

Hyperstyle Error Codes Analysis

Qualitative Analysis

Introduction

After comparing the outputs from Pylint and Hyperstyle on the same dataset, it was very obvious that Hyperstyle generates a larger quantity of more specific error messages. For example, Pylint would generate an error code such as W0612(Variables unused) while on the other hand, Hyperstyle's version of this error code would be B007(Loop control variables unused).

Mapping multiple Hyperstyle error codes to a single Pylint error code

Although the error codes in Hyperstyle are more detailed, some of them are actually very similar to each other and could be mapped to the same Pylint error code. Eg: the Hyperstyle error codes E111(Used when indentation is not a multiple of a certain number), E113(Used when there is unexpected indentation), E999(Used when there's an indentation error) could be mapped to the Pylint error code W0311(Used when an unexpected number of indentation's tabulations or spaces has been found).

Merging Pylint and Hyperstyle error codes

For many instances, Hyperstyle and Pylint display error codes that are different from each other but by definition, they are the same. Later on, we have decided that these error codes should be merged to the same error code for more consistency.

Conclusion

In conclusion, Hyperstyle catches more error codes as it goes into more details. The advantage of this is that sometimes we can catch more errors that are related to refactoring (too many elif branches). However, some of the error codes are arguably petty and unnecessary, which should be filtered.

Quantitative Analysis

Categorizing Pylint output using grouping convention from Stegman/

Category	Frequency
Names	2
Comments	3
Layout	4
Formatting	7
Flow	2
Expressions	8
Decomposition	0
Modularization	0
Best Practice	8
Error	3

Categorizing Hyperstyle output using grouping convention from Stegman

Category	Frequency
Names	4
Comments	4
Layout	0
Formatting	13
Flow	3
Expressions	4
Decomposition	0
Modularization	0
Best Practice	3
Error	3

It should be noted that the last 2 rows are not categories from Stegman. While tabulating the data, it has come to our notice that some error codes just do not belong in any of the 8 categories mentioned by Stegman. Therefore, they have been added to the last 2 rows of the table.