

In this section, we provide preliminary findings related to IaC metrics and defects. In this study, we used a randomly-selected Mozilla Release Engineering repository that contained Puppet files.

A. Correlation of IaC code metrics with count of defects

First we present which metrics are more correlated with IaC defects, using Spearman correlation coefficient. We observe from Table 1, two IaC metrics namely count of developers and churn show relatively stronger relationship with count of defects.

Table 1: List of IaC metrics that correlate with defect count

Name of metric	Correlation Co-efficient
Size of body text	0.46
Count of lines	0.44
Churn	0.52
Count of developers	0.79

B. Difference in code metrics

Next, we describe which of the code metrics are significantly different for IaC scripts that have at least one defect, and IaC scripts that have no defects at all. The results are shown in Table 2. We used the Wilcoxon rank sum test with a p-value of 0.05 to determine significance. We also recorded the effect size using Cliff's Delta, for each of the presented code metrics. As shown in Table 2, the difference in count of developers between files that have at least one defect and no defects at all, is significant and large.

Table 2: List of IaC metrics that are significantly different

Name of metric	Effect size
Size of body text	0.42**
Count of lines	0.39**
Churn	0.35**
Count of developers	0.50***

*** indicates large effect size
** indicates medium effect size

Findings from Table 1 and 2 indicate that a subset of metrics extracted from IaC scripts are more associated with count of defects, and can be potentially used to create IaC defect prediction models.

C. Categories of defects

Finally in the third subsection, we categorize the identified defects, using a defect classification scheme proposed by Thung et al. [5]. The percentage of defects that belong to each category is provided in Table 3.

Table 3: Categories of defects for Puppet repository

Category	Percentage of Defects
Control and data flow	64.97%
Non functional	13.32%
Structural	21.71%