МИНИСТЕРСТВО НАУКИ И ВЫСШЕГО ОБРАЗОВАНИЯ РОССИЙСКОЙ ФЕДЕРАЦИИ ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ АВТОНОМНОЕ ОБРАЗОВАТЕЛЬНОЕ УЧРЕЖДЕНИЕ ВЫСШЕГО ОБРАЗОВАНИЯ

«Национальный исследовательский ядерный университет «МИФИ» (НИЯУ МИФИ)

Институт Финансовых Технологий и Экономической Безопасности Кафедра Финансового мониторинга

**Лабораторная работа №1 по курсу**

**«Эконометрика»**

|  |  |
| --- | --- |
| **Выполнил студент группы С19-702:** | Буданцев А.А. |
| **Проверил:** | Домашова Д.В. |

# 1. Постановка задачи

По показателям экономического развития регионов Российской Федерации провести регрессионный анализ, в качестве результативного признака выступает ожидаемая продолжительность жизни мужчины. В качестве объясняющих признаков:

1) Х1 – Число больничных коек на 1000 человек населения

2) Х2 – Численность врачей всех специальностей всего человек на 1000 человек населения

3) Х3 – Среднедушевые денежные доходы населения (в месяц; рублей)

4) Х4 – Численность рабочей силы, тыс. человек на 10000 человек населения

5) Х5 – Зарегистрировано преступлений по ч. 4, ст. 111 УК РФ - умышленное причинение тяжкого вреда здоровью, повлекшее по неосторожности смерть потерпевшего на 100000 человек населения

6) Х6 – Предварительно расследовано преступлений, совершенных в состоянии алкогольного опьянения на 1000 человек

7) Х7 – Общие коэффициенты смертности (число умерших на 1000 человек населения)

8) Х8 – Выбросы загрязняющих веществ в атмосферный воздух, отходящих от стационарных источников (тысяч тонн) на 100000 человек населения

9) Х9 – Расходы на охрану окружающей среды (млн руб) на 1000 человек

Для этого:

1. Оценить функцию регрессии.

2. Исследовать уравнение регрессии на значимость.

3. Для значимой модели регрессии исследовать значимость коэффициентов.

4. Построить доверительные интервалы для значимых параметров связи.

5. Провести экономический анализ результатов.

Исходные данные приведены в приложении 1.

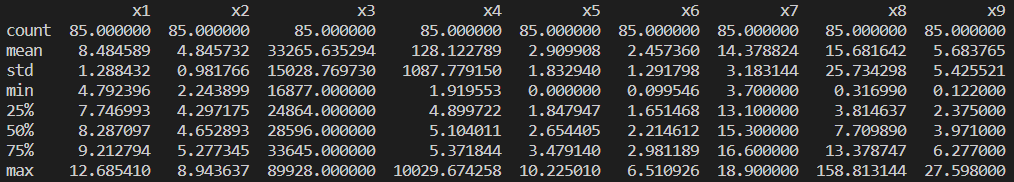
# 2. Оценка коэффициентов классической линейной модели множественной регрессии

Найдем оценки основных характеристик объясняющих переменных. Результаты расчетов в пакете STATISTICA представлены в таблице 1.

Таблица 1. Оценки основных характеристик объясняющих переменных

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Variable | Descriptive Statistics (данные) | | | | |
| |  | | --- | | Valid N | | |  | | --- | | Mean | | |  | | --- | | Minimum | | |  | | --- | | Maximum | | |  | | --- | | Std.Dev. | |
| |  | | --- | | X1 | | 85 | 8,48 | 4,79 | 12,69 | 1,29 |
| |  | | --- | | X2 | | 85 | 4,85 | 2,24 | 8,94 | 0,98 |
| |  | | --- | | X3 | | 85 | 33265,64 | 16877,00 | 89928,00 | 15028,77 |
| |  | | --- | | X4 | | 85 | 128,12 | 1,92 | 10029,67 | 1087,78 |
| |  | | --- | | X5 | | 85 | 2,91 | 0,00 | 10,23 | 1,83 |
| |  | | --- | | X6 | | 85 | 2,46 | 0,10 | 6,51 | 1,29 |
| |  | | --- | | X7 | | 85 | 14,38 | 3,70 | 18,90 | 3,18 |
| |  | | --- | | X8 | | 85 | 15,68 | 0,32 | 158,81 | 25,73 |
| |  | | --- | | X9 | | 85 | 5,68 | 0,12 | 27,60 | 5,43 |

Результаты расчётов в Python:

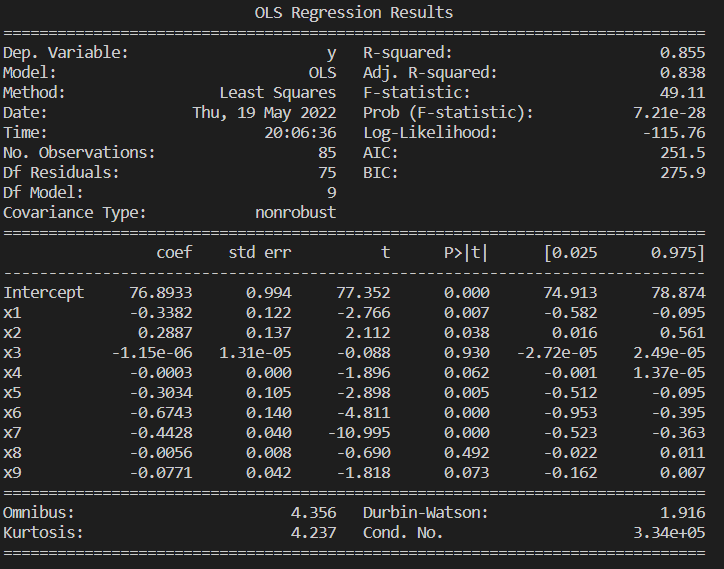


Найдем оценки коэффициентов уравнения регрессии, используя пакет STATISTICA, результаты представлены в таблице 2.

Таблица 2. Оценка коэффициентов уравнения регрессии

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| N=85 | Regression Summary for Dependent Variable: Y (данные) R= ,92463062 R?= ,85494178 Adjusted R?= ,83753479 F(9,75)=49,115 p<0,0000 Std.Error of estimate: 1,0055 | | | | | |
| |  | | --- | | b\* | | |  | | --- | | Std.Err. of b\* | | |  | | --- | | b | | |  | | --- | | Std.Err. of b | | |  | | --- | | t(75) | | |  | | --- | | p-value | |
| |  | | --- | | Intercept | |  |  | 76,89332 | 0,994074 | 77,3517 | 0,000000 |
| |  | | --- | | X1 | | -0,174686 | 0,063160 | -0,33822 | 0,122289 | -2,7658 | 0,007145 |
| |  | | --- | | X2 | | 0,113637 | 0,053814 | 0,28875 | 0,136739 | 2,1117 | 0,038046 |
| |  | | --- | | X3 | | -0,006926 | 0,078854 | -0,00000 | 0,000013 | -0,0878 | 0,930245 |
| |  | | --- | | X4 | | -0,117578 | 0,062026 | -0,00027 | 0,000142 | -1,8956 | 0,061865 |
| |  | | --- | | X5 | | -0,222908 | 0,076919 | -0,30338 | 0,104687 | -2,8980 | 0,004922 |
| |  | | --- | | X6 | | -0,349155 | 0,072577 | -0,67427 | 0,140156 | -4,8108 | 0,000008 |
| |  | | --- | | X7 | | -0,565017 | 0,051390 | -0,44281 | 0,040274 | -10,9947 | 0,000000 |
| |  | | --- | | X8 | | -0,058204 | 0,084354 | -0,00564 | 0,008177 | -0,6900 | 0,492324 |
| |  | | --- | | X9 | | -0,167732 | 0,092245 | -0,07712 | 0,042414 | -1,8183 | 0,073006 |

Результаты расчётов в Python:



Таким образом, оценка уравнения регрессии выглядит следующим образом:

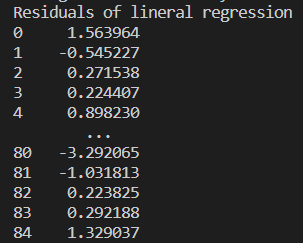
Далее необходимо найти модельное значение результативного признака, вычислить оценку вектора регрессионных остатков.



Таблица 3. Оценка вектора регрессионных остатков

|  |  |  |  |
| --- | --- | --- | --- |
|  | Predicted & Residual Values Y | | |
| |  | | --- | | Observed Value | | |  | | --- | | Predicted Value | | |  | | --- | | Residual | |
| |  | | --- | | Алтайский край | | 65,180000 | 63,616035 | 1,563965 |
| |  | | --- | | Амурская область | | 62,270000 | 62,815228 | -0,545227 |
| |  | | --- | | Архангельская область | | 65,709999 | 65,438461 | 0,271538 |
| |  | | --- | | Астраханская область | | 67,510002 | 67,285591 | 0,224411 |
| |  | | --- | | Белгородская область | | 67,489998 | 66,591774 | 0,898224 |
| |  | | --- | | Брянская область | | 65,070000 | 65,165115 | -0,095116 |
| |  | | --- | | Владимирская область | | 64,589996 | 64,361794 | 0,228203 |
| |  | | --- | | Волгоградская область | | 67,099998 | 65,481667 | 1,618332 |
| |  | | --- | | Вологодская область | | 64,889999 | 64,994804 | -0,104805 |
| |  | | --- | | Воронежская область | | 66,410004 | 66,355728 | 0,054276 |
| |  | | --- | | г. Москва | | 72,389999 | 70,397530 | 1,992470 |
| |  | | --- | | г. Санкт-Петербург | | 69,300003 | 69,773964 | -0,473961 |
| |  | | --- | | г. Севастополь | | 68,699997 | 67,448784 | 1,251213 |
| |  | | --- | | Еврейская автономная область | | 62,410000 | 62,430084 | -0,020084 |
| |  | | --- | | Забайкальский край | | 62,869999 | 63,217815 | -0,347816 |
| |  | | --- | | Ивановская область | | 65,139999 | 65,108589 | 0,031410 |
| |  | | --- | | Иркутская область | | 62,630001 | 63,681732 | -1,051731 |
| |  | | --- | | Кабардино-Балкарская | | 70,160004 | 70,568993 | -0,408989 |
| |  | | --- | | Калининградская область | | 68,389999 | 67,872040 | 0,517960 |
| |  | | --- | | Калужская область | | 65,019997 | 64,953384 | 0,066612 |
| |  | | --- | | Камчатский край | | 63,980000 | 65,004524 | -1,024525 |
| |  | | --- | | Карачаево-Черкесская | | 70,510002 | 69,067467 | 1,442535 |
| |  | | --- | | Кемеровская область | | 63,220001 | 62,841228 | 0,378773 |
| |  | | --- | | Кировская область | | 65,989998 | 64,546936 | 1,443062 |
| |  | | --- | | Костромская область | | 65,730003 | 64,808167 | 0,921837 |
| |  | | --- | | Краснодарский край | | 67,330002 | 67,887505 | -0,557503 |
| |  | | --- | | Красноярский край | | 64,459999 | 64,341194 | 0,118805 |
| |  | | --- | | Курганская область | | 64,480003 | 63,305485 | 1,174519 |
| |  | | --- | | Курская область | | 65,110001 | 65,277435 | -0,167435 |
| |  | | --- | | Ленинградская область | | 66,320000 | 67,049622 | -0,729622 |
| |  | | --- | | Липецкая область | | 65,160004 | 65,040077 | 0,119926 |
| |  | | --- | | Магаданская область | | 63,560001 | 64,342995 | -0,782993 |
| |  | | --- | | Московская область | | 66,739998 | 66,942688 | -0,202690 |
| |  | | --- | | Мурманская область | | 64,330002 | 65,046631 | -0,716629 |
| |  | | --- | | Ненецкий автономный | | 63,439999 | 63,454891 | -0,014893 |
| |  | | --- | | Нижегородская область | | 64,680000 | 64,864342 | -0,184341 |
| |  | | --- | | Новгородская область | | 63,770000 | 64,735733 | -0,965733 |
| |  | | --- | | Новосибирская область | | 65,160004 | 66,237038 | -1,077034 |
| |  | | --- | | Омская область | | 65,220001 | 66,000587 | -0,780586 |
| |  | | --- | | Оренбургская область | | 64,760002 | 65,204834 | -0,444832 |
| |  | | --- | | Орловская область | | 64,720001 | 66,146187 | -1,426186 |
| |  | | --- | | Пензенская область | | 66,099998 | 65,539131 | 0,560867 |
| |  | | --- | | Пермский край | | 64,089996 | 65,533279 | -1,443283 |
| |  | | --- | | Приморский край | | 64,589996 | 64,339912 | 0,250084 |
| |  | | --- | | Псковская область | | 63,880001 | 63,922874 | -0,042873 |
| |  | | --- | | Республика Адыгея | | 68,699997 | 68,615776 | 0,084221 |
| |  | | --- | | Республика Алтай | | 63,560001 | 65,037689 | -1,477688 |
| |  | | --- | | Республика Башкортостан | | 65,239998 | 65,690376 | -0,450378 |
| |  | | --- | | Республика Бурятия | | 65,320000 | 64,988640 | 0,331360 |
| |  | | --- | | Республика Дагестан | | 73,330002 | 72,665771 | 0,664230 |
| |  | | --- | | Республика Ингушетия | | 77,430000 | 74,648567 | 2,781433 |
| |  | | --- | | Республика Калмыкия | | 67,919998 | 68,377670 | -0,457672 |
| |  | | --- | | Республика Карелия | | 63,709999 | 64,809319 | -1,099319 |
| |  | | --- | | Республика Коми | | 64,750000 | 63,564899 | 1,185101 |
| |  | | --- | | Республика Крым | | 66,220001 | 66,581635 | -0,361633 |
| |  | | --- | | Республика Марий Эл | | 65,239998 | 66,132065 | -0,892067 |
| |  | | --- | | Республика Мордовия | | 66,800003 | 66,279114 | 0,520889 |
| |  | | --- | | Республика Саха (Якутия) | | 66,059998 | 65,007004 | 1,052994 |
| |  | | --- | | Республика Северная | | 68,839996 | 69,786324 | -0,946327 |
| |  | | --- | | Республика Татарстан | | 67,349998 | 67,319618 | 0,030380 |
| |  | | --- | | Республика Тыва | | 61,950001 | 62,956768 | -1,006767 |
| |  | | --- | | Республика Хакасия | | 64,699997 | 64,142059 | 0,557938 |
| |  | | --- | | Ростовская область | | 67,410004 | 66,938637 | 0,471367 |
| |  | | --- | | Рязанская область | | 65,739998 | 65,987404 | -0,247406 |
| |  | | --- | | Самарская область | | 65,379997 | 66,200821 | -0,820824 |
| |  | | --- | | Саратовская область | | 66,309998 | 65,180016 | 1,129982 |
| |  | | --- | | Сахалинская область | | 64,620003 | 63,210598 | 1,409405 |
| |  | | --- | | Свердловская область | | 64,709999 | 64,950722 | -0,240723 |
| |  | | --- | | Смоленская область | | 64,639999 | 64,703362 | -0,063362 |
| |  | | --- | | Ставропольский край | | 68,510002 | 68,176811 | 0,333191 |
| |  | | --- | | Тамбовская область | | 66,599998 | 64,925713 | 1,674286 |
| |  | | --- | | Тверская область | | 64,290001 | 64,593704 | -0,303703 |
| |  | | --- | | Томская область | | 65,959999 | 66,137199 | -0,177200 |
| |  | | --- | | Тульская область | | 64,900002 | 64,889824 | 0,010178 |
| |  | | --- | | Тюменская область | | 67,540001 | 68,328880 | -0,788879 |
| |  | | --- | | Удмуртская Республика | | 65,300003 | 65,770172 | -0,470169 |
| |  | | --- | | Ульяновская область | | 65,910004 | 65,388802 | 0,521202 |
| |  | | --- | | Хабаровский край | | 63,230000 | 65,354584 | -2,124584 |
| |  | | --- | | Ханты-Мансийский автономный округ | | 68,760002 | 68,974930 | -0,214928 |
| |  | | --- | | Челябинская область | | 64,970001 | 64,625847 | 0,344154 |
| |  | | --- | | Чеченская Республика | | 69,489998 | 72,782066 | -3,292068 |
| |  | | --- | | Чувашская Республика | | 65,459999 | 66,491814 | -1,031815 |
| |  | | --- | | Чукотский автономный округ | | 62,549999 | 62,326176 | 0,223824 |
| |  | | --- | | Ямало-Ненецкий автономный округ | | 67,540001 | 67,247810 | 0,292191 |
| |  | | --- | | Ярославская область | | 65,400002 | 64,070961 | 1,329041 |

Результаты расчётов в Python:



# 3. Анализ вариации результативного признака Y

Рассчитав с помощью STATISTICA (рис.3), получим:

R2 =0,85; R = 0,92; SEE = 0,99

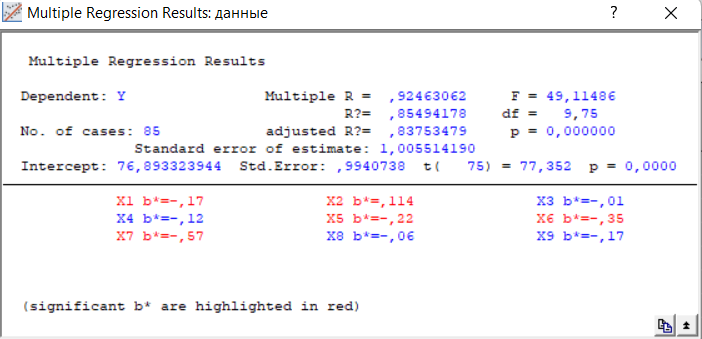


Рисунок 1. Оценка коэффициента детерминации в пакете STATISTICA

# 4. Проверка гипотезы о нормальном характере распределения регрессионных остатков

Дальнейшее изучение свойств оценок КЛММР проводится при дополнительном предположении и нормальном характере распределения регрессионных остатков:



Это предположение необходимо проверить.

Выдвинем гипотезы:

Гипотеза H0: Распределение регрессионных остатков не отличается от нормального.

Гипотеза H1: Распределение регрессионных остатков отличается от нормального.

Для проверки гипотезы воспользуемся критерием Колмогорова-Смирнова. В пакете программ «Statistica» получим следующий результат:

Таблица 4. Проверка гипотезы о нормальном законе распределения регрессионных остатков

|  |  |  |  |
| --- | --- | --- | --- |
| Variable | Tests of Normality (данные) | | |
| |  | | --- | | N | | |  | | --- | | max D | | |  | | --- | | K-S p | |
| |  | | --- | | Residual | | 85 | 0,077491 | p > .20 |

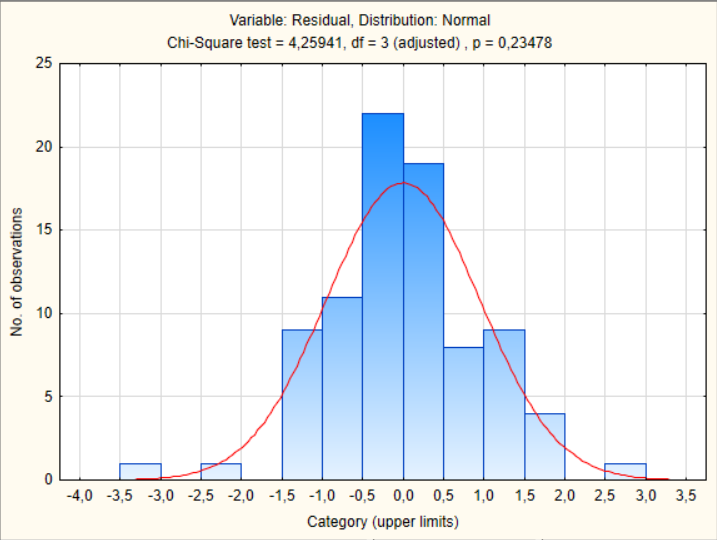
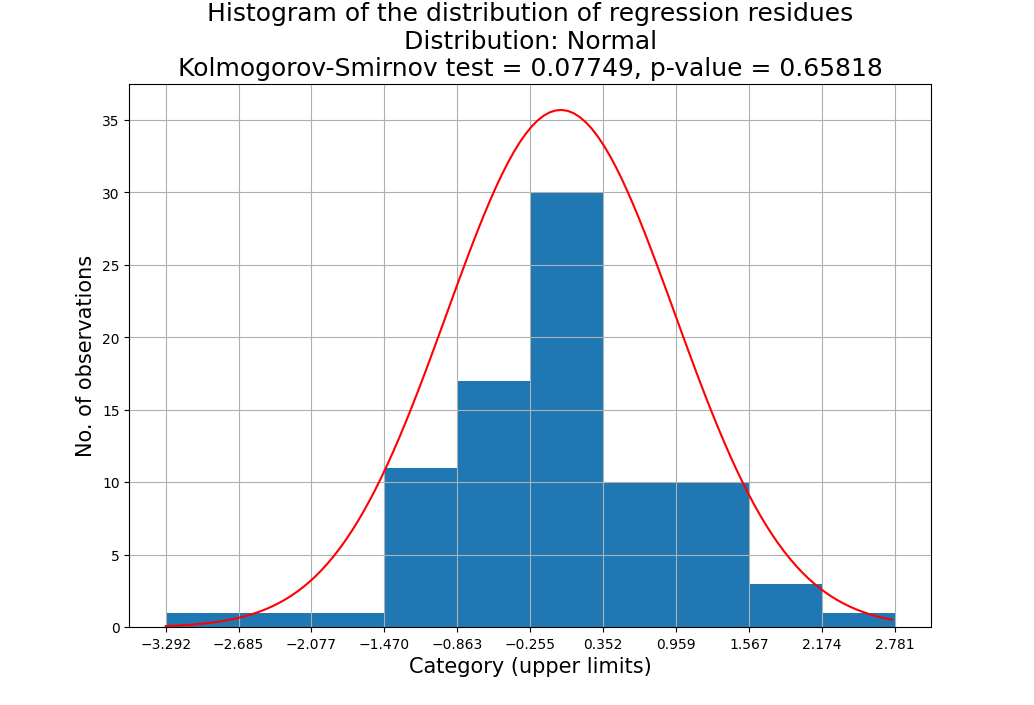
Так как p > 0.05 для теста Колмогорова-Смирнова => распределение регрессионных остатков является нормальным.

Рисунок 2. Гистограмма распределения регрессионных остатков

Результаты расчётов в Python:

# 5. Проверка значимости уравнения регрессии и значимости коэффициентов

Проверим гипотезу о не значимости ЛММР:

Н0: *β1=β2=β3= β4= β5 = β6= β7 = β8 = β9=0*

Альтернативная гипотеза Н1:

Н1: .

По итогам проверки в пакете STATISTICA (см. рис. 1) получили, что F(9,75) = 49,11 уровень значимости р=0,00<0.05. Таким образом, построенная ЛММР значима.

# 6. Проверка гипотез о значимости коэффициента ЛММР

**Проверим гипотезы о значимости коэффициентов ЛММР**

Таблица 5. Проверка значимости коэффициентов уравнения регрессии

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| N=85 | Regression Summary for Dependent Variable: Y (данные) R= ,92463062 R?= ,85494178 Adjusted R?= ,83753479 F(9,75)=49,115 p<0,0000 Std.Error of estimate: 1,0055 | | | | | |
| |  | | --- | | b\* | | |  | | --- | | Std.Err. of b\* | | |  | | --- | | b | | |  | | --- | | Std.Err. of b | | |  | | --- | | t(75) | | |  | | --- | | p-value | |
| |  | | --- | | Intercept | |  |  | 76,89332 | 0,994074 | 77,3517 | 0,000000 |
| |  | | --- | | X1 | | -0,174686 | 0,063160 | -0,33822 | 0,122289 | -2,7658 | 0,007145 |
| |  | | --- | | X2 | | 0,113637 | 0,053814 | 0,28875 | 0,136739 | 2,1117 | 0,038046 |
| |  | | --- | | X3 | | -0,006926 | 0,078854 | -0,00000 | 0,000013 | -0,0878 | 0,930245 |
| |  | | --- | | X4 | | -0,117578 | 0,062026 | -0,00027 | 0,000142 | -1,8956 | 0,061865 |
| |  | | --- | | X5 | | -0,222908 | 0,076919 | -0,30338 | 0,104687 | -2,8980 | 0,004922 |
| |  | | --- | | X6 | | -0,349155 | 0,072577 | -0,67427 | 0,140156 | -4,8108 | 0,000008 |
| |  | | --- | | X7 | | -0,565017 | 0,051390 | -0,44281 | 0,040274 | -10,9947 | 0,000000 |
| |  | | --- | | X8 | | -0,058204 | 0,084354 | -0,00564 | 0,008177 | -0,6900 | 0,492324 |
| |  | | --- | | X9 | | -0,167732 | 0,092245 | -0,07712 | 0,042414 | -1,8183 | 0,073006 |

1) Н0: *β0=0*;

Н1: *β00*.

Нулевая гипотеза не принимается (р=0,00<0,05), коэффициент *β0* значимо отличен от нуля. t(75) = 77,35

2) Н0: *β1=0*;

Н1: *β10*.

Нулевая гипотеза не принимается (р=0,01<0,05), коэффициент *β1* значимо отличен от нуля. t(75) = -2,76

3) Н0: *β2=0*;

Н1: *β20*.

Нулевая гипотеза не принимается (р=0,04<0,05), коэффициент *β2* значимо отличен от нуля. t(75) = 2,11

4) Н0: *β3=0*;

Н1: *β30*.

Нулевая гипотеза принимается (р=0,93>0,05), коэффициент *β3* незначим.

t(75) = -0,08

5) Н0: *β4=0*;

Н1: *β40*.

Нулевая гипотеза принимается (р=0,06>0,05), коэффициент *β4* незначим.

t(75) = -1,89

6) Н0: *β5=0*;

Н1: *β50*.

Нулевая гипотеза не принимается (р=0,00<0,05), коэффициент *β5* значимо отличен от нуля. t(75) = -2,89

7) Н0: *β6=0*;

Н1: *β60*.

Нулевая гипотеза не принимается (р=0,00<0,05), коэффициент *β6* значимо отличен от нуля. t(75) = -4,81

8) Н0: *β7=0*;

Н1: *β70*.

Нулевая гипотеза не принимается (р=0,00<0,05), коэффициент *β7* значимо отличен от нуля. t(75) = -10,99

9) Н0: *β8=0*;

Н1: *β80*.

Нулевая гипотеза принимается (р=0,49>0,05), коэффициент *β8* незначим. t(75) = -0,69

10) Н0: *β9=0*;

Н1: *β90*.

Нулевая гипотеза принимается (р=0,07>0,05), коэффициент *β9* незначим. t(75) = -1,81

# 7. Построение доверительных интервалов для значимых коэффициентов КЛМНР

Для коэффициентов уравнения регрессии значимо отличных от нуля находим доверительные интервалы, используя статистику



имеющую распределение Стьюдента с степенями свободы.



# 

# 8. Выявление мультиколлинеарности.

Перейдем к рассмотрению критериев по выявлению мультиколлинеарности:

1. С помощью внешних признаков
2. С помощью формальных признаков

Проверка внешних признаков:

Таблица 6. Проверка внешних признаков мультиколлениарности

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| N=85 | Regression Summary for Dependent Variable: Y (данные) R= ,92463062 R?= ,85494178 Adjusted R?= ,83753479 F(9,75)=49,115 p<0,0000 Std.Error of estimate: 1,0055 | | | | | |
| |  | | --- | | b\* | | |  | | --- | | Std.Err. of b\* | | |  | | --- | | b | | |  | | --- | | Std.Err. of b | | |  | | --- | | t(75) | | |  | | --- | | p-value | |
| |  | | --- | | Intercept | |  |  | 76,89332 | 0,994074 | 77,3517 | 0,000000 |
| |  | | --- | | X1 | | -0,174686 | 0,063160 | -0,33822 | 0,122289 | -2,7658 | 0,007145 |
| |  | | --- | | X2 | | 0,113637 | 0,053814 | 0,28875 | 0,136739 | 2,1117 | 0,038046 |
| |  | | --- | | X3 | | -0,006926 | 0,078854 | -0,00000 | 0,000013 | -0,0878 | 0,930245 |
| |  | | --- | | X4 | | -0,117578 | 0,062026 | -0,00027 | 0,000142 | -1,8956 | 0,061865 |
| |  | | --- | | X5 | | -0,222908 | 0,076919 | -0,30338 | 0,104687 | -2,8980 | 0,004922 |
| |  | | --- | | X6 | | -0,349155 | 0,072577 | -0,67427 | 0,140156 | -4,8108 | 0,000008 |
| |  | | --- | | X7 | | -0,565017 | 0,051390 | -0,44281 | 0,040274 | -10,9947 | 0,000000 |
| |  | | --- | | X8 | | -0,058204 | 0,084354 | -0,00564 | 0,008177 | -0,6900 | 0,492324 |
| |  | | --- | | X9 | | -0,167732 | 0,092245 | -0,07712 | 0,042414 | -1,8183 | 0,073006 |

Внешние признаки:

Как видно из таблицы 6, модель значима. Однако:

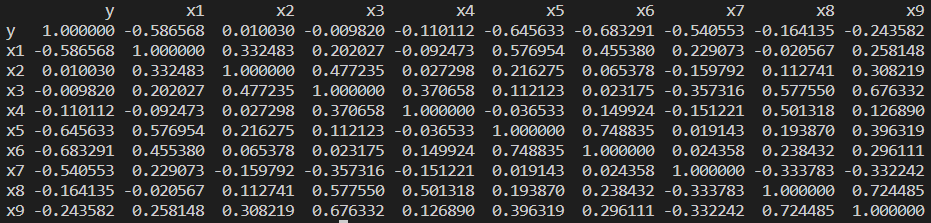
1. В уравнении регрессии присутствуют незначимые коэффициенты при переменных X3, Х4, Х8, Х9;
2. Стандартная ошибка оценки коэффициентов при переменных X3, Х8 превосходит значения самих коэффициентов;

Формальные признаки:

Таблица 7. Проверка формальных признаков мультиколлинеарности

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Variable | Correlations (данные) | | | | | | | | | |
| |  | | --- | | X1 | | |  | | --- | | X2 | | |  | | --- | | X3 | | |  | | --- | | X4 | | |  | | --- | | X5 | | |  | | --- | | X6 | | |  | | --- | | X7 | | |  | | --- | | X8 | | |  | | --- | | X9 | | |  | | --- | | Y | |
| |  | | --- | | X1 | | 1,000000 | 0,332483 | 0,202027 | -0,092473 | 0,576954 | 0,455380 | 0,229073 | -0,020567 | 0,258148 | -0,586568 |
| |  | | --- | | X2 | | 0,332483 | 1,000000 | 0,477235 | 0,027298 | 0,216275 | 0,065378 | -0,159792 | 0,112741 | 0,308219 | 0,010030 |
| |  | | --- | | X3 | | 0,202027 | 0,477235 | 1,000000 | 0,370658 | 0,112123 | 0,023175 | -0,357316 | 0,577550 | 0,676332 | -0,009820 |
| |  | | --- | | X4 | | -0,092473 | 0,027298 | 0,370658 | 1,000000 | -0,036533 | 0,149924 | -0,151221 | 0,501318 | 0,126890 | -0,110112 |
| |  | | --- | | X5 | | 0,576954 | 0,216275 | 0,112123 | -0,036533 | 1,000000 | 0,748835 | 0,019143 | 0,193870 | 0,396319 | -0,645633 |
| |  | | --- | | X6 | | 0,455380 | 0,065378 | 0,023175 | 0,149924 | 0,748835 | 1,000000 | 0,024358 | 0,238432 | 0,296111 | -0,683291 |
| |  | | --- | | X7 | | 0,229073 | -0,159792 | -0,357316 | -0,151221 | 0,019143 | 0,024358 | 1,000000 | -0,333783 | -0,332242 | -0,540553 |
| |  | | --- | | X8 | | -0,020567 | 0,112741 | 0,577550 | 0,501318 | 0,193870 | 0,238432 | -0,333783 | 1,000000 | 0,724485 | -0,164135 |
| |  | | --- | | X9 | | 0,258148 | 0,308219 | 0,676332 | 0,126890 | 0,396319 | 0,296111 | -0,332242 | 0,724485 | 1,000000 | -0,243582 |
| |  | | --- | | Y | | -0,586568 | 0,010030 | -0,009820 | -0,110112 | -0,645633 | -0,683291 | -0,540553 | -0,164135 | -0,243582 | 1,000000 |

Результаты расчётов в Python:



Коэффициент корреляции между X3 и Х9 равен 0,676332

Коэффициент корреляции между X5 и Х6 равен 0,748835

Коэффициент корреляции между X8 и Х9 равен 0,724485

Построим ЛММР для каждой из объясняющих переменных на остальные объясняющие переменные:

Для Х1

Для Х2

Для Х3

Для X4

Для Х5

Для Х6

Для Х7

Для Х8

Для Х9

Анализ оценок коэффициентов детерминации показал наличие заметной связи между объясняющими переменными X3, X5, X6, X8, X9 и всеми остальными признаками соответственно. Таким образом, можно сделать вывод о наличии мультиколлинеарности.

# 9. Устранение мультиколлинеарности

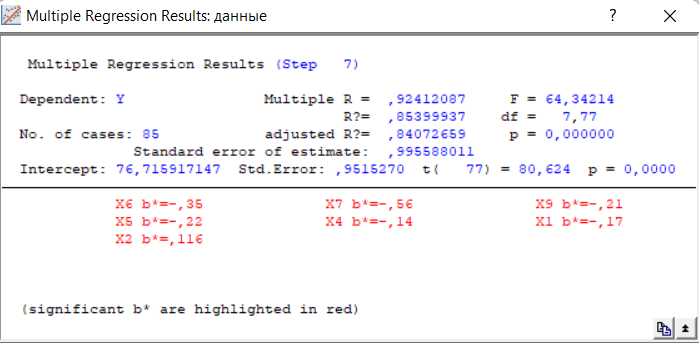
1. С помощью пошаговой регрессии с включением переменных:

Рисунок 3.Результаты пошаговой регрессии с включением переменных

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| N=85 | Regression Summary for Dependent Variable: Y (данные) R= ,92412087 R?= ,85399937 Adjusted R?= ,84072659 F(7,77)=64,342 p<0,0000 SEE = 0,99559 | | | | | |
| |  | | --- | | b\* | | |  | | --- | | Std.Err. of b\* | | |  | | --- | | b | | |  | | --- | | Std.Err. of b | | |  | | --- | | t(77) | | |  | | --- | | p-value | |
| |  | | --- | | Intercept | |  |  | 76,71592 | 0,951527 | 80,6240 | 0,000000 |
| |  | | --- | | X6 | | -0,350616 | 0,069634 | -0,67709 | 0,134472 | -5,0352 | 0,000003 |
| |  | | --- | | X7 | | -0,563736 | 0,050132 | -0,44180 | 0,039289 | -11,2450 | 0,000000 |
| |  | | --- | | X9 | | -0,213886 | 0,052724 | -0,09834 | 0,024243 | -4,0567 | 0,000118 |
| |  | | --- | | X5 | | -0,221876 | 0,075776 | -0,30197 | 0,103132 | -2,9280 | 0,004484 |
| |  | | --- | | X4 | | -0,142299 | 0,046338 | -0,00033 | 0,000106 | -3,0709 | 0,002948 |
| |  | | --- | | X1 | | -0,166249 | 0,059213 | -0,32189 | 0,114647 | -2,8076 | 0,006320 |
| |  | | --- | | X2 | | 0,115942 | 0,049269 | 0,29461 | 0,125190 | 2,3533 | 0,021163 |

Таблица 8. Результаты оценивания параметров линейной модели множественной регрессии

Результаты расчётов в Python:



1. С помощью пошаговой регрессии с исключением переменных:

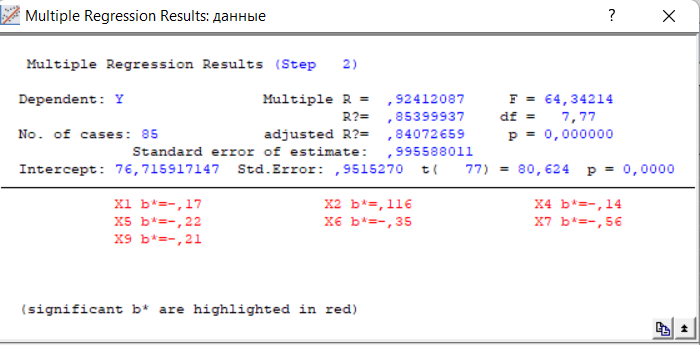
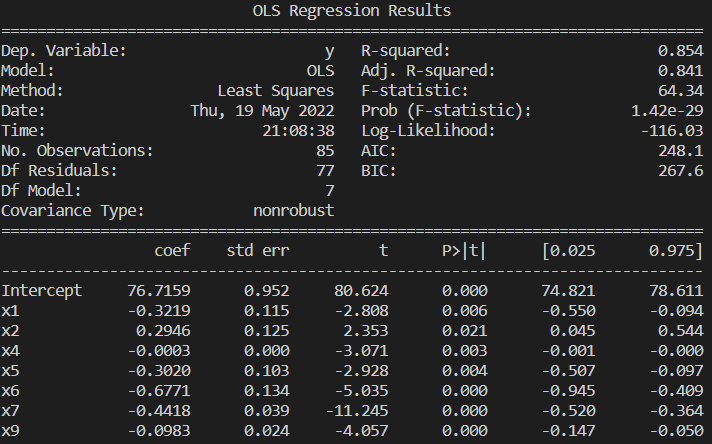


Рисунок 4.Результаты пошаговой регрессии с исключением переменных

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| N=85 | Regression Summary for Dependent Variable: Y (данные) R= ,92412087 R?= ,85399937 Adjusted R?= ,84072659 F(7,77)=64,342 p<0,0000 SEE = 0,99 | | | | | | |
| |  | | --- | | b\* | | |  | | --- | | Std.Err. of b\* | | |  | | --- | | b | | |  | | --- | | Std.Err. of b | | |  | | --- | | t(77) | | | |  | | --- | | p-value | |
| |  | | --- | | Intercept | |  |  | 76,71592 | 0,951527 | 80,6240 | | 0,000000 |
| |  | | --- | | X1 | | -0,166249 | 0,059213 | -0,32189 | 0,114647 | -2,8076 | 0,006320 | |
| |  | | --- | | X2 | | 0,115942 | 0,049269 | 0,29461 | 0,125190 | 2,3533 | 0,021163 | |
| |  | | --- | | X4 | | -0,142299 | 0,046338 | -0,00033 | 0,000106 | -3,0709 | 0,002948 | |
| |  | | --- | | X5 | | -0,221876 | 0,075776 | -0,30197 | 0,103132 | -2,9280 | 0,004484 | |
| |  | | --- | | X6 | | -0,350616 | 0,069634 | -0,67709 | 0,134472 | -5,0352 | 0,000003 | |
| |  | | --- | | X7 | | -0,563736 | 0,050132 | -0,44180 | 0,039289 | -11,2450 | 0,000000 | |
| |  | | --- | | X9 | | -0,213886 | 0,052724 | -0,09834 | 0,024243 | -4,0567 | 0,000118 | |

Таблица 9. Результаты оценивания параметров линейной модели множественной регрессии

Результаты расчётов в Python:

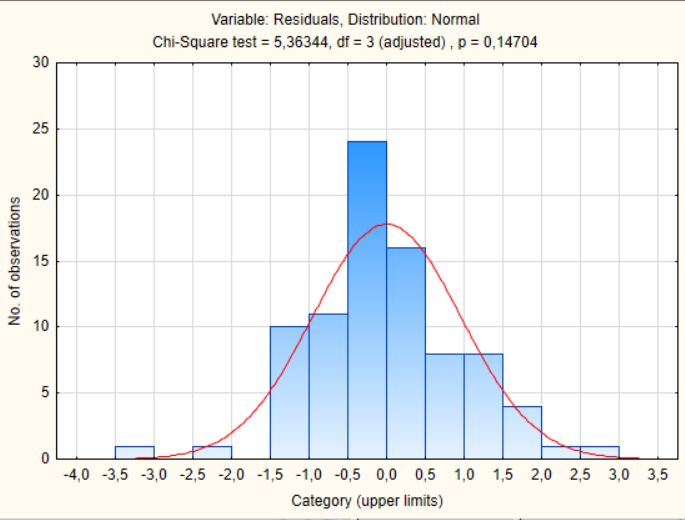


**Выводы**

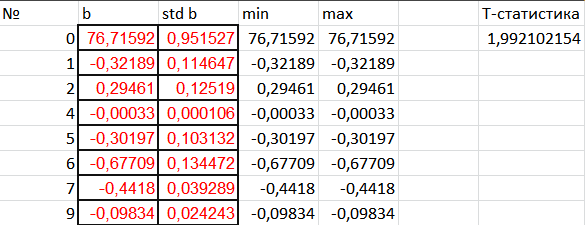
Оценка уравнения регрессии имеет вид:

Доля вариации ожидаемой продолжительности жизни мужчины, обусловленная изменениями объясняющих признаков составляет 85,39%, т.е. на 14,61% вариация заболеваемости зависит от неучтенных факторов.

По итогам проверки в пакете STATISTICA получили, что уровень значимости р<0,05, значение статистики F(7,77)= 64,342. Таким образом, построенная КЛММР значима.

Также были исследованы регрессионные остатки, анализ которых показал нормальность их распределения.

Для коэффициентов уравнения регрессии значимо отличных от нуля были найдены доверительные интервалы:



Данное уравнение регрессии имеет значимые коэффициенты , , , , .

* При увеличении число больничных коек на 1000 человек населения на 1 единицу ожидаемая продолжительность жизни мужчины уменьшается на 0,32;
* При увеличении численности врачей всех специальностей всего человек на 1000 человек населения на 1 единицу ожидаемая продолжительность жизни мужчины увеличится на 0,29;
* При увеличении численности рабочей силы, тыс. человек на 10000 человек населения на 1 единицу ожидаемая продолжительность жизни мужчины уменьшится на 0,0003;
* При увеличении зарегистрированных преступлений по ч. 4, ст. 111 УК РФ - умышленное причинение тяжкого вреда здоровью, повлекшее по неосторожности смерть потерпевшего на 100000 человек населения на 1 единицу ожидаемая продолжительность жизни мужчины уменьшится на 0,3;
* При увеличении предварительно расследованных преступлений, совершенных в состоянии алкогольного опьянения на 1000 человек на 1 единицу ожидаемая продолжительность жизни мужчины уменьшится на 0,68;
* При увеличении общего коэффициента смертности (число умерших на 1000 человек населения) на 1 единицу ожидаемая продолжительность жизни мужчины уменьшится на 0,44;
* При увеличении расходов на охрану окружающей среды (млн руб) на 1000 человек на 1 единицу ожидаемая продолжительность жизни мужчины уменьшится на 0,1;

# Приложение 1. Исходные данные

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Y** | **X1** | **X2** | **X3** | **X4** | **X5** | **X6** | **X7** | **X8** | **X9** |
| 65,18 | 9,213 | 4,394 | 23 864 | 4,889 | 4,897 | 3,736 | 16,4 | 7,703 | 1,184 |
| 62,27 | 9,621 | 5,608 | 35 499 | 5,208 | 6,866 | 3,885 | 16,1 | 17,567 | 6,098 |
| 65,71 | 8,157 | 5,350 | 36 779 | 4,754 | 2,246 | 2,968 | 14,6 | 17,114 | 12,356 |
| 67,51 | 8,742 | 6,268 | 25 199 | 5,094 | 1,617 | 2,350 | 13,6 | 11,270 | 3,494 |
| 67,49 | 6,999 | 4,072 | 32 841 | 5,423 | 2,222 | 1,339 | 15,5 | 10,946 | 7,374 |
| 65,07 | 7,648 | 4,104 | 28 596 | 5,037 | 3,340 | 2,438 | 16,9 | 4,059 | 1,675 |
| 64,59 | 8,571 | 3,703 | 25 922 | 5,344 | 3,479 | 2,021 | 18,2 | 4,508 | 2,179 |
| 67,1 | 8,833 | 4,618 | 24 864 | 5,091 | 2,453 | 2,323 | 15,8 | 7,137 | 4,936 |
| 64,89 | 7,797 | 3,654 | 29 522 | 4,891 | 2,722 | 2,817 | 15,5 | 33,117 | 6,555 |
| 66,41 | 7,977 | 5,277 | 32 078 | 5,167 | 2,144 | 1,651 | 16,5 | 5,094 | 2,932 |
| 72,39 | 7,174 | 7,006 | 77 283 | 5,820 | 0,372 | 0,478 | 11,8 | 0,473 | 4,398 |
| 69,3 | 8,287 | 8,944 | 49 207 | 5,781 | 0,986 | 0,518 | 13,5 | 1,246 | 2,69 |
| 68,7 | 7,747 | 4,297 | 29 957 | 4,352 | 1,698 | 1,765 | 13,9 | 1,302 | 2,102 |
| 62,41 | 11,827 | 3,832 | 28 048 | 5,016 | 5,205 | 3,780 | 15,8 | 10,734 | 4,556 |
| 62,87 | 9,850 | 5,285 | 27 046 | 5,015 | 5,849 | 5,482 | 13,5 | 13,154 | 4,081 |
| 65,14 | 8,307 | 4,593 | 26 277 | 5,378 | 3,073 | 2,166 | 17,6 | 2,510 | 0,899 |
| 62,63 | 9,441 | 4,956 | 27 571 | 4,883 | 5,007 | 3,453 | 14,9 | 27,808 | 10,56 |
| 70,16 | 7,778 | 4,803 | 21 957 | 5,205 | 0,345 | 0,753 | 9,9 | 0,379 | 0,717 |
| 68,39 | 7,716 | 4,620 | 29 518 | 5,258 | 1,656 | 1,835 | 13,1 | 2,026 | 2,049 |
| 65,02 | 8,245 | 4,234 | 32 442 | 5,316 | 3,138 | 2,134 | 17,2 | 2,569 | 4,063 |
| 63,98 | 11,161 | 5,241 | 55 373 | 5,801 | 3,522 | 3,198 | 12,7 | 12,454 | 8,338 |
| 70,51 | 8,374 | 4,240 | 18 949 | 4,525 | 0,862 | 1,627 | 10,6 | 3,363 | 1,596 |
| 63,22 | 9,270 | 4,472 | 25 433 | 4,831 | 5,300 | 3,721 | 16,2 | 61,906 | 6,973 |
| 65,99 | 8,964 | 5,031 | 24 192 | 5,103 | 2,754 | 3,223 | 16,7 | 7,040 | 3,819 |
| 65,73 | 9,753 | 3,732 | 25 780 | 4,966 | 1,289 | 2,675 | 16,6 | 7,283 | 3,203 |
| 67,33 | 7,300 | 4,525 | 36 838 | 4,955 | 1,848 | 1,010 | 14,3 | 7,304 | 2,392 |
| 64,46 | 8,195 | 5,082 | 32 832 | 5,137 | 4,216 | 2,929 | 14,1 | 89,216 | 15,663 |
| 64,48 | 8,893 | 2,941 | 21 860 | 4,538 | 2,361 | 4,159 | 17,3 | 4,896 | 2,518 |
| 65,11 | 8,785 | 5,341 | 29 786 | 5,219 | 2,403 | 2,215 | 17,3 | 4,926 | 3,115 |
| 66,32 | 5,910 | 3,677 | 33 149 | 5,224 | 1,940 | 1,253 | 15 | 11,224 | 9,446 |
| 65,16 | 8,427 | 4,175 | 32 226 | 5,372 | 1,616 | 1,842 | 17,7 | 28,907 | 5,667 |
| 63,56 | 11,161 | 6,311 | 70 864 | 6,311 | 4,363 | 2,981 | 12,5 | 28,939 | 19,223 |
| 66,74 | 8,341 | 4,613 | 47 046 | 5,330 | 2,627 | 1,111 | 14,6 | 2,618 | 4,932 |
| 64,33 | 10,481 | 5,174 | 46 355 | 5,732 | 2,762 | 1,995 | 13,4 | 27,479 | 19,049 |
| 63,44 | 7,396 | 5,103 | 84 147 | 10029,674 | 2,248 | 4,114 | 10,1 | 133,534 | 12,027 |
| 64,68 | 9,036 | 4,898 | 33 645 | 5,527 | 2,738 | 2,010 | 17,3 | 3,887 | 6,212 |
| 63,77 | 8,231 | 4,304 | 26 268 | 5,092 | 1,196 | 2,753 | 18,2 | 10,816 | 3,19 |
| 65,16 | 9,566 | 5,614 | 31 563 | 5,167 | 3,130 | 1,709 | 15,3 | 5,915 | 1,221 |
| 65,22 | 8,246 | 5,188 | 27 354 | 5,324 | 2,927 | 2,096 | 15,5 | 7,844 | 4,669 |
| 64,76 | 7,926 | 4,701 | 24 719 | 4,900 | 2,758 | 2,815 | 16,2 | 21,068 | 3,993 |
| 64,72 | 7,927 | 4,596 | 26 843 | 4,899 | 0,982 | 1,380 | 18,1 | 3,815 | 1,25 |
| 66,1 | 8,168 | 4,564 | 24 118 | 5,186 | 1,886 | 2,210 | 17,4 | 2,561 | 1,275 |
| 64,09 | 7,505 | 5,090 | 30 215 | 4,879 | 3,405 | 2,706 | 15,5 | 10,990 | 6,126 |
| 64,59 | 9,833 | 5,156 | 37 304 | 5,389 | 5,590 | 2,660 | 15,3 | 9,122 | 4,623 |
| 63,88 | 9,560 | 3,283 | 26 436 | 5,140 | 1,633 | 2,382 | 18,9 | 5,976 | 1,933 |
| 68,7 | 6,792 | 3,938 | 23 927 | 4,359 | 1,080 | 1,250 | 13,1 | 1,619 | 1,405 |
| 63,56 | 7,276 | 4,553 | 30 293 | 4,399 | 4,065 | 6,174 | 11,3 | 3,071 | 3,317 |
| 65,24 | 7,992 | 4,495 | 21 677 | 4,782 | 2,799 | 2,674 | 14,9 | 11,030 | 5,957 |
| 65,32 | 8,503 | 4,491 | 30 249 | 439,221 | 4,075 | 4,969 | 11,8 | 9,577 | 3,971 |
| 73,33 | 6,955 | 4,109 | 26 092 | 4,321 | 0,349 | 0,234 | 6,2 | 0,317 | 0,256 |
| 77,43 | 4,792 | 4,403 | 27 661 | 5,071 | 0,000 | 0,315 | 3,7 | 0,344 | 0,122 |
| 67,92 | 7,805 | 5,031 | 16 877 | 5,039 | 1,869 | 2,568 | 11,1 | 0,935 | 1,152 |
| 63,71 | 7,923 | 5,348 | 19 811 | 4,888 | 2,158 | 3,614 | 16,5 | 21,396 | 5,27 |
| 64,75 | 9,762 | 5,167 | 32 583 | 5,101 | 3,486 | 4,752 | 13,4 | 43,625 | 13,461 |
| 66,22 | 8,346 | 4,566 | 36 677 | 4,900 | 1,848 | 1,327 | 16,1 | 1,928 | 2,189 |
| 65,24 | 8,154 | 3,663 | 22 950 | 4,927 | 2,982 | 2,110 | 14,3 | 8,200 | 4,253 |
| 66,8 | 8,157 | 5,527 | 21 264 | 5,468 | 1,430 | 2,156 | 16,4 | 7,710 | 3,013 |
| 66,06 | 8,579 | 5,977 | 20 631 | 5,054 | 7,875 | 3,976 | 9,2 | 28,903 | 17,878 |
| 68,84 | 9,711 | 7,048 | 46 338 | 4,560 | 0,291 | 0,410 | 12 | 1,469 | 1,511 |
| 67,35 | 6,405 | 4,441 | 35 635 | 5,224 | 1,621 | 1,999 | 13,9 | 8,354 | 7,849 |
| 61,95 | 10,604 | 5,025 | 18 972 | 3,834 | 10,225 | 6,511 | 9,2 | 1,594 | 2,636 |
| 64,7 | 7,541 | 4,034 | 23 837 | 4,760 | 4,921 | 4,505 | 14 | 20,859 | 6,344 |
| 67,41 | 7,886 | 3,902 | 31 427 | 5,137 | 1,855 | 1,242 | 15,4 | 4,217 | 1,726 |
| 65,74 | 7,808 | 5,474 | 27 312 | 4,795 | 2,214 | 1,447 | 17,9 | 7,068 | 2,587 |
| 65,38 | 7,515 | 4,981 | 29 893 | 5,391 | 1,885 | 1,633 | 16,6 | 8,158 | 6,273 |
| 66,31 | 9,859 | 5,224 | 24 046 | 4,957 | 3,181 | 2,076 | 16,4 | 5,374 | 2,611 |
| 64,62 | 11,555 | 6,475 | 60 770 | 5,681 | 5,370 | 3,934 | 13,6 | 13,135 | 15,49 |
| 64,71 | 8,787 | 4,568 | 37 374 | 4,974 | 3,192 | 2,576 | 15,7 | 18,397 | 6,277 |
| 64,64 | 9,568 | 5,170 | 28 152 | 5,298 | 3,412 | 2,419 | 17 | 6,394 | 2,375 |
| 68,51 | 8,993 | 4,634 | 23 911 | 4,945 | 1,440 | 0,968 | 12,9 | 3,953 | 2,077 |
| 66,6 | 7,506 | 4,309 | 27 889 | 5,142 | 2,654 | 2,466 | 17,8 | 7,044 | 3,247 |
| 64,29 | 9,325 | 4,653 | 27 681 | 5,271 | 2,279 | 1,944 | 18,4 | 7,862 | 3,398 |
| 65,96 | 9,050 | 5,323 | 28 857 | 5,123 | 3,088 | 2,743 | 13,3 | 16,528 | 5,553 |
| 64,9 | 8,722 | 4,041 | 29 385 | 5,547 | 2,376 | 1,321 | 18,7 | 8,330 | 3,193 |
| 67,54 | 7,172 | 2,244 | 50 059 | 1,920 | 1,339 | 1,159 | 9,3 | 57,332 | 14,253 |
| 65,3 | 7,721 | 5,156 | 25 449 | 5,104 | 3,033 | 3,755 | 14,1 | 8,823 | 2,915 |
| 65,91 | 8,227 | 4,351 | 24 590 | 5,062 | 2,744 | 2,124 | 16,8 | 2,552 | 2,989 |
| 63,23 | 8,303 | 5,760 | 41 740 | 5,464 | 4,082 | 2,625 | 15,4 | 9,343 | 6,014 |
| 68,76 | 7,211 | 5,701 | 54 443 | 5,396 | 1,938 | 2,022 | 7,6 | 67,091 | 17,714 |
| 64,97 | 8,052 | 4,275 | 26 628 | 5,424 | 3,454 | 3,050 | 15,9 | 13,379 | 6,823 |
| 69,49 | 6,215 | 3,176 | 24 596 | 4,315 | 0,066 | 0,100 | 6,2 | 1,081 | 0,754 |
| 65,46 | 8,618 | 5,101 | 21 155 | 5,089 | 2,337 | 1,771 | 15,4 | 2,079 | 2,585 |
| 62,55 | 12,685 | 7,257 | 89 541 | 6,204 | 7,953 | 4,891 | 10,4 | 34,398 | 22,788 |
| 67,54 | 7,507 | 6,035 | 89 928 | 5,764 | 4,342 | 2,589 | 6 | 158,813 | 27,598 |
| 65,4 | 9,269 | 5,589 | 29 514 | 5,380 | 6,362 | 1,813 | 17,2 | 6,859 | 5,945 |