Lab: Splunk and Log Analysis

Using Splunk at **http:// asecuritysite.com:8000** determine the following. You will be allocated a login.

Demo: https://youtu.be/Q7J_Fg_4zVI

Requirement	Answer
What is the start date of the log?	
How many log events are in d:\\buttercup\\mailsv\\secure.log:	
How many log events are in d:\\buttercup\\www1\\access.log:	
How many log events are in d:\\buttercup\\www1\\secure.log:	
How many log events are in d:\\buttercup\\www2\\access.log:	
What is the first username in the security log that gave an incorrect password (Hint: failed password | reverse)	
What is the first IP address in the security log that gave an incorrect password (Hint: failed password | reverse)	
Refer to the Splunk analysis. How many accesses were accessed by a "Chrome" browser and a "GET" method request (Hint - "chrome" AND method=GET)	
How many accesses were accessed by a "Chrome" browser or a "GET" method request (Hint - "chrome" OR method=GET)	
How many accesses were accessed by a "Chrome" browser and a "POST" method request (Hint - "chrome" AND method=POST)	
How many accesses were access by a "Chrome" browser or a "POST" method request (Hint - "chrome" OR method=POST)	
When was the peak accesses by a "Chrome" browser or a "POST" method request (Hint - "chrome" OR method=POST)	
How many accesses are there from a Safari browser (Hint: "safari"):	
How many accesses are there from a Chrome browser (Hint: "chrome"):	
How many accesses are there from a Mozilla browser (Hint: "mozilla"):	
On what day is there most activity in the secure logs (Hint: sourcetype=secure*):	
For the access.log from www1, which is the most popular HTTP response value (Hint - source="d:\\buttercup\\www1\\access.log"| top limit=5 status)	

For the access.log from www1, which is the second most popular HTTP response value (Hint = source="d:\\buttercup\\www1\\access.log"| top limit=5 status)	
For the access.log from www1, which is the most popular IP address for accesses (Hint - source="d:\\buttercup\\www1\\access.log"| top limit=5 clientip)	
For the access.log from www1, which is the second most popular IP address for accesses (Hint - source="d:\\buttercup\\www1\\access.log"| top limit=5 clientip)	
For the access.log from www1, which is the most popular action (Hint - source="d:\\buttercup\\www1\\access.log"| top limit=5 action):	
Refer to the Splunk analysis. For the access.log from www1, which is the second most popular action (source="d:\\buttercup\\www1\\access.log"| top limit=5 action)	
For the access.log from www1, estimate the number of iPad accesses (Hint - source="d:\\buttercup\\www1\\access.log" ipad)	
For the access.log from www1, what is the top refer domain (Hint - source=" d:\\buttercup\\www1\\access.log " | top limit=20 referer)	
Which is the first time for a refer from google.com (Hint - source="d:\\buttercup\\www1\\access.log" referer="http://www.google.com" | reverse)	
Which is the IP address of the client which is first referred from google.com (source="d:\\buttercup\\www1\\access.log" referer="http://www.google.com" | reverse)	
Are there any successful accesses to signals.zip (Hint - signals.zip status=200)	
Refer to the Splunk analysis for secure*.log. How many failed password attempts were there from 194.8.74.23 (Hint - sourcetype=secure* 194.8.74.23 failed)	
Refer to the Splunk analysis for secure*.log. What day of the week had the most failed password attempts from 194.8.74.23 (Hint - sourcetype=secure* 194.8.74.23 failed)	
Refer to the Splunk analysis for access*.log. What day had the most successful purchases (Hint - action=purchase status=200)	
Refer to the Splunk analysis for access*.log. What day had the fewest purchases (Hint - action=purchase status=200)	
Refer to the Splunk analysis for access*.log. What day had the most purchases which were not successfully processed (Hint - action=purchase status!=200)	
Refer to the Splunk analysis for access*.log. How many STRATEGY games have been successfully purchased (Hint - categoryId=STRATEGY action=purchase status=200)	
Refer to the Splunk analysis for access*.log. Which file access always produces a 404 return message	
Refer to the Splunk analysis for access*.log. Which file access always produces a 404 return message: anna_nicole.html, productscreen.html, numa.html, cart.do or oldlink	
Refer to the Splunk analysis for access*.log. How many ARCADE games have been successfully purchased (Hint - categoryId=ARCADE action=purchase status=200)	

Refer to the Splunk analysis for access*.log. How many TEE games have been successfully purchased? (Hint - categoryId=TEE action=purchase status=200)	
Refer to the Splunk analysis for access*.log. How many SIMULATION games have been successfully purchased? (Hint - categoryId=TEE action=purchase status=200)	
Refer to the Splunk analysis for access*.log. How many SHOOTER games have been successfully purchased? (Hint - categoryId=SHOOTER action=purchase status=200)	
Refer to the Splunk analysis for secure*.log. What day of the week had the least failed password attempts from 194.8.74.23? (Hint - failed password 194.8.74.23)	
Refer to the Splunk analysis for access*.log. For an HTTP GET request, which is the most popular return code (Hint - sourcetype="access*" method="GET"| top limit=20 status)	
Refer to the Splunk analysis for access*.log. For an HTTP GET request, which is the 2nd most popular return code (Hint - sourcetype="access*" method="GET"| top limit=20 status)	

B Regular Expression Searches

We can use regular expressions to find information. For example, to find the number of accesses from an IP address which starts with “182.”, we can use:

```
get | regex _raw="182\.\d{1,3}\.\d{1,3}\.\d{1,3}"
```

Determine the number of accesses for GET from any address which begins with 182:

The security team search for an address that is ending with .22, and do a search with:

```
get | regex _raw="\d{1,3}.\d{1,3}.\d{1,3}.22"
```

But it picks up logs which do not include addresses with .22 at the end. What is the problem with the request, and how would you modify the request:

You are told that there’s accesses to a file which ends in “a.html”. Using a regular expression, such as:

```
get | regex _raw="[a]+\.html"
```

Outline three HTML files which end with the characters ‘a’, or an ‘e’, and have ‘.html’ as an extension:

A simple domain name check is:

```
get | regex _raw="[a-zA-Z\.]+\.(com|net|uk)"
```

If we now try:

```
get | regex _raw="[a-zA-Z0-9\-\.]+\.(com|org|net|mil|edu|COM|ORG|NET|MIL|EDU|UK)"
```

we will return events with domain names:

Outline which ones have been added:

We can search for email addresses with:

```
get | regex _raw="(?<email>[\w\d\.\-]+\@[\w\d\.]+)"
```

Which email addresses are present:

We can search for times using regular expressions, such as:

```
get | regex _raw="[0-9]{2}\:22\:[0-9]{2}"
```

How many GET requests where there at 22 minutes past the hour:

How many GET requests were made at 14 seconds past the minute:

C Investigation

The incident response team wants a report of unusual HTTP requests. Produce a report of the HTTP response codes, and what they identify. Which ones could be malicious?

1. The security team would like a report on the most popular user names that have failed on the security logs. What would you report?
2. The company are worried about the sales of some of their games. Which game category has the least amount of sales, and which is the best seller?
3. The Web design department have been told that there are missing files on the Web site. Investigate the files/pages that are missing on the site (Hint: 404 codes). . What would you report?
4. You have been asked to investigate accesses to the file named passwords.pdf in the Buttercup Games Splunk trace. Investigate any accesses related to it, and outline any possible significant evidence of malicious activity related to these accesses. . What would you report?

F Test

Now perform the following test:

http://asecuritysite.com/tests/tests?sortby=siem