UNIWERSYTET EKONOMICZNY

**Paulina Rejda**

Analityka gospodarcza

Analityk decyzji biznesowych

**Analiza płodności w szwajcarskich prowincjach w 1888 r.**

Narzędzia Komputerowe w Analizie Danych

Spis treści

[**Problem badawczy i wprowadzenie**](#_o4hjw2oomhij) **3**

[**Podsumowanie zmiennych za pomocą miar statystycznych**](#_3znysh7) **3**

[**Testowanie istotności współczynnika korelacji**](#_1t3h5sf) **7**

[**Wizualizacja rozkładów zmiennych**](#_17dp8vu) **9**

[**Tworzenie modelu**](#_3rdcrjn) **10**

[5.1 Test Fishera](#_9r8s9zv34b5o) 12

[5.2 Test T- Studenta](#_2jxsxqh) 12

[5.3 Próba usunięcia najmniej istotnej zmiennej](#_z337ya) 13

[**Macierz kowariancji**](#_3j2qqm3) **13**

[**Test normalności Shapiro-Wilka**](#_4i7ojhp) **14**

[**Wartości teoretyczne**](#_2xcytpi) **14**

[**Reszty modelu**](#_1ci93xb) **15**

[**Analiza skupień**](#_3as4poj) **16**

[**Podsumowanie. Wnioski**](#_23ckvvd) **20**

# **Problem badawczy i wprowadzenie**

Pod koniec XIX wieku miało miejsce w Szwajcarii przejście demograficzne, czyli okres przejścia od wysokiego do niskiego poziomu urodzeń, zgonów i przyrostu naturalnego stabilizującego rozmiar populacji, typowego dla krajów rozwiniętych. Postanowiono zbadać:

**Jakie czynniki mają wpływ na spadek płodności wśród mieszkańców szwajcarskich prowincji w 1888 roku?**

Analiza została przeprowadzona w celu sprawdzenia kierunku oraz siły zależności między indeksem płodności a zmiennymi dotyczącymi zjawisk społeczno-ekonomicznych. Zaczęto od zaprezentowania podstawowych statystyk dotyczących danych, stworzono również model ekonometryczny, przeprowadzona została również analiza wielowymiarowa z wyróżnieniem powstałych grup. Badanie to pozwala zrozumieć kształtowanie się rozwoju demograficznego ówczesnej Szwajcarii.

Do stworzenia modelu posłużą nam następujące dane składające się z ***47 obserwacji*** dla ***6 zmiennych***:

* Płodność (fertility) – wyrażone za pomocą indeksu (na podstawie zestandaryzowanych wartości)
* Rolnictwo (agriculture) - odsetek mężczyzn pracujących w rolnictwie
* Badania (examination) - odsetek poborowych otrzymujących najlepsze wyniki w egzaminach do wojska
* Wykształcenie (education) - odsetek poborowych o wykształceniu wyższym niż szkoła podstawowa
* Katolik (catholic) - odsetek osób wyznających wiarę katolicką
* Śmiertelność niemowląt (infant mortality) - odsetek niemowląt o długości życia krótszej niż 1 rok

Wszystkie zmienne, za wyjątkiem zmiennej Płodność, wyrażone są jako odsetek badanej populacji.

# **Podsumowanie zmiennych za pomocą miar statystycznych**

Fertility Agriculture Examination Education Catholic Infant.Mortality

Min. :35.00 Min. : 1.20 Min. : 3.00 Min. : 1.00 Min. : 2.150 Min. :10.80

1st Qu.:64.70 1st Qu.:35.90 1st Qu.:12.00 1st Qu.: 6.00 1st Qu.: 5.195 1st Qu.:18.15

Median :70.40 Median :54.10 Median :16.00 Median : 8.00 Median : 15.140 Median :20.00

Mean :70.14 Mean :50.66 Mean :16.49 Mean :10.98 Mean : 41.144 Mean :19.94

3rd Qu.:78.45 3rd Qu.:67.65 3rd Qu.:22.00 3rd Qu.:12.00 3rd Qu.: 93.125 3rd Qu.:21.70

Max. :92.50 Max. :89.70 Max. :37.00 Max. :53.00 Max. :100.000 Max. :26.60

Tabela 1.

W powyższym zestawieniu (Tabela 1.) zaprezentowane zostały podstawowe statystyki dotyczące zmiennych: wartość minimalna i maksymalna, pierwszy i trzeci kwartyl, mediana oraz średnia. Warto zwrócić uwagę w szczególności na statystyki zmiennej objaśnianej - Płodności, które przedstawione zostały w Tabeli 2.

Min. 1st Qu. Median Mean 3rd Qu. Max.

35.00 64.70 70.40 70.14 78.45 92.50

Tabela 2.

Zmienna przyjmuje wartości od 35 do 92.5. Średnia wynosi 70.14, pomimo dość niskiej wartości minimalnej. Rozstęp wynosi 57.50. Więcej statystyk dla zmiennej Płodność znajdziemy poniżej:

Wariancja:

156.0425

Odchylenie standardowe:

156.0425

Kwartyle:

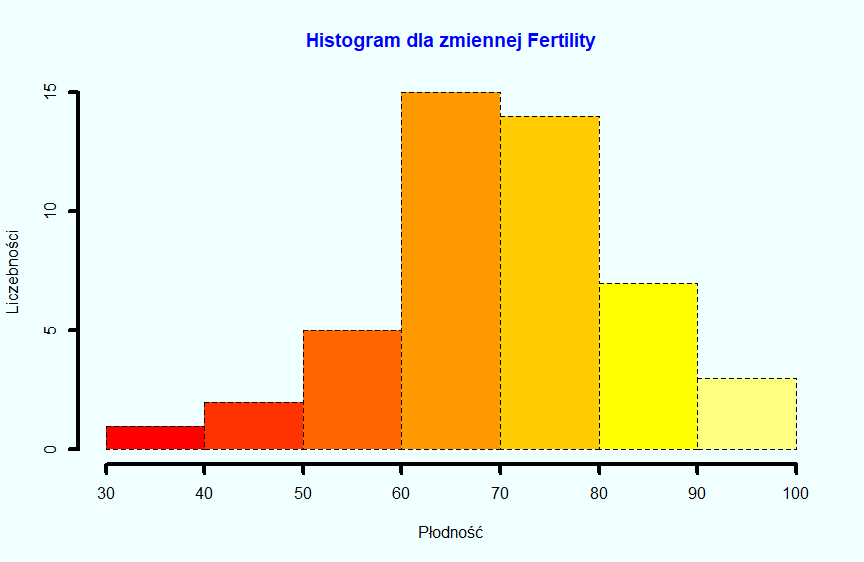
0% 25% 50% 75% 100%

35.00 64.70 70.40 78.45 92.50

Rozstęp międzykwartylowy:

13.75

## Poniższy histogram (Obrazek 1.) prezentuje liczebność obserwacji należących do 7 przedziałów stworzonych dla zmiennej Płodność.



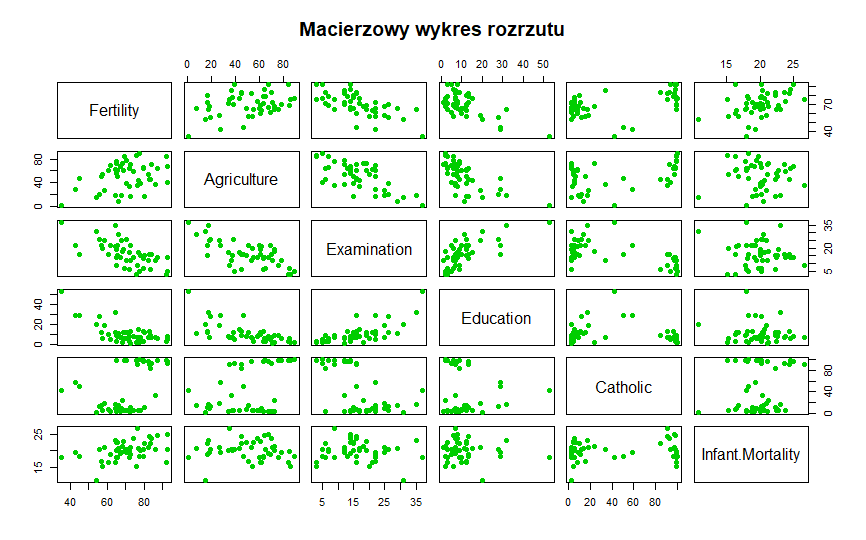
Obrazek 1.

Na podstawie wykresu możemy wywnioskować, że dla większości prowincji indeks płodności mieścił się w przedziale 60-80.

## 

## Macierzowy wykres rozrzutu zaprezentowany został na Obrazku 2. Prezentuje on zależności między wszystkimi zmiennymi. Do dokładniejszej analizy tych zależności wykorzystano również macierz korelacji (Tabela 3.). Pozwala ona ocenić, które zmienne są od siebie istotnie zależne.

## 



Obrazek2.

## Macierz korelacji:

Fertility Agriculture Examination Education Catholic Infant.Mortality

Fertility 1.0000000 0.35307918 -0.6458827 -0.66378886 0.4636847 0.41655603

Agriculture 0.3530792 1.00000000 -0.6865422 -0.63952252 0.4010951 -0.06085861

Examination -0.6458827 -0.68654221 1.0000000 0.69841530 -0.5727418 -0.11402160

Education -0.6637889 -0.63952252 0.6984153 1.00000000 -0.1538589 -0.09932185

Catholic 0.4636847 0.40109505 -0.5727418 -0.15385892 1.0000000 0.17549591

Infant.Mortality 0.4165560 -0.06085861 -0.1140216 -0.09932185 0.1754959 1.00000000

Tabela 3.

Wstępnie przyjęto, że istotne są korelacje między następującymi zmiennymi:

Płodność – Badania   
 Płodność - Wykształcenie  
 Rolnictwo – Badania   
 Rolnictwo – Wykształcenie  
 Badania – Wykształcenie  
 Katolik – Badania  
 Katolik – Płodność  
 Śmiertelność niemowląt – jedynie skorelowana ze zmienną Płodność; dodatkowo nie jest ona silna

# **Testowanie istotności współczynnika korelacji**

Sprawdzono istotność współczynnika korelacji dla wszystkich par zmiennych wymienionych powyżej oraz korelacji wszystkich zmiennych ze zmienną Płodność. Otrzymano następujące wyniki:

Płodność – Rolnictwo

data: swiss$Fertility and swiss$Agriculture

t = 2.5316, df = 45, p-value = 0.01492

alternative hypothesis: true correlation is not equal to 0

95 percent confidence interval:

0.07334947 0.58130587

sample estimates:

cor

0.3530792

Płodność – Badania

data: swiss$Fertility and swiss$Examination

t = -5.6753, df = 45, p-value = 9.45e-07

alternative hypothesis: true correlation is not equal to 0

95 percent confidence interval:

-0.7870674 -0.4403995

sample estimates:

cor

-0.6458827

Płodność - Wykształcenie

data: swiss$Fertility and swiss$Education

t = -5.9536, df = 45, p-value = 3.659e-07

alternative hypothesis: true correlation is not equal to 0

95 percent confidence interval:

-0.7987075 -0.4653206

sample estimates:

cor

-0.6637889

Płodność – Katolik

data: swiss$Fertility and swiss$Catholic

t = 3.5107, df = 45, p-value = 0.001029

alternative hypothesis: true correlation is not equal to 0

95 percent confidence interval:

0.2036326 0.6626204

sample estimates:

cor

0.4636847

Płodność – Śmiertelność niemowląt

data: swiss$Fertility and swiss$Infant.Mortality

t = 3.0737, df = 45, p-value = 0.003585

alternative hypothesis: true correlation is not equal to 0

95 percent confidence interval:

0.1469699 0.6285366

sample estimates:

cor

0.416556

Rolnictwo – Badania

data: swiss$Agriculture and swiss$Examination

t = -6.3341, df = 45, p-value = 9.952e-08

alternative hypothesis: true correlation is not equal to 0

95 percent confidence interval:

-0.8133545 -0.4974484

sample estimates:

cor

-0.6865422

Rolnictwo – Wykształcenie

data: swiss$Agriculture and swiss$Education

t = -5.5804, df = 45, p-value = 1.305e-06

alternative hypothesis: true correlation is not equal to 0

95 percent confidence interval:

-0.7829085 -0.4316231

sample estimates:

cor

-0.6395225

Badania – Wykształcenie

data: swiss$Examination and swiss$Education

t = 6.5463, df = 45, p-value = 4.811e-08

alternative hypothesis: true correlation is not equal to 0

95 percent confidence interval:

0.5144218 0.8209342

sample estimates:

cor

0.6984153

Katolik –Badania

data: swiss$Catholic and swiss$Examination

t = -4.687, df = 45, p-value = 2.588e-05

alternative hypothesis: true correlation is not equal to 0

95 percent confidence interval:

-0.7384534 -0.3417900

sample estimates:

cor

-0.5727418

Na uznanie korelacji za istotną pozwoliła bardzo niska wartość p-value, na podstawie której odrzucono hipotezę zerową wskazującą na brak korelacji. Poniżej (Tabela 4.) zaprezentowano poprawioną macierz korelacji, gdzie oznaczono kolorem czerwonym statystycznie istotne korelacje.

Fertility Agriculture Examination Education Catholic Infant.Mortality

Fertility 1.0000000 0.35307918 -0.6458827 -0.66378886 0.4636847 0.41655603

Agriculture 0.3530792 1.00000000 -0.6865422 -0.63952252 0.4010951 -0.06085861

Examination -0.6458827 -0.68654221 1.0000000 0.69841530 -0.5727418 -0.11402160

Education -0.6637889 -0.63952252 0.6984153 1.00000000 -0.1538589 -0.09932185

Catholic 0.4636847 0.40109505 -0.5727418 -0.15385892 1.0000000 0.17549591

Infant.Mortality 0.4165560 -0.06085861 -0.1140216 -0.09932185 0.1754959 1.00000000

Tabela 4.

Z przedstawionych danych wynika, że dla zmiennej objaśnianej istnieje istotna korelacja ze zmienną Wykształcenie (-0,66) oraz Badania (-0,65). Wizualizacje tych zależności możemy zobaczyć na poniższych wykresach rozrzutu (Wykres 1. I Wykres 2.).

# **Tworzenie modelu**

## Poniżej zaprezentowano podsumowanie dla modelu ze wszystkimi zmiennymi (Tabela 5.).

Call:

lm(formula = Fertility ~ ., data = swiss)

Residuals:

Min 1Q Median 3Q Max

-15.2743 -5.2617 0.5032 4.1198 15.3213

Coefficients:

Estimate Std. Error t value Pr(>|t|)

(Intercept) 66.91518 10.70604 6.250 1.91e-07 \*\*\*

Agriculture -0.17211 0.07030 -2.448 0.01873 \*

Examination -0.25801 0.25388 -1.016 0.31546

Education -0.87094 0.18303 -4.758 2.43e-05 \*\*\*

Catholic 0.10412 0.03526 2.953 0.00519 \*\*

Infant.Mortality 1.07705 0.38172 2.822 0.00734 \*\*

---

Signif. codes: 0 ‘\*\*\*’ 0.001 ‘\*\*’ 0.01 ‘\*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 7.165 on 41 degrees of freedom

Multiple R-squared: 0.7067, Adjusted R-squared: 0.671

F-statistic: 19.76 on 5 and 41 DF, p-value: 5.594e-10

Tabela 5.

Za pomocą selekcji postępującej dobrano zmienne do modelu. (Tabela 6.)

Start: AIC=238.35

Fertility ~ 1

Df Sum of Sq RSS AIC

+ Education 1 3162.7 4015.2 213.04

+ Examination 1 2994.4 4183.6 214.97

+ Catholic 1 1543.3 5634.7 228.97

+ Infant.Mortality 1 1245.5 5932.4 231.39

+ Agriculture 1 894.8 6283.1 234.09

<none> 7178.0 238.34

Step: AIC=213.04

Fertility ~ Education

Df Sum of Sq RSS AIC

+ Catholic 1 961.07 3054.2 202.18

+ Infant.Mortality 1 891.25 3124.0 203.25

+ Examination 1 465.63 3549.6 209.25

<none> 4015.2 213.04

+ Agriculture 1 61.97 3953.3 214.31

Step: AIC=202.18

Fertility ~ Education + Catholic

Df Sum of Sq RSS AIC

+ Infant.Mortality 1 631.92 2422.2 193.29

+ Agriculture 1 486.28 2567.9 196.03

<none> 3054.2 202.18

+ Examination 1 2.46 3051.7 204.15

Step: AIC=193.29

Fertility ~ Education + Catholic + Infant.Mortality

Df Sum of Sq RSS AIC

+ Agriculture 1 264.176 2158.1 189.86

<none> 2422.2 193.29

+ Examination 1 9.486 2412.8 195.10

Step: AIC=189.86

Fertility ~ Education + Catholic + Infant.Mortality + Agriculture

Df Sum of Sq RSS AIC

<none> 2158.1 189.86

+ Examination 1 53.027 2105.0 190.69

Tabela 6.

Przeprowadzona selekcja pozwoliła na dobór następujących zmiennych objaśnianych do modelu:

* Wykształcenie
* Katolik
* Śmiertelność niemowląt
* Rolnictwo

## 

## W Tabeli 7. zaprezentowano podsumowanie modelu z powyższymi zmiennymi.

Call:

lm(formula = Fertility ~ Agriculture + Infant.Mortality + Catholic +

Education, data = swiss)

Residuals:

Min 1Q Median 3Q Max

-14.6765 -6.0522 0.7514 3.1664 16.1422

Coefficients:

Estimate Std. Error t value Pr(>|t|)

(Intercept) 62.10131 9.60489 6.466 8.49e-08 \*\*\*

Agriculture -0.15462 0.06819 -2.267 0.02857 \*

Infant.Mortality 1.07844 0.38187 2.824 0.00722 \*\*

Catholic 0.12467 0.02889 4.315 9.50e-05 \*\*\*

Education -0.98026 0.14814 -6.617 5.14e-08 \*\*\*

---

Signif. codes: 0 ‘\*\*\*’ 0.001 ‘\*\*’ 0.01 ‘\*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 7.168 on 42 degrees of freedom

Multiple R-squared: 0.6993, Adjusted R-squared: 0.6707

F-statistic: 24.42 on 4 and 42 DF, p-value: 1.717e-10

Tabela 7.

Model będzie mieć zatem następującą postać:

**Y = 62,101 – 0,155 Rolnictwo + 1,078 Śmiertelność niemowląt + 0,125 Katolik – 0,98 Wykształcenie  
 (9,605) (0,068) (0,382) (0,029) (0,148)**

Możemy zinterpretować go w następujący sposób – jeżeli indeks płodności wzrośnie o 1 jednostkę, odsetek mężczyzn pracujących w rolnictwie spadnie o 0,15%, śmiertelność niemowląt wzrośnie o 1,8%, odsetek osób o wierze katolickiej wzrośnie o 0,13%, a odsetek osób o wykształceniu wyższym spadnie o 0,98%. Funkcja zwróciła również współczynnik determinacji o wartości 0,7 i współczynnik zbieżności wynoszący 0,3. Zmienna nie została objaśniona przez model w 30%.

## 

## **5.1 Test Fishera**

Wartość statystyki F wynosi 24,42. Bardzo małe p-value (1.717e-10) pozwala na odrzucenie hipotezy zerowej. Oznacza to, że przynajmniej jedna zmienna wpływa istotnie na zmienną objaśnianą.

## **5.2 Test T- Studenta**

Odczytawszy wartości t oraz p-value dla każdej zmiennej, można stwierdzić, że każda zmienna wpływa istotnie na zmienną objaśnianą w większym lub mniejszym stopniu.

## **5.3 Próba usunięcia najmniej istotnej zmiennej**

## Postanowiono sprawdzić, jak usunięcie najmniej istotnej zmiennej (Rolnictwo) wpłynie na model. Wyniki zaprezentowano w Tabeli 8.

Call:

lm(formula = Fertility ~ Infant.Mortality + Catholic + Education,

data = swiss)

Residuals:

Min 1Q Median 3Q Max

-14.4781 -5.4403 -0.5143 4.1568 15.1187

Coefficients:

Estimate Std. Error t value Pr(>|t|)

(Intercept) 48.67707 7.91908 6.147 2.24e-07 \*\*\*

Infant.Mortality 1.29615 0.38699 3.349 0.00169 \*\*

Catholic 0.09607 0.02722 3.530 0.00101 \*\*

Education -0.75925 0.11680 -6.501 6.83e-08 \*\*\*

---

Signif. codes: 0 ‘\*\*\*’ 0.001 ‘\*\*’ 0.01 ‘\*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 7.505 on 43 degrees of freedom

Multiple R-squared: 0.6625, Adjusted R-squared: 0.639

F-statistic: 28.14 on 3 and 43 DF, p-value: 3.15e-10

Tabela 8.

Okazało się, iż model po usunięciu tej zmiennej gorzej (bo w 66%) wyjaśnia badane zjawisko.

# **Macierz kowariancji**

Macierz kowariancji dla stworzonego modelu zaprezentowano w Tabeli 9.

(Intercept) Agriculture Infant.Mortality Catholic Education

(Intercept) 92.25383713 -0.4037117745 -3.261892671 0.0736621463 -0.7914397576

Agriculture -0.40371177 0.0046498653 0.006547147 -0.0008601074 0.0066467529

Infant.Mortality -3.26189267 0.0065471470 0.145821802 -0.0027766377 0.0124239813

Catholic 0.07366215 -0.0008601074 -0.002776638 0.0008348346 -0.0008256439

Education -0.79143976 0.0066467529 0.012423981 -0.0008256439 0.0219444757

Tabela 9.

# **Test normalności Shapiro-Wilka**

W celu sprawdzenia normalności rozkładu przeprowadzono test Shapiro-Wilka. Wyniki zaprezentowano w Tabeli 10.

Shapiro-Wilk normality test

data: model1$residuals

W = 0.97657, p-value = 0.459

Tabela 10.

Hipoteza zerowa zakłada normalność rozkładu. P-value osiąga dość wysoką wartość. Stwierdzono brak podstaw do odrzucenia hipotezy zerowej. Można zatem stwierdzić, że rozkład jest normalny.

# **Wartości teoretyczne**

Wartości teoretyczne modelu zaprezentowano w Tabeli 11.

Courtelary Delemont Franches-Mnt Moutier Neuveville Porrentruy Broye Glane Gruyere

72.89274 80.82380 84.49005 75.69829 63.53068 89.75906 81.41183 82.74193 81.82180

Sarine Veveyse Aigle Aubonne Avenches Cossonay Echallens Grandson Lausanne

80.07518 84.96209 59.60831 65.68402 65.98577 67.00343 74.79546 70.98245 54.94859

La Vallee Lavaux Morges Moudon Nyone Orbe Oron Payerne Paysd'enhaut

52.06106 63.91476 63.11651 75.36169 62.36555 64.87869 73.05877 71.59485 69.07341

Rolle Vevey Yverdon Conthey Entremont Herens Martigwy Monthey St Maurice

61.43896 64.17333 71.63105 75.57410 76.87260 78.47373 77.38741 83.15511 73.08919

Sierre Sion Boudry La Chauxdfnd Le Locle Neuchatel Val de Ruz ValdeTravers V. De Geneve

76.05781 71.19277 66.99384 73.95507 68.55708 54.92513 71.61428 74.45610 34.65212

Rive Droite Rive Gauche

52.38306 57.47648

Tabela 11.

## **Reszty modelu**

Reszty modelu przedstawione zostały poniżej w Tabeli 12. i na Wykresie 3.

Courtelary Delemont Franches-Mnt Moutier Neuveville Porrentruy Broye Glane Gruyere

7.30725823 2.27619862 8.00994889 10.10171339 13.36931929 -13.65906440 2.38817305 9.65806794 0.57819514

Sarine Veveyse Aigle Aubonne Avenches Cossonay Echallens Grandson Lausanne

2.82482470 2.13791317 4.49168514 1.21597750 2.91422650 -5.30342833 -6.49545502 0.71755115 0.75141306

La Vallee Lavaux Morges Moudon Nyone Orbe Oron Payerne Paysd'enhaut

2.23894265 1.18524354 2.38348819 -10.36169321 -5.76554893 -7.47868656 -0.55876753 2.60514622 2.92658536

Rolle Vevey Yverdon Conthey Entremont Herens Martigwy Monthey St Maurice

-0.93896195 -5.87333311 -6.23104912 -0.07410455 -7.57260493 -1.17372629 -6.88740542 -3.75511010 -8.08919332

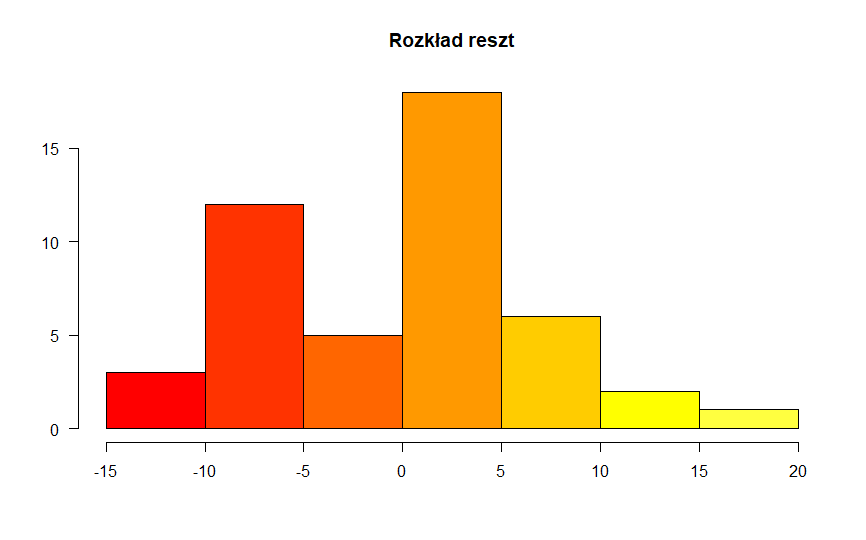
Sierre Sion Boudry La Chauxdfnd Le Locle Neuchatel Val de Ruz ValdeTravers V. De Geneve

16.14219254 8.10723156 3.40616473 -8.25506883 4.14291632 9.47487347 5.98571740 -6.85610435 0.34787822

Rive Droite Rive Gauche

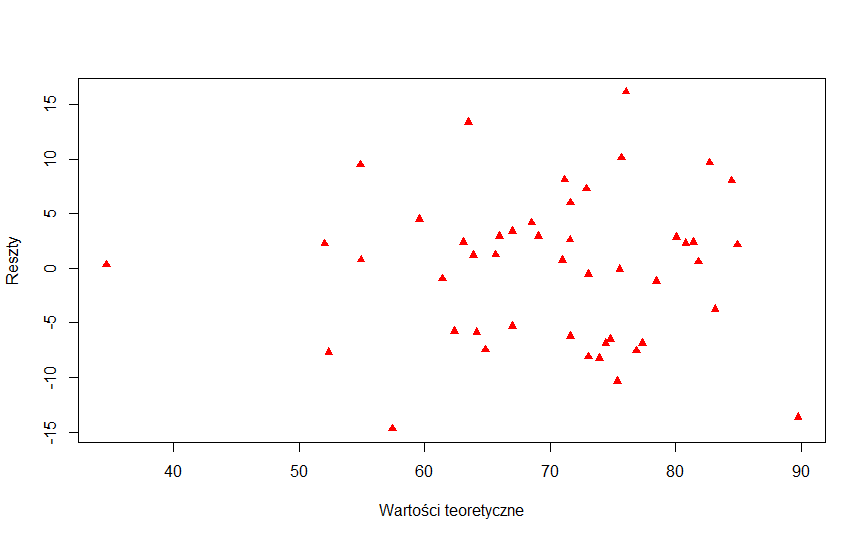
-7.68305930 -14.67648071

## Tabela 12.



Wykres 3.

Wykres 4. przedstawia wykres reszt względem wartości teoretycznych.



Wykres 4.

Przedstawiony wykres jest dość “chaotyczny”, nie wykazuje żadnych wyraźnych wzorców, co dobrze o nim świadczy (nie podejrzewamy wystąpienia np. heteroskedastyczności).

# **Analiza skupień**

* **Standaryzacja i macierz odległości**

Kolejnym etapem analizy było przeprowadzenie analizy skupień.

Dane poddano standaryzacji. W Tabeli 13. przedstawiono macierz odległości.

## 

Courtelary Delemont Franches-Mnt Moutier Neuveville Porrentruy Broye Glane Gruyere Sarine

Delemont 2.4856523

Franches-Mnt 2.9230303 1.1511610

Moutier 1.4496356 1.6504704 1.7793833

Neuveville 1.3807448 2.5482991 3.0703772 1.4727019

Porrentruy 2.7588023 1.7279340 2.6245852 2.6982243 3.2046480

Broye 3.1682167 1.7626459 2.3655291 2.4019300 2.8066100 2.1389420

Glane 3.3729623 1.8751543 2.3479396 2.6564346 3.2251551 2.1259579 0.8767479

Gruyere 2.7550491 1.0033764 1.3825273 1.7397279 2.5306859 2.1793003 1.2758690 1.7073698

Sarine 2.4488664 1.5295739 2.3074229 2.1750904 2.5042832 1.4887819 1.2972882 1.3966101 1.4679608

Veveyse 3.1978760 1.6332953 2.2040407 2.4700045 3.0582668 1.8409413 0.5665621 0.5142568 1.3794413 1.2080017

Aigle 3.1596773 3.6984111 4.0547601 2.8089232 2.0124982 4.5533026 3.6484652 4.3560378 3.2851713 3.8104653

Aubonne 2.7442168 2.9590643 3.4541996 2.2260691 1.6915094 3.7494589 3.0019152 3.6496854 2.7661542 3.3065602

Avenches 2.1952517 2.8746455 3.6498666 2.2759364 1.2899659 3.0795406 2.5690439 3.1172189 2.7664860 2.5615395

Cossonay 3.2045935 3.6847330 4.1594712 2.8766007 2.1619846 4.2487751 3.3512241 4.0981996 3.2638616 3.6882246

Echallens 2.8910734 2.8097033 3.3875211 2.3395178 2.0518741 3.2557722 2.2962177 3.0135942 2.4373016 2.8275370

Grandson 1.3620622 2.7160342 3.1412631 1.4947250 0.9605528 3.2645093 3.1054180 3.6052016 2.6641406 2.8417713

Lausanne 3.0003375 4.4490200 5.0918949 3.8070114 2.6745948 4.5645374 4.6384325 5.1502134 4.3508911 3.9659299

La Vallee 4.9369299 6.1177753 6.2267551 5.0942156 4.4160428 6.8688937 6.3911353 7.0277252 5.7344197 6.1181006

Lavaux 2.9136685 3.2676002 3.8643977 2.5889467 1.7542113 3.8347058 2.9371030 3.6300412 2.9617141 3.2607444

Morges 2.8002736 3.4871459 3.9123224 2.5404283 1.6756973 4.1882560 3.3360565 4.0349680 3.0910532 3.4801779

Moudon 2.2822023 2.7206904 3.4255678 2.1658537 1.8026170 2.9072333 2.7565465 3.3264592 2.7207007 2.8630517

Veveyse Aigle Aubonne Avenches Cossonay Echallens Grandson Lausanne La Vallee Lavaux

Delemont

Franches-Mnt

Moutier

Neuveville

Porrentruy

Broye

Glane

Gruyere

Sarine

Veveyse

Aigle 4.0953188

Aubonne 3.3788505 1.4030095

Avenches 2.9007789 2.1804542 1.5194648

Cossonay 3.7995129 1.1292035 1.1168782 1.7419617

Echallens 2.6988289 2.0734094 1.1751995 1.3653164 1.2839421

Grandson 3.3343260 1.9427208 1.6036548 1.5900537 1.9355015 1.9464552

Lausanne 4.9234840 2.9585941 3.5383594 2.9551436 3.3683884 3.8760229 2.7752858

La Vallee 6.7631483 3.3133294 4.5617109 5.0029213 4.1337777 5.1799556 4.1520944 3.4064574

Lavaux 3.3752293 1.3651606 0.7817486 1.1592809 0.7837653 1.0216069 1.8170746 3.2961687 4.5577031

Morges 3.7689379 0.5935944 1.1739107 1.6924762 0.7739543 1.6508788 1.5641390 2.8518894 3.6236049 0.9827633

Moudon 3.0013994 2.4198707 1.3344750 1.1990709 1.7677264 1.1481047 1.4974728 3.5568603 5.2297944 1.4436020

Morges Moudon Nyone Orbe Oron Payerne Paysd'enhaut Rolle Vevey Yverdon

Delemont

Franches-Mnt

Moutier

Neuveville

Porrentruy

Broye

Glane

Gruyere

Sarine

Veveyse

Aigle

Aubonne

Avenches

Cossonay

Echallens

Grandson

Lausanne

La Vallee

Lavaux

Morges

Moudon 1.9652603

Conthey Entremont Herens Martigwy Monthey St Maurice Sierre Sion Boudry La Chauxdfnd

Delemont

Franches-Mnt

Moutier

Neuveville

Porrentruy

Broye

Glane

Gruyere

Sarine

Veveyse

Aigle

Aubonne

Avenches

Cossonay

Echallens

Grandson

Lausanne

La Vallee

Lavaux

Morges

Moudon

Le Locle Neuchatel Val de Ruz ValdeTravers V. De Geneve Rive Droite

Delemont

Franches-Mnt

Moutier

Neuveville

Porrentruy

Broye

Glane

Gruyere

Sarine

Veveyse

Aigle

Aubonne

Avenches

Cossonay

Echallens

Grandson

Lausanne

La Vallee

Lavaux

Morges

Moudon

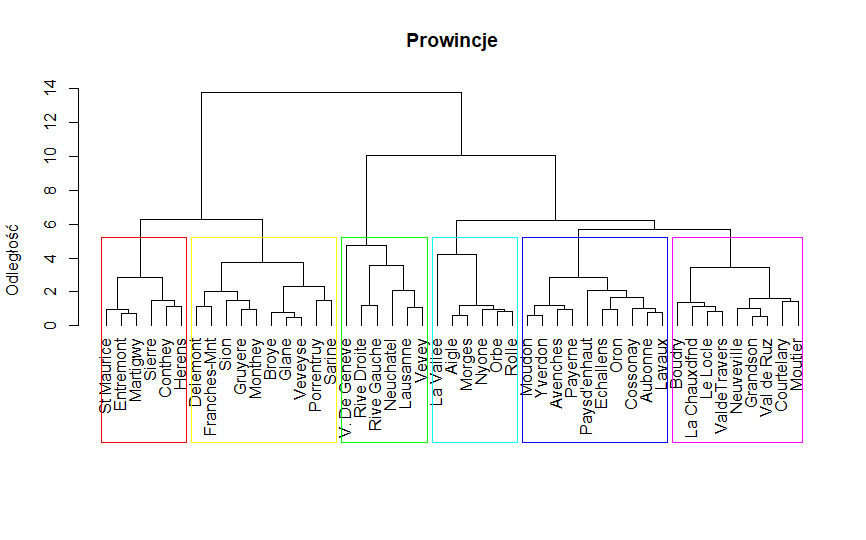
[ osiągnięto getOption("max.print") -- pominięto 25 wierszy]

Tabela 13.

## 

## **Grupowanie metodą Warda. Dendrogram**

Do grupowania prowincji wykorzystano metodę Warda. Wyniki zaprezentowano na Wykresie 5. i Wykresie 6. Przedstawia się dwa wykresy w zarówno pionowej jak i poziomej orientacji. Obserwacje zostały podzielone na 6 skupień.



Wykres 5.

## 

## 

## 

## Wykres 6.

## 

## **Przynależność do skupień**

W Tabeli 14. zaprezentowano przynależność prowincji do konkretnych skupień.

Courtelary Delemont Franches-Mnt Moutier Neuveville Porrentruy Broye Glane Gruyere

1 2 2 1 1 2 2 2 2

Sarine Veveyse Aigle Aubonne Avenches Cossonay Echallens Grandson Lausanne

2 2 3 4 4 4 4 1 5

La Vallee Lavaux Morges Moudon Nyone Orbe Oron Payerne Paysd'enhaut

3 4 3 4 3 3 4 4 4

Rolle Vevey Yverdon Conthey Entremont Herens Martigwy Monthey St Maurice

3 5 4 6 6 6 6 2 6

Sierre Sion Boudry La Chauxdfnd Le Locle Neuchatel Val de Ruz ValdeTravers V. De Geneve

6 2 1 1 1 5 1 1 5

Rive Droite Rive Gauche

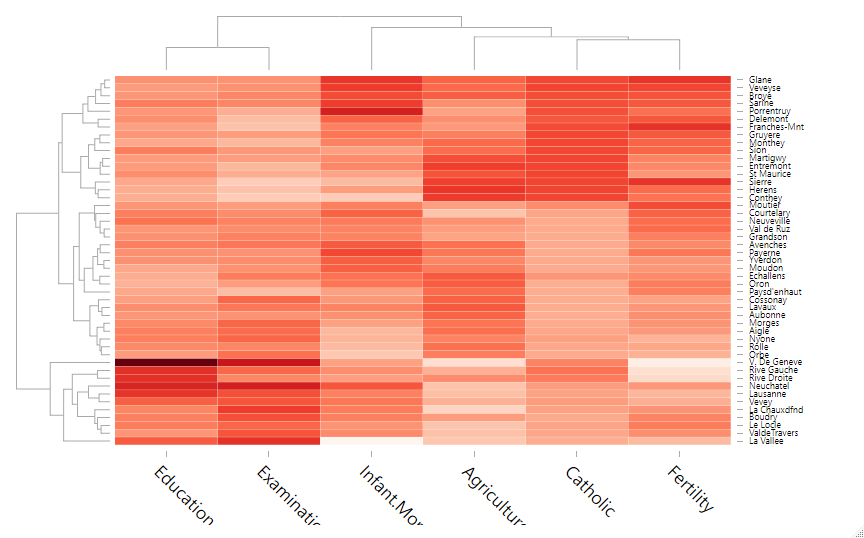
5 5

Tabela 14.

## **Mapa ciepła**

Poniżej przedstawiono mapę ciepła (Wykres 7.) prezentującą wartości zmiennych dla konkretnych prowincji wraz z dendrogramem. Im bardziej intensywny kolor, tym wyższe wartości osiągnęła zmienna dla danej prowincji.

Na pierwszy rzut oka zauważyć można na przykład, że V. De Geneve ma największy odsetek wykształconych obywateli i najmniejszy odsetek osób utrzymujących się z rolnictwa. Jednocześnie płodność w tej prowincji osiąga najmniejszą wartość.



Wykres 7.

# **Podsumowanie. Wnioski**

Przejście demograficzne to okres przejścia od wysokiego do niskiego poziomu urodzeń, zgonów i przyrostu naturalnego stabilizującego rozmiar populacji, typowego dla krajów rozwiniętych. Analizując problem badawczy udało się dowiedzieć, jakie zmienne wpływają istotnie na płodność w XIX-wiecznej Szwajcarii.

Największy wpływ na wysokość płodności mają edukacja i wyniki testów. Są to zależności ujemne, co oznacza, że im wyższe wykształcenie i wyniki osiągają obywatele, tym niższą płodność można zaobserwować.

Osoby utrzymujące się z rolnictwa osiągają niższe wyniki na testach i są gorzej wykształceni. Wykształcenie pozytywnie wpływa na wyniki testów, z którymi statystycznie gorzej radzą sobie Katolicy.

Model najlepiej objaśniający płodność zawierał zmienne dotyczące wykształcenia, wiary, śmiertelności niemowląt oraz pozyskiwania środków do życia z rolnictwa.

Przeprowadzone testy dowiodły, że każda ze zmiennych jest istotna, a usunięcie choć jednej negatywnie wpływa na model.

W analizie skupień, prowincje zostały podzielone na 6 grup po przeprowadzeniu grupowania z wykorzystaniem metody Warda.