1. Regroupement par année et analyse de distribution.

Après analyse du jeu de données, on constate qu’il ne semble pas y avoir de valeur aberrante. La moyenne semble donc être un bon choix pour la standardisation. On décide alors de grouper les données par année, tous pays confondus :

Une image contenant texte, capture d’écran, nombre, Police

Description générée automatiquement

Puis on étudie la distribution des variables pour constater qu’il y a bien moins de valeurs en 2005 que pour les autres années. Cela nous conduit à supprimer les données de 2005 pour l’analyse globale par année.

Une image contenant texte, capture d’écran, Police, conception

Description générée automatiquement



1. Analyse des corrélations

On analyse ensuite les corrélations pour le jeu de données global (df\_global2) et le jeu de données par année :

Une image contenant texte, Police, capture d’écran

Description générée automatiquement

Une image contenant capture d’écran, texte, carré, Caractère coloré

Description générée automatiquement

On constate les corrélations semblent globalement identiques entre les deux jeux de données. Cependant, on constate également une opposition entre les corrélations pour la variable : "Negative affect".

En effet, les corrélations observées pour cette variable dans le second graphique semblent contrintuitives. C'est d'autant plus marqué que l’on n’observe pas la même chose avec le premier jeu de données. Plus précisément, on voit les points suivants dans le deuxième graphique :

* Log GDP per capita et Negative affect : 0.67
* Healthy life exceptancy at birth et Negative affect : 0.76
* Freedom to make life choices et Negative affect : 0.86
* Perceptions of corruption : -0.76

Cela indiquerait donc que l'amélioration de variables "objectives", comme le PIB et l’espérance de vie en bonne santé, ne serait pas corrélée avec les sentiments négatifs mais plutôt qu'elle y serait négativement corrélée.

De même, cela laisserait penser que plus la corruption augmente, moins celle-ci semble peser sur le moral de la population comme si les gens la considéreraient comme une forme de normalité. Ce point mériterait une analyse plus poussée.

Sachant cela, on réaffiche les matrices de corrélations :

Une image contenant texte, capture d’écran, nombre, Police

Description générée automatiquement

En dehors de la variable de négativité, on remarque les points suivants au niveau de la matrice de corrélation du jeu par année:

* 1. Fortes corrélations positives :
* **Healthy life expectancy et Log GDP per capita** : Avec une corrélation de 0.969, cela suggère qu'il existe une très forte relation positive entre l'espérance de vie en bonne santé à la naissance et le PIB par habitant.

Les pays avec un PIB par habitant plus élevé ont généralement une espérance de vie en bonne santé plus longue.

Cela peut être dû à de meilleurs soins de santé, à une meilleure nutrition et à d'autres facteurs sociaux et économiques dans les pays plus riches.

* **Life Ladder et Log GDP per capita** : Le coefficient de corrélation est de 0.866, ce qui indique une forte corrélation positive.

Cela signifie que, en général, lorsque le PIB par habitant augmente, le Life Ladder tend également à augmenter.

C'est cohérent avec l'idée que des niveaux de revenu plus élevés sont souvent associés à un bien-être plus élevé.

Cela reflète souvent une meilleure qualité de vie, une sécurité accrue, et d'autres avantages associés à une économie plus forte.

* **Positive affect et Social support** : cette corrélation (0.746) indique que lorsque les personnes ressentent un soutien social plus élevé, elles ont également tendance à éprouver des affects positifs plus fréquemment.

Le soutien social peut inclure des relations familiales et amicales solides, des interactions sociales régulières, etc.

* **Freedom to make life choices et Log GDP per capita** : Une corrélation de 0.733 montre que la liberté de faire des choix de vie tend à augmenter avec le PIB par habitant.
  1. Fortes corrélations négatives :
* **Freedom to make life choices et Perceptions of corruption** : Cette forte corrélation négative (-0.904) suggère que dans les endroits où les perceptions de corruption sont élevées, la liberté de faire des choix de vie est généralement réduite.

La corruption peut entraver la liberté individuelle et la confiance dans les institutions.

* **Healthy life expectancy at birth et Perceptions of corruption** (-0.831) : les pays avec des perceptions élevées de corruption ont tendance à avoir une espérance de vie en bonne santé plus courte.

La corruption peut affecter la qualité des soins de santé, l'accès à des services médicaux de qualité et d'autres facteurs importants pour la santé.

* **Log GDP per capita et Perceptions of corruption** (-0.780) : une économie plus forte (mesurée par le PIB par habitant) est généralement associée à des perceptions plus faibles de la corruption. Les pays avec un PIB par habitant plus élevé peuvent avoir des institutions plus solides, une meilleure gouvernance et une réglementation plus stricte pour contrer la corruption.

Si l'on exclut la perception de la corruption de l'analyse, les corrélations négatives principales sont les suivantes :

* **Freedom to make life choices et Generosity** (-0.524) : cette corrélation négative indique que lorsqu'il y a une plus grande liberté de faire des choix de vie, la générosité tend à être plus faible.

C'est une observation intéressante et pourrait suggérer que dans les sociétés où les individus ont plus de liberté personnelle, il pourrait y avoir moins d'incitatifs ou de pressions sociales pour des actes généreux.

Cependant, il est important de noter que la causalité n'est pas claire, et d'autres facteurs pourraient influencer cette relation.

* **Healthy life expectancy at birth et Generosity** (-0.455) : les pays ayant une espérance de vie en bonne santé plus longue semblent avoir des niveaux de générosité plus bas. Cela pourrait être dû à divers facteurs socio-économiques, comme une plus grande sécurité économique réduisant la nécessité de dons charitables, mais cela nécessite une investigation plus approfondie.
* **Log GDP per capita et Generosity** (-0.333) : il semble y avoir une corrélation négative modérée entre le PIB par habitant et la générosité. Cela pourrait suggérer que les pays plus riches (en termes de PIB par habitant) sont moins généreux.

Cependant, cela pourrait aussi être dû à la manière dont la "générosité" est mesurée, car les dons charitables dans les pays riches pourraient prendre des formes différentes ou être influencés par des facteurs fiscaux.

* **Life Ladder et Generosity** (-0.323) : cette corrélation suggère que le bien-être ou le bonheur (mesuré par Life Ladder) est légèrement négativement corrélé avec la générosité.

Cela pourrait signifier que dans les régions où les gens sont plus heureux ou satisfaits de leur vie, il y a moins d'actes de générosité. Encore une fois, cela pourrait être lié à la manière dont la générosité est mesurée ou à des facteurs culturels ou socio-économiques.

En résumé, ces corrélations suggèrent que le PIB par habitant, l'espérance de vie en bonne santé, et la liberté de faire des choix de vie sont étroitement liés et influencent positivement le bien-être des individus.

En revanche, la perception de la corruption semble avoir des effets négatifs sur ces mêmes facteurs.

* 1. Normalisation

Une image contenant texte, capture d’écran, Police

Description générée automatiquement

Une image contenant texte, capture d’écran, nombre, Police

Description générée automatiquement

On affiche ensuite ces données sous forme de graphique :

Une image contenant texte, Police, capture d’écran

Description générée automatiquement

Une image contenant ligne, diagramme, Tracé, texte

Description générée automatiquement

Voici quelques observations que l'on peut tirer de ce graphique :

* **Life Ladder** : il semble y avoir une tendance générale à la hausse du Life Ladder au fil des ans, avec quelques baisses intermédiaires. La valeur atteint son maximum en 2020. Cela indique donc que la perception de la qualité de vie a suivi une tendance haussière depuis 2006.
* **Log GDP per capita** : cette variable montre également une tendance à la hausse, culminant en 2020. Il y a une forte corrélation entre le PIB par habitant et le Life Ladder, ce qui est cohérent avec l'idée que les pays avec des revenus plus élevés ont généralement un plus grand bien-être.
* **Social support** : après un pic en 2006, le soutien social a diminué jusqu'en 2008, puis a fluctué au fil des ans, atteignant un autre pic en 2020.
* **Healthy life expectancy at birth** : cette variable montre une tendance claire à la hausse, suggérant que l'espérance de vie en bonne santé s'améliore au fil du temps.
* **Freedom to make life choices** : il y a une tendance générale à la hausse avec quelques fluctuations.
* **Generosity** : Cette variable a connu des fluctuations significatives au fil des ans. Elle atteint sa valeur la plus basse en 2018.
* **Perceptions of corruption** : après avoir atteint un pic en 2007, les perceptions de la corruption ont généralement diminué, atteignant leur point le plus bas en 2020.
* **Positive affect** : il y a eu des fluctuations au fil des ans, avec un pic notable en 2020.
* **Negative affect** : Cette variable montre une tendance générale à la hausse, atteignant son maximum en 2020.

Voici des réflexions globales réalisées à partir des points précédents :

Les années récentes, en particulier 2020, semblent avoir des valeurs maximales pour plusieurs variables, notamment Life Ladder, Log GDP per capita, Social support, Healthy life expectancy, Freedom to make life choices et Positive affect.

On peut supposer que la diminution globale constatée en 2021 est due à la pandémie de COVID-19 et les confinements. On a pu par exemple constater à cette occasion une envolée des démissions ou "big quit" (https://fr.wikipedia.org/wiki/Grande\_D%C3%A9mission).

Cependant, il serait intéressant de mener une étude sur la période 2022 - 2023 pour vérifier la tendance.

Concernant la perception de la corruption, on constate une tendance globale à la baisse entre 2006 (apogée) et 2020 (étiage). Étant donné ce qui a été constaté concernant les corrélations, cette baisse s'explique en grande partie par l'amélioration des variables qui lui sont fortement négativement corrélées : la liberté de choix, l'espérance de vie en bonne santé et le PIB.

En outre, on s'aperçoit que les périodes où elle est le plus forte sont des périodes de crise : 2008 (crise des *Subprime*) et 2020 (COVID).

On constate une hausse progressive de la négativité sur toute la période. La positivité augmente également mais de manière moins directe. Comme vu précédemment, il n'y a pas de corrélation forte entre ces deux variables. Toutefois, nous n'avons pas trouvé de point permettant d'expliquer cette tendance.

Enfin, on constate que la générosité est la variable qui affiche la plus grande variation sur la période bien que tendanciellement en baisse entre 2006 et 2020. Apparaissant comme corrélée négativement à la majorité des autres variables, il serait intéressant de comprendre quelles sont les éléments qui l'influencent le plus.

On pourrait émettre l'hypothèse que les crises comme celle de 2008 impactent fortement les budgets alloués aux dons aux associations par exemple. Ainsi, les gens auraient eu moins tendance à donner alors.