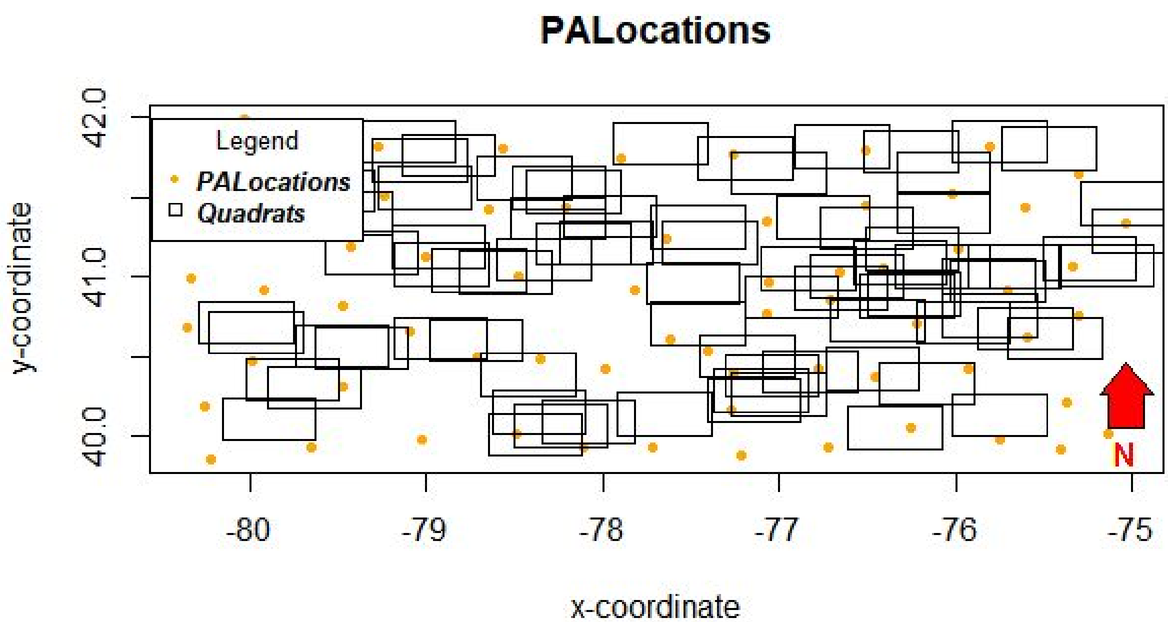
Spatial Data Analytics Mid-Term

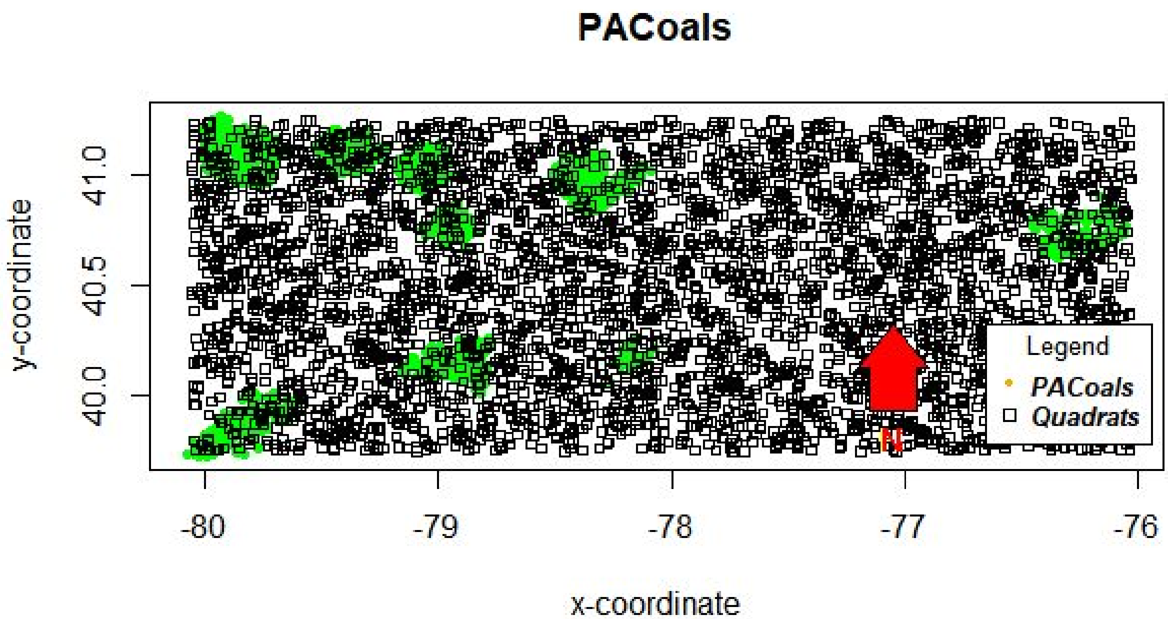
Part - A

1) Map

PALocations:-



PACoals:-



2) Table

PALocations:-

	Number of Events(K)	Number of Quadrants(X)	K-Mean	(K-Mean)^2	X(K-Mean)^2
1	1	49	0.1625	0.02640625	1.293906
2	2	29	1.1625	1.35140625	39.190781
3	3	2	2.1625	4.67640625	9.352813

PACoals:-

	Number of Events(K)	Number of Quadrants(X)	K-Mean	(K-Mean)^2	X(K-Mean)^2
1	1	28	-0.0265	7.022500e-04	0.019663
2	2	3659	0.9735	9.477023e-01	3467.642533
3	3	35	1.9735	3.894702e+00	136.314579
4	4	16	2.9735	8.841702e+00	141.467236
5	5	25	3.9735	1.578870e+01	394.717556
6	6	21	4.9735	2.473570e+01	519.449747
7	7	17	5.9735	3.568270e+01	606.605938
8	8	13	6.9735	4.862970e+01	632.186129
9	9	17	7.9735	6.357670e+01	1080.803938
10	10	16	8.9735	8.052370e+01	1288.379236
11	11	11	9.9735	9.947070e+01	1094.177725
12	12	11	10.9735	1.204177e+02	1324.594725
13	13	4	11.9735	1.433647e+02	573.458809
14	14	14	12.9735	1.683117e+02	2356.363831
15	15	7	13.9735	1.952587e+02	1366.810916
16	16	13	14.9735	2.242057e+02	2914.674129
17	17	8	15.9735	2.551527e+02	2041.221618
18	18	8	16.9735	2.880997e+02	2304.797618
19	19	8	17.9735	3.230467e+02	2584.373618
20	20	9	18.9735	3.599937e+02	3239.943320
21	21	3	19.9735	3.989407e+02	1196.822107
22	22	7	20.9735	4.398877e+02	3079.213916
23	23	9	21.9735	4.828347e+02	4345.512320
24	24	7	22.9735	5.277817e+02	3694.471916
25	25	2	23.9735	5.747287e+02	1149.457405
26	26	7	24.9735	6.236757e+02	4365.729916

	Number of Events(K)	Number of Quadrants(X)	K-Mean	(K-Mean)^2	X(K-Mean)^2
27	28	3	26.9735	7.275697e+02	2182.709107
28	29	2	27.9735	7.825167e+02	1565.033405
29	30	1	28.9735	8.394637e+02	839.463702
30	31	4	29.9735	8.984107e+02	3593.642809
31	32	4	30.9735	9.593577e+02	3837.430809
32	33	2	31.9735	1.022305e+03	2044.609405
33	35	2	33.9735	1.154199e+03	2308.397404
34	37	2	35.9735	1.294093e+03	2588.185405
35	40	1	38.9735	1.518934e+03	1518.933702
36	41	1	39.9735	1.597881e+03	1597.880702
37	43	1	41.9735	1.761775e+03	1761.774702
38	44	1	42.9735	1.846722e+03	1846.721702
39	48	1	46.9735	2.206510e+03	2206.509702

3)

Report:

```
# number of x-quadrats in PALocs
qXlocs <- 10
# number of y-quadrats in PALocs
qYlocs <- 8

# number of x-quadrats in PACoals
qXcoals <- 100
# number of y-quadrats in PACoals
qYcoals <- 40
```

```
VMR Value of PALocs using Random Quadrat Sampling Approach is: 0.763083317589269
```

```
VMR Value of PACoals using Random Quadrat Sampling Approach is: 17.9758780416
```

Shown above is the VMR values for 10 x 8 and 100 x 40 quadrat sizes for PALocations and PACoals respectively. As we can see above, VMR value of above 1 for PACoals indicates high variability among the quadrat counts, implying that more quadrats contain very few or very many events than would be expected by chance. This is suggestive of clustering. Thus, we can safely say that these set of points of PACoals is clustered.

And, VMR value of below 1 for PALocs indicates an evenly spaced arrangement. Generally, a VMR greater than 1.0 indicates a tendency towards clustering in the pattern, and a VMR less than 1.0 indicates an evenly scattered arrangement. Thus PALocs is an evenly spaced arrangement.

Strength of Quadrat Count Method:

A relatively easy way to analyze the spatial point pattern.

Weakness of Quadrat Count Method:

● Quadrat size affects results. If the quadrat size is too large or too small, the point pattern may not be detected, thus leading to a different conclusion. Instead of using 10 x 8 quadrat size for PALocs, if we use 30 x 20, the VMR for Regular Quadrant count rose sharply to 4.33.

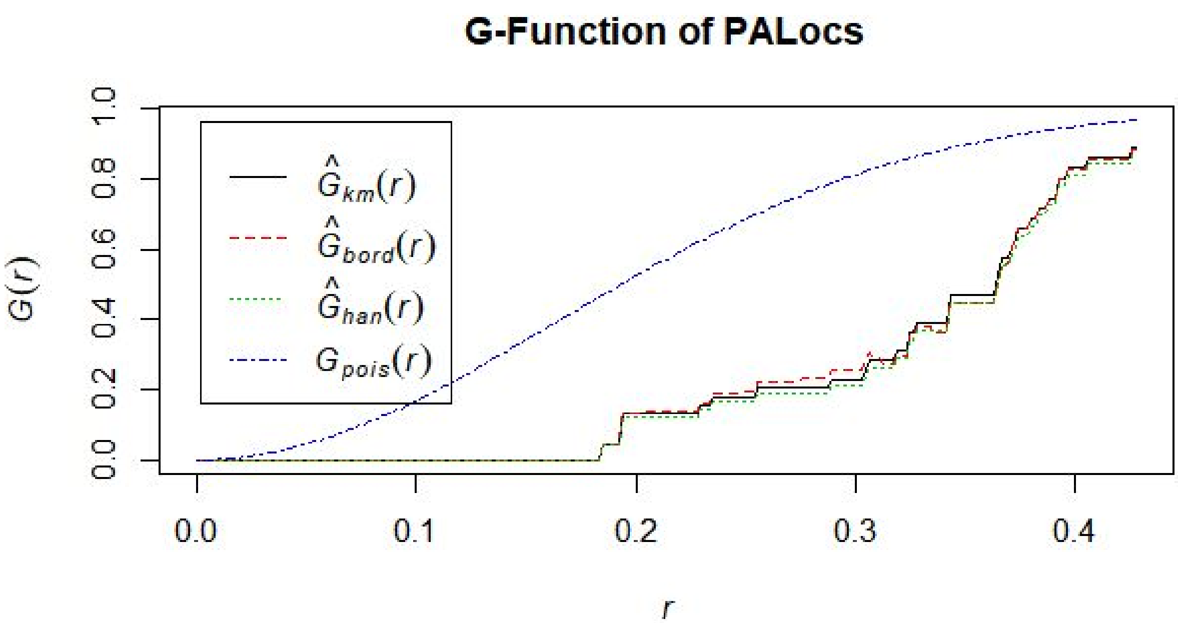
● Gives markedly different point patterns which can give rise to identical frequency distributions of points by quadrats.

● Quadrat Count is actually a measure of dispersion, and not really pattern, because it is based primarily on the density of points, and not their arrangement in relation to one another. This results in a single measure for the entire distribution, so variations within the region are not recognized.

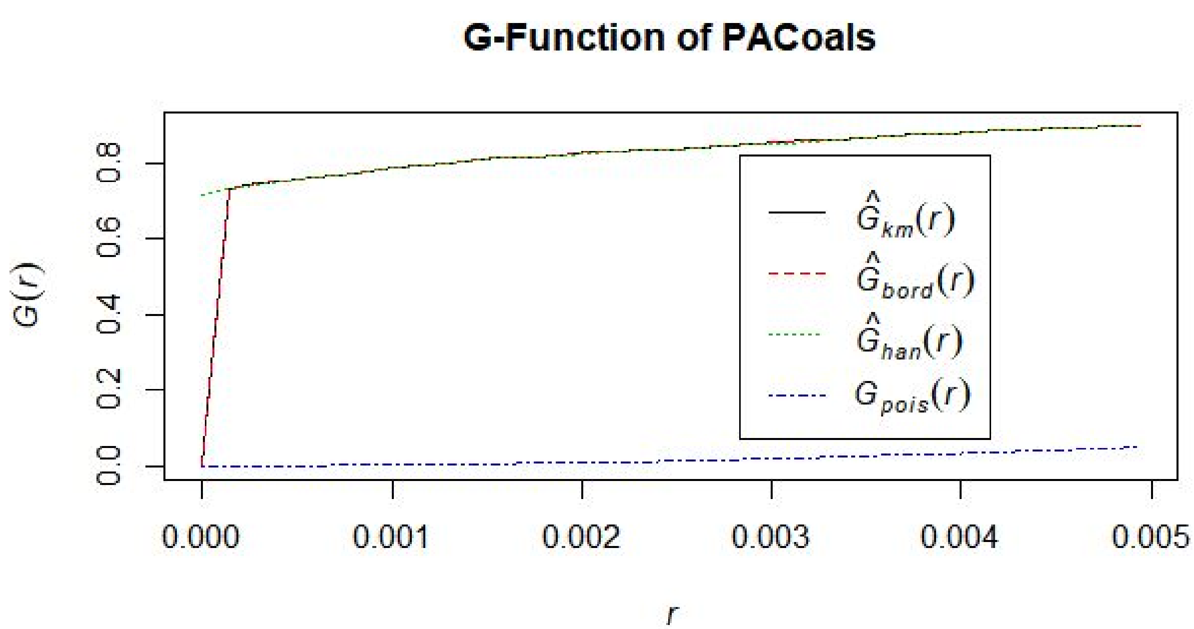
PART - B

1)

A plot showing the result of G function of PALocs:

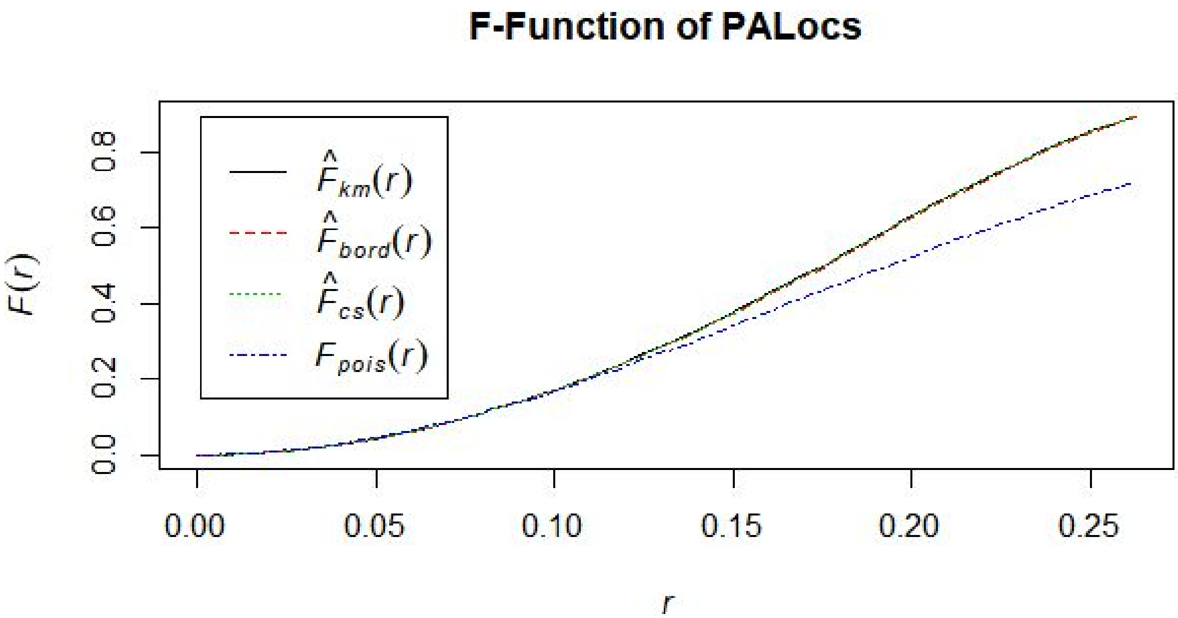


A plot showing the result of G function of PACoals:

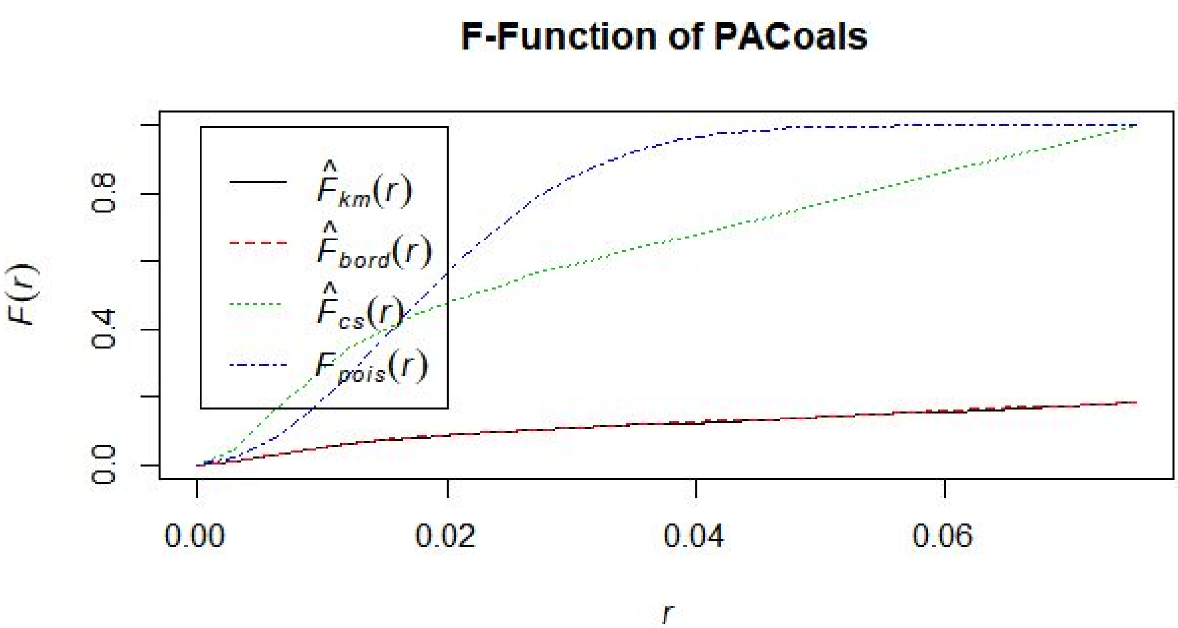


2)

A plot showing the result of F function of PALocs:



A plot showing the result of F function of PACoals:



3) Report:

Shown above is the G and F function plots for PALocations and PACoals respectively.

As can be observed, in the case of PACoals, the G function rises rapidly at short distances. In contrast, the F function rises steadily across a range of distances. The F function rises slowly at first, but more rapidly at longer distances. This is indicative of clustering in PACoals.

And, in the case of PALocs, the G function does not rise at all until the critical spacing of about 0.2, after which it rises quickly, reaching almost 100% by a distance of slightly above 0.4. The F function rises smoothly in this case. F rises quickly at low r of about 0.15. This is indicative of evenly spaced arrangement in PALocs.

While G shows how close together events in the pattern are, F relates to how far events are from arbitrary locations in the study area. So, if events are clustered in a corner of the study region, G rises sharply at short distances because many events have a very close nearest neighbor. The F function, on the other hand, is likely to rise slowly at first, but more rapidly at longer distances, because a good proportion of the study area is fairly empty, so that many locations are at quite long distances from the nearest event in the pattern. For evenly spaced patterns, the opposite is true. Most locations in the dataset are relatively close to an event, so that F rises quickly at low d. However, events are relatively far from each other, so that G initially increases slowly and rises more quickly at longer distances.