**НАЦІОНАЛЬНИЙ ТЕХНІЧНИЙ УНІВЕРСИТЕТ УКРАЇНИ**

**«КИЇВСЬКИЙ ПОЛІТЕХНІЧНИЙ ІНСТИТУТ ІМЕНІ ІГОРЯ СІКОРСЬКОГО»**

**Інститут прикладного системного аналізу**

**Кафедра системного проектування**

**ЗВІТ**

з виконання лабораторної роботи

з дисципліни **«Еколого-економічна організація виробництва»**

на тему: **«Кореляційно-регресійний аналіз впливу соціально-економічних чинників на рівень доходів населення за регіонами України»**

Виконав:

студент групи ДА-82

ННК «ІПСА»

Муравльов Андрій

Викладач: Караєва Н.В.

Київ — 2021

# **ЗМІСТ**

[**ЗМІСТ** 2](#_Toc1)

[**ТЕОРИТИЧНІ ДАНІ** 2](#_Toc2)

[**ТАБЛИЦІ ДАНИХ** 3](#_Toc3)

[**АЛГОРИТМ РОЗРАХУНКУ КОЕФІЦІЄНТА КОРЕЛЯЦІЇ** 8](#_Toc4)

[**ВИСНОВКИ** 9](#_Toc5)

# **ТЕОРИТИЧНІ ДАНІ**

**Кореляційно-регресійний аналіз —** це побудова та аналіз економіко-математичної моделі у вигляді рівняння регресії (рівняння кореляційного зв’язку), що виражає залежність результативної ознаки від однієї або кількох ознак-факторів і дає оцінку міри щільності зв’язку.

Кореляційно-регресійний аналіз складається з таких етапів:

* попередній (апріорний) аналіз,
* збирання інформації та її первинна обробка,
* побудова моделі (рівняння регресії),
* оцінка й аналіз моделі.

**Мультиколінеарність** моделі означає існування лінійної залежності або сильної кореляції між двома чи більше факторами. Мультиколінеарність між факторами Хі та Хj ( і = j) називається строгою, якщо існує лінійна залежність   
Xj = сХi.

На практиці економічні фактори часто пов'язані між собою і це істотно впливає на якість економетричного моделювання.

**Задачі роботи**

* Обґрунтувати вибір вхідним параметрів і сформувати таблицю з вхідними даними.
* Побудувати кореляційну матрицю засобами Excel.
* Розробити алгоритм розрахунку коефіцієнту кореляції.
* Проаналізувати результати розрахунку.

Ми вважаємо, що рівень доходів (РД) населення залежить від **заробітної плати** (ЗП), **прибутку та змішаного доходу** (ПЗД), **доходів від власності** (ДВ), **соціальних допомог** (СД) та **соціальних трансфертів в натурі** (СТН).

Тепер з допомогою даних Держкомстату України створимо таблиці 1, 2 та 3 з вхідними індикаторами, які характеризують рівень доходів населення у різних регіонах України. Будемо досліджувати 2013, 2016 і 2018 роки.

# **ТАБЛИЦІ ДАНИХ**

Таблиця 1. Значення вхідних індикаторів за 2013 рік (у млн грн).

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Область** | **РД** | **ЗП** | **ПЗД** | **ДВ** | **СД** | **СТН** |
| Автономна Республіка Крим | 57 324 | 21 010 | 8 914 | 2 389 | 12 573 | 8 442 |
| Вінницька | 46 157 | 15 135 | 10 869 | 2 119 | 10 640 | 6 349 |
| Волинська | 26 907 | 8 607 | 5 446 | 998 | 6 726 | 4 701 |
| Дніпропетровська | 124 594 | 57 783 | 16 924 | 7 707 | 25 394 | 13 947 |
| Донецька | 166 366 | 74 002 | 25 361 | 8 594 | 38 234 | 16 750 |
| Житомирська | 34 947 | 11 935 | 6 670 | 1 272 | 9 012 | 5 340 |
| Закарпатська | 29 102 | 9 192 | 5 850 | 790 | 6 782 | 5 347 |
| Запорізька | 62 671 | 25 941 | 11 360 | 3 370 | 12 906 | 7 681 |
| Івано-Франківська | 37 310 | 10 873 | 8 654 | 1 221 | 8 579 | 6 319 |
| Київська | 58 894 | 25 054 | 9 910 | 2 701 | 13 540 | 6 938 |
| Кіровоградська | 27 695 | 9 331 | 5 220 | 1 578 | 6 678 | 4 307 |
| Луганська | 71 485 | 29 941 | 9 070 | 3 073 | 18 978 | 8 759 |
| Львівська | 75 762 | 28 509 | 13 123 | 3 496 | 16 405 | 11 868 |
| Миколаївська | 35 125 | 13 704 | 5 704 | 1 845 | 7 732 | 4 897 |
| Одеська | 78 285 | 28 916 | 11 828 | 3 956 | 14 918 | 10 636 |
| Полтавська | 46 984 | 19 313 | 7 427 | 2 934 | 10 339 | 6 074 |
| Рівненська | 31 811 | 10 692 | 6 296 | 1 094 | 7 723 | 5 457 |
| Сумська | 33 469 | 12 284 | 6 486 | 1 618 | 7 672 | 4 692 |
| Тернопільська | 26 345 | 7 982 | 5 267 | 1 122 | 6 273 | 4 604 |
| Харківська | 91 333 | 35 883 | 16 218 | 4 744 | 17 932 | 13 409 |
| Херсонська | 29 489 | 9 061 | 6 312 | 1 290 | 6 583 | 4 280 |
| Хмельницька | 36 770 | 11 780 | 8 148 | 1 773 | 8 725 | 5 407 |
| Черкаська | 35 024 | 12 632 | 5 296 | 2 179 | 8 897 | 5 231 |
| Чернівецька | 22 408 | 6 431 | 4 794 | 795 | 5 210 | 3 926 |
| Чернігівська | 30 393 | 11 004 | 5 382 | 1 642 | 7 633 | 4 317 |
| м. Київ | 218 747 | 118 529 | 16 202 | 22 922 | 24 176 | 23 218 |
| м. Севастополь | 13 336 | 5 210 | 937 | 730 | 2 863 | 1 944 |

Таблиця 2. Значення вхідних індикаторів за 2016 рік (у млн грн).

«…» — відсутність даних з області.

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Область** | **РД** | **ЗП** | **ПЗД** | **ДВ** | **СД** | **СТН** |
| Автономна Республіка Крим | … | … | … | … | … | … |
| Вінницька | 69 654 | 23 458 | 19 043 | 2 447 | 12 522 | 10 746 |
| Волинська | 39 359 | 13 537 | 9 255 | 1 071 | 8 064 | 6 733 |
| Дніпропетровська | 184 138 | 86 057 | 33 836 | 7 462 | 29 886 | 21 892 |
| Донецька | 111 547 | 55 007 | 11 260 | 3 370 | 27 512 | 11 588 |
| Житомирська | 51 920 | 18 436 | 11 822 | 1 424 | 10 509 | 8 634 |
| Закарпатська | 42 235 | 14 501 | 10 474 | 804 | 8 034 | 7 243 |
| Запорізька | 94 160 | 37 880 | 22 191 | 3 282 | 16 472 | 11 402 |
| Івано-Франківська | 54 492 | 16 483 | 15 608 | 1 219 | 10 155 | 8 950 |
| Київська | 87 937 | 39 426 | 17 543 | 2 252 | 16 030 | 11 529 |
| Кіровоградська | 40 427 | 14 247 | 9 073 | 1 999 | 7 899 | 6 418 |
| Луганська | 38 022 | 17 685 | 3 094 | 1 193 | 10 434 | 5 014 |
| Львівська | 112 697 | 44 323 | 24 725 | 3 707 | 19 520 | 17 117 |
| Миколаївська | 50 728 | 20 881 | 9 976 | 1 718 | 8 977 | 7 191 |
| Одеська | 115 025 | 44 524 | 21 667 | 3 675 | 17 654 | 13 606 |
| Полтавська | 69 789 | 28 707 | 12 483 | 4 335 | 12 319 | 10 502 |
| Рівненська | 45 716 | 16 201 | 11 027 | 1 133 | 8 966 | 7 609 |
| Сумська | 50 951 | 18 803 | 11 858 | 1 943 | 9 030 | 8 256 |
| Тернопільська | 38 727 | 12 275 | 8 922 | 1 072 | 7 330 | 7 506 |
| Харківська | 131 681 | 52 212 | 28 455 | 4 281 | 23 011 | 18 727 |
| Херсонська | 42 707 | 13 768 | 10 773 | 1 391 | 7 756 | 5 694 |
| Хмельницька | 55 542 | 18 123 | 14 690 | 2 180 | 10 433 | 8 786 |
| Черкаська | 51 710 | 18 901 | 8 840 | 3 004 | 10 792 | 8 929 |
| Чернівецька | 32 397 | 9 664 | 8 471 | 825 | 6 215 | 5 544 |
| Чернігівська | 44 283 | 16 288 | 8 934 | 1 797 | 8 985 | 7 566 |
| м. Київ | 333 927 | 185 379 | 34 193 | 17 868 | 29 268 | 33 378 |
| м. Севастополь | … | … | … | … | … | … |

Таблиця 3. Значення вхідних індикаторів за 2018 рік.

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Область** | **РД** | **ЗП** | **ПЗД** | **ДВ** | **СД** | **СТН** |
| Автономна Республіка Крим | … | … | … | … | … | … |
| Вінницька | 114 480 | 45 927 | 28 588 | 3 593 | 16 185 | 17 361 |
| Волинська | 63 810 | 26 013 | 13 537 | 1 572 | 10 347 | 10 967 |
| Дніпропетровська | 307 844 | 157 772 | 53 211 | 7 725 | 43 590 | 34 490 |
| Донецька | 174 771 | 92 064 | 13 630 | 4 433 | 39 928 | 19 934 |
| Житомирська | 84 830 | 34 727 | 18 881 | 2 130 | 13 869 | 13 276 |
| Закарпатська | 69 194 | 28 437 | 16 030 | 846 | 10 306 | 11 825 |
| Запорізька | 147 627 | 67 656 | 30 873 | 3 932 | 22 430 | 18 620 |
| Івано-Франківська | 87 479 | 31 557 | 23 376 | 1 316 | 13 509 | 14 292 |
| Київська | 150 606 | 74 615 | 26 930 | 3 402 | 22 387 | 20 797 |
| Кіровоградська | 63 999 | 26 431 | 13 222 | 2 382 | 10 347 | 10 157 |
| Луганська | 58 880 | 29 470 | 3 837 | 1 193 | 14 971 | 8 638 |
| Львівська | 189 077 | 83 401 | 39 619 | 4 144 | 26 831 | 28 501 |
| Миколаївська | 81 581 | 37 991 | 14 839 | 2 031 | 12 226 | 11 009 |
| Одеська | 193 923 | 80 374 | 35 080 | 4 254 | 23 826 | 23 741 |
| Полтавська | 114 656 | 53 539 | 19 263 | 5 502 | 16 977 | 16 738 |
| Рівненська | 73 661 | 30 444 | 16 485 | 1 426 | 11 610 | 12 206 |
| Сумська | 79 848 | 33 411 | 17 306 | 2 663 | 12 234 | 12 521 |
| Тернопільська | 61 731 | 23 505 | 13 107 | 1 434 | 9 363 | 11 384 |
| Харківська | 216 227 | 96 555 | 41 142 | 6 859 | 31 726 | 32 606 |
| Херсонська | 68 064 | 25 098 | 15 128 | 1 582 | 10 318 | 10 128 |
| Хмельницька | 86 821 | 34 060 | 19 781 | 3 025 | 13 701 | 13 976 |
| Черкаська | 82 600 | 34 194 | 13 771 | 4 077 | 14 357 | 13 801 |
| Чернівецька | 52 108 | 18 522 | 12 561 | 830 | 7 957 | 9 043 |
| Чернігівська | 69 247 | 29 818 | 12 779 | 2 082 | 11 732 | 11 528 |
| м. Київ | 555 666 | 333 786 | 59 089 | 18 731 | 45 049 | 58 654 |
| м. Севастополь | … | … | … | … | … | … |

Маючи дані, можемо побудувати матриці кореляції в Excel.   
На рис. 1 наведені результати побудови матриці кореляції за даними 2013 року.

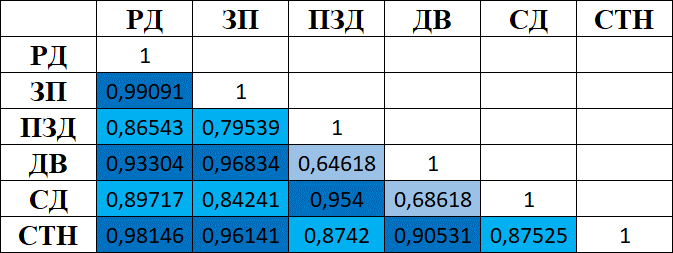


Рис. 1. Результати побудови матриці кореляції за даними 2013 року

Коефіцієнти кореляції виділені різними кольорами. Покажемо межі значень:

* 0 … 0.3 — зв'язку немає (червоний);
* 0.3 … 0.5 — слабкий зв'язок (жовтий);
* 0.5 … 0.7 — середній зв'язок (світло-синій);
* 0.7 … 0.9 — сильний зв'язок (блакитний);
* 0.9 … 1 — дуже сильний зв'язок (синій).

Знаючи це, поглянемо на рис.1. Аж 7 пар індикаторів дають результат більше за 0.9, тобто дуже сильний зв'язок між ними. Між ЗП та РД коефіцієнт максимально наближений до 1 (0,990914718), це можна пояснити тим, що зарплата є основним джерелом доходів населення. Дві пари показують середній зв’язок (див. рис. 1). Інші пари мають коефіцієнти 0.7…0.9, що каже нам про сильний зв'язок між ними. Нийнижчий коєфіцієнт — 0,646182572 між ДВ і ПЗД.

Отже, існує залежність між індикаторами на високому рівні.

Перейдемо до наступної таблиці даних.

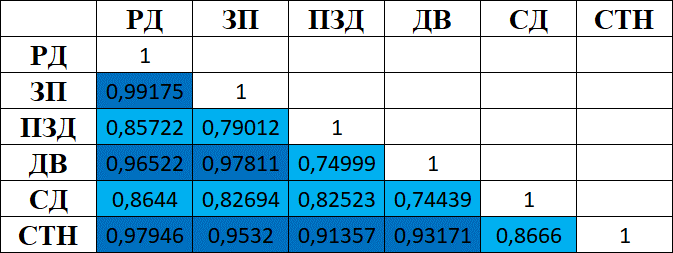


Рис. 2. Результати побудови матриці кореляції за даними 2016 року

На рис. 2 (дані за 2016 рік), кореляція між ЗП та РД знову найвища (0,991748727). Знову бачимо 7 пар величин, що дуже сильно між собою корелюють (хоч і дещо інших порівняно з попередньою матрицею): показники більші за 0.9. Інші пари показують значення кореляції 0.7–0.9, що показує сильний зв'язок між ними. Нийнижчий коєфіцієнт — 0,744390433 між ДВ і СД.

Так, залежність між індикаторами є на дуже високому рівні (особливо порівнюючи з 2013 роком).

Розглянемо останній набір даних — за 2018 рік.

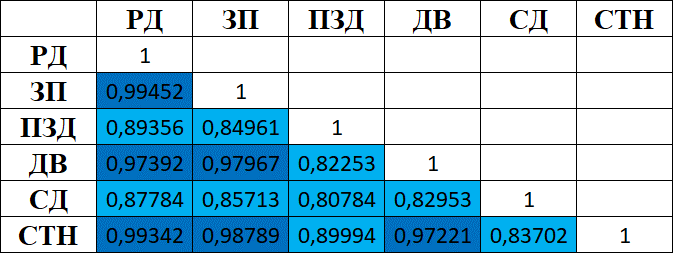
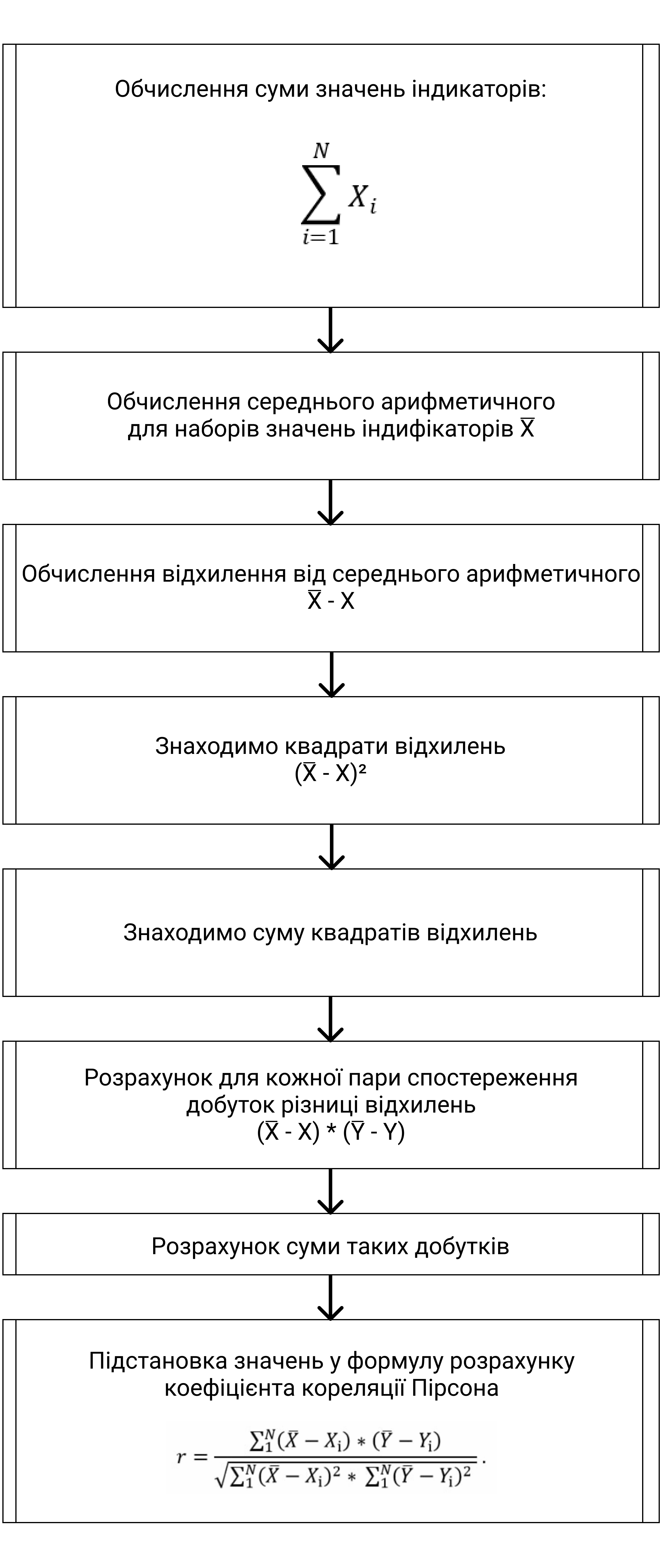


Рис. 3. Результати побудови матриці кореляції за даними 2018 року

Тут 6 пар мають коефіцієнти більші за 0.9, тобто мають дуже сильний зв'язок між собою. Вже третє пара ЗП та РД має коефіцієнт майже 1 (0,9945195). До речі, тут вже з’явився конкурент у найкращої пари, а саме — пара СТН/РД (0,993422925). Серед інших пар показники у межах 0.8–0.9, що вказує на сильний зв'язок між ними. Нийнижчий коєфіцієнт — 0,807839911 має пара СД/ПЗД.

Отже, залежність залишається на високому рівні.

# **АЛГОРИТМ РОЗРАХУНКУ КОЕФІЦІЄНТА КОРЕЛЯЦІЇ**



# **ВИСНОВКИ**

Результати, показані на рис. 1–3 свідчать, що у всіх взятих роках (а саме 2013, 2016 і 2018) найбільший вплив на доходи населення мала заробітна плата (ЗП), а найменший — прибуток та змішаний дохід (ПЗД) у 2012 і 2016 роках та соціальні допомоги (СД) у 2018 році. Провівши аналіз результатів кореляції можна помітити, що вплив майже усіх показників є дуже високим (коефіцієнти більші за 0.85). Вплив ЗП протягом усіх років був найсильнішим: коефіцієнт вище за 0.99. Варто зазначити, що на рівні з ЗП у 2018 році на РД впливали й СТН (0,993422925).

Вплив прибутку та змішаних доходів і соціальних допомог був нижче за вплив інших індикаторів. Це прослідковується постійно, тому можна сказати про це як про закономірність.

Результати розрахунків матриць на рис. 1–3 показують, що між показниками впливу існує сильний кореляційний зв'язок, тому для побудови оптимальної економетричної прогнозної моделі небідно надалі провести мультиколінеарний аналіз.

Протокол, Excel таблицю та алгоритм також викладено на системі контролю версій GitHub

<https://github.com/lakub-muravlov/fourth-course-projects/tree/main/EEOV/Lab1/Matrix>