

Nome: Hugo Muniz Albuquerque

Disciplina: Estatística Espacial

Questão 1:

O cientista de dados poderia ter feito um mapa de pontos, com fundo do google maps, de furtos para a região analisada. A fim de observar a concentração dos pontos espacialmente. Outra opção seria ter feito um mapa para cada tipo de carro roubado para analisar algum padrão entre bairros e tipos de carros. Além disso, poderia ter feito a análise da distribuição espacial dos eventos para investigar se existe um comportamento de agrupamento ou repulsão ou aleatoriedade espacial. Assim, teria que fazer um análise da média do processo espacial (efeito de primeira ordem) e a análise da dependência espacial (efeito de segunda ordem). A análise dos efeitos de primeira ordem serviria para observar a intensidade do padrão de pontos. Já os efeitos da segunda ordem poderiam revelar algum padrão de dependência espacial. Também poderia ter feitos a análise da função F que leva em conta a distância entre um ponto e o evento mais próximo.

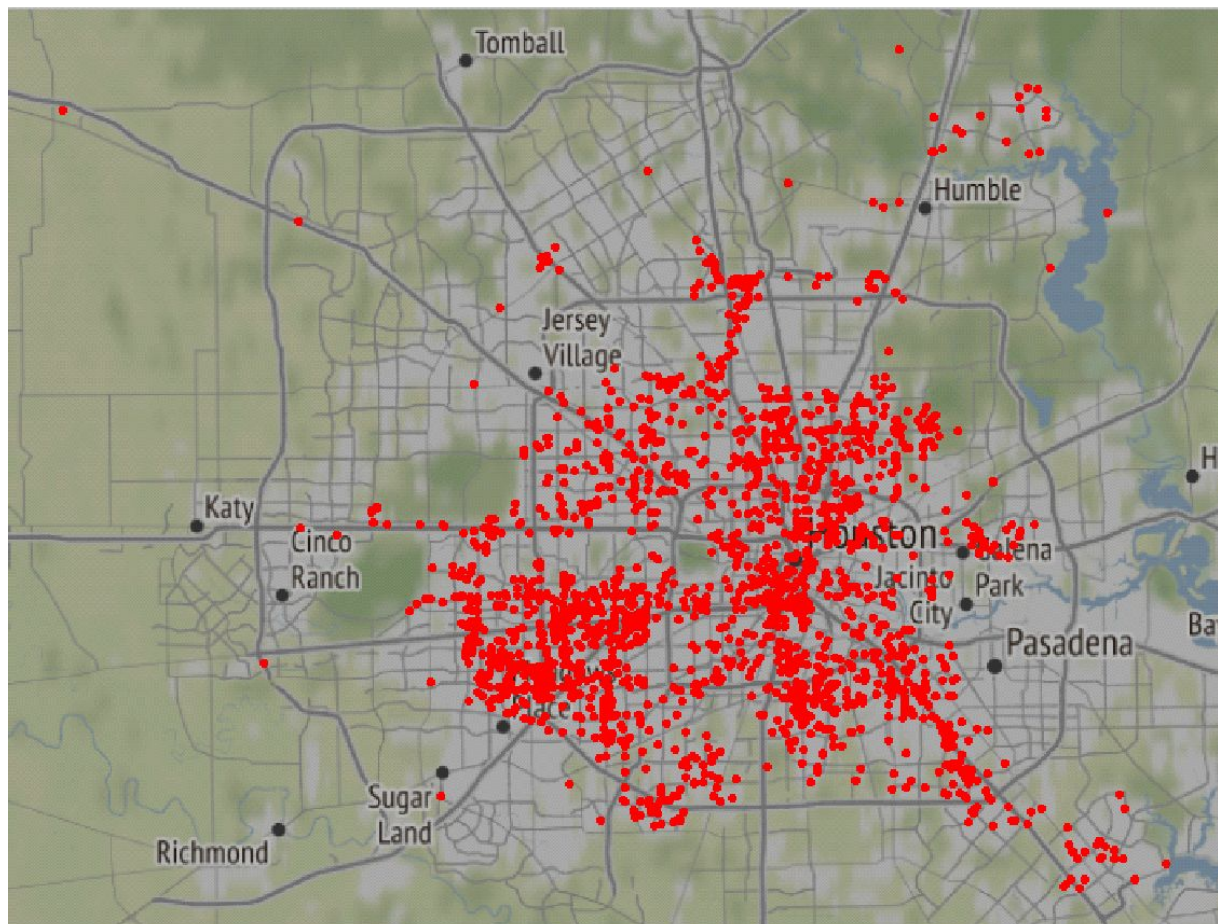
Para validar as análises poderia ter sido feito um teste de hipótese de completa aleatoriedade espacial (CSR). O teste de Clark-Evans é indicado para este processo, onde um p-valor muito baixo poderia indicar que o padrão de pontos não é CSR, enquanto que um p-valor relativamente baixo indicaria um padrão de agrupamento, e um p-valor igual a 1 um padrão de regularidade. além disso, poderia ter considerado mais de um critério para a matriz de vizinhança W e poderia ter calculado o Índice de Moran Local.

A conclusão dele está errada. Ao contrário do que ele falou, os bairros mais escuros indicam que existe um maior número de furtos de carros. Outro ponto errado foi que um p-valor de 0.008 indica que a hipótese nula foi rejeitada, logo não é correto afirmar que não existe padrão espacial, como ele concluiu. Além disso, um Índice Moran Global de 0.61 indica que existe autocorrelação espacial, indicando um padrão de agrupamento, o que contradiz o cientista de dados.

Questão 2:

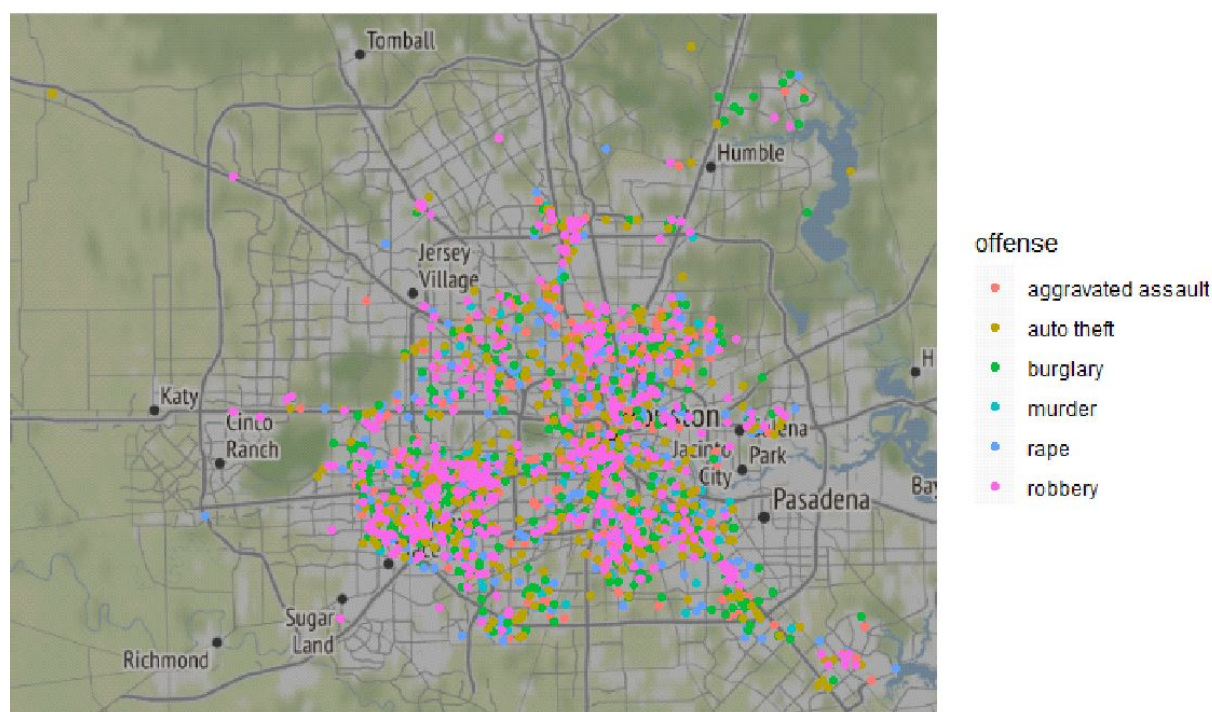
O comportamento da intensidade de crimes parece ocorrer de forma similar para todos os seus tipos para a cidade de Houston ou os padrões tendem a ser diferentes?

Mapa de pontos dos crimes em Houston



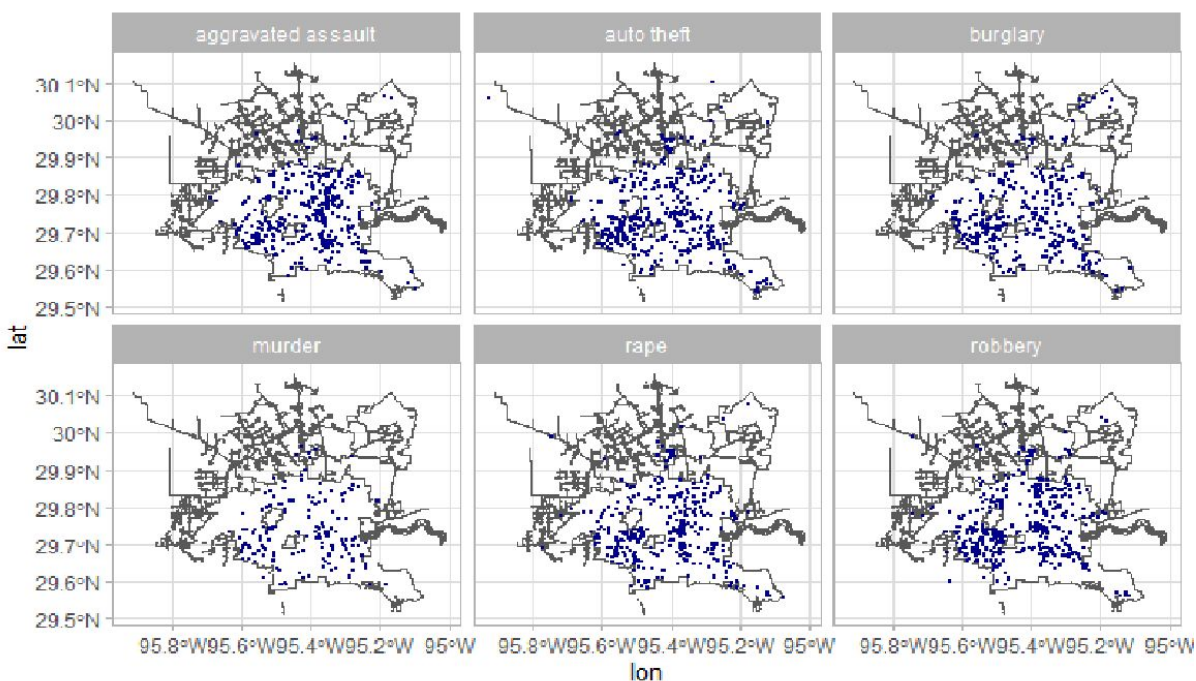
Este gráfico mostra os pontos em que foram localizados os crimes. Neste é possível notar uma concentração no centro da cidade, porém será preciso analisar por tipos de crimes para ter uma noção melhor da distribuição espacial dos crimes. Assim, primeiramente foi feito a coloração dos pontos de acordo com o tipo de crime (utilizando a variável offense). O mapa encontra-se abaixo:

Mapa de pontos coloridos por tipo de crime (offense)



Neste mapa pode-se observar uma concentração maior de pontos rosas e verdes, porém não é uma visualização suficientemente adequada para a análise dos padrões de pontos de cada tipo de crime. Portanto, foi feito um mapa facetado por tipo utilizando a função `facet_wrap(~offense)` cujo resultado está no mapa abaixo.

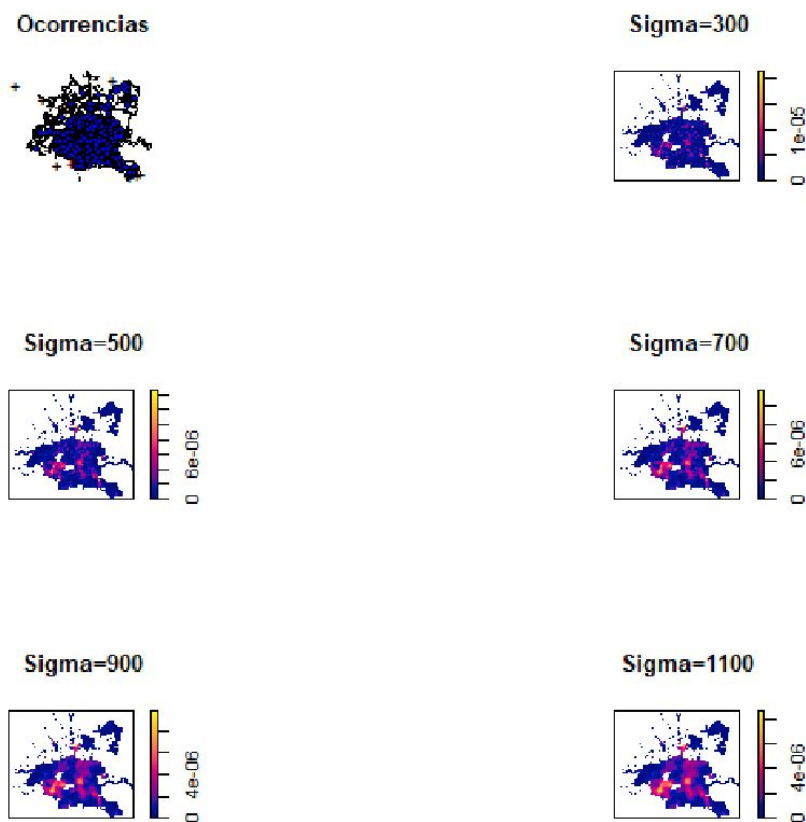
Mapa facetado por tipos de crimes (offense)



Todos os tipos de crime se concentram aproximadamente em torno do centro do mapa de Houston, mas alguns apresentam concentrações maiores em algumas regiões e direções. O tipo aggravated assault é o que parece apresentar a maior concentração em regiões específicas. Este apresenta uma concentração maior a leste e sudeste, e um pouco menor a sudoeste e oeste. Enquanto que o tipo robbery na direção sudoeste, sudeste e no centro-leste. O tipo rape apresenta uma concentração maior na região centro-sul e sudoeste e leste, porém destaca-se casos mais distantes a sudeste. O tipo auto-theft a sudoeste, centro-sul, centro-norte e sudeste até regiões

distantes. O tipo burglary apresenta uma concentração maior a sudoeste e um pouco menos a sudeste, observa-se casos distantes no nordeste e alguns casos no centro-norte. O tipo murder é o que possui menor concentração em uma região específica, se concentrando a sudeste, centro e sudoeste, com uma pequena presença no norte. Para melhorar a análise, foi feito um mapa da intensidade dos pontos como um todo e facetado por tipo de crime, a fim de observar melhor as regiões e direções com maiores concentrações. Para isso foi preciso definir um raio adequado para a função de densidade do padrão de pontos. Este segue abaixo:

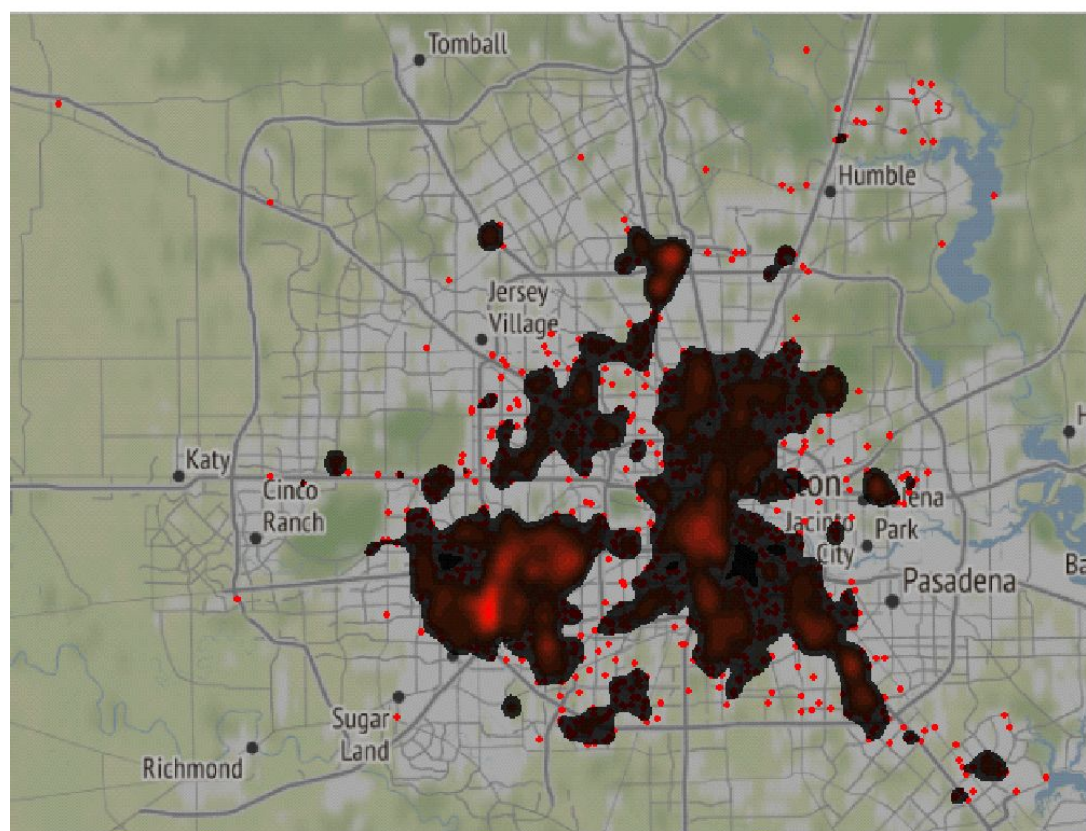
Mapa de intensidade com diferentes valores de raio/sigma



Aqui neste mapas foram aplicados diferentes raios/sigmas à função de densidade, onde foi escolhido um raio de 500m e um kernel Gaussiano para estimar as

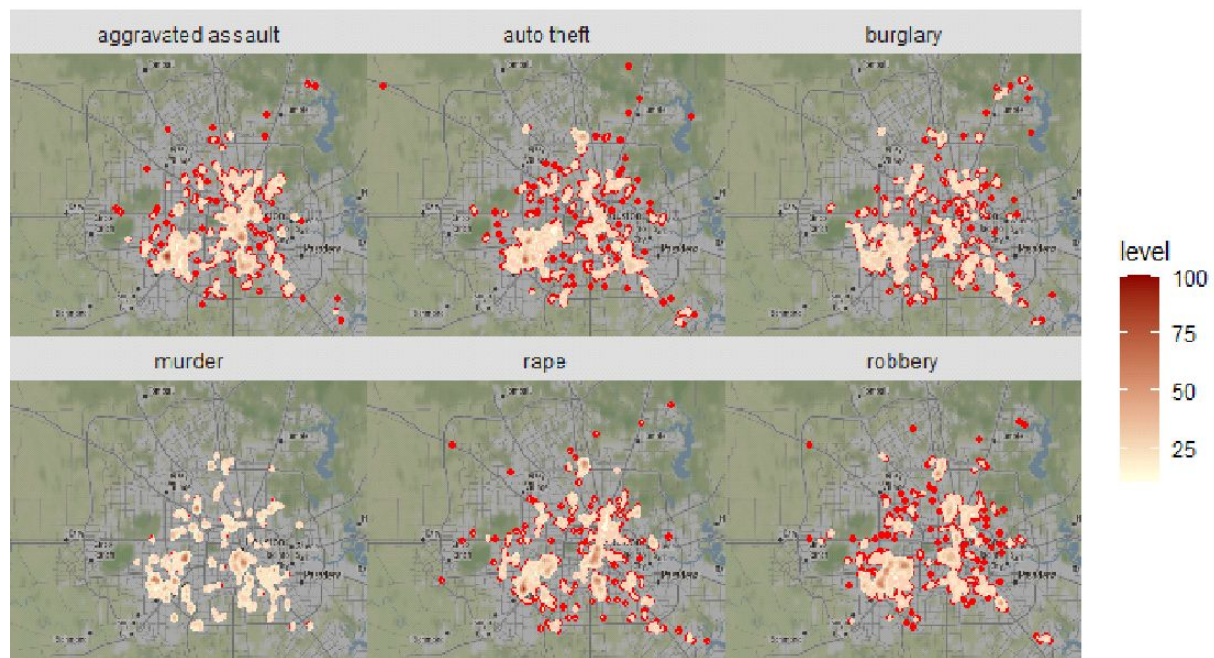
funções densidades do padrão de pontos a ser plotado. Posteriormente, foi criado um mapa de intensidade dos padrões de pontos a ser plotado.

Mapa da intensidade dos padrões de pontos de crimes em Houston



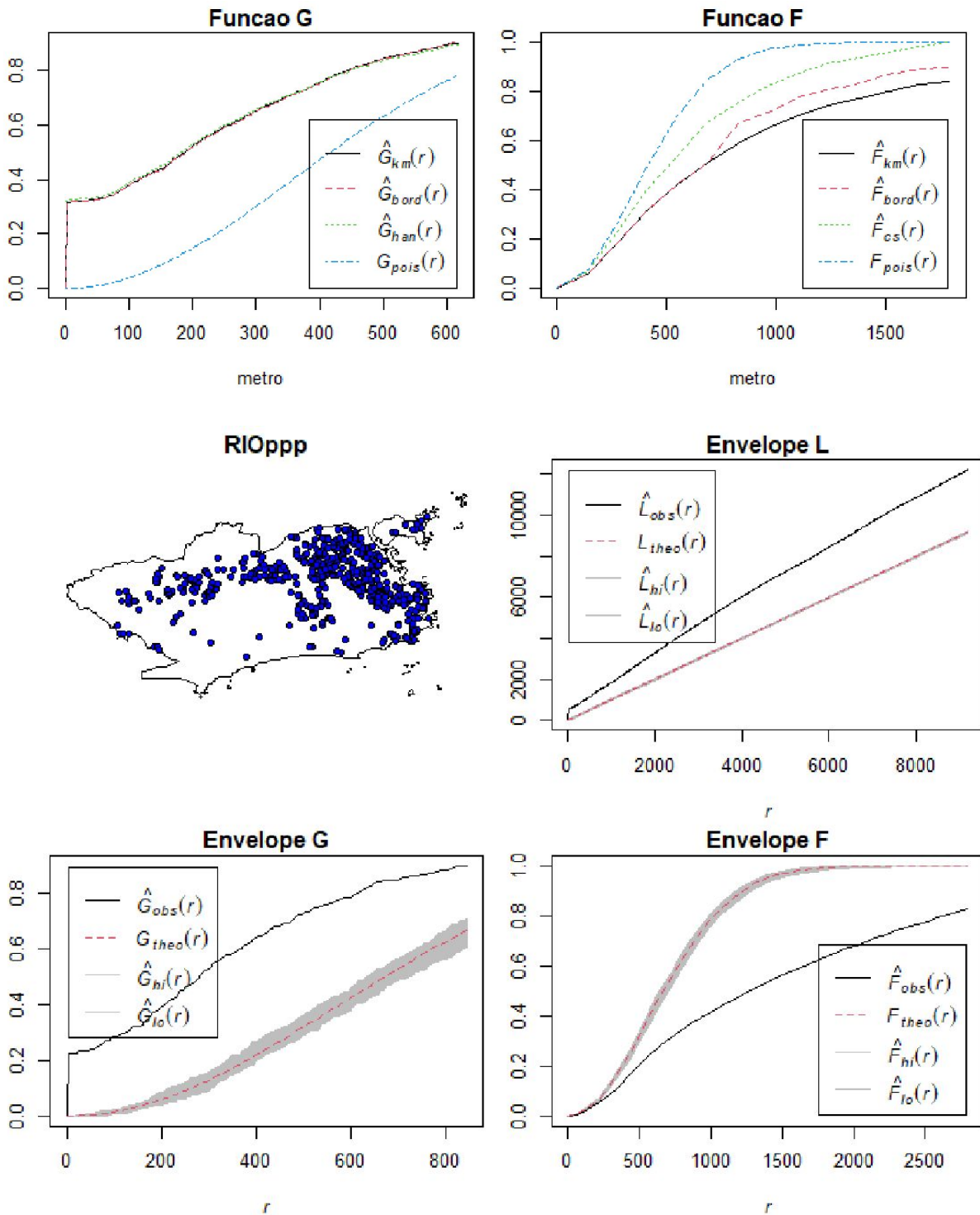
Este mapa oferece uma melhor visualização das concentrações de pontos no espaço da cidade de Houston. Nele é possível observar uma concentração maior na região compreendendo o sudeste, centro-sul e sudoeste em relação ao centro da cidade. O mapa abaixo fornecerá informações visuais mais apuradas das concentrações dos pontos por tipo de crime.

Mapa da intensidade dos pontos facetado por time de crime



O mapa de intensidade corrobora com o que foi analisado com o mapa de pontos para cada tipo de crimes, além de oferecer uma caracterização melhor do padrão de agrupamento presente em cada um dos tipos de crime. Apesar da análise de mapa fornecer boas informações, é preciso realizar testes de hipóteses para confirmar se existe um padrão de agrupamento nestes dados.

Mapa da função G e F



O gráfico das funções G e F e do envelope das funções G e F revelam um padrão de agrupamento. A curva da função G está acima curva teórica de pontos

totalmente aleatórios (curva azul) e a curva da função F está abaixo da curva azul, logo observa-se um possível padrão de agrupamento. Em complemento, para verificar a não aleatoriedade espacial completa foi feito o teste de Clark-Evans, cujo resultado foi o seguinte:

Clark-Evans test

No edge correction

Z-test

data: RIOppp

$R = 0.52561$, **p-value < 2.2e-16**

alternative hypothesis: two-sided

Como o p-valor está menor do que o nível de significância de 5% a hipótese nula de que a distribuição é de completa aleatoriedade espacial é rejeitada, logo não existe completa aleatoriedade espacial (CSR) nestes dados. Além disso, foi realizado o Clark-Evans para testar o agrupamento. Segue abaixo o resultado:

Clark-Evans test

No edge correction

Z-test

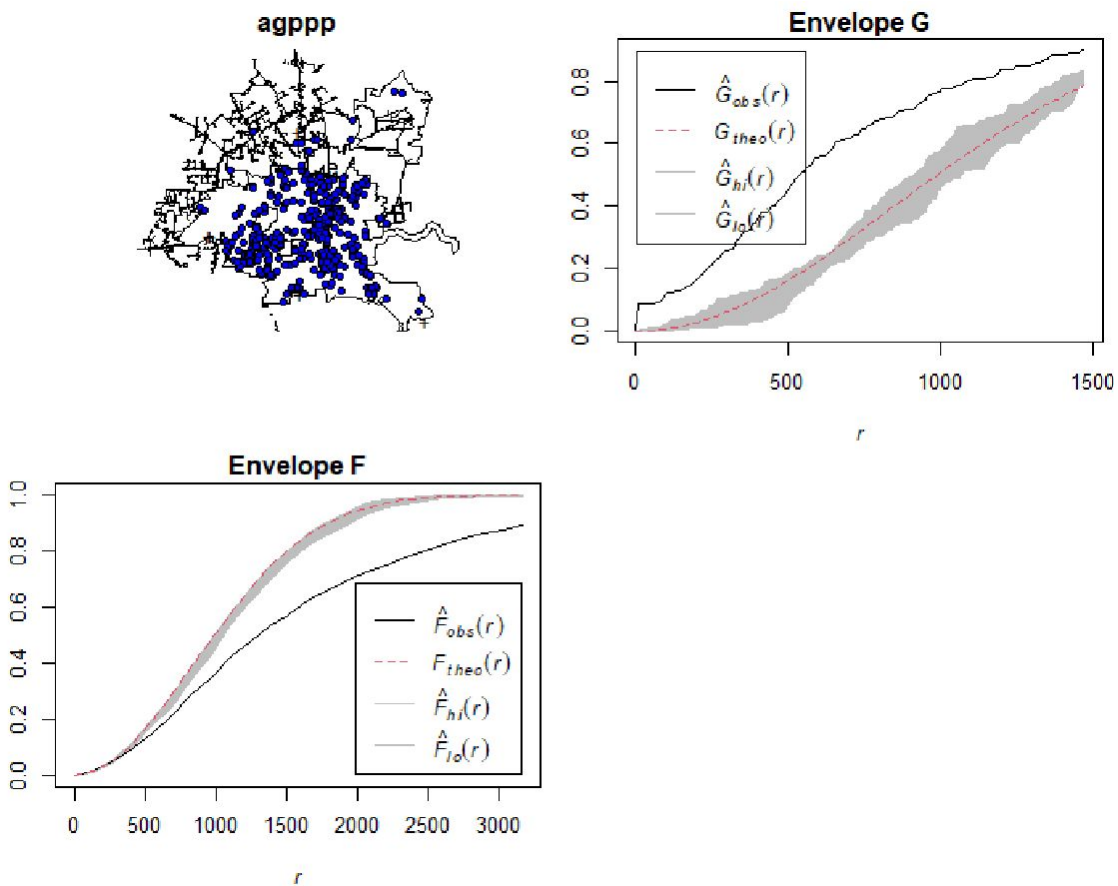
data: RIOppp

$R = 0.52561$, **p-value < 2.2e-16**

alternative hypothesis: clustered ($R < 1$)

Neste teste, o p-valor está indicando que se deve rejeitar a hipótese nula de que não existe um padrão de agrupamento. Além desses testes, foram conduzidos teste de Clark-Evans para cada tipo de crime. Segue abaixo os resultados:

Envelope para Aggravated assault



O gráfico envelope F e G sugerem um padrão de agrupamento.

Clark-Evans test

No edge correction

Z-test

data: agppp

$R = 0.71516$, **p-value < 2.2e-16**

alternative hypothesis: two-sided

Clark-Evans test

No edge correction

Z-test

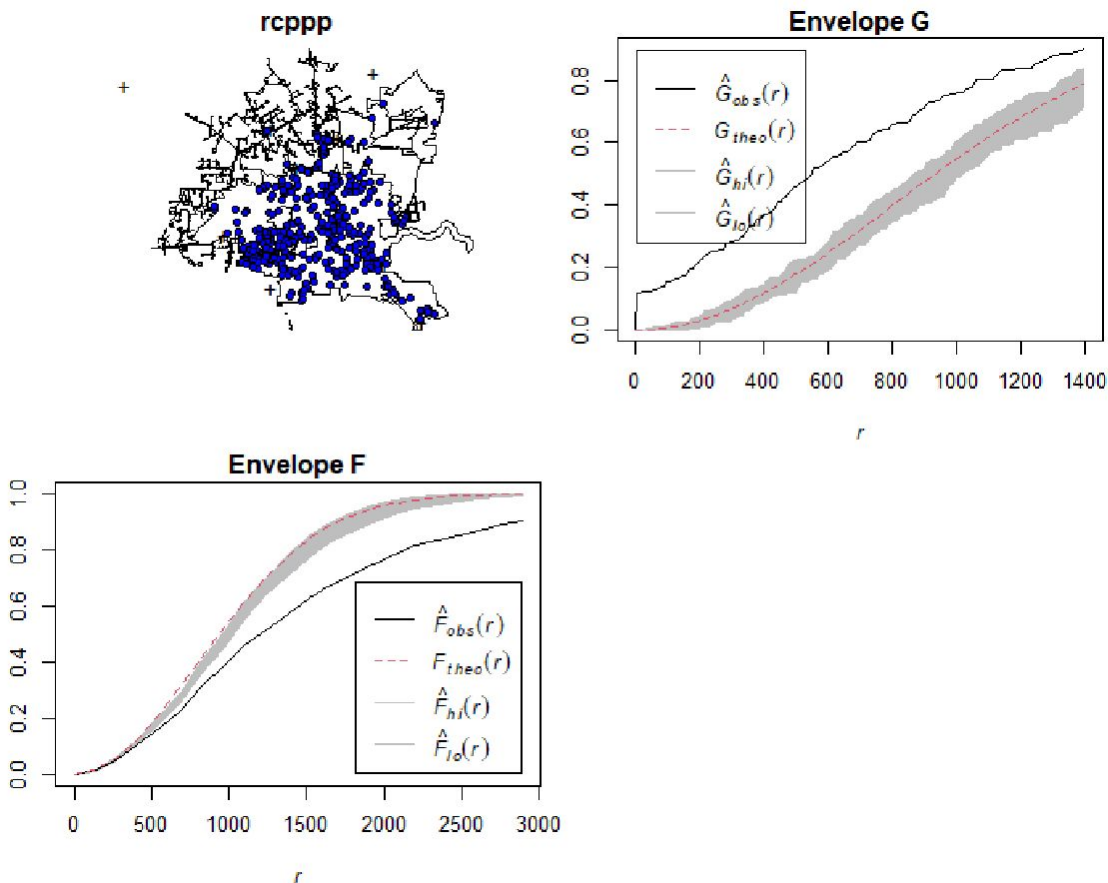
data: agppp

$R = 0.71516$, **p-value < 2.2e-16**

alternative hypothesis: clustered ($R < 1$)

Um pvalor menor do que o nivel de significancia, nos dois testes, indica que não existe CSR e que existe um padrão de agrupamento para o tipo aggravated assault. Agora vamos para o o tipo auto_theft:

Envelope para Auto_theft



O gráfico envelope F e G sugerem um padrão de agrupamento.

Clark-Evans test

No edge correction

Z-test

10/08/2020

data: rcppp

$R = 0.74949$, **p-value < 2.2e-16**

alternative hypothesis: two-sided

Clark-Evans test

No edge correction

Z-test

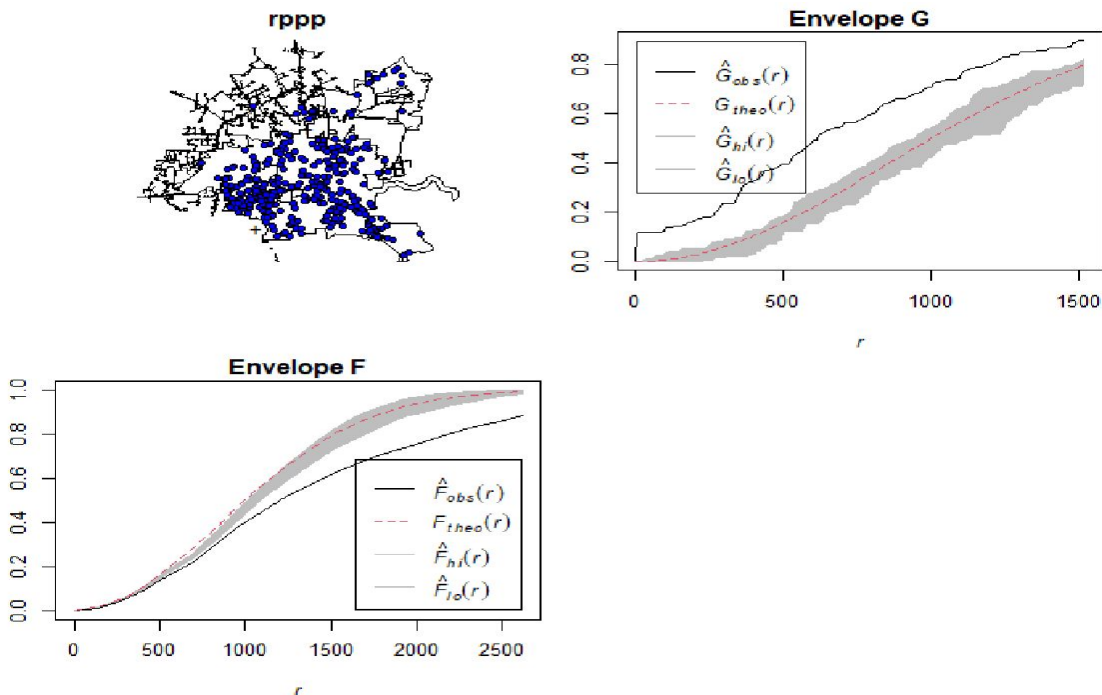
data: rcppp

$R = 0.74949$, **p-value < 2.2e-16**

alternative hypothesis: clustered ($R < 1$)

Os dois testes indicam respectivamente que não existe CSR e que existe um padrão de agrupamento para o tipo auto_theft. Agora vamos para o o tipo burglary:

Envelope para burglary



O gráfico envelope F e G sugerem um padrão de agrupamento.

10/08/2020

Clark-Evans test

No edge correction

Z-test

data: rppp

$R = 0.77057$, **p-value < 2.2e-16**

alternative hypothesis: two-sided

Clark-Evans test

No edge correction

Z-test

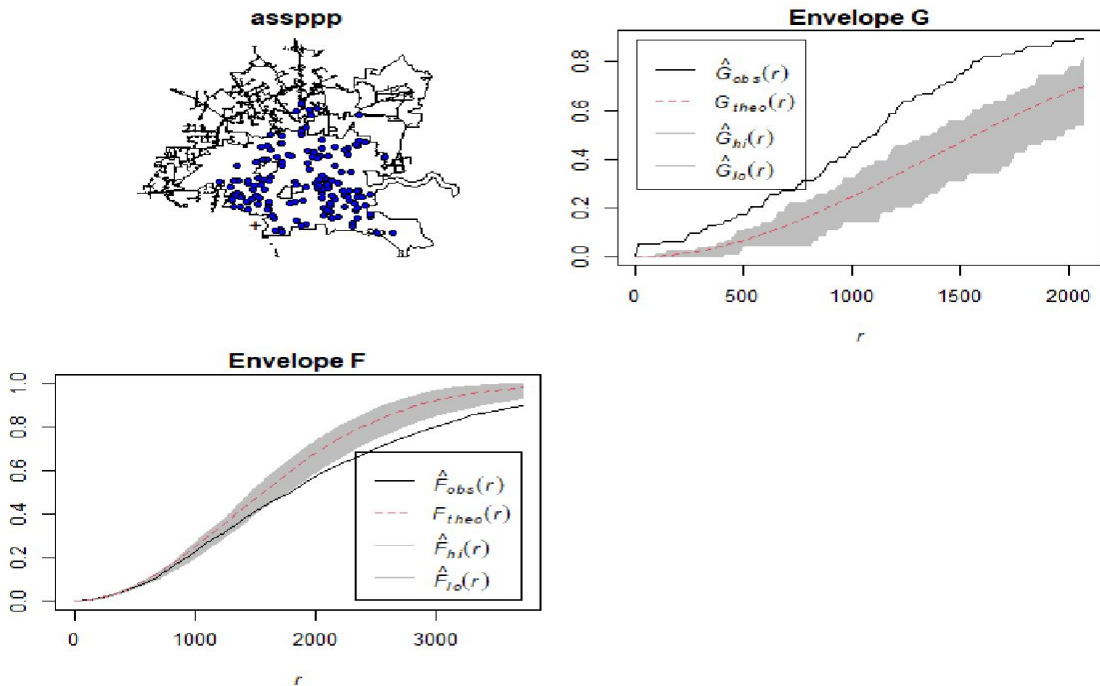
data: rppp

$R = 0.77057$, **p-value < 2.2e-16**

alternative hypothesis: clustered ($R < 1$)

Os dois testes indicam respectivamente que não existe CSR e que existe um padrão de agrupamento para o tipo burglary. O próximo será o tipo murder:

Envelope para murder



O gráfico envelope F e G sugerem um padrão de agrupamento.

Clark-Evans test

No edge correction

Z-test

data: assppp

$R = 0.78445$, **p-value = 2.838e-07**

alternative hypothesis: two-sided

Clark-Evans test

No edge correction

Z-test

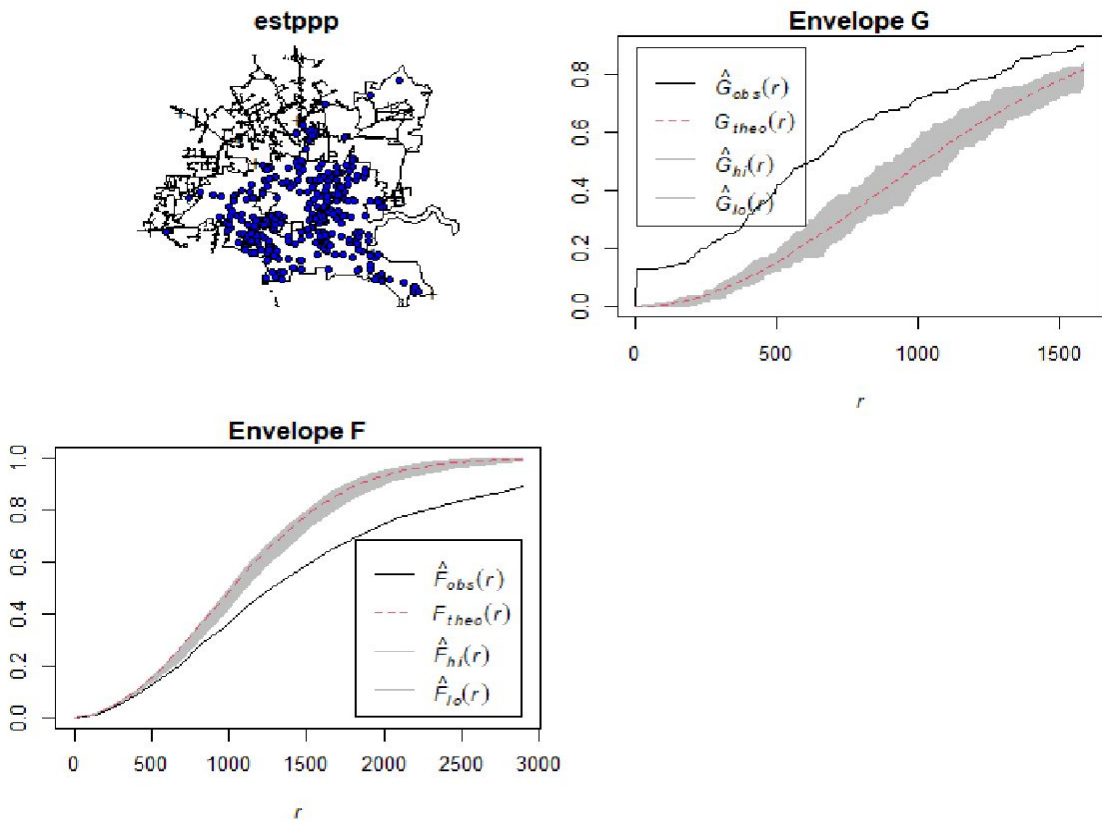
data: assppp

$R = 0.78445$, **p-value = 1.419e-07**

alternative hypothesis: clustered ($R < 1$)

Os dois testes indicam respectivamente que não existe CSR e que existe um padrão de agrupamento para o tipo murder. O próximo será o tipo rape:

Envelope para rape



O gráfico envelope F e G sugerem um padrão de agrupamento.

Clark-Evans test

No edge correction

Z-test

data: estppp

R = 0.81332, **p-value = 6.869e-12**

alternative hypothesis: two-sided

Clark-Evans test

No edge correction

Z-test

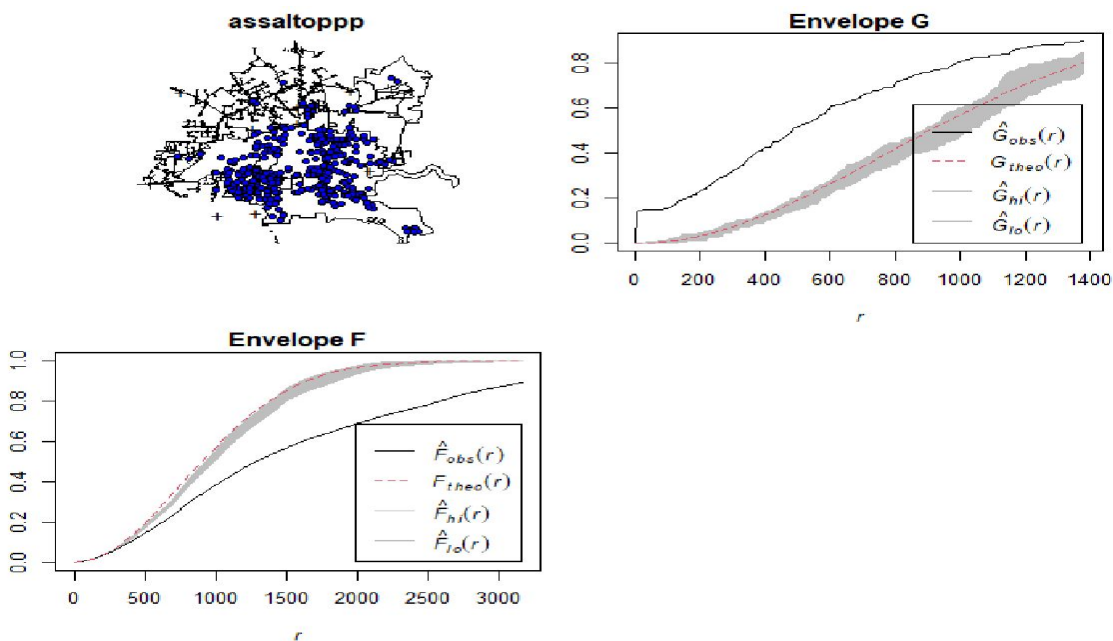
data: estppp

$R = 0.81332$, **p-value = $3.434e-12$**

alternative hypothesis: clustered ($R < 1$)

Os dois testes indicam respectivamente que não existe CSR e que existe um padrão de agrupamento para o tipo rape. O próximo será o tipo robbery:

Envelope para robbery



O gráfico envelope F e G sugerem um padrão de agrupamento.

Clark-Evans test

No edge correction

Z-test

data: assaltoppp

$R = 0.66885$, **p-value < $2.2e-16$**

alternative hypothesis: two-sided

Clark-Evans test

No edge correction

Z-test

data: assaltoppp

$R = 0.66885$, **p-value < 2.2e-16**

alternative hypothesis: clustered ($R < 1$)

Os dois testes indicam respectivamente que não existe CSR e que existe um padrão de agrupamento para o tipo robbery. Assim, podemos concluir que todos os tipos possuem um padrão de agrupamento e de não aleatoriedade espacial completa. Além disso, os tipos rape e murder apresentam um pvalor maior do que os outros tipos de crime, o que pode indicar um padrão de agrupamento mais fraco nestes tipos de crime. Por fim, os padrões para cada tipo apesar de terem uma pequena semelhança, não são similares por apresentarem padrões de agrupamento em regiões específicas a cada tipo.