

# 项目 1

## 1. 提取数据

- 1) 在工作区中输入如下图所示的代码（使用 ORDER BY 是为了更方便的查找中国的城市），然后将数据导出为 CSV 文档。根据 city\_list 文档中的城市名找到离我最近的城市为广州。

The screenshot shows a data workspace interface. On the left, there is a sidebar with a 'Schema' section containing a tree view with nodes: 'city\_data', 'city\_list', and 'global\_data'. The 'city\_list' node is selected. The main area is titled 'Input' and contains a SQL query editor with the following code:

```
1 SELECT *
2 FROM city_list
3 ORDER BY country
```

Below the query editor, there is a green 'Success!' message and a blue 'EVALUATE' button. The 'Output' section shows '345 results' and a 'Download CSV' button. The results are displayed in a table with two columns: 'city' and 'country'.

city	country
Kabul	Afghanistan
Tirana	Albania
Algiers	Algeria

- 2) 在工作区编写一个 SQL 查询来提取城市的气温数据，如下图所示（由于查询到的数据过多，所以使用 WHERE city = 'Guangzhou' 对数据进行筛选，这样可以直接获得广州的气温数据）并导出为 CSV 文档。

Input		HISTORY ▾	MENU ▾
SCHEMA	↻	<pre> 1  SELECT * 2  FROM city_data 3  WHERE city = 'Guangzhou' </pre>	
city_data	▾		
city_list	▾		
global_data	▾		
		Success!	EVALUATE
Output		174 results	Download CSV
year	city	country	avg_temp
1840	Guangzhou	China	20.98
1841	Guangzhou	China	21.02
1842	Guangzhou	China	21.16

3) 编写一个 SQL 查询来提取全球平均温度数据，如下图所示导出为 CSV 文档。

Input		HISTORY ▾	MENU ▾
SCHEMA	↻	<pre> 1  SELECT * 2  FROM global_data 3 </pre>	
city_data	▾		
city_list	▾		
global_data	▾		
		Success!	EVALUATE
Output		266 results	Download CSV
year	avg_temp		
1750	8.72		
1751	7.98		
1752	5.78		

## 2. 计算移动平均线

- 1) 在得到了广州的气温数据和全球气温数据后，我们把 CSV 文档通过 Excel 打开并进行了简单的编辑后得到了如下的表格（因为表格过长，所以只选取了最上面的一小部分）。由于广州只在 1840 年到 2013 年有数据，所以就删除了全球平均气温无法与其对应的部分。

year	Guangzhou's avg_temp	global avg_temp
1840	20.98	7.8
1841	21.02	7.69
1842	21.16	8.02
1843	21.25	8.17
1844	20.86	7.65
1845	20.84	7.85
1846	21.46	8.55
1847	21.07	8.09
1848	20.75	7.98
1849	21.03	7.98
1850	21.21	7.9
1851	21.33	8.18
1852	21.28	8.1
1853	21.78	8.04
1854	21.64	8.21

- 2) 在整理好数据后，新建一个名为 5-year MA of Guangzhou's temp 的列，使用 AVERAGE() 函数来计算前五年的平均温度，如下图所示。然后通过复制粘贴得到 1844 年到 2013 年所有年份的广州年平均温度的 5 年移动平均值。

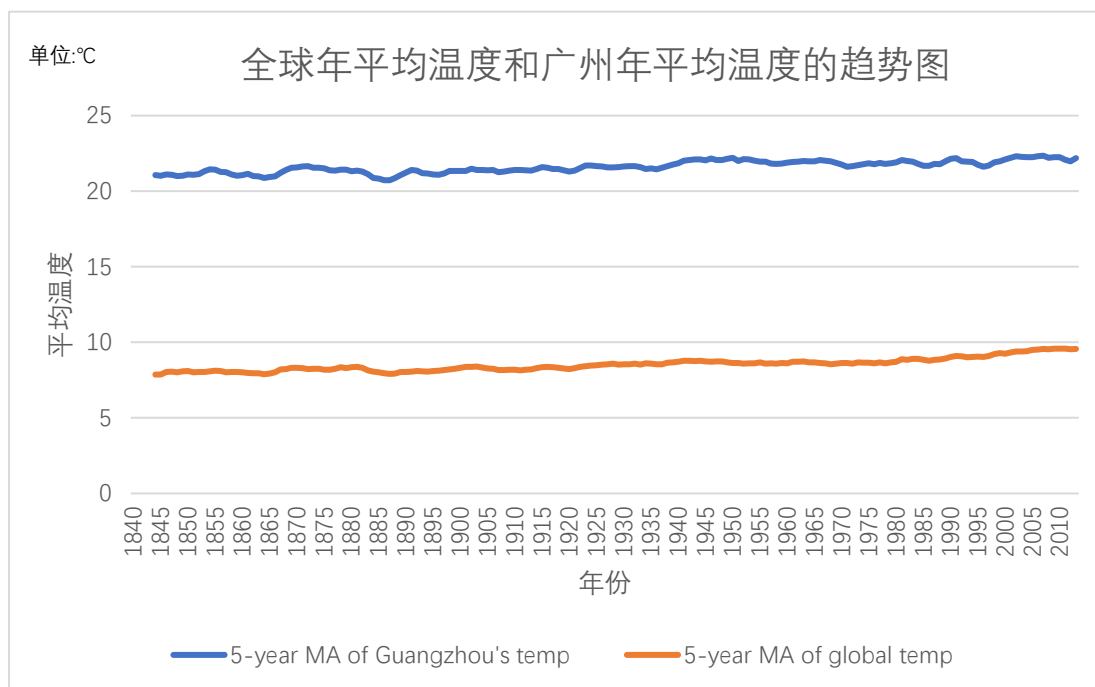
year	Guangzhou's avg_temp	5-year MA of Guangzhou's temp
1840	20.98	
1841	21.02	
1842	21.16	
1843	21.25	
1844	20.86	=AVERAGE(B2:B6)

3) 使用同样的方法得到 1844 年到 2013 年所有年份的全球平均气温的 5 年移动平均值, 如下图所示 (因表格太长, 所以只截取了一部分)。

year	5-year MA of Guangzhou's temp	5-year MA of global temp
1840		
1841		
1842		
1843		
1844	21.054	7.866
1845	21.026	7.876
1846	21.114	8.048
1847	21.096	8.062
1848	20.996	8.024
1849	21.03	8.09
1850	21.104	8.1
1851	21.078	8.026
1852	21.12	8.028
1853	21.326	8.04
1854	21.448	8.086
1855	21.424	8.128
1856	21.278	8.092
1857	21.246	8.024
1858	21.1	8.036
1859	21.028	8.044
1860	21.066	8.014

### 3. 可视化数据

考虑到要将广州的平均温度和全球的平均温度的变化趋势进行判断分析, 所以折线图是一种合理的选择。所以根据上面表格的数据, 可以通过 Excel 对表格的数据生成一张如下折线图。



#### 4. 数据分析

根据上图，可以得到以下几点发现：

- 1) 与全球平均温度气温相比，广州的平均温度要高  $13^{\circ}\text{C}$  左右，且这种差异在 173 年间基本保持一致。
- 2) 全球的平均温度在 173 年间不是保持在一个恒定的值，而是在一个较小的范围内波动。广州的平均温度也符合这一规律。
- 3) 全球的平均温度虽然在一定时间内（几年内）出现过下降的趋势，但整体上从 1940 年到 2013 年上升了  $1^{\circ}\text{C}$  左右。广州的平均温度在这期间总体上也有一定的上升。
- 4) 广州的平均温度波动的程度比全球平均温度波动的程度大，可能是因为全球平均温度是平均各个地区的平均温度后得出的数值，所以波动会比较小。