Lecture 10:
Iterative Tasks in R

www.terragalleria.com

Announcements

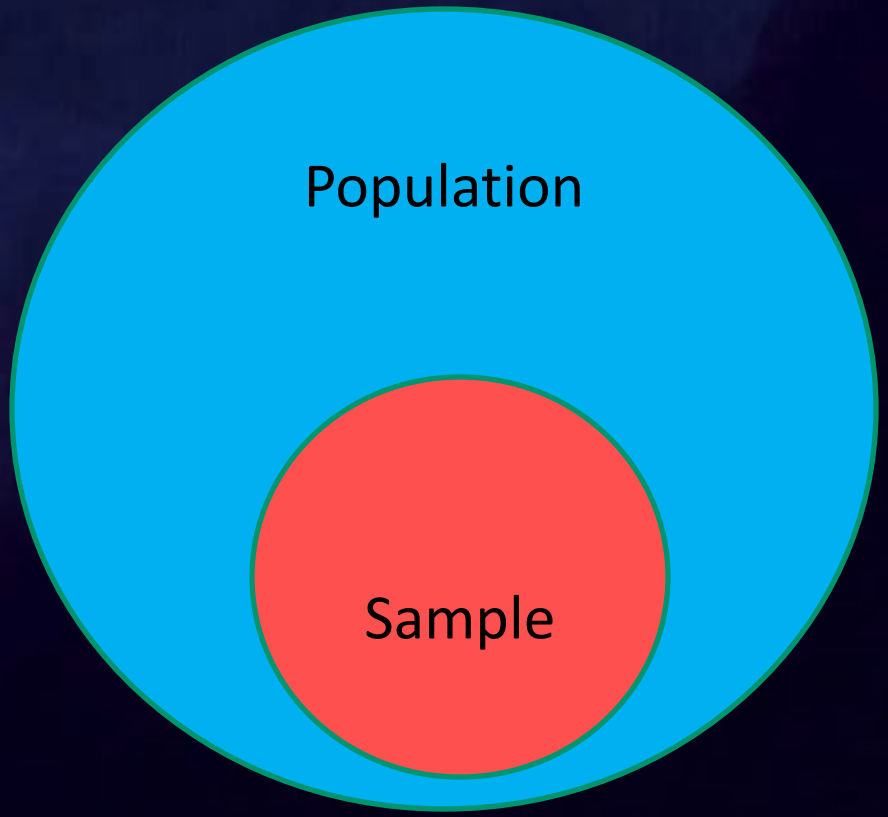
1. Assignment #1 is graded. Your grade is on eCampus. I provided feedback on the submitted eCampus assignment, so if you would like to see the feedback, contact the group member who submitted it.

2. Assignment #3 is now available on eCampus. It is due in one week (Tuesday, Feb. 26) so start making arrangements with your group of three.

Sample Size

- Number of samples needed depends on:
  - Type of Sample
  - Degree of precision
  - Level of confidence

- Remember that as sample size increases, time and effort also increase.
- → No significant benefit to sampling more than is statistically required!

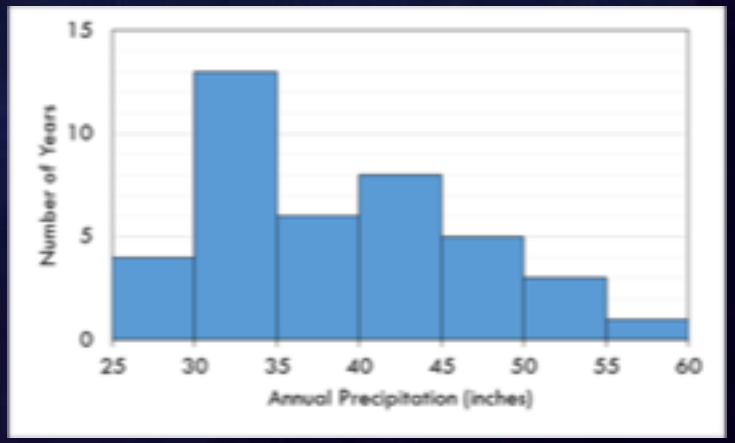


Sample Size

- How much **error (*E*)** are we willing to tolerate in our sample?
  - In our example: Say we want accuracy within 0.5 inches.
  - How may samples do we need?
- What level of confidence (*Z*) do we want to have?
- What is the standard deviation of the sample?

Example: Annual Precipitation in Washington, DC over 40 years

Ordered Sample			
26”	35”	39”	45”
26”	35”	40”	46”
28”	35”	40”	47”
29”	36”	41”	47”
32”	36”	41”	48”
32”	36”	41”	50”
33”	36”	41”	51”
33”	38”	41”	51”
34”	39”	43”	54”
35”	39”	43”	57”



Mean = 39.9 inches

Standard Deviation = 7.5 inches

Sample Size

- How many samples if we want 0.5” inches of error at 90% confidence?
- Standard Deviation = 7.5 inches
- Z = 1.65 (90% confidence)

$$n = \left(\frac{Zs}{E} \right)^2 = \left(\frac{1.65 * 7.5}{0.5} \right)^2 = 612.5$$

- Our sample does not have enough samples (40 samples < 612.5)
- Only enough if we are willing to tolerate 2 inches of error
- Note: *E* is independent of confidence intervals

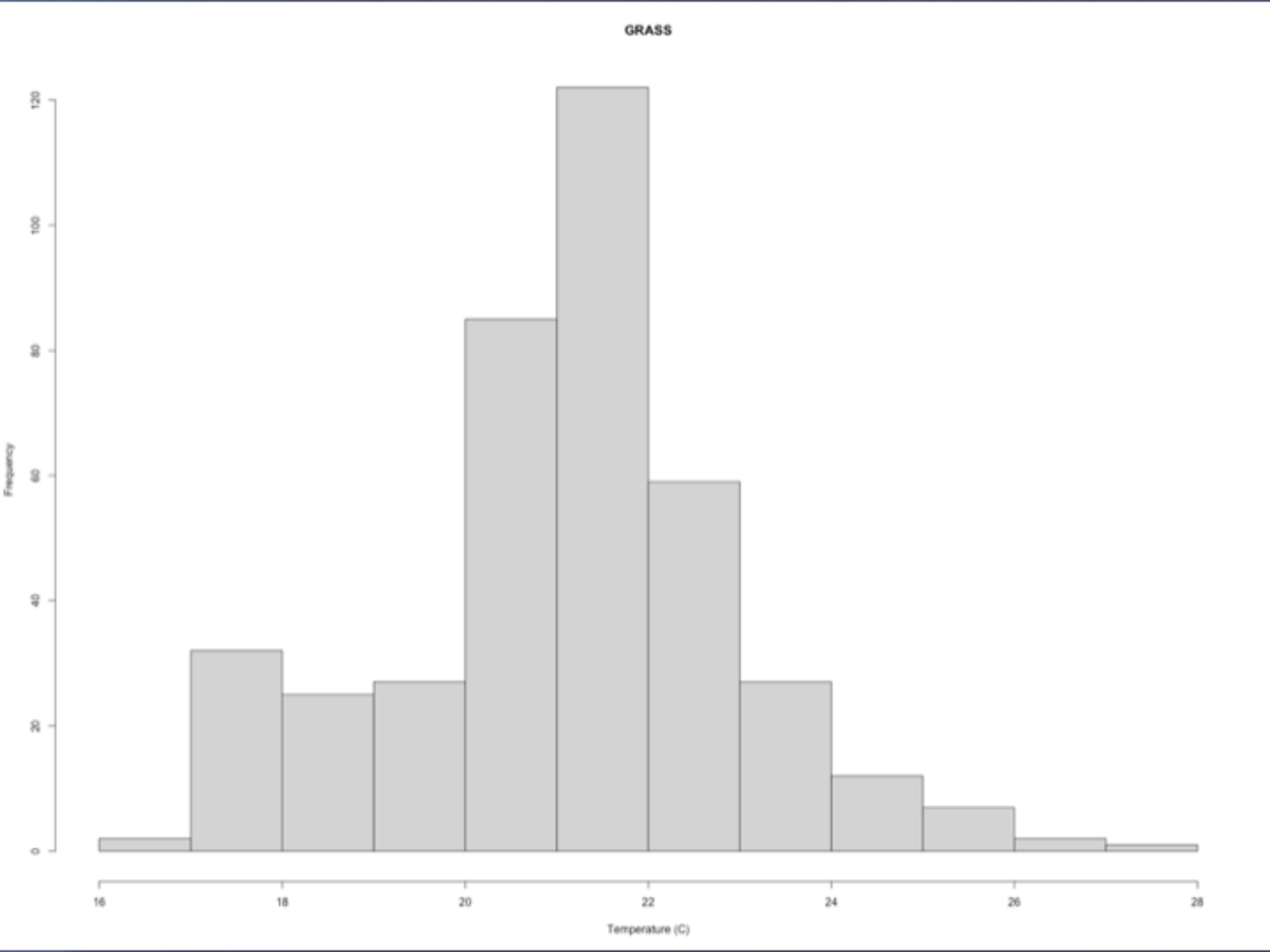
Strategy: Take a Pre-Sample

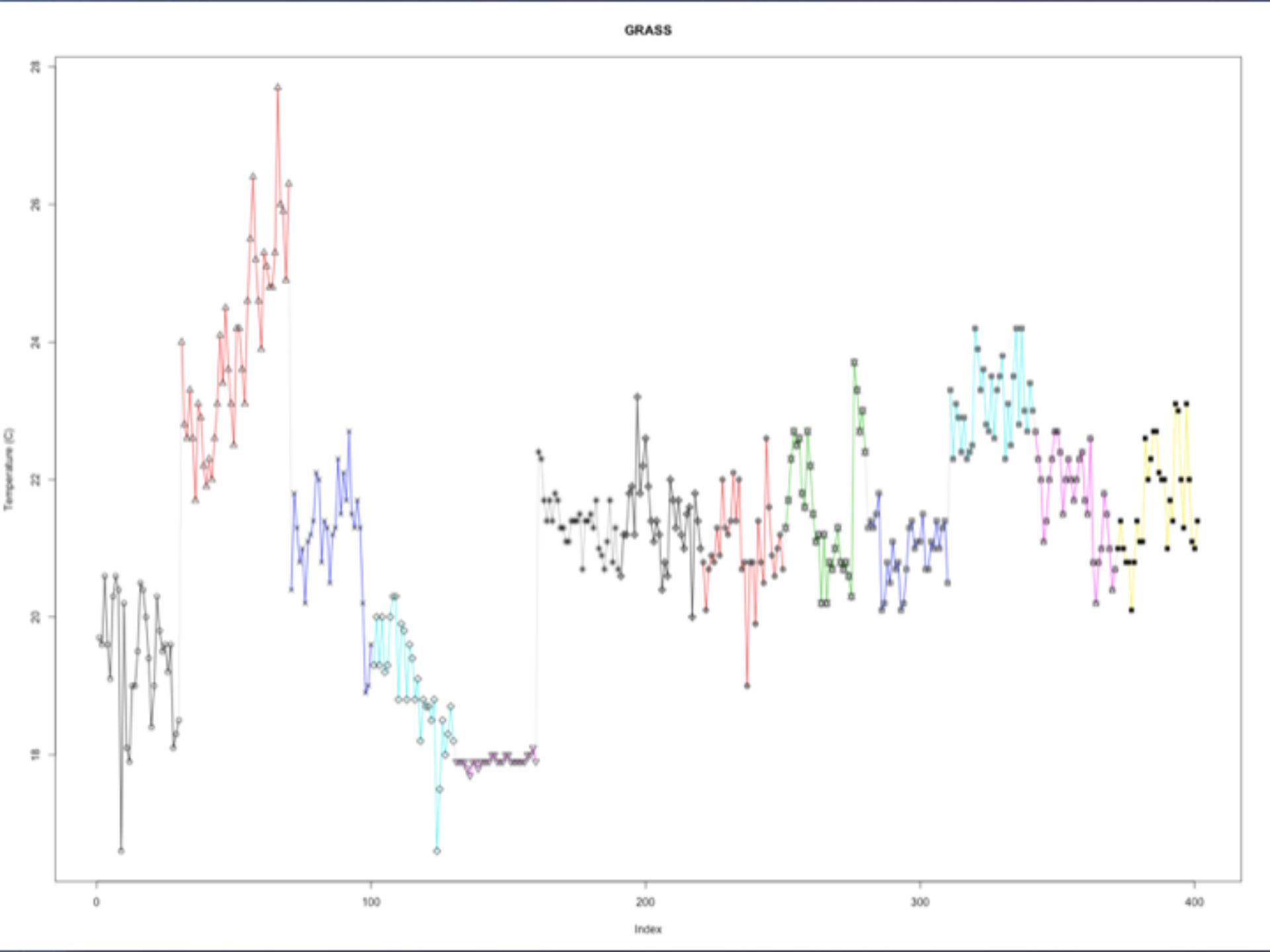
- Use a small set of samples to get an idea of the standard deviation
  - If possible, a pre-sample should include at least **30 samples**

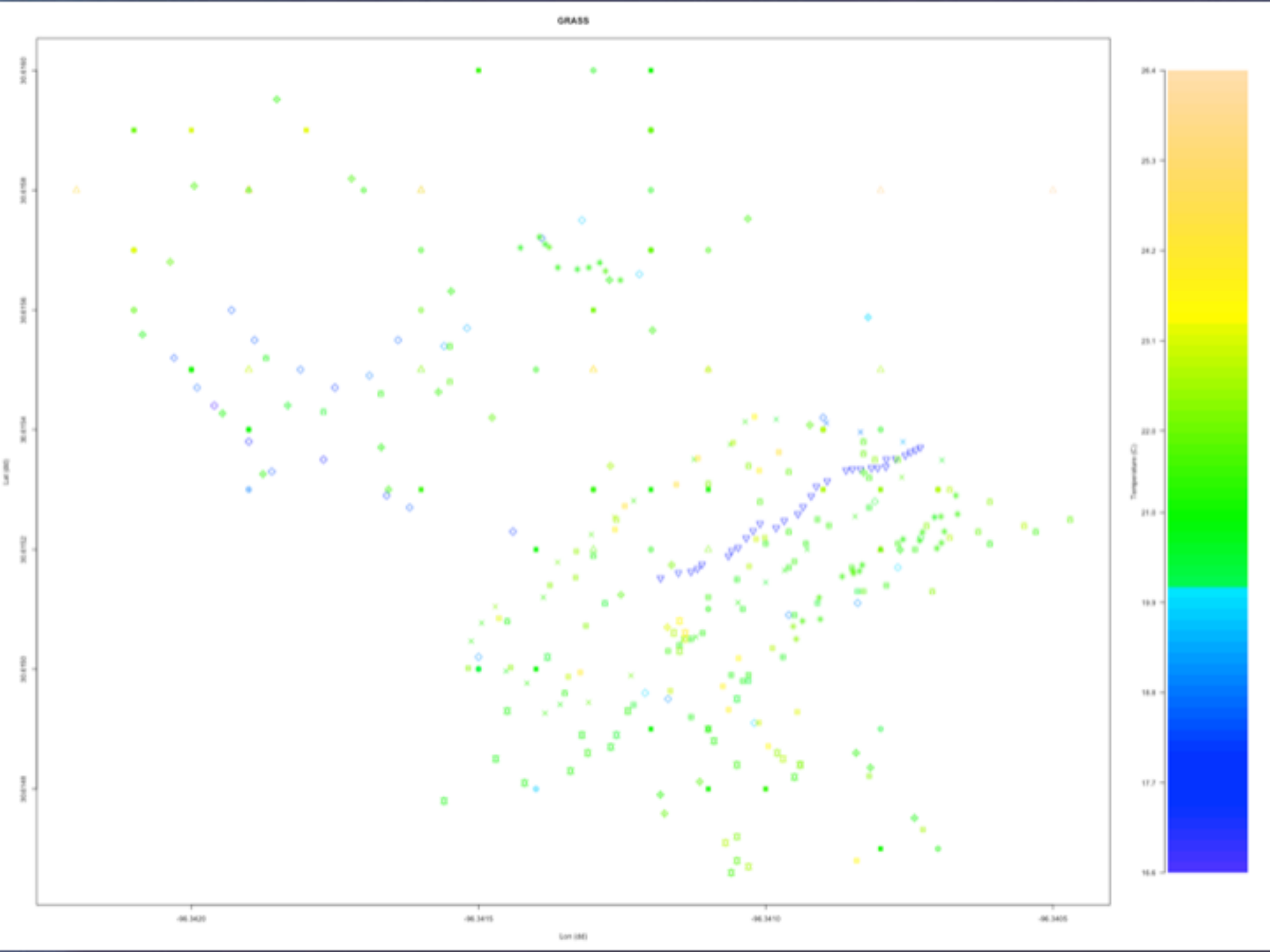
- **Compute error to determine if adequate**

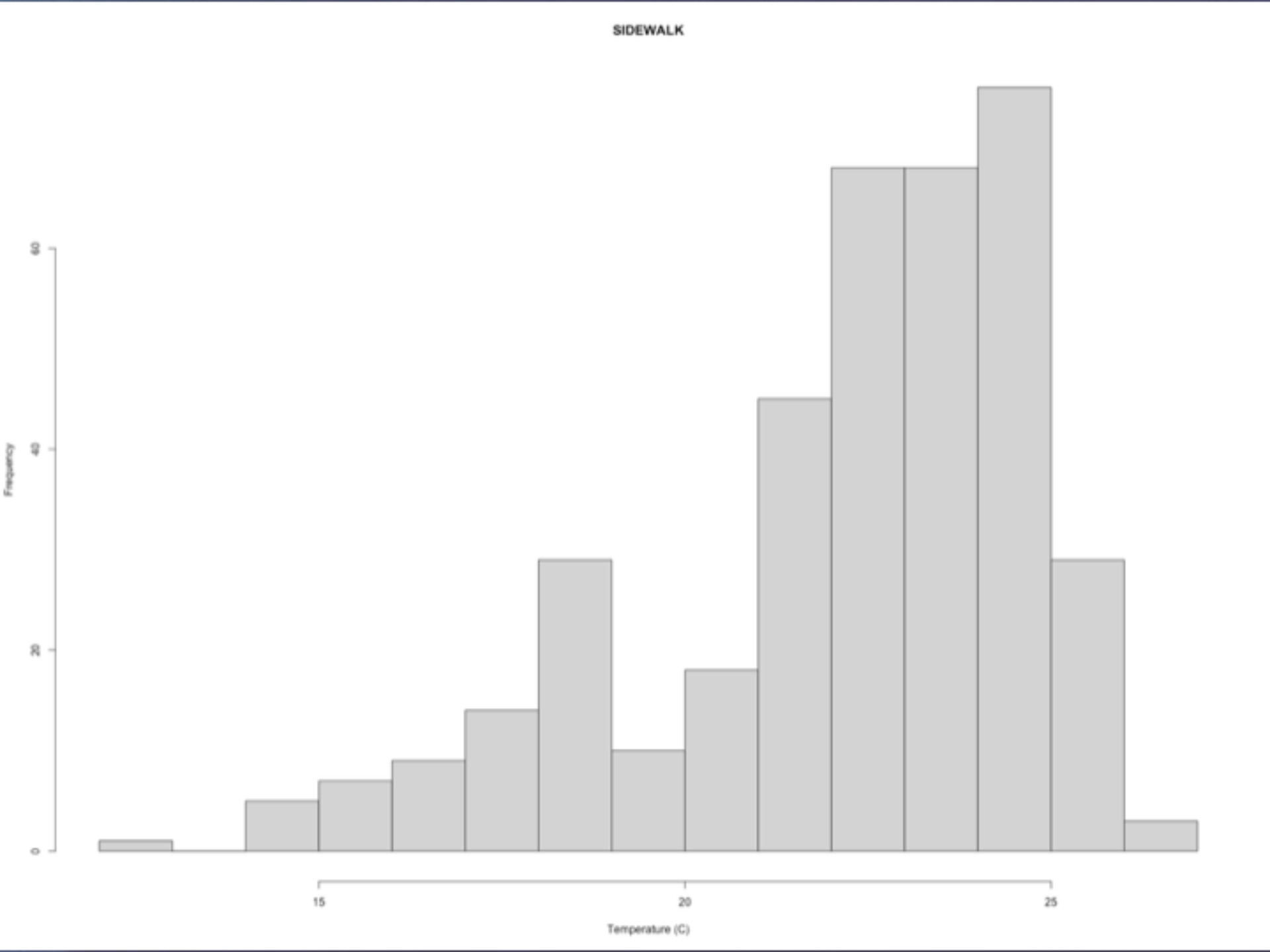
Temperature Data from Assignment #1

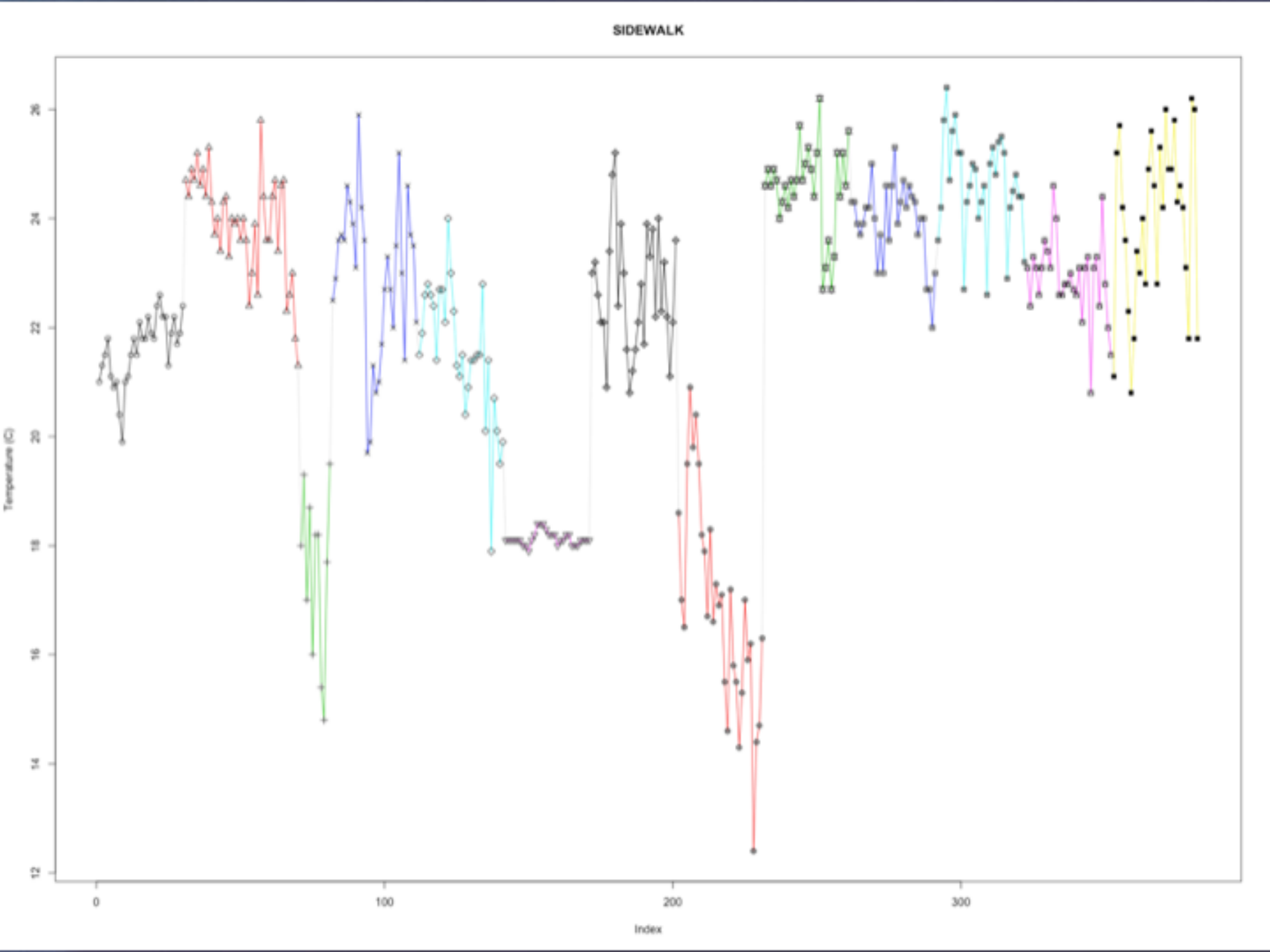
Let’s open up R…

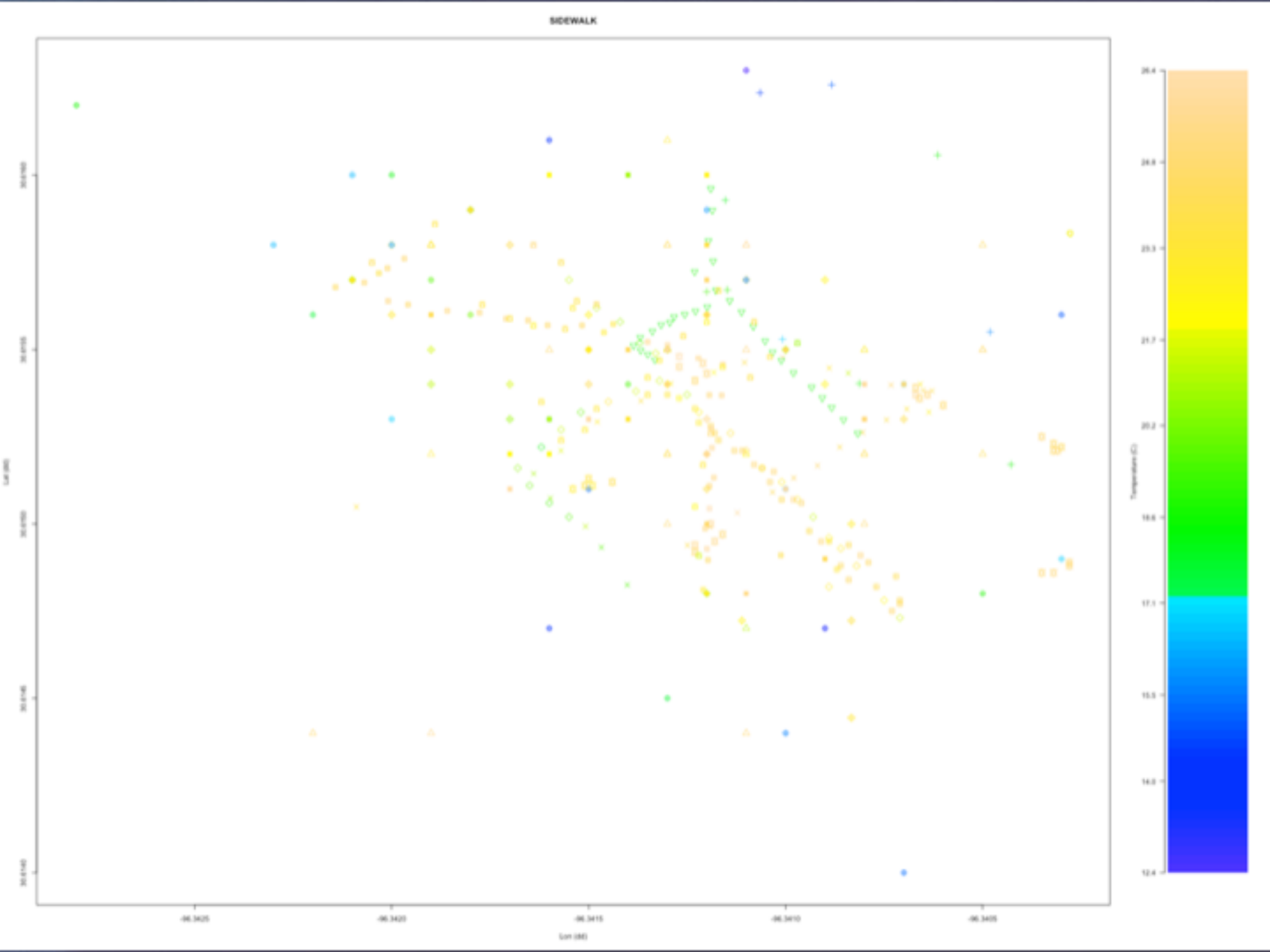

GRASS
Frequency
0
20
40
60
80
100
120
16
18
20
22
24
26
28
Temperature (C)

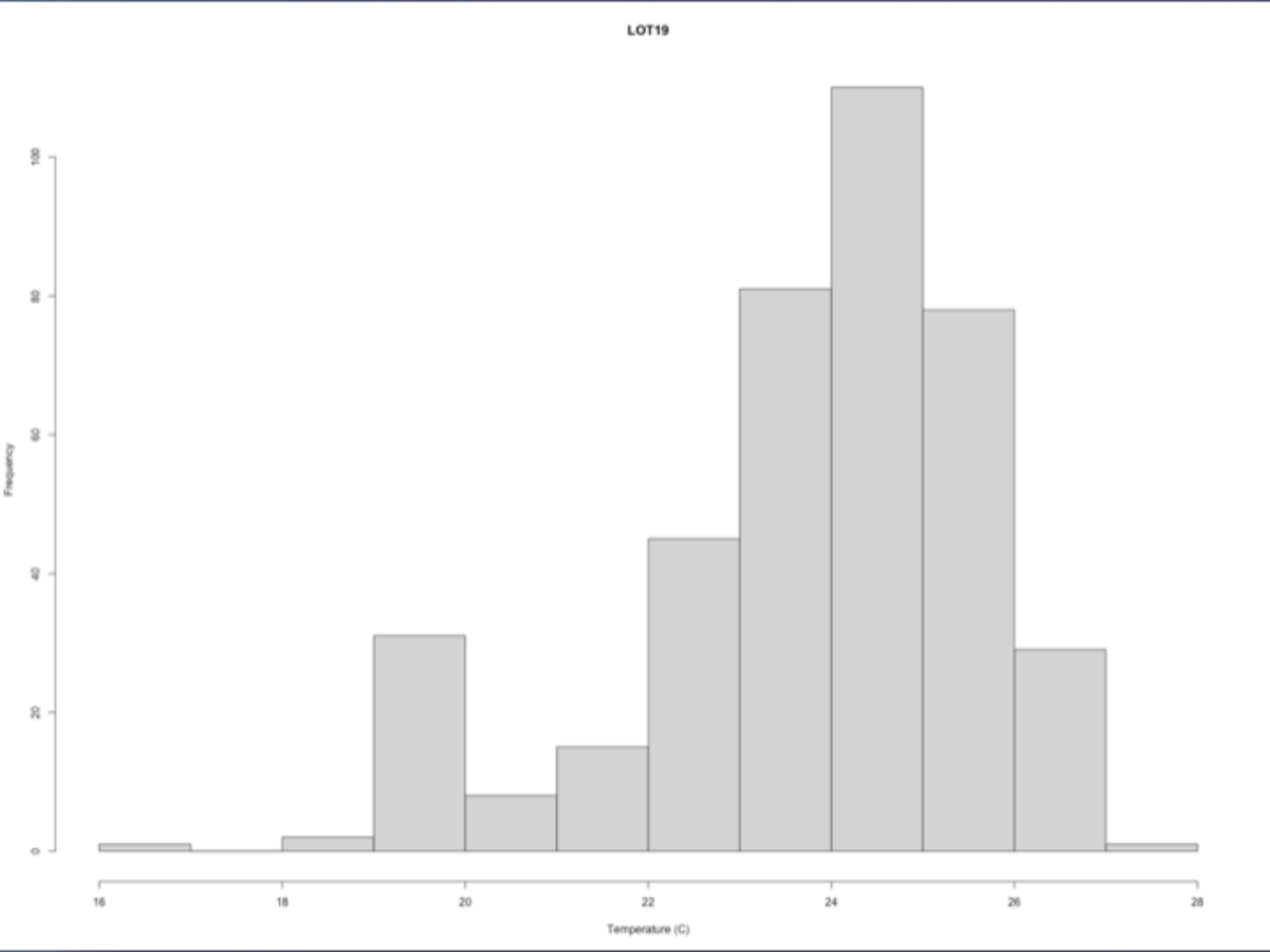

GRASS
Temperature (C)
28
26
24
22
20
18
0
100
200
300
400
Index

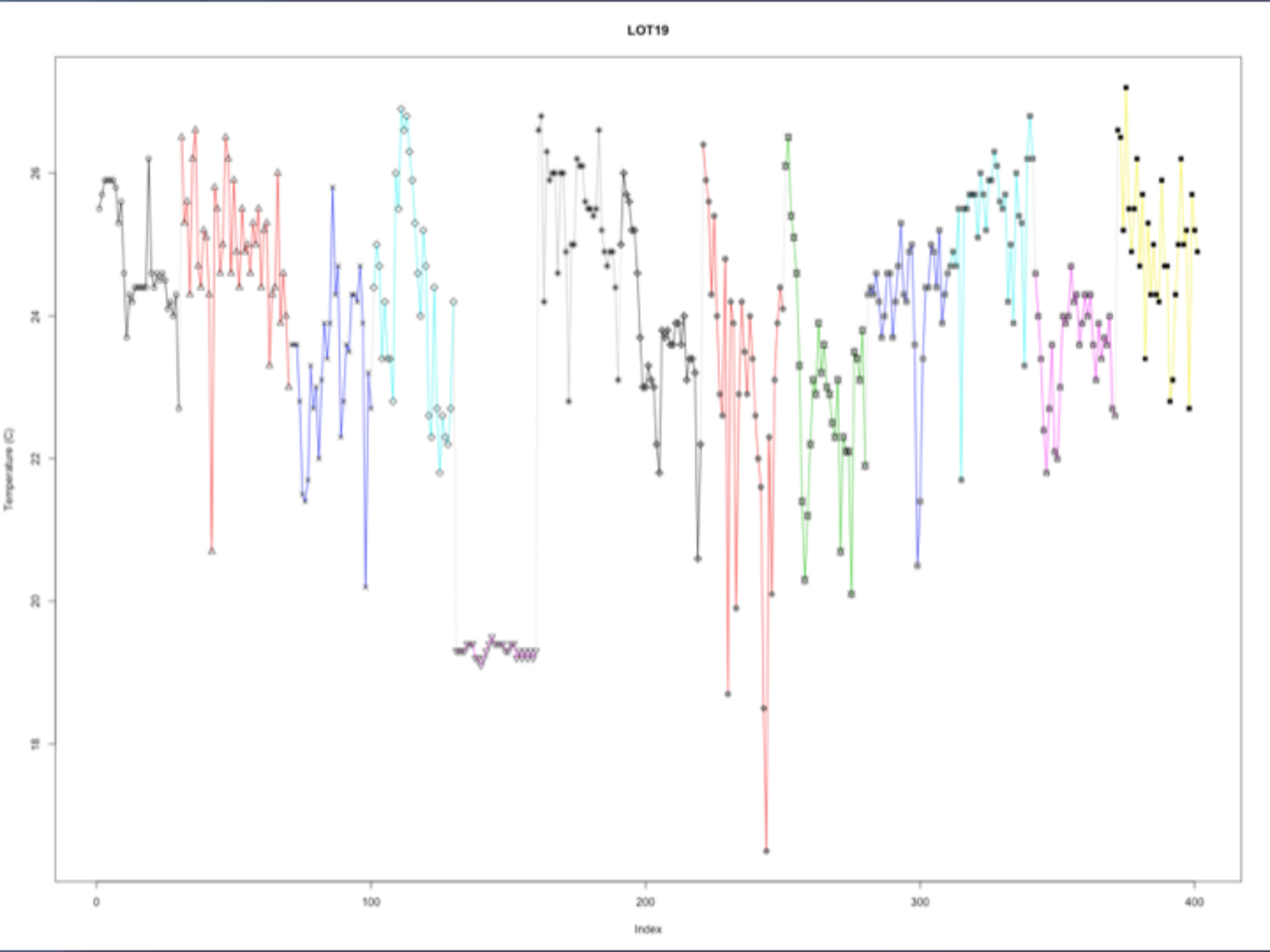

GRASS
30.6160
30.6158
30.6156
30.6154
30.6152
30.6150
30.6148
Lat (dd)
-86.3420
-86.3415
-86.3410
-86.3405
Lon (dd)
26.4
25.3
24.2
23.1
22.0
21.0
19.9
18.8
17.7
16.6
Temperature (C)

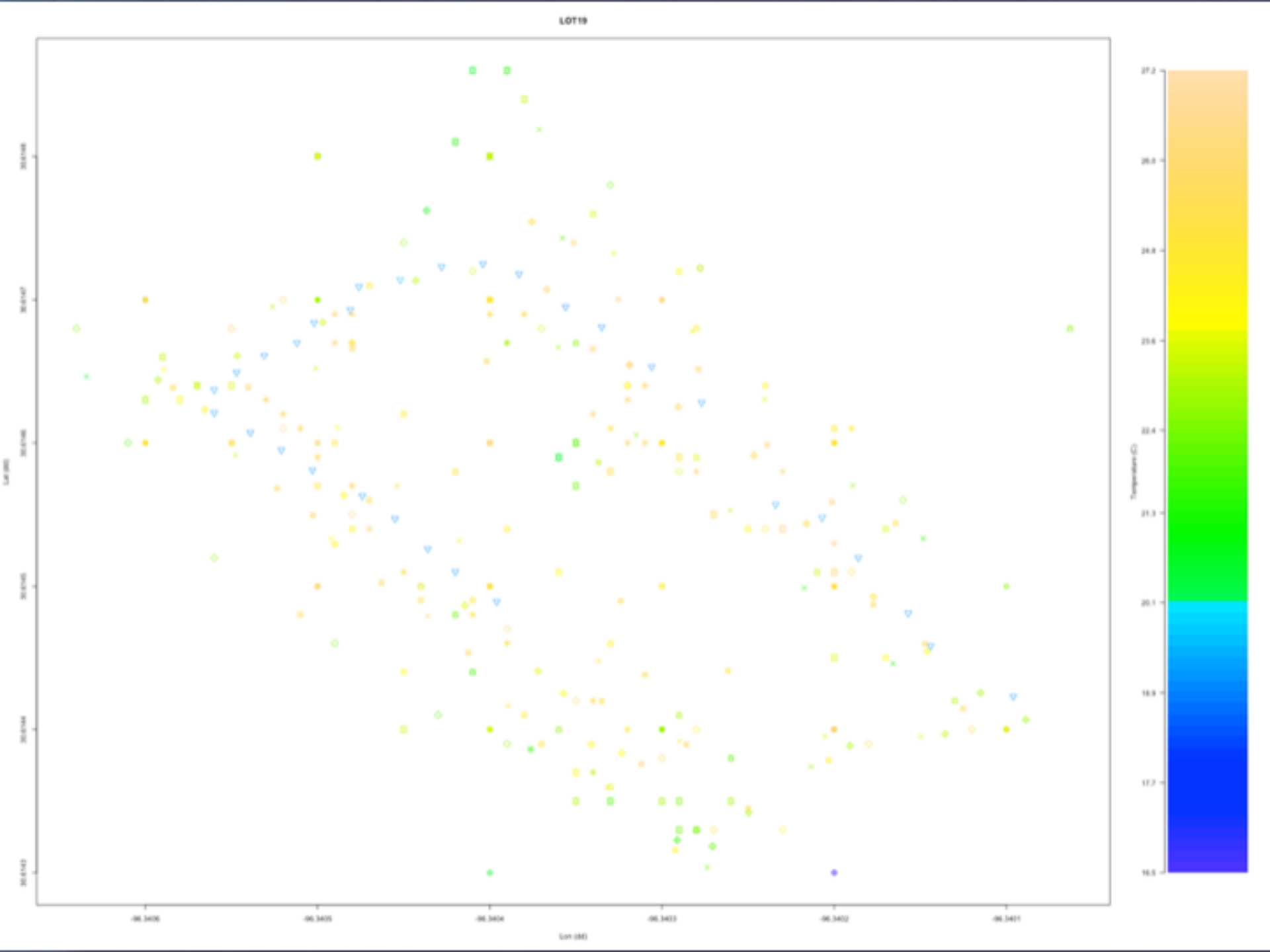

SIDEWALK
Frequency
0
20
40
60
15
20
25
Temperature (C)

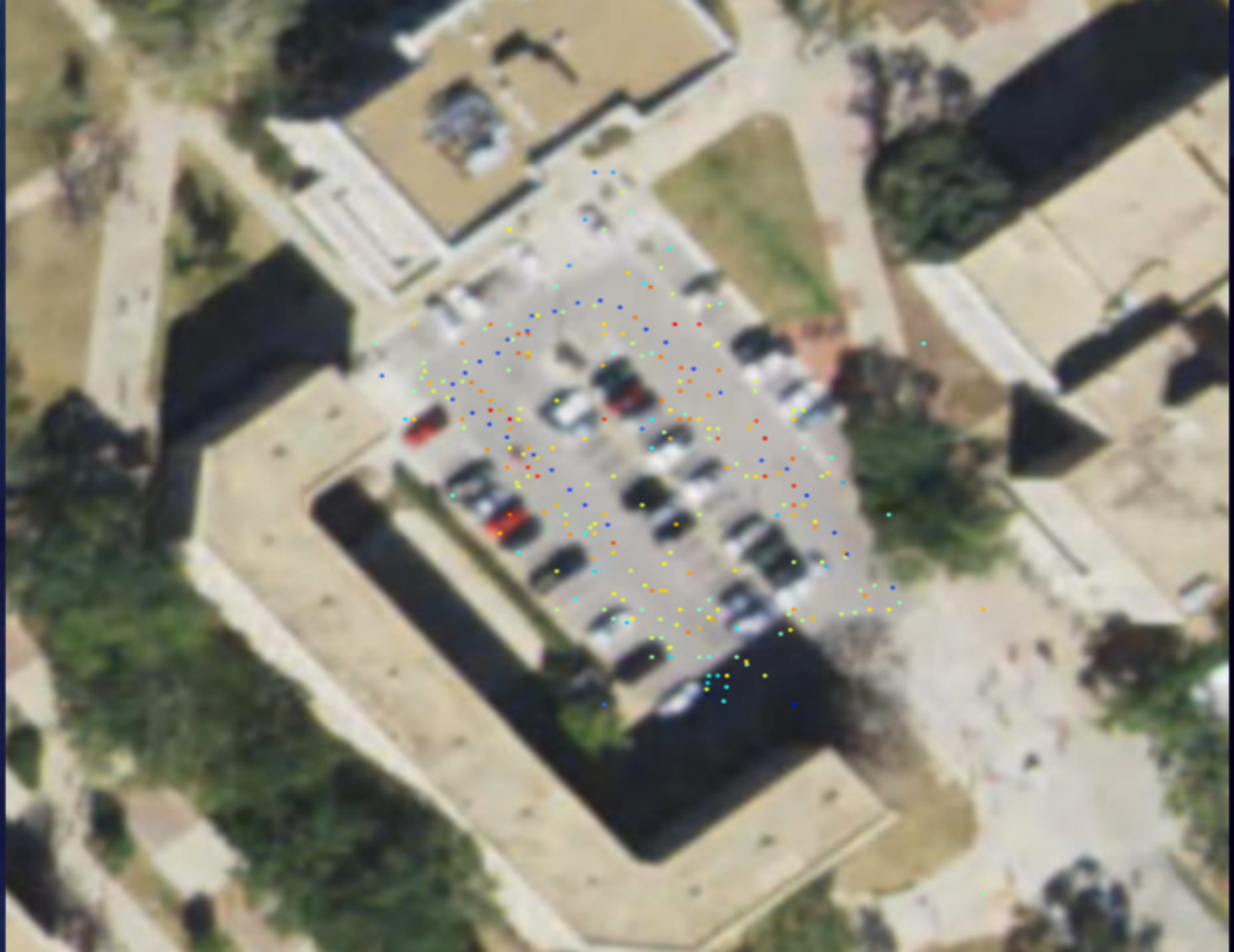

SIDEWALK
Temperature (C)
26
24
22
20
18
16
14
12
0
100
200
300
Index


SIDEWALK
30.6160
30.6155
30.6150
30.6145
30.6140
Lat (dd)
-96.3425
-96.3420
-96.3415
-96.3410
-96.3405
Lon (dd)
26.4
24.8
23.3
21.7
20.2
18.6
17.1
15.5
14.0
12.4
Temperature (C)


LOT19
Frequency
0
20
40
60
80
100
16
18
20
22
24
26
28
Temperature (C)


LOT19
Temperature (C)
18
20
22
24
26
0
100
200
300
400
Index


LOT19
30.6143
30.6144
30.6145
30.6146
30.6147
30.6148
Lat (dd)
-96.3406
-96.3405
-96.3404
-96.3403
-96.3402
-96.3401
Lon (dd)
27.2
26.0
24.8
23.6
22.4
21.3
20.1
18.9
17.7
16.5
Temperature (C)



Lecture 10:
Iterative Tasks in R

www.terragalleria.com