

Dane zostały pozyskane z publicznie dostępnych źródeł i oczyszczone w programie Excel („data munging”). Z uwagi na niewielkie populacje usunąłem dane z małych krajów wyspiarskich. Z analizy zostały wyłączone także dane z krajów znanych jako raje podatkowe lub specjalne enklawy (n.p. Barbados, Luxemburg, Monaco), gdyż większość rejestrowanych tam patentów pochodzi z podmiotów zagranicznych. W sumie dane obejmują 152 kraje, w tym 59 z dostępną informacją o typowych dla danej nacji cechach przywództwa (GLOBE).

Analizy zostały wykonane w języku R v4. Rozkład danych i podstawowe założenia korelacji i regresji liniowej zostały sprawdzone za pomocą histogramów, wykresów pudełkowych i testu Shapiro-Wilk. Korel

Wybor modeli. W badaniu zastosowałem hierarchiczne podejście do budowy modelu, rozpoczynając proces od addytywnego modelu liniowego uwzględniającego sześć zmiennych predykcyjnych. Zaletą tego podejścia jest względna prostota, odpowiadająca stosunkowo niewielkiej liczbie obserwacji: 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 GLM. Dla równowagi zastosowałem zatem złożony model liniowy Beta, odpowiedni dla przewidywania wartości ułamkowych o wysokiej dyspersji (średnia, wariancja). W celu normalizacji, liczba patentów per capita została pomnożona przez tysiąc i poddana transformacji logarytmicznej z offsetem dla wartości zerowych [ $\log(\text{patents}/\text{population} * 1000 + 0.0001)$ ] (Shapiro-Wilk  $W=0.955$ ). 11 modeli o różnych wariantach interakcji między predyktorami zostało dopasowanych i porównanych na podstawie AIC. Aby uwzględnić możliwe odchylenia na skutek bardzo niskiej liczby patentów w krajach ubogich, badanie zostało powtórzone na grupie XYZ krajów o wysokim wskaźniku rozwoju ( $HDI > 0.7$ ), a także na grupie GHJ krajów o wysokiej liczbie patentów ( $PPT > 0.1$ ).

Oba podejścia zawołały zbliżonym wynikiem. 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). Optymalny model Regresja Beta również określa GERD jako jedyny istotny predyktor: modele uwzględniające typy przywództwa okazują się natomiast wyjątkowo źle dopasowane. TABELA

Wszystkie przeanalizowane modele sugerują podobne tempo wzrostu innowacyjności: zwiększenie nakładów na R&D o jeden punkt procentowy owocuje wzrostem liczby patentów od 0.8 (regresja liniowa, 152 kraje) do 1.16 patentu na tysiąc mieszkańców (regresja Beta, 152 kraje). Modele uwzględniające tylko kraje rozwinięte sugerują wartości w tym przedziale.

Polecenia w języku R są dostępne pod adresem: [github](#)

Wykresy wykonano pakietem ggplot2.