

Will Rogers phenomenon

Following the geographical migration in the United States at the time of the pre-war economic depression, American comedian Will Rogers noted the following:

When the [Okies](#) left Oklahoma and moved to California, they raised the average intelligence level in both states.

That was obviously a joke, but the phenomenon is actually possible. For example, consider the following two series of integers: $A = [5, 6, 7, 8, 9]$ $B = [1, 2, 3, 4]$ If we move the element 5 from series A to series B , the average of both series increases.

A similar phenomenon occurs in practice, for example, when groups of patients with [carcinoma](#) - divided into stages according to the [TNM system](#) - are compared. Felsenstein et al. compared two groups of patients with lung cancer who had been diagnosed in 1953-54 and 1977, respectively. While the distribution of patients over the TNM stages I-III in both groups was similar, it was revealed that the survival at 6 months for all stages was better in the group from 1977. In 1977, however, modern methods such as computed tomography, ultrasound and isotope research had been used frequently to determine the stage. If the patients would have been classified without the use of these techniques, in 1977, a significant number of them would have ended in a more favorable stage. If in that case, the survival at 6 months was calculated again, this would be no different from the treated patients in 1953-54. The improved survival in 1977 was not a result of improved therapy, but was the result of a more accurate classification of stages by using of new diagnostic techniques. This study shows the danger of conclusions based on comparisons with historical control groups, even if a seemingly equal stage classification is used.

Assignment

In this assignment, a series of integers is represented as a list or a tuple of integers. You may also assume that all sequences (lists or tuples) used in this statement are not empty. Asked:

- Write a function `average` to which a series of integers must be passed. The function must return the average value of the given numbers, represented as a *floating point* number.
- Write a function `move1` to which three arguments should be passed. The first two arguments are **lists** of integers. The third argument contains a set of integers that were selected from the first **list**. The function should return a value of `None`, but after calling the function all elements of the third argument must be removed from the first argument, and added in the specified order at the back of the second argument. If an element occurs several times in the first argument, then the first occurrence of it must be removed when deleting.
- Write a function `move2` to which three arguments should be passed. The first two arguments are **sets** of integers. The third argument contains a set of integers that were selected from the first **set**. The function cannot modify any argument, but must return two new lists. The first list contains all the elements of the first argument, in the specified order, but from which all elements of the third argument were removed. If an element occurs several times in the list, then the first occurrence must be removed. The second list contains all the elements of the second argument, to which all the elements of the third argument were added at the end

in the specified sequence.

- Write a function `willrogers` to which three arguments should be passed. The first two arguments are **sets** of integers. The third argument contains a set of integers that were selected from the first **set**. The function cannot modify any argument, and must return a Boolean value that indicates whether the average values of the first two arguments both rise after the elements of the third argument were removed from the first series and added to the second series.

Example

```
>>> average((5, 6, 7, 8, 9))
```

```
7.0
```

```
>>> average([1, 2, 3, 4])
```

```
2.5
```

```
>>> sequence1 = [5, 6, 7, 8, 9]
```

```
>>> sequence2 = [1, 2, 3, 4]
```

```
>>> move1(sequence1, sequence2, [5])
```

```
>>> sequence1
```

```
[6, 7, 8, 9]
```

```
>>> sequence2
```

```
[1, 2, 3, 4, 5]
```

```
>>> sequence1 = (5, 6, 7, 8, 9)
```

```
>>> sequence2 = [1, 2, 3, 4]
```

```
>>> move2(sequence1, sequence2, [5])
```

```
[(6, 7, 8, 9), [1, 2, 3, 4, 5])
```

```
>>> sequence1
```

```
(5, 6, 7, 8, 9)
```

```
>>> sequence2
```

```
[1, 2, 3, 4]
```

```
>>> willrogers([5, 6, 7, 8, 9], [1, 2, 3, 4], [5])
```

```
True
```

```
>>> willrogers((5, 6, 7, 8, 9), (1, 2, 3, 4), (7, 9))
```

```
False
```

Sources

- **Feinstein AR, Sosin DM, Wells CK (1985).** The Will Rogers phenomenon. Stage migration and new diagnostic techniques as source of misleading statistics for survival in cancer. *The New England Journal of Medicine* **312(25)**, 1604-1608. [↗](#)
- **Sormani MP, Tinorè M, Rovaris M, Rovira A, Vidal X, Bruzzi P, Filippi M, Montalban X (2008).** Will Rogers phenomenon in multiple sclerosis. *Annals of Neurology* **64(4)**, 428-433. [↗](#)

Naar aanleiding van de geografische migratie in de Verenigde Staten ten tijde van de vooroorlogse economische depressie, merkte de Amerikaanse komiek Will Rogers het volgende op:

When the [Okies](#) left Oklahoma and moved to California, they raised the average intelligence level in both states.

Dat was natuurlijk als een grap bedoeld, maar het fenomeen is daadwerkelijk mogelijk.

Beschouw bijvoorbeeld de volgende twee reeksen van natuurlijke getallen: $A = [5, 6, 7, 8, 9]$ en $B = [1, 2, 3, 4]$. Als we het element 5 verplaatsen van reeks A naar reeks B , dan stijgt het gemiddelde van beide reeksen.

Een soortgelijk fenomeen komt in de praktijk bijvoorbeeld voor wanneer groepen patiënten met een [carcinoom](#) — in stadia ingedeeld volgens het [TNM-systeem](#) — met elkaar vergeleken worden. Felsenstein et al. vergeleken twee groepen patiënten met longcarcinoom, bij wie de diagnose respectievelijk gesteld was in 1953-54 en 1977. Terwijl de verdeling van de patiënten over de TNM-stadia I-III in beide groepen gelijk was, bleek de overleving na 6 maanden voor alle stadia beter te zijn in de groep uit 1977. Feit was echter dat in 1977 ten behoeve van de stadiumindeling ook veelvuldig gebruik was gemaakt van moderne methoden zoals computertomografie, echografie en isotopenonderzoek. Indien de patiënten in 1977 zouden ingedeeld zijn zonder gebruik te maken van deze technieken, dan zou een aanzienlijk aantal in een gunstiger stadium terecht zijn gekomen. Als in dat laatste geval de overleving na 6 maanden opnieuw werd berekend, verschilde deze niet van de in 1953-54 behandelde patiënten. De verbeterde overleving in 1977 bleek dus geen gevolg van een verbeterde therapie, maar was het resultaat van een meer nauwkeurige stadiumindeling door gebruik te maken van nieuwe diagnostische technieken. Dit onderzoek toont het gevaar aan van conclusies gebaseerd op vergelijkingen met historische controlegroepen, zelfs indien gebruik gemaakt wordt van een schijnbaar gelijke stadiumindeling.

Opgave

In deze opgave wordt een reeks natuurlijke getallen voorgesteld als een lijst of een tuple van natuurlijke getallen. Je mag er ook van uitgaan dat alle reeksen (lijsten of tuples) die in deze opgave gebruikt worden niet leeg zijn. Gevraagd wordt:

- Schrijf een functie `gemiddelde` waaraan een reeks natuurlijke getallen moet doorgegeven worden. De functie moet de gemiddelde waarde van de gegeven getallen teruggeven, voorgesteld als een *floating point* getal.
- Schrijf een functie `verplaats1` waaraan drie argumenten moeten doorgegeven worden. De eerste twee argumenten zijn **lijsten** natuurlijke getallen. Het derde argument bevat een reeks natuurlijke getallen die geselecteerd werden uit de eerste **lijst**. De functie moet de waarde `None` teruggeven, maar na het aanroepen van de functie moeten alle elementen van het derde argument verwijderd zijn uit het eerste argument en in de opgegeven volgorde achteraan toegevoegd zijn aan het tweede argument. Als een element verschillende keren in het eerste argument voorkomt, dan moet bij het verwijderen het eerste voorkomen ervan verwijderd worden.
- Schrijf een functie `verplaats2` waaraan drie argumenten moeten doorgegeven worden. De eerste twee argumenten zijn **reeksen** natuurlijke getallen. Het derde argument bevat een reeks natuurlijke getallen die geselecteerd werden uit de eerste **reeks**. De functie mag geen enkel argument wijzigen, maar moet twee nieuwe lijsten teruggeven. De eerste lijst bevat alle elementen van het eerste argument, in de opgegeven volgorde, maar waaruit alle elementen van het derde argument verwijderd werden. Als een element verschillende keren in de lijst voorkomt, dan moet bij het verwijderen het eerste voorkomen ervan verwijderd worden. De tweede lijst bevat alle elementen van het tweede argument, waaraan alle elementen van het derde argument in de opgegeven volgorde achteraan toegevoegd werden.
- Schrijf een functie `willrogers` waaraan drie argumenten moeten doorgegeven worden. De

eerste twee argumenten zijn **reeksen** natuurlijke getallen. Het derde argument bevat een reeks natuurlijke getallen die geselecteerd werden uit de eerste **reeks**. De functie mag geen enkel argument wijzigen, en moet een Booleaanse waarde teruggeven die aangeeft of de gemiddelde waarden van de eerste twee argumenten beide stijgen nadat de elementen van het derde argument uit de eerste reeks werden verwijderd en aan de tweede reeks werden toegevoegd.

Voorbeeld

```
>>> gemiddelde((5, 6, 7, 8, 9))
```

```
7.0
```

```
>>> gemiddelde([1, 2, 3, 4])
```

```
2.5
```

```
>>> sequentie1 = [5, 6, 7, 8, 9]
```

```
>>> sequentie2 = [1, 2, 3, 4]
```

```
>>> verplaats1(sequentie1, sequentie2, [5])
```

```
>>> sequentie1
```

```
[6, 7, 8, 9]
```

```
>>> sequentie2
```

```
[1, 2, 3, 4, 5]
```

```
>>> sequentie1 = (5, 6, 7, 8, 9)
```

```
>>> sequentie2 = [1, 2, 3, 4]
```

```
>>> verplaats2(sequentie1, sequentie2, [5])
```

```
(6, 7, 8, 9), [1, 2, 3, 4, 5])
```

```
>>> sequentie1
```

```
(5, 6, 7, 8, 9)
```

```
>>> sequentie2
```

```
[1, 2, 3, 4]
```

```
>>> willrogers([5, 6, 7, 8, 9], [1, 2, 3, 4], [5])
```

```
True
```

```
>>> willrogers((5, 6, 7, 8, 9), (1, 2, 3, 4), (7, 9))
```

```
False
```

Bronnen

- **Feinstein AR, Sosin DM, Wells CK (1985).** The Will Rogers phenomenon. Stage migration and new diagnostic techniques as source of misleading statistics for survival in cancer. *The New England Journal of Medicine* **312(25)**, 1604-1608. [↗](#)
- **Sormani MP, Tinorè M, Rovaris M, Rovira A, Vidal X, Bruzzi P, Filippi M, Montalban X (2008).** Will Rogers phenomenon in multiple sclerosis. *Annals of Neurology* **64(4)**, 428-433. [↗](#)