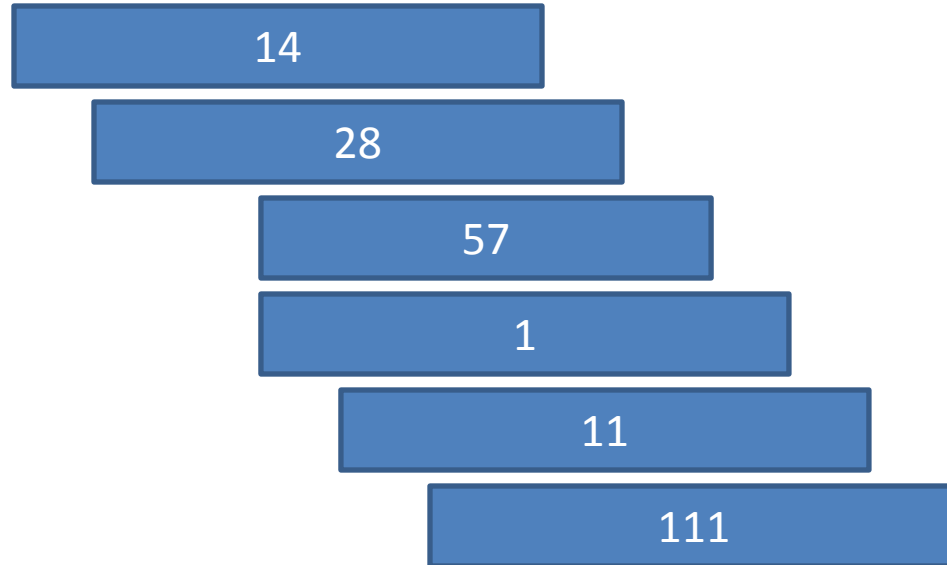


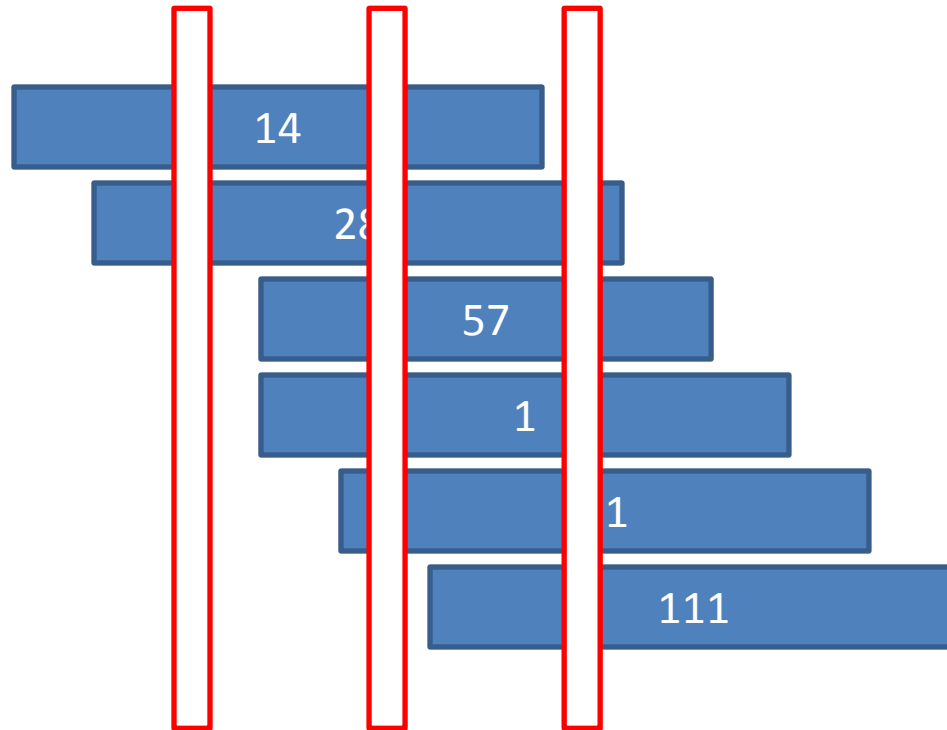
GSBS Bootstrappers: Bedtools Workshop #4 Map

Michael Purcaro
November 4, 2015

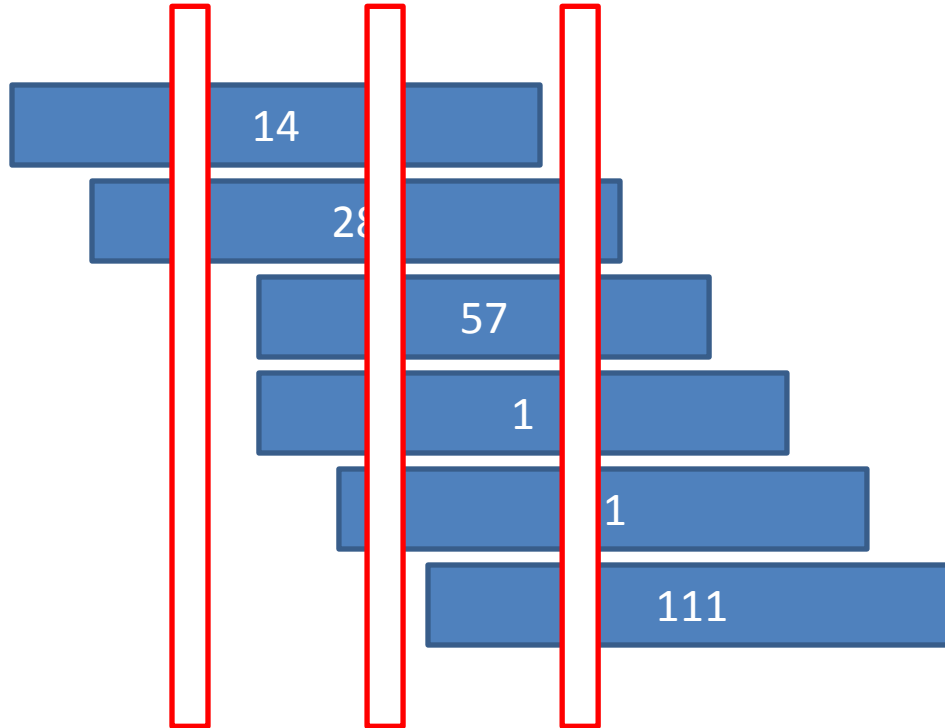
Bedtools: Map



Bedtools: Map

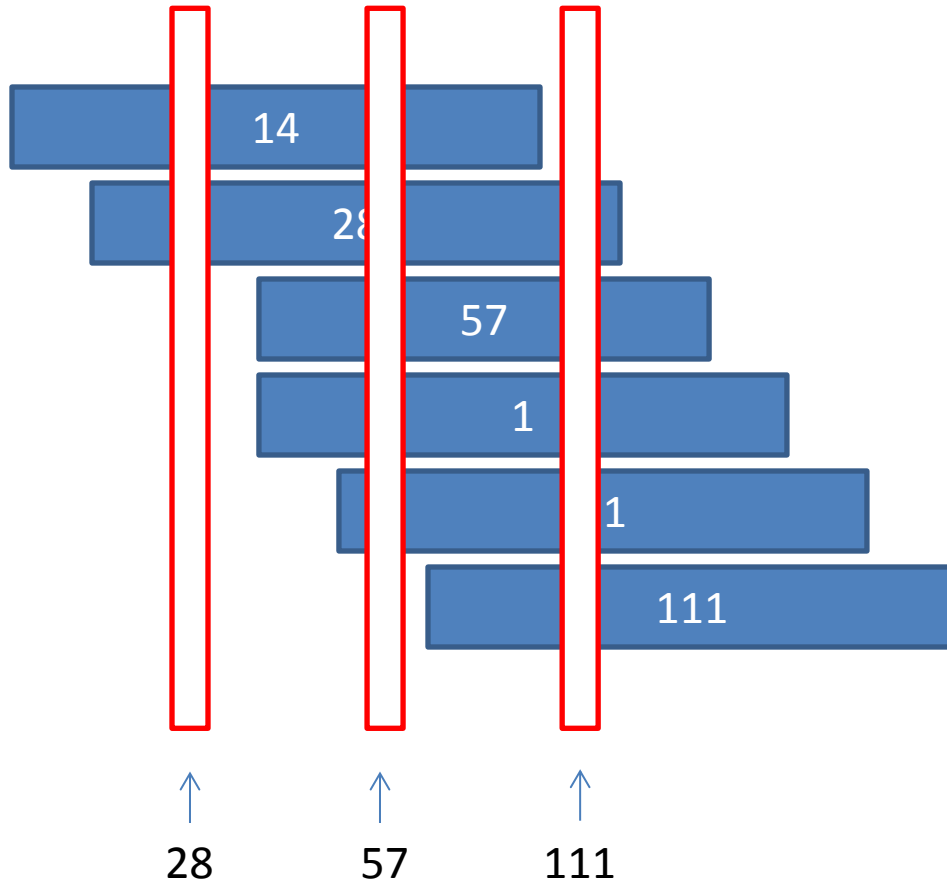


Bedtools: Map



- sum** - *numeric only*
- count** - *numeric or text*
- count_distinct** - *numeric or text*
- min** - *numeric only*
- max** - *numeric only*
- absmin** - *numeric only*
- absmax** - *numeric only*
- mean** - *numeric only*
- median** - *numeric only*
- antimode** - *numeric or text*
- collapse** - *numeric or text*
- distinct** - *numeric or text*
- concat** - *numeric or text*

Bedtools: Map



sum - numeric only
count - numeric or text
count_distinct - numeric or text
min - numeric only
max - numeric only
absmin - numeric only
absmax - numeric only
mean - numeric only
median - numeric only
antimode - numeric or text
collapse - numeric or text
distinct - numeric or text
concat - numeric or text

Bedtools: Map

- bedtools map
 - a regionsOfInterest.bed
 - b data.bed
 - c column # in data.bed to use
 - o function to use (mean, max, etc.)

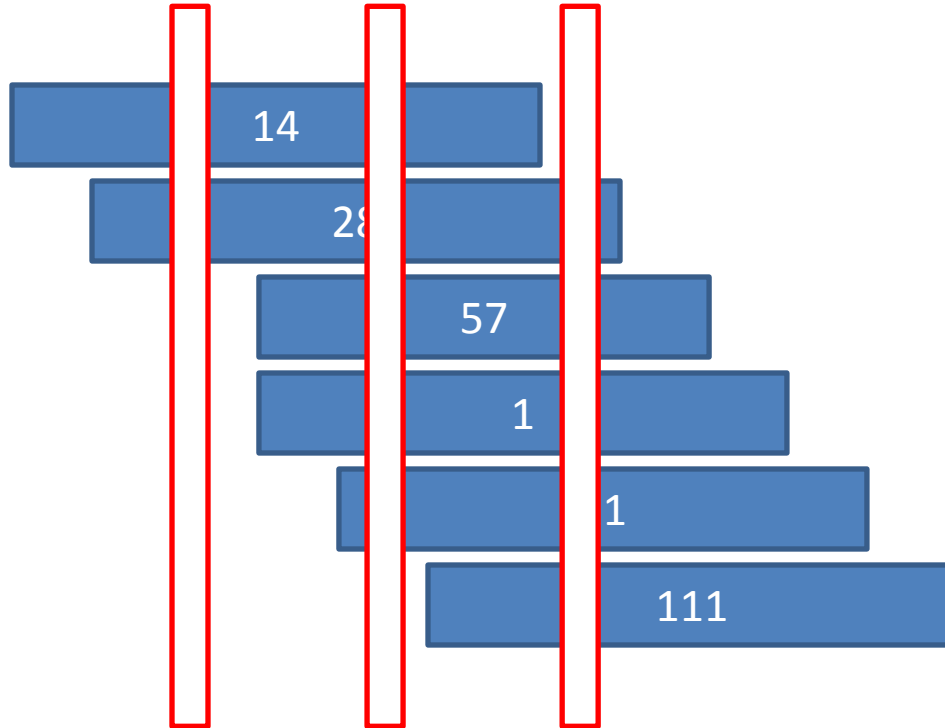
Exercise 1

Find mean/median/min/max values for columns 5 and 6 from “fimo.bed” for the following regions of interest:

chr7 10000 11000

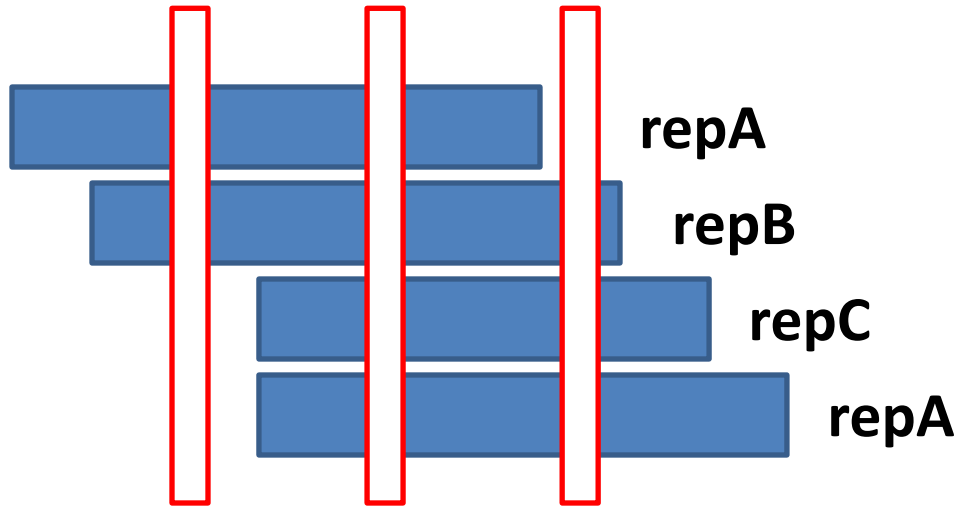
chr7 25000 30000

Bedtools: Map



- sum** - *numeric only*
- count** - *numeric or text*
- count_distinct** - *numeric or text*
- min** - *numeric only*
- max** - *numeric only*
- absmin** - *numeric only*
- absmax** - *numeric only*
- mean** - *numeric only*
- median** - *numeric only*
- antimode** - *numeric or text*
- collapse** - *numeric or text*
- distinct** - *numeric or text*
- concat** - *numeric or text*

Bedtools: Map



```
chr7 10000 11000 . repA
chr7 10200 11200 . repB
chr7 10400 11400 . repC
chr7 10400 11600 . repA
```

Bedtools: Map

```
chr7 10000 11000 . repA
chr7 10200 11200 . repB
chr7 10400 11400 . repC
chr7 10400 11600 . repA
```

```
$ bedtools map -a fimo.regions_of_interest.bed
               -b text.bed
               -o collapse
```

```
chr7 10000 11000 repA,repB,repC,repA
chr7 25000 30000 .
```

Bedtools: Map

```
chr7 10000 11000 . repA
chr7 10200 11200 . repB
chr7 10400 11400 . repC
chr7 10400 11600 . repA
```

```
$ bedtools map -a fimo.regions_of_interest.bed
               -b text.bed
               -o distinct
```

```
chr7 10000 11000 repA, repB, repC
chr7 25000 30000 .
```

Exercise 2

The strongest DNase-seq peaks across 80 different cell types and developmental stages were selected and placed in a file “master_peaks.bed”.

For each master peak, list which other peaks (from the “all_peaks.bed” file) are present at the peak.

Exercise 2

Master peaks:

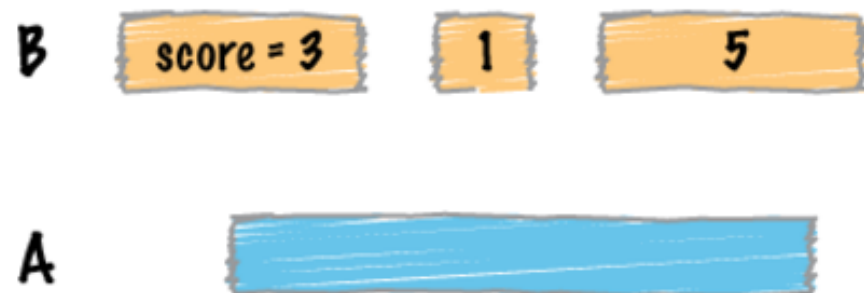
```
chr1 3002720 3002870 4.37932 ENCFF001YVC  
chr1 3025600 3025750 8.0632 ENCFF001YLB
```

All peaks:

```
chr1 3002720 3002870 1921300-ENCFF001YVC-4.37932  
chr1 3002740 3002890 4406931-ENCFF001YNU-6.60995  
chr1 3002740 3002890 562121-ENCFF001YPN-19.466
```



map



$$B_{\text{score}} \text{ map } A \text{ (mean)} = \text{mean}(3, 1, 5) = 3$$

$$B_{\text{score}} \text{ map } A \text{ (max)} = \text{max}(3, 1, 5) = 5$$

Table Of Contents

map

Usage and option summary

Default behavior - compute

mean Compute the mean of

collapse List each value of

distinct List each *unique*

-s Only include intervals that

-S Only include intervals that

Multiple operations and col