
TP3

IFT3913

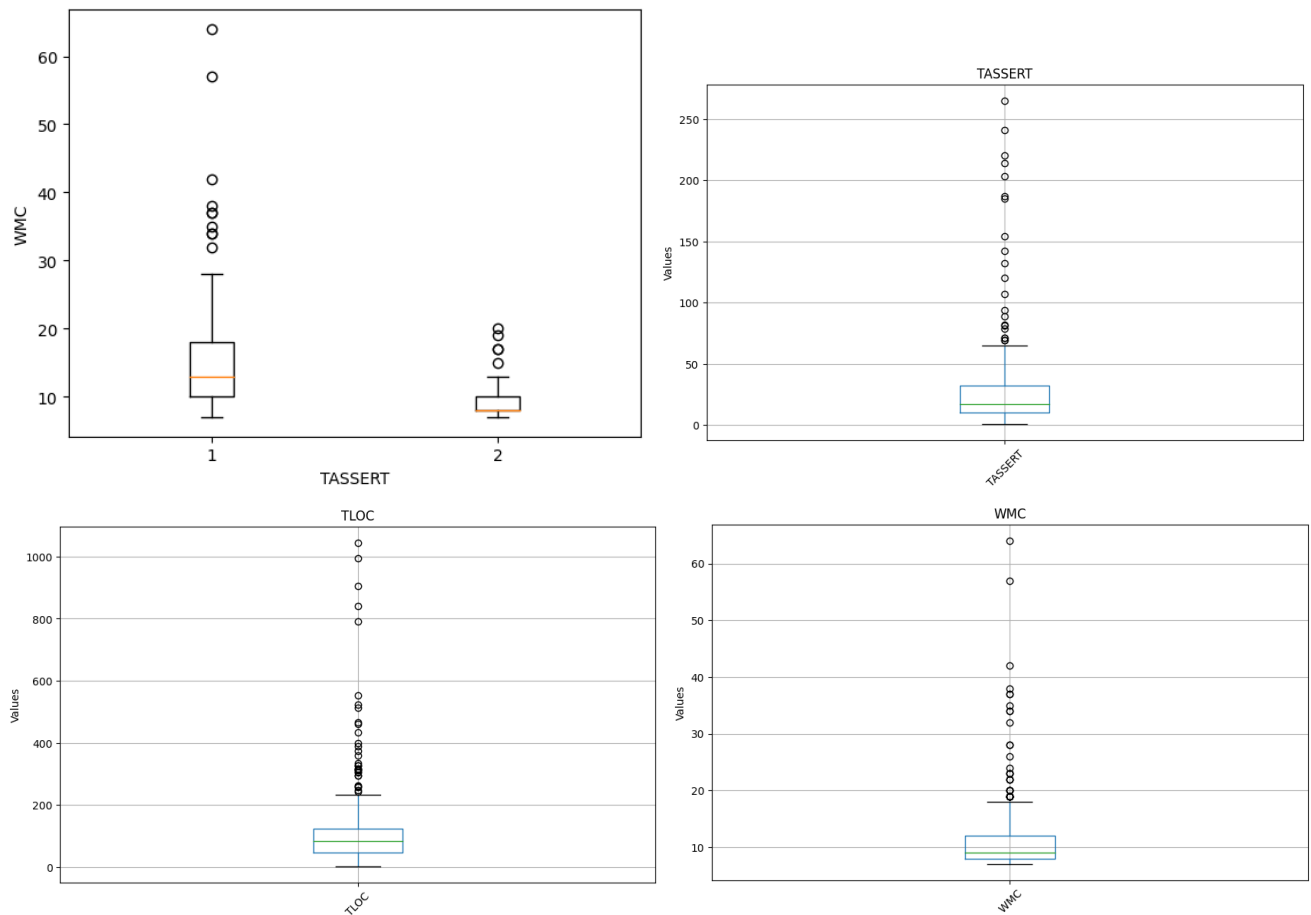
Author

Sen Zeng 20009293

Yuning Sun 20141918

T1

1.0.1 Boîte à Moustache



1.0.2 Métriques

	TLOC	WMC	TASSERT
Minimum	3	7	1
Q1	47.5	8.0	10.0
Median	83.0	9.0	17.0
Mean	115.133	11.584	27.194
Q3	124.5	12.0	32.0
Maximum	1045	64	265
Lower Limit	-68.0	2.0	-23.0
Upper Limit	240.0	18.0	65.0
Number of Outliers	31	32	21

Table 1: Statistics for TLOC, WMC, and TASSERT

Nous pouvons observer que la limite inférieure a une valeur de -68, alors que le minimum est 3. Ceci sous-entend qu'il n'y a pas de 'outliers' inférieur. Nous observons cependant que le Maximum est 1045 alors que la limite supérieur est seulement de 240, ce qui prouve que

les 31 'Outliers' se trouve au niveau supérieur. De plus, nous remarquons que la médiane est inférieur à la moyenne avec un grand écart ce qui montre que la distribution normale est incliné vers la droite.

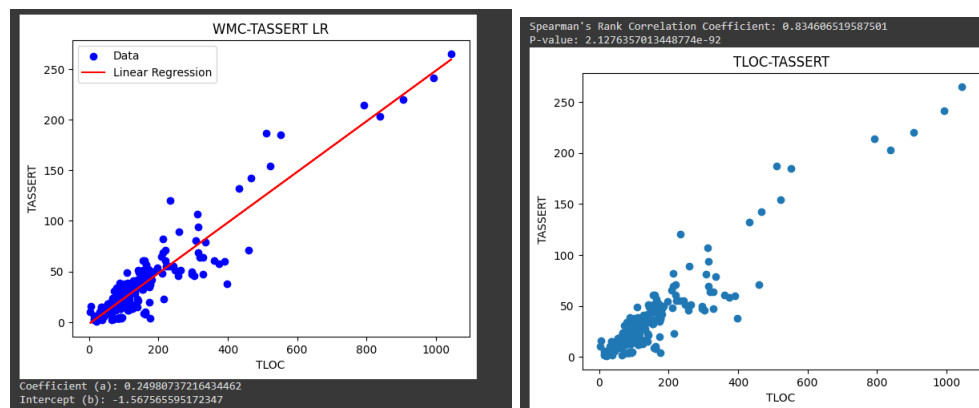
Avec un minimum de 7 et un maximum de 64, nous pouvons voir que la distribution des valeurs est plus petite que celle de TLOC. Avec les données de WMC, nous pourrions voir que malgré sa distribution plus petite il existe tout de même 32 outliers. Comparativement à 31 de la distribution de TLOC. Encore une fois, en comparant le minimum et le lower limite on observe que les outliers se trouve au niveau supérieur. La limite supérieur est de 18, mais le maximum est 64. Encore une fois, la distribution normale est shifté vers la droite puisque la médiane est inférieur à la moyenne.

Dans le cas de TASSERT, nous observons un écart significatif entre la médiane et la moyenne ce qui mène à une distribution normale incliné vers la droite. Avec une valeur de quartile supérieur à 32 et une limite supérieur de 65 on voit que le quartile est loin du maximum qui est de 265, ce qui porte donc à dire que les valeurs aberrantes se trouve au niveau supérieur avec une plus grande importance que WMC.

Les trois métriques indique une tendance de la distribution vers la droite, de plus les données aberrantes sont également prononcé vers le haut ce qui explique que la moyenne est supérieure à la médiane.

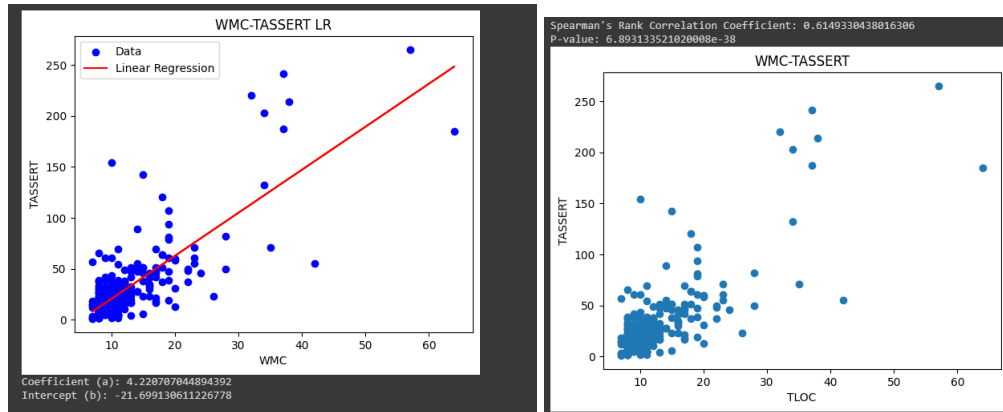
2 T2

2.1 Graphes de corrélation



Selon la régression linéaire, nous avons $y = 0.25x - 1.57$. Une corrélation directe entre TLOC et TASSERT, lorsque TLOC incrémente, TASSERT incrémente avec un taux de 0.25, ajusté par -1.57.

Le coefficient de 0.8346 indique la forte corrélation positive entre TLOC et TASSERT. TLOC augmente TASSERT augmente également ce qui donne la même conclusion que la régression linéaire. La p-value extrêmement faible indique que c'est statistiquement significative, donc il existe une réelle corrélation entre les TLOC et TASSERT.



Selon la régression linéaire, nous avons $y = 4.22x - 21.7$. Une corrélation directe entre WMC et TASSERT, lorsque WMC incrémente, TASSERT incrémente avec un taux de 4.22. Ce taux est plus prononcé que le coefficient entre TLOC et TASSERT, mais il est ajusté par -21.7. Il suffit d'incrémenter WMC avec une petite valeur, TASSERT sera incrémenté 4.22 fois,

Le coefficient de 0.6149 indique une corrélation positive modérée entre WMC et TASSERT, donc il existe également une tendance à ce que TASSERT augmente avec WMC, cependant cette corrélation n'est pas aussi forte que celle entre TLOC et TASSERT. Encore une fois p-value très faible indique que la corrélation est statistiquement significative, donc cette corrélation n'est pas du tout due au hasard.

3 T3

3.1 Choix d'étude

Quasi-expérience

3.2 Énoncé des hypothèses

« les classes qui contiennent plus de 20 assertions sont plus complexes que celles contiennent moins de 20 assertions »

3.3 Définition des variables

3.3.1 Variable d'état

Le nombre d'assertion dans une classe, donné par TASSERT.

3.3.2 Variable dépendant

La complexité de classe, donnée par WMC.

3.4 Conception de quasi-expérience

Les données dans `jfreechart-test-stats.csv` seront séparés en deux parties: les classes avec plus de 20 assertions et celles avec moins de 20. Ensuite, la complexité de ces deux parties sera étudiée et évaluée. Facultativement, le résultat sera comparé avec le coefficient de corrélation Spearman entre WMC et TASSERT pour évaluer l'exactitude du résultat.

3.5 Interprétation et généraliser le résultat

3.5.1 Certains données importants

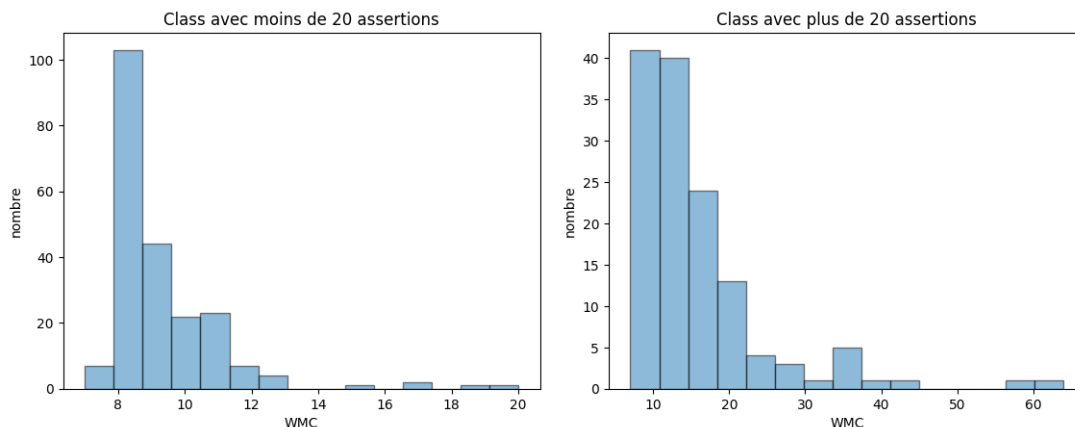
article booktabs

	Groupe 1	Groupe 2
Median	13.0	8.0
Mean	15.496	9.144
Min	7	7

Table 2: Statistics for Groupe 1 and Groupe 2

3.5.2 Boîte à Moustache

3.5.3 Histogramme



3.5.4 Discussion sur le résultat

Même si les exceptions se présentent, les classes qui contiennent plus de 20 assertions tendent d'être plus complexes. Mais il y a une très grande partie de ces classes qui sont situées autour de 10 (WMC), tandis que pour le cas des classes qui contiennent moins de 20 assertions, ce chiffre est entre 8 et 10, qui n'est pas très loin de l'autre partie. En plus, le WMC minimum de ces deux parties sont tous à 7, mais le maximum de la partie avec moins de 20 assertions va jusqu'à 20, ce qui dépasse une grande partie de celle avec plus de 20 assertions.

Il faut noter que cette quasi-expérience est limitée à un certain degré: la largeur des classes n'est pas prise en mesure dans l'hypothèse, ce qui peut affecter énormément le résultat obtenu.