

IX Szkoła Przywództwa Instytutu Wolności

Esej zaliczeniowy

Krzysztof Kozak

kmkozak87@gmail.com

## **ZAŁĄCZNIK: MODELOWANIE STATYSTYCZNE**

[https://github.com/evo-eco-gen/Zarzadzanie\\_Innowacje](https://github.com/evo-eco-gen/Zarzadzanie_Innowacje)

### **DANE**

Wskaźniki HDI, GERD, RQ oraz GLOBE dla poszczególnych państw zostały pozyskane z publicznie dostępnych źródeł (ONZ, Bank Światowy, GLOBE) i połączone („*data munging*”) w programie Excel (wersja 16.88). Z uwagi na niewielkie populacje wykluczyłem małe kraje wyspiarskie. Z analizy zostały wyłączone także państwa znane jako raje podatkowe lub specjalne enklawy (np. Barbados, Luxemburg, Monako), gdyż większość rejestrowanych tam patentów pochodzi z podmiotów zagranicznych. W sumie dane obejmują 152 kraje, w tym 59 z dostępnymi pomiarami cech przywództwa (badanie GLOBE 2004). Rozkłady zmiennych zostały sprawdzone za pomocą histogramów i wykresów pudełkowych, a podstawowe założenia korelacji i prostej regresji liniowej potwierdzone testem Shapiro-Wilk.

Analizy zostały wykonane w języku R (wersja 4.3.3) i zwizualizowane w RStudio (wersja 2023.06.0) za pomocą pakietu *ggplot2*. Mapa liczby patentów na tysiąc mieszkańców (Wykres 1) powstała z użyciem pakietów *rnaturalearth* i *countrycode*.

### **DOPASOWANIE MODELI**

W badaniu zastosowałem hierarchiczne podejście do budowy modelu, rozpoczynając proces od prostego, addytywnego modelu liniowego, uwzględniającego wszystkie sześć zmiennych predykcyjnych (HDI, GERD, RQ, autokratyczność, autonomiczność, samozachowawczość). W celu normalizacji (test Shapiro-Wilk po transformacji:  $W=0.955$ ), liczba patentów *per capita* została przemnożona przez tysiąc i poddana transformacji logarytmicznej z minimalnym *offsetem* dla wartości zerowych ( $\varepsilon = 0.0001$ ):

$$\log\left(\frac{\text{liczba patentów}}{\text{populacja}} * 1000 + \varepsilon\right)$$

Zaletą zwykłej regresji liniowej jest względna prostota, stosowna przy niewielkiej liczbie obserwacji ( $N=152$ ): badanie złożonych interakcji między predyktorami może być mylące, gdy brak danych wejściowych. Należy jednak uwzględnić, że dane zawierają kilka zmiennych o radykalnie innych rozkładach i skalach, co uzasadnia zastosowanie uogólnionego modelu liniowego (GLM) o bardziej złożonej konstrukcji. Dla równowagi zastosowałem zatem złożony model: regresję beta (w pakiecie *betareg*), odpowiednią do przewidywania wartości ułamkowych o wysokiej dyspersji (średnia zmiennej przewidywanej  $PPT = 0.338$ , wariancja =  $0.719$ ). 11 modeli o różnych wariantach interakcji między predyktorami zostało dopasowanych i porównanych na podstawie Kryterium Informacji Akaikego (AIC) oraz współczynnika dopasowania regresji beta ( $\phi$ ). Aby zrekompensować możliwe odchylenia na skutek znikomej liczby patentów w krajach ubogich, badanie zostało powtórzone na grupie 101 krajów o wysokim wskaźniku rozwoju ( $HDI > 0.7$ ), a także osobno na grupie 46 krajów o wysokiej liczbie patentów ( $PPT > 0.1$ ). Polska należy do obu kategorii.

Oba typy regresji zaowocowały zbliżonymi wynikami. Prosta regresja liniowa wskazuje na GERD jako jedyny istotny predyktor liczby patentów na tysiąc mieszkańców, nawet po uwzględnieniu zamożności i wykształcenia mieszkańców (HDI), oraz stopnia biurokratyzacji (RQ) (Tabela Z1). Najlepszy z jedenastu modeli porównanych regresją beta również określa GERD jako jedyny istotny predyktor: modele uwzględniające typy przywództwa okazują się natomiast wyjątkowo źle dopasowane. Wszystkie przeanalizowane modele sugerują podobne tempo wzrostu innowacyjności: zwiększenie nakładów na R&D o jeden punkt procentowy skutkuje wzrostem na poziomie pomiędzy 0.8 (regresja liniowa, 152 kraje) a 1.16 patentu na tysiąc mieszkańców (optymalna regresja beta, 152 kraje). Modele uwzględniające tylko kraje rozwinięte także sugerują wartości w tym przedziale.

Model	Dopasowanie (-A/C)	Dopasowanie ( <i>phi</i> )	Istotny predyktor	Współczynnik	<i>p</i>
HDI + GERD + RQ	650	22.2	GERD	1.16	<0.0001
GERD	634	18.3	GERD	1.46	<0.0001
HDI	553	2.53	HDI	5.22	<0.0001
RQ	550	5.71	RQ	1.03	<0.0001
wszystkie predyktory, interakcje	196	19.2	HDI:GERD:RQ	13.4	<0.0001
wszystkie predyktory, addytywne	158	4.26	GERD	1.01	<0.0001
zarządzanie samozachowawcze	114	1.34	samozachowawczość	-1.20	0.005
zarządzanie autokratyczne	109	1.18	n/a	n/a	0.118
zarządzanie autonomiczne	108	1.16	n/a	n/a	0.148

Tabela Z1. Rozszerzone wyniki dopasowania modeli regresji beta do danych ze 152 krajów.