

Practical Outline 3

```
library(tidyverse)
library(IRanges)

ir <- IRanges(c(4, 5, 8, 15, 19, 28, 40), width=c(15, 6, 7, 12, 9, 3, 6))
names(ir) = letters[1:7]
mcols(ir) <- mtcars[1:7, 1:3]
ir
```

```
## IRanges object with 7 ranges and 3 metadata columns:
##      start      end      width |      mpg      cyl      disp
##      <integer> <integer> <integer> | <numeric> <numeric> <numeric>
##  a         4        18        15 |      21.0         6      160
##  b         5        10         6 |      21.0         6      160
##  c         8        14         7 |      22.8         4      108
##  d        15        26        12 |      21.4         6      258
##  e        19        27         9 |      18.7         8      360
##  f        28        30         3 |      18.1         6      225
##  g        40        45         6 |      14.3         8      360
```

Exercise K

What is the mean `disp` for ranges whose `start` is less than 20? **209.2**

```
mean(mcols(ir[start(ir) < 20])$disp)
```

```
## [1] 209.2
```

What is the maximum width for ranges whose `cyl` is 6? **15**

```
max(width(ir[mcols(ir)$cyl==6]))
```

```
## [1] 15
```

Exercise L

```
load('./data/data_practical_3.RData')
```

How many overlaps are there between `ir1` and `ir2`? **26**

```
nolaps <- countOverlaps(ir2, ir1)
sum(nolaps)
```

```
## [1] 26
```

What is the name of the range in `ir2` with the most overlaps with `ir1`? **r has 6 overlaps**

```
nolaps[nolaps == max(nolaps)]
```

```
## r
## 6
```

Subset `ir2` to show only ranges that have more than 2 overlaps with `ir1`

```
mcols(ir2)$Overlaps <- nolaps
ir2[mcols(ir2)$Overlaps > 2]
```

```
## IRanges object with 4 ranges and 1 metadata column:
```

```
##      start      end      width | Overlaps
##      <integer> <integer> <integer> | <integer>
##  q         11         15         5 |         3
##  r         19         25         7 |         6
##  t         27         35         9 |         3
##  w         36         48        13 |         4
```

```
library(Homo.sapiens)
hg <- genes(Homo.sapiens)
hg
```

```
## GRanges object with 23056 ranges and 1 metadata column:
```

```
##      seqnames      ranges strand |      GENEID
##      <Rle>      <IRanges> <Rle> | <CharacterList>
##      1      chr19 58858172-58874214 - |          1
##     10      chr8 18248755-18258723  + |         10
##    100     chr20 43248163-43280376  - |        100
##   1000     chr18 25530930-25757445  - |       1000
##  10000      chr1 243651535-244006886 - |      10000
##     ...      ...      ...      ... |      ...
##   9991      chr9 114979995-115095944 - |      9991
##   9992     chr21 35736323-35743440  + |      9992
##   9993     chr22 19023795-19109967  - |      9993
##   9994      chr6 90539619-90584155  + |      9994
##   9997     chr22 50961997-50964905  - |      9997
##  -----
##  seqinfo: 93 sequences (1 circular) from hg19 genome
```

Exercise M

```
hg <- keepStandardChromosomes(hg, pruning.mode="coarse")
```

How many base pairs does chromosomes 15 span? **102.5 million**

```
seqinfo(hg)["chr15"]
```

```
## Seqinfo object with 1 sequence from hg19 genome:
```

```
##  seqnames seqlengths isCircular genome
##  chr15      102531392      NA      hg19
```

Can you make a table of the number of genes in all of the standard human chromosomes?

```
table(seqnames(hg))
```

```
##
## chr1 chr2 chr3 chr4 chr5 chr6 chr7 chr8 chr9 chr10 chr11 chr12 chr13
## 2326 1464 1271 873 1022 1000 1108 818 945 903 1439 1173 449
## chr14 chr15 chr16 chr17 chr18 chr19 chr20 chr21 chr22 chrX chrY chrM
## 779 791 938 1337 341 1607 647 296 535 918 53 0
```
