

Excel中OFFSET函数（二）——“一个矩阵转换为三列数据”

原创 孤独的马铃薯 于 2019-09-26 00:21:31 发布 阅读量8.4k 收藏 26 点赞数 6

版权

分类专栏：[数据预处理](#) [# Excel](#)



数据预处理 同时被 2 个专栏收录 ▾

0 订阅 4 篇文章

[订阅专栏](#)

在我之前写的“Excel中OFFSET函数（一）——“实现矩阵转置””，已经简单介绍了OFFSET函数。阅读本篇文章中如有任何不解之处，请参考上篇文章。

以下是中国2018年8月~2019年7月的各省CPI数据，由于我要利用该SPSS软件进行统计，需要将其化为三列很多行的数据（行数=省市数*时间数，一行表示一个对应关系），由此引出了我对excel中OFFSET函数的学习。

原数据排列格式：

| 时间 | 北京市 | 天津市 | 河北省 | 山西省 | 内蒙古自治区 | 辽宁省 | 吉林省 | 黑龙江省 | 上海市 | 江苏省 |
|----------|-------|-------|-------|-------|--------|-------|-------|-------|-------|-------|
| 2019年7月 | 102.6 | 102.3 | 102.9 | 103 | 102.6 | 102.4 | 103.3 | 103.1 | 102.8 | 102.9 |
| 2019年6月 | 102.6 | 102.2 | 103.2 | 102.8 | 102.3 | 102.4 | 103.3 | 102.8 | 102.7 | 102.9 |
| 2019年5月 | 101.8 | 102.3 | 103.3 | 102.9 | 102.4 | 102.2 | 102.9 | 102.5 | 102.6 | 103 |
| 2019年4月 | 101.6 | 102.9 | 102.9 | 102.7 | 102.3 | 101.9 | 102.7 | 102.6 | 102.4 | 102.7 |
| 2019年3月 | 101.9 | 102.4 | 102.5 | 102.5 | 102 | 101.6 | 102.3 | 101.9 | 102.2 | 102.5 |
| 2019年2月 | 101.6 | 101.8 | 101.7 | 101.6 | 101.4 | 100.8 | 101.5 | 101.2 | 101.7 | 101.8 |
| 2019年1月 | 101.9 | 101.7 | 101.9 | 101.8 | 101.6 | 100.8 | 101.4 | 101 | 101.1 | 102 |
| 2018年12月 | 102 | 101.8 | 102 | 101.9 | 101.6 | 101.3 | 101.7 | 101.2 | 101.2 | 102.2 |
| 2018年11月 | 102.4 | 102 | 102.4 | 102.4 | 101.9 | 101.5 | 102.1 | 101.6 | 101.9 | 102.6 |
| 2018年10月 | 103 | 102.2 | 102.7 | 102.9 | 102.3 | 101.8 | 102.5 | 102 | 102.2 | 102.9 |
| 2018年9月 | 102.8 | 102.3 | 103 | 102.8 | 102.2 | 102 | 102.5 | 101.8 | 101.5 | 103 |
| 2018年8月 | 102.9 | 102.2 | 102.9 | 101.8 | 101.8 | 103.4 | 102.9 | 101.7 | 101.8 | 102.7 |

整理后数据排列格式：



孤独的马铃薯

[关注](#)

👍 6



★ 26

💬 1



[专栏目录](#)

| | | |
|-----|----------|-------|
| 北京市 | 2018年8月 | 102.9 |
| 北京市 | 2018年9月 | 102.8 |
| 北京市 | 2018年10月 | 103 |
| 北京市 | 2018年11月 | 102.4 |
| 北京市 | 2018年12月 | 102 |
| 北京市 | 2019年1月 | 101.9 |
| 北京市 | 2019年2月 | 101.6 |
| 北京市 | 2019年3月 | 101.9 |
| 北京市 | 2019年4月 | 101.6 |
| 北京市 | 2019年5月 | 101.8 |
| 北京市 | 2019年6月 | 102.6 |
| 北京市 | 2019年7月 | 102.3 |
| 天津市 | 2018年8月 | 102.2 |
| 天津市 | 2018年9月 | 102.3 |
| 天津市 | 2018年10月 | 102.2 |
| 天津市 | 2018年11月 | 102 |
| 天津市 | 2018年12月 | 101.8 |
| 天津市 | 2019年1月 | 101.7 |
| 天津市 | 2019年2月 | 101.8 |
| 天津市 | 2019年3月 | 102.4 |
| 天津市 | 2019年4月 | 102.9 |
| 天津市 | 2019年5月 | 102.3 |
| 天津市 | 2019年6月 | 102.2 |
| 天津市 | 2019年7月 | 102.9 |
| 河北省 | 2018年8月 | 102.9 |

1.省市每个都纵向输出12次(因为共有12个时间点)

这个操作有点类似于我们之前在“矩阵转置”中的“行转列”。上篇文章中的“行转列”——`Offset(A1,0,row(A1)-1)`，主要利用了ROW()函数，将下拉操作造成的行数的变化反应到了offset()函数中右移量cols的变化。

在现在面对的这个问题中，我们依然需要利用row()函数的这一特性。但在此基础上我们还需要增加一些其他的操作，以满足连续输出12次的要求。

纵向连续输出12次同一个目标格内的内容，只需要满足下面的条件：

当下拉引起的行标变化范围在1~12时——offset函数中右移量cols保持不变

简单



孤独的马铃薯

关注

6



26



1



专栏目录

即，使得第一个12都是北京市，第2个12都是天津市，.....，第i个12都是**省(市)。

找到规律了吧？使用除法“/”嘛。分子，分母又该如何确定呢？

分母当然毫无疑问地选择12嘛，不解释。

分子呢？

首先，分子里肯定要用到row函数的，理由看上面。那么需要确定是就是“row(A?)”中？部分的值（因为row返回行数，列标是A或B.....都无所谓）。

我们需要知道的是：当“/”放入函数中时，遇到无法整除的情况，其返回的是所得商的整数部分。例如：row(A1)/12 = 0; row(A11)/12 = 0; row(12)/12=1.....（类似于C语言中的int类型做除法）

从上面的例子中，我们可以看到，如果我们使用row(A1)开始，那么一直到row(A11)都是0。可以将起点设在“北京市”，这样第一次输出的就是“北京市”，但是输出的次数却是11次。这是因为在Excel中没有A0，使得坐标的行标不可能从0开始递增。

因此我们选择以A1作为起点，以row(A12)/12作为offset的右移参数cols。对应的函数为：=OFFSET(\$A\$1,0,ROW(A12)/12)

| | A | B | C | D | E | F |
|----|----------|-------|-------|-------|-------|-------|
| 2 | 2019年7月 | 102.6 | 102.3 | 102.9 | 103 | 102.4 |
| 3 | 2019年6月 | 102.6 | 102.2 | 103.2 | 102.8 | 102.7 |
| 4 | 2019年5月 | 101.8 | 102.3 | 103.3 | 102.9 | 102.4 |
| 5 | 2019年4月 | 101.6 | 102.9 | 102.9 | 102.7 | 102.3 |
| 6 | 2019年3月 | 101.9 | 102.4 | 102.5 | 102.5 | 102.4 |
| 7 | 2019年2月 | 101.6 | 101.8 | 101.7 | 101.6 | 101.4 |
| 8 | 2019年1月 | 101.9 | 101.7 | 101.9 | 101.8 | 101.6 |
| 9 | 2018年12月 | 102 | 101.8 | 102 | 101.9 | 101.6 |
| 10 | 2018年11月 | 102.4 | 102 | 102.4 | 102.4 | 101.9 |
| 11 | 2018年10月 | 103 | 102.2 | 102.7 | 102.9 | 102.3 |
| 12 | 2018年9月 | 102.8 | 102.3 | 103 | 102.8 | 102.3 |
| 13 | 2018年8月 | 102.9 | 102.2 | 102.9 | 101.8 | 101.8 |
| 14 | | | | | | |
| 15 | | | | | | |
| 16 | | | | | | |
| 17 | | | 北京市 | | | |
| 18 | | | 北京市 | | | |
| 19 | | | 北京市 | | | |
| 20 | | | 北京市 | | | |
| 21 | | | 北京市 | | | |
| 22 | | | 北京市 | | | |
| 23 | | | 北京市 | | | |
| 24 | | | 北京市 | | | |
| 25 | | | 北京市 | | | |
| 26 | | | 北京市 | | | |
| 27 | | | 北京市 | | | |
| 28 | | | 北京市 | | | |
| 29 | | | 天津市 | | | |



孤独的马铃薯

关注

6



26

1



专栏目录

小结:

因此, 想让某点(以A1为例)及其右方的数据, 在指定单元格 (以C17为例) 下方每个都输出n次, 只需要在指定的单元格内写入函数:

OFFSET(\$A\$1,0,ROW(An)/n)

里面的n自己代入值, A1自己选, 指定单元格C17我也不管。(*.*)

2. 在一个省市的表(12行)中, 日期递增分布(原来是19年到18年——变为18年到19年)

在这个操作中, 如果不考虑将反向放置, 其实相当于将A2到A13的数据, 一次又一次得复制粘贴到一个又一个省市的右方对应位置。那么现在我们所需要做的, 就是通过Offset函数, 将这些重复性的工作用一次下拉来完成。

(1) 有了之前的基础之后, 对于将一个省市的日期反向放置 (从上到下为18年到19年) 的操作是比较容易实现的。只需要将参考系 (起点) 定在A13, 并在OFFSET中 rows参数前加负号 (rows值为负表示向上移动) 即可。【起点定在2018年8月, 从下向上移动】

| | | | | | | |
|-----|----------|-------|----------|---------------------------------|-------|------|
| D17 | | | | =OFFSET(\$A\$13,-(ROW(A1)-1),0) | | |
| | A | B | C | D | E | F |
| 5 | 2019年4月 | 101.6 | 102.9 | 102.9 | 102.7 | 102. |
| 6 | 2019年3月 | 101.9 | 102.4 | 102.5 | 102.5 | 10 |
| 7 | 2019年2月 | 101.6 | 101.8 | 101.7 | 101.6 | 101. |
| 8 | 2019年1月 | 101.9 | 101.7 | 101.9 | 101.8 | 101. |
| 9 | 2018年12月 | 102 | 101.8 | 102 | 101.9 | 101. |
| 10 | 2018年11月 | 102.4 | 102 | 102.4 | 102.4 | 101. |
| 11 | 2018年10月 | 103 | 102.2 | 102.7 | 102.9 | 102. |
| 12 | 2018年9月 | 102.8 | 102.3 | 103 | 102.8 | 102. |
| 13 | 2018年8月 | 102.9 | 102.2 | 102.9 | 101.8 | 101. |
| 14 | | | | | | |
| 15 | | | | | | |
| 16 | | | | | | |
| 17 | | 北京市 | 2018年8月 | | | |
| 18 | | 北京市 | 2018年9月 | | | |
| 19 | | 北京市 | 2018年10月 | | | |
| 20 | | 北京市 | 2018年11月 | | | |
| 21 | | 北京市 | 2018年12月 | | | |
| 22 | | 北京市 | 2019年1月 | | | |
| 23 | | 北京市 | 2019年2月 | | | |
| 24 | | 北京市 | 2019年3月 | | | |
| 25 | | 北京市 | 2019年4月 | | | |
| 26 | | 北京市 | 2019年5月 | | | |
| 27 | | 北京市 | 2019年6月 | | | |
| 28 | | 北京市 | 2019年7月 | | | |
| 29 | | 天津市 | 时间 | | | |
| 30 | | 天津市 | #REF! | | | |
| 31 | | 天津市 | #REF! | | | |
| 32 | | 天津市 | #REF! | | | |

对应的OFFSET函数为: OFFSET(\$A\$13,-(ROW(A1)-1),0)

解释: 使用ROW(A1)函数——得到下拉的行数。 (



孤独的马铃薯

关注

6



26



1



专栏目录

小结:

因此, 想让某一行中的数据, 在指定单元格 (以D17为例) 下方循环输出, 只需要在指定的单元格内写入函数:

OFFSET(\$A\$13,MOD(ROW(A1)-1,n),0)

里面的n自己代入值, **A13**以及**ROW(A1)**之后是否需要减一, 由你的起点位置决定(这里选择了A13). =。=!

3.数据对应分布(难点在于将不同列的放在同一列内)

(1) 先将北京市的CPI对应放入

在2中我们知道了, 利用**OFFSET(\$A\$13,MOD(ROW(A1)-1,n),1)**可以将一行循环放入指定单元格下方, 所以要实现我们当前需求, 只要将上面函数的右移量参数cols由**0**——>**1**即可。

(2) 实现输完12个数据后, 自动换列继续输出

通过(1)我们只能在北京市的CPI一行不停循环输出, 而我们需要的是: 当北京市的数据输出完成(此时定位在B2)后, 函数能够自动转入天津市的第一个数据(由B2——>C13)。

此时便需要关注: **OFFSET(reference,rows,cols,height,width)**中的cols——右移量参数。在1 (省市输出12次后**换列**) 中已经给出了换列的方法——**ROW(An)/n**。



孤独的马铃薯

关注



6



26



1



专栏目录

| | | | | | | | |
|-----|---|---|-----|----------|-------|---|---|
| E17 | | | | | | | |
| | A | B | C | D | E | F | G |
| 17 | | | 北京市 | 2018年8月 | 102.9 | | |
| 18 | | | 北京市 | 2018年9月 | 102.8 | | |
| 19 | | | 北京市 | 2018年10月 | 103 | | |
| 20 | | | 北京市 | 2018年11月 | 102.4 | | |
| 21 | | | 北京市 | 2018年12月 | 102 | | |
| 22 | | | 北京市 | 2019年1月 | 101.9 | | |
| 23 | | | 北京市 | 2019年2月 | 101.6 | | |
| 24 | | | 北京市 | 2019年3月 | 101.9 | | |
| 25 | | | 北京市 | 2019年4月 | 101.6 | | |
| 26 | | | 北京市 | 2019年5月 | 101.8 | | |
| 27 | | | 北京市 | 2019年6月 | 102.6 | | |
| 28 | | | 北京市 | 2019年7月 | 102.6 | | |
| 29 | | | 天津市 | 2018年8月 | 102.2 | | |
| 30 | | | 天津市 | 2018年9月 | 102.3 | | |
| 31 | | | 天津市 | 2018年10月 | 102.2 | | |
| 32 | | | 天津市 | 2018年11月 | 102 | | |
| 33 | | | 天津市 | 2018年12月 | 101.8 | | |
| 34 | | | 天津市 | 2019年1月 | 101.7 | | |
| 35 | | | 天津市 | 2019年2月 | 101.8 | | |
| 36 | | | 天津市 | 2019年3月 | 102.4 | | |
| 37 | | | 天津市 | 2019年4月 | 102.9 | | |
| 38 | | | 天津市 | 2019年5月 | 102.3 | | |
| 39 | | | 天津市 | 2019年6月 | 102.2 | | |
| 40 | | | 天津市 | 2019年7月 | 102.3 | | |
| 41 | | | 河北省 | 2018年8月 | 102.9 | | |
| 42 | | | 河北省 | 2018年9月 | 103 | | |
| 43 | | | 河北省 | 2018年10月 | 102.7 | | |
| 44 | | | 河北省 | 2018年11月 | 102.4 | | |
| 45 | | | 河北省 | 2018年12月 | 102 | | |
| 46 | | | 河北省 | 2019年1月 | 101.9 | | |

=OFFSET(\$A\$13,-MOD(ROW(A1)-1,12),ROW(A12)/12)

因此将 (1) (2) 中的方法结合使用，得到——**OFFSET(\$A\$13,-**

MOD(ROW(A1)-1,12),ROW(A12)/12);

最终效果图如右：

小结：

因此，想让某点(以A1为例)下方的n个数据连续输出，并在输出完n个后自动换列输出第二列的n个数据——即将一个矩阵的每一列，顺序连接拼成一个列向量。只需要在指定的单元格内写入函数：

OFFSET(\$x\$y,-MOD(ROW(A1)-1,n),ROW(An)/n);

xy(起点，这里是A13)——定在矩阵的第一个元素之前；n为每一列的分量个数。

OFFSET函数-Excel最难的函数，太抽象了

03-25

今天跟大家讲解下Excel大神必备的Excel函数offset，offset函数的用法十分广泛，动态图表，动态表格，以及动态区域的计算等。很多人都觉得offset函数十分难理解，但是今天我要带你们吃透这个...

把正方形矩阵，比如相关系数矩阵，转化为3列以便做

-*- coding:utf-8 -*- import pandas as pd import numpy as np



孤独的马铃薯

关注

6

1

26

1

专栏目录

513

d...