
Looking at SAS code as data

Christof Binder, F. Hoffmann-La Roche



Introduction

- Working with a big code base can be difficult
- Traditional approach
 - Code review
 - Reading documentation
 - Reviewing test cases

Code review

Roche



Looking from a different angle

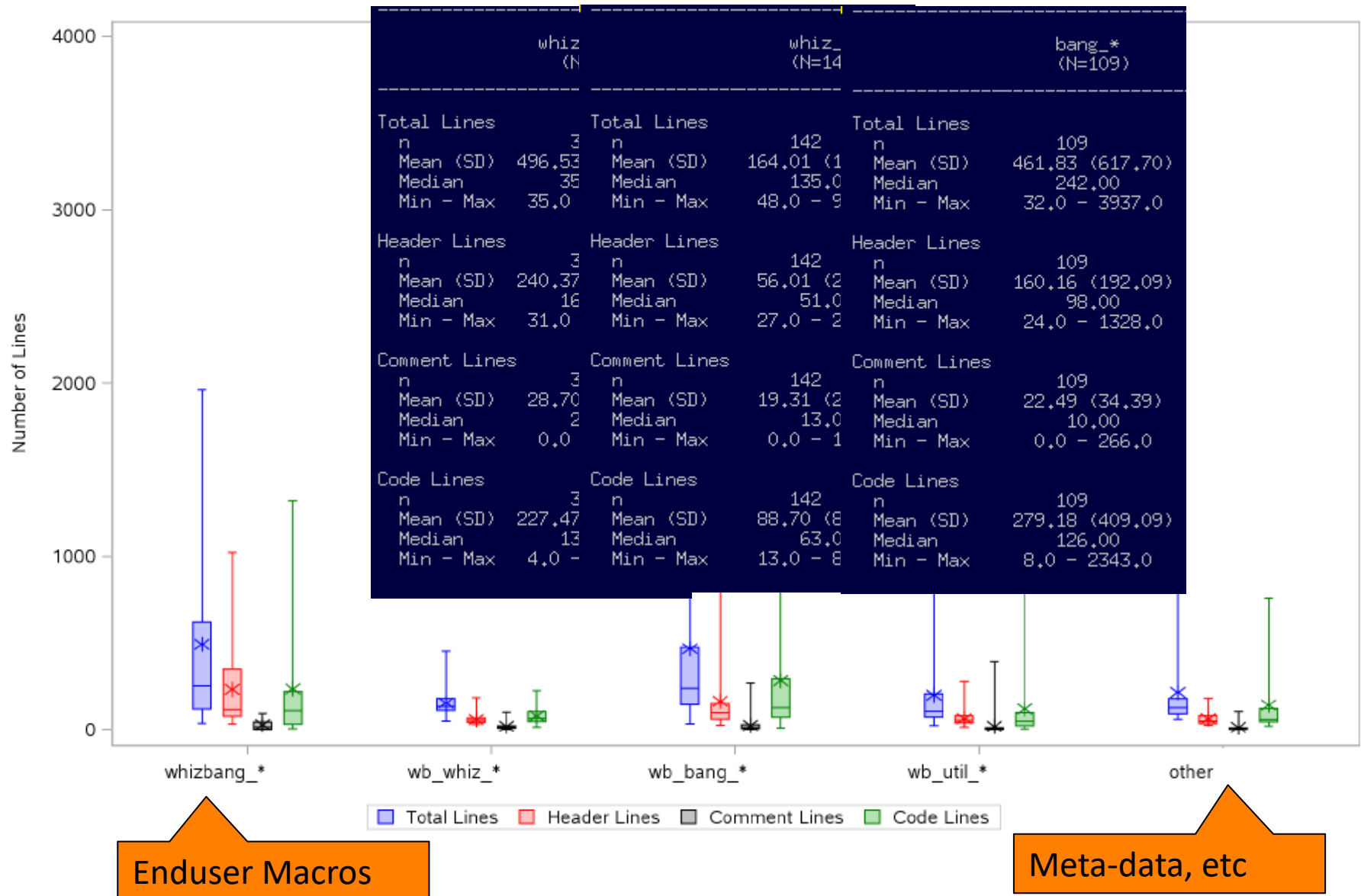


- A bigish system with two components
- Whiz - part
- Bang - part
- A few end-user macros and a lot of utility/internal macros

Look at basic statistics

- Definitions
 - Header: first comment in the file: `/* ... */`
 - Comment: any `/* .. */`, `* ... ;` or `%* ... ;` line after the header
 - Code: any non-header, non-comment, non-blank line

Basic Statistics



Dynamic view

- Methodology
 - Using perl/python scripts to generate the raw data
 - Using D3 javascript library for the visualisation



Summary

- Taking over a system can be daunting (even if it's your own after a while)
- Code review of a large system might not be feasible
- Viewing code as data can open new approaches
- Static analysis can point to macros that warrant a closer look
- Dynamic presentation can give a good feel about the code and focus the attention on clusters of interest, i.e. generate questions

Thank
you



***Doing now what patients need
next***