

# Scripting Languages

Python – algorithm exercises

- caterpillar
- leader



Find the number of contiguous subsequences with the given sum (caterpillar algorithm)

										_	
s=12	4	5	3	5	6	3	3	8	1	9	
	0	1	2	3	4	5	6	7	8	9	
answer=3		12				12		12			

- all numbers are positive integers 0 < n <= 100</li>
- subsequences may overlap
- Read data (as lists) from pickle-files:
  - "caterpillar\_1e1.pkl" ... 1e5, 1e6, 1e7  $\rightarrow$  n = 10, 100000, 1000000, 10000000
- Find out the algorithm time complexity relation between number of elements (n) and execution time t

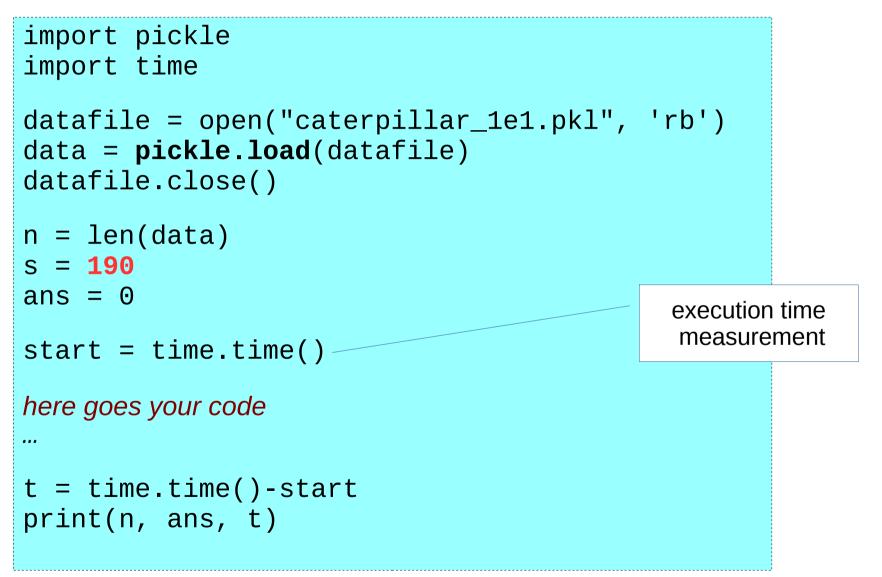
```
Results for s = 190:

1e1 \rightarrow 1, 1e5 \rightarrow 2011, 1e6 \rightarrow 19990, 1e7 \rightarrow 199883

t(n) ~ n
```



### Program structure





Find the leader

- The leader is the candidate whose name occurs more than n/2 times in n-votes
- The list of candidates is unknown
- Sequence may have only one leader or none

candidates =  $4 \rightarrow [3,4,5,6]$ votes = 10leader = "5"

- if the leader exists, the sequence may still contain up to n/2 different candidates
- the sequence of votes may be huge both execution time and memory consumption are important factors



- Solutions:
  - naive: identify all different candidates, count the votes for each candidate separately and confirm its leadership
    - time(n) ~ n<sup>2</sup>
    - mem(n) ~ n
  - better: make the table of votes for all candidates, choose the candidate with highest vote number and confirm its leadership
    - time(n) ~ n
    - mem(n) ~ n
  - good: sort (in place) the table of votes, take the candidate name from the middle position (n/2 ±1), count its votes and confirm its leadership
    - time(n) ~  $n \cdot \log(n)$
    - mem(n) ~ n



#### Find the best solution:

- time(n) ~ n
- mem(n) ~ 1
- read data (as lists) from pickle-files:
  - "leader\_1e3.pkl" ... 1e5, 1e7
     → n = 1000, 100000, 10000000
- candidates: positive integers
- print leader value (or -1 if none)

**Results:** 

 $1e3 \rightarrow leader = 14 (537 \text{ votes})$  $1e5 \rightarrow leader = 37820 (votes 50214)$  $1e7 \rightarrow leader = 4740208 (5001628 \text{ votes})$ 



### Program structure

```
import pickle
import time
datafile = open("leader_1e1.pkl", 'rb')
data = pickle.load(datafile)
datafile.close()
n = len(data)
leader = -1
start = time.time()
here goes your code
...
t = time.time()-start
print(n, leader, t)
```