Architektura dużych projektów bioinformatycznych. Zadanie 1.

**bedGraph type**: The [bedGraph](http://genome.ucsc.edu/goldenpath/help/bedgraph.html) format is also an older format used to display [sparse](http://genome.ucsc.edu/goldenpath/help/wiggle.html#sparse) data or data that contains elements of varying size.

description:

first line – single line, all parameters separated by spaces

**track type=bedGraph name=***track\_label* **description=***center\_label*

**visibility=***display\_mode* **color=***r,g,b* **altColor=***r,g,b*

**priority=***priority* **autoScale=***on|off* **alwaysZero=***on|off*

**gridDefault=***on|off* **maxHeightPixels=***max:default:min*

**graphType=***bar|points* **viewLimits=***lower:upper*

**yLineMark=***real-value* **yLineOnOff=***on|off*

**windowingFunction=***maximum|mean|minimum* **smoothingWindow=**off|2-16

following track line data in 4 columns:

*chromA* *chromStartA* *chromEndA* *dataValueA*

*chromB* *chromStartB* *chromEndB* *dataValueB*

Bedgraph track data values can be integer or real, positive or negative values. Chromosome positions are specified as 0-relative. The first chromosome position is 0. The last position in a chromosome of length N would be N - 1. Only positions specified have data

**Wig format**: The wiggle (WIG) format is an older format for display of dense, continuous data such as GC percent, probability scores, and transcriptome data. Wiggle data elements must be equally sized

**the first line must be a**[**track definition line**](http://genome.ucsc.edu/goldenpath/help/customTrack.html#_blank)**(*i.e.*, track type=wiggle\_0)**, which designates the track as a wiggle track and adds a number of options for controlling the default display.

Wiggle format is composed of declaration lines and data lines, and require a separate wiggle track definition line. There are two options for formatting wiggle data: **variableStep** and**fixedStep**. These formats were developed to allow the file to be written as compactly as possible.

* **variableStep** is for data with irregular intervals between new data points and is the more commonly used wiggle format. After the wiggle track definition line, variableStep begins with a declaration line and is followed by two columns containing chromosome positions and data values:
* **variableStep** **chrom=***chrN* **[span=***windowSize***]**
* *chromStartA* *dataValueA*
* *chromStartB* *dataValueB*

*... etc ...* *... etc ...*

The declaration line starts with the word **variableStep** and is followed by a specification for a chromosome. The optional **span** parameter (default: span=1) allows data composed of contiguous runs of bases with the same data value to be specified more succinctly. The span begins at each chromosome position specified and indicates the number of bases that data value should cover. For example, this variableStep specification:

variableStep chrom=chr2

300701 12.5

300702 12.5

300703 12.5

300704 12.5

300705 12.5

is equivalent to:

variableStep chrom=chr2 span=5

300701 12.5

Both versions display a value of 12.5 at position 300701-300705 on chromosome 2

|  |  |
| --- | --- |
| Converter | |
| type = bedGraph | type = wiggle\_0 |
| Other parameters from track line should be unchanged | |
| In wig file there is only one track line a no comments line | |
|  | fixedStep and variable Step !!!! |
| Chromosome positions are specified as 0-relative. The first chromosome position is 0. The last position in a chromosome of length N would be N - 1 | Chromosome positions are specified as 1-relative  For a chromosome of length N, the first position is 1 and the last position is N. |
| One parameter – path to input file | |