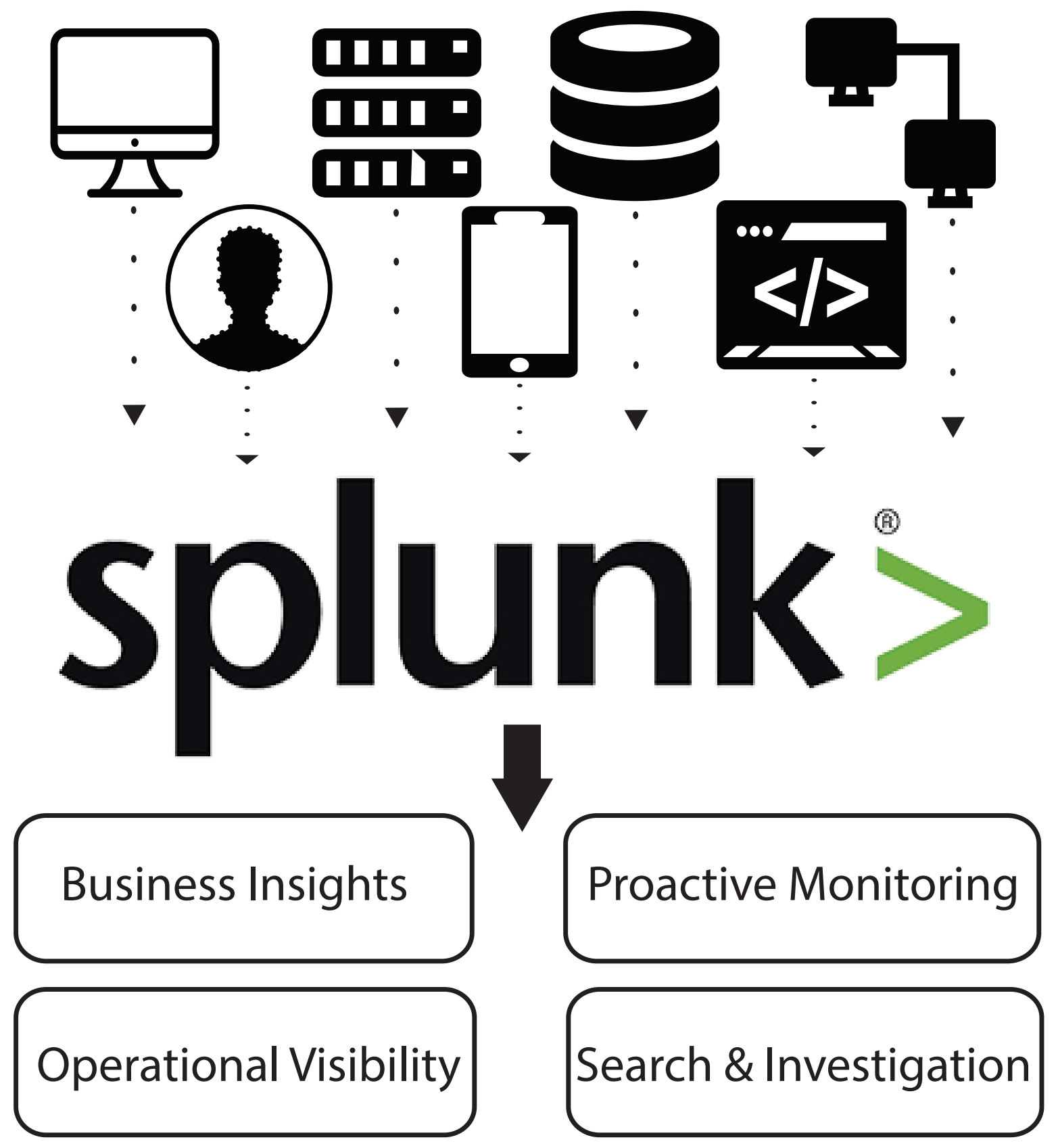
intelligentIR | Discovering the incident in Incident Response

STARBUCKS®

MORE INFO MORE PROBLEMS

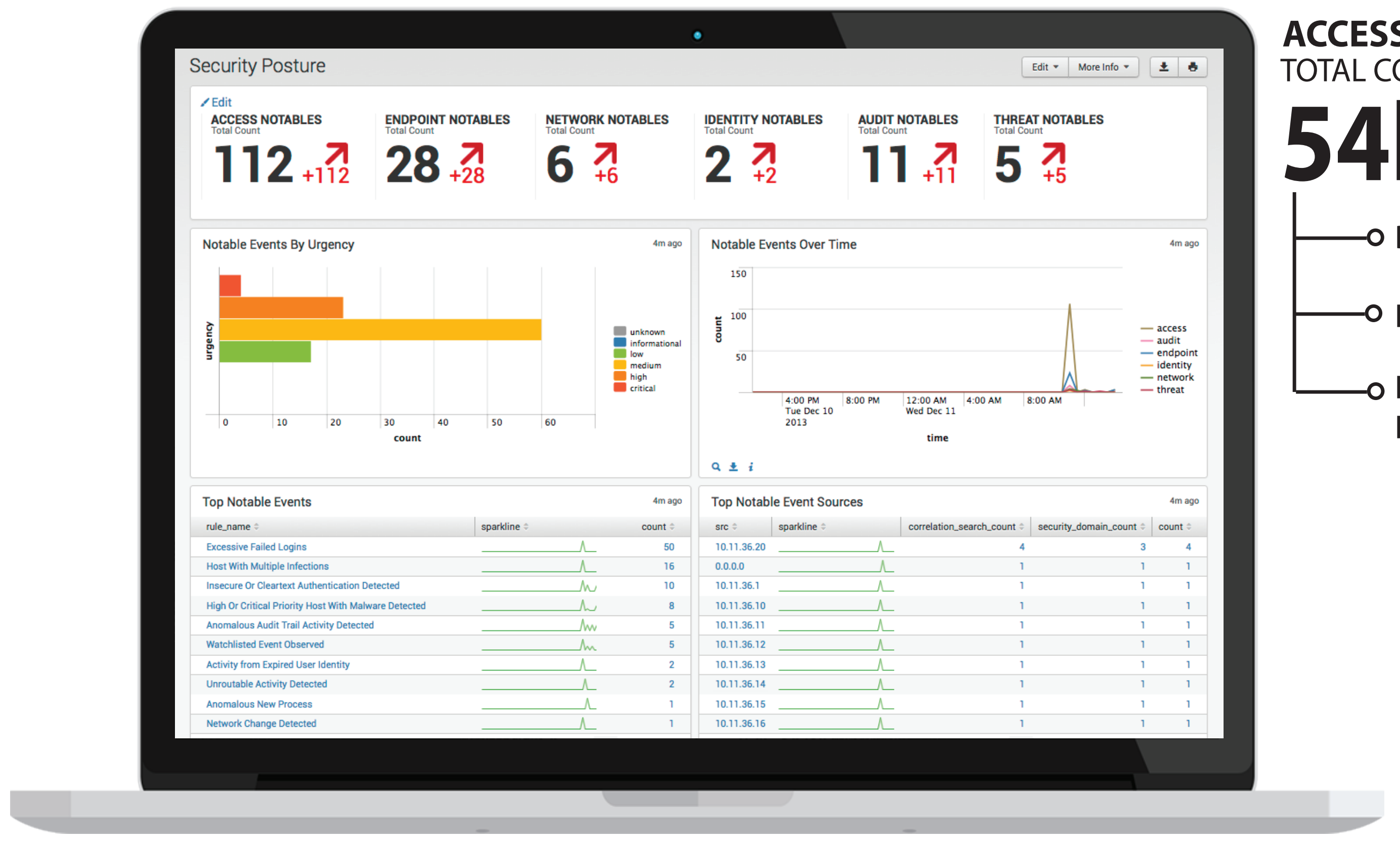
Starbucks currently uses Splunk as a Security Information and Event Management tool. Splunk aggregates machine data from all across the organization and indexes it to make it searchable for analysis and reporting



1 CURRENT LANDSCAPE

2 SECURITY EVENT MONITORING

The Splunk Enterprise Security application features various dashboards that communicate information about “notable” events. These events are defined by custom search queries that correlate events across differnent indexes or information sources



Security Posture

Edit | More Info

ACCESS NOTABLES Total Count 112 +112
ENDPOINT NOTABLES Total Count 28 +28
NETWORK NOTABLES Total Count 6 +6
IDENTITY NOTABLES Total Count 2 +2
AUDIT NOTABLES Total Count 11 +11
THREAT NOTABLES Total Count 5 +5

Top Notable Events (4m ago)

rule_name	sparkline	count
Excessive Failed Logins		50
Host With Multiple Infections		16
Insecure Or Cleartext Authentication Detected		10
High Or Critical Priority Host With Malware Detected		8
Anomalous Audit Trail Activity Detected		5
Watchlisted Event Observed		5
Activity from Expired User Identity		2
Unroutable Activity Detected		2
Anomalous New Process		1
Network Change Detected		1

Top Notable Event Sources (4m ago)

src	sparkline	correlation_search_count	security_domain_count	count
10.11.36.20		4	3	4
0.0.0.0		1	1	1
10.11.36.1		1	1	1
10.11.36.10		1	1	1
10.11.36.11		1	1	1
10.11.36.12		1	1	1
10.11.36.13		1	1	1
10.11.36.14		1	1	1
10.11.36.15		1	1	1
10.11.36.16		1	1	1

ACCESS NOTABLES
TOTAL COUNT
54k +2k

- Brute Force Access Behavior Detected
- Excessive Failed logins
- High or Critical Priority Individual Logging in to Infected Machine

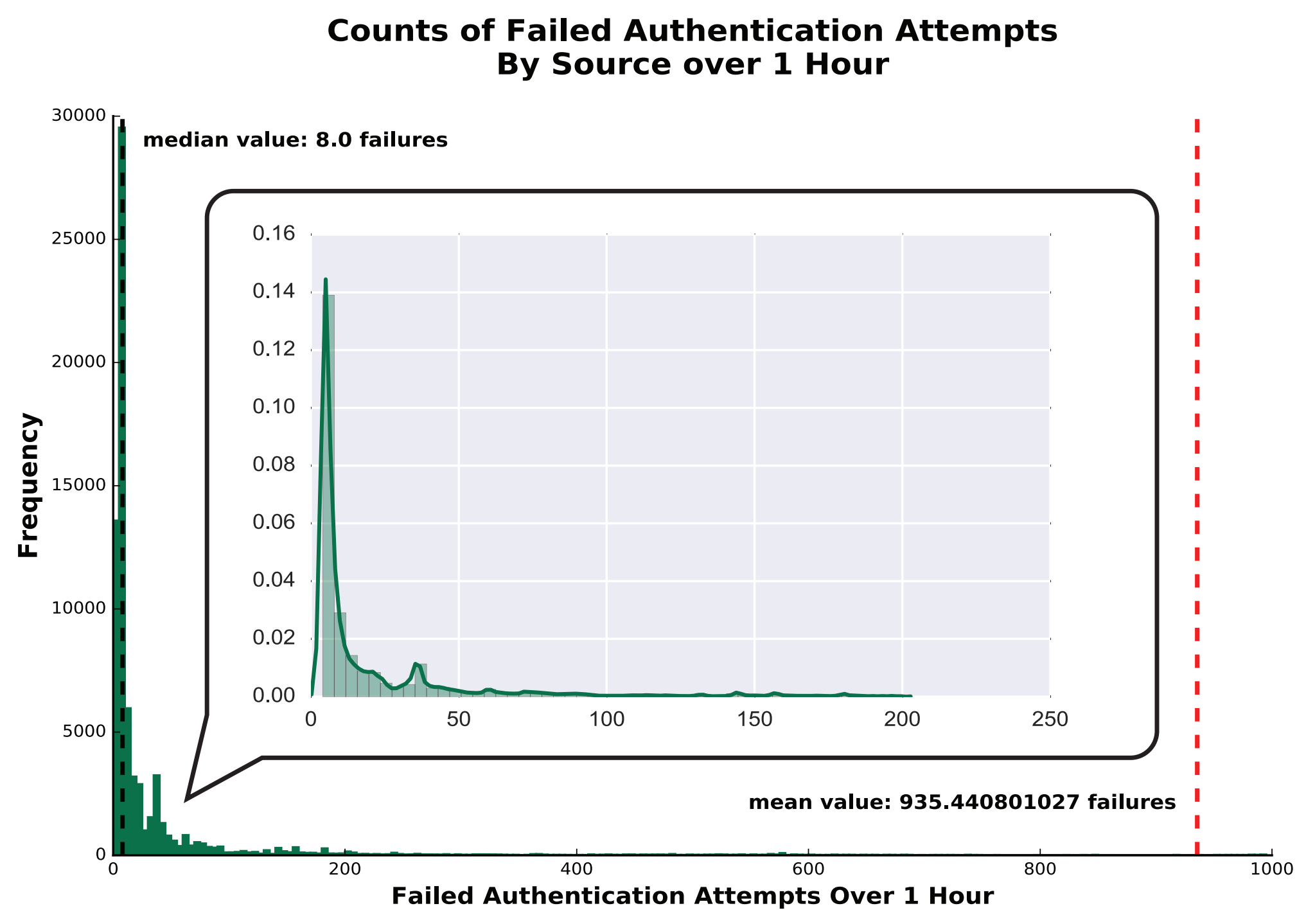
Too many notable events are being generated. Which ones do we prioritize for investgation and response?

INFORMATION PROBLEM 3

“IT’S NOT INFORMATION OVERLOAD. IT’S FILTER FAILURE” -CLAY SHIRKY

New Search | last 7 days

```
index="win" OR index="virtual" OR index="web"
| bucket _time span=1h
| stats values(src_ip) as ip, count as auth_attempts,
count(eval(action=="failure")) as failure,
count(eval(action=="success"))
as success, dc(app) as count_auth_apps, values(app) as auth_apps,
dc(user) as unique_users_count, values(user) as unique_user_list,
dc(dest) as dest_count, values(dest) as dest_list, values(signature)
as signature BY src _time
| WHERE success>0
| xsFindBestConcept failure from failures_by_src_count_1h in
authentication as concept | WHERE concept="medium" OR
concept="high" OR concept="extreme"
| eval hour=strftime(_time,"%H")
| eval proportion=failure/auth_attempts
```



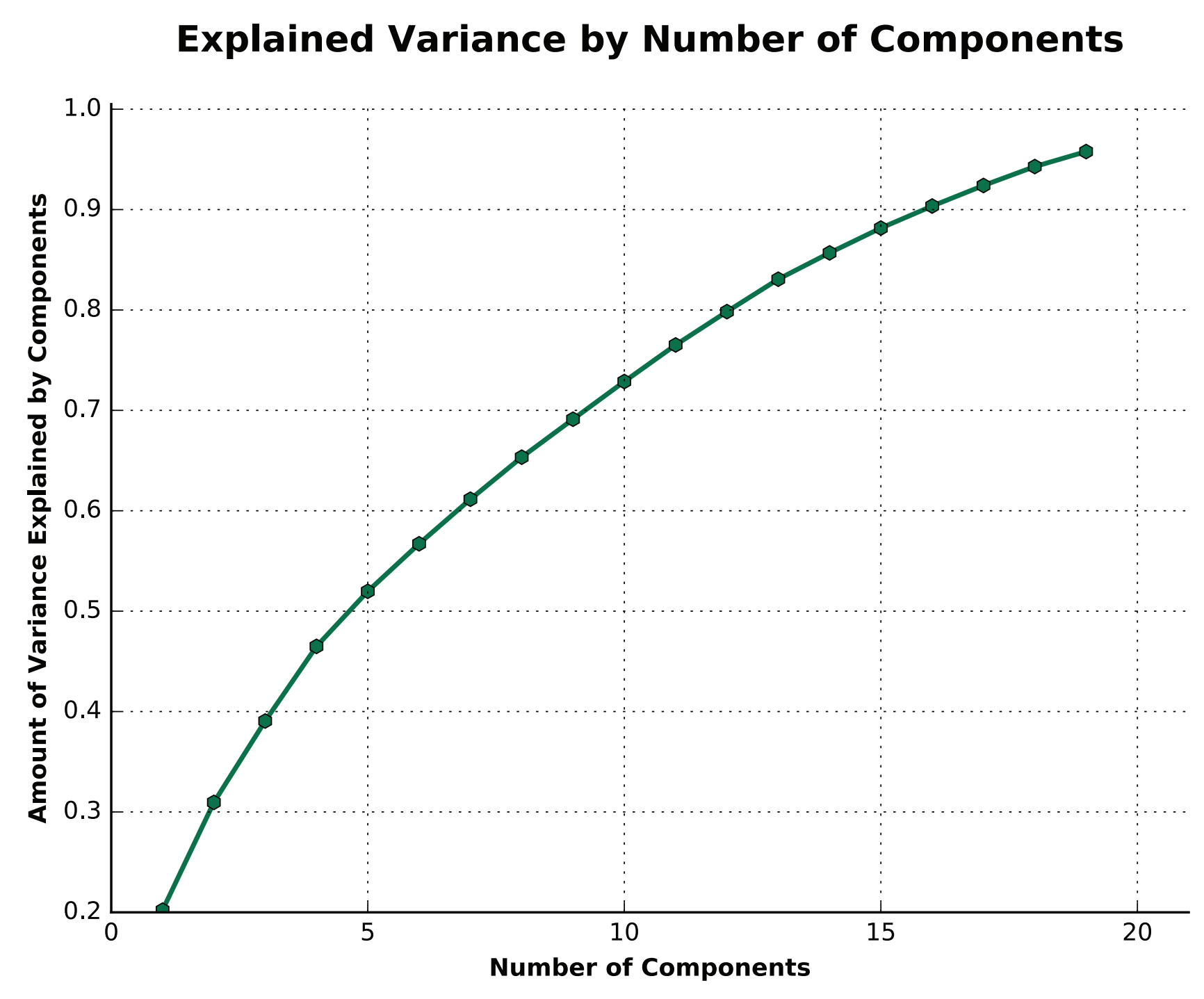
70k+ observations; 37 features; 4,321 unique sources

4 EXPLORING THE DATA & DEFINING A USE CASE

Machines produce the data; machines can learn from the data

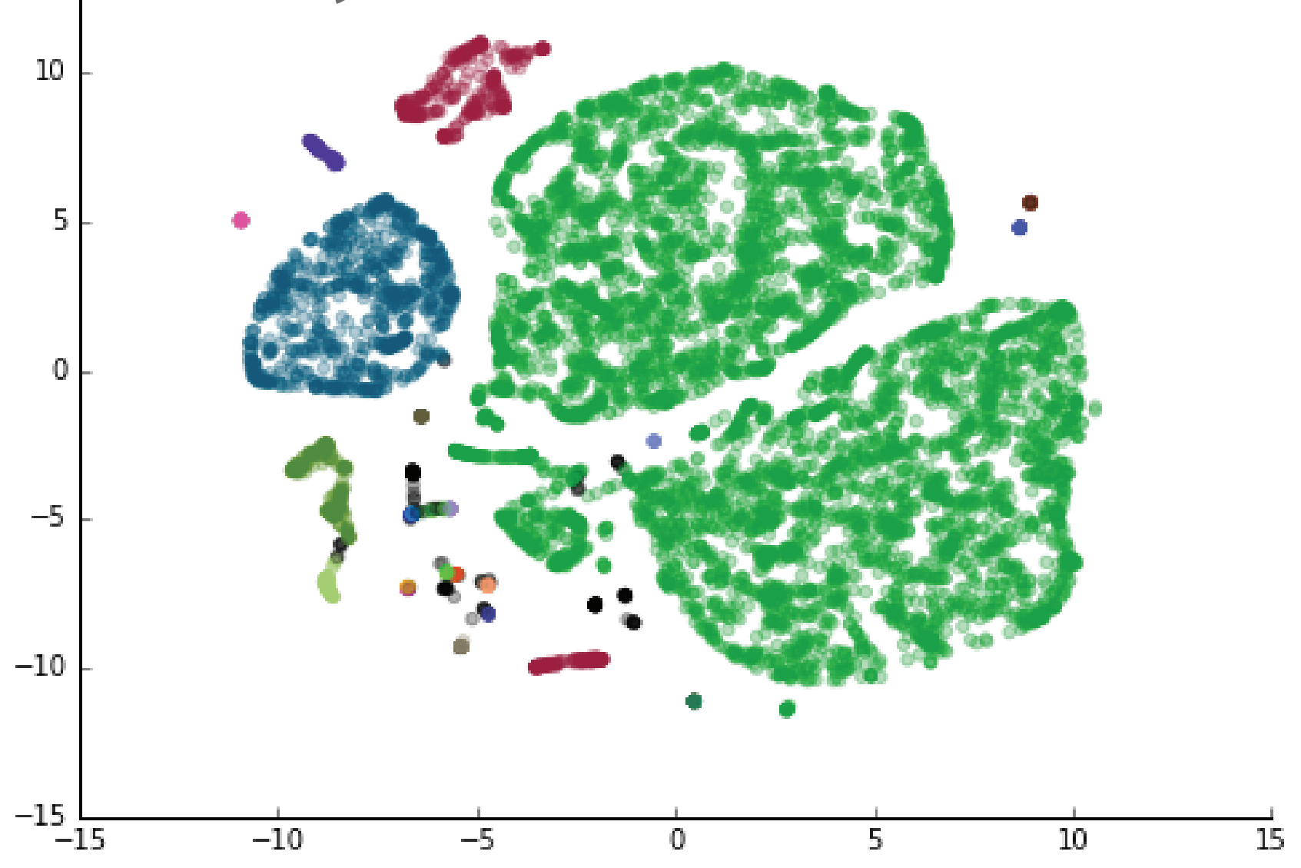
- Dimensionality Reduction with PCA
- Cluster Analysis with DBSCAN
- Random Forest Classification

5 PRINCIPAL COMPONENT ANALYSIS

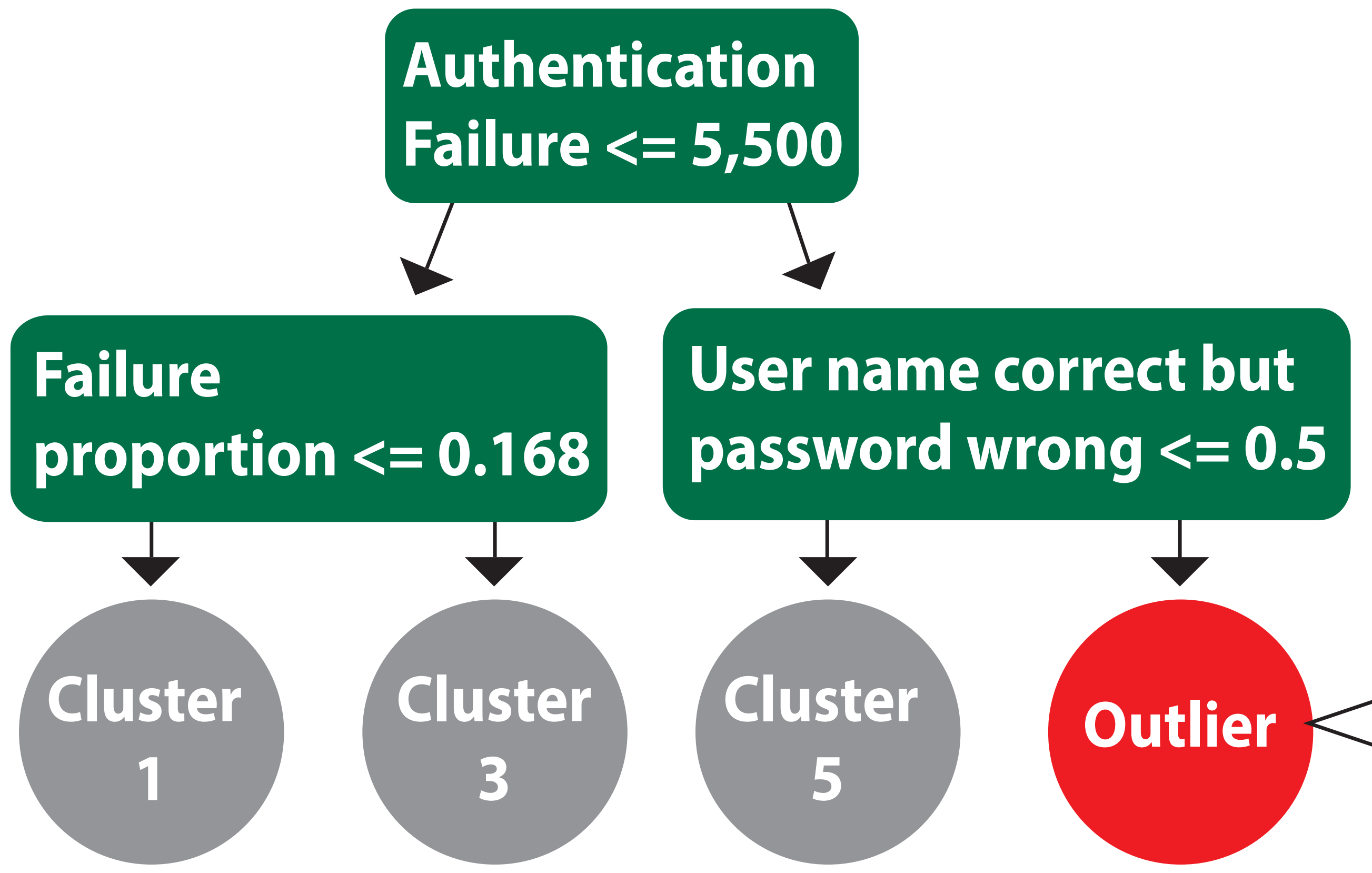


15 components capture nearly 90% of the variation in the data

Clustering identifies groups of events with similar characteristics. Can be used to identify and label events of interest



6 EVENT CLUSTERING


Authentication Failure <= 5,500
Failure proportion <= 0.168
User name correct but password wrong <= 0.5
Cluster 1
Cluster 3
Cluster 5
Outlier

Piroritize outliers because they represent behavior that is unlike other notable events. Underlying assumption is that most notable events are actually normal behavior

src	failure	unique_users	signatures	z_sco
src ip 1	16 (72%)	user x, user y, user x	a user account was created, user name does not exist	1.863

7 EVENT CLASSIFICATION

WHATS NEXT?

- The power of this analysis is that it identifies outliers with regard to other source behavior as well as a single source’s historical behavior. However, it is highly sensitive to the parameters used in clustering. Therefore exploring the characteristics of clusters is vital.
- A natural next step in this process involves identifying times when automated action could be taken. For example, if an event is classified as an outlier and meets various conditions we might suspend a user’s account.
- Machine learning can be applied to a wide range of security use cases. A particular use case of interest at Starbucks is using nal language processing to identify DNS exfiltration attacks

FURTURE WORK 8

Kyle Estlick
Mid-Career MSIM
Data Science | Information Security

Information School
UNIVERSITY of WASHINGTON