

tsv-utils之调整矩阵形状: reshape

一、tsv-utils reshape介绍

功能描述:

`tsv-utils reshape` 根据 `key/value` 对, 合并属于相同 `value` 的字符串。

命令行接口:

```
1 $ tsv-utils reshape
2
3 Usage: tsv-utils reshape <tsv> <map>
```

二、使用场景实例及其用法

经典使用案例

1. 16S 数据分析, 将同一分组的特征归并到一行, 选择特征进行数据可视化

示例演示

示例文件: `genus.freqs.txt`, `maps.txt`。

```
1 $ cat genus.freqs.txt | head -n 6
```

| | | | | | | | | |
|---|-------------------|-----------|----------|-----|-----------|----------|-----------|----------|
| 1 | #level | A-1 | A-2 | B-1 | B-2 | C-1 | C-2 | |
| 2 | Parvibaculum | 0.0003635 | | | 0.0002873 | | 5.318e-05 | 0 |
| | | 0.004845 | 0.004367 | | | | | |
| 3 | Agromyces | 0 | 0 | 0 | 0 | 0.009585 | | 0.0114 |
| 4 | Pigmentiphaga | 0.0001091 | | | 0.0001642 | | 0.002765 | 0.001841 |
| | | 5.209e-05 | 0 | | | | | |
| 5 | Mangrovibacterium | | 0 | 0 | | 0.002499 | 0.001785 | 0 |
| | | 0 | | | | | | |
| 6 | Pseudonocardia | 0.00189 | 0.001683 | | 0 | 0 | 0 | 0 |

```
1 $ cat mapping_file.txt
```

| | | |
|---|-----------|-------------|
| 1 | #SampleID | Description |
| 2 | A-1 | A |
| 3 | A-2 | A |
| 4 | B-1 | B |
| 5 | C-2 | C |
| 6 | B-2 | B |
| 7 | C-1 | C |
| 8 | C-1 | C |
| 9 | C-1 | C |

运行命令:

```
1 | $ tsv-utils reshape genus.freqs.txt mapping_file.txt | head -n6
```

| | | | | |
|---|--------------|---|-----------|-----------|
| 1 | Parvibaculum | A | 0.0003635 | 0.0002873 |
| 2 | Parvibaculum | B | 5.318e-05 | 0 |
| 3 | Parvibaculum | C | 0.004845 | 0.004367 |
| 4 | Agromyces | A | 0 | 0 |
| 5 | Agromyces | B | 0 | 0 |
| 6 | Agromyces | C | 0.009585 | 0.0114 |

注意事项: 该操作保持特征信息不变, 汇总统一分组元素.

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: 2020-08-10 11:56 AM