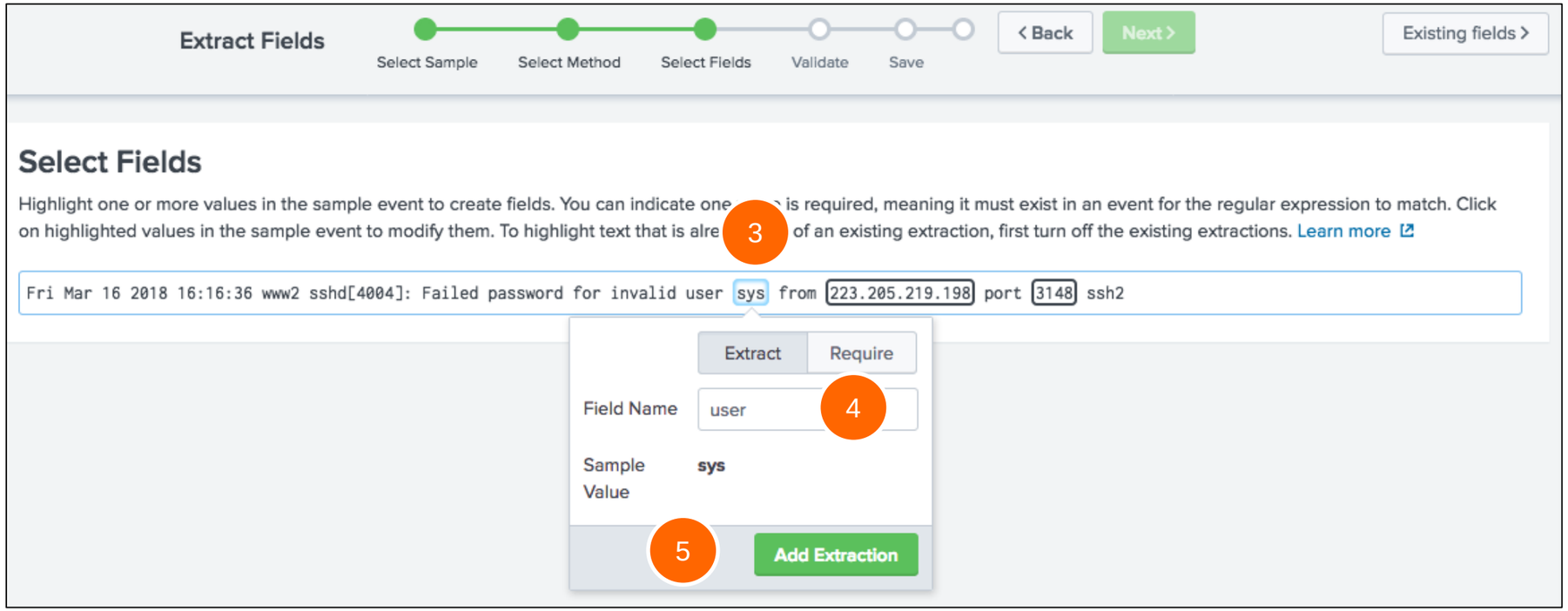
Copyright © 2018 Splunk, Inc. All rights reserved | 14 June 2018

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/7.0.0/Knowledge/AboutSplunkregularexpressions

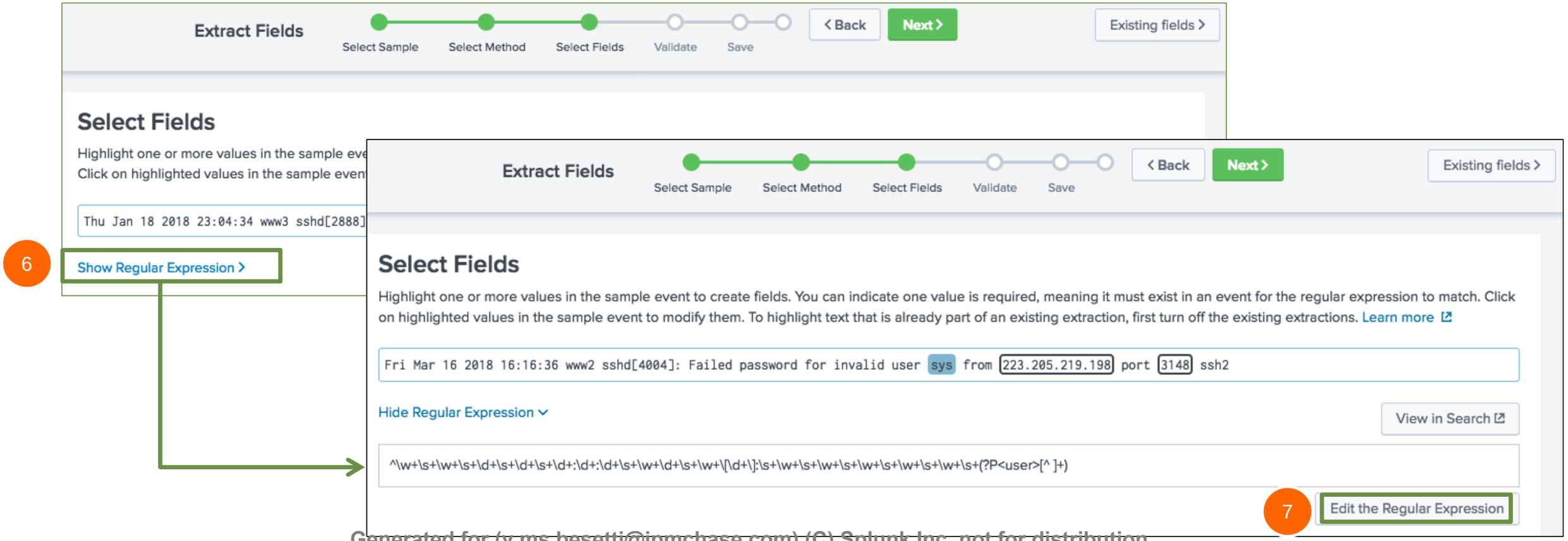


Generated for (v.ms.besetti@jpmchase.com) (C) Splunk Inc, not for distribution

Copyright © 2018 Splunk, Inc. All rights reserved | 14 June 2018

Editing Regex for Field Extractions – Show Regex

6. Click Show Regular Expression >
7. Click Edit the Regular Expression



Generated for (v.ms.besetti@jpmchase.com) (C) Splunk Inc, not for distribution

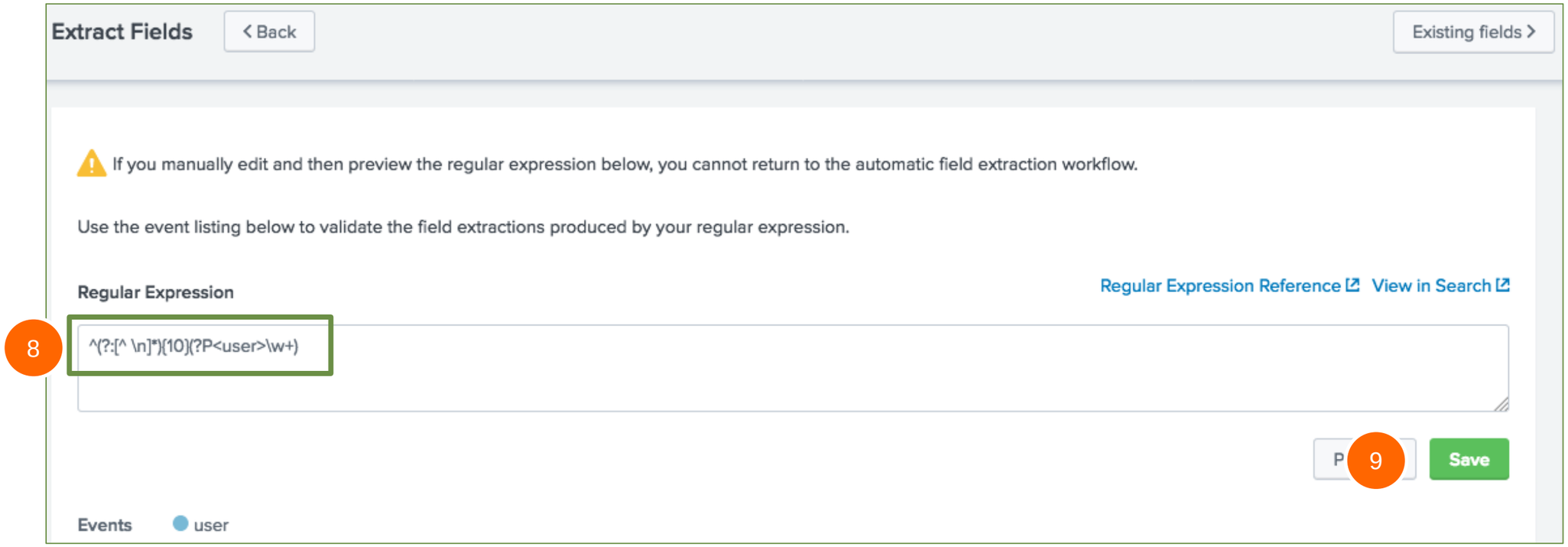
Copyright © 2018 Splunk, Inc. All rights reserved | 14 June 2018

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.

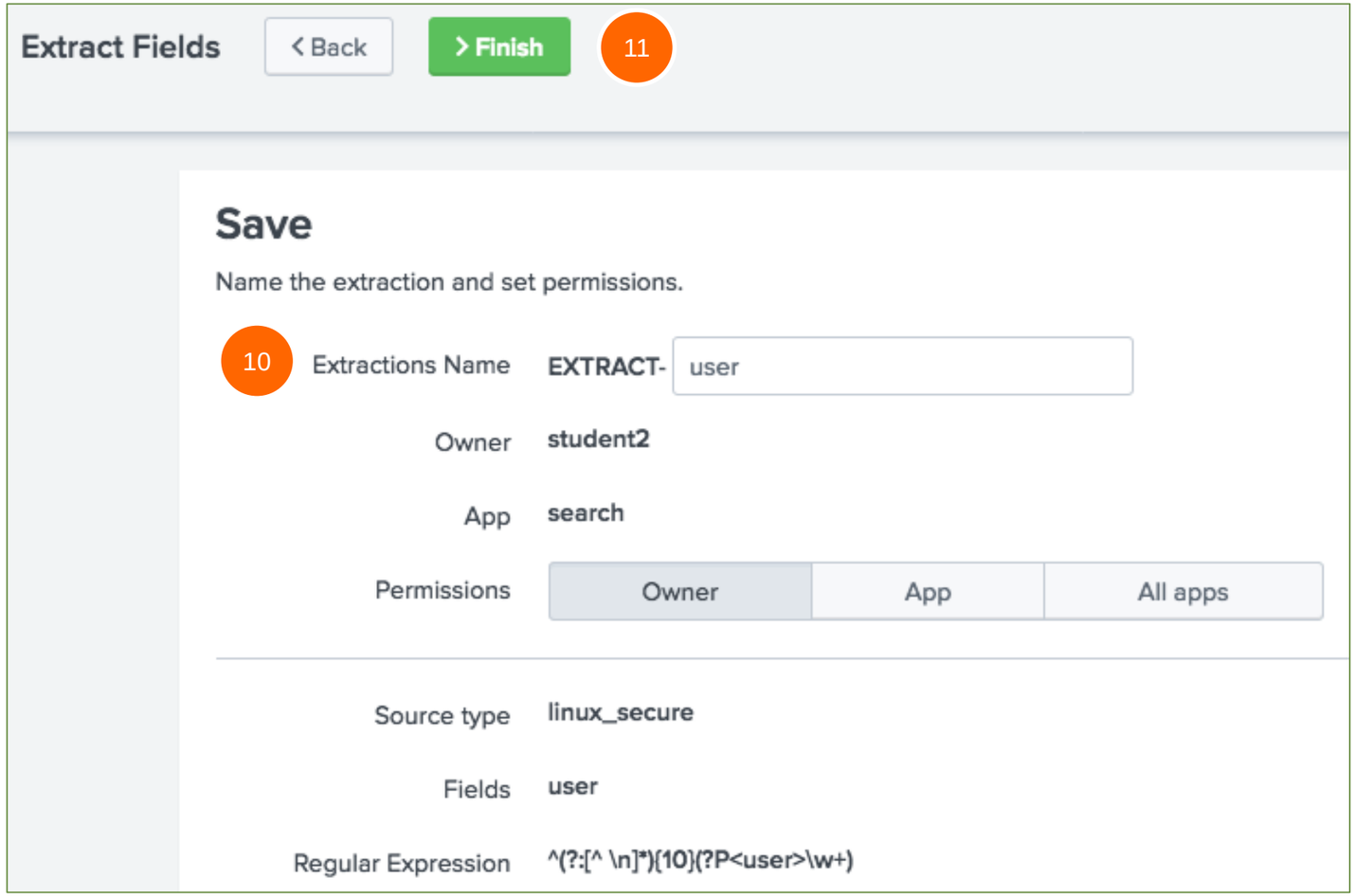


Generated for (v.ms.besetti@jpmchase.com) (C) Splunk Inc, not for distribution

Copyright © 2018 Splunk, Inc. All rights reserved | 14 June 2018

Editing Regex for Field Extractions - Save

10. Review the Extractions Name and set permissions
11. Click >Finish



Note

For more information about directly editing these objects go to:
http://docs.splunk.com/Documentation/Splunk/7.0.0/Admin/Propsconf

http://docs.splunk.com/Documentation/Splunk/7.0.0/Admin/Transformsconf

Generated for (v.ms.besetti@jpmchase.com) (C) Splunk Inc, not for distribution

Copyright © 2018 Splunk, Inc. All rights reserved | 14 June 2018

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 (v.ms.besetti@jpmchase.com) (C) Splunk Inc, not for distribution

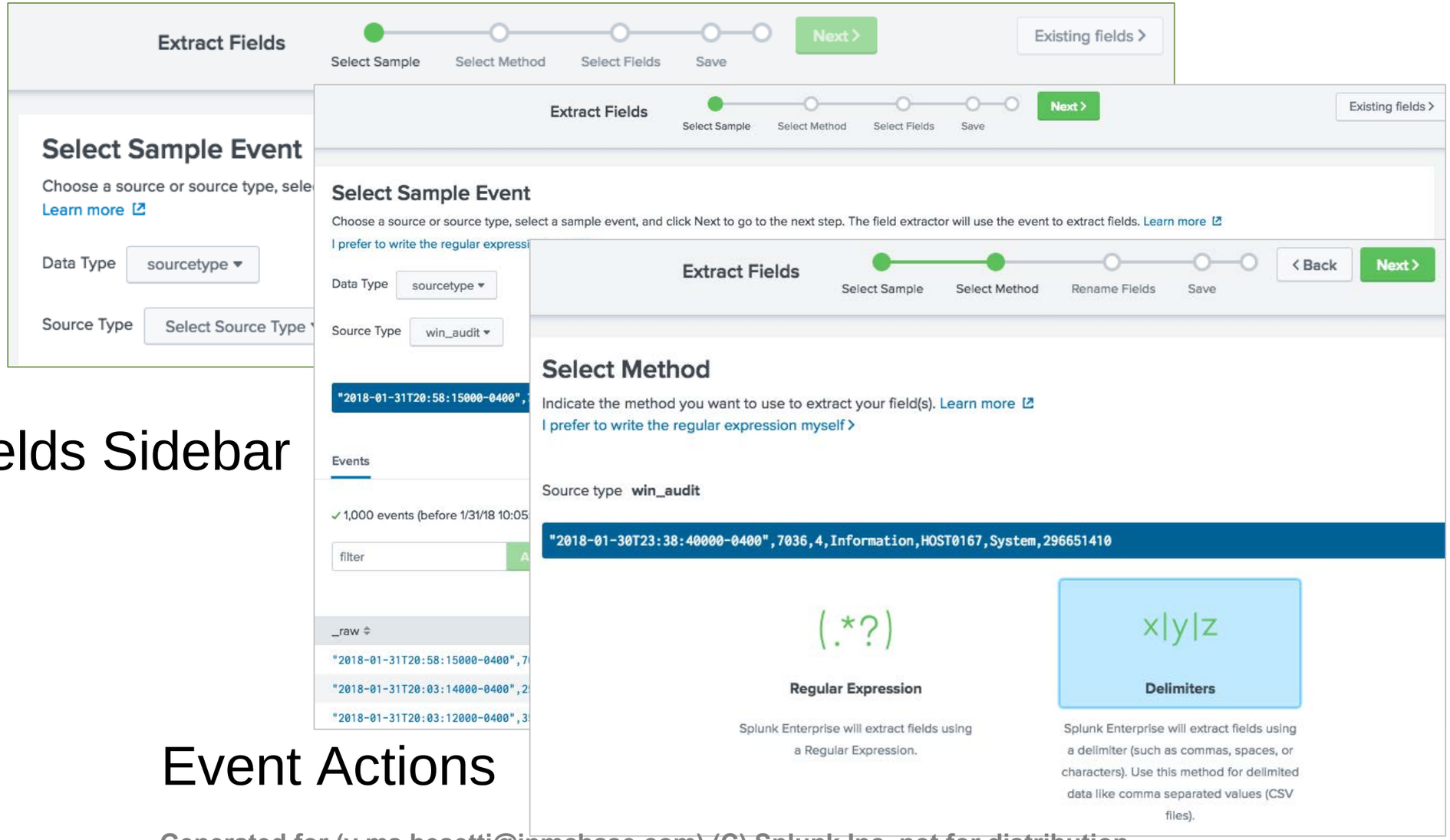
Copyright © 2018 Splunk, Inc. All rights reserved | 14 June 2018

Field Extraction Workflows – Delimiters

Settings

Fields Sidebar

Event Actions

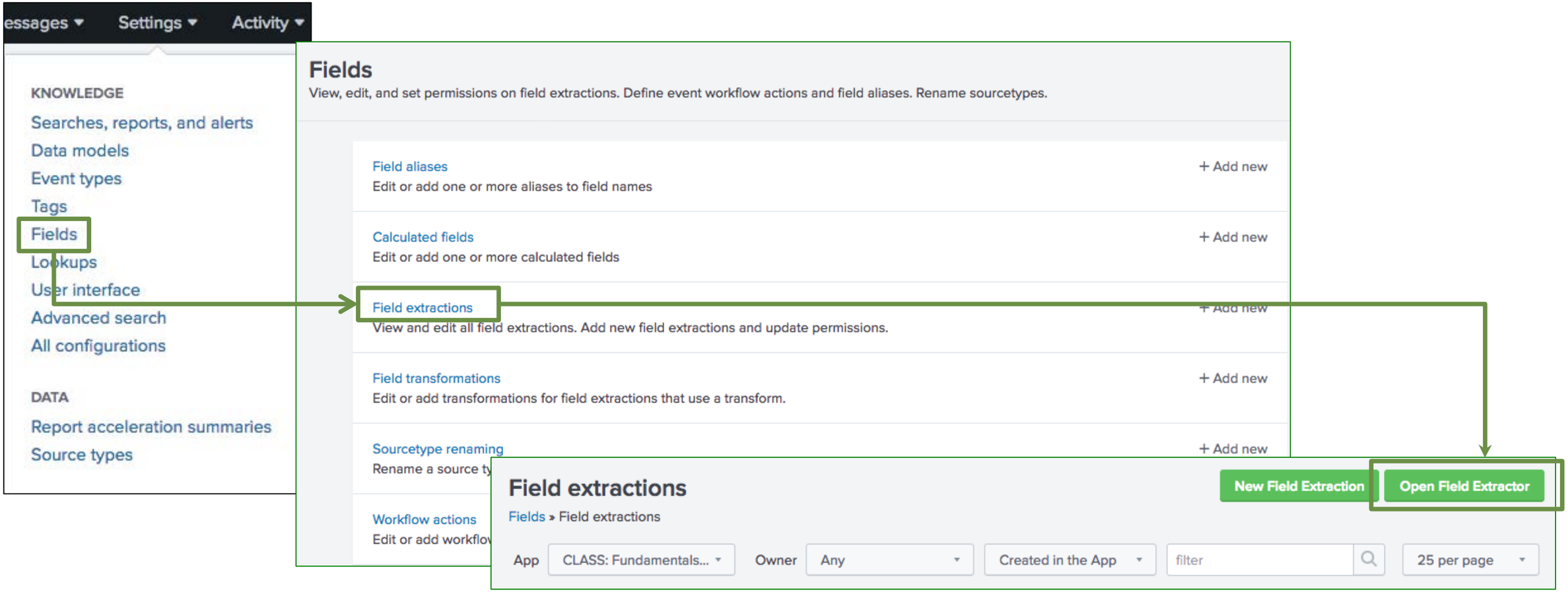


Generated for (v.ms.besetti@jpmchase.com) (C) Splunk Inc, not for distribution

Copyright © 2018 Splunk, Inc. All rights reserved | 14 June 2018

Delimited Field Extractions from Settings

Settings > Fields > Field extractions > Open Field Extractor

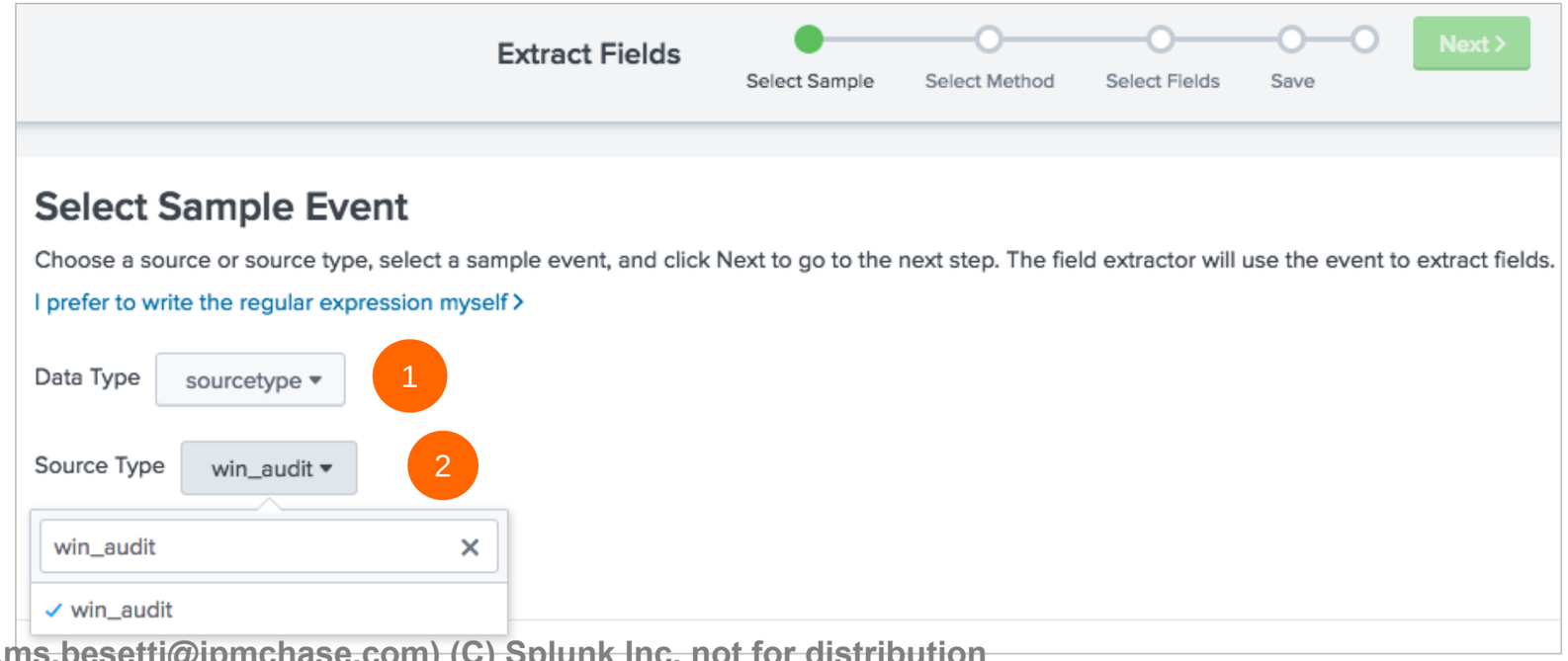


Generated for (v.ms.besetti@jpmchase.com) (C) Splunk Inc, not for distribution

Copyright © 2018 Splunk, Inc. All rights reserved | 14 June 2018

Delimited Field Extractions (Settings) – Select Sample

1. Select the Data Type
   - sourcetype
   - source
2. Select the Source Type

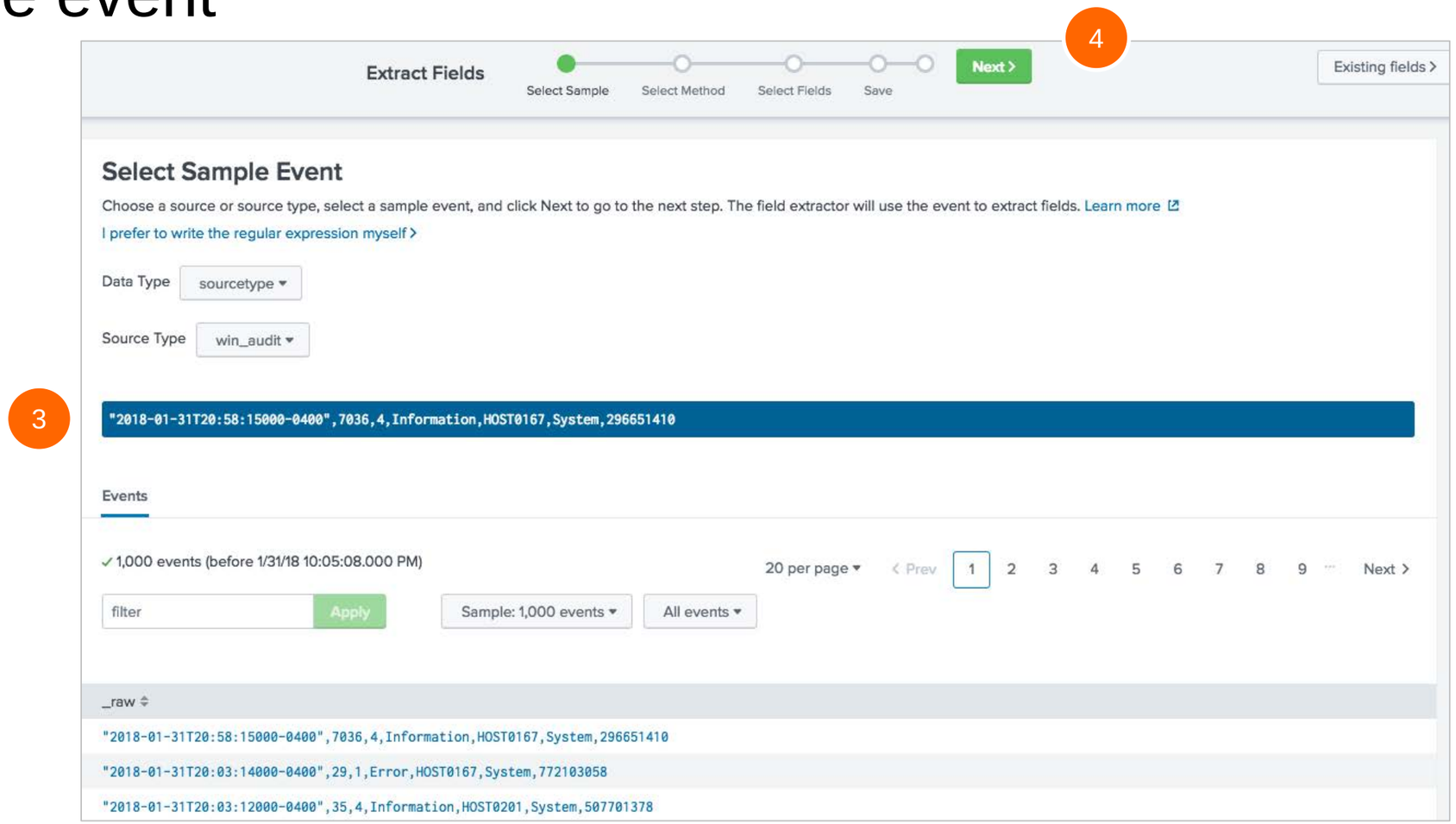


Generated for (v.ms.besetti@jpmchase.com) (C) Splunk Inc, not for distribution

Copyright © 2018 Splunk, Inc. All rights reserved | 14 June 2018

Delimited Field Extractions (Settings) – Select Event

3. Select a sample event
4. Click Next >

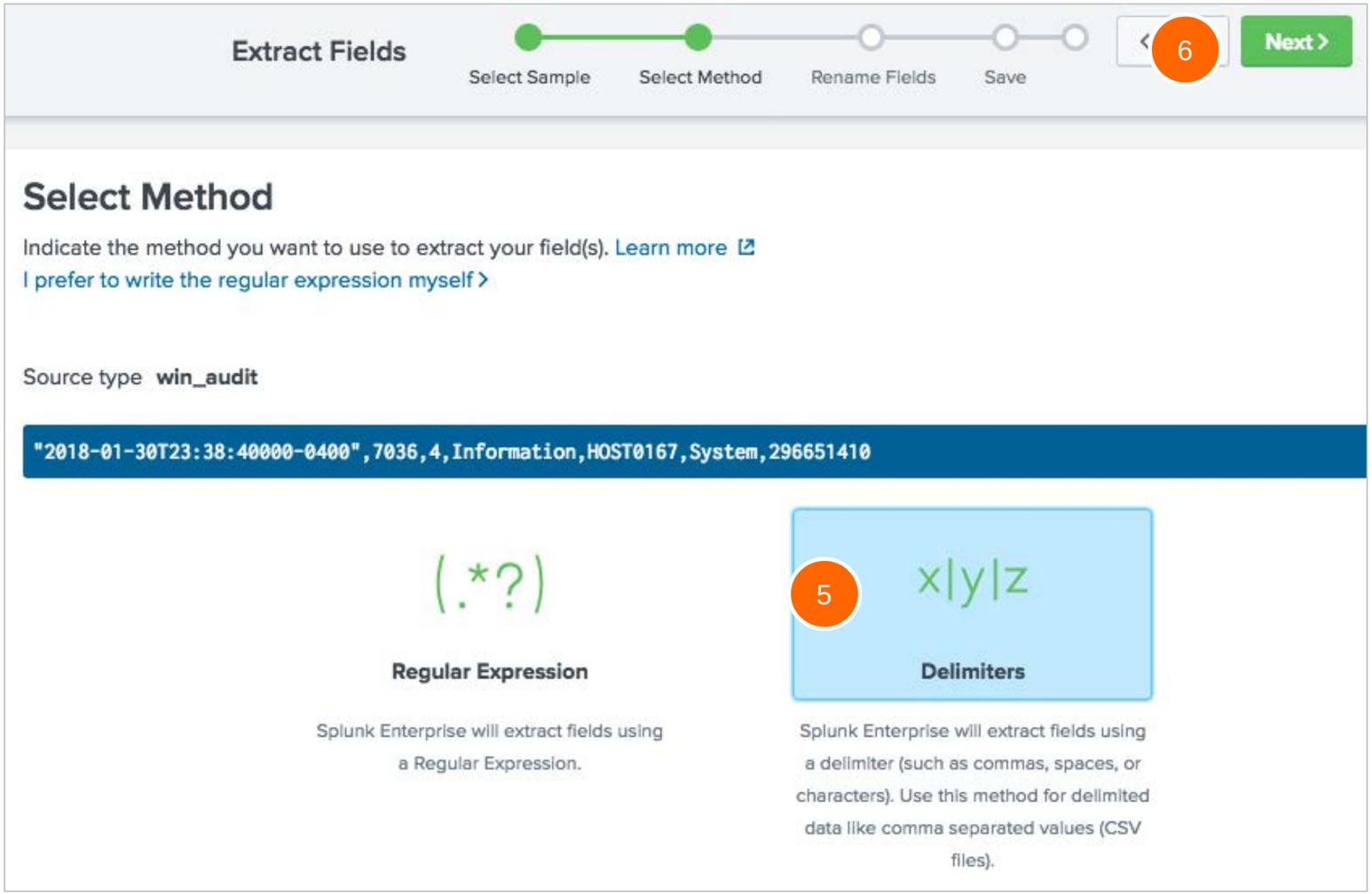


Generated for (v.ms.besetti@jpmchase.com) (C) Splunk Inc, not for distribution

Copyright © 2018 Splunk, Inc. All rights reserved | 14 June 2018

Delimited Field Extractions (Settings) – Select Method

5. Select Delimiters
6. Click Next >

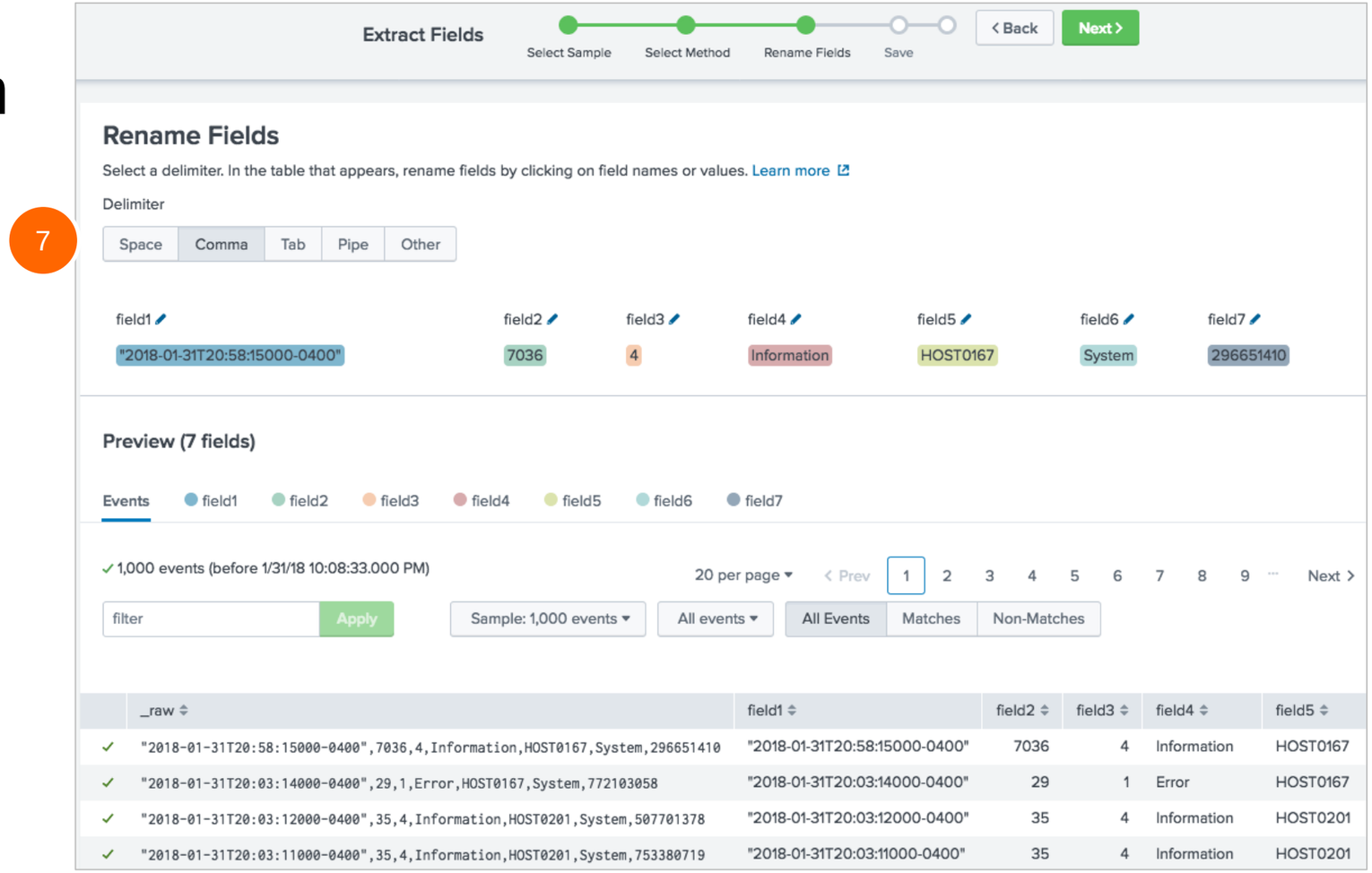


Generated for (v.ms.besetti@jpmchase.com) (C) Splunk Inc, not for distribution

Copyright © 2018 Splunk, Inc. All rights reserved

Delimited Field Extractions (Settings) - Select Delimiter

7. Select the Delimiter used in your event

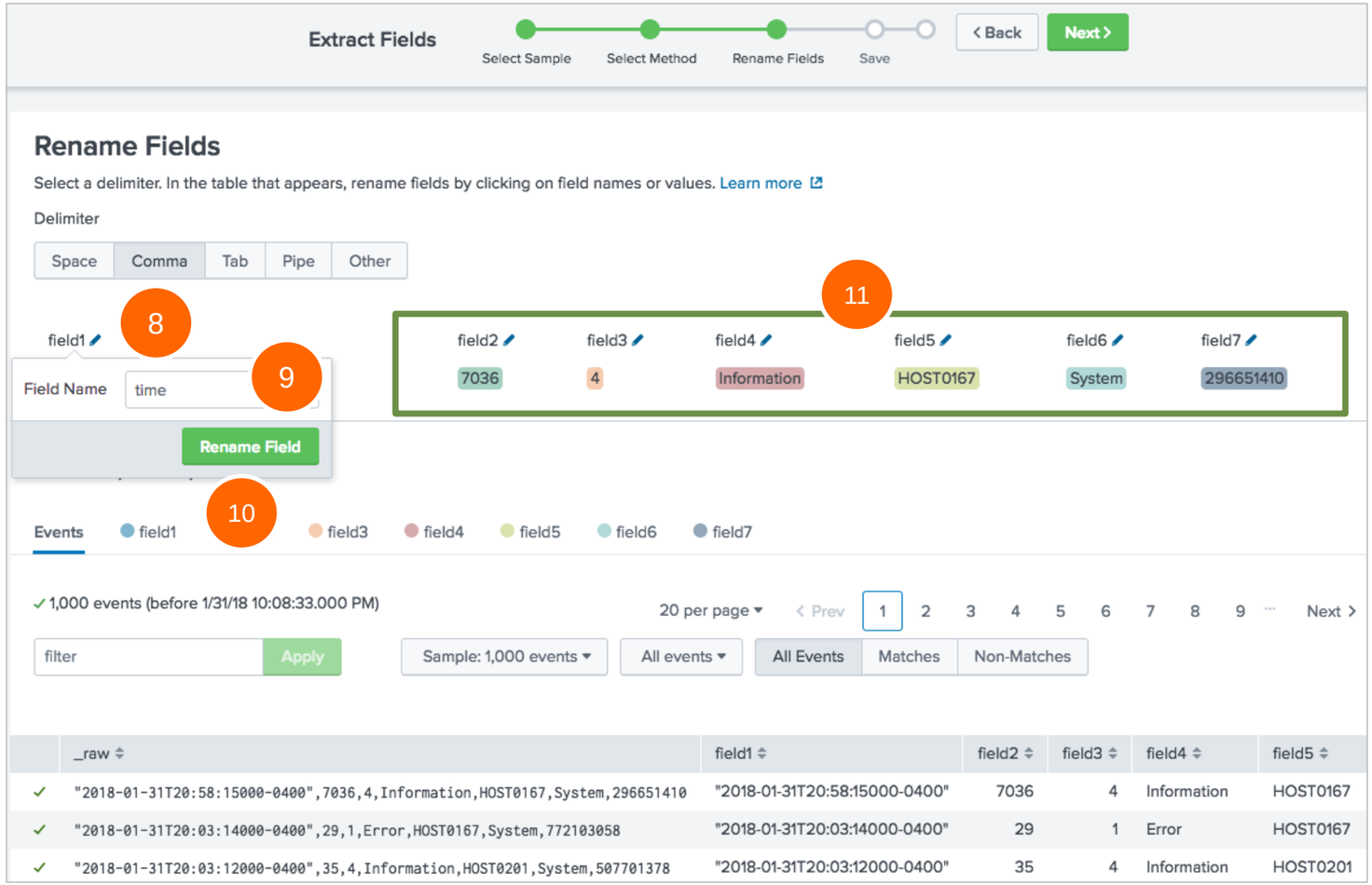


Generated for (v.ms.besetti@jpmchase.com) (C) Splunk Inc, not for distribution

Copyright © 2018 Splunk, Inc. All rights reserved | 14 June 2018

Delimited Field Extractions (Settings) – Rename Field

8. Click the icon next to the default field name
9. Enter a new field name
10. Click Rename Field
11. Repeat these steps for all fields

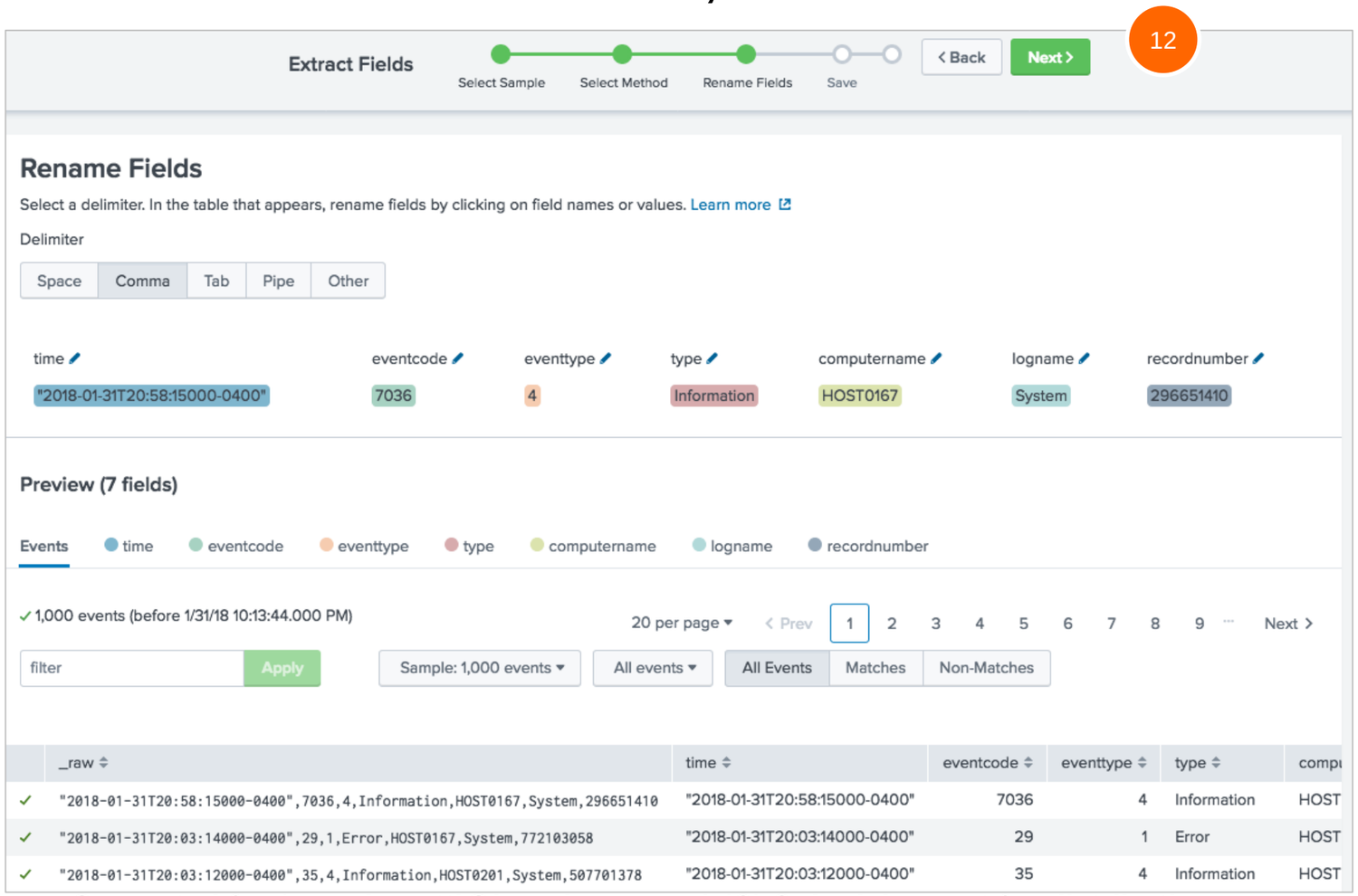


Generated for (v.ms.besetti@jpmchase.com) (C) Splunk Inc, not for distribution

Copyright © 2018 Splunk, Inc. All rights reserved | 14 June 2018

Delimited Field Extractions (Settings) – Rename Field (cont.)

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

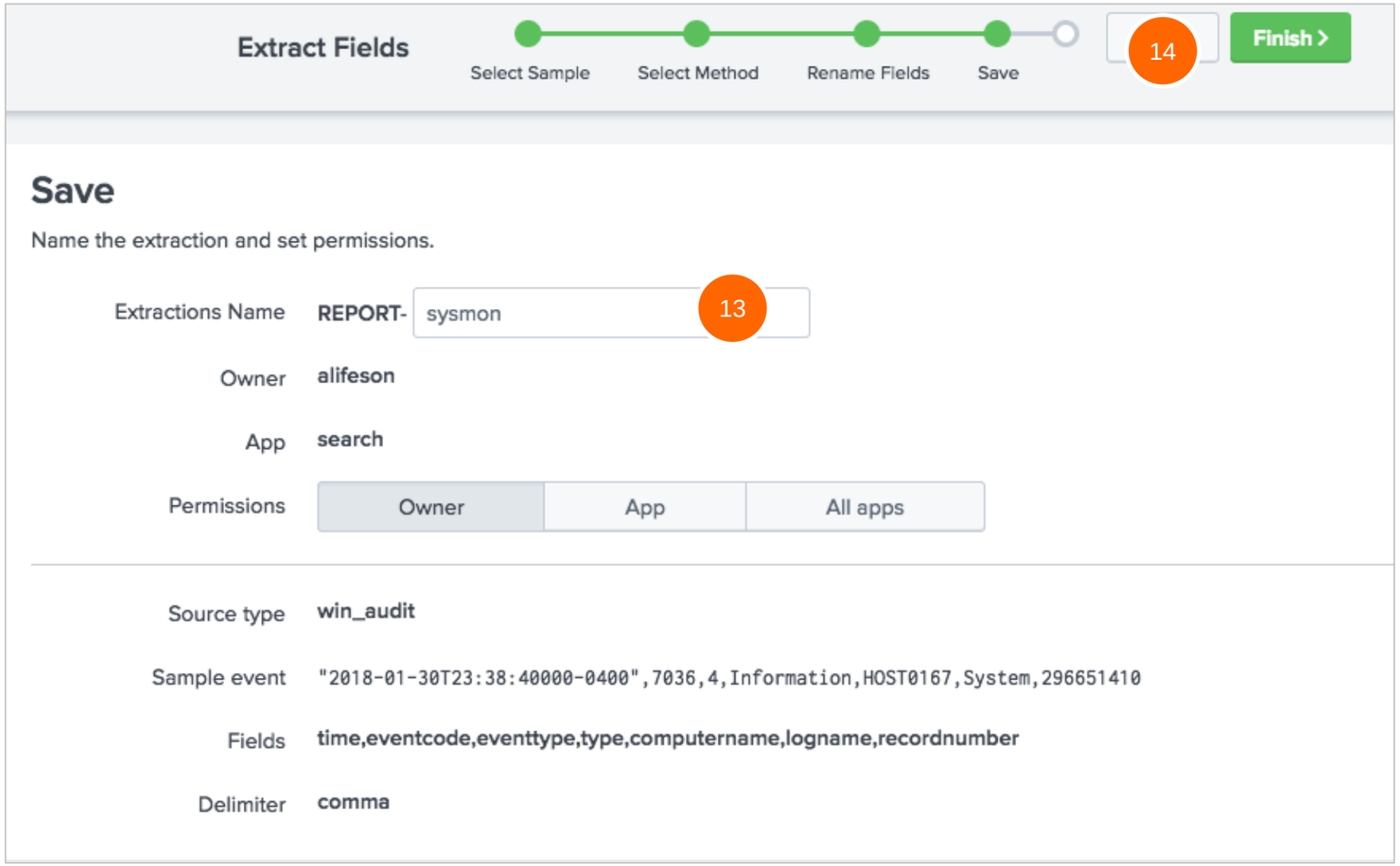


Generated for (v.ms.besetti@jpmchase.com) (C) Splunk Inc, not for distribution

Copyright © 2018 Splunk, Inc. All rights reserved | 14 June 2018

Delimited Field Extractions (Settings) – Save

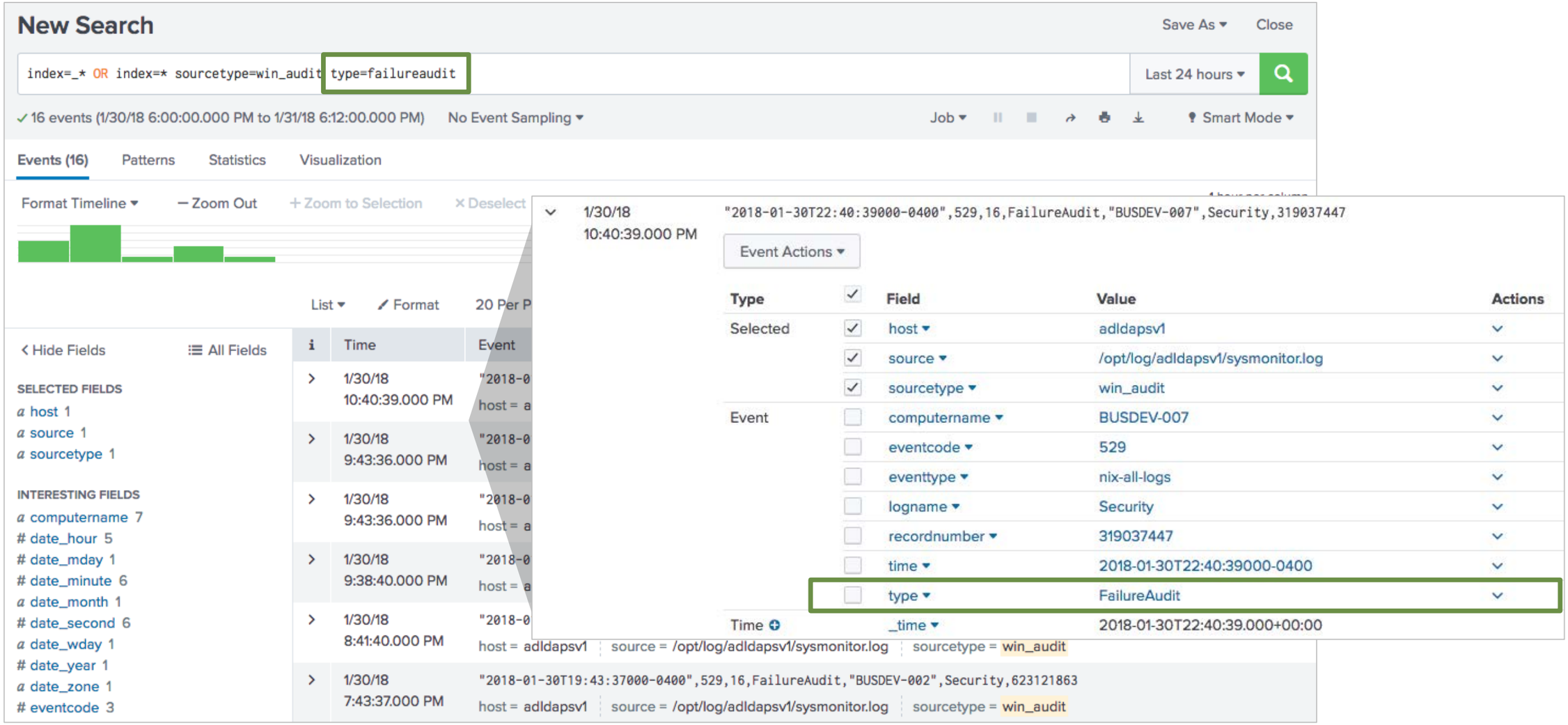
13. Name your extraction
14. Click Finish>



Generated for (v.ms.besetti@jpmchase.com) (C) Splunk Inc, not for distribution

Copyright © 2018 Splunk, Inc. All rights reserved | 14 June 2018

Using a Delimited Field Extraction



Generated for (v.ms.besetti@jpmchase.com) (C) Splunk Inc, not for distribution

Copyright © 2018 Splunk, Inc. All rights reserved | 14 June 2018

Module 9: Creating Field Aliases and Calculated Fields

Generated for (v.ms.besetti@jpmchase.com) (C) Splunk Inc, not for distribution

Copyright © 2018 Splunk, Inc. All rights reserved

Module Objectives

- Create and use field aliases
- Create calculated fields

Generated for (v.ms.besetti@jpmchase.com) (C) Splunk Inc, not for distribution

Copyright © 2018 Splunk, Inc. All rights reserved

Field Aliases

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

Generated for (v.ms.besetti@jpmchase.com) (C) Splunk Inc, not for distribution

Copyright © 2018 Splunk, Inc. All rights reserved | 14 June 2018

Field Alias Example

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

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

sourcetype=cisco_firewall

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

sourcetype=cisco_wsa_squid

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

sourcetype=winauthentication_security

Generated for (v.ms.besetti@jpmchase.com) (C) Splunk Inc, not for distribution

Copyright © 2018 Splunk, Inc. All rights reserved | 14 June 2018

Creating a Field Alias

Settings > Fields > Field Aliases > New Field Alias

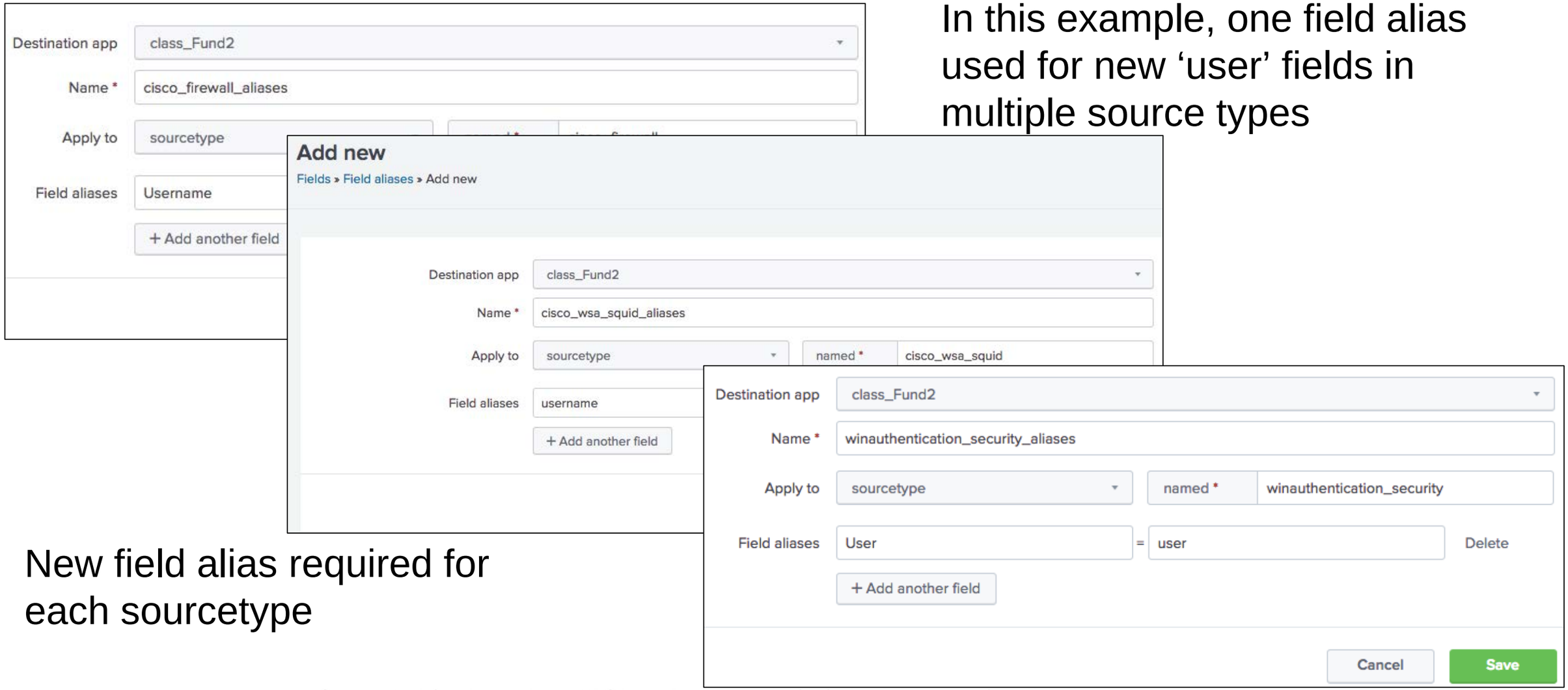
1. Select the app associated with the field alias
2. Enter a Name for the field alias
3. Apply the field alias to a default field:
   - Host
   - Source
   - Sourcetype
4. Enter the name for the existing field and the new alias

Generated for (v.ms.besetti@jpmchase.com) (C) Splunk Inc, not for distribution

Copyright © 2018 Splunk, Inc. All rights reserved | 14 June 2018

Creating a Field Alias (cont.)

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



New field alias required for each sourcetype

Generated for (v.ms.besetti@jpmchase.com) (C) Splunk Inc, not for distribution

Copyright © 2018 Splunk, Inc. All rights reserved | 14 June 2018

Testing the Field Alias

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

New Search

Save As Close

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

Job Smart Mode

Events (96) Patterns Statistics Visualization

Format Timeline Zoom Out

Hide Fields All Fields

SELECTED FIELDS

a host 2

a source 3

a sourcetype 3

INTERESTING FIELDS

a action 3

a bcg_ip 1

a bcg_workstation 1

bytes_in 75

a c_ip 57

sourcetype

3 Values, 100% of events

Selected Yes No

Reports

Top values Top values by time Rare values

Events with this field

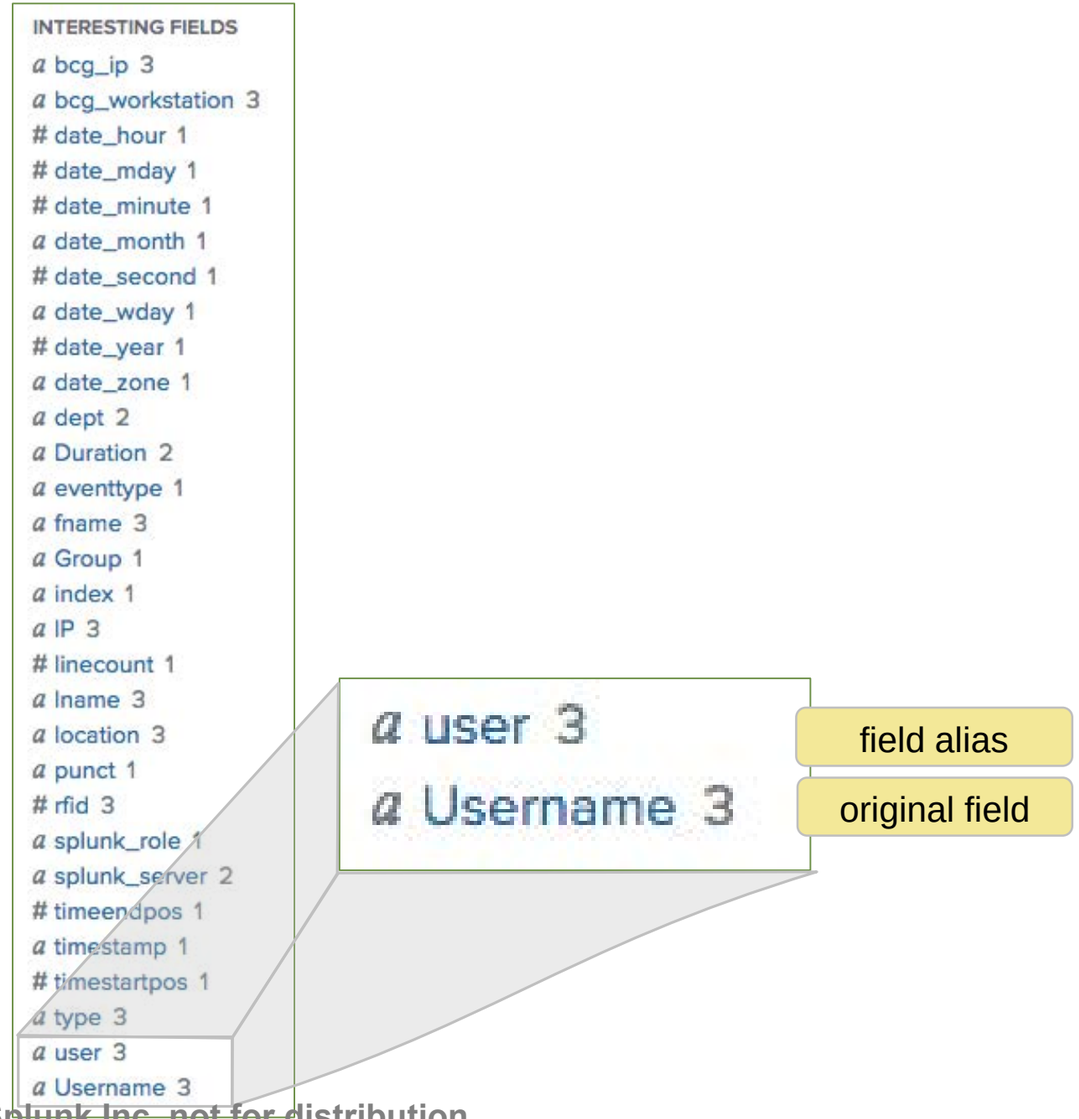
Values	Count	%
cisco_wsa_squid	78	81.25%
winauthentication_security	16	16.667%
cisco_firewall	2	2.083%

Generated for (v.ms.besetti@jpmchase.com) (C) Splunk Inc, not for distribution

Copyright © 2018 Splunk, Inc. All rights reserved | 14 June 2018

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 (v.ms.besetti@jpmchase.com) (C) Splunk Inc, not for distribution

Copyright © 2018 Splunk, Inc. All rights reserved | 14 June 2018

Field Aliases and Lookups

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

employees.csv

```
rfid,fname,lname,user,email,dept,location,ip
108423575302,Alle,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,iking@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 (v.ms.besetti@jpmchase.com) (C) Splunk Inc, not for distribution

Copyright © 2018 Splunk, Inc. All rights reserved | 14 June 2018

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

Save As ▾ Close

```
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 ▾ Smart Mode ▾

Events Patterns **Statistics (5)** Visualization

100 Per Page ▾ Format Preview ▾

usage	Megabytes
Violation	0.01721763610839843 80
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.

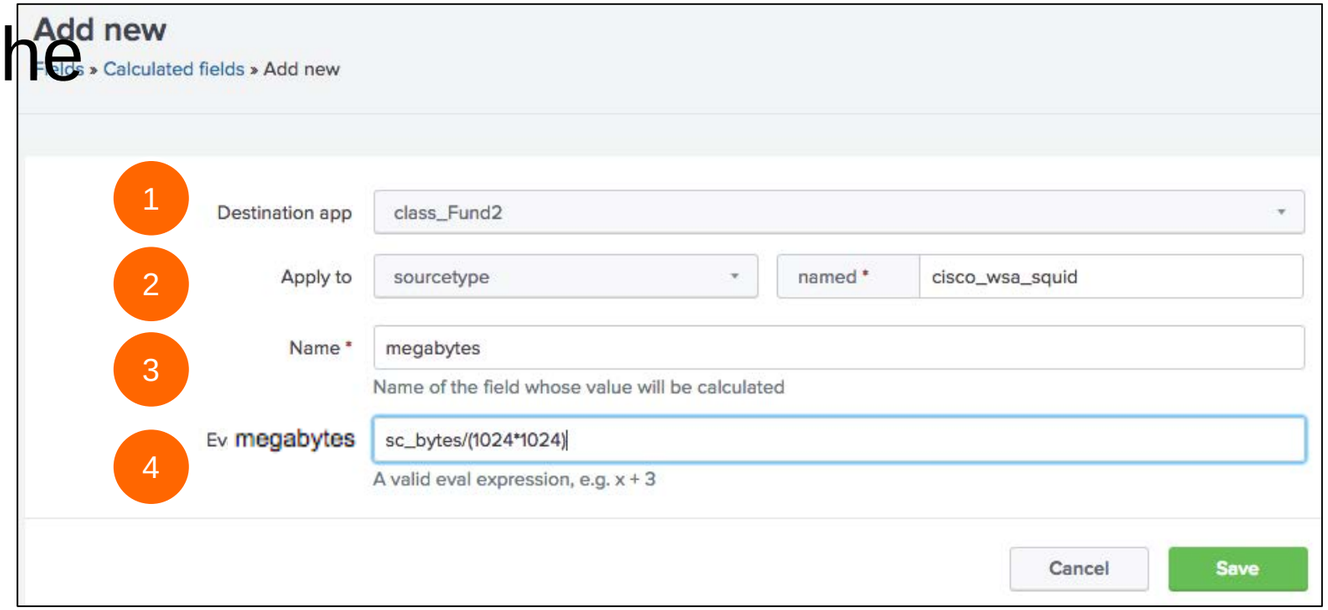
Generated for (v.ms.besetti@jpmchase.com) (C) Splunk Inc, not for distribution

Copyright © 2018 Splunk, Inc. All rights reserved | 14 June 2018

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

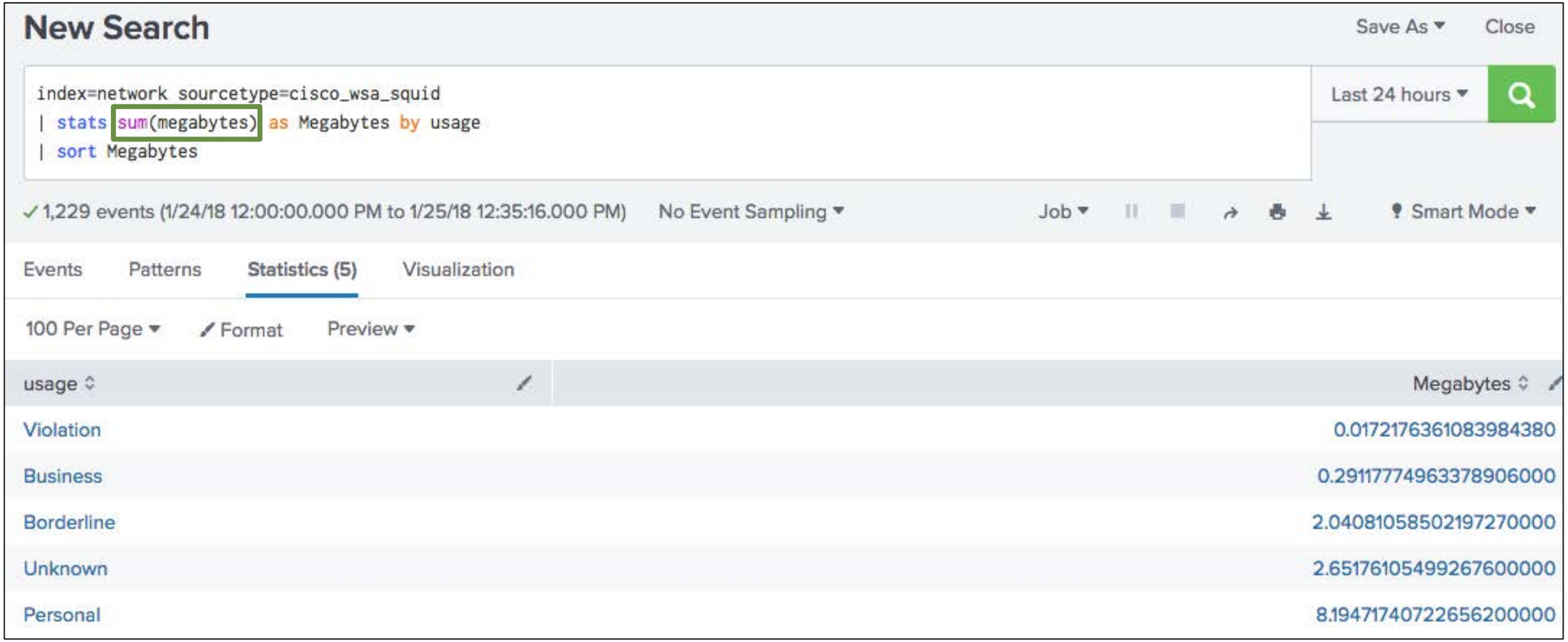


Generated for (v.ms.besetti@jpmchase.com) (C) Splunk Inc, not for distribution

Copyright © 2018 Splunk, Inc. All rights reserved | 14 June 2018

Using a Calculated Field

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



Generated for (v.ms.besetti@jpmchase.com) (C) Splunk Inc, not for distribution

Copyright © 2018 Splunk, Inc. All rights reserved | 14 June 2018

Module 10: Working with Tags and Event Types

Generated for (v.ms.besetti@jpmchase.com) (C) Splunk Inc, not for distribution

Copyright © 2018 Splunk, Inc. All rights reserved

Module Objectives

- Create and use tags
- Describe event types and their uses
- Create an event type

Generated for (v.ms.besetti@jpmchase.com) (C) Splunk Inc, not for distribution

Copyright © 2018 Splunk, Inc. All rights reserved | 14 June 2018

Describing Tags

- Tags are like nicknames 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

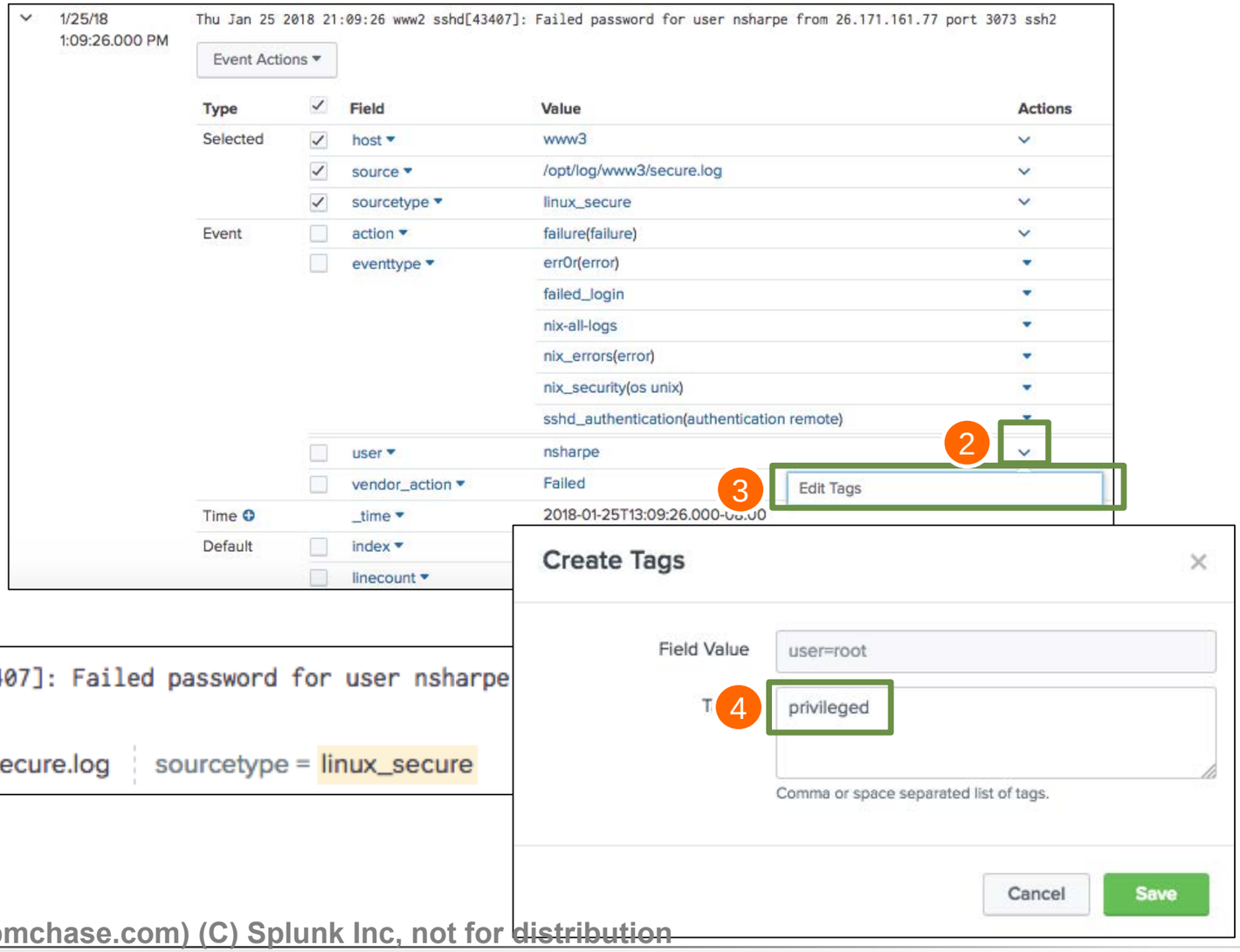
Generated for (v.ms.besetti@jpmchase.com) (C) Splunk Inc, not for distribution

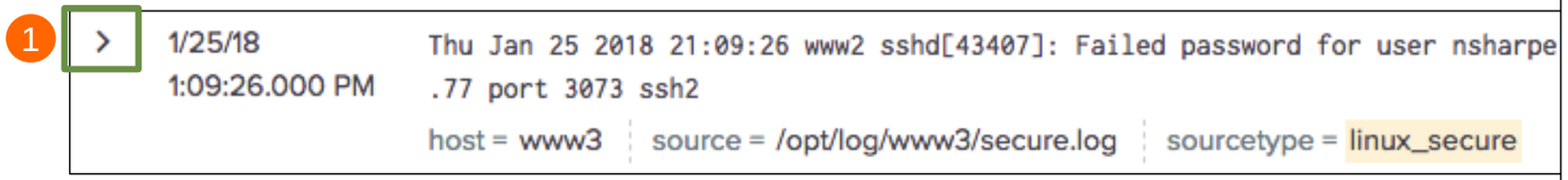
Copyright © 2018 Splunk, Inc. All rights reserved | 14 June 2018

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



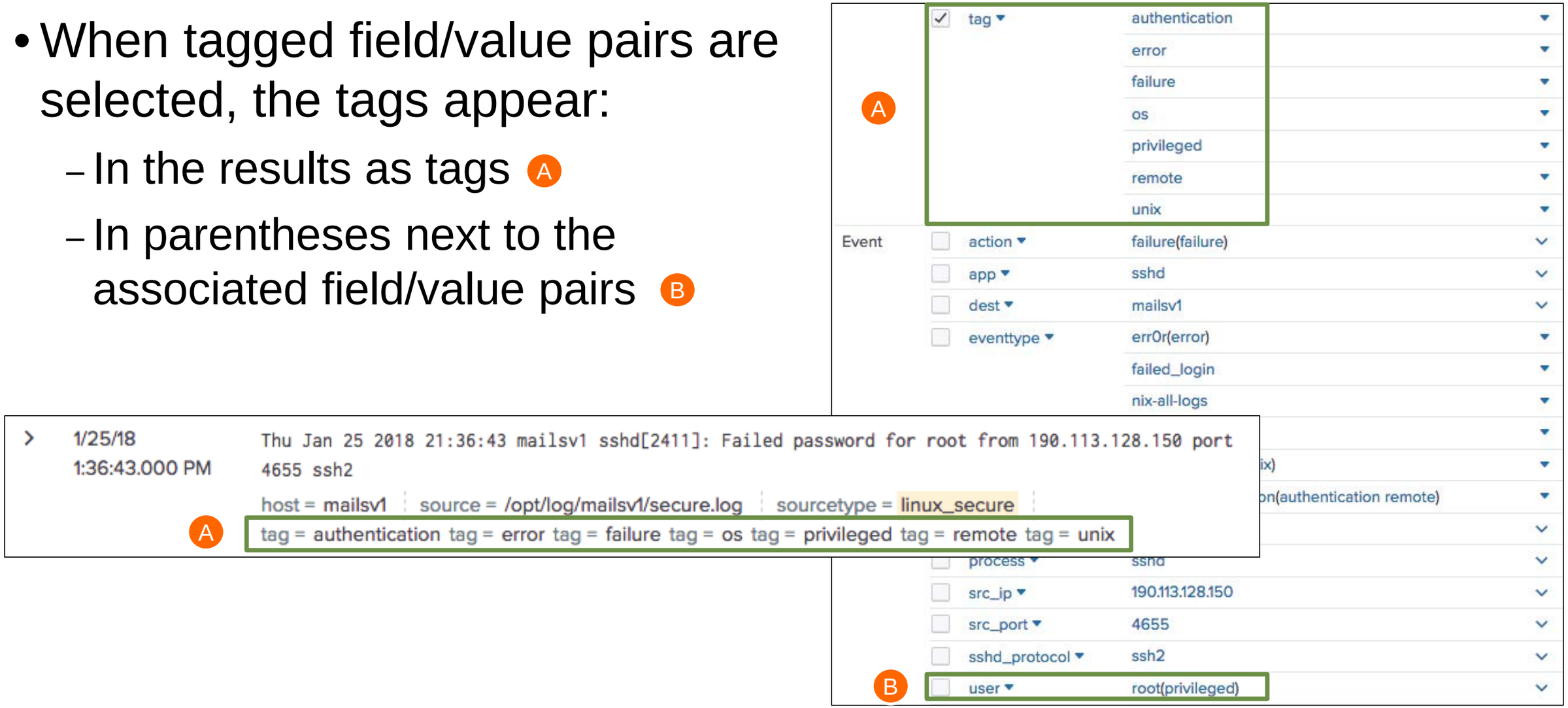


Generated for (v.ms.besetti@jpmchase.com) (C) Splunk Inc, not for distribution

Copyright © 2018 Splunk, Inc. All rights reserved | 14 June 2018

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

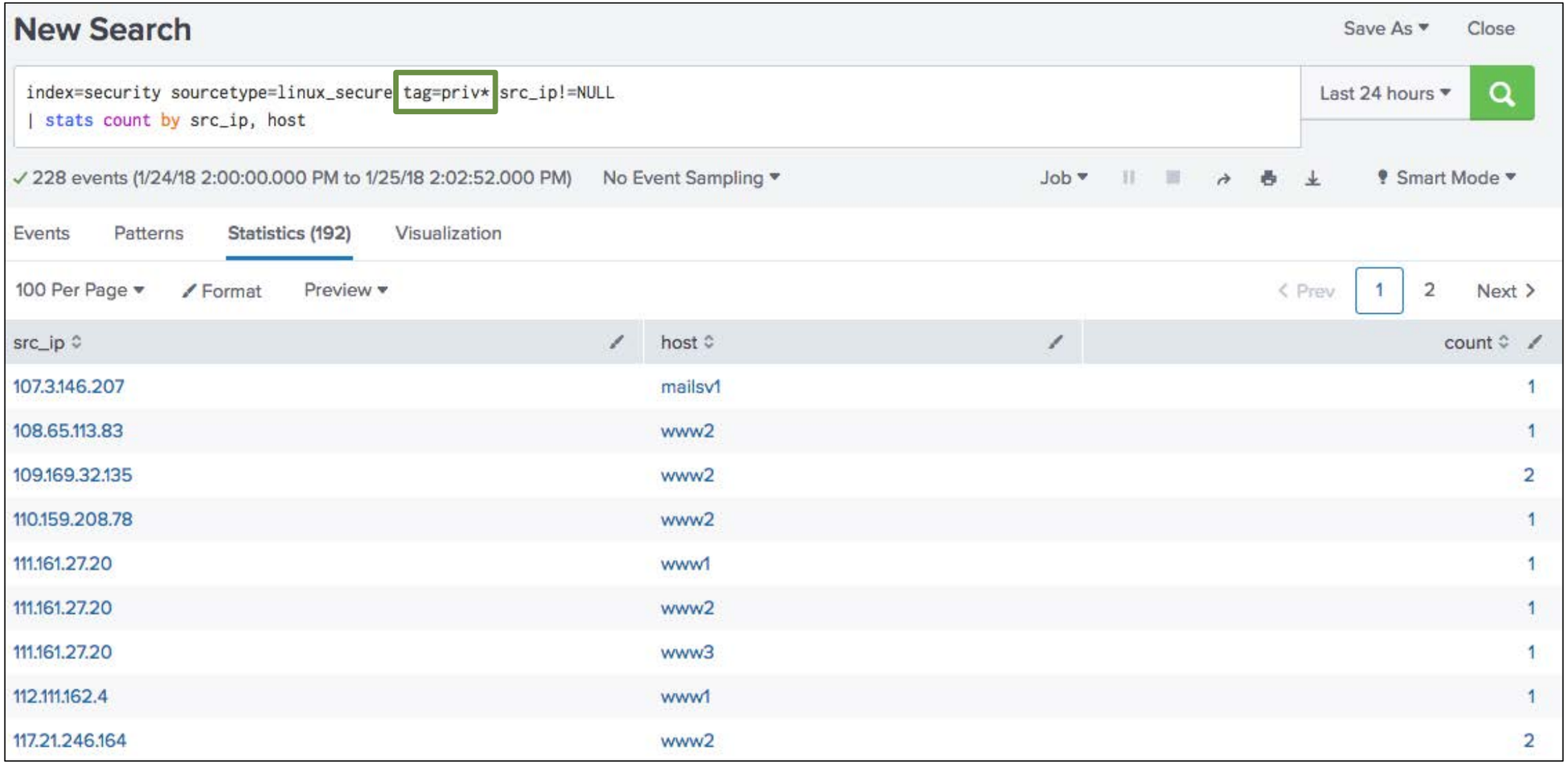


Generated for (v.ms.besetti@jpmchase.com) (C) Splunk Inc, not for distribution

Copyright © 2018 Splunk, Inc. All rights reserved | 14 June 2018

Using Tags

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



Generated for (v.ms.besetti@jpmchase.com) (C) Splunk Inc, not for distribution

Copyright © 2018 Splunk, Inc. All rights reserved | 14 June 2018

Searching for Tags

- To search for a tag associated with a value:
  - tag=<tagname>

```
tag=privileged
```

- To search for a tag associated with a value on a specific field:
  - tag::<field>=<tagname>

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

- To search for a tag using a partial field value:
  - Use (*) wildcard

```
tag=p*
```

Note

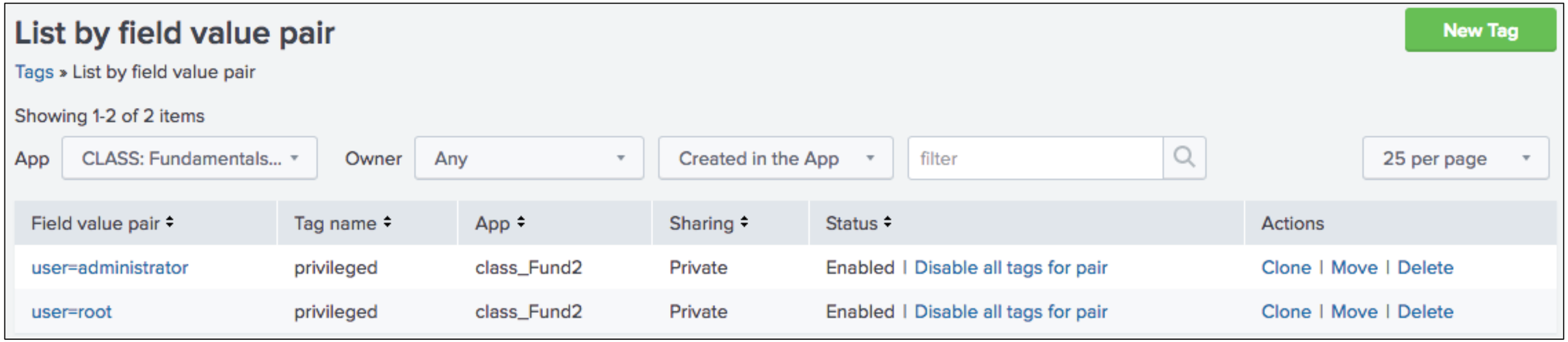
Tag names are case sensitive.

Generated for (v.ms.besetti@jpmchase.com) (C) Splunk Inc, not for distribution

Copyright © 2018 Splunk, Inc. All rights reserved | 14 June 2018

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

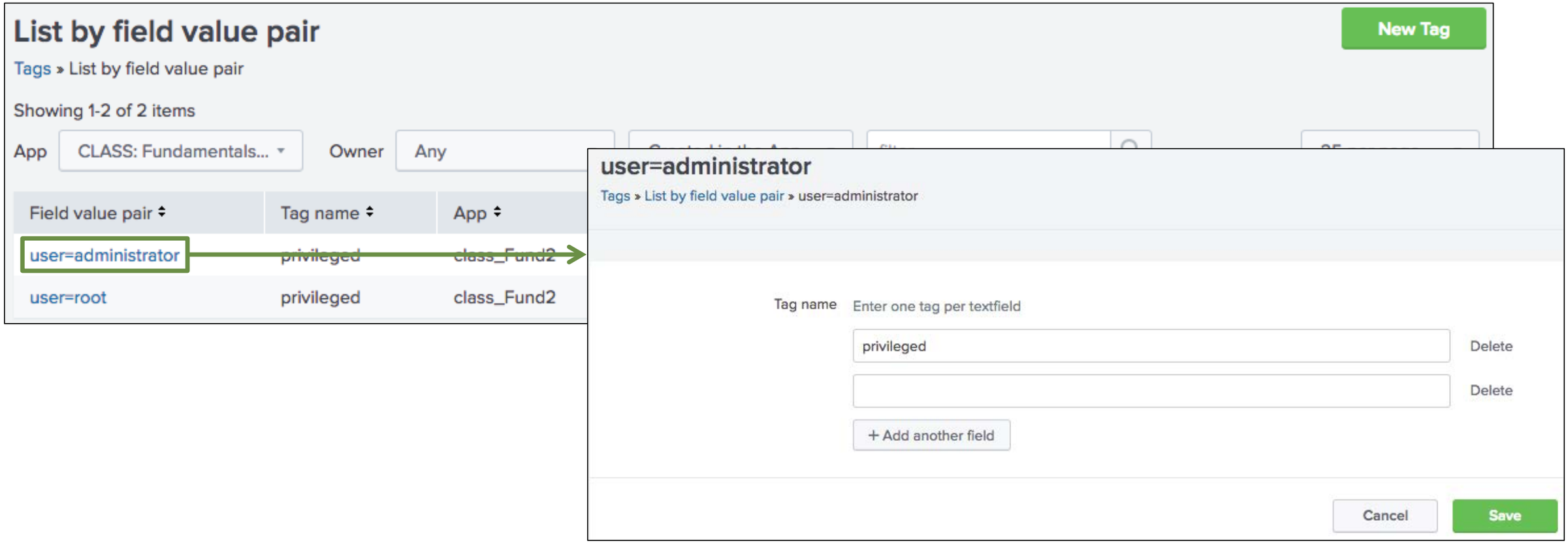


Generated for (v.ms.besetti@jpmchase.com) (C) Splunk Inc, not for distribution

Copyright © 2018 Splunk, Inc. All rights reserved | 14 June 2018

Adding/Changing the Tag Name

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

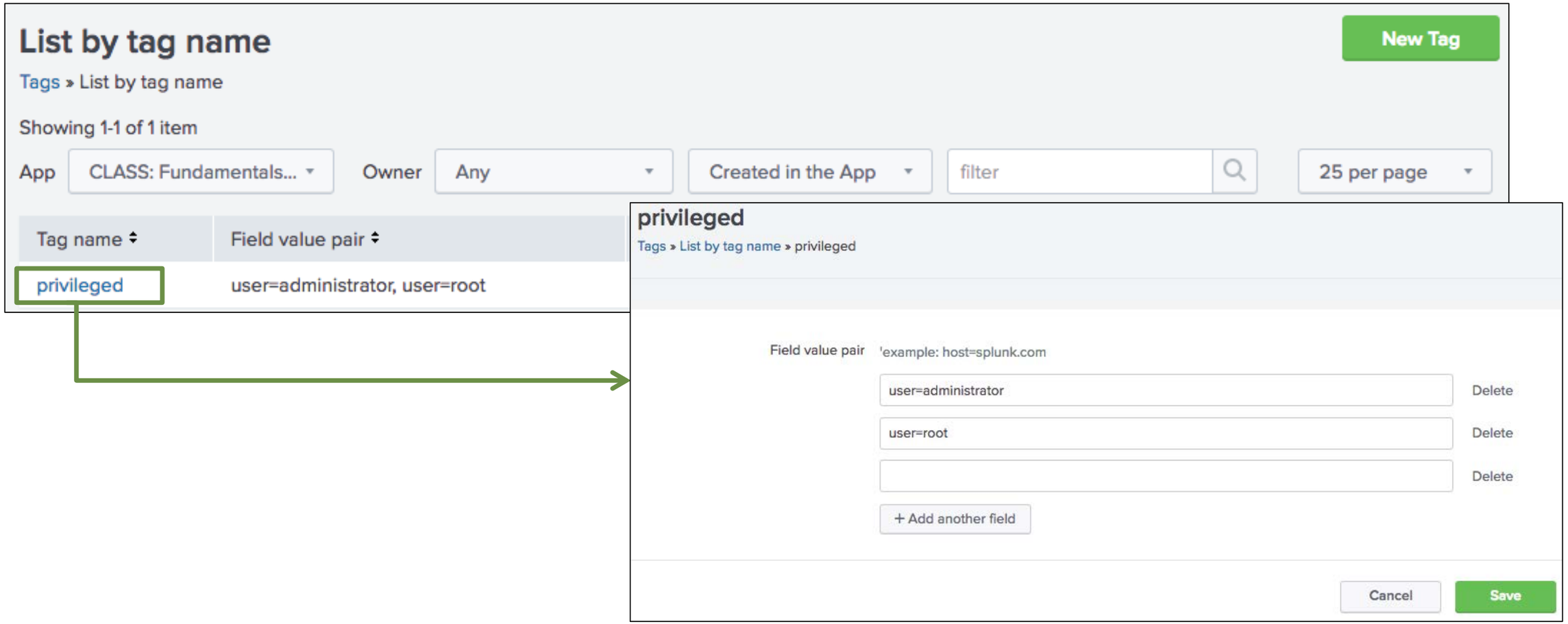


Generated for (v.ms.besetti@jpmchase.com) (C) Splunk Inc, not for distribution

Copyright © 2018 Splunk, Inc. All rights reserved | 14 June 2018

Adding/Changing the Field Value Pair

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



Generated for (v.ms.besetti@jpmchase.com) (C) Splunk Inc, not for distribution

Copyright © 2018 Splunk, Inc. All rights reserved | 14 June 2018

Describing Event Types

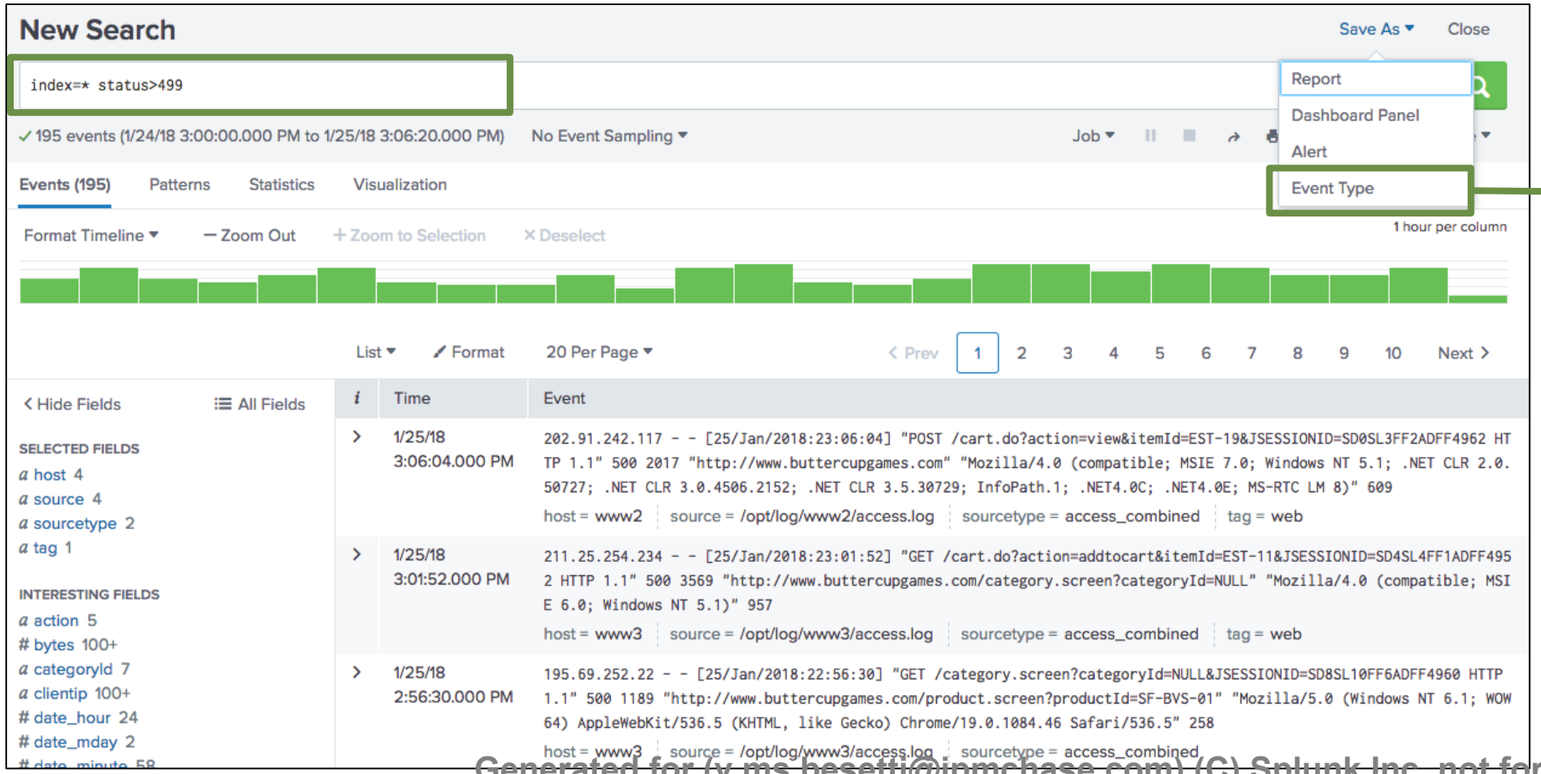
- A method of categorizing events based on a search
- A useful method for institutional knowledge capturing and sharing
- Can be tagged to group similar types of events

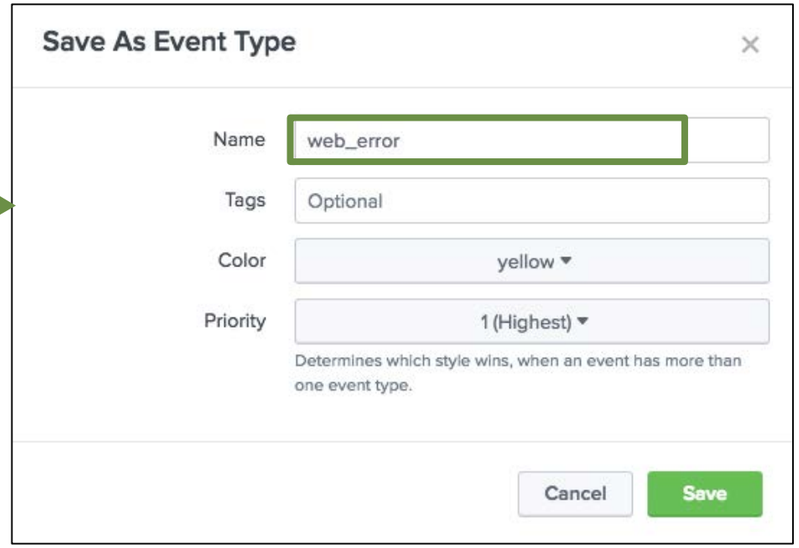
Generated for (v.ms.besetti@jpmchase.com) (C) Splunk Inc, not for distribution

Copyright © 2018 Splunk, Inc. All rights reserved | 14 June 2018

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)





Note

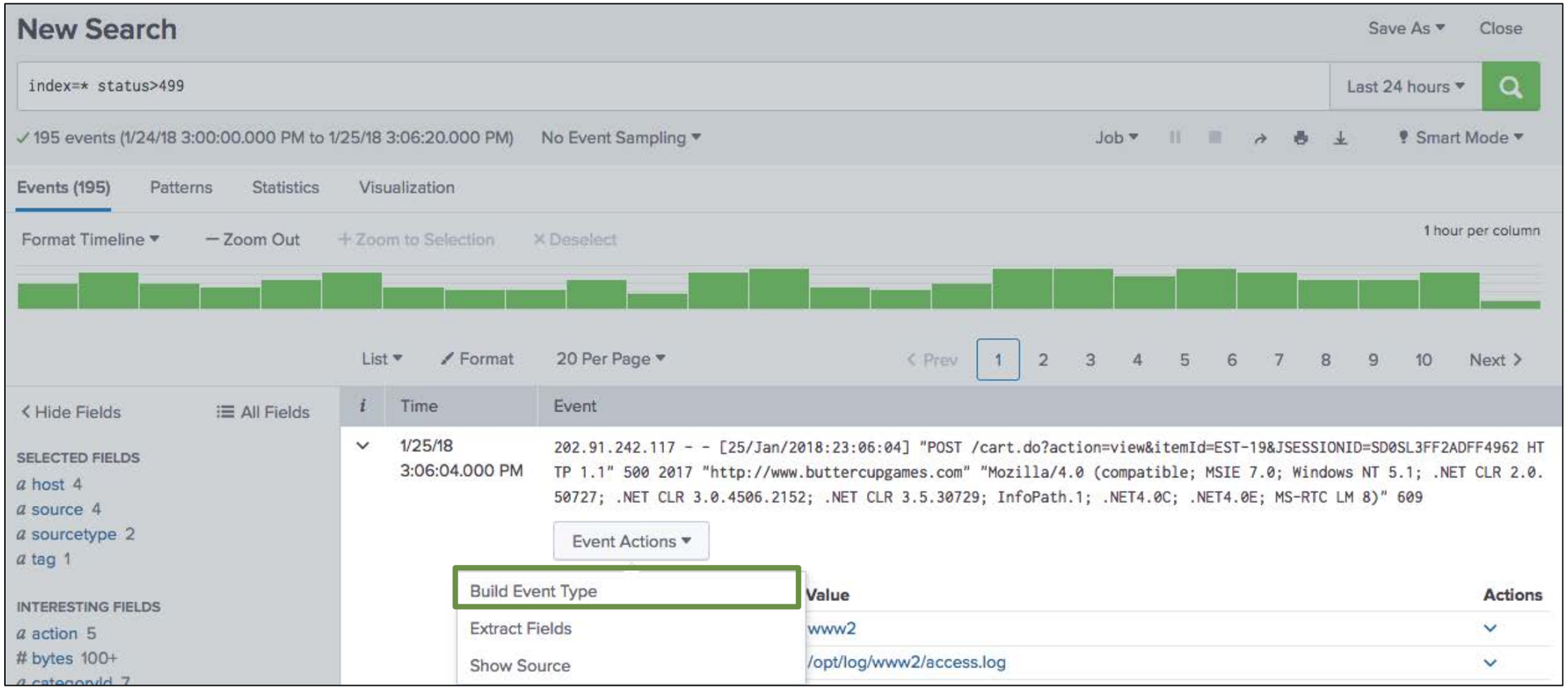
Must be a basic search (cannot contain pipes or subsearches).

Generated for (v.ms.besetti@jpmchase.com) (C) Splunk Inc, not for distribution

Copyright © 2018 Splunk, Inc. All rights reserved | 14 June 2018

Using the Event Type Builder

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

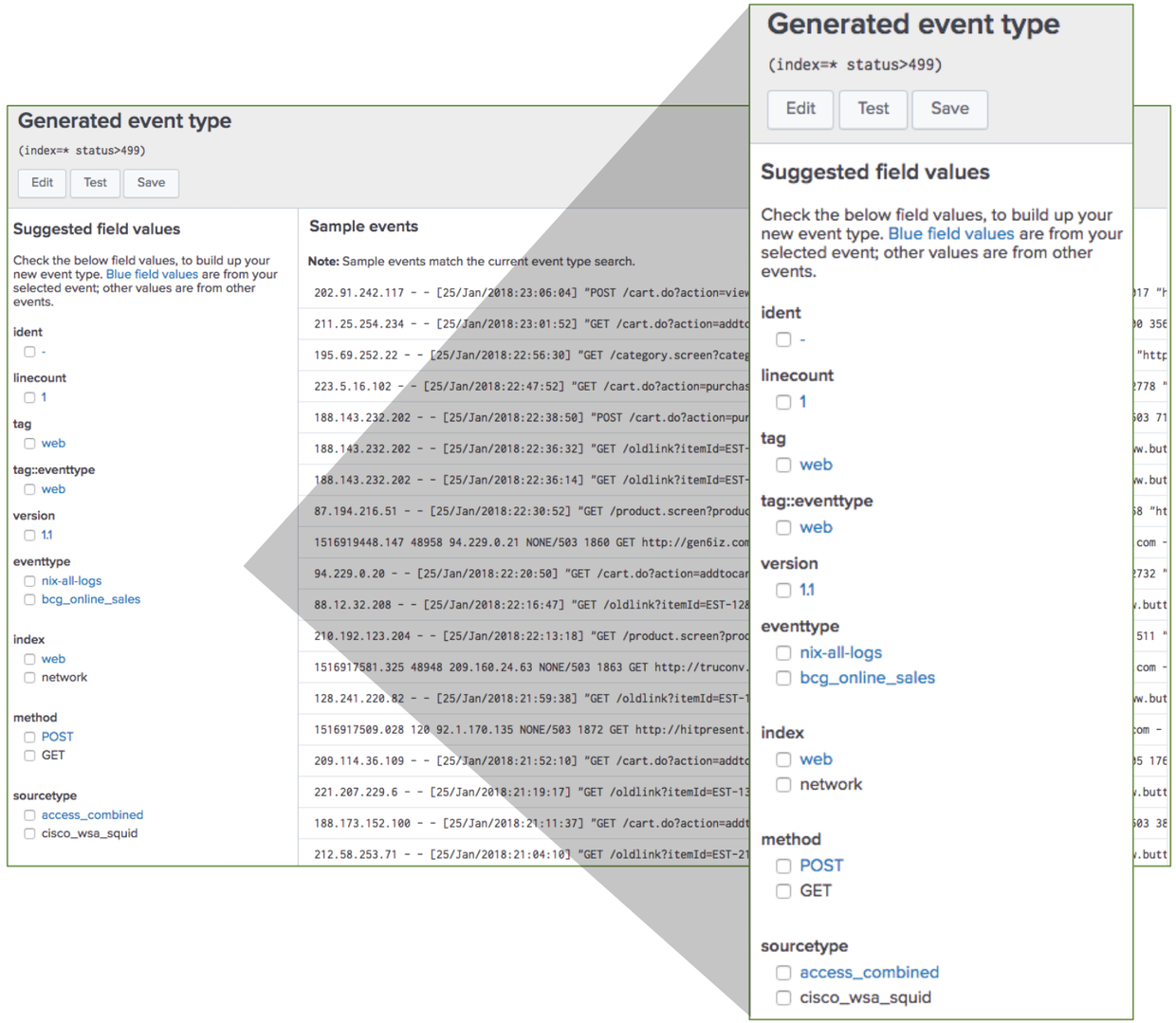


Generated for (v.ms.besetti@jpmchase.com) (C) Splunk Inc, not for distribution

Copyright © 2018 Splunk, Inc. All rights reserved | 14 June 2018

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

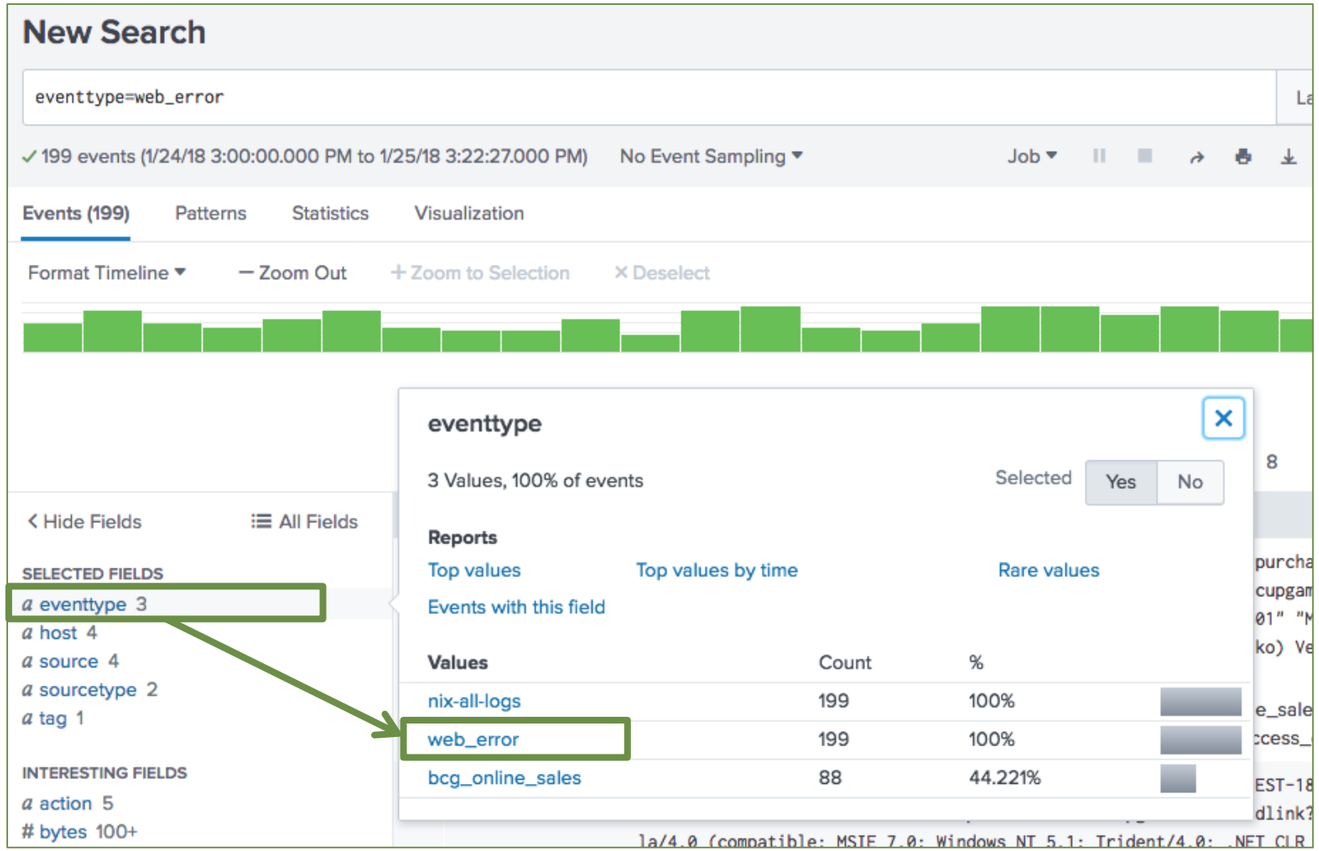


Generated for (v.ms.besetti@jpmchase.com) (C) Splunk Inc, not for distribution

Copyright © 2018 Splunk, Inc. All rights reserved | 14 June 2018

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



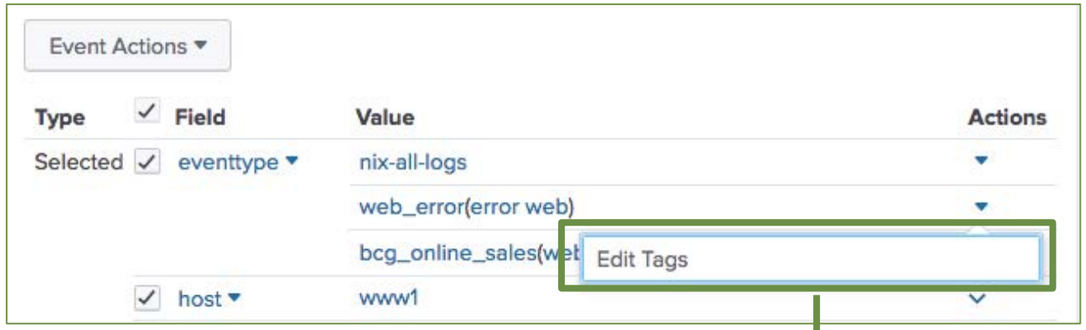
Generated for (v.ms.besetti@jpmchase.com) (C) Splunk Inc, not for distribution

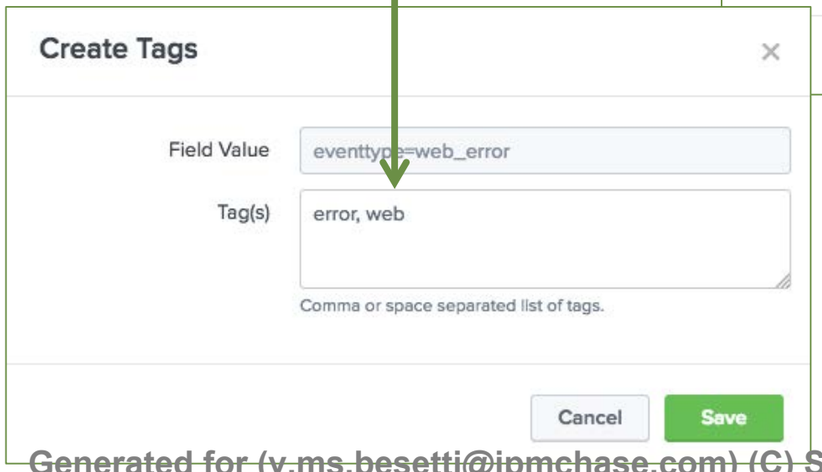
Copyright © 2018 Splunk, Inc. All rights reserved | 14 June 2018

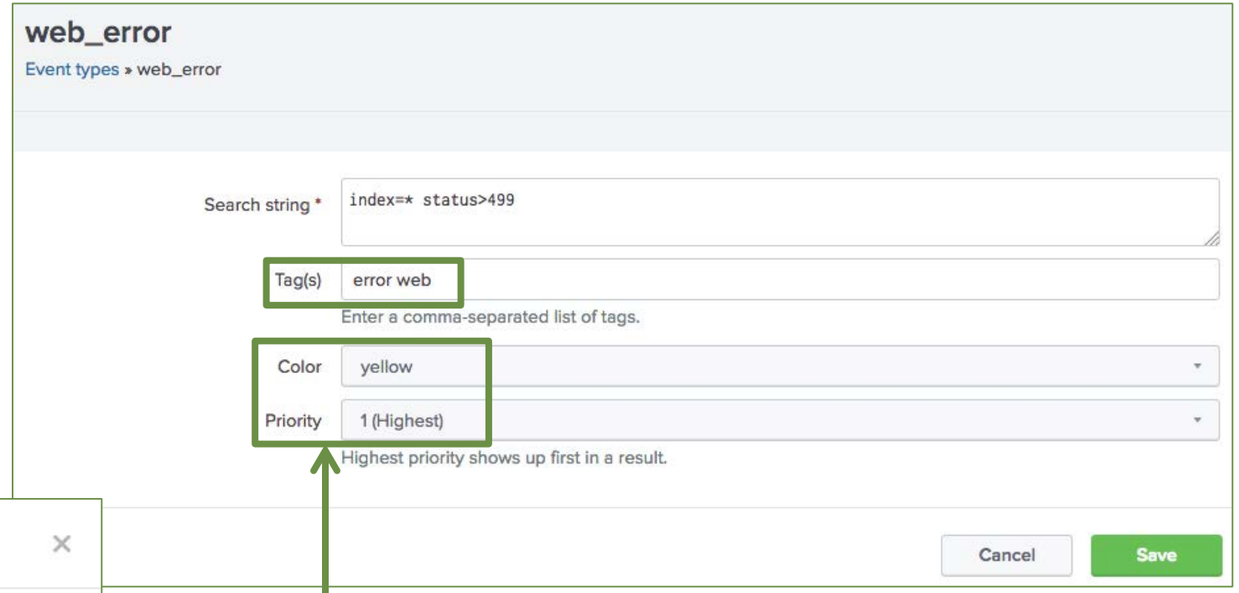
Tagging Event Types

You can tag event types two ways:

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







Priority determines which event type color displays for an event

Generated for (v.ms.besetti@jpmchase.com) (C) Splunk Inc, not for distribution

Copyright © 2018 Splunk, Inc. All rights reserved | 14 June 2018

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

Generated for (v.ms.besetti@jpmchase.com) (C) Splunk Inc, not for distribution

Copyright © 2018 Splunk, Inc. All rights reserved

Module 11: Creating and Using Macros

Generated for (v.ms.besetti@jpmchase.com) (C) Splunk Inc, not for distribution

Copyright © 2018 Splunk, Inc. All rights reserved

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

Generated for (v.ms.besetti@jpmchase.com) (C) Splunk Inc, not for distribution

Copyright © 2018 Splunk, Inc. All rights reserved | 14 June 2018

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

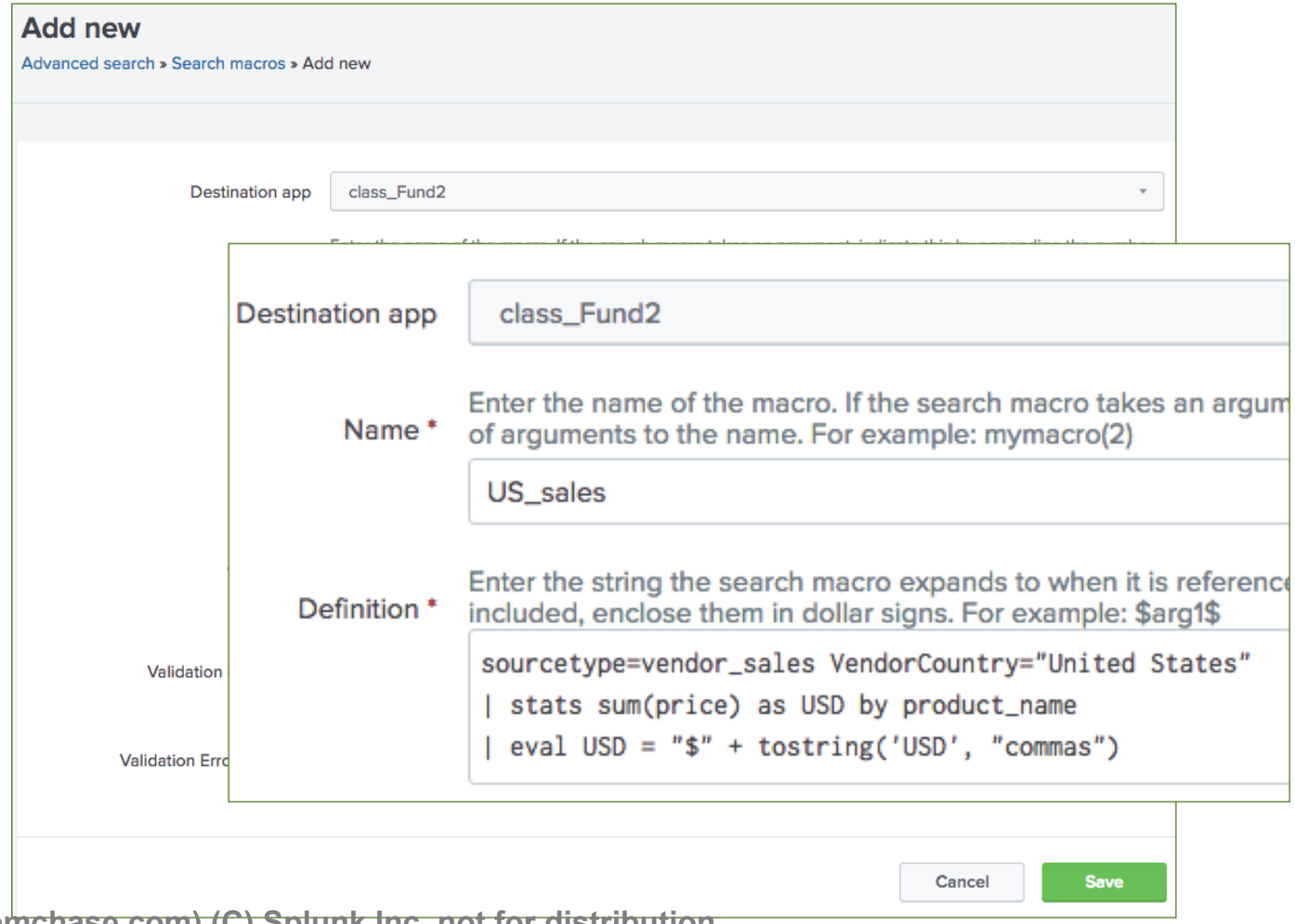
Generated for (v.ms.besetti@jpmchase.com) (C) Splunk Inc, not for distribution

Copyright © 2018 Splunk, Inc. All rights reserved | 14 June 2018

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



Note

Test your search string before saving your macro. You can check the contents of your macro with keyboard shortcuts (Command-Shift-E on Mac OS or Control-Shift-E on Linux or Windows) from the Search bar in the Search page.

Generated for (v.ms.besetti@jpmchase.com) (C) Splunk Inc, not for distribution

Copyright © 2018 Splunk, Inc. All rights reserved | 14 June 2018

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

```
`US_sales`
```

```
sourcetype=vendor_sales
VendorCountry="United States"
| stats sum(price) as USD by product_name
| eval USD = "$" + tostring('USD’,"commas")
```

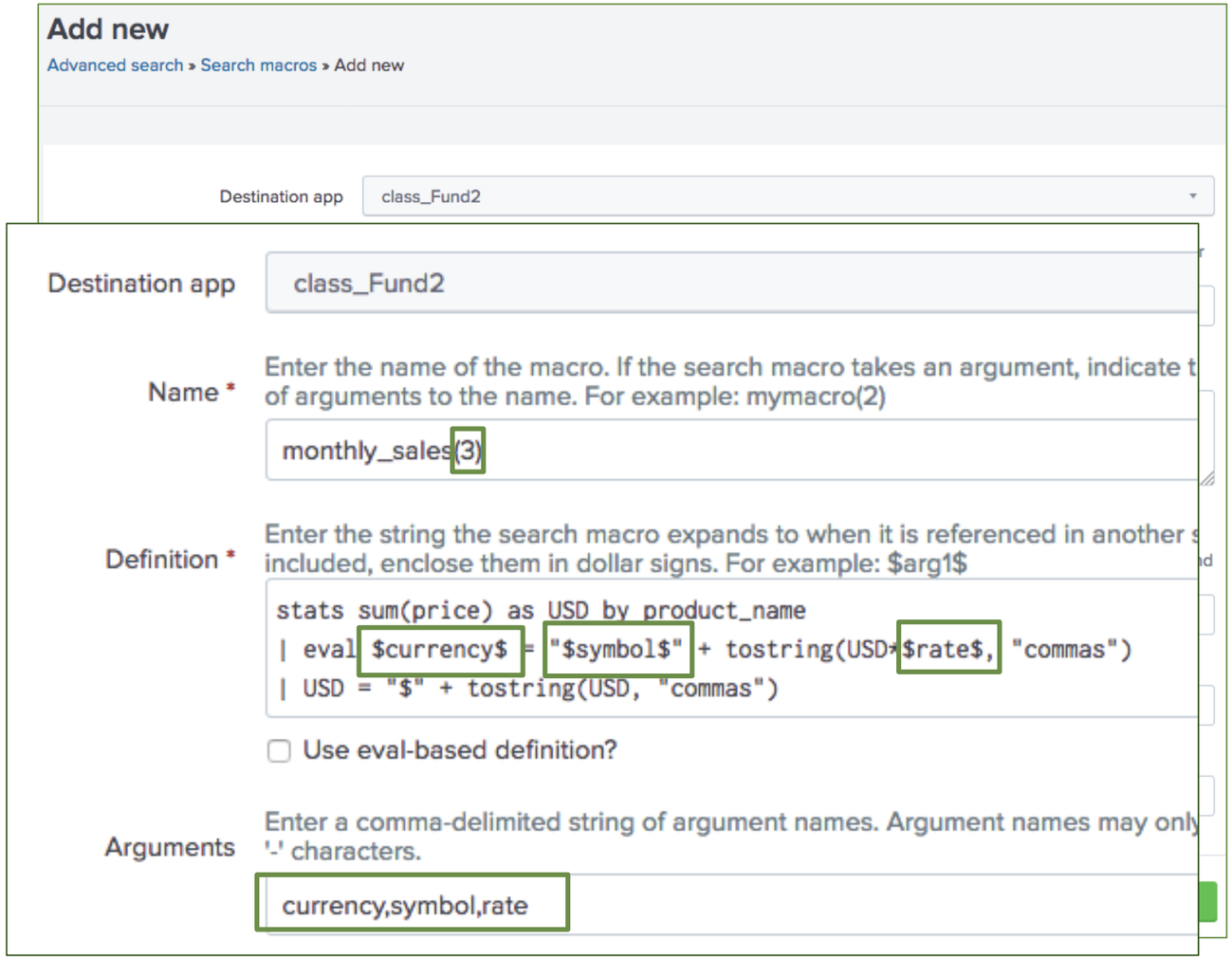
product_name	USD
Benign Space Debris	$349.86
Curling 2014	$779.61
Dream Crusher	$2,959.26
Final Sequel	$824.67
Fire Resistance Suit of Provolone	$207.48
Holy Blade of Gouda	$173.71
Manganiello Bros.	$1,879.53
Manganiello Bros. Tee	$599.40
Mediocre Kingdoms	$799.68

Generated for (v.ms.besetti@jpmchase.com) (C) Splunk Inc, not for distribution

Copyright © 2018 Splunk, Inc. All rights reserved | 14 June 2018

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



Generated for (v.ms.besetti@jpmchase.com) (C) Splunk Inc, not for distribution

Copyright © 2018 Splunk, Inc. All rights reserved | 14 June 2018

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

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

```
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")
| eval 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 (v.ms.besetti@jpmchase.com) (C) Splunk Inc, not for distribution

Copyright © 2018 Splunk, Inc. All rights reserved | 14 June 2018

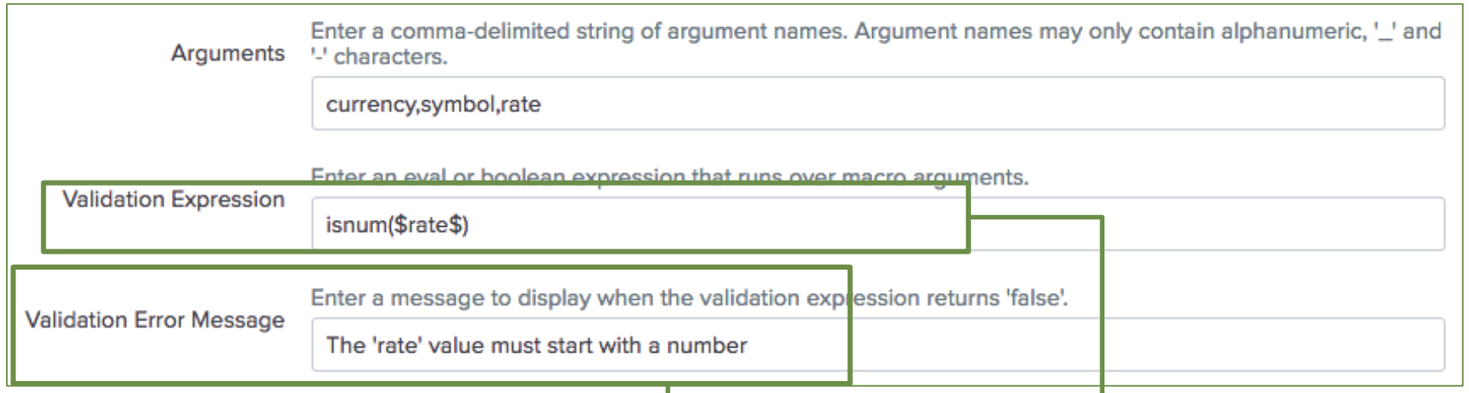
Validating Macros

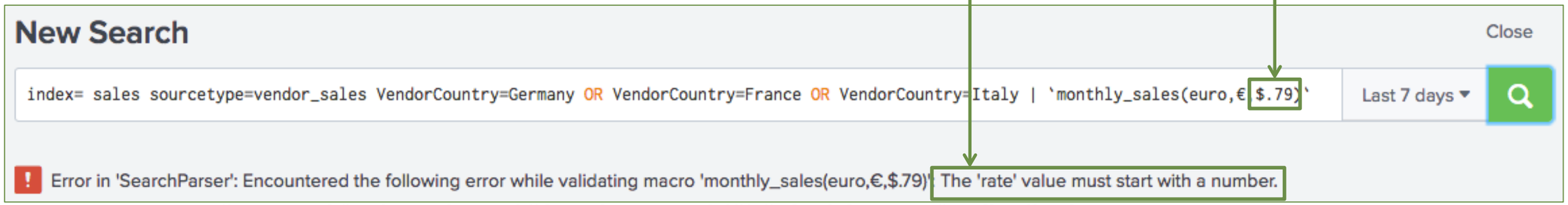
You can validate argument values in your macro

- 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

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.





Generated for (v.ms.besetti@jpmchase.com) (C) Splunk Inc, not for distribution

Copyright © 2018 Splunk, Inc. All rights reserved | 14 June 2018

Module 12: Creating and Using Workflow Actions

Generated for (v.ms.besetti@jpmchase.com) (C) Splunk Inc, not for distribution

Copyright © 2018 Splunk, Inc. All rights reserved

Module Objectives

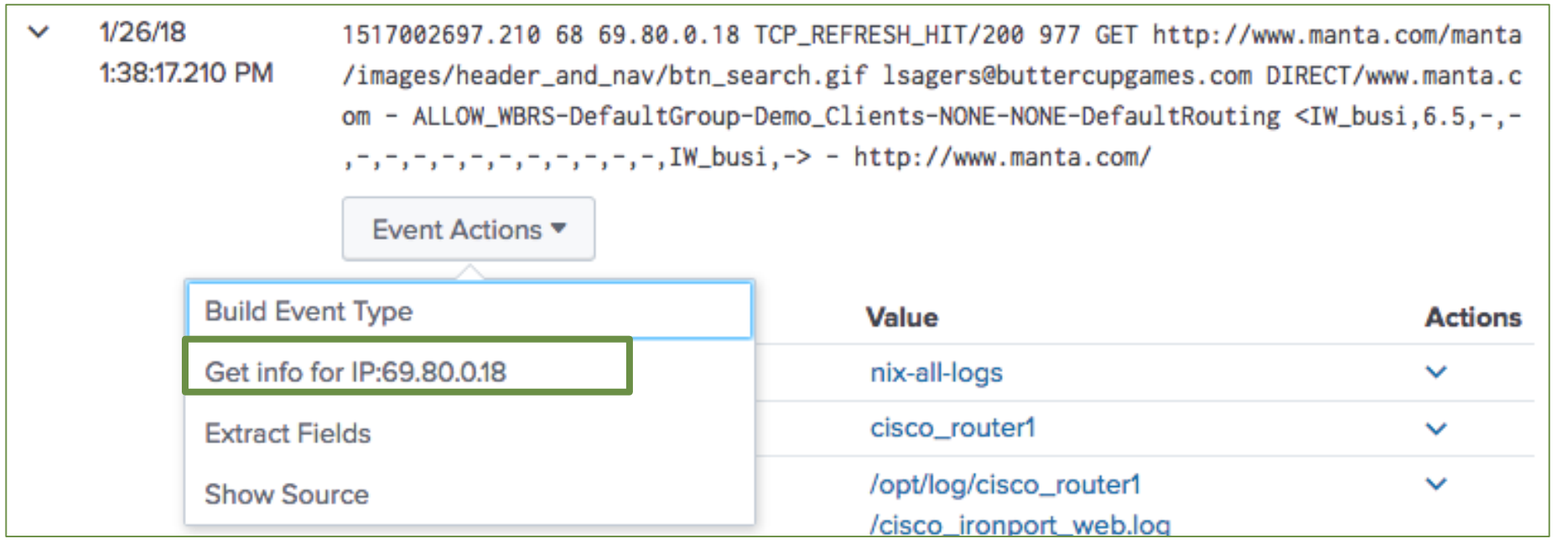
- Create a GET workflow action
- Create a POST workflow action
- Create a Search workflow action

Generated for (v.ms.besetti@jpmchase.com) (C) Splunk Inc, not for distribution

Copyright © 2018 Splunk, Inc. All rights reserved | 14 June 2018

What are Workflow Actions?

- Execute workflow actions from an event 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



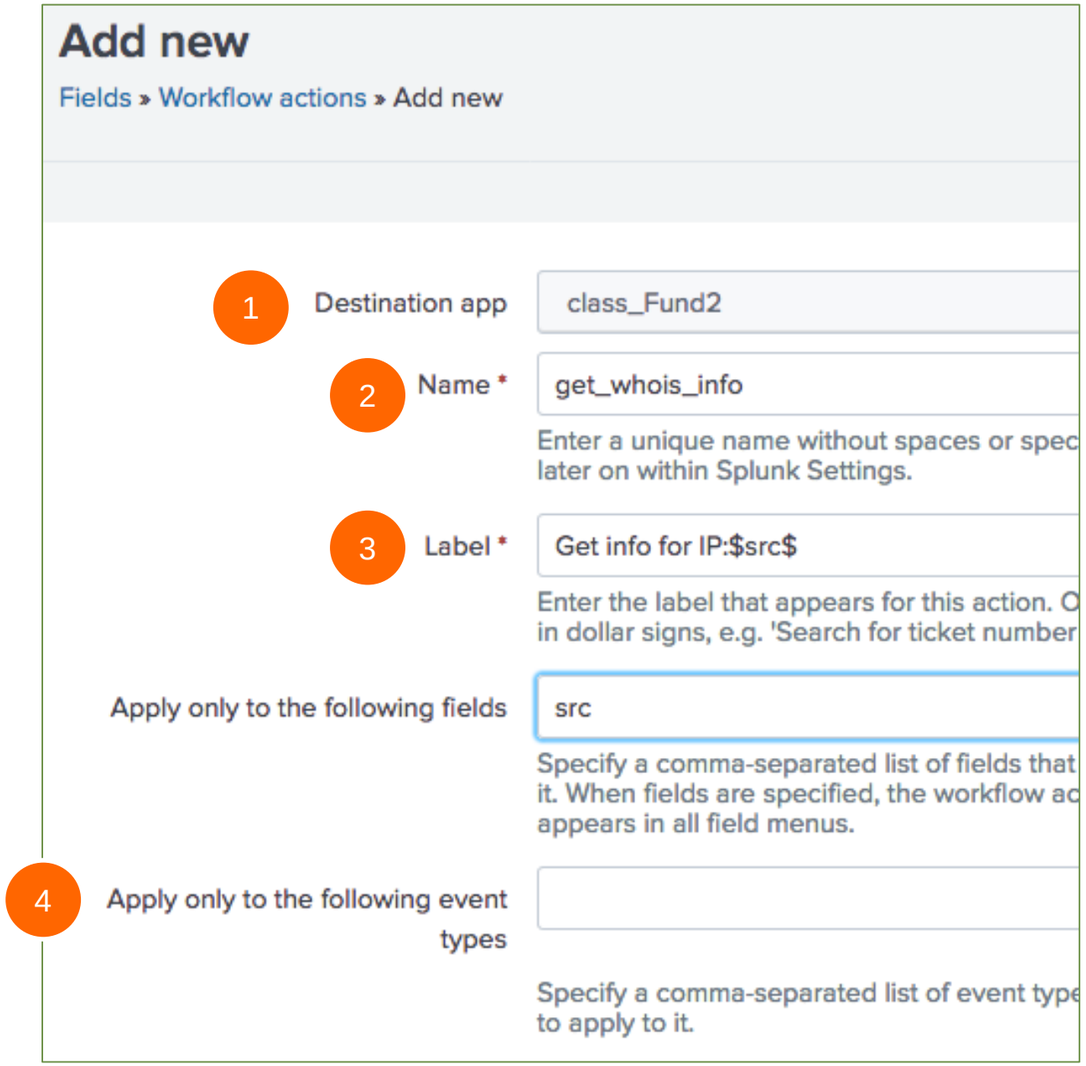
Generated for (v.ms.besetti@jpmchase.com) (C) Splunk Inc, not for distribution

Copyright © 2018 Splunk, Inc. All rights reserved | 14 June 2018

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 menu
4. Determine if your workflow action applies to a field or event type

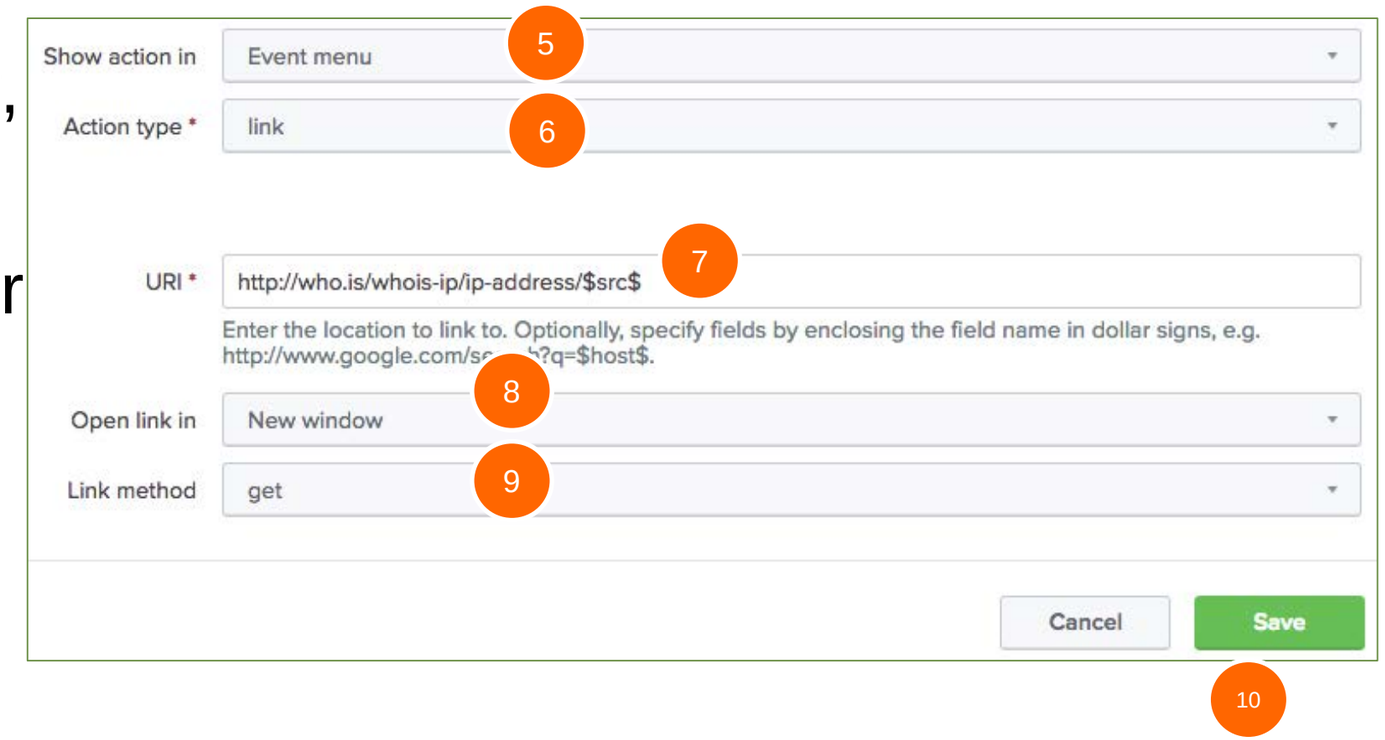


Generated for (v.ms.besetti@jpmchase.com) (C) Splunk Inc, not for distribution

Copyright © 2018 Splunk, Inc. All rights reserved | 14 June 2018

Creating a GET Workflow Action (cont.)

5. From the **Show action in** drop-down list, select **Event menu**
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



Generated for (v.ms.besetti@jpmchase.com) (C) Splunk Inc, not for distribution

Copyright © 2018 Splunk, Inc. All rights reserved | 14 June 2018

Testing the GET Workflow Action

1/26/18
1:38:17.210 PM

1517002697.210 68 69.80.0.18 TCP_REFRESH_HIT/200 977 GET http://www.manta.com/manta/images/header_and_nav/btn_search.gif lsagers@buttercupgames.com DIRECT/www.manta.com - ALLOW_WBRS-DefaultGroup-Demo_Clients-NONE-NONE,-,-,-,-,-,-,-,-,-,-,IW_busi,-> - http://www.mant

Event Actions

- Build Event Type
- Get info for IP:69.80.0.18
- Extract Fields
- Show Source

Value
~~nix-all-logs~~
cisco_router1
/opt/log/cisco_ /cisco_ironport

IP Information for 69.80.0.18

Quick Stats

IP Location	Barbados Bridgetown Cable & Wireless (barbados) Limited
ASN	AS14813 BB-COLUMBUS - Columbus Telecommunications (Barbados) Limited, BB (registered Mar 15, 2007)
Whois Server	whois.arin.net
IP Address	69.80.0.18

```
NetRange:       69.80.0.0 - 69.80.63.255
CIDR:           69.80.0.0/18
NetName:        BDS-NET9
NetHandle:      NET-69-80-0-0-1
Parent:         NET69 (NET-69-0-0-0-0)
NetType:        Direct Allocation
OriginAS:       AS11139
Organization:   Cable & Wireless (Barbados) Limited (CWBL-1)
RegDate:        2010-03-23
Updated:        2012-03-02
Ref:            https://whois.arin.net/rest/net/NET-69-80-0-0-1
```

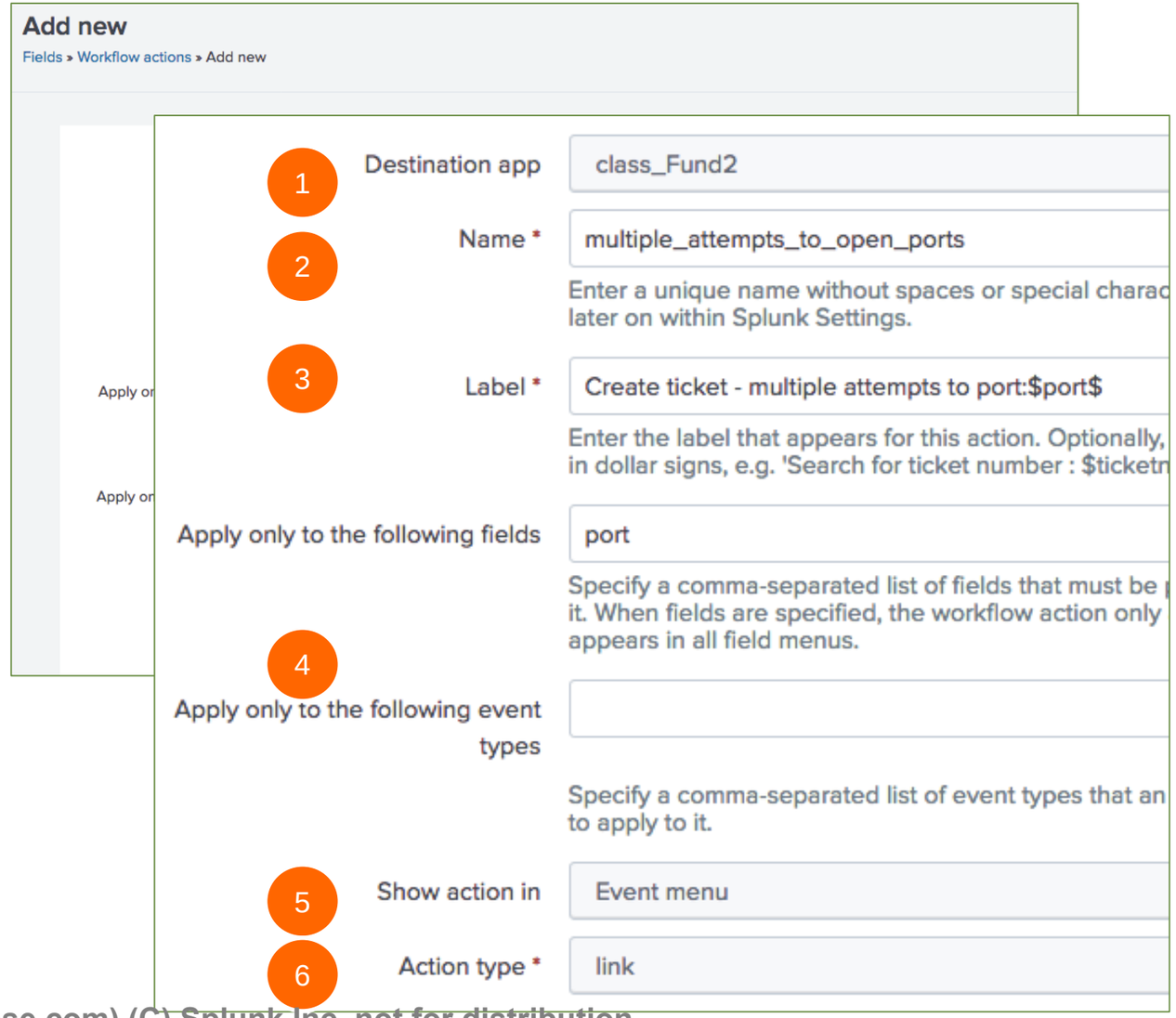
Generated for (v.ms.besetti@jpmchase.com) (C) Splunk Inc, not for distribution

Copyright © 2018 Splunk, Inc. All rights reserved | 14 June 2018

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

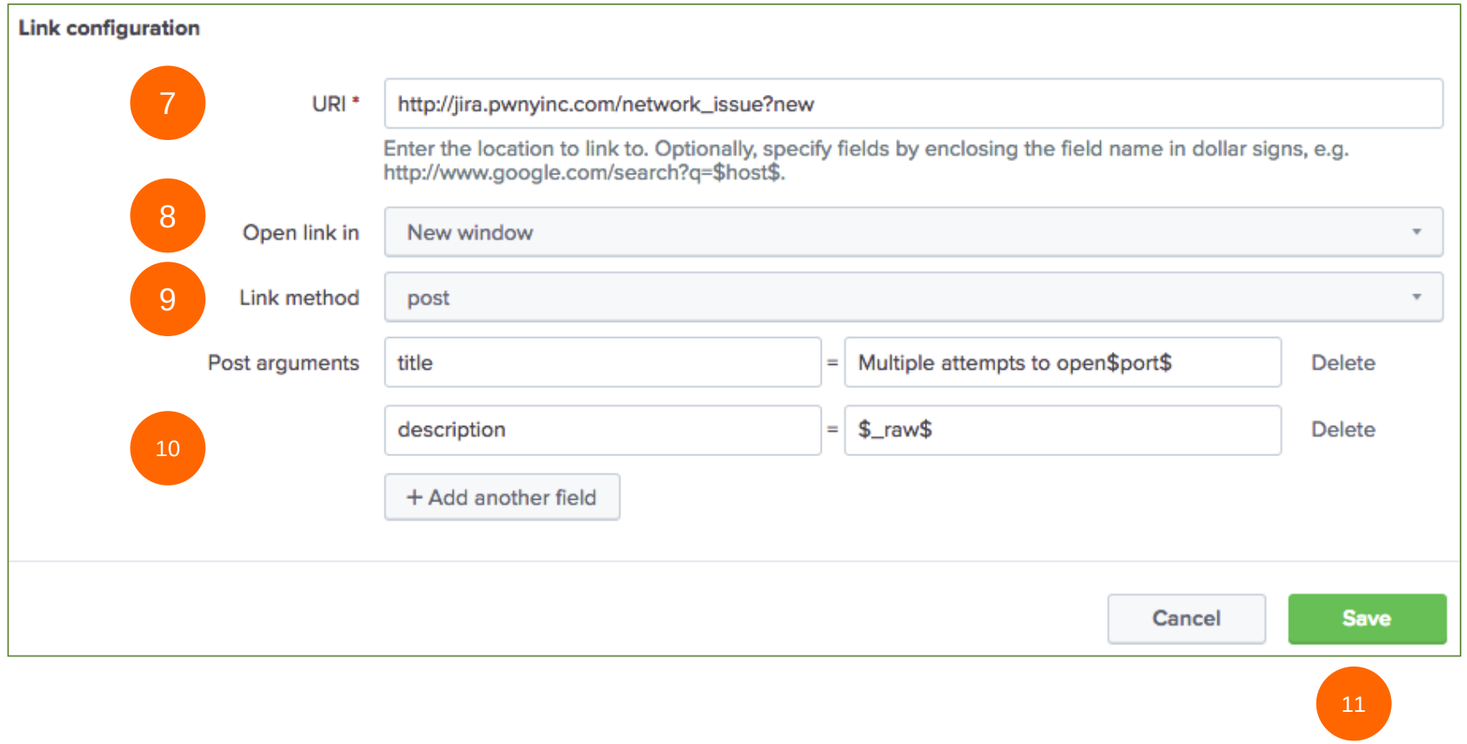


Generated for (v.ms.besetti@jpmchase.com) (C) Splunk Inc, not for distribution

Copyright © 2018 Splunk, Inc. All rights reserved | 14 June 2018

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



Generated for (v.ms.besetti@jpmchase.com) (C) Splunk Inc, not for distribution

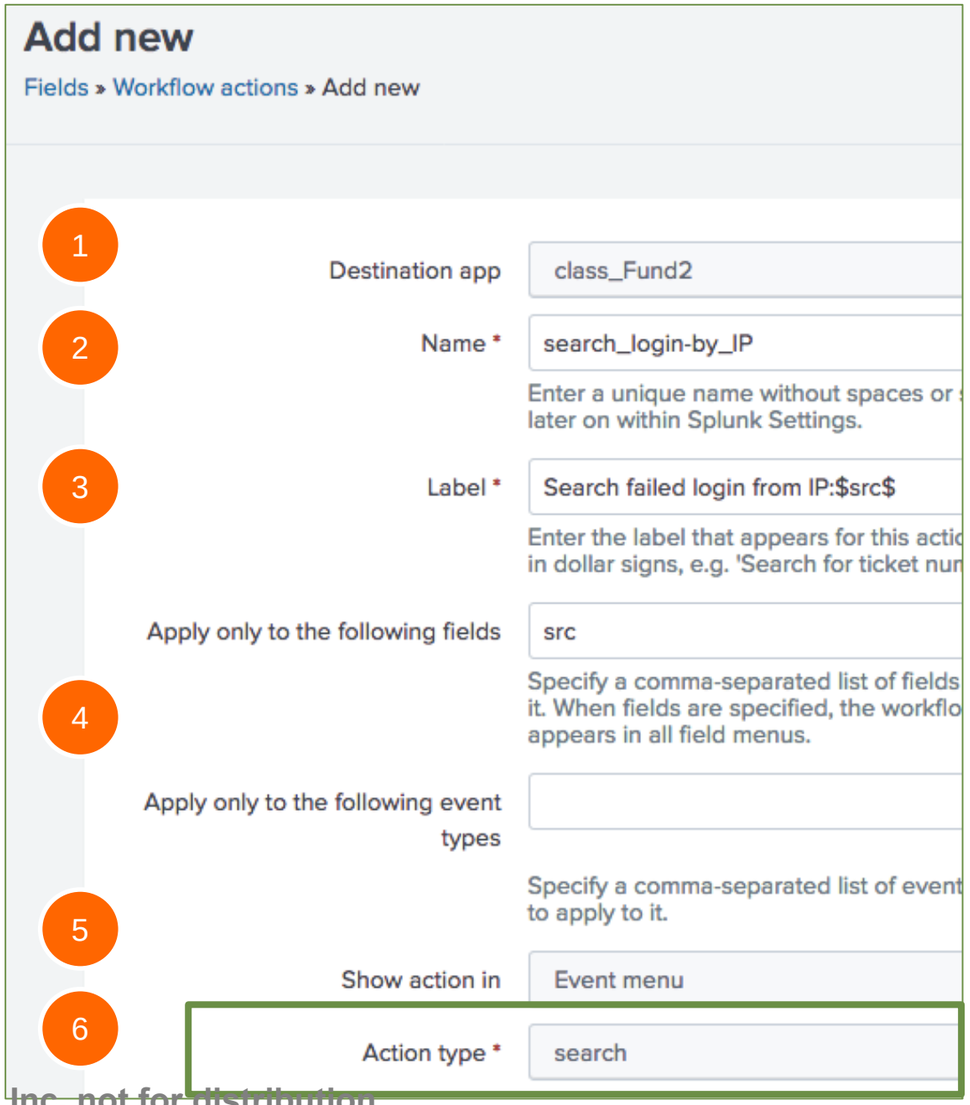
Copyright © 2018 Splunk, Inc. All rights reserved | 14 June 2018

Creating a Search Workflow Action

Settings > Fields > Workflow actions> New

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

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

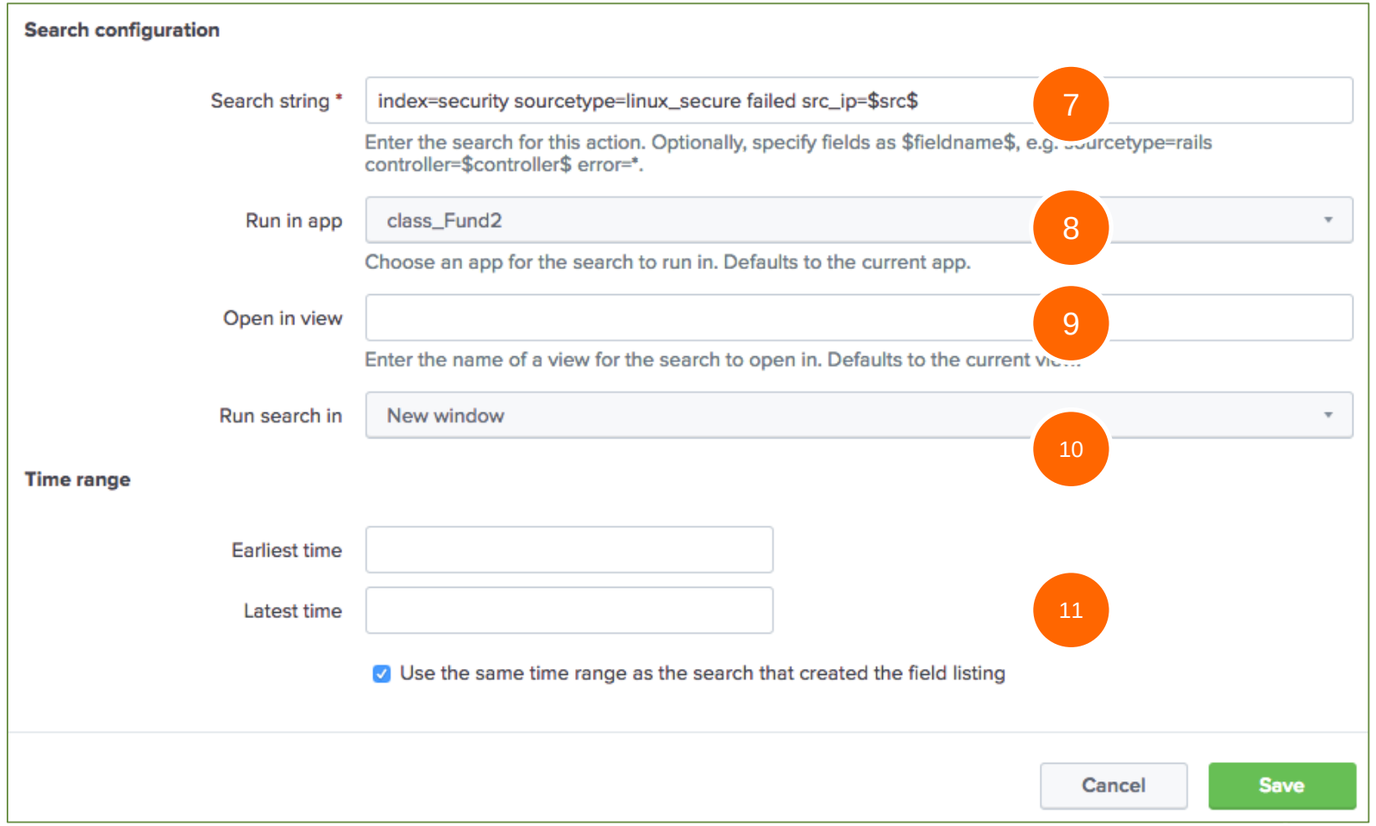


Generated for (v.ms.besetti@jpmchase.com) (C) Splunk Inc, not for distribution

Copyright © 2018 Splunk, Inc. All rights reserved | 14 June 2018

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. Enter the time range for the search or choose to use the same time range as the search
12. Save



12

Generated for (v.ms.besetti@jpmchase.com) (C) Splunk Inc, not for distribution

Copyright © 2018 Splunk, Inc. All rights reserved | 14 June 2018

Testing the Search Workflow Action

1/26/18
2:34:02.508 PM

1517006042.508 92 188.173.152.100 TCP_REFRESH_HIT/200 3677 GET http://www.hit-country-music-lyrics.com/imgs/concert-tickets.gif rerde@buttercupgames.com DIRECT/www.hit-country-music-lyrics.com image/gif DEFAULT_CASE-DefaultGroup-Demo_Clients-NONE-NONE-DefaultRouting <IW_art,ns,0,-,-,-,-,0,-,-,-,-,-,-,-,IW_art,-> - http://www.hit-country-music-lyrics.com/

Event Actions

- Build Event Type
- Extract Fields
- Search failed login from IP:188.173.152.100
- Show Source

New Search

Save As | Close

index=security sourcetype=linux_secure failed src_ip=188.173.152.100

Last 24 hours

✓ 22 events (1/25/18 2:00:00.000 PM to 1/26/18 2:34:57.000 PM) | No Event Sampling | Job | Smart Mode

Events (22) | Patterns | Statistics | Visualization

Format Timeline | – Zoom Out | + Zoom to Selection | × Deselect | 1 hour per column

List | Format | 20 Per Page | < Prev | 1 | 2 | Next >

< Hide Fields | All Fields

SELECTED FIELDS
- a eventtype 6
- a host 4
- a source 4
- a sourcetype 1
- a tag 7

INTERESTING FIELDS
- a action 1
- a app 1
- # date_hour 9
- # date_mday 2
- # date_minute 20

i	Time	Event
>	1/26/18 2:09:38.000 PM	Fri Jan 26 2018 22:09:38 www2 sshd[1287]: Failed password for invalid user administrator from 188.173.152.100 port 3454 ssh2 eventtype = err0r 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
>	1/26/18 1:59:13.000 PM	Fri Jan 26 2018 21:59:13 www2 sshd[2005]: Failed password for invalid user rdb from 188.173.152.100 port 1623 ssh2 eventtype = err0r 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 (v.ms.besetti@jpmchase.com) (C) Splunk Inc, not for distribution

Copyright © 2018 Splunk, Inc. All rights reserved | 14 June 2018

Module 13: Creating Data Models

Generated for (v.ms.besetti@jpmchase.com) (C) Splunk Inc, not for distribution

Copyright © 2018 Splunk, Inc. All rights reserved

Objectives

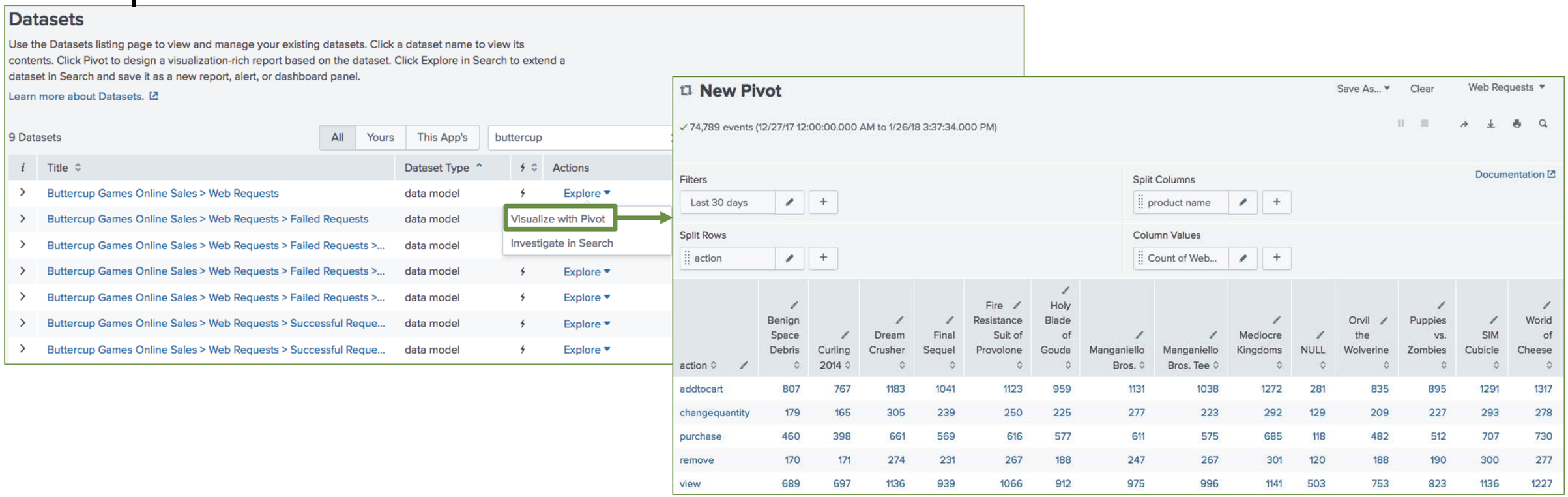
- Describe the relationship between data models and Pivot
- Identify data model datasets
- Identify dataset fields
- Create a data model
- Use a data model in Pivot

Generated for (v.ms.besetti@jpmchase.com) (C) Splunk Inc, not for distribution

Copyright © 2018 Splunk, Inc. All rights reserved | 14 June 2018

Reviewing Pivot

- Used for creating reports and dashboards (discussed in Splunk Fundamentals 1)
- As a knowledge manager, you're responsible for building the data model that provides the datasets for Pivot

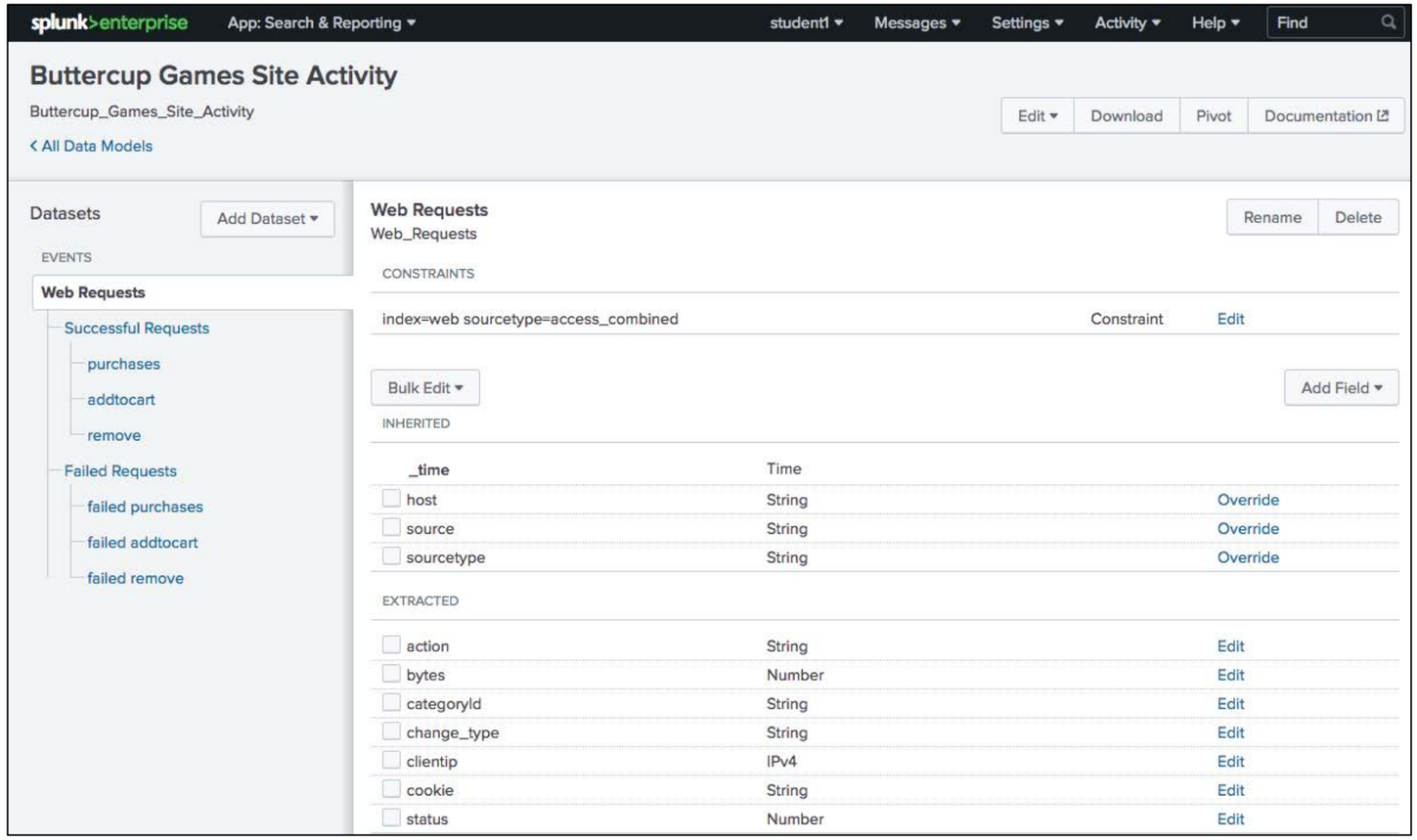


Generated for (v.ms.besetti@jpmchase.com) (C) Splunk Inc, not for distribution

Copyright © 2018 Splunk, Inc. All rights reserved | 14 June 2018

Overview of Data Models

- Hierarchically structured datasets that generate searches and drive Pivot
- Pivot reports are created based on datasets
- Each event, search or transaction is saved as a separate dataset



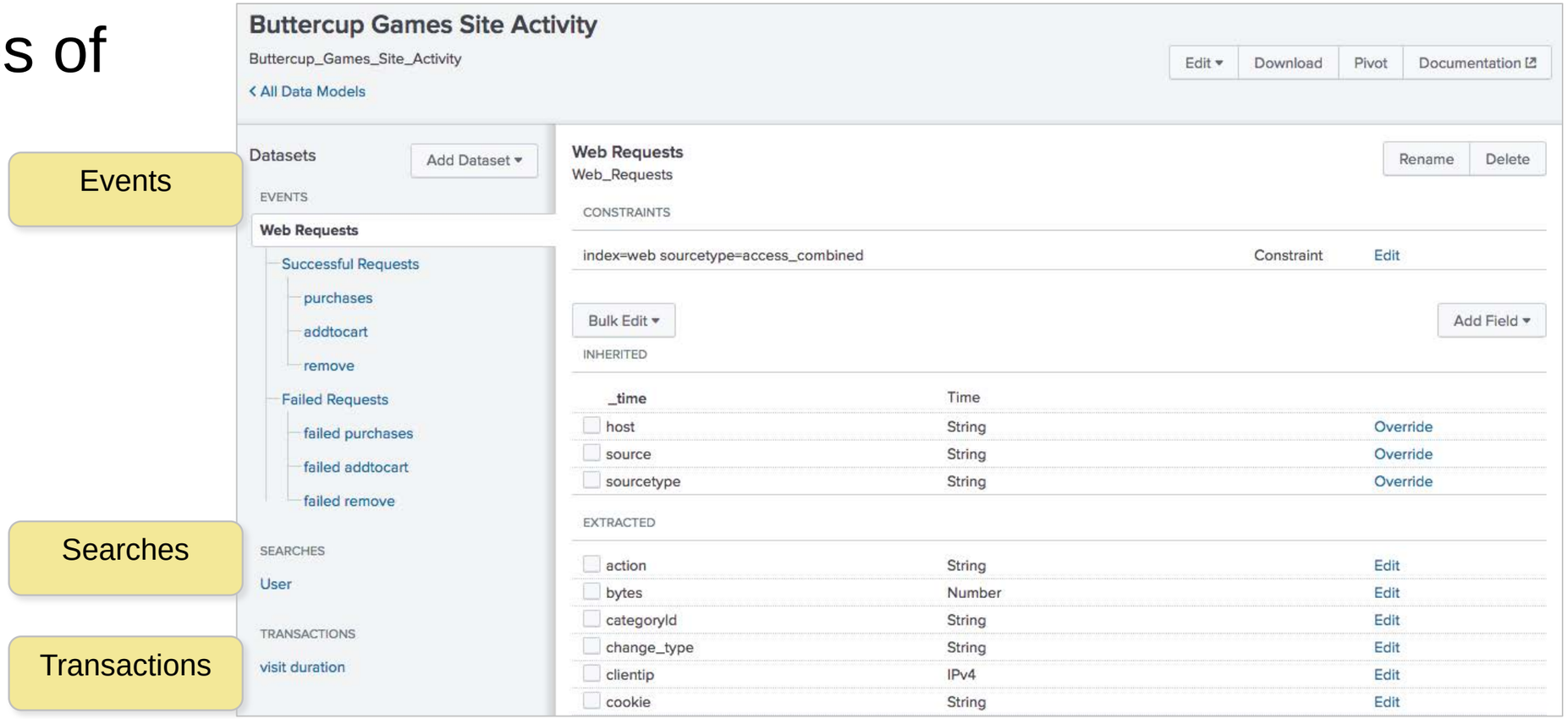
Generated for (v.ms.besetti@jpmchase.com) (C) Splunk Inc, not for distribution

Copyright © 2018 Splunk, Inc. All rights reserved | 14 June 2018

Data Model Dataset Types

A data model can consist of 3 types of datasets

- Events
- Searches
- Transactions



Events

Searches

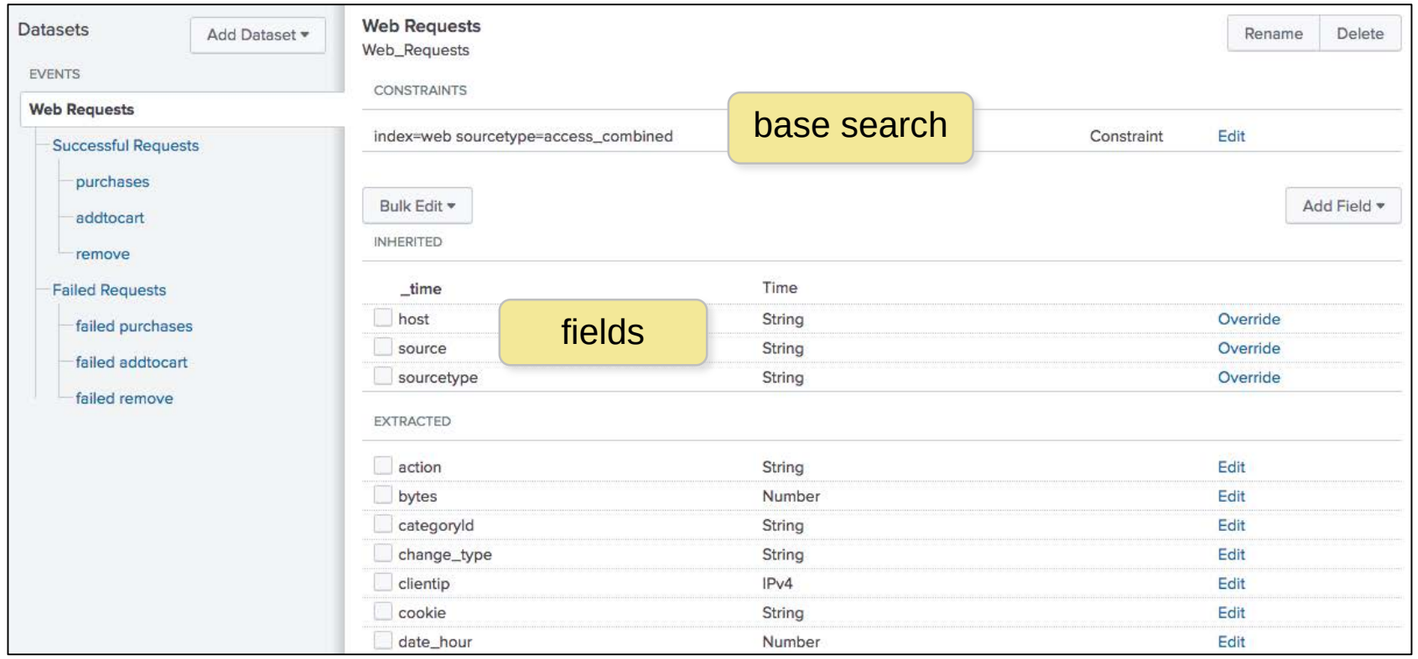
Transactions

Generated for (v.ms.besetti@jpmchase.com) (C) Splunk Inc, not for distribution

Copyright © 2018 Splunk, Inc. All rights reserved | 14 June 2018

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



Generated for (v.ms.besetti@jpmchase.com) (C) Splunk Inc, not for distribution

Copyright © 2018 Splunk, Inc. All rights reserved | 14 June 2018

Event Object Hierarchy and Constraints

Datasets | Add Dataset
EVENTS
- Web Requests
  - Successful Requests
    - purchases
    - addtocart
    - remove
  - Failed Requests
    - failed purchase
    - failed addtocar
    - failed remove

Web Requests
Web_Requests
CONSTRAINTS
1 index=web sourcetype=access_combined — Constraint — Edit

base search – all access_combined events

Each constraint inherits the parent search string

Datasets | Add Dataset
EVENTS
- Web Requests
  - Successful Requests
    - purchases
    - addtocart
    - remove
  - Failed Requests
    - failed purchases
    - failed addtocart
    - failed remove

Successful Requests
Successful_Requests
Rename | Delete
CONSTRAINTS
2 index=web sourcetype=access_combined
status<400
Edit

successful requests

Datasets | Add Dataset
EVENTS
- Web Requests
  - Successful Requests
    - purchases
    - addtocart
    - remove
  - Failed Requests
    - failed purchases
    - failed addtocart
    - failed remove

purchases
purchases
Rename | Delete
CONSTRAINTS
3 index=web sourcetype=access_combined — Inherited
status<400
action=purchase productId=* — Constraint — Edit

successful requests for purchases

Bulk Edit | Add Field
INHERITED

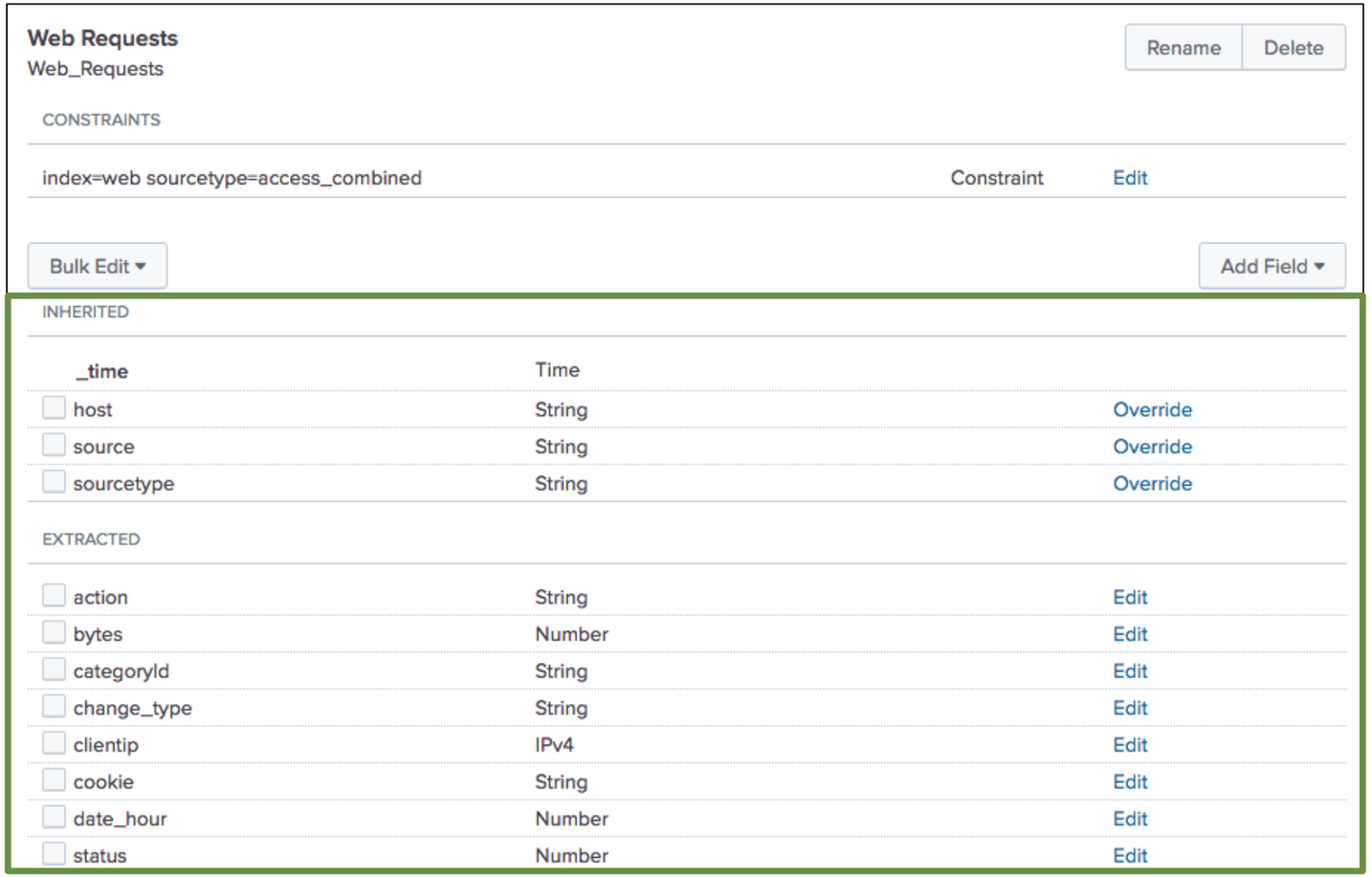
_time	Time	
action	String	Override
bytes	Number	Override

Generated for (v.ms.besetti@jpmchase.com) (C) Splunk Inc, not for distribution

Copyright © 2018 Splunk, Inc. All rights reserved | 14 June 2018

Dataset Fields

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



Generated for (v.ms.besetti@jpmchase.com) (C) Splunk Inc, not for distribution

Copyright © 2018 Splunk, Inc. All rights reserved | 14 June 2018

Creating a Data Model

Settings > Data Models

splunk>enterprise Apps ▾ student1 ▾ Messages ▾ Settings ▾ Activity ▾ Help ▾ Find

Data Models

Data models enable users to easily create reports in the Pivot tool. Learn More

Upload Data Model | New Data Model

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

i	Title	Type	Actions	App
>	Alerts	data model	Clone Pivot	Splunk_SA_C
>	Application State	data model		
>	Authentication	data model		
>	Buttercup Games Online Sales	data model		

ID is automatically populated from Title, but can be overridden

Choose app context

New Data Model

Title: Buttercup Games Site Activity

ID: Buttercup_Games_Site_Activity

Can only contain letters, numbers and underscores.

App: Search & Reporting ▾

Description: optional

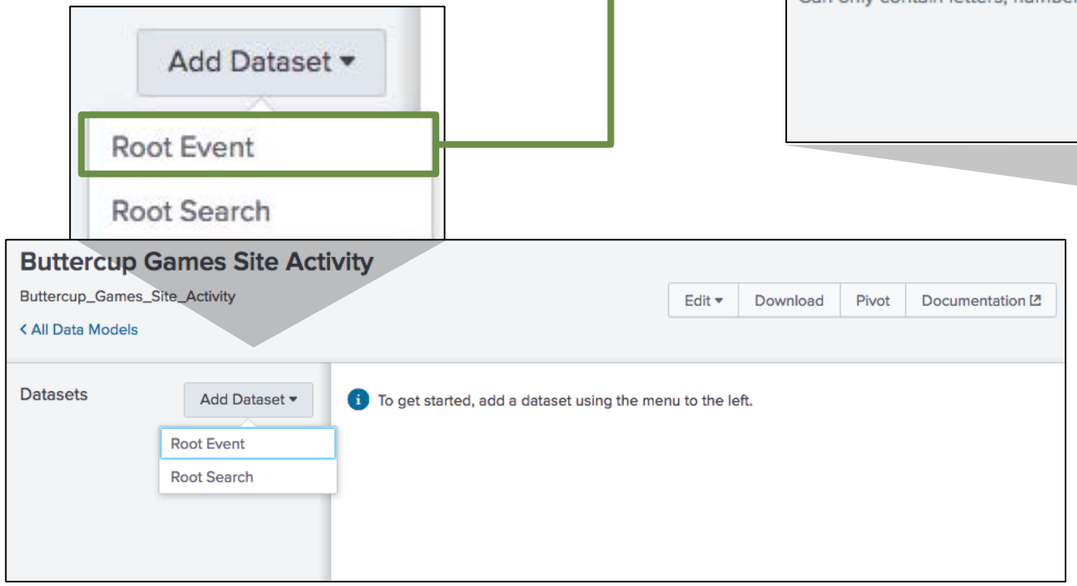
Cancel | Create

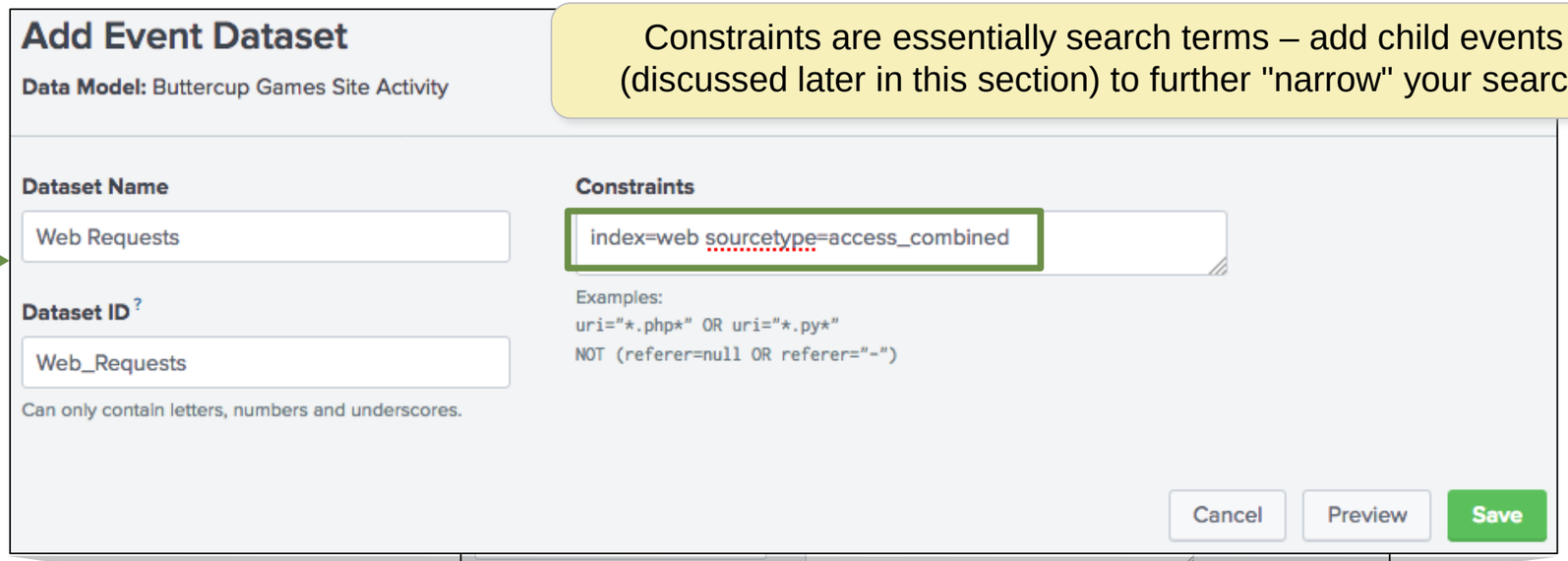
Generated for (v.ms.besetti@jpmchase.com) (C) Splunk Inc, not for distribution

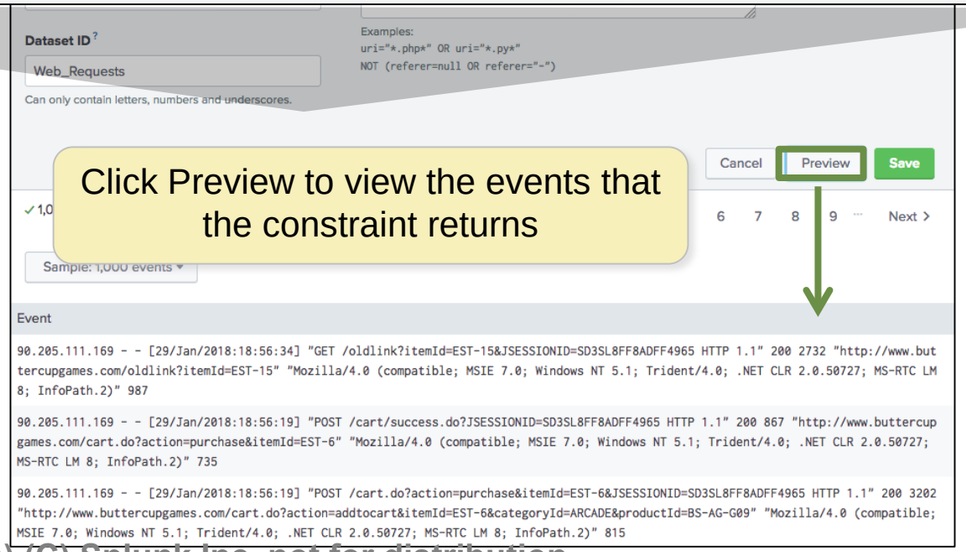
Copyright © 2018 Splunk, Inc. All rights reserved | 14 June 2018

Adding a Root Event

Constraints are essentially search terms – add child events (discussed later in this section) to further "narrow" your search







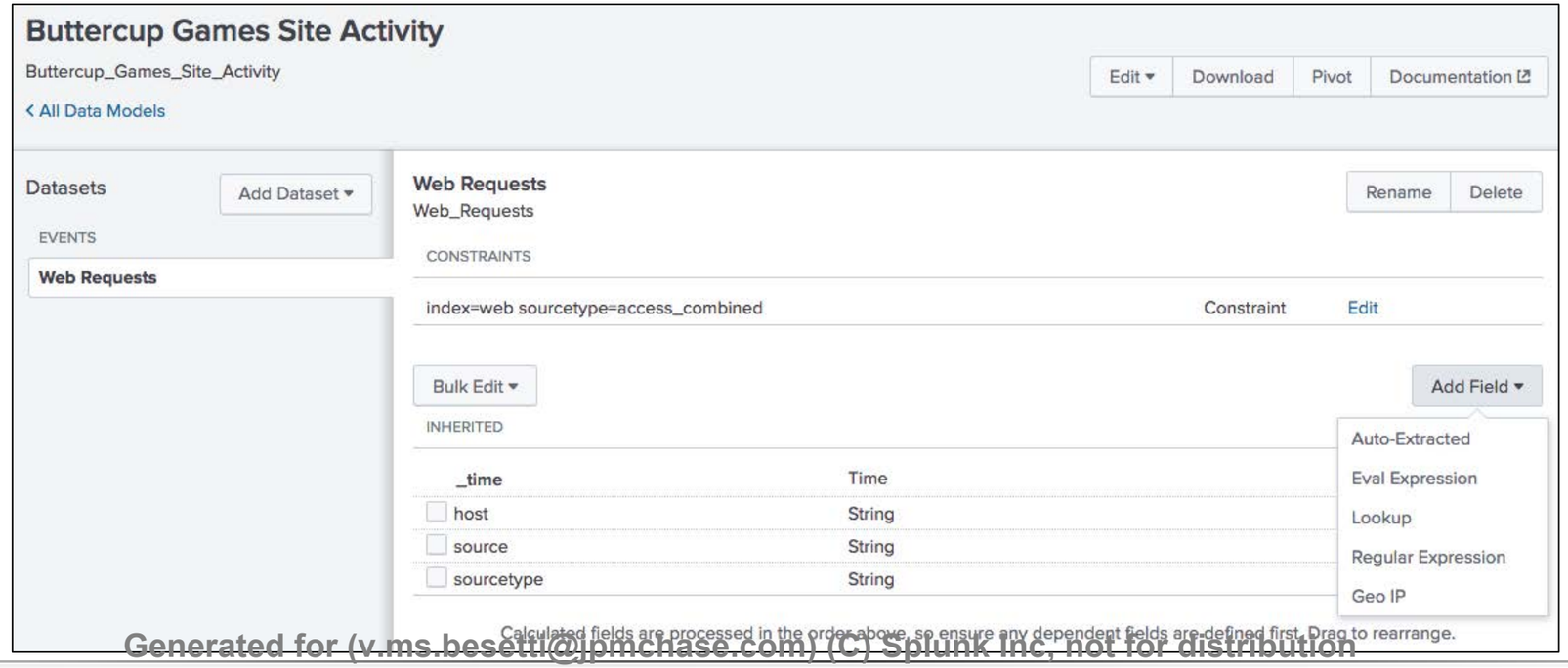
Click Preview to view the events that the constraint returns

Generated for (v.ms.besetti@jpmchase.com) (C) Splunk Inc, not for distribution

Copyright © 2018 Splunk, Inc. All rights reserved | 14 June 2018

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

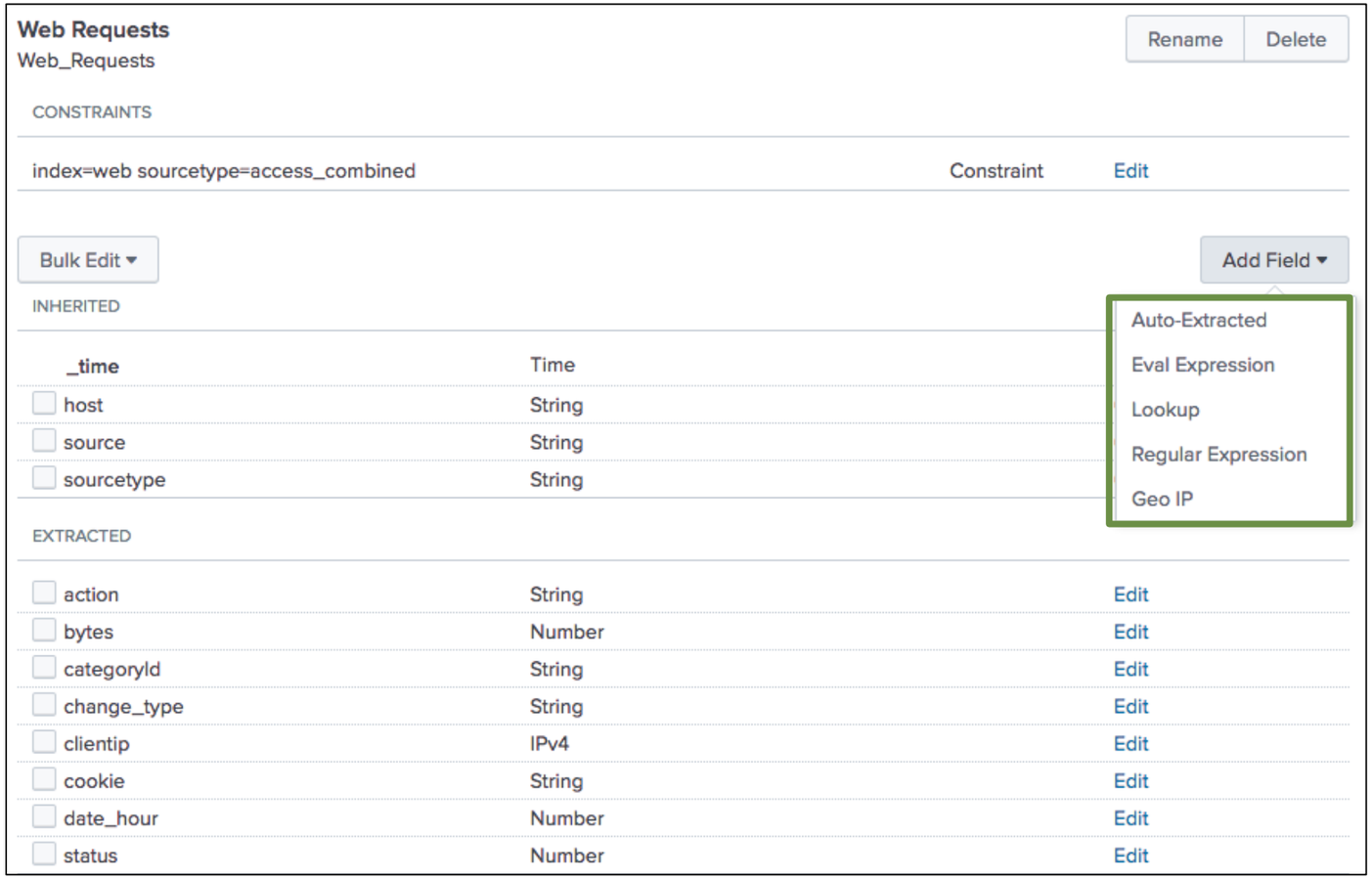


Generated for (v.ms.besetti@jpmchase.com) (C) Splunk Inc, not for distribution

Copyright © 2018 Splunk, Inc. All rights reserved | 14 June 2018

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.

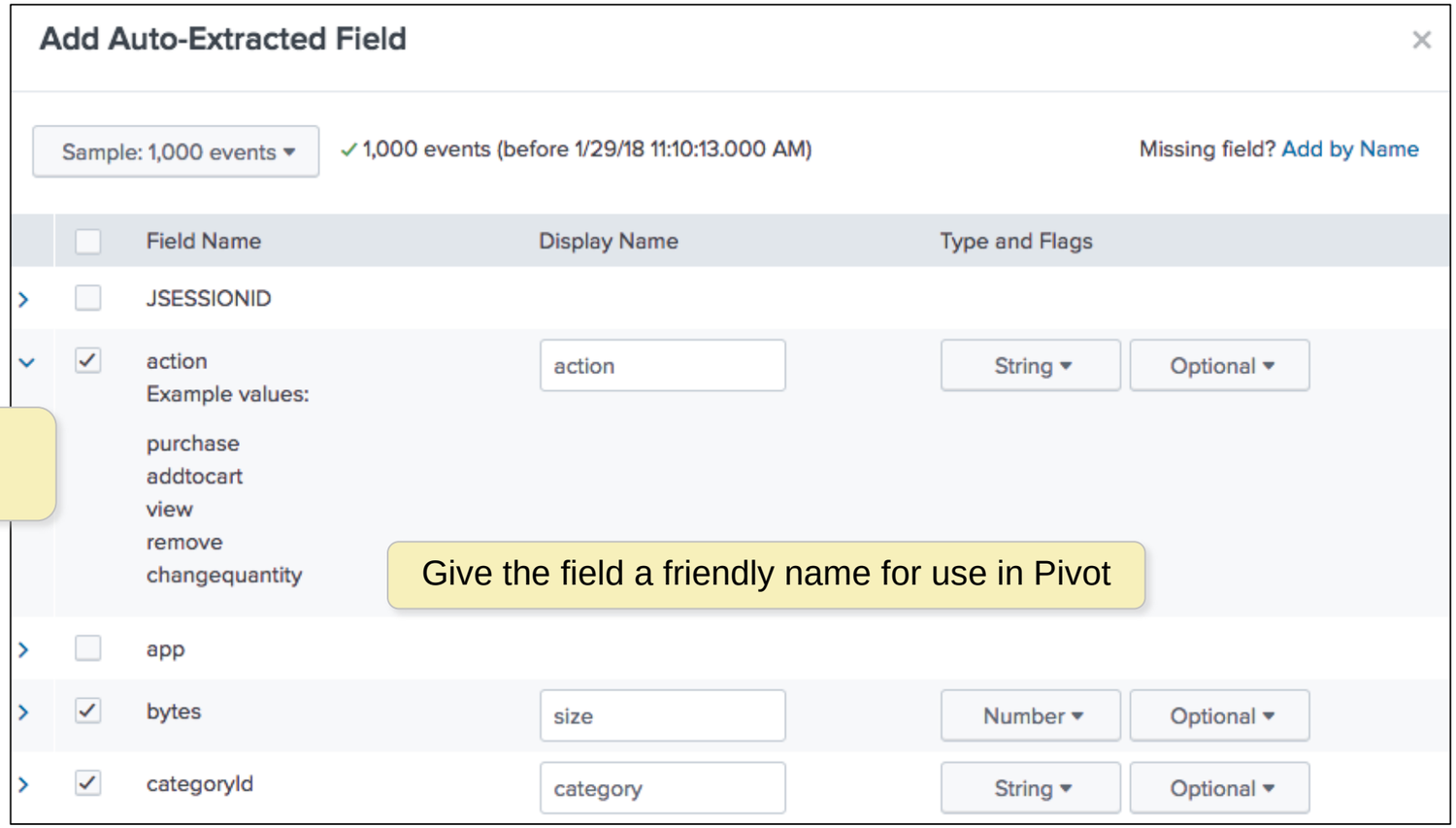


Generated for (v.ms.besetti@jpmchase.com) (C) Splunk Inc, not for distribution

Copyright © 2018 Splunk, Inc. All rights reserved | 14 June 2018

Adding Fields – Auto-Extracted

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



View a field's example values

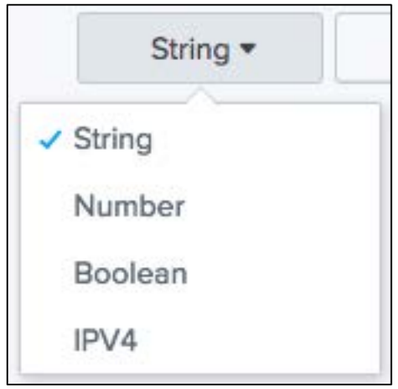
Give the field a friendly name for use in Pivot

Generated for (v.ms.besetti@jpmchase.com) (C) Splunk Inc, not for distribution

Copyright © 2018 Splunk, Inc. All rights reserved | 14 June 2018

Field Types

- String: Field values are recognized as alpha-numeric
- 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

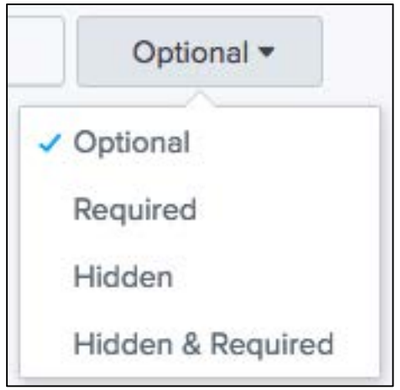


Generated for (v.ms.besetti@jpmchase.com) (C) Splunk Inc, not for distribution

Copyright © 2018 Splunk, Inc. All rights reserved | 14 June 2018

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

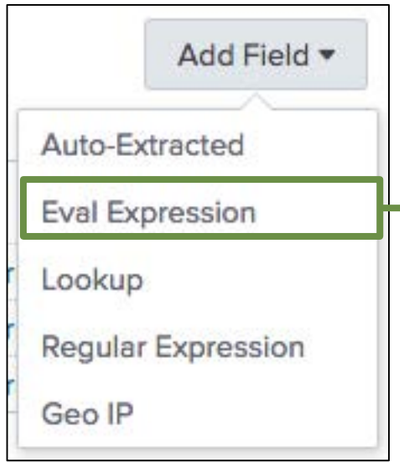


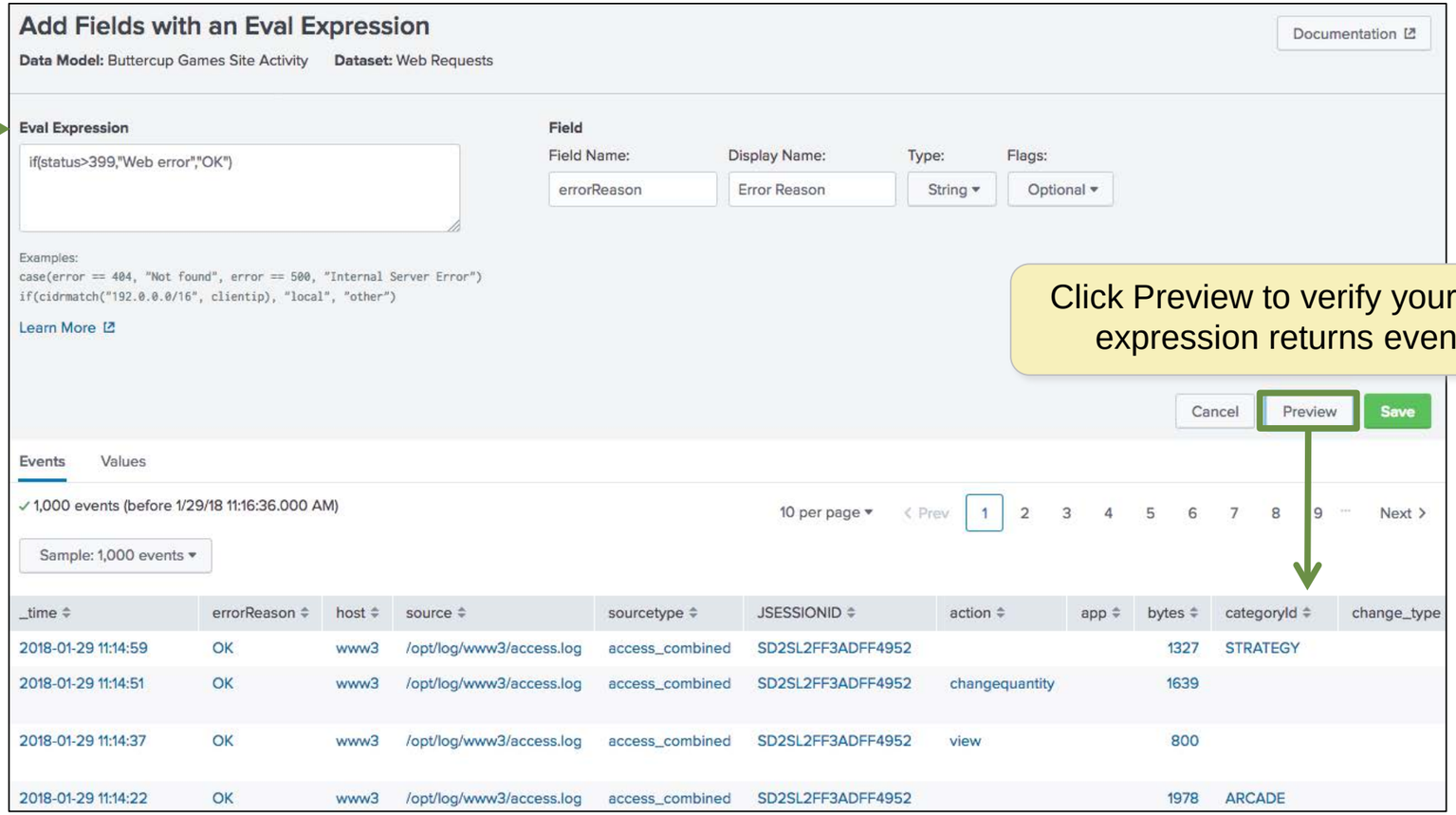
Generated for (v.ms.besetti@jpmchase.com) (C) Splunk Inc, not for distribution

Copyright © 2018 Splunk, Inc. All rights reserved | 14 June 2018

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





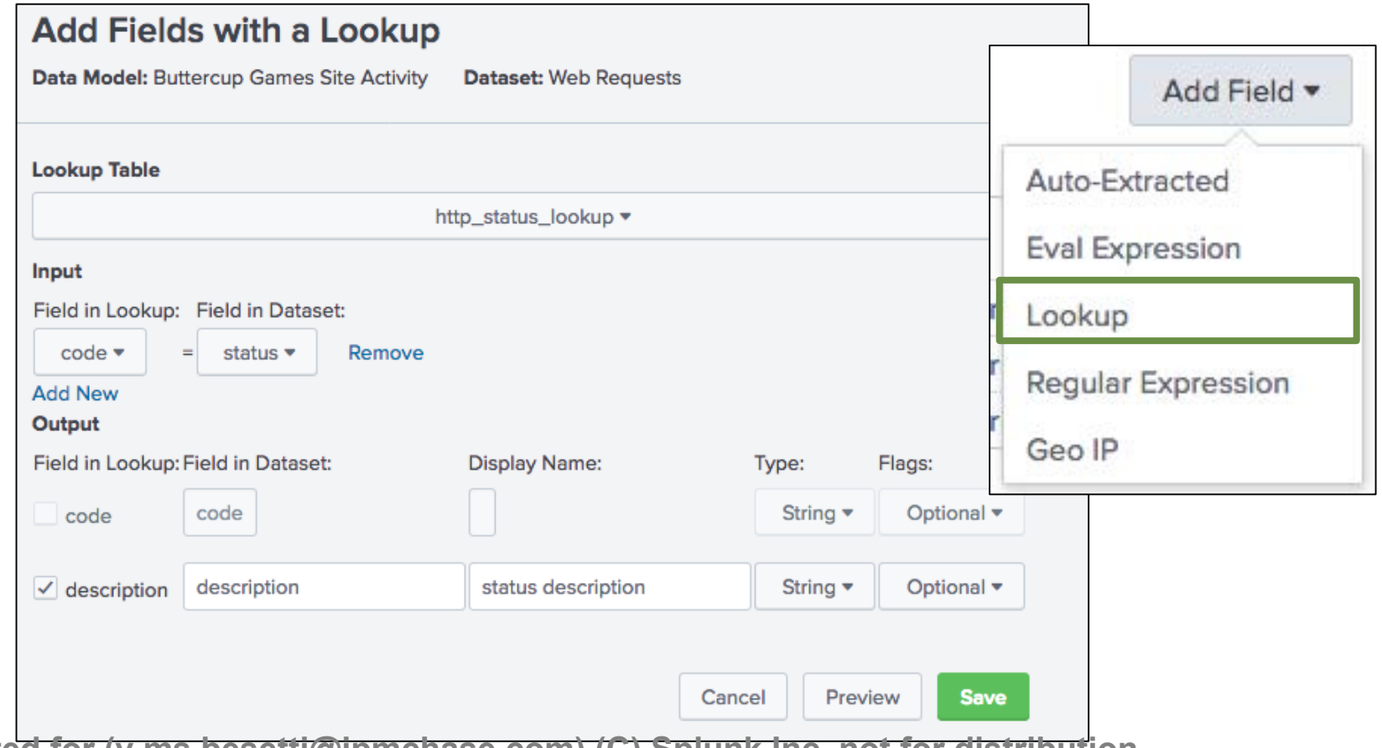
Click Preview to verify your eval expression returns events

Generated for (v.ms.besetti@jpmchase.com) (C) Splunk Inc, not for distribution

Copyright © 2018 Splunk, Inc. All rights reserved | 14 June 2018

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

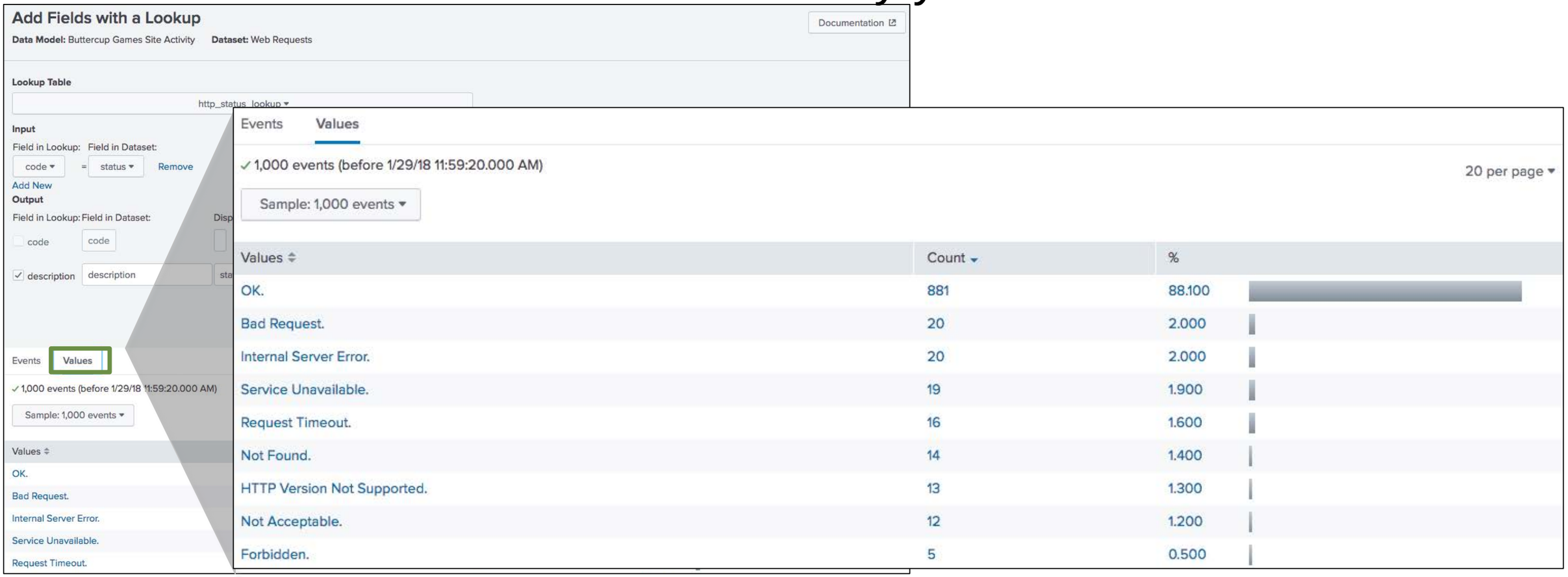


Generated for (v.ms.besetti@jpmchase.com) (C) Splunk Inc, not for distribution

Copyright © 2018 Splunk, Inc. All rights reserved | 14 June 2018

Adding Fields – Lookups (cont.)

- Use Preview to test your lookup settings
- Use the Events and Values tab to verify your results

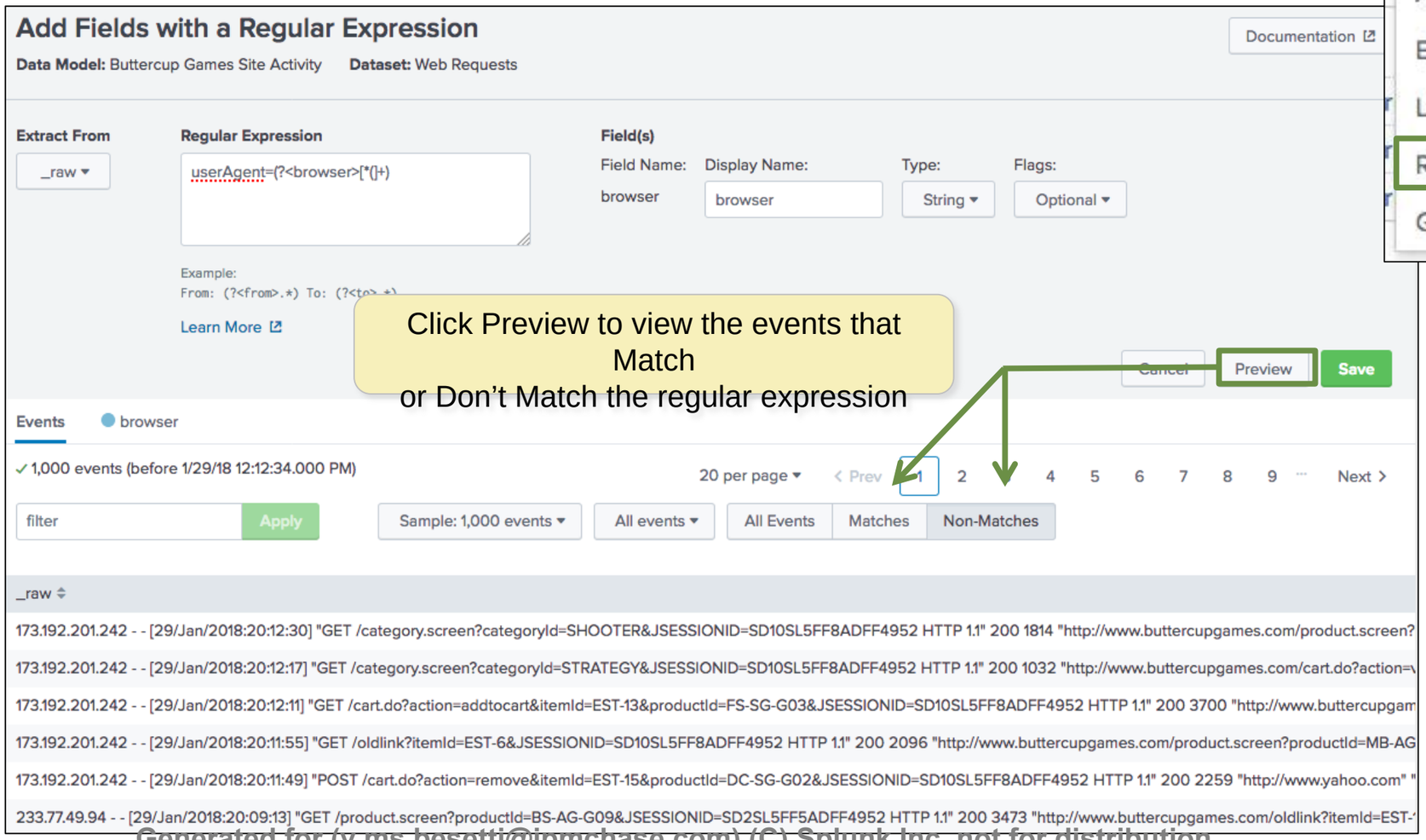


Generated for (v.ms.besetti@jpmchase.com) (C) Splunk Inc, not for distribution

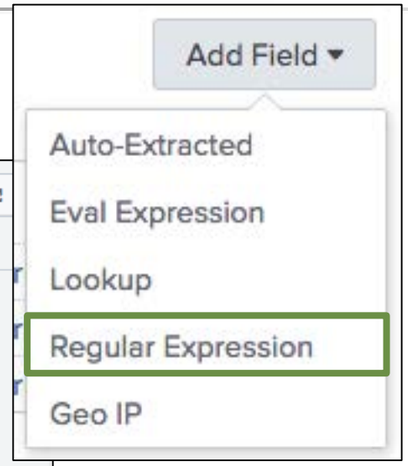
Copyright © 2018 Splunk, Inc. All rights reserved | 14 June 2018

Adding Fields – Regular Expression

You can define a new field using a regular expression



Click Preview to view the events that Match
or Don't Match the regular expression



Generated for (v.ms.besetti@jpmchase.com) (C) Splunk Inc, not for distribution

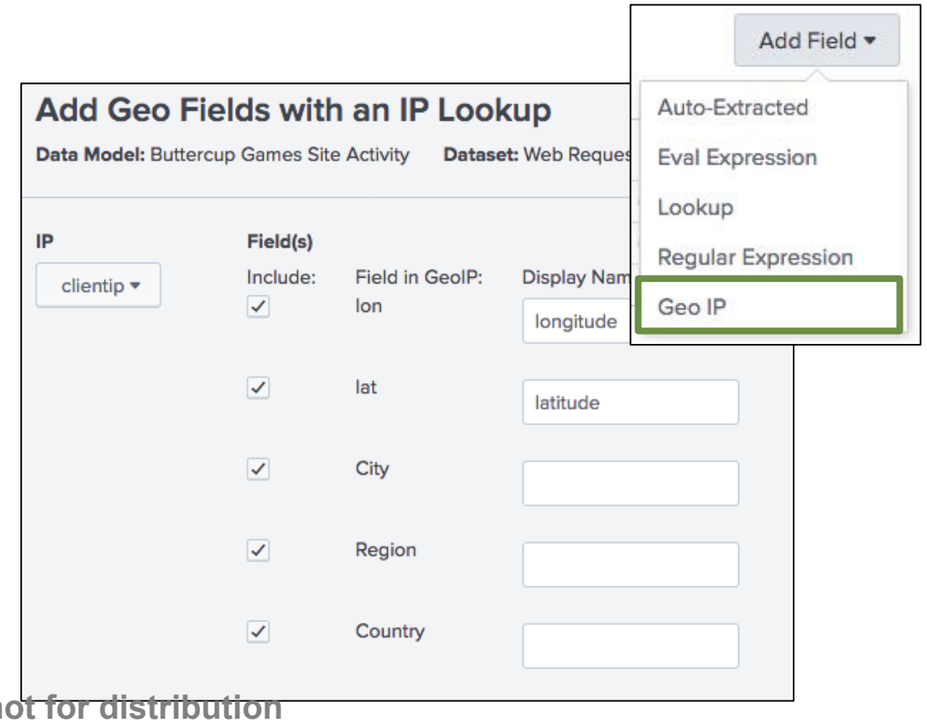
Copyright © 2018 Splunk, Inc. All rights reserved | 14 June 2018

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

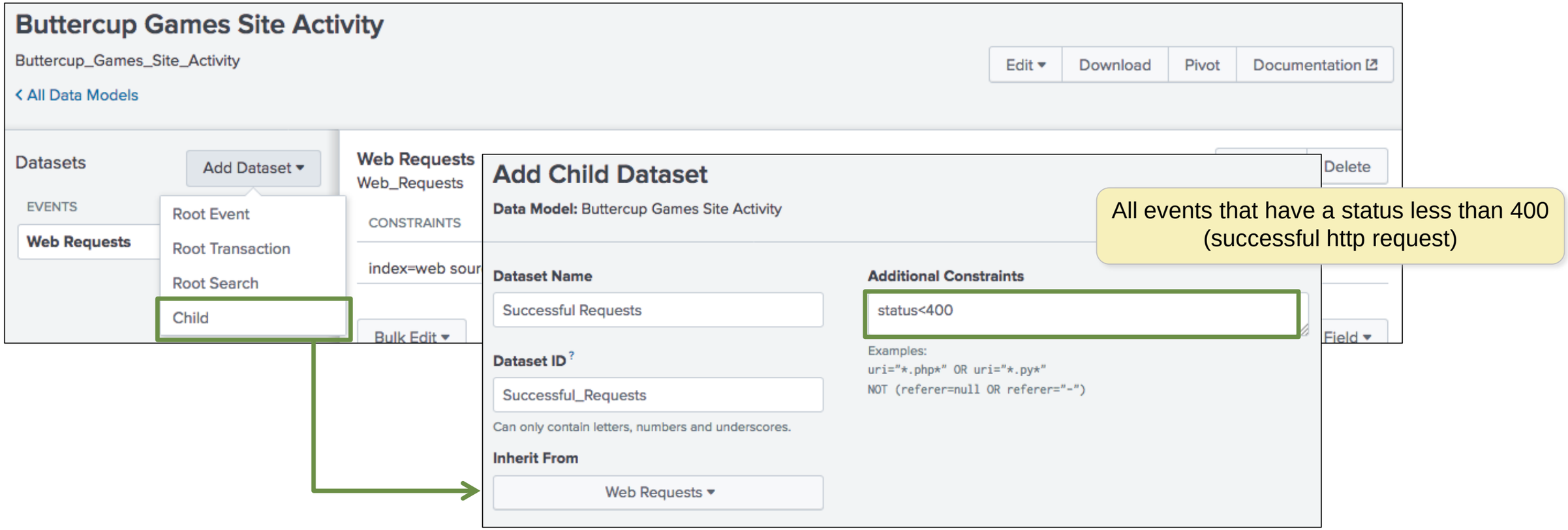


Generated for (v.ms.besetti@jpmchase.com) (C) Splunk Inc, not for distribution

Copyright © 2018 Splunk, Inc. All rights reserved | 14 June 2018

Adding Child Datasets

When you create a new child dataset, you give it one or more additional constraints



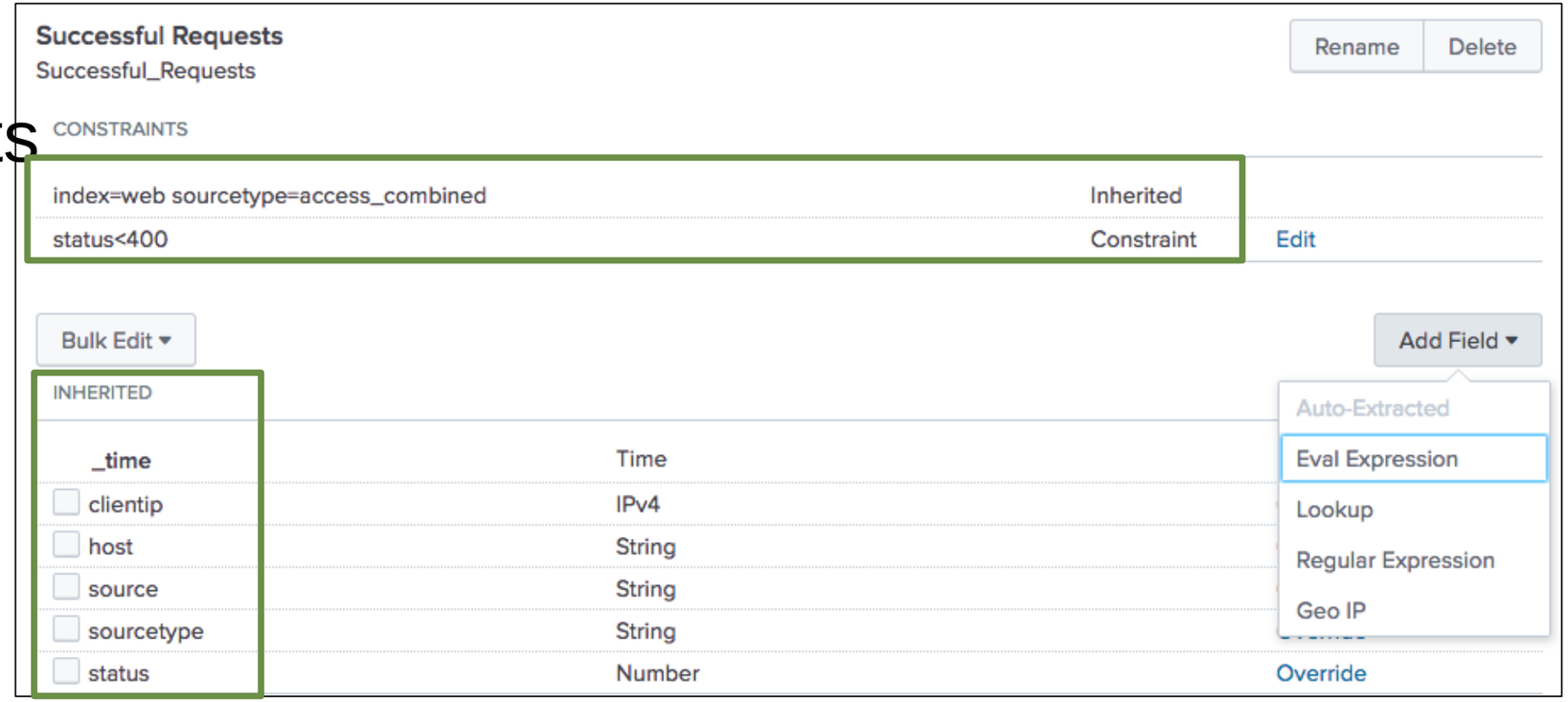
All events that have a status less than 400 (successful http request)

Generated for (v.ms.besetti@jpmchase.com) (C) Splunk Inc, not for distribution

Copyright © 2018 Splunk, Inc. All rights reserved | 14 June 2018

Adding Child Datasets (cont.)

- Child datasets inherit all fields from the parent events
- You can add more fields to child datasets

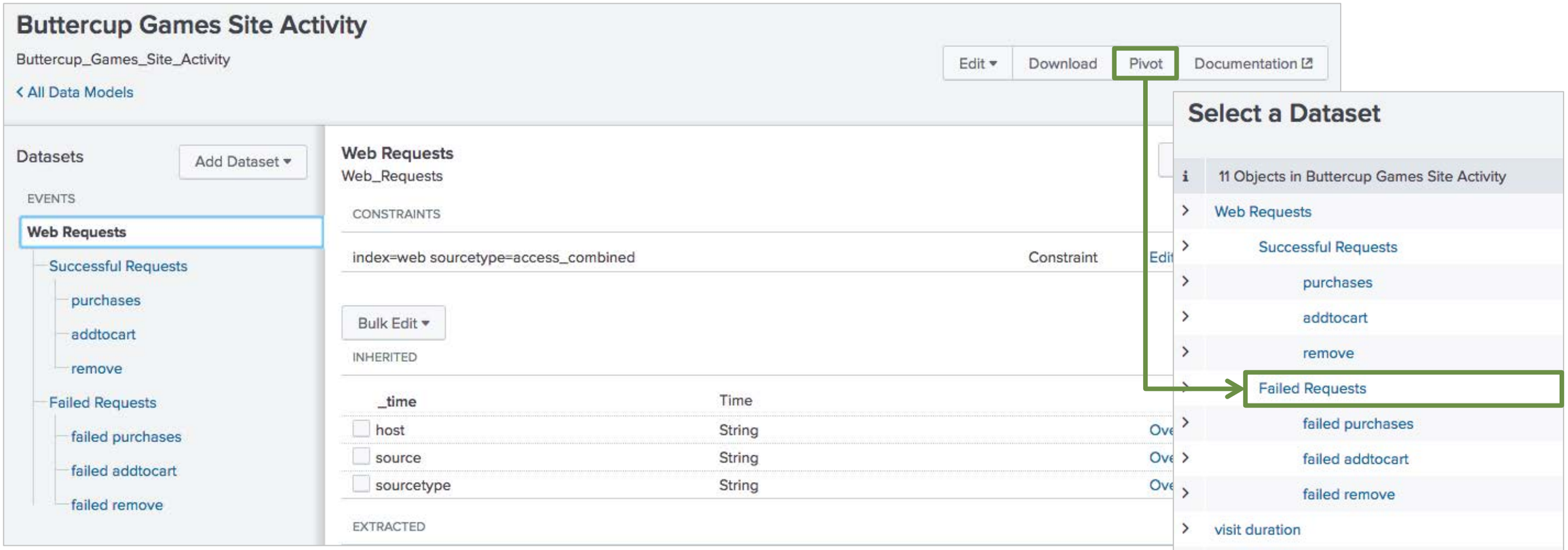


Generated for (v.ms.besetti@jpmchase.com) (C) Splunk Inc, not for distribution

Copyright © 2018 Splunk, Inc. All rights reserved | 14 June 2018

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



Generated for (v.ms.besetti@jpmchase.com) (C) Splunk Inc, not for distribution

Copyright © 2018 Splunk, Inc. All rights reserved | 14 June 2018

Using the Data Model in Pivot

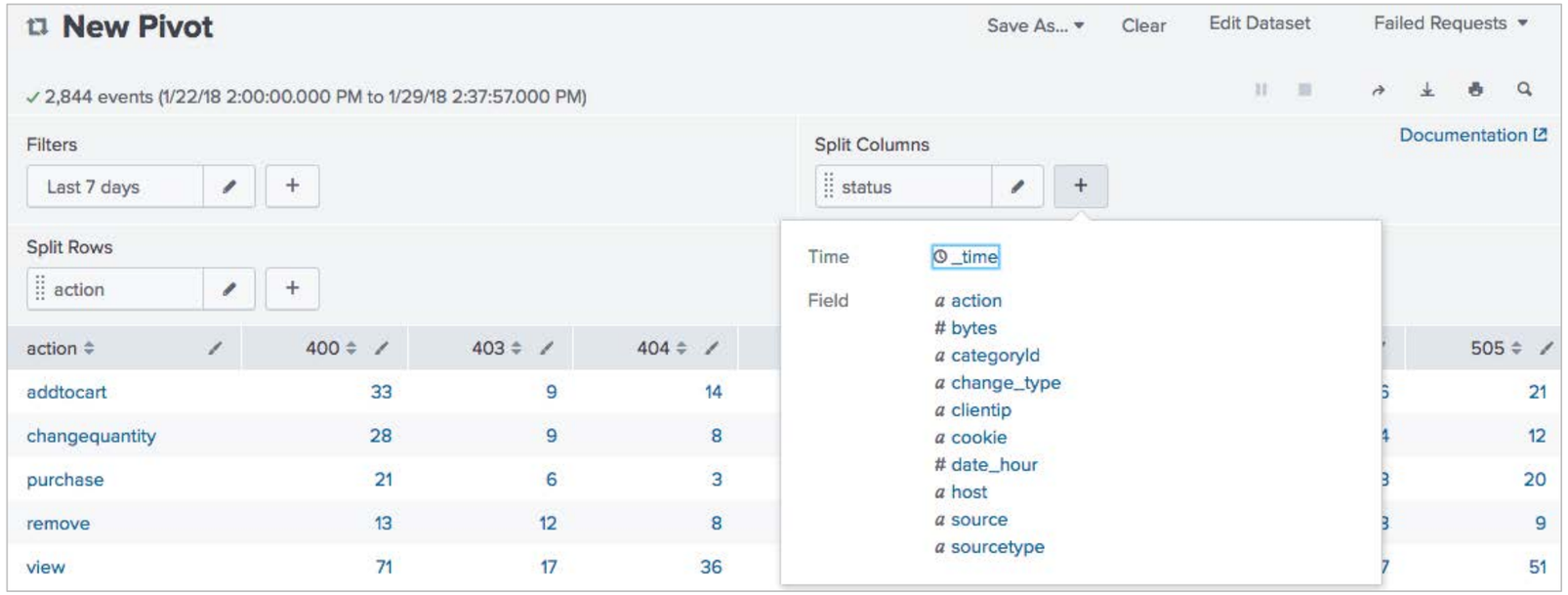
The New Pivot window automatically populates with a count of events for the selected dataset

Generated for (v.ms.besetti@jpmchase.com) (C) Splunk Inc, not for distribution

Copyright © 2018 Splunk, Inc. All rights reserved | 14 June 2018

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

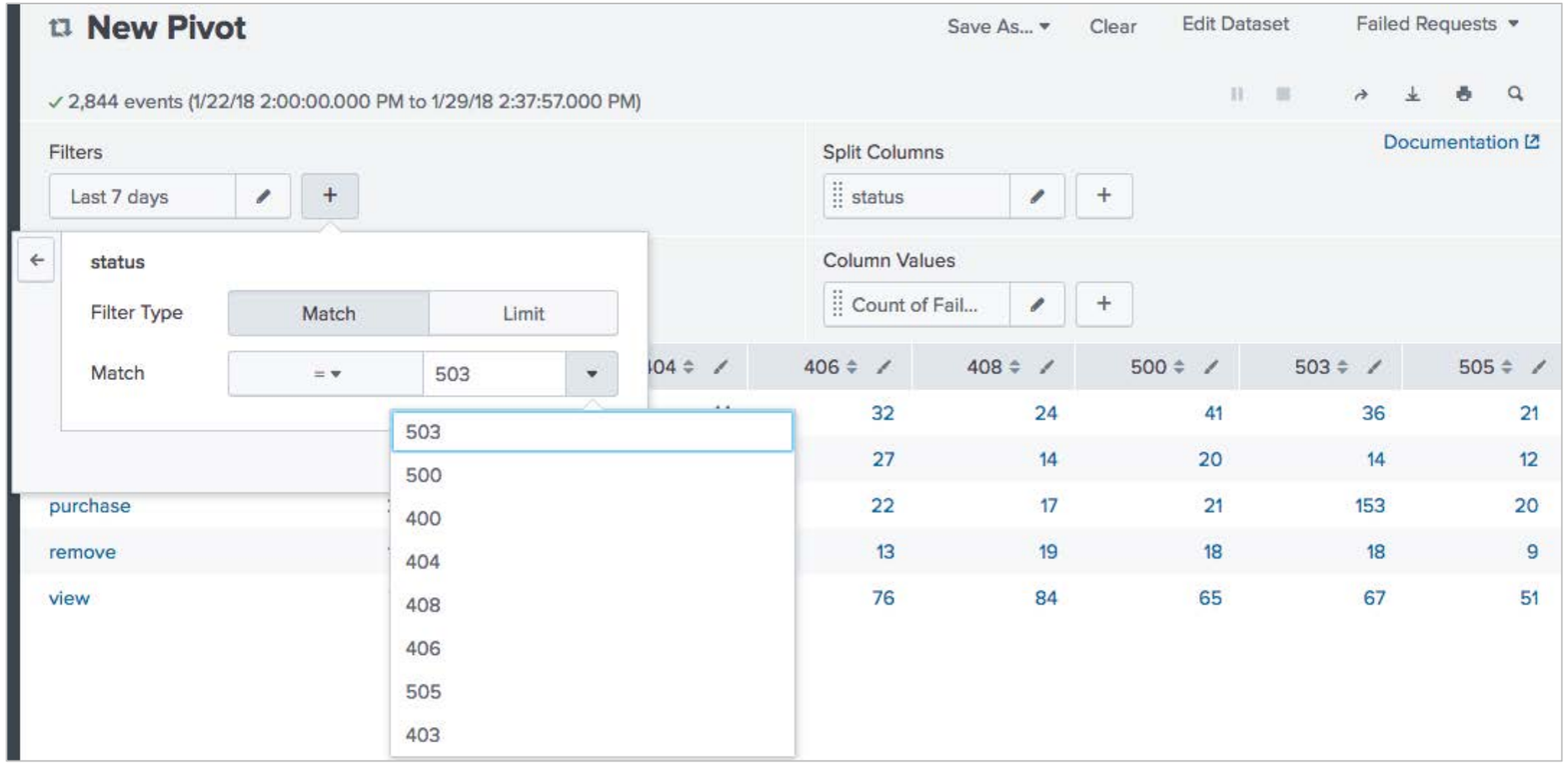


Generated for (v.ms.besetti@jpmchase.com) (C) Splunk Inc, not for distribution

Copyright © 2018 Splunk, Inc. All rights reserved | 14 June 2018

Pivot – Using Fields (cont.)

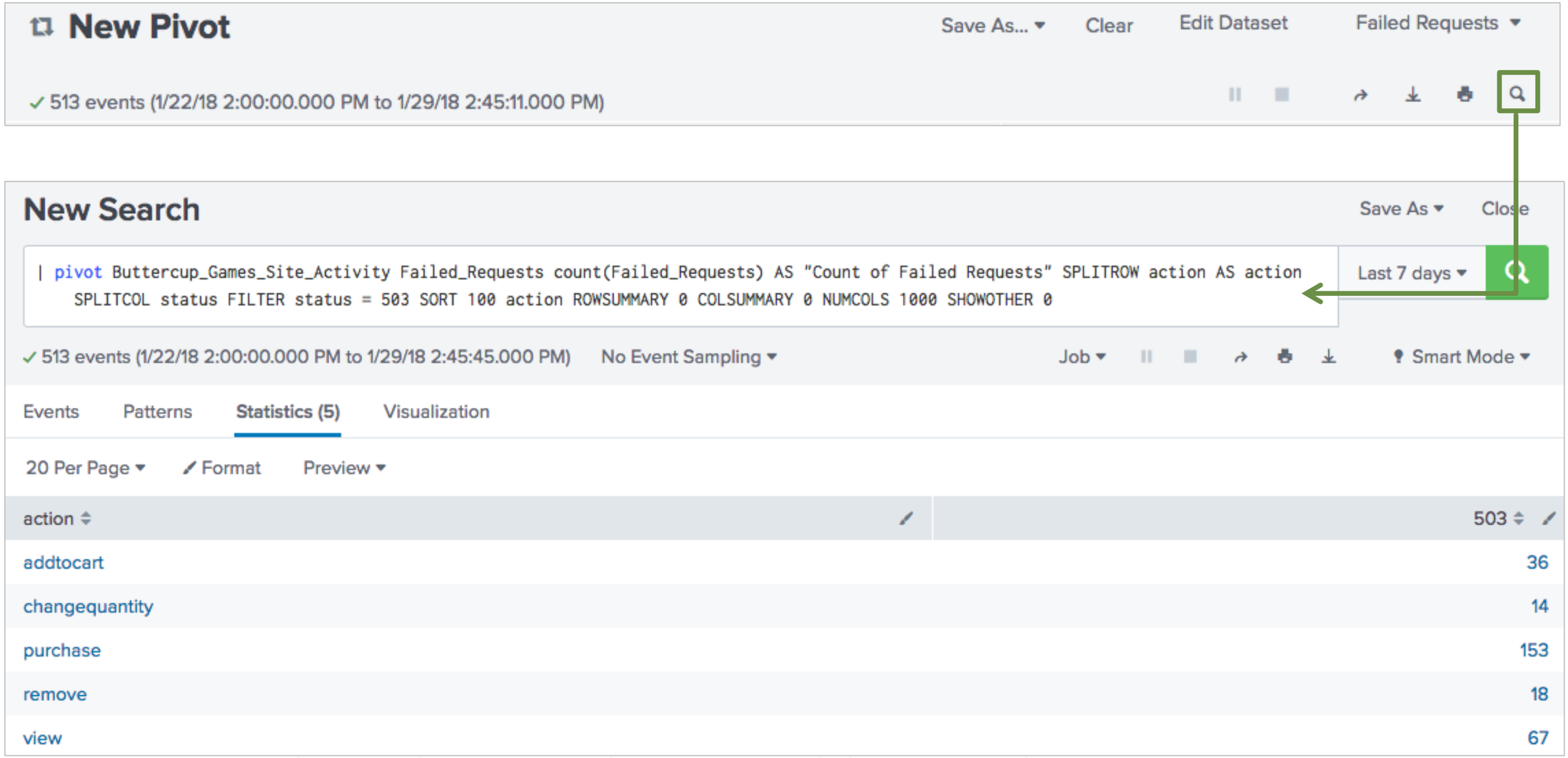
- Fields can also be used to filter events in the Pivot interface
- In this example, the Pivot report is filtered to only return results where status=503



Generated for (v.ms.besetti@jpmchase.com) (C) Splunk Inc, not for distribution

Copyright © 2018 Splunk, Inc. All rights reserved | 14 June 2018

Underlying Search



Generated for (v.ms.besetti@jpmchase.com) (C) Splunk Inc, not for distribution

Copyright © 2018 Splunk, Inc. All rights reserved | 14 June 2018

Underlying Search (cont.)

Data model name

Object name

```
| pivot Buttercup_Games_Site_Activity failed_request
count(failed_request) 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

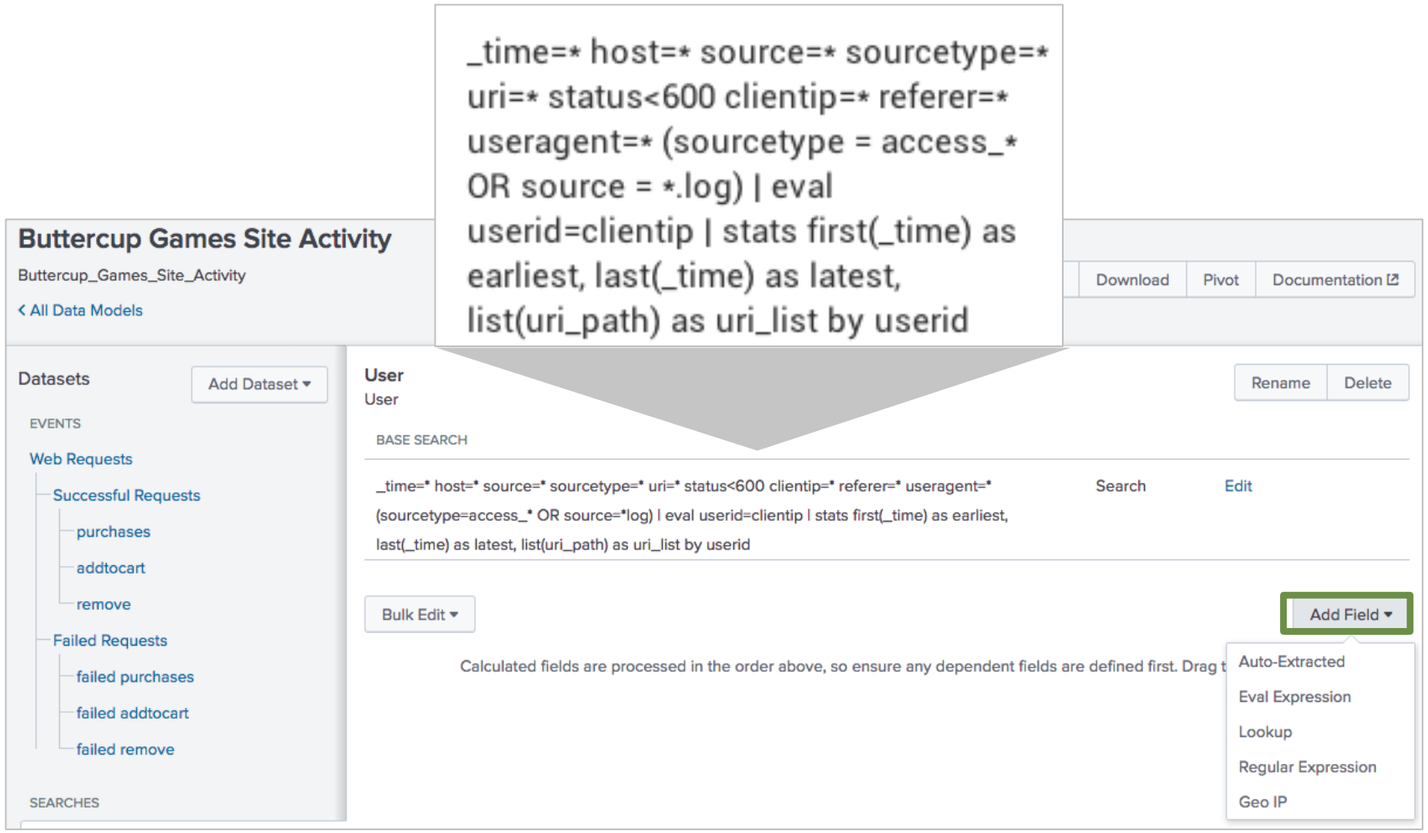
```
SPLITCOL status
FILTER status = 503
```

Generated for (v.ms.besetti@jpmchase.com) (C) Splunk Inc, not for distribution

Copyright © 2018 Splunk, Inc. All rights reserved | 14 June 2018

Data Model Search Datasets

- Arbitrary searches that include transforming commands to define the dataset that they represent
- Search datasets can also have fields, which are added via the **Add Field** button

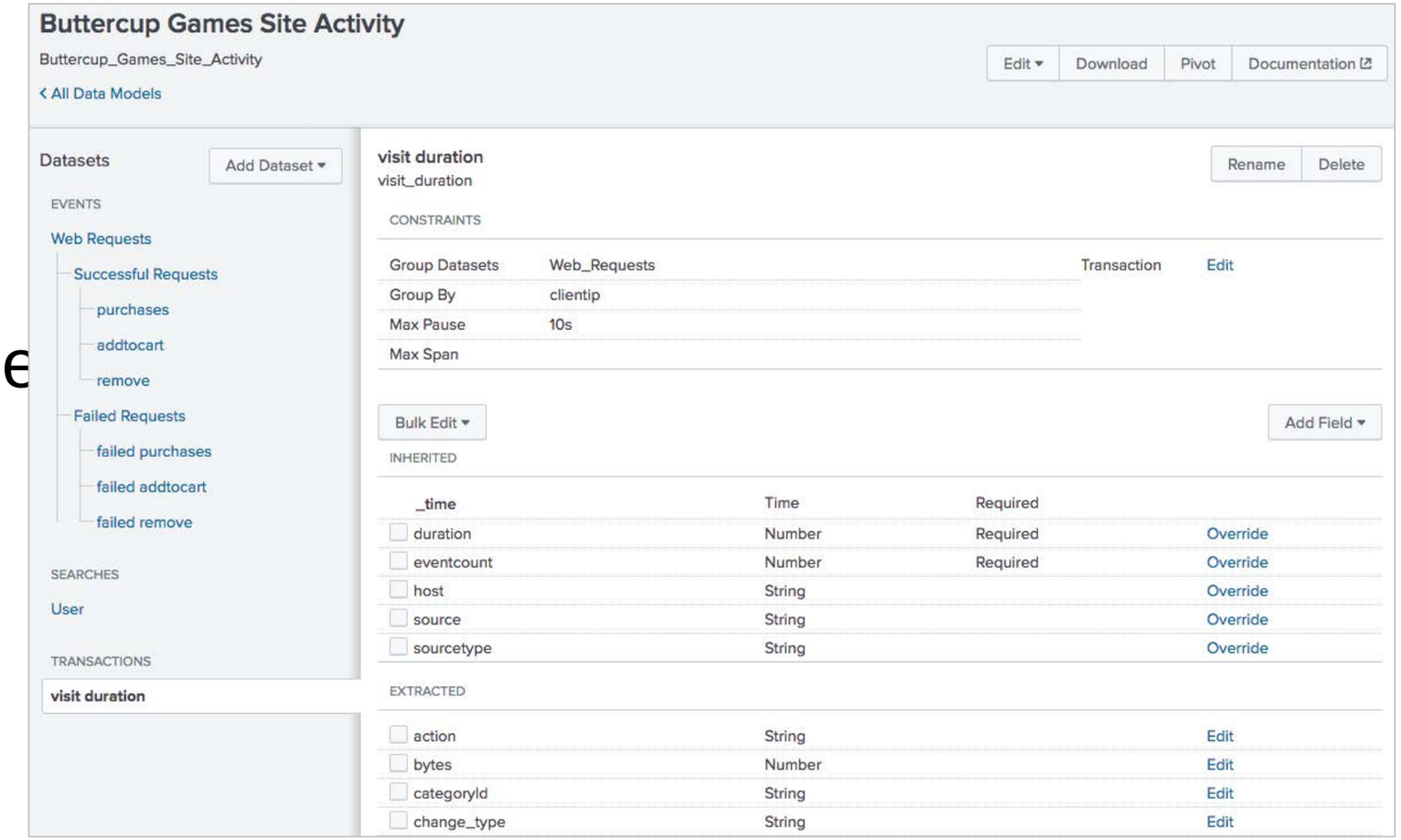


Generated for (v.ms.besetti@jpmchase.com) (C) Splunk Inc, not for distribution

Copyright © 2018 Splunk, Inc. All rights reserved | 14 June 2018

Data Model Transaction Datasets

- Enable the creation of datasets that represent transactions
- Use fields that have already been added to the model using event or search datasets

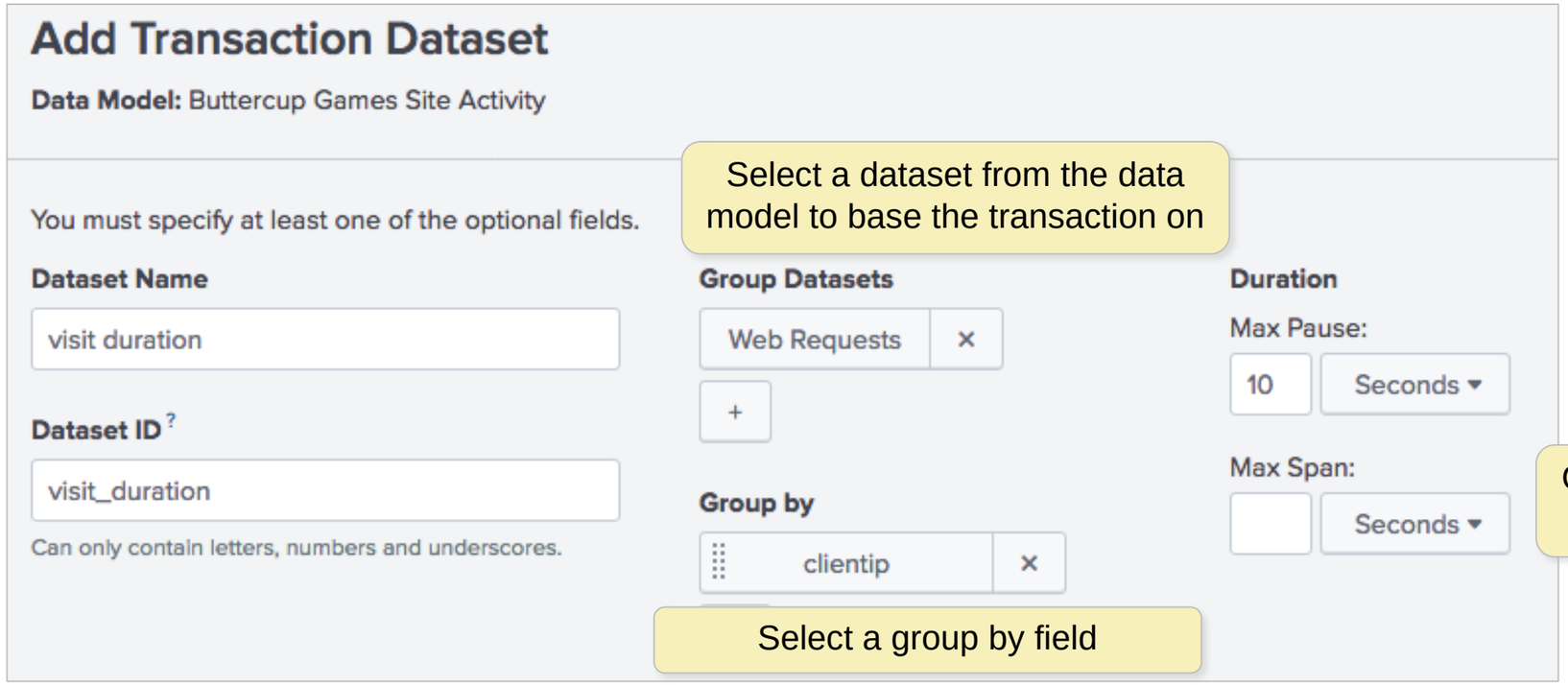


Generated for (v.ms.besetti@jpmchase.com) (C) Splunk Inc, not for distribution

Copyright © 2018 Splunk, Inc. All rights reserved | 14 June 2018

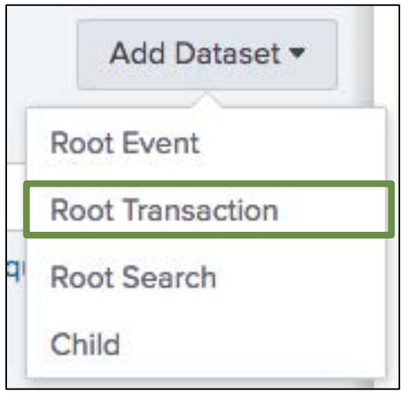
Adding a Transaction

- You can add a transaction to the data model
- The transaction dataset below would equate to the search:
  `sourcetype=access_* | transaction clientip maxpause=10s`



Select a dataset from the data model to base the transaction on

Select a group by field



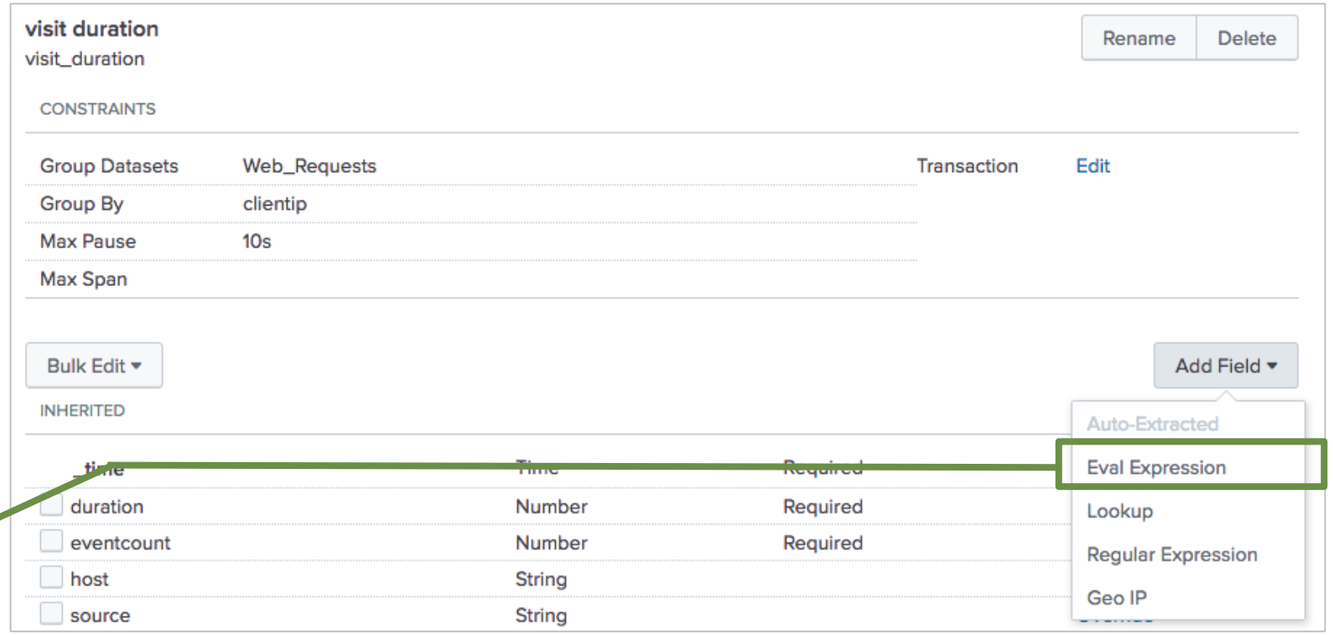
Optionally select max pause and max span

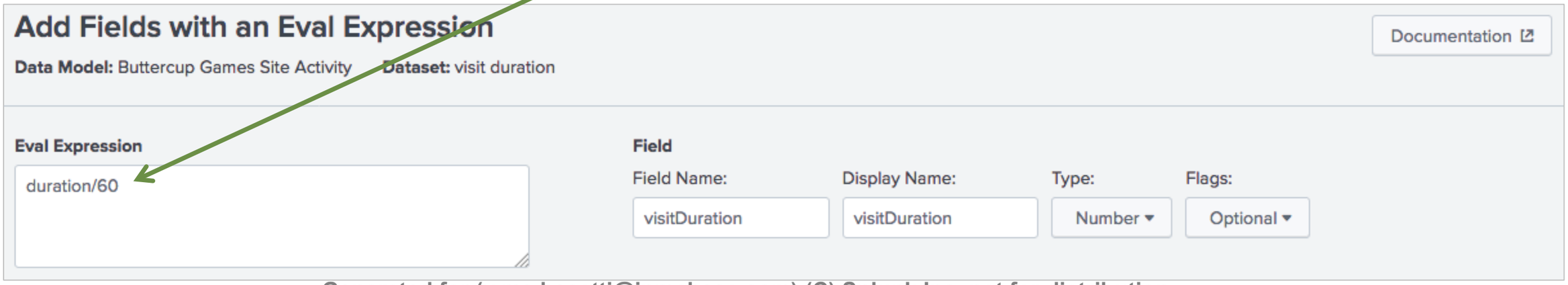
Generated for (v.ms.besetti@jpmchase.com) (C) Splunk Inc, not for distribution

Copyright © 2018 Splunk, Inc. All rights reserved | 14 June 2018

Adding a Transaction (cont.)

- You can then 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





Generated for (v.ms.besetti@jpmchase.com) (C) Splunk Inc, not for distribution

Copyright © 2018 Splunk, Inc. All rights reserved | 14 June 2018

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

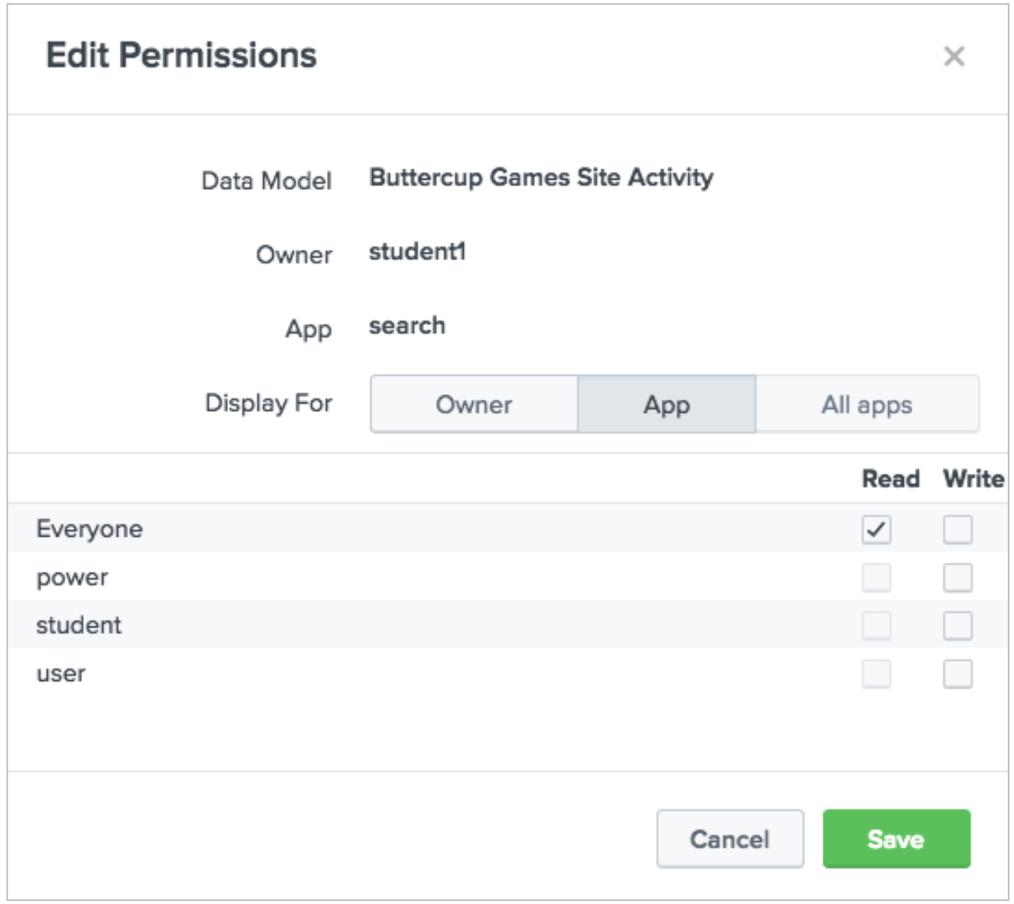
Data model acceleration is discussed later in this module.

Generated for (v.ms.besetti@jpmchase.com) (C) Splunk Inc, not for distribution

Copyright © 2018 Splunk, Inc. All rights reserved | 14 June 2018

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



Generated for (v.ms.besetti@jpmchase.com) (C) Splunk Inc, not for distribution

Copyright © 2018 Splunk, Inc. All rights reserved | 14 June 2018

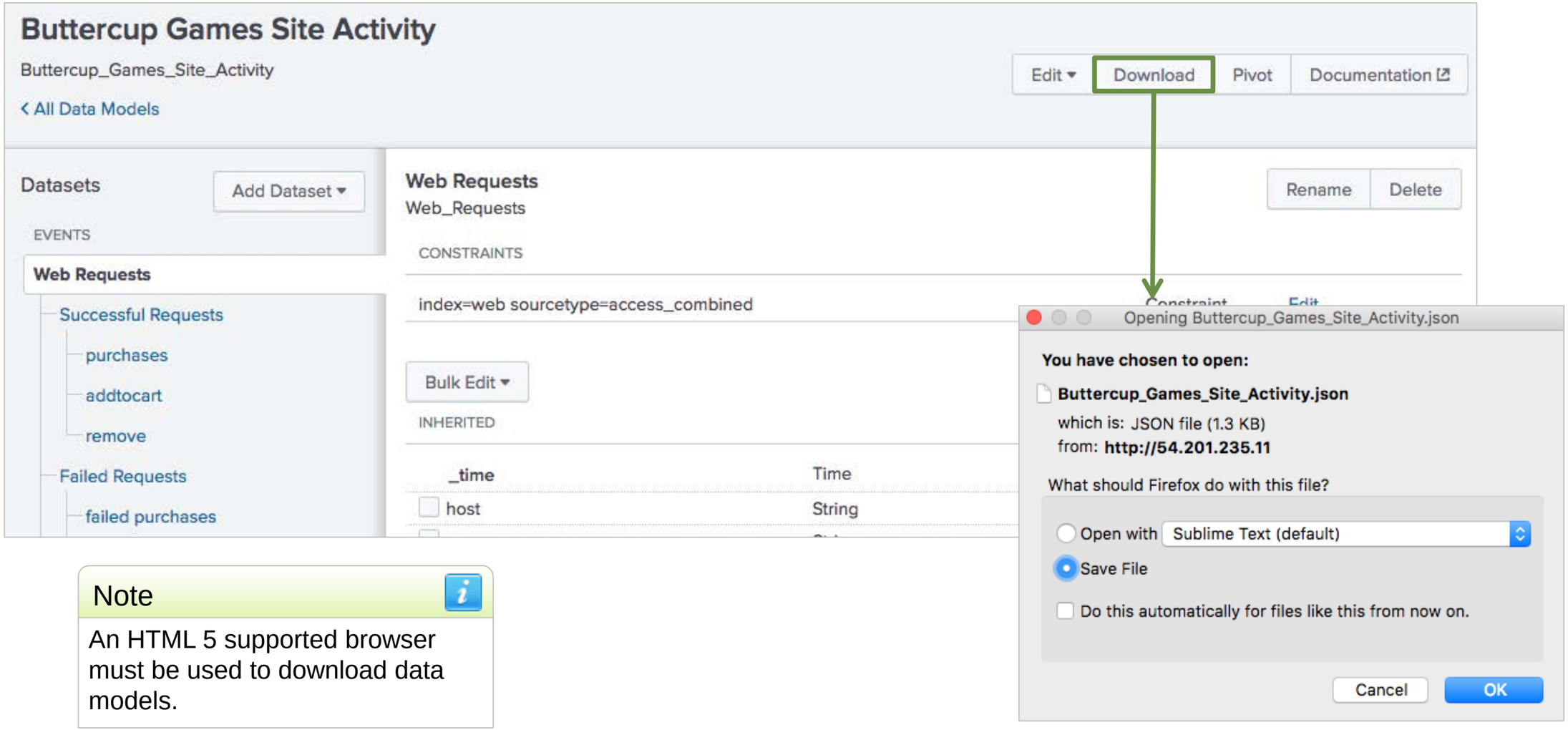
Download and Upload Data Models

- 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

Generated for (v.ms.besetti@jpmchase.com) (C) Splunk Inc, not for distribution

Copyright © 2018 Splunk, Inc. All rights reserved

Downloading a Data Model



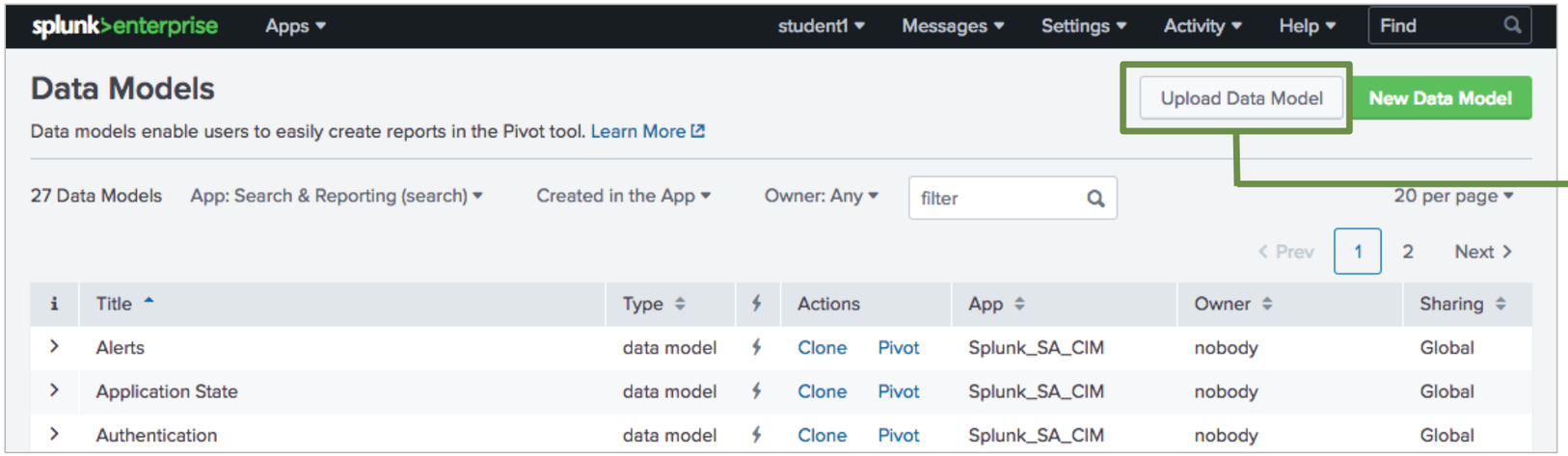
Note

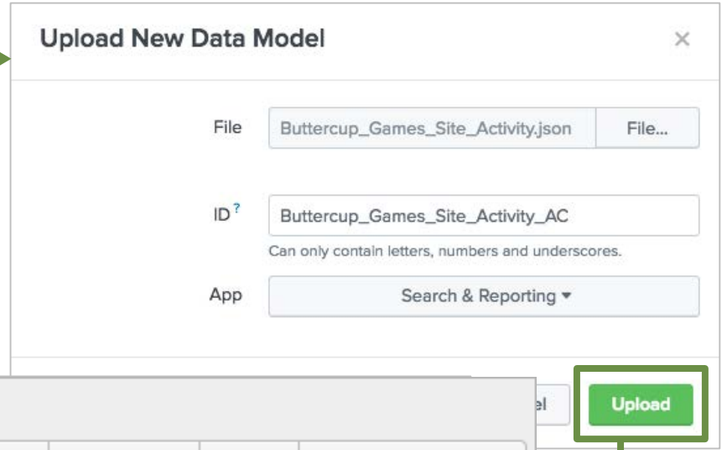
An HTML 5 supported browser must be used to download data models.

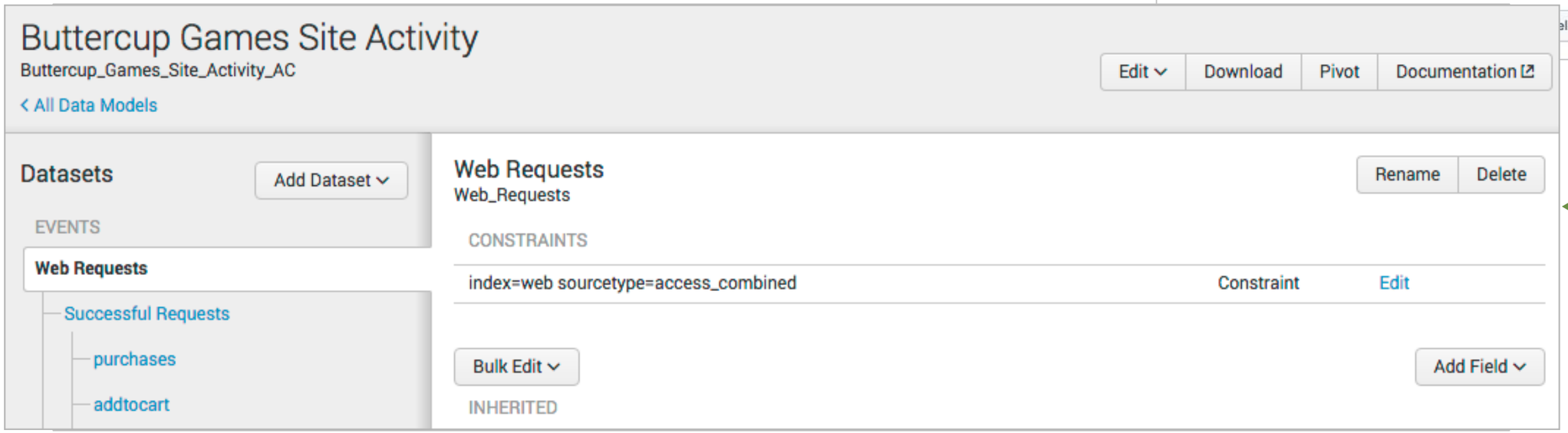
Generated for (v.ms.besetti@jpmchase.com) (C) Splunk Inc, not for distribution

Copyright © 2018 Splunk, Inc. All rights reserved | 14 June 2018

Uploading a Data Model







Generated for (v.ms.besetti@jpmchase.com) (C) Splunk Inc, not for distribution

Copyright © 2018 Splunk, Inc. All rights reserved | 14 June 2018

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 *Advanced Searching and Reporting*

Generated for (v.ms.besetti@jpmchase.com) (C) Splunk Inc, not for distribution

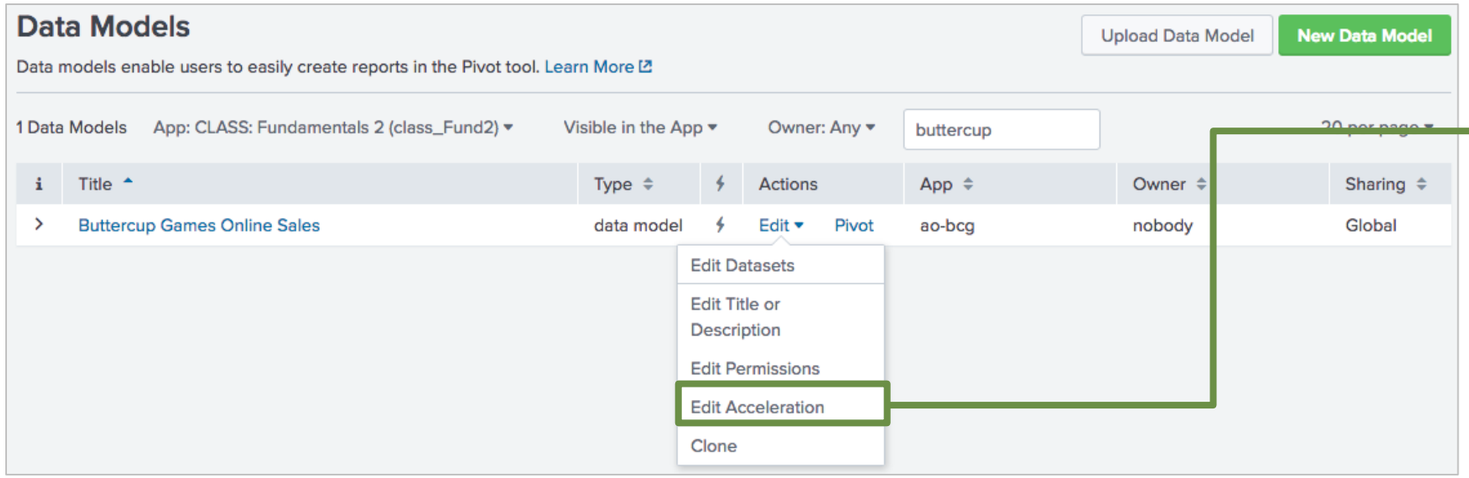
Copyright © 2018 Splunk, Inc. All rights reserved | 14 June 2018

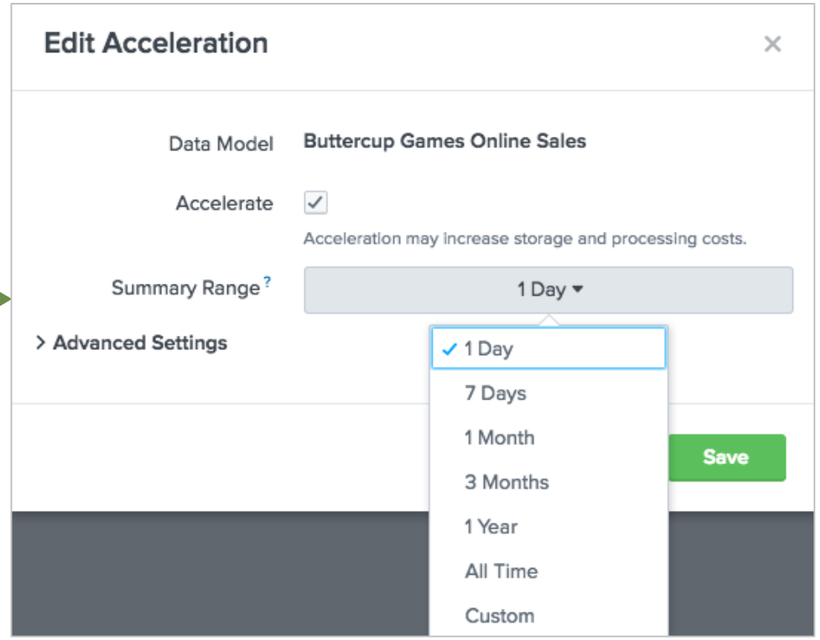
Accelerating a Data Model

- 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

Note

Only root events can be accelerated. If there are multiple root events, only the first root event is accelerated.





Generated for (v.ms.besetti@jpmchase.com) (C) Splunk Inc, not for distribution

Copyright © 2018 Splunk, Inc. All rights reserved | 14 June 2018

Module 14: Using the Common Information Model (CIM) Add-On

Generated for (v.ms.besetti@jpmchase.com) (C) Splunk Inc, not for distribution

Copyright © 2018 Splunk, Inc. All rights reserved

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

Generated for (v.ms.besetti@jpmchase.com) (C) Splunk Inc, not for distribution

Copyright © 2018 Splunk, Inc. All rights reserved | 14 June 2018

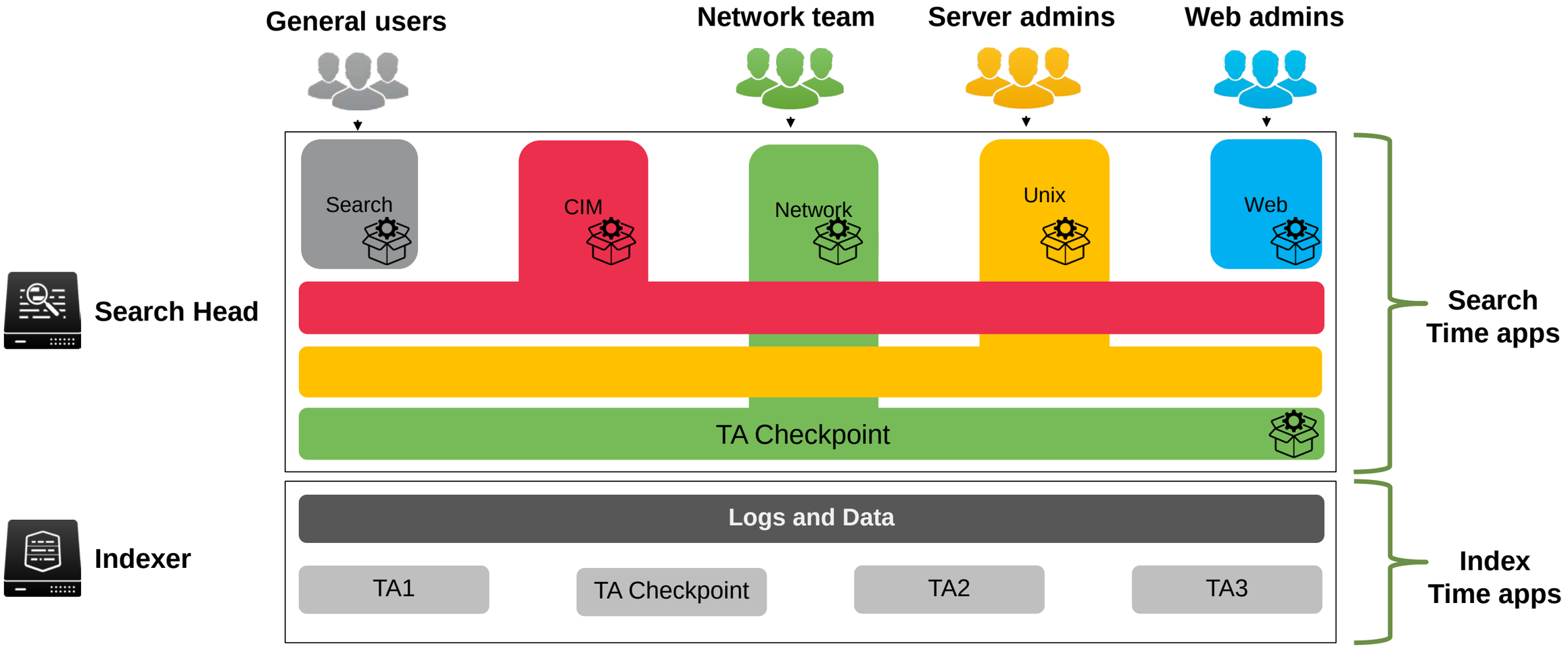
What is the Common Information Model (CIM)?

- 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

Generated for (v.ms.besetti@jpmchase.com) (C) Splunk Inc, not for distribution

Copyright © 2018 Splunk, Inc. All rights reserved | 14 June 2018

How the Splunk CIM Works



Generated for (v.ms.besetti@jpmchase.com) (C) Splunk Inc, not for distribution

Copyright © 2018 Splunk, Inc. All rights reserved | 14 June 2018

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.	

Generated for (v.ms.besetti@jpmchase.com) (C) Splunk Inc, not for distribution

Copyright © 2018 Splunk, Inc. All rights reserved | 14 June 2018

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 (v.ms.besetti@jpmchase.com) (C) Splunk Inc, not for distribution

Copyright © 2018 Splunk, Inc. All rights reserved | 14 June 2018

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 (v.ms.besetti@jpmchase.com) (C) Splunk Inc, not for distribution

Copyright © 2018 Splunk, Inc. All rights reserved | 14 June 2018

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/Howtousethesereferencetables

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

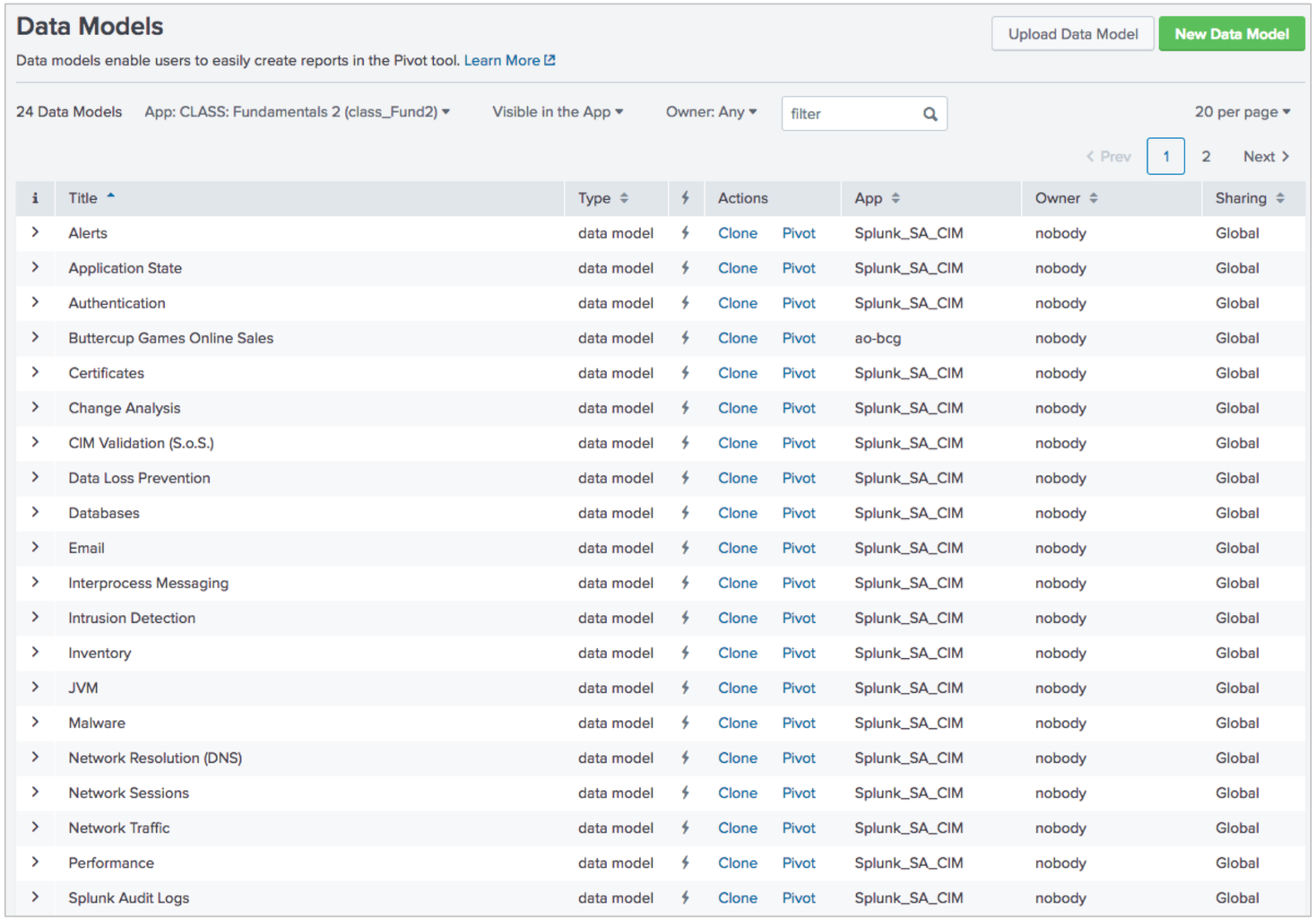
The data models included in the CIM add-on are configured with data model acceleration turned off.

Generated for (v.ms.besetti@jpmchase.com) (C) Splunk Inc, not for distribution

Copyright © 2018 Splunk, Inc. All rights reserved | 14 June 2018

Using the CIM Add-on

1. Examine your data
   - Go to **Settings > Data Models**
   - Identify a data model relevant to your dataset



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

Best Practice

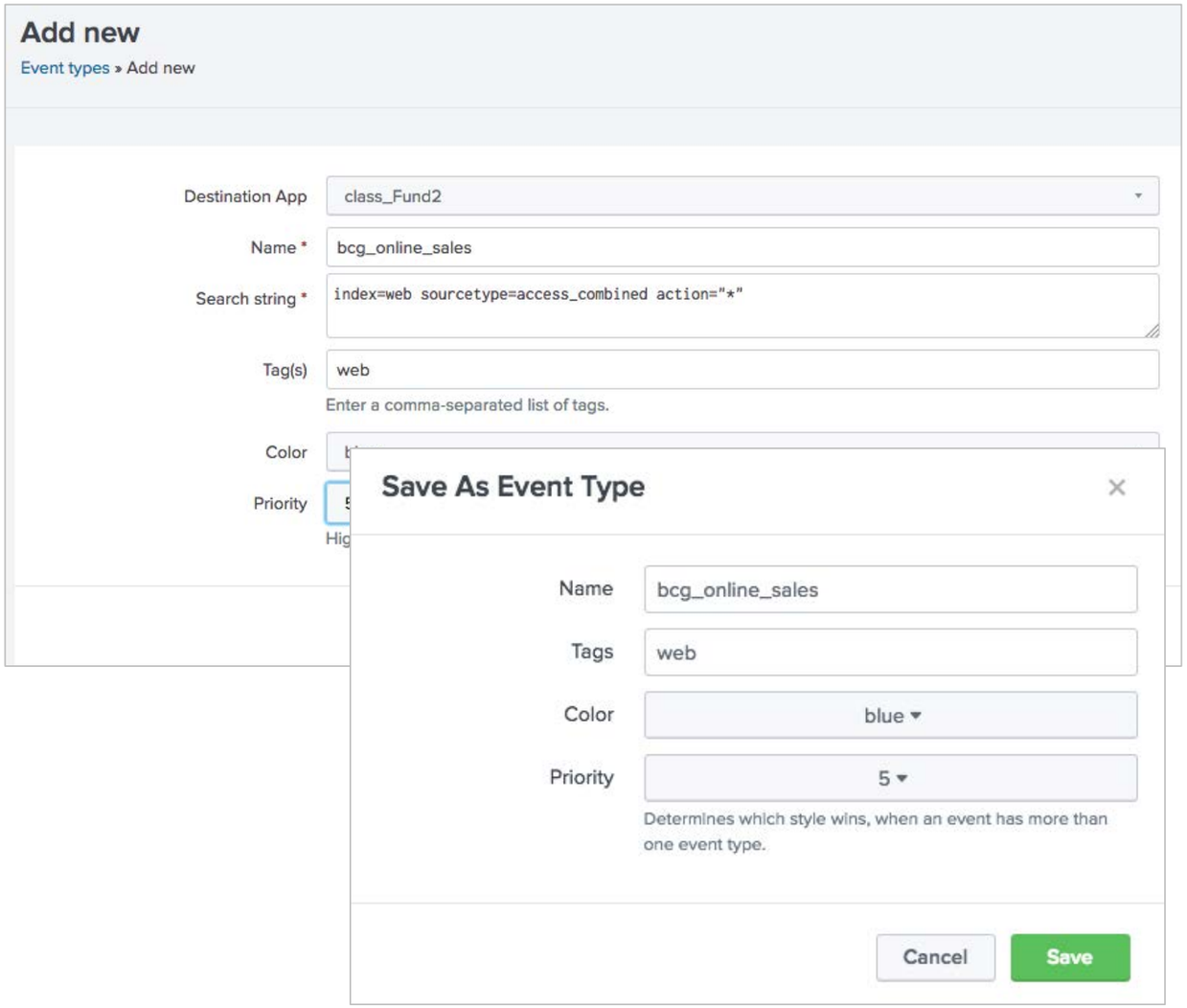
Keep the CIM Reference Tables in Splunk Docs page open in a separate tab.

Generated for (v.ms.besetti@jpmchase.com) (C) Splunk Inc, not for distribution

Copyright © 2018 Splunk, Inc. All rights reserved | 14 June 2018

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

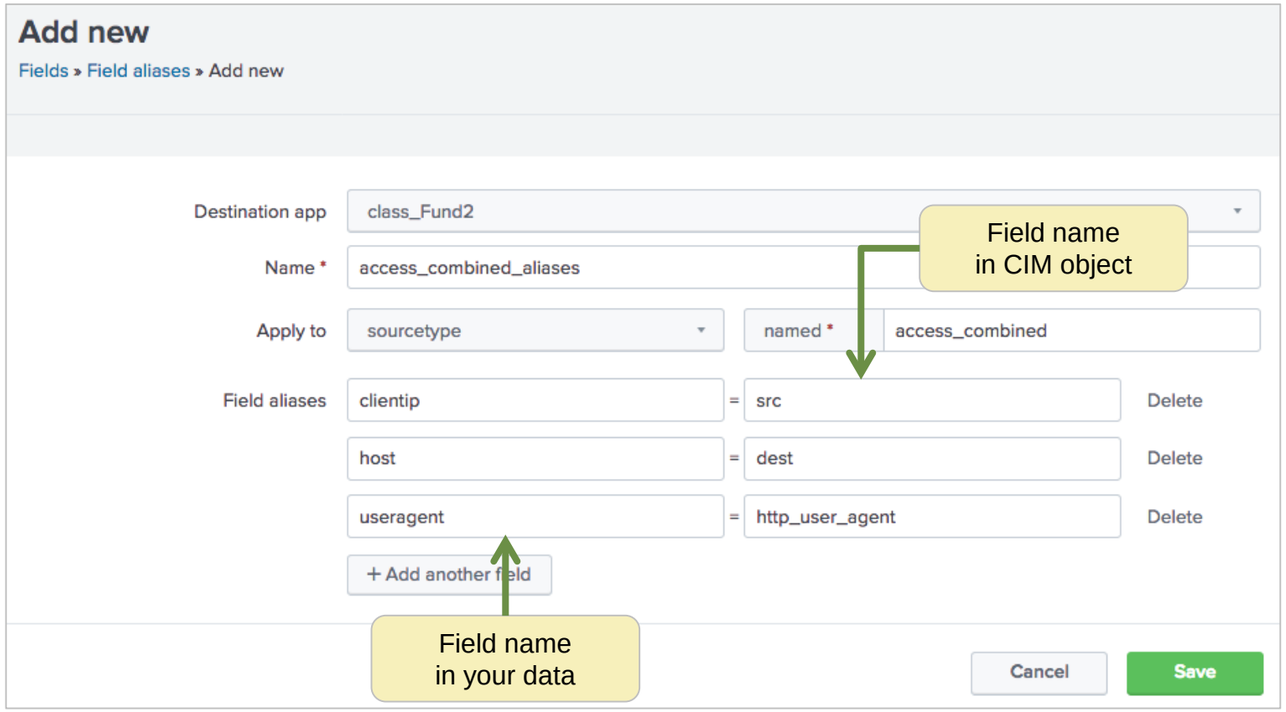


Generated for (v.ms.besetti@jpmchase.com) (C) Splunk Inc, not for distribution

Copyright © 2018 Splunk, Inc. All rights reserved | 14 June 2018

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

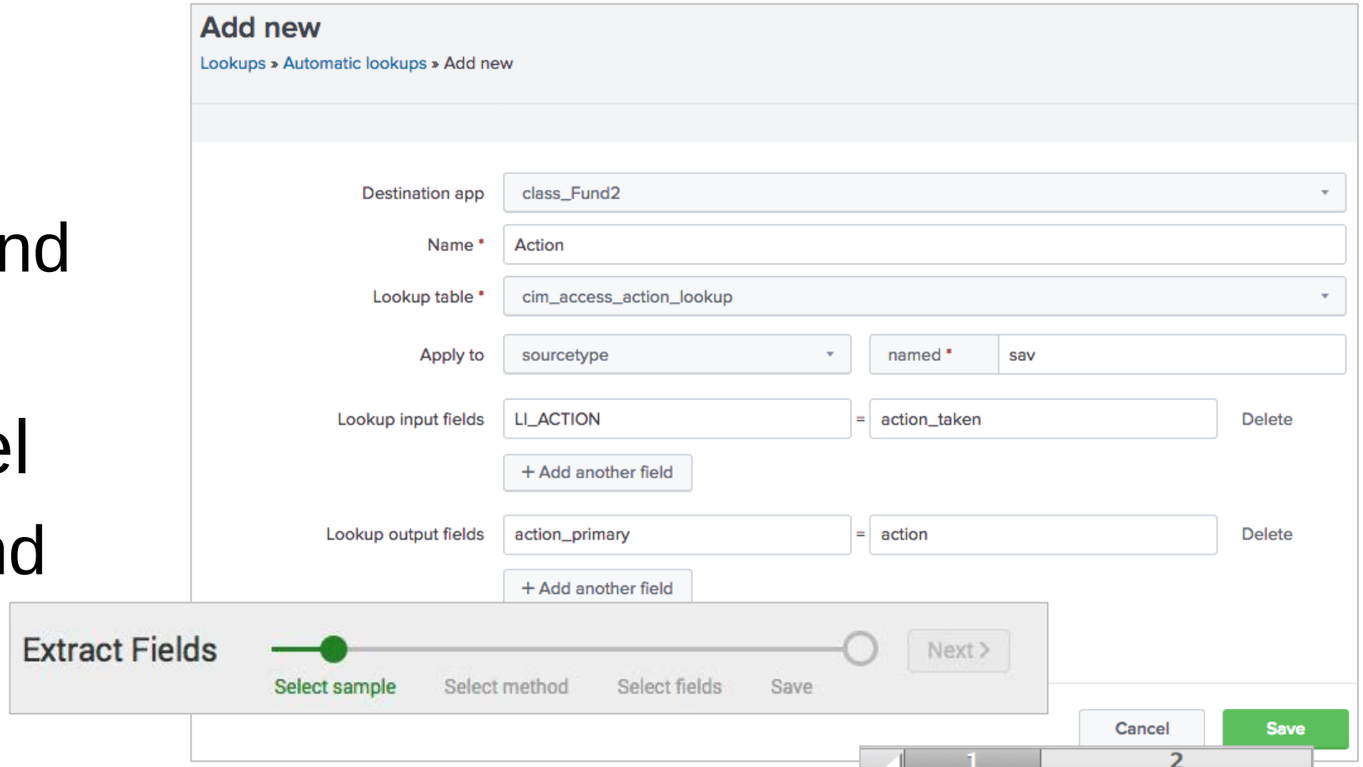


Generated for (v.ms.besetti@jpmchase.com) (C) Splunk Inc, not for distribution

Copyright © 2018 Splunk, Inc. All rights reserved | 14 June 2018

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



	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

Note

For more information, see the *Common Information Model Add-on Manual*: http://docs.splunk.com/Documentation/CIM/latest/User/Overview

Generated for (v.ms.besetti@jpmchase.com) (C) Splunk Inc, not for distribution

Copyright © 2018 Splunk, Inc. All rights reserved | 14 June 2018

`datamodel` Command

- Search against a specified data model object
- Return a description of all or a specified data model and its objects
- Is a generating command and should be the first command in the pipeline

datamodel Learn More

Allows user to examine data models and run the search for a datamodel object.

Example:
`| datamodel`

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.

Generated for (v.ms.besetti@jpmchase.com) (C) Splunk Inc, not for distribution

Copyright © 2018 Splunk, Inc. All rights reserved | 14 June 2018

datamodel Command – Example

```
| datamodel Web Web search | fields Web*
```

A B C D E

- A Command
- B Data model name
- C Data model dataset name
- D Command
- E Find field names with Web prefix

Dataset name prepended to field names in your data

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

Note

When using the `datamodel` command, the data model name and dataset name are case-sensitive.

Generated for (v.ms.besetti@jpmchase.com) (C) Splunk Inc, not for distribution

Copyright © 2018 Splunk, Inc. All rights reserved | 14 June 2018

`from` Command

- Retrieves data from a data model or named dataset
- Must be the first command in a search
- Different than just using `datamodel`
  - `datamodel` returns **all** fields prepended with data model name
  - `from datamodel` returns specified fields only

```
| from datamodel:"internal_server.splunkdaccess"
```

- `from` can also retrieve data from saved searches, reports, or lookup files

```
| from savedsearch:mysecurityquery
```

Generated for (v.ms.besetti@jpmchase.com) (C) Splunk Inc, not for distribution

Copyright © 2018 Splunk, Inc. All rights reserved | 14 June 2018

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

Generated for (v.ms.besetti@jpmchase.com) (C) Splunk Inc, not for distribution

Copyright © 2018 Splunk, Inc. All rights reserved | 14 June 2018