

Przetwarzanie Dużych Zbiorów Danych

Założenia projektu

Mariusz Bańcer, 208746 Maciej Kolański, 209806

Marzec 2018

1 Cel projektu

Celem projektu jest umożliwienie porównania wydatków na edukację poszczególnych państw z uwzględnieniem ich wielkości, przyrostu naturalnego oraz stanu i tempa rozwoju gospodarczego reprezentowanego przez wartość i zmianę Produktu Krajowego Brutto.

2 Struktura danych wejściowych

1. Wielkość populacji państw w poszczególnych latach

<http://datahub.io/core/population#resource-population>

Nazwa	Typ
Państwo/Region	tekst
Skrót nazwy	tekst
Rok	rok
Wielkość populacji	liczba

2. Produkt Krajowy Brutto państw w poszczególnych latach

<http://datahub.io/core/gdp>

Nazwa	Typ
Państwo/Region	tekst
Skrót nazwy	tekst
Rok	rok
Produkt Krajowy Brutto	liczba

3. Wydatki na edukację państw w poszczególnych latach

<https://data.worldbank.org/indicator/SE.XPD.TOTL.GD.ZS>

Nazwa	Typ
Państwo	tekst
Skrót nazwy	tekst
Rok	rok
Całkowite wydatki rządowe na edukację jako procent PKB	liczba

3 Opis danych wyjściowych

Rysunek 1: Nazwy parametrów wyjściowych wraz z przykładowymi danymi

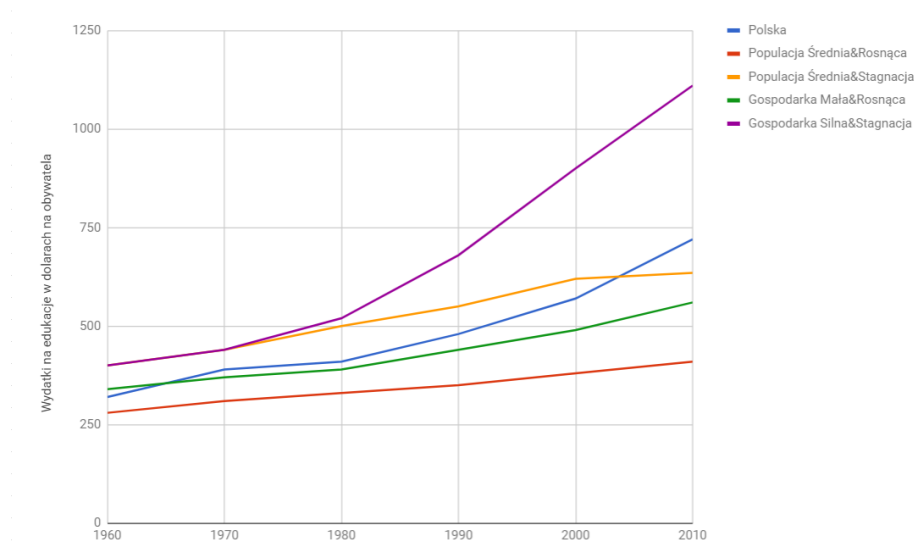
Rok	Państwo	Region	Segment wielkości	Segment gospodarczy	Edukacja per capita	Wydatki na edukację względem średniej			
						Region	Segment wielkości	Segment gospodarczy	Ogółem
2015	Polska	Europa wschodnia	Średni&Stagnacja	Średni&Rosnący	654	+10%	-15%	-2%	4%
1970	Polska	Europa wschodnia	Średni&Rosnący	Mały&Rosnący	350	+4%	-11%	8%	-5%

- **Rok** - rok w oparciu o który wyznaczone zostaną pozostałe wartości
- **Państwo** - nazwa analizowanego kraju
- **Region** - przynależność geograficzna państwa
- **Segment wielkości** - klasa wyznaczona na podstawie wielkości populacji oraz średniej zmiany z okresu poprzednich 5 lat
- **Segment gospodarczy** - klasa wyznaczona na podstawie wielkości PKB oraz średniej zmiany z okresu poprzednich 5 lat
- **Edukacja per capita** - wartość wydatków na edukację na obywatela wyrażona w dolarach (z wykorzystaniem kursu przyjętego w zbiorze danych o PKB) wyznaczona na podstawie PKB, populacji oraz wydatków na edukację
- **Wydatki na edukację względem średniej** - rozbieżność względem średniej wartości wydatków na edukację wyrażona w procentach

Dodatkowo prezentowane będą miejsca w rankingach wydatków w poszczególnych segmentach, analogicznie do różnic procentowych.

4 Graficzna reprezentacja danych wyjściowych

Rysunek 2: Zestawienie wydatków na edukację Polski i powiązanych segmentów



Zestawienie wydatków państw i segmentów pozwala na obserwację trendu i porównanie państw o podobnej strukturze gospodarczej czy społecznej. W klarowny sposób prezentuje pozycję państwa w segmencie oraz umożliwia predykcję sytuacji na podstawie widocznych zmian w wysokości wydatków.