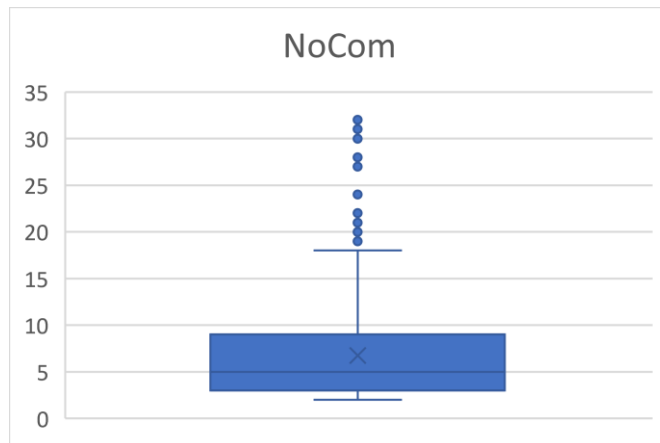


IFT 3913 TP3

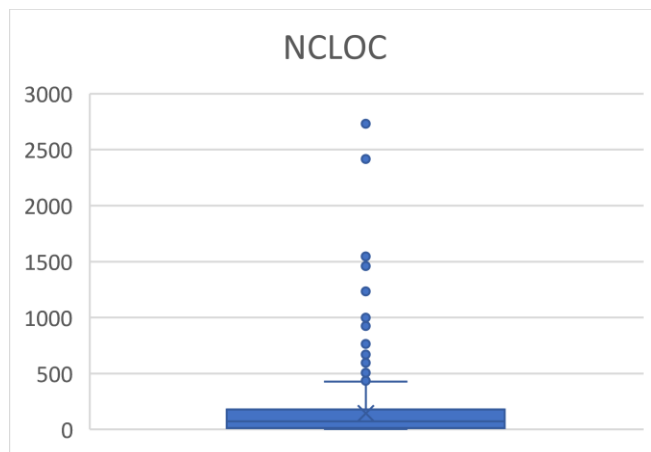
Par Hugo Carrier - 20197563 et Maggie Robert - 20182443

18 novembre 2022

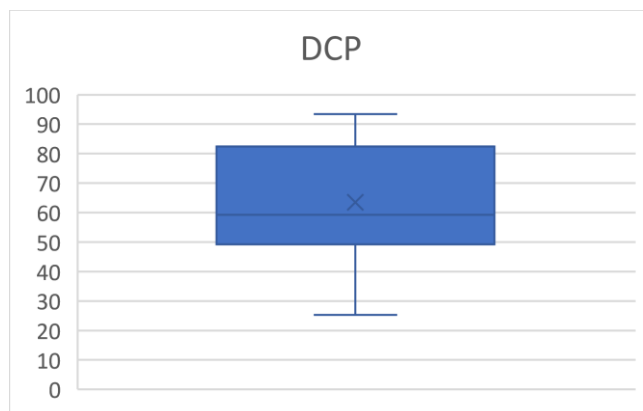
1. Visualiser les données



m	Médian	5
u	Quartil1	3
l	Quartil3	9
	Min	2
	Max	32
d	Longueur	6
s	Limite supérieure	12
i	Limite inférieure	2

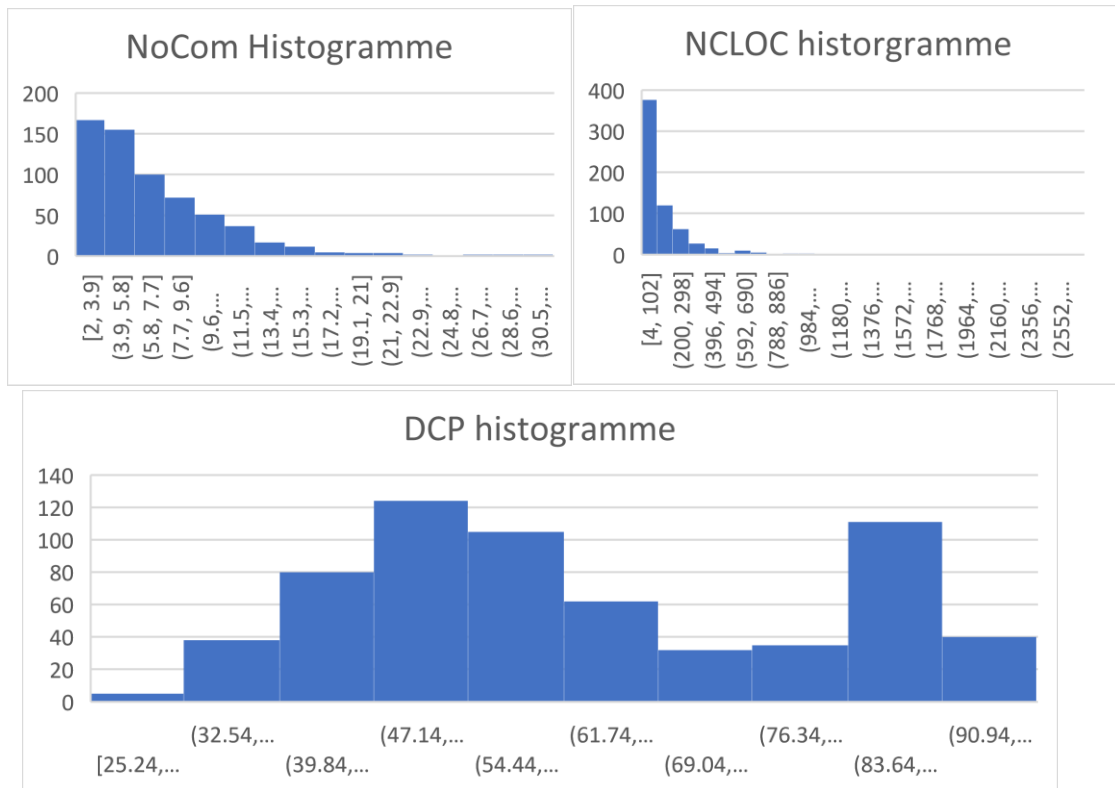


m	Médian	71.5
u	Quartil	12
l	Quartil	180
	Min	4
	Max	2732
d	Longueur	168
s	Limite supérieure	264
i	Limite inférieure	4

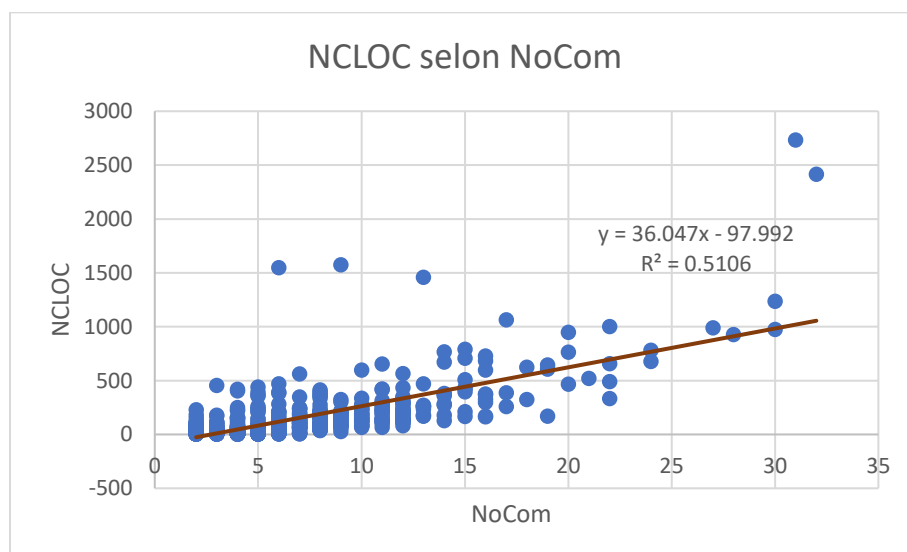


m	Médian	59.21
u	Quartil	49.27
l	Quartil	82.36
	Min	25.24
	Max	93.44
d	Longueur	33.09
s	Limite supérieure	98.905
i	Limite inférieure	32.725

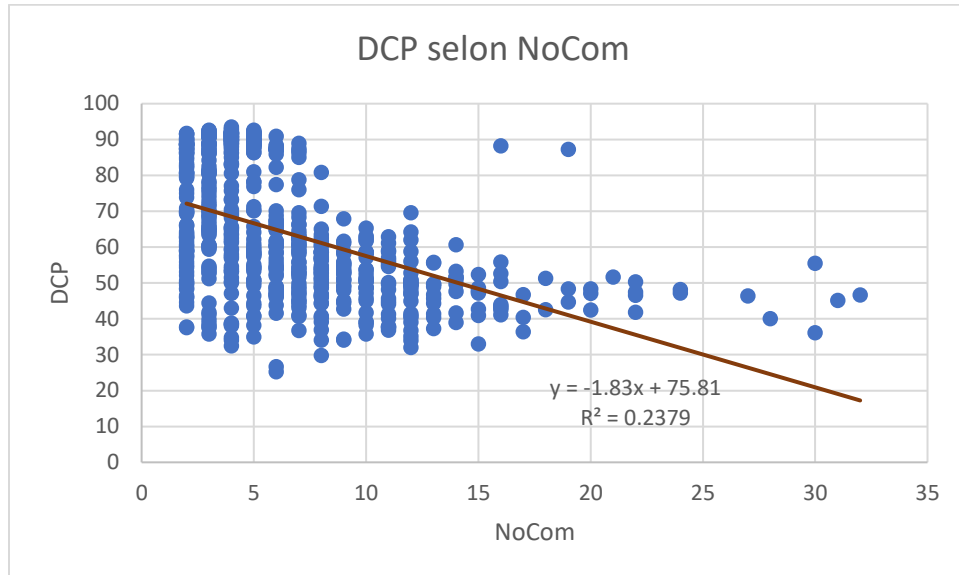
Nous avons utilisé le test de Jarque-Bera afin de déterminer rapidement si les données sont distribuées de façon normale. Dans les 3 cas, elles ne le sont pas. Nous avons un résultat de 491 pour NoCom, 37861 pour NCLOC et 471 pour DCP. On s'attendait à une valeur de moins de 0.62 pour la taille de nos données. Visuellement, nous pouvons voir que les variables sont soit très biaisées (NoCom et NCLOC) ou aplaties (DCP).



2. Étudier les corrélations



Spearman	0.688037
----------	----------



Spearman	-0.53352
----------	----------

Étant donné que nos distributions ne semblent pas être normales, nous ne pouvons pas utiliser la corrélation de Pearson. Nous avons utilisé celle de Spearman. Elle est aussi appelée la corrélation de rang et elle compare les positions relatives. Selon les corrélations mentionnées plus haut, NCLOC semble varié dans le même sens que NoCom et DPC semble varié dans le sens contraire de DCP. On peut faire cette même remarque en observant les pentes des régressions et nous avons aussi sortie le coefficient de détermination (r^2). On remarque que les deux régressions semblent être de mauvaise qualité. Logiquement, les données n'étaient pas normales et lorsqu'observée, elles ne semblaient pas avoir la même distribution. Ainsi, on s'attendait à de mauvaise régression.

3. Est-ce que les classes qui ont été modifiées plus de 10 fois sont mieux commentées que celles qui ont été modifiées moins de 10 fois ?

Afin de répondre à cette question, nous allons prendre nos données et les séparer en deux groupes : un groupe qui contient les classes qui ont été modifiées plus de 10 fois et un groupe qui ont été 10 fois et moins. Une fois ces groupes créer, nous allons faire un test statique. Parce que nous avons des données qui n'ont pas une distribution normale, nous avons pris

le test de Kruskal-Wallis. C'est un test qui utilise les *rank* pour savoir si en moyenne des groupes sont significativement différentes. Notre hypothèse nulle est la qualité des commentaires des classes qui ont été modifiées plus de 10 fois n'est pas de différente des classes qui ont été modifiées moins de 10 fois. Notre hypothèse alternative est que les classe ayant été modifié plus de 10 fois ont une qualité de commentaire supérieur. Afin de vérifier cette hypothèse nous allons regarder la densité des commentaires selon le groupe qu'ils ont été affectés en regardant leur nombre de commit. Nous allons prendre un seuil de 5 % comme valeur significative. Nous voulons regarder un côté de la courbe car l'hypothèse indique que les classes modifiées plus de 10 fois sont mieux commentées. Si on avait dit commenté de façon différente, on aurait eu un « *two-tail test* ».

Le test qui suit nous montre un *p-value* de 0.13 . Ainsi, au seuil de 5%, on ne peut pas refuser l'hypothèse nulle. Il n'est donc pas possible de confirmer que les classes qui ont été modifiées plus de 10 fois sont mieux commentées que celles qui ont été modifiées moins de 10 fois.

Les menaces à la qualité de nos résultats sont l'utilisation de proxy comme variable. La densité de commentaires nous donne une bonne image pour décrire la qualité des commentaires, mais ne reflète pas entièrement la vérité. On suppose que mieux commenté égal plus commenté, hors, pour chacune des classes, il y a probablement un niveau de commentaire parfait qui différente de l'une à l'autre. Aussi, il pourrait y avoir trop de commentaires ou pas assez. Ceci est quelque chose que nous ne regardons pas avec le test.

Nous avons aussi regardé le t-test (par erreur). On supposait que nos données étaient présentées sous forme normal. Dans ce cas, nous aurions pu refuser l'hypothèse nulle et conclure qu'au risque de se tromper une fois sur 20, que les classes qui ont été modifiées plus de 10 fois sont mieux commentées que celles qui ont été modifiées moins de 10 fois. Mais parce que nos données ne sont pas normales, nous ne pouvons pas utiliser cette conclusion.