

Współczynnik dzietności w wybranych regionach świata w latach 2005 - 2010

Alicja Wiączkoska

2023-02-05

Contents

Opis bazy	1
Podstawowe statystyki	2
Zależności między danymi	2

Opis bazy

Analizie poddano fragment bazy danych sporządzonej przez **Organizację Narodów Zjednoczonych**. Jest to zestawienie wskaźników dotyczących liczby żywych urodzeń w zależności od wieku matek w poszczególnych latach oraz regionach świata.

Przygotowanie ramki danych Baza została pobrana ze strony <https://population.un.org/wpp/Download/Standard/CSV/>. Następnie zapisano ją jako ramkę danych i usunięto zbędne fragmenty, a także zmieniono typu kolumn na **factor** tam, gdzie było to konieczne.

Nazwy kolumn badanej bazy

```
## 'data.frame': 162 obs. of 6 variables:
## $ Location: Factor w/ 3 levels "Europe and Northern America",...: 3 3 3 3 3 3 3 3 3 ...
## $ Time : Factor w/ 6 levels "2005","2006",...: 1 1 1 1 1 1 1 1 2 ...
## $ AgeGrp : Factor w/ 9 levels "10-14","15-19",...: 1 2 3 4 5 6 7 8 9 1 ...
## $ ASFR : num 8.36 124.31 230.81 244.07 218.37 ...
## $ PASFR : num 0.754 11.211 20.816 22.012 19.693 ...
## $ Births : num 381 4974 7921 6896 5048 ...
```

Analizowana ramka danych składa się z 6 kolumn i 162 wierszy.

Poszczególne kolumny to:

- **Location** - nazwa analizowanego regionu,
- **Time** - lata z przedziału 2005 - 2010,
- **AgeGrp** - przedziały wiekowe kobiet: 10-14, 15-19, 20-24, 25-29, 30-34, 35-39, 40-44, 45-49, 50-54,
- **ASFR** (*Age-Specific Fertility Rate*) - współczynnik dzietności (według wieku) obliczany ze wzoru:

$$\frac{\text{liczba dzieci urodzonych przez matki w podanym przedziale wiekowym}}{\text{liczba kobiet w podanym przedziale wiekowym}} \cdot 1000,$$

- PASFR (*Percent Age-specific Fertility Rate*) - procentowy współczynnik dzietności (według wieku) obliczany dla poszczególnych lat ze wzoru:

$$\frac{\text{liczba dzieci urodzonych przez matki w podanym przedziale wiekowym}}{\text{liczba wszystkich dzieci urodzonych w danym roku}} \cdot 100\%$$

(PASFR dla poszczególnego regionu w danym roku sumuje się do 100),

- Births - liczba żywych dzieci (w tysiącach) urodzonych przez matki w danym przedziale wiekowym.

Jaké można odczyta ze struktury ramki, kolumny Location, Time oraz AgeGrp są typu factor, a pozostałe typu numeric.

Podstawowe statystyki

```
##                               Location    Time      AgeGrp      ASFR
## Europe and Northern America :54  2005:27  10-14 :18  Min.   : 0.012
## Latin America and the Caribbean:54 2006:27  15-19 :18  1st Qu.: 3.814
## Sub-Saharan Africa           :54  2007:27  20-24 :18  Median : 45.202
##                               2008:27  25-29 :18  Mean   : 69.649
##                               2009:27  30-34 :18  3rd Qu.:107.414
##                               2010:27  35-39 :18  Max.   :244.073
##                               (Other):54
##
##      PASFR      Births
## Min.   : 0.0030  Min.   : 0.468
## 1st Qu.: 0.7285  1st Qu.: 104.810
## Median :10.0250  Median :1461.544
## Mean   :11.1111  Mean   :2047.122
## 3rd Qu.:20.8212  3rd Qu.:3164.995
## Max.   :29.4140  Max.   :8798.518
##
```

Jak można odczytać, W kolumnie Location znajdują się 3 regiony: Europe and Northern America, Latin America and the Caribbean, Sub-Saharan Africa. Do każdego z nich są przyporządkowane dokładnie 54 wiersze.

Wskaźnik ASFR przyjmuje najmniejszą wartość 0.012 dla przedziału wiekowego 50-54 w regionie Europe and Northern America w roku 2008. Natomiast największa liczba dzieci przypadających na 1000 kobiet została osiągnięta w 2008 roku, a wynosi 244.073 dla grupy wiekowej 25-29 i obszaru Europe and Northern America. Średnia wartość występująca w kolumnie ASFR to 69.6485247, a mediana 45.2015.

Natomiast najmniejszą wartością wskaźnika PASFR jest 0.003% przyjmowane dla przedziału wiekowego 50-54 w 2010 roku w regionie Europe and Northern America. Największy procentowy udział liczby urodzeń wynosi 29.414% osiągnięty w grupie wiekowej 25-29 obliczany na terenie Europe and Northern America w 2010 roku. Średnia wartość występująca w kolumnie PASFR to 11.1111296, a mediana 10.025.

Najmniej dzieci w liczbie 0.468 urodziło się wśród kobiet w wieku 50-54 w regionie Europe and Northern America roku 2010. Z kolei maksymalną liczbę urodzeń odnotowano w 2010 roku wynoszącą 0.468 tys. wśród matek w wieku 20-24 zamieszkujących obszar Sub-Saharan Africa. Średnia wartość występująca w kolumnie Births to 2047.1219074, a mediana 1461.544.

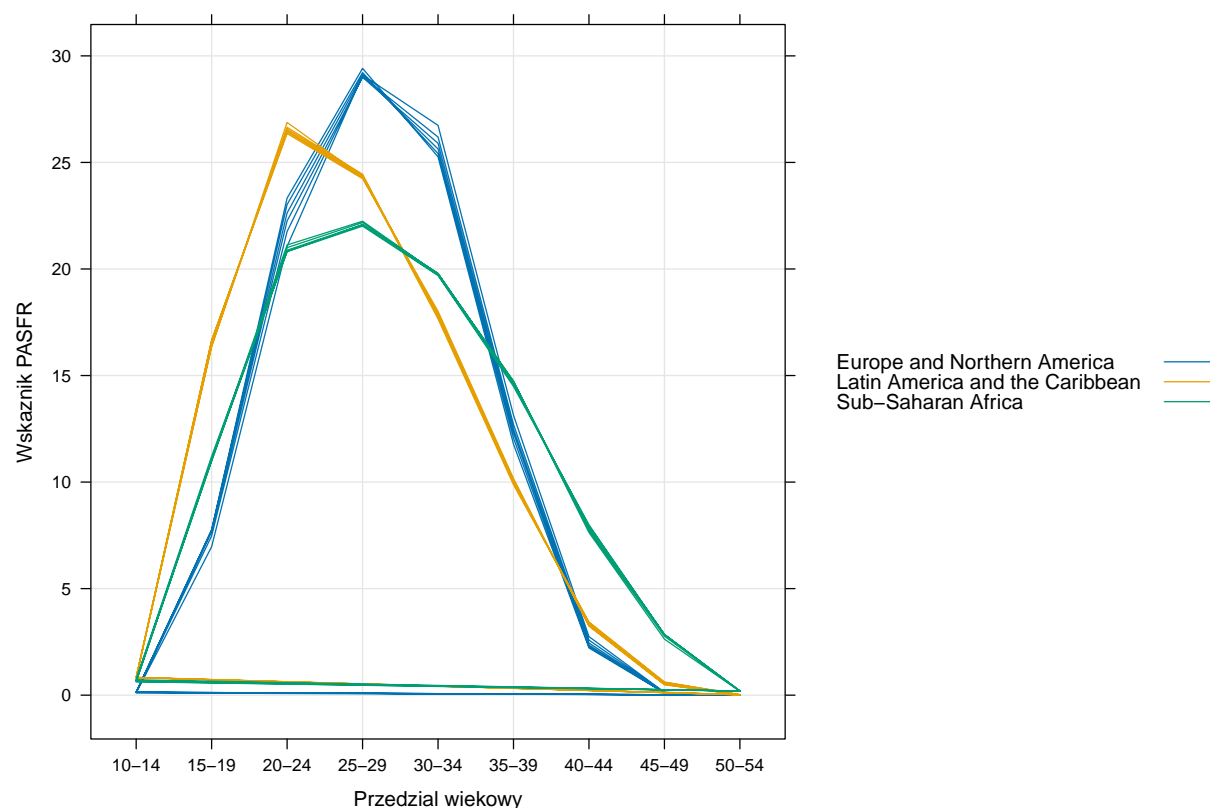
Zależności między danymi

Różnice między regionami pod względem wysokości wskaźnika PASFR w celu zbadania, w których grupach wiekowych rodzi się najwięcej dzieci, utworzono nową kolumnę o nazwie PASFR_lvl1s informującą o wielkości procentowego udziału dzieci urodzonych przez matki w danej grupie wiekowej w

określonym roku i miejscu. Aby grupy były porównywalnej wielkości obliczono medianę oraz kwartyle dla wszystkich danych wskaźnika PASFR, które wyznaczyły cztery kategorie: low, medium-low, medium-high, high. Wyniki zapisano w tabeli poniżej.

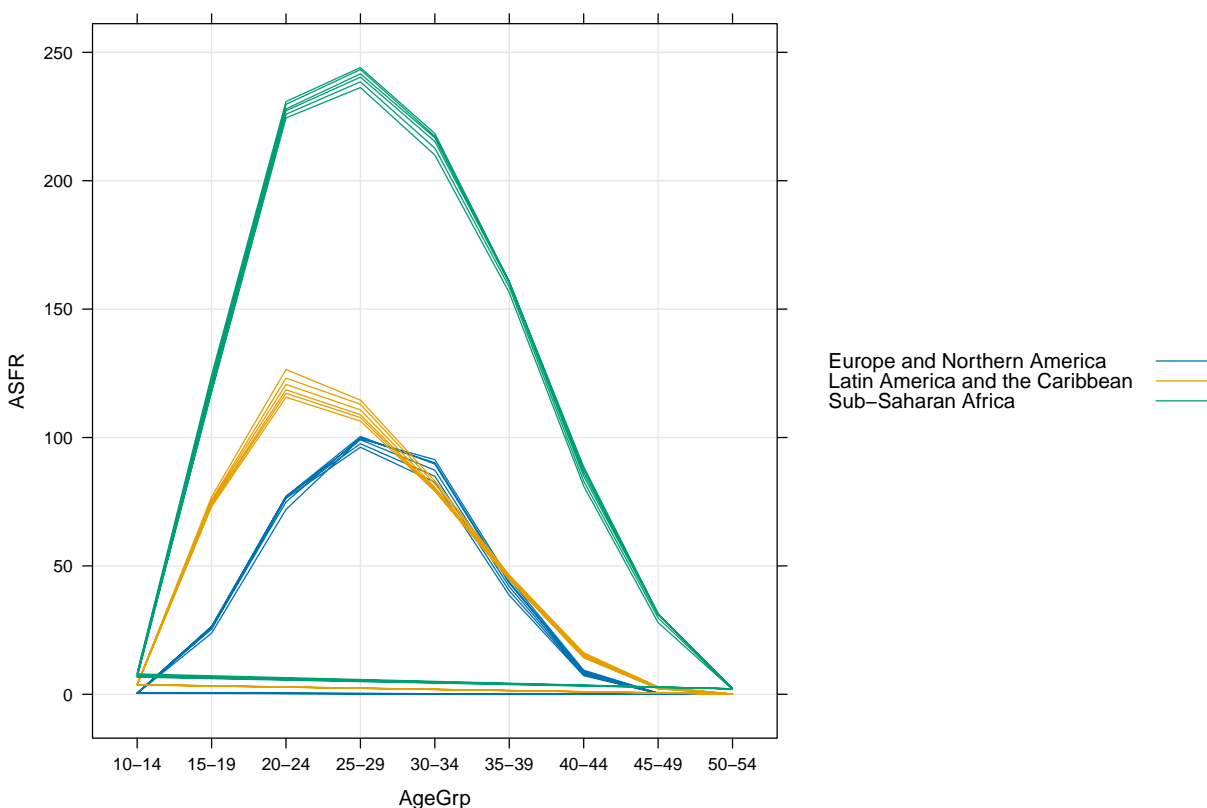
Można wyciągnąć wniosek, że w badanych regionach najrzadziej matkami zostają kobiety na końcu okresu rozrodczego (45-54 lata). Najwięcej dzieci z kolei rodzi się wśród kobiet między 20. a 29. rokiem życia.

Wskaźnik PASFR a grupa wiekowa matek w zależności od regionu Na podstawie zgromadzonych danych stworzono wykres zależności procentowej liczby urodzeń od grupy wiekowej matek z wyróżnieniem obszarów ich zamieszkania. Każda z linii odpowiada informacjom pochodzącym z jednego roku.



Można z niego odczytać, że wskaźnik PASFR jest dla Europy i Północnej Ameryki jest najwyższy w przedziale wiekowym 25-29 lat i niewiele niższy w grupie 30-34 lat. Z kolei W Ameryce Łacińskiej i na Karaibach dominujący wiek matek to 20-24 r.ż i tuż za nim 25-29. W Afryce Subsaharyjskiej jest podobna sytuacja, lecz w odwrotnej kolejności. Można też wywnioskować, że procentowo więcej kobiet rodzi dziecko po 40 r.ż. W Afryce niż w pozostałych regionach.

Wskaźnik ASFR a grupa wiekowa matek w zależności od regionu

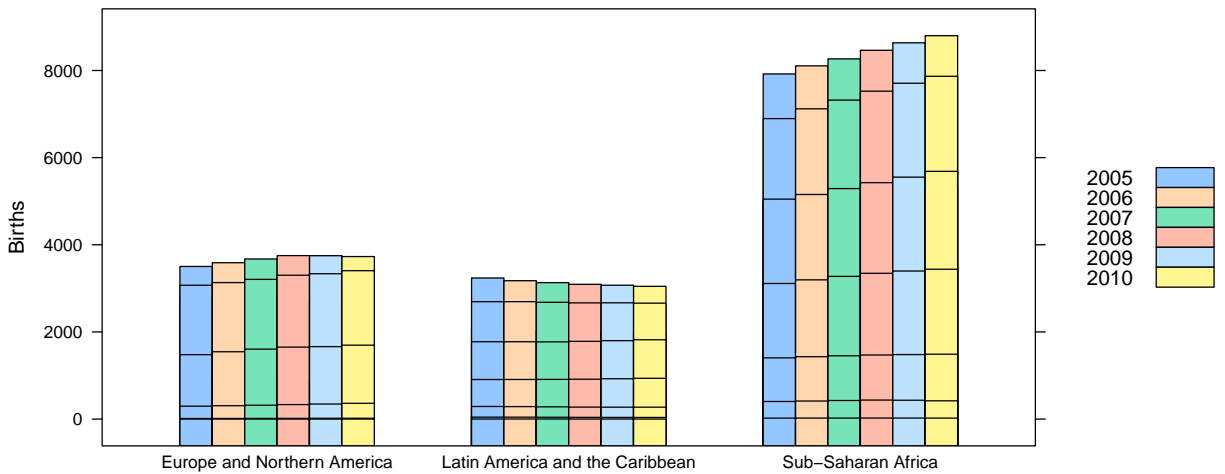


W tym przypadku zamiast procentowego udziału urodzeń, wzięta została ich liczba. Z tego też powodu bardzo widoczna jest dominacja Afryki w liczbie urodzeń w każdej grupie wiekowej. Kształty łamanych (łącznie z maksymalnymi wartościami dla poszczególnych grup wiekowych) pozostają niezmienione.

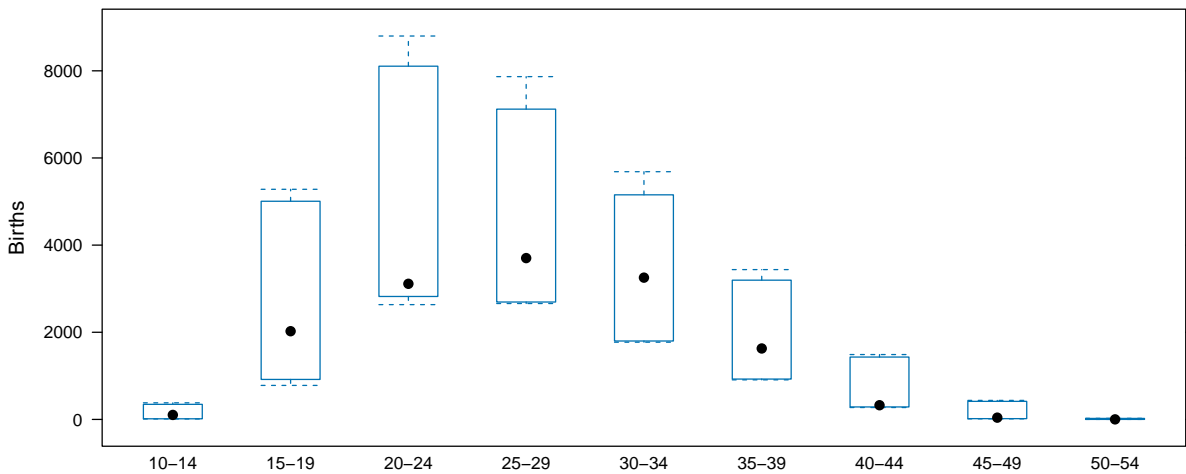
Liczba urodzeń w poszczególnych regionach

Region	Suma urodzeń w latach 2005-2010 (w tys.)
Europe and Northern America	75160.08
Latin America and the Caribbean	65743.01
Sub-Saharan Africa	190730.66

Z powyższej ramki danych stworzonej na podstawie analizowanej bazy łatwo można zauważyć, że ponad dwukrotnie więcej dzieci urodziło się w Sub-Saharan Africa niż w pozostałych regionach. Najmniej urodzeń odnotowano w sumie w regionie Latin America and the Caribbean, lecz liczba ta jest porównywalna z wynikiem dla Europe and Northern America. Poniższy wykres słupkowy ukazujący zmiany liczby urodzeń na przestrzeni lat w badanych regionach również obrazuje te tendencje.



Ogólna Liczba urodzeń w poszczególnych grupach wiekowych



Z Powyższego wykresu odczytać można nie tylko uśrednioną liczbę urodzeń w każdej z grup wiekowych, ale też odchylenia od tej wartości. Możemy zauważyć że liczba urodzeń na początku i końcu okresu rozrodczego nie zmieniała się zbyt w zależności od czasu lub regionu. Możemy za to dostrzec duże różnice w dzietności kobiet w wieku 20-29 lat.

Histogramy Na koniec zbadano jakie wartości pojawiały się najczęściej w badanej bazie danych dla wskaźnika ASFR oraz liczby urodzeń. Okazało się, że w obu przypadkach najwięcej było stosunkowo niskich wartości, a najmniej tych na końcu skali.

