

Udacity

项目一：

探索未来气候发展趋势

学员姓名：徐振文

时间：2018/4/13

步骤一：从数据库中提取数据

首先，我需要从与数据库相连的指定工作区中导出世界气温数据以及最接近你居住地的大城市气温数据。我在这一步骤中使用了 SQL 进行数据提取。

- 编写 SQL 查询，提取城市数据，导出到 CSV 文件。

```
SELECT * /*这行表示我需要提取列表中所有列的数据*/  
FROM city_data /*这行表示我需要从 city_data 这个表中提取*/  
WHERE city = 'Wuhan' AND country = 'China'; /*这行表示我指定提取  
国家是中国的武汉这座城市的所有信息*/
```

由于我居住在中国武汉，所以我使用了上面三行代码提取出了武汉从 1841 年到 2013 年每年的平均气温。

- 编写一个 SQL 查询来提取全球数据，并导出到 CSV 文件。

```
SELECT * /*这行表示我需要提取列表中所有列的数据*/  
FROM global_data; /*这行表示我需要从 global_data 这个表中提取*/
```

由于 global_data 这个表中只有年份和平均气温，所以我仅使用上面两行代码提取出了世界从 1750 年到 2015 年每年的平均气温。

提取完数据以后，我使用工作区中的下载按钮将两次提取的数据下载成 .csv 文件以待之后使用。

步骤二：用喜爱使用的工具打开下载的 CSV 文件

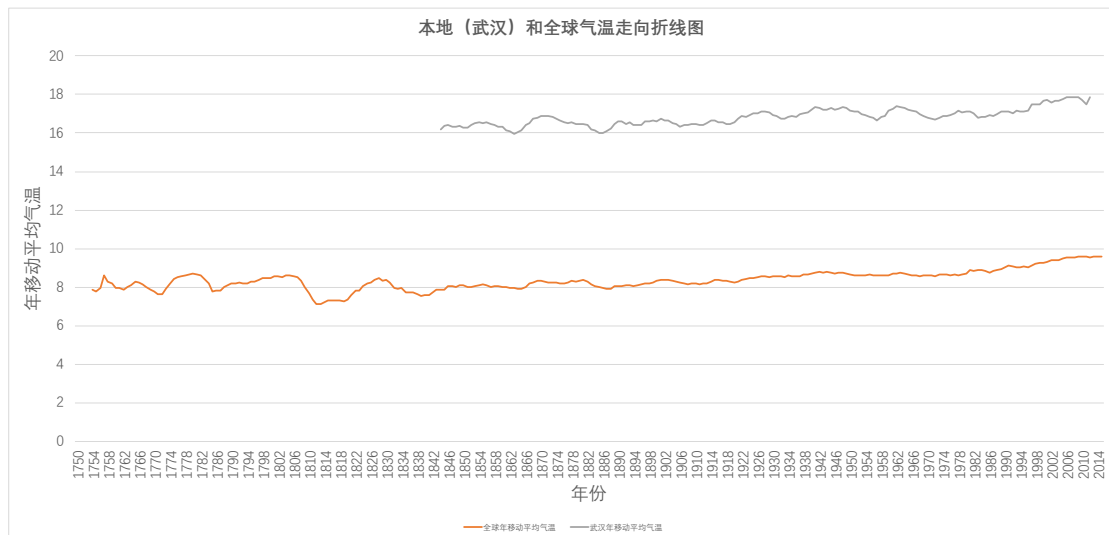
我在这一步骤中选择使用 Excel 打开了下载的 CSV 文件。

步骤三：创建一个线条图

在打开两个 .csv 文件之后，我打算以每年的前 5 年为一个集群计算移动平均值。首先我在提取的城市年平均气温的列表文件中的平均气温的后一列中取名“5-year_avg”，并在此列第 6 行的空格中调用 AVERAGE 函数，具体输入为“=AVERAGE(D2:D6)”，然后我将选中框下拉直到 2013 年，之后对应每一年的移动平均数便都出来了。这一步骤我使用的是 Excel。

同理，我在提取的世界年平均气温的列表文件中也计算了从 1754 年开始直到 2015 年的每年的移动平均数。

最后我将两组数据统一，绘制了如下的折线图：



步骤四：观察

如上图所示，1.根据武汉的年移动平均气温折线图走向可以看出，年平均气温呈现出整体缓慢上升的趋势；2.而根据全球的移动平均气温折线图走向可以看出，年平均气温先呈现出基本保持平衡的趋势，然后呈现出缓慢上升的趋势；3.两折线图的趋势基本相近，但是武汉的折线图的波动更大；4.在 1850 年至 2013 年，武汉和全球的平均气温都是在 2 度以内进行波动。这一步骤我仍旧使用的是 Excel。

- 与全球平均气温相比，你所在城市平均气温是比较热还是比较冷？长期气温差异是否一致？

我所在的城市平均气温比较热，但是长期气温差异基本一致。

- “长期以来，你所在城市气温变化与全球平均气温变化相比如何？”

长期以来，我所在的城市的气温变化的总体趋势与全球平均气温的变化趋势基本一致，都呈现出缓慢上升的状态；但是我所在的城市的气温变化波动较全球的气温变化更大。

- 整体趋势如何？世界越来越热还是越来越冷了？气温走向与过去几百年的走向是否一致？

整体呈现出气温上升的趋势。世界越来越热了。从 1750 年到 1850 年，世界的平均气温在 8 度上下波动，从 1850 年到 1950 年，世界的平均气温仍旧在 8 度左右，但是波动更小，从 1950 年到现在，世界的平均气温一直呈现上升趋势，快要接近 10 度。

在决定如何可视化气温走向时，我主要考虑的因素是年份的范围，本地和全球平均气温的范围，以及能否将两组数据放入同一折线图进行比较。