



SAMM Benchmark Top 10 Lessons Learned

SAMM User Day
September 25, 2024



SAMM Benchmark

Benchmark tab in Excel Toolbox and SAMMY tool

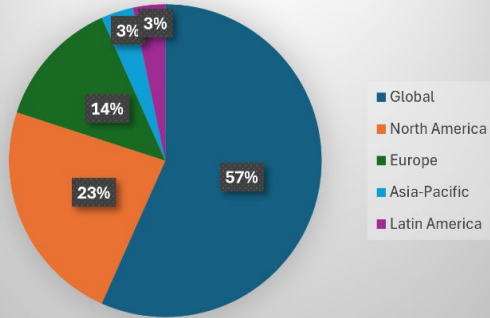


<https://bit.ly/sammbenchmarksubmission>

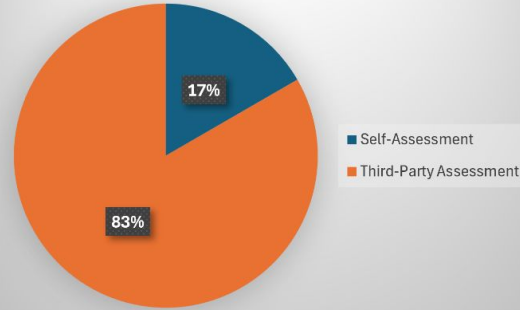


Demographics

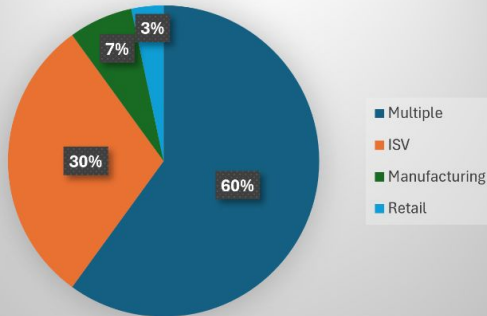
Geographic Region



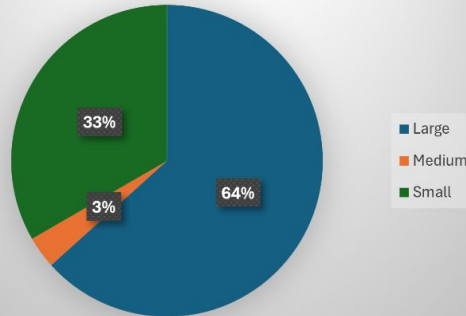
Assessment Type



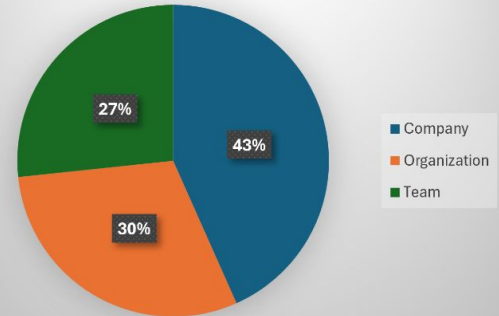
Industry



Company Size



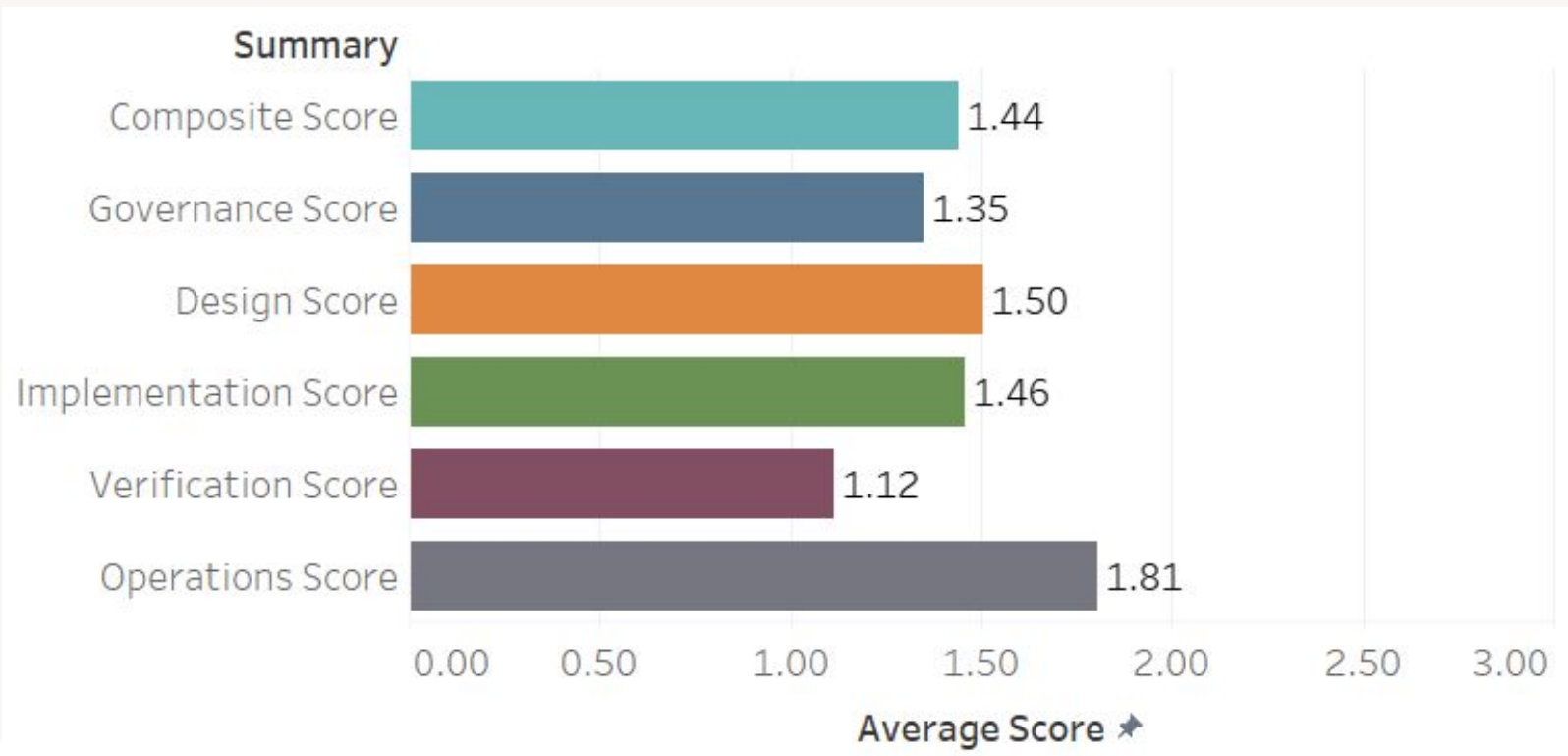
Scope of Assessment



Demographics Highlights

- 30 datasets
 - Too few to provide more granular results
- Most assessments are done by a reputable third party
 - Higher quality data
- The majority of the companies are large multinationals
 - Mid-sized companies are underrepresented
- Results averaging problems
 - Governance and Operations in multinationals vs small
- Assessment scope represents a nice mix

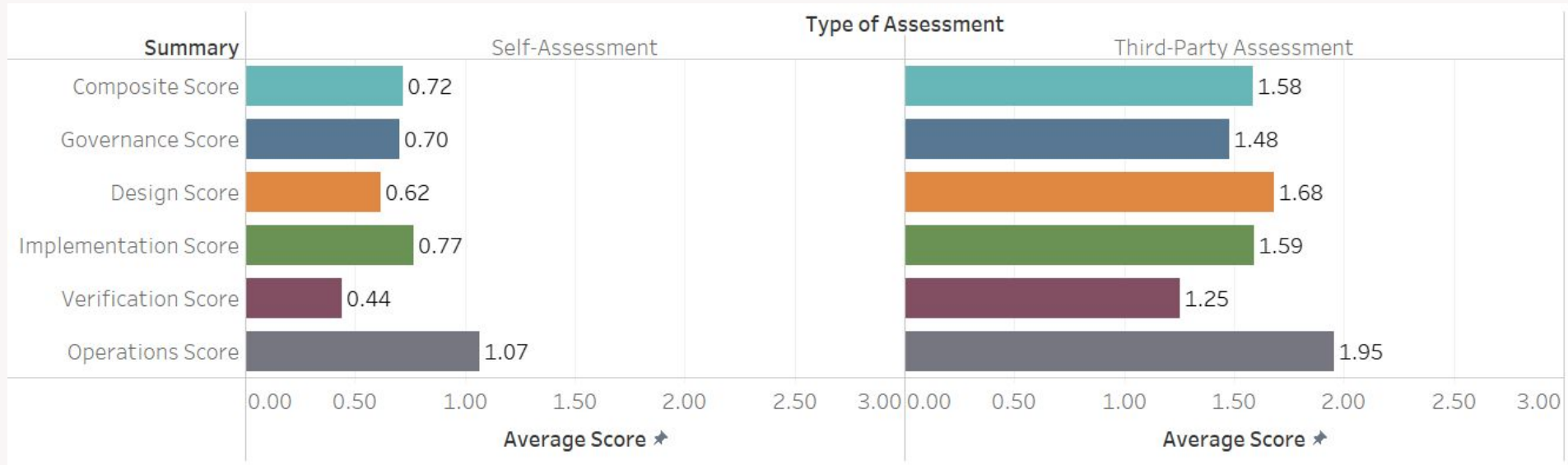
Overall Results: 30 datasets



Overall Results

- Higher score on Operations
 - Expected especially for large multinationals
- Higher score for Implementation
 - The success of the Dev(Sec)Ops paradigm
- Lower Governance score is surprising
 - Arguably due averaging skewing
- Higher score on Design
 - “Shift left”
- Verification is too low

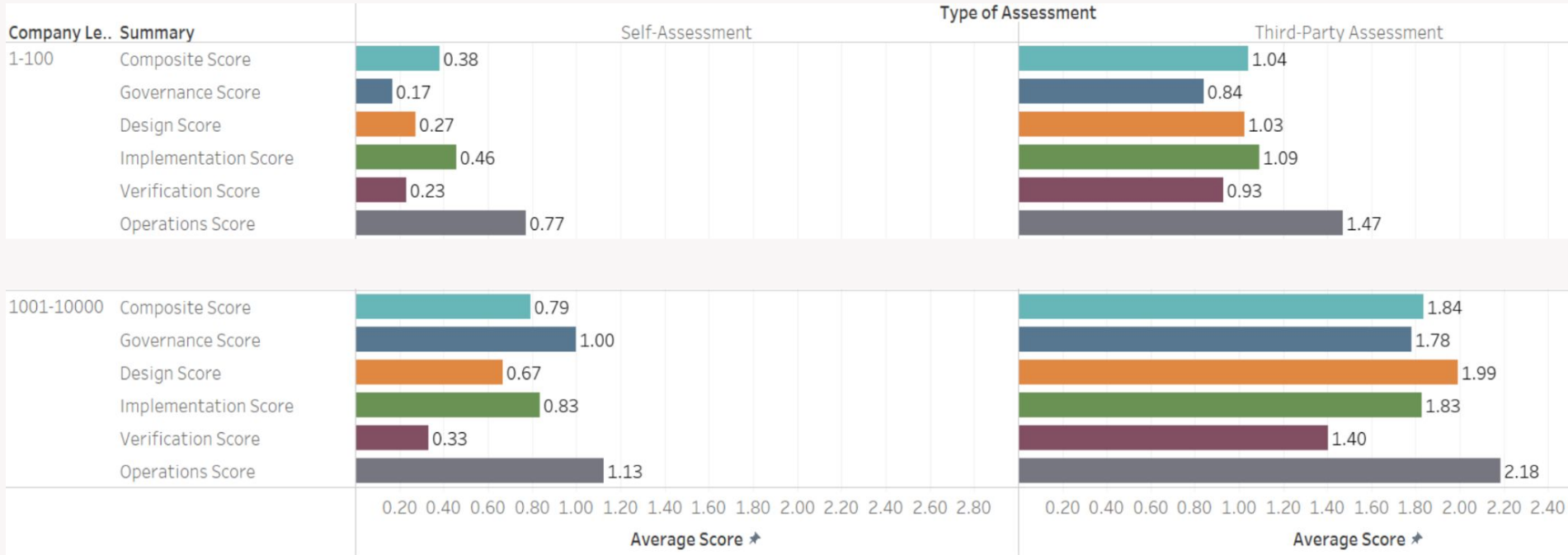
Overall results vs assessment type



Self vs third-party assessments

- Surprisingly low scores for self-assessments
 - Prepped third party assessments
 - Subjectivity factor in self-assessments
- Generally third-party assessments are considered to be more accurate or representative
- Third-party assessments will vary based on skill/knowledge
- Self-assessments will vary based on skill/knowledge and honesty

Overall results vs company size

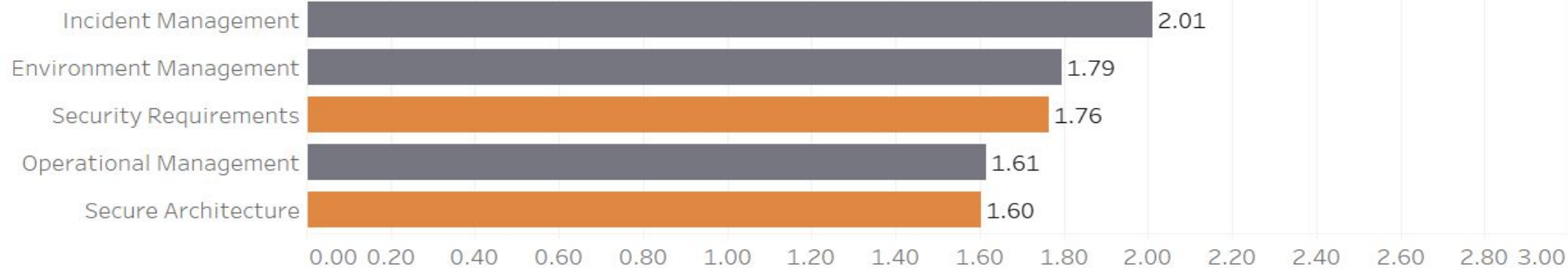


Large vs small companies

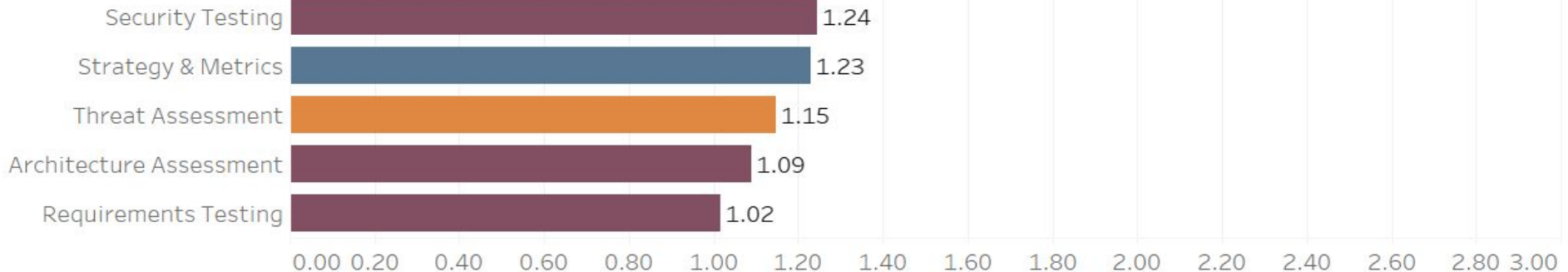
- Expected outcomes
 - Small companies score worse than large
 - Low score on governance for small companies
- Surprises
 - High score on Operations for small companies
 - Self-assessments have a much lower verification score

Top vs Bottom Scoring Activities

Summary



Summary



Avg. Score (Summary Data V) ★

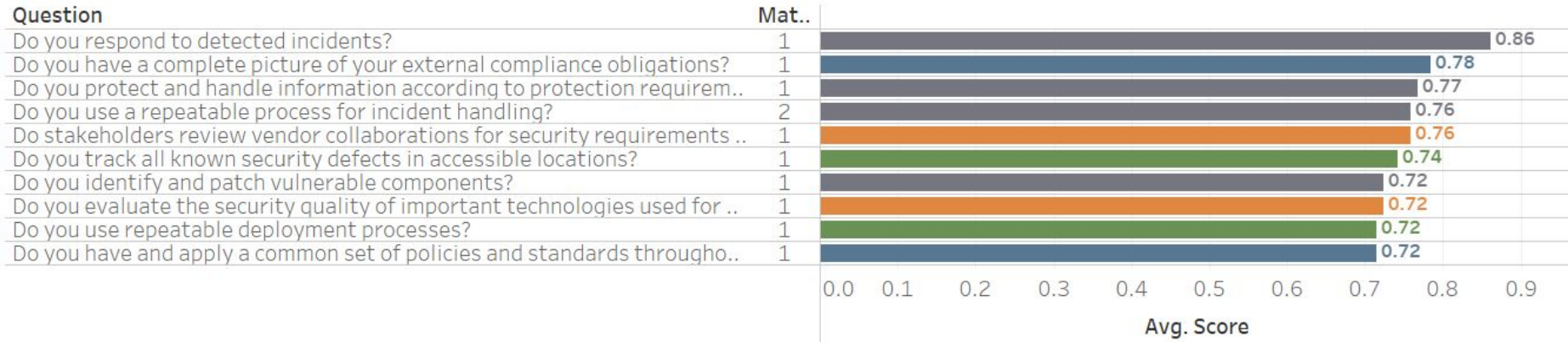
Top vs Bottom Scoring Activities

- Incident & Environment Management are historically handled well at large multinationals
- Deployment is thanks to Dev(Sec)Ops successes
- Security Requirements and Secure Architecture are probably thanks to the “Shift Left” paradigm

Top vs Bottom Scoring Activities

- Low scores on Requirements Testing and Architecture Assessment are surprising in combination with “Shift Left”
 - Did we do the right thing?
- Threat Assessment is historically a low scoring activity
- Low scores on Security Testing is surprising
 - Best practices for SAST/DAST usage
 - Pen testing lessons learned (L3)
- Low scores on Strategy & Metrics is perhaps surprising
 - Averaging issues
 - Metrics is a problematic topic in SAMM

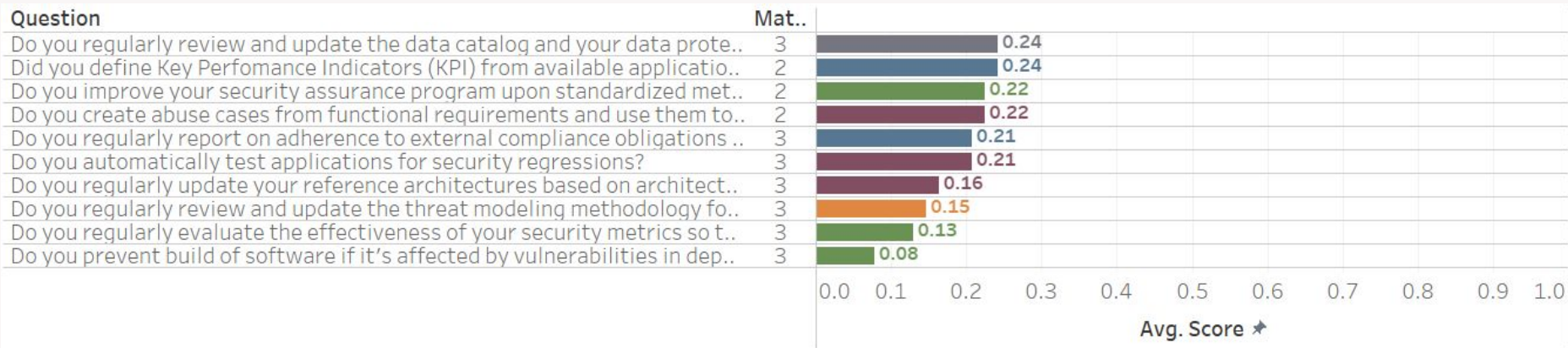
Top 10 Scoring Questions



Top Scoring Questions

- No surprises here
 - 9 top scoring questions are from maturity level 1
 - 1 is from maturity level 2 and related to incident handling

Bottom 10 Scoring Questions



Bottom Scoring Questions

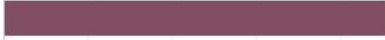




- Most bottom scoring questions are from maturity level 3
- Surprising finding
 - We don't define KPIs, but we improve our AppSec program based on metrics and KPIs.
- Metrics and feedback is a low scorer for both levels 2 and 3
 - Setting up an effective metrics program is hard
- "Do you prevent build of software if it's affected by vulnerabilities in dependencies?"
 - Only few teams enforce "known good" dependencies

Top 5 Ranked Maturity 2 & 3 Questions

Question	Mat..	Rank..		
Do you use a repeatable process for incident handling?	2	4		0.76
Are vendors aligned with standard security controls and software devel..	3	11		0.71
Do you have a dedicated incident response team available?	3	14		0.70
Do vendors meet the security responsibilities and quality measures of se..	2	14		0.70
Do you keep an overview of the state of security defects across the orga..	2	18		0.68

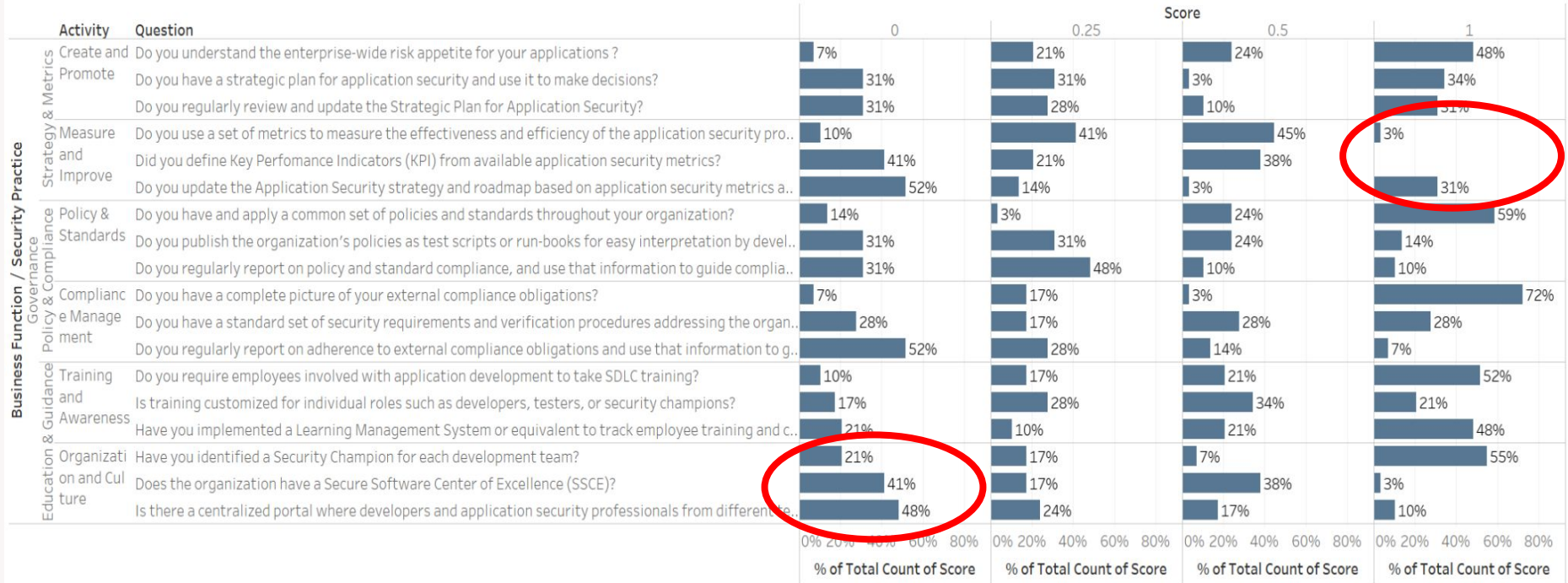
- Looking at what at Maturity level 2 and 3 is being done
 - Incident Response is no surprise, as it's typically inherited from Ops
 - Vendor management is a bit of a surprise, but that may be from the current limited dataset
 - Defect management is a nice addition to this list

Bottom 5 Ranked Maturity 1 Questions

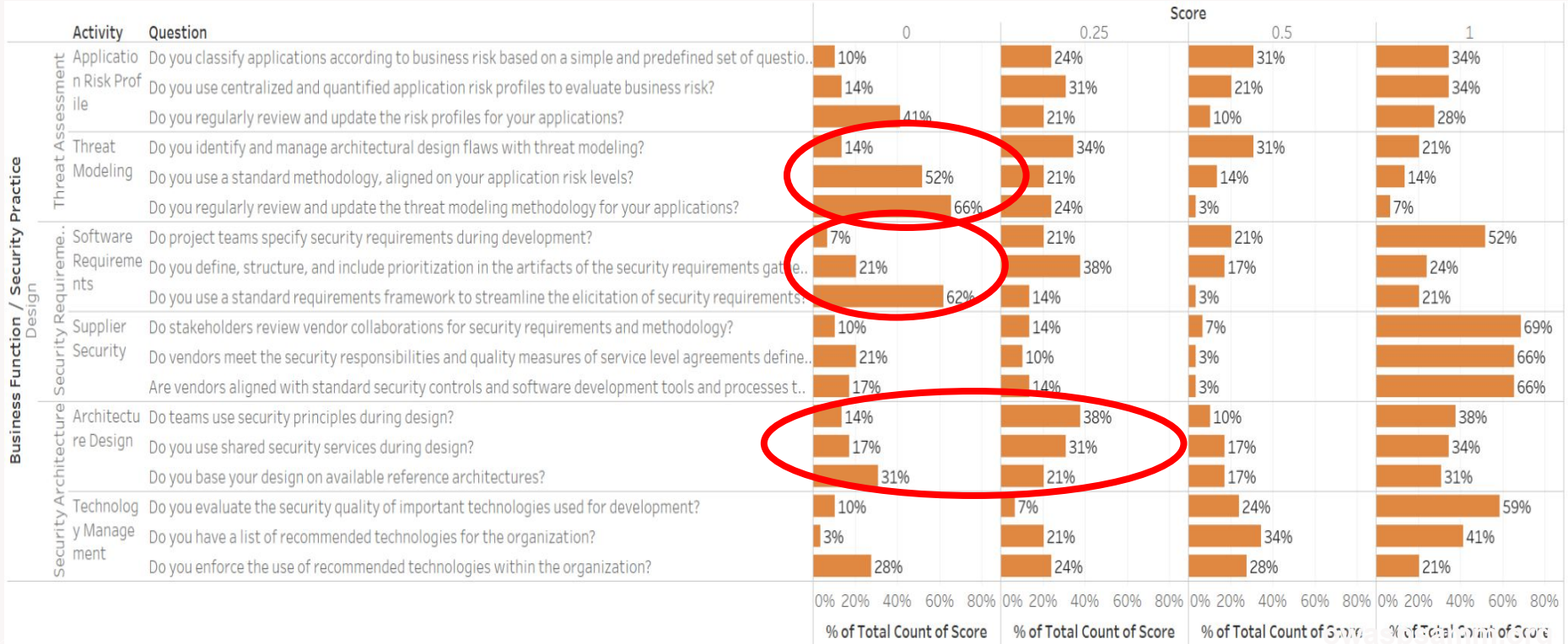
Question	Mat..	Rank..							
Do you review the application architecture for mitigations of typical thre..	1	48		0.46					
Do you identify and manage architectural design flaws with threat modeli..	1	50		0.45					
Do you manually review the security quality of selected high-risk compon..	1	53		0.44					
Do you use a set of metrics to measure the effectiveness and efficiency of ..	1	64		0.36					
Do you test applications using randomization or fuzzing techniques?	1	78		0.25					

- Looking at what at Maturity level 1 is not being done
 - Randomized testing and fuzzing is largely not done
 - Basic Security Metrics are apparently a real struggle
 - Manual reviews of high-risk components
 - Lack of basic threat modeling is still present
 - Reviewing the architecture after deployment is uncommon
 - Basic metrics is just above this list at rank 46

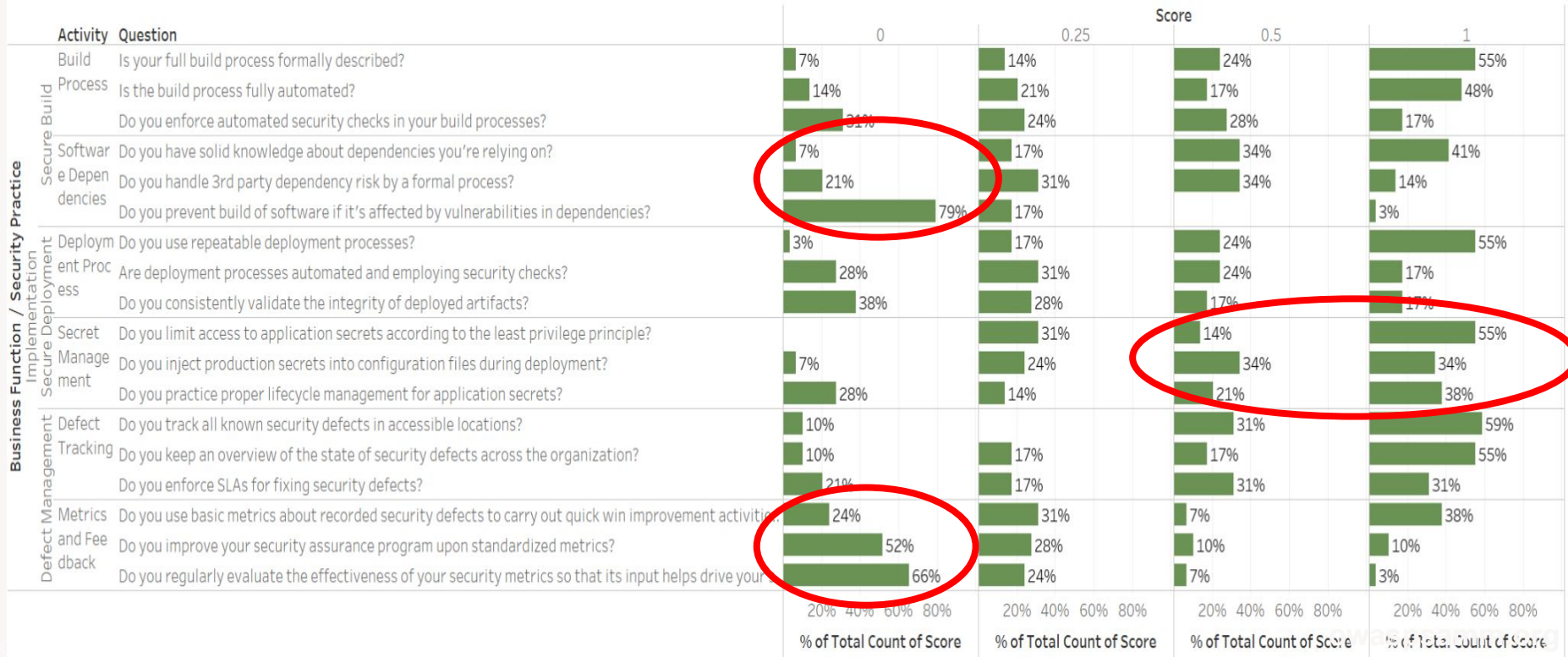
Governance Insights



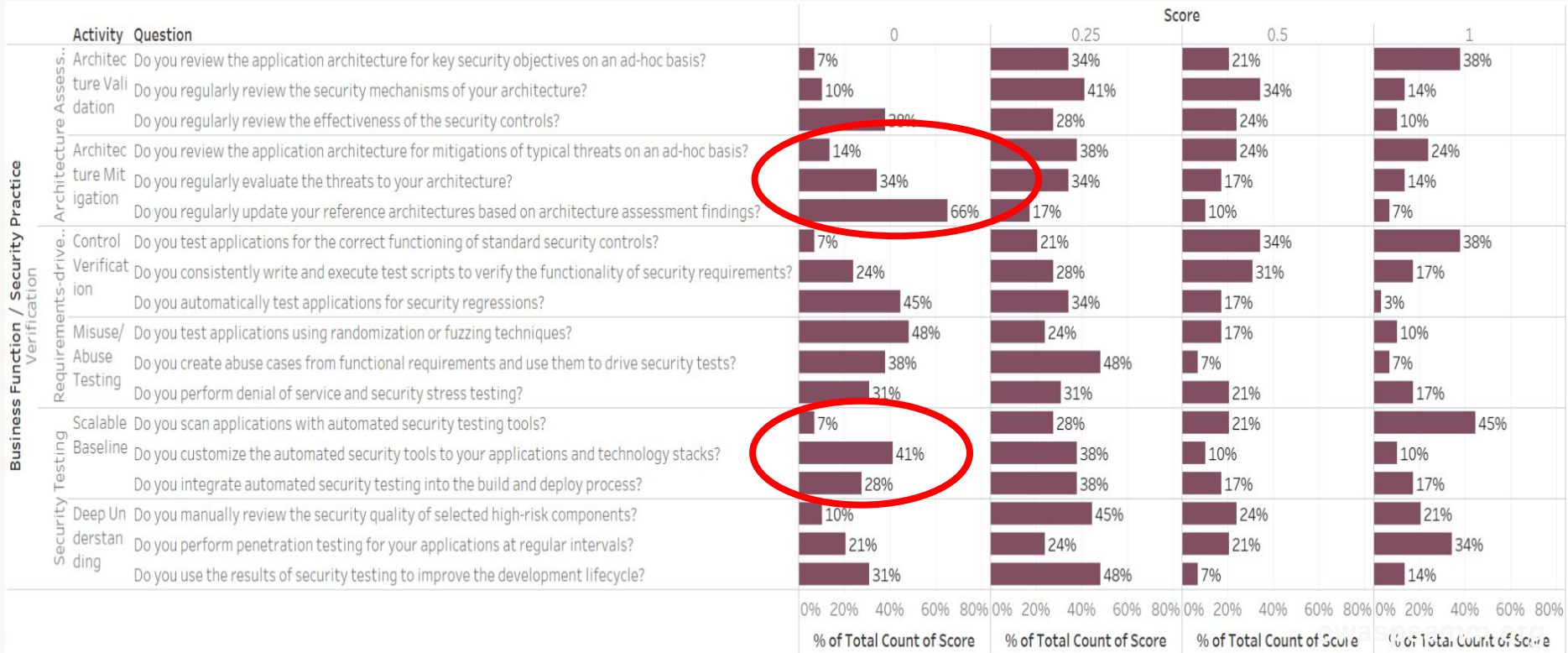
Design Insights



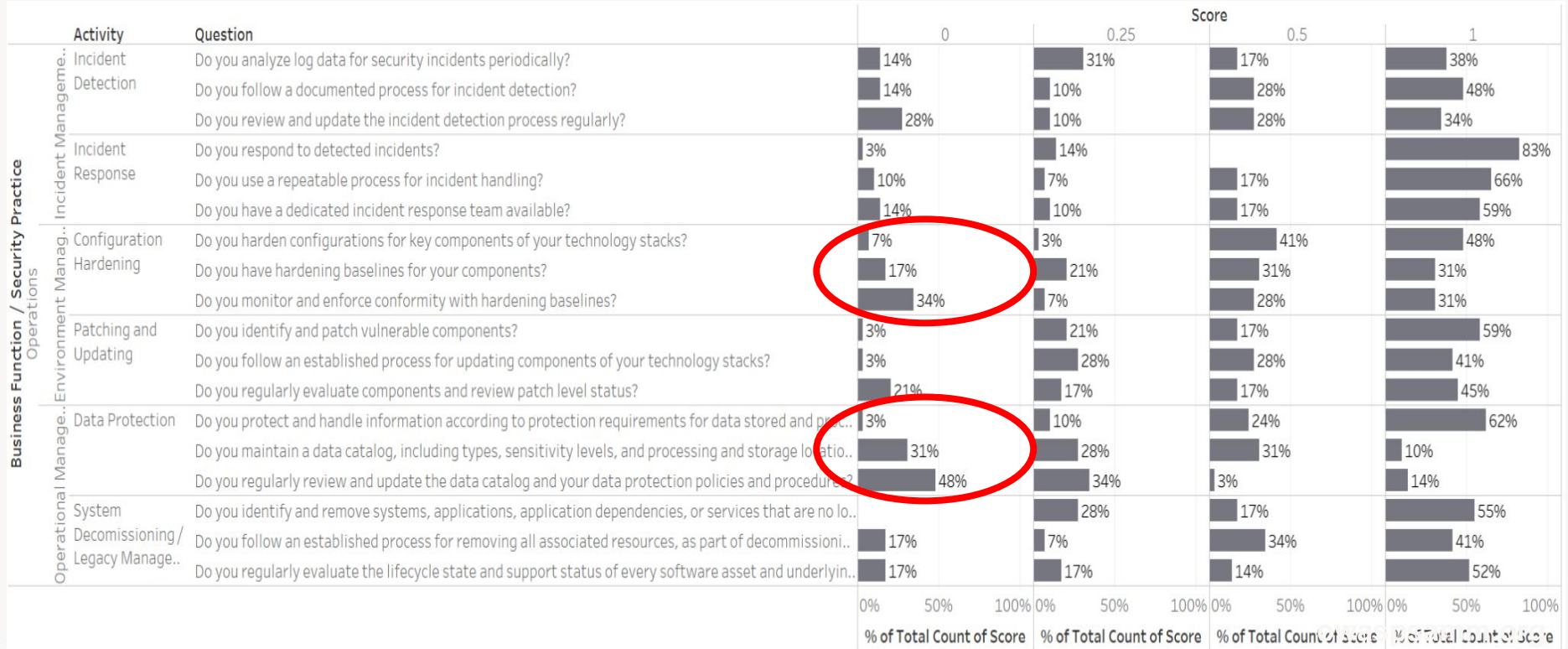
Implementation Insights



Verification Insights



Operations Insights





Thank you!

owaspsamm.org