



POLITECHNIKA RZESZOWSKA  
im. Ignacego Łukasiewicza  
WYDZIAŁ MATEMATYKI i FIZYKI STOSOWANEJ

Projektowanie modeli łączenia źródeł danych

## **Analiza frekwencji w kinach w Polsce w latach 2013-2023**

---

Małgorzata Radomska 173203  
Adrianna Rapa 173204

Inżynieria i Analiza Danych

Rzeszów 2024

# Spis treści

|       |                                                                    |    |
|-------|--------------------------------------------------------------------|----|
| 1     | Wstęp.....                                                         | 3  |
| 1.1   | Cel analizy .....                                                  | 3  |
| 1.2   | Opis danych .....                                                  | 3  |
| 1.3   | Przegląd danych.....                                               | 4  |
| 2     | Statystyki opisowe i wizualizacje danych .....                     | 4  |
| 2.1   | Parametry statystyczne.....                                        | 4  |
| 2.2   | Wizualizacje danych .....                                          | 5  |
| 3     | Dobór zmiennych do modelu .....                                    | 7  |
| 3.1   | Macierz współczynników korelacji .....                             | 7  |
| 3.2   | Metoda Hellwiga – metoda wskaźników pojemności informacyjnej ..... | 8  |
| 3.3   | Porównanie modeli .....                                            | 9  |
| 4     | Estymacja parametrów modelu .....                                  | 12 |
| 5     | Weryfikacja modelu .....                                           | 13 |
| 5.1   | Dopasowanie modelu.....                                            | 13 |
| 5.2   | Istotność modelu jako całości .....                                | 13 |
| 5.3   | Istotność poszczególnych zmiennych .....                           | 13 |
| 5.4   | Współczynniki i interpretacja .....                                | 13 |
| 5.5   | Błąd standardowy modelu.....                                       | 14 |
| 5.6   | Współliniowość zmiennych .....                                     | 14 |
| 5.7   | Wyrazistość modelu.....                                            | 15 |
| 5.8   | Koincydencja .....                                                 | 16 |
| 6     | Testy statystyczne.....                                            | 16 |
| 6.1   | Testy normalności.....                                             | 16 |
| 6.1.1 | Test Shapiro-wilka.....                                            | 16 |
| 6.1.2 | Test Kołmogorowa-Smirnowa.....                                     | 17 |
| 6.2   | Testy autokorelacji składnika losowego .....                       | 18 |
| 6.2.1 | Test Durbin-Watsona.....                                           | 18 |
| 6.2.2 | Test Breuscha-Godfrey'a .....                                      | 19 |
| 6.3   | Testy heteroskedastyczności .....                                  | 20 |
| 6.3.1 | Test Breuscha-Pagana.....                                          | 20 |
| 6.3.2 | Test White'a .....                                                 | 21 |
| 6.4   | Test serii.....                                                    | 22 |
| 6.5   | Test symetrii.....                                                 | 23 |
| 7     | Podsumowanie i wnioski.....                                        | 25 |

# 1 Wstęp

Projekt realizowany jest w ramach przedmiotu „Projektowanie modeli łączenia źródeł danych” na kierunku Inżynieria i Analiza Danych, semestr czwarty, grupa L4. Celem projektu jest zastosowanie zdobytej wiedzy teoretycznej i praktycznej do analizy danych oraz budowy modelu ekonometrycznego dla danych dotyczących frekwencji w kinach na terenie Polski w latach 2013-2023.

## 1.1 Cel analizy

Celem niniejszej analizy jest określenie, jakie czynniki wpływają na liczbę widzów w kinach w Polsce, oraz opracowanie modelu ekonometrycznego opisującego tę zależność. Analiza opiera się na danych z lat 2013–2023 i obejmuje zmienne związane z infrastrukturą kinową, czynnikami demograficznymi oraz kosztami dla konsumentów. Uzyskane wyniki pozwolą zrozumieć badane zjawisko oraz posłużyć do prognozowania liczby widzów w kinach, co może być przydatne dla branży rozrywkowej w Polsce.

## 1.2 Opis danych

Dane zostały pobrane ze strony Banku Danych Lokalnych dostępnej pod linkiem [bdl.stat.gov.pl](http://bdl.stat.gov.pl). Zawierają informacje o 17 cechach opisujących zjawisko frekwencji w polskich kinach. Dane są roczne. Data pobrania danych to styczeń 2025.

Zestaw obejmuje dane z zakresu lat 2013-2023. Przedział czasowy został wybrany w taki sposób, aby uniknąć braków informacji.

- **Zmienna zależna (Y):**

Y – liczba widzów w kinach w Polsce

- **Zmienne niezależne (X):**

X1 - liczba kin w Polsce

X2 - liczba sal kinowych w Polsce

X3 - liczba miejsc na widowni w kinach w Polsce

X4 - liczba seansów w kinach w Polsce

X5 - liczba ludności w wieku przedprodukcyjnym w miastach w Polsce

X6 - liczba ludności w wieku produkcyjnym w miastach w Polsce

X7 - przeciętne miesięczne wynagrodzenie brutto w Polsce [zł]

X8 - liczba ludności aktywnej zawodowo pracującej w Polsce [tys.]

X9 - liczba abonentów telewizji w Polsce

X10 - liczba konsumentów korzystających z łączności abonenckich w Polsce [tys.]

X11 - przeciętna cena detaliczna biletu do kina w Polsce [zł]

X12 - liczba imprez masowych artystyczno-rozrywkowych w Polsce

X13 - przeciętne miesięczne wydatki na 1 osobę na rekreację i kulturę w Polsce [zł]

X14 - ludność w miastach w % ogółu ludności w Polsce

X15 - liczba osób uczestniczących w imprezach masowych w Polsce

X16 - liczba samochodów osobowych w Polsce

## X17 - liczba przewozów pasażerskich komunikacją miejską na 1 mieszkańca w Polsce

### 1.3 Przegląd danych

- Liczba cech: 17
- Liczba rekordów: 11
- Zakres lat: 2013-2023

Tabela zawierająca dane:

| x1    | x2    | x3      | x4        | x5        | x6         | x7         | x8     | x9         |
|-------|-------|---------|-----------|-----------|------------|------------|--------|------------|
| 469   | 1 243 | 271 781 | 1 645 637 | 3 890 766 | 14 763 159 | 3 877,43   | 15 110 | 6 718 004  |
| 463   | 1 243 | 266 479 | 1 756 954 | 3 875 671 | 14 568 489 | 4 003,99   | 15 454 | 6 696 615  |
| 444   | 1 276 | 271 011 | 1 792 660 | 3 869 484 | 14 358 387 | 4 150,86   | 15 723 | 6 665 101  |
| 484   | 1 364 | 283 400 | 1 988 399 | 3 881 164 | 14 142 759 | 4 290,52   | 15 915 | 6 639 706  |
| 491   | 1 416 | 287 948 | 2 080 183 | 3 912 079 | 13 933 976 | 4 527,89   | 16 190 | 6 639 470  |
| 497   | 1 454 | 290 642 | 2 205 346 | 3 932 445 | 13 729 602 | 4 834,76   | 16 283 | 6 567 510  |
| 523   | 1 513 | 293 584 | 2 203 964 | 3 953 178 | 13 542 227 | 5 181,63   | 16 744 | 6 508 247  |
| 511   | 1 511 | 290 710 | 1 098 976 | 3 847 634 | 13 408 468 | 5 523,32   | 16 734 | 6 401 356  |
| 523   | 1 539 | 294 252 | 1 355 464 | 3 847 374 | 13 218 522 | 6 001,02   | 17 161 | 6 272 113  |
| 532   | 1 565 | 296 948 | 2 130 995 | 3 828 264 | 13 043 706 | 6 705,62   | 17 251 | 4 703 846  |
| 535   | 1 585 | 298 224 | 2 131 204 | 3 784 453 | 12 901 595 | 7 595,30   | 17 307 | 4 526 677  |
| x10   | x11   | x12     | x13       | x14       | x15        | x16        | x17    | y          |
| 4 318 | 17,06 | 5 730   | 69,22     | 60,42     | 21 133 415 | 19 389 446 | 94,0   | 36 936 592 |
| 4 332 | 17,65 | 6 070   | 70,13     | 60,34     | 23 253 936 | 20 003 863 | 96,4   | 41 030 024 |
| 3 991 | 18,23 | 6 280   | 73,48     | 60,27     | 23 297 417 | 20 723 423 | 95,5   | 45 681 501 |
| 3 738 | 18,73 | 6 480   | 78,06     | 60,18     | 24 435 646 | 21 675 388 | 98,0   | 51 823 896 |
| 3 270 | 19,48 | 6 876   | 81,71     | 60,13     | 26 201 281 | 22 503 579 | 97,3   | 56 718 898 |
| 2 771 | 19,80 | 7 074   | 76,93     | 60,05     | 27 615 218 | 23 429 016 | 98,3   | 59 177 816 |
| 2 336 | 20,02 | 6 949   | 82,53     | 60,01     | 27 826 562 | 24 360 166 | 100,7  | 61 710 894 |
| 2 054 | 19,34 | 1 564   | 69,47     | 59,80     | 4 915 276  | 25 113 862 | 59,5   | 19 490 969 |
| 1 780 | 19,56 | 2 964   | 79,55     | 59,68     | 10 982 909 | 25 869 804 | 65,8   | 27 464 839 |
| 1 570 | 20,82 | 5 247   | 89,62     | 59,55     | 23 018 676 | 26 457 659 | 80,9   | 41 416 969 |
| 1 409 | 21,78 | 6 313   | 100,82    | 59,45     | 26 942 644 | 27 227 691 | 86,1   | 49 741 313 |

## 2 Statystyki opisowe i wizualizacje danych

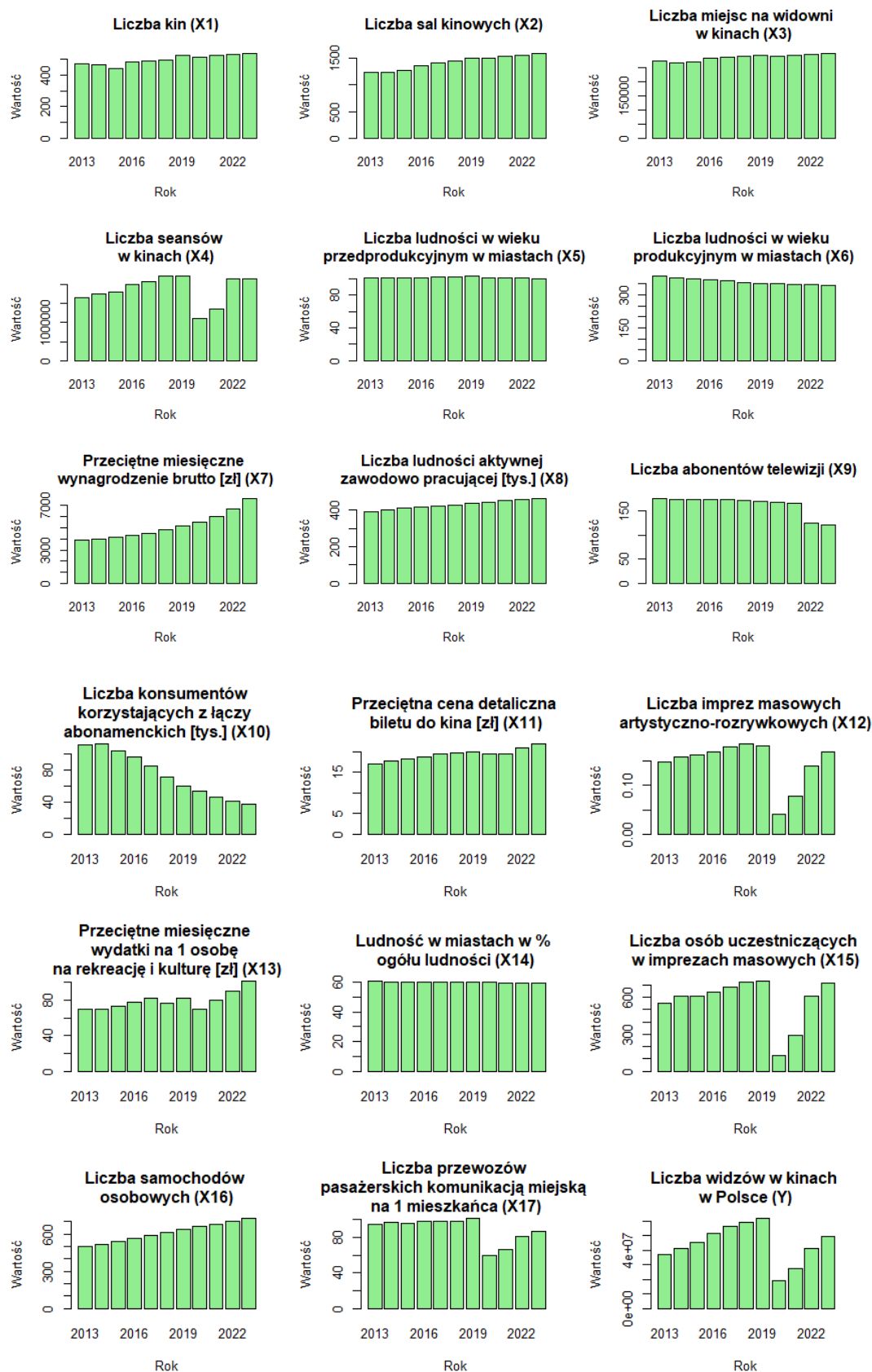
### 2.1 Parametry statystyczne

Obliczono podstawowe parametry statystyczne, tj. odchylenie standardowe, średnią arytmetyczną, współczynnik zmienności, medianę wartości minimalną i maksymalną.

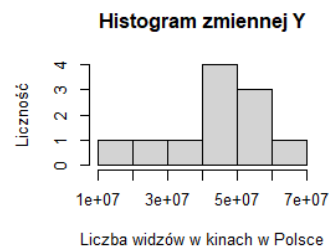
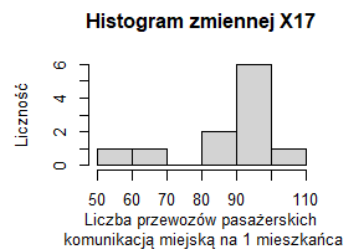
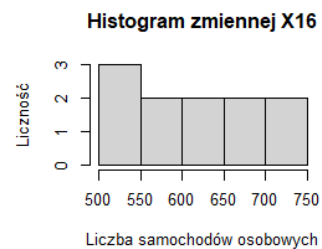
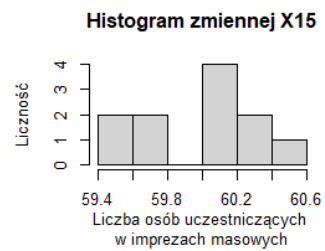
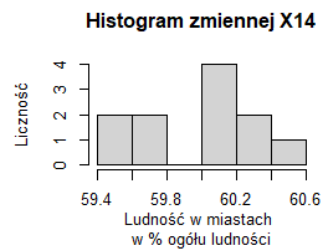
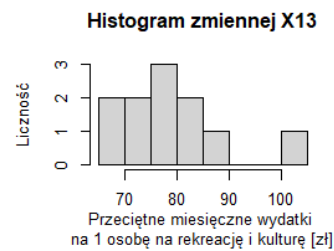
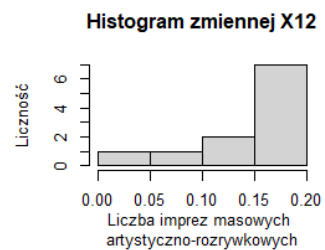
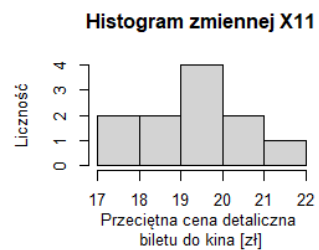
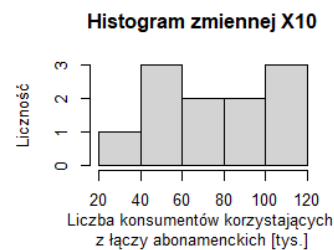
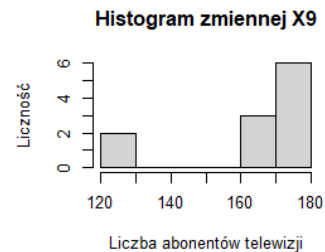
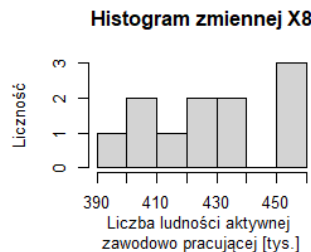
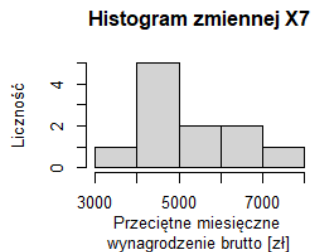
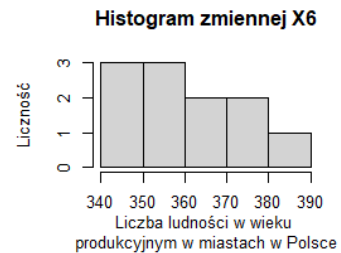
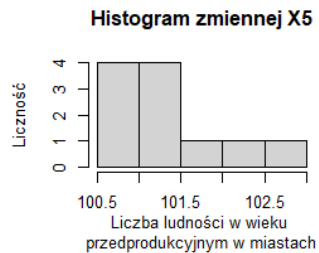
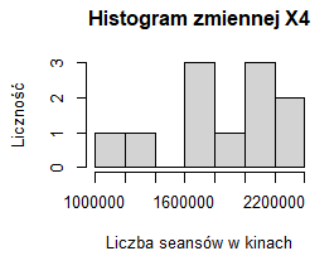
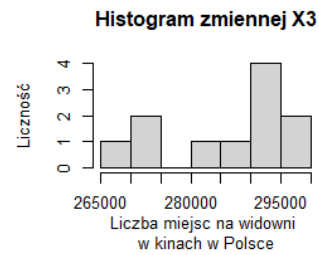
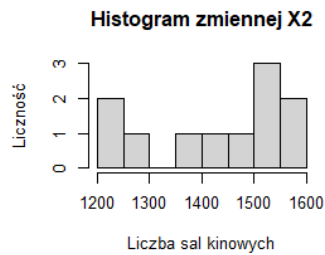
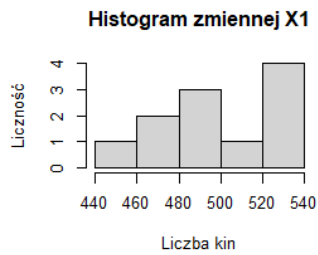
| x1          | x2          | x3          | x4          | x5          | x6         | x7         | x8         | x9         |                        |
|-------------|-------------|-------------|-------------|-------------|------------|------------|------------|------------|------------------------|
| 28,87075789 | 122,9644846 | 10689,88121 | 349532,463  | 45725,03004 | 595754,505 | 1144,95109 | 716,105503 | 764574,29  | Odchylenie standardowe |
| 497,4545455 | 1428,090909 | 285907,1818 | 1853616,545 | 3874773,818 | 13782808,2 | 5153,84909 | 16352      | 6212604,09 | Średnia                |
| 5,80%       | 8,61%       | 3,74%       | 18,86%      | 1,18%       | 4,32%      | 22,22%     | 4,38%      | 12,31%     | Vz - wsp. zmienności   |
| 497         | 1454        | 290642      | 1988399     | 3875671     | 13729602   | 4834,76    | 16283      | 6567510    | Mediana                |
| 444         | 1243        | 266479      | 1098976     | 3784453     | 12901595   | 3877,43    | 15110      | 4526677    | Min                    |
| 535         | 1585        | 298224      | 2205346     | 3953178     | 14763159   | 7595,3     | 17307      | 6718004    | Max                    |
| x10         | x11         | x12         | x13         | x14         | x15        | x16        | x17        | y          |                        |
| 1060,773599 | 1,298183091 | 1677,81697  | 9,09151498  | 0,310847947 | 6949735,83 | 2566971,58 | 13,3958918 | 12568358,1 | Odchylenie standardowe |
| 2869,909091 | 19,31545455 | 5595,181818 | 79,22909091 | 59,98909091 | 21783907,3 | 23341263,4 | 88,4090909 | 44653973,7 | Średnia                |
| 36,96%      | 6,72%       | 29,99%      | 11,47%      | 0,52%       | 31,90%     | 11,00%     | 15,15%     | 28,15%     | Vz - wsp. zmienności   |
| 2771        | 19,48       | 6280        | 78,06       | 60,05       | 23297417   | 23429016   | 95,5       | 45681501   | Mediana                |
| 1409        | 17,06       | 1564        | 69,22       | 59,45       | 4915276    | 19389446   | 59,5       | 19490969   | Min                    |
| 4332        | 21,78       | 7074        | 100,82      | 60,42       | 27826562   | 27227691   | 100,7      | 61710894   | Max                    |

## 2.2 Wizualizacje danych

- Wykresy słupkowe zmiennych



- Histogramy zmiennych



### 3 Dobór zmiennych do modelu

W celu dobrania zmiennych do modelu ekonometrycznego opisującego analizowane zjawisko wykorzystano następujące metody: macierz współczynników korelacji, metoda Hellwiga oraz analizę wartości regresji liniowej.

#### 3.1 Macierz współczynników korelacji

Na podstawie zebranych danych utworzono macierz współczynników korelacji. W kolejnym kroku

obliczono wartość statystyki T-studenta oraz wartość krytyczną współczynnika korelacji  $r^* = \sqrt{\frac{T^2}{T^2 + n - 2}}$ .

Następnie zastosowano poniższy algorytm:

1. Ze zbioru zmiennych eliminowane są te, których wartość bezwzględna korelacji ze zmienną zależną jest mniejsza bądź równa wartości krytycznej:

$$|r_{x_iy}| \leq r^*.$$

2. Spośród zmiennych, które nie zostały wyeliminowane w poprzednim kroku, wybierana jest ta, której wartość bezwzględna korelacji ze zmienną zależną jest największa:

$$|r_{x_{h_y}}| = \max\{|r_{x_iy}|\}.$$

3. Wśród zbioru zmiennych niezależnych wybierane są te, których wartość bezwzględna korelacji ze zmienną wyselekcjonowaną wcześniej jest większa od wartości krytycznej;

$$|r_{x_hx_i}| > r^*.$$

Macierz współczynników korelacji:

|     | x1       | x2       | x3       | x4       | x5        | x6       | x7       | x8       | x9       | x10      | x11      | x12      | x13      | x14          | x15          | x16      | x17          | y        |
|-----|----------|----------|----------|----------|-----------|----------|----------|----------|----------|----------|----------|----------|----------|--------------|--------------|----------|--------------|----------|
| x1  | 1        | 0,958969 | 0,943837 | 0,135831 | -0,318065 | -0,92906 | 0,874507 | 0,921763 | -0,69065 | -0,944   | 0,861713 | -0,28143 | 0,705274 | -0,893504139 | -0,133021089 | 0,93639  | -0,471445682 | -0,0392  |
| x2  | 0,958969 | 1        | 0,978933 | 0,120258 | -0,33727  | -0,98558 | 0,890278 | 0,978255 | -0,67594 | -0,98497 | 0,922913 | -0,30713 | 0,708586 | -0,933936598 | -0,165853612 | 0,984094 | -0,513781077 | -0,02797 |
| x3  | 0,943837 | 0,978933 | 1        | 0,216623 | -0,225869 | -0,94071 | 0,824079 | 0,92728  | -0,61791 | -0,93677 | 0,910693 | -0,20204 | 0,713449 | -0,870167726 | -0,072843622 | 0,935089 | -0,408473849 | 0,090061 |
| x4  | 0,135831 | 0,120258 | 0,216623 | 1        | 0,285165  | -0,09657 | 0,110032 | 0,081392 | -0,29097 | -0,03248 | 0,407304 | 0,890366 | 0,576283 | -0,024287    | 0,947280604  | 0,063179 | 0,758748102  | 0,925906 |
| x5  | -0,31806 | -0,33727 | -0,22587 | 0,285165 | 1         | 0,463657 | -0,67229 | -0,44428 | 0,738807 | 0,44207  | -0,37703 | 0,446361 | -0,46008 | 0,639996961  | 0,344753237  | -0,48225 | 0,585946495  | 0,535262 |
| x6  | -0,92906 | -0,98558 | -0,94071 | -0,09657 | 0,4636574 | 1        | -0,93851 | -0,99366 | 0,750626 | 0,990985 | -0,93525 | 0,32975  | -0,74224 | 0,975104106  | 0,179993056  | -0,99854 | 0,55040322   | 0,081774 |
| x7  | 0,874507 | 0,890278 | 0,824079 | 0,110032 | -0,672286 | -0,93851 | 1        | 0,921778 | -0,90698 | -0,93159 | 0,900192 | -0,27973 | 0,827841 | -0,979930036 | -0,114216855 | 0,946244 | -0,516979496 | -0,13505 |
| x8  | 0,921763 | 0,978255 | 0,92728  | 0,081392 | -0,444285 | -0,99366 | 0,921778 | 1        | -0,72352 | -0,9838  | 0,915801 | -0,33943 | 0,729906 | -0,964273785 | -0,192119223 | 0,991309 | -0,553171138 | -0,08869 |
| x9  | -0,69065 | -0,67594 | -0,61791 | -0,29097 | 0,7388074 | 0,750626 | -0,90698 | -0,72352 | 1        | 0,728671 | -0,79158 | 0,060607 | -0,8533  | 0,842084367  | -0,110268385 | -0,75515 | 0,296082969  | 0,035522 |
| x10 | -0,944   | -0,98497 | -0,93677 | -0,03248 | 0,44207   | 0,990985 | -0,93159 | -0,9838  | 0,728671 | 1        | -0,89962 | 0,389701 | -0,69037 | 0,966713879  | 0,241257053  | -0,99496 | 0,597754808  | 0,14474  |
| x11 | 0,861713 | 0,922913 | 0,910693 | 0,407304 | -0,377031 | -0,93525 | 0,900192 | 0,915801 | -0,79158 | -0,89962 | 1        | 0,001159 | 0,885472 | -0,895499736 | 0,152652748  | 0,92316  | -0,248531915 | 0,2341   |
| x12 | -0,28143 | -0,30713 | -0,20204 | 0,890366 | 0,446361  | 0,32975  | -0,27973 | -0,33943 | 0,060607 | 0,389701 | 0,001159 | 1        | 0,26212  | 0,388061645  | 0,983501477  | -0,35972 | 0,960739206  | 0,930124 |
| x13 | 0,705274 | 0,708586 | 0,713449 | 0,576283 | -0,460078 | -0,74224 | 0,827841 | 0,729906 | -0,8533  | -0,69037 | 0,885472 | 0,26212  | 1        | -0,756634932 | 0,40891013   | 0,731586 | 0,004649691  | 0,369975 |
| x14 | -0,8935  | -0,93394 | -0,87017 | -0,02429 | 0,639997  | 0,975104 | -0,97993 | -0,96427 | 0,842084 | 0,966714 | -0,8955  | 0,388062 | -0,75663 | 1            | 0,230434354  | -0,97969 | 0,614455358  | 0,20364  |
| x15 | -0,13302 | -0,16585 | -0,07284 | 0,947281 | 0,3447532 | 0,179993 | -0,11422 | -0,19212 | -0,11027 | 0,241257 | 0,152653 | 0,983501 | 0,40891  | 0,230434354  | 1            | -0,20925 | 0,904394379  | 0,930139 |
| x16 | 0,93639  | 0,984094 | 0,935089 | 0,063179 | -0,482255 | -0,99854 | 0,946244 | 0,991309 | -0,75515 | -0,99496 | 0,92316  | -0,35972 | 0,731586 | -0,979689219 | -0,209245446 | 1        | -0,576719689 | -0,118   |
| x17 | -0,47145 | -0,51378 | -0,40847 | 0,758748 | 0,5859465 | 0,550403 | -0,51698 | -0,55312 | 0,296083 | 0,597755 | -0,24853 | 0,960739 | 0,00465  | 0,614455358  | 0,904394379  | -0,57672 | 1            | 0,860225 |
| y   | -0,0392  | -0,02797 | 0,090061 | 0,925906 | 0,5352616 | 0,081774 | -0,13505 | -0,08869 | 0,035522 | 0,14474  | 0,2341   | 0,930124 | 0,369975 | 0,203640458  | 0,930138659  | -0,118   | 0,860225043  | 1        |

Statystyka T-studenta dla poziomu istotności  $\alpha = 0,05$  i  $n - 2$  stopni swobody oraz wartość  $r^*$ :

|            |          |    |          |
|------------|----------|----|----------|
| n          | 11       | r* | 0,602069 |
| alfa       | 0,05     |    |          |
| T_alfa,n-2 | 2,262157 |    |          |

Postępowanie zgodnie z algorytmem w krokach:

1. Eliminacja zmiennych: X1, X2, X3, X5, X6, X7, X8, X9, X10, X11, X13, X14 i X16.
2. Wybór zmiennej X15.
3. Eliminacja zmiennych: X4, X12 i X17.

Po zastosowaniu algorytmu na macierzy współczynników korelacji otrzymano model postaci:

$$y = a_1X_{15} + b.$$

### 3.2 Metoda Hellwiga – metoda wskaźników pojemności informacyjnej

W tej metodzie skupiono się na doborze takiej kombinacji zmiennych do modelu, która posiada maksymalny integralny wskaźnik pojemności informacyjnej. Wykorzystano do tego następujący kod w języku R:

```
2 #@Autor, napisał, opisał, przetestował: Napora Jarosław
3
4 #dane najpierw x później y
5
6
7 #----- Hellwig z danych -----
8 library(readxl)
9 Dane=read_excel("C:/Users/gosia/Desktop/Studia/Semestr V/Projektowanie modeli
10                 łączenia źródeł danych/Projekt/Kino.xlsx", sheet=2, range="B33:S44")
11
12 N=length(Dane)-1      #oblicza ilość zmiennych objaśnianych
13 M=2^N-1               #oblicza ilość kombinacji 0-1
14
15 str(Dane)
16 zm_obj=Dane[,1:N]     #tworzy macierz zmiennych objaśnianych
17 r=cor(zm_obj, use = "pairwise.complete.obs") #tworzy macierz korelacji między zmiennymi
18 r=as.matrix(abs(r))   #tworzy wartości bezwzględne i macierz
19 R=cor(zm_obj,Dane[,N+1]) #tworzy wektor korelacji Y z każdą ze zmiennych
20 R=as.vector(R)        #zapisuje wektor jako wektor
21
22 tab=as.matrix(expand.grid(rep(list(0:1), N))[-1,]) #tworzy macierz 0-1 kombinacji
23 colnames(tab)=colnames(Dane)[1:N]
24 wyniki=matrix(0,M,N) #tworzy macierz 0 na wyniki czastkowe pojemnosci
25 colnames(wyniki)=colnames(Dane)[1:N]
26 for(i in 1:M)
27 {
28   for(j in 1:N)
29   {
30     if(tab[i,j]!=0){wyniki[i,j]=(R[j]^2)/(tab[i,j]%*(as.vector(r[,j])))}
31   }
32 }
33
34 maks=which.max(rowSums(wyniki))
35 tab[maks,]
36
37 #Podgląd kilku najlepszych wynik?w
38 wynikis=cbind(wyniki,0)
39 wynikis[, (N+1)]=rowSums(wyniki)
40 nazwy=colnames(as.data.frame(wyniki))
41 colnames(wynikis)=c(nazwy,"hellwigP")
42
43 ind=order(wynikis[, (N+1)],decreasing = TRUE)[1:15] #zwróci 15 najleprzych
44 najlepsze15=wynikis[ind,]
45
46
```

W wyniku działania powyższego skryptu otrzymujemy zestaw zmiennych objaśniających, który daje maksymalną wartość integralnego wskaźnika pojemności.

```
> tab[maks,]
  x1  x2  x3  x4  x5  x6  x7  x8  x9 x10 x11 x12 x13 x14 x15 x16 x17
  0   0   0   0   1   0   0   0   0   0   0   1   1   0   0   0   1
```



Otrzymujemy również 15 najlepszych kombinacji:

|    | x1 | x2 | x3 | x4        | x5        | x6 | x7 | x8 | x9 | x10 | x11        | x12        | x13        | x14        | x15        | x16 | x17       | hellwigP  |
|----|----|----|----|-----------|-----------|----|----|----|----|-----|------------|------------|------------|------------|------------|-----|-----------|-----------|
| 1  | 0  | 0  | 0  | 0.0000000 | 0.1666791 | 0  | 0  | 0  | 0  | 0   | 0.00000000 | 0.3464645  | 0.10249630 | 0.00000000 | 0.00000000 | 0   | 0.3340285 | 0.9496684 |
| 2  | 0  | 0  | 0  | 0.0000000 | 0.1425089 | 0  | 0  | 0  | 0  | 0   | 0.00000000 | 0.2486846  | 0.07744923 | 0.00000000 | 0.2402140  | 0   | 0.2380572 | 0.9469139 |
| 3  | 0  | 0  | 0  | 0.2139298 | 0.1194048 | 0  | 0  | 0  | 0  | 0   | 0.00000000 | 0.2115027  | 0.00000000 | 0.00000000 | 0.2097117  | 0   | 0.1915804 | 0.9461294 |
| 4  | 0  | 0  | 0  | 0.2360384 | 0.1328197 | 0  | 0  | 0  | 0  | 0   | 0.00000000 | 0.2548802  | 0.07159975 | 0.00000000 | 0.00000000 | 0   | 0.2488113 | 0.9441495 |
| 5  | 0  | 0  | 0  | 0.2805527 | 0.1359191 | 0  | 0  | 0  | 0  | 0   | 0.00000000 | 0.2783023  | 0.00000000 | 0.00000000 | 0.00000000 | 0   | 0.2492009 | 0.9439751 |
| 6  | 0  | 0  | 0  | 0.2638948 | 0.1353435 | 0  | 0  | 0  | 0  | 0   | 0.00000000 | 0.2756868  | 0.00000000 | 0.00000000 | 0.2682407  | 0   | 0.0000000 | 0.9431658 |
| 7  | 0  | 0  | 0  | 0.2757711 | 0.1356366 | 0  | 0  | 0  | 0  | 0   | 0.00000000 | 0.00000000 | 0.00000000 | 0.00000000 | 0.2760253  | 0   | 0.2544020 | 0.9418350 |
| 8  | 0  | 0  | 0  | 0.1870335 | 0.1170062 | 0  | 0  | 0  | 0  | 0   | 0.00000000 | 0.1976961  | 0.05840520 | 0.00000000 | 0.1896956  | 0   | 0.1913500 | 0.9411866 |
| 9  | 0  | 0  | 0  | 0.0000000 | 0.1706730 | 0  | 0  | 0  | 0  | 0   | 0.00000000 | 0.3860717  | 0.00000000 | 0.00000000 | 0.3817748  | 0   | 0.0000000 | 0.9385195 |
| 10 | 0  | 0  | 0  | 0.0000000 | 0.1662545 | 0  | 0  | 0  | 0  | 0   | 0.00000000 | 0.00000000 | 0.09240899 | 0.00000000 | 0.3314154  | 0   | 0.3434400 | 0.9335190 |
| 11 | 0  | 0  | 0  | 0.3732247 | 0.1569599 | 0  | 0  | 0  | 0  | 0   | 0.00000000 | 0.4012122  | 0.00000000 | 0.00000000 | 0.00000000 | 0   | 0.0000000 | 0.9313967 |
| 12 | 0  | 0  | 0  | 0.0000000 | 0.1460831 | 0  | 0  | 0  | 0  | 0   | 0.00000000 | 0.2709322  | 0.00000000 | 0.00000000 | 0.2732389  | 0   | 0.2384138 | 0.9286680 |
| 13 | 0  | 0  | 0  | 0.2326446 | 0.1325500 | 0  | 0  | 0  | 0  | 0   | 0.00000000 | 0.00000000 | 0.06652680 | 0.00000000 | 0.2423649  | 0   | 0.2539960 | 0.9280823 |
| 14 | 0  | 0  | 0  | 0.0000000 | 0.1478147 | 0  | 0  | 0  | 0  | 0   | 0.03823207 | 0.3821609  | 0.00000000 | 0.00000000 | 0.3542842  | 0   | 0.0000000 | 0.9224919 |
| 15 | 0  | 0  | 0  | 0.2619366 | 0.1352936 | 0  | 0  | 0  | 0  | 0   | 0.00000000 | 0.2469634  | 0.00000000 | 0.02597319 | 0.2520119  | 0   | 0.0000000 | 0.9221786 |

Kombinację zmiennych z największym integralnym wskaźnikiem pojemności wykorzystano do stworzenia modelu postaci:

$$y = a_1X_5 + a_2X_{12} + a_3X_{13} + a_4X_{17} + b.$$

Sprawdzone zostały korelacje pomiędzy wybranymi zmiennymi objaśniającymi za pomocą macierzy współczynników korelacji:

|     | x5       | x12      | x13         | x17       | y         |
|-----|----------|----------|-------------|-----------|-----------|
| x5  | 1        | 0,446361 | -0,46007751 | 0,5859465 | 0,5352616 |
| x12 | 0,446361 | 1        | 0,262120014 | 0,9607392 | 0,9301239 |
| x13 | -0,46008 | 0,26212  | 1           | 0,0046497 | 0,3699755 |
| x17 | 0,585946 | 0,960739 | 0,004649691 | 1         | 0,860225  |
| y   | 0,535262 | 0,930124 | 0,369975487 | 0,860225  | 1         |

Zauważono, że zmienne X12 i X17 są ze sobą mocno skorelowane. Aby zapewnić niezależność zmiennych objaśniających stworzono dwa kolejne modele – pierwszy z wyeliminowaniem zmiennej X17, a drugi z usunięciem zmiennej X12. Modele te można przedstawić wzorami:

$$y = a_1X_5 + a_2X_{13} + a_3X_{12} + b;$$

$$y = a_1X_5 + a_2X_{12} + a_3X_{17} + b.$$

### 3.3 Porównanie modeli

Po stworzeniu czterech modeli z różnymi kombinacjami zmiennych objaśniających dokonano porównania, które miało na celu wyselekcjonowania ostatecznej postaci modelu opisującego zmienną Y. Do tego działania wykorzystano analizę regresji liniowej: wartości  $R^2$  ( $> 0,75$ ), dopasowanego  $R^2$  ( $> 0,75$ ), istotności F ( $< 0,05$ ) oraz p-value ( $< 0,05$ ). Dodatkowo wykorzystano funkcję AIC w języku R do porównania wartości kryterium informacyjnego Akaikego.

|                                                |                                                                                  |
|------------------------------------------------|----------------------------------------------------------------------------------|
| Model I - macierz korelacji                    | $y = a \cdot x_{15} + b$                                                         |
| Model II - metoda Hellwiga                     | $y = a_1 \cdot x_5 + a_2 \cdot x_{12} + a_3 \cdot x_{13} + a_4 \cdot x_{17} + b$ |
| Model III - metoda Hellwiga (bez zmiennej X17) | $y = a_1 \cdot x_5 + a_2 \cdot x_{12} + a_3 \cdot x_{13} + b$                    |
| Model IV - metoda Hellwiga (bez zmiennej X12)  | $y = a_1 \cdot x_5 + a_2 \cdot x_{13} + a_3 \cdot x_{17} + b$                    |

|         |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
|---------|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|
| Model I |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
|---------|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|

Model I jest modelem istotnym, ponieważ wartość  $R^2 > 0,75$ . Dodatkowo wartości istotności statystyki F oraz p-value dla współczynnika zmiennej X15 wynoszą poniżej 0,05.

[illegible]

Wartości  $R^2$  oraz istotności F wskazują, że model II jest istotny. Jednak wartości p-value dla współczynników zmiennych X12, X13 i X17 zdecydowanie przekraczają 0,05. Model nie jest więc brany pod uwagę w dalszej analizie.

[illegible]

W przypadku modelu III wartości wszystkich analizowanych parametrów są zgodne z założeniami. Dodatkowo wartości  $R^2$  dla tego modelu są większe niż dla modelu I.

| Model IV  |        |       |            |                       |    |             |             |             |                |  |  |  |  |
|-----------|--------|-------|------------|-----------------------|----|-------------|-------------|-------------|----------------|--|--|--|--|
| x5        | x13    | x17   | y          | SUMMARY OUTPUT        |    |             |             |             |                |  |  |  |  |
| 3 890 766 | 69,22  | 94,0  | 36 936 592 |                       |    |             |             |             |                |  |  |  |  |
| 3 875 671 | 70,13  | 96,4  | 41 030 024 | Regression Statistics |    |             |             |             |                |  |  |  |  |
| 3 869 484 | 73,48  | 95,5  | 45 681 501 | Multiple R            |    |             |             |             |                |  |  |  |  |
| 3 881 164 | 78,06  | 98,0  | 51 823 896 | R Square              |    |             |             |             |                |  |  |  |  |
| 3 912 079 | 81,71  | 97,3  | 56 718 898 | Adjusted R Square     |    |             |             |             |                |  |  |  |  |
| 3 932 445 | 76,93  | 98,3  | 59 177 816 | Standard Error        |    |             |             |             |                |  |  |  |  |
| 3 953 178 | 82,53  | 100,7 | 61 710 894 | Observations          |    |             |             |             |                |  |  |  |  |
| 3 847 634 | 69,47  | 59,5  | 19 490 969 |                       |    |             |             |             |                |  |  |  |  |
| 3 847 374 | 79,55  | 65,8  | 27 464 839 | ANOVA                 |    |             |             |             |                |  |  |  |  |
| 3 828 264 | 89,62  | 80,9  | 41 416 969 |                       |    |             |             |             |                |  |  |  |  |
| 3 784 453 | 100,82 | 86,1  | 49 741 313 |                       |    |             |             |             |                |  |  |  |  |
|           |        |       |            |                       | df | SS          | MS          | F           | Significance F |  |  |  |  |
|           |        |       |            | Regression            | 3  | 1,67655E+15 | 5,5885E+14  | 64,07922467 | 1,86658E-05    |  |  |  |  |
|           |        |       |            | Residual              | 7  | 6,10487E+13 | 8,72124E+12 |             |                |  |  |  |  |
|           |        |       |            | Total                 | 10 | 1,7376E+15  |             |             |                |  |  |  |  |
|           |        |       |            |                       |    |             |             |             |                |  |  |  |  |
|           |        |       |            |                       |    |             |             |             |                |  |  |  |  |
|           |        |       |            |                       |    |             |             |             |                |  |  |  |  |
|           |        |       |            |                       |    |             |             |             |                |  |  |  |  |
|           |        |       |            |                       |    |             |             |             |                |  |  |  |  |
|           |        |       |            |                       |    |             |             |             |                |  |  |  |  |
|           |        |       |            |                       |    |             |             |             |                |  |  |  |  |
|           |        |       |            |                       |    |             |             |             |                |  |  |  |  |
|           |        |       |            |                       |    |             |             |             |                |  |  |  |  |
|           |        |       |            |                       |    |             |             |             |                |  |  |  |  |
|           |        |       |            |                       |    |             |             |             |                |  |  |  |  |
|           |        |       |            |                       |    |             |             |             |                |  |  |  |  |
|           |        |       |            |                       |    |             |             |             |                |  |  |  |  |
|           |        |       |            |                       |    |             |             |             |                |  |  |  |  |
|           |        |       |            |                       |    |             |             |             |                |  |  |  |  |
|           |        |       |            |                       |    |             |             |             |                |  |  |  |  |
|           |        |       |            |                       |    |             |             |             |                |  |  |  |  |
|           |        |       |            |                       |    |             |             |             |                |  |  |  |  |
|           |        |       |            |                       |    |             |             |             |                |  |  |  |  |
|           |        |       |            |                       |    |             |             |             |                |  |  |  |  |
|           |        |       |            |                       |    |             |             |             |                |  |  |  |  |
|           |        |       |            |                       |    |             |             |             |                |  |  |  |  |
|           |        |       |            |                       |    |             |             |             |                |  |  |  |  |
|           |        |       |            |                       |    |             |             |             |                |  |  |  |  |
|           |        |       |            |                       |    |             |             |             |                |  |  |  |  |
|           |        |       |            |                       |    |             |             |             |                |  |  |  |  |
|           |        |       |            |                       |    |             |             |             |                |  |  |  |  |
|           |        |       |            |                       |    |             |             |             |                |  |  |  |  |
|           |        |       |            |                       |    |             |             |             |                |  |  |  |  |
|           |        |       |            |                       |    |             |             |             |                |  |  |  |  |
|           |        |       |            |                       |    |             |             |             |                |  |  |  |  |
|           |        |       |            |                       |    |             |             |             |                |  |  |  |  |
|           |        |       |            |                       |    |             |             |             |                |  |  |  |  |
|           |        |       |            |                       |    |             |             |             |                |  |  |  |  |
|           |        |       |            |                       |    |             |             |             |                |  |  |  |  |
|           |        |       |            |                       |    |             |             |             |                |  |  |  |  |
|           |        |       |            |                       |    |             |             |             |                |  |  |  |  |
|           |        |       |            |                       |    |             |             |             |                |  |  |  |  |
|           |        |       |            |                       |    |             |             |             |                |  |  |  |  |
|           |        |       |            |                       |    |             |             |             |                |  |  |  |  |
|           |        |       |            |                       |    |             |             |             |                |  |  |  |  |
|           |        |       |            |                       |    |             |             |             |                |  |  |  |  |
|           |        |       |            |                       |    |             |             |             |                |  |  |  |  |
|           |        |       |            |                       |    |             |             |             |                |  |  |  |  |
|           |        |       |            |                       |    |             |             |             |                |  |  |  |  |
|           |        |       |            |                       |    |             |             |             |                |  |  |  |  |
|           |        |       |            |                       |    |             |             |             |                |  |  |  |  |
|           |        |       |            |                       |    |             |             |             |                |  |  |  |  |
|           |        |       |            |                       |    |             |             |             |                |  |  |  |  |
|           |        |       |            |                       |    |             |             |             |                |  |  |  |  |
|           |        |       |            |                       |    |             |             |             |                |  |  |  |  |
|           |        |       |            |                       |    |             |             |             |                |  |  |  |  |
|           |        |       |            |                       |    |             |             |             |                |  |  |  |  |
|           |        |       |            |                       |    |             |             |             |                |  |  |  |  |
|           |        |       |            |                       |    |             |             |             |                |  |  |  |  |
|           |        |       |            |                       |    |             |             |             |                |  |  |  |  |
|           |        |       |            |                       |    |             |             |             |                |  |  |  |  |
|           |        |       |            |                       |    |             |             |             |                |  |  |  |  |
|           |        |       |            |                       |    |             |             |             |                |  |  |  |  |
|           |        |       |            |                       |    |             |             |             |                |  |  |  |  |
|           |        |       |            |                       |    |             |             |             |                |  |  |  |  |
|           |        |       |            |                       |    |             |             |             |                |  |  |  |  |
|           |        |       |            |                       |    |             |             |             |                |  |  |  |  |
|           |        |       |            |                       |    |             |             |             |                |  |  |  |  |
|           |        |       |            |                       |    |             |             |             |                |  |  |  |  |
|           |        |       |            |                       |    |             |             |             |                |  |  |  |  |
|           |        |       |            |                       |    |             |             |             |                |  |  |  |  |
|           |        |       |            |                       |    |             |             |             |                |  |  |  |  |
|           |        |       |            |                       |    |             |             |             |                |  |  |  |  |
|           |        |       |            |                       |    |             |             |             |                |  |  |  |  |
|           |        |       |            |                       |    |             |             |             |                |  |  |  |  |
|           |        |       |            |                       |    |             |             |             |                |  |  |  |  |
|           |        |       |            |                       |    |             |             |             |                |  |  |  |  |
|           |        |       |            |                       |    |             |             |             |                |  |  |  |  |
|           |        |       |            |                       |    |             |             |             |                |  |  |  |  |
|           |        |       |            |                       |    |             |             |             |                |  |  |  |  |
|           |        |       |            |                       |    |             |             |             |                |  |  |  |  |
|           |        |       |            |                       |    |             |             |             |                |  |  |  |  |

Dla modelu IV także wszystkie badane parametry mają odpowiednie wartości. Ponadto zaobserwowano, że wartości  $R^2$  również są większe niż dla modelu I, ale mniejsze niż dla modelu III.

Bazując na analizie parametrów regresji liniowej modelem najlepiej opisującym zmienną objaśniającą jest **model III**.

```
> AIC(model_macierz)
[1] 375.3436
> AIC(model_hellwig)
[1] 368.0652
> AIC(model3)
[1] 366.247
> AIC(model4)
[1] 366.9383
```

Analizując wartości kryterium informacyjnego Akaikego wybór modelu opierany jest na najmniejszym wyniku funkcji AIC. W powyższym przypadku najmniejszą wartość otrzymujemy dla modelu III.

Podsumowując, zarówno dzięki analizie parametrów regresji liniowej, jak i wartości kryterium informacyjnego Akaikego, wyciągnięto wniosek, że modelem najlepiej opisującym zmienną Y jest **model III**. Na podstawie wyników porównania zdecydowano, że poniższy model zostanie wykorzystany do dalszej analizy badanego zjawiska.

$$y = a_1X_5 + a_2X_{13} + a_3X_{12} + b,$$

gdzie:

Y – liczba widzów w kinach w Polsce,

X5 - liczba ludności w wieku przedprodukcyjnym w miastach w Polsce,

X12 - liczba imprez masowych artystyczno-rozrywkowych w Polsce,

X13 - przeciętne miesięczne wydatki na 1 osobę na rekreację i kulturę w Polsce [zł].

## 4 Estymacja parametrów modelu

W celu obliczenia współczynników  $a_1, a_2, a_3$  i  $b$  zastosowano metodę najmniejszych kwadratów. Skorzystano ze wzoru  $\beta = X^T X^{-1} X^T y$  i wykonano potrzebne operacje na macierzach. Metoda najmniejszych kwadratów została zastosowana w celu minimalizacji sumy kwadratów różnic między rzeczywistymi wartościami a wartościami teoretycznymi. Wynikające z tego współczynniki określają wpływ każdej zmiennej objaśniającej na zmienną zależną. W wyniku otrzymano poniższe wartości współczynników ( $b, a_1, a_2, a_3$ ):

|            |              |
|------------|--------------|
| <b>B =</b> | -503125788,3 |
|            | 122,9739833  |
|            | 4661,185244  |
|            | 570538,0638  |

Wartość współczynnika  $B_1 = 122,97$  oznacza, że wzrost zmiennej  $x_5$  o jednostkę powoduje wzrost zmiennej  $y$  (liczba widzów) o 122,97 tys. osób, przy założeniu stałości pozostałych zmiennych. Podobnie współczynniki  $B_2$  i  $B_3$  można interpretować jako wpływ zmiennych  $x_{12}$  i  $x_{13}$ .

Następnie podstawiono obliczone wartości do wzoru modelu i obliczono wartości teoretyczne zmiennej objaśnianej  $\hat{y}_t$ . W kolejnym kroku obliczono reszty ze wzoru  $\xi_t = y_t - \hat{y}_t$  oraz podniesiono je do kwadratu.

| y          | x5        | x12   | x13    | y^          | e          | e^2                |
|------------|-----------|-------|--------|-------------|------------|--------------------|
| 36 936 592 | 3 890 766 | 5 730 | 69,22  | 41538441,2  | -4 601 849 | 21 177 016 060 443 |
| 41 030 024 | 3 875 671 | 6 070 | 70,13  | 41786141,54 | -756 118   | 571 713 738 075    |
| 45 681 501 | 3 869 484 | 6 280 | 73,48  | 43915452,92 | 1 766 048  | 3 118 925 812 453  |
| 51 823 896 | 3 881 164 | 6 480 | 78,06  | 48897090,43 | 2 926 806  | 8 566 190 853 041  |
| 56 718 898 | 3 912 079 | 6 876 | 81,71  | 56627124,41 | 91 774     | 8 422 391 329      |
| 59 177 816 | 3 932 445 | 7 074 | 76,93  | 57327355,29 | 1 850 461  | 3 424 204 836 395  |
| 61 710 894 | 3 953 178 | 6 949 | 82,53  | 62489339,89 | -778 446   | 605 978 002 014    |
| 19 490 969 | 3 847 634 | 1 564 | 69,47  | 16958464,14 | 2 532 505  | 6 413 580 858 048  |
| 27 464 839 | 3 847 374 | 2 964 | 79,55  | 29203173,93 | -1 738 335 | 3 021 808 328 212  |
| 41 416 969 | 3 828 264 | 5 247 | 89,62  | 43239945,32 | -1 822 976 | 3 323 242 669 461  |
| 49 741 313 | 3 784 453 | 6 313 | 100,82 | 49211181,92 | 530 131    | 281 038 960 227    |
|            |           |       |        | <b>Suma</b> | 0          | 50 512 122 509 700 |

Na tej podstawie wygenerowano wykres porównujący rzeczywiste wartości zmiennej  $Y$  oraz wartości teoretyczne, a co za tym idzie – porównujący opisywane zjawisko ze stworzonym modelem.



Wykres porównawczy pokazuje dopasowanie modelu do rzeczywistych danych. Model w dużej mierze naśladuje przebieg rzeczywistego zjawiska, co sugeruje, że zmienne objaśniające skutecznie opisują zmienną zależną.

## 5 Weryfikacja modelu

|                                                                      |           |       |        |                |              |                |                |
|----------------------------------------------------------------------|-----------|-------|--------|----------------|--------------|----------------|----------------|
| Model: $y = a_1 \cdot x_5 + a_2 \cdot x_{12} + a_3 \cdot x_{13} + b$ |           |       |        | SUMMARY OUTPUT |              |                |                |
| y                                                                    | x5        | x12   | x13    |                |              |                |                |
| 36 936 592                                                           | 3 890 766 | 5 730 | 69,22  |                |              |                |                |
| 41 030 024                                                           | 3 875 671 | 6 070 | 70,13  |                |              |                |                |
| 45 681 501                                                           | 3 869 484 | 6 280 | 73,48  |                |              |                |                |
| 51 823 896                                                           | 3 881 164 | 6 480 | 78,06  |                |              |                |                |
| 56 718 898                                                           | 3 912 079 | 6 876 | 81,71  |                |              |                |                |
| 59 177 816                                                           | 3 932 445 | 7 074 | 76,93  |                |              |                |                |
| 61 710 894                                                           | 3 953 178 | 6 949 | 82,53  |                |              |                |                |
| 19 490 969                                                           | 3 847 634 | 1 564 | 69,47  |                |              |                |                |
| 27 464 839                                                           | 3 847 374 | 2 964 | 79,55  |                |              |                |                |
| 41 416 969                                                           | 3 828 264 | 5 247 | 89,62  |                |              |                |                |
| 49 741 313                                                           | 3 784 453 | 6 313 | 100,82 |                |              |                |                |
|                                                                      |           |       |        | ANOVA          |              |                |                |
|                                                                      |           |       |        | df             | SS           | MS             | F              |
|                                                                      |           |       |        |                |              |                | Significance F |
|                                                                      |           |       |        | Regression     | 3            | 1,68709E+15    | 5,62363E+14    |
|                                                                      |           |       |        | Residual       | 7            | 5,05121E+13    | 7,21602E+12    |
|                                                                      |           |       |        | Total          | 10           | 1,7376E+15     |                |
|                                                                      |           |       |        |                |              |                |                |
|                                                                      |           |       |        |                | Coefficients | Standard Error | t Stat         |
|                                                                      |           |       |        |                |              |                | P-value        |
|                                                                      |           |       |        |                |              |                | Lower 95%      |
|                                                                      |           |       |        |                |              |                | Upper 95%      |
|                                                                      |           |       |        |                |              |                | Lower 95%      |
|                                                                      |           |       |        |                |              |                | Upper 95%      |
|                                                                      |           |       |        | Intercept      | -503125788,3 | 107438438,5    | -4,682921636   |
|                                                                      |           |       |        | x5             | 122,9739833  | 26,60793153    | 4,62170399     |
|                                                                      |           |       |        | x12            | 4661,185228  | 667,1615411    | 6,986591614    |
|                                                                      |           |       |        | x13            | 570538,0651  | 124090,1559    | 4,597770555    |

### 5.1 Dopasowanie modelu

- $R^2 = 0.9703$

Model wyjaśnia 97,03% wariancji liczby widzów w kinach, co oznacza bardzo dobre dopasowanie. Wskazuje to, że zmienne X5 (liczba ludności przedprodukcyjnej w miastach), X12 (liczba imprez masowych) i X13 (wydatki na kulturę) mają silny związek z liczbą widzów.

- $\text{Adjusted } R^2 = 0.9585$

Skorygowany współczynnik determinacji jest również wysoki, co oznacza, że model dobrze radzi sobie nawet po uwzględnieniu liczby zmiennych objaśniających.

### 5.2 Istotność modelu jako całości

- $F = 77,92$  (bardzo wysoka wartość statystyki F)
  - $p\text{-value}$  dla modelu =  $9,64E-06$  (bliskie zero)
- Oznacza to, że model jest istotny statystycznie – co najmniej jedna zmienna objaśniająca ma wpływ na liczbę widzów w kinach.

### 5.3 Istotność poszczególnych zmiennych

- Intercept:  $p = 0,0023$
- $x_5$ :  $p = 0,0090$  (Istotny wpływ – im więcej dzieci i młodzieży w miastach, tym więcej widzów w kinach)
- $x_{12}$ :  $p = 0,0004$  (Bardzo istotny wpływ – więcej wydarzeń kulturalnych sprzyja frekwencji w kinach)
- $x_{13}$ :  $p = 0,0025$  (Bardzo istotny wpływ – wzrost wydatków na kulturę przekłada się na większą liczbę widzów)

Wszystkie zmienne mają  $p < 0,01$ , co sugeruje, że każda z nich ma istotny wpływ na zmienną y.

### 5.4 Współczynniki i interpretacja

- X5 (liczba ludności przedprodukcyjnej):**  
Każdy dodatkowy mieszkaniec w wieku przedprodukcyjnym w miastach powoduje wzrost liczby widzów o ok. 123 osoby rocznie.
- X12 (liczba imprez masowych):**  
Każda dodatkowa impreza artystyczna/rozrywkowa zwiększa liczbę widzów w kinach o ok. 4,7 tys. osób.
- X13 (wydatki na kulturę):**  
Wzrost średnich miesięcznych wydatków na kulturę o 1 zł na osobę powoduje wzrost liczby widzów w kinach o 570 tys. rocznie.

## 5.5 Błąd standardowy modelu

Błąd standardowy = 2 686 264,6 widzów

Oznacza to, że przeciętne odchylenie prognozowanej liczby widzów od rzeczywistej wynosi ok. 2,7 mln osób rocznie.

W porównaniu do całkowitej liczby widzów (ok. 40-50 mln rocznie), błąd wynosi 5-7%, co jest stosunkowo akceptowalnym poziomem w kontekście prognozowania.

## 5.6 Współliniowość zmiennych

Współczynnik VIF (Variance Inflation Factor) służy do wykrywania współliniowości pomiędzy zmiennymi objaśniającymi w modelu regresji. Jeśli zmienne objaśniające są silnie skorelowane, mogą powodować niestabilność modelu, co utrudnia interpretację wyników i wpływa na jakość prognoz.

## Interpretacja VIF

- **VIF  $< 5$**   $\rightarrow$  Brak problemów współliniowości
- **VIF  $5 - 10$**   $\rightarrow$  Umiarkowana współliniowość, warto sprawdzić korelacje
- **VIF  $> 10$**   $\rightarrow$  Silna współliniowość, należy usunąć lub połączyć zmienne.

Aby sprawdzić współliniowość, dla każdej zmiennej objaśniającej przeprowadzono regresję, w której była ona zmienną zależną, a pozostałe zmienne były objaśniające. Następnie obliczono  $R^2$  i współczynnik VIF według wzoru:

$$VIF = \frac{1}{1 - R^2}$$

|                      |          |        |                       |              |                |              |                |              |              |  |
|----------------------|----------|--------|-----------------------|--------------|----------------|--------------|----------------|--------------|--------------|--|
| x5 - zmienna zależna |          |        | SUMMARY OUTPUT        |              |                |              |                |              |              |  |
| x5                   | x12      | x13    |                       |              |                |              |                |              |              |  |
| 3 890 766            | 5 730    | 69,22  | Regression Statistics |              |                |              |                |              |              |  |
| 3 875 671            | 6 070    | 70,13  | Multiple R            |              |                |              |                |              |              |  |
| 3 869 484            | 6 280    | 73,48  | R Square              |              |                |              |                |              |              |  |
| 3 881 164            | 6 480    | 78,06  | Adjusted R Square     |              |                |              |                |              |              |  |
| 3 912 079            | 6 876    | 81,71  | Standard Error        |              |                |              |                |              |              |  |
| 3 932 445            | 7 074    | 76,93  | Observations          |              |                |              |                |              |              |  |
| 3 953 178            | 6 949    | 82,53  |                       |              |                |              |                |              |              |  |
| 3 847 634            | 1 564    | 69,47  | ANOVA                 |              |                |              |                |              |              |  |
| 3 847 374            | 2 964    | 79,55  |                       |              |                |              |                |              |              |  |
| 3 828 264            | 5 247    | 89,62  | df                    | SS           | MS             | F            | Significance F |              |              |  |
| 3 784 453            | 6 313    | 100,82 | Regression            | 2            | 12806188145    | 6403094073   | 5,02579211     | 0,038574354  |              |  |
|                      |          |        | Residual              | 8            | 10192373950    | 1274046744   |                |              |              |  |
|                      |          |        | Total                 | 10           | 22998562096    |              |                |              |              |  |
| VIF =                | 2,256448 |        |                       |              |                |              |                |              |              |  |
| Brak współliniowości |          |        |                       |              |                |              |                |              |              |  |
|                      |          |        |                       | Coefficients | Standard Error | t Stat       | P-value        | Lower 95%    | Upper 95%    |  |
|                      |          |        | Intercept             | 4028860,707  | 95127,0766     | 42,35240744  | 1,06472E-10    | 3809497,275  | 4248224,139  |  |
|                      |          |        | x12                   | 16,59101127  | 6,646738052    | 2,496113302  | 0,037166454    | 1,263605847  | 31,91841669  |  |
|                      |          |        | x13                   | -3116,489289 | 1226,639336    | -2,540672875 | 0,034674437    | -5945,124668 | -287,8539097 |  |
|                      |          |        |                       |              |                |              |                | -5945,124668 | -287,8539097 |  |

Dla zmiennej X5 (Ludność przedprodukcyjna) uzyskano  $R^2 = 0.5568$ , co odpowiada  $VIF = 2.2564$ . Wynik ten wskazuje, że nie występuje problem współliniowości.

|                       |           |        |                       |              |                |              |                |              |              |
|-----------------------|-----------|--------|-----------------------|--------------|----------------|--------------|----------------|--------------|--------------|
| x12 - zmienna zależna |           |        | SUMMARY OUTPUT        |              |                |              |                |              |              |
| x12                   | x5        | x13    |                       |              |                |              |                |              |              |
| 5 730                 | 3 890 766 | 69,22  | Regression Statistics |              |                |              |                |              |              |
| 6 070                 | 3 875 671 | 70,13  | Multiple R            |              |                |              |                |              |              |
| 6 280                 | 3 869 484 | 73,48  | 0,690257456           |              |                |              |                |              |              |
| 6 480                 | 3 881 164 | 78,06  | R Square              |              |                |              |                |              |              |
| 6 876                 | 3 912 079 | 81,71  | 0,476455356           |              |                |              |                |              |              |
| 7 074                 | 3 932 445 | 76,93  | Adjusted R Square     |              |                |              |                |              |              |
| 6 949                 | 3 953 178 | 82,53  | 0,345569195           |              |                |              |                |              |              |
| 1 564                 | 3 847 634 | 69,47  | Standard Error        |              |                |              |                |              |              |
| 2 964                 | 3 847 374 | 79,55  | 1423,550218           |              |                |              |                |              |              |
| 5 247                 | 3 828 264 | 89,62  | Observations          |              |                |              |                |              |              |
| 6 313                 | 3 784 453 | 100,82 | 11                    |              |                |              |                |              |              |
|                       |           |        | ANOVA                 |              |                |              |                |              |              |
|                       |           |        | df                    | SS           | MS             | F            | Significance F |              |              |
|                       |           |        | Regression            | 2            | 14753805,84    | 7376902,921  | 3,640227143    | 0,075130259  |              |
|                       |           |        | Residual              | 8            | 16211961,79    | 2026495,224  |                |              |              |
|                       |           |        | Total                 | 10           | 30965767,64    |              |                |              |              |
|                       |           |        |                       |              |                |              |                |              |              |
|                       |           |        |                       | Coefficients | Standard Error | t Stat       | P-value        | Lower 95%    | Upper 95%    |
|                       |           |        | Intercept             | -105329,218  | 43068,30487    | -2,445631846 | 0,040209738    | -204644,907  | -6013,528912 |
|                       |           |        | x5                    | 0,026389617  | 0,010572283    | 2,496113302  | 0,037166454    | 0,002009888  | 0,050769345  |
|                       |           |        | x13                   | 109,4371296  | 53,17243228    | 2,058155419  | 0,073570484    | -13,17871899 | 232,0529782  |
|                       |           |        |                       |              |                |              |                |              |              |
| VIF = 1,9100568       |           |        |                       |              |                |              |                |              |              |
| Brak współliniowości  |           |        |                       |              |                |              |                |              |              |

Dla zmiennej X12 (Liczba imprez masowych) uzyskano  $R^2 = 0.4765$ , co odpowiada  $VIF = 1.9101$ . Oznacza to brak współliniowości w modelu.

| x13 - zmienna zależna |           |       | SUMMARY OUTPUT        |    |              |                |              |                |              |              |              |              |
|-----------------------|-----------|-------|-----------------------|----|--------------|----------------|--------------|----------------|--------------|--------------|--------------|--------------|
| x13                   | x5        | x12   |                       |    |              |                |              |                |              |              |              |              |
| 69,22                 | 3 890 766 | 5 730 | Regression Statistics |    |              |                |              |                |              |              |              |              |
| 70,13                 | 3 875 671 | 6 070 | Multiple R            |    |              |                |              |                |              |              |              |              |
| 73,48                 | 3 869 484 | 6 280 | R Square              |    |              |                |              |                |              |              |              |              |
| 78,06                 | 3 881 164 | 6 480 | Adjusted R Square     |    |              |                |              |                |              |              |              |              |
| 81,71                 | 3 912 079 | 6 876 | Standard Error        |    |              |                |              |                |              |              |              |              |
| 76,93                 | 3 932 445 | 7 074 | Observations          |    |              |                |              |                |              |              |              |              |
| 82,53                 | 3 953 178 | 6 949 |                       |    |              |                |              |                |              |              |              |              |
| 69,47                 | 3 847 634 | 1 564 | ANOVA                 |    |              |                |              |                |              |              |              |              |
| 79,55                 | 3 847 374 | 2 964 |                       | df | SS           | MS             | F            | Significance F |              |              |              |              |
| 89,62                 | 3 828 264 | 5 247 | Regression            | 2  | 440,5898275  | 220,2949137    | 3,760724676  | 0,070571746    |              |              |              |              |
| 100,82                | 3 784 453 | 6 313 | Residual              | 8  | 468,6222635  | 58,57778293    |              |                |              |              |              |              |
|                       |           |       | Total                 | 10 | 909,2120909  |                |              |                |              |              |              |              |
| VIF = 1,9401812       |           |       |                       |    |              |                |              |                |              |              |              |              |
| Brak współliniowości  |           |       |                       |    |              |                |              |                |              |              |              |              |
|                       |           |       |                       |    | Coefficients | Standard Error | t Stat       | P-value        | Lower 95%    | Upper 95%    | Lower 95%    | Upper 95%    |
|                       |           |       | Intercept             |    | 616,7422966  | 214,8415308    | 2,870684705  | 0,020807035    | 121,3168384  | 1112,167755  | 121,3168384  | 1112,167755  |
|                       |           |       | x5                    |    | -0,000143289 | 5,63981E-05    | -2,540672875 | 0,034674437    | -0,000273343 | -1,32349E-05 | -0,000273343 | -1,32349E-05 |
|                       |           |       | x12                   |    | 0,003163385  | 0,001537       | 2,058155419  | 0,073570484    | -0,000380943 | 0,006707713  | -0,000380943 | 0,006707713  |

Dla zmiennej X13 (Przeciętne wydatki na kulturę) uzyskano  $R^2 = 0.4848$ , co odpowiada  $VIF = 1.9418$ . Również w tym przypadku współliniowość nie stanowi problemu.

Wszystkie wartości VIF są mniejsze niż 5, co oznacza, że zmienne nie są nadmiernie skorelowane i nie powodują problemów współliniowości w modelu. Model regresji można uznać za stabilny i dobrze skonstruowany pod względem zależności między zmiennymi objaśnianymi.

## 5.7 Wyrazistość modelu

Współczynnik zmienności jest jedną z metod oceny wyrazistości modelu, czyli tego, jak dobrze model wyjaśnia zmienność zmiennej zależnej. V pozwala na określenie, czy model dostatecznie precyzyjnie opisuje zależności w danych.

Wzór na współczynnik zmienności losowej (V):

$$V = \left( \frac{S_e}{\bar{Y}} \right) \times 100\%$$

gdzie:

- $S_e$  – błąd standardowy reszt.
- $\bar{Y}$  – średnia wartość zmiennej zależnej  $Y$ .

### Interpretacja wyniku

- Jeśli  $V < 30\%$  → model jest dobrze dopasowany, ma niską zmienność losową.
- Jeśli  $V \geq 30\%$  → model nie jest wystarczająco precyzyjny, może wymagać poprawek.

| Wyrazistość modelu:        |            |
|----------------------------|------------|
| $V = S_e / y_{sr} * 100\%$ |            |
| Se                         | 2686264,6  |
| y_sr                       | 44 653 974 |
| V =                        | 6%         |

Wartość  $V = 6\%$  oznacza, że model jest bardzo dobrze dopasowany i jego zmienność losowa jest niewielka.



- Jeśli  $p \geq 0.05$ , brak podstaw do odrzucenia hipotezy zerowej – rozkład reszt może być normalny.
- Jeśli  $p < 0.05$ , odrzucamy hipotezę normalności – reszty nie są normalnie rozłożone.



```

5 # Dopasowanie modelu regresji
6 model <- lm(y ~ x5 + x12 + x13, data = dane)
7
8 # Pobranie reszt modelu
9 reszty <- residuals(model)
10
11 # Test Shapiro-Wilka
12 shapiro_result <- shapiro.test(reszty)
13
14 # wyniki testu
15 cat("Test Shapiro-Wilka\n")
16 cat("Statystyka W:", shapiro_result$statistic, "\n")
17 cat("wartość p:", shapiro_result$p.value, "\n")
18
19 # wniosek
20 if (shapiro_result$p.value < 0.05) {
21   cat("Odrzucamy hipotezę zerową – reszty nie mają rozkładu normalnego na poziomie istotności 0.05.\n")
22 } else {
23   cat("Brak podstaw do odrzucenia hipotezy zerowej – reszty mogą pochodzić z rozkładu normalnego.\n")
24 }

```

```

Test Shapiro-Wilka
> cat("Statystyka W:", shapiro_result$statistic, "\n")
Statystyka W: 0.9495303
> cat("wartość p:", shapiro_result$p.value, "\n")
wartość p: 0.6379563
> # wniosek
> if (shapiro_result$p.value < 0.05) {
+   cat("Odrzucamy hipotezę zerową – reszty nie mają rozkładu normalnego na poziomie istotności 0.05.\n")
+ } else {
+   cat("Brak podstaw do odrzucenia hipotezy zerowej – reszty mogą pochodzić z rozkładu normalnego.\n")
+ }
Brak podstaw do odrzucenia hipotezy zerowej – reszty mogą pochodzić z rozkładu normalnego.

```

### Wyniki testu Shapiro-Wilka

- Statystyka W: 0.9495
- Wartość p: 0.6379

### Interpretacja:

Ponieważ wartość  $p = 0.6379$  jest większa niż przyjęty poziom istotności 0.05, brak podstaw do odrzucenia hipotezy zerowej. Oznacza to, że reszty mogą pochodzić z rozkładu normalnego, co sugeruje, że model spełnia to kluczowe założenie.

### 6.1.2 Test Kołmogorowa-Smirnowa

Test Kołmogorowa-Smirnowa bada zgodność rozkładu reszt z dowolnym teoretycznym rozkładem, w tym przypadku z rozkładem normalnym. Jest mniej czuły dla małych prób, ale dobrze sprawdza się przy większych zbiorach danych.

### Hipotezy testowe:

- $H_0$ : Rozkład reszt nie różni się istotnie od rozkładu normalnego.
- $H_1$ : Rozkład reszt różni się istotnie od rozkładu normalnego.

### Interpretacja:

- Jeśli  $p \geq 0.05$ , brak podstaw do odrzucenia hipotezy zerowej – reszty mogą pochodzić z rozkładu normalnego.
- Jeśli  $p < 0.05$ , odrzucamy hipotezę normalności – reszty różnią się od normalnych.

```

27 # Test Kołmogorowa-Smirnowa
28 ks_result <- ks.test(reszty, "pnorm", mean=mean(reszty), sd=sd(reszty))
29
30 # wyniki testu
31 cat("Test Kołmogorowa-Smirnowa\n")
32 cat("Statystyka D:", ks_result$statistic, "\n")
33 cat("wartość p:", ks_result$p.value, "\n")
34
35 # wniosek
36 if (ks_result$p.value < 0.05) {
37   cat("Odrzucamy hipotezę zerową – reszty nie są zgodne z rozkładem normalnym na poziomie istotności 0.05.\n")
38 } else {
39   cat("Brak podstaw do odrzucenia hipotezy zerowej – reszty mogą pochodzić z rozkładu normalnego.\n")
40 }

```

```

Test Kołmogorowa-Smirnowa
> cat("Statystyka D:", ks_result$statistic, "\n")
Statystyka D: 0.1476402
> cat("wartość p:", ks_result$p.value, "\n")
wartość p: 0.9423125
> # wniosek
> if (ks_result$p.value < 0.05) {
+   cat("Odrzucamy hipotezę zerową – reszty nie są zgodne z rozkładem normalnym na poziomie istotności 0.05.\n")
+ } else {
+   cat("Brak podstaw do odrzucenia hipotezy zerowej – reszty mogą pochodzić z rozkładu normalnego.\n")
+ }
Brak podstaw do odrzucenia hipotezy zerowej – reszty mogą pochodzić z rozkładu normalnego.

```

## Wyniki testu Kołmogorowa-Smirnowa

- Statystyka D: 0.1476
- Wartość p: 0.9423

### Interpretacja:

Wartość  $p = 0.9423$  jest znacznie większa niż 0.05, co oznacza, że brak podstaw do odrzucenia hipotezy zerowej.

Oznacza to, że rozkład reszt nie różni się istotnie od rozkładu normalnego.

## 6.2 Testy autokorelacji składnika losowego

Autokorelacja reszt w modelu regresji liniowej oznacza, że wartości reszt są ze sobą skorelowane, czyli występuje zależność między kolejnymi obserwacjami. W klasycznej regresji liniowej zakłada się, że reszty powinny być losowe i nieskorelowane – w przeciwnym razie model może być błędnie oszacowany, a testy istotności parametrów mogą dawać niewiarygodne wyniki.

### 6.2.1 Test Durbin-Watsona

Test Durbin-Watsona (DW) służy do wykrywania autokorelacji pierwszego rzędu, czyli sytuacji, w której bieżąca wartość reszt jest zależna od wartości reszt z poprzedniej obserwacji:

$$e_t = \rho e_{t-1} + u_t$$

gdzie:

- $e_t$  – reszta w momencie  $t$ ,
- $e_{t-1}$  – reszta z poprzedniego okresu,
- $u_t$  – losowy składnik zakłócający.

### Hipotezy testowe:

- $H_0$ : Brak autokorelacji pierwszego rzędu ( $\rho = 0$ ).
- $H_1$ : Występuje autokorelacja pierwszego rzędu ( $\rho \neq 0$ ).

### Interpretacja statystyki Durbin-Watsona

Statystyka Durbin-Watsona (DW) przyjmuje wartości w zakresie od 0 do 4. Interpretacja wyniku zależy od jego wartości:

- Jeśli statystyka DW jest bliska 2, oznacza to, że brak autokorelacji reszt, co jest pożądane w modelu regresji.
- Jeśli statystyka DW jest mniejsza niż 1,5, sugeruje to autokorelację dodatnią, co oznacza, że kolejne wartości reszt są do siebie podobne – mogą występować systematyczne błędy w modelu.
- Jeśli statystyka DW jest większa niż 2,5, wskazuje to na autokorelację ujemną, co oznacza, że wartości reszt mają tendencję do naprzemiennego przyjmowania wartości dodatnich i ujemnych.

```
49 # Test Durbin-Watsona
50 dw_result <- durbinwatsonTest(model)
51
52 # wyniki testu
53 cat("Test Durbin-Watsona\n")
54 cat("Statystyka DW:", dw_result$dw, "\n")
55 cat("wartość p:", dw_result$p, "\n")
56
57 # Interpretacja wyników
58 if (dw_result$p < 0.05) {
59   cat("Odrzucamy hipotezę zerową - występuje autokorelacja reszt.\n")
60 } else {
61   cat("Brak podstaw do odrzucenia hipotezy zerowej - brak autokorelacji reszt.\n")
62 }
```

Test Durbin-Watsona  
> cat("Statystyka DW:", dw\_result\$dw, "\n")  
Statystyka DW: 1.490466  
> cat("wartość p:", dw\_result\$p, "\n")  
wartość p: 0.054  
> # Interpretacja wyników  
> if (dw\_result\$p < 0.05) {  
+ cat("Odrzucamy hipotezę zerową - występuje autokorelacja reszt.\n")  
+ } else {  
+ cat("Brak podstaw do odrzucenia hipotezy zerowej - brak autokorelacji reszt.\n")  
+ }  
Brak podstaw do odrzucenia hipotezy zerowej - brak autokorelacji reszt.

#### Wyniki testu Durbin-Watsona:

- Statystyka DW: 1.4905
- Wartość p: 0.054

#### Interpretacja:

Wartość DW zbliżona do 1,5 może sugerować lekką autokorelację dodatnią, jednak wartość  $p = 0.054$  jest większa niż 0.05, co oznacza, że brak podstaw do odrzucenia hipotezy zerowej. Ostatecznie nie wykryto istotnej statystycznie autokorelacji reszt pierwszego rzędu.

#### 6.2.2 Test Breuscha-Godfrey'a

Test Breuscha-Godfrey'a (BG) jest bardziej ogólny niż test Durbin-Watsona, ponieważ pozwala wykryć autokorelację wyższych rzędów (czyli nie tylko pierwszego rzędu, ale także drugiego, trzeciego itd.). Działa również poprawnie w modelach zawierających zmienne opóźnione.

#### Hipotezy testowe:

- $H_0$ : Brak autokorelacji wyższych rzędów.
- $H_1$ : Występuje autokorelacja reszt wyższego rzędu.

### Interpretacja wyników:

Test Breuscha-Godfrey'a zwraca wartość p, którą interpretujemy następująco:

- $p \geq 0.05$ : Brak podstaw do odrzucenia hipotezy zerowej - brak autokorelacji reszt
- $p < 0.05$ : Odrzucamy hipotezę zerową - występuje autokorelacja reszt

```
69 # Test Breuscha-Godfrey'a
70 bg_result <- bgtest(model, order = 2)
71
72 # wyniki testu
73 cat("Test Breuscha-Godfrey'a\n")
74 cat("Statystyka LM:", bg_result$statistic, "\n")
75 cat("wartość p:", bg_result$p.value, "\n")
76
77 # Interpretacja wyników
78 if (bg_result$p.value < 0.05) {
79   cat("Odrzucamy hipotezę zerową - występuje autokorelacja wyższych rzędów.\n")
80 } else {
81   cat("Brak podstaw do odrzucenia hipotezy zerowej - brak autokorelacji wyższych rzędów.\n")
82 }
```

```
Test Breuscha-Godfrey'a
> cat("Statystyka LM:", bg_result$statistic, "\n")
Statystyka LM: 0.1201798
> cat("wartość p:", bg_result$p.value, "\n")
wartość p: 0.9416799
> # Interpretacja wyników
> if (bg_result$p.value < 0.05) {
+   cat("Odrzucamy hipotezę zerową - występuje autokorelacja wyższych rzędów.\n")
+ } else {
+   cat("Brak podstaw do odrzucenia hipotezy zerowej - brak autokorelacji wyższych rzędów.\n")
+ }
Brak podstaw do odrzucenia hipotezy zerowej - brak autokorelacji wyższych rzędów.
```

### Wyniki testu Breuscha-Godfrey'a:

- Statystyka LM: 0.1208
- Wartość p: 0.9417

### Interpretacja:

Wartość  $p = 0.9417$  jest znacznie większa niż 0.05, co oznacza, że brak podstaw do odrzucenia hipotezy zerowej. Oznacza to, że w modelu nie wykryto autokorelacji wyższych rzędów.

## 6.3 Testy heteroskedastyczności

Heteroskedastyczność oznacza, że wariancja reszt modelu regresji nie jest stała – innymi słowy, zmienność błędów nie pozostaje na podobnym poziomie dla różnych wartości zmiennych objaśniających. Jest to naruszenie jednego z kluczowych założeń klasycznej regresji liniowej, które zakłada homoskedastyczność, czyli stałość wariancji reszt.

### 6.3.1 Test Breuscha-Pagana

Test Breuscha-Pagana bada, czy wariancja reszt jest zależna od wartości zmiennych objaśniających. Jeśli wariancja rośnie lub maleje w miarę wzrostu wartości zmiennych objaśniających, oznacza to, że występuje heteroskedastyczność liniowa.

### Hipotezy testowe:

- $H_0$ : Reszty są homoskedastyczne (wariancja reszt jest stała).
- $H_1$ : Reszty są heteroskedastyczne (wariancja reszt zależy od wartości zmiennych objaśniających).

### Interpretacja wyników:

- Jeśli  $p \geq 0.05$ , brak podstaw do odrzucenia hipotezy zerowej - brak heteroskedastyczności (model spełnia założenie stałej wariancji reszt).
- Jeśli  $p < 0.05$ , odrzucamy hipotezę zerową - występuje heteroskedastyczność (wariancja reszt zmienia się w zależności od wartości zmiennych objaśniających).

```
85 # Testy heteroskedastyczności
86
87 # Instalacja pakietu
88 install.packages("lmtest")
89 library(lmtest)
90
91 # Test Breuscha-Pagana
92 bp_result <- bptest(model)
93
94 # wyniki testu
95 cat("Test Breuscha-Pagana (heteroskedastyczność)\n")
96 cat("Statystyka testowa:", bp_result$statistic, "\n")
97 cat("wartość p:", bp_result$p.value, "\n")
98
99 # Interpretacja wyników
100 if (bp_result$p.value < 0.05) {
101   cat("Odrzucamy hipotezę zerową - występuje heteroskedastyczność reszt.\n")
102 } else {
103   cat("Brak podstaw do odrzucenia hipotezy zerowej - model spełnia założenie homoskedastyczności.\n")
104 }
```

Test Breuscha-Pagana (heteroskedastyczność)

```
> cat("Statystyka testowa:", bp_result$statistic, "\n")
Statystyka testowa: 2.857951
> cat("wartość p:", bp_result$p.value, "\n")
wartość p: 0.4140487
> # Interpretacja wyników
> if (bp_result$p.value < 0.05) {
+   cat("Odrzucamy hipotezę zerową - występuje heteroskedastyczność reszt.\n")
+ } else {
+   cat("Brak podstaw do odrzucenia hipotezy zerowej - model spełnia założenie homoskedastyczności.\n")
+ }
```

Brak podstaw do odrzucenia hipotezy zerowej - model spełnia założenie homoskedastyczności.

### Wyniki testu Breuscha-Pagana

- Statystyka testowa: **2.85791**
- Wartość p: **0.4140**

### Interpretacja:

Wartość  $p = 0.4140$  jest większa niż 0.05, co oznacza, że brak podstaw do odrzucenia hipotezy zerowej. Oznacza to, że model spełnia założenie homoskedastyczności, czyli wariancja reszt jest stała i nie zależy od wartości zmiennych objaśniających.

### 6.3.2 Test White'a

Test White'a jest bardziej ogólny niż test Breuscha-Pagana, ponieważ pozwala wykryć nieliniową heteroskedastyczność, a także interakcje między zmiennymi objaśniającymi. Dzięki temu jest bardziej uniwersalnym narzędziem do diagnozy problemów związanych z heteroskedastycznością.

### Hipotezy testowe:

- $H_0$ : Reszty są homoskedastyczne (wariancja reszt jest stała).
- $H_1$ : Reszty są heteroskedastyczne (wariancja reszt zmienia się w sposób bardziej złożony, np. nieliniowy).

### Interpretacja wyników:

- Jeśli  $p \geq 0.05$ , brak podstaw do odrzucenia hipotezy zerowej - brak heteroskedastyczności.
- Jeśli  $p < 0.05$ , odrzucamy hipotezę zerową - występuje heteroskedastyczność (nieliniowa zależność wariancji reszt od zmiennych objaśniających).

```

112 # Test white'a
113 white_result <- bptest(model, varformula = ~ fitted(model) + I(fitted(model)^2))
114
115 # wyniki testu
116 cat("Test white'a (heteroskedastyczność)\n")
117 cat("Statystyka testowa:", white_result$statistic, "\n")
118 cat("wartość p:", white_result$p.value, "\n")
119
120 # Interpretacja wyników
121 if (white_result$p.value < 0.05) {
122   cat(" Odrzucamy hipotezę zerową - występuje heteroskedastyczność reszt.\n")
123 } else {
124   cat("Brak podstaw do odrzucenia hipotezy zerowej - model spełnia założenie homoskedastyczności.\n")
125 }

```

```

Test white'a (heteroskedastyczność)
> cat("Statystyka testowa:", white_result$statistic, "\n")
Statystyka testowa: 1.364037
> cat("wartość p:", white_result$p.value, "\n")
wartość p: 0.5055955
> # Interpretacja wyników
> if (white_result$p.value < 0.05) {
+   cat(" Odrzucamy hipotezę zerową - występuje heteroskedastyczność reszt.\n")
+ } else {
+   cat("Brak podstaw do odrzucenia hipotezy zerowej - model spełnia założenie homoskedastyczności.\n")
+ }
Brak podstaw do odrzucenia hipotezy zerowej - model spełnia założenie homoskedastyczności.

```

### Wyniki testu White'a

- Statystyka testowa: **1.3640**
- Wartość p: **0.5059**

### Interpretacja:

Podobnie jak w teście Breuscha-Pagana, wartość  $p = 0.5059$  jest znacznie większa niż 0.05, co oznacza, że brak podstaw do odrzucenia hipotezy zerowej. Oznacza to, że w modelu nie występuje heteroskedastyczność, nawet w bardziej złożonej (nieliniowej) formie.

## 6.4 Test serii

Test serii bada, czy znaki kolejnych reszt są ułożone w losowy sposób. Jeśli reszty mają tendencję do występowania w sekwencjach (np. dłuższe serie wartości dodatnich lub ujemnych), oznacza to, że mogą one być systematycznie skorelowane, co może wskazywać na błędy modelu.

### Hipotezy testowe

- $H_0$ : Reszty są losowe – brak systematycznych wzorców.
- $H_1$ : Reszty nie są losowe – występują zależności między wartościami reszt.

Jeśli test odrzuci hipotezę zerową, oznacza to, że reszty modelu wykazują pewne regularności, które mogą świadczyć o problemach w modelu.

```

134 # Test serii|
135
136 # Podział reszt na wartości dodatnie i ujemne
137 znaki <- ifelse(reszty > 0, 1, 0)
138
139 # Test serii
140 runs_result <- runs.test(as.factor(znaki))
141
142 # Wyniki testu
143 cat("Test Serii (badanie losowości reszt)\n")
144 cat("Statystyka testowa:", runs_result$statistic, "\n")
145 cat("Wartość p:", runs_result$p.value, "\n")
146
147 # Interpretacja wyników
148 if (runs_result$p.value < 0.05) {
149   cat("Odrzucamy hipotezę zerową – reszty NIE są losowe, możliwe błędy modelu.\n")
150 } else {
151   cat("Brak podstaw do odrzucenia hipotezy zerowej – reszty są losowe.\n")
152 }

```

```

Test Serii (badanie losowości reszt)
> cat("Statystyka testowa:", runs_result$statistic, "\n")
Statystyka testowa: -0.2916059
> cat("Wartość p:", runs_result$p.value, "\n")
Wartość p: 0.7705879
> # Interpretacja wyników
> if (runs_result$p.value < 0.05) {
+   cat("Odrzucamy hipotezę zerową – reszty NIE są losowe, możliwe błędy modelu.\n")
+ } else {
+   cat("Brak podstaw do odrzucenia hipotezy zerowej – reszty są losowe.\n")
+ }
Brak podstaw do odrzucenia hipotezy zerowej – reszty są losowe.

```

### Wyniki testu serii

- Statystyka testowa: **-0.2916**
- Wartość p: **0.7706**

### Interpretacja:

Wartość  $p = 0.7706$  jest znacznie większa niż 0.05, co oznacza, że brak podstaw do odrzucenia hipotezy zerowej. Oznacza to, że reszty są losowe, co jest zgodne z założeniami poprawnie skonstruowanego modelu.

## 6.5 Test symetrii

Test ten bada, czy rozkład reszt jest równomierny po obu stronach mediany. W przypadku regresji liniowej, reszty powinny być równie często dodatnie, jak i ujemne oraz rozkładać się w sposób symetryczny.

### Hipotezy testowe

- $H_0$ : Rozkład reszt jest symetryczny względem mediany.
- $H_1$ : Rozkład reszt nie jest symetryczny.

Jeśli test odrzuci hipotezę zerową, oznacza to, że rozkład reszt jest asymetryczny, co może wskazywać na problemy w modelu.



```

157 # Test symetrii
158 symmetry_result <- symmetry.test(reszty, boot = FALSE)
159
160 # wyniki testu
161 cat("Test symetrii (badanie rozkładu reszt)\n")
162 cat("Statystyka testowa:", symmetry_result$statistic, "\n")
163 cat("wartość p:", symmetry_result$p.value, "\n")
164
165 # Interpretacja wyników
166 if (symmetry_result$p.value < 0.05) {
167   cat("Odrzucamy hipotezę zerową – reszty NIE są symetryczne, model może być błędny.\n")
168 } else {
169   cat("Brak podstaw do odrzucenia hipotezy zerowej – reszty są symetryczne.\n")
170 }

```

```

Test symetrii (badanie rozkładu reszt)
> cat("Statystyka testowa:", symmetry_result$statistic, "\n")
Statystyka testowa: -0.183175
> cat("wartość p:", symmetry_result$p.value, "\n")
wartość p: 0.8546607
> # Interpretacja wyników
> if (symmetry_result$p.value < 0.05) {
+   cat("Odrzucamy hipotezę zerową – reszty NIE są symetryczne, model może być błędny.\n")
+ } else {
+   cat("Brak podstaw do odrzucenia hipotezy zerowej – reszty są symetryczne.\n")
+ }
Brak podstaw do odrzucenia hipotezy zerowej – reszty są symetryczne.

```

### Wyniki testu symetrii

- Statystyka testowa: -0.1832
- Wartość p: 0.8547

### Interpretacja:

Wartość  $p = 0.8547$  jest znacznie większa niż 0.05, co oznacza, że brak podstaw do odrzucenia hipotezy zerowej. Oznacza to, że rozkład reszt modelu jest symetryczny względem mediany, co jest zgodne z założeniami regresji liniowej.



## 7 Podsumowanie i wnioski

Celem przeprowadzonej analizy było stworzenie modelu ekonometrycznego opisującego zależność między liczbą widzów w kinach w Polsce a wybranymi czynnikami społeczno-ekonomicznymi. W ramach projektu zastosowano regresję liniową oraz przeprowadzono szereg testów statystycznych w celu oceny poprawności modelu.

Wyniki analizy wykazały, że opracowany model bardzo dobrze wyjaśnia zmienność liczby widzów w kinach, co potwierdza wysoka wartość współczynnika determinacji  $R^2$ . Kluczowe zmienne objaśniające obejmowały liczbę ludności w wieku przedprodukcyjnym, liczbę imprez masowych oraz przeciętne miesięczne wydatki na rekreację i kulturę.

Przeprowadzone testy statystyczne wykazały, że model spełnia podstawowe założenia klasycznej regresji liniowej. Nie stwierdzono problemów związanych z autokorelacją reszt, heteroskedastycznością ani współliniowością zmiennych objaśniających. Dodatkowo, testy normalności, losowości oraz symetrii reszt potwierdziły poprawność estymacji parametrów modelu. Analiza współczynnika zmienności wykazała, że model charakteryzuje się wysoką precyzją prognozowania.

Podsumowując, opracowany model można uznać za stabilny i poprawnie skonstruowany.