

UNIVERSITÉ DE MONTRÉAL

TP3

PAR

CHARLES-ANTOINE TREMBLAY ET HASSAN ABASS

BAC. EN INFORMATIQUE

FACULTÉ DES ARTS ET SCIENCE

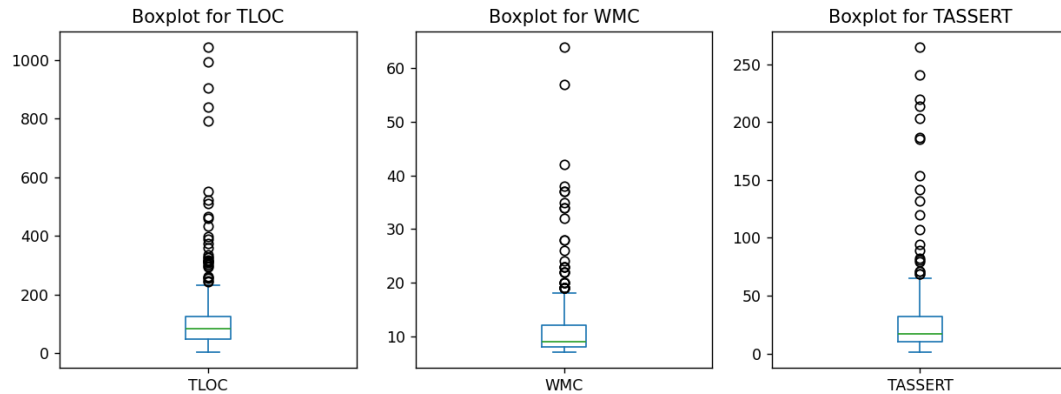
TRAVAIL PRÉSENTÉ À MICHALIS FAMELIS

DANS LE CADRE DU COURS IFT-3913

QUALITÉ DU LOGICIEL ET MÉTRIQUES

NOVEMBRE 2023

T1. Voici les résultats obtenus pour la boîte à moustache :



	TLOC	WMC	TASSERT
count	351.000000	351.000000	351.000000
mean	115.133903	11.584046	27.193732
std	130.874932	6.533271	34.796175
min	3.000000	7.000000	1.000000
25%	47.500000	8.000000	10.000000
50%	83.000000	9.000000	17.000000
75%	124.500000	12.000000	32.000000
max	1045.000000	64.000000	265.000000

TLOC (Total Lines of Code):

Count (nombre d'observations) : Il y a 351 observations dans l'échantillon.

Moyenne (mean) : La moyenne de TLOC est d'environ 115.13 lignes de code par observation.

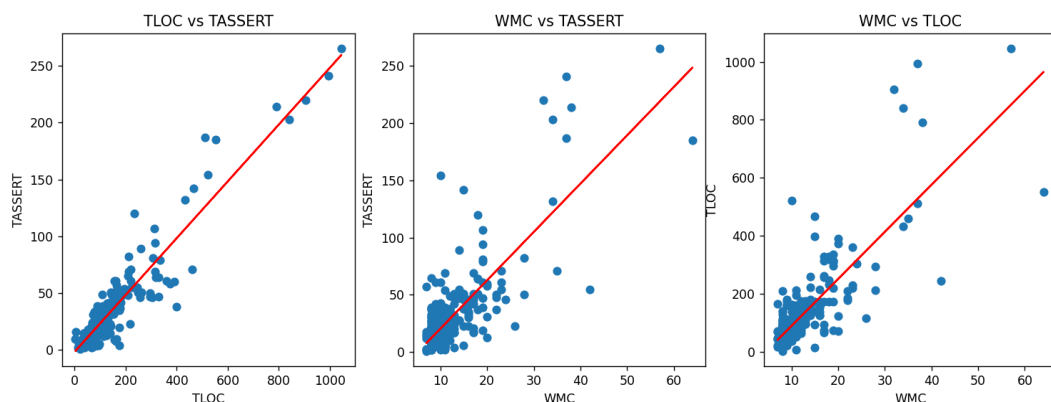
Écart-type (std) : L'écart-type de TLOC est d'environ 130.87, indiquant une dispersion considérable autour de la moyenne.

Minimum et Maximum : La plus petite observation a 3 lignes de code, et la plus grande a 1045 lignes de code.

Quartiles (25%, 50%, 75%) : Ces valeurs divisent les observations en quartiles. Par exemple, 25% des observations ont 47.5 lignes de code ou moins, 50% ont 83 lignes de code ou moins, et 75% ont 124.5 lignes de code ou moins.

Et ceci est le même principe pour les 2 autres métriques.

T2. Voici les résultats pour les corrélations entre TLOC, TASSERT et WMC



```
Corrélation entre TLOC et TASSERT : 0.9395723187973003  
Corrélation entre WMC et TASSERT : 0.7924728265231529  
Corrélation entre WMC et TLOC : 0.8090954919068969
```

Corrélation entre TLOC et TASSERT : Une corrélation de 0.94 entre TLOC et TASSERT indique une forte corrélation positive. Cela signifie que, en général, à mesure que le nombre de lignes de code

augmente, le nombre d'assertions augmente également. En d'autres termes, il y a une relation linéaire positive forte entre la taille d'une classe et le nombre total d'assertions.

Corrélation entre WMC et TASSERT : Une corrélation de 0.79 entre WMC et TASSERT indique également une corrélation positive significative. Cela suggère qu'à mesure que la complexité d'une classe augmente, le nombre d'assertions tend également à augmenter. Cependant, la corrélation n'est pas aussi forte que celle entre TLOC et TASSERT.

Corrélation entre WMC et TLOC : Une corrélation de 0.81 entre WMC et TLOC indique une corrélation positive forte. Cela signifie qu'à mesure que la complexité d'une classe augmente, la taille de la classe en termes de lignes de code a également tendance à augmenter. Cette corrélation renforce la relation entre la complexité et la taille des classes.

Interprétation : Les fortes corrélations positives suggèrent des relations linéaires entre ces métriques dans notre échantillon de données. Les résultats sont cohérents avec les attentes théoriques, car il est logique de s'attendre à ce que des classes plus grandes (en termes de lignes de code ou de complexité des méthodes) aient un plus grand nombre d'assertions.

T3.

Choix d'étude :

- Nous avons choisi le type d'expérience suivant : Expériences contrôlées

Avec les caractéristiques suivantes:

- Population : L'ensemble des classes dans le projet.
- Échantillonnage : Sélection aléatoire ou stratifiée des classes dans le projet.
- Variable indépendante : Le nombre d'assertions dans une classe (plus de 20 ou moins de 20).
- Variable dépendante : Le niveau de complexité de la classe (Évalué avec TLOC, WMC et TASSERT).

Énoncé des hypothèses :

- Hypothèse nulle (H0) : Il n'y a pas de différence significative dans la complexité entre les classes avec plus de 20 assertions et celles avec moins de 20 assertions.
- Hypothèse alternative (H1) : Les classes avec plus de 20 assertions sont plus complexes que celles avec moins de 20 assertions.

Définition des variables :

Variable indépendante :

- Catégorie 1 : Classes avec plus de 20 assertions.
- Catégorie 2 : Classes avec moins de 20 assertions.

Variables dépendantes :

- TLOC
- WMC
- TASSERT

Méthodologie :

Programme T3 nous permet de vérifier en utilisant le test de Student, si les classes ayant + de 20 assertions sont plus complexes que celles en ayant 20 ou moins.

Voici comment on procède:

- Chargement des données : Les données sont chargées à partir du fichier CSV fourni (jfreechart-test-stats.csv).

- Création des groupes : Deux groupes sont créés en fonction du nombre d'assertions dans chaque classe. groupe_plus_20_assertions contient les classes avec plus de 20 assertions, tandis que groupe_moins_20_assertions contient celles avec 20 assertions ou moins.

- Test t de Student : Le test t de Student est utilisé pour comparer les moyennes des métriques TLOC et WMC entre les deux groupes. Le résultat du test est la statistique de test (t_stat_TLOC et t_stat_WMC) et la p-valeur correspondante (p_value_TLOC et p_value_WMC).

- Interprétation des résultats : En fonction du seuil de signification alpha (fixé à 0.05 dans ce cas), le code affiche si la différence observée est statistiquement significative. Si la p-valeur est inférieure à alpha, on rejette l'hypothèse nulle selon laquelle il n'y a pas de différence significative.

Résultats:

Test t de Student pour TLOC : Statistique=11.63357726030889, P-valeur=1.1701139364069334e-26

Test t de Student pour WMC : Statistique=10.058866057637966, P-valeur=4.5415661292742075e-21

Il y a une différence significative dans TLOC entre les deux groupes. ==> complexité plus élevée pour les classes ayant plus de 20 assertions

Il y a une différence significative dans WMC entre les deux groupes. ==> complexité plus élevée pour les classes ayant plus de 20 assertions

Interpolation et généralisation des résultats:

TLOC:

- Stat du test: 11.63
- P-value: $1.17e-26$ (proche de 0)

--> Il y a une différence significative dans TLOC entre les classes avec plus de 20 assertions et celles avec moins de 20 assertions. Donc nous rejetons l'hypothèse H_0 .

WMC:

- Stat du test: 10.06
- P-value: $4.54e-21$ (proche de 0)

--> Il y a une différence significative dans WMC entre les classes avec plus de 20 assertions et celles avec moins de 20 assertions. Donc nous rejetons l'hypothèse H_0 .

Limitations de l'étude

- Choix des métriques : Les métriques TLOC et WMC sont utilisées pour mesurer la complexité, mais elles pourraient ne pas capturer tous les aspects de la complexité logicielle. D'autres métriques pertinentes pourraient être incluses dans des études futures pour une compréhension plus complète de la complexité des classes comme la profondeur de l'héritage ou le nombre de ligne de code par méthode par exemple.

- Généralisation des résultats : Les résultats de cette étude peuvent être spécifiques au contexte des classes logicielles étudiées. Les classes dans d'autres projets ou domaines pourraient réagir différemment. On ne peut donc pas assumer que la généralisation s'applique pour tous les logiciels.

Recommandation future:

- Élargissement des métriques : Pour obtenir une vue plus complète de la complexité logicielle, il pourrait être intéressant d'inclure d'autres métriques pertinentes, telles que la profondeur de l'héritage ou le nombre de ligne de code par méthode.

Applicabilité des résultats:

Les résultats de cette étude peuvent être applicables à d'autres contextes de développement logiciel, mais leur transférabilité dépend de divers facteurs. Ces facteurs incluent la similarité des projets en termes de taille, de domaine et de technologie.

Par exemple, les conclusions peuvent être plus directement pertinentes pour des projets similaires qui partagent les caractéristiques suivantes :

- Taille du projet : Les résultats peuvent être plus applicables à des projets de taille similaire en termes de nombre de lignes de code, de nombre de classes, etc.

- Technologie utilisée : La nature de la technologie sous-jacente (langage de programmation, architecture logicielle, etc.) peut influencer la généralisation des résultats. Des projets utilisant des technologies similaires peuvent bénéficier davantage des conclusions de cette étude.