

- Prenom : Salim
- Nom : El Rhilani
- Prenom : Mamadou Daye
- Nom : Bah

Tache 1:

- voir annexe github dossier task1 information pertinentes a calculer sont:
- l = quartile inferieur
- m = median
- u = quartile superieur
- $d = u - l$, longueur de la boite
- s = limite superieure, $s = u + 1.5d$

- $i = l - 1.5s$

- Pour DCP:

- $l = 49.25$
- $m = 59.21$
- $u = 82.41$
- $d = 82.41 - 49.25 = 33.16$
- $s = 82.41 + 1.5 * 33.16 = 132.15$
- $i = l - 1.5s \Leftrightarrow 49.25 - 1.5 * 33.16 = -0.49$

- Pour NCLOC:

- $l = 12$
- $m = 71.5$
- $u = 180$
- $d = u - l = 168$
- $s = u + 1.5 * d = 180 + 1.5 * 168 = 432$
- $i = l - 1.5s = 12 - 1.5 * 432 = -598.75$

- Pour NOCom:

- $l = 3$
- $m = 5$
- $u = 9$
- $d = 9 - 3 = 6$
- $s = u + 1.5 * d = 9 + 1.5 * 6 = 18$
- $i = l - 1.5 * s = 3 - 1.5 * 18 = -24$

Conclusions:

- Metrique DCP: Avec le test QQPLOT on constate que les points suivent grandement une ligne droite Avec le test Boite a moustache on constate que la partie inferieur de la boite represente presque le double de la longue de l'extremite superieur de la boite a moustache Avec la respresentation en histogramme de la distribution on constate la distribution est loin detre normal tel que nous ne remarquons pas une symetrie On ne peut accepter l'hypothese que la distribution des donnes de cette metrique soit normal
- Metrique NCLOC: Avec le test boite a moustache, on constate la partie inferieur de la boite a moustache est presque nul, on est loin d'avoir une distribution normale Avec le test QQPLOT, on constate que les poins s'ecartement totalement de la ligne droite Avec la representation histogramme, nous sommes loin d'avoir une symetrie. On ne peut accepter l'hypothese que la distribution des donnes de cette metrique soit normal
- Metrique NoCom: Avec le test boite a moustache, Il y a une inegale repartion, des longueurs de la boite. Avec le test QQPLOT, au debut les points semblent suivrent une ligne droite puis s'ecartent de la ligne Avec la representation histogramme, On constate que la distribution n'est pas symetrique On ne peut accepter l'hypothese que la distribution des donnes de cette metrique soit normal

Tache 2 :

- voir annexe github dossier task2

Étudiez les corrélations entre NoCom et NCLOC, NoCom et DCP. Visualisez les données, calculez les droites de régression, et les coefficients de corrélation qui ont du sens.

- Comme dans la tâche précédente on a conclu que les métriques sont vraisemblablement pas normalement distribuées, nous allons utiliser la corrélation robuste pour analyser la relation entre nos métriques par le coefficient de Spearman.
- On observe que la corrélation entre NoCom et NCLOC est d'environ égale à 0.688; Celle entre NoCom et DCP de -0.534 (voir images et sources pour plus de détails).
- Comme l'intervalle des valeurs ainsi que leur interprétation sont les mêmes que pour r , le résultat est donc toujours compris entre -1 et 1 .
- Cependant, $\rho \neq 1$, $\rho \neq -1$ et $\rho \neq 0$ pour nos deux résultats obtenus. On ne peut donc pas dire que ces métriques ont une parfaite corrélation (linéaire ou négative) ou qu'elles ne sont pas du tout corrélées.
- En ce qui concerne la régression linéaire entre NoCom et NCLOC, on trouve que $a = 36.047$ et que $b = -97.9925$ – Ce qui nous donne une droite d'équation $\hat{Y} = -97.9925 + 36.047x$.
- Entre NoCom et DCP, $a = -1.83$ et $b = 75.8095 \implies \hat{Y} = 75.8095 - 1.83x$.
- Lorsqu'on trace les deux droites, on retrouve la présence de quelques points extrêmes (voir NoCom_NCLOC_4.jpeg & NoCom_DCP_4.jpeg), ce qui pourrait influencer significativement l'estimation de a et b .

Tache 3 :

- voir annexe github dossier task3
- - Nous voulons évaluer l'hypothèse : « les classes qui ont été modifiées plus de 10 fois sont mieux commentées que celles qui ont été modifiées moins de 10 fois ». Décrivez la conception d'une quasi-expérience qui vous permettra de le faire. Ensuite, évaluez l'hypothèse, discutez les résultats et décrivez vos conclusions.

Rappel :

Quasi-expériences • Impossible de tester physiquement (concrètement) les hypothèses • Impossible de former un « vrai » groupe de contrôle pour des raisons pratiques et/ou éthiques • Pas d'équivalence entre groupe d'étude et groupe de contrôle • Affectation aux groupes contrôlée • Traitement non contrôlé

- choix d'étude : Expériences (voir chap6 p.17)

Quasi-expérience : - On sélectionne différentes classes de nos projets ou d'autres trouvées sur internet. - Parmi ces classes, il y en a qui auront eu plus de 10 modifications / commits, et d'autres moins, nous allons soumettre ces classes à une évaluation sur une échelle à différentes personnes avec un niveau en programmation variable, cela peut aller d'une personne avec quelques jours d'expérience, à plusieurs mois. - Nous allons attribuer à chacun des classes que nous aurons nous-même choisies, ils auront la possibilité de noter ces classes sur une échelle de 1 à 5. - Nous ferons ensuite une moyenne et des comparatifs pour arriver à une conclusion satisfaisante.

Variable : - La classe - Nombre de modification - Appréciation : Note attribuée par l'utilisateur à la qualité des commentaires - Complexité : Complexité de la classe selon l'utilisateur sur une échelle de 1 à 5

Interprétation :

- Est-ce que la taille de la classe impacte sur les résultats ?
 - Après avoir réalisé l'expérience on constate que la taille d'une classe n'a pas forcément un impact sur les résultats,
- Est-ce que beaucoup de commentaires implique une bonne qualité de commentaires ? Non, les petites classes AGE.java et utils.java par exemple contiennent une ligne de commentaire par fonction ce qui est largement suffisant pour la

comprehension

- L'utilisateur arrive a comprendre le code sans jamais l'avoir vu/utilise avant ? Pour les petites classes, ayant un petit nombre de fonction l'utilisateur arrive a comprendre le role des fonctions de la classe. Par ailleurs l'utilisateur comprend sans difficulte le fichier IFT1015tp2.py malgre sa taille consequente grace a la bonne documentation. Le fichier cryption est difficile de comprehension pour les utilisateurs en general car on ne sait pas exactement ce que fait le fichier malgre les commentaires.
- Discussion :
 - Apres avoir realise notre quasi-experience on constate que le nombre de commits n'a pas forcement un impact sur la qualite/nombre de commentaire. Il est difficile tout de meme de juger ceci car les commentaires restent quelque chose de subjectif. Par ailleurs, On remarque que les fichiers avec le plus de commits ont une moyenne plus elevee que les fichiers avec moins de 10 commits, ce qui pourrait tout de meme nous laisser penser que plus des commits sont effectues sur un fichier plus la qualite des commentaires pourrait augmenter. Enfin, cet avis reste mitige car la classe Age.java avec moins de 10 commits possede une moyenne plus elevee que la classe utils.java. Les commentaires dependent de plusieurs parametres, tel que la complexite du projet, taille du fichier etc..