

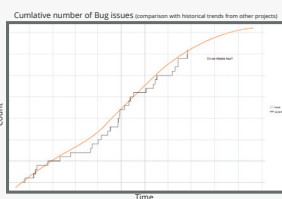
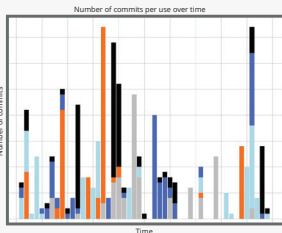
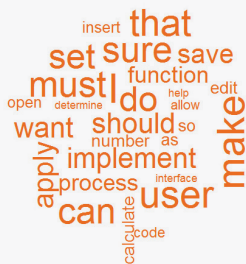
Software development by the numbers

Adnan Fiaz Data Scientist, Mango Solutions
Email afiaz@mango-solutions.com

There are numerous software development methods out there and the evolution of these methods still continues. However, this evolution has not yet embraced data or employed analytics.

EMBRACING DATA

The software development process is a rich source of data. We often generate data with most business processes but rarely do we store it or store it in a structured format.



EMPLOYING ANALYTICS

Data on its own won't improve your process, but using data to employ analytics can. Depending on the process step there are different use cases.

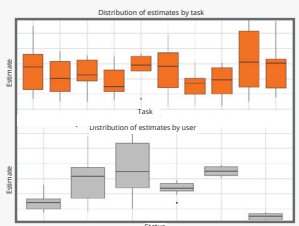
PLANNING



Any process will have a planning step and this step too generates data:

- Requirements
- Time estimates
- Team/people

- Apply text mining to uncover what makes a good user story
- Use past estimates to assess estimation accuracy and improve future decisions



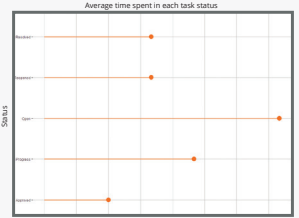
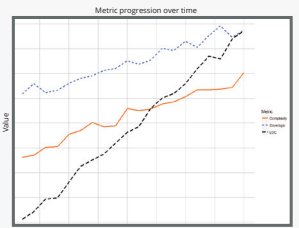
IMPLEMENTATION



When it comes to writing code and managing tests there is an established toolset that generates structured data:

- Code
- Source control
- Task Management

- Give code feedback using packages such as goodpractice, lint, covr, to improve coding ability.
- Anticipate potential problems by tracking and predicting code metrics
- Relate code commits to other data (e.g. defects or build failures)
- Use data on task progress to uncover potential bottlenecks



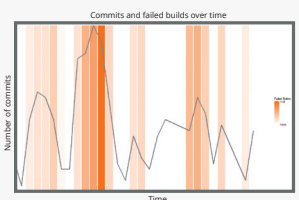
DEPLOYMENT



The software development process doesn't end after a product is deployed. It just enters into a new cycle of development, but we can now also consider data from two other sources:

- Defects
- Builds (CI)

- Track defects and compare with past projects to anticipate future defects
- Predict the probability a project will overrun or fail



SUMMARY: Regardless of the software development methodology used, data is generated at every stage in the development cycle. This can be structured data (in the form of code, commits or tasks) or unstructured data (in the form of requirements). Either way its data that we can use to improve our process.

When it comes to developing software this improvement means planning more efficiently, writing better code, preventing bugs or build failures, and delivering satisfying software products.



2 Methuen Park, Chippenham, Wiltshire, SN14 0GB UK
+44 (0)1249 705 450
info@mango-solutions.com
mango-solutions.com