
Ekonometria Przestrzenna Projekt

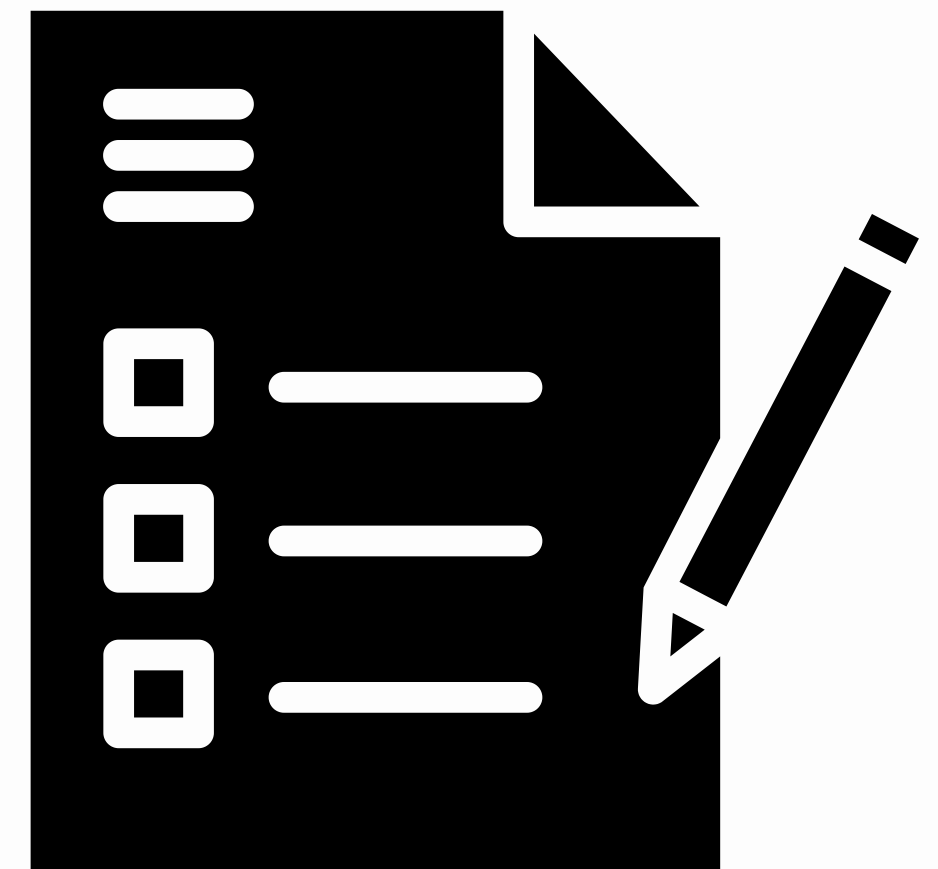


POLITECHNIKA GDAŃSKA

PRZYGOTOWALI: JULIA URBAN, JULIA SULKA, BARBARA MICHALSKA, JAKUB PILARSKI

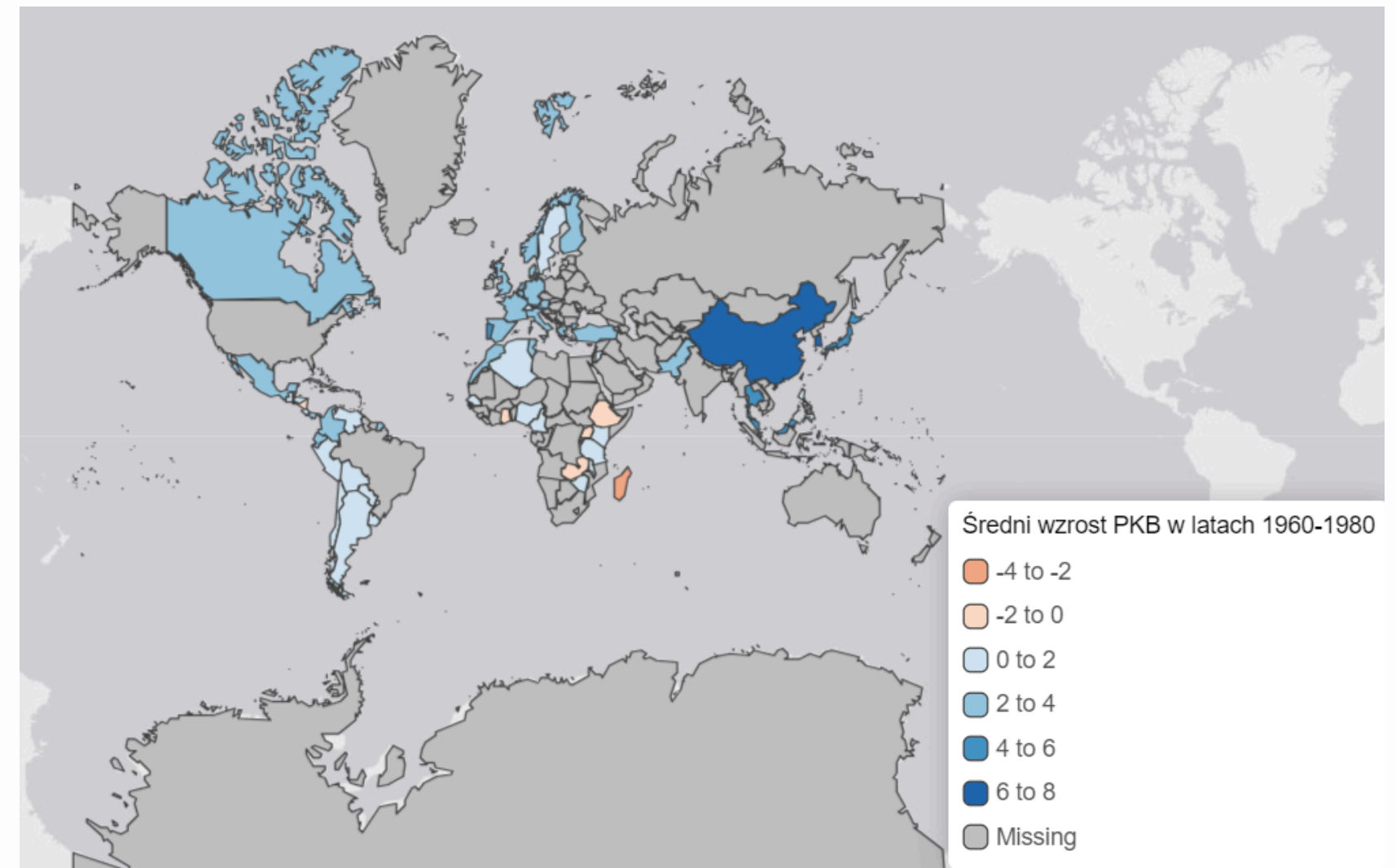
Agenda

01. Wstęp
02. Opis danych
03. Analiza opisowa
04. Ekspoloracyjna analiza danych
05. Modelowanie ekonometryczne
06. Wnioski



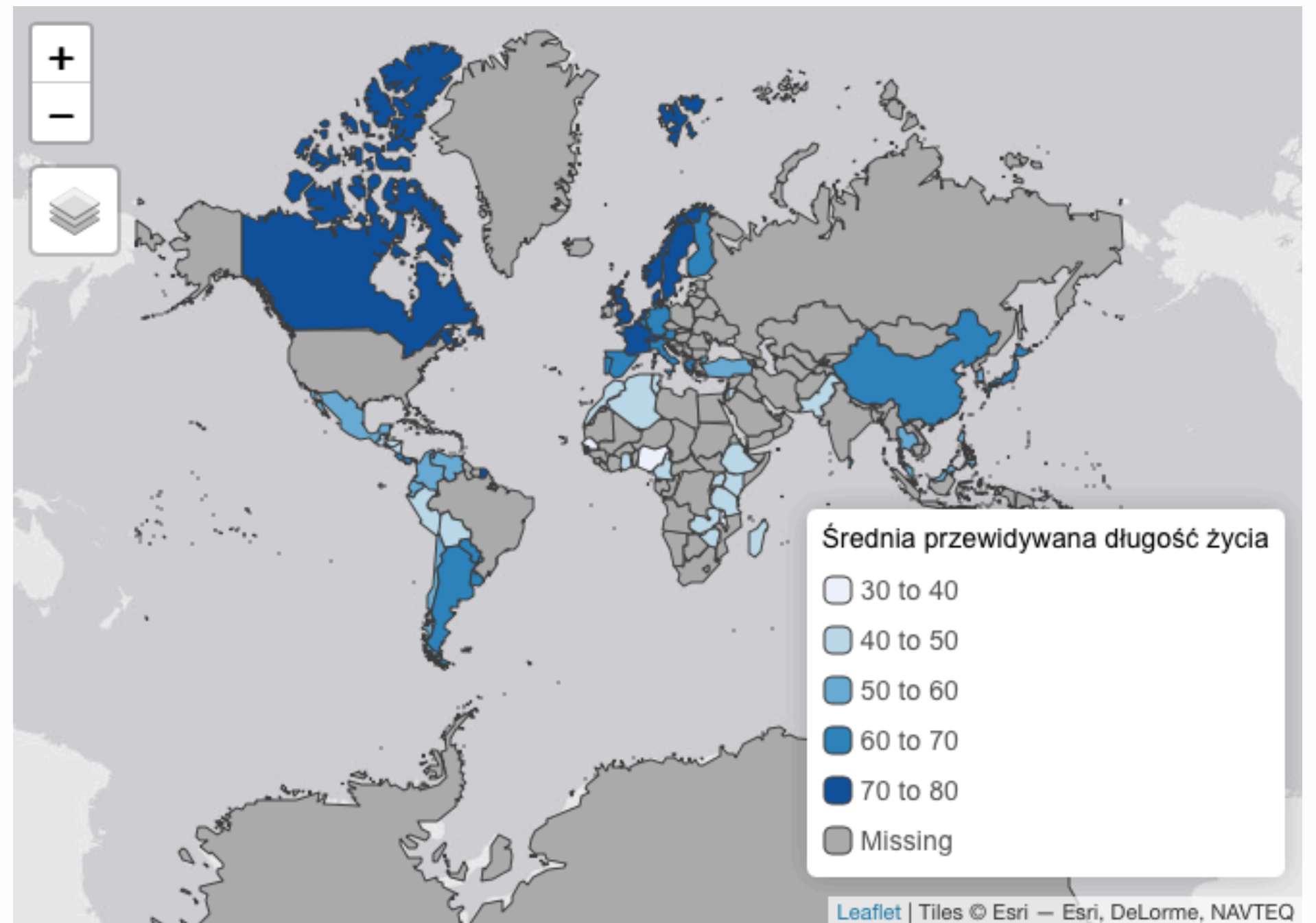
Zmienna zależna

Zmienną objaśnianą w niniejszej analizie jest średnioroczna stopa wzrostu produktu krajowego brutto (PKB) per capita w latach 1960–1980 (gdp_growth_1960_1980). Wybór tej zmiennej wynika z chęci zbadania długookresowych determinant wzrostu gospodarczego oraz ich potencjalnych zależności przestrzennych. Wskaźnik ten pozwala ocenić tempo rozwoju ekonomicznego poszczególnych krajów w badanym okresie oraz stanowi podstawę do identyfikacji wzorców przestrzennych i regionalnych różnic.



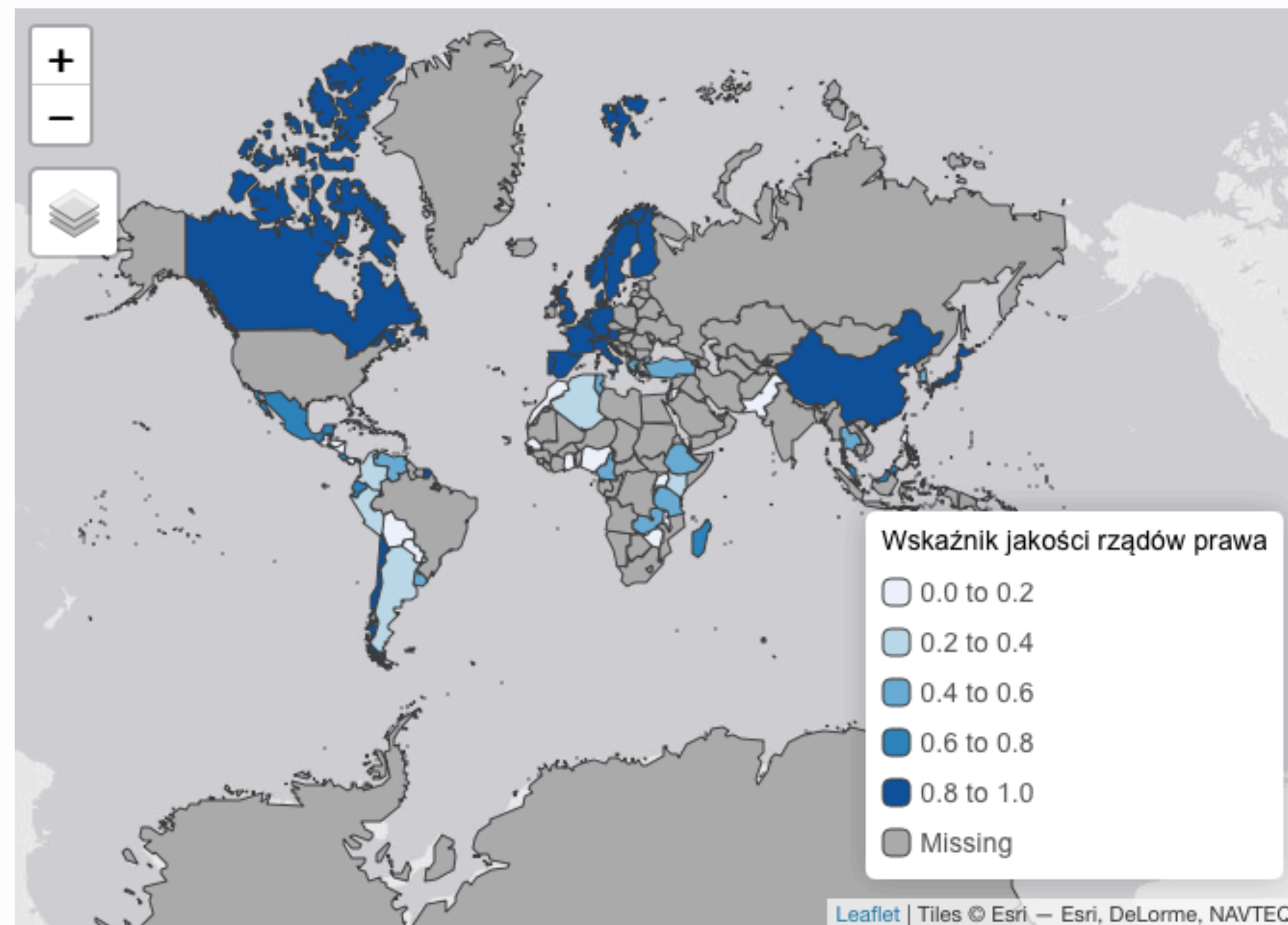
Średnia przewidywana długość życia

Przeciętna liczba lat, jaką statystycznie oczekuje się, że przeżyje osoba urodzona w danym kraju. Wysoka wartość zwykle wskazuje na lepszy poziom zdrowia i warunki życia.



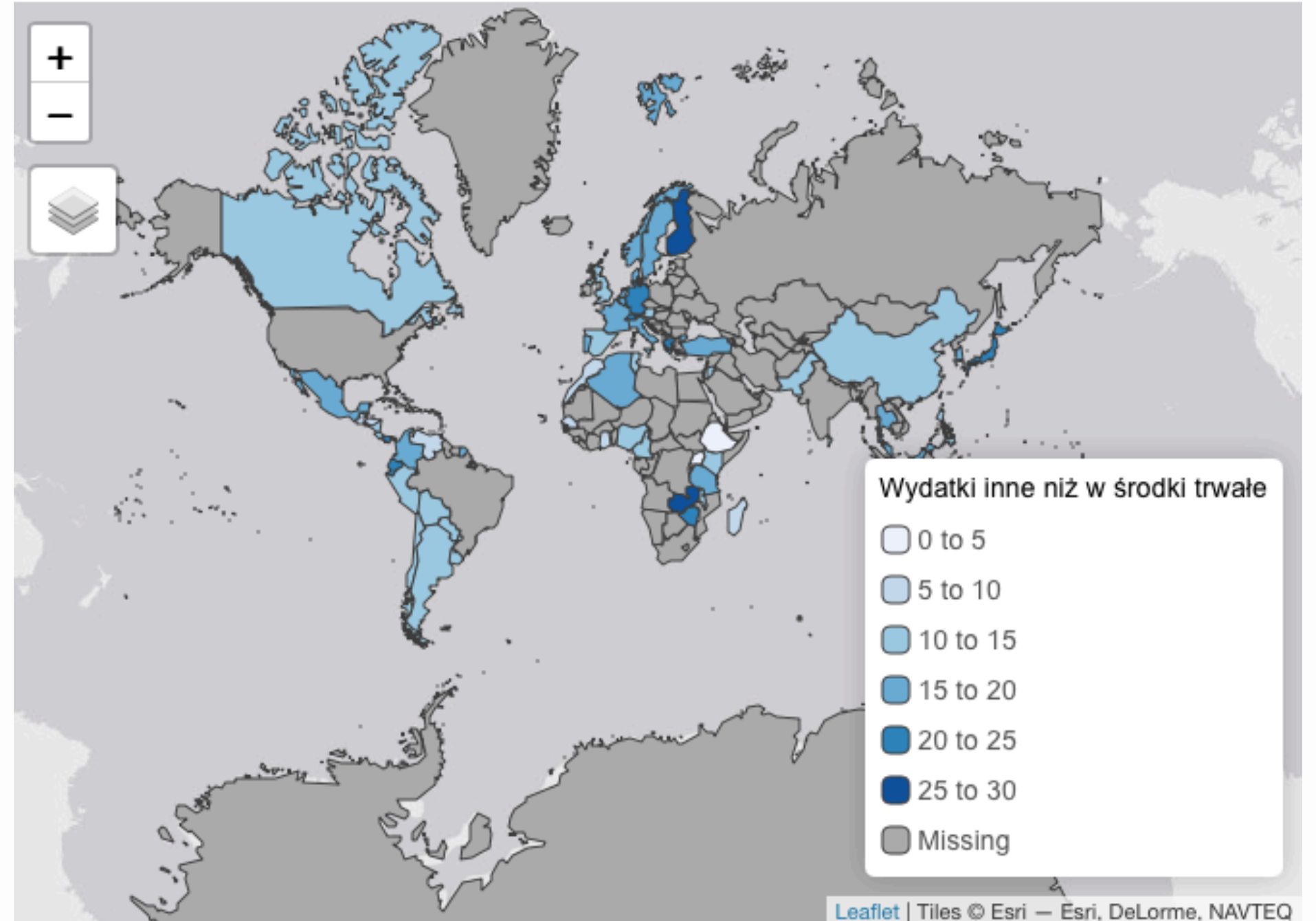
Wskaźnik jakości rządów prawa

Mierzy stopień przestrzegania prawa, stabilność systemu prawnego oraz skuteczność wymiaru sprawiedliwości w danym kraju. Wyższe wartości oznaczają lepszą jakość instytucji i często korelują z wyższym rozwojem gospodarczym.



Wydatki inne niż środki trwałe

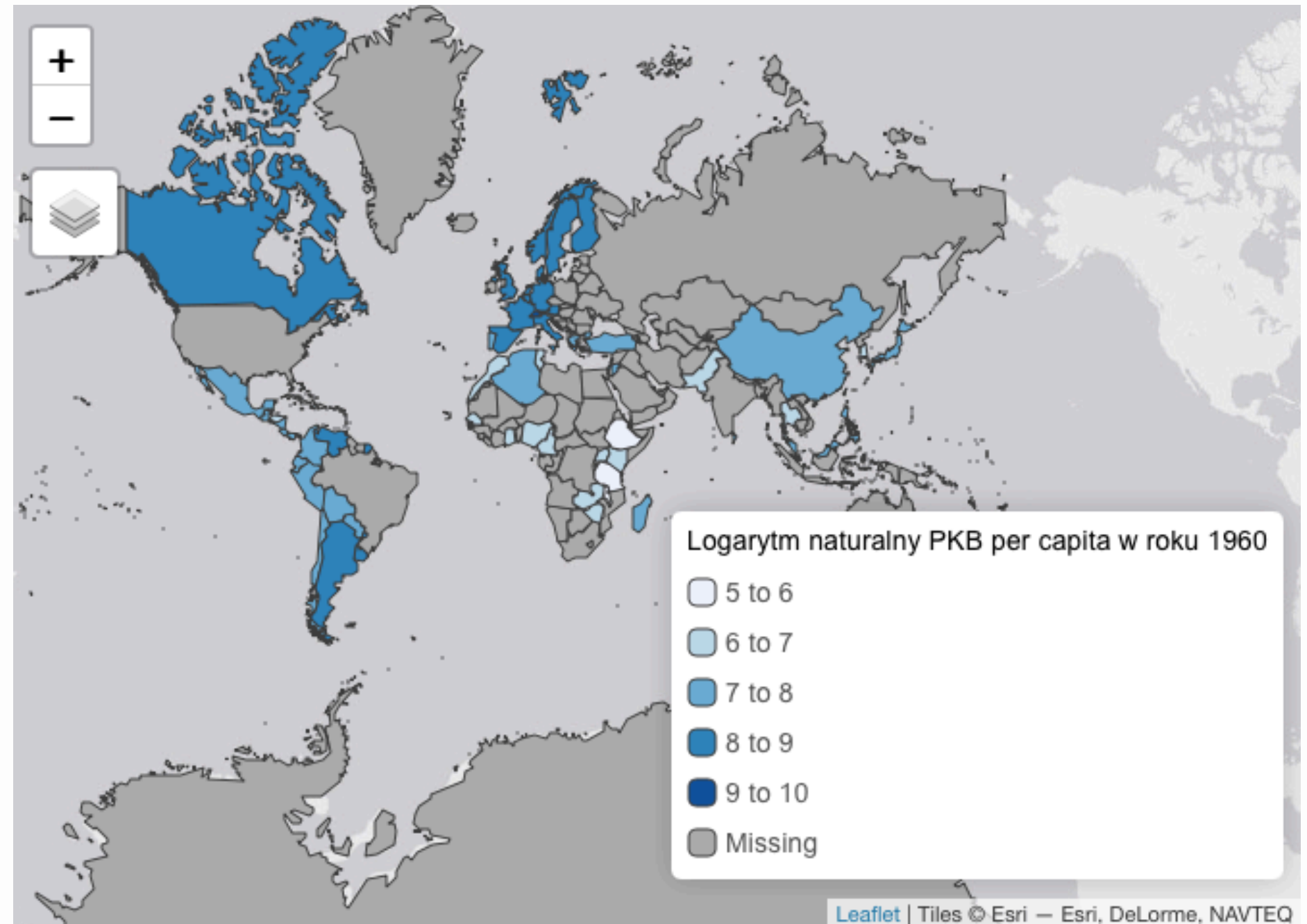
Procent wydatków innych niż w środki trwałe, np. w kapitał ludzki, infrastrukturę społeczną.



Logarytm naturalny PKB per capita w roku 1960

Ilościowa zmienna ciągła wyrażająca logarytm naturalny PKB per capita w roku 1960.

Wartości tej zmiennej odzwierciedlają poziom zamożności krajów w początkowym roku analizy. Dzięki logarytmicznej transformacji możliwe jest zmniejszenie wpływu wartości odstających oraz uzyskanie bardziej symetrycznego rozkładu, co sprzyja modelowaniu regresyjnemu. Wysokie wartości GDPsh560 oznaczają wyższy poziom dochodu per capita w 1960 r., co może mieć wpływ na tempo późniejszego rozwoju gospodarczego.



Macierz wag przestrzennych

Została ona oparta na relacjach sąsiedztwa pomiędzy krajami uwzględnionymi w zbiorze danych. W pierwszej kolejności utworzono reprezentację przestrzenną krajów w postaci punktów z przypisanymi współrzędnymi geograficznymi. Następnie określono, które kraje są sobie geograficznie bliskie i mogą być traktowane jako sąsiedzi — na tej podstawie zbudowano graf sąsiedztwa.

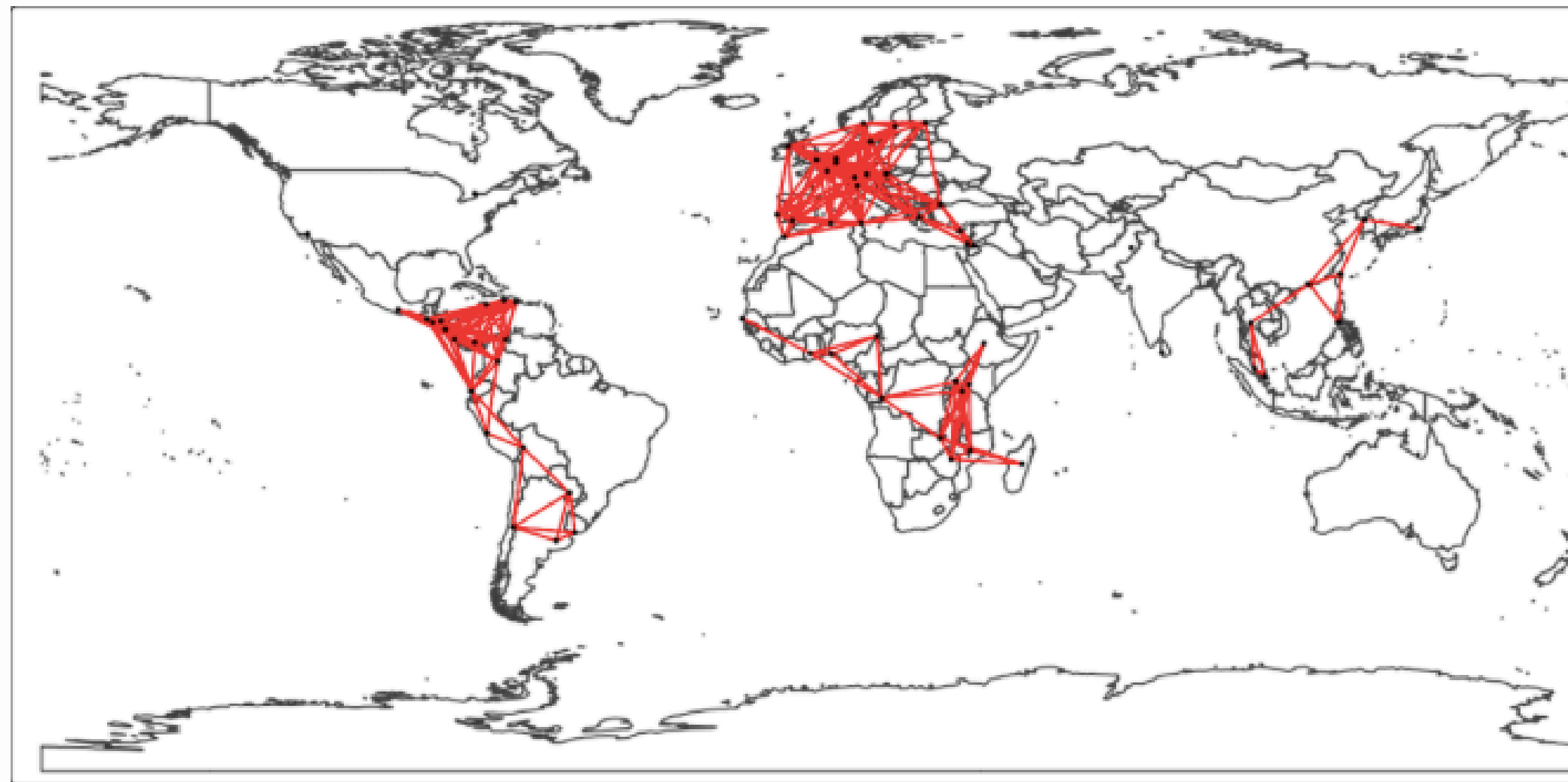


Tabela z przypisanymi sąsiadami

Zaprezentowano tabelę z przypisanymi sąsiadami dla wybranych krajów, co ułatwia interpretację powiązań przestrzennych.

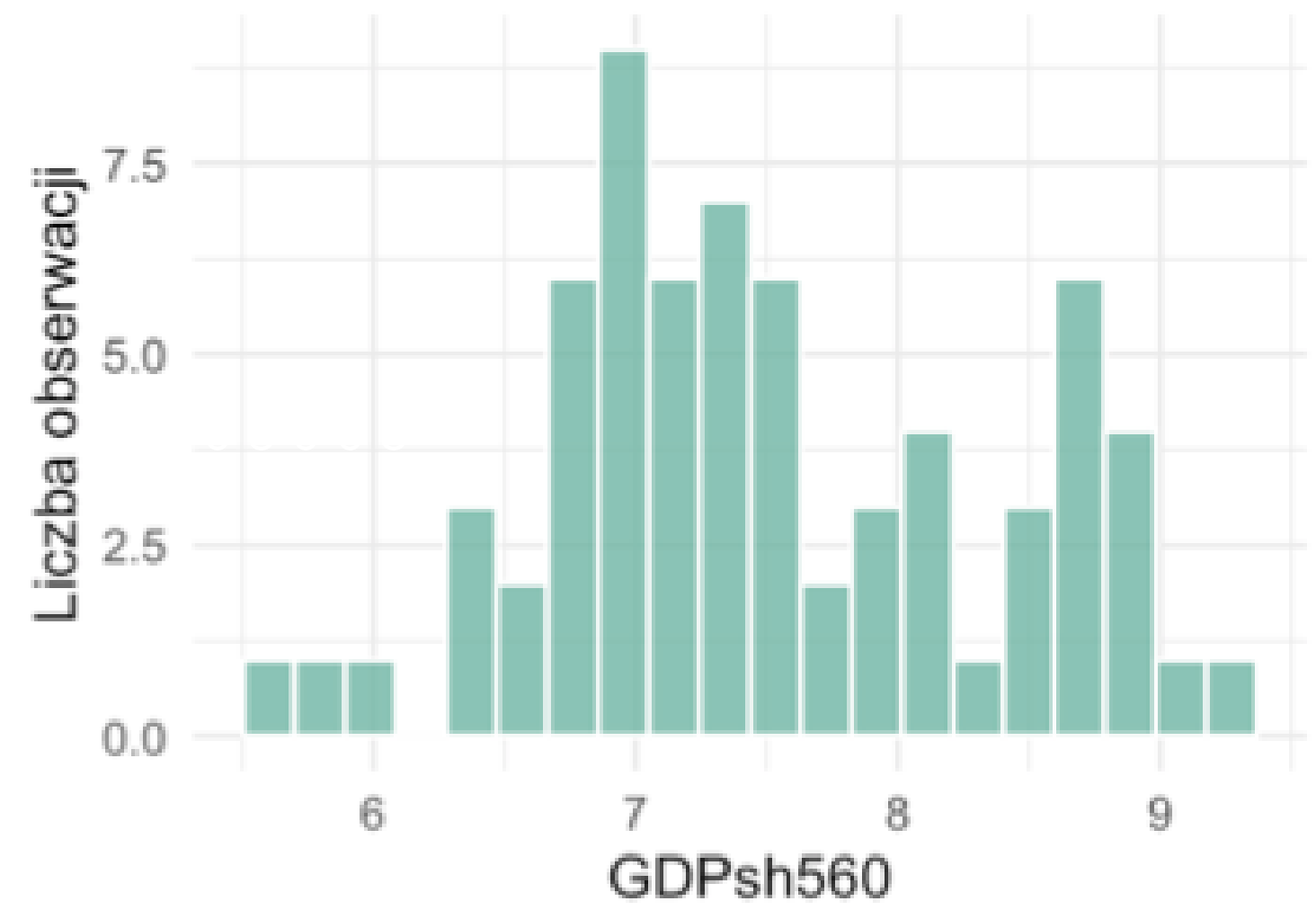
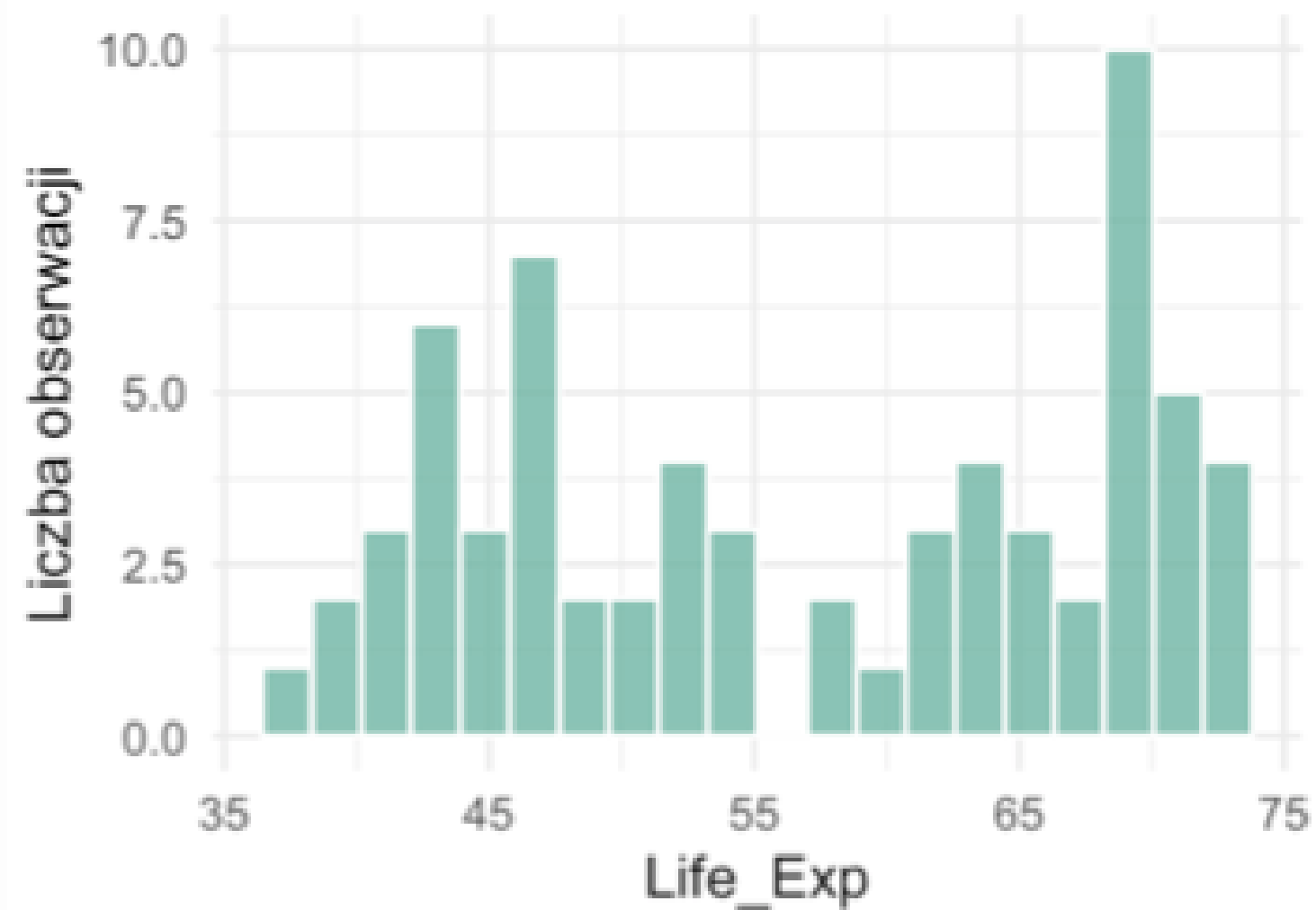
country	neighbors
Algeria	Morocco, Tunisia, Austria, Belgium, France, Germany, Italy, Netherlands, Portugal, Spain, Switzerland, United Kingdom
Cameroon	Congo - Brazzaville, Ghana, Nigeria
Congo - Brazzaville	Cameroon, Ghana, Nigeria, Tanzania, Uganda, Zambia
Ethiopia	Kenya, Tanzania, Uganda
Ghana	Cameroon, Congo - Brazzaville, Nigeria, Senegal
Kenya	Ethiopia, Malawi, Tanzania, Uganda, Zambia, Zimbabwe
Madagascar	Malawi, Zambia, Zimbabwe
Malawi	Kenya, Madagascar, Tanzania, Uganda, Zambia, Zimbabwe
Morocco	Algeria, Tunisia, France, Portugal, Spain, United Kingdom
Nigeria	Cameroon, Congo - Brazzaville, Ghana

Statystyki opisowe

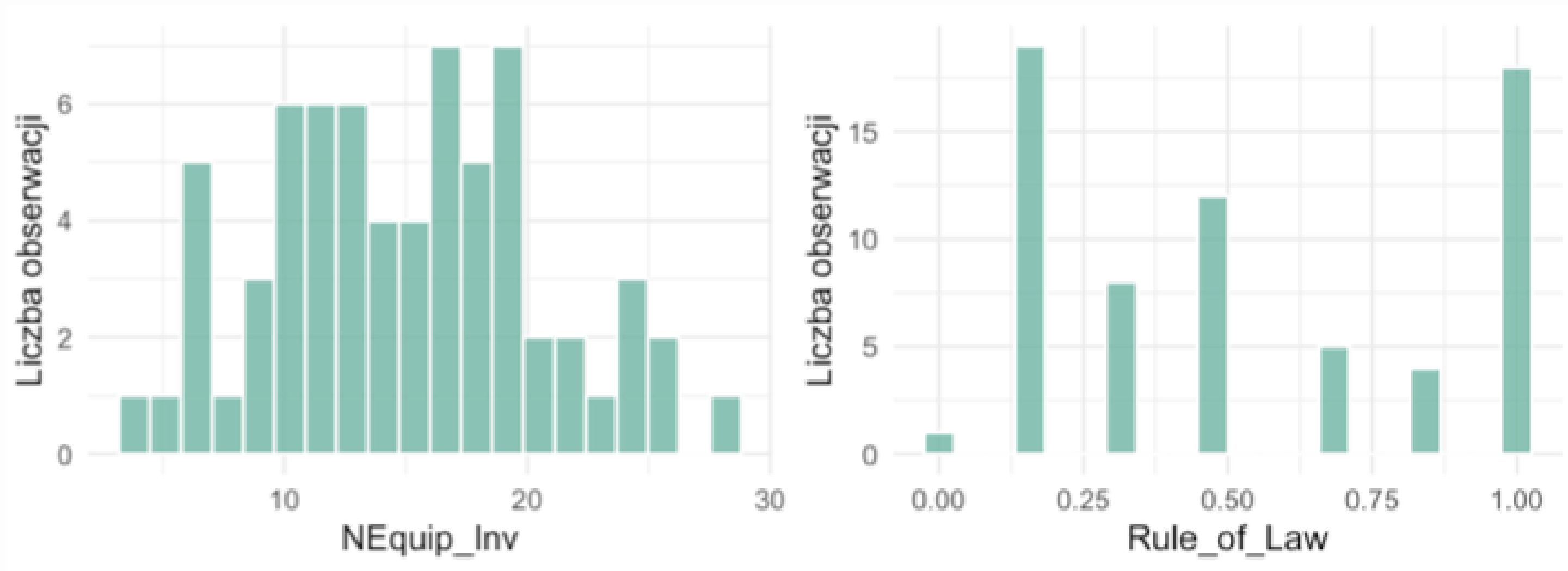
Podstawowe statystyki opisowe wybranych zmiennych

Zmienna	Średnia	Odchylenie standardowe	Minimum	Mediana	Maksimum
Oczekiwana długość życia	57	11	38	57	73
Logarytm naturalny PKB per capita w roku 1960	8	1	6	7	9
Procent wydatków innych niż w środki trwałe	15	5	4	15	28
Jakość rządów prawa	1	0	0	0	1

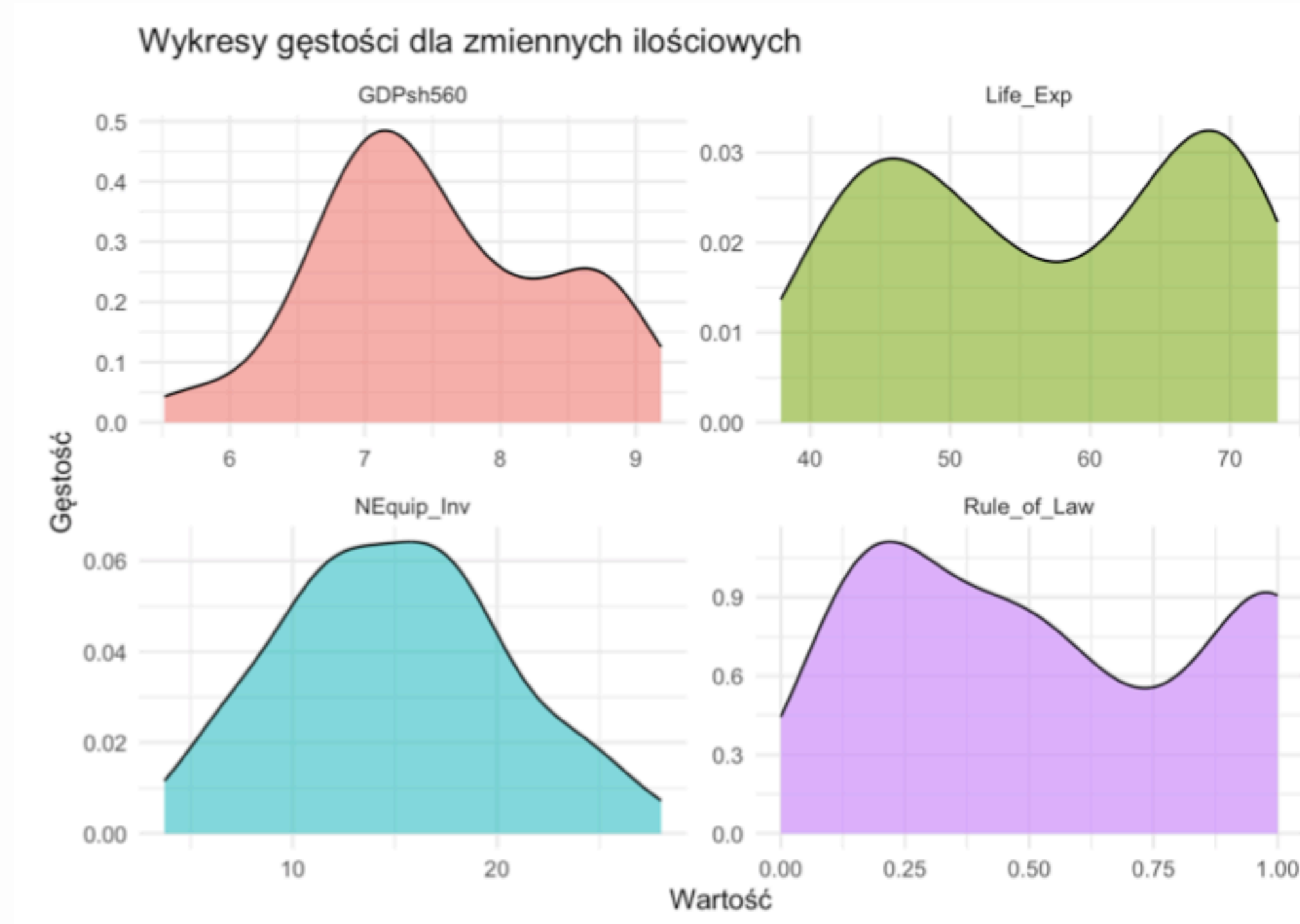
Histogramy



Histogramy



Wykresy gęstości



Korelacja zmiennych ilościowych

Macierz korelacji

	Life_Exp	GDPsh560	NEquip_Inv	Rule_of_Law
Life_Exp	1.000	0.859	0.456	0.711
GDPsh560	0.859	1.000	0.371	0.691
NEquip_Inv	0.456	0.371	1.000	0.415
Rule_of_Law	0.711	0.691	0.415	1.000

Zmienna jakościowa

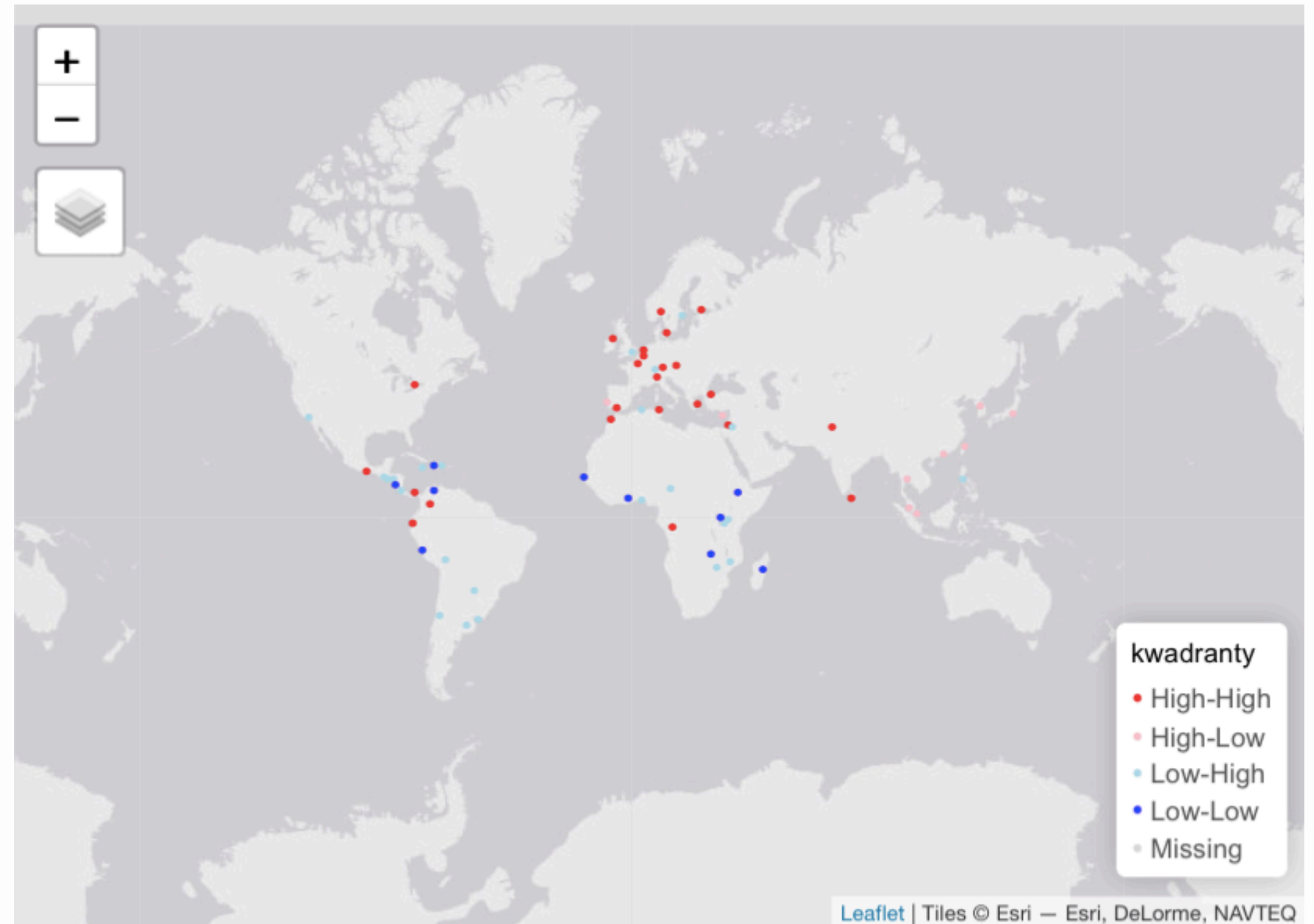
Zmienna SubSahara dzieli analizowane państwa na dwie grupy: należące do regionu Afryki Subsaharyjskiej (1) oraz pozostałe (0). W zestawieniu danych: 80,6% państw (54 obserwacje) nie należy do Afryki Subsaharyjskiej, natomiast 19,4% państw (13 obserwacji) pochodzi z tego regionu. Taki rozkład wskazuje na znaczną przewagę państw spoza Afryki Subsaharyjskiej w analizowanej próbie, co warto uwzględnić przy interpretacji wyników modeli przestrzennych czy statystyk opisowych.

Rozkład zmiennej SubSahara

Kategoria	Liczba obserwacji	Procent
Nie	54	80.6
Tak (Afryka Subsaharyjska)	13	19.4

Wskaźniki LISA i ich wizualizacja

Wskaźniki lokalne autokorelacji przestrzennej (LISA — Local Indicators of Spatial Association) służą do wykrywania przestrzennych wzorców korelacji lokalnych, czyli np. skupisk wartości wysokich lub niskich w danych geograficznych.



Klasyczny model OLS

Model OLS został oszacowany w celu zbadania czynników wpływających na średni wzrost PKB w latach 1960–1980. Wyniki wskazują, że wszystkie uwzględnione zmienne są statystycznie istotne ($p < 0,05$), a dopasowanie modelu jest wysokie ($R^2 = 0,724$).

Life_Exp (średnia długość życia): ma istotny dodatni wpływ na wzrost gospodarczy. Sugeruje to, że wyższy poziom zdrowia i warunków życia sprzyja rozwojowi gospodarczemu.

NEquip_Inv (wydatki inne niż w środki trwałe): również wpływa pozytywnie – inwestycje w kapitał ludzki i infrastrukturę społeczną mogą stymulować wzrost.

Rule_of_Law (praworządność): istotny pozytywny wpływ sugeruje, że lepsze instytucje i stabilność prawna wspierają rozwój gospodarczy.

GDPsh560 (PKB per capita w 1960): wykazuje istotny negatywny wpływ, co można interpretować jako efekt doganiania – kraje biedniejsze w 1960 rozwijały się szybciej.

	Estimate	Std. Error	t value	Pr(> t)	
(Intercept)	11.61455	1.72527	6.732	8.25e-09	***
Life_Exp	0.10708	0.02510	4.267	7.42e-05	***
NEquip_Inv	0.06478	0.02695	2.404	0.0194	*
Rule_of_Law	2.81040	0.56725	4.954	6.62e-06	***
GDPsh560	-2.34865	0.30176	-7.783	1.41e-10	***
SubSahara	-2.52823	0.43942	-5.753	3.46e-07	***

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1					
Residual standard error: 1.003 on 58 degrees of freedom					
(179 observations deleted due to missingness)					
Multiple R-squared: 0.7238, Adjusted R-squared: 0.7					
F-statistic: 30.4 on 5 and 58 DF, p-value: 4.899e-15					

Model przestrzenny SAR

Pseudo $R^2 \approx 0.73$ – bardzo dobra zgodność modelu z danymi.

Rho (ρ) = 0,052 – dodatni współczynnik opóźnienia przestrzennego sugeruje, że wzrost PKB w danym kraju może być nieznacznie powiązany z sytuacją gospodarczą w krajach sąsiednich.

Test LR (test ilorazu wiarygodności): wartość testu = 0,365; p-value = 0,546

Test Walda (z-test dla ρ): $z = 0,627$; p-value = 0,531

W testach brak podstaw do odrzucenia hipotezy zerowej ($H_0: \rho = 0$), co oznacza, że efekt przestrzennego opóźnienia nie jest statystycznie istotny. Innymi słowy, nie ma silnych dowodów na przestrzenną zależność pomiędzy wzrostem PKB w krajach sąsiadujących.

```
Call: lagsarlm(formula = gdp_growth_1960_1980 ~ Life_Exp + NEquip_Inv +  
  Rule_of_Law + GDPsh560 + SubSahara, data = data_sar, listw = lw,  
  zero.policy = TRUE)
```

Residuals:

Min	1Q	Median	3Q	Max
-2.468093	-0.654171	-0.013437	0.638347	2.770030

Type: lag

Regions with no neighbours included:

47 55 57 64 68 82 119 137 138 150 157 188 204

Coefficients: (asymptotic standard errors)

	Estimate	Std. Error	z value	Pr(> z)
(Intercept)	11.542909	1.637030	7.0511	1.775e-12
Life_Exp	0.109463	0.023970	4.5667	4.955e-06
NEquip_Inv	0.063978	0.025591	2.5000	0.01242
Rule_of_Law	2.746156	0.542639	5.0607	4.176e-07
GDPsh560	-2.363660	0.288718	-8.1868	2.220e-16
SubSahara	-2.447860	0.425349	-5.7549	8.667e-09

Rho: 0.052005, LR test value: 0.36506, p-value: 0.54571

Asymptotic standard error: 0.082926

z-value: 0.62713, p-value: 0.53058

Wald statistic: 0.39329, p-value: 0.53058

Log likelihood: -87.66412 for lag model

ML residual variance (sigma squared): 0.90525, (sigma: 0.95144)

Number of observations: 64

Number of parameters estimated: 8

AIC: 191.33, (AIC for lm: 189.69)

LM test for residual autocorrelation

test value: 0.65999, p-value: 0.41656

Test Breusch-Pagan i Moran I

Test Breuscha–Pagana wskazuje na obecność ogólnej heteroskedastyczności w modelu — oznacza to, że wariancja reszt nie jest stała i zmienia się w zależności od wartości zmiennych objaśniających. Z kolei test Morana I nie wykazuje przestrzennej autokorelacji reszt, co oznacza brak przestrzennego skupiania się dużych lub małych wartości reszt, czyli brak przestrzennej heteroskedastyczności.

Wyniki obu testów się nie wykluczają — przeciwnie: model cierpi na ogólną heteroskedastyczność, ale dobrze uwzględnia przestrzenną strukturę danych.

```
studentized Breusch-Pagan test
```

```
data:  ols_model  
BP = 11.252, df = 5, p-value = 0.04661
```

```
Moran I test under randomisation
```

```
data:  residuals_sar  
weights: lw  
n reduced by no-neighbour observations
```

```
Moran I statistic standard deviate =  
0.64025, p-value = 0.261
```

```
alternative hypothesis: greater
```

```
sample estimates:
```

```
Moran I statistic      Expectation
```

```
Variance
```

```
0.07500337
```

```
-0.02000000
```

```
0.02201833
```

Model przestrzenny SEM

Lambda = 0,1613 ze statystyką LR = 1,1685 i p-wartością 0,2797

Odrzucamy więc hipotezę o obecności istotnej autokorelacji przestrzennej w składniku błędu, ponieważ $p > 0,05$. Wskazuje to, że przestrzenna korelacja reszt nie jest statystycznie istotna — czyli model SEM nie wykazuje silnej autokorelacji przestrzennej w błędzie.

Świadczy to o braku potrzeby uwzględniania przestrzennej autokorelacji w błędach — zmienne wyjaśniające dobrze opisują zjawisko, a reszty nie są obciążone strukturą przestrzenną.

Residuals:

Min	1Q	Median	3Q	Max
-2.499158	-0.633651	-0.073223	0.640066	2.747212

Type: error

Regions with no neighbours included:

47 55 57 64 68 82 119 137 138 150 157 188 204

Coefficients: (asymptotic standard errors)

	Estimate	Std. Error	z value	Pr(> z)
(Intercept)	11.064232	1.682660	6.5754	4.851e-11
Life_Exp	0.112246	0.023563	4.7637	1.901e-06
NEquip_Inv	0.059982	0.025521	2.3502	0.01876
Rule_of_Law	2.687696	0.539717	4.9798	6.364e-07
GDPsh560	-2.298133	0.290100	-7.9219	2.442e-15
SubSahara	-2.365595	0.448071	-5.2795	1.295e-07

Lambda: 0.16127, LR test value: 1.1685, p-value: 0.27971

Asymptotic standard error: 0.12356

z-value: 1.3052, p-value: 0.19183

Wald statistic: 1.7035, p-value: 0.19183

Log likelihood: -87.2624 for error model

ML residual variance (sigma squared): 0.88477, (sigma: 0.94062)

Number of observations: 64

Number of parameters estimated: 8

AIC: 190.52, (AIC for lm: 189.69)

Model przestrzenny SLX

Wyniki modelu regresji przestrzennej SLX pokazują, że zmienne niezależne bezpośrednio wpływają na tempo wzrostu PKB w latach 1960–1980, natomiast żadna ze zmiennych przestrzennych (lagorytmowanych) nie jest istotna statystycznie ($p > 0,1$), co oznacza, że wpływ cech sąsiadów na wzrost PKB jest znikomy.

Reszty modelu charakteryzują się standardowym błędem 1,009, co wskazuje na umiarkowany rozrzut obserwowanych wartości względem wartości prognozowanych.

Test F dla całego modelu jest istotny ($F = 15,46$, $p < 0,001$), co świadczy o tym, że model jest statystycznie istotny i zmienne objaśniające razem dobrze tłumaczą zmienność wzrostu PKB.

Residuals:

Min	1Q	Median	3Q	Max
-2.1192	-0.4919	0.1434	0.5603	2.4401

Coefficients:

	Estimate	Std. Error	t value	Pr(> t)	
(Intercept)	12.01427	1.86784	6.432	3.69e-08	***
Life_Exp	0.10591	0.02660	3.982	0.00021	***
NEquip_Inv	0.06603	0.02820	2.342	0.02298	*
Rule_of_Law	2.88499	0.59071	4.884	9.98e-06	***
GDPsh560	-2.37123	0.34640	-6.845	8.00e-09	***
SubSahara	-2.22364	0.76838	-2.894	0.00551	**
Life_Exp_lag	-0.03383	0.03620	-0.934	0.35436	
NEquip_Inv_lag	0.04965	0.03842	1.292	0.20184	
Rule_of_Law_lag	0.23021	0.82078	0.280	0.78020	
GDPsh560_lag	0.10810	0.26060	0.415	0.67994	
SubSahara_lag	-0.86298	0.86680	-0.996	0.32397	

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 1.009 on 53 degrees of freedom
Multiple R-squared: 0.7447, Adjusted R-squared: 0.6965
F-statistic: 15.46 on 10 and 53 DF, p-value: 1.842e-12

Model przestrzenny Durbin

Parametr przestrzennej autoregresji ($\rho = 0,221$) jest istotny statystycznie (test Wald: $p = 0,034$), co wskazuje na istnienie dodatniego wpływu wzrostu PKB w sąsiednich regionach na wzrost w danym regionie.

Test LM wykazał brak przestrzennej autokorelacji reszt ($p = 0,926$), co potwierdza poprawną specyfikację modelu.

Model Durbin pozwolił na bardziej szczegółowe uchwycenie zależności przestrzennych, jednak głównym czynnikiem różnicującym wzrost gospodarczy pozostają lokalne cechy regionów, a nie ich przestrzenne otoczenie.

```
Residuals:
    Min       1Q   Median       3Q      Max
-2.10483 -0.50645  0.12848  0.51948  2.56948

Type: mixed
Regions with no neighbours included:
 47 55 57 64 68 82 119 137 138 150 157 188 204
Coefficients: (asymptotic standard errors)
      Estimate Std. Error z value Pr(>|z|)
(Intercept)  10.974309   1.683922   6.5171 7.167e-11
Life_Exp      0.109377   0.023424   4.6695 3.019e-06
NEquip_Inv    0.063251   0.024857   2.5446 0.010941
Rule_of_Law   2.777868   0.517198   5.3710 7.830e-08
GDPsh560     -2.250192   0.305487  -7.3659 1.759e-13
SubSahara    -2.109297   0.674269  -3.1283 0.001758
lag.Life_Exp  -0.053271   0.032646  -1.6318 0.102726
lag.NEquip_Inv 0.047329   0.034618   1.3672 0.171568
lag.Rule_of_Law -0.044403   0.729991  -0.0608 0.951498
lag.GDPsh560   0.217358   0.235202   0.9241 0.355419
lag.SubSahara -0.617015   0.770323  -0.8010 0.423142

Rho: 0.22128, LR test value: 3.6751, p-value: 0.05523
Asymptotic standard error: 0.10443
      z-value: 2.1189, p-value: 0.034101
Wald statistic: 4.4896, p-value: 0.034101

Log likelihood: -83.4938 for mixed model
ML residual variance (sigma squared): 0.7783, (sigma: 0.88221)
Number of observations: 64
Number of parameters estimated: 13
AIC: 192.99, (AIC for lm: 194.66)
LM test for residual autocorrelation
test value: 0.0087093, p-value: 0.92565
```

Wnioski

Analizując modele SAR, SDM, SEM oraz SLX, można stwierdzić, że dominujące znaczenie mają czynniki lokalne nad efektami sąsiedztwa. Modele SDM i SLX nie wykazały istotności statystycznej zmiennych przestrzennych, co potwierdza, że wzrost PKB głównie zależy od wewnętrznych cech kraju. Z kolei modele SAR i SEM, pomimo uwzględnienia efektów przestrzennych, również wskazały na umiarkowane znaczenie zmiennych lokalnych w wyjaśnianiu wzrostu gospodarczego.

Testy Breuscha–Pagana potwierdziły homoskedastyczność modeli, co oznacza stabilność wariantywności reszt.

Wniosek: Polityki gospodarcze powinny skupić się na wzmocnieniu wewnętrznych czynników wzrostu, takich jak inwestycje w kapitał ludzki, przy równoczesnym monitorowaniu wpływu sąsiednich krajów na aspekty gospodarcze.

Bibliografia

1. Wleklińska Dagma, Materiały z Kursu Ekonometria przestrzenna (2024_2025_NST_online) - projekt zespołowy



Dziękujemy
za uwagę!

