

Szczęście na świecie

Zuzanna Kotlińska

18 01 2022

Wstęp

Raport opisuje analizę eksploracyjną zbioru danych dotyczącego wskaźnika szczęścia na świecie w roku 2015. Zawiera on podstawowe informacje o budowanie danego zbioru, a także rozkładach i zależnościach między zmiennymi. Dane pochodzą ze strony Kaggle.

Kilka informacji o zbiorze danych

Opisywana ramka danych zawiera 12 kolumn oraz 158 wierszy. Pięć pierwszych obserwacji wygląda następująco:

Country	Region	Happiness.Rank	Happiness.Score	Standard.Error	Economy..GDP.per.Capita.
Switzerland	Western Europe	1	7.587	0.03411	1.39651
Iceland	Western Europe	2	7.561	0.04884	1.30232
Denmark	Western Europe	3	7.527	0.03328	1.32548
Norway	Western Europe	4	7.522	0.03880	1.45900
Canada	North America	5	7.427	0.03553	1.32629

Family	Health..Life.Expectancy.	Freedom	Trust..Government.	Corruption.	Generosity	Dystopia.Residual
1.34951	0.94143	0.66557		0.41978	0.29678	2.51738
1.40223	0.94784	0.62877		0.14145	0.43630	2.70201
1.36058	0.87464	0.64938		0.48357	0.34139	2.49204
1.33095	0.88521	0.66973		0.36503	0.34699	2.46531
1.32261	0.90563	0.63297		0.32957	0.45811	2.45176

Poniżej opis każdej z kolumn.

- **Country** - badany pod względem współczynnika szczęścia kraj,
- **Region** - region, do którego dany kraj należy,
- **Happiness.rank** - pozycja w rankingu, zależna od wysokości wartości kolumny **Happiness.Score**,

- **Happiness.Score** - średnia ocena mieszkańców danego kraju, opisująca poziom szczęścia, gdzie 0 - możliwie najgorsze pod względem szczęścia życie, 10 - możliwie najlepsze życie.
- **Standard.Error** - błąd standardowy dla kolumny **Happiness.Score**,

Następne kolumny opisują siedem czynników, które mogą tłumaczyć pozycję danego kraju w rankingu, ale nie mają wpływu na wartość kolumny **Score**.

- **Economy (GDP.per.capita)** - wskaźnik opisujący produkt krajowy brutto danego kraju podzielony przez liczbę jego mieszkańców,
- **Family** - wskaźnik roli rodziny w danym państwie,
- **Health (Life.expectancy)** - wskaźnik lat zdrowego życia,
- **Freedom** - wskaźnik wolności podejmowania życiowych decyzji,
- **Trust (Government Corruption)** - wskaźnik, który opisuje pozycję danego państwa pod względem korupcji w rządzie,
- **Generosity** - wskaźnik hojności obywateli danego państwa,
- **Dystopia.residual** - wskaźnik szczęścia dystopii,

Poniższa tabela przedstawia własności każdej zmiennej liczbowej. Pominęłam **Overall.rank**, z którego nie ma sensu liczyć między innymi średniej czy liczby unikalnych wartości. To samo dotyczy kolumn **Country** i **Region**, tym bardziej, że ich zmienne nie są numeryczne.

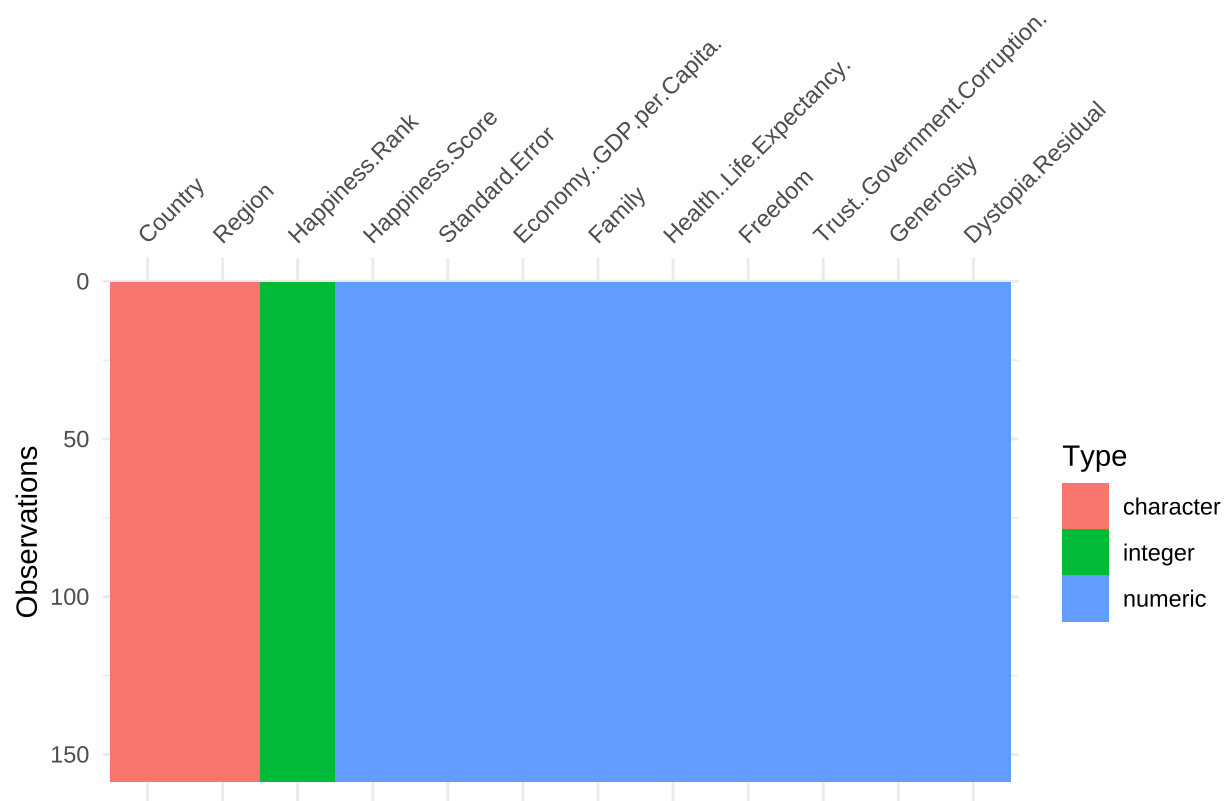
Zmienna	Liczność	Liczba NA	Średnia	Odchylenie standardowe
Happiness.Score	158	0	5.3757342	1.1450101
Standard.Error	158	0	0.0478847	0.0171462
Economy..GDP.per.Capita.	158	0	0.8461372	0.4031208
Family	158	0	0.9910459	0.2723691
Health..Life.Expectancy.	158	0	0.6302594	0.2470778
Freedom	158	0	0.4286149	0.1506928
Trust..Government.Corruption.	158	0	0.1434218	0.1200341
Generosity	158	0	0.2372955	0.1266849
Dystopia.Residual	158	0	2.0989768	0.5535498

Zmienna	Wartość minimalna	Wartość maksymalna	Unikalne wartości
Happiness.Score	2.83900	7.58700	157
Standard.Error	0.01848	0.13693	153
Economy..GDP.per.Capita.	0.00000	1.69042	158
Family	0.00000	1.40223	158
Health..Life.Expectancy.	0.00000	1.02525	157
Freedom	0.00000	0.66973	158
Trust..Government.Corruption.	0.00000	0.55191	157
Generosity	0.00000	0.79588	158
Dystopia.Residual	0.32858	3.60214	158

Analiza zmiennych

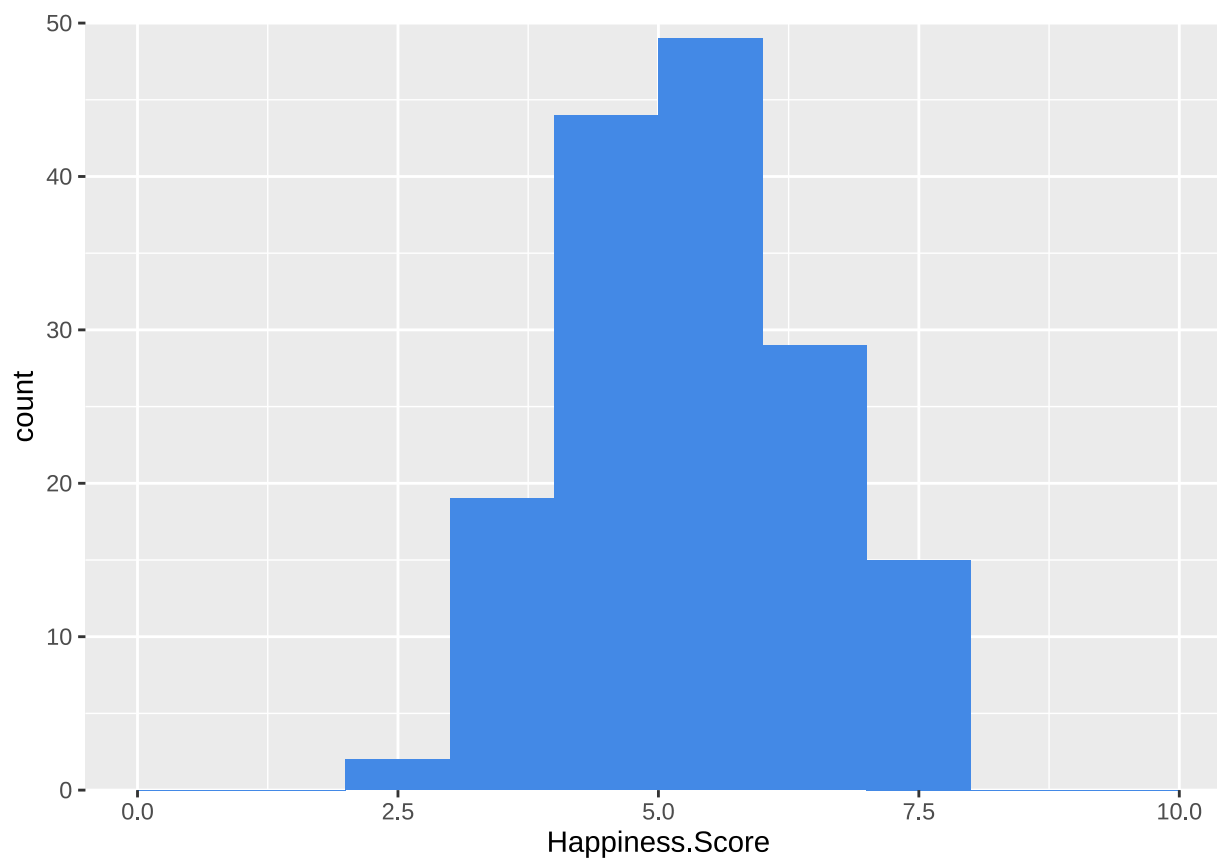
Typy zmiennych

Spójrzmy najpierw, jakiego typu są zmienne dla każdej z analizowanych kolumn.



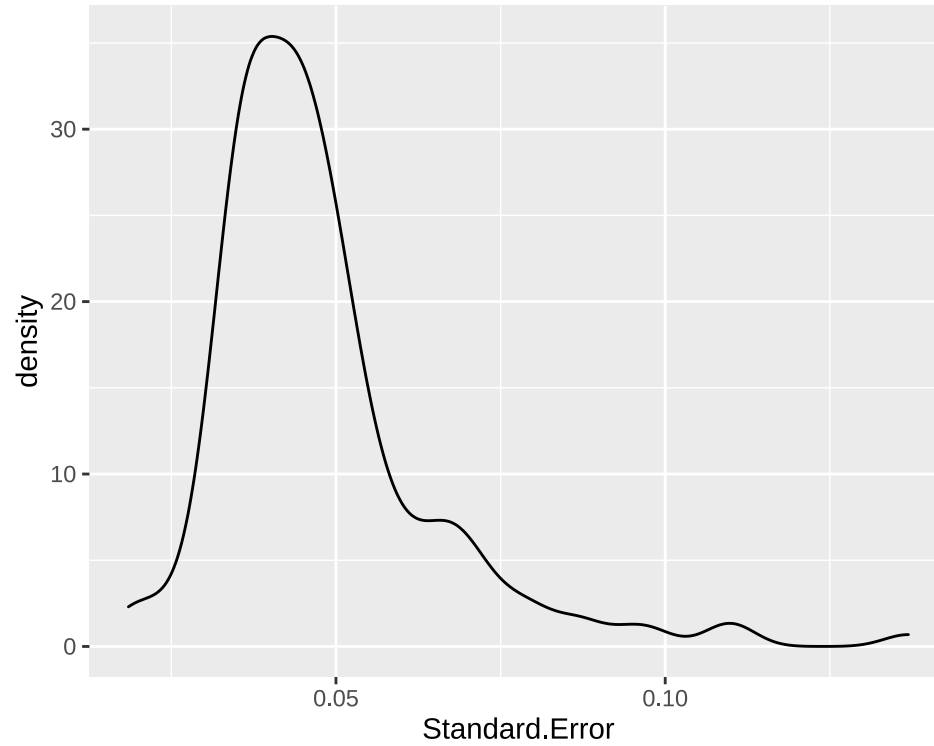
Rozkłady zmiennych

1. Happiness Score

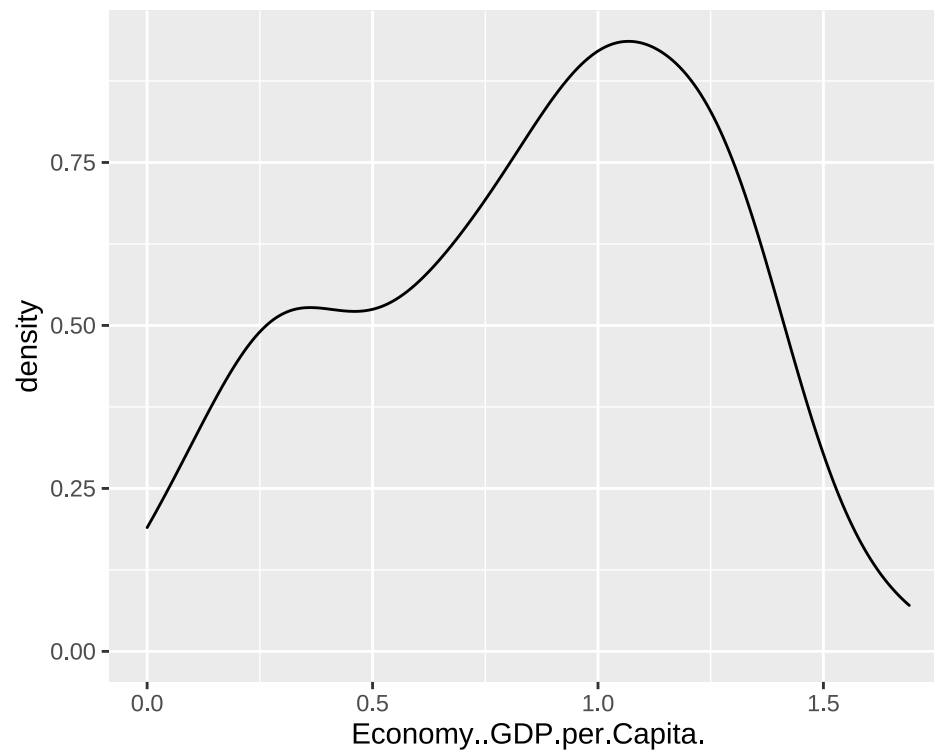


Jak widać, najwięcej krajów przyjmuje wskaźnik szczęścia na poziomie ok. 6 w skali 0-10. Co ciekawe, dla żadnego kraju nie przekracza on 8.

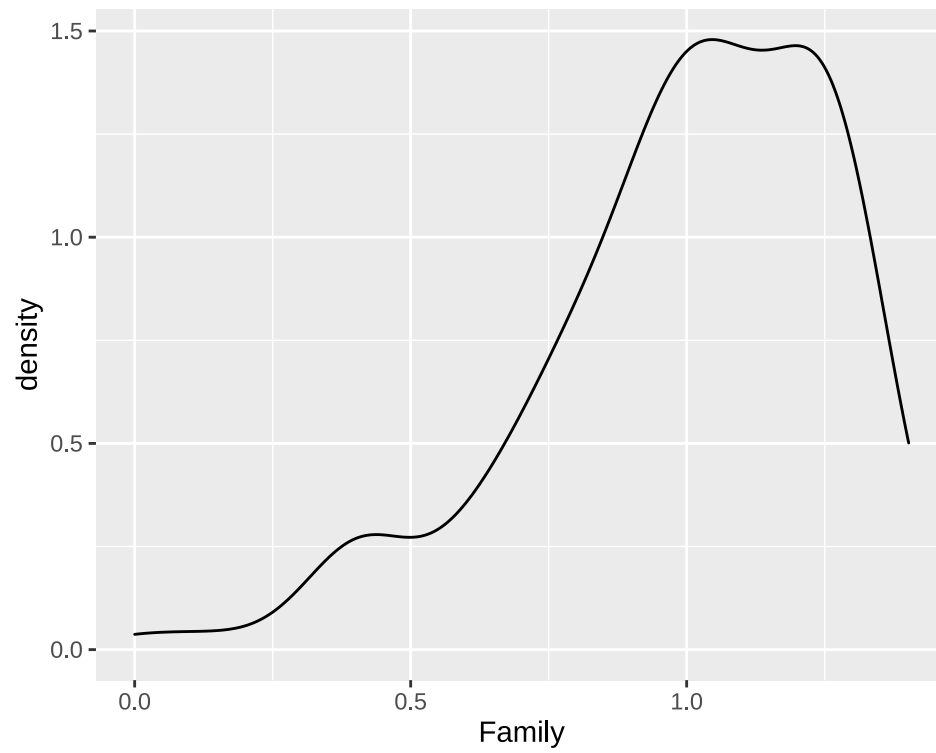
2. Standard Error



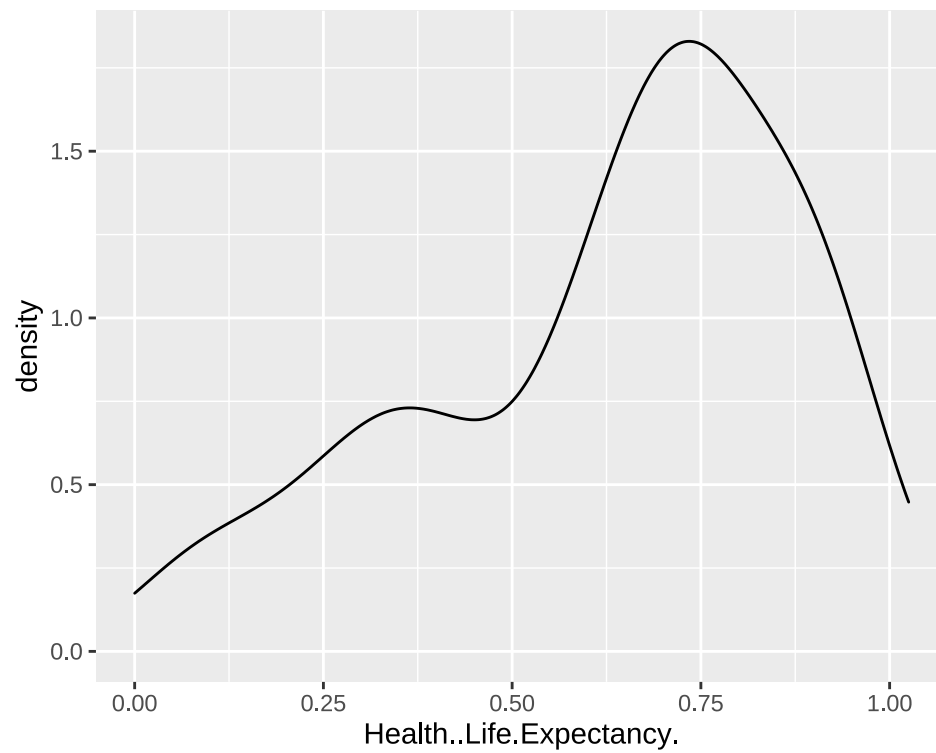
3. GDP per capita



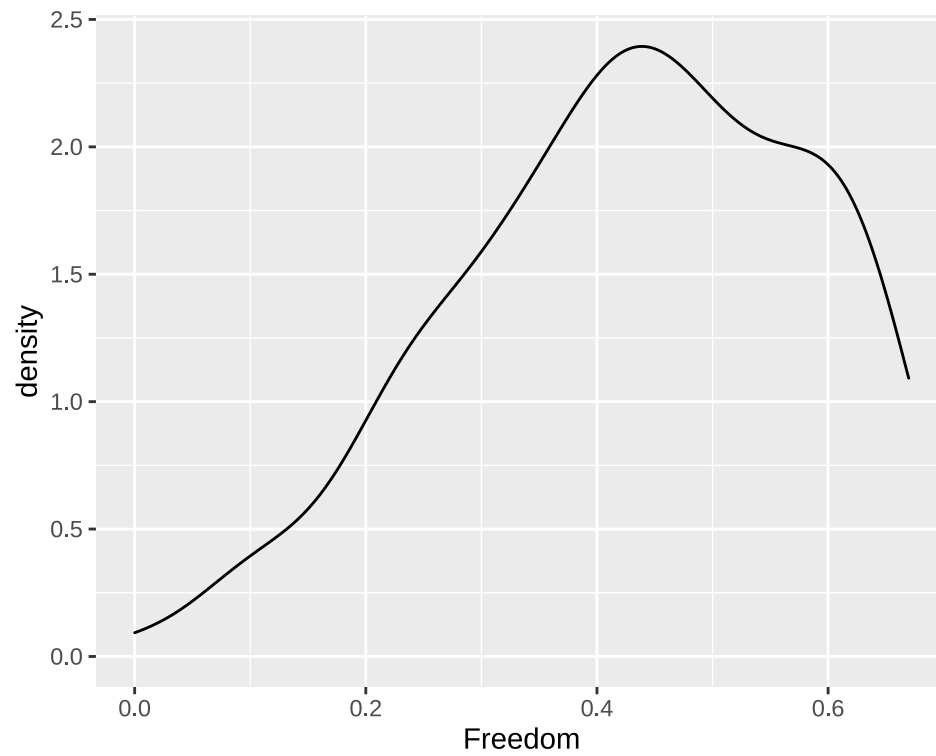
4. Family



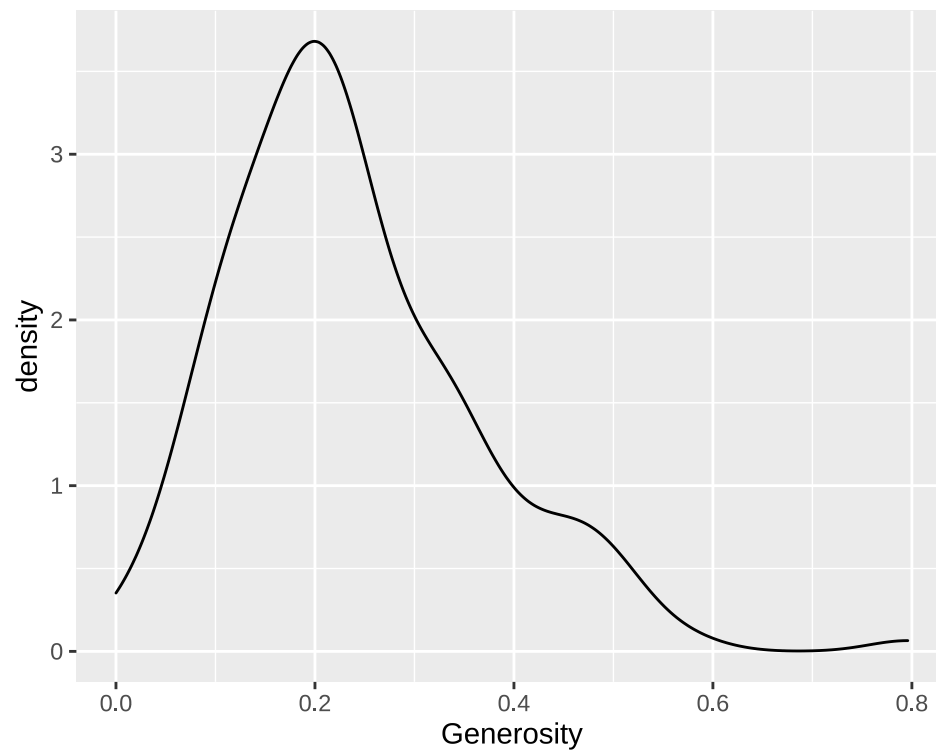
5. Health



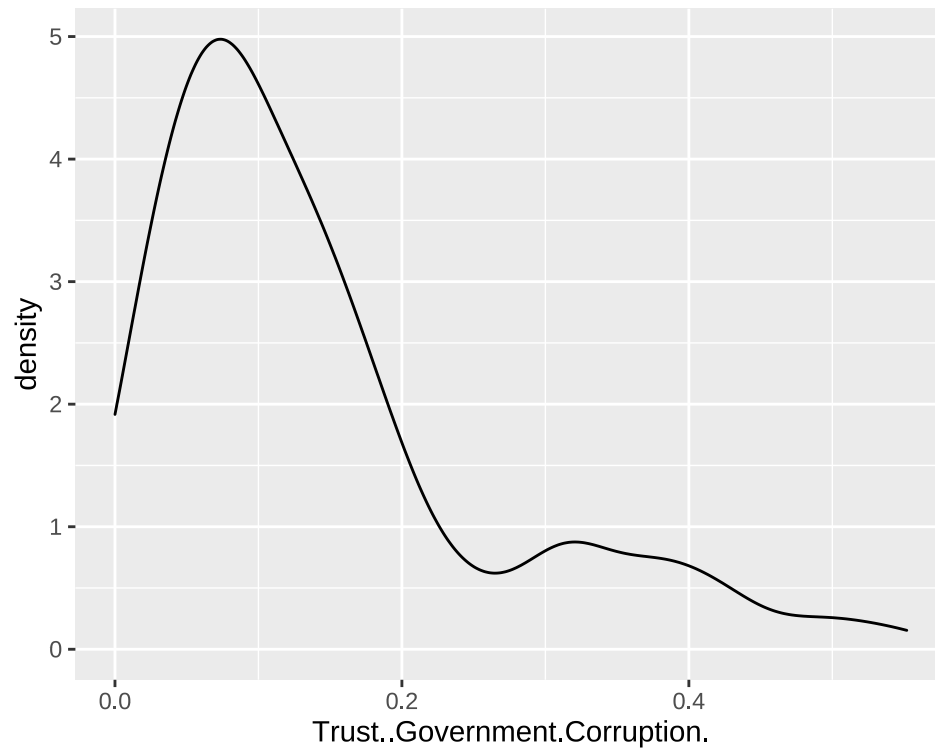
6. Freedom



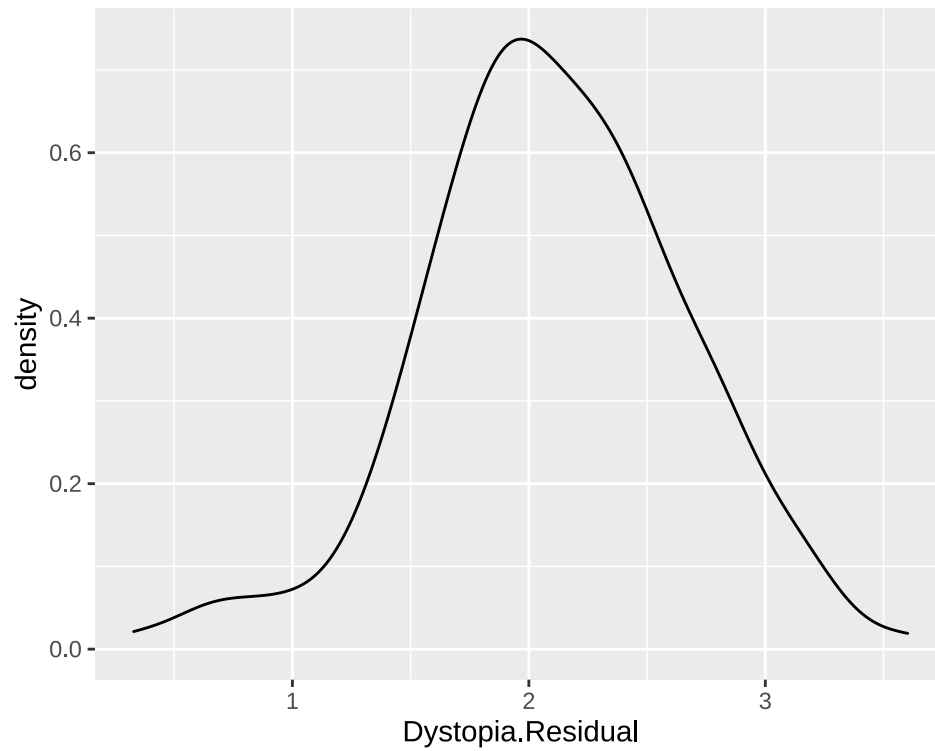
7. Generosity



8. Perceptions of corruption



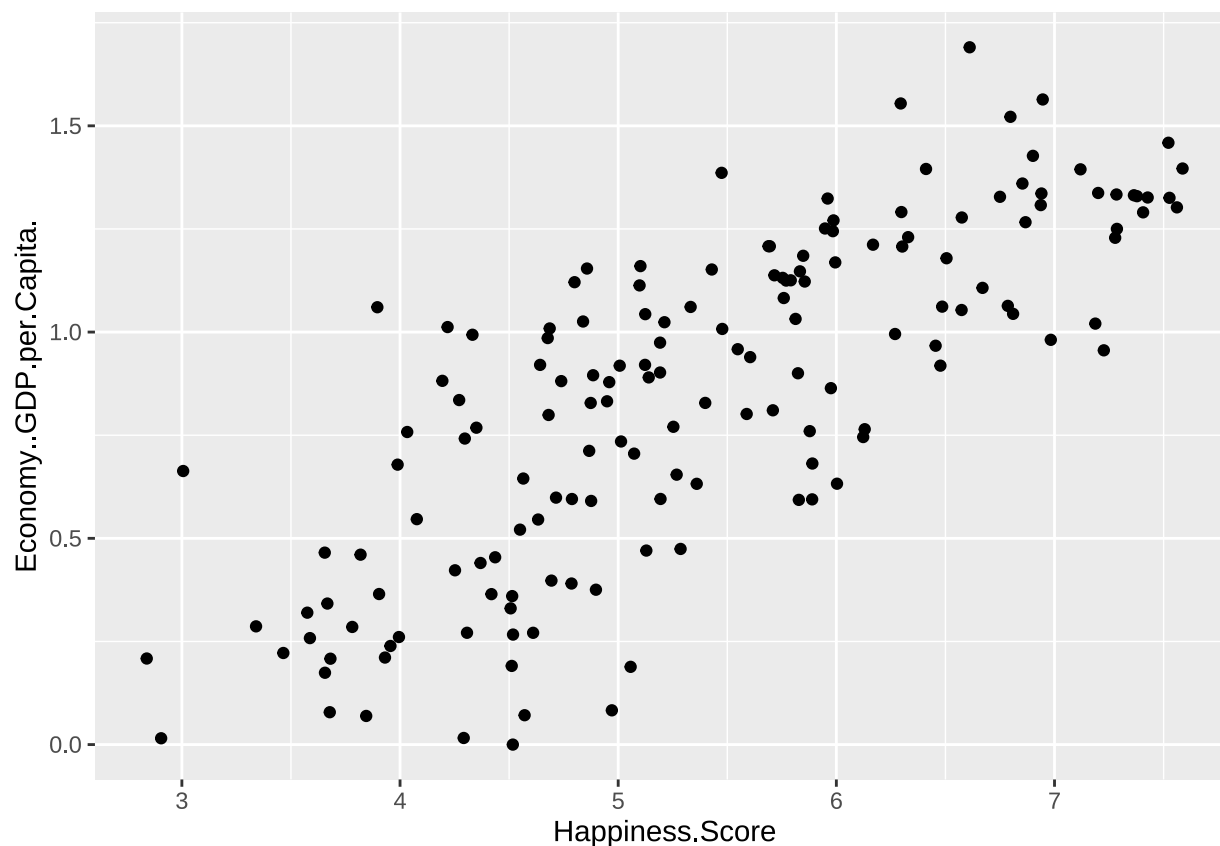
9. Dystopia Residual



Zależności między zmiennymi

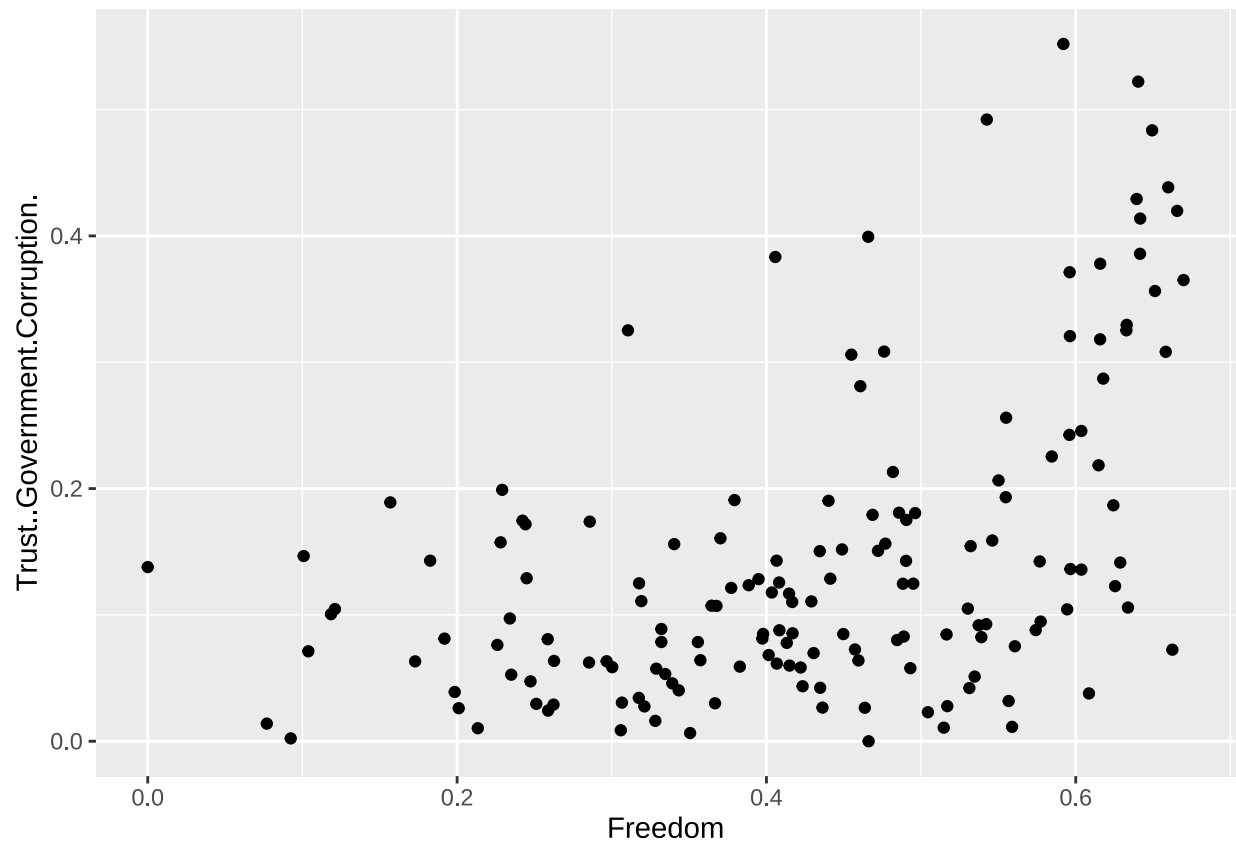
Jak zostało wcześniej wspomniane, większość z opisywanych czynników nie wpływa w żaden sposób na pozycję państwa w rankingu, która jest zależna jedynie od czynnika `Score`. Sprawdźmy jednak, czy ocena obywateli danego państwa ma związek z wartościami niektórych z czynników.

Zależność między Happiness Score i GDP per capita

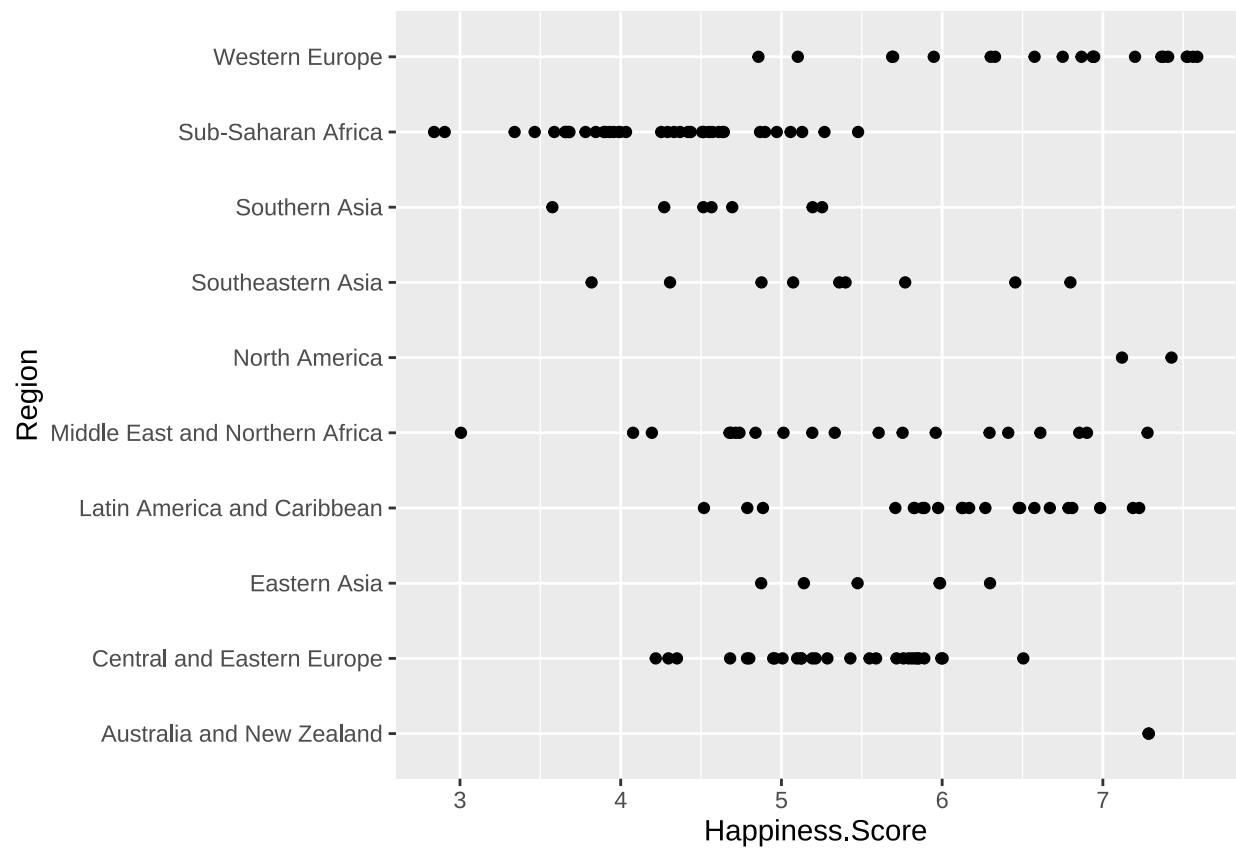


Jak widać, bogatsze państwa mają z reguły większą wartość czynnika `Happiness Score`, a więc ich obywatele są statystycznie szczęśliwsi od obywateli państw z mniejszym współczynnikiem `Economy (GDP per capita)`.

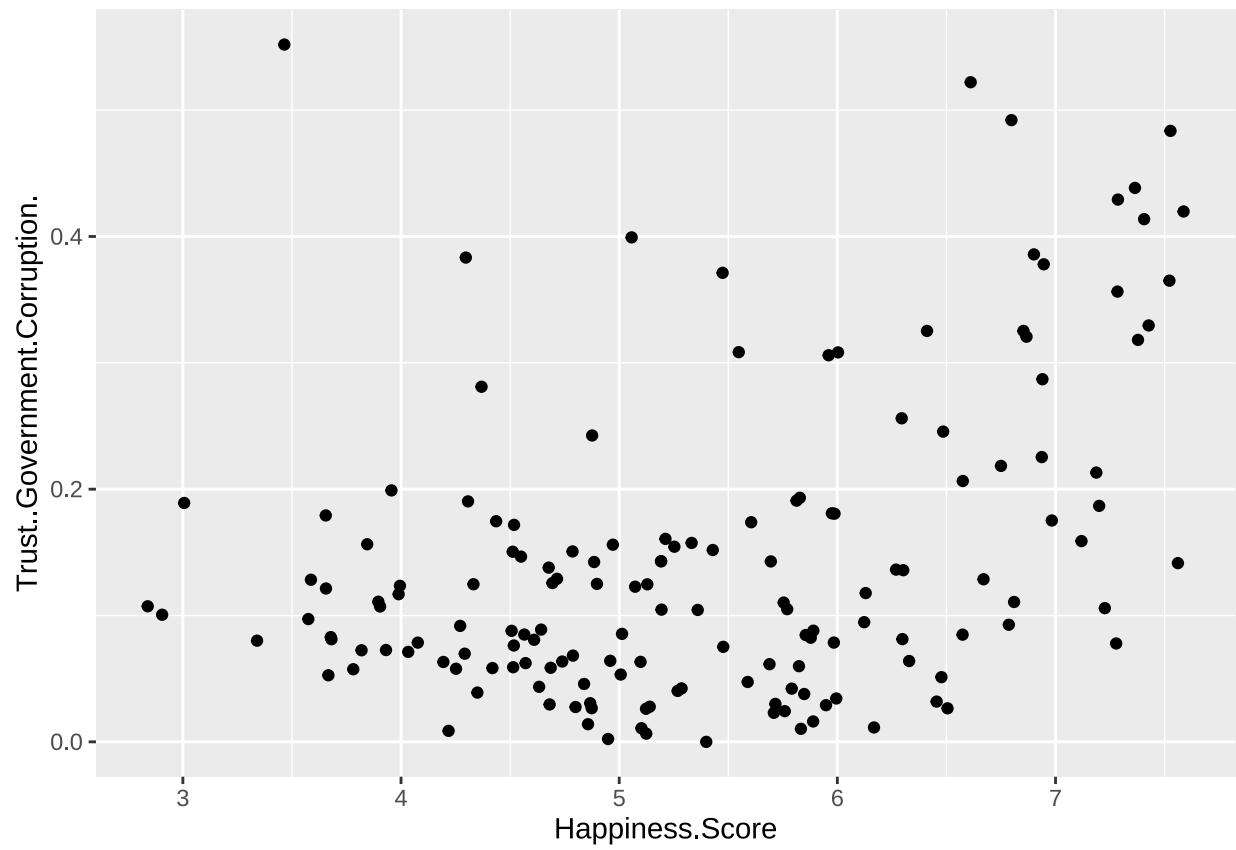
Zależność między Freedom i Trust



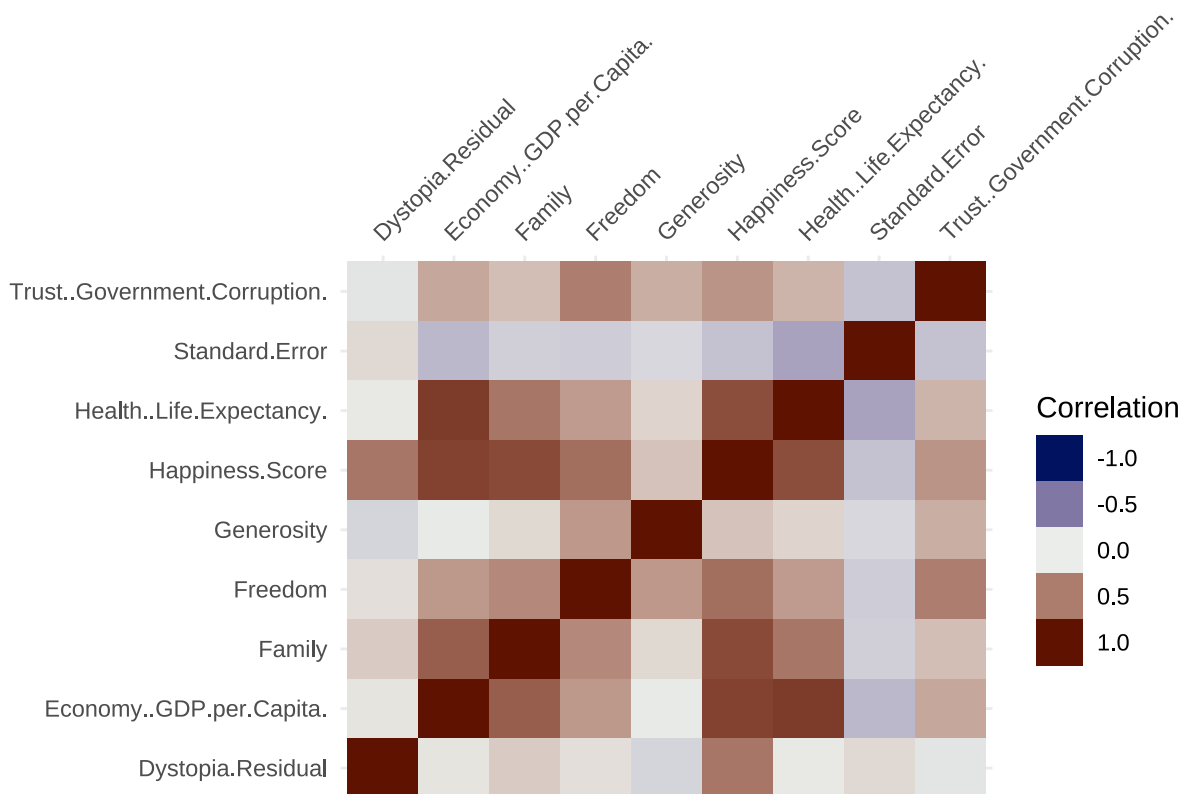
Zależność między Region i Happiness Score



Zależność między Trust i Happiness Score



Korelacje między zmiennymi



Wnioski

Na podstawie analizy danego zbioru, można wywnioskować, że najszczęśliwsi byli mieszkańcy Europy Zachodniej, Ameryki Północnej oraz Australii i Nowej Zelandii, zaś najniższy współczynnik szczęścia odnotowano w Afryce Północnej i Środkowowschodniej. Od razu nasuwa się tutaj pewna zależność, że im bogatsze państwa, tym większy współczynnik szczęścia, co udowadnia również umieszczony wyżej wykres. Warto także zauważyć, że najbardziej wolne kraje mają wyższy wskaźnik zaufania do rządu w sprawie korupcji niż reszta i to one zwykle przyjmują wyższy współczynnik **Happiness Score**.