Going Long on Temperature

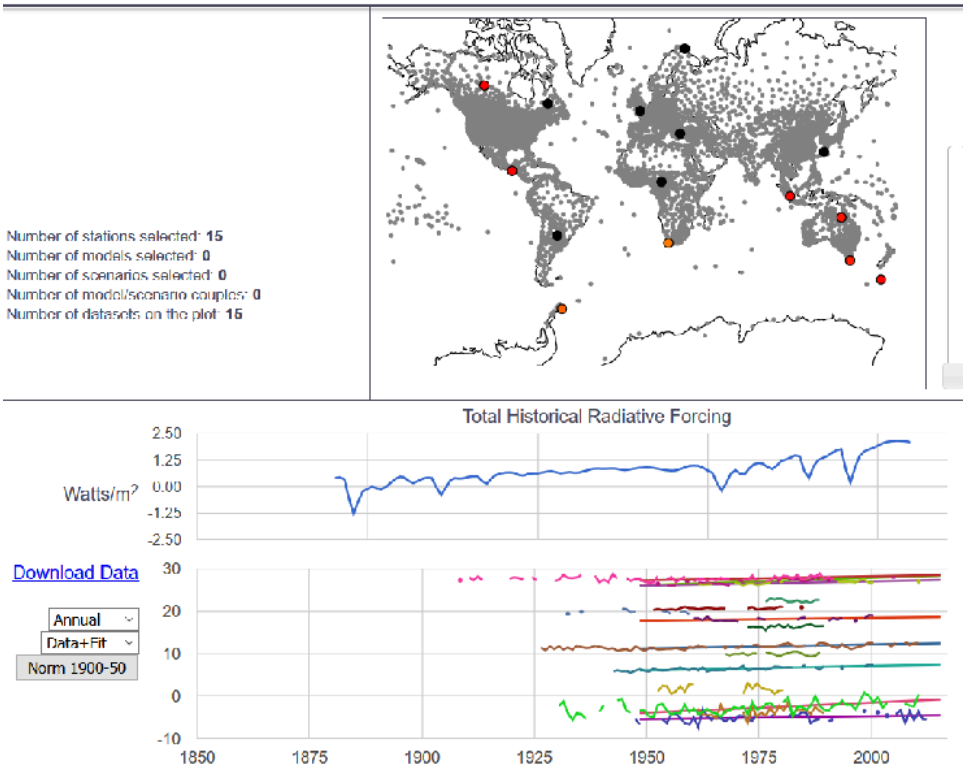
Simon Crase

27 October 2016

Introduction

Idea

My project started with the idea of showing trends in temperatures from randomly selected stations. I downloaded the list of station names from the Time Series Browser, and wrote a short Python script to select station names at random. I then manually selected stations, see link here. However, there is a problem as the figure makes clear: *stations are not distributed uniformly, so the selection will also be skewed.*



I therefore decided to sample uniformly across the globe: select positions at random, and take the readings from the nearest stations with adequate data. See Methodology Section below.

References

- Code is stored in my Github repository - user name 'weka11'
- The algorithm for uniformly sampling points on the surface of a globe is documented in Statistical Mechanics: Algorithms and Computations, Werner Krauth, PDF here - *algorithm 1-22, direct-surface.*
- Time Series Browser
- NOAA National Climatic Data Center

Methodology

I decided against using the Time Series Browser, as it was designed for accessing data by station name, not latitude and longitude. I could have written a script to generate random locations, and then screen scrape the data, but past experince with other websites has taught me that this is fraught with difficulty (typically this requires the exploitation of undocumented features in the website: *here be Dragons!*). I decided to use the *data behind the Time Series Browser* instead. The analyses are perfomed in R, and this document has been generarated by R Markdown - see References above. **NB: as the R code uses a random number generator, the results of each run will be different.**

- Download stations and monthly average temperature readings from NOAA National Climatic Data Center
- Randomly sample 25 locations, uniformly distributed distributed on the surface of the Globe, using the algorithm from Werner Karuth's book, above.
- Filter the list of stations so they are restricted to those with readings in the time interval of interest (currently 1950 to the present)
- For each of the 25 locations, find the nearest station.
- Tabulate the stations, and plot the time annual average temperatures, along with a regression line

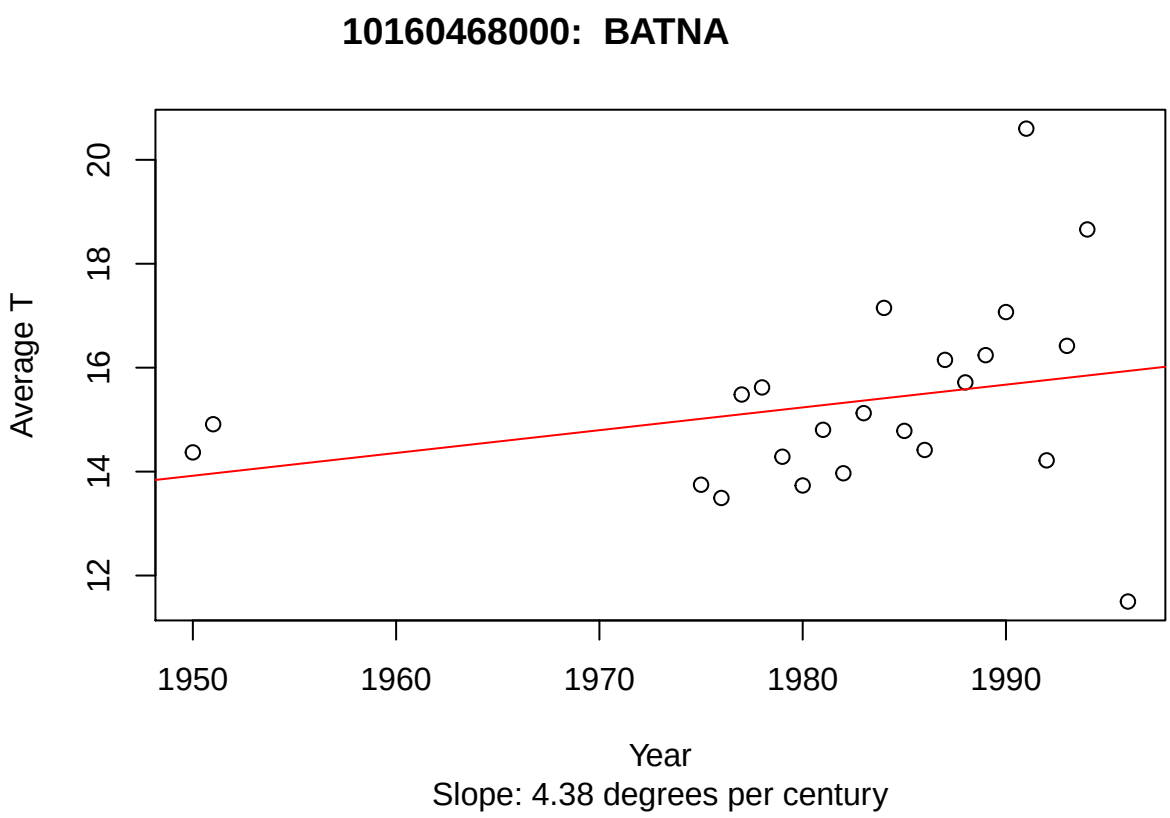
Results

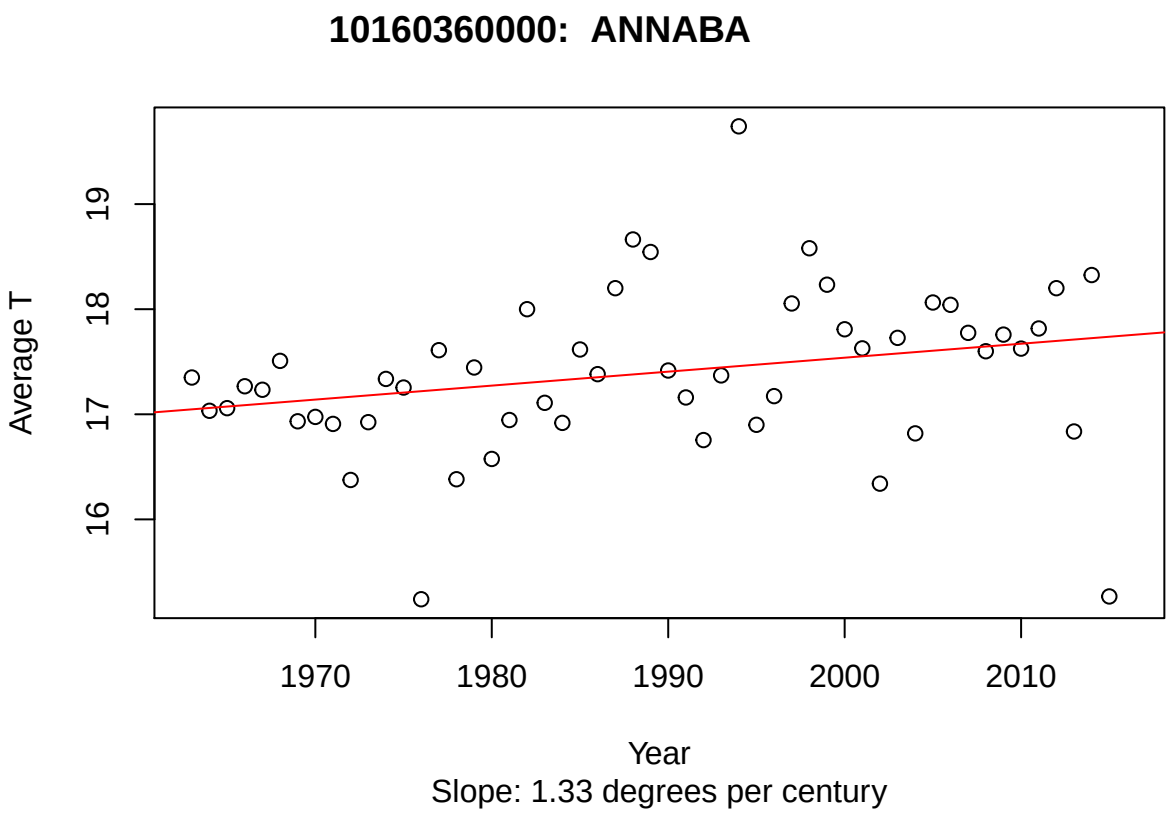
Stations and their locations

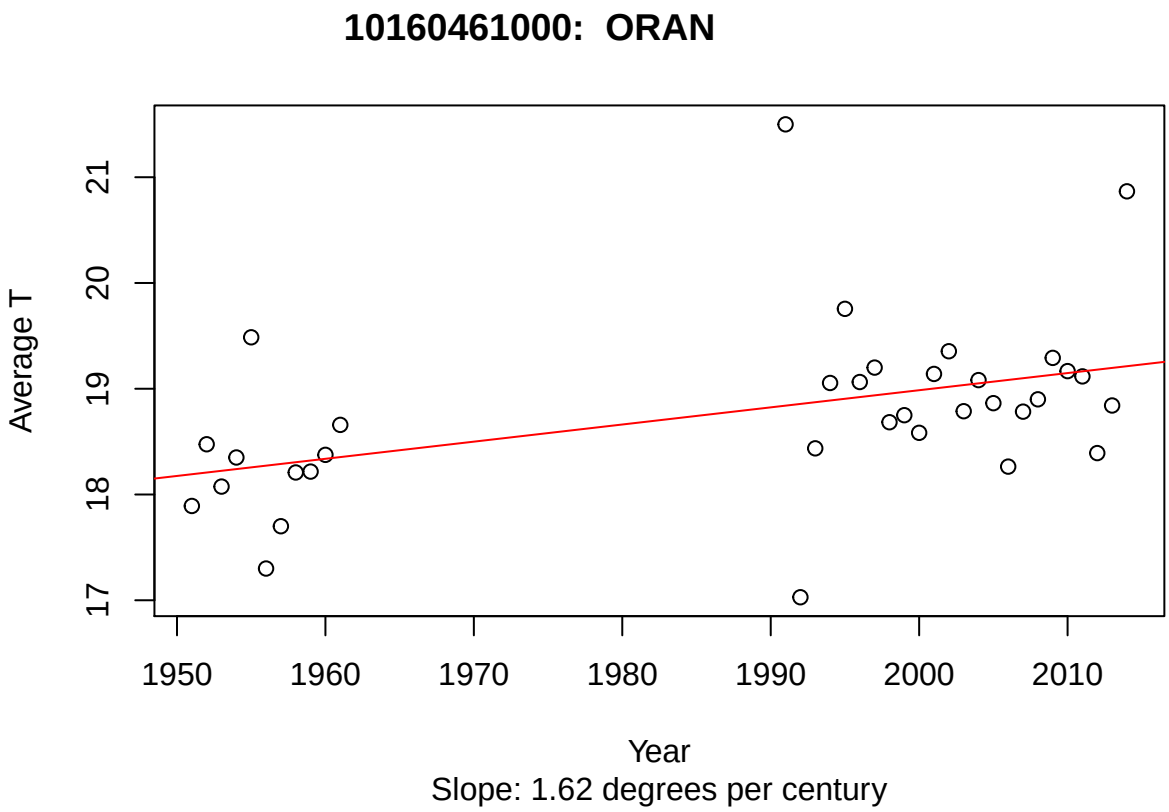
ID	NAME	LATITUDE	LONGITUDE
10160468000	BATNA	35.55	6.18
10160360000	ANNABA	36.83	7.82
10160360000	ANNABA	36.83	7.82
10160461000	ORAN	35.70	-0.65
10160425000	CHLEF	36.22	1.33
10160468000	BATNA	35.55	6.18
10160360000	ANNABA	36.83	7.82
10160461000	ORAN	35.70	-0.65
10160468000	BATNA	35.55	6.18
10160360000	ANNABA	36.83	7.82
10160360000	ANNABA	36.83	7.82
10160360000	ANNABA	36.83	7.82
10160360000	ANNABA	36.83	7.82
10160390000	DAR-EL-BEIDA	36.72	3.25
10160355000	SKIKDA	36.93	6.95
10160360000	ANNABA	36.83	7.82
10160360000	ANNABA	36.83	7.82

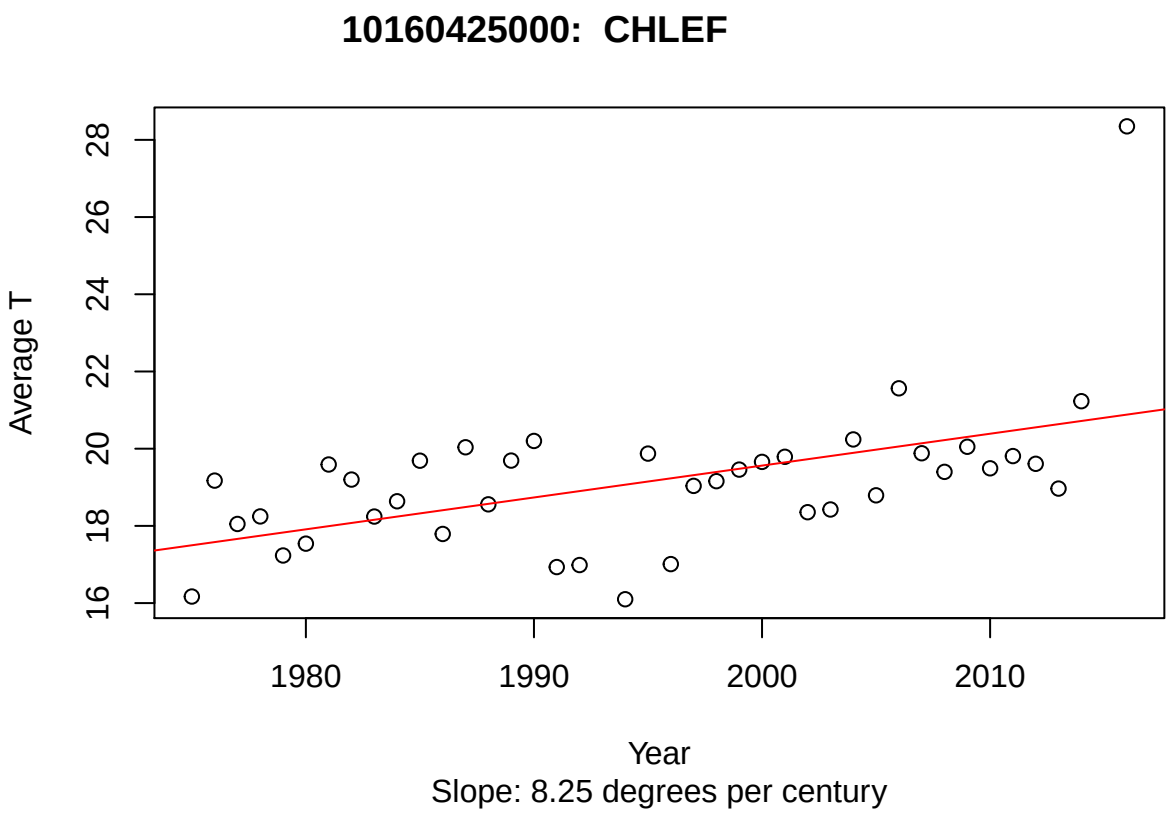
ID	NAME	LATITUDE	LONGITUDE
10160360000	ANNABA	36.83	7.82
10160461000	ORAN	35.70	-0.65
10160390000	DAR-EL-BEIDA	36.72	3.25
10160468000	BATNA	35.55	6.18
10160461000	ORAN	35.70	-0.65
10160360000	ANNABA	36.83	7.82
10160360000	ANNABA	36.83	7.82
10160461000	ORAN	35.70	-0.65

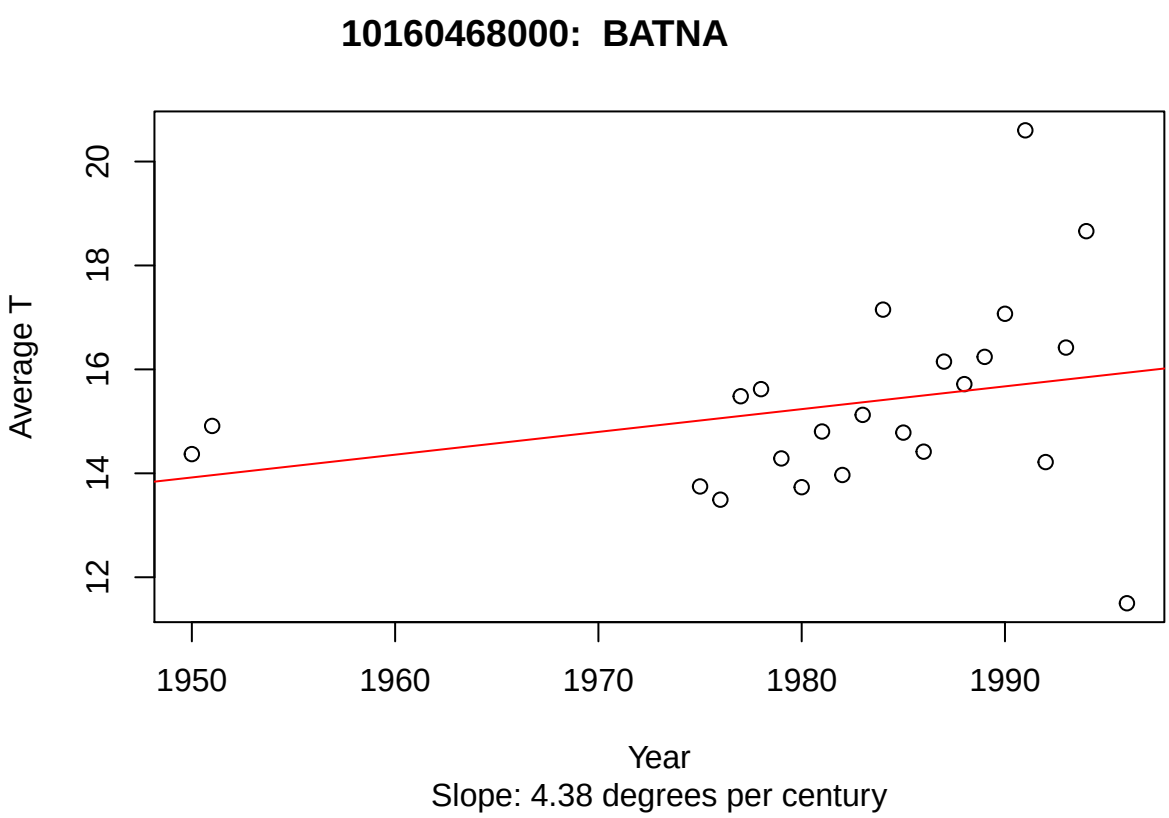
Details from each Station

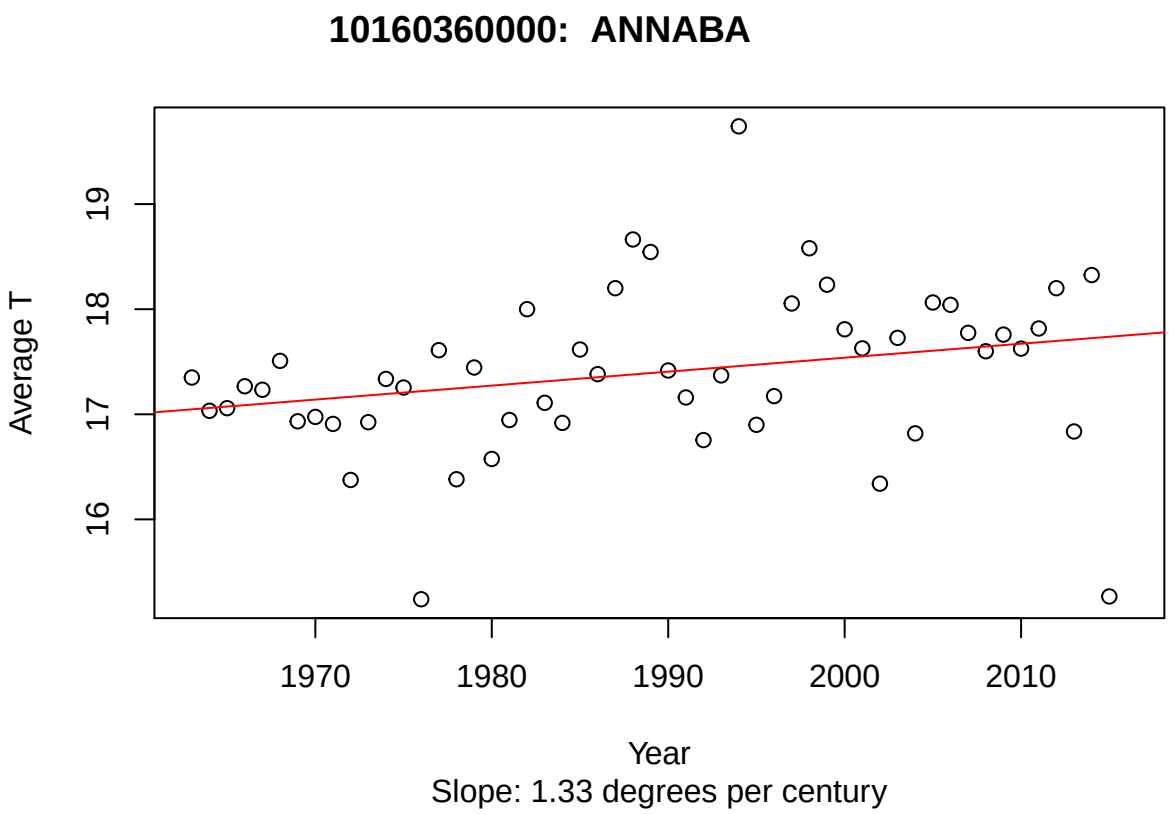


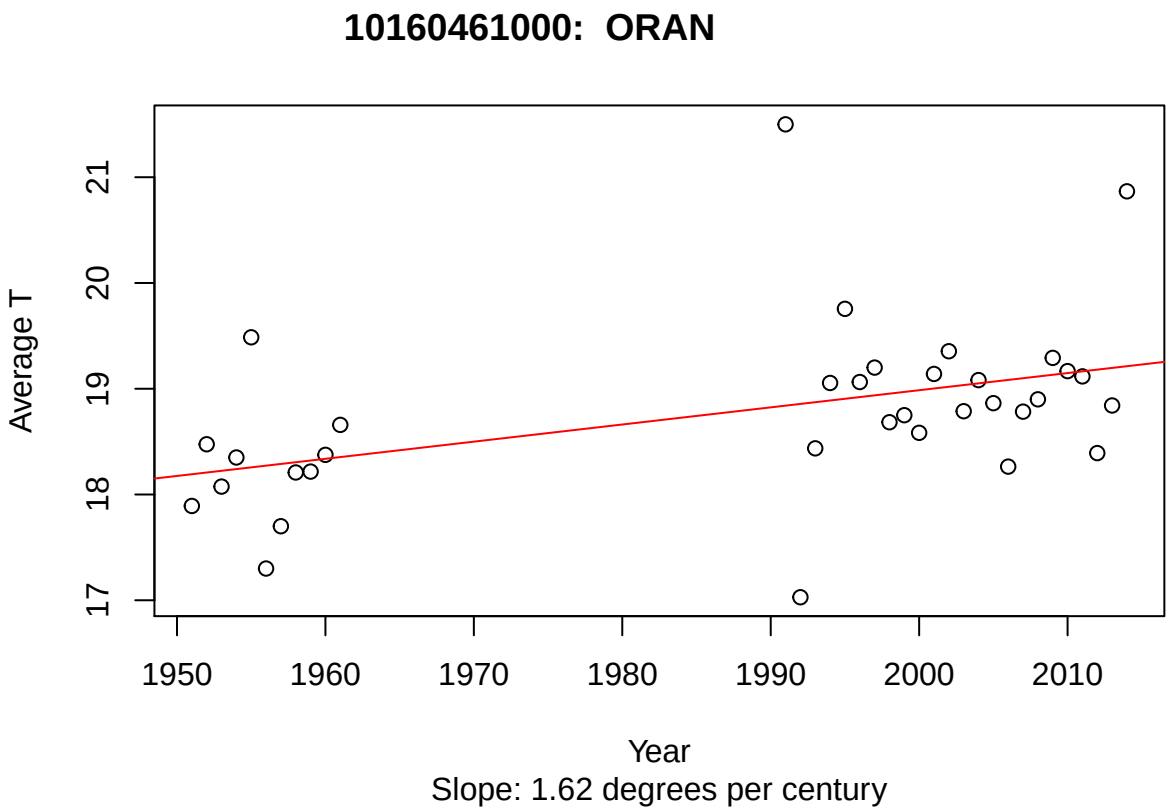

10160360000: ANNABA
Average T
16
17
18
19
1970
1980
1990
2000
2010
Year
Slope: 1.33 degrees per century

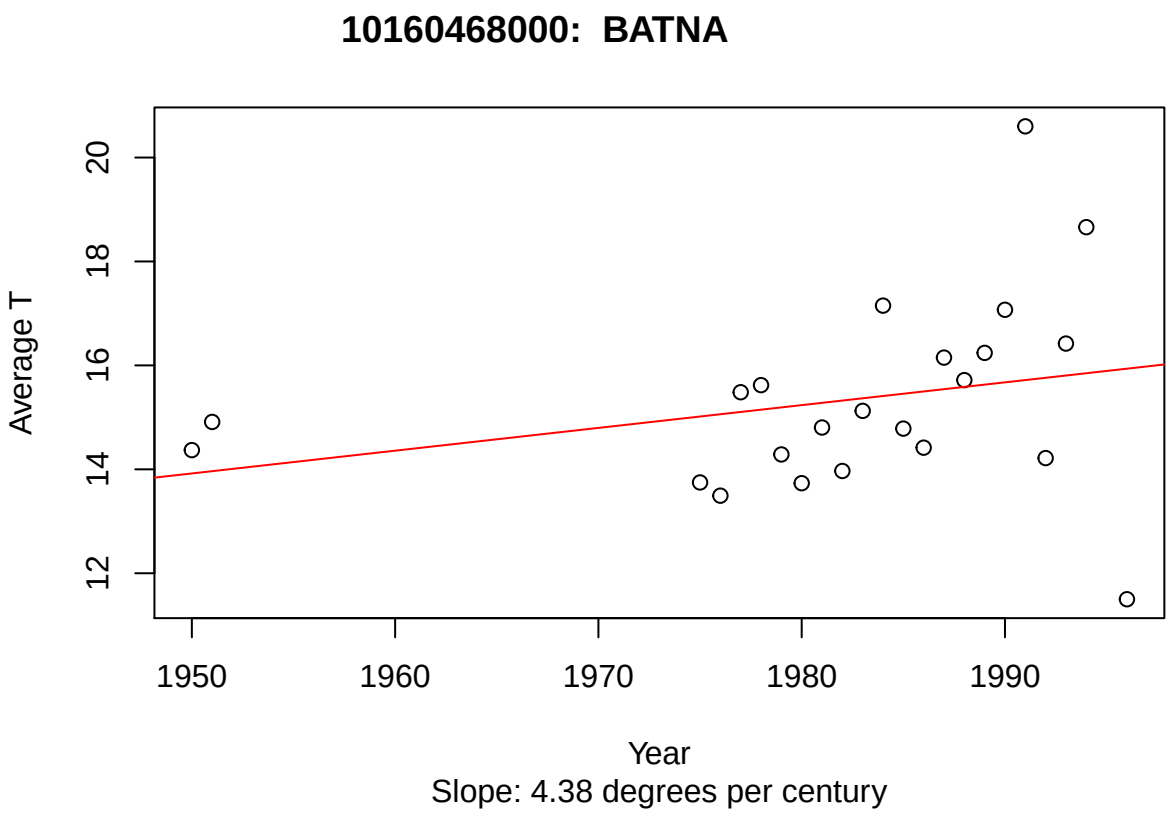

10160461000: ORAN
Average T
17
18
19
20
21
1950
1960
1970
1980
1990
2000
2010
Year
Slope: 1.62 degrees per century

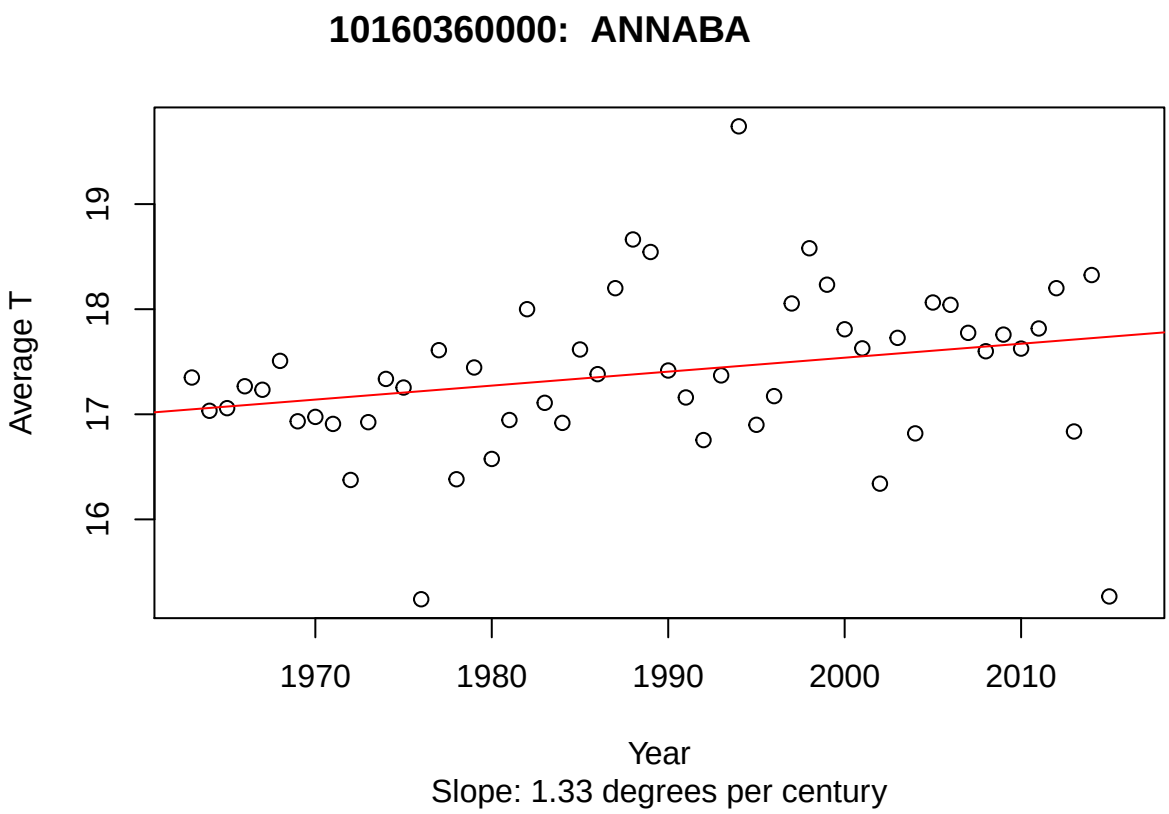

10160425000: CHLEF
Average T
16
18
20
22
24
26
28
1980
1990
2000
2010
Year
Slope: 8.25 degrees per century

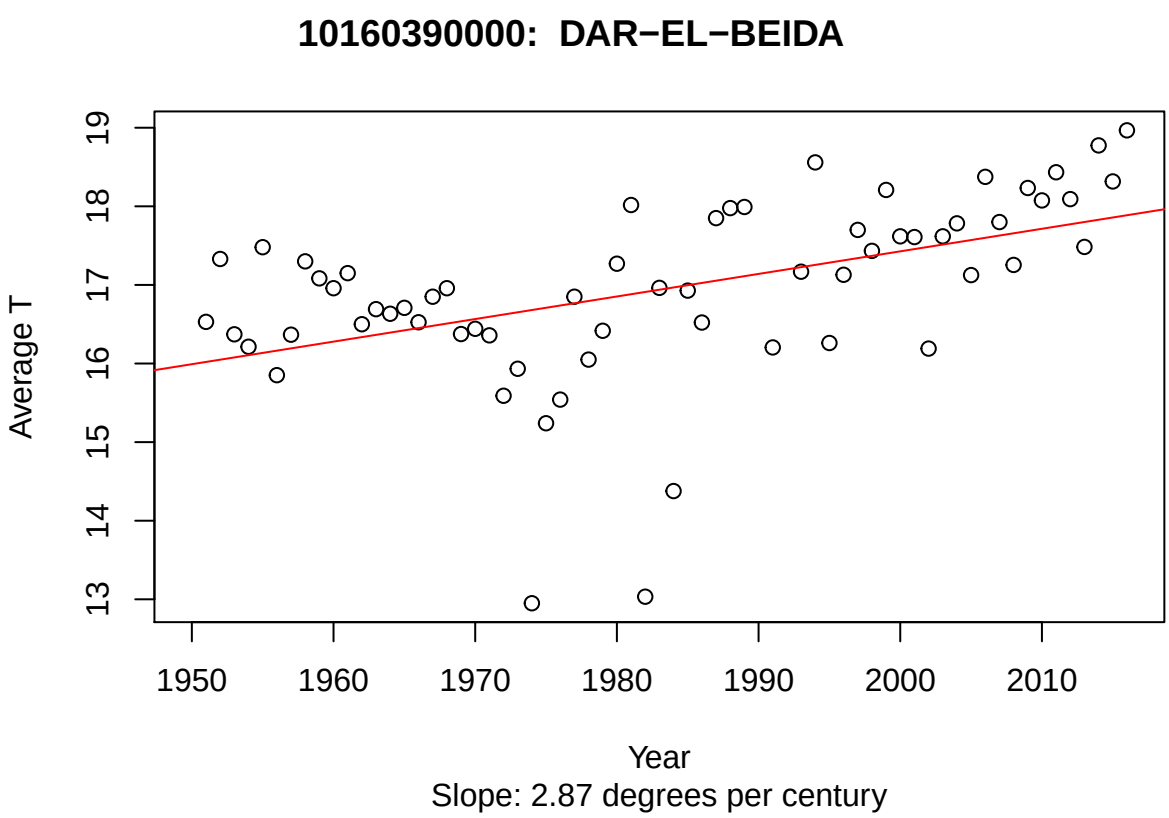

10160468000: BATNA
Average T
12
14
16
18
20
1950
1960
1970
1980
1990
Year
Slope: 4.38 degrees per century

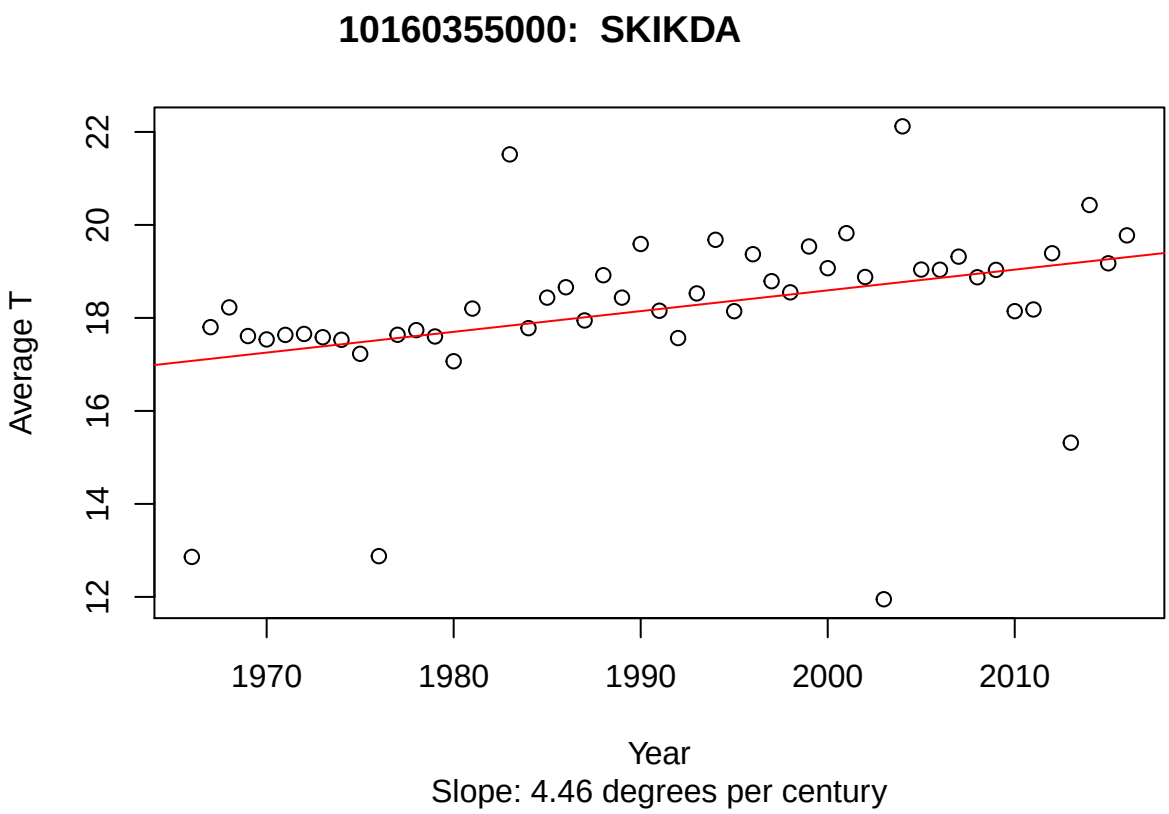

10160360000: ANNABA
Average T
16
17
18
19
1970
1980
1990
2000
2010
Year
Slope: 1.33 degrees per century

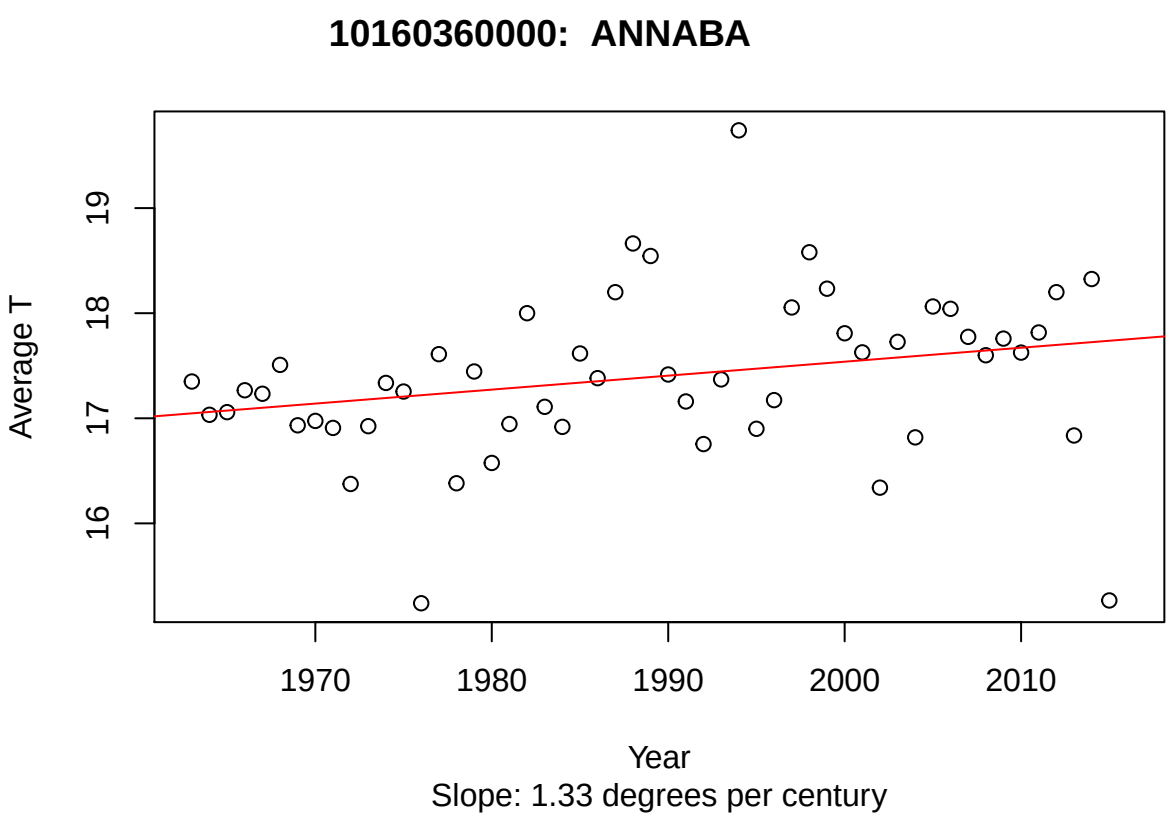

10160461000: ORAN
Average T
17
18
19
20
21
1950
1960
1970
1980
1990
2000
2010
Year
Slope: 1.62 degrees per century

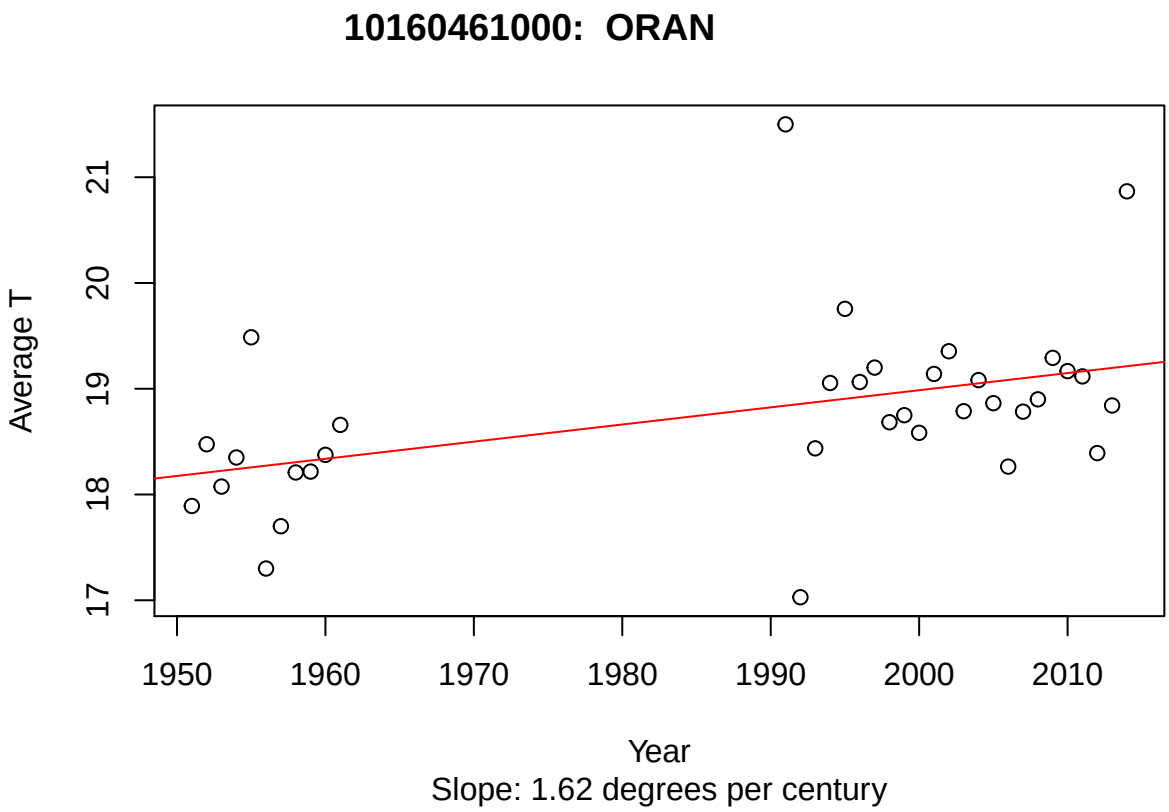

10160468000: BATNA
Average T
12
14
16
18
20
1950
1960
1970
1980
1990
Year
Slope: 4.38 degrees per century

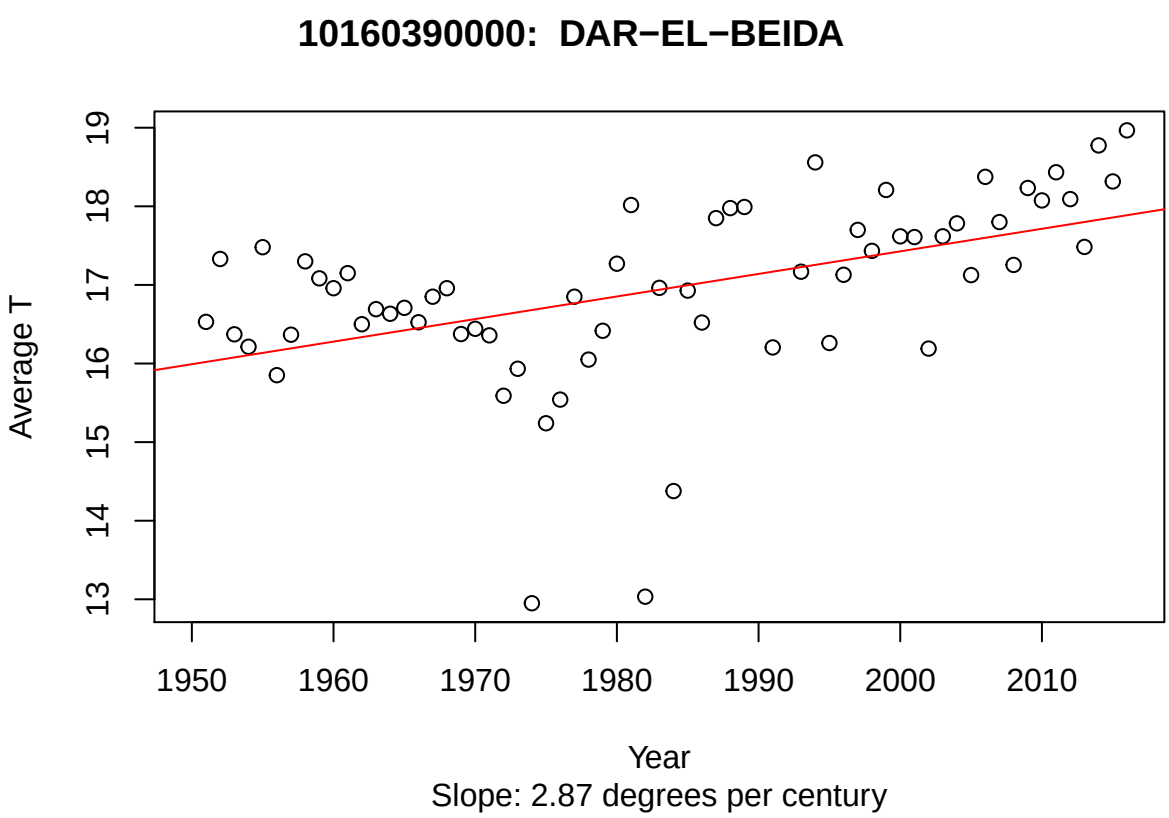

10160360000: ANNABA
Average T
16
17
18
19
1970
1980
1990
2000
2010
Year
Slope: 1.33 degrees per century

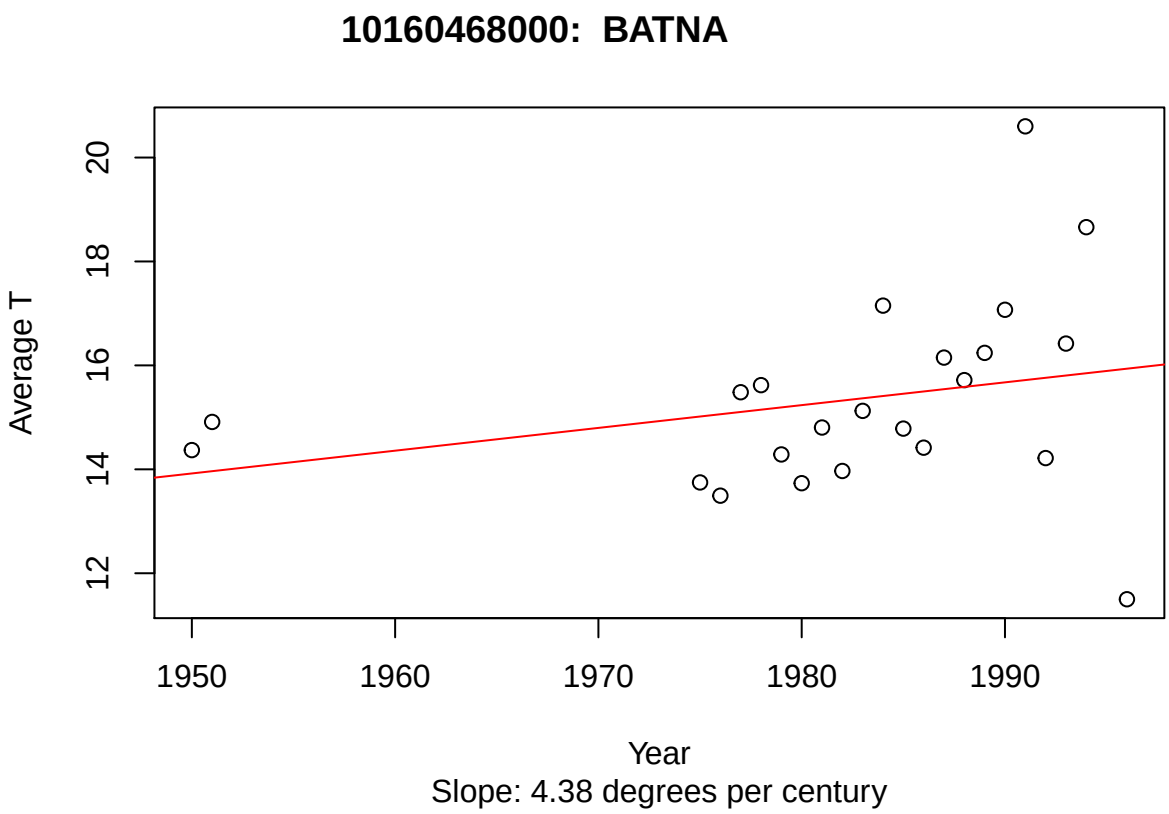

10160390000: DAR−EL−BEIDA
Average T
13
14
15
16
17
18
19
1950
1960
1970
1980
1990
2000
2010
Year
Slope: 2.87 degrees per century

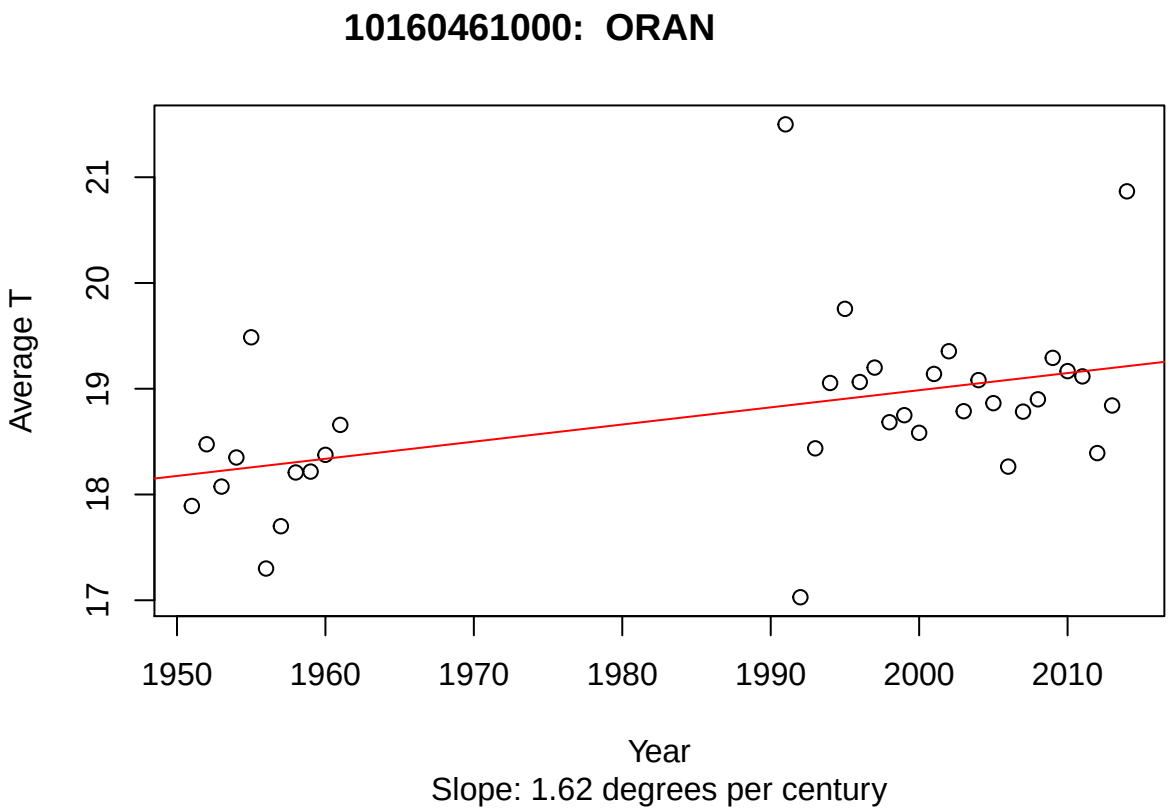

10160355000: SKIKDA
Average T
12
14
16
18
20
22
1970
1980
1990
2000
2010
Year
Slope: 4.46 degrees per century

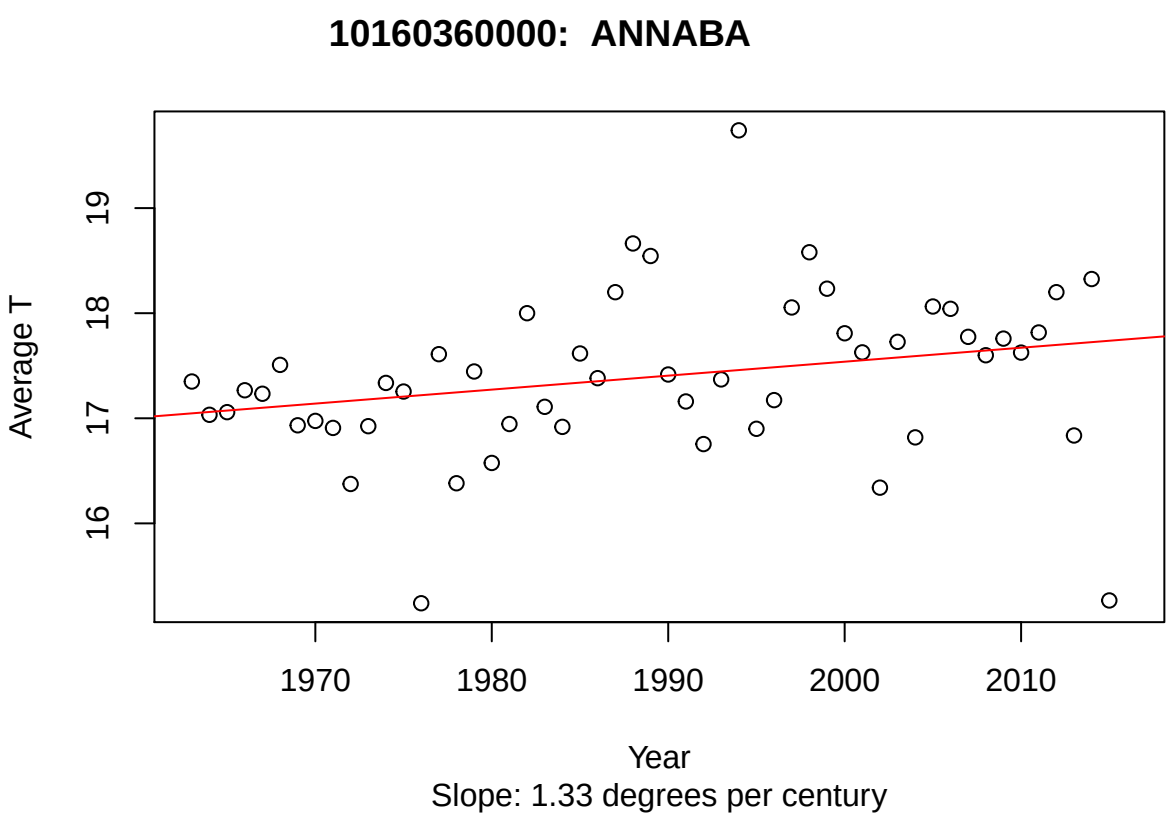

10160360000: ANNABA
Average T
16
17
18
19
1970
1980
1990
2000
2010
Year
Slope: 1.33 degrees per century

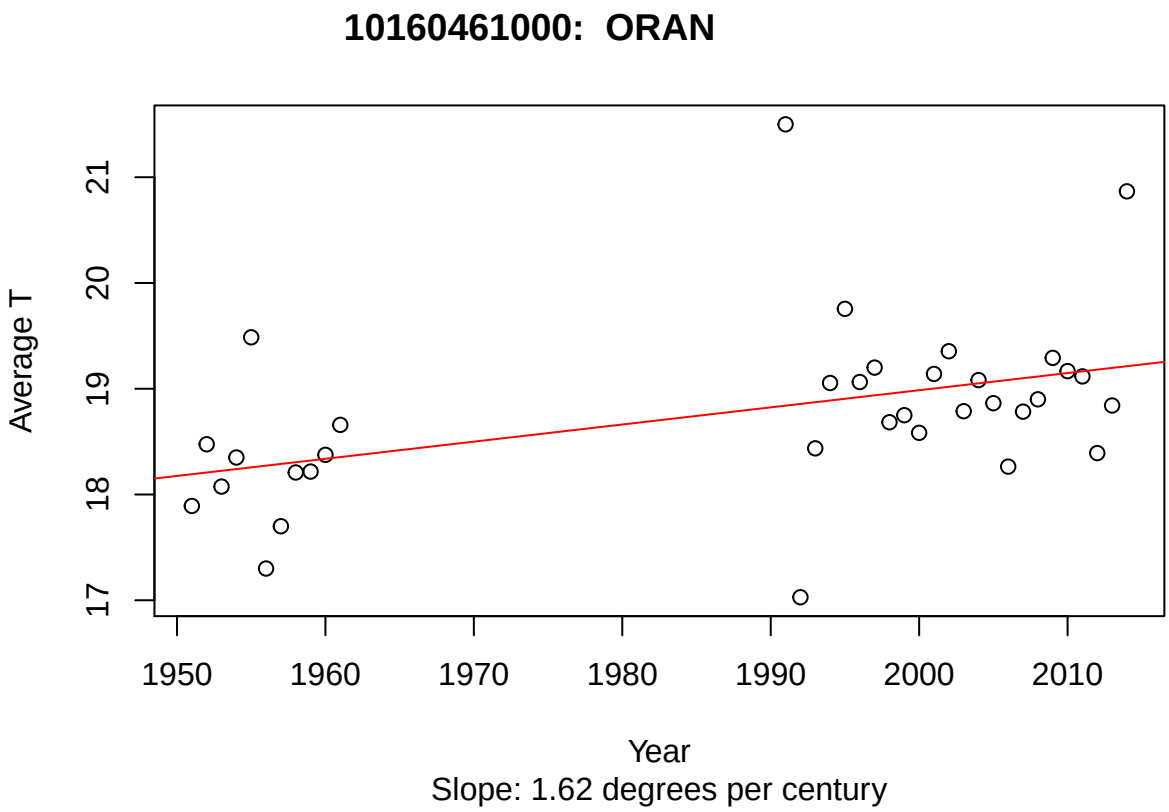

10160461000: ORAN
Average T
17
18
19
20
21
1950
1960
1970
1980
1990
2000
2010
Year
Slope: 1.62 degrees per century


10160390000: DAR−EL−BEIDA
Average T
13
14
15
16
17
18
19
1950
1960
1970
1980
1990
2000
2010
Year
Slope: 2.87 degrees per century


10160468000: BATNA
Average T
12
14
16
18
20
1950
1960
1970
1980
1990
Year
Slope: 4.38 degrees per century


10160461000: ORAN
Average T
17
18
19
20
21
1950
1960
1970
1980
1990
2000
2010
Year
Slope: 1.62 degrees per century


10160360000: ANNABA
Average T
16
17
18
19
1970
1980
1990
2000
2010
Year
Slope: 1.33 degrees per century



Conclusion

The temperature increased at 25 stations and decreased at 0 stations. This is in line with the results I have seen each time I have generated this report for a new random selection: there are far more stations with increases than decreases.

Future work

- Allow user to specify number of results needed since 1950 (currently script just requires some results).
- Allow user sp specify a threshold for increases, so we would only consider temperature to increasing if the slopw exceeded some value - say 0.5 degrees per century.
- Include a summary plot for all stations.
- Allow program to download latest data from NOAA, insetad of using canned data.