PRACTICAL SOC METRICS

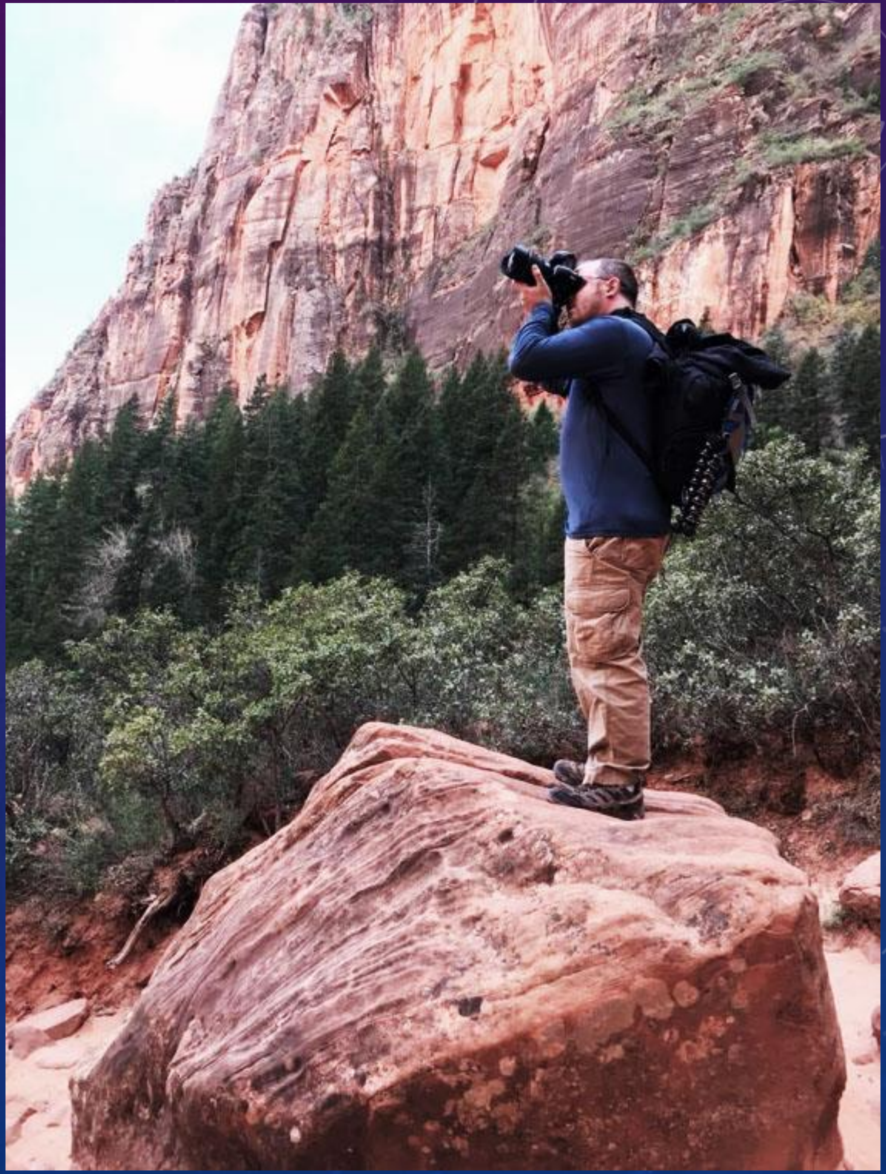
PRESENTED BY CARSON ZIMMERMAN

IN COLLABORATION WITH CHRIS CROWLEY

FIRST 2019

ABOUT CARSON

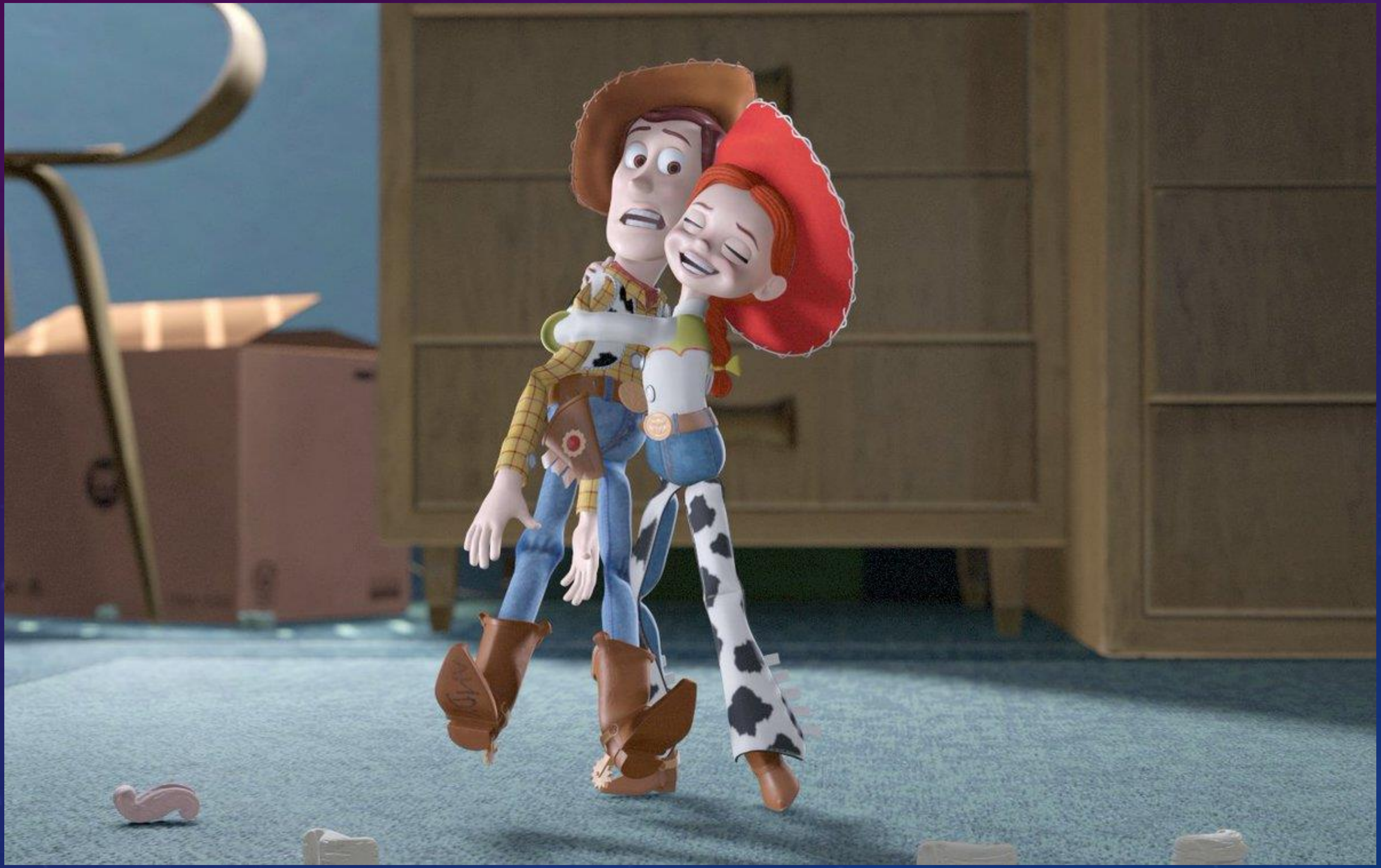
- Worked in Security Operations for ~15 years
- SOC Engineering Team Lead @ Microsoft
- Previously SOC engineer, analyst & consultant @ MITRE
- Checkout my book if you haven't already: https://www.mitre.org/publications/all/ten-strategies-of-a-world-class-cybersecurity-operations-center



ABOUT CHRIS

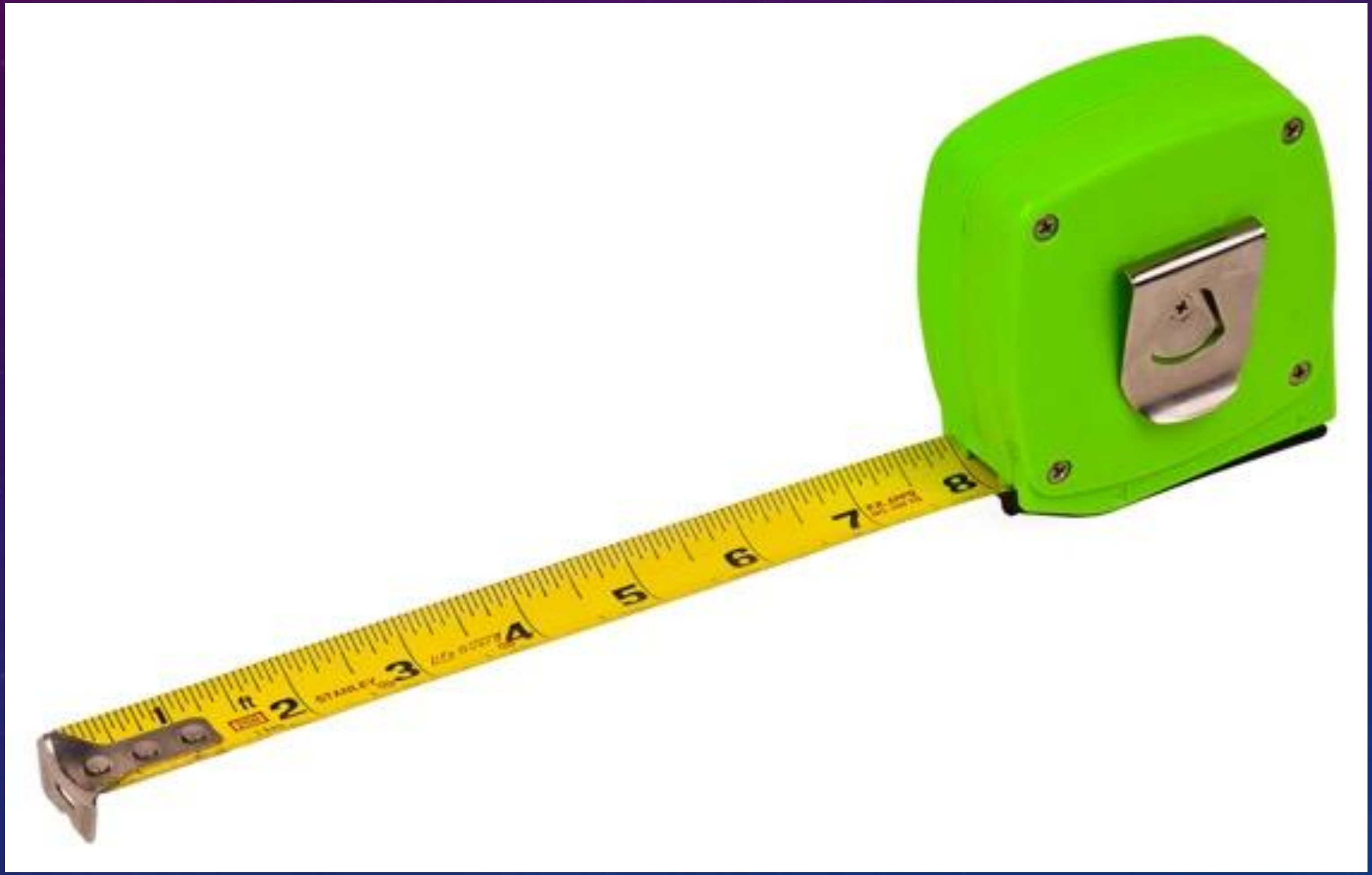
- Independent Consultant (Montance.com)
- SANS Institute
  - Senior Instructor & Course Author
  - SOC Survey Author (2017, 2018, 2019)
  - Security Operations Summit Chair
- SOC-class.com – Security Operations Class on building & running a SOC
- Engagements with Defense, Education, Energy, Financial, IT, Manufacturing, Science, Software Development, …

PICK SOMETHING YOU LOVE…



http://disney.wikia.com/wiki/File:TS2_Jessie_hugs_Woody.jpg

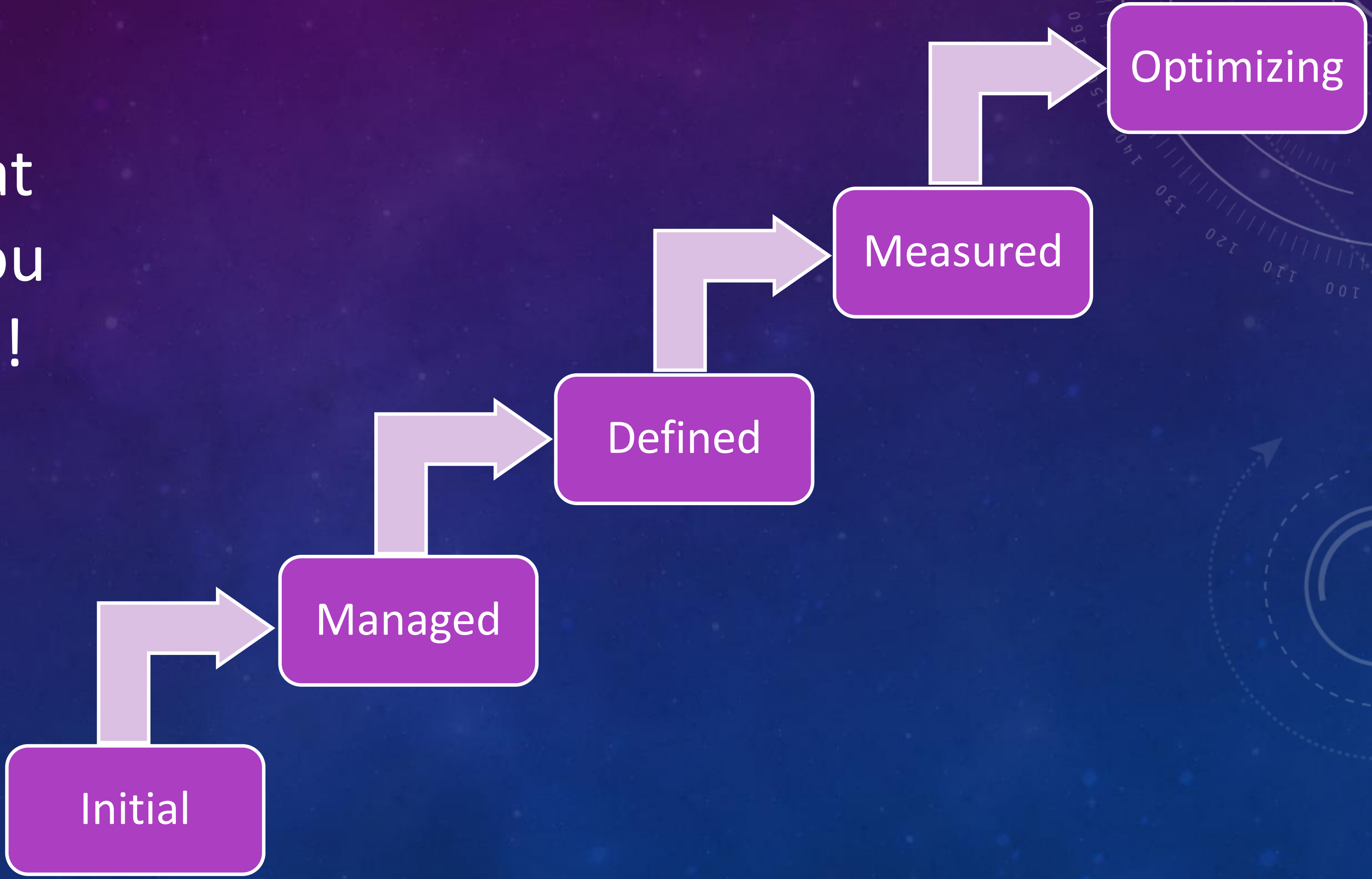
…AND MEASURE IT



https://en.wikipedia.org/wiki/Tape_measure#/media/File:Measuring-tape.jpg

MEASURING THINGS USUALLY DRIVES CHANGE

Even if you're not at CMM level >= 3, you can still get started!



METRICS ARE LIKE LIGHTSABERS

https://www.maxpixel.net/Laser-Sword-Lightsaber-Green-Science-Fiction-Space-1675211

THEY CAN BE USED FOR GOOD…

https://www.scifinow.co.uk/blog/top-5-star-wars-scenes-we-want-to-see-on-blu-ray/

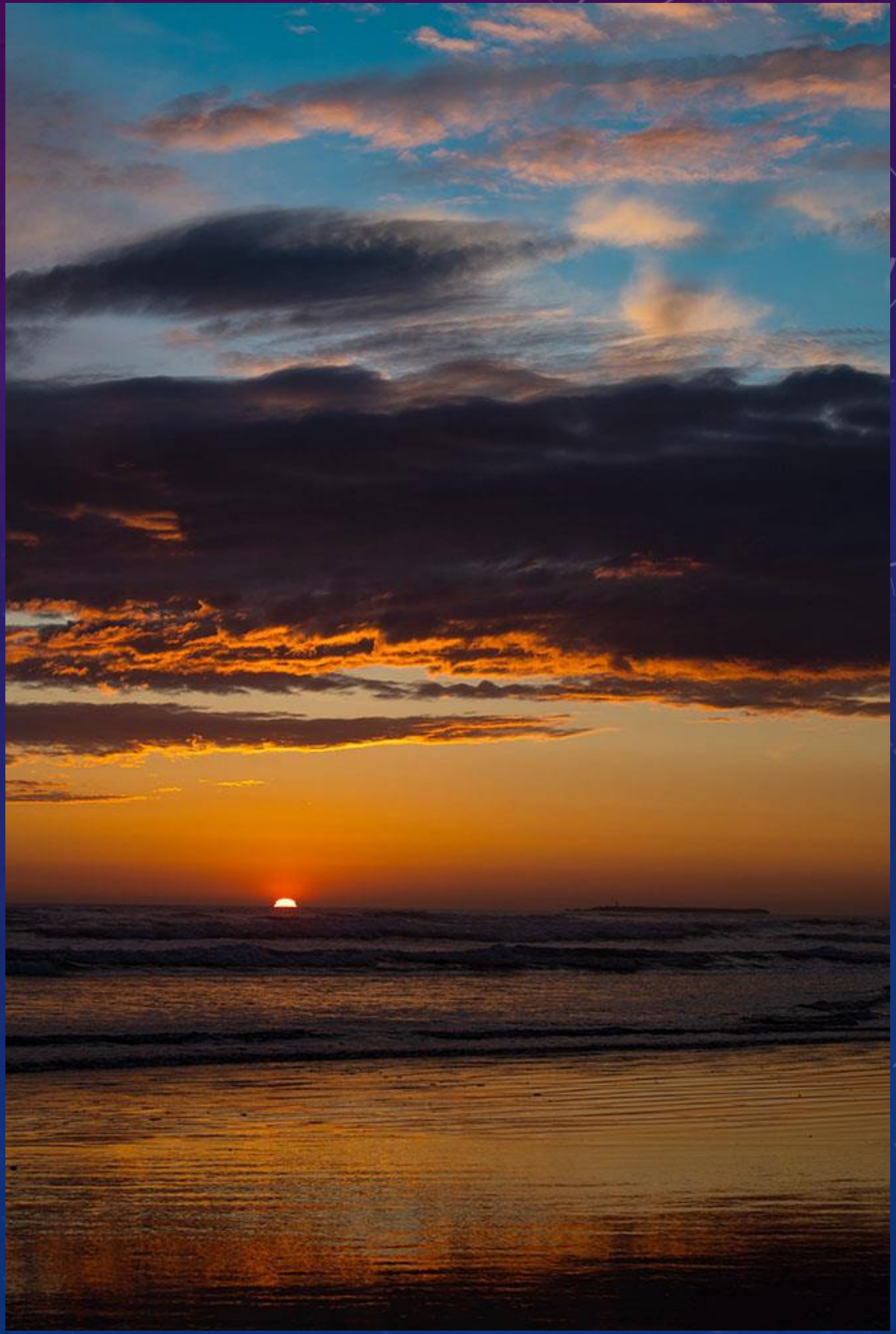
…AND FOR EVIL

http://starwars.wikia.com/wiki/File:UnidentifiedClan-RotS.jpg

SOME DEFINITIONS

- Metrics: things you can objectively measure
  - Input: behaviors and internal mechanisms
  - Output: results, typically customer-facing
- Service level agreements (SLAs): agreement/ commitment between provider and customer
- Service level objectives (SLOs): performance metric or benchmark associated with an SLA

https://searchcio.techtarget.com/answer/Whats-the-difference-between-SLO-and-SLA



TOP TIPS

- Metric data should be free and easy to calculate
  - ½ of all SOCs collect metrics according to SANS SOC survey 2017 & 2018
- There should be a quality measure that compensates for perversion anytime there’s a time based metric
- Metrics aren’t (necessarily) SLOs
  - The metric is there to help screen, diagnose, and assess performance
  - Don’t fall into a trap of working to some perceived metric objective
  - Any metric should have an intended effect, and realize the measurement and calculation isn’t always entirely valid
- Expectations, messaging, objectives- all distinct!

DATA SOURCES

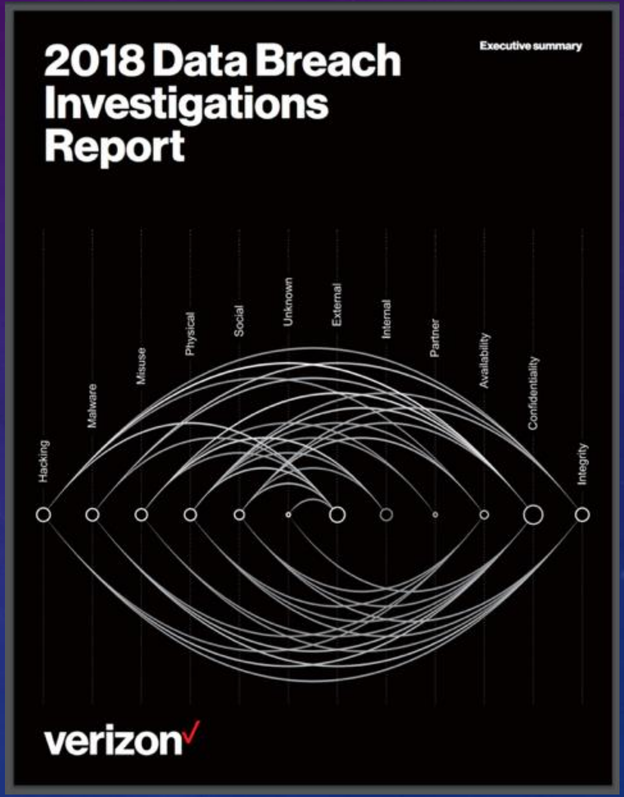
- SOC Ticketing/case management system
- SIEM / analytic platform / EDR- anywhere analysts create detections, investigate alerts
- SOC code repository
- SOC budget
  - CAPEX including hardware & software
  - OPEX including people & cloud
- Enterprise asset management systems
- Vulnerability management

https://video-images.vice.com/articles/5b02e43f187df600095f5e7c/lede/1526917810059-GettyImages-159825349.jpeg

EXISTING RESOURCES

- SOC CMM: measure your SOC top to bottom
- VERIS Framework: track your incidents well
- SANS SOC Survey: recent polls from your peers

https://www.fireeye.com/content/dam/collateral/en/mtrends-2018.pdf

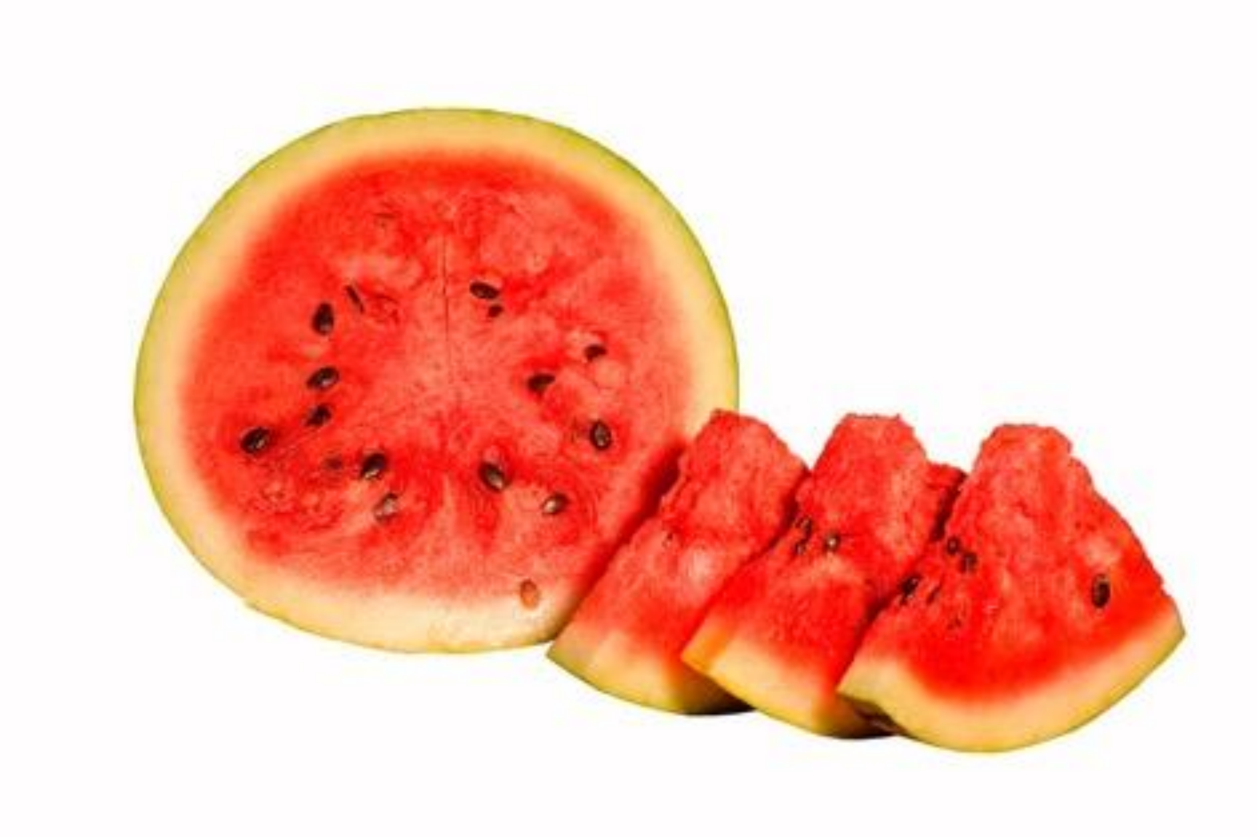


https://www.verizonenterprise.com/resources/reports/rp_DBIR_2018_Report_execsummary_en_xg.pdf

EXAMPLE METRICS

METRIC FOCUS 1: DATA FEED HEALTH

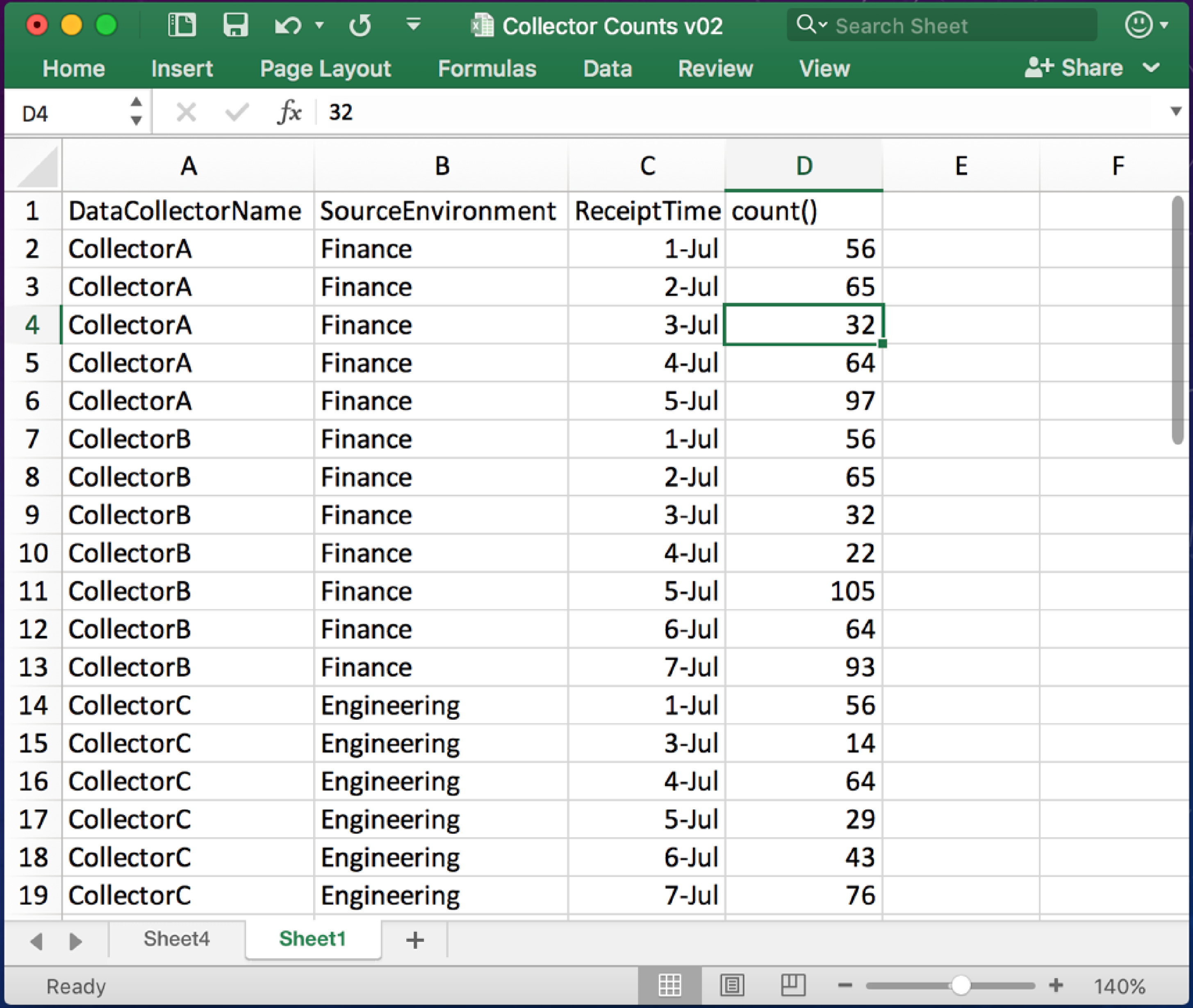
- Is it “green”
- What is green anyway?
- Just because it’s up doesn’t mean all is well
  - Delays in receipt
  - Drops
    - Temporary
    - Permanent
  - Blips



https://en.wikipedia.org/wiki/Watermelon#/media/File:Watermelon_cross_BNC.jpg

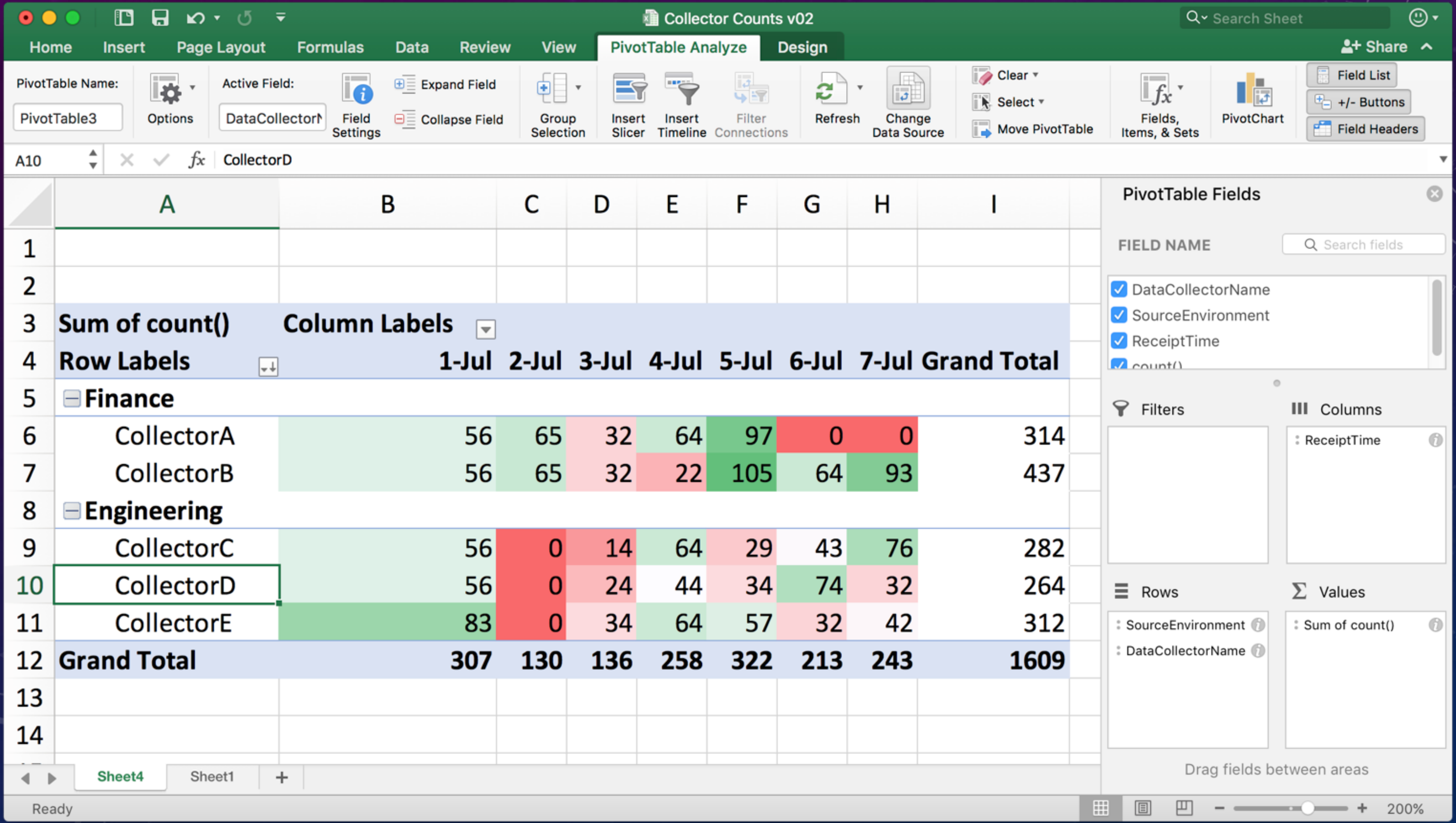
HOW MANY EVENTS ARE WE RECEIVING?

Select count(*) | group by DataCollectorName, SourceEnvironment, bin(ReceiptTime, day)



	A	B	C	D
1	DataCollectorName	SourceEnvironment	ReceiptTime	count()
2	CollectorA	Finance	1-Jul	56
3	CollectorA	Finance	2-Jul	65
4	CollectorA	Finance	3-Jul	32
5	CollectorA	Finance	4-Jul	64
6	CollectorA	Finance	5-Jul	97
7	CollectorB	Finance	1-Jul	56
8	CollectorB	Finance	2-Jul	65
9	CollectorB	Finance	3-Jul	32
10	CollectorB	Finance	4-Jul	22
11	CollectorB	Finance	5-Jul	105
12	CollectorB	Finance	6-Jul	64
13	CollectorB	Finance	7-Jul	93
14	CollectorC	Engineering	1-Jul	56
15	CollectorC	Engineering	3-Jul	14
16	CollectorC	Engineering	4-Jul	64
17	CollectorC	Engineering	5-Jul	29
18	CollectorC	Engineering	6-Jul	43
19	CollectorC	Engineering	7-Jul	76

3 MINUTES LATER…



PivotTable Name: PivotTable3

Active Field: DataCollectorN

A10 | CollectorD

Sum of count()	Column Labels							
Row Labels	1-Jul	2-Jul	3-Jul	4-Jul	5-Jul	6-Jul	7-Jul	Grand Total
⊟ Finance								
CollectorA	56	65	32	64	97	0	0	314
CollectorB	56	65	32	22	105	64	93	437
⊟ Engineering								
CollectorC	56	0	14	64	29	43	76	282
CollectorD	56	0	24	44	34	74	32	264
CollectorE	83	0	34	64	57	32	42	312
Grand Total	307	130	136	258	322	213	243	1609

PivotTable Fields

FIELD NAME

☑ DataCollectorName
☑ SourceEnvironment
☑ ReceiptTime
☑ count()

Filters

Columns: ReceiptTime

Rows: SourceEnvironment, DataCollectorName

Values: Sum of count()

Drag fields between areas

ADVANCED: AUTO DETECTION OF OUTAGES

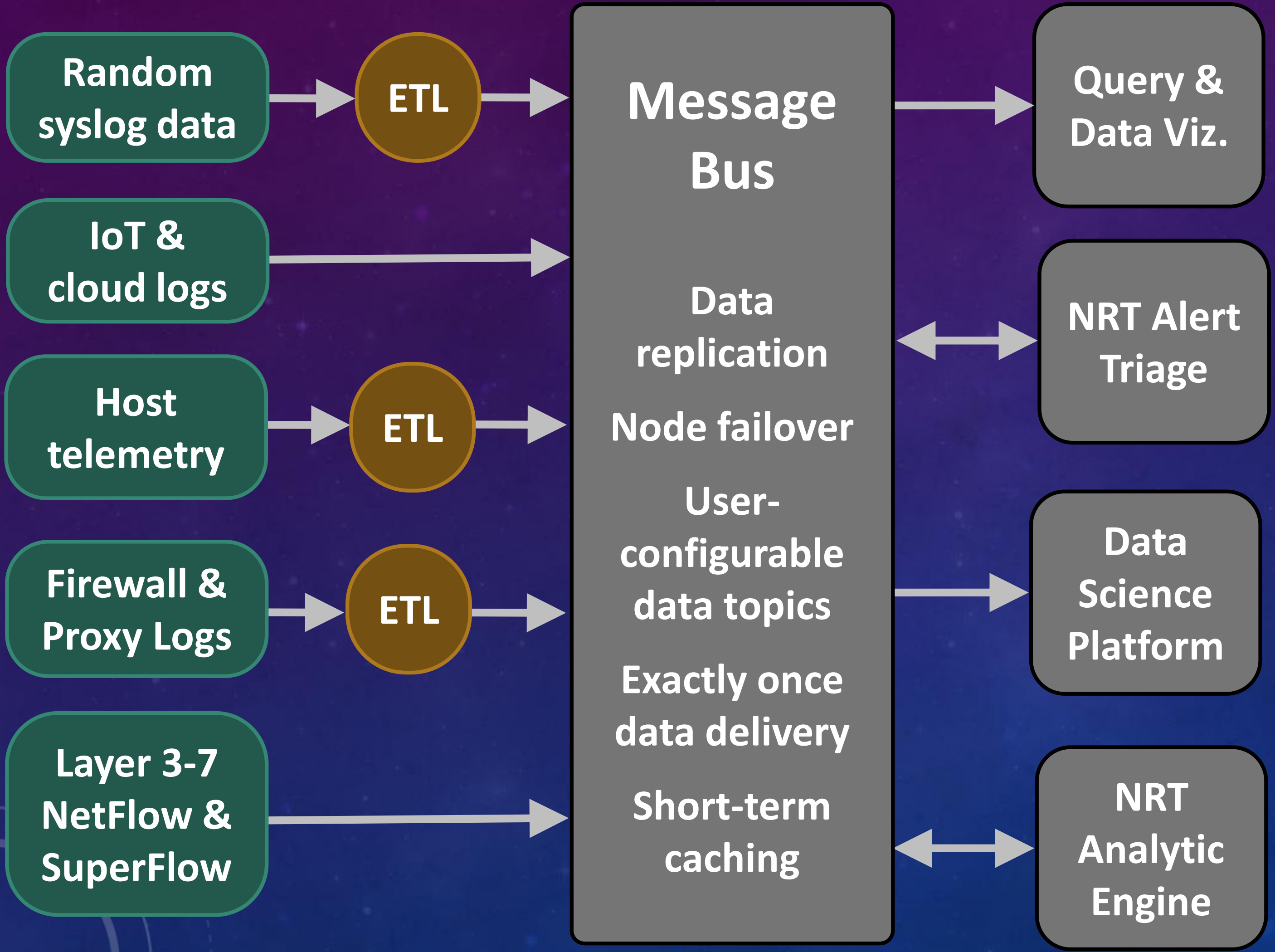
```
OldCounts = Select OldCount=count(*)/7, OldDevices= distinct(deviceHostName)
| where ReceiptTime < ago(1 day) and ReceiptTime > ago(8 days)
| group by DataCollectorName, SourceEnvironment;

NewCounts = Select NewCount=count(*), NewDevices= distinct(deviceHostName)
| where ReceiptTime > ago(1 day)
| group by DataCollectorName, SourceEnvironment;

Join NewCounts on OldCounts by DataCollectorName, SourceEnvironment
| project CountRatio = NewCount/OldCount,
DeviceRatio = NewDevices/OldDevices
| IsBroken = OR( CountRatio < 25%, DeviceRatio < 50%)
```

RESULT

	OldCount	NewCount	OldDevices	NewDevices	IsBroken
Collector A	2230	2120	1002	934	No
Collector B	1203	1190	894	103	Yes
Collector C	3203	3305	342	325	No
Collector D	1120	305	569	234	Yes
Collector E	342	102	502	496	Yes

- Detection of dead, slow or lagging collectors or sensors is fully automated
- Consider human eyes on: weekly or monthly

ADVANCED: MEASURE TIME EVERYWHERE



Latency as a factor of:

1. Clock skew
2. Systems rejoining the network & network outages
3. Lack of capacity:
   a. Ingest & parsing
   b. Decoration / enrichment
   c. NRT analytics & correlation
   d. Batched query

METRIC FOCUS 2: COVERAGE

Dimensions:

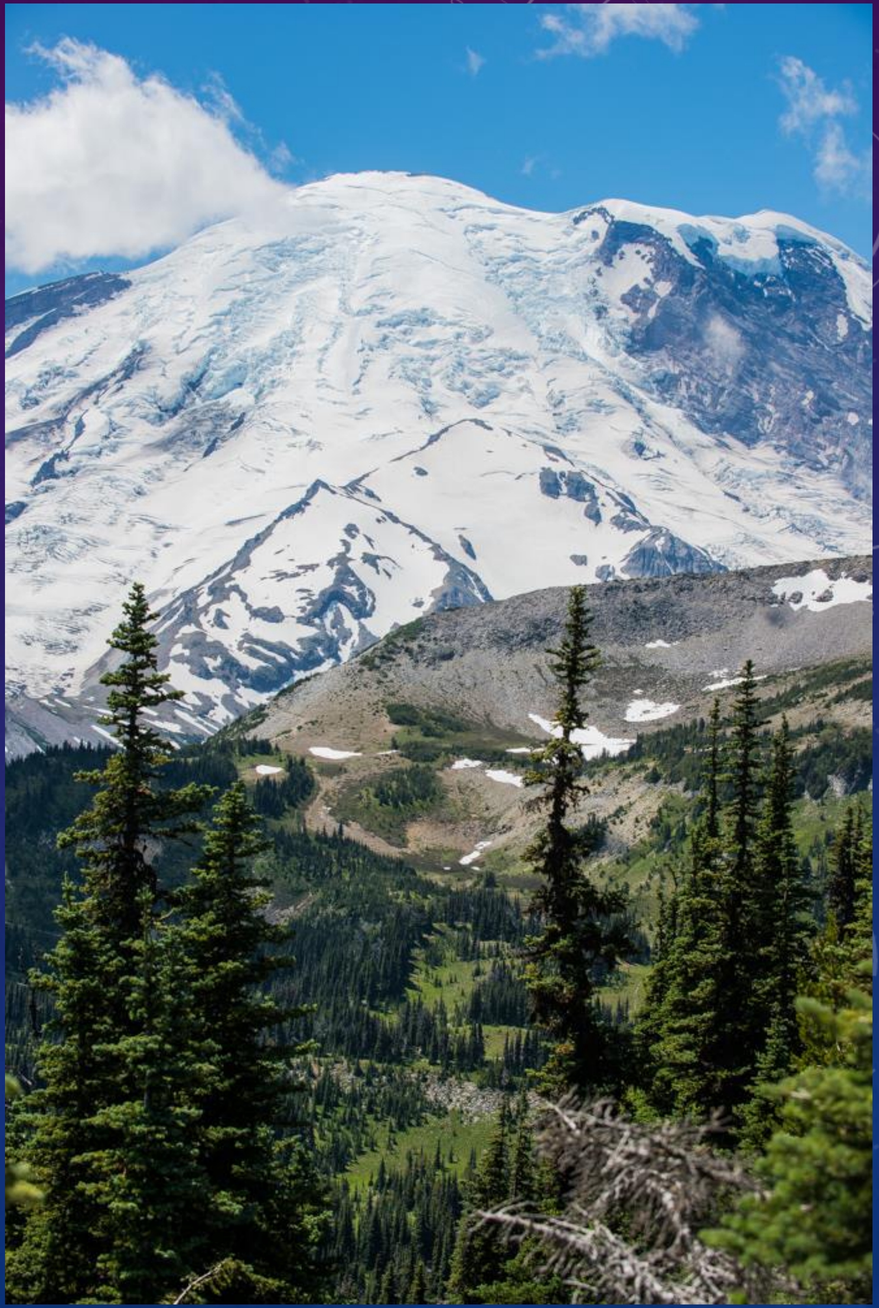
1. Absolute number *and* percentage of coverage per compute environment/enclave/domain
2. Kill chain or ATT&CK cell
3. Layer of the compute stack (network, OS, application, etc.)
4. Device covered (Linux, Windows, IoT, network device)

Tips:

1. Never drive coverage to 100%
   a. You don’t know what you don’t know
   b. Always a moving target
2. There is always another environment to cover, customer to serve
3. There will always be more stones to turn over; don’t ignore any of these dimensions

MANAGED VS WILDERNESS

- Percentage of systems “managed”:
  - Inventoried?
  - Tied to an asset/business owner?
  - Tied to a known business/mission function?
  - Subject to configuration management?
  - Assigned to a responsible security team/POC?
  - Risk assessed?
- If all are yes: it’s managed
- If not: it’s “wilderness”
- SOC observed device counts help identify “unknown unknowns” in the wilderness



VALIDATING DATA FEED & DETECTION COVERAGE

1. Expected heartbeat & true activity from every sensor and data feed
2. Detection triggers
    a. Injected late into pipeline as synthetic events: consider “unit” tests for each of your detections
    b. Injected early into pipeline as fake “bad” activity on hosts or networks
3. Blue/purple/red teaming: strong way to test your SOC!

MONITORING SLAS/SLOS

- SLA: Agreement = monetary (or other penalty) for failing to meet
- SLO: Objective = no specific penalty agreed to for failing to meet
- Institution & missions specific where these need to be set in place
- Don’t monitor everything the same way!
  - Instrumentation, custom detections, response times, retention

Basic Service

- Host EDR
- Network logs
- Standard mix of detections
- Yearly engagement

Advanced Service

- Basic, plus:
- 3 application logs
- 1 focused detection/quarter
- Quarterly engagement

METRIC FOCUS 3: SCANNING AND SWEEPING

Basic

- # + % of known on prem & cloud assets scanned for vulns
- Amount of time it took to compile vulnerability/risk status on covered assets during last high CVSS score “fire drill”
- Number of people needed to massage & compile these numbers monthly

Advanced

- Time to sweep and compile results for a given vuln or IOC:
  - A given domain/forest identity plane
  - Everything Internet-facing
  - All user desktop/laptops
  - Everything
- # + % of assets you can’t/don’t cover (IoT, network devices, etc.)

METRIC FOCUS 4: YOUR ANALYTICS

Basics:

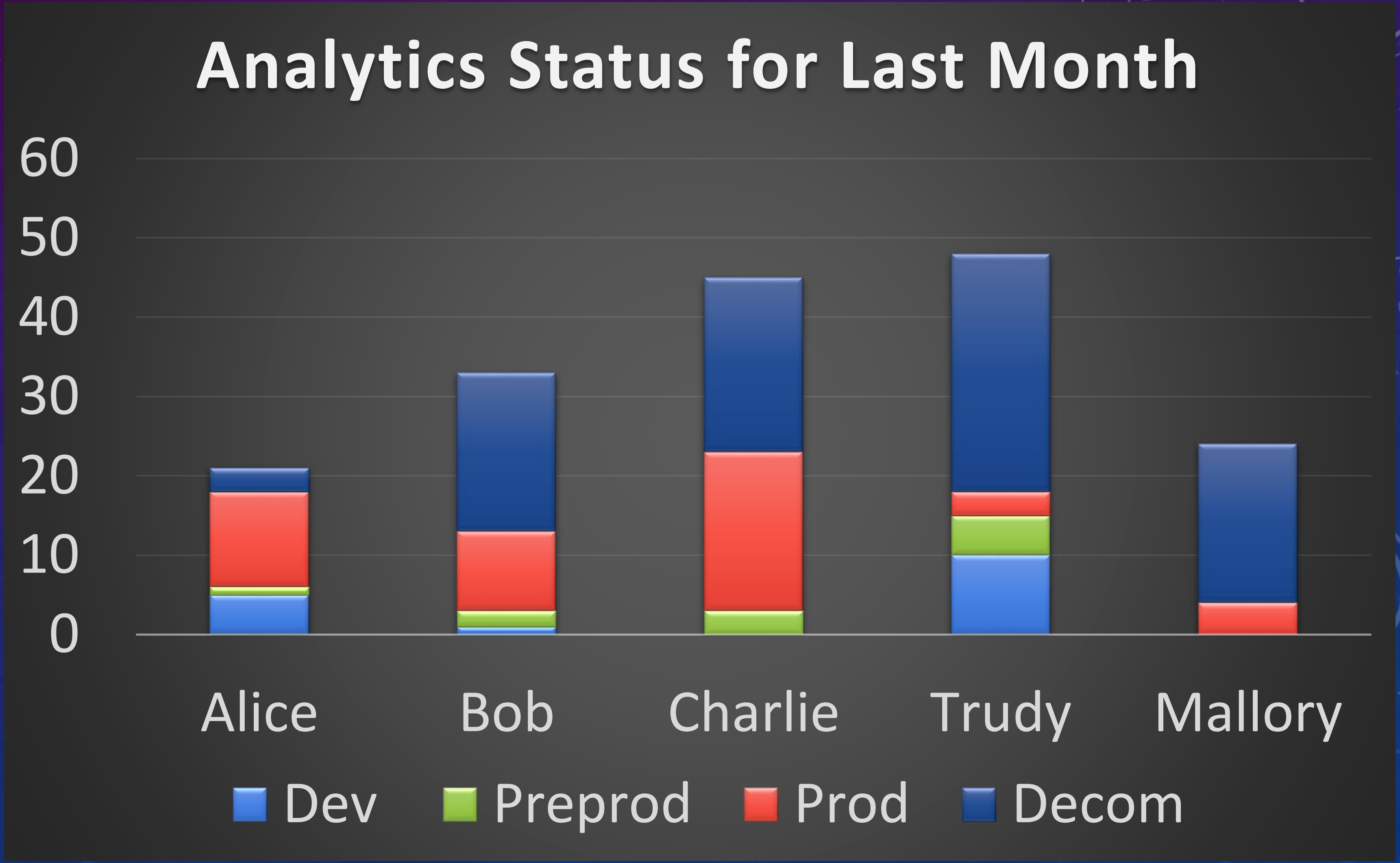
1. Name
2. Description
3. Kill chain mapping
4. ATT&CK cell mapping
5. Depends on which data type(s) (OS logs, Netflow, etc.)
6. Covers which environments/enclave
7. Created- who, when

Advanced:

8. Runs in what framework (Streaming, batched query, etc.)
9. Last modified- who, when
10. Last reviewed- who, when
11. Status- dev, preprod, prod, decom
12. Output routes to… (analyst triage, automated notification, etc.)

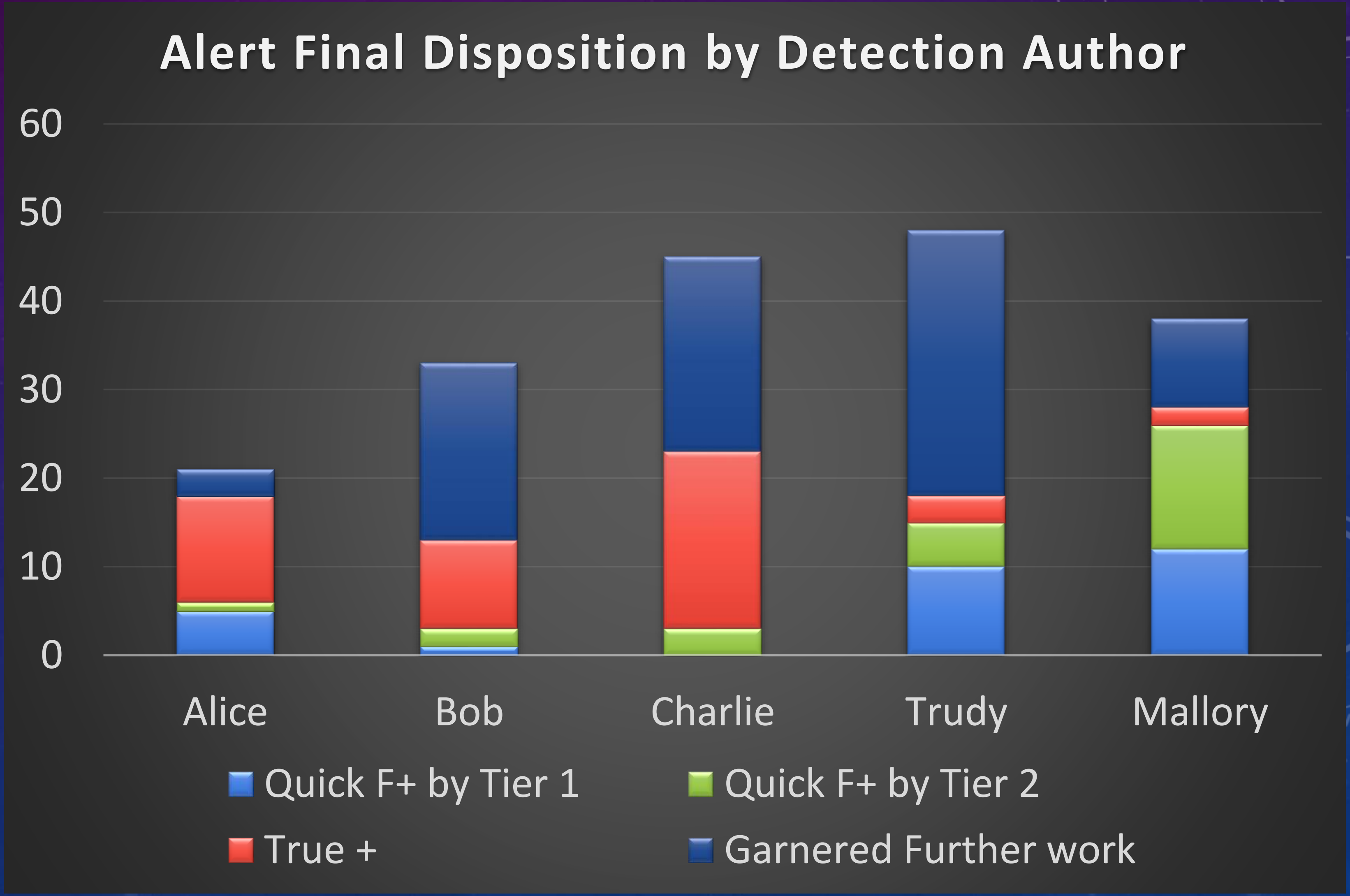
MEASURE ANALYST PRODUCTIVITY

- Is this good or evil?
- Can this be gamed?

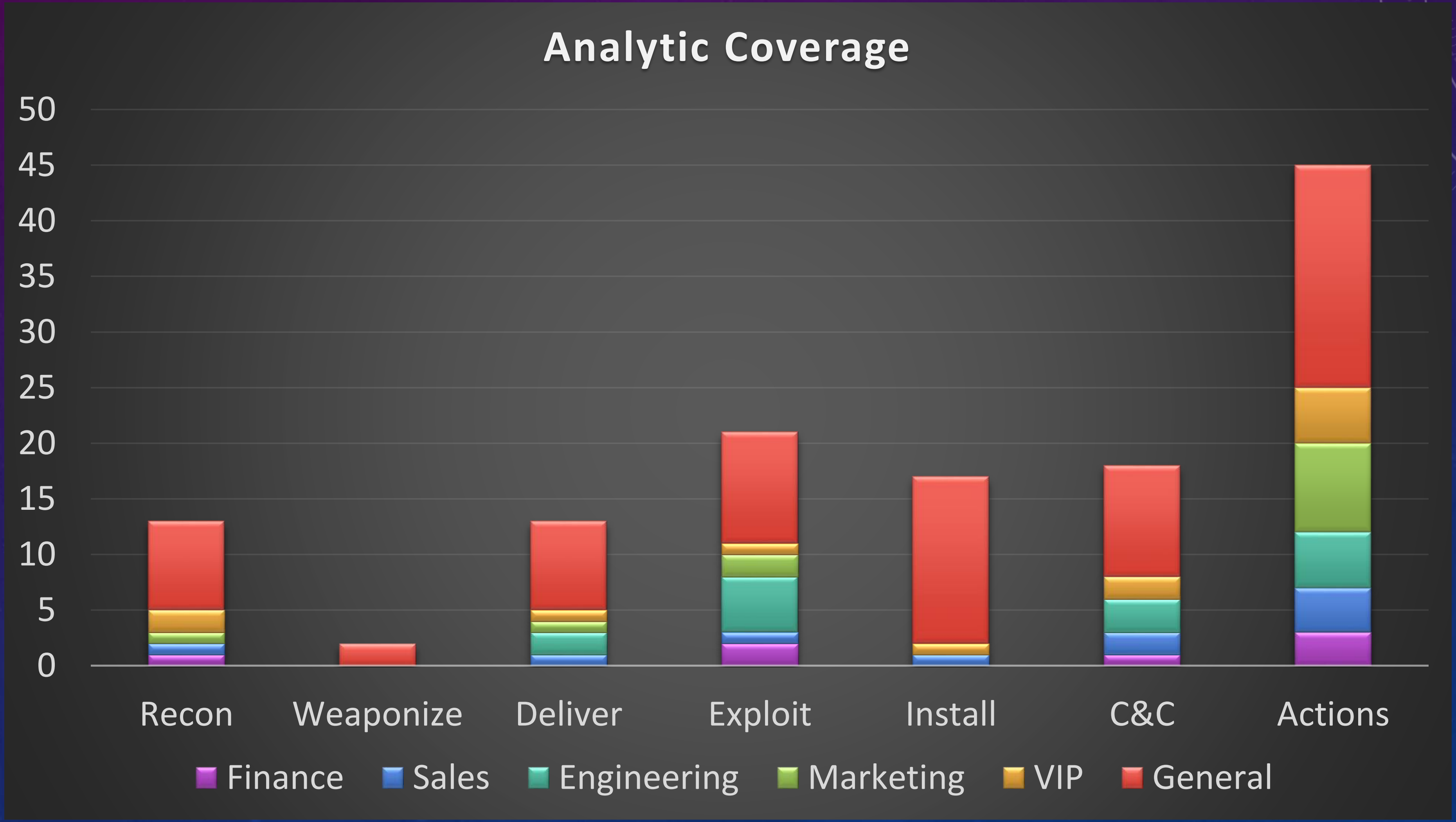


HOW FRUITFUL ARE EACH AUTHOR’S DETECTIONS?

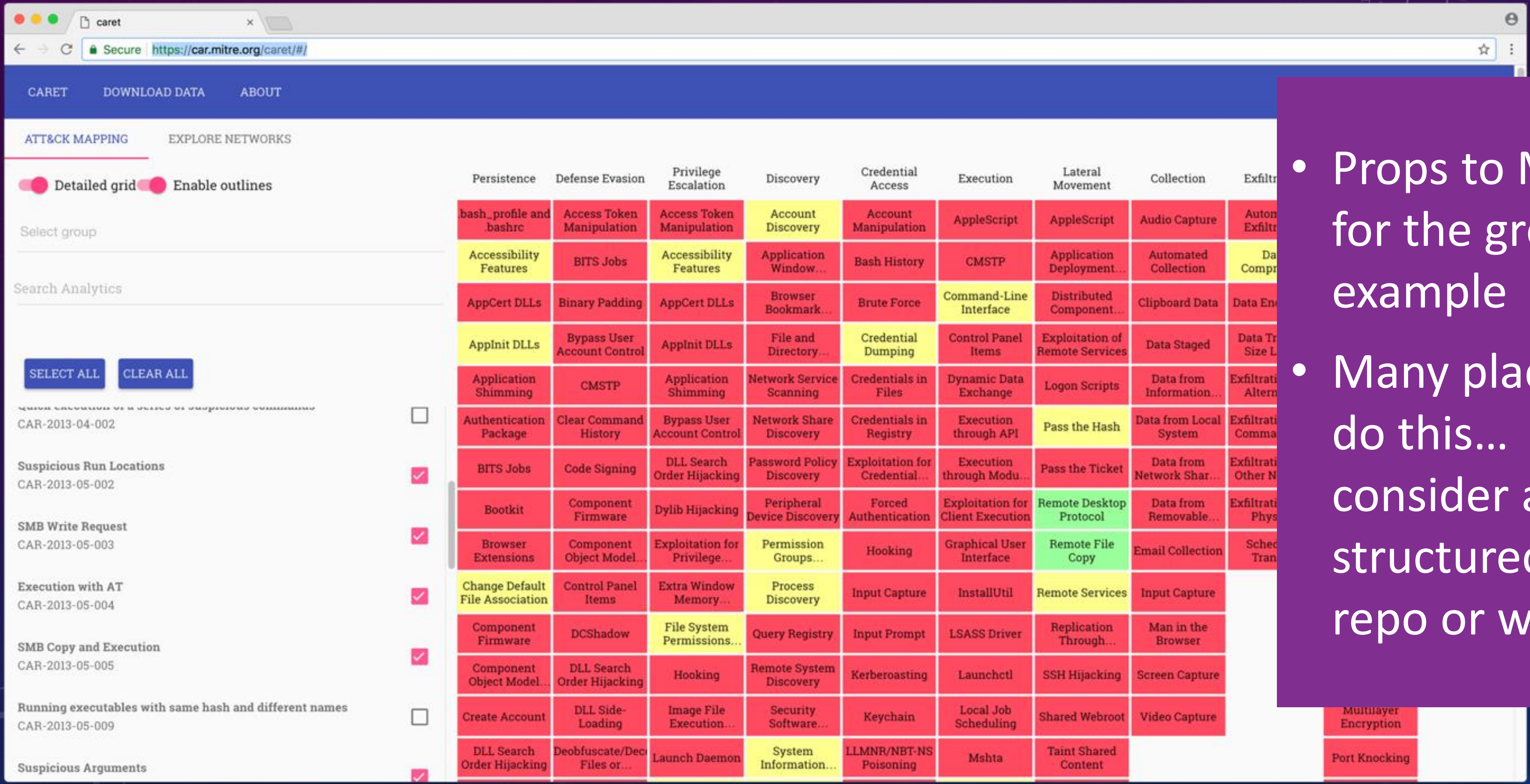
- # of times a detection or analytic fired, *attributed to the detection author*
- Is this evil?
- How can this be gamed?



HOW ARE YOU SUPPORTING YOUR CUSTOMERS?



MAP YOUR ANALYTICS TO ATT&CK



Quick execution of a series of suspicious commands
CAR-2013-04-002 □

Suspicious Run Locations
CAR-2013-05-002 ☑

SMB Write Request
CAR-2013-05-003 ☑

Execution with AT
CAR-2013-05-004 ☑

SMB Copy and Execution
CAR-2013-05-005 ☑

Running executables with same hash and different names
CAR-2013-05-009 □

Suspicious Arguments ☑

Persistence	Defense Evasion	Privilege Escalation	Discovery	Credential Access	Execution	Lateral Movement	Collection	Exfiltr	
bash_profile and .bashrc	Access Token Manipulation	Access Token Manipulation	Account Discovery	Account Manipulation	AppleScript	AppleScript	Audio Capture	Autom Exfiltr	
Accessibility Features	BITS Jobs	Accessibility Features	Application Window...	Bash History	CMSTP	Application Deployment...	Automated Collection	Da Compr	
AppCert DLLs	Binary Padding	AppCert DLLs	Browser Bookmark...	Brute Force	Command-Line Interface	Distributed Component...	Clipboard Data	Data En	
AppInit DLLs	Bypass User Account Control	AppInit DLLs	File and Directory...	Credential Dumping	Control Panel Items	Exploitation of Remote Services	Data Staged	Data Tr Size L	
Application Shimming	CMSTP	Application Shimming	Network Service Scanning	Credentials in Files	Dynamic Data Exchange	Logon Scripts	Data from Information...	Exfiltrati Altern	
Authentication Package	Clear Command History	Bypass User Account Control	Network Share Discovery	Credentials in Registry	Execution through API	Pass the Hash	Data from Local System	Exfiltrati Comma	
BITS Jobs	Code Signing	DLL Search Order Hijacking	Password Policy Discovery	Exploitation for Credential...	Execution through Modu...	Pass the Ticket	Data from Network Shar...	Exfiltrati Other N	
Bootkit	Component Firmware	Dylib Hijacking	Peripheral Device Discovery	Forced Authentication	Exploitation for Client Execution	Remote Desktop Protocol	Data from Removable...	Exfiltrati Phys	
Browser Extensions	Component Object Model...	Exploitation for Privilege...	Permission Groups...	Hooking	Graphical User Interface	Remote File Copy	Email Collection	Sche Tran	
Change Default File Association	Control Panel Items	Extra Window Memory...	Process Discovery	Input Capture	InstallUtil	Remote Services	Input Capture		
Component Firmware	DCShadow	File System Permissions...	Query Registry	Input Prompt	LSASS Driver	Replication Through...	Man in the Browser		
Component Object Model...	DLL Search Order Hijacking	Hooking	Remote System Discovery	Kerberoasting	Launchctl	SSH Hijacking	Screen Capture		
Create Account	DLL Side-Loading	Image File Execution...	Security Software...	Keychain	Local Job Scheduling	Shared Webroot	Video Capture		Multilayer Encryption
DLL Search Order Hijacking	Deobfuscate/Dec Files or...	Launch Daemon	System Information...	LLMNR/NBT-NS Poisoning	Mshta	Taint Shared Content			Port Knocking

- Props to MITRE for the great example
- Many places to do this… consider any structured code repo or wiki

https://car.mitre.org

METRIC FOCUS 5: ANALYST PERFORMANCE

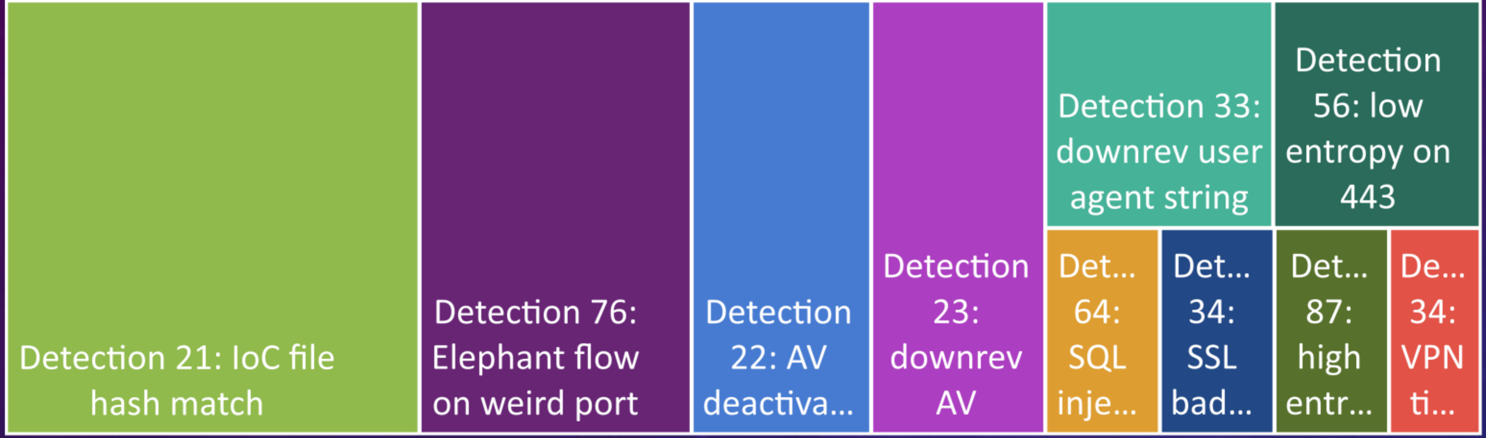
1. Name
2. Join date
3. Current role & time in role
4. Number of alerts triaged in last 30 days
5. % true positive rate for escalations
6. % response rate for customer escalations
7. Number of escalated cases handled in last 30 days
8. Mean time to close a case
9. Number of analytics/detections created that are currently in production
10. Number of detections modified that are currently in production
11. Total lines committed to SOC code repo in last 90 days
12. Success/fail rate of queries executed in last 30 days
13. Median run time per query
14. Mean lexical/structural similarity in queries run

Analyst Baseball Card

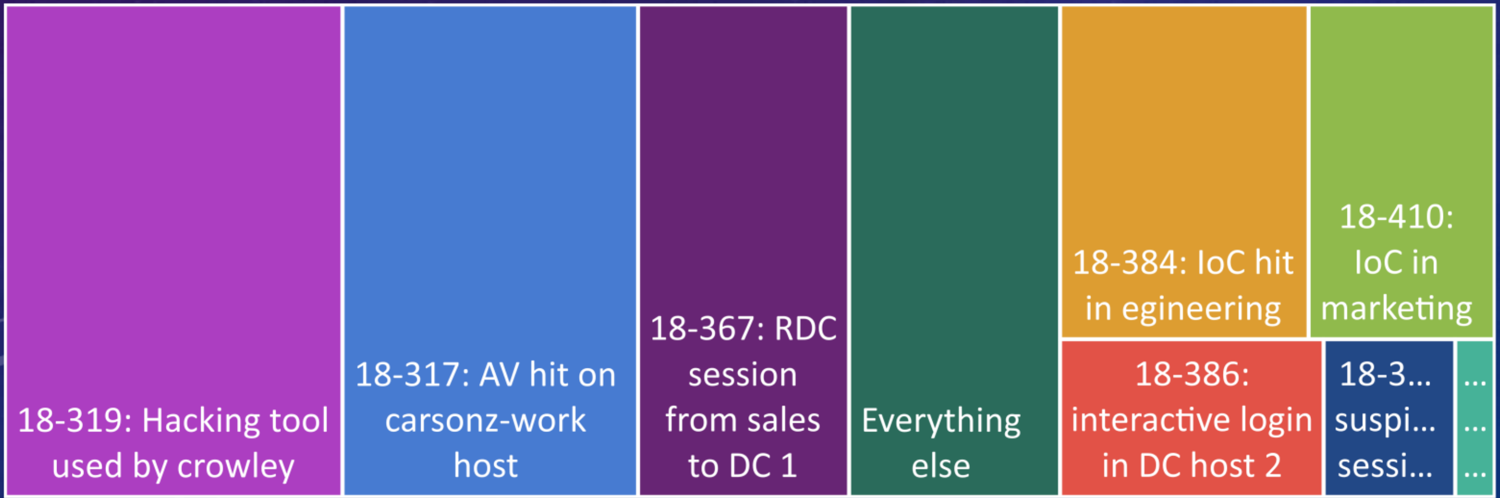
Christopher Crowley	Name
Chris	Preferred first name
TwoGuns	Callsign
2015-11-17	Join Date
NSM Analyst - Senior	Current Role
1 year, 1 month	Time in Role
38	Alerts Triaged in last 30 days
91.40%	Percent True Positive Rate
82.70%	Response rate percent for customer escalation
19	Escalated cases handled in last 30 days
1:34	Mean time to close case
7	Number analytics created currently in production
28	Number detection modified currently in production
423	Total lines committed to SOC code repository in last 90 days
91.40%	Success rate of queries against SIEM in last 30 days
0:09	Median run time per query
0.23	Mean lexical structure similarity in queries run in last 30 days

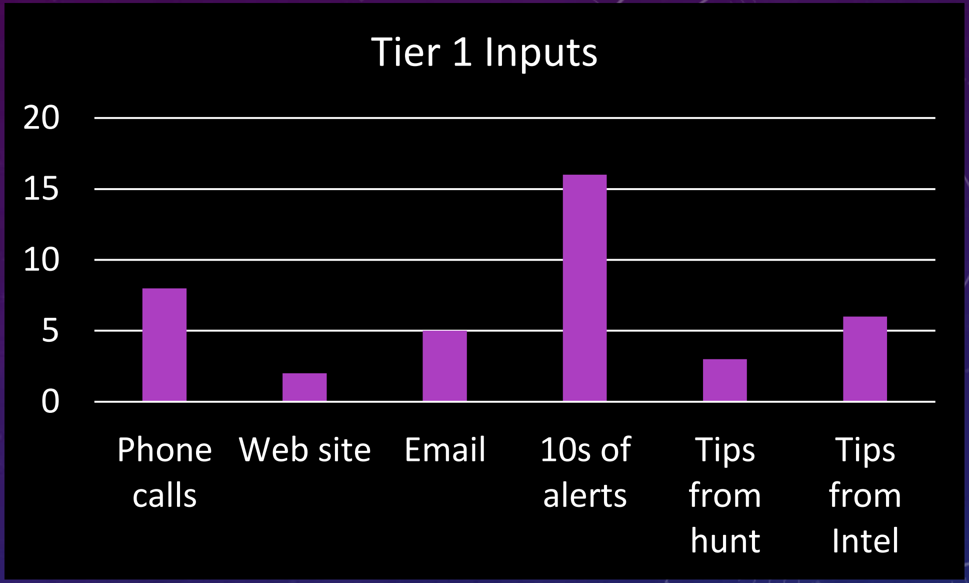
DAILY REVIEW DASHBOARD

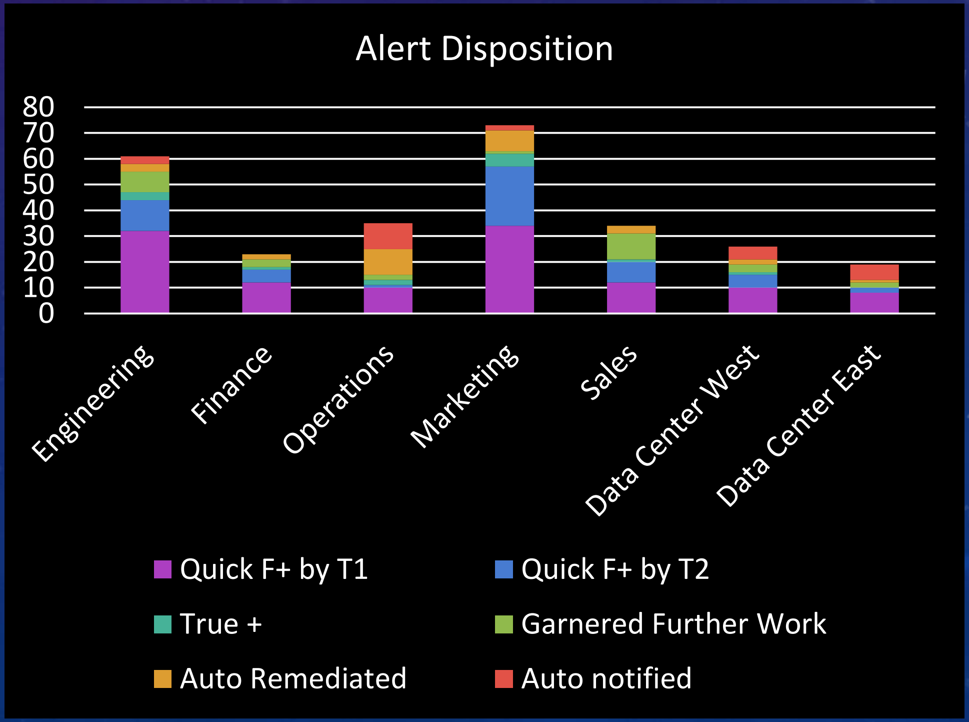
Top firing detections



Top time spent per case







METRIC FOCUS 6: INCIDENT HANDLING

- Mean/median adversary dwell time
- Mean and median time to…
  - Triage & Escalate
  - Identify
  - Contain
  - Eradicate & recover
- Divergence from SLA/SLO?
- Insufficient eradication?
- Threat attributed?

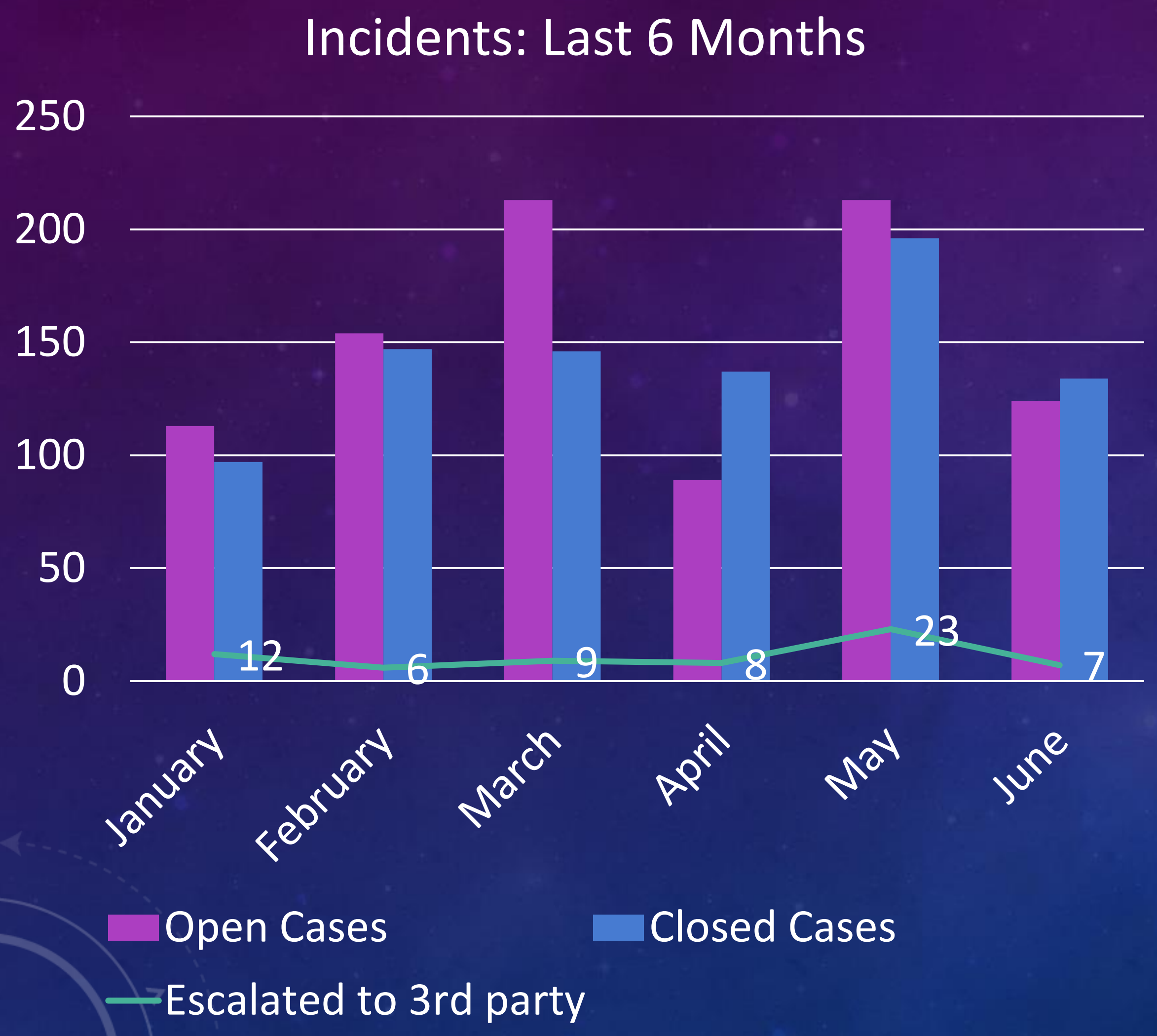
Top sources of confirmed incidents

- Proactive? Reactive?
- User reports? SOC monitoring?

Data & ”anecdata”: unforced errors and impediments

- Time waiting on other teams to do things
- No data/bad data/ data lost
- Incorrect/ambiguous conclusions
- Time spent arguing with other parties

TYPICAL INCIDENT METRICS



More ideas:

- Mean/median time to respond
- Cases left open > time threshold
- Cases left open by initial reporting/detection type
- Stacked bar chart by case type

INCIDENT IMPACT

Low

- Few systems (or only a specific type)
- Unimportant systems
- Unimportant data

Moderate

- More systems (or many common types)
- Important or high value person's, account, or system
- Important data at risk

High

- Most systems (or almost all types)
- Highest level accounts, users, and systems
- Business critical data

INCIDENT IMPACT CATEGORY

Functional

- Low – minimal function disruption
- Moderate – substantial disruption
- High – complete disruption

Informational

- Intellectual Property (L/M/H)
- Integrity Manipulation (L/M/H)
- Privacy violated (such as PII / PHI)

Recoverable

- Regular – predictable using resources on hand
- Supplemented – predictable with augmented resources
- Unrecoverable – data breach which cannot be undone

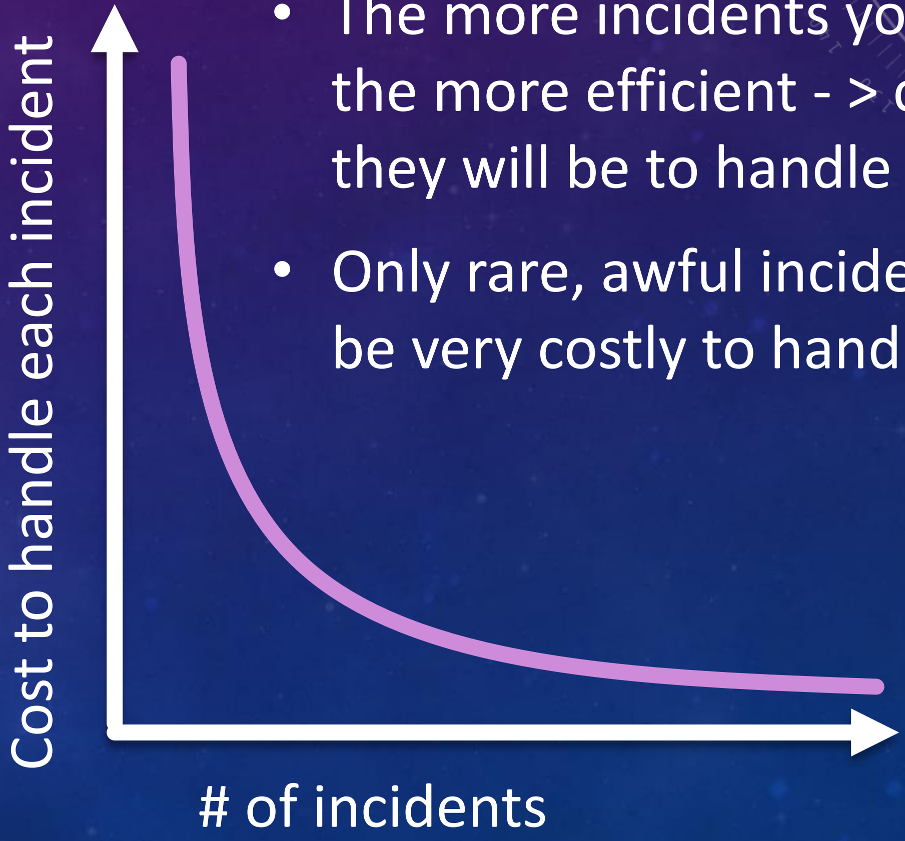
See more here: https://www.us-cert.gov/incident-notification-guidelines#impact-category-descriptions

INCIDENT AVOIDABILITY

- The vast majority of incidents are avoidable… everyone realizes this
  - Collect metrics on *how* avoidable, what could have been done to prevent
- Crowley's Incident Avoidability metric
  1. A measure, already available in the environment, is applied to other systems/networks, but wasn't applied -> resulting in the incident
  2. A measure is available (generally) and something (economic, political) prevents implementing it within the organization
  3. Nothing is available to prevent that method of attack
- Attribution for measure/mechanism in 1 & 2 is critical

METRIC FOCUS 7: INCIDENT FINANCIALS: COST

- $ for handling, $ for actual loss
- Routine handling
  - All alerts & reports fielded
  - Per escalated event to tier 2
  - True positives
- Consider:
  - Cost of people
  - Technology
  - Proportion of time spent



- The more incidents you handle, the more efficient - > cheaper they will be to handle
- Only rare, awful incidents should be very costly to handle

INCIDENT FINANCIALS: VALUE

- Start with standard impact value assigned to each incident
- $ saved/loss prevented
  - Routine incidents: standard calculation
  - Escalated & customized handling: often speculate
- What to do?
  - Past incidents
  - Reporting from other orgs, news
  - Iterate with execs

Example implied value: loss prevention

- Incidents that were escalated to legal counsel, law enforcement
- Incidents handled that clobbered competitors
- Direct value of IP caught in exfil
- Value of systems not being bricked from EFI bootkit

METRIC FOCUS 8: TOP RISK AREAS & HYGIENE

- Make vulnerability management data available to customers
  - Self service model
  - Scan results down to asset & item scanned
- But don’t beat them over the head with every measure!
  - Pick classic ones they will always be measured on
  - Scanning, monitoring, patching

- Pick top risk items from own incident avoidability metrics and public intel reporting to focus on each year, semester, or quarter
  - Internet-exposed devices
  - Code signing enforcement
  - EDR deployment
  - Single factor auth
  - Non-managed devices & cloud resources

CONCLUSION

SUMMARY: INTERNAL METRICS

- Analyst baseball card
  - Raw output / productivity
  - Technical & operational quality
  - Pedigree, training, growth
  - Kudos, ”saves”
- Data feed health
  - Up/down
  - Latency
- Daily alert volume & FP rate
- Weekly intel & IOC processing volume
- Weekly forensics/malware volume
- Analytic coverage
  - Kill chain & ATT&CK cell
  - Dependencies: source, detection framework
  - Written by whom
  - Volume & success rates
  - Customer coverage

SUMMARY: EXTERNAL METRICS

Key themes: **Cost – Value – Risk**

Always be ready to answer: "what have you done for me lately?"

- Managed vs unmanaged assets
- Monitoring & scanning coverage
- Top risk areas & hygiene
  - Top issues that are leading to incidents
- Custom detections & value add

- Incidents handled
  - Cost incurred & avoided
  - Causes & impediments
- Mean/median dwell time
- Mean/median time to identify, contain, eradicate, recover
- Mean/median time to respond to a data call, such as an IOC sweep

SUMMARY: SLAS / SLOS

Key themes:

For written agreements, select only the SLAs necessary to suit mission objectives

Examples:

- Response initiation within 4 hours
- Reporting / Notification frequency at minimum daily regarding any active incident rated at moderate severity
- If less that 90%, 5% "Managed Systems" percentage increase quarterly (improvement in asset tracking and identification as well as business coordination), above 90%, 1% increase quarterly
- Increased performance on repeated incidents of the same nature on the same systems (demonstrated improvement in proficiency)

CLOSING

- Whatever you do, measure something
- You can do it, regardless of how mature, old, or big your SOC is
- Pick your investments carefully
- Iterate constantly

http://memeshappen.com/meme/custom/you-can-do-it-18134

QUESTIONS

"THERE ARE LIES, DAMN LIES, AND STATISTICS." -- UNKNOWN