

# atlas-utils之根据分组信息处理OTU表： group\_by

## 一、atlas-utils group\_by介绍

功能描述：

`atlas-utils group_by` 命令可以根据分组信息对 `zotu` 表进行汇总处理, 比如将同一个分组的 `zotu` 表的数值求平均,或者加和处理。

命令行接口：

```
1 Usage: atlas-utils group_by <tsv> <map>
2
3 Options:
4   -t STR operation. 1:sum, 2: average.
5                        default: [1];
6   -r round value to nearest intege;
```

可选参数：

```
1   -t 字符串    操作，1：总和，2：平均值，默认为 1；
2   -r          将数值转四舍五入为整数；
```

## 二、使用场景实例及其用法

示例演示：

示例文件： `zotu_table.txt` , `mapping_file.txt`

```
1 $ cat zotu_table.txt | head
```

1	#OTU ID	A-1	A-2	B-1	B-2	C-1	C-2
2	ZOTU_1	0	0	87	278	1829	3608
3	ZOTU_2	223	447	1268	1583	52	69
4	ZOTU_3	0	0	162	159	1116	2021
5	ZOTU_4	0	0	99	50	1250	2172
6	ZOTU_5	0	0	1	8	1216	2143
7	ZOTU_6	1	2	9	16	1155	1938
8	ZOTU_7	0	0	73	41	1044	1883
9	ZOTU_8	0	0	1353	1231	2	1
10	ZOTU_9	0	0	1172	1022	1	2

```
1 $ cat mapping_file.txt | head
```

1	#SampleID	Description
2	A-1	A
3	A-2	A
4	B-1	B
5	B-2	B
6	C-1	C
7	C-2	C

运行命令：

使用默认参数对分组求和

```
1 | $ atlas-utils group_by zotu_table.txt mapping_file.txt | head
```

1	#OTU ID	A	B	C
2	ZOTU_1	0	365	5437
3	ZOTU_2	670	2851	121
4	ZOTU_3	0	321	3137
5	ZOTU_4	0	149	3422
6	ZOTU_5	0	9	3359
7	ZOTU_6	3	25	3093
8	ZOTU_7	0	114	2927
9	ZOTU_8	0	2584	3
10	ZOTU_9	0	2194	3

选项参数1: 使用 `-t`，设置进行分组求平均。

```
1 | $ atlas-utils group_by -t 2 zotu_table.txt mapping_file.txt | head
```

1	#OTU ID	A	B	C
2	ZOTU_1	0	182.5	2718
3	ZOTU_2	335	1426	60.5
4	ZOTU_3	0	160.5	1568
5	ZOTU_4	0	74.5	1711
6	ZOTU_5	0	4.5	1680
7	ZOTU_6	1.5	12.5	1546
8	ZOTU_7	0	57	1464
9	ZOTU_8	0	1292	1.5
10	ZOTU_9	0	1097	1.5

选项参数2: 使用 `-r`，将数值转四舍五入为整数。

```
1 | $ atlas-utils group_by -t 2 -r zotu_table.txt mapping_file.txt | head
```

1	#OTU	ID	A	B	C
2	ZOTU_1	0		183	2719
3	ZOTU_2	335		1426	61
4	ZOTU_3	0		161	1569
5	ZOTU_4	0		75	1711
6	ZOTU_5	0		5	1680
7	ZOTU_6	2		13	1547
8	ZOTU_7	0		57	1464
9	ZOTU_8	0		1292	2
10	ZOTU_9	0		1097	2

本文材料为 **BASE (Biostack Applied bioinformatic SEies )** 课程 **Linux Command Line Tools for Life Scientists** 材料，版权归 上海逻捷信息科技有限公司 所有。

Last Update: 2020-09-11 11:56 AM