NIDDEL

Determining the Fit and Impact of CTI Indicators on Your Monitoring Pipeline (#tiqtest2)

Alex Pinto - Chief Data Scientist – Niddel
(now a part of Verizon)
@alexcpsec
@NiddelCorp

Who am I?

- Brazilian Immigrant
- Security Data Scientist
- Capybara Enthusiast
- Co-Founder at Niddel (@NiddelCorp)
- Founder of MLSec Project (@MLSecProject)
- What is **Niddel**? – Niddel is a security vendor that provides a SaaS-based Autonomous Threat Hunting System
- We are now a part of Verizon, but this is not what this talk is about, so hit me up later!

This Talk Contains

- 1 Fair Warning
- 1 Witty Metaphor
- 3 Novel(-ish) Ideas
- 2 Hopeful Dreams
- 1 Enlightening Conclusion
- Several Self-Serving Callbacks

- At least 1 Capybara

Nutrition Facts

Serving Size 125g

Amount Per Serving

Calories 65 Calories from Fat 2

	% Daily Value*
Total Fat 0g	0%
Saturated Fat 0g	0%
Trans Fat	
Cholesterol 0mg	0%
Sodium 1mg	0%
Total Carbohydrate 17g	6%
Dietary Fiber 3g	12%
Sugars 13g	
Protein 0g	

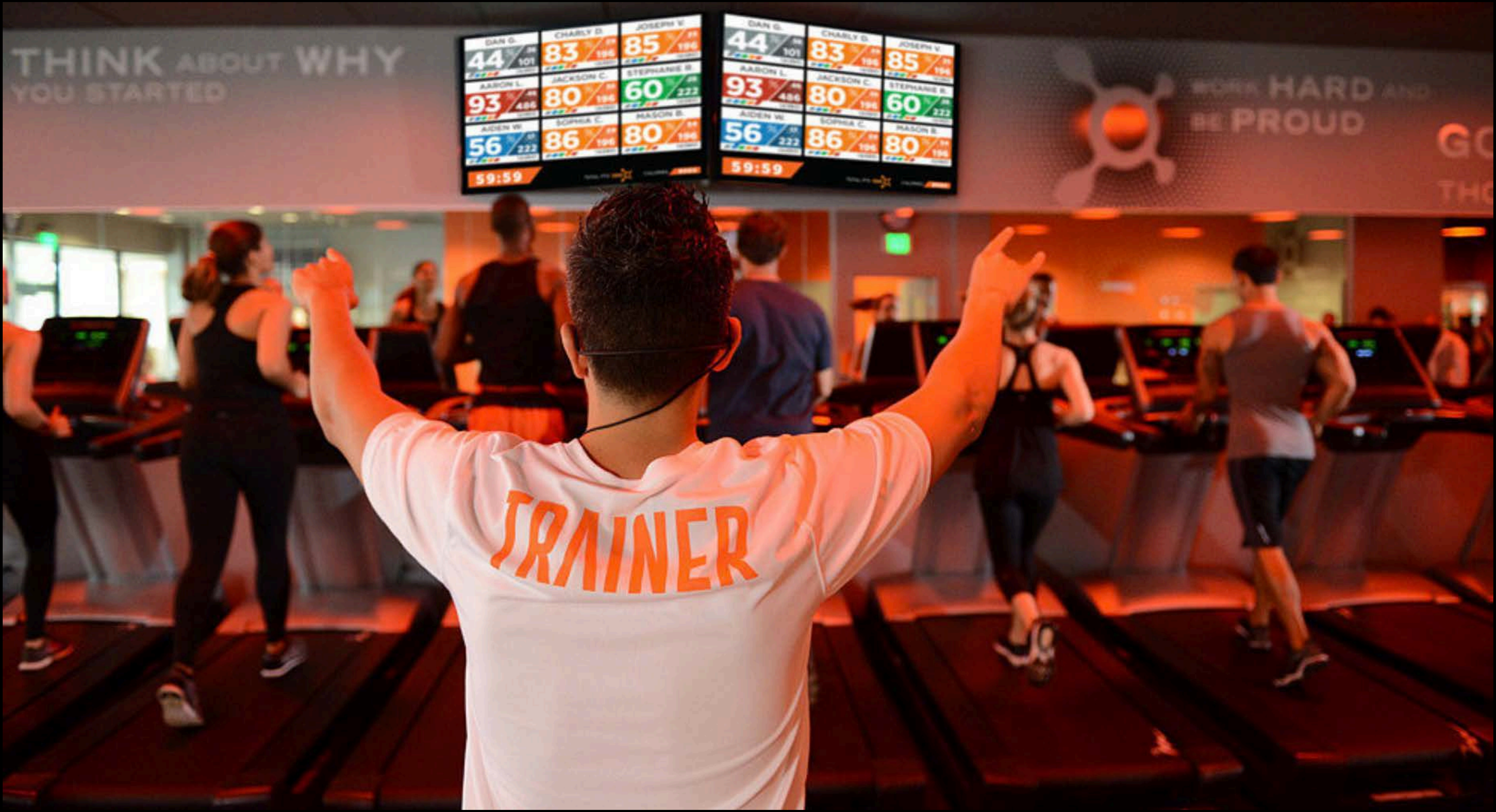
Vitamin A	1%	Vitamin C	10%
Calcium	1%	Iron	1%

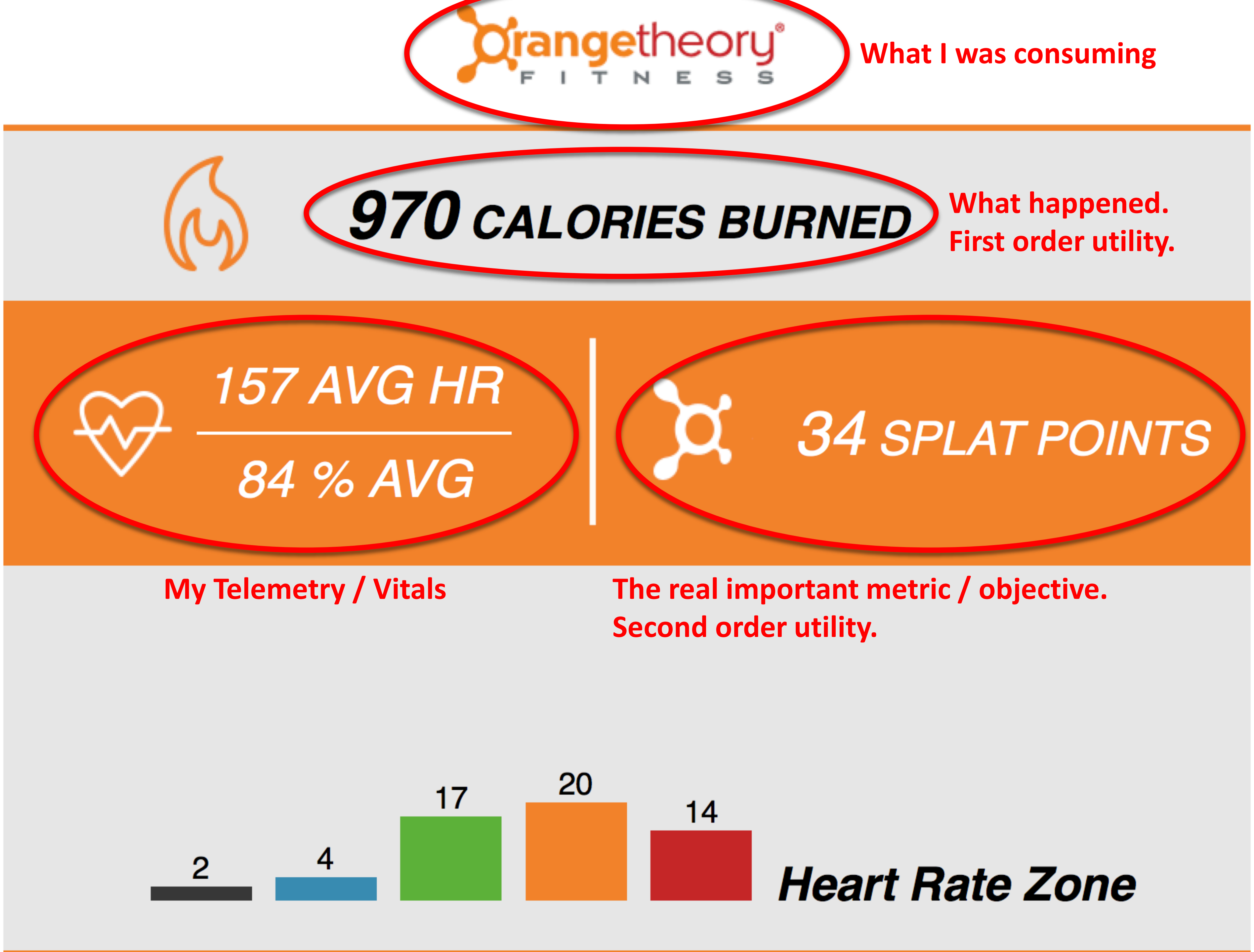
*Percent Daily Values are based on a 2,000 calorie diet. Your daily values may be higher or lower depending on your calorie needs.

Fair Warning

- This is a presentation about Metrics
  - Please hold your applause
  - Data Scientists like data at scale (duh)
  - Only by measuring the impact we can have, we will be able to have effective “supply chain management” and “industrialization” of threat intel
  - Data QA and analysis is 95% of any ML effort

Metrics on What?


THINK ABOUT WHY
YOU STARTED
DAN G. 44 101
CHARLY D. 83 196
JOSEPH V. 85 196
AARON L. 93 486
JACKSON C. 80 196
STEPHANIE R. 60 222
AIDEN W. 56 222
SOPHIA C. 86 196
MASON B. 80 196
59:59
DAN G. 44 101
CHARLY D. 83 196
JOSEPH V. 85 196
AARON L. 93 486
JACKSON C. 80 196
STEPHANIE R. 60 222
AIDEN W. 56 222
SOPHIA C. 86 196
MASON B. 80 196
59:59
WORK HARD AND
BE PROUD
GO
THO
TRAINER



orangetheory FITNESS

What I was consuming

970 CALORIES BURNED

What happened.
First order utility.

157 AVG HR

84 % AVG

34 SPLAT POINTS

My Telemetry / Vitals

The real important metric / objective.
Second order utility.

Taking Diminishing Returns into Account

TIQ-TEST Classic™

Mostly inside baseball.

- NOVELTY – How often do feeds update themselves?
- AGING – How long does an indicator sit on a feed?

- OVERLAP – How do they compare to what you got?
- UNIQUENESS – How many indicators are found in only one feed?

Insights on what we are consuming.

- POPULATION – How does this population distribution compare to another one ?

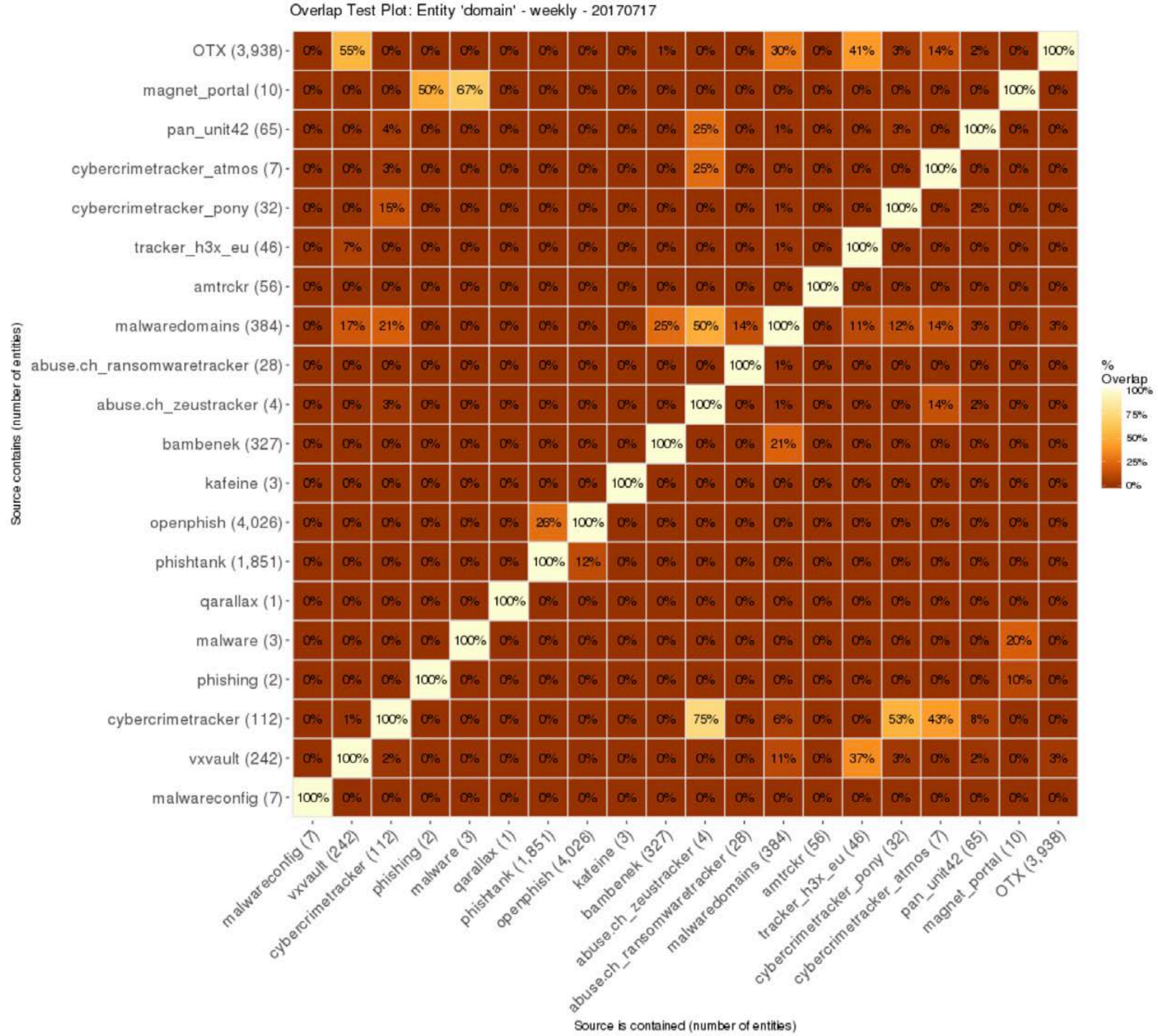
Insights on first order utility, how the data affects us.

Coverage Test

Coverage Test (aka Overlap 2.0)

- Our interpretation of Coverage:
  - Are you getting the data you need from the myriad feeds you consume?
  - How much unique data does the feed contain?
  - What actual DETECTION and CONTEXT opportunities arise from the data you have available?

Overlap Classic™ is still too much inside baseball


Overlap Test Plot: Entity 'domain' - weekly - 20170717

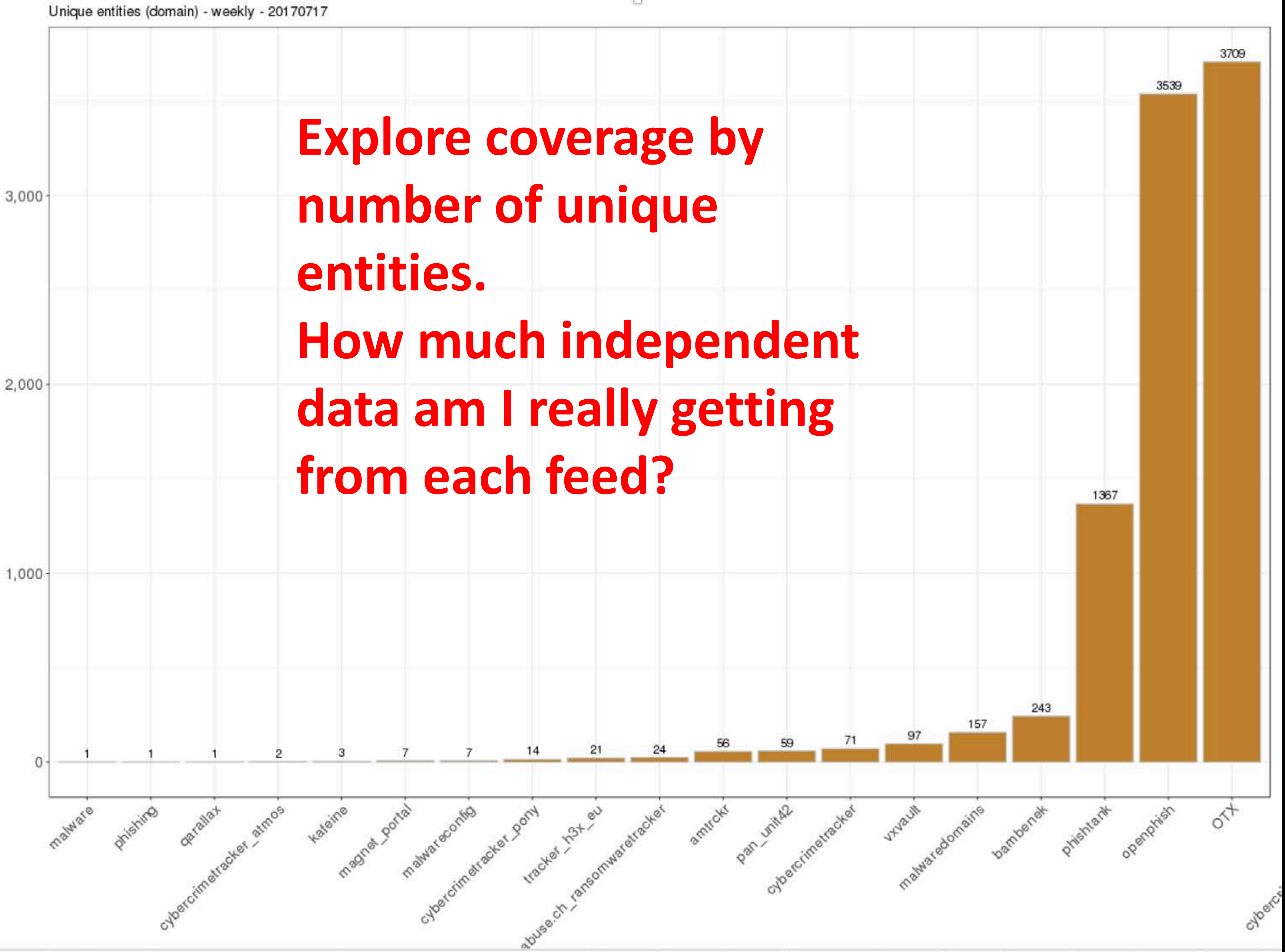
	malwareconfig (7)	vxvault (242)	cybercrimetracker (112)	phishing (2)	malware (3)	qarallax (1)	phishtank (1,851)	openphish (4,026)	kafeine (3)	bambenek (327)	abuse.ch_zeustracker (4)	abuse.ch_ransomwaretracker (28)	malwaredomains (384)	amtrckr (56)	tracker_h3x_eu (46)	cybercrimetracker_pony (32)	cybercrimetracker_atmos (7)	pan_unit42 (65)	magnet_portal (10)	OTX (3,938)
OTX (3,938)	0%	55%	0%	0%	0%	0%	0%	0%	0%	1%	0%	0%	30%	0%	41%	3%	14%	2%	0%	100%
magnet_portal (10)	0%	0%	0%	50%	67%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	100%	0%
pan_unit42 (65)	0%	0%	4%	0%	0%	0%	0%	0%	0%	0%	25%	0%	1%	0%	0%	3%	0%	100%	0%	0%
cybercrimetracker_atmos (7)	0%	0%	3%	0%	0%	0%	0%	0%	0%	0%	25%	0%	0%	0%	0%	0%	100%	0%	0%	0%
cybercrimetracker_pony (32)	0%	0%	15%	0%	0%	0%	0%	0%	0%	0%	0%	0%	1%	0%	0%	100%	0%	2%	0%	0%
tracker_h3x_eu (46)	0%	7%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	1%	0%	100%	0%	0%	0%	0%	0%
amtrckr (56)	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	100%	0%	0%	0%	0%	0%	0%
malwaredomains (384)	0%	17%	21%	0%	0%	0%	0%	0%	0%	25%	50%	14%	100%	0%	11%	12%	14%	3%	0%	3%
abuse.ch_ransomwaretracker (28)	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	100%	1%	0%	0%	0%	0%	0%	0%	0%
abuse.ch_zeustracker (4)	0%	0%	3%	0%	0%	0%	0%	0%	0%	0%	100%	0%	1%	0%	0%	0%	14%	2%	0%	0%
bambenek (327)	0%	0%	0%	0%	0%	0%	0%	0%	0%	100%	0%	0%	21%	0%	0%	0%	0%	0%	0%	0%
kafeine (3)	0%	0%	0%	0%	0%	0%	0%	0%	100%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
openphish (4,026)	0%	0%	0%	0%	0%	0%	26%	100%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
phishtank (1,851)	0%	0%	0%	0%	0%	0%	100%	12%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
qarallax (1)	0%	0%	0%	0%	0%	100%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
malware (3)	0%	0%	0%	0%	100%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	20%	0%
phishing (2)	0%	0%	0%	100%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	10%	0%
cybercrimetracker (112)	0%	1%	100%	0%	0%	0%	0%	0%	0%	0%	75%	0%	6%	0%	0%	53%	43%	8%	0%	0%
vxvault (242)	0%	100%	2%	0%	0%	0%	0%	0%	0%	0%	0%	0%	11%	0%	37%	3%	0%	2%	0%	3%
malwareconfig (7)	100%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%

Coverage Test

- For each feed you have available:

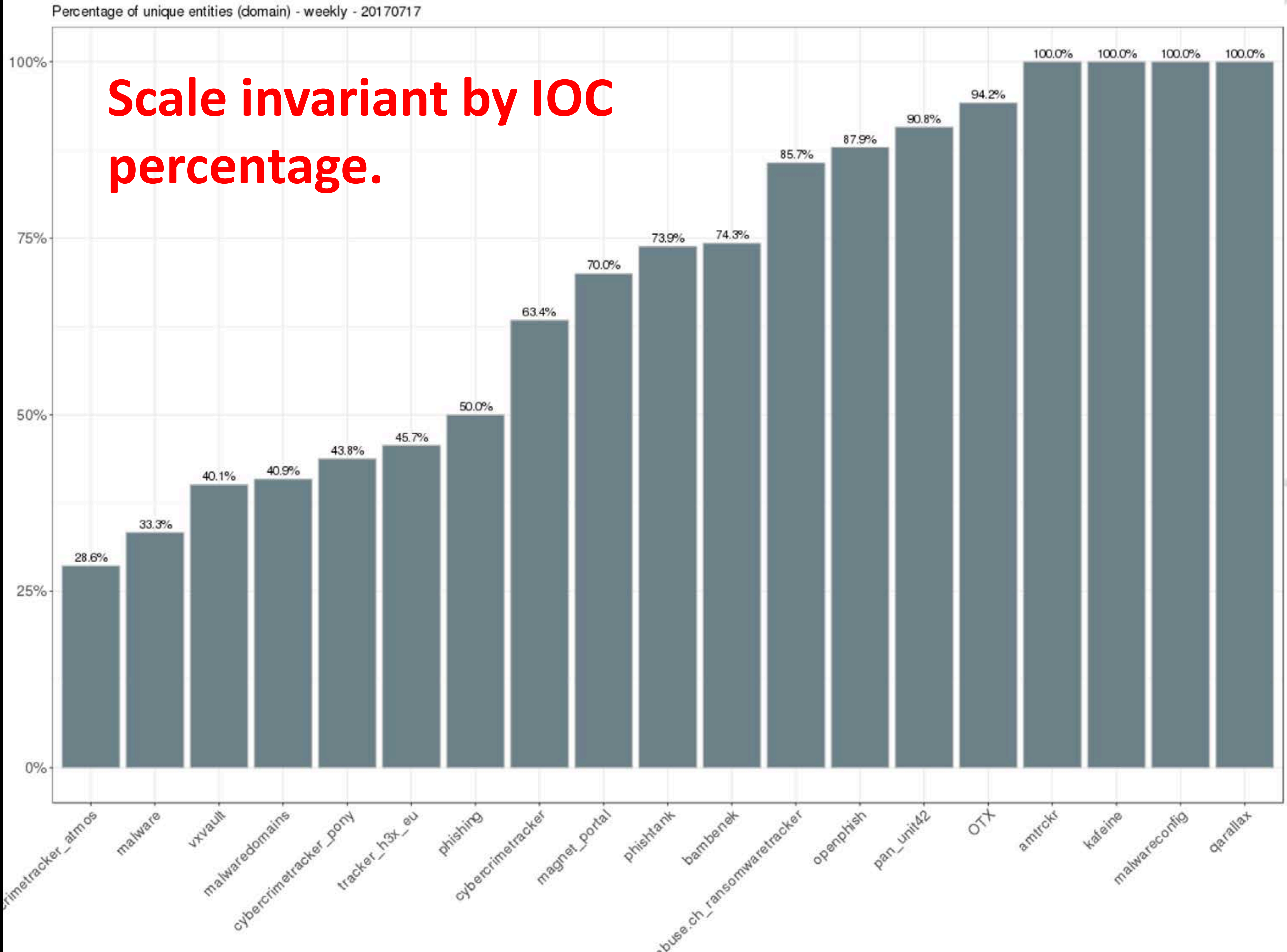
$$Coverage_{Feed} = setdiff(IOC_{Feed}, IOC_{ALL})$$

$$Coverage\%_{Feed} = \frac{setdiff(IOC_{Feed}, IOC_{ALL})}{IOC_{Feed}}$$



Explore coverage by number of unique entities.

How much independent data am I really getting from each feed?



Scale invariant by IOC percentage.

Coverage Test - Caveats

- Too much uniqueness could mean a lot of FPs!
- Having overlap is NOT BAD
  - Confidence + different workflow mapping

- This is not related to “CTI Generation” coverage, as in source and methods utilization and actor tracking
  - Aaron Shelmire did some work on that
  - Ex: Dridex -> Locky -> GlobeImp -> Dridex from same actors

Fitness Test

Fitness Test (aka Population 2.0)

- The original Population test was too concerned in using fancy statistics to be useful.

- Trends and population comparisons ARE COOL, and a good way to drive detection engines, but a bad way to evaluate clearly if a feed has a relationship to your environment.

- Detection power of feeds only matter of they “fit” your telemetry

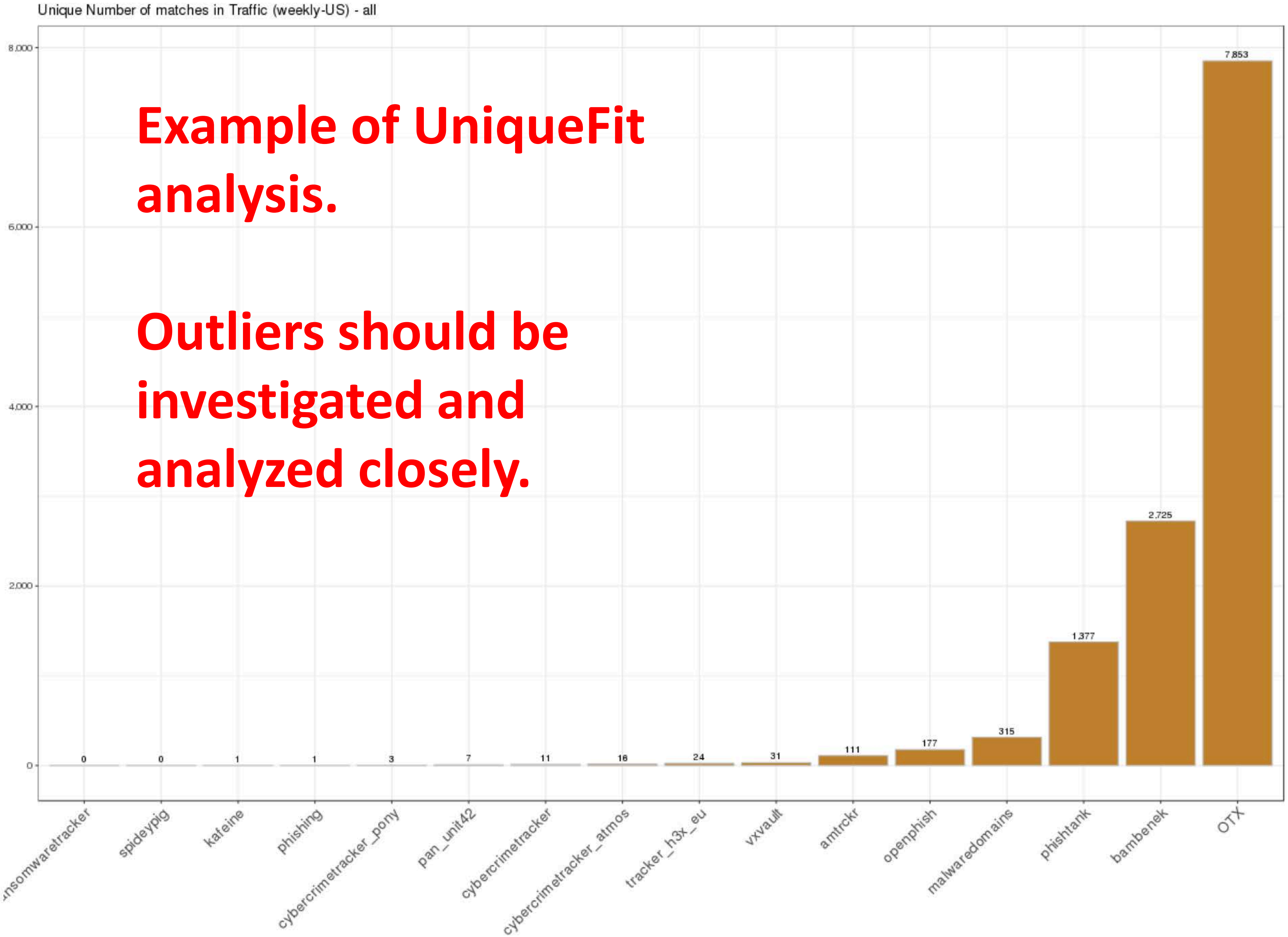
Fitness Test

- For each feed you have available:

$$Fit_{Feed} = intersect(\,IOC_{Feed}, Telemetry)$$

$$UniqueFit_{Feed} = intersect(setdiff(\,IOC_{Feed}, IOC_{ALL}), Telemetry)$$

Unique IOCs per feed from our Coverage test



Example of UniqueFit analysis.

Outliers should be investigated and analyzed closely.

Fitness Test - Caveats

- A bad Fit does NOT mean a bad Feed. Best ICS / OT feed data will probably “not fit” the telemetry of a small credit union.

- A Fitness value that is too high could also mean a high number of false positives, unless the feeds themselves are too different.

- Sharing communities: Fitness answers the “am I the only one?” question perfectly.

Impact Test

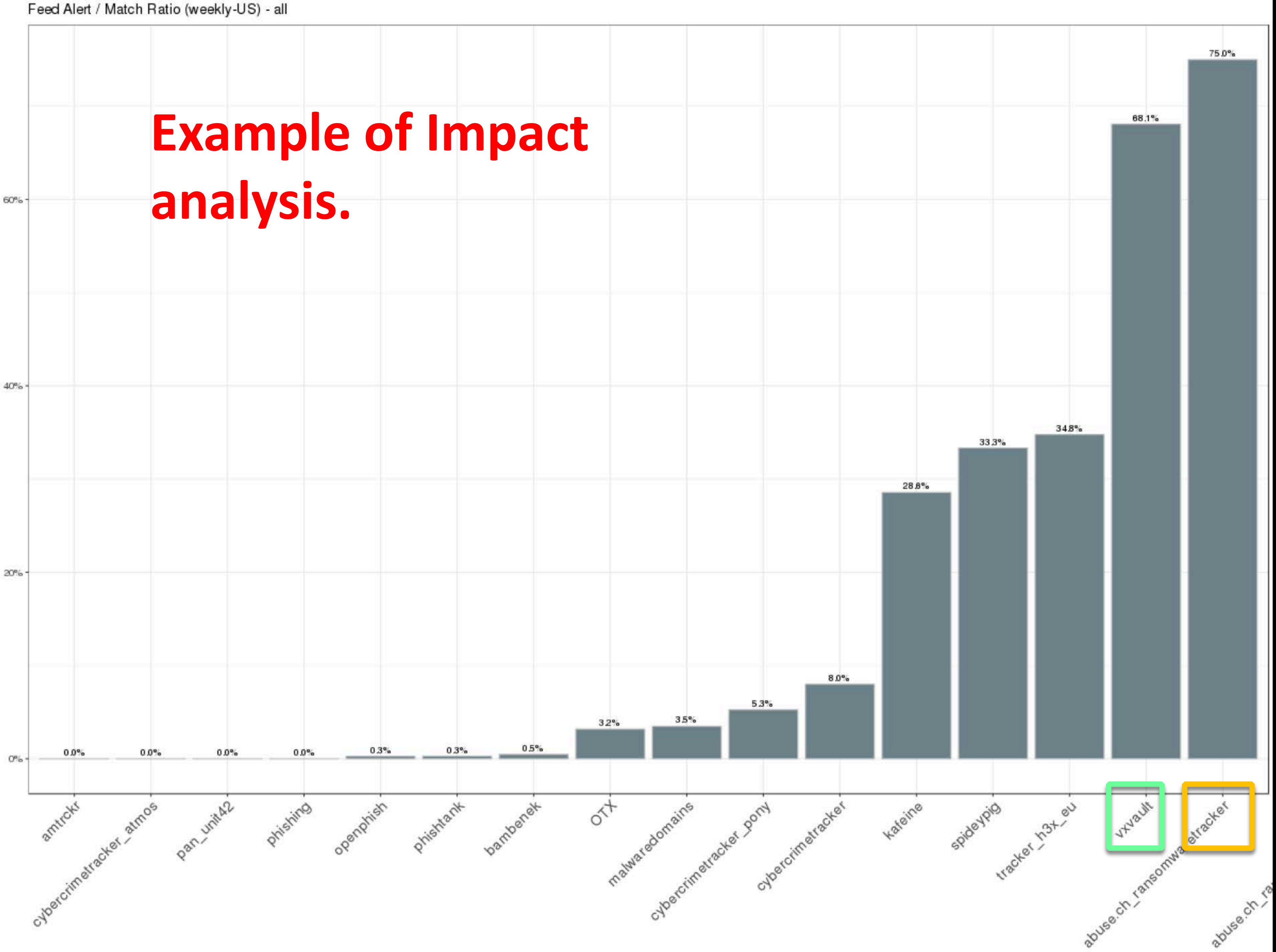
Impact Test (our Splat points)

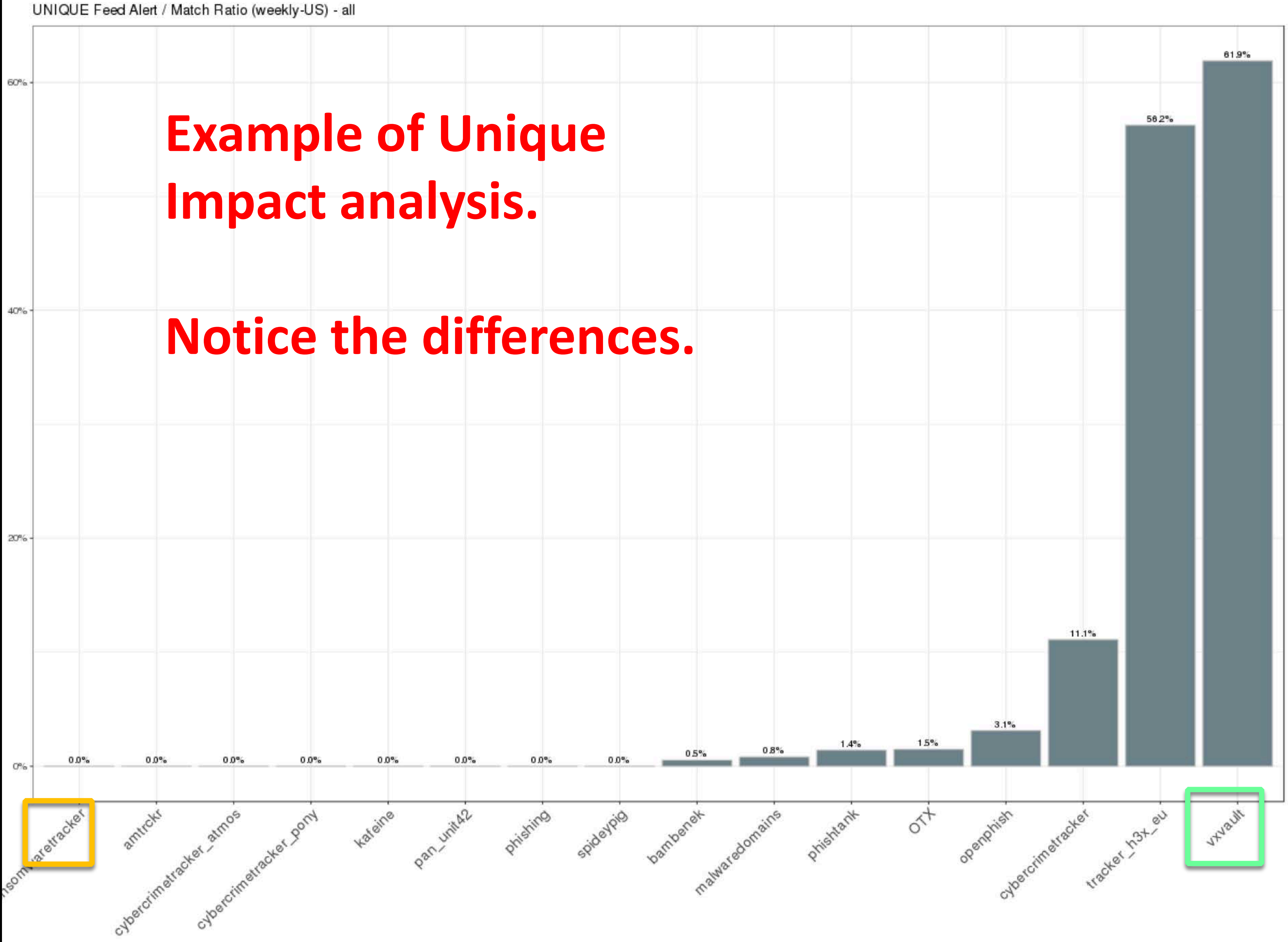
- “How much detection are we getting out of this?”

$$Impact_{Feed} = \frac{good\_alerts(intersect(\ IOC_{Feed}, Telemetry))}{intersect(\ IOC_{Feed}, Telemetry)}$$

- What is a “good alert”? What is a “false positive”?
- <u>Good alert</u>: An alert that was “correct” even if it had been alerted by something else is not a false positive.

Example of Impact analysis.



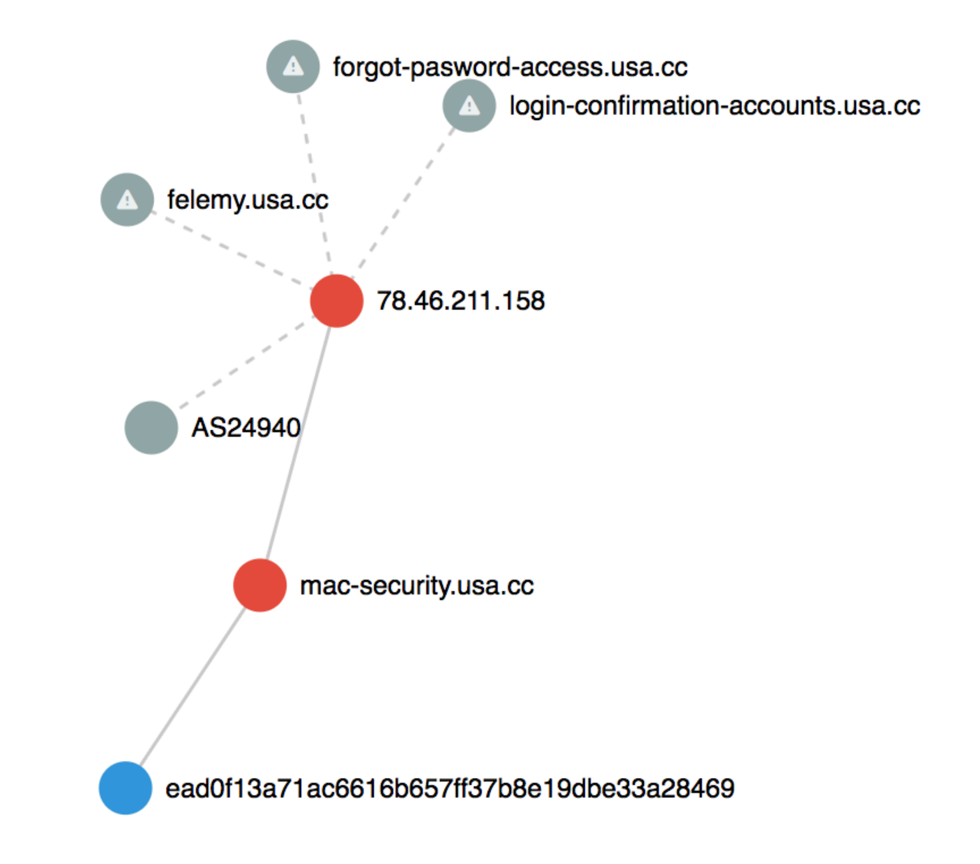


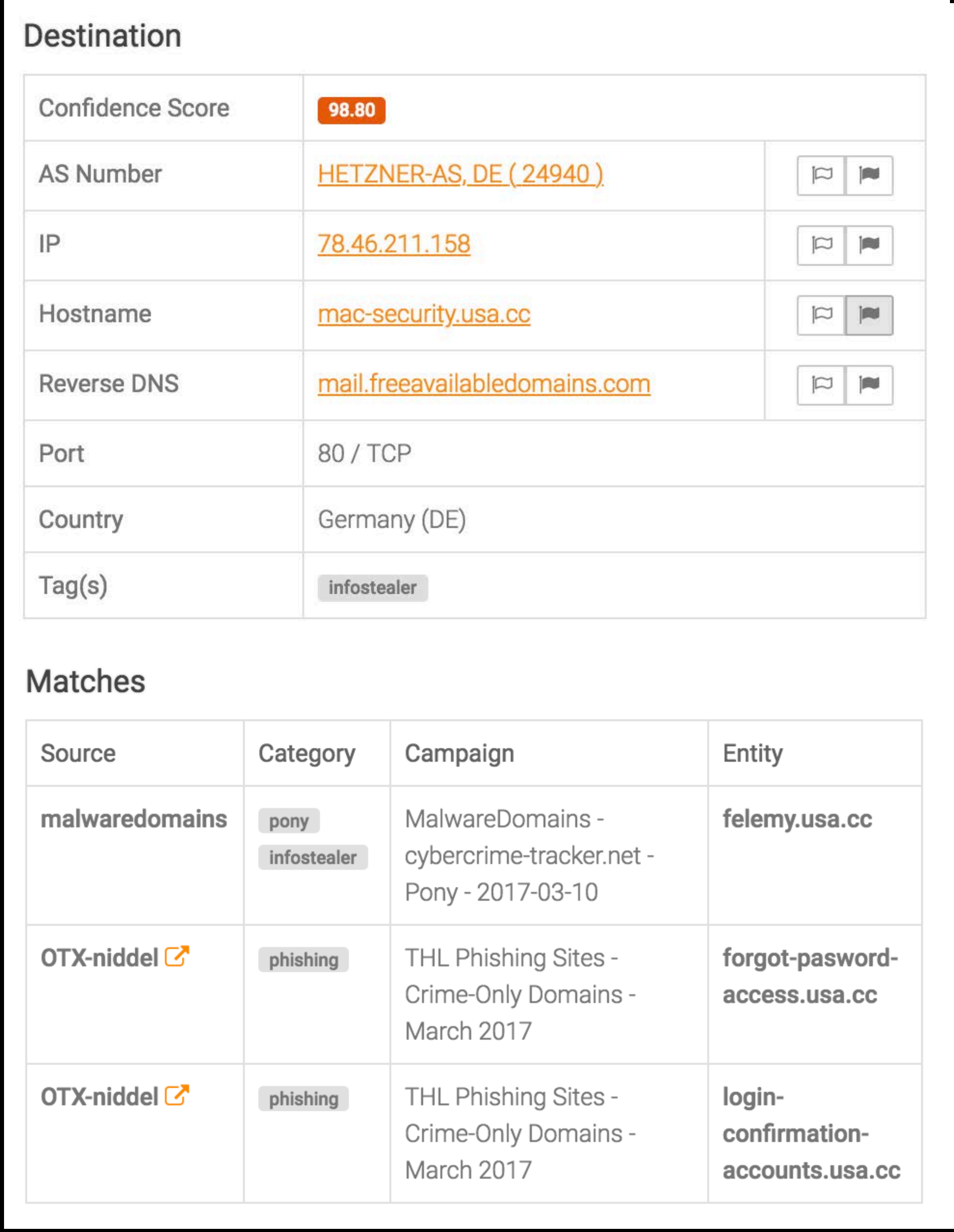
Example of Unique Impact analysis.

Notice the differences.

Deep Impact Test

- What if it's not a direct IOC match but we learned from it?
- Best usage from CTI is "climbing the pyramid", and learning TTPs
- Not so simple to account for correctly


forgot-pasword-access.usa.cc
login-confirmation-accounts.usa.cc
felemy.usa.cc
78.46.211.158
AS24940
mac-security.usa.cc
ead0f13a71ac6616b657ff37b8e19dbe33a28469

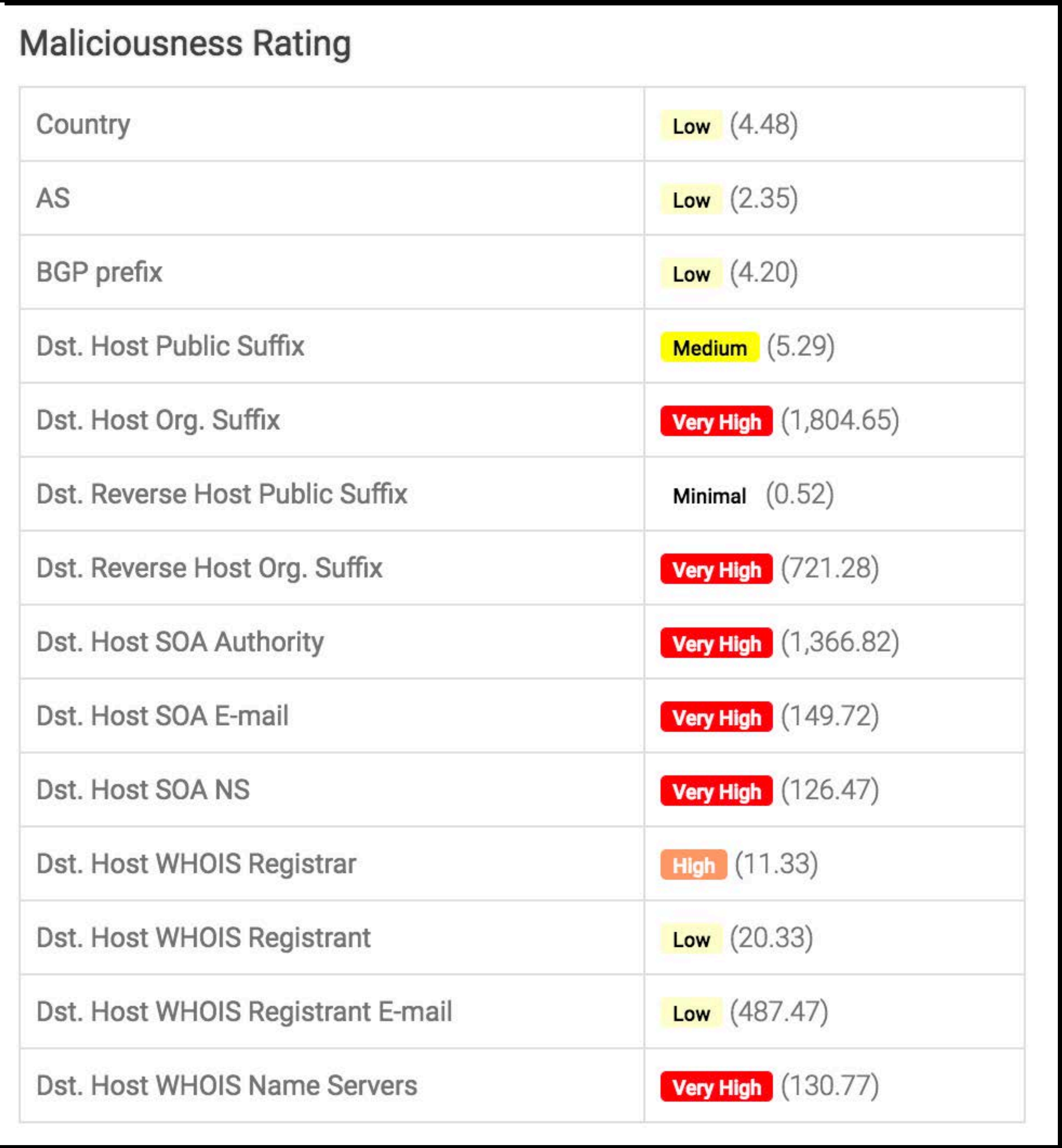


Destination

Confidence Score	98.80	
AS Number	HETZNER-AS, DE (24940)	⚐ ⚑
IP	78.46.211.158	⚐ ⚑
Hostname	mac-security.usa.cc	⚐ ⚑
Reverse DNS	mail.freeavailabledomains.com	⚐ ⚑
Port	80 / TCP	
Country	Germany (DE)	
Tag(s)	infostealer	

Matches

Source	Category	Campaign	Entity
malwaredomains	pony infostealer	MalwareDomains - cybercrime-tracker.net - Pony - 2017-03-10	**felemy.usa.cc**
OTX-niddel ↗	phishing	THL Phishing Sites - Crime-Only Domains - March 2017	**forgot-pasword-access.usa.cc**
OTX-niddel ↗	phishing	THL Phishing Sites - Crime-Only Domains - March 2017	**login-confirmation-accounts.usa.cc**



Maliciousness Rating

Country	Low (4.48)
AS	Low (2.35)
BGP prefix	Low (4.20)
Dst. Host Public Suffix	Medium (5.29)
Dst. Host Org. Suffix	Very High (1,804.65)
Dst. Reverse Host Public Suffix	Minimal (0.52)
Dst. Reverse Host Org. Suffix	Very High (721.28)
Dst. Host SOA Authority	Very High (1,366.82)
Dst. Host SOA E-mail	Very High (149.72)
Dst. Host SOA NS	Very High (126.47)
Dst. Host WHOIS Registrar	High (11.33)
Dst. Host WHOIS Registrant	Low (20.33)
Dst. Host WHOIS Registrant E-mail	Low (487.47)
Dst. Host WHOIS Name Servers	Very High (130.77)

TIQ-Test 3.0? 🔮🤖

Ideas from a Metric Filled Future

- BENEFIT – “By using this feed / combination of feeds correctly correctly, you are likely to have ~10 actionable alerts per week”

- ASSURANCE – “By using this feed / combination of feeds correctly, you will have the capability to detect threat actor / malware family X within an SLA of 24 hours”

In Summary…

You can't buy capyness.

In Summary

- To avoid diminishing returns from buying / ingesting new CTI feeds you must be continually working hard to make them work for you.

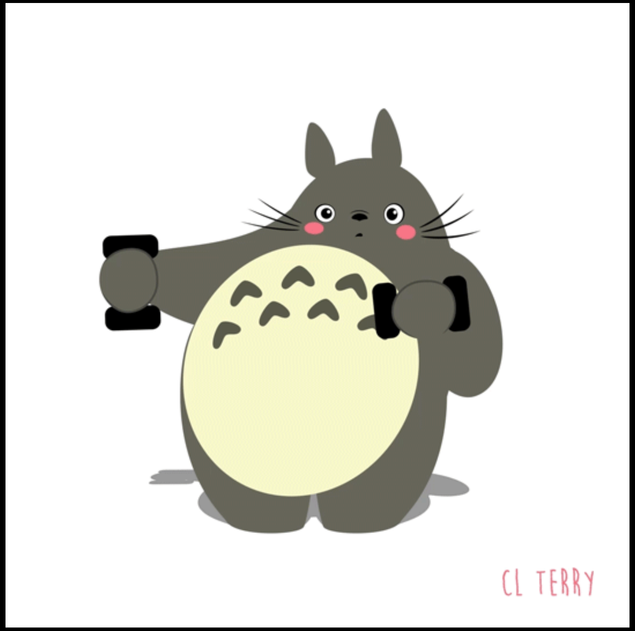
- Failing to understand the caveats of proper usage and selection of feeds as your org matures will lead you to a “detection plateau” where more feeds are not making you more secure.

Questions?

NIDDEL

Share, like, subscribe.
Q&A and Feedback please!

Alex Pinto – alexcp@niddel.com
@alexcpsec
@NiddelCorp



“If you can't measure it, you can't improve it.” - Peter Drucker