bedtools

> A swiss-army knife of tools for genomic-analysis tasks.
> Used to intersect, group, convert and count data in BAM, BED, GFF/GTF, VCF format.

- Intersect two files with respect to the sequences' strand and save the result to {{path/to/output_file}}:

```
bedtools intersect -a {{path/to/file_1}} -b {{path/to/file_2}} -s > {{path/to/output_file}}
```

- Intersect two files with a left outer join, i.e. report each feature from {{file_1}} and NULL if no overlap with {{file_2}}:

```
bedtools intersect -a {{path/to/file_1}} -b {{path/to/file_2}} -lof > {{path/to/output_file}}
```

- Using more efficient algorithm to intersect two pre-sorted files:

```
bedtools intersect -a {{path/to/file_1}} -b {{path/to/file_2}} -sorted > {{path/to/output_file}}
```

- Group file {{path/to/file}} based on the first three and the fifth column and summarize the sixth column by summing it up:

```
bedtools groupby -i {{path/to/file}} -c 1-3,5 -g 6 -o sum
```

- Convert bam-formated file to a bed-formated one:

```
bedtools bamtobed -i {{path/to/file}}.bam > {{path/to/file}}.bed
```

- Find for all features in {{file_1}}.bed the closest one in {{file_2}}.bed and write their distance in an extra column (input files must be sorted):

```
bedtools closest -a {{path/to/file_1}}.bed -b {{path/to/file_2}}.bed -d
```