HW7 DATA 120

1.

	Full Data				Sample Data			
	X	Y	Z	Intensity	X	Y	Z	Intensity
count	14542343.000000	14542343.000000	14542343.000000	14542343.000000	1000.000000	1000.000000	1000.000000	1000.000000
mean	1188547.759321	1871290.642155	619.603483	14745.508808	1188566.028730	1871305.531060	618.906800	14576.535000
std	628.497279	704.344796	49.383413	14996.561529	623.551864	697.669716	45.977396	15062.408272
min	1187500.000000	1870000.000000	574.650000	0.000000	1187500.970000	1870000.250000	579.440000	43.000000
25%	1188030.440000	1870702.320000	591.000000	3820.000000	1188056.665000	1870743.367500	591.267500	3707.250000
50%	1188495.620000	1871331.660000	608.800000	7598.000000	1188513.760000	1871325.590000	609.160000	7435.500000
75%	1188995.190000	1871892.620000	626.920000	23246.000000	1189016.092500	1871900.782500	627.440000	22900.750000
max	1190000.000000	1872500.000000	953.140000	64877.000000	1189996.230000	1872497.080000	931.160000	57768.000000

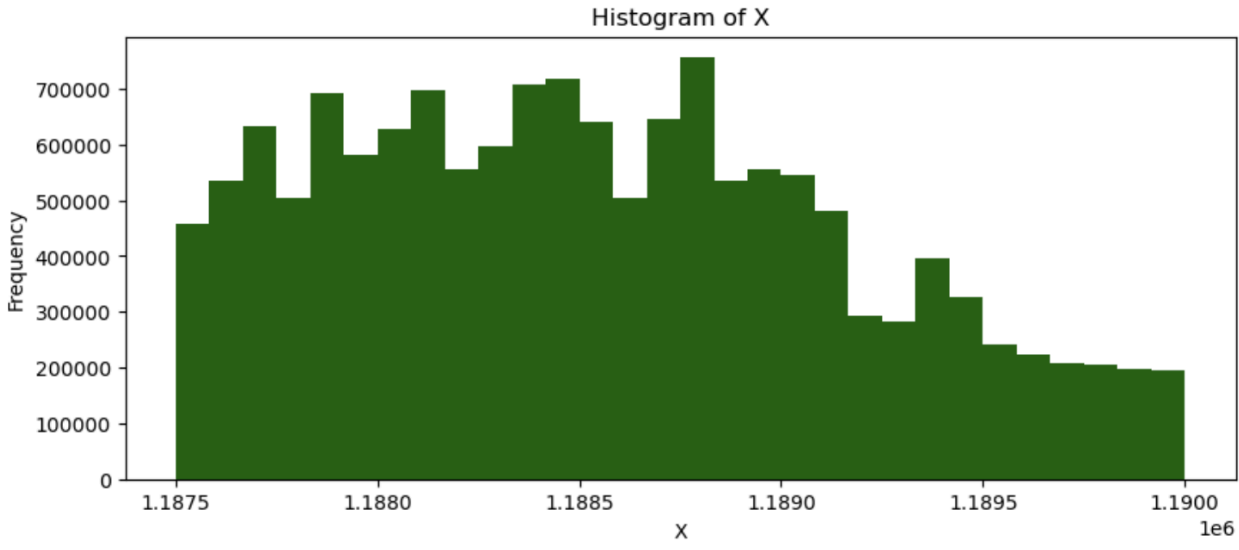
2.

```
Correlation Matrix:
                   X         Y         Z  Intensity
X           1.000000 -0.082429 -0.250978  -0.014268
Y          -0.082429  1.000000  0.111968  -0.133896
Z          -0.250978  0.111968  1.000000  -0.124210
Intensity  -0.014268 -0.133896 -0.124210   1.000000
```

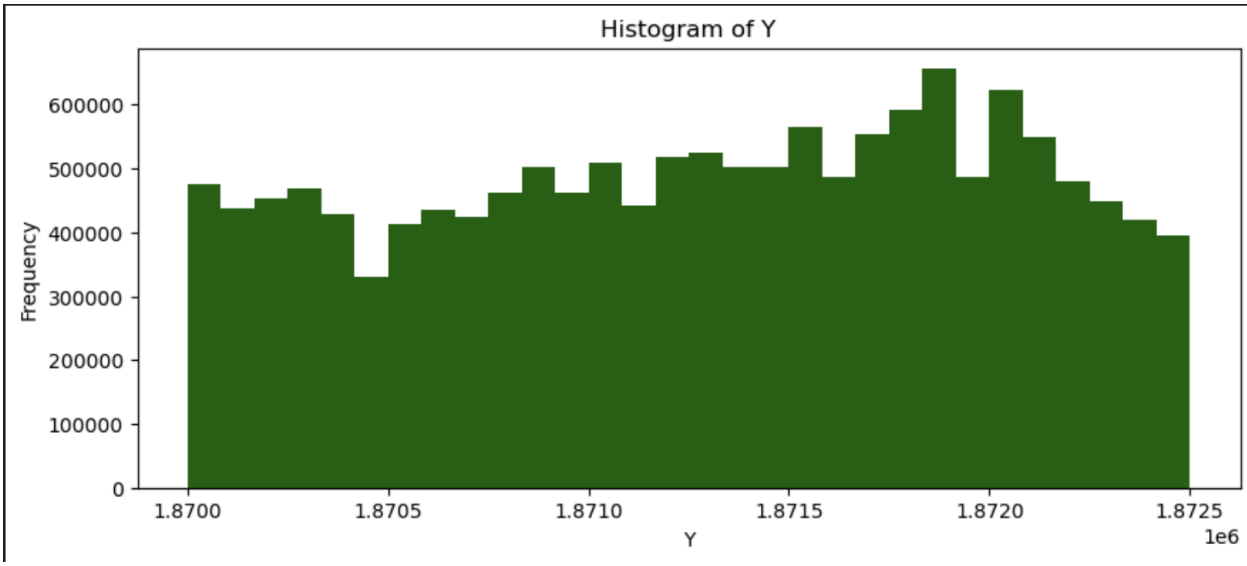
   a. There are no clearly significant correlations. None of the values are more than 0.5 from 0.

3.

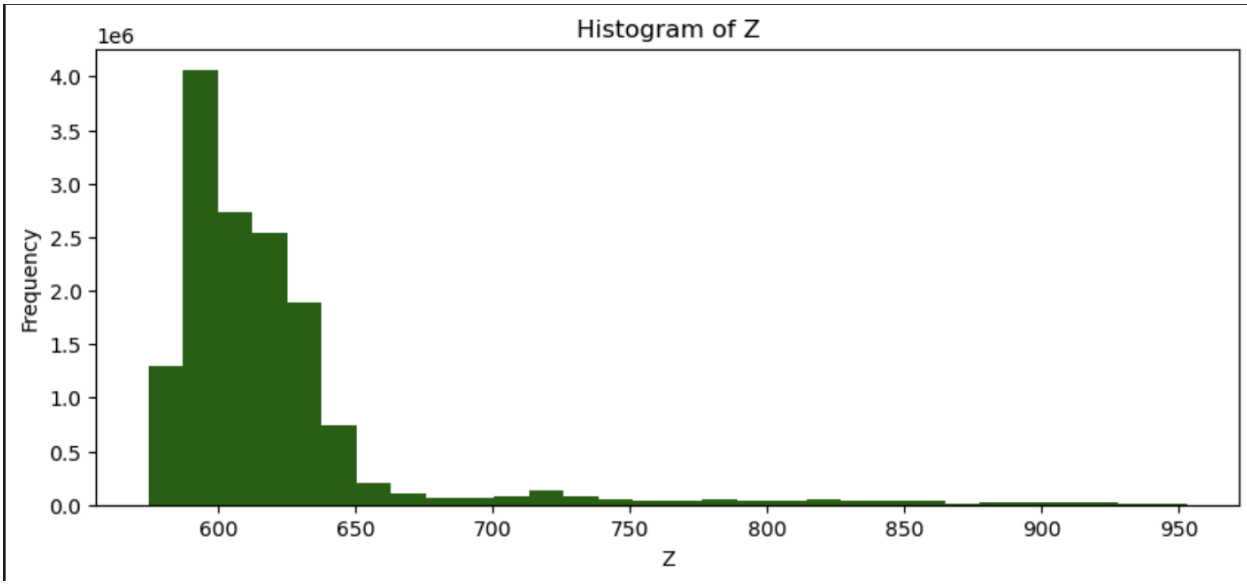
a.

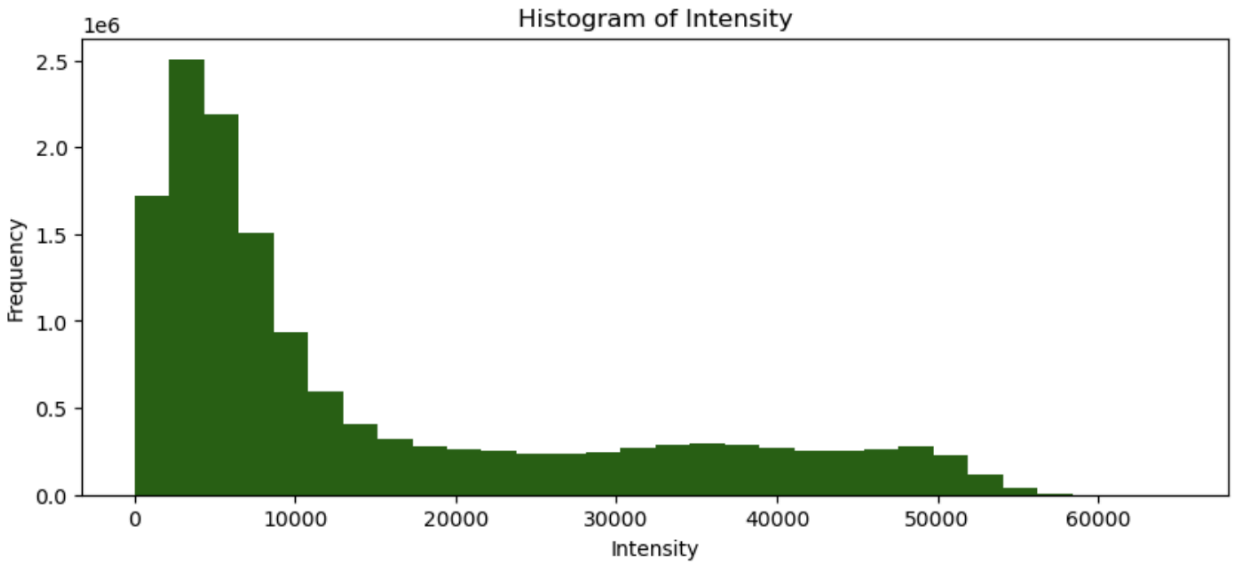


b.


Histogram of Y
Frequency
600000
500000
400000
300000
200000
100000
0
1.8700
1.8705
1.8710
1.8715
1.8720
1.8725
1e6
Y

c.

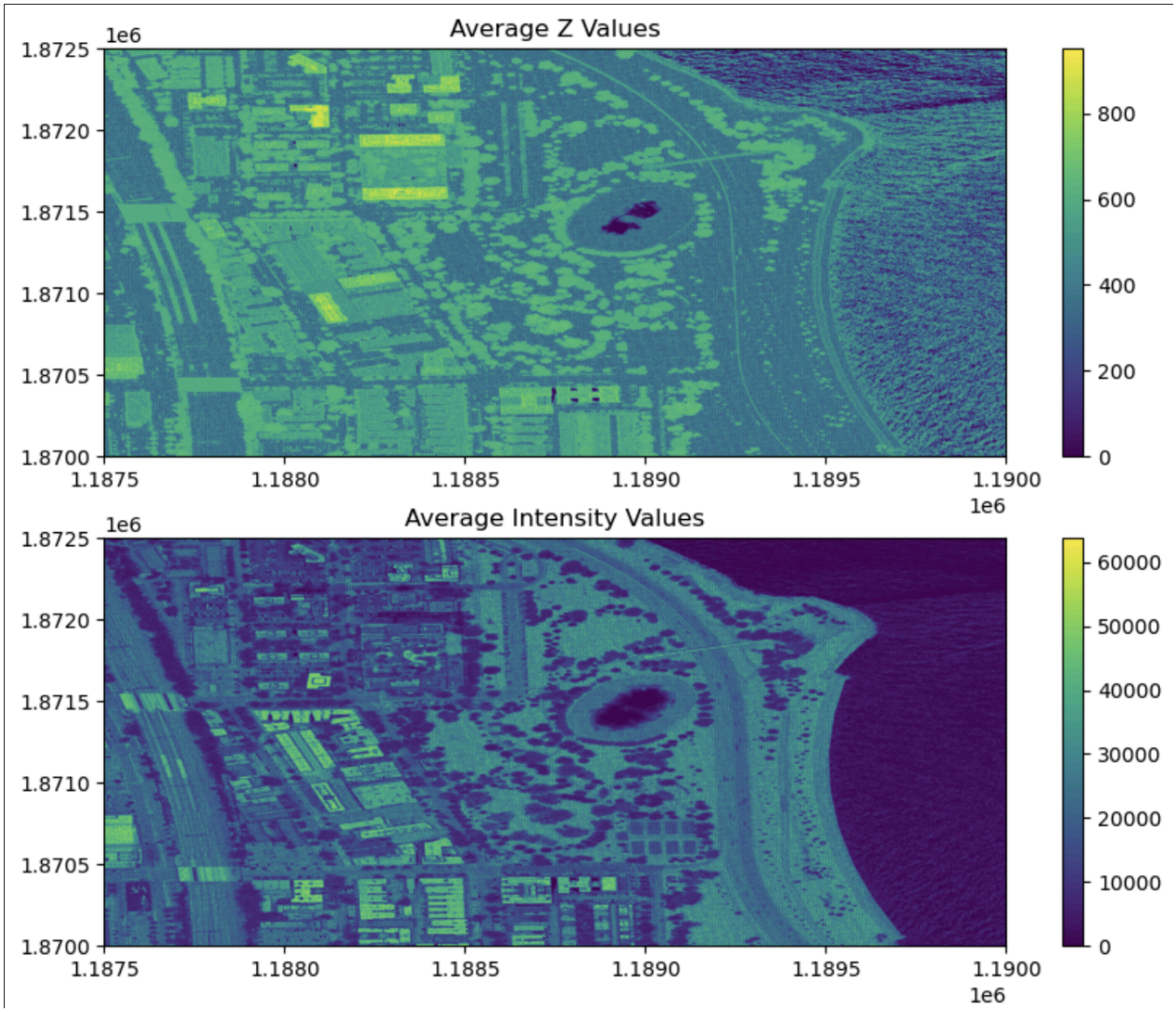

Histogram of Z
1e6
Frequency
4.0
3.5
3.0
2.5
2.0
1.5
1.0
0.5
0.0
600
650
700
750
800
850
900
950
Z



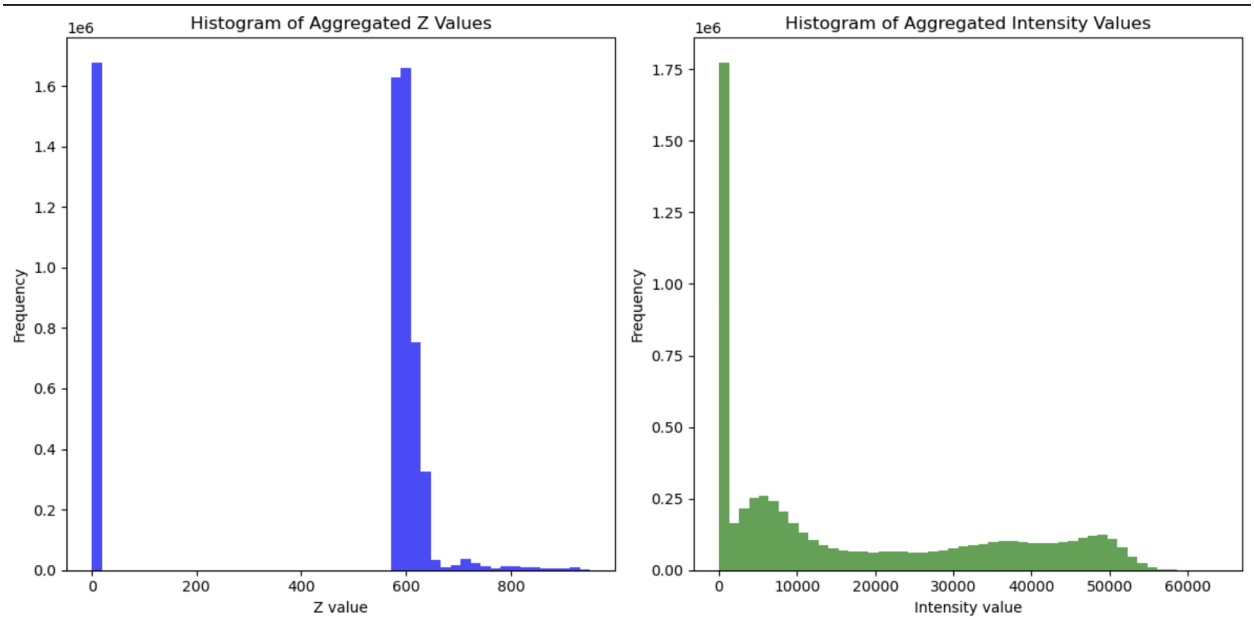
d.

e. The distributions of Z and Intensity are highly skewed to the right. The distribution of X is mildly skewed to the right. The distribution of Y is essentially uniform.

4.


Average Z Values
1e6
1.8725
1.8720
1.8715
1.8710
1.8705
1.8700
1.1875
1.1880
1.1885
1.1890
1.1895
1.1900
1e6
800
600
400
200
0
Average Intensity Values
1e6
1.8725
1.8720
1.8715
1.8710
1.8705
1.8700
1.1875
1.1880
1.1885
1.1890
1.1895
1.1900
1e6
60000
50000
40000
30000
20000
10000
0

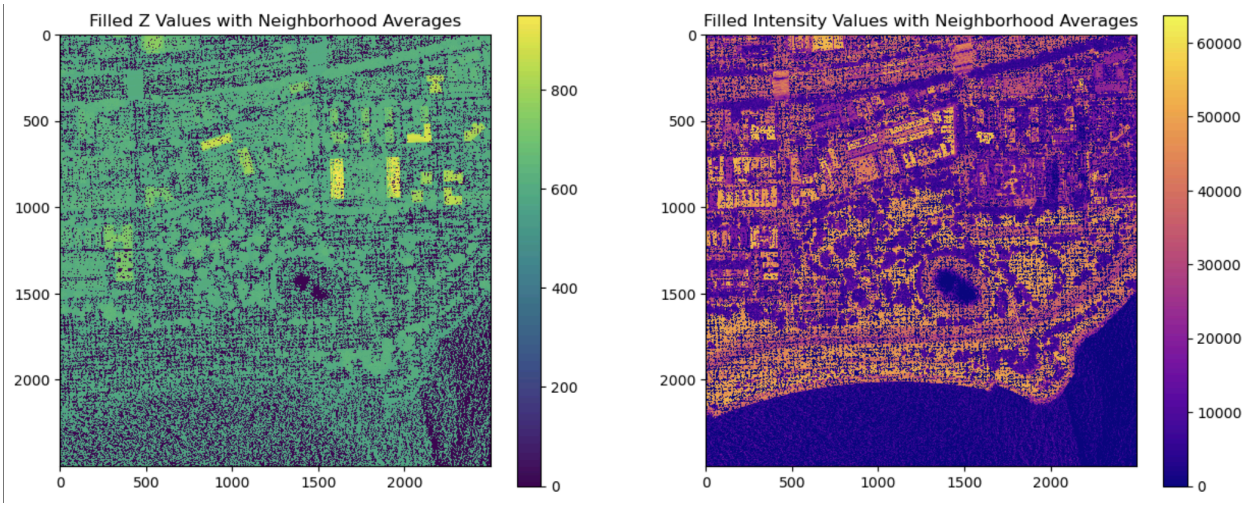
5.



a. Very skewed to the right. There are definitely a lot of 0 values in both distributions.

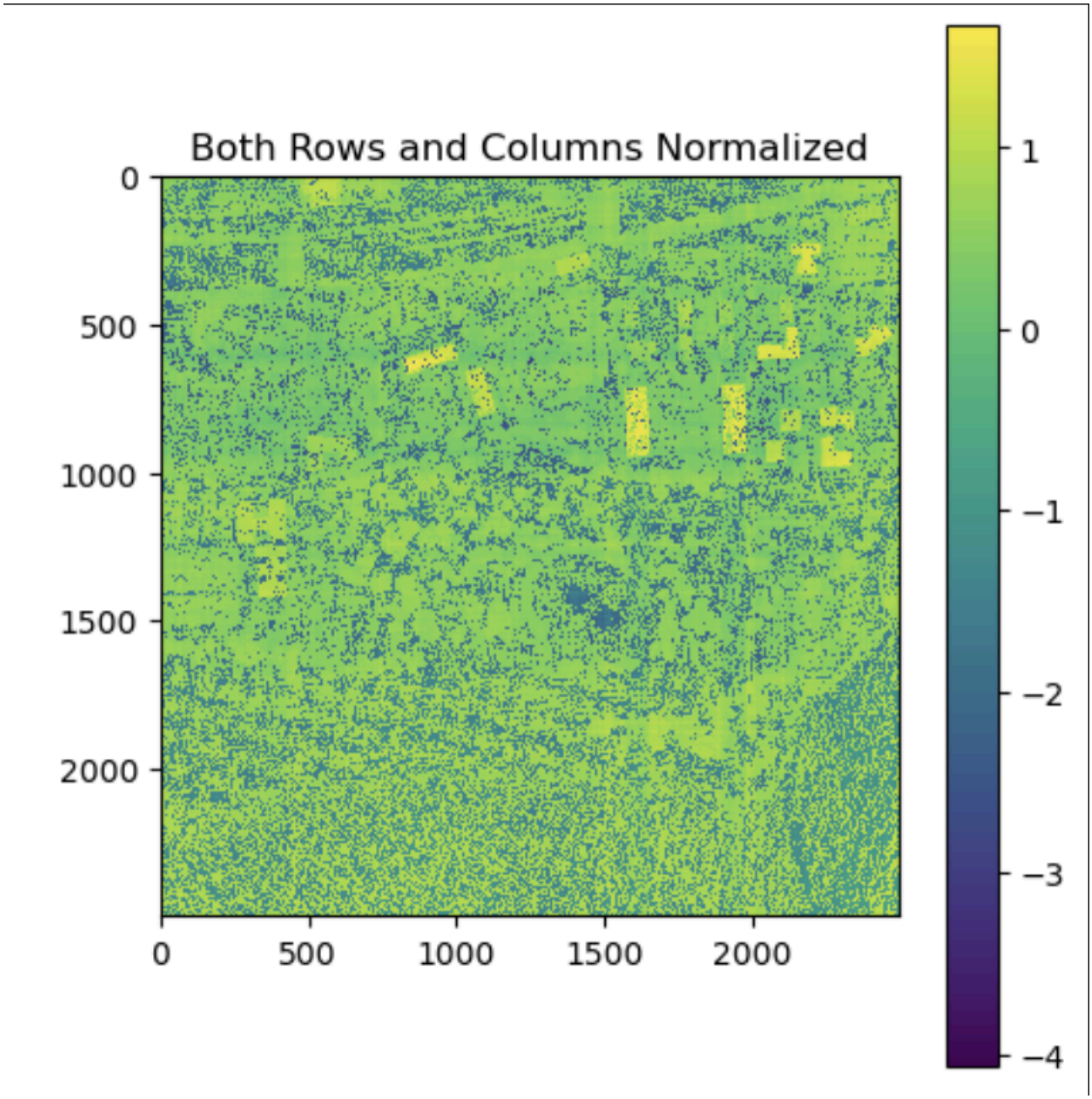
6. Have filled in any data points with NaN with the average of the 4 data points around it.

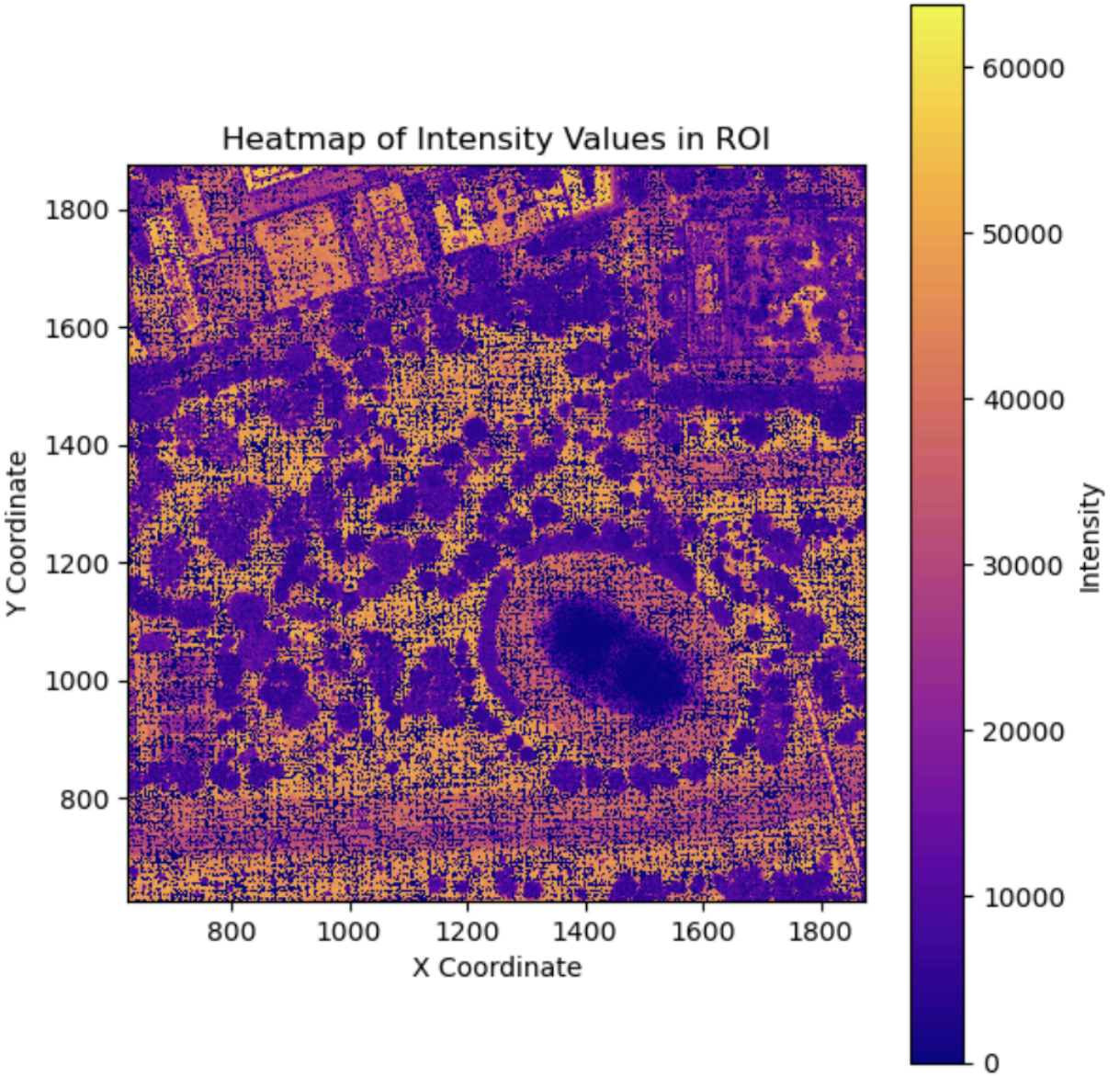
7.



a. It might have been a sharper image without filling it in using my averaging method.

8.


Both Rows and Columns Normalized
0
500
1000
1500
2000
0
500
1000
1500
2000
1
0
−1
−2
−3
−4



9.

10. In a TXT file in the repository.