

ПРОСТОРОВИЙ АНАЛІЗ ТА СТАТИСТИЧНЕ МОДЕЛЮВАННЯ В ГЕОЛОГІЇ

Практичне робота 3 (2023)

Частина 1 Регресія

Завдання 1 :

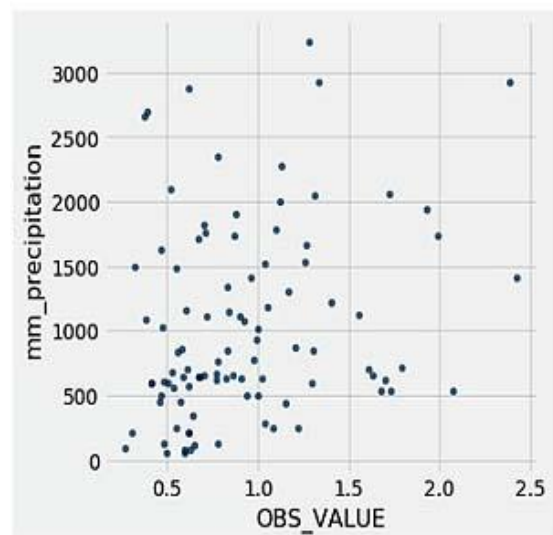
Таблиця водних ресурсів `water`, що наведена нижче, містить один рядок для кожної країни з даними за 2014 рік. Стовпець `OBS_VALUE` представляє приблизну ціну 1,5-літрової пляшки мінеральної води в цій країні, а стовпець `mm_precipitation` містить середню кількість опадів у цій країні (у міліметрах/рік).

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

Expression	Values
<code>np.average(water.column('OBS_VALUE'))</code>	0.919016
<code>np.std(water.column('OBS_VALUE'))</code>	0.464763
<code>np.average(water.column('mm_precipitation'))</code>	1010.4
<code>np.std(water.column('mm_precipitation'))</code>	752.475
<code>correlation(water, 'OBS_VALUE', 'mm_precipitation')</code>	0.262079



Яке значення кореляції `correlation(water, 'mm_precipitation', 'OBS_VALUE')`?

0.268079

Яким є нахил (*slope*) рівняння лінії регресії за даними таблиці для знаходження значення `OBS_VALUE` y з використанням `mm_precipitation` x ? (Підказка: $slope = r \frac{SD_y}{SD_x}$)

0.00016187198548390312

Яким є залишок (перетин) (*intercept*) рівняння лінії регресії за даними таблиці для знаходження значення *OBS_VALUE* *y* з використанням *mm_precipitation* *x*? (Підказка: $intercept = \mu_y - slope * \mu_x$)

0.7554605458670643

Використовуючи рівняння лінії регресії, яким би ми очікували *OBS_VALUE* у 2014 році для країни, яка мала в середньому 700 мм опадів? (Підказка: $\hat{y} = slope * x + intercept$)

0.8687709357057966

Завдання 2 :

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

Так

Ні

Частина 2 Середньо-квадратична похибка

Завдання 3 :

Оберіть Вашу відповідь для наступних тверджень:

Лінійна кореляція передбачає причинно-наслідковий зв'язок?

Так

Ні

Причинно-наслідковий зв'язок передбачає лінійну кореляцію?

Так

Ні

Метод найменших квадратів - це спосіб створення лінії регресії?

Так

Ні

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

Так

Ні

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

Так

Ні

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

Так

Ні

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 розмістіть його на GitHub надавши посилання нижче:

[Вставити посилання на Вашу Практичну роботу 3 \(2023\) завдання 3 4 на GitHub](#)