

House of representatives

The *United States House of Representatives* consists of 435 voting members.



Seal of the House of Representatives

Each voting member is elected in a congressional district that is established on the basis of data from the census. On the basis of the decennial census it is determined how many representatives each state will receive in the House of Representatives (in the census of 2000 any seat represents an average of 646,952 inhabitants). Each state gets at least one seat. Every ten years, the states revise the number of congressional districts and the corresponding borders to reach conformity with the number of delegates they can send to the House of Representatives. One member is appointed for a period of two years. Furthermore, the House has six non-voting members from Guam, American Samoa, the Northern Mariana Islands, the Virgin Islands, Puerto Rico and the District of Columbia.

Since 1940, the number of representatives per state is determined using the Huntington-Hill method. This method was developed by Edward Huntington and Joseph Hill. In this method, the percentage differences between the sizes of the districts are minimized. The method begins by allocating one seat to each state. Then, the next step is repeated until there are no more seats left:

1. For each state, the geometric average is calculated of the ratio of the population to the current number of seats and the ratio of the population to the number of seats if the next seat would go to that state. This average is given by the following formula: $\frac{p}{\sqrt{n(n+1)}}$ where p is the population and n the current number of seats.
2. The subsequent seat goes to the state for which the average is the highest.

At the first census in 1790 there were only 15 states, and there were only 105 seats in the House of Representatives. At that time the Huntington-Hill method was not used, but if that would have been the case, the distribution would have looked like this.

State	Inhabitants	Seats
Connecticut	236841	7
Delaware	55540	2
Georgia	70835	2
Kentucky	68705	2
Maryland	278514	8
Massachusetts	475327	14

State	Inhabitants	Seats
New Hampshire	141822	4
New Jersey	179570	5
New York	331589	10
North Carolina	353523	10
Pennsylvania	432879	12
Rhode Island	68446	2
South Carolina	206236	6
Vermont	85533	3
Virginia	630560	18

Assignment

The file [uscensus2010.csv](#) contains the results of the censuses in the United States since 1910. The file [uscensus1790.csv](#) contains the results of the census in the US in 1790. The first element of each row, is the state with which that row corresponds. For the header the first element is empty. The elements in the header indicate which column corresponds to which year.

Asked:

1. Write a function `census` to which the name of a file must be passed. This must be a CSV file with in the first column the names of the states and regions, and in subsequent columns numbers the censuses. The function also has a second optional argument that specifies which column in the CSV file will be used. By default, the second column is used, and the column parameter has the value 1. This function then returns a dictionary linking the names of the states with their population, the number of which is, of course, an integer.
2. Write a function `seatdistribution` to which a dictionary must be passed as an argument. This dictionary is the same size as those returned by the `census` function. The function must return a new dictionary, which connects each state with the number of seats for that state. The function also has an optional second parameter `total` of which the default equals 435 and that shows how many seats are distributed in total.

Example

```
>>> census('uscensus2010.csv')
{'Mississippi': 1797114, 'Oklahoma': 1657155, ..., 'Maine': 742371}
>>> census('uscensus2010.csv', column=11)
{'Mississippi': 2967297, 'Oklahoma': 3751351, ..., 'Maine': 1328361}
>>> inhabitants = census('uscensus2010.csv', column=11)
>>> seatdistribution(inhabitants)
{'Mississippi': 4, 'Oklahoma': 5, ..., 'Maine': 2}
```

Het Huis van Afgevaardigden van de Verenigde Staten (in het Engels: *United States House of Representatives*) is het lagerhuis van het Amerikaans Congres en bestaat uit 435 stemgerechtigde leden.



Zegel van het Huis van Afgevaardigden

Elk stemgerechtigd lid wordt gekozen in een congresdistrict dat vastgesteld wordt op basis van de gegevens van de volkstelling (*census*). Aan de hand van de tienjaarlijkse volkstelling wordt bepaald hoeveel vertegenwoordigers iedere staat krijgt toebedeeld in het Huis van Afgevaardigden (in de volkstelling van 2000 staat elke zetel voor gemiddeld 646.952 inwoners). Elke staat krijgt minstens één zetel. Iedere tien jaar herzien de staten het aantal congresdistricten en de bijbehorende grenzen om die in overeenstemming te bereiken met het aantal afgevaardigden dat ze naar het Huis van Afgevaardigden mogen sturen. Een lid zit een termijn van twee jaar. Verder heeft het Huis nog zes leden zonder stemrecht, uit Guam, Amerikaans-Samoa, de Noordelijke Marianen, de Amerikaanse Maagdeneilanden, Puerto Rico en het District of Columbia.

Sinds 1940 wordt het aantal afgevaardigden per staat bepaald volgens de Huntington-Hill-methode. Deze methode werd ontwikkeld door Edward Huntington en Joseph Hill. In deze methode worden percentageverschillen tussen de grootte van de districten geminimaliseerd. De methode begint met aan elke staat één zetel toe te kennen. Vervolgens wordt de volgende stap herhaald totdat er geen zetels meer over zijn:

1. Voor elke staat wordt het geometrisch gemiddelde berekend van de verhouding van het bevolkingsaantal tot het huidig aantal zetels en de verhouding van het bevolkingsaantal tot het aantal zetels als de volgende zetel naar die staat zou gaan. Dit gemiddelde wordt gegeven door de volgende formule: $\sqrt[n]{\frac{p}{n}}$ waarbij p het bevolkingsaantal is en n het huidig aantal zetels.
2. De volgende zetel gaat naar de staat waarvoor dit gemiddelde het hoogste is.

Bij de eerste volkstelling in 1790 waren er maar 15 staten en er waren er maar 105 zetels in het Huis van Afgevaardigden. Op dat moment werd de Huntington-Hill-methode nog niet gebruikt, maar als dat wel het geval was geweest, dan had de verdeling er als volgt uitgezien.

Staat	Inwoners	Zetels
Connecticut	236841	7
Delaware	55540	2
Georgia	70835	2
Kentucky	68705	2
Maryland	278514	8
Massachusetts	475327	14
New Hampshire	141822	4
New Jersey	179570	5

Staat	Inwoners	Zetels
New York	321589	19
North Carolina	353523	10
Pennsylvania	432879	12
Rhode Island	68446	2
South Carolina	206236	6
Vermont	85533	3
Virginia	630560	18

Opgave

Het bestand [uscensus2010.csv](#) bevat de resultaten van de volkstellingen in de Verenigde Staten sinds 1910. Het bestand [uscensus1790.csv](#) bevat de resultaten van de volkstelling in de Verenigde Staten van 1790. Het eerste element van elke rij, is de staat waarmee die rij correspondeert. Voor de hoofding is het eerste element leeg. De elementen in de hoofding geven aan welke kolom overeenkomt met welk jaar.

Gevraagd wordt:

1. Schrijf een functie `volkstelling` waaraan de naam van een bestand moet worden meegegeven. Dit moet een CSV-bestand zijn waarbij in de eerste kolom de namen van de staten en de regio's staan en in de daaropvolgende kolommen de aantallen van de volkstellingen. De functie heeft ook nog een tweede optioneel argument dat aangeeft welke kolom uit het CSV-bestand er gebruikt zal worden. Standaard wordt de tweede kolom gebruikt, en heeft de parameter `kolom` dus de waarde 1. Deze functie geeft dan een dictionary terug die de namen van de staten verbindt met hun bevolkingsaantal, waarbij dit aantal natuurlijk een geheel getal is.
2. Schrijf een functie `zetelverdeling` waaraan een dictionary als argument moet doorgegeven worden. Deze dictionary is van hetzelfde formaat als deze die wordt teruggegeven door de functie `volkstelling`. De functie moet een nieuwe dictionary teruggeven, die elke staat verbindt met het aantal zetels voor die staat. De functie heeft ook nog een optionele tweede parameter `totaal` die standaard gelijk is aan 435 en die aangeeft hoeveel zetels er in totaal verdeeld worden.

Voorbeeld

```
>>> volkstelling('uscensus2010.csv')
{'Mississippi': 1797114, 'Oklahoma': 1657155, ..., 'Maine': 742371}
>>> volkstelling('uscensus2010.csv', kolom=11)
{'Mississippi': 2967297, 'Oklahoma': 3751351, ..., 'Maine': 1328361}
>>> inwoners = volkstelling('uscensus2010.csv', kolom=11)
>>> zetelverdeling(inwoners)
{'Mississippi': 4, 'Oklahoma': 5, ..., 'Maine': 2}
```