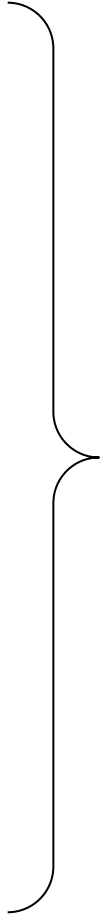


# External Sorting with N-way Merge



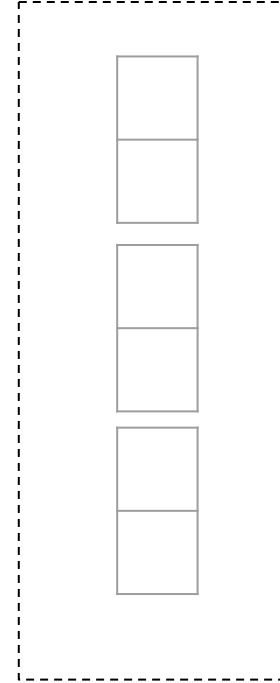
34
82
45
3
78
21
75
16
8
22
56
49



Data File.  
Disk Block size =2 records

34
82
45
3
78
21
75
16
8
22
56
49

Data File.  
Disk Block size = 2 records



Main Memory available  
 $M=3$  (i.e., 3 disk blocks can be stored in Main Memory)

34
82
45
3
78
21
75
16
8
22
56
49

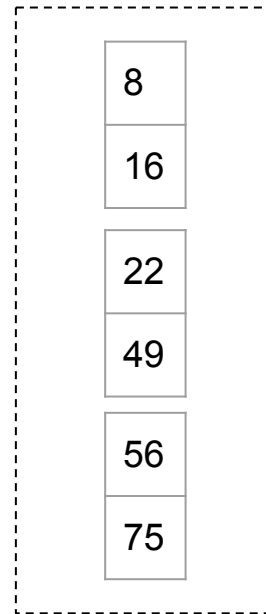
Data File.  
Disk Block size = 2 records

3
21
34
45
75
78

Create Runs (sort 3 disk blocks  
at a time)



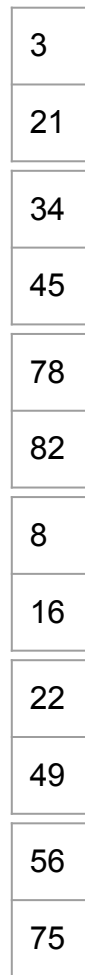
Data File.  
Disk Block size = 2 records



Create Runs (sort 3 disk blocks  
at a time)

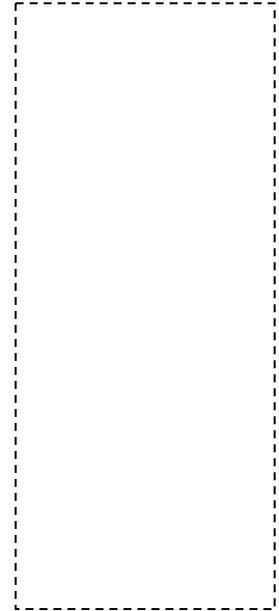


Data File.  
Disk Block size = 2 records



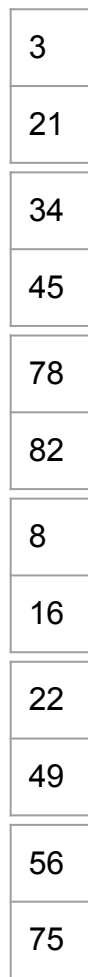
Sorted Run 1

Sorted Run 2



Create Runs (sort 3 disk blocks  
at a time)

**CASE 1:  $M > \#Runs$**

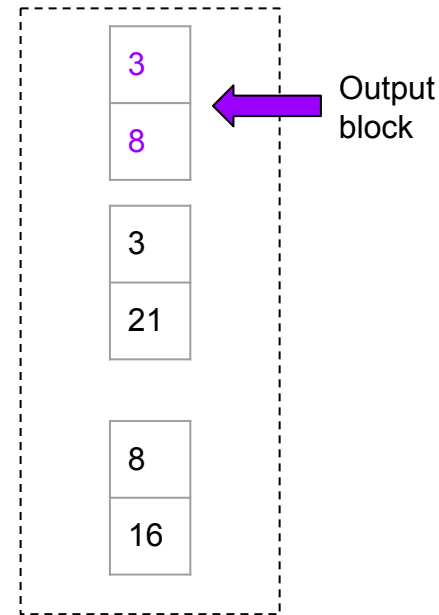


Sorted Run 1

Sorted Run 2

**Given that  $M > \# \text{ runs}$   $\rightarrow$  we can merge  $\text{Min}(\text{runs}, M-1)$  runs in a single shot. We use N-way merge operation**

**CASE:  $M > \# \text{Runs}$**



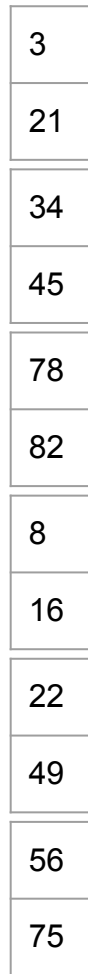
Merge the Runs. Bring in one block at a time from the runs and keep writing the output block.

**A single N-way merge operation is enough. No passes needed**





Data File.  
Disk Block  
size =2  
records

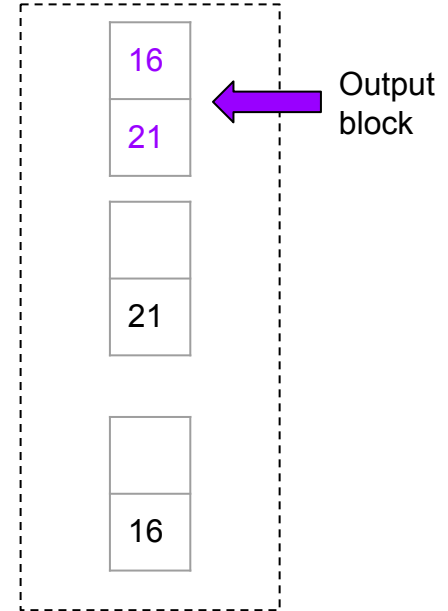


Sorted Run 1

Sorted Run 2



**CASE:  $M > \#Runs$**

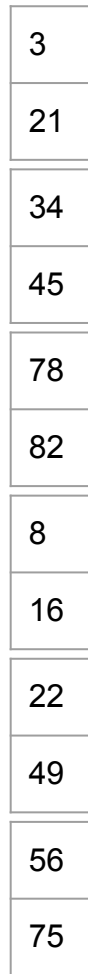


Merge the Runs. Bring in one block at a time from the runs and keep writing the output block.

**A single N-way merge operation is enough. No passes needed**

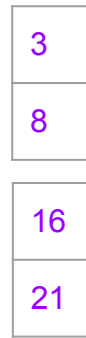


Data File.  
Disk Block  
size =2  
records

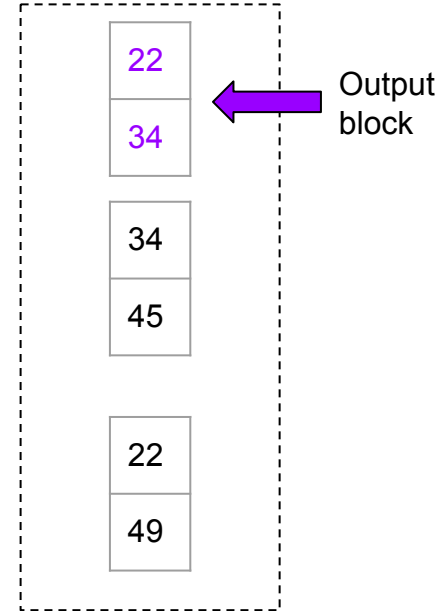


Sorted Run 1

Sorted Run 2



**CASE:  $M > \#Runs$**

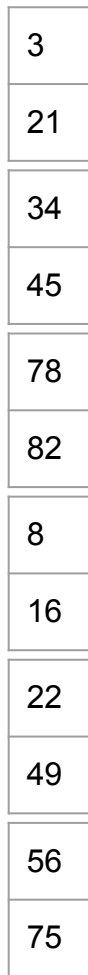


Merge the Runs. Bring in one block at a time from the runs and keep writing the output block.

**A single N-way merge operation is enough. No passes needed**



Data File.  
Disk Block  
size =2  
records

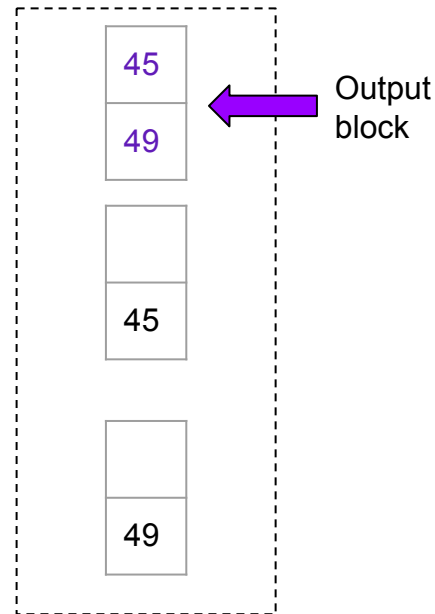


Sorted Run 1

Sorted Run 2



**CASE:  $M > \#Runs$**

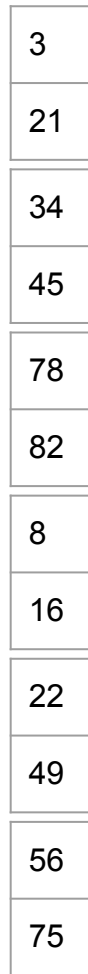


Merge the Runs. Bring in one block at a time from the runs and keep writing the output block.

**A single N-way merge operation is enough. No passes needed**



Data File.  
Disk Block  
size =2  
records

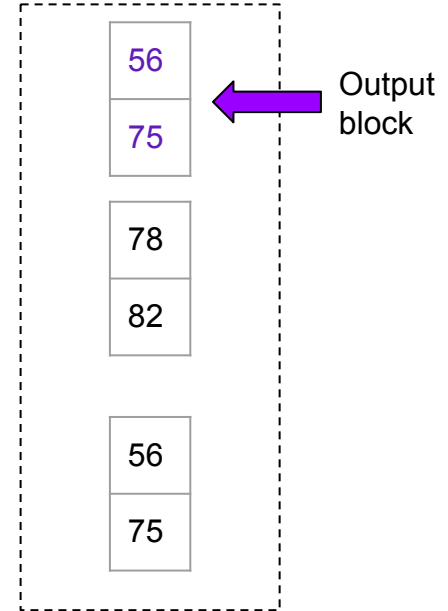


Sorted Run 1

Sorted Run 2



**CASE:  $M > \#Runs$**

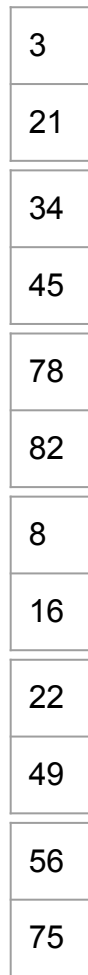


Merge the Runs. Bring in one block at a time from the runs and keep writing the output block.

**A single N-way merge operation is enough. No passes needed**

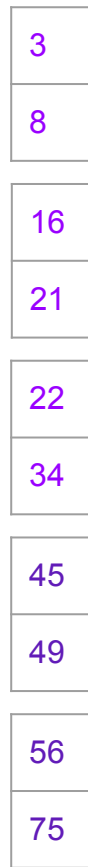


Data File.  
Disk Block  
size =2  
records

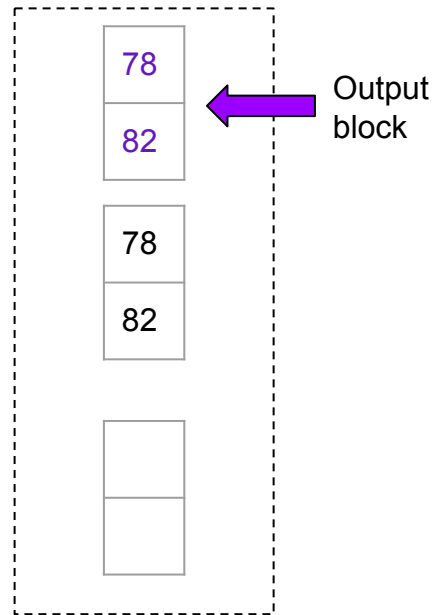


Sorted Run 1

Sorted Run 2



**CASE:  $M > \#Runs$**

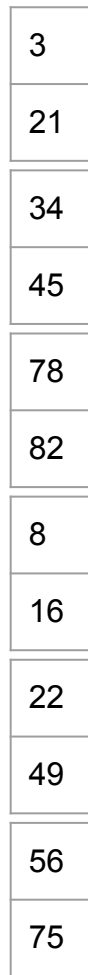


Merge the Runs. Bring in one block at a time from the runs and keep writing the output block.

**A single N-way merge operation is enough. No passes needed**

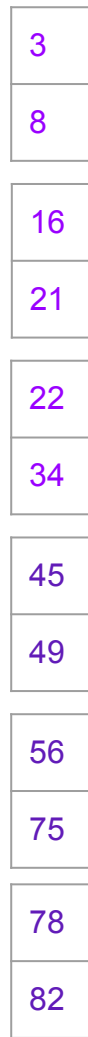


Data File.  
Disk Block  
size =2  
records

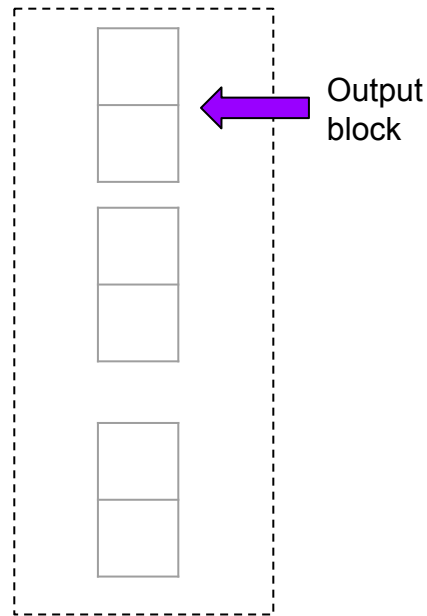


Sorted Run 1

Sorted Run 2



**CASE:  $M > \#Runs$**



Merge the Runs. Bring in one block at a time from the runs and keep writing the output block.

**A single N-way merge operation is enough. No passes needed**

**CASE 1:  $M \leq \#Runs$**

Data File.  
Disk Block  
size = 1  
records

g	24
a	19
d	31
c	33
b	14
e	16
r	16
d	21
m	3
p	2
d	7
a	14

initial  
relation

a	19
d	31
g	24

Main Memory  
size  $M=3$  disk  
blocks

Read 3 disk  
blocks at a time  
and create runs



Data File.  
Disk Block  
size =1  
records

g	24
a	19
d	31
c	33
b	14
e	16
r	16
d	21
m	3
p	2
d	7
a	14

a	19
d	31
g	24

initial  
relation      runs  
create  
runs

b	14
c	33
e	16

Main Memory  
size M=3 disk  
blocks

Read 3 disk  
blocks at a time  
and create runs

Data File.  
Disk Block  
size = 1  
records

g	24
a	19
d	31
c	33
b	14
e	16
r	16
d	21
m	3
p	2
d	7
a	14

a	19
d	31
g	24

b	14
c	33
e	16

initial  
relation      runs  
create  
runs

d	21
m	3
r	16

Main Memory  
size  $M=3$  disk  
blocks

Read 3 disk  
blocks at a time  
and create runs

Data File.  
Disk Block  
size =1  
records

g	24
a	19
d	31
c	33
b	14
e	16
r	16
d	21
m	3
p	2
d	7
a	14

initial  
relation

a	19
d	31
g	24

b	14
c	33
e	16

d	21
m	3
r	16

runs

create  
runs

a	14
d	7
p	2

Main Memory  
size  $M=3$  disk  
blocks

Read 3 disk  
blocks at a time  
and create runs

Data File.  
Disk Block  
size =1  
records

g	24
a	19
d	31
c	33
b	14
e	16
r	16
d	21
m	3
p	2
d	7
a	14

initial  
relation

a	19
d	31
g	24

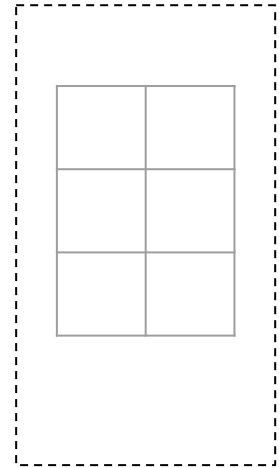
b	14
c	33
e	16

d	21
m	3
r	16

a	14
d	7
p	2

runs

create  
runs



Main Memory  
size  $M=3$  disk  
blocks

In general, we can merge  
 $\text{Min}(\# \text{runs}, M-1)$  runs at a  
time. As  $M=3$  in our case,  
we merge 2 runs at a time.

We merge two runs at a  
time. One block in main  
memory is saved for  
output result. Each Run  
spans across blocks

Data File.  
Disk Block  
size =1  
records

g	24
a	19
d	31
c	33
b	14
e	16
r	16
d	21
m	3
p	2
d	7
a	14

initial  
relation

create  
runs

a	19
d	31
g	24

b	14
c	33
e	16

d	21
m	3
r	16

a	14
d	7
p	2

runs

a	19
a	19
b	14

Output  
block

Main Memory  
size  $M=3$  disk  
blocks

Read one block each  
from each of the two  
runs. Save one block for  
output. Perform Merge  
operation

Data File.  
Disk Block  
size =1  
records

g	24
a	19
d	31
c	33
b	14
e	16
r	16
d	21
m	3
p	2
d	7
a	14

initial  
relation

create  
runs

a	19
d	31
g	24

b	14
c	33
e	16

d	21
m	3
r	16

a	14
d	7
p	2

runs

a	19
---	----

b	14
d	31
b	14

Output  
block

Main Memory  
size  $M=3$  disk  
blocks

Read one block each  
from each of the two  
runs. Save one block for  
output. Perform Merge  
operation

Data File.  
Disk Block  
size =1  
records

g	24
a	19
d	31
c	33
b	14
e	16
r	16
d	21
m	3
p	2
d	7
a	14

initial  
relation

create  
runs

a	19
d	31
g	24

b	14
c	33
e	16

d	21
m	3
r	16

a	14
d	7
p	2

runs

a	19
---	----

b	14
---	----

c	33
d	31
c	33

Output  
block

Main Memory  
size  $M=3$  disk  
blocks

Read one block each  
from each of the two  
runs. Save one block for  
output. Perform Merge  
operation

Data File.  
Disk Block  
size =1  
records

g	24
a	19
d	31
c	33
b	14
e	16
r	16
d	21
m	3
p	2
d	7
a	14

initial  
relation

create  
runs

a	19
d	31
g	24

b	14
c	33
e	16

d	21
m	3
r	16

a	14
d	7
p	2

runs

a	19
---	----

b	14
---	----

c	33
---	----

d	31
d	31
e	16

Output  
block

Main Memory  
size  $M=3$  disk  
blocks

Read one block each  
from each of the two  
runs. Save one block for  
output. Perform Merge  
operation



Data File.  
Disk Block  
size =1  
records

g	24
a	19
d	31
c	33
b	14
e	16
r	16
d	21
m	3
p	2
d	7
a	14

initial  
relation

create  
runs

a	19
d	31
g	24

b	14
c	33
e	16

d	21
m	3
r	16

a	14
d	7
p	2

runs

a	19
---	----

b	14
---	----

c	33
---	----

d	31
---	----

e	16
g	24
e	16

Output  
block

Main Memory  
size  $M=3$  disk  
blocks

Read one block each  
from each of the two  
runs. Save one block for  
output. Perform Merge  
operation

Data File.  
Disk Block  
size =1  
records

g	24
a	19
d	31
c	33
b	14
e	16
r	16
d	21
m	3
p	2
d	7
a	14

initial  
relation

a	19
d	31
g	24

b	14
c	33
e	16

d	21
m	3
r	16

a	14
d	7
p	2

runs

create  
runs

a	19
---	----

b	14
---	----

c	33
---	----

d	31
---	----

e	16
---	----

g	24
g	24
	1

Output  
block

Main Memory  
size  $M=3$  disk  
blocks

Read one block each  
from each of the two  
runs. Save one block for  
output. Perform Merge  
operation

Data File.  
Disk Block  
size =1  
records

g	24
a	19
d	31
c	33
b	14
e	16
r	16
d	21
m	3
p	2
d	7
a	14

initial  
relation

a	19
d	31
g	24

b	14
c	33
e	16

d	21
m	3
r	16

a	14
d	7
p	2

runs

create  
runs

a	19
---	----

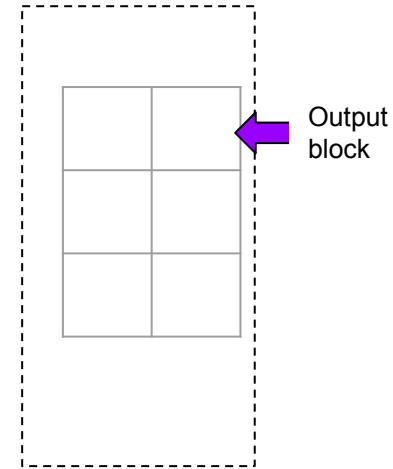
b	14
---	----

c	33
---	----

d	31
---	----

e	16
---	----

g	24
---	----



Main Memory  
size  $M=3$  disk  
blocks

Read one block each  
from each of the two  
runs. Save one block for  
output. Perform Merge  
operation

Data File.  
Disk Block  
size =1  
records

g	24
a	19
d	31
c	33
b	14
e	16
r	16
d	21
m	3
p	2
d	7
a	14

initial  
relation

create  
runs

a	19
d	31
g	24

b	14
c	33
e	16

d	21
m	3
r	16

a	14
d	7
p	2

runs

a	19
---	----

b	14
---	----

c	33
---	----

d	31
---	----

e	16
---	----

g	24
---	----

Merge Run 3 and Run 4  
as well

d	21
a	14

Output  
block

Main Memory  
size  $M=3$  disk  
blocks

Read one block each  
from each of the two  
runs. Save one block for  
output. Perform Merge  
operation

Data File.  
Disk Block  
size =1  
records

g	24
a	19
d	31
c	33
b	14
e	16
r	16
d	21
m	3
p	2
d	7
a	14

initial  
relation

a	19
d	31
g	24

b	14
c	33
e	16

d	21
m	3
r	16

a	14
d	7
p	2

runs

create  
runs

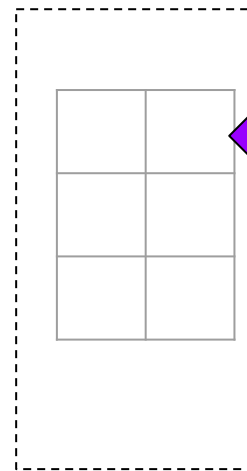
a	19
b	14
c	33
d	31
e	16
g	24

a	14
d	7
d	21
m	3
p	2
r	16

runs

merge  
pass-1

merge  
pass-2



Output  
block

Main Memory  
size  $M=3$  disk  
blocks

Read one block each  
from each of the two  
runs. Save one block for  
output. Perform Merge  
operation

Data File.  
Disk Block  
size =1  
records

g	24
a	19
d	31
c	33
b	14
e	16
r	16
d	21
m	3
p	2
d	7
a	14

initial  
relation

a	19
d	31
g	24

b	14
c	33
e	16

d	21
m	3
r	16

a	14
d	7
p	2

runs

create  
runs

a	19
b	14
c	33
d	31
e	16
g	24

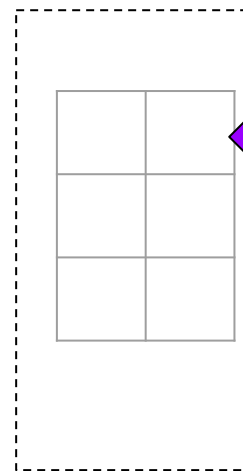
a	14
d	7
d	21
m	3
p	2
r	16

runs

merge  
pass-1

merge  
pass-2

Merge these  
runs which  
got created in  
the previous  
pass



Main Memory  
size  $M=3$  disk  
blocks

Read one block each  
from each of the two  
runs. Save one block for  
output. Perform Merge  
operation

Data File.  
Disk Block  
size =1  
records

g	24
a	19
d	31
c	33
b	14
e	16
r	16
d	21
m	3
p	2
d	7
a	14

initial  
relation

a	19
d	31
g	24

b	14
c	33
e	16

d	21
m	3
r	16

a	14
d	7
p	2

runs

create  
runs

a	19
b	14
c	33
d	31
e	16
g	24

a	14
d	7
d	21
m	3
p	2
r	16

runs

merge  
pass-1

merge  
pass-2

a	14
a	19
b	14
c	33
d	7
d	21
d	31
e	16
g	24
m	3
p	2
r	16

sorted  
output

In this example, ties were  
broken according to the  
values in the second field.  
However, such things  
depend on the  
specification. Algorithm is  
independent of such  
aspects