

Report Deliverable 1

MARCO CALAVARO

MATRICOLA: 0295233

RoadMap

- Introduzione p.3
- Progettazione p.4
 - Codice
 - Analisi dei valori
- Variabili p.6
- Risultati p.7
- Discussione dei risultati p.8
- Risultati progetto durante il corso p.10
- Link p.11

Introduzione

L'obiettivo di questo primo deliverable è stato quello di produrre un process control chart che misuri la stabilità di un attributo di progetto.

L'attributo studiato: numero di ticket di tipo fixed al variare dei mesi

Il recupero di tali valori è stato effettuato sul portale apache dedicato agli issues in Jira ([link](#))

Progetti analizzati: BookKeeper , ZooKeeper.

(Durante il corso è stato effettuato lo studio sul progetto STDCXX)

Ambiente di sviluppo: sistema operativo Windows 10 e Eclipse IDE.

Deploy: il progetto è stato caricato su GitHub, con collegati sia TravisCI che SonarCloud.

- I [link](#) sono disponibili alla pagina finale.

Progettazione Codice

Per poter realizzare lo scopo del progetto era necessario ottenere delle misurazioni sull'attributo ricercato.

È stato quindi realizzato un codice che sfrutta le API di Jira per ottenere tutti i ticket con resolution pari a fixed, per il tipo di ticket ricercato non è stato necessario aggiungere altri filtri.

La chiamata alle suddette API restituisce l'elenco dei ticket in formato Json, non tutti i ticket vengono considerati nello studio.

Ai fini dello studio sono stati considerati solo i ticket per i quali è stato trovato un commit associato sul version control system del progetto.

Ad esempio se il ticket ha come chiave rappresentativa STDCXX-501 vi deve essere un commit che nel commento contenga tale chiave per essere considerato valido ai fini dell'analisi.

Ottenuto l'elenco effettivo di ticket validi si passa a salvare i risultati, nello specifico la chiave del ticket e la resolution date, in un file formato csv che rappresenta il dataset sul quale è stata effettuata l'analisi.

Progettazione Analisi dei valori

Il codice ha avuto il solo scopo di creare un dataset di partenza, ma il resto delle computazioni e la creazione dei grafici è stata realizzata attraverso Microsoft Excel. In particolare attraverso le tabelle pivot si è calcolato il numero di ticket in ogni mese. Tramite tale tabella sono stati ricavati i valori di media e varianza della distribuzione dei ticket .

Nel periodo di osservazione non tutti i mesi avevano dei commit per questo, tramite le impostazioni di tale tabella, sono stati aggiunti i valori mancanti (numero ticket pari a 0).

Sono stati poi calcolati i valori di:

- Media (mean)
- Deviazione standard (STDV)
- Upper limit : $\text{mean} + 3 * \text{STDV}$
- Lower limit : $\text{mean} - 3 * \text{STDV}$

Variabili

Per studiare i risultati delle misurazioni e i grafici ottenuti è necessario prima di tutto descrivere tutte le variabili considerate nel dataset e in Excel.

Dataset:

- Tickets key: rappresenta il nome univoco del ticket (NOMEPROGETTO-id)
- Resolution date: data associata alla risoluzione del ticket

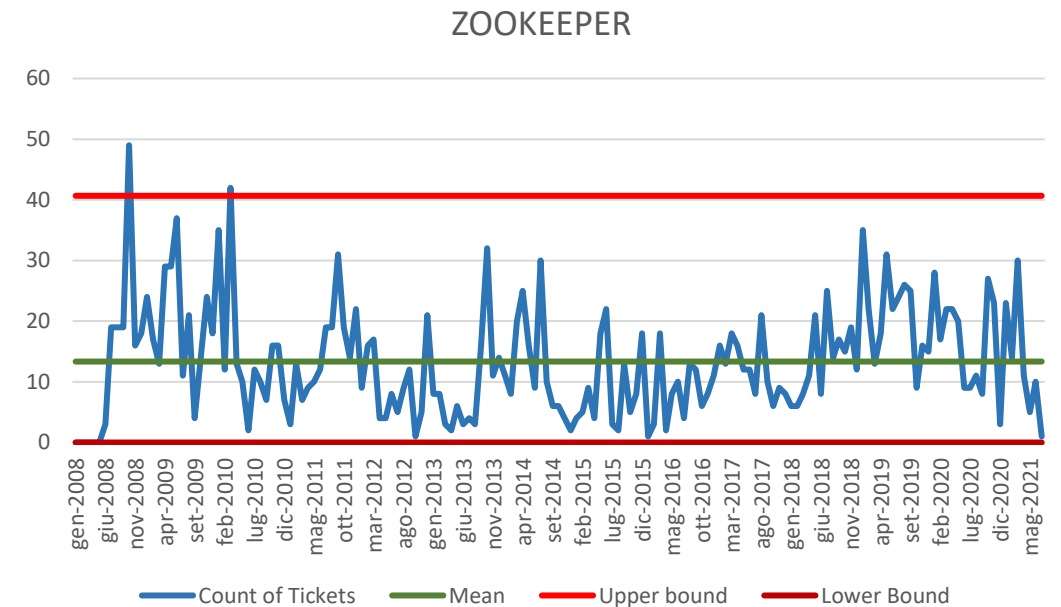
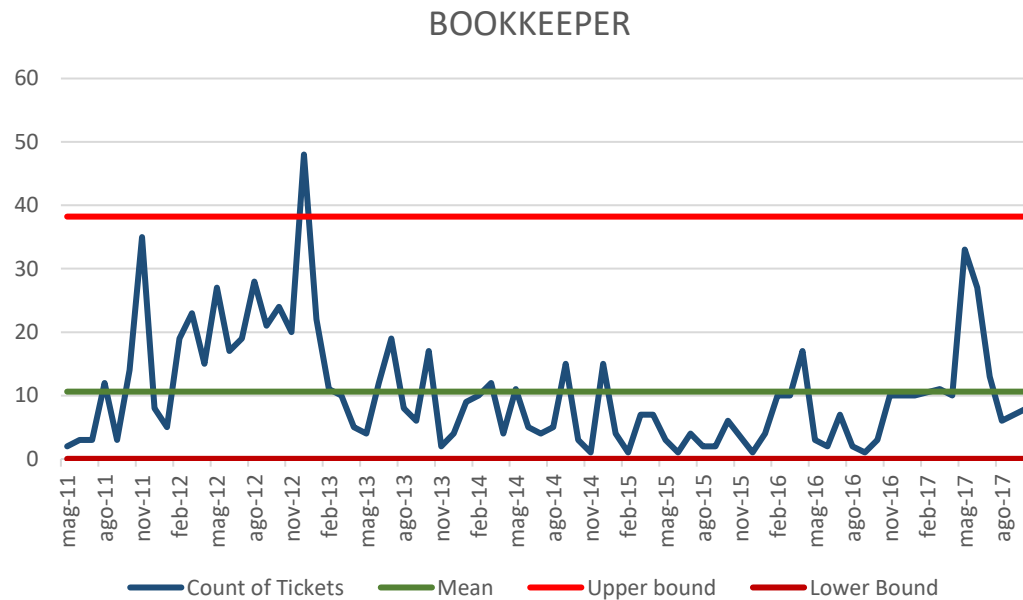
Tali valori sono stati trasformati in una nuova tabella mediante la funzione pivot di Excel:

- Date: corrisponde ad un mese
- Count of Tickets: il numero di ticket misurati nel mese corrispondente a Date
- Mean: media di Count of Tickets
- STDV: deviazione standard di Count of Tickets
- Upper limit e lower limit calcolati come descritto in precedenza

Risultati

Completata l'analisi si è creato il process control chart:

- Sull'asse x vi è la Date
- Sull'asse y vi è il conteggio dei tickets



I grafici sono stati riportati senza disegnare i punti per una maggior chiarezza visiva (aggiunti nell'Excel)

Discussione

Osservando i due grafici si può subito notare come l'intervallo temporale in cui è stato possibile effettuare un'osservazione e trarne le opportune misurazioni è molto differente per i due progetti.

ZooKeeper è un progetto apache molto seguito con un continuo sviluppo, che comporta molti più ticket registrati e molta oscillazione del valore osservato

BookKeeper invece è un progetto secondario che non gode della stessa organizzazione nello sviluppo rispetto a ZooKeeper. Inoltre indagando sui ticket presenti in Jira per tale progetto, si è scoperto che l'attività legata agli issues è stata trasferita su GitHub a partire dall'ottobre del 2017 ([BOOKKEEPER-1107](#)) ecco perché l'intervallo di osservazione è inferiore.

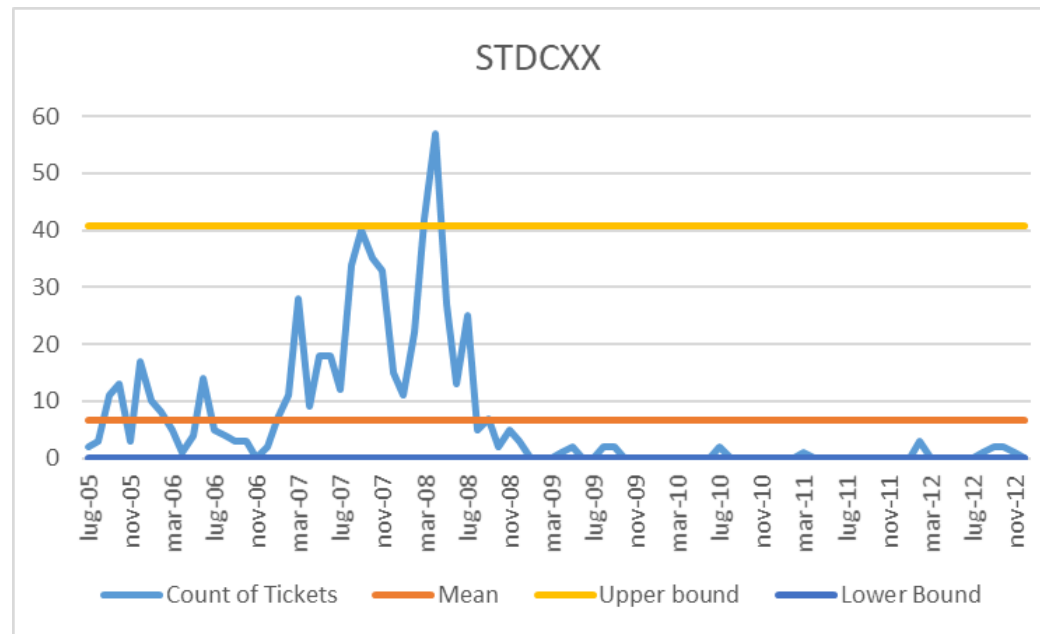
Discussione

Scendendo ora nell'analisi più dettagliata dei grafici risultanti si osserva che:

- Il lower limit per entrambi i progetti è impostato a zero poiché il valore restituito dalla formula è negativo e naturalmente non è significativo. Il numero minimo di tickets in un mese non può essere minore di zero.
- L'upper limit per entrambi i progetti è abbastanza simile ed è approssimativamente 40
 - Entrambi i progetti hanno superato almeno una volta tale bound, in particolare nelle fasi iniziali del progetto
 - Si osserva inoltre che la curva è tendenzialmente sopra la media nella prima parte del periodo di osservazione. Questo è abbastanza comprensibile poiché le prime fasi di un progetto sono caratterizzate dalla risoluzione dei problemi (si sta misurando il numero di tickets fixed, cioè che sono stati risolti), mentre con il passare del tempo si presuppone che si raggiunga una maggiore stabilità.
 - Questo fenomeno è stato riscontrato maggiormente nel progetto BookKeeper.

Risultati STDCXX

Di seguito è riportato il grafico risultante effettuato durante il corso per il progetto della libreria standard c++.



Link

- GitHub: <https://github.com/IlConteCvma/Deliverable1>
- Travis : <https://travis-ci.com/github/IlConteCvma/Deliverable1>
- SonarCloud: https://sonarcloud.io/dashboard?id=IlConteCvma_Deliverable1