**ПРОСТОРОВИЙ АНАЛІЗ ТА СТАТИСТИЧНЕ МОДЕЛЮВАННЯ В ГЕОЛОГІЇ**

***Практичне робота 3 (2023)***

***Частина 1 Регресія***

***Завдання 1 :***

Таблиця водних ресурсів water, що наведена нижче, містить один рядок для кожної країни з даними за 2014 рік. Стовпець OBS\_VALUE представляє приблизну ціну 1,5-літрової пляшки мінеральної води в цій країні, а стовпець mm\_precipitation містить середню кількість опадів у цій країні (у міліметрах/рік).

|  |  |  |
| --- | --- | --- |
| ***Таблиця водних ресурсів*** | | |
| *COUNTRY* | OBS\_VALUE | mm\_precipitation |
| *Albania* | 0.55 | 1485 |
| *Algeria* | 0.027 | 89 |
| *Angola* | 1 | 1010 |
| *Argentina* | 1.29667 | 591 |
| *Armenia* | 0.5325 | 562 |
| *Australia* | 2.07302 | 534 |
| *Austria* | 0.72 | 1110 |
| *Azerbaijan* | 0.576 | 447 |
| *Bangladesh* | 0.374 | 2666 |
| *... (90 rows omitted)* | | |

Наступна таблиця нижче містить розрахунки разом із діаграмою розсіювання за даними таблиці

|  |  |
| --- | --- |
| Table of calculations from data in the water table. These are the expressions and values. np.average(water.column('OBS_VALUE')) = 0.919016; np.std(water.column('OBS_VALUE')) = 0.464763; np.average(water.column('mm_precipitation')) = 1010.4; p.std(water.column('mm_precipitation')) = 752.475; correlation(water, 'OBS_VALUE', 'mm_precipitation') = 0.262079 |  |

*Яке значення кореляції correlation(water, ‘mm\_precipitation’, ‘OBS\_VALUE’)?*

*Яким є нахил (slope) рівняння лінії регресії за даними таблиці для знаходження значення* OBS\_VALUE *y з використанням* mm\_precipitation *x? (Підказка: )*

*Яким є залишок (перетин) (intercept) рівняння лінії регресії за даними таблиці для знаходження значення* OBS\_VALUE *y з використанням* mm\_precipitation *x? (Підказка: )*

*Використовуючи рівняння лінії регресії, яким би ми очікували OBS\_VALUE у 2014 році для країни, яка мала в середньому 700 мм опадів? (Підказка: )*

***Завдання 2 :***

Чи підходить знайдене рівняння лінії регресії для прогнозування ціни на основі опадів?

***Частина 2 Середньо-квадратична похибка***

***Завдання 3 :***

Оберіть Вашу відповідь для наступних тверджень:

Лінійна кореляція передбачає причинно-наслідковий зв’язок?

Причинно-наслідковий зв'язок передбачає лінійну кореляцію?

Метод найменших квадратів - це спосіб створення лінії регресії?

Ми можемо розрахувати «найкращу апроксимацію» лінії регресії, якщо мінімізуємо середню помилку?

Лінія регресії є єдиною лінією, яка мінімізує середню квадратичну помилку?

Незалежно від форми діаграми розсіювання, існує унікальна лінія, яка мінімізує середню квадратичну помилку оцінки?

***Завдання 4:***

Виконайте інструкції наведені в файлі *pr3\_4.ipynb* з директорії pr3\_4 (архіву pr3\_4.zip). Приєднайте до вашого проекту JUPYTER NOTEBOOK систему контролю версій Git та розмістить його на GitHub надавши посилання нижче:

*Вставити посилання на Вашу Практичну робота 3 (2023) завдання 3\_4 на GitHub: …*