

Le bonheur dans le monde

Maud AGNES, Evann BONNAVENTURE & Charlotte SARTER

19/06/2022

Introduction

Ce projet est réalisé dans le cadre de la matière **Visualisation de données** enseignée à l'Université de Technologie de Troyes.

Notre analyse porte sur le **World Happiness Report**, une enquête annuelle qui récolte des données du monde entier pour indiquer **comment les gens évaluent leur propre vie** dans plus de 150 pays du globe. Le but de notre projet est de concevoir les meilleures visualisations possibles pour analyser les données du World Happiness Report et pour mettre en valeur les informations qui en découlent.



Lien des datasets

- [World Happiness Report](#) (de 2015 à 2019)
- [World Happiness Report](#) (2020)
- [World Happiness Report](#) (2021)
- [Healthy Lifestyle Cities Report 2021](#)

Le dernier dataset apporte des **indicateurs supplémentaires** sur ce qui est susceptible de provoquer le bonheur des populations (prix de la vie, soleil, obésité...). Cependant, il ne comporte uniquement les données de 2021, ainsi nous pourrons seulement le comparer avec le dataset qui comporte les données du World Happiness Report de 2021. Cela nous permettra d'apporter une nouvelle dimension à notre analyse.

Données

Le jeu de données que nous avons choisi d'étudier est issu d'une **enquête de référence sur l'état du bonheur dans le monde**. Il comporte des données **de 2015 à 2021**, et classe **155 pays** selon leur niveau de bonheur. Pour établir ce classement, l'ONU demande à un **échantillon d'environ 1 000 personnes** de "répondre à une série de questions sur la perception de leur qualité de vie sur une échelle de 1 à 10" (échelle de Cantril). Il permet d'établir une corrélation entre différents critères (liberté, corruption, cadre de vie...) et le niveau de bonheur qui en découle.

Les jeux de données initiaux comportent 13 variables :

Variable	Class	Description
country	character	Nom du pays
region	character	Région à laquelle le pays appartient
hapiness rank	integer	Classement du pays sur la base du score du bonheur
hapiness score	double	Un indicateur mesuré chaque année en posant aux personnes de l'échantillon la question suivante : "Comment évaluez-vous votre bonheur sur une échelle de 0 à 10 où 10 est le plus heureux"
lower confidence interval	double	Intervalle de confiance inférieur du score de bonheur
upper confidence interval	double	Intervalle de confiance supérieur du score de bonheur
economy (GPD per capita)	double	La mesure dans laquelle le PIB contribue au calcul du score du bonheur
family	double	La mesure dans laquelle la famille contribue au calcul du score du bonheur
health (life expectancy)	double	La mesure dans laquelle l'espérance de vie a contribué au calcul du score du bonheur
freedom	double	La mesure dans laquelle la liberté a contribué au calcul du score du bonheur
trust (governement corruption)	double	La mesure dans laquelle la perception de la corruption contribue au score de bonheur
generosity	double	La mesure dans laquelle la générosité a contribué au calcul du score de bonheur
dystopia residual	double	"Résidu" correspondant à l'écart entre le modèle théorique et la réalité, auquel on ajoute un score de dystopie (score d'un pays hypothétique moins bien classé que tous les autres)

Par la suite, nous avons ajouté plusieurs données nécessaires à la création de cartes : les **coordonnées géographiques des pays**, ainsi que leur **géométrie**, afin de pouvoir colorier les surfaces des pays sur les cartes. Nous avons également ajouté une colonne indiquant le **continent** sur lequel se trouve chaque pays.

L'ensemble des données utilisées est disponible dans le dossier [/data/](#)

Ces données nous semblent pertinentes dans le cadre d'une analyse car :

- Elles sont analysables dans le **temps**
- Elles sont analysables **géographiquement**
- Elles permettent d'étudier **différents facteurs de contribution au bonheur** en fonction des régions du monde, des cultures...

Plan d'analyse

1. **Classements** : Quels sont les pays et régions du monde les mieux classés en matière de bonheur global ?
2. **Analyse temporelle** : Comment évolue le bonheur moyen au fil des années ? (en regroupant par région, en utilisant des facet charts pour visualiser les différentes années en même temps)
3. **Analyse géographique** : Y'a-t-il des régions du monde moins heureuses que d'autres ? Pourquoi ?
→ Utiliser les facteurs de contribution du score pour mettre en évidence des causes de disparité
4. **Analyse factorielle** : Le niveau de bonheur est-il directement corrélé à la liberté des individus ? → Question extensible à la richesse, à l'espérance de vie...
5. Quelle combinaison de facteurs hauts entraîne une hausse du bonheur ? Quelle combinaison de facteurs bas entraîne une diminution de celui-ci ? (par exemple: avoir une espérance de vie élevée ET un PIB élevé ET un taux de liberté élevé implique-t-il nécessairement un haut niveau de bonheur, au-dessus d'un certain seuil ?)

Nettoyage des données

Le **nettoyage des jeux de données** est la première étape du projet, et certainement l'une des plus importantes. Nous nous en sommes rendus compte après avoir essayé de concevoir des graphiques avec les jeux de données bruts : nous avons été très vite restreints, d'un côté car il était impossible d'analyser les données temporellement car les données étaient séparées dans des datasets différents, et de l'autre parce que les noms des variables comportaient des espaces et différaient d'un dataset à l'autre.

Afin de pouvoir travailler efficacement, nous avons opéré le nettoyage suivant :

1. Nous avons commencé par **analyser tous les datasets** pour voir si les données nommées de la même manière correspondent à la même chose (ce qui n'était pas toujours le cas). Parfois, la variable "freedom" d'un dataset correspondait à la variable "explained_by_freedom" d'un autre dataset.
2. Nous avons ensuite **renommé toutes les colonnes de tous les datasets** de la même manière selon la **convention** définie ci-dessous.
3. Nous avons **fusionné l'ensemble des datasets** de manière à travailler sur un seul tableau propre à l'aide de filtres par la suite. Nous avions au préalable **ajouté une variable Year** car cette donnée n'était pas fournie à l'intérieur des datasets et nous aurions mélangé toutes les données lors de la fusion des datasets.
4. Afin de pouvoir travailler avec des cartes, nous avons été contraints d'ajouter pour chaque pays des **données géométriques et GPS**.

Convention de nommage des colonnes

- Le nom des colonnes commence par une **Majuscule**
- Les espaces sont remplacés par un underscore "**_**"

À savoir

Avant de commencer l'analyse détaillée du jeu de données, il y a quelques informations à connaître.

1. Comme nous allons analyser le score de bonheur des régions, nous avons jugé utile de présenter la **liste des pays de chaque région** :

Region	Countries
Australia and New Zealand	Australia, New Zealand
Central and Eastern Europe	Albania, Armenia, Azerbaijan, Belarus, Bosnia and Herzegovina, Bulgaria, Croatia, Czech Republic, Estonia, Georgia, Hungary, Kazakhstan, Kosovo, Kyrgyzstan, Latvia, Lithuania, Macedonia, Moldova, Montenegro, Poland, Romania, Russia, Serbia, Slovakia, Slovenia, Tajikistan, Turkmenistan, Ukraine, Uzbekistan
Eastern Asia	China, Hong Kong, Japan, Mongolia, South Korea, Taiwan
Latin America and Caribbean	Argentina, Bolivia, Brazil, Chile, Colombia, Costa Rica, Dominican Republic, Ecuador, Guatemala, Haiti, Honduras, Jamaica, Mexico, Nicaragua, Panama, Paraguay, Peru, Salvador, Suriname, Trinidad and Tobago, Uruguay, Venezuela
Middle East and Northern Africa	Algeria, Bahrain, Egypt, Iran, Iraq, Israel, Jordan, Kuwait, Lebanon, Libya, Morocco, Oman, Palestinian Territories, Qatar, Saudi Arabia, Syria, Tunisia, Turkey, United Arab Emirates, Yemen
North America	Canada, United States
Southeastern Asia	Cambodia, Indonesia, Laos, Malaysia, Myanmar, Philippines, Singapore, Thailand, Vietnam
Southern Asia	Afghanistan, Bangladesh, Bhutan, India, Nepal, Pakistan, Sri Lanka
Sub-Saharan Africa	Angola, Benin, Botswana, Burkina Faso, Burundi, Cameroon, Central African Republic, Chad, Comoros, Congo (Brazzaville), Congo (Kinshasa), Djibouti, Ethiopia, Gabon, Ghana, Guinea, Ivory Coast, Kenya, Lesotho, Liberia, Madagascar, Malawi, Mali, Mauritania, Mauritius, Mozambique, Niger, Nigeria, Rwanda, Senegal, Sierra Leone, Somaliland region, South Africa, Sudan, Swaziland, Tanzania, Togo, Uganda, Zambia, Zimbabwe
Western Europe	Austria, Belgium, Cyprus, Denmark, Finland, France, Germany, Greece, Iceland, Ireland, Italy, Luxembourg, Malta, Netherlands, North Cyprus, Norway, Portugal, Spain, Sweden, Switzerland, United Kingdom

2. Qu'est ce que le bonheur ?

Avant d'étudier le World Happiness Report, il est important de **définir ce qu'est le bonheur**, car il peut s'agir d'une notion assez subjective.

Le bonheur est un **état émotionnel agréable, équilibré et durable** dans lequel se trouve quelqu'un qui estime être parvenu à la satisfaction des aspirations et désirs qu'il juge importants. Il perçoit alors sa propre situation de manière positive et ressent un **sentiment de plénitude et de sérénité**, sans stress ni inquiétude. Cette impression ressentie, indispensable à la survie des mammifères, est principalement le résultat de la production de sérotonine, réduisant la prise de risques et poussant ainsi l'individu à maintenir une situation qui lui est favorable. Le bonheur ne doit pas être confondu avec la **sensation passagère de plaisir**, issue principalement de la production de dopamine et non de sérotonine, mais représente au contraire un état d'équilibre, agréable, qui dure dans le temps.

Source : [Wikipedia](#)

3. Est-ce que les données récoltées sont vraiment objectives ?

La principale question posée lors des enquêtes sur le bonheur est : "Imaginez une échelle avec des barreaux numérotés de zéro en bas à dix en haut. Le barreau du haut représente la meilleure vie possible pour vous, et le barreau du bas la pire vie possible pour vous. Sur quel barreau pensez-vous vous tenir à ce moment de votre vie ?". Connue sous le nom d'**échelle de Cantril**, cette évaluation du bien-être, ainsi que la formulation précise de la question, constitue un standard commun pour les recherches sur le **bien-être subjectif**. En effet, le bonheur est une notion très subjective, et peut rapidement être faussée par les petits événements du quotidien. Si on pose la question à quelqu'un qui vient d'avoir un F en IF36 et qui a loupé son train pour rentrer chez ses parents à Toulouse, il y a beaucoup de chances que sa réponse soit influencée par ses troubles actuels. Pourtant, ils ne définissent pas le bonheur sur le long terme de l'individu, seulement il est difficile pour l'homme d'avoir une vision d'ensemble sur sa vie. Ainsi, l'échelle de Cantril semble être selon les experts la mesure la plus fructueuse du bonheur des individus.

Il faut donc garder à l'esprit que les données constituent une **auto-évaluation des individus**, ainsi elles sont par défaut subjectives. Cependant, elles tendent à être comparables grâce à l'échelle de Cantril, qui nous offre la mesure la plus universelle possible du bonheur.

Sources : [Cairn.info](#) & [Our World in Data](#)

I) **Classements** : Quels sont les pays et régions du monde les mieux classés en matière de bonheur ?

Dans cette partie, nous identifierons **les régions du monde et les pays où la population est la plus heureuse**. Nous pourrons comparer ces classements avec les informations que nous obtiendrons suite à notre analyse factorielle, de sorte à observer si les populations les plus riches, libres, généreuses, etc. sont aussi les plus heureuses.

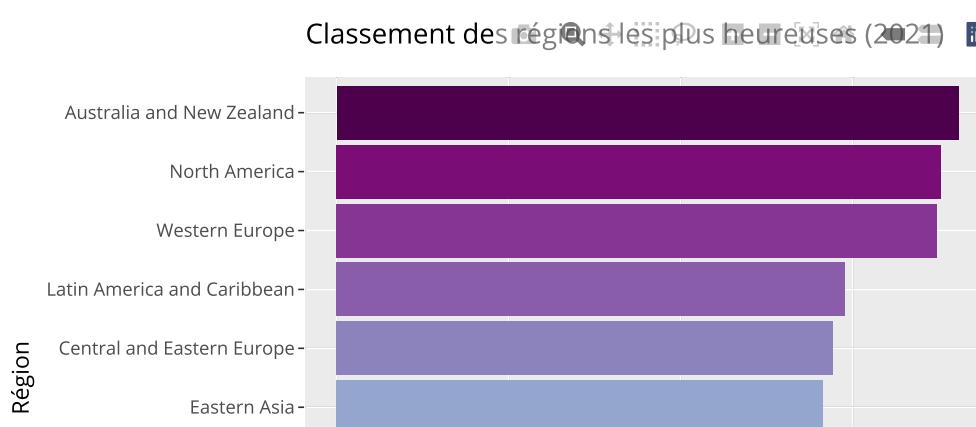
Les visualisations réalisées

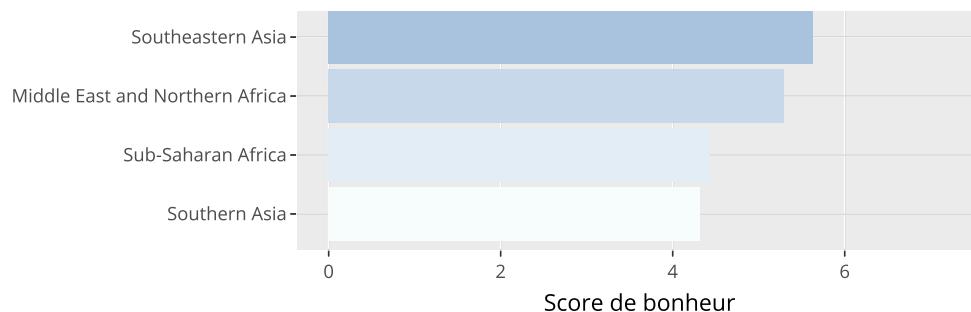
1. Le classement des régions en fonction de leur score de bonheur de 2021
2. Les 10 pays les plus heureux en 2021
3. Les 10 pays les moins heureux en 2021

L'ensemble des visualisations sont réalisées avec les données de l'**année 2021**, car ce sont les données les plus récentes que nous possédons à ce jour.

1. Classement des régions en fonction de leur score de bonheur

La première visualisation est un **classement des régions en fonction de la moyenne des scores de bonheur de leurs pays**. Certaines régions, comme l'Australie et la Nouvelle Zélande et l'Amérique du Nord ne comptent que très peu de pays, ainsi ce classement ne reflète pas forcément les mêmes informations que le classement des pays les plus heureux, que nous allons découvrir par la suite.





Sans trop de surprise, on observe que **les 3 régions du monde les plus heureuses sont les régions où les pays sont les plus développés**, soit L'Australie et la Nouvelle Zélande, l'Amérique du Nord et l'Europe de l'Ouest.

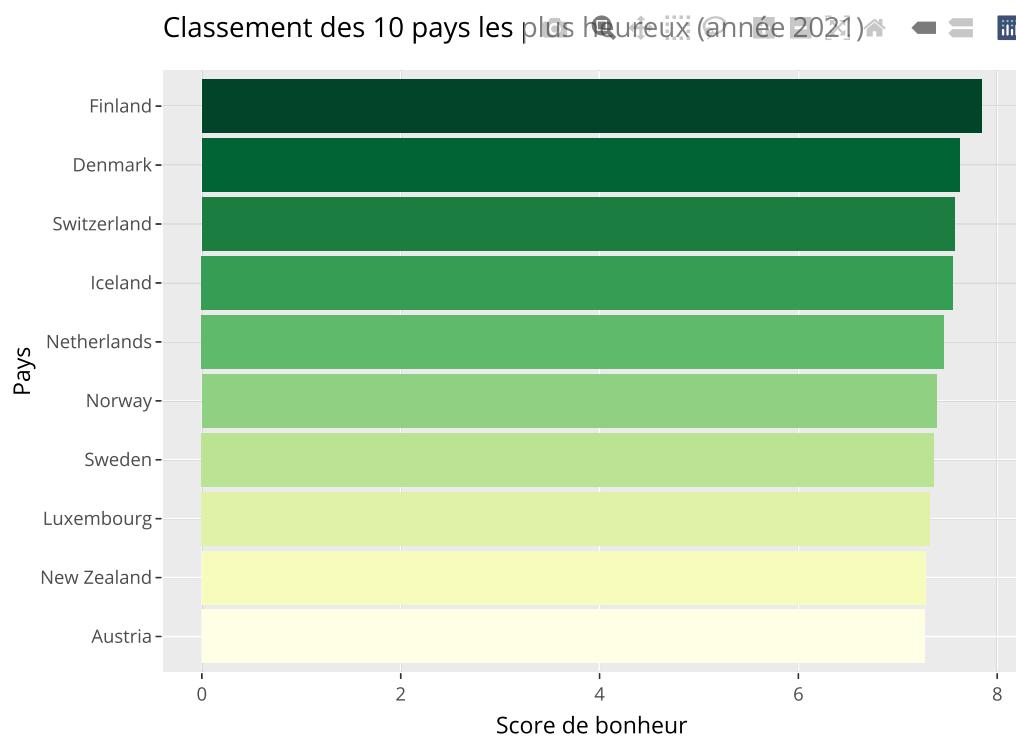
Parallèlement, on constate que **les régions en bas du classement**, soit l'Asie du Sud Ouest et l'Afrique Sub-saharienne, **correspondent aux régions où les pays sont sous développés**.

Ainsi, les **conditions de vie** respectives dans les différentes régions du monde semblent avoir une **forte influence sur le bonheur des populations**. Même si l'argent ne fait pas le bonheur, on constate ici qu'il y contribue fortement.

1. Classement des pays en fonction de leur score de bonheur

1. Top 10 des pays les plus heureux

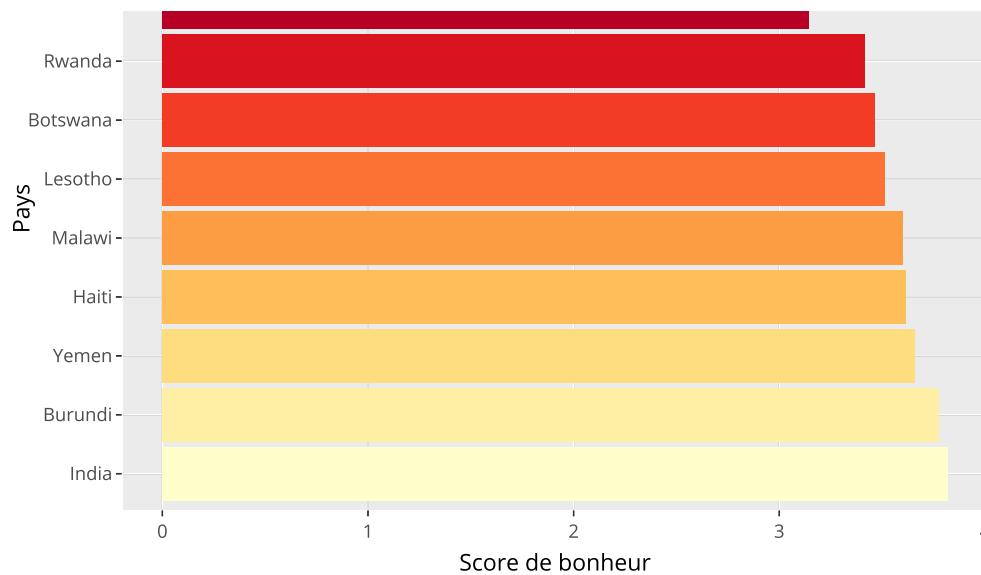
Ce deuxième graphique en barres permet de visualiser les **10 pays les plus heureux en 2021**.



Contrairement à la visualisation précédente, on constate que la majorité des pays de ce Top 10 n'appartiennent pas aux deux régions les mieux classées (sans compter la Nouvelle Zélande qui occupe la 9ème position du classement). En effet, **9 des 10 pays du classement font partie de l'Europe de l'Ouest**, pourtant elle n'occupe que la 3ème place du classement par région. Cela montre qu'il doit y avoir des disparités entre les pays de cette région.

2. Top 10 des pays les moins heureux





Comme nous nous y attendions, le pays le moins heureux du monde en 2021 est l'Afghanistan. Son score si bas s'explique par la **situation actuelle du pays** : prise de pouvoir par les talibans, crimes de guerre, non respect des droits de l'Homme, famine...

En ce qui concerne les autres pays du classement, leur position s'explique par un **climat dangereux** : risques d'attentats élevés (Zimbabwe), présence de gangs violents (Haiti), taux de criminalité élevé (Lesotho), vols et cambriolages nombreux (Rwanda)...

Ainsi, le **climat d'insécurité** dans lequel vivent les populations de ces pays influe beaucoup sur le bonheur des habitants.

II) Analyse temporelle : Comment évolue le bonheur moyen au fil des années ?

Cette deuxième partie s'attachera à répondre à la question "Comment a évolué le bonheur au fil des années entre 2015 à 2021 ?". Le déroulement de l'analyse nous amènera par la suite à diviser cette question en plusieurs sous-questions moins vastes.

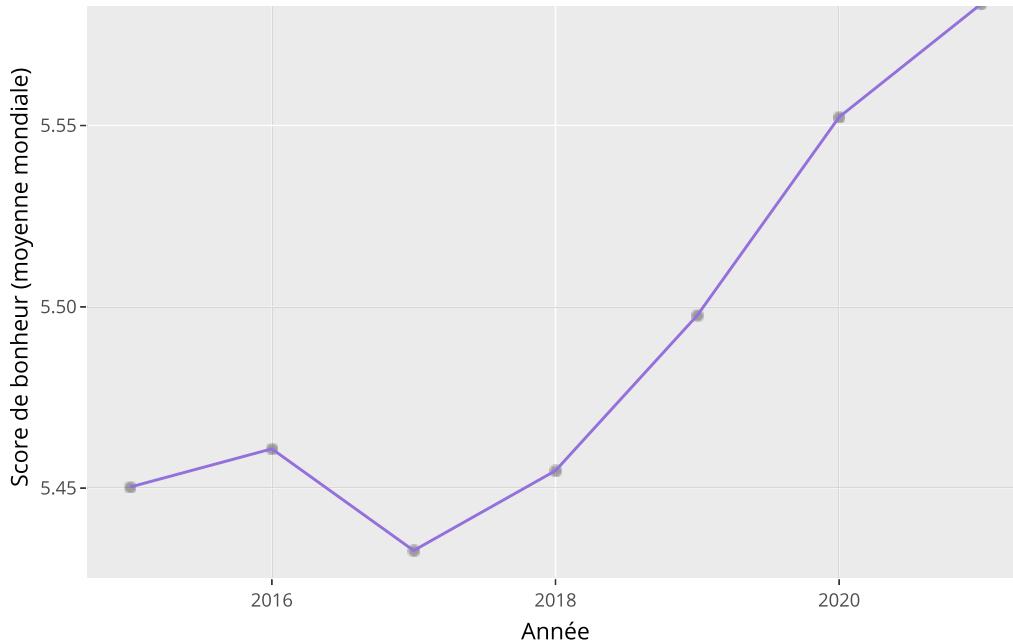
Les visualisations réalisées

1. La courbe d'évolution du bonheur dans le monde entre 2015 et 2021
2. Les courbes d'évolution du bonheur dans chacune des régions du monde entre 2015 et 2021
3. Les Top 10 des pays dont le score de bonheur a le plus augmenté / diminué entre 2015 et 2021
4. L'évolution du score de bonheur moyen des régions chaque année entre 2015 et 2021 (bar chart)

1. L'évolution du bonheur moyen dans le monde

Nous commencerons cette analyse par la **visualisation du bonheur moyen au niveau mondial**. La visualisation suivante est donc construite à partir de la moyenne globale des scores de bonheur de tous les pays chaque année.

Attention : On peut penser à première vue que la variation du bonheur moyen au fil des ans est importante. Il est important de prendre en compte l'échelle de l'axe des ordonnées : **la variation est en réalité relativement faible**, cette échelle ayant précisément été choisie pour qu'elle soit plus visuelle.



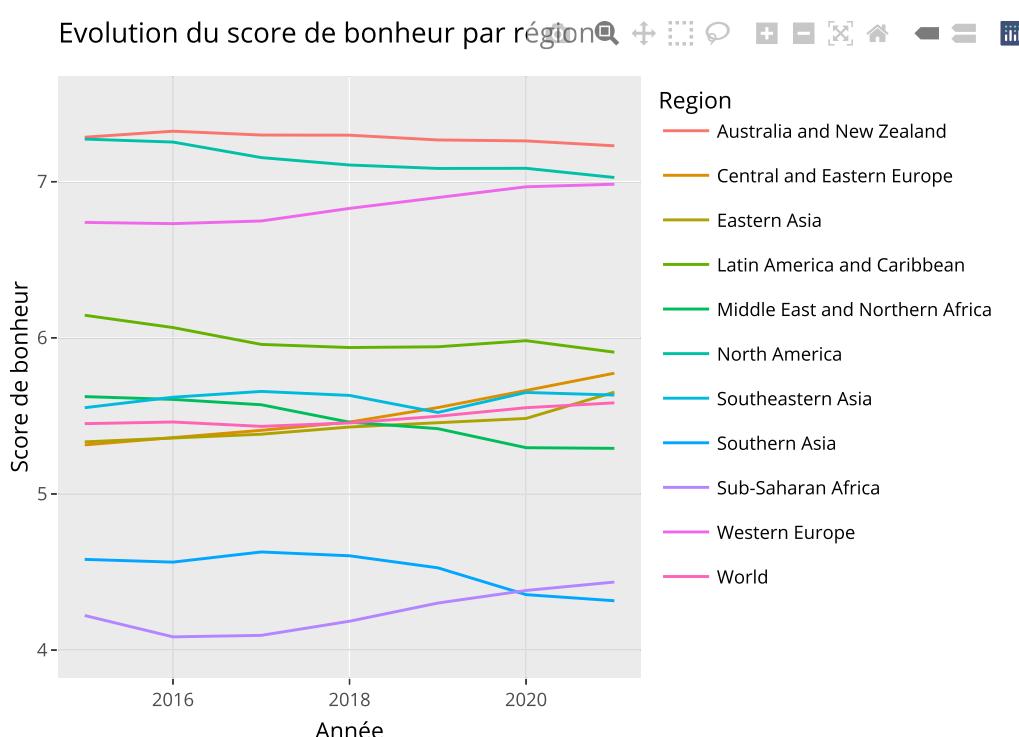
Ce graphique nous indique que le niveau moyen de bonheur dans le monde observe une **tendance à la hausse**, hormis une légère baisse en 2017.

Fait intéressant, **on n'observe pas de baisse en 2020, année d'apparition du Covid-19 et de confinements généralisés dans de nombreux pays**. On peut cependant s'interroger sur la représentativité de ce résultat au vu de la moyenne globale effectuée. Qu'en est-il lorsque nous nous plaçons à l'échelle des différentes **régions du monde** ?

2. L'évolution du bonheur dans les différentes régions du monde

Cette seconde visualisation s'attache à répondre à la question plus précise de l'**évolution du score de bonheur moyen par région**, chaque ligne représentant les données d'une région. On transforme pour cela nos données de sorte à les regrouper par année ET par région, puis on calcule la moyenne du score de bonheur par région pour chaque année.

On peut donc observer l'évolution des scores de bonheur obtenus via ce regroupement par la visualisation suivante :



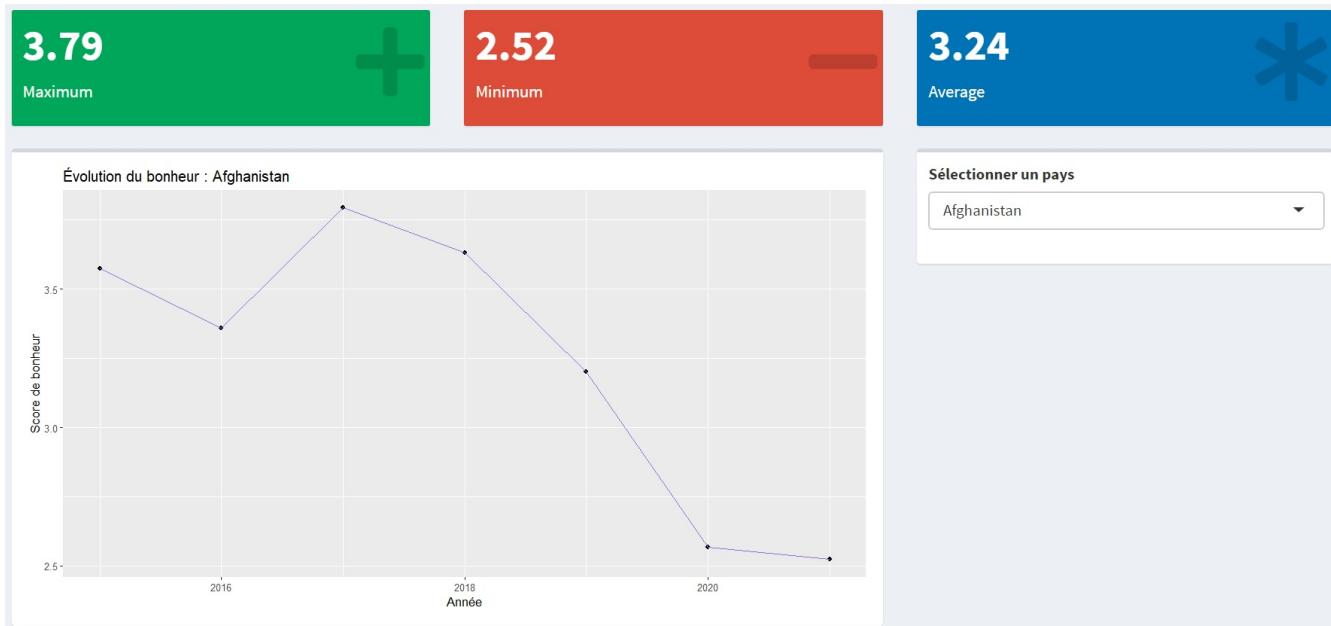
On a ajouté à ces données triées par région la **moyenne mondiale**. On peut donc effectuer

simplement des **comparaisons** (en cliquant sur l'étiquette d'une région pour la faire apparaître/disparaître notamment) entre les différentes régions du monde.

On se rend ici aisément compte qu'il ne fallait pas se contenter d'une moyenne globale, et qu'une moyenne par région nous montre des **disparités bien plus importantes** quant à l'évolution du niveau de bonheur dans le monde au cours des dernières années.

L'Asie du Sud, par exemple, observe une tendance à la baisse de son score de bonheur depuis 2017, tandis que l'Europe occidentale observe une tendance à la hausse sur cette même période.

Pour terminer sur cet aspect de l'évolution du score de bonheur, une visualisation est disponible sur notre application Rshiny pour détailler l'évolution du score de bonheur à l'échelle des pays :

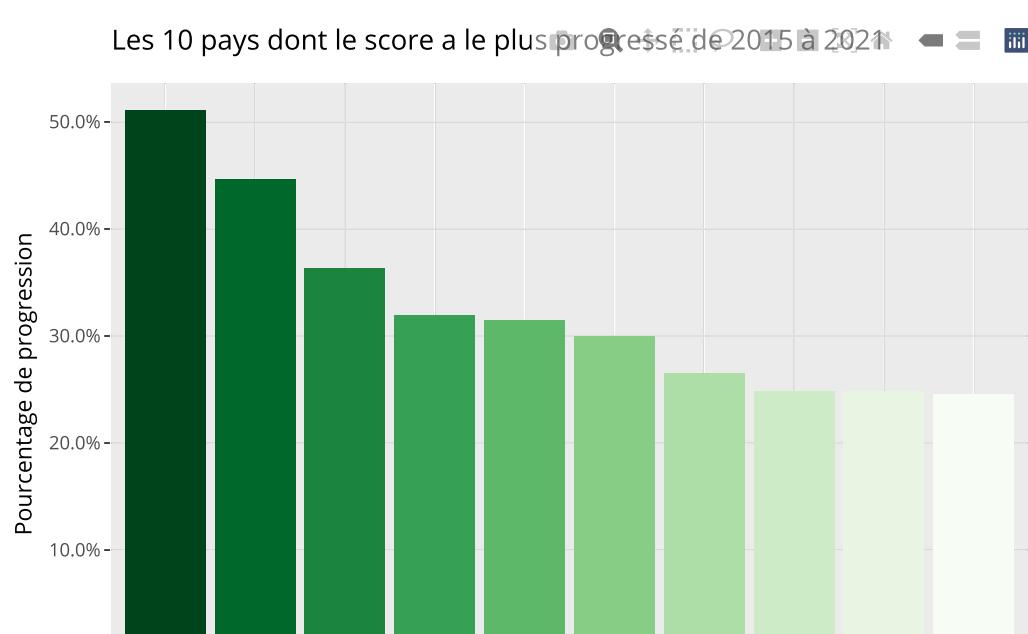


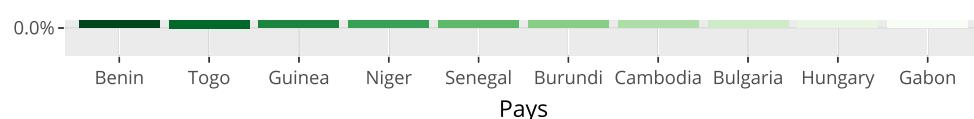
3. Les pays dont le niveau de bonheur a le plus varié au cours des dernières années

a) Les pays avec la plus grande augmentation de leur score de bonheur entre 2015 et 2021

La visualisation suivante s'attache aux **pays ayant connu la plus grande hausse de leur score de bonheur entre 2015 et 2021**. Elle représente les 10 taux d'accroissement du score de bonheur les plus importants, en prenant le score de 2015 comme référence.

Un taux d'accroissement de 20% signifie donc qu'en 2021, le score de bonheur du pays est 20% plus élevé qu'en 2015.



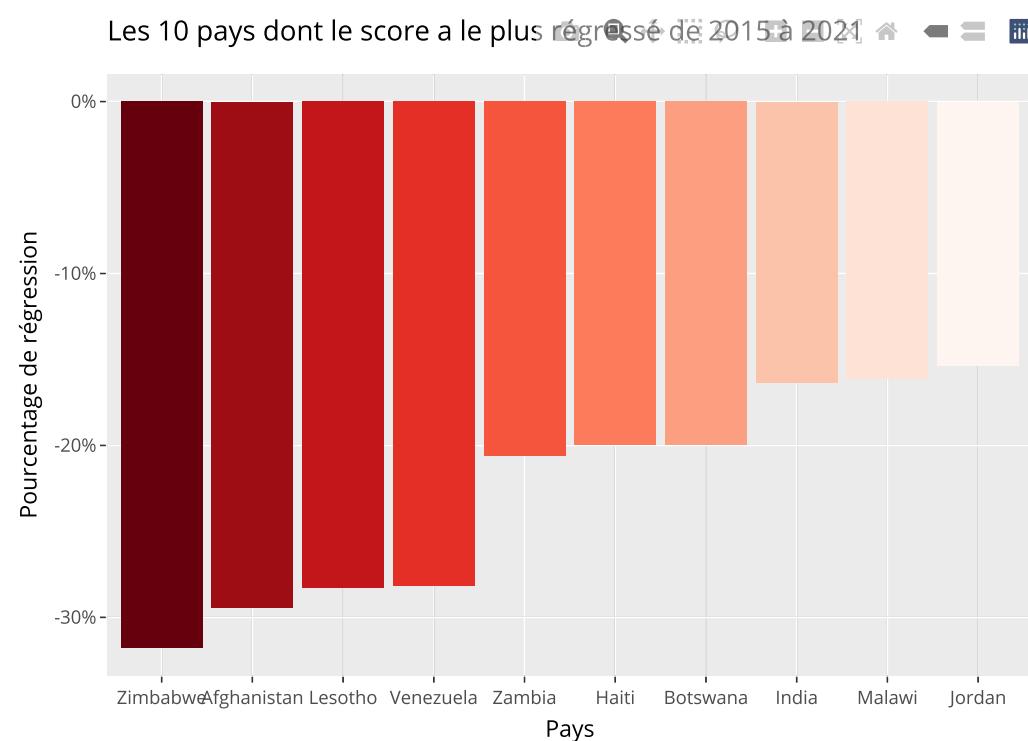


Nous sommes **mitigés** face aux informations qui ressortent de ce graphique. D'un côté, il nous paraît logique que les pays avec le plus fort taux d'accroissement sont des **pays dont le score de bonheur est assez bas, et peut donc facilement s'améliorer**. D'un autre côté, lorsqu'on s'intéresse à la situation actuelle au Bénin, on constate que le pays est situé dans une région soumise à beaucoup d'actes terroristes, et que la cybercriminalité se développe beaucoup dans le pays. On peut alors s'interroger sur les **limites du World Happiness Report** : quelle partie de la population du pays est interrogée ? **Est-ce que l'échantillon interrogé est vraiment représentatif de l'avis de la population ?**

Remarque : L'échantillon interrogé compte environ 1000 individus, ce qui ne représente que **0,008% de la population** du Bénin, qui compte un peu plus de 12 millions d'habitants en 2020.

b) Les pays avec la plus grande régression de leur score de bonheur entre 2015 et 2021

Voyons maintenant le résultat pour **les pays dont le score a le plus chuté entre 2015 et 2021** :



Sans surprise, les pays de ce Top 10 sont quasiement les mêmes que les **pays les moins heureux dans le monde en 2021**. Cette baisse du score de bonheur s'explique principalement par une **dégradation des conditions de vie** dans le pays et un **climat d'insécurité croissant**.

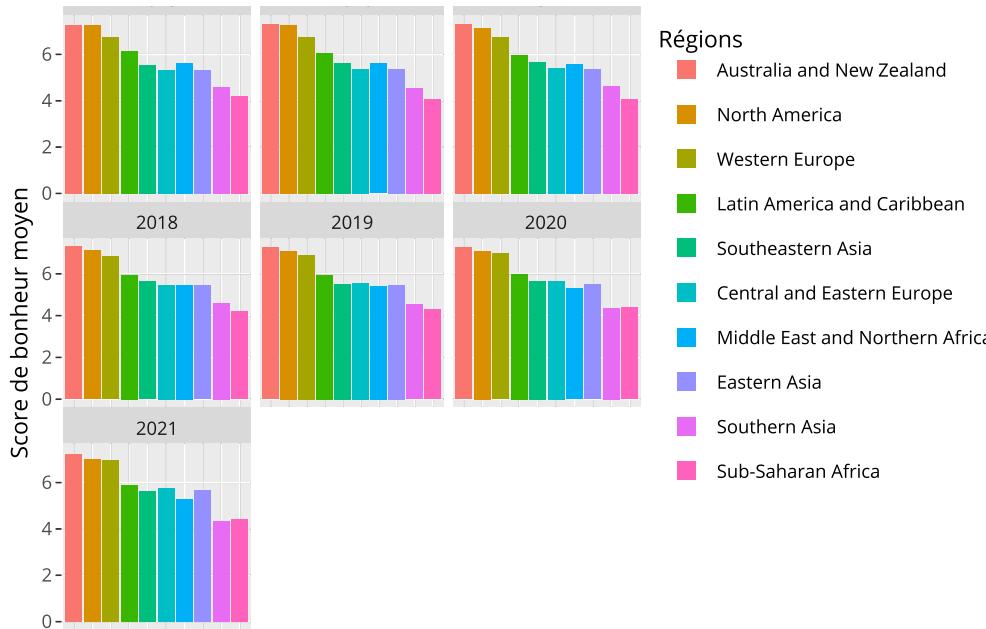
4. Le score de bonheur moyen par région chaque année

La visualisation suivante représente l'**évolution du score de bonheur moyen de chaque région au fil des années**.

Non, ces graphes ne sont pas identiques ;)

L'intérêt principal de ce graphique est de constater en un coup d'oeil que **les scores moyens par région restent très stables**. Globalement, chaque région du monde conserve un score de bonheur similaire, et l'allure globale de ces graphes évolue peu au cours des années.





En effet, le classement des quatre régions les plus heureuses, soit l'Australie et la Nouvelle Zélande, l'Amérique du Nord, l'Europe de l'Ouest et l'Amérique Latine n'a pas évolué entre 2015 et 2021.

On constate cependant que l'Europe de l'Est a devancé l'Asie du Sud-Est, que l'Asie de l'Est a devancé l'Afrique du Nord et le Moyen-Orient, et que l'Afrique sub-saharienne a devancé l'Asie du Sud-Ouest.

Ce dernier constat est sûrement dû au fait que le score de bonheur de l'Afghanistan (Asie du Sud-Ouest) a beaucoup chuté, puisqu'il est passé de 3,58 à 2,52 en l'espace de 6 ans.

Il faut donc faire attention à l'**interprétation des informations** : la situation de crise d'un seul pays peut être responsable de la chute d'une région dans le classement.

III) Analyse géographique : Y'a-t-il des régions du monde moins heureuses que d'autres ?

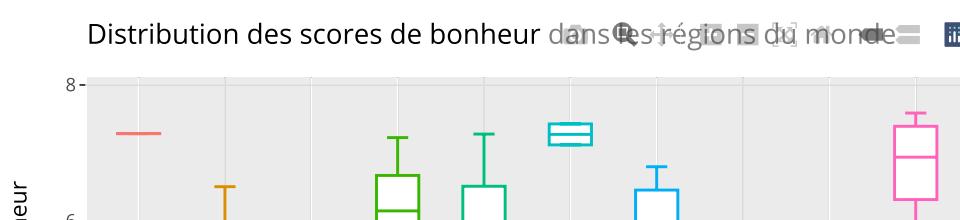
Dans cette partie, le but est de **mieux visualiser les préférences géographiques du bonheur** : et pour cela, rien de mieux que l'utilisation de **cartes choroplèthes**.

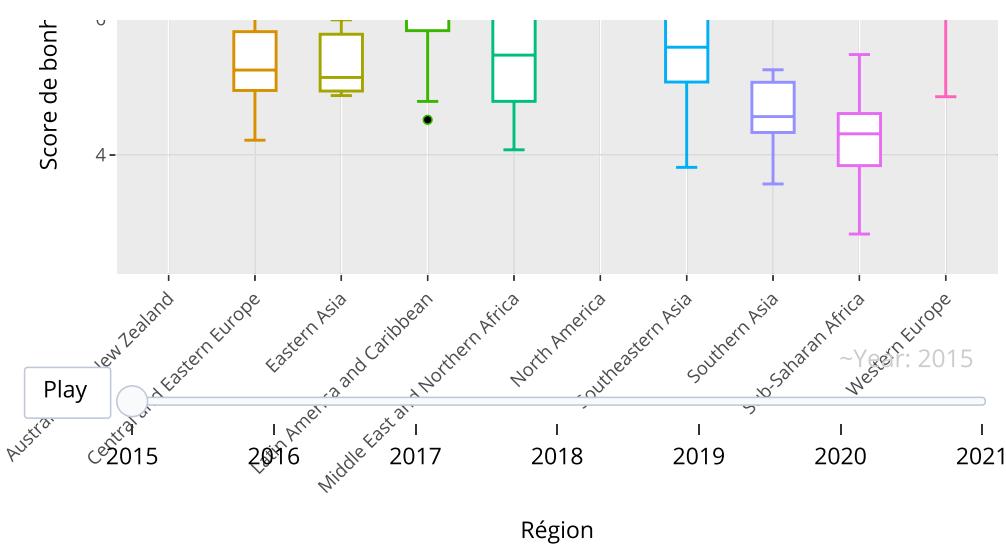
Les visualisations réalisées

1. Un boxplot qui indique la distribution des scores de bonheur de chaque région chaque année
2. Un bar chart qui indique la répartition des pays en fonction de leur score de bonheur
3. Une carte du monde colorée par pays en fonction de leur score de bonheur (avec Leaflet)
4. Les cartes réalisées avec **Tableau** : la contribution des différents facteurs au score de bonheur

1. La distribution des scores de bonheur de chaque région chaque année

La visualisation en boxplot nous permet de **situer rapidement les scores de bonheurs des pays d'une même région**. On constate en un coup d'œil que **certaines régions font l'objet de beaucoup de disparités**.



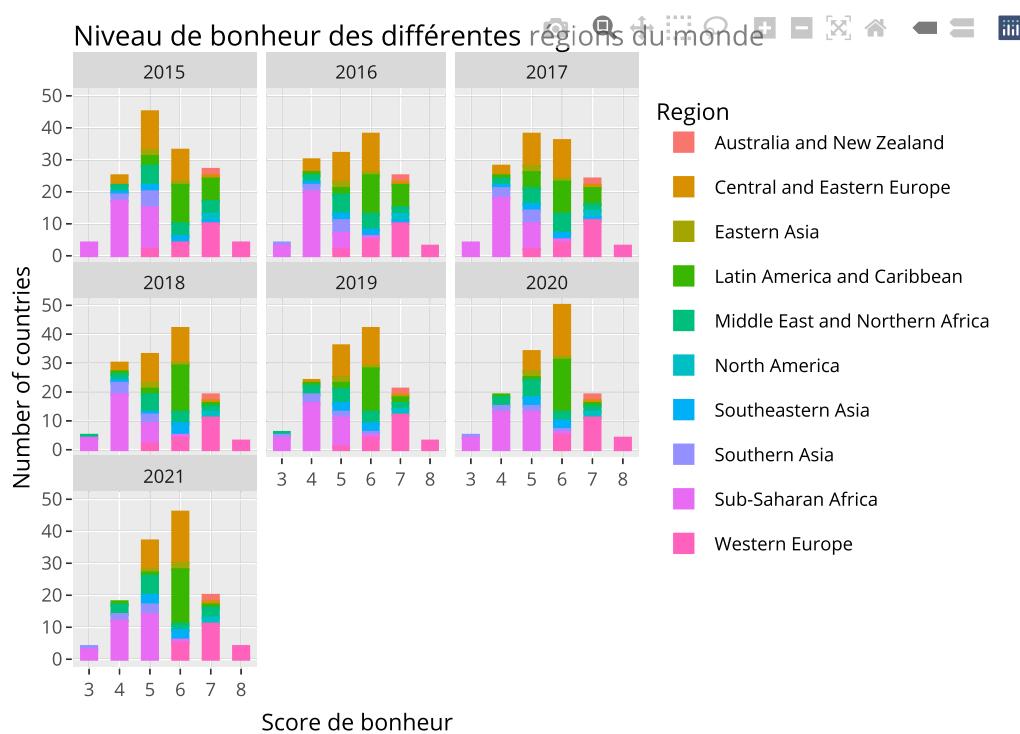


En effet, l'Afrique du Nord et le Moyen-Orient et l'Asie du Sud-Est présentent un écart de score de bonheur respectif de 3.2 et 2.98 entre leur pays le plus heureux et leur pays le moins heureux (2015). Cela montre que **les inégalités** (de développement, de conditions de vie, de richesse, de santé...) que l'on observe à l'échelle mondiale **s'observent aussi plus localement à l'échelle des différentes régions du monde.**

2. La répartition du niveau de bonheur des différentes régions du monde chaque année

Ce diagramme en barres permet de visualiser d'une autre manière la répartition du bonheur des pays d'une même région.

Le score de bonheur de chaque pays est arrondi à l'entier le plus proche. Le graphique prend en compte chacun des scores de bonheur individuel de chaque pays, puis les regroupe par régions.



On remarque que la plupart des régions du monde voient les scores de bonheur des différents pays répartis sur trois entiers consécutifs. Cela montre un **lien évident entre le score de bonheur d'un pays et la région du monde dans laquelle il se situe.**

On constate aussi que depuis 2018, en Amérique latine, le score de bonheur avoisine 6 pour une très grande majorité des pays. La région est donc assez équilibrée et égalitaire.

L'Europe de l'Est, quant à elle, est toujours la seule à avoir dans certains pays un score de bonheur à 8. Les régions les moins heureuses d'après les graphiques sont depuis 2015, l'Afrique subsaharienne et

l'Asie du Sud-Est. L'Afrique du Nord et le Moyen-Orient et de l'Afrique de l'Est semblent voir leurs scores de bonheur augmenter au fil des années.

Cette visualisation peut paraître redondante avec les deux précédentes, mais elle permet de mieux visualiser la **répartition des pays (et des régions) sur l'échelle de Cantril**.

Si on regarde l'évolution globale de la hauteur des barres, on constate que la répartition du bonheur en 2016 était beaucoup plus uniforme qu'à l'heure actuelle.

En 2016, il y avait autant de pays avec un score de 4 que de pays avec un score de 5 ou 6.

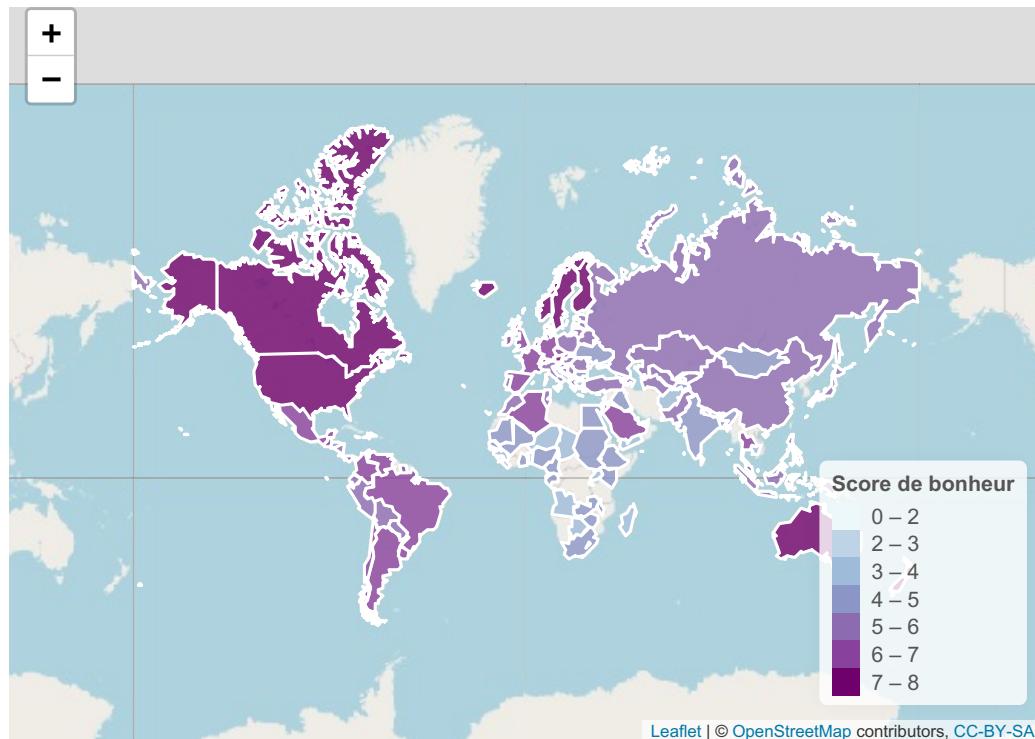
En 2021, la répartition des pays est beaucoup plus hétérogène.

Cela montre que l'on tend vers une **hétérogénéité croissante mais positive** (puisque de plus en plus de pays sont heureux et de moins en moins sont malheureux) **de la répartition du bonheur dans le monde**.

3. Le bonheur dans le monde : visualisation sur une carte

C'est une des visualisations les plus attendues au vu de notre jeu de données : **une carte** ! Rien de mieux pour visualiser et situer géographiquement les pays et régions les plus heureux.

Cette carte a été réalisée avec le package Leaflet de R. Elle est **interactive** : lorsque vous survolez un pays avec votre souris, le nom du pays, son classement et son score de bonheur précis s'affichent.



En un coup d'oeil, on visualise directement les pays et régions les plus heureux : ce sont ceux les plus foncés sur la carte, soit l'Amérique du Nord, l'Europe du Nord et de l'Ouest et l'Australie et la Nouvelle-Zélande.

On observe également que l'Afrique semble être le continent le moins heureux, avec un score de bonheur qui ne semble pas dépasser 5.

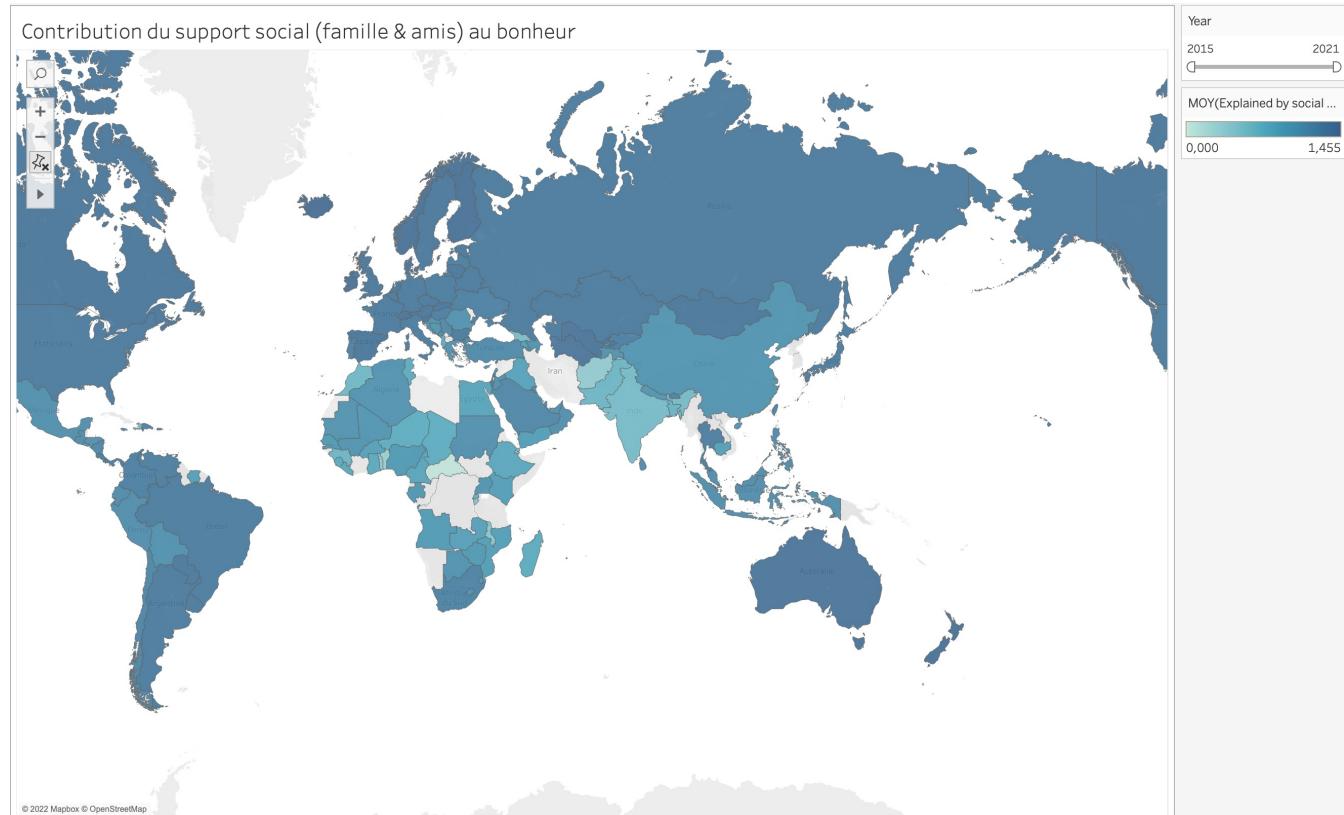
Enfin, l'Amérique latine et l'Asie sont dans la moyenne, avec un score de bonheur qui se situe pour la majorité des pays entre 5 et 7.

Cela correspond aux **tendances observées précédemment**, notamment via les classements des pays et régions en fonction de leur score de bonheur.

4. Les cartes réalisées avec **Tableau** : un aperçu de la contribution des facteurs au bonheur

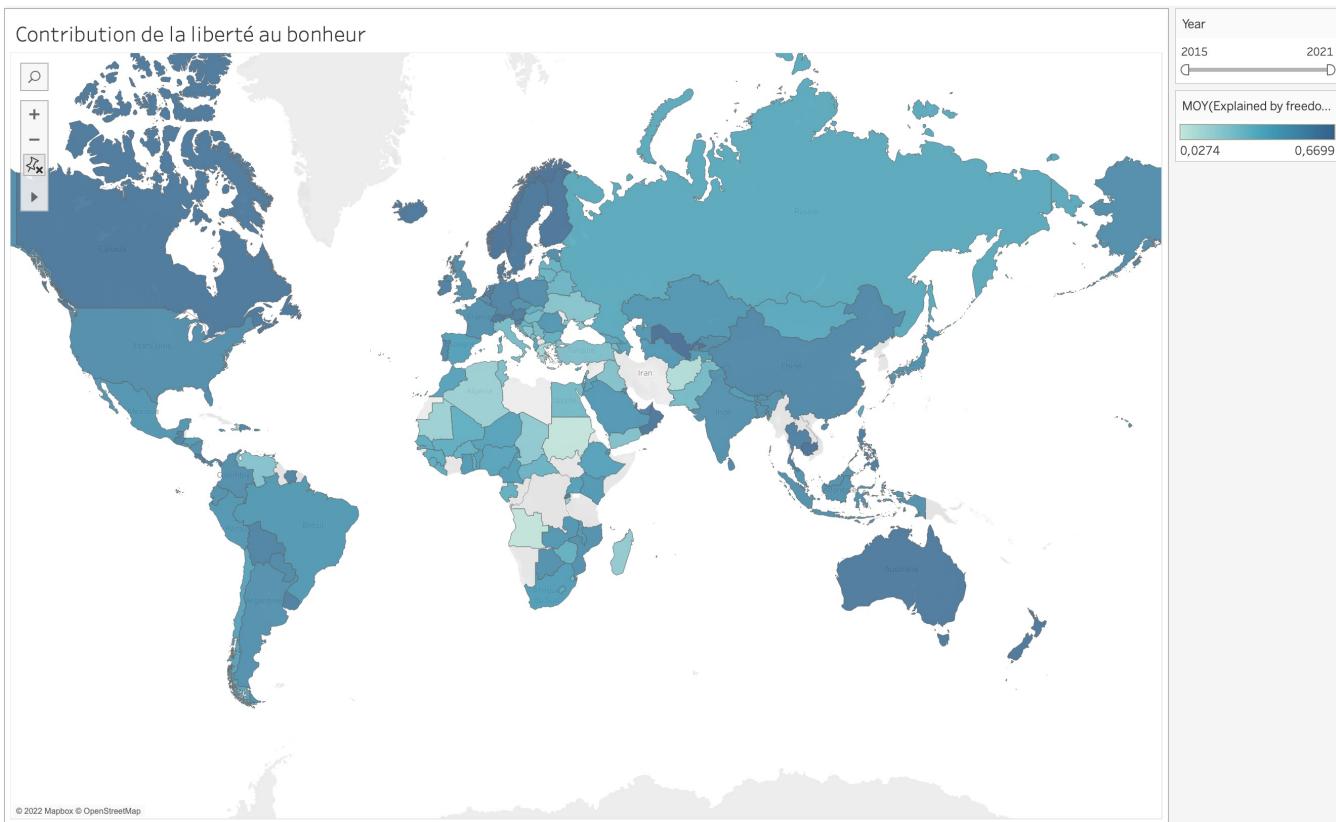
Les cartes suivantes ont toutes été réalisées sur **Tableau**, et intégrées au rapport via des fichiers .png trouvables dans le dossier _data. L'intégralité de nos réalisations sur Tableau sont également disponibles dans le fichier “visualisations_tableau.twb”.

La carte suivante permet de visualiser la **contribution du support social au score de bonheur dans le monde**, moyennée de 2015 à 2021 :



On constate globalement que le **support social contribue énormément au bonheur dans beaucoup de pays**. Nous essaierons de voir si cette tendance se confirme à l'aide d'un graphique hiérarchisant la contribution de chaque facteur au bonheur.

Nous pourrions penser que la **liberté des individus** est un élément essentiel à leur bonheur. Voyons dans quelle mesure ce facteur y contribue :



On constate cette fois des **niveaux de contribution plus faibles**, bien qu'existant. Notamment, le contraste en Russie avec la précédente carte est saisissant, mais finalement relativement logique au vu du régime politique en place dans ce pays...

IV) Analyse factorielle : Quels facteurs contribuent le plus au bonheur dans le monde ?

Les scores de bonheur précédemment mis en lumière s'expliquent par **différents facteurs** :

- La **liberté**
- L'**espérance de vie**
- Le **PIB** par habitant
- La **générosité**
- La **confiance dans le gouvernement** en place
- Le **support social** (par la famille et les amis)

Le jeu de données possède des champs décrivant le niveau de contribution de chacun de ces facteurs au score du bonheur (tous les champs dont le nom commence par Explained_by_...).

En plus de ces champs, les jeux de données des années 2020 et 2021 possèdent des champs donnant des valeurs pour ces facteurs (Generosity, Freedom, Life_expectancy...).

Cette partie s'attachera donc à tenter d'**observer les corrélations pouvant exister entre ces facteurs et un score de bonheur élevé ou non.**

Pour chacun des facteurs étudiés, nous nous attendons à observer une **corrélation linéaire positive** : plus une population est riche, en bonne santé, libre, entourée et généreuse, plus elle est censée être heureuse, n'est-ce pas ?

Les visualisations réalisées

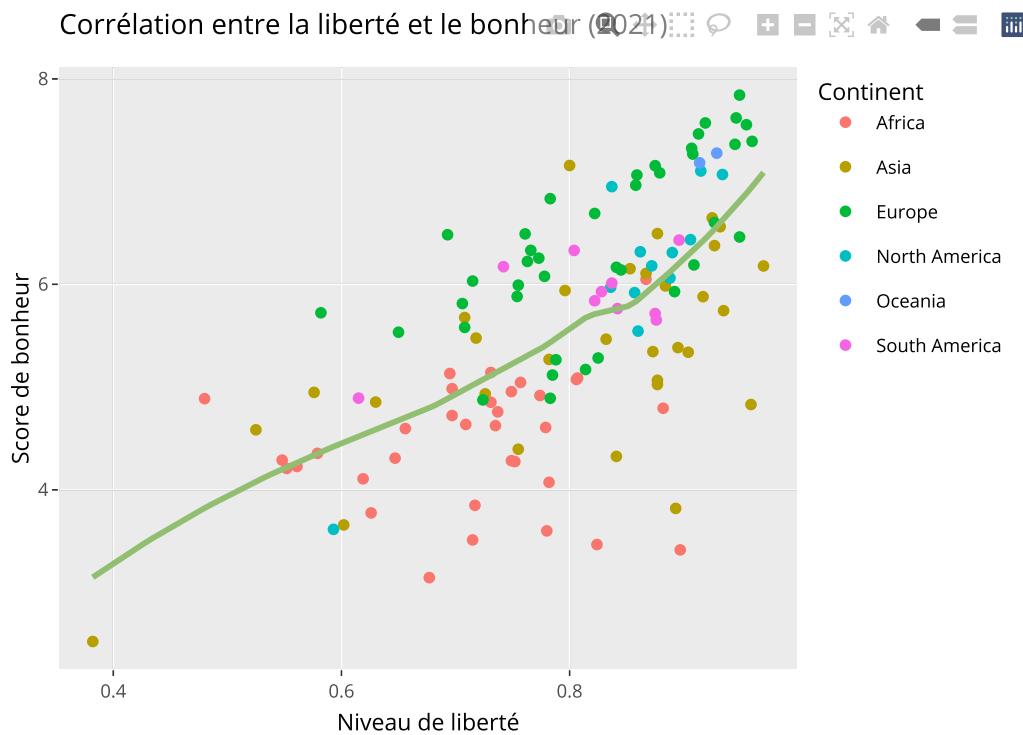
1. La corrélation entre le score de bonheur des pays et chacun des facteurs étudiés

2. L'analyse de facteurs supplémentaires

1. La corrélation de chaque facteur et le score de bonheur des pays du monde

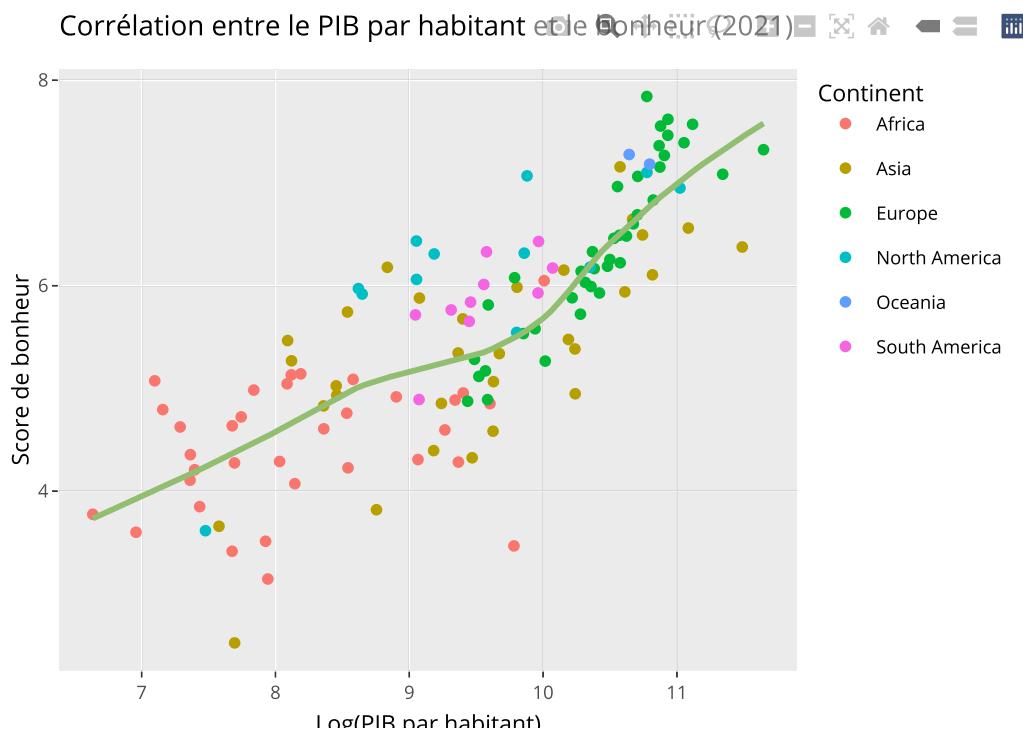
L'ensemble des visualisations suivantes sont réalisées avec les données de l'**année 2021**, car ce sont les données les plus récentes que nous possédons à ce jour.

Commençons par le facteur "**Liberté**": la visualisation suivante est un graphique en nuage de points, présentant en abscisse les valeurs enregistrées pour le facteur "liberté" et en ordonnée les valeurs de score de bonheur correspondantes pour chaque pays.

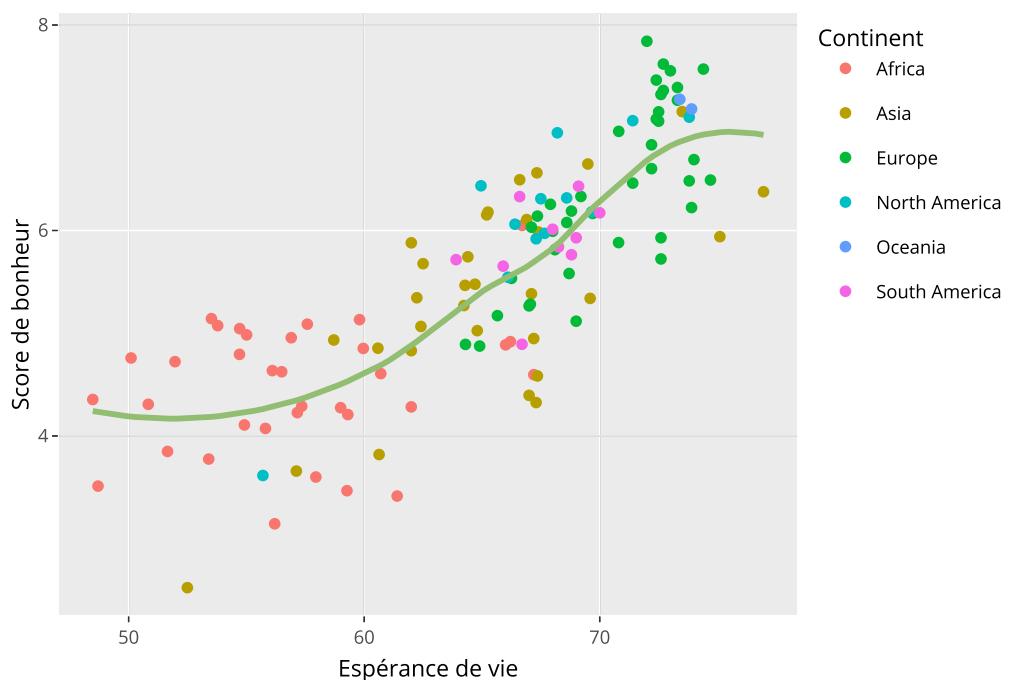


On constate immédiatement à la visualisation de ce graphique que le sentiment de liberté, comme on pouvait légitimement s'y attendre, est bien corrélé positivement au score de bonheur des différents pays. **Plus les habitants d'un pays se sentent libres, plus ils sont heureux.**

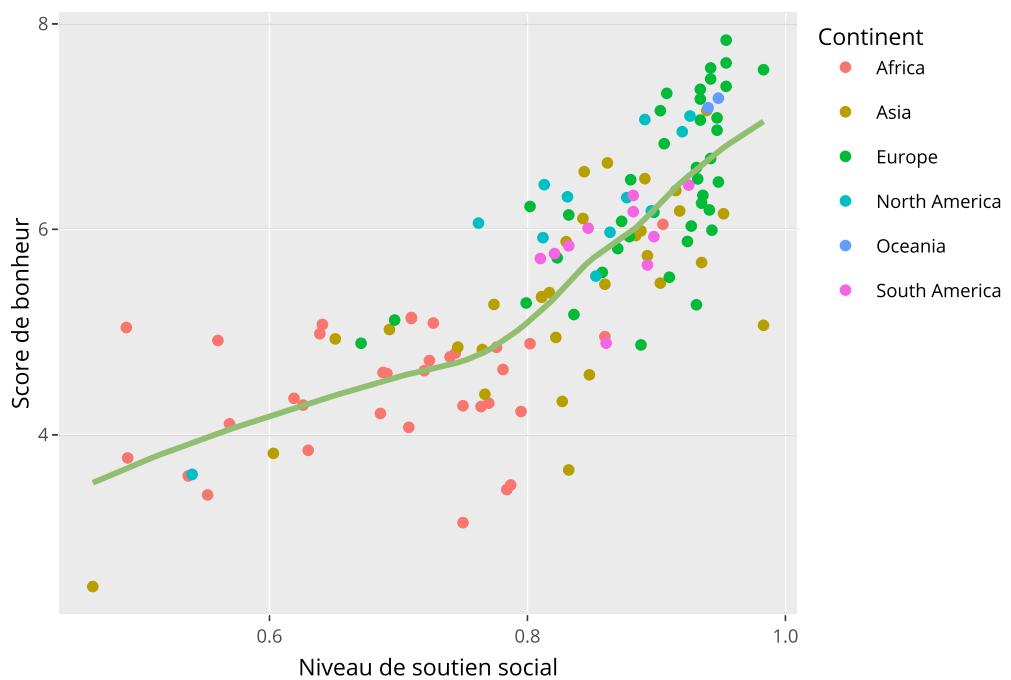
Effectuons la même analyse pour le **PIB par habitant**, l'**espérance de vie** et le **soutien social** :



Corrélation entre espérance de vie et bonheur



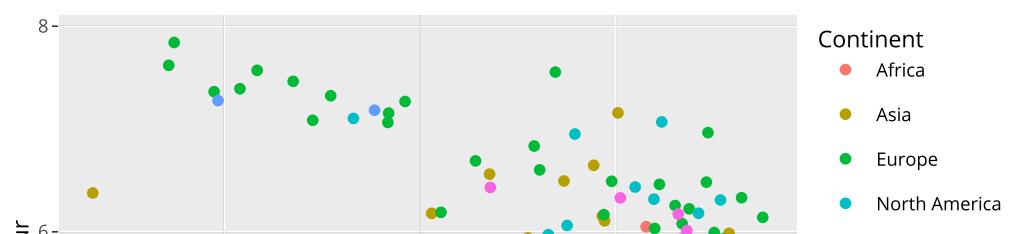
Corrélation entre soutien social et bonheur

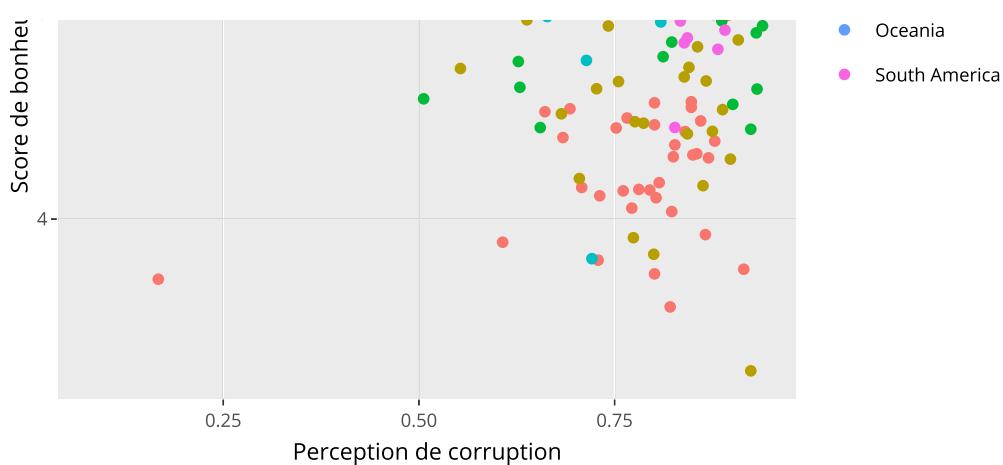


On peut encore une fois observer une **corrélation positive** entre ces derniers facteurs et le score de bonheur.

Concernant le facteur de **confiance en le gouvernement**, nous avons été surpris de constater à première vue une corrélation négative : des habitants en accord avec leur régime politique ne seraient pas plus heureux qu'une population soumise à une dictature ? Cela ne fait pas beaucoup de sens, ainsi nous nous sommes rendus compte que certains jeux de données comprenaient non pas le facteur "Confiance dans le gouvernement" mais "**Perception de corruption**". Nous avons alors harmonisé toutes les données pour construire la visualisation suivante :

Score de bonheur en fonction de la perception de corruption

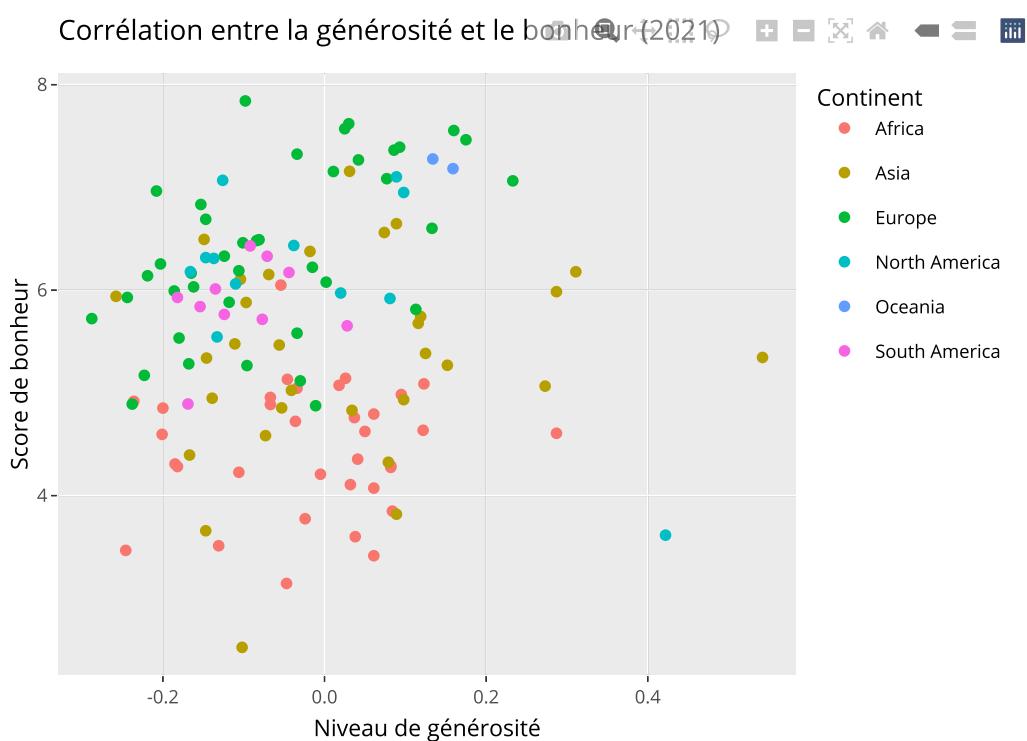




On constate que **la plupart des pays**, peu importe leur localisation géographique et leur niveau de bonheur, **présentent une forte perception de corruption**. Ainsi, on ne peut pas vraiment établir de corrélation entre le niveau de bonheur des pays et leur confiance en leur gouvernement.

De plus, un second facteur, plus étonnant cette fois, ne contribue par non plus grandement au score de bonheur.

En effet, voici le graphique de corrélation obtenu pour le facteur "**Générosité**" :



On constate cette fois-ci, et à l'inverse des graphes précédemment étudiés, que les valeurs sont relativement éparpillées et **ne permettent pas d'établir une corrélation entre la générosité et le score de bonheur**. Ce facteur est donc a priori moins impactant sur les variations de score entre les différents pays.

2. La corrélation entre les différents facteurs proposés dans un dataset supplémentaire et les scores de bonheur de l'année 2021

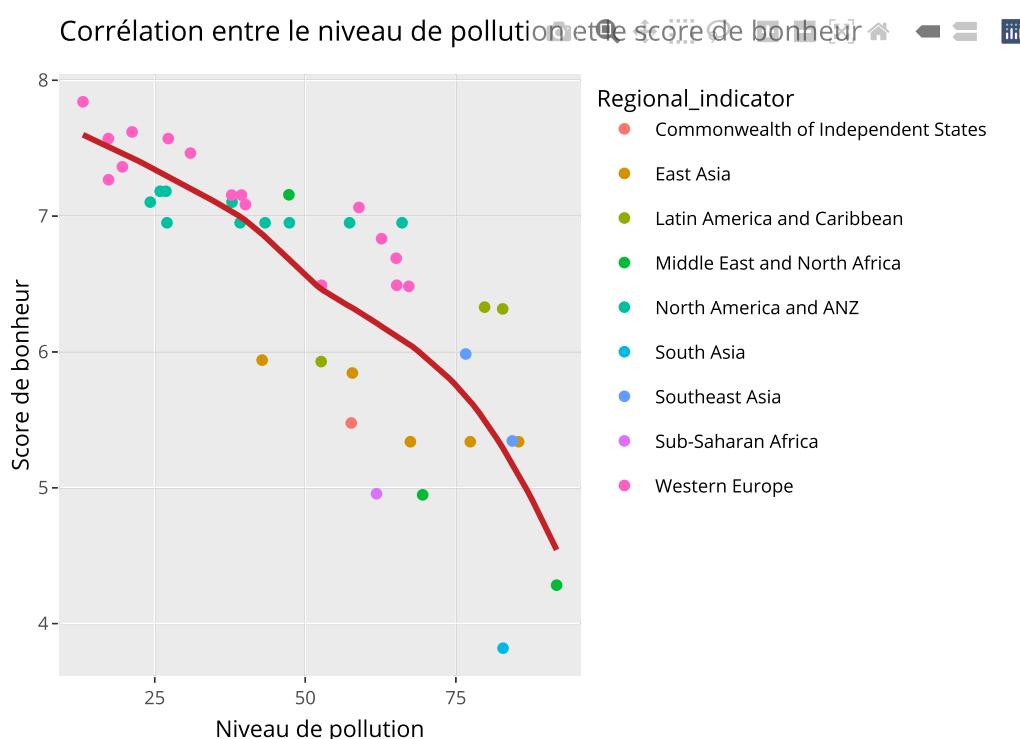
Lorsque l'on se renseigne sur les facteurs qui propulsent la Finlande en tête du classement des pays les plus heureux, on observe que ce ne sont pas forcément ceux que nous venons d'étudier. En effet, les facteurs qui ressortent sont : une planification urbaine intelligente, la présence de beaucoup d'espaces verts, de bons systèmes de santé, d'éducation, de transports, d'imposition...

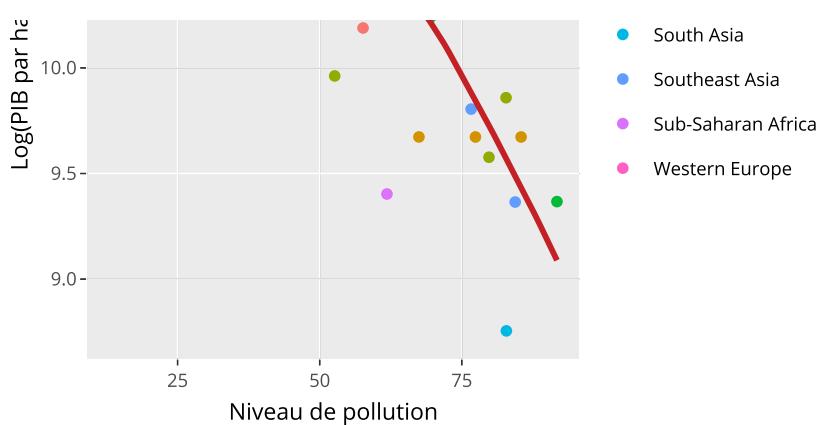
Nous avons donc recherché un **dataset supplémentaire, afin de compléter et d'enrichir notre analyse**.

Celui que nous avons trouvé comporte les données de **44 villes** dans le monde, et plusieurs **nouveaux facteurs** dont :

- Le **temps d'ensoleillement** annuel de la ville
- Le **taux de pollution** de la ville
- Le **prix d'une bouteille d'eau** dans la ville
- Le **niveau d'obésité** du pays
- Le **temps de travail moyen** des habitants dans l'année
- Le **nombre d'activités en extérieur** proposées

Le graphique qui suit montre la corrélation éventuelle entre le score de bonheur et le niveau de pollution d'un ville. Les villes sont rassemblées par régions du monde.





Le graphique ci-dessus montre qu'il existe bien une corrélation entre le niveau de pollution et le PIB par habitant. En effet, **les pays ayant un PIB élevé sont ceux ayant un score de pollution plus faible.**

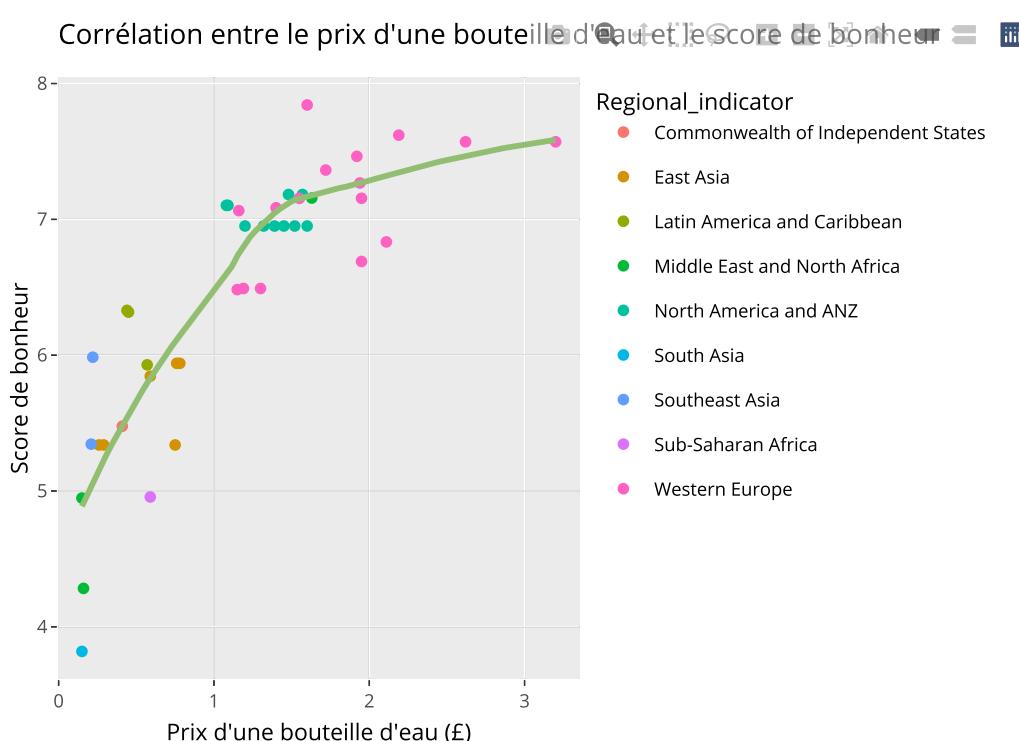
Cette découverte paraît étrange au vu de ce que l'on sait : la Chine et les États-Unis sont les pays qui polluent le plus. Cependant, ce n'est visiblement pas aux États-Unis que se fait le plus ressentir la pollution. Ainsi, les pays développés ont bien un niveau de pollution plus faible que les pays en voie de développement, ce qui ne permet pas vraiment de tirer une conclusion de notre analyse précédente.

En effet, il paraît aussi assez plausible que les pays en voie de développement ou sous-développés ont d'autres préoccupations plus importantes que leur impact environnemental, ce qui peut expliquer le phénomène observé.

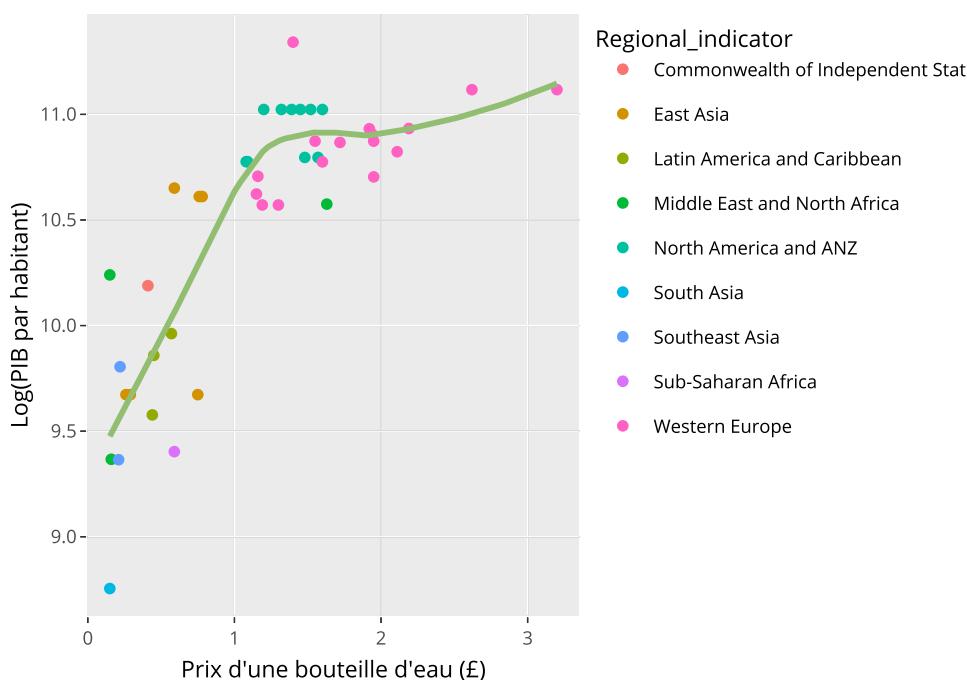
Le dataset complémentaire propose également une variable qui correspond au **prix d'une bouteille d'eau** dans les différentes villes. À nouveau, si on compare ce chiffre avec le score de bonheur des villes, nous pouvons voir une corrélation assez surprenante.

Les lieux où le prix de l'eau est le plus faible sont les lieux avec le score de bonheur le plus faible.

On peut alors penser que plus le prix de l'eau est élevé, plus le pays est riche et développé. Cela rejoint notre remarque précédente : **le score de bonheur est bien corrélé au niveau de développement des pays, mais pas forcément au prix de la vie ou au niveau de pollution**, qui sont des conséquences du niveau de développement.



Afin de vérifier l'observation précédente, il peut être intéressant d'observer la **corrélation entre le PIB par habitant et le prix d'une bouteille d'eau**.



Le graphique est en accord avec ce qui avait été observé. En effet, plus le PIB est élevé, plus le prix d'une bouteille d'eau augmente, ce qui montre une **corrélation entre le prix d'e l'eau et le niveau de développement des pays.**

Ici, nous étudions le prix d'une bouteille d'eau, mais celui-ci reflète d'une manière plus générale le **prix de la vie** dans le pays.

V) Quels sont les facteurs qui contribuent le plus au bonheur ?

Notre dernier axe d'analyse concerne le **poids des différents facteurs dans les scores de bonheur**. Nous allons répondre aux questions suivantes :

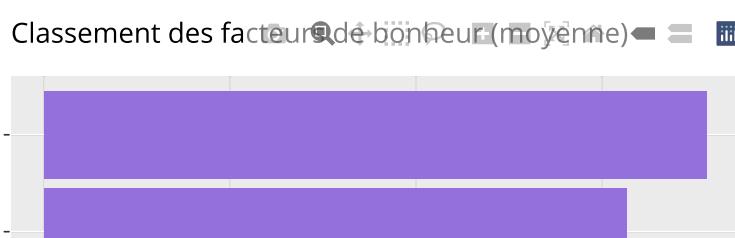
- Est-ce que les facteurs contribuent tous autant au bonheur ?
- Est-ce que les facteurs qui contribuent le plus au bonheur sont les mêmes dans toutes les régions du monde ?

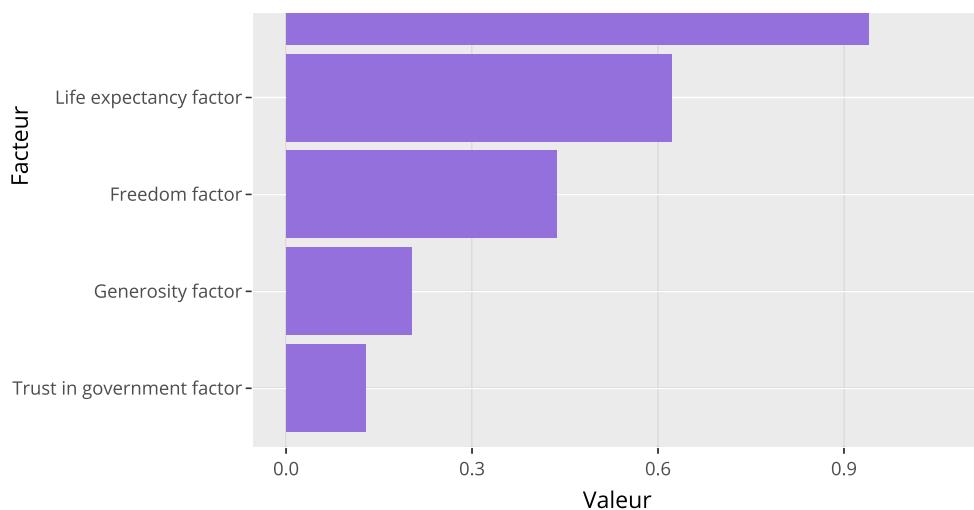
Les visualisations réalisées

1. Le classement des facteurs en fonction de leur contribution au bonheur mondial
2. Le classement des facteurs en fonction de leur contribution au bonheur dans les différentes régions du monde (2021)

Les facteurs qui contribuent le plus au bonheur dans le monde depuis 2015

Nous avons voulu connaître l'**influence de chaque facteur au niveau mondial**. Le graphique ci-dessous regroupe toutes les régions et toutes les années, et donne ainsi la valeur moyenne de la contribution des différents facteurs au bonheur.





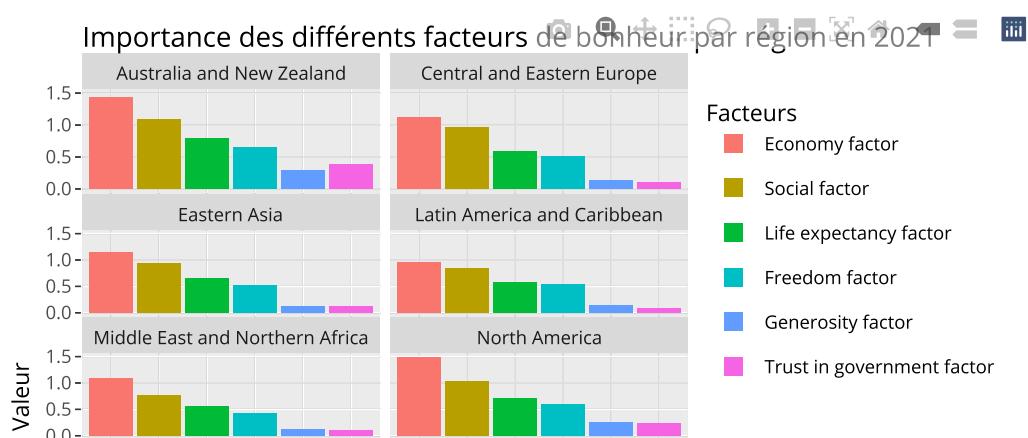
En moyenne et sur les 6 dernières années, c'est le **facteur social qui contribue le plus au bonheur**, suivi de l'**économie**. L'entourage d'une personne est donc un facteur universellement corrélé au bonheur des populations.

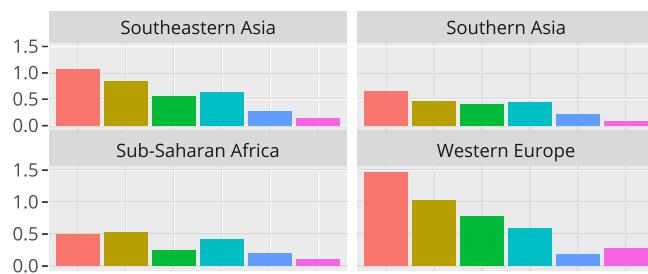
La **confiance dans le gouvernement** arrive en dernier dans ce classement. Cela confirme l'observation faite dans la partie précédente. En effet, la perception de corruption est assez élevée pour tout le monde, et ne contribue pas vraiment à rendre les gens heureux... Il est donc cohérent que la confiance dans le gouvernement ne soit pas un facteur d'influence important.

La contribution des facteurs au bonheur dans les différentes régions du monde (2021)

Nous avons voulu préciser l'analyse obtenue via le précédent graphique, en nous demandant quelles étaient les tendances de contribution des facteurs **dans chaque région** en 2021. Plusieurs constats nous ont alors frappé :

- Dans les régions développées, qui sont aussi globalement les plus heureuses, la tendance de répartition de la contribution des facteurs reste semblable.
- Dans les régions les moins heureuses, le niveau de contribution au bonheur des différents facteurs est plus bas (ce qui est logique), et le rôle de chacun des facteurs est plus équilibré.
- Enfin, et c'est le constat le plus important et l'observation la plus intéressante, on remarque un contraste non négligeable entre le précédent graphique et celui-ci. En effet, le graphe précédent place le **support social** comme facteur contribuant le plus au bonheur **en moyenne** de 2015 à 2021. Cependant, on constate ici que c'est globalement le **facteur économique** et non social qui prédomine en 2021. Ce phénomène peut aisément s'expliquer par la **pandémie** qui a frappé la plupart des régions du monde en 2020, et les confinements qui en ont résulté. Les interactions sociales entre les gens ont alors été grandement réduites, ce qui explique leur importance moindre dans le score de bonheur de 2021.





Conclusion

En conclusion, le dataset du **World Happiness Report** nous a permis d'étudier le niveau de bonheur à travers le monde. Nous avons d'abord **constaté** puis **analysé** les **variations temporelles et géographiques** du score de bonheur, puis nous avons tenté de **mettre en lumière les raisons de ces variations** en **contextualisant** au maximum les données.

Nous avons débuté notre analyse en établissant le **classement des différentes régions du monde en fonction de leur score de bonheur**. Sans trop de surprise, les **régions les plus développées arrivent en tête du classement**, et inversement, les régions les moins développées arrivent en bas du classement. Les régions ayant un score de bonheur plus faible sont aussi souvent des régions où **l'atmosphère globale est assez médiocre** (conflits, climat dangereux, insécurité...).

Ensuite, nous avons analysé l'**évolution du score de bonheur dans chacune des régions du monde depuis 2015**. Nous avons observé que les pays voyant leur score de bonheur augmenter ces dernières années étaient, pour la plupart, des pays en bas du classement. Le Bénin notamment, est le pays dont le score de bonheur a le plus augmenté ces dernières années. Nous nous sommes alors posé la question de la **pertinence de l'échantillon d'individus interrogé pour l'enquête** puisque la situation actuelle au Bénin ne semble pas aller de pair avec un climat "favorable" à une hausse du score de bonheur.

Cette observation nous a également amené à nous questionner sur les buts du World Happiness report. Cette étude contribue au **Réseau des solutions de développement durable des Nations unis**. Elle permettrait de donner une plus grande visibilité aux recherches qui analysent la construction et les moyens de favoriser le bonheur à travers le monde.

En effet, le climat global d'un pays peut être étudié par le biais de **différents facteurs, dont nous avons étudié l'impact sur les scores de bonheur**. Notre analyse factorielle a confirmé certaines de nos idées préconçues : le niveau de liberté, le PIB par habitant ou encore le niveau de soutien social sont fortement corrélés au niveau de bonheur. D'autres données pourraient sûrement nous apporter plus d'informations quant aux facteurs influant sur le niveau de bonheur. Par exemple, les habitants de la Finlande (premier pays du classement) ne justifient pas seulement leur niveau de bonheur par les facteurs étudiés. En effet, beaucoup l'expliquent par la **qualité de l'environnement** (espaces verts, réseau de transport fiable, bon niveau d'éducation...).

Notre avis sur le projet

"J'ai trouvé ce projet très intéressant et enrichissant. J'ai pu découvrir deux logiciels, Tableau et RStudio, et un nouveau langage de programmation dont l'utilité diffère des langages que je connais déjà comme le Java ou le C. J'ai trouvé que le R est un langage qui peut être difficile à débugger, ainsi j'ai beaucoup appris à mieux formuler mes problèmes pour être plus efficace lors de mes recherches de solutions sur internet. J'aurais aimé plus développer la **création de cartes interactives avec Leaflet**, notamment l'idée que j'avais en tête d'introduire un graphique dynamique dans les info-bulles des pays tracés sur la carte." Charlotte

"Ce projet m'a permis de découvrir le langage R par le biais d'une étude assez libre. J'ai eu

du mal par moment à aller au bout des graphiques que j'avais en tête. J'aurais aimé que l'on puisse centrer une partie de l'étude sur un pays en particulier afin de voir les disparités à l'échelle d'un pays. Malheureusement, la plupart de nos données sont à plus grandes échelles et ne permettaient donc pas cette étude." Maud

"Me destinant à un métier de data analyst, cette UE m'a permis de découvrir le langage R, puissant et simple d'utilisation dans une optique de faire des dashboard. J'ai notamment beaucoup apprécié la partie Rshiny dans cette optique, et aurais aimé l'approfondir un peu plus. De plus, la partie du cours s'attardant sur la bonne façon de représenter une donnée, et les différents biais inhérents à certaines méthodes de représentation apporte une réelle plus-value." Evann