探索未来气候发展趋势项目

概览

在这个项目中，我将分析*广州*和*全球*的气温数据，并比较*广州*气温走向与*全球*气温走向，形成可视化的分析报告。

说明

本次分析将按照：

- 采集数据
- 整理数据
- 创建图像
- 观察分析
- 得出结论

的步骤进行研究分析，试用SQL语言和r语言画图工具进行数据可视化。

步骤详解

1.采集数据(SQL)

- 确认城市 查询数据库中的中国城市。试用如下代码在网页中查询相关数据：

  ```
  SELECT *
  FROM city_list
  WHERE country = 'China'
  ```

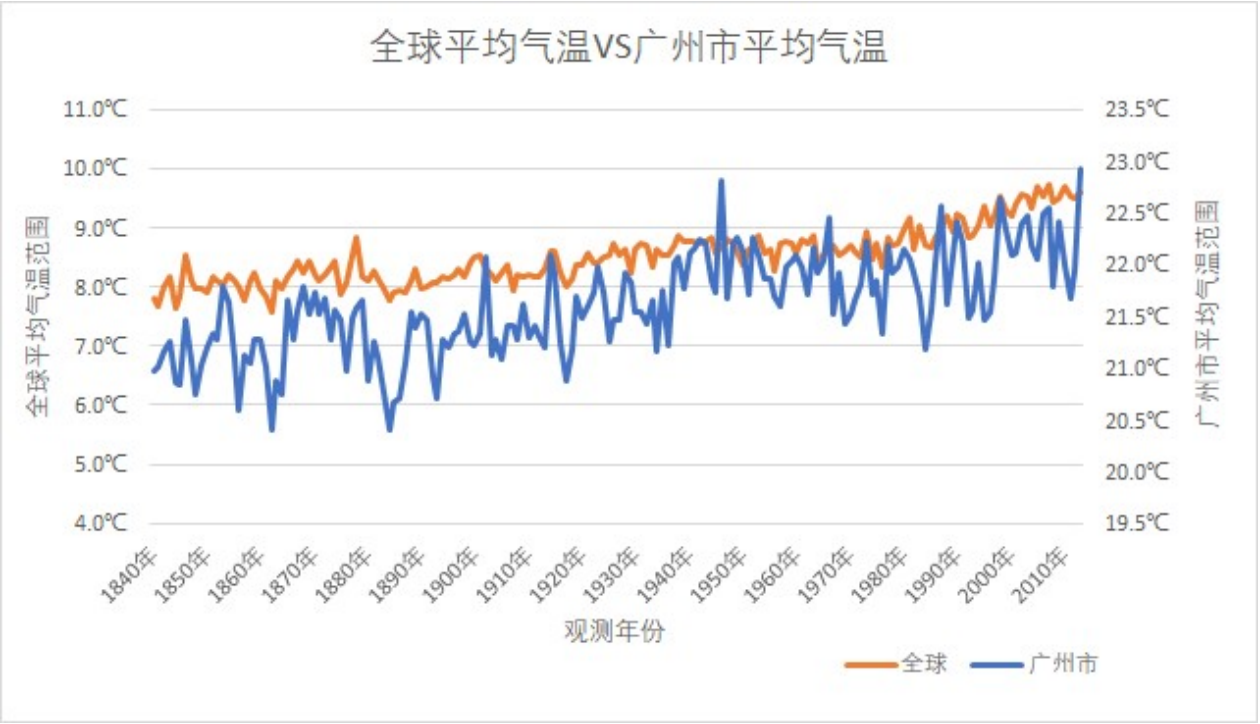
  由于答主所在城市为深圳，纵观全表，广州为表中距离最近城市，所以将广州选为全球气候的对比城市。
- 提取广州市和全球相同年份平均气温值

  ```
  SELECT c.year, c.avg_temp as city_temp, g.avg_temp as global_temp
  FROM city_data c, global_data g
  WHERE c.year = g.year
  AND c.city = 'Guangzhou'
  ```

  保存数据到result.xls文件。

2. 整理数据(excel)

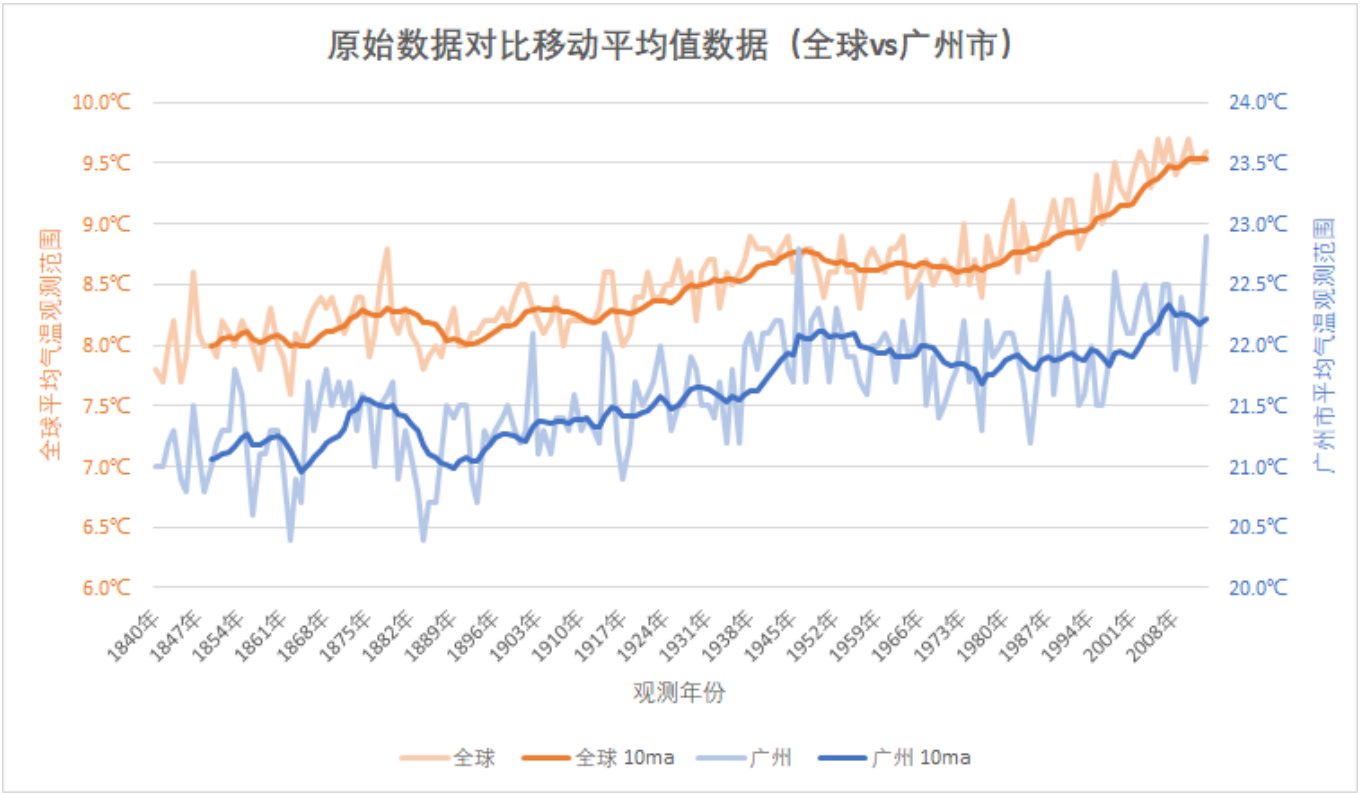
我们将生成的xls文件通过excel形成初步的观测图像，如下（为突出两组数据的走向，右侧增加了次坐标轴作为参考）：



由图中可观测到，两组数据在未经过处理前，形成的曲线波动较大，难以看出明显的运动趋势； 这里我们选择使用移动平均法对两组数据的曲线进行平滑处理，以使曲线运动趋势更加明显，更易观测； 我们知道，移动平均数的N值越大，曲线越平滑，但也使折线图的一些细节被埋没。根据经验，我们取N=10，作为移动平均值的处理周期，既保证观察到更多局部的波动，又突出长远趋势。 通过excel中的average函数对两组原始数据进行处理。

3. 形成图表(excel)

经过整理后，形成了两幅图表的10年移动平均值图像。如下：



4. 观察分析

观测平滑处理后的折线图，我可以初步得出以下结论：

- 整体而言，全球平均气温逐年递增，世界是越来越热的；

- 整体而言，广州市平均气温逐年递增，广州市也是越来越热的；
- 广州市的平均气温高于同期全球平均气温；

同时我又产生了以下几个问题，它们是：

- 广州市平均气温与全球气温变动有一定的相关性吗？（从移动平均值曲线基本平行推断得出）；
- 在2019年，广州市平均气温大约会达到多少？

5.进一步验证（excel）

- 针对第一个问题，我使用excel中的【数据分析】加载项中的【相关系数】，对全球平均气温和广州市平均气温原始数据进行分析，得到以下内容：

	广州	全球
广州	1	
全球	0.758083	1

可知，两组原始数据的相关系数约为0.758,证明两组数据具有较高的相关性。

- 针对第二个问题：通过之前的观测发现，数据具有明显的上升趋势，明显的趋势变动的时间序列，于是我选择通过holt指数平滑法进行分析，主要使用excel的【数据分析-回归】和【规划求解】功能进行。涉及到公式为：

$$L_{t+1}=\alpha D_{t+1}+(1-\alpha)(L_t+Tt)$$

$$T_{t+1}=\beta(L_{t+1}-L_t)+(1-\beta)T_t$$

$$F_{t+1}=L_t+T_t$$

$$F_{t+n}=L_t+nT_t$$

其中：

F代表预测值，L代表截距(level)，T 代表斜率（trend），D 表示实际值；

α和β表示两个参数。

1) 整理数据，将日期栏按照序列加工为“1、2、3、4、……174”，通过回归计算得以下内容：

SUMMARY OUTPUT

回归统计	
Multiple R	0.70251685
R Square	0.493529924
Adjusted R	0.490585331
标准误差	0.343997199
观测值	174

方差分析

	df	SS	MS	F	ignificance F
回归分析	1	19.83344	19.83344	167.6055	3.36E-27
残差	172	20.35346	0.118334		
总计	173	40.1869			

	Coefficients	标准误差	t Stat	P-value	Lower 95%	Upper 95%	下限 95.0%	上限 95.0%
Intercept	21.02565278	0.052382	401.3881	1.5E-257	20.92226	21.12905	20.92226	21.12905
X Variable	0.006721604	0.000519	12.94625	3.36E-27	0.005697	0.007746	0.005697	0.007746

2) 确认初始trend和level(分别对应Coefficients的1、2行)，然后根据以上公式计算出forcast.ft值，如图：

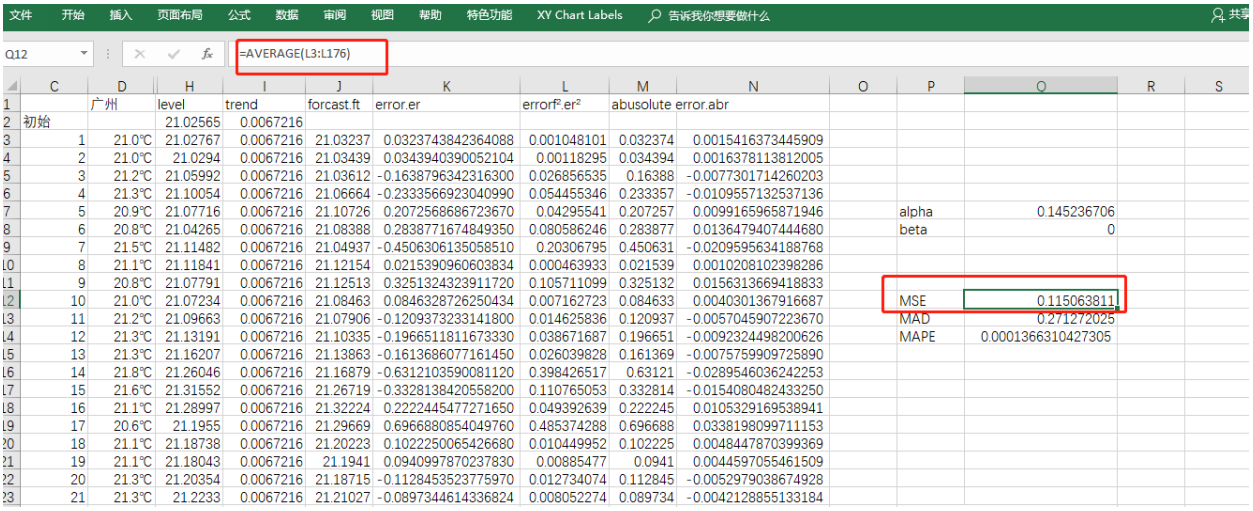
	广州	level	trend	forcast.ft
初始		21.02565	0.0067216	
1	21.0℃	21.02767	0.0067216	21.03237
2	21.0℃	21.0294	0.0067216	21.03439
3	21.2℃	21.05992	0.0067216	21.03612
4	21.3℃	21.10054	0.0067216	21.06664
5	20.9℃	21.07716	0.0067216	21.10726
6	20.8℃	21.04265	0.0067216	21.08388
7	21.5℃	21.11482	0.0067216	21.04937
8	21.1℃	21.11841	0.0067216	21.12154
9	20.8℃	21.07791	0.0067216	21.12513
10	21.0℃	21.07234	0.0067216	21.08463
11	21.2℃	21.09663	0.0067216	21.07906
12	21.3℃	21.13191	0.0067216	21.10335
13	21.3℃	21.16207	0.0067216	21.13863
14	21.8℃	21.26046	0.0067216	21.16879
15	21.6℃	21.31552	0.0067216	21.26719
16	21.1℃	21.28997	0.0067216	21.32224
17	20.6℃	21.1955	0.0067216	21.29669
18	21.1℃	21.18738	0.0067216	21.20223
19	21.1℃	21.18043	0.0067216	21.1941
20	21.3℃	21.20354	0.0067216	21.18715
21	21.3℃	21.2233	0.0067216	21.21027
22	21.0℃	21.19661	0.0067216	21.23002
23	20.4℃	21.08666	0.0067216	21.20333
24	20.9℃	21.0653	0.0067216	21.09338
25	20.7℃	21.01799	0.0067216	21.07202
26	21.7℃	21.12279	0.0067216	21.02471
27	21.3℃	21.15427	0.0067216	21.12951
28	21.6℃	21.22475	0.0067216	21.16099
29	21.8℃	21.31404	0.0067216	21.23147
30	21.5℃	21.3468	0.0067216	21.32077
31	21.7℃	21.40384	0.0067216	21.35352
32	21.5℃	21.42355	0.0067216	21.41056
33	21.7℃	21.46945	0.0067216	21.43027
34	21.3℃	21.45058	0.0067216	21.47617
35	21.6℃	21.47803	0.0067216	21.4573
36	21.5℃	21.48697	0.0067216	21.48475
37	21.0℃	21.42199	0.0067216	21.49369
38	21.5℃	21.43906	0.0067216	21.42871
39	21.6℃	21.46818	0.0067216	21.44578
40	21.7℃	21.5076	0.0067216	21.4749
41	20.9℃	21.4251	0.0067216	21.51432
42	21.3℃	21.41267	0.0067216	21.43182
43	21.1℃	21.37301	0.0067216	21.41939
44	20.8℃	21.29553	0.0067216	21.37973
45	20.4℃	21.17121	0.0067216	21.30225
46	20.7℃	21.10852	0.0067216	21.17793
47	20.7℃	21.05493	0.0067216	21.11524
48	21.1℃	21.06722	0.0067216	21.06165
49	21.5℃	21.13582	0.0067216	21.07395
50	21.4℃	21.17994	0.0067216	21.14255
51	21.5℃	21.23217	0.0067216	21.18666
52	21.5℃	21.27681	0.0067216	21.23889
53	20.9℃	21.22783	0.0067216	21.28353
54	20.7℃	21.15692	0.0067216	21.23455

55	21.3℃	21.18344	0.0067216	21.16364
56	21.2℃	21.19159	0.0067216	21.19016
57	21.3℃	21.21308	0.0067216	21.19831
58	21.4℃	21.24598	0.0067216	21.2198
59	21.5℃	21.28861	0.0067216	21.2527
60	21.3℃	21.29601	0.0067216	21.29534
61	21.2℃	21.28781	0.0067216	21.30273
62	21.3℃	21.29533	0.0067216	21.29454
63	22.1℃	21.41794	0.0067216	21.30205
64	21.1℃	21.37751	0.0067216	21.42466
65	21.3℃	21.372	0.0067216	21.38423
66	21.1℃	21.33824	0.0067216	21.37872
67	21.4℃	21.35296	0.0067216	21.34496
68	21.4℃	21.36553	0.0067216	21.35968
69	21.3℃	21.36176	0.0067216	21.37225
70	21.6℃	21.40211	0.0067216	21.36848
71	21.3℃	21.39302	0.0067216	21.40883
72	21.4℃	21.39978	0.0067216	21.39974
73	21.3℃	21.39104	0.0067216	21.4065
74	21.2℃	21.36904	0.0067216	21.39776
75	22.1℃	21.48094	0.0067216	21.37576
76	21.9℃	21.54755	0.0067216	21.48767
77	21.2℃	21.50282	0.0067216	21.55427
78	20.9℃	21.42101	0.0067216	21.50954
79	21.2℃	21.39466	0.0067216	21.42773
80	21.7℃	21.44475	0.0067216	21.40138
81	21.5℃	21.45852	0.0067216	21.45147
82	21.6℃	21.48481	0.0067216	21.46524
83	21.7℃	21.52181	0.0067216	21.49154
84	22.0℃	21.59701	0.0067216	21.52853
85	21.7℃	21.61771	0.0067216	21.60373
86	21.3℃	21.57731	0.0067216	21.62443
87	21.5℃	21.57183	0.0067216	21.58404
88	21.5℃	21.56714	0.0067216	21.57855
89	21.9℃	21.62123	0.0067216	21.57386
90	21.8℃	21.65294	0.0067216	21.62795
91	21.5℃	21.63647	0.0067216	21.65966
92	21.5℃	21.6224	0.0067216	21.6432
93	21.4℃	21.59584	0.0067216	21.62912
94	21.7℃	21.61672	0.0067216	21.60256
95	21.2℃	21.56194	0.0067216	21.62344
96	21.8℃	21.60226	0.0067216	21.56866
97	21.2℃	21.54958	0.0067216	21.60898
98	22.0℃	21.62074	0.0067216	21.5563
99	22.1℃	21.6961	0.0067216	21.62747
100	21.8℃	21.71693	0.0067216	21.70282
101	22.1℃	21.77831	0.0067216	21.72365
102	22.1℃	21.83078	0.0067216	21.78503
103	22.2℃	21.89015	0.0067216	21.8375
104	22.2℃	21.9409	0.0067216	21.89687
105	21.8℃	21.92618	0.0067216	21.94762
106	21.7℃	21.89907	0.0067216	21.9329
107	22.8℃	22.03567	0.0067216	21.9058
108	21.7℃	21.99266	0.0067216	22.04239
109	22.2℃	22.02852	0.0067216	21.99938
110	22.3℃	22.07369	0.0067216	22.03524

111	22.1℃	22.08326	0.0067216	22.08042
112	21.7℃	22.03334	0.0067216	22.08998
113	22.3℃	22.07782	0.0067216	22.04006
114	22.1℃	22.08678	0.0067216	22.08454
115	21.9℃	22.0654	0.0067216	22.0935
116	21.9℃	22.04712	0.0067216	22.07212
117	21.7℃	22.00245	0.0067216	22.05385
118	21.6℃	21.94975	0.0067216	22.00918
119	22.0℃	21.96279	0.0067216	21.95647
120	22.0℃	21.97394	0.0067216	21.96951
121	22.1℃	21.998	0.0067216	21.98066
122	22.0℃	22.00403	0.0067216	22.00472
123	21.7℃	21.96562	0.0067216	22.01075
124	22.2℃	22.00541	0.0067216	21.97234
125	21.9℃	21.99584	0.0067216	22.01213
126	22.0℃	22.00219	0.0067216	22.00256
127	22.5℃	22.08024	0.0067216	22.00891
128	21.5℃	22.00171	0.0067216	22.08696
129	21.9℃	21.99268	0.0067216	22.00843
130	21.4℃	21.91235	0.0067216	21.99941
131	21.5℃	21.85821	0.0067216	21.91907
132	21.7℃	21.84098	0.0067216	21.86493
133	21.8℃	21.84077	0.0067216	21.8477
134	22.2℃	21.89869	0.0067216	21.84749
135	21.7℃	21.87558	0.0067216	21.90541
136	21.8℃	21.87035	0.0067216	21.8823
137	21.3℃	21.79326	0.0067216	21.87707
138	22.2℃	21.85808	0.0067216	21.79998
139	21.9℃	21.86991	0.0067216	21.8648
140	22.0℃	21.89455	0.0067216	21.87663
141	22.1℃	21.93013	0.0067216	21.90127
142	22.1℃	21.96055	0.0067216	21.93686
143	21.9℃	21.9575	0.0067216	21.96727
144	21.7℃	21.92585	0.0067216	21.96422
145	21.2℃	21.82617	0.0067216	21.93257
146	21.6℃	21.79907	0.0067216	21.8329
147	22.0℃	21.834	0.0067216	21.80579
148	22.6℃	21.95099	0.0067216	21.84072
149	21.6℃	21.90576	0.0067216	21.95772
150	22.1℃	21.93972	0.0067216	21.91248
151	22.4℃	22.01231	0.0067216	21.94644
152	22.2℃	22.04532	0.0067216	22.01904
153	21.5℃	21.97186	0.0067216	22.05204
154	21.6℃	21.9236	0.0067216	21.97858
155	22.0℃	21.94044	0.0067216	21.93032
156	21.5℃	21.88222	0.0067216	21.94716
157	21.5℃	21.83245	0.0067216	21.88894
158	21.9℃	21.84801	0.0067216	21.83917
159	22.6℃	21.96297	0.0067216	21.85473
160	22.3℃	22.01766	0.0067216	21.96969
161	22.1℃	22.03537	0.0067216	22.02439
162	22.1℃	22.0505	0.0067216	22.04209
163	22.4℃	22.10701	0.0067216	22.05722
164	22.5℃	22.16983	0.0067216	22.11373
165	22.2℃	22.17996	0.0067216	22.17655
166	22.1℃	22.17409	0.0067216	22.18668

167	22.5℃	22.22717	0.0067216	22.18081
168	22.5℃	22.27254	0.0067216	22.23389
169	21.8℃	22.20965	0.0067216	22.27926
170	22.4℃	22.24304	0.0067216	22.21638
171	22.0℃	22.21349	0.0067216	22.24977
172	21.7℃	22.14466	0.0067216	22.22021
173	22.0℃	22.12939	0.0067216	22.15138
174	22.9℃	22.24706	0.0067216	22.13612
				22.25378
				22.2605
				22.26722
				22.27395
				22.28067
				22.28739

3) 根据结果计算出，计算预测值与实际值差异error.er，差异平方error².er²，绝对差异，以及平均绝对差异，再根据差异值计算出均方误差MSE，如图：

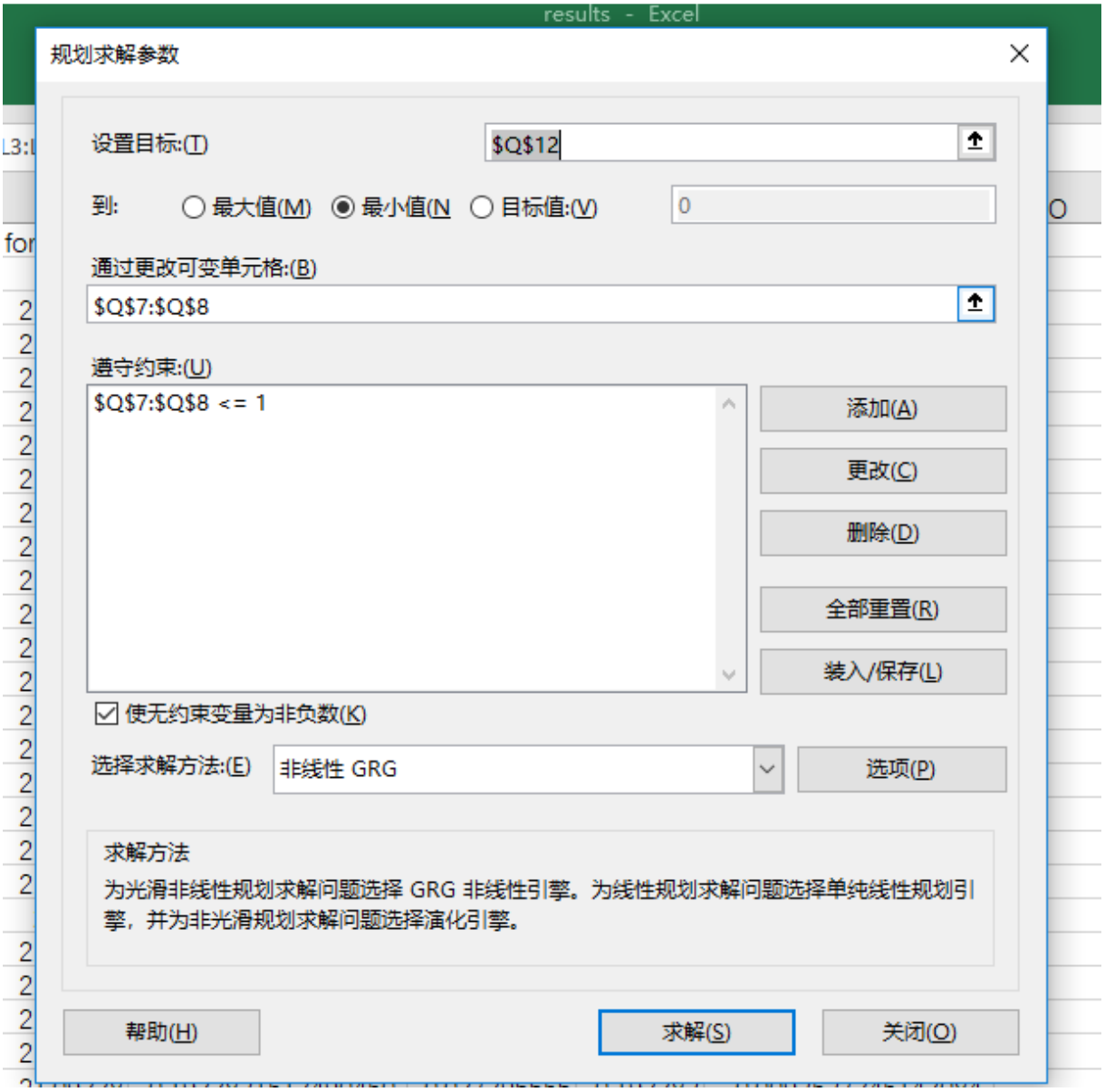


文件 开始 插入 页面布局 公式 数据 审阅 视图 帮助 特色功能 XY Chart Labels 告诉我你想要做什么

Q12 =AVERAGE(L3:L176)

	C	D	H	I	J	K	L	M	N	O	P	Q	R	S
1		广州	level	trend	forcast.ft	error.er	errorf².er²	abusolute	error.abr					
2	初始		21.02565	0.0067216										
3	1	21.0℃	21.02767	0.0067216	21.03237	0.0323743842364088	0.001048101	0.032374	0.0015416373445909					
4	2	21.0℃	21.0294	0.0067216	21.03439	0.0343940390052104	0.00118295	0.034394	0.0016378113812005					
5	3	21.2℃	21.05992	0.0067216	21.03612	-0.1638796342316300	0.026856535	0.16388	-0.0077301714260203					
6	4	21.3℃	21.10054	0.0067216	21.06664	-0.2333566923040990	0.054455346	0.233357	-0.0109557132537136					
7	5	20.9℃	21.07716	0.0067216	21.10726	0.2072568686723670	0.04295541	0.207257	0.0099165965871946		alpha	0.145236706		
8	6	20.8℃	21.04265	0.0067216	21.08388	0.2838771674849350	0.080586246	0.283877	0.0136479407444680		beta	0		
9	7	21.5℃	21.11482	0.0067216	21.04937	-0.4506306135058510	0.20306795	0.450631	-0.0209595634188768					
10	8	21.1℃	21.11841	0.0067216	21.12154	0.0215390960603834	0.000463933	0.021539	0.0010208102398286					
11	9	20.8℃	21.07791	0.0067216	21.12513	0.3251324323911720	0.105711099	0.325132	0.0156313669418833					
12	10	21.0℃	21.07234	0.0067216	21.08463	0.0846328726250434	0.007162723	0.084633	0.0040301367916687		MSE	0.115063811		
13	11	21.2℃	21.09663	0.0067216	21.07906	-0.1209373233141800	0.014625836	0.120937	-0.0057045907223670		MAD	0.271272025		
14	12	21.3℃	21.13191	0.0067216	21.10335	-0.1966511811673330	0.038671687	0.196651	-0.0092324498200626		MAPE	0.0001366310427305		
15	13	21.3℃	21.16207	0.0067216	21.13863	-0.1613686077161450	0.026039828	0.161369	-0.0075759909725890					
16	14	21.8℃	21.26046	0.0067216	21.16879	-0.6312103590081120	0.398426517	0.63121	-0.0289546036242253					
17	15	21.6℃	21.31552	0.0067216	21.26719	-0.3328138420558200	0.110765053	0.332814	-0.0154080482433250					
18	16	21.1℃	21.28997	0.0067216	21.32224	0.2222445477271650	0.049392639	0.222245	0.0105329169538941					
19	17	20.6℃	21.1955	0.0067216	21.29669	0.6966880854049760	0.485374288	0.696688	0.0338198099711153					
20	18	21.1℃	21.18738	0.0067216	21.20223	0.1022250065426680	0.010449952	0.102225	0.0048447870399369					
21	19	21.1℃	21.18043	0.0067216	21.1941	0.0940997870237830	0.00885477	0.0941	0.0044597055461509					
22	20	21.3℃	21.20354	0.0067216	21.18715	-0.1128453523775970	0.012734074	0.112845	-0.0052979038674928					
23	21	21.3℃	21.2233	0.0067216	21.21027	-0.0897344614336824	0.008052274	0.089734	-0.0042128855133184					

4) 最后通过规划求解，求得MSE最小时的α和β组合，注意，这里设定α和β<=1，如图：



得到在MSE最小时，α≈0.1452，β= 0，此时预测到的2019年广州市平均气温为22.29℃。如图

167	22.5℃	22.22717	0.0067216	22.18081	-0.3191896513015320	0.101882033	0.31919	-0.0141862067245125
168	22.5℃	22.27254	0.0067216	22.23389	-0.2661099941040350	0.070814529	0.26611	-0.0118271108490682
169	21.8℃	22.20965	0.0067216	22.27926	0.4792605485326650	0.229690673	0.479261	0.0219844288317736
170	22.4℃	22.24304	0.0067216	22.21638	-0.1836240711059070	0.033717799	0.183624	-0.0081975031743709
171	22.0℃	22.21349	0.0067216	22.24977	0.2497664877927690	0.062383298	0.249766	0.0113530221723986
172	21.7℃	22.14466	0.0067216	22.22021	0.5202128295559270	0.270621388	0.520213	0.0239729414541902
173	22.0℃	22.12939	0.0067216	22.15138	0.1513804355231090	0.022916036	0.15138	0.0068809288874140
174	22.9℃	22.24706	0.0067216	22.13612	-0.7638839565771890	0.583518699	0.763884	-0.0333573780164711
175				22.25378				
176				22.2605				
177				22.26722				
178				22.27395				
179				22.28067				
180（2019年）				22.28739				

6. 得出结论

- 整体而言，全球平均气温逐年递增，世界是越来越热的；
- 整体而言，广州市平均气温逐年递增，广州市也是越来越热的；
- 广州市的平均气温高于同期全球平均气温；
- 广州市平均气温与全球气温变动有一定的相关性，相关系数约为0.758；