Part A:

Below are steps I used to achieve the Moran's I value which gives us a measure of autocorrelation:

- Created a square matrix which holds the distance between all the set of points which was calculated by using Latitude and Longitude.
- Calculated the median of all values in the above square matrix which came out to be 0.158
- After which I calculated the weight matrix whose elements are either 0 if their corresponding distance matrix element was greater than median value and 1 if their corresponding distance matrix element was less than the median value.
- The above formulated weight matrix was then used to evaluate the Moran's I expression.

The numeric value of Moran I expression I got was 0.13 which suggests a positive autocorrelation though a weak one which suggests that there is positive spatial autocorrelation observed for the attribute Particulate Matter but the degree of this autocorrelation is small as its value is less than 0.3.

(*R code attached in separate file*)

Part B:Industrial Mining Dataset

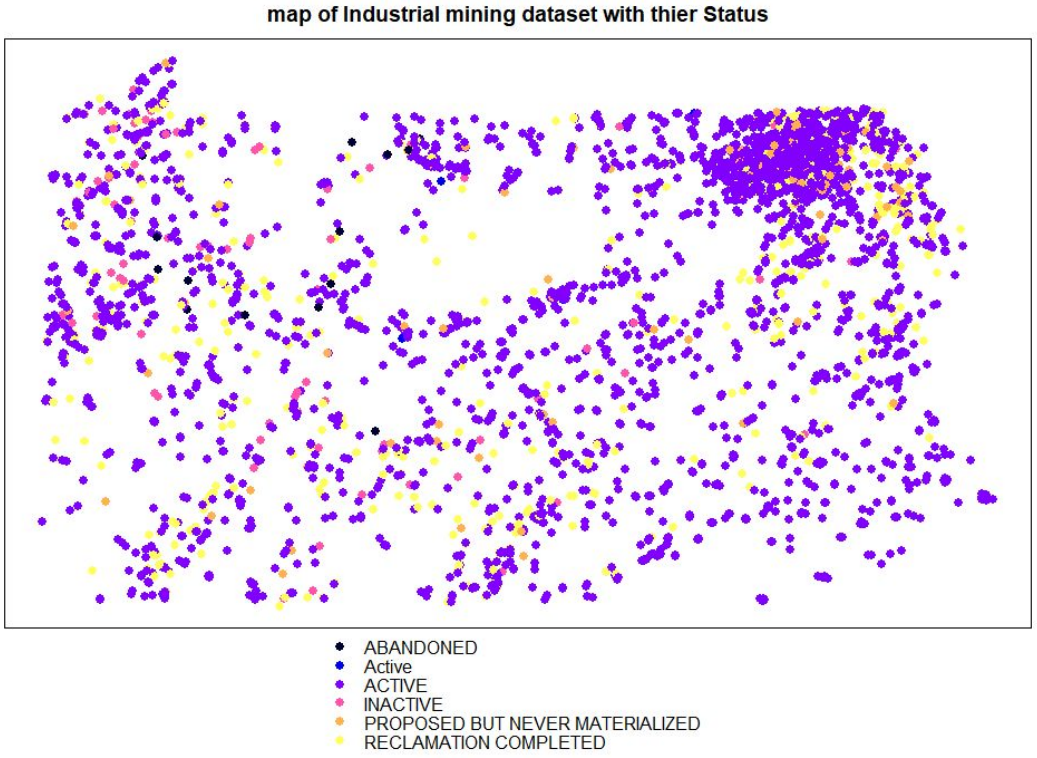
(Below is the analysis of plots followed my maps and plots)

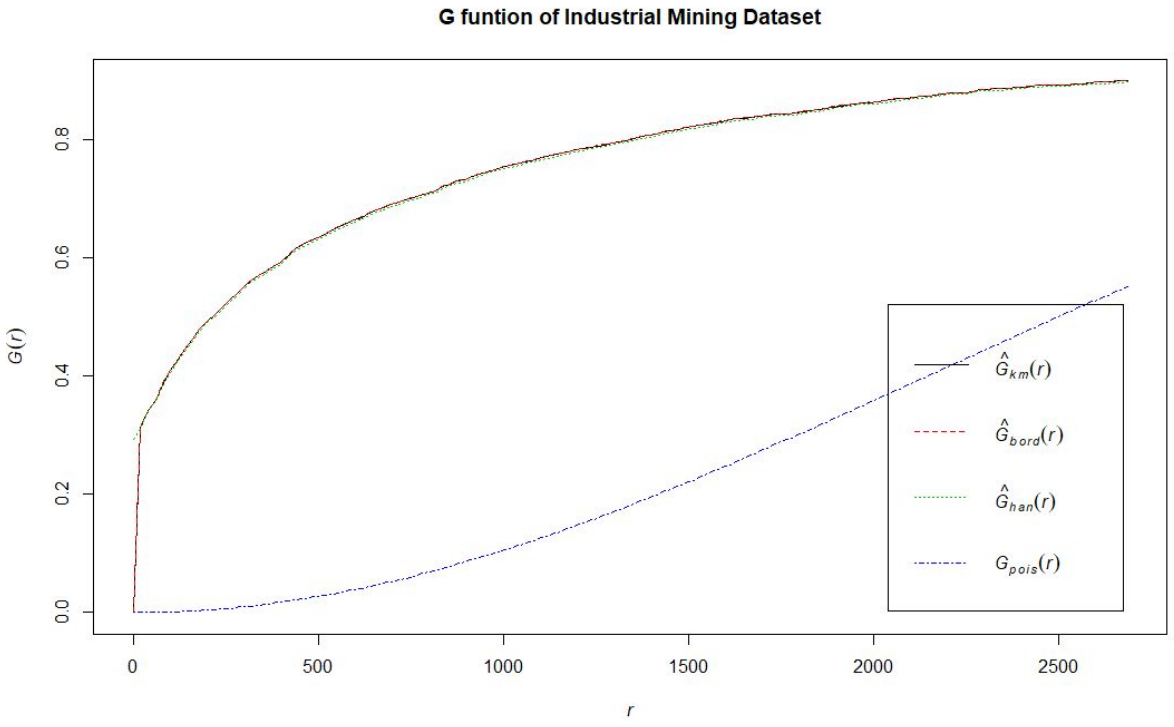
As the data for Industrial mining dataset contains evenly spaced points and most locations in Pattern are relatively close to an event, so that F rises quickly at low distances. However, as the events are relatively far from each other, so that G initially increases slowly and rises more quickly at longer distances and the same is evident from the G & F plots below figures below.

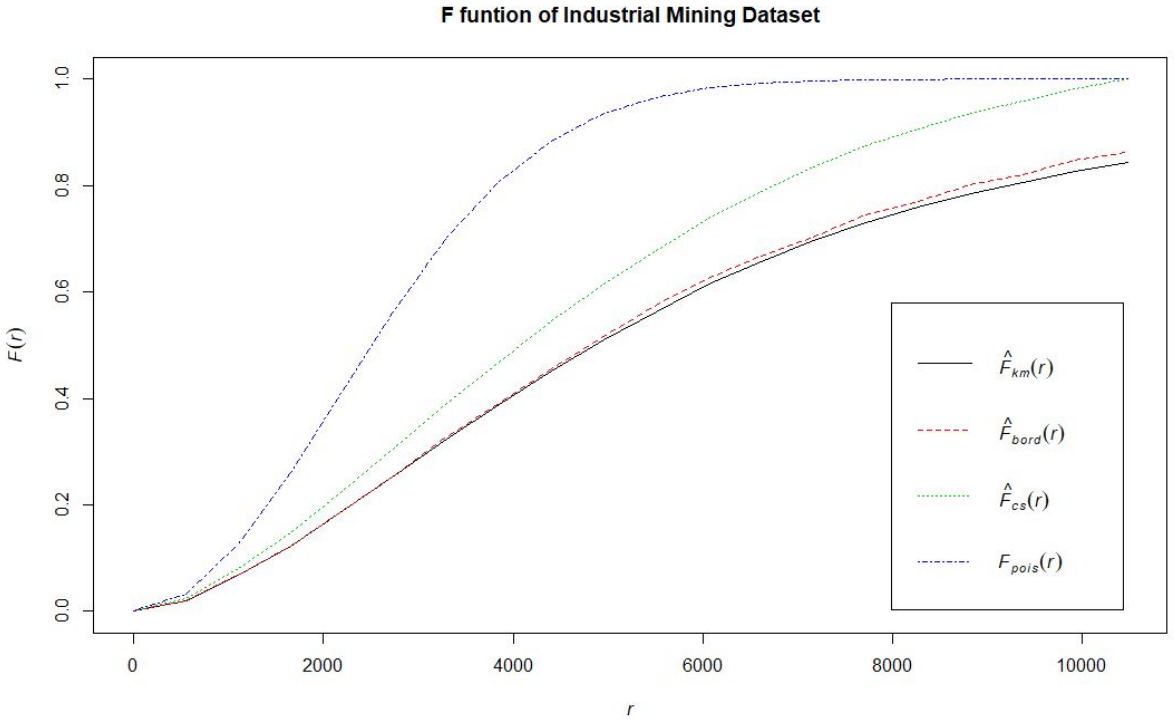
As the data for Oil Gas Location dataset is clustered the G function rises sharply at shorter distances because many events have a very close nearest neighbour whereas the F function rises slowly at first and more rapidly at larger distances and same fact is reflected from G & F plot of oil gas dataset.

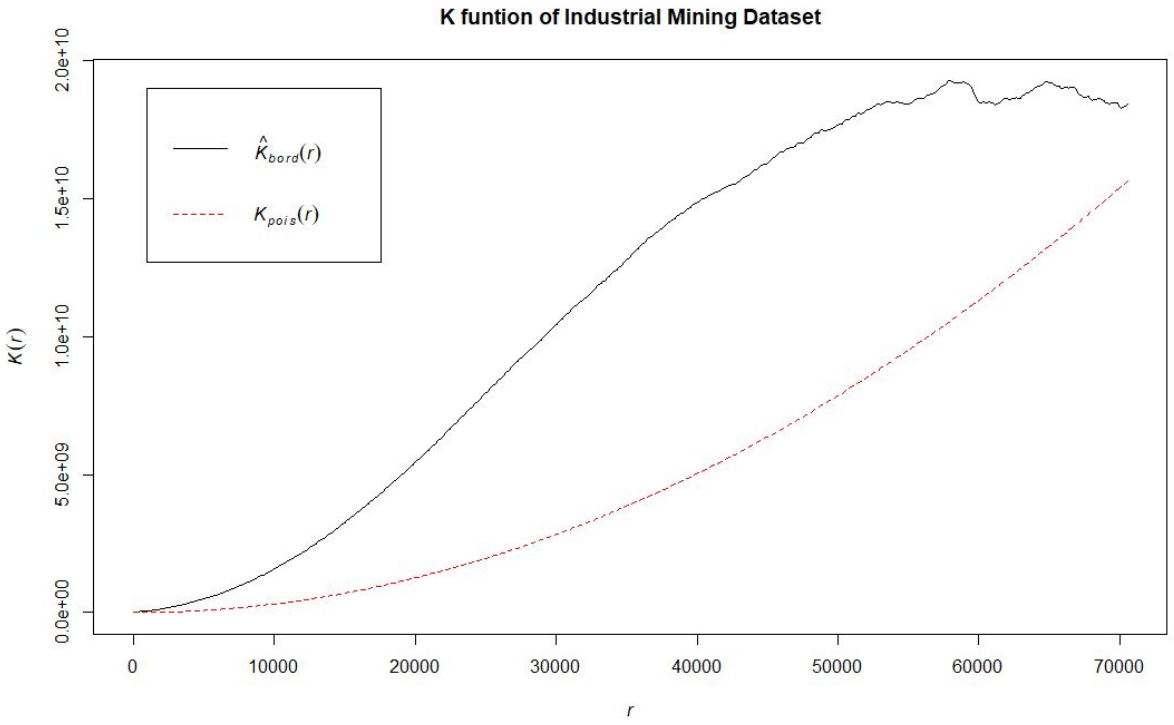
Because of different nature of two datasets G & F function behave contrastingly in both the cases.

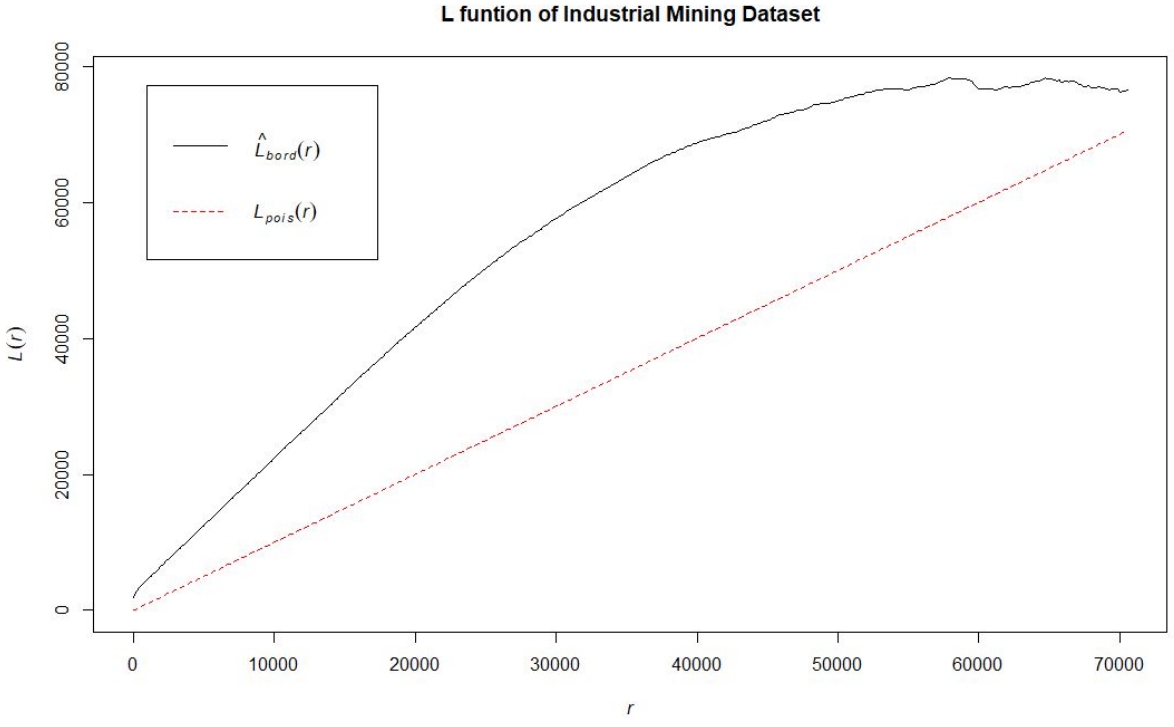
Where L function is above zero, there are more events at the corresponding spacing than would be expected under CSR; when it is negative, there are fewer events than expected. In both the datasets L function for all values of d other than ones close to zero are greater than zero,indicating that there are more events at these spacings than expected under CSR. For both the datasets, L function is close to zero for very small d than rises rapidly.The L funtion for the Industrial mining dataset suggests that there is a little degree of clustering but it is very small and same can be interpreted from the map.The K function depends on the square of distance so its value rises rapidly in both the datasets.It should be noted that the value of K and L functions are to a certain degree distorted by edge effects and methods like guard zone can be used to reduce it.



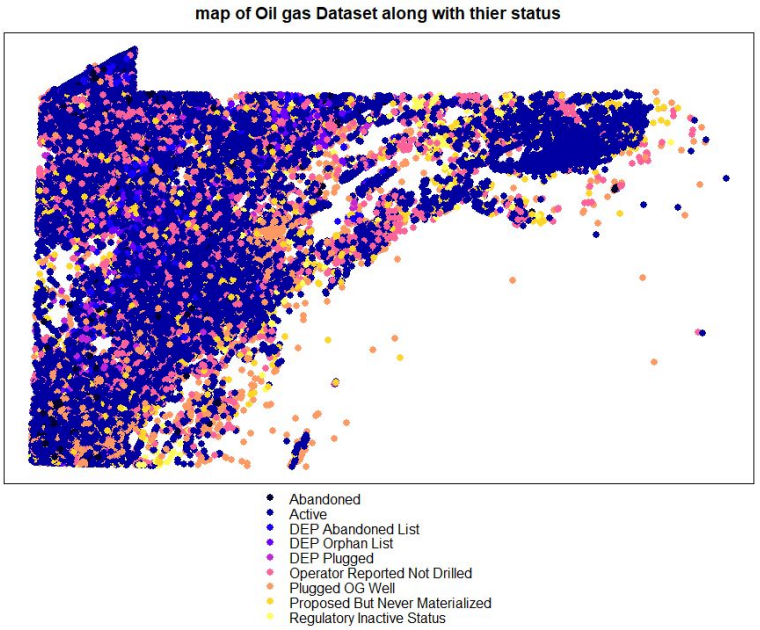

G funtion of Industrial Mining Dataset
G(r)
0.0
0.2
0.4
0.6
0.8
0
500
1000
1500
2000
2500
r
$\hat{G}_{km}(r)$
$\hat{G}_{bord}(r)$
$\hat{G}_{han}(r)$
$G_{pois}(r)$

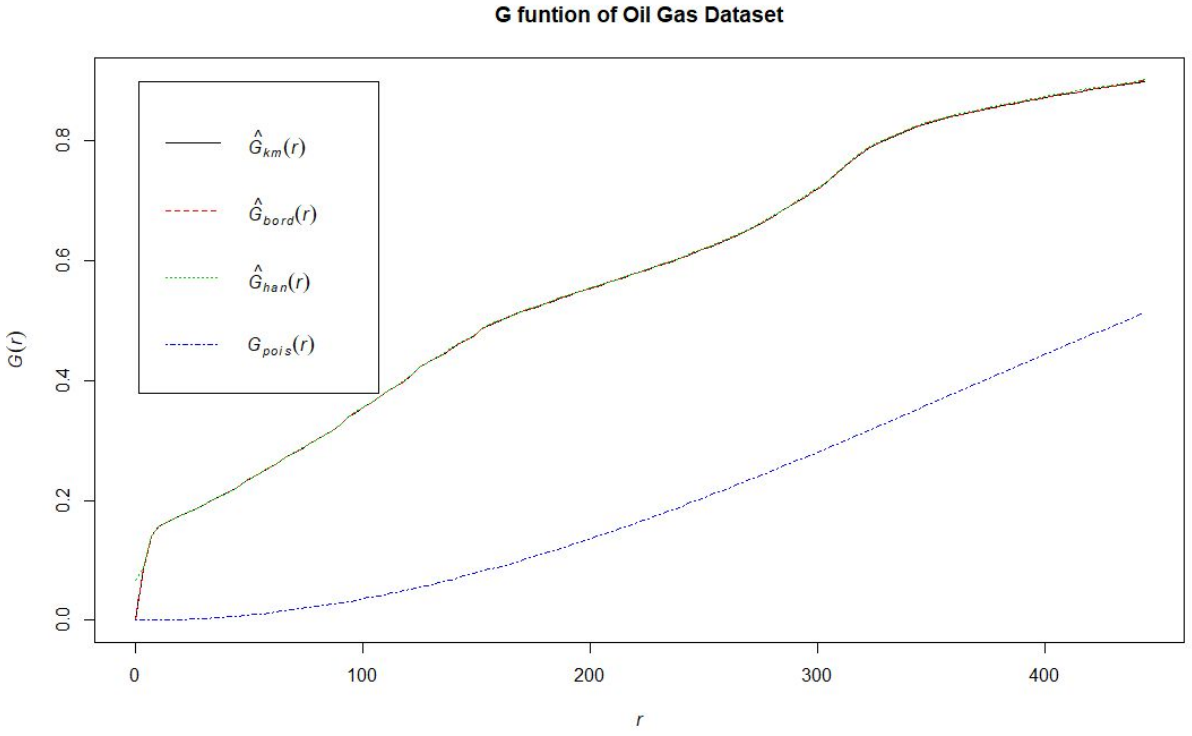

F funtion of Industrial Mining Dataset
F(r)
0.0
0.2
0.4
0.6
0.8
1.0
0
2000
4000
6000
8000
10000
r
$\hat{F}_{km}(r)$
$\hat{F}_{bord}(r)$
$\hat{F}_{cs}(r)$
$F_{pois}(r)$

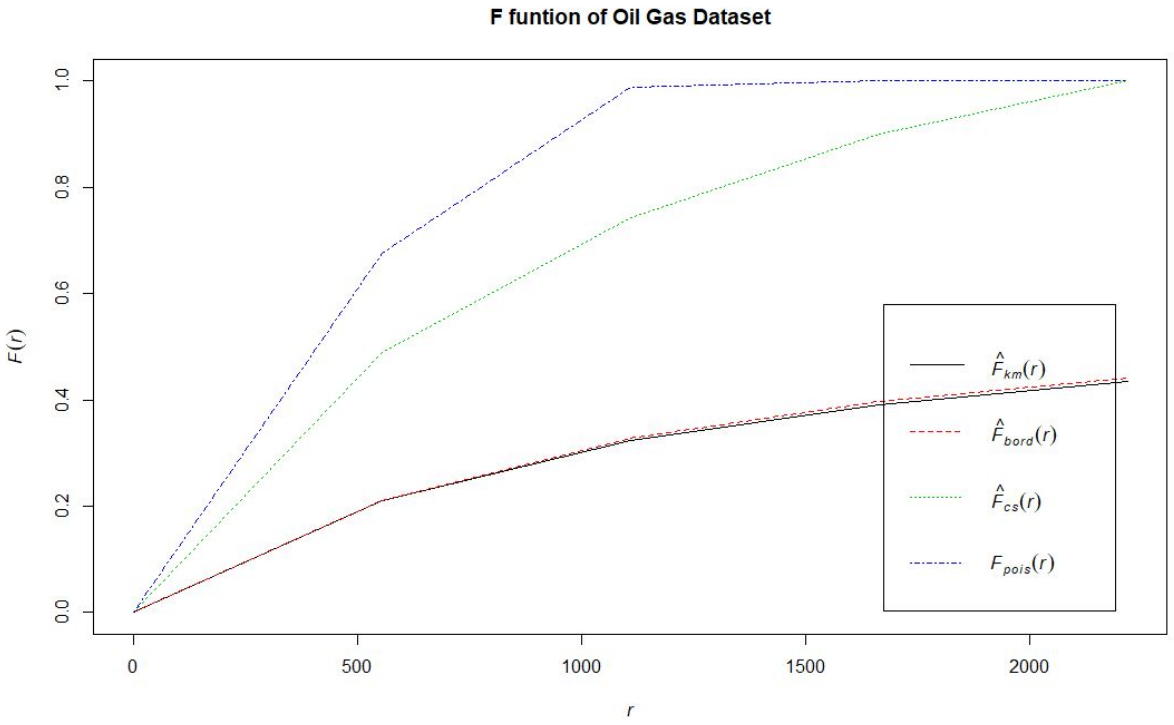


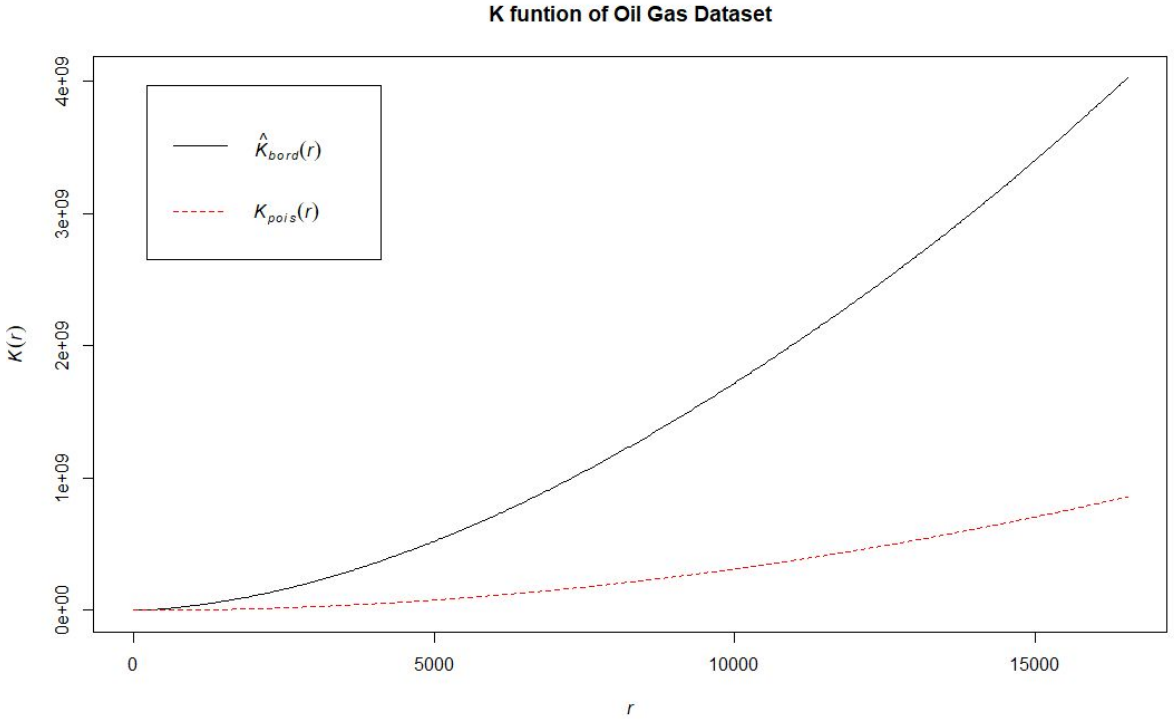


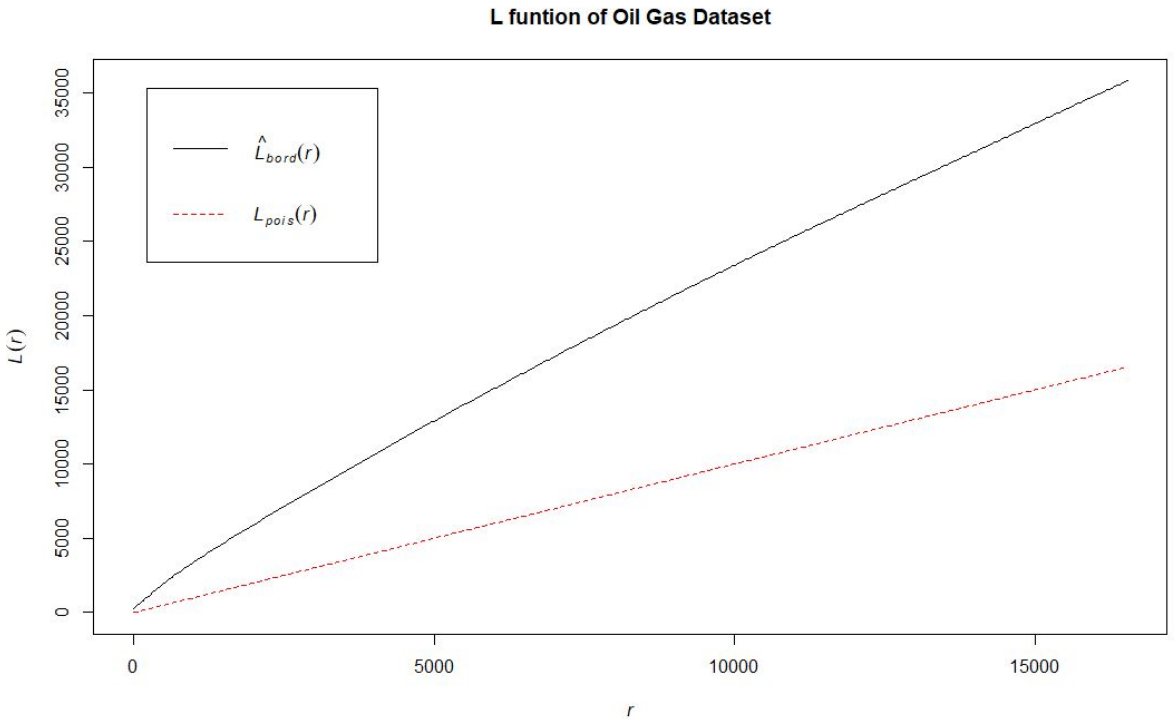
Oil Gas Location:


map of Oil gas Dataset along with thier status
Abandoned
Active
DEP Abandoned List
DEP Orphan List
DEP Plugged
Operator Reported Not Drilled
Plugged OG Well
Proposed But Never Materialized
Regulatory Inactive Status


G funtion of Oil Gas Dataset
$\hat{G}_{km}(r)$
$\hat{G}_{bord}(r)$
$\hat{G}_{han}(r)$
$G_{pois}(r)$
G(r)
0.0
0.2
0.4
0.6
0.8
0
100
200
300
400
r


F funtion of Oil Gas Dataset
F(r)
r
0.0
0.2
0.4
0.6
0.8
1.0
0
500
1000
1500
2000
$\hat{F}_{km}(r)$
$\hat{F}_{bord}(r)$
$\hat{F}_{cs}(r)$
$F_{pois}(r)$


K funtion of Oil Gas Dataset
K(r)
r
0e+00
1e+09
2e+09
3e+09
4e+09
0
5000
10000
15000
$\hat{K}_{bord}(r)$
$K_{pois}(r)$


L funtion of Oil Gas Dataset
$\hat{L}_{bord}(r)$
$L_{pois}(r)$
L(r)
0
5000
10000
15000
20000
25000
30000
35000
r