

Defensive Security Project

by: Peerapat Phatpanichot (Win)

Table of Contents

This document contains the following resources:

01

**Monitoring
Environment**

02

Attack Analysis


03

**Project Summary
& Future
Mitigations**

Monitoring Environment

Scenario

In this project, we developed and implemented a defensive monitoring solution using Splunk to protect Virtual Space Industries (VSI) from potential cyberattacks by their competitor, JobeCorp, focusing on analyzing and alerting for suspicious activities on Windows and Apache servers.

The background of the slide is a dark red color with a complex geometric pattern. This pattern is composed of numerous overlapping triangles and squares, creating a mosaic-like effect. The colors are various shades of dark red and maroon, giving it a textured, crystalline appearance.

["Splunk Machine Learning Tool Kit " App]

Title: Splunk Machine Learning Toolkit (MLTK)

Content:

- **Description:** The Splunk Machine Learning Toolkit provides powerful machine learning capabilities within Splunk. It includes SPL commands, custom visualizations, and various ML assistants that enable you to apply machine learning to your own data.
- **Features:**
 - Predictive analytics (linear and logistic regression).
 - Anomaly detection (numeric and categorical outliers).
 - Time series forecasting.
 - Clustering and other advanced ML algorithms.
- **Why MLTK?:** Enhances Splunk’s native monitoring capabilities with advanced predictive and anomaly detection features, making it invaluable for security and operational monitoring.

Showcase


Welcome to the Machine Learning Toolkit Showcase. Watch and learn from interactive end-to-end examples using real datasets. Click on an example to pre-populate the Assistant with the sample dataset and its settings. Inspect the Search Processing Language as well as the underlying code of these examples to see how it all works.

View examples by

ML OperationIndustry

Filter Examples


Predict Fields



View examples that predict the value of a numeric or categorical field using the values from other fields in the event.

15 Examples Available


Detect Outliers



View examples that detect numeric and categorical values that differ significantly from values in the rest of the data. Identified outliers are indicative of interesting, unusual, and possibly dangerous events.

14 Examples Available


Forecast Time Series



View examples that predict the next value in a sequence of time series data by using past time series data.

9 Examples Available

Cluster Events



View examples that partition events with multiple fields into groups of events based on the values of those fields.

8 Examples Available

>

Python for Scientific Computing (for Linux 64-bit)

Python for Scientific Computing (for Windows 64-bit)

Install

This add-on contains a Python interpreter bundled with the following scientific and machine learning libraries: numpy, scipy, pandas, scikit-learn, and statsmodels. With this add-on, you can import these powerful libraries in your own custom search commands, custom rest endpoints, modular inputs, and so forth.

This add-on is available for Linux (64-bit), Windows (64-bit) and Mac. Make sure you install the appropriate one for your Splunk deployment.

Less

Category: [Artificial Intelligence](#), [Utilities](#) | Author: [Splunk Inc.](#) | Downloads: 124257 |

Released: 25 days ago | Last Updated: 22 days ago | [View on Splunkbase](#)

>

Python for Scientific Computing (for Windows 64-bit)

Python for Scientific Computing (for Windows 64-bit)

Install

This add-on contains a Python interpreter bundled with the following scientific and machine learning libraries: numpy, scipy, pandas, scikit-learn, and statsmodels. With this add-on, you can import these powerful libraries in your own custom search commands, custom rest endpoints, modular inputs, and so forth.

This add-on is available for Linux (64-bit), Windows (64-bit) and Mac. Make sure you install the appropriate one for your Splunk deployment.

Less

Category: [Artificial Intelligence](#), [Utilities](#) | Author: [Splunk Inc.](#) | Downloads: 32205 |

Released: 25 days ago | Last Updated: 22 days ago | [View on Splunkbase](#)

>

Python for Scientific Computing (for Mac 64-bit)

Python for Scientific Computing (for Mac 64-bit)

Open App

This add-on contains a Python interpreter bundled with the following scientific and machine learning libraries: numpy, scipy, pandas, scikit-learn, and statsmodels. With this add-on, you can import these powerful libraries in your own custom search commands, custom rest endpoints, modular inputs, and so forth.

This add-on is available for Linux (64-bit and 32-bit), Windows (64-bit) and Mac. Make sure you install the appropriate one for your Splunk deployment.

Less

Category: [Artificial Intelligence](#), [Utilities](#) | Author: [Splunk Inc.](#) | Downloads: 13064 |

Released: 25 days ago | Last Updated: 22 days ago | [View on Splunkbase](#)

>

Python for Scientific Computing (for Mac ARM64)

Python for Scientific Computing (for Mac ARM64)

Install

This add-on contains a Python interpreter bundled with the following scientific and machine learning libraries: numpy, scipy, pandas, scikit-learn, and statsmodels. With this add-on, you can import these powerful libraries in your own custom search commands, custom rest endpoints, modular inputs, and so forth.

This add-on is available for Linux (64-bit), Windows (64-bit) and Mac. Make sure you install the appropriate one for your Splunk deployment.

Less

Category: [Artificial Intelligence](#), [Utilities](#) | Author: [Splunk Inc.](#) | Downloads: 589 |

Released: 25 days ago | Last Updated: 22 days ago | [View on Splunkbase](#)

6

Splunk Machine Learning Toolkit (MLTK)

Use Case: Forecasting web traffic using Apache logs to predict potential future spikes or drops in request volume, helping to anticipate and mitigate potential server overloads or downtime.

- **Scenario Overview:**
 - **Problem:** Sudden traffic changes can lead to performance degradation or outages.
 - **Solution:** Using the Smart Forecasting Assistant, VSI can forecast traffic patterns and prepare for expected high-traffic periods.
 - **Benefit:** Proactively scaling resources or adjusting security measures based on forecasted data, enhancing reliability and performance.

Smart Forecasting: Apache_logsDraft

Cancel

Next >

Forecast **request_log** for the next **24 hour(s)**, based on **3 days** of data, with a confidence interval of **95%** and **10 hour(s)** of data held back.

Define

Learn

Review

Operationalize

Define Data Source

View history

Q Search

Datasets

Metrics

source="apache_logs.txt" | timechart span=1h count as request_log

All time

Q

✓ 10,000 events (3/17/20 9:05:00.000 PM to 9/4/24 1:38:10.000 AM)

Job

||

Smart Mode

Data Preview

Visualization

20 Per Page

« Prev

1

2

3

4

5

Next »

_time	request_log
2020-03-17 21:00	74
2020-03-17 22:00	111
2020-03-17 23:00	115
2020-03-18 00:00	118
2020-03-18 01:00	120
2020-03-18 02:00	125
2020-03-18 03:00	126
2020-03-18 04:00	123
2020-03-18 05:00	118
2020-03-18 06:00	121
2020-03-18 07:00	129

7

Splunk Machine Learning Toolkit (MLTK)

Smart Forecasting: Apache_logs Draft

Cancel

Save

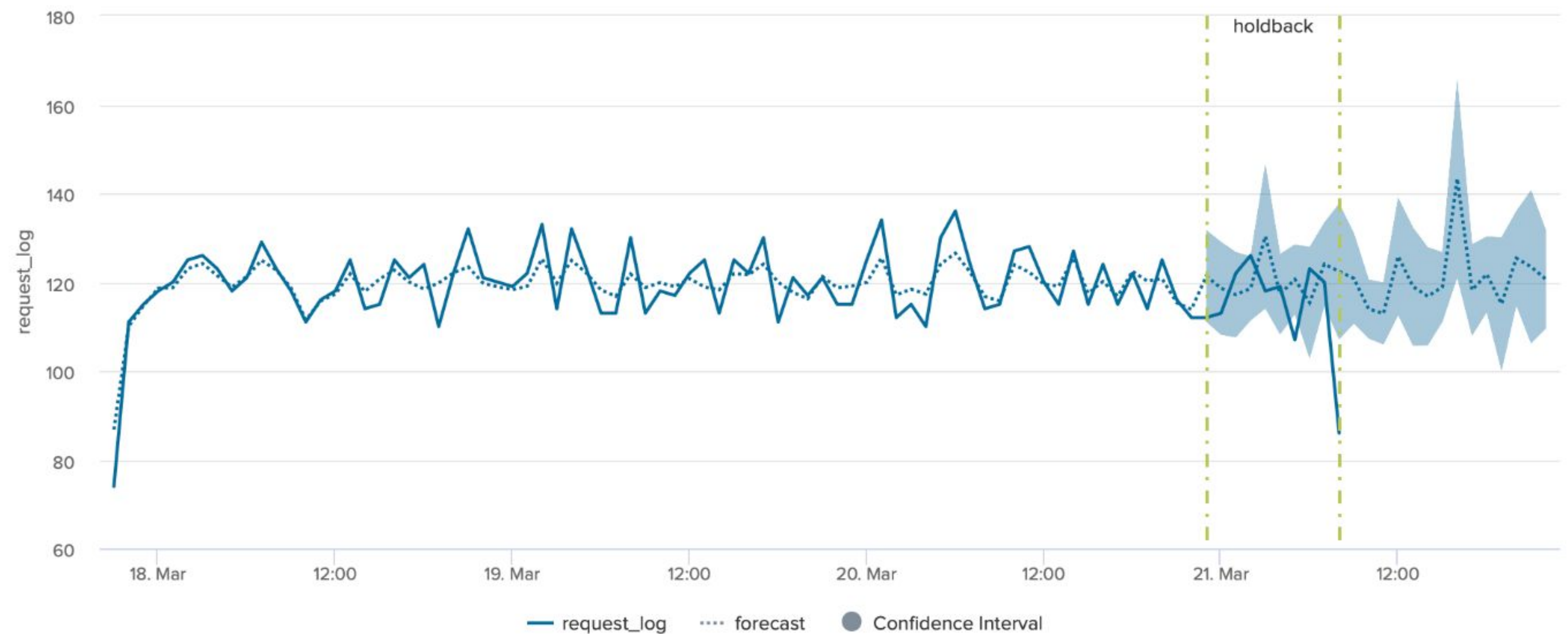
[< Back](#)

Next >

Forecast **request_log** for the next **24 hour(s)**, based on **3 days** of data, with a confidence interval of **95%** and **10 hour(s)** of data held back.

[View history](#)

SPL

☒ Confidence Interval

Learn Data

Fields to forecast

request_log (1) ▼

Holdback period

10 Hour(s)

Future timespan 

24 Hour(s)

Confidence interval

1  99 95  

Special days field

(Optional)

Period

(Optional)

Notes

(optional)



Define

Learn

 Review



eration

Operationalize

Splunk Machine Learning Toolkit (MLTK)

Detect Numeric Outliers

Find values that differ significantly from previous values.

Assistant Settings

Enter a search

source="apache_logs.txt" | timechart span=1h count as request_count

All time

Q

✓ 10,000 events (3/17/20 9:05:00.000 PM to 9/4/24 12:31:14.000 AM)

Job ▾ || ■ Smart Mode ▾

Field to analyze

request_count ▾

Threshold method

Standard Deviation ▾

Threshold multiplier

2

Sliding window (# of values)

(optional)

☒ Include current point

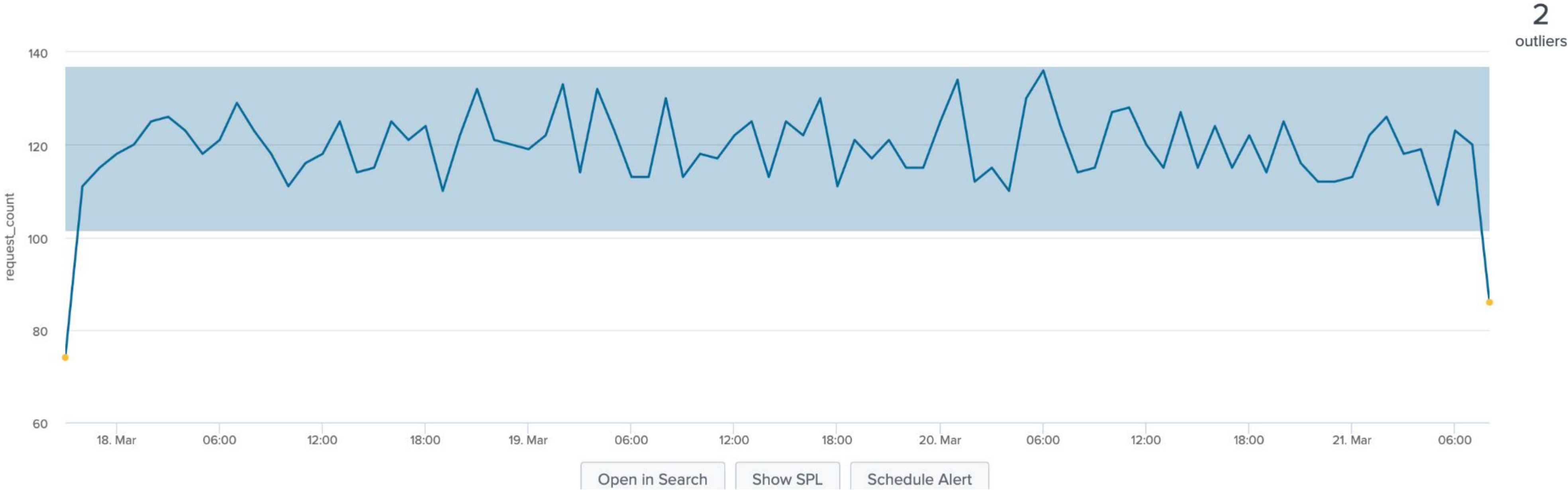
Fields to split by

Detect Outliers

Open in Search

Show SPL

Data and Outliers



Logs Analyzed

1

Windows Logs

- Contains security events and logs from VSI's Windows server.
- Logs include fields like `signature_id`, `user`, `status`, and `severity`.
- Data captures login attempts, system changes, and security breaches.
- Used to establish baselines and monitor for abnormal activities.
- Helps identify failed logins, account deletions, and unauthorized access.

2

Apache Logs

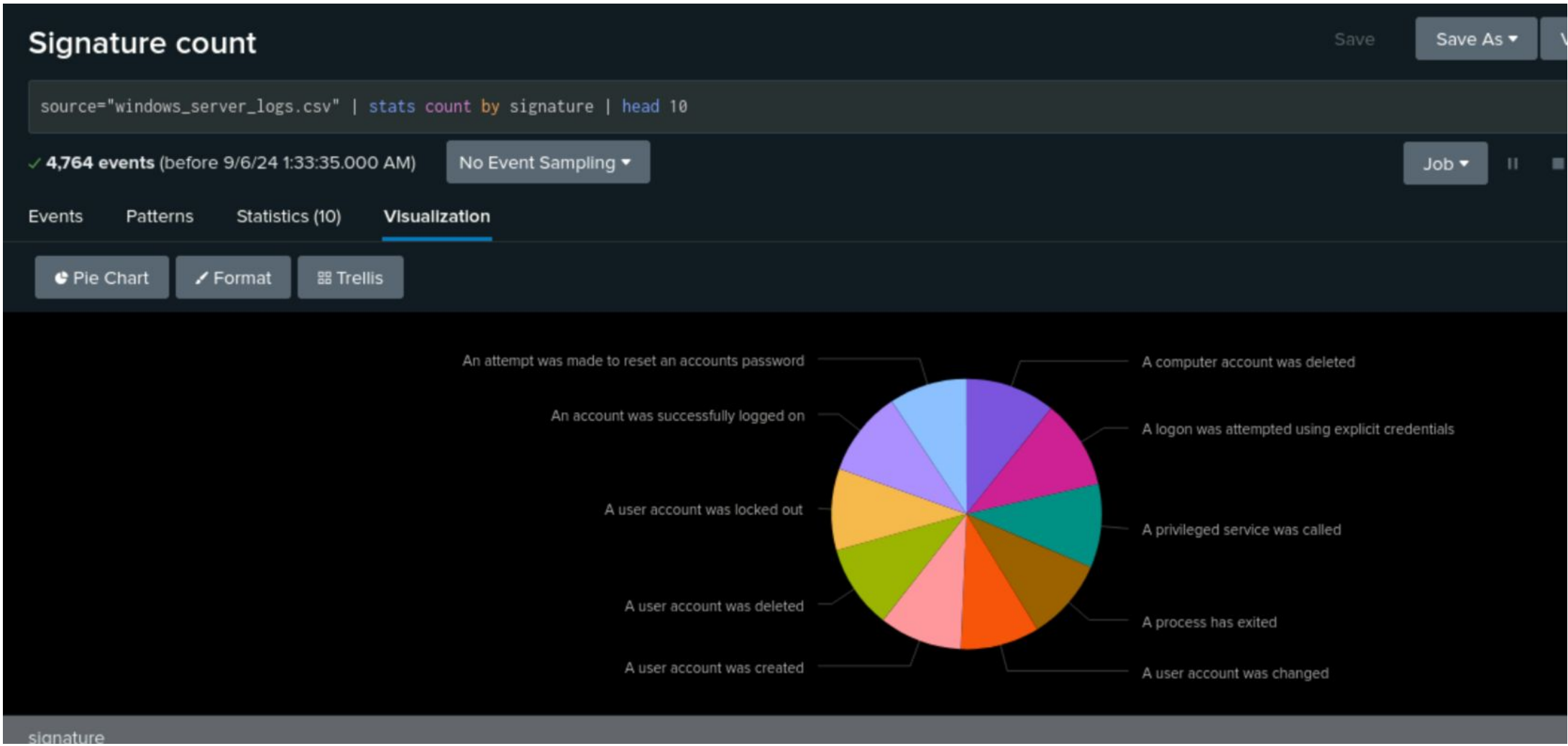
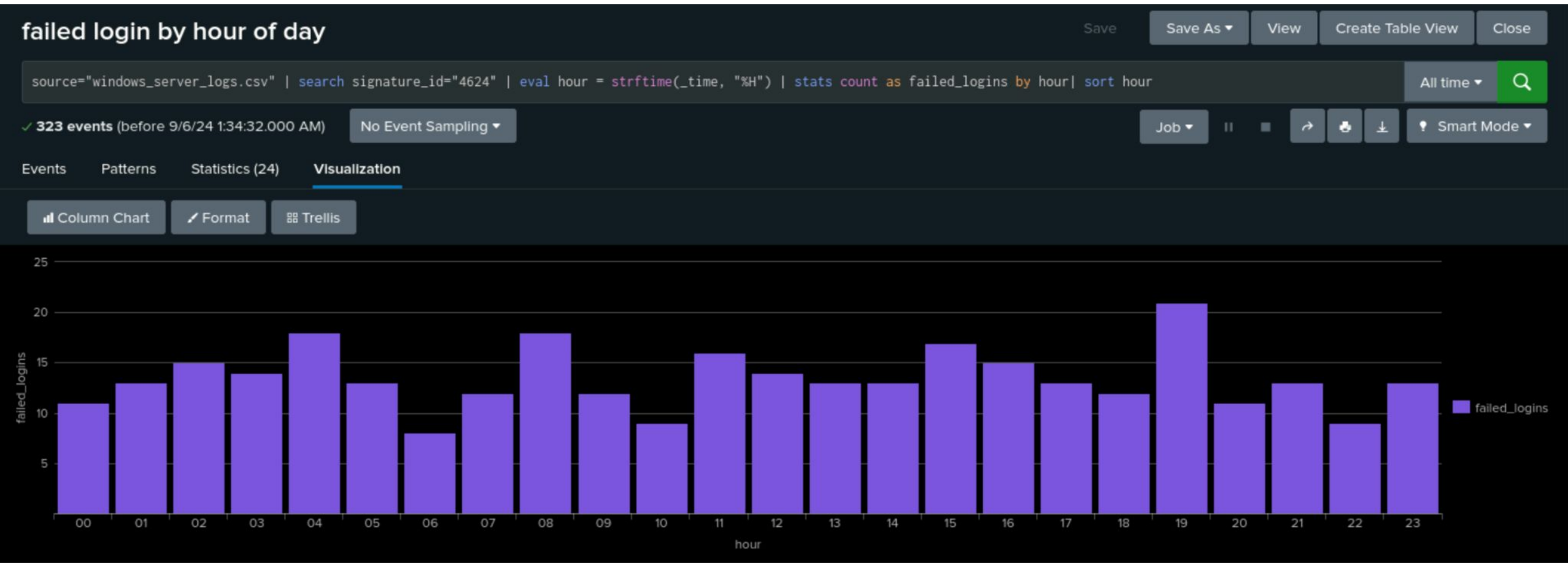
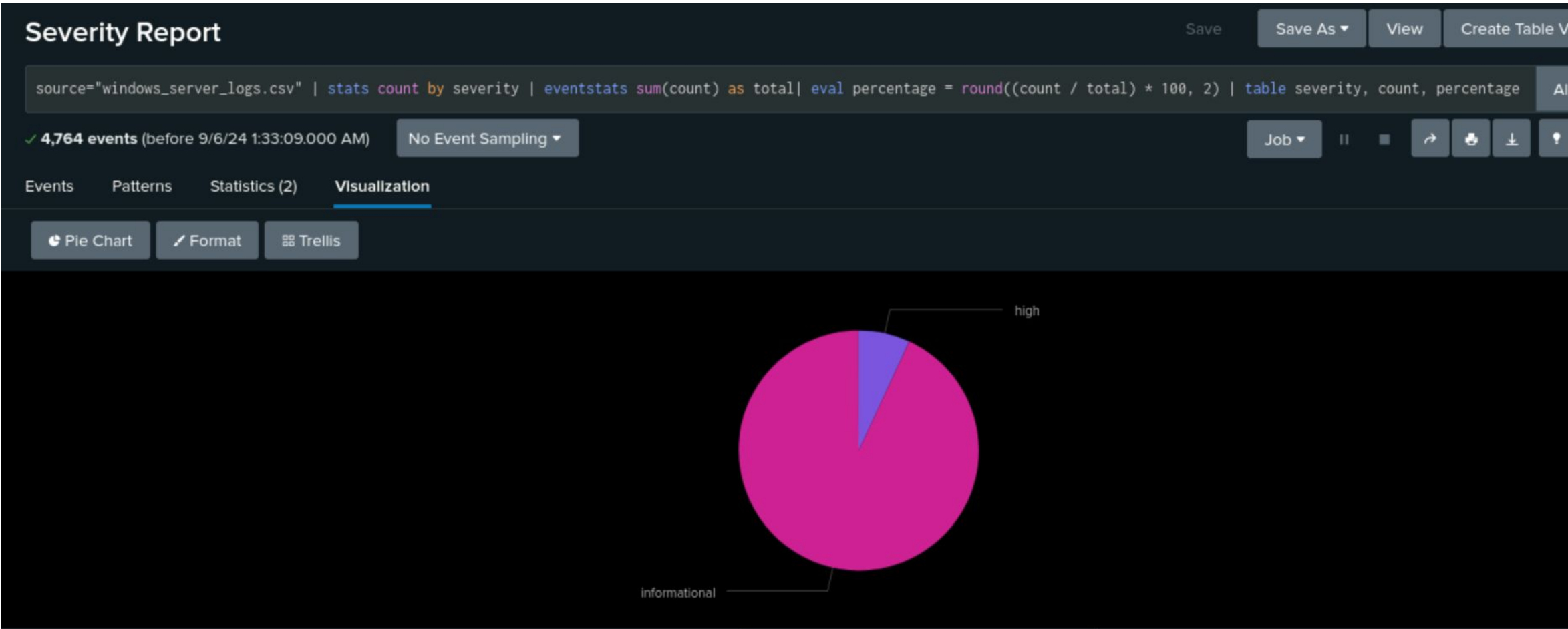
- Captures HTTP requests to VSI's public-facing website.
- Logs include fields like `method`, `status`, `clientip`, and `useragent`.
- Tracks GET and POST requests, indicating user interactions and access points.
- Monitors referrer domains to detect potential malicious traffic sources.
- Provides insights into response codes, identifying server errors and attack attempts(`apache_logs`).

Windows Logs

Reports—Windows

Report Name	Report Description
Severity Report	Displays the distribution of events based on severity, highlighting high and informational severity events.
Failed Login by Hour of Day	Shows the count of failed login attempts for each hour of the day, identifying peak times of unsuccessful access attempts.
Top 10 Most Lockout Accounts	Identifies the top 10 user accounts with the highest number of lockouts, indicating potential security concerns or brute force attempts.
Signature Count	Provides a count of various event signatures, including password reset attempts, successful logins, account deletions, and more, to identify frequent activities.

Images of Reports—Windows



Alerts—Windows

Alert Name	Alert Description	Alert Baseline	Alert Threshold
Alert for Hourly Failed Window Activity	Monitors failed login attempts on Windows systems to detect abnormal or potentially malicious activity.	5.92 failed logins per hour (calculated average of failed logins).	> 11 failed logins per hour.

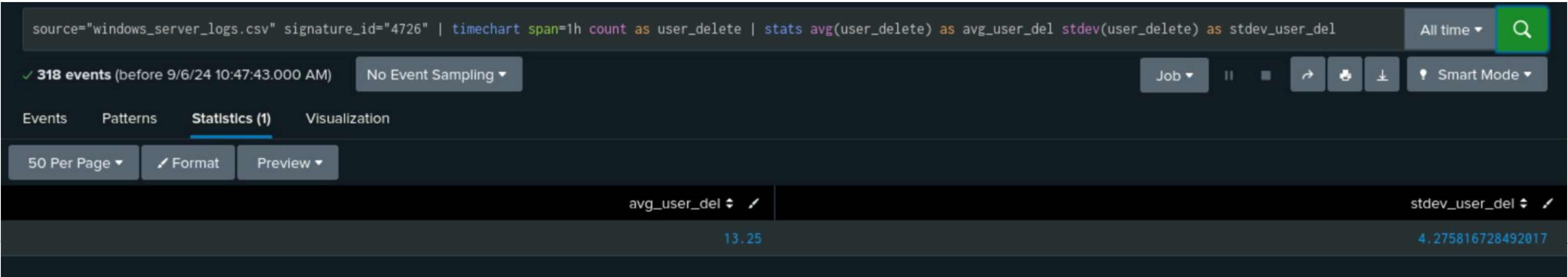
JUSTIFICATION: The baseline was determined based on the calculated average of failed logins per hour, which is 5.92. The threshold is set at 11 using the formula $\text{avg} + (2 \times \text{stdev})$.



Alerts—Windows

Alert Name	Alert Description	Alert Baseline	Alert Threshold
ALERT for hourly count of a user account delete	This alert monitors the hourly count of user account deletions to detect suspicious activities that exceed normal thresholds.	13.25 (average count of user deletions per hour)	>22 deletions per hour

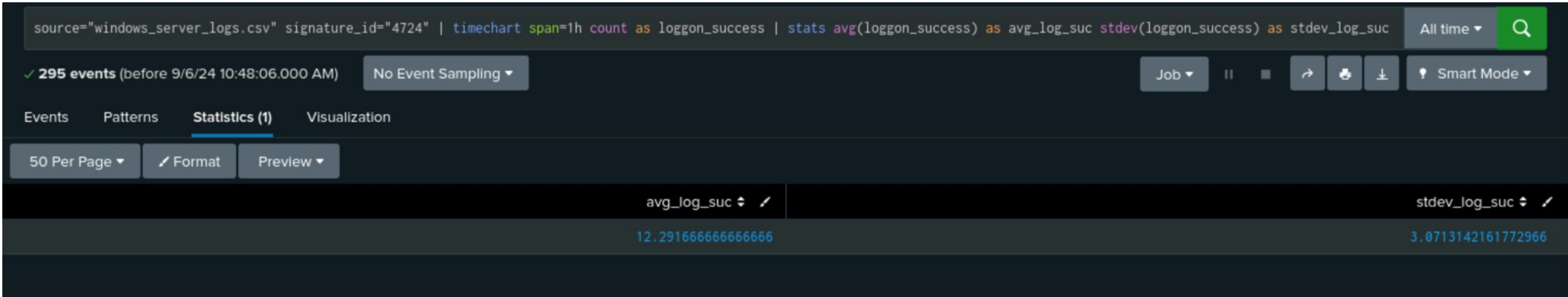
JUSTIFICATION: The baseline of 13.25 was calculated based on the average hourly deletions, providing a typical activity measure. The threshold is set at a higher value (22) to identify unusual spikes using the formula $\text{avg} + (2 \times \text{stdev})$.



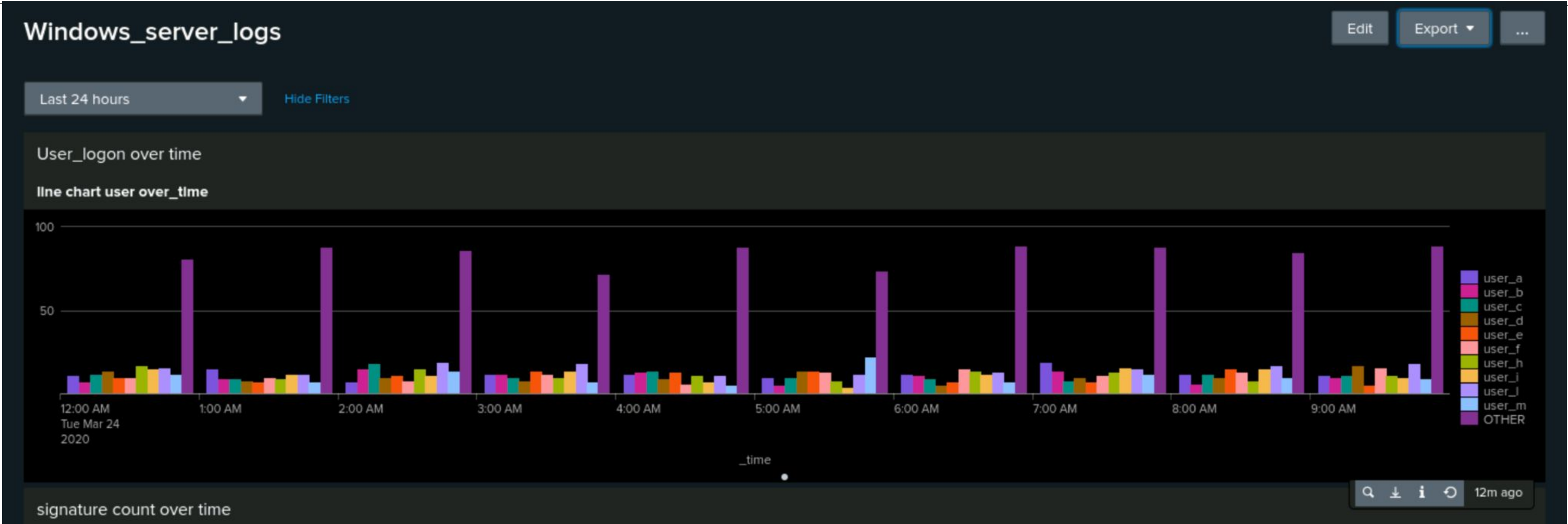
Alerts—Windows

Alert Name	Alert Description	Alert Baseline	Alert Threshold
An account successfully logged on ALERT WINDOWS	This alert monitors the number of successful login events on Windows servers.	13	19

JUSTIFICATION: The baseline value of 12.29 was calculated as the average number of successful login events per hour, and the threshold of 19 is set to alert when the count significantly exceeds the baseline using the formula $\text{avg} + (2 \times \text{stdev})$.



Dashboards—Windows



Dashboards—Windows

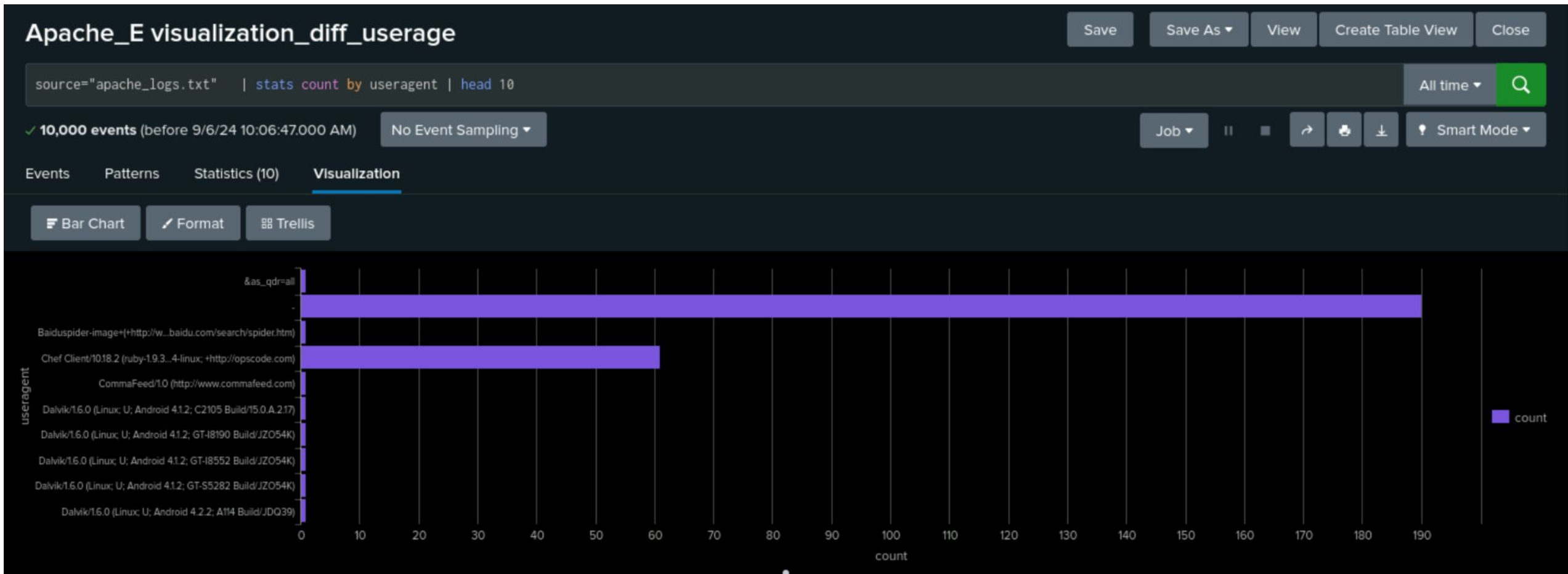
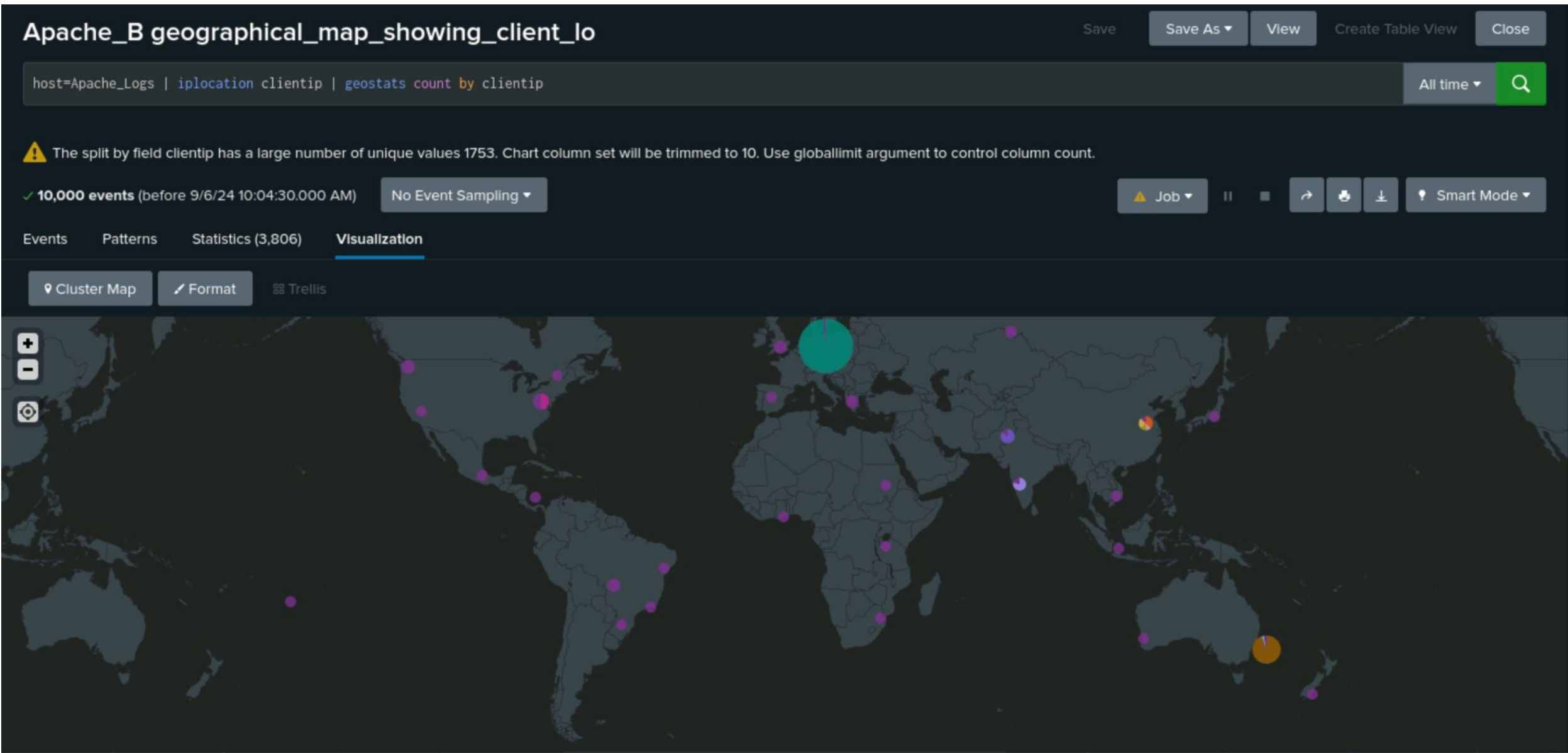
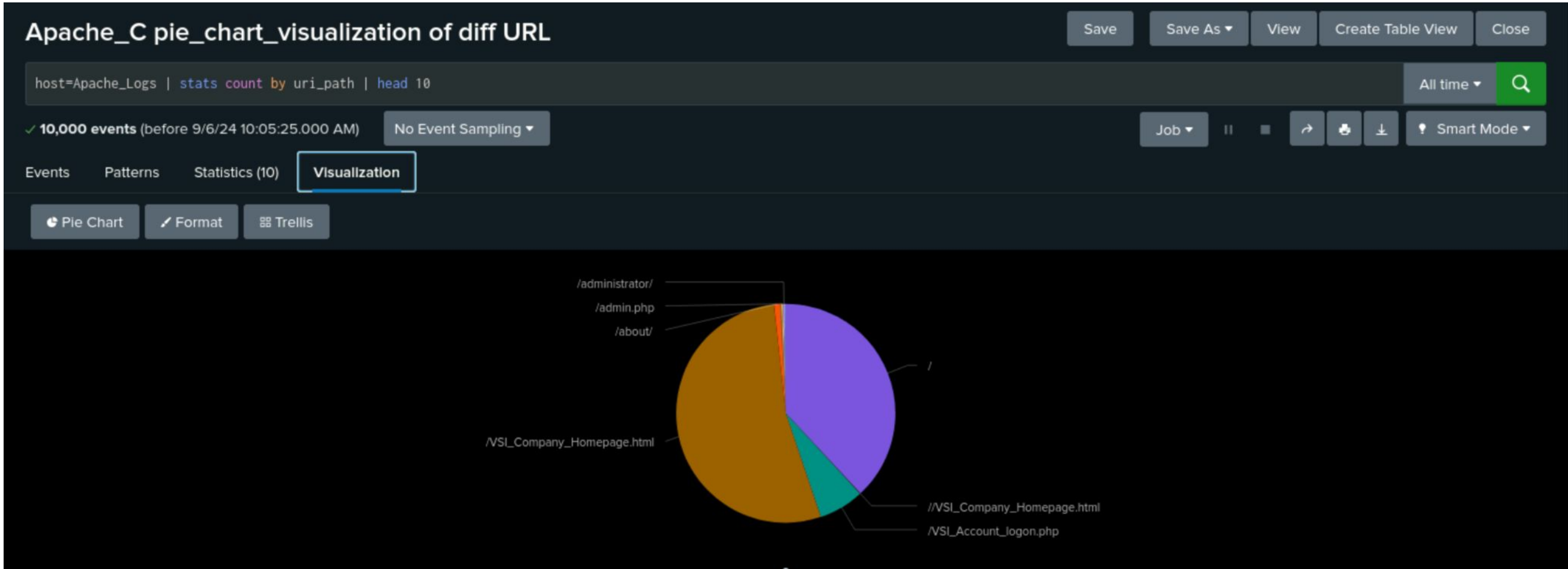


Apache Logs

Reports—Apache

Report Name	Report Description
Apache_A Line Chart HTTP Method Overtime	Displays the count of different HTTP methods (GET, POST, HEAD, OPTIONS) over time, highlighting periods of unusual activity indicative of a potential attack.
Apache_C Pie Chart Visualization of Diff URL	Shows the distribution of requests to various URIs, identifying which pages were targeted the most, indicating possible attack vectors on sensitive endpoints.
Apache_B Geographical Map Showing Client Locations	Maps the geographical distribution of client IP addresses, highlighting suspicious international activity and identifying regions with high request volumes.
Apache_E Visualization Diff Userage	Bar chart showing user access counts, helping identify abnormal user activity patterns that may indicate unauthorized access attempts or brute force attacks.

Images of Reports—Apache



Alerts—Apache

Alert Name	Alert Description	Alert Baseline	Alert Threshold
APACHElog_Hourly_Activity_from_anyCOUNTRY_NOT_US	This alert monitors hourly activity from countries outside the United States, triggering when abnormal levels of traffic are detected.	100	> 120.

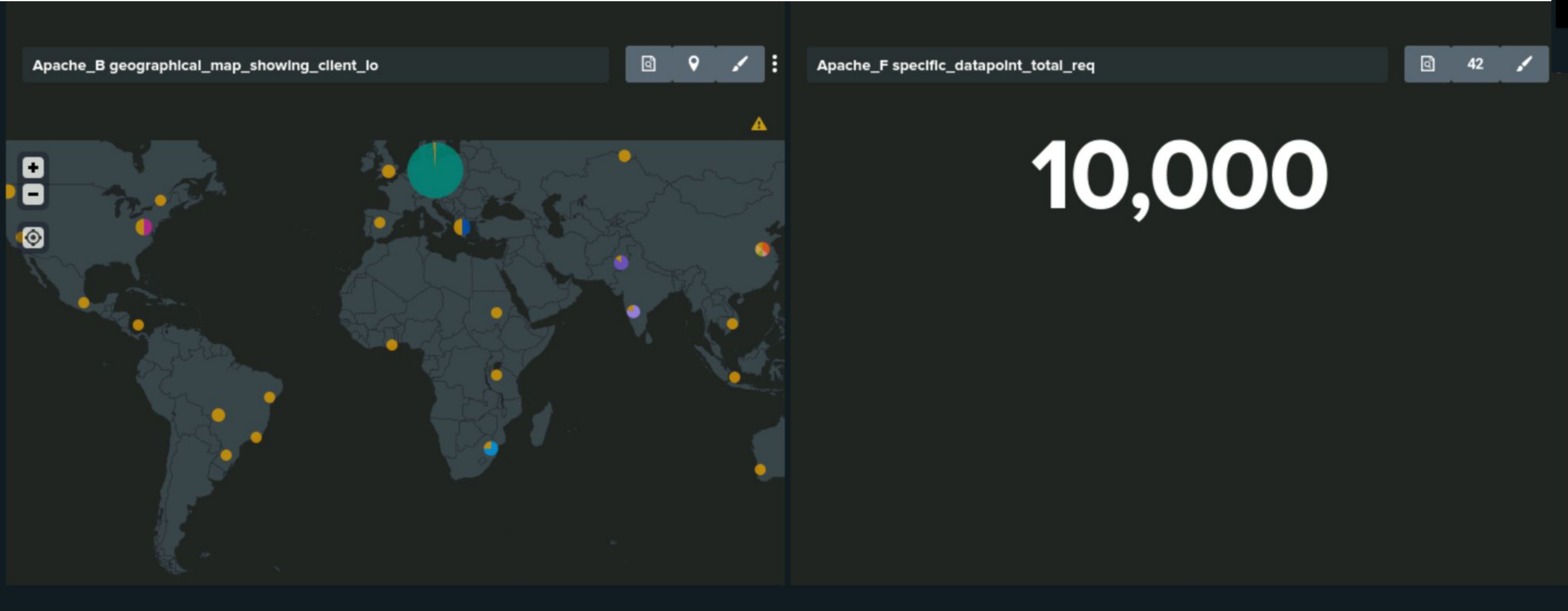
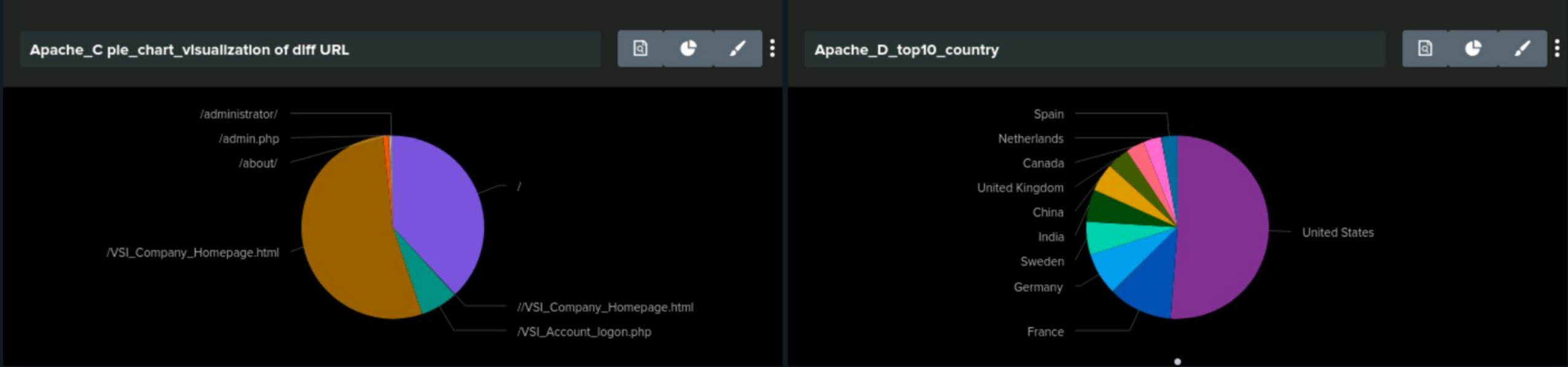
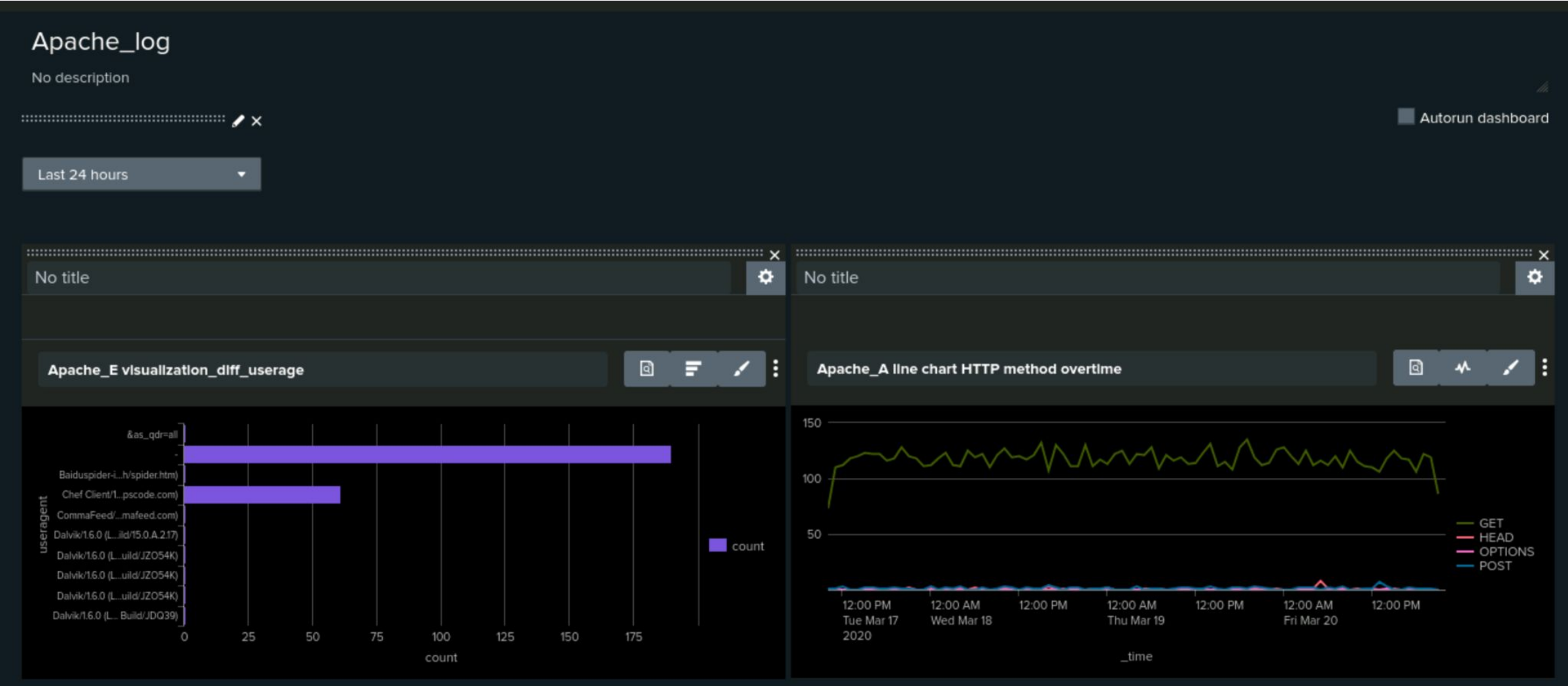
JUSTIFICATION: The baseline is set based on typical hourly traffic from non-US countries. The threshold of 120 after skimming through the calculations activity counts

Alerts—Apache

Alert Name	Alert Description	Alert Baseline	Alert Threshold
Apache hourly count of HTTP POST method	This alert monitors the hourly count of HTTP POST requests to detect unusual activity that may indicate potential attacks or data exfiltration attempts.	2	> 4

JUSTIFICATION: The baseline is set according to typical HTTP POST traffic, and the threshold of 4 was chosen to capture anomalies as the average was 1.2 and stdev was also around 1.2 so after using the formula $\text{avg} + (2 * \text{stdev})$, I came up with this conclusion using avg as baseline and threshold as the result of the formula

Dashboards—Apache



Attack Analysis

Attack Summary—Windows

The reports highlighted a significant volume of failed login attempts, account deletions, and suspicious user activities, indicating potential security breaches and unauthorized access attempts mainly from user_a and user_k.

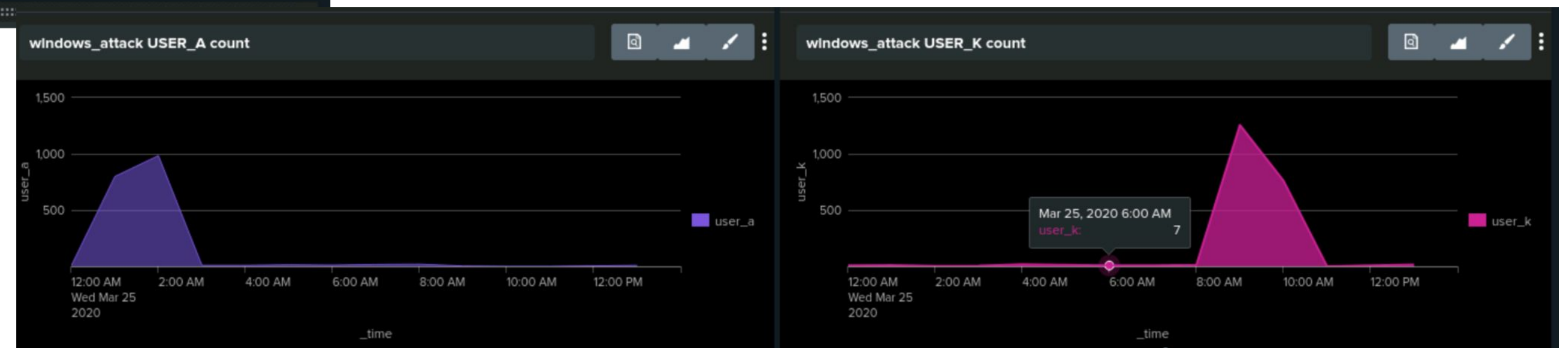
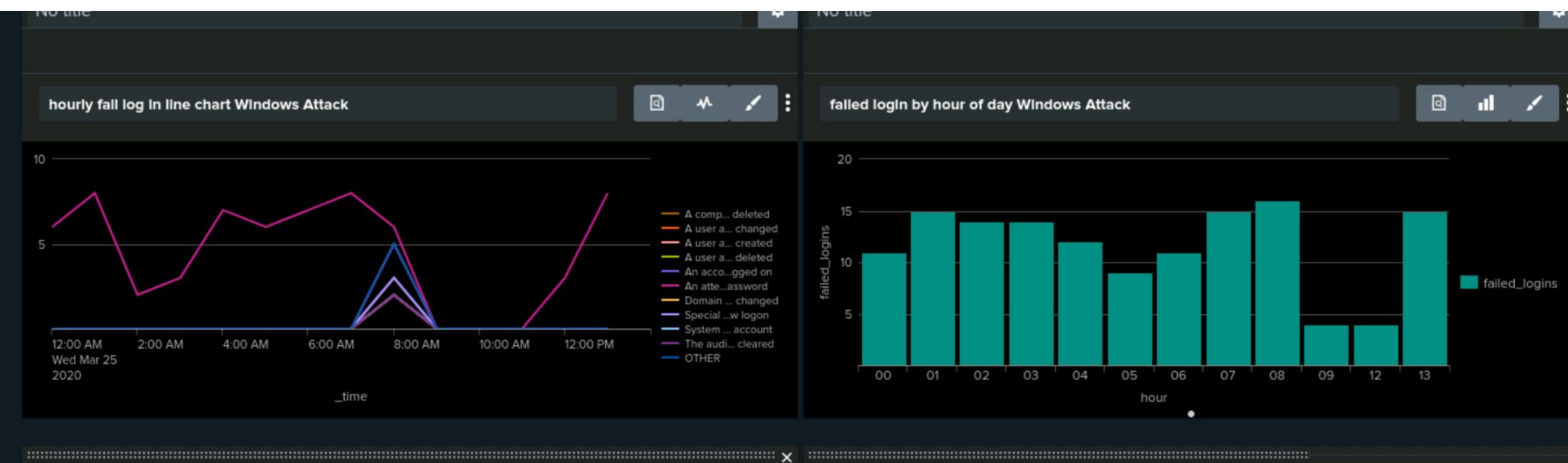
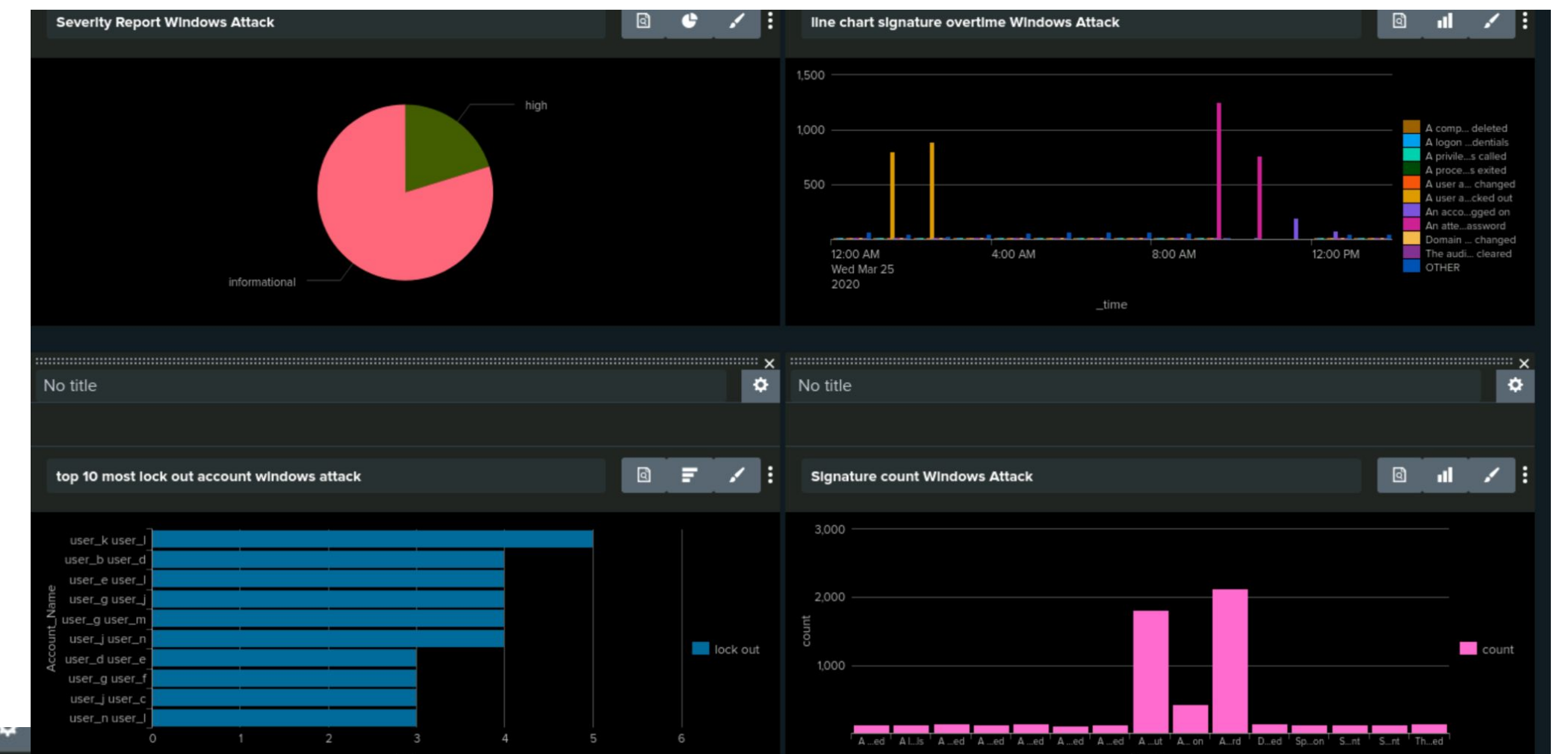
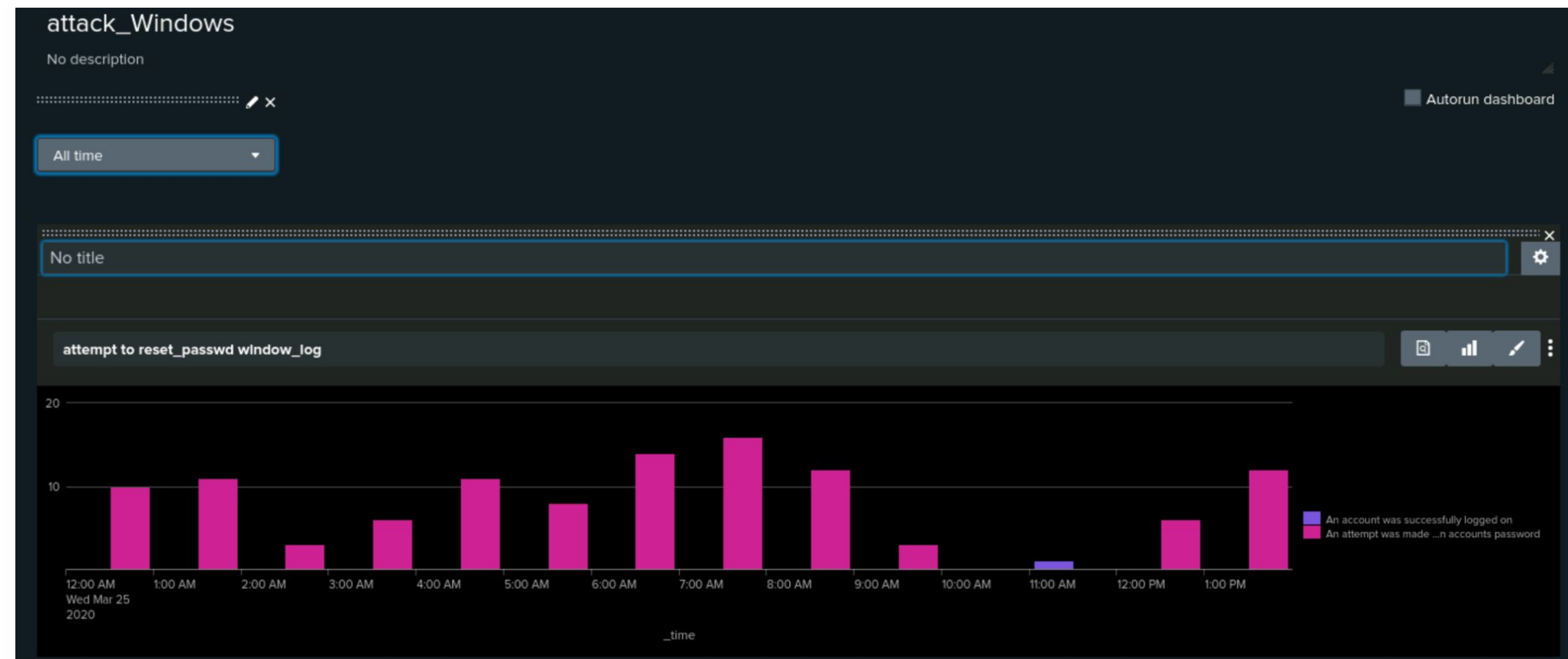


Attack Summary—Windows

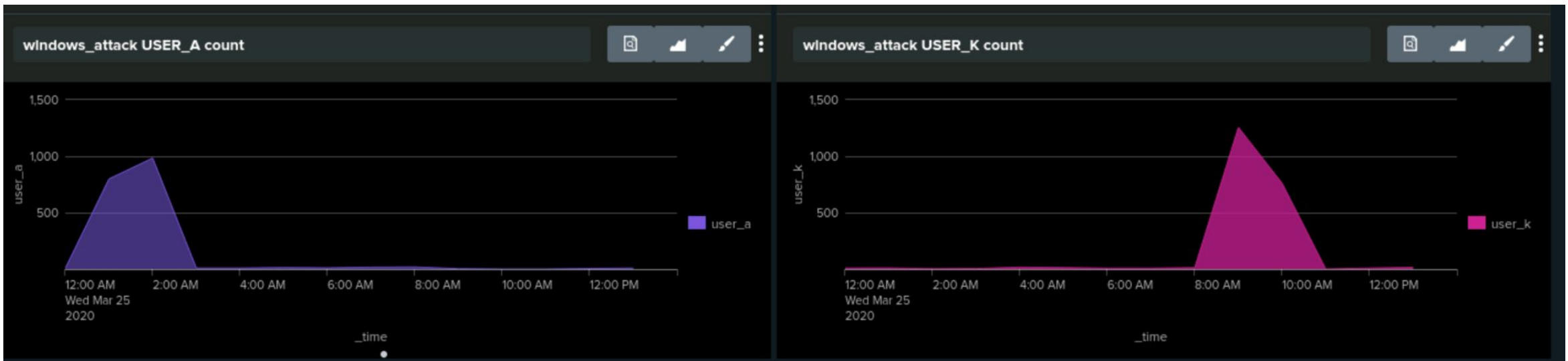
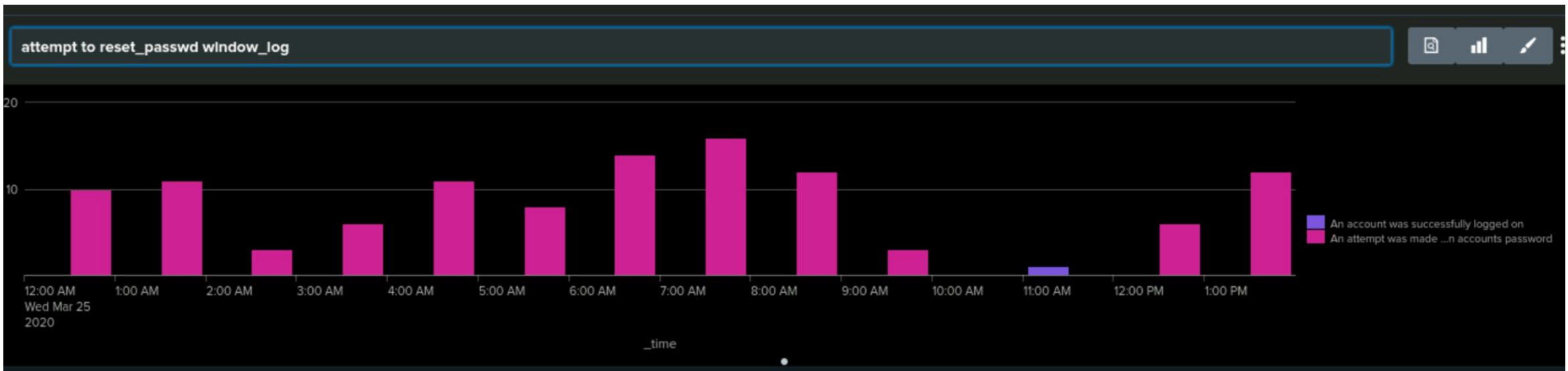
Summarize your findings from your alerts when analyzing the attack logs. Were the thresholds correct?

- The alerts effectively detected anomalies like high failed logins and account deletions. The thresholds were generally accurate, capturing abnormal activities, but some could be fine-tuned for better precision.

Attack Summary—Windows-Dashboard



Screenshots of Attack Logs



Attack Summary—Apache

The attack focused on exploiting web application vulnerabilities using high volumes of GET and POST requests targeting sensitive pages, indicating unauthorized access attempts.

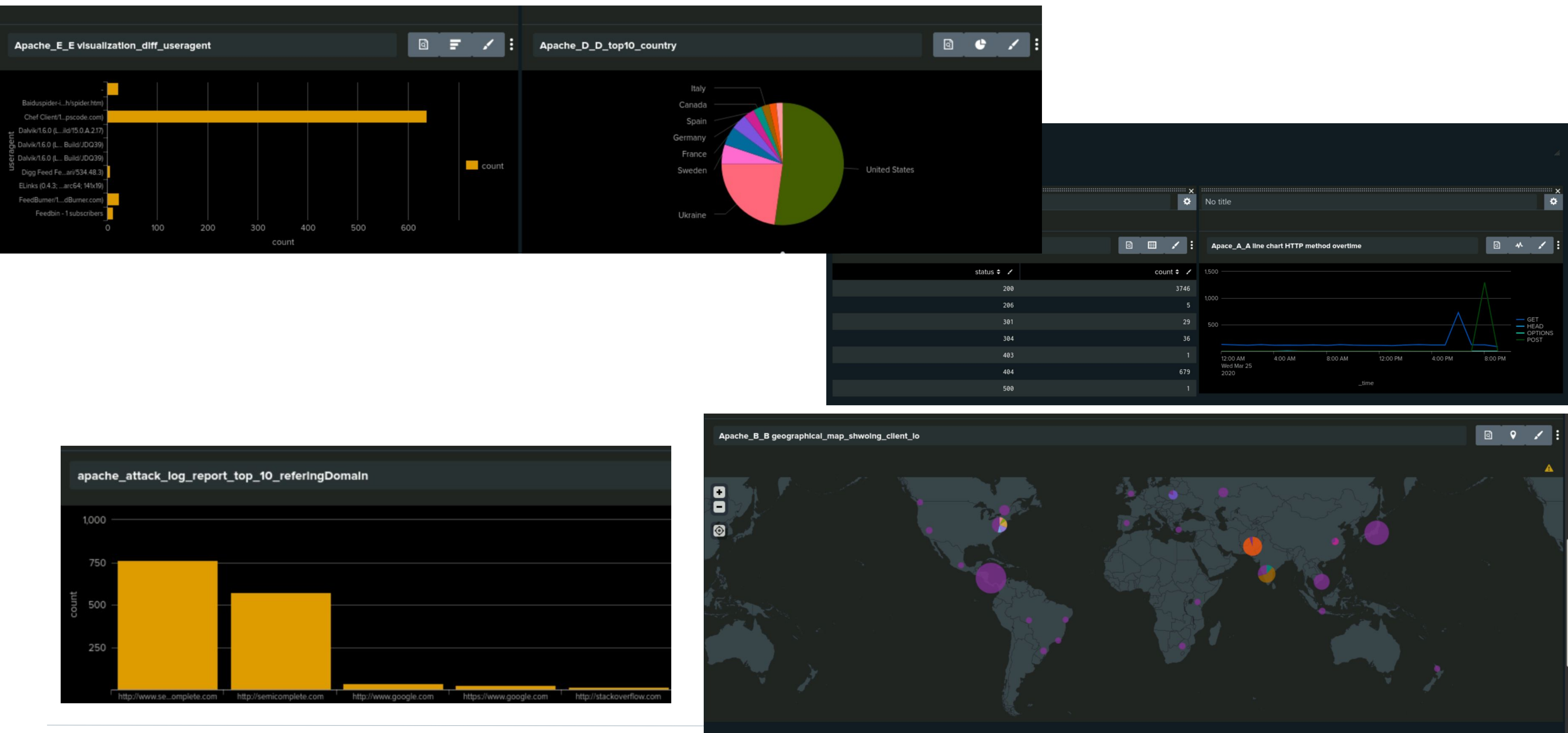
Attack Summary—Apache

Summarize your findings from your alerts when analyzing the attack logs. Were the thresholds correct?

The alerts identified unusual spikes in HTTP POST and non-US activity, validating that the thresholds were appropriately set to detect suspicious behavior.

The attack involved a high volume of HTTP POST and GET requests originating from multiple international locations, indicating a potential automated attack targeting specific endpoints on the server.

Attack Summary—Apache



Summary and Future Mitigations

Project 3 Summary

- What were your overall findings from the attack that took place?

The attack primarily targeted sensitive URIs such as administrative and login pages using methods like GET and POST at specific times, indicating attempts at unauthorized access and exploitation of web vulnerabilities.

- To protect VSI from future attacks, what future mitigations would you recommend?

To protect VSI, implement strict access controls on sensitive URIs, enhance monitoring and alerting for unusual activity, enforce strong authentication measures (e.g., MFA), and regularly update and patch web applications to close known vulnerabilities.