Catégoriser les réactions face à l'adversité : exploration de nouvelles méthodologies.

SYNTHÈSE DE STAGE - M1 SSD

Garance Malnoë encadrée par Jan Höltge

Resilience Researche Centre — Dalhousie University

Mai - Août 2025

Table des matières

1	Introduction	1
2	Analyse transversale	1
	2.1 Données: Projet RYSE	1
	2.2 Résidus de régression linéaire et statistiques d'influence	2
	2.3 Méthodes de classification des résidus	2
	2.4 Sélection d'une méthode de classification	3
3	Analyse longitudinale	4
	3.1 Méthode spike-and-slab	4
	3.2 Données : Projet LORA	4
	3.3 Résultats	5
4	Discussion	5
\mathbf{R}	Léférences	6
\mathbf{A}	nnexes	7





1 Introduction

La notion de résilience peut-être définie de plusieurs manières [SMa+21], elle peut notamment être vue, lors d'une exposition à une adversité significative (ex : décès d'un proche, divorce, perte d'emploi), comme à la fois la capacité des individus à accéder aux ressources psychologiques, sociales, culturelles et physiques qui soutiennent leur bien-être et comme la capacité individuelle et collective à négocier pour que ces ressources soient fournies et vécues de manière culturellement significative. Pendant longtemps, la résilience a été vue comme un trait inhérent à chaque individu et immuable. Cependant, cette vision a été fortement remise en cause et il existe maintenant un soutien important [UT20] en faveur d'une vision de la résilience comme le résultat de divers processus issus de différents systèmes interconnectés (génétique, environnemental, social, ...).

Une part importante de la recherche quantitative en résilience s'appuie sur la modélisation des trajectoires typiques après l'exposition à une adversité, notamment en observant l'évolution d'indicateur(s) psychologique(s) au cours du temps. Dans la majorité des études, 4 trajectoires ressortent [GHB18] : la résilience (faible détresse psychologique dans la durée) en moyenne 65% des individus, le rétablissement (détresse psychologique prolongée mais qui finit par s'estomper) en moyenne 23%, la détresse chronique (détresse psychologique ne disparaissant pas avec le temps) en moyenne 11%, et la symptomatologie tardive (apparition de symptomes au dessus du seuil de diagnostique après un délais important) en moyenne 10%, bien que parfois d'autres trajectoires puissent également apparaître et les prévalences légèrement changer. Cependant, les résultats de certaines de ces études sont remis en cause [IL16] pour leur utilisation d'hypothèses assez restrictives sur la variance intra-classe alors que lorsque ces hypothèses sont relaxés le nombre de classes trouvées et les prévalences des classes sont significativement affectés.

Le centre de recherche sur la résilience de l'Université Dalhousie, où j'ai réalisé mon stage avec Jan Höltge, s'inscrit dans cette remise en question de la vision de la résilience comme la réponse commune ou normale face à l'adversité, pour la voir comme un évènement hors normes. L'objectif principal de mon stage a été d'explorer différentes méthodes statistiques permettant de classer les individus en trois groupes : résilients, vulnérables et dans la norme. D'abord à partir de données transversales en classifiant les résidus de modèles linéaires avec diverses méthodes puis à partir de données longitudinales en appliquant une méthode s'appuyant sur des statistiques bayésiennes pour sélectionner les effets aléatoires dans des modèles de mixture.

2 Analyse transversale

2.1 Données : Projet RYSE

Les données utilisées pour l'analyse transversale proviennent du projet Resilient Youth in Stressed Environments (RYSE) [Dal]. Ce projet avait pour but d'étudier les facteurs et processus multisystémiques soutenant la résilience chez des adolescents (14–24 ans) vivant dans des communautés dépendantes des industries pétrolière et gazière au Canada et en Afrique du Sud. Lors des collectes de données en 2018, 2019 et 2020, ces communautés ont connu leurs pires récessions historiques, entraînant des taux élevés de chômage, des tensions familiales, ainsi qu'une augmentation de l'anxiété et de la dépression chez les adolescents. Les deux communautés étant très différentes sur les plans culturel et économique, seules les données sud-africaines ont été utilisées pour cette analyse. De plus, afin de maximiser la taille de l'échantillon (entre 300 et 500 participants selon les varaibles considérées), l'analyse transversale a porté sur les données de la première collecte en 2018 [THU23].

De nombreuses variables ont été mesurées lors de la collecte, mais seule une partie a été utilisée ici. Elles ont été traitées selon trois catégories : variables résultat, variables d'adversité, et variables explicatives pour la sélection des méthodes de classification. Pour les questionnaires, les participants disposaient de plusieurs niveaux de réponse (numérotés différemment selon les questionnaires), permettant d'obtenir un score final par addition des réponses.

Nom	Valeurs	Description	Type
Age	14-24	Age du participant au moment de la réponse	Explicative
Sexe	Femme,	Sexe du participant	Explicative
	homme, autre		
WES	9-63	Work Engagement Scale: engagement au travail	Résultat
SES	33-165	School Engagement Scale : engagement à l'école	Résultat
SF-15	0-100	Short-Form health survey 15 : questionnaire sur l'état de santé	Résultat et explicative
BDI-II	0-63	Beck's Depression Inventory : mesure de la sévérité de la dépression	Résultat et adversité
CPTS	0-80	Childhood Post-traumatic stress : mesure du stress lié a des évènements traumatiques de l'enfance	Adversité et explicative
CYRM	28-140	Child and Youth Resilience Measure : mesure de la résilience socio-écologique	Explicative
PoNS	10-40	Perception of Neighborhood Scale : questionnaire sur l'environnement où habite le participant	Explicative
FAS	0-10	Family Adversity Scale : mesure des adversités au sein de la famille	Explicative
BCE	0-10	Benevolent Childhood Experiences : évènement positifs ou protecteurs durant l'enfance	Explicative

Table 1 – Varaiables utilisées issues du projet RYSE

2.2 Résidus de régression linéaire et statistiques d'influence

La première étape de l'analyse a consisté à obtenir une mesure de la résilience pour chaque participant. Une méthode couramment utilisée et dont l'utilisation est validée dans ce contexte [CHC22] est la résidualisation. Elle consiste à modéliser un résultat (par exemple un score de bien-être) en fonction d'une mesure d'adversité, puis à utiliser les résidus de cette régression comme indicateur de résilience. Autrement dit, on mesure à quel point un individu s'écarte de ce qui est prédit par le modèle. Les individus ayant des résidus proches de zéro (donc proches de la droite de régression) sont considérés comme "dans la norme", tandis que ceux avec des résidus éloignés de zéro sont classés comme résilients ou vulnérables, selon le signe du résidu et la sens du lien entre adversité et résultat.

Afin de limiter l'influence d'individus ayant un impact disproportionné sur le modèle, des statistiques d'influence ont été utilisées. Conformément aux recommandations de [HU22], les individus avec une distance de Cook strictement supérieure à 4/n (où n est le nombre total de participants) ont été exclus de la construction du modèle, mais ont été inclus dans le calcul des résidus finaux. Voir l'Annexe A (4) pour un exemple.

2.3 Méthodes de classification des résidus

La seconde étape a consisté à explorer différentes méthodes pour classer les individus en trois groupes : résilient, vulnérable et dans la norme.

- Une approche **basique** considérant tout individu avec un résidu non nul comme résilient ou vulnérable selon le signe. Elle a été utilisée uniquement comme point de comparaison.
- L'utilisation d'intervalles de prédiction (valeurs probables pour une observation individuelle) et d'intervalles de confiance (valeurs probables pour la moyenne). Les individus à l'intérieur de l'intervalle sont considérés comme normaux, les autres comme résilients ou vulnérables. Des seuils à 95% et 99% (confiance) et à 50%, 60% et 75% (prédiction) ont été testés.
- L'utilisation d'intervalles de crédibilité bayésiens (probabilité sur la moyenne). Même logique que précédemment, avec des intervalles de 50%, 60%, 70%, 80%, 90%, 99% et 99.9%.
- Une méthode par **quantiles**, classant comme résilients ou vulnérables les individus dans les n% extrêmes des résidus (n variant de 5 à 35 par pas de 5).
- Une approche basée sur l'écart-type par niveau d'adversité : pour chaque plage d'un score d'adversité (par exemple les seuils cliniques du BDI-II), on calcule l'écart-type des résidus correspondant

à un point dont l'adversité se trouve dans cette plage, et on considère comme atypiques (résilients ou vulnérables) les individus situés en dehors de 0.5, 1 ou 2 écarts-types.

• L'application de **k-means** avec trois clusters. Le groupe avec la moyenne intermédiaire est considéré comme "normal", les autres comme résilient ou vulnérable.

Ces différentes méthodes permettent d'obtenir des classes de tailles variables. Par exemple, la méthode par quantiles produit un groupe normal très large lorsque l'on choisit un faible pourcentage, tandis que les intervalles de crédibilité tendent à définir des groupes résilient et vulnérable plus étendus, rendant le groupe normal bien plus restreint. De plus, certaines méthodes, comme celle basée sur l'écart-type ou les intervalles de confiance, permettent d'ajuster les seuils en fonction de l'adversité, ce qui introduit une forme d'adaptativité dans la classification. Voir l'annexe B (4) pour quelques visualisations.

2.4 Sélection d'une méthode de classification

Une question se pose alors : quelle méthode de classification privilégier ? Il n'existe pas de seuil universellement accepté définissant la taille ou la forme que devrait prendre le groupe normal, et le choix de la méthode dépend en réalité surtout des objectifs poursuivis. Dans le contexte de la recherche en résilience, deux critères principaux ont guidé l'évaluation des méthodes : la capacité à faire émerger des profils de résilience et la capacité à prédire l'appartenance à un groupe à partir de variables explicatives.

2.4.1 Capacité à faire émerger des profils de résilience : LPA et LCA

Une hypothèse courante en résilience postule de l'existence de profils différenciés, que, par exemple, certains individus présenteraient peu de symptômes dépressifs mais une mauvaise santé physique, tandis que d'autres, très affectés psychologiquement, maintiendraient un fort engagement scolaire. Pour explorer cette idée, une Latent Profile Analysis (LPA) a été appliquée à trois dimensions de la résilience (santé physique SF-15, engagement scolaire SES, santé mentale BDI-II avec le stress CPTS comme adversité), en fixant les résidus à zéro pour les individus classés comme normaux. Les modèles ont été comparés via l'AIC, le BIC et l'entropie.

Les meilleurs scores AIC et BIC ont été systématiquement obtenus pour des modèles avec un groupe normal très large et un grand nombre de classes (plus de 8). Cependant, les classes obtenues manquent de lisibilité, suggérant que la forte concentration des résidus autour de zéro, due à la transformation, pousse l'algorithme à créer des classes sur des variations mineures et donc peu pertinentes.

Pour contourner ce biais, une Latent Class Analysis (LCA) a aussi été testée, cette méthode utilisant des variables catégorielles plutôt que continues permettant ainsi d'utiliser directement les classifications obtenues. Les résultats obtenus sont similaires, voire plus marqués : les meilleurs AIC et BIC favorisent encore les méthodes avec un groupe normal aussi large que possible. Cette performance semble surtout refléter une simplification des données, plus facile à modéliser, sans garantir une réelle interprétabilité.

2.4.2 Capacité à prédire l'appartenance à un groupe : classifieurs

Un enjeu central en recherche en résilience est d'identifier les facteurs associés à la résilience. Les méthodes de classification ont ainsi été comparées selon leur capacité à construire un classifieur de type random forest performant pour prédire la classe des résidus de l'engagement scolaire ou professionnel (SES ou WES selon l'âge), en prenant la dépression (BDI-II) comme variable d'adversité et en utilisant l'âge, le sexe, CYRM, SF-15, PoNS, CPTS, FAS et BCE comme variables prédictives.

La comparaison des classifieurs est rendue complexe par le fait que chaque méthode produit un jeu d'entrainement différent. Deux indicateurs ont été d'abord été observés : l'accuracy sur les données d'entraînement et l'écart d'accuracy avec le modèle nul attribuant toujours la classe majoritaire. On observe que, lorsque la taille du groupe normal diminue, l'accuracy globale baisse, mais que l'écart avec le modèle nul augmente, suggérant une meilleure détection des classes résilientes et vulnérables et l'analyse des métriques par classe

(précision et rappel) confirmant cette tendance.

Cela conduit à interroger les priorités de la recherche en résilience. Il a été choisi de focaliser l'attention sur la précision et le rappel des résilients, tout en maintenant un groupe normal majoritaire. En appliquant ces critères, seules quelques méthodes restent valides : k-means, les quantiles entre 15% et 30%, les seuils à 0,5 et 1 écart-type, ainsi que les intervalles de prédiction à 60% et 75%. Parmi elles, le seuil à 30% de quantiles apparaît comme la méthode la plus performante.

3 Analyse longitudinale

Les modèles mixtes, couramment utilisés en psychologie, permettent de modéliser les variations interindividuelles à l'aide d'effets aléatoires spécifiques à chaque participant. La méthode *spike-and-slab* proposée par Rodriguez, Williams et Rast [RWR24] permet d'évaluer si ces effets diffèrent significativement de l'effet moyen. L'idée est donc d'appliquer cette approche aux résidus de régressions linéaires entre une variable résultat et une mesure d'adversité répétées plusieurs fois dans le temps, afin d'identifier les individus dont la réponse face à l'adversité s'écarte régulièrement de la réponse commune.

3.1 Méthode spike-and-slab

La méthode spike-and-slab est une approche bayésienne modélisant chaque effet aléatoire comme un mélange de deux composantes : une spike (probabilité concentrée en zéro, représentant un effet aléatoire nul) et une slab (distribution plus large autorisant des effets non nuls). Lors d'itérations MCMC, un indicateur binaire, suivant d'une loi de Bernoulli, décide si l'effet provient de la spike ou de la slab. La proportion d'itérations dans la slab constitue alors la probabilité d'inclusion postérieure (PIP), estimant la probabilité qu'un individu ait un effet aléatoire différent de 0. L'indicateur binaire dépend d'une probabilité a priori d'inclusion dans la slab, fixée entre 0 et 1. Une valeur de 0.5 indique une absence d'hypothèse préférentielle; des valeurs plus proches de 1 traduisent une croyance initiale plus forte de la présence d'un effet aléatoire.

3.2 Données : Projet LORA

Les données utilisées proviennent du projet **LOngitudinal Resilience Assessment (LORA)** [ACh+21], visant à étudier les mécanismes de résilience face aux stress du quotidien à travers une approche quantitative. L'échantillon initial comprend 1191 adultes allemands, âgés de 18 à 50 ans à l'inclusion. L'étude s'étend sur cinq ans, avec des questionnaires en ligne tous les trois mois et des évaluations approfondies tous les 18 mois.

Bien que de nombreuses variables aient été collectées, seules trois ont été retenues ici, correspondant à des scores issus de questionnaires. Chaque score est obtenu en additionnant les réponses à des items sur des échelles ordinales.

Nom	Valeurs	Description	\mathbf{Type}
PSS	14-24	Perceived Stress Scale : mesure du stress perçu	Adversité
DH	14-24	Daily Hassles: Tracas quotidiens	Adversité
$_{ m GHQ}$	Variable	General Health Questionnaire : outil de dépistage du	Résultat
		bien-être psychologique	

Table 2 – Variables utilisées issues du projet LORA

Pour limiter les biais liés aux données manquantes, l'analyse se concentre sur les 515 participants ayant complété l'ensemble des 12 premières mesures.

L'analyse a porté sur les modèles mixtes $GHQ \sim DH$ et $GHQ \sim PSS$ avec un effet aléatoire sur l'intercept, afin de détecter les écarts moyens des individus par rapport au niveau moyen de résilience.

3.3 Résultats

La méthode spike-and-slab s'est bien révélée applicable aux données, produisant des résultats cohérents, voir Annexe C (4).

Les deux modèles présentent des tendances similaires : plus le seuil est élevé, plus le groupe normal (effet aléatoire nul) devient dominant. Par exemple, avec un seuil de 95%, 13% des individus sont classés comme anormaux dans le modèle $GHQ \sim PSS$ (5% résilients, 8% vulnérables), et 19% dans le modèle $GHQ \sim DH$ (8,5% résilients, 10,5% vulnérables).

Comme pour l'analyse transversale, le principal enjeu réside dans le choix du seuil de PIP au-delà duquel un effet est considéré comme significatif. Ce seuil influence directement la taille des groupes obtenus.

4 Discussion

L'approche transversale visait surtout à proposer un panorama des approches possible et d'explorer leur impact sur la formation des trois groupes plutôt que d'établir une nouvelle méthode de référence. Cependant, les résultats suggèrent que la méthode fondée sur les quantiles à 30% constitue un compromis intéressant, mais ce constat demanderait à être répliqué sur d'autres données pour en évaluer la robustesse.

Du côté longitudinal, la démarche était également avant tout exploratoire, avec l'idée de tester l'applicabilité de la méthode spike-and-slab à la classification d'individus lorsque l'on dispose de plusieurs mesures dans le temps. Les résultats sont intéressants mais la problématique du choix du seuil reste la même que dans le cas transversal. Un seuil bas permet d'identifier plus d'individus atypiques, mais au risque d'avoir une vision de la résilience comme quelque chose de plus commun. À l'inverse, un seuil élevé garantit de s'intéresser aux individu réellement hors-normes mais au risque d'exclure des cas potentiellement toujours pertinents. Dans le cas d'une application, ce choix devrait être défini selon les priorités de recherche ou les contraintes d'intervention. Une piste d'exploration pertinente serait d'examiner l'impact de la modification des priors selon la classification observée à chaque mesure (par exemple, un individu classé comme normal à chaque vague pourrait avoir un prior plus proche autour de zéro qu'un individu systématiquement atypique). Cette approche soulève néanmoins plusieurs contraintes : la nécessité de disposer de données longitudinales riches, dont la collecte est coûteuse et longue, ainsi que des temps de calcul importants liés à l'inférence bayésienne et aux itérations MCMC.

Ainsi, la fin du stage sera consacrée à l'exploration de l'impact de la modification des priors sur les résultats dans le cadre de l'analyse longitudinale, la rédaction d'un article en collaboration avec mon tuteur de stage. Et, si le temps le permet, une exploration complémentaire sera menée à travers l'application de méthodes d'analyse de réseaux aux données du projet LORA.

Ce stage, en lui-même, a un impact écologique relativement faible, limité à l'utilisation de mon ordinateur portable personnel et de GitHub. Cependant, bien que tous les trajets entre mon domicile et le lieu de travail aient été (et seront) effectués à pied, venir faire un stage à Halifax, au Canada, est loin d'être anodin sur le plan environnemental : 3,2 tonnes de CO2 réparties sur six vols, soit presque autant que mon empreinte carbone annuelle, alors même qu'aucune étape du stage ne nécessitait ma présence sur place. Cela soulève une vraie question sur le sens et la valeur ajoutée d'un tel déplacement. Il peut peut-être, en partie, se justifier par l'utilité sociale de la recherche en résilience, qui vise à mieux comprendre les mécanismes de protection et de vulnérabilité face à l'adversité. Cette compréhension a des implications importantes pour les politiques publiques de santé mentale : certains travaux influents, concluant que la majorité des personnes exposées à des traumas étaient peu affectées, ont parfois été interprétés comme des arguments pour limiter les interventions. Or, comme exposé en introduction, cette lecture est aujourd'hui remise en question. En choisissant de considérer la résilience et la vulnérabilité comme des écarts à la moyenne, ce stage contribue, très modestement, à cette réflexion critique.

Références

- [IL16] Frank J Infurna et Suniya S Luthar. « Resilience to Major Life Stressors Is Not as Common as Thought ». In: Perspect Psychol Science 11 (2016), p. 175-194. DOI: 10.1177/1745691615621271...
- [GHB18] Isaac Galatzer-Levy, Sandy H. Huang et George Bonanno. « Trajectories of resilience and dysfunction following potential trauma: A review and statistical evaluation ». In: Clinical Psychology Review 63 (2018), p. 41-55. DOI: 10.1016/j.cpr.2018.05.008.
- [UT20] Michael UNGAR et Linda THERON. « Resilience and mental health: how multisystemic processes contribute to positive outcomes ». In: Lancet Psychiatry 7 (2020), p. 441-448. DOI: 10.1016/.
- [ACh+21] A.Chmitorz et al. « Longitudinal determination of resilience in humans to identify mechanisms of resilience to modern-life stressors: the longitudinal resilience assessment (LORA) study ». In: European Archives of Psychiatry and Clinical Neuroscience 271 (2021), p. 1035-1051. DOI: 10.1007/s00406-020-01159-2.
- [SMa+21] Ann S.MASTEN et al. « Resilience in Development and Psychopathology: Multisystem Perspectives ». In: Annual Review of Clinical Psychology (2021). DOI: 10.1146/annurev-clinpsy-081219-120307.
- [CHC22] Stephanie Cahill, Reinmar Hager et Tarani Chandola. « The validity of the residuals approach to measuring resilience to adverse childhood experiences ». In: *Child and Adolescent Psychiatry and Mental Health* 16 (2022). Doi: 10.1186/s13034-022-00449-y.
- [HU22] Jan HÖLTGE et Michael UNGAR. « Quantifying Resilience as an Outcome : Advancing the Residual Approach with Influence Statistics to Derive More Adequate Thresholds of Resilience ». In : Adversity and Resilience Science 3 (2022), p. 381-390. DOI : 10.1007/s42844-022-00078-6.
- [THU23] Linda Theron, Jan Höltge et Michael Ungar. « Multisystemic supports and adolescent resilience to depression over time: A South African mixed methods study ». In: Development and Psychopathology 35 (2023), p. 2365-2383. DOI: 10.1017/S0954579423000494.
- [RWR24] Josue E. RODRIGUEZ, Donald R. WILLIAMS et Philippe RAST. « Who Is and Is Not "Average"? Random Effects Selection With Spike-and-Slab Priors ». In: *Psychological Methods* 29 (2024), p. 117-136. DOI: 10.1037/met0000535.
- [Dal] DALHOUSIE. RYSE Project. https://ryseproject.org/. Visité le 18/06/25.

Annexe A

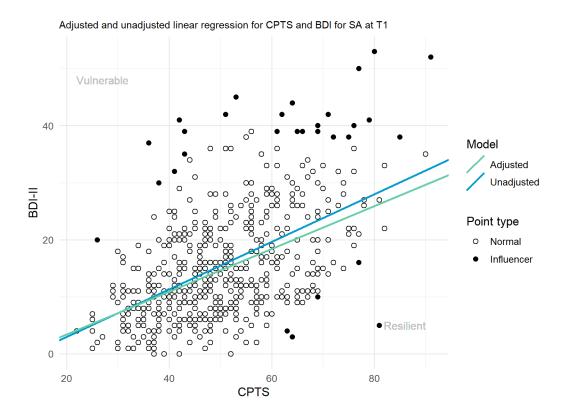


FIGURE 1 – Régressions linéaires ajusté et non-ajusté de la variable BDI-II selon la variable CPTS

Annexe B

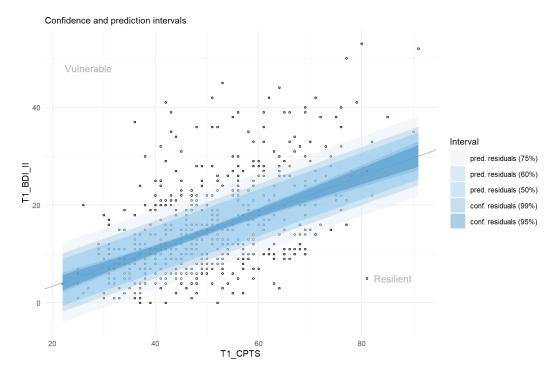


FIGURE 2 – Intervalles de prédiction (75%,60%,50%) et de confiance (99%,90%) pour la régression de BDI-II selon CPTS

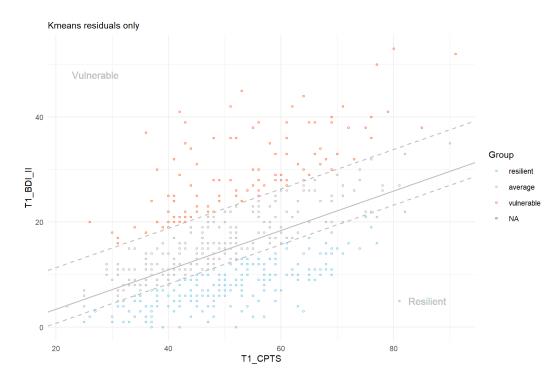
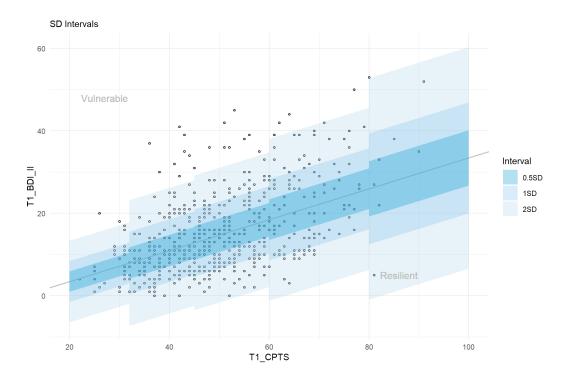


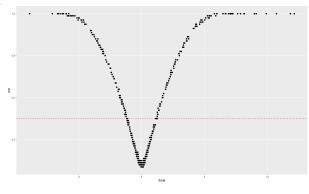
FIGURE 3 – Classification obtenue avec l'algorithme k-means pour la régression de BDI-II selon CPTS



 $\mbox{Figure $4-$Intervalles pour la méthode des écart-types pour les valeurs 0.5, 1 et 2 pour la régression de BDI-II selon CPTS } \\$

Annexe C $\label{eq:continuous}$ General Health Questionnaire and Perceived Stress Scale

PIP	pct_non_average	pct_resilient	pct_vulnerable
0.50	43.69	20.78	22.91
0.55	39.81	18.64	21.17
0.60	35.73	15.92	19.81
0.65	33.40	14.56	18.83
0.70	30.29	13.20	17.09
0.75	27.18	12.04	15.15
0.80	24.08	10.87	13.20
0.85	19.61	8.16	11.46
0.90	16.31	6.60	9.71
0.95	13.59	5.63	7.96
1.00	0.00	0.00	0.00



General Health Questionnaire and Daily Hassles

0.50 51.46 27.96	23.50
0.55 47.38 26.41	20.97
0.60 43.11 24.27	18.83
0.65 40.19 22.72	17.48
0.70 37.09 20.58	16.50
0.75 35.15 19.61	15.53
0.80 32.04 17.48	14.56
0.85 28.74 14.95	13.79
0.90 24.85 12.23	12.62
0.95 19.03 8.54	10.49
1.00 0.00 0.00	0.00

