# Introduction of Splunk

**Splunk** is a powerful platform for analyzing machine data. In simple words, Splunk is a software or engine which can be used to Search, Monitor, Investigate, Troubleshoot, Alert and Report any data.

Splunk works in three phases,

1. **Data Identification:** Identify the machine data which can answer the requirement.
2. **Data Transformation:** Transform the machine data into results which can answer given requirement.
3. **Data Visualization:** Visualize data in the form of reports, interactive graphs and charts to get valuable insights.

# Installation

This section is going to cover Splunk installation. We can download free enterprise version of Splunk from below URL:

<https://www.splunk.com/en_us/download/splunk-enterprise.html>

To start Splunk from Unix or Windows, Direct to Splunk installation folder/bin and execute below command,

1. **./splunk start:** This command is used to start Splunk instance.
2. **./splunk stop:** This command is used to stop Splunk instance.
3. **./splunk restart:** This command is used to restart(stop,start) Splunk instance.

# Splunk Processing Components

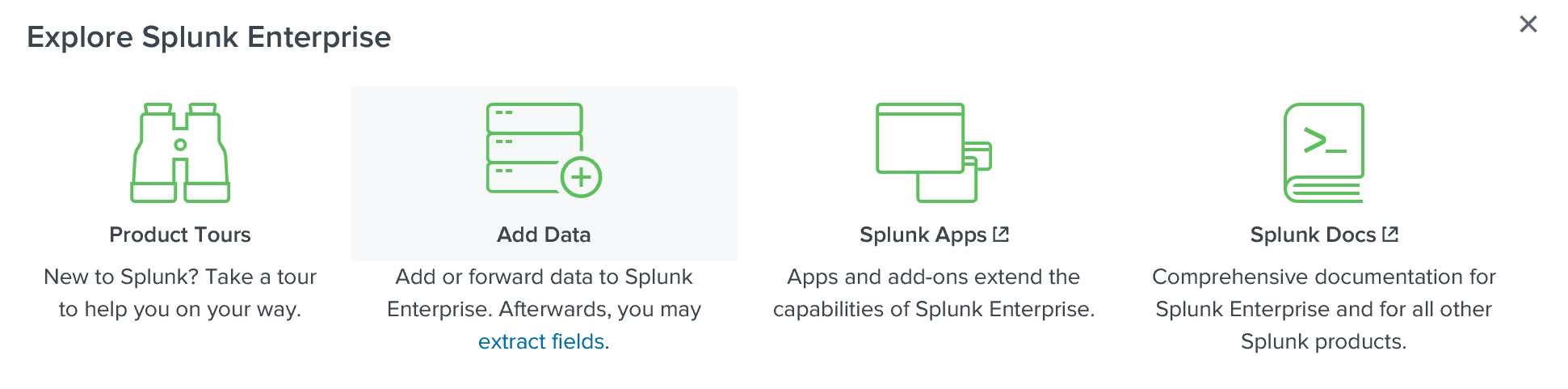
There are three processing components that Splunk use to make data available, accessible and visible.

1. **Indexer:** Splunk Indexer is a database of Splunk where it stores all the logs in smart way. It process incoming machine data and store the results in indexes as events. It creates number of files organized in sets of directories by age.
2. **Search Head:** Splunk search head is front end for Splunk, Commonly accessed via web browser. This allows Splunk search language to search the data. All the incoming requests from users will be distributed across all the indexes. This also provide user’s to create Dashboards, Alerts, and Reports. Search head can run searches across multiple indexes at the same time.
3. **Forwarder:** Forwarders are Splunk Enterprise instances which consume data from different sources and distribute to indexes. These act as Splunk agent to collect logs from remote machines. Splunk forwarders are of two types,
4. Universal Forwarder (UF): This is a Splunk Enterprise agent installed on non Splunk system to gather the logs. This cannot parse or index the data.
5. Heavy Weight Forwarder (HWF): This is a full instance of Splunk Enterprise agent with advanced capabilities like parsing and possible data filtering. It works as a remote collector.

# Searching with Splunk

The Goal of the search is to filter, visualize and summarize large amount of data, in order to answer user requirements about data. Few points on Splunk Search Dashboard,

1. **Add/Forward Data:** Data can be added to Splunk in multiple ways. One way is manually adding the files using Splunk Web Interface. Other way is configuring Splunk Universal Forwarder/ Heavy Forwarder. We can add data in Splunk Enterprise using “Add Data” section.

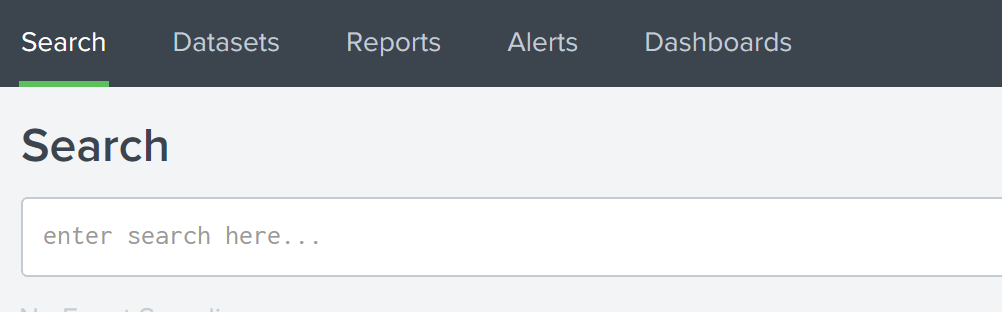


1. **Universal Forwarder Configuration:** There four configurations we have consider in order to setup universal forwarder. They are,
   1. Inputs.conf: Controls how the forwarder collects data.
   2. Outputs.conf: Controls how the forwarder sends data to an indexer or other forwarder.
   3. Server.conf: For connection and performance tuning.
   4. Deploymentclient.conf: For connecting to a deployment server.

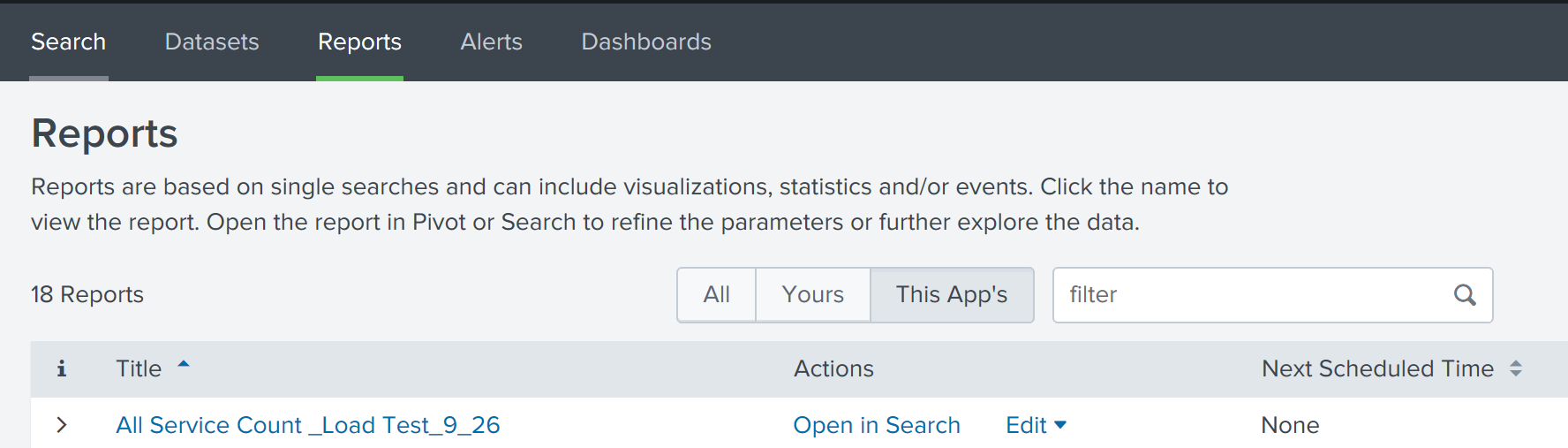
Once universal forwarder is installed, we can configure this using Command Line Interface. The list of commands can send the data to indexes using forwarder:

* 1. splunk add forward-server <host name or ip address>:<listening port>
  2. splunk set deploy-poll <host name or ip address>:<management port>
  3. splunk add monitor C:\Testing\Log\test.log
  4. splunk restart

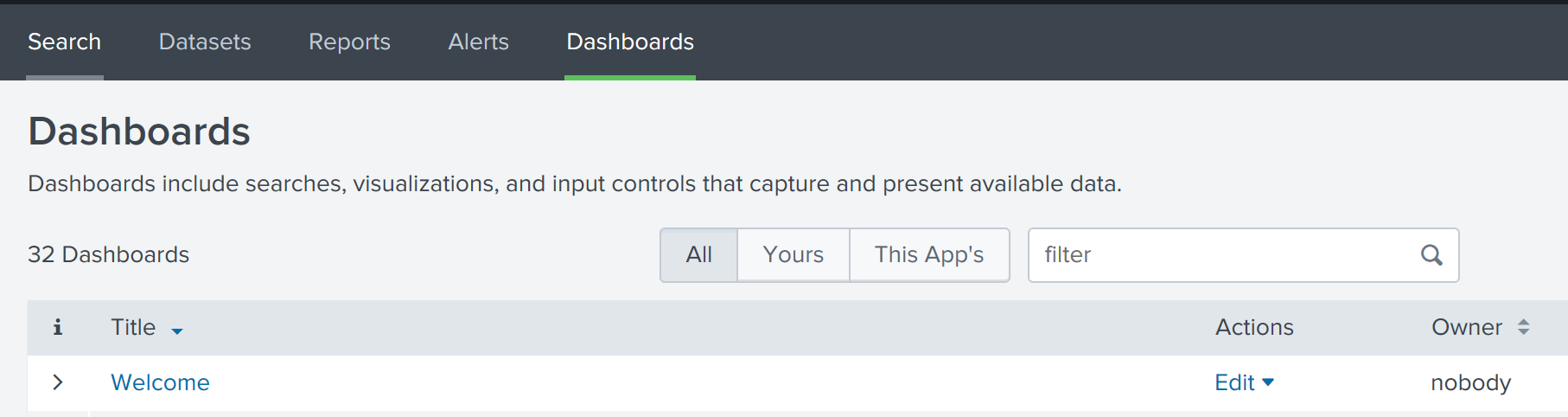
1. **Search:** This is similar to google search box, where we can enter a search query to retrieve the data. For example, if we want to get all the rows having a string “Apple” in it, we can just type do search like “index=\* Apple”. This search is going to return us all the indexes(data) having this string Apple. We can enter complex queries into this search box using operators like AND, OR, NOT etc. We will discuss advanced operations in the later part of this document.



1. **Datasets:** This is a place where we can view all the uploaded datasets into Splunk. We can open them by clicking on specific dataset file.
2. **Reports:** Splunk reports are based on single searches. These can include visualizations, statistics and/or events. We can see all the created reports in the Reports Section.



1. **Alerts:** Alerts set a condition which triggers an action, such as sending an email that contains the results of the triggering search to a list of people or execute a script. We can create alerts using Search head and all the alerts will be saved in Alerts section.
2. **Dashboards:** Dashboard is group of reports arranged as panels. Dashboards include searches, visualizations, and input controls the capture and present available data.



# SPL: Search Processing Language

Splunk helps us to query massive datasets using Search Processing Language. These queries convert the indexed events into meaningful answers. There are five components used in Splunk Search Language. They are,

1. **Search Terms:** These are general search terms we want to identify in the events. Example search terms are “error”, “success”, “exception” etc.
2. **Commands:** Splunk commands tells Splunk, what we want to do with the search result. Example commands are “stats”, “dedup”, “evalstats” etc.
3. **Functions:** Splunk functions are used along with commands to format data in given way. Example functions are “list”, “count” etc
4. **Arguments:** Arguments are passed to Splunk functions. Example arguments are “list(product-name)”, count(\*) etc
5. **Clauses:** Clauses are used to group the results or combine them. Example clauses are “as”, “where” etc

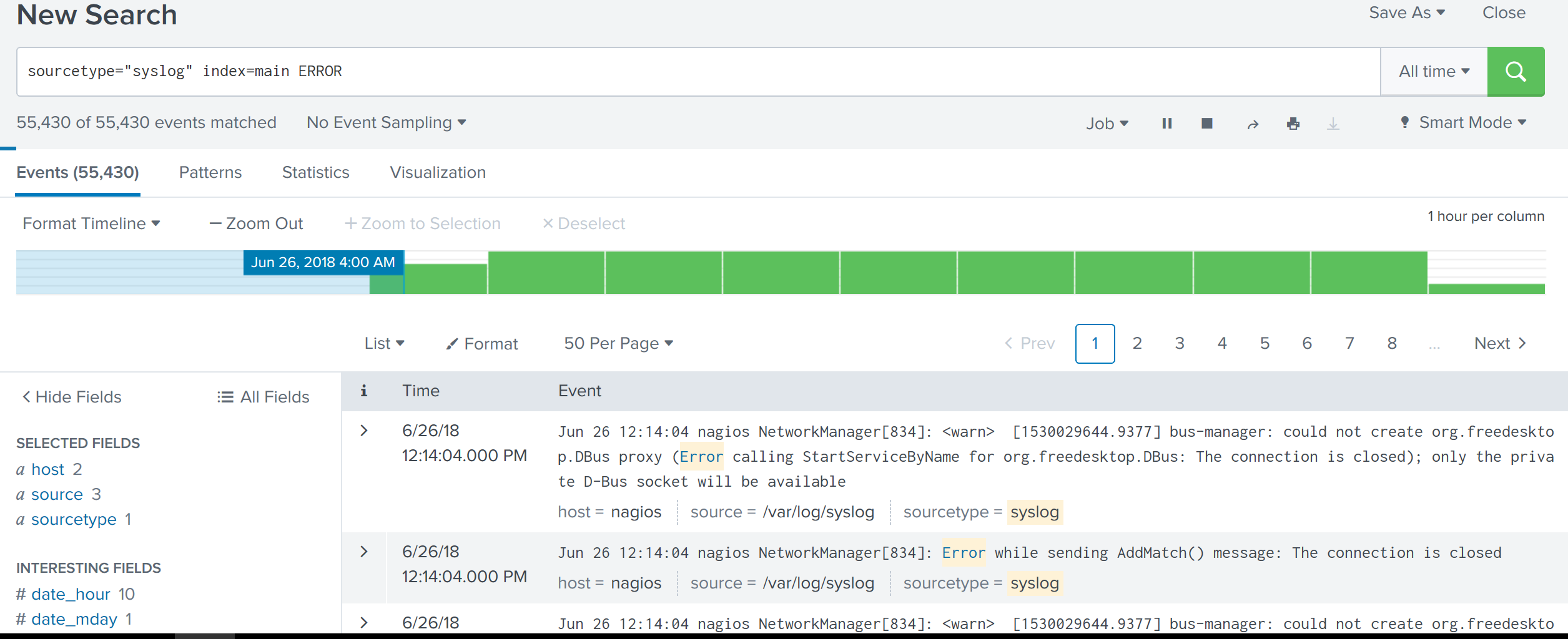
Complete query with all the five components,

“sourcename=account status=200 |stats list(productname) as game”

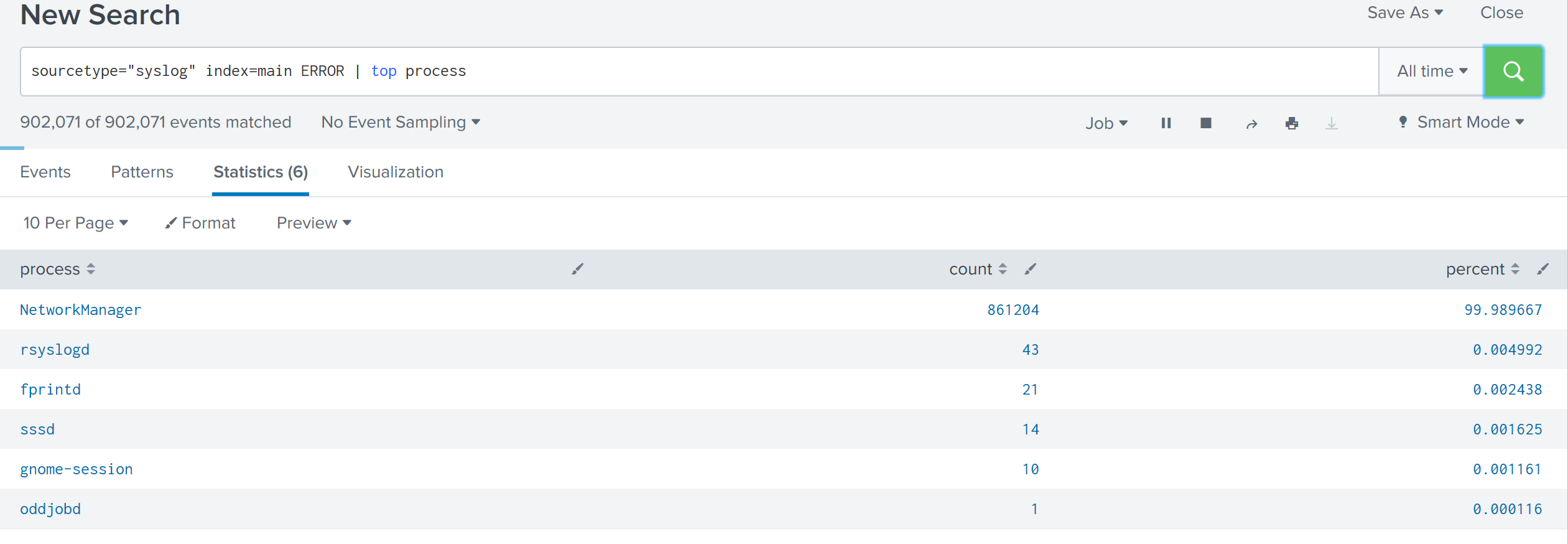
In the above example we have PIPE(|) which is used to separated individual commands that make up the search. Observe the sequence of operations below for given search query,

**Search Query:** “sourcetype="syslog" index=main ERROR | top process | fields - percent”

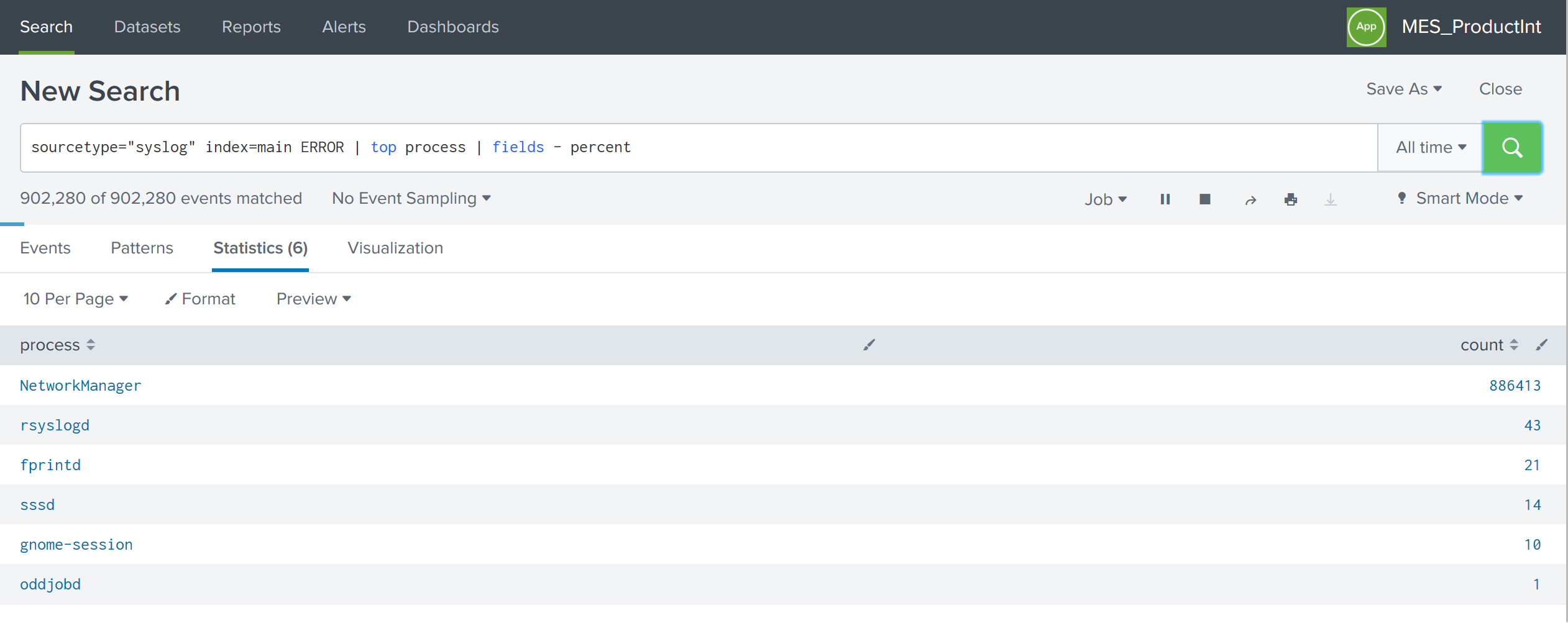
First part of the query returns all the events having ERROR.



Second part of the query is going to return top 10 most common values of a given field.



Third part of the query, with an argument(- percent) will remove a field from returned search results.



# Splunk Commands

Splunk commands tell Splunk, what we want to do with the search result. Some of Splunk commands are discussed below,

**Fields:** This command is used to either include or exclude a specific field from search results.  
Include Example: sourcetype=access-combined | field stats clients : This will include stats & clients  
Exclude Example: sourcetype=access-combined | field - stats clients: This will exclude stats

**Table:** Table command is used to display search results in tabular format.  
Example: sourcetype=access-combined | table stats, client, status

**Rename:** Rename command is used to rename field names.  
Example: sourcetype=access-combined | rename stats as “statistics” | table statistics, client

**Dedup:** Dedup command is used to remove events with duplicate entries.  
Example: sourcetype=access-combined | dedup clients | table stats,clients

**Sort:** Sort command is used to sort the events using a field name.  
Desc Example: sourcetype=access-combined | sort – sales\_price vendor : This will order in descending  
Asc Example: sourcetype=access-combined | sort sales\_price vendor : This will order in descending  
  
**Where:** Where command is used like SQL where. It is used to define a condition.  
Example: sourcetype=access-combined | where sales\_price > 10

**Head:** Head command is used to return the first count results. Using Head permits a search to stop retrieving events from disk when it finds the desired number of results.  
Example: sourcetype=access-combined | where sales\_price > 10 | head 10

**Top:** Top command is used to find most common values of a given field. By default it will product count and percentage of results.  
Example: sourctype=access-combined | top vendor limit=20

**Stats:** Stats command is used to produce statistics of search result. Stats command returns a table of results where each row represents a single unique combination of values of the group by fields.  
Example1: sourcetype="syslog" index=main ERROR | stats count by process  
Example2: sourcetype=”syslog” index=main ERROR | stats dc(product\_name) as prod\_name

**Chart:** Chart command is used to create tabular data output suitable to create charts. We specify the x-axis variable using over or by  
Example: sourcetype=”syslog” index=main ERROR|chart max(delay) OVER foo BY bar

**Replace:** Replace command is used to find and replace of specific field values with replacement values.  
Example: sourcetype=”syslog” | replace Aug WITH AUGUST IN Month

**Eval:** Eval command calculates an expression and puts the resulting value into new field.  
Example: sourcetype=”syslog” | eval status=if(logvalue=0,success,failed)

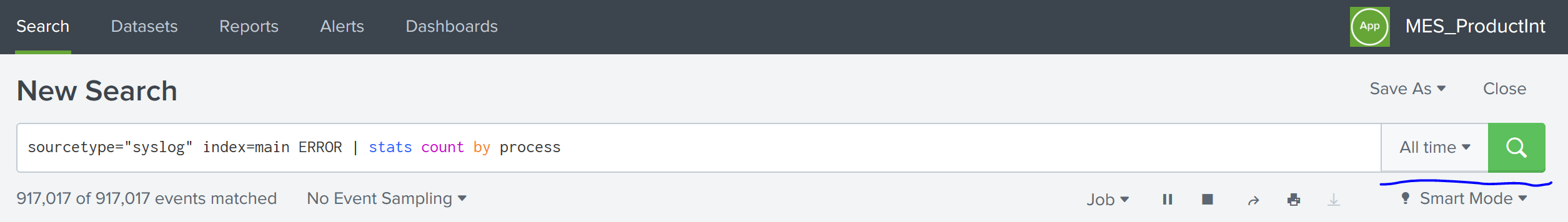
**Lookup:** lookup command manually invokes field lookups from a lookup table, enabling us to add field values from external sources.

# Splunk Reports

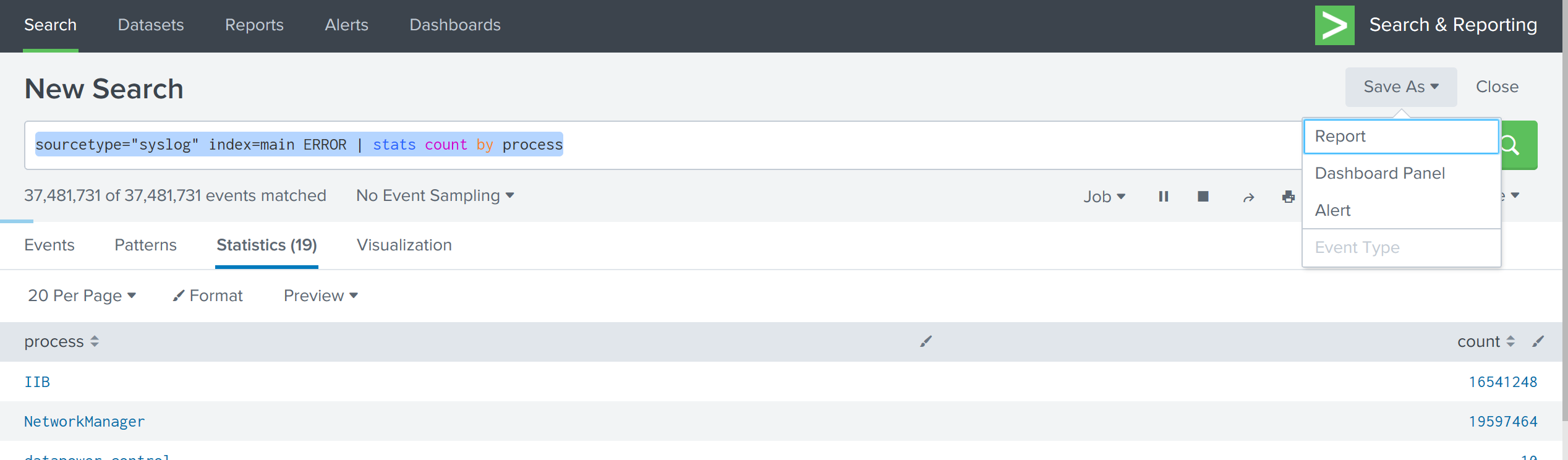
Splunk reports are based on single searches. These can include visualizations, statistics and/or events.

Steps to create Report in Splunk:

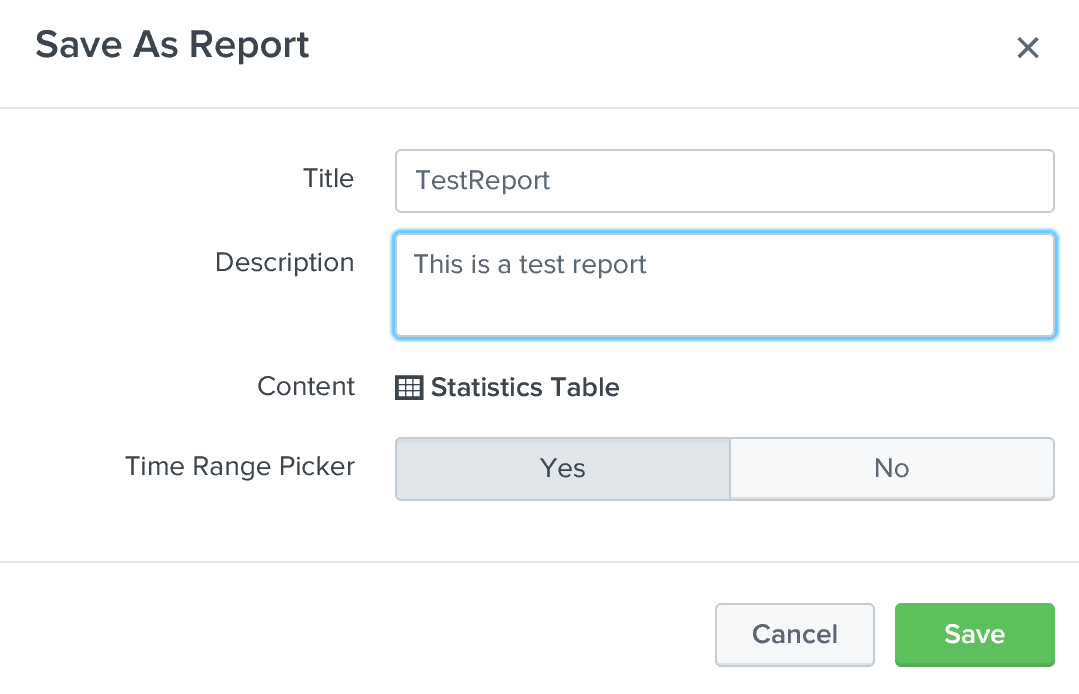
1. Enter Search query in the Search Head and click on Search (green button)



1. We can select time ranges by clicking on drop down at All time. Once search query returns results, Click on Save As -> select report.



1. Save the report with proper name and description.

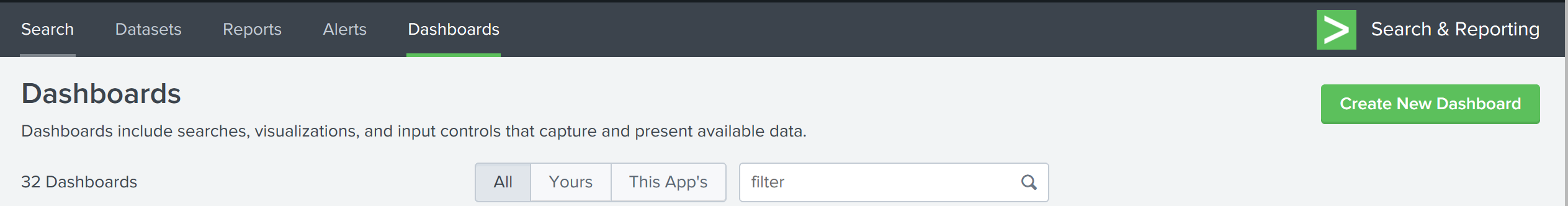


# Splunk Dashboard

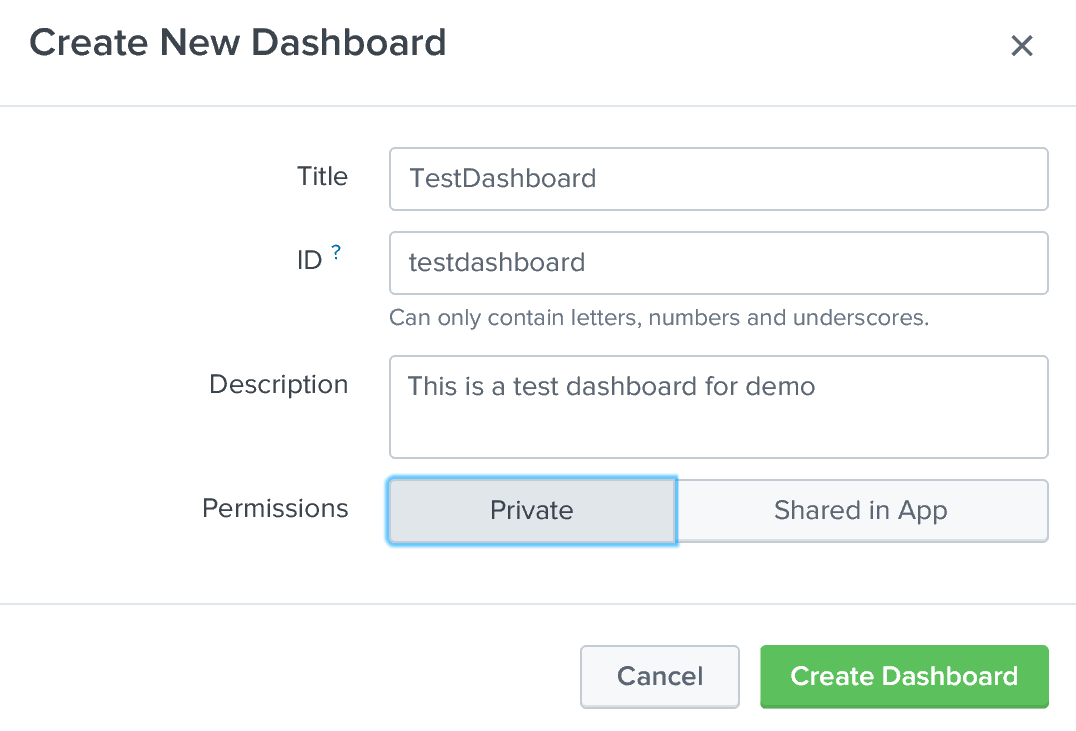
Splunk dashboard is a group of reports arranged as panels. Dashboards include searches, visualizations, and input controls that capture and present available data.

Steps to create Dashboard in Splunk:

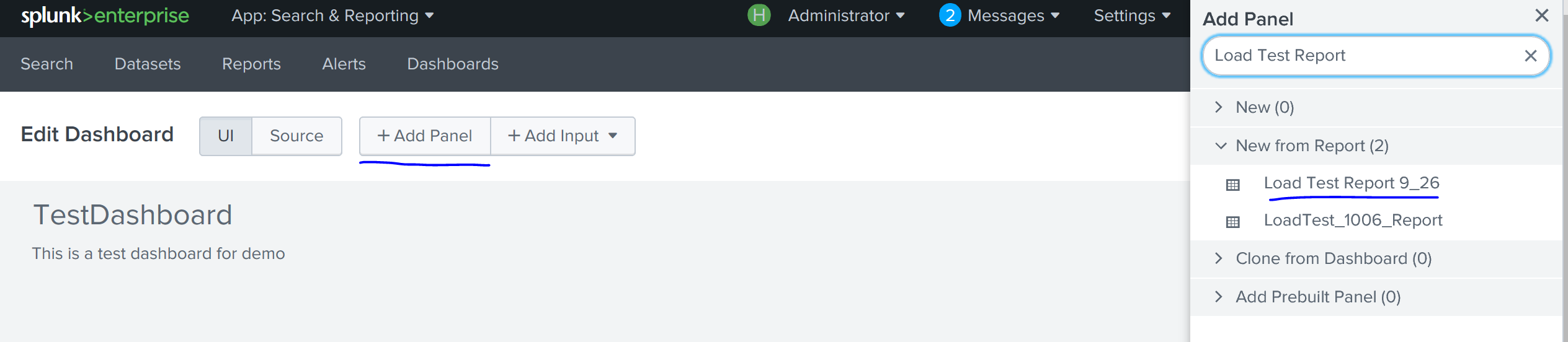
1. Navigate to Dashboards tab and click on “Create New Dashboard”.



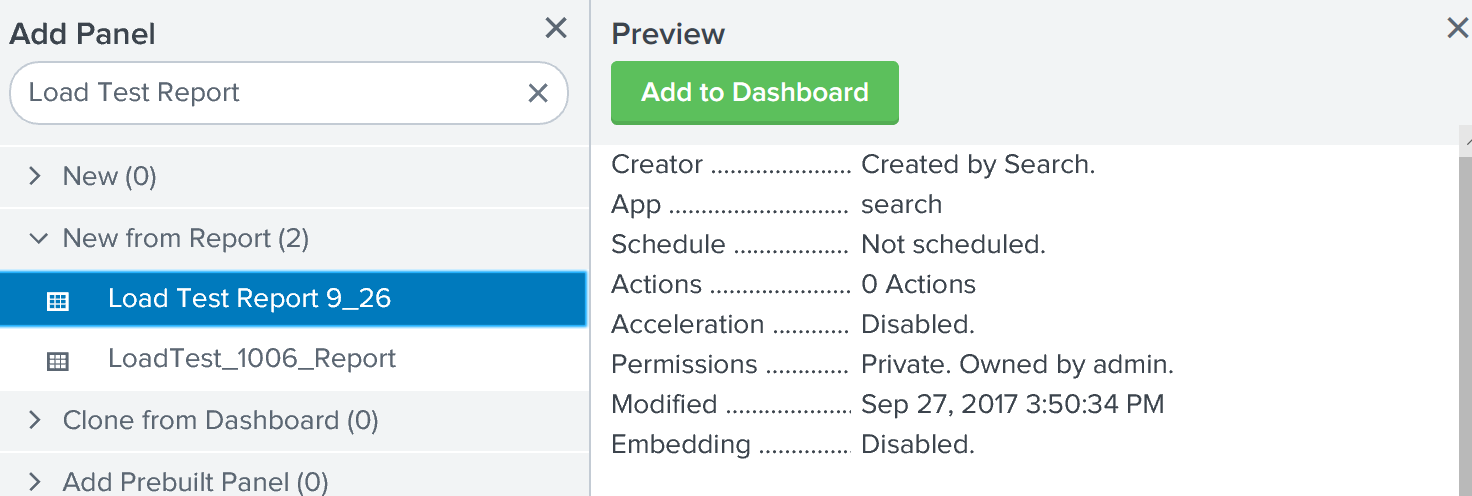
1. Save the Dashboard with appropriate name and description then click Create Dashboard.



1. Click on “Add Panel” from Edit Dashboard Section. This will open Search window to select a report. Enter the report name which we want to add it to Dashboard.



1. Select the report name which will give us the preview of the search query.



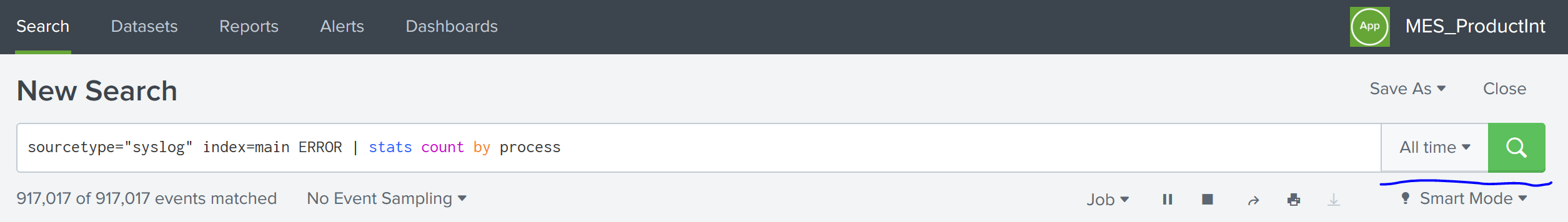
1. Click on Add to Dashboard and click on Save. Repeat the same process to add more panels to the Dashboard.

# Splunk Alerts:

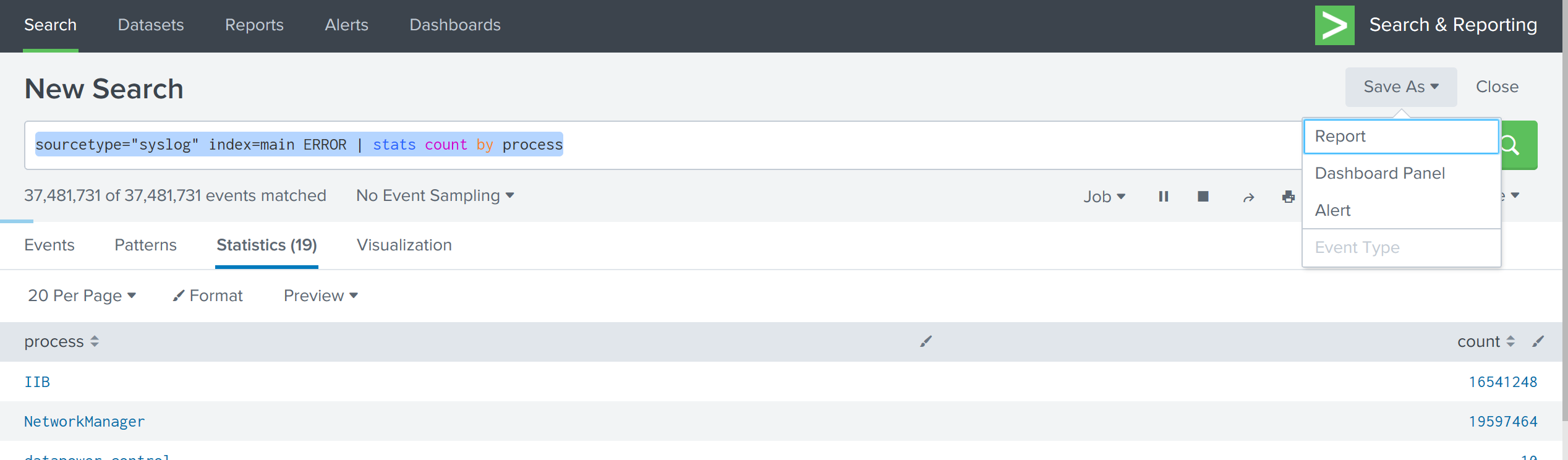
Splunk Alerts set a condition which triggers an action, such as sending an email that contains the results of the triggering search to a list of people or execute a script. We can create alerts using Search head and all the alerts will be saved in Alerts section.

**Steps to create Alerts in Splunk:**

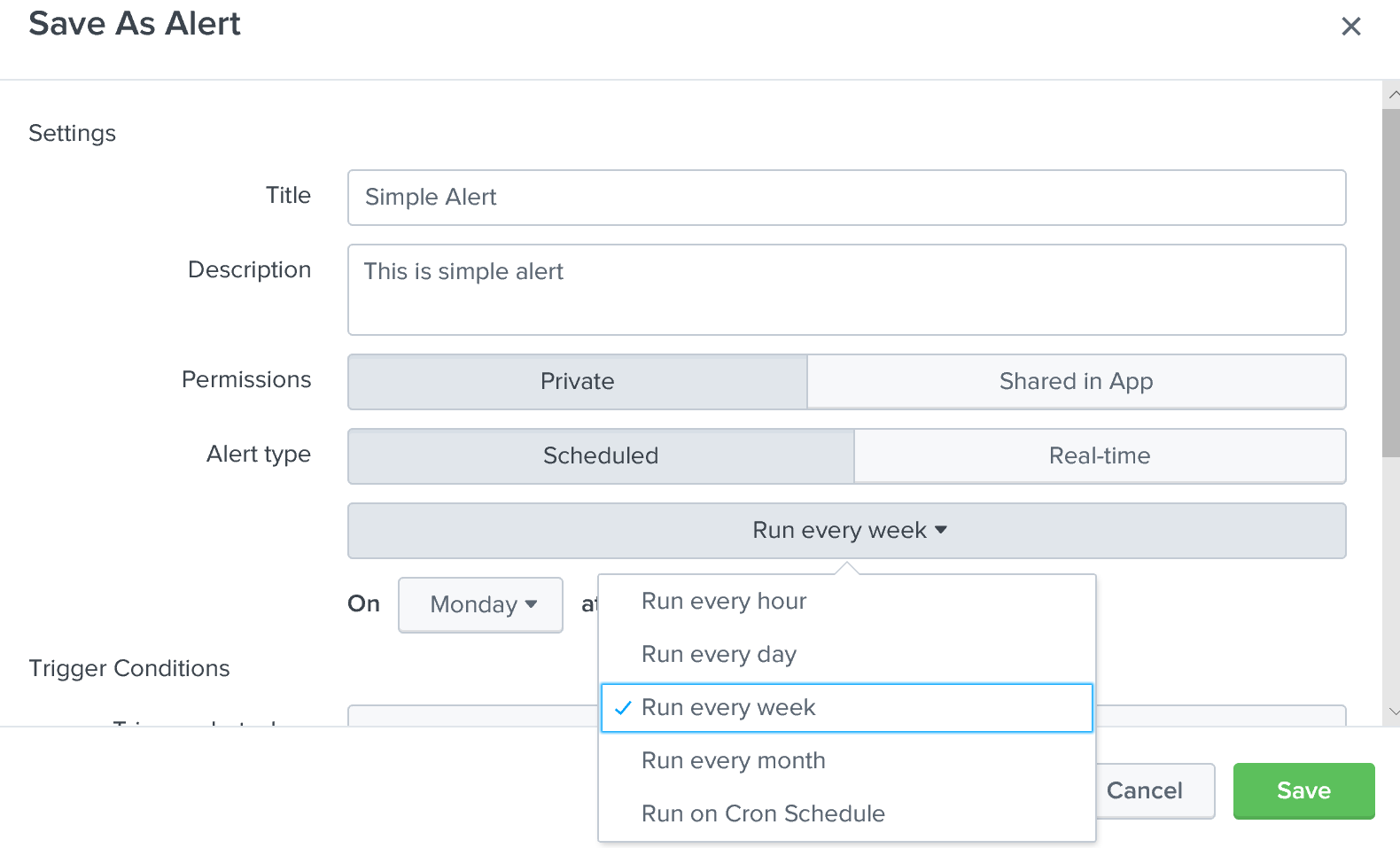
1. Enter Search query in the Search Head and click on Search (green button)



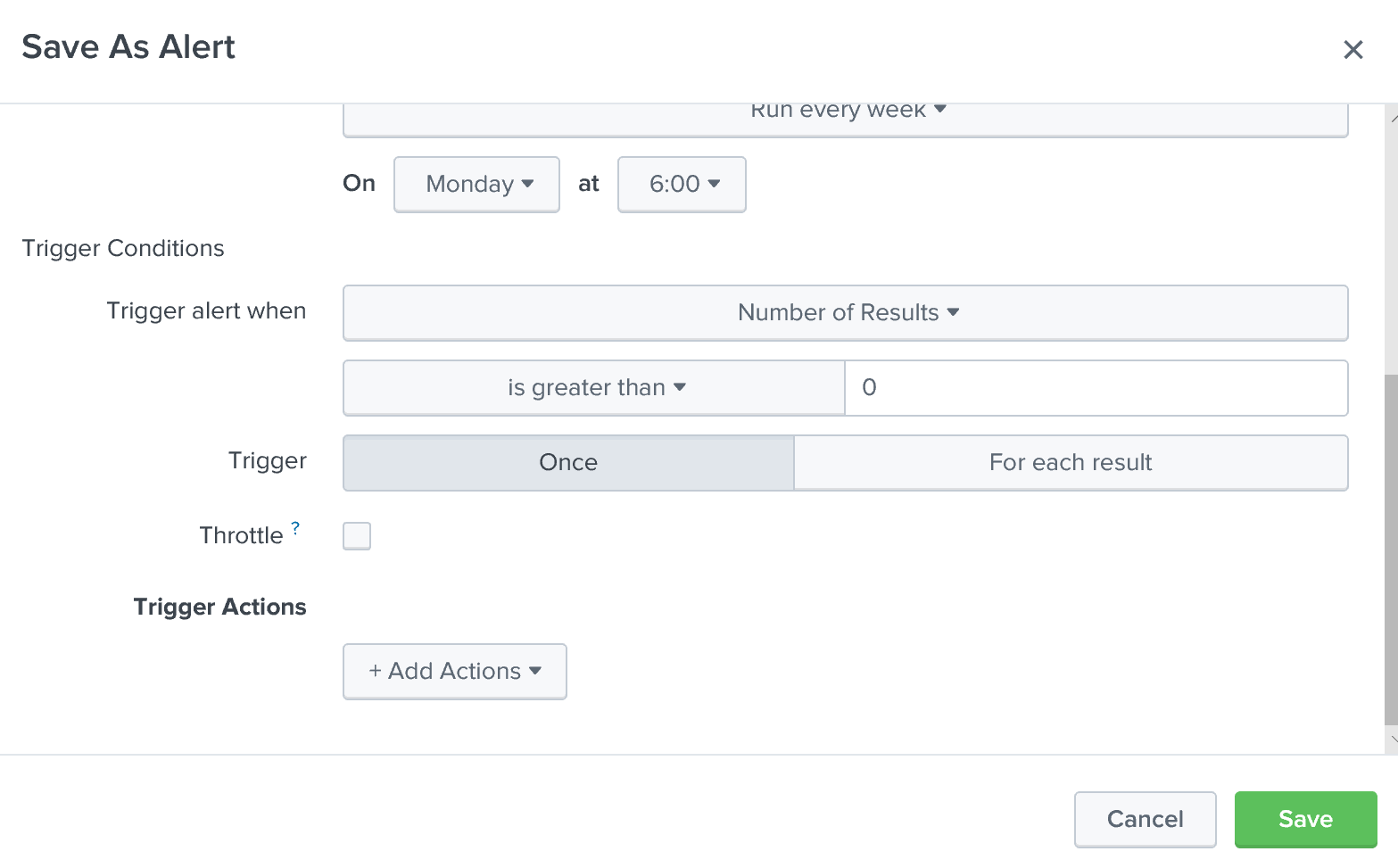
1. We can select time ranges by clicking on drop down at All time. Once search query returns results, Click on Save As -> select Alert.



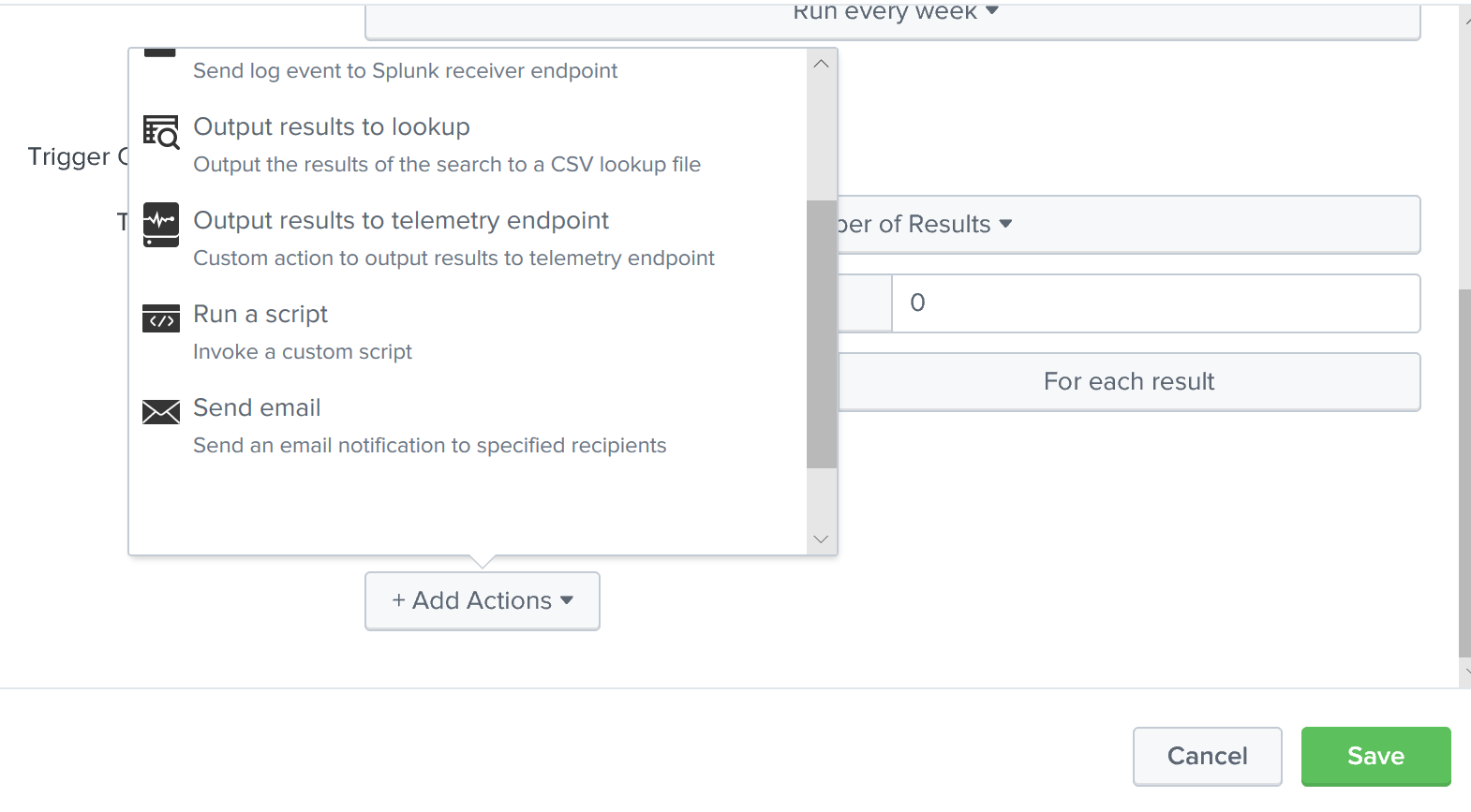
1. Provide proper Title, Description, Permissions and Alert Type. If the alert type is Scheduled then we can select when we want the alert to run. If it is real-time then alert will run always and whenever the search returns result it will be triggered.



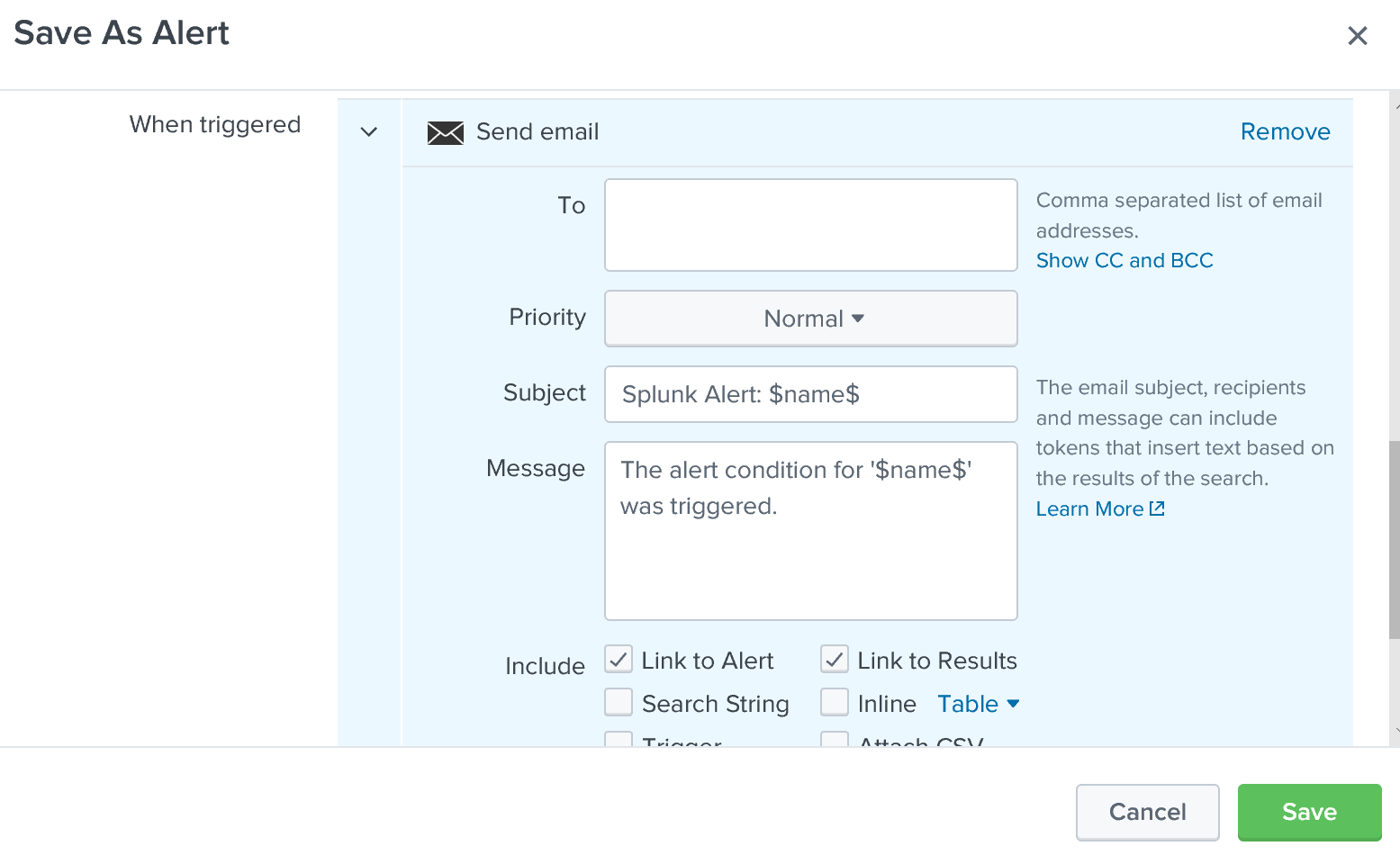
1. We can also select the Trigger conditions on when to trigger the alert.



1. Under Trigger Actions we can select what action we want to take for the Alert. Some of the actions are Run a script OR Send Email OR Output results to lookup.



1. Selecting Send Email will pop up another window where we have to enter receivers email address, Priority, Subject, Message and Search results.

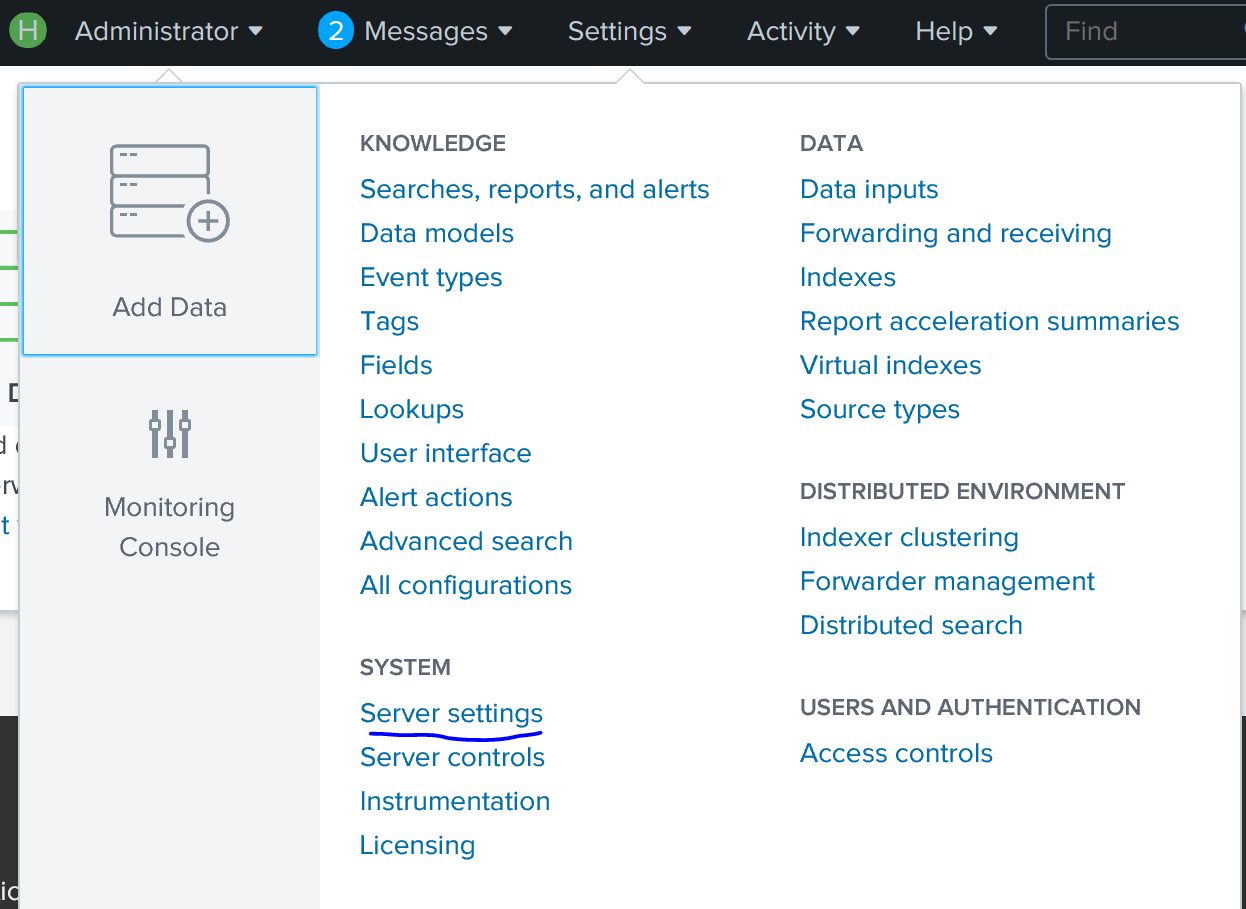


1. Once all the options are selected, we can save the alert.

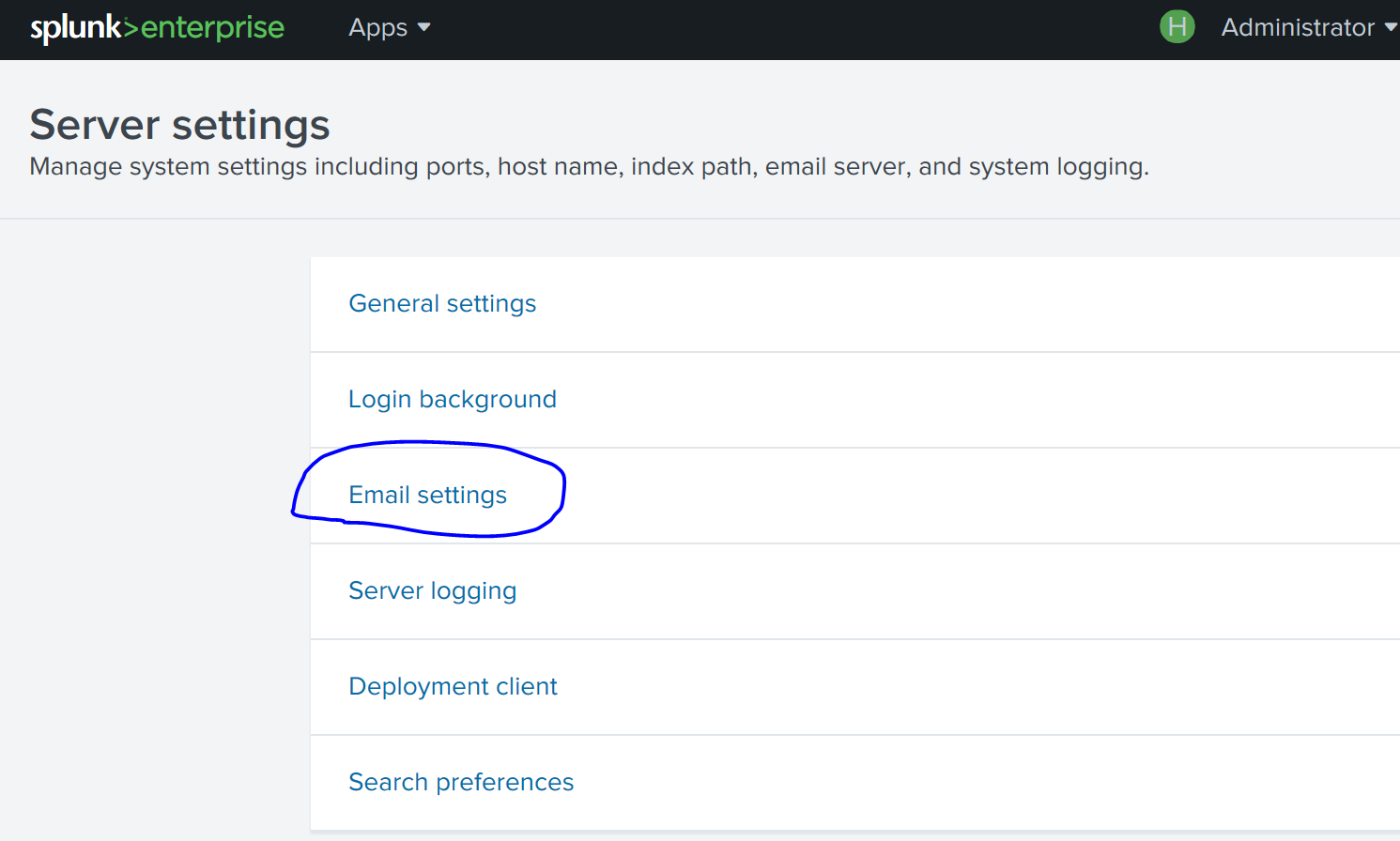
**Mail server settings to create Alerts:**

We have to setup email server settings in Splunk to send emails whenever an alert is triggered. Follow the below steps to configure the mail settings.

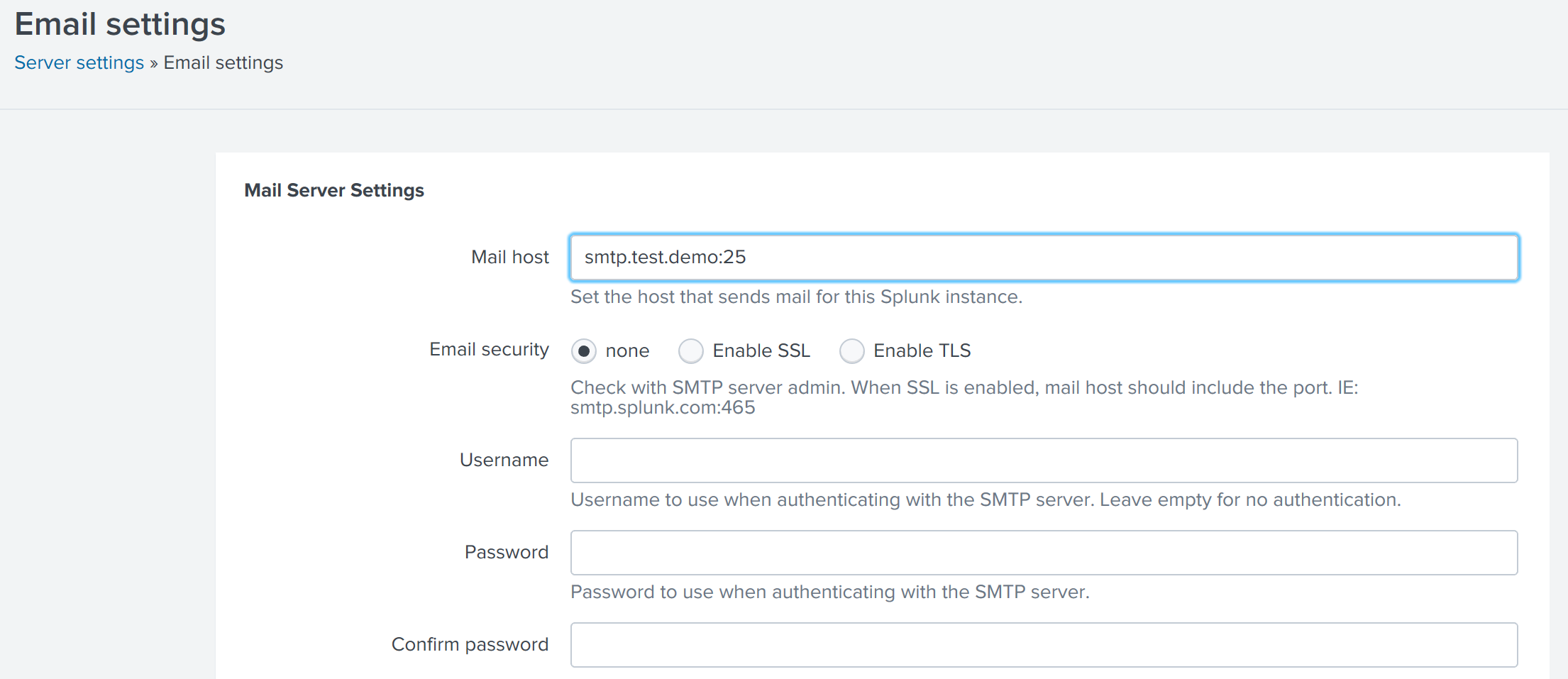
1. Click on Administrator and Settings under SYSTEM section.



1. Click on Email settings to open the configuration window.



1. Under mail server settings, we have to provide the Mail host, Email Security, Username and Password. Email format and PDF report settings are optional.



**Reference:**

http://docs.splunk.com/Documentation/Splunk/7.1.1/SearchReference/WhatsInThisManual