

### Análise espacial de pontos:

#### Distribuição de espaços culturais e de entretenimento no estado de São Paulo

Nicolas Byung Kwan Cho

O estudo realizado busca entender como estão distribuídos os equipamentos culturais na cidade de São Paulo e como tal organização influencia no consumo destes espaços para habitantes da cidade. Para a análise realizada, os espaços culturais foram divididos em 4 classificações, sendo estas: bibliotecas, espaços culturais, museus e teatros, cinemas e shows. Os dados utilizados foram retirados do portal GeoSampa, que oferece uma quantidade de informações georreferenciadas.

Para uma inspeção inicial, carrega-se os dados para visualização em mapa. Os resultados obtidos podem ser visualizados na figura 1.

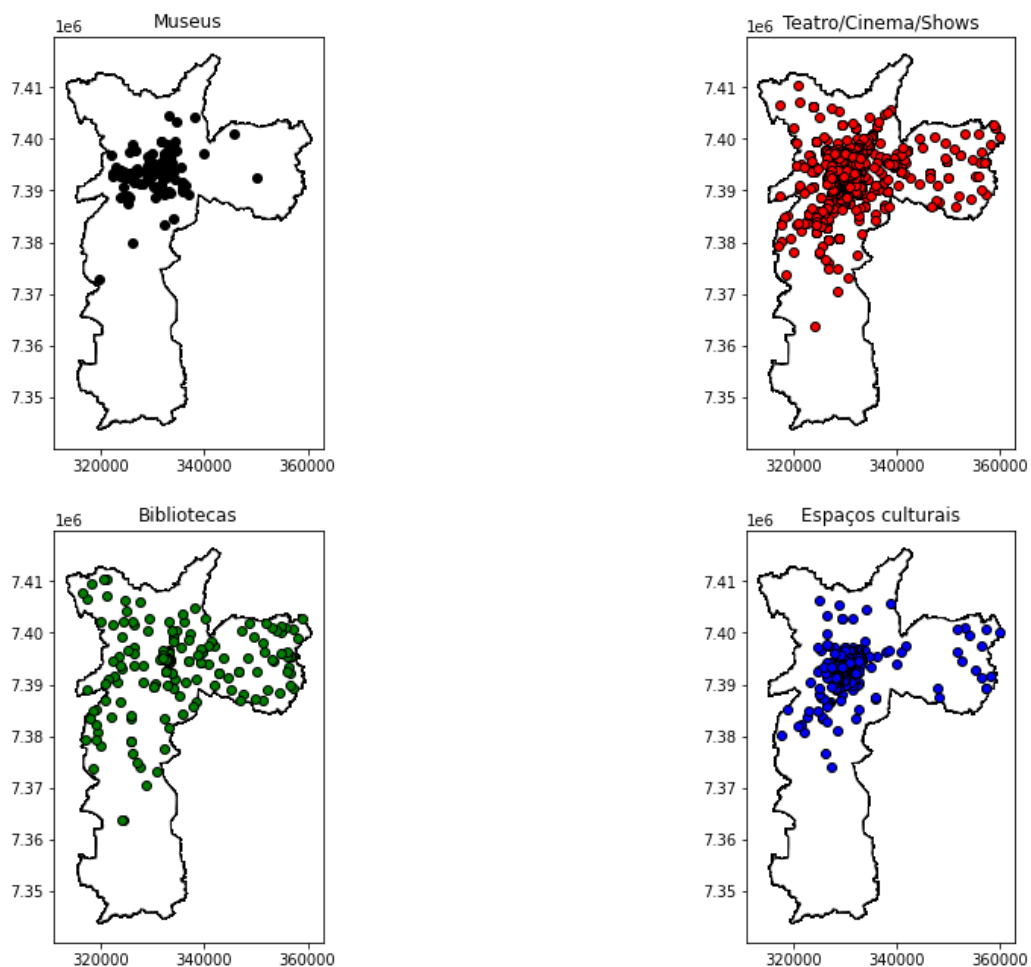
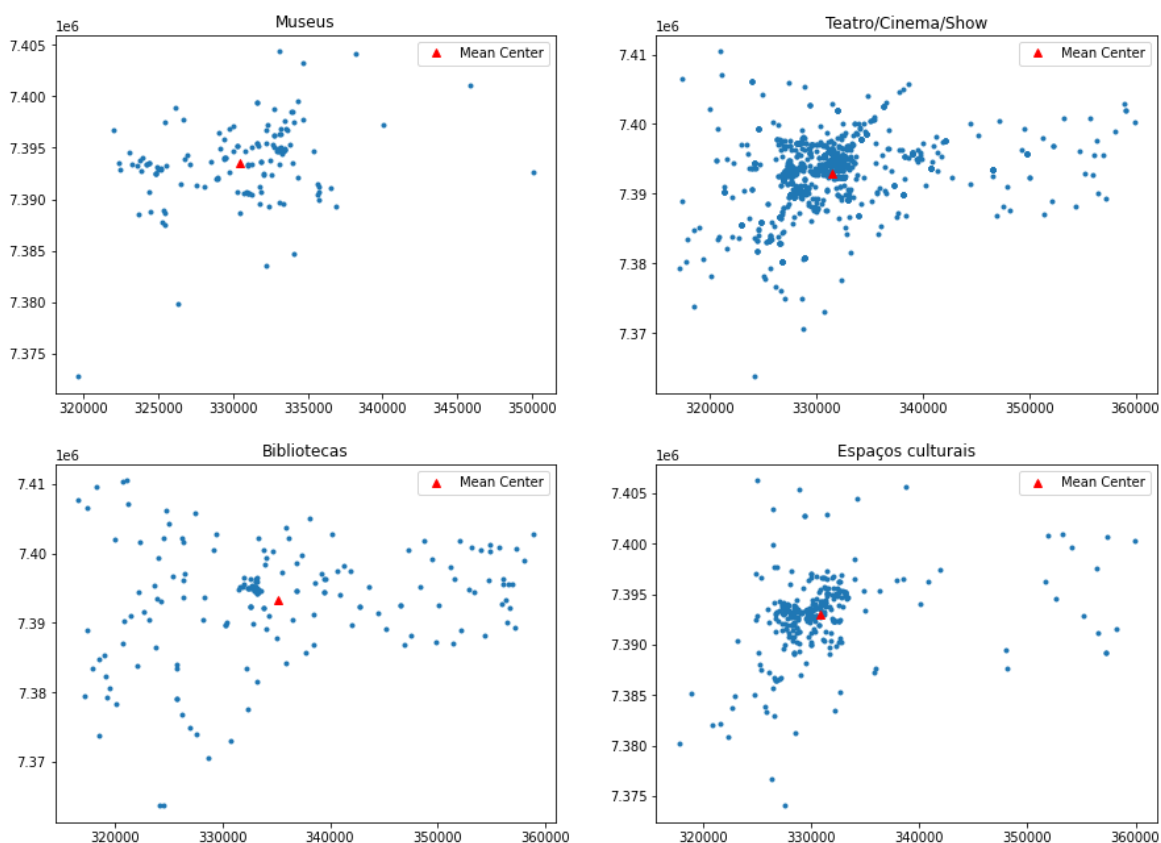


Figura 1-Mapeamento dos equipamentos culturais na cidade de São Paulo

Em uma análise inicial, é possível verificar uma concentração dos espaços em questão na região central da cidade, considerando também uma baixa frequência nas regiões mais periféricas da cidade (como o extremo sul ou leste). Dos equipamentos culturais em questão, verifica-se a melhor distribuição de bibliotecas em relação aos outros espaços.

Determinado os pontos no mapa da cidade de São Paulo, é possível calcular o ponto médio que representa cada classificação analisada, como verificado na imagem 2.



*Figura 2-Centro médio de cada equipamento cultural analisado*

Assim, da figura 2 percebe-se que os centros médios se localizam na região central da cidade. A distância padrão também é calculada para cada espaço analisado, sendo os resultados obtidos (em metros):

Distância padrão Museus: 6013.938569810965  
 Distância padrão Teatro/Cinema/Show: 8000.005341337328  
 Distância padrão Bibliotecas: 13844.584888657462  
 Distância padrão Espaços culturais: 7192.230957756014

A distância padrão é uma medida que funciona de forma similar ao desvio padrão. É uma medida que representa a dispersão de pontos em relação ao centro médio. Assim, valores maiores de distância padrão representa uma distribuição mais dispersa, enquanto valores menores representa uma distribuição mais compacta. Dos resultados obtidos, observa-se que a biblioteca é o espaço com maior dispersão entre os quatro.

Com o centro médio e a distância padrão, é possível visualizar melhor a dispersão da distribuição por meio de um círculo ou elipse de distância padrão. O círculo padrão tem como centro o centro médio dos pontos e utiliza como raio a distância padrão, sendo útil para visualizar a dispersão conforme o tamanho do círculo. Já a elipse é útil para identificar uma tendência de alongamento em certa direção com maior frequência de pontos. Os resultados dos círculos e elipses de distância padrão são identificados na figura 3 e 4.

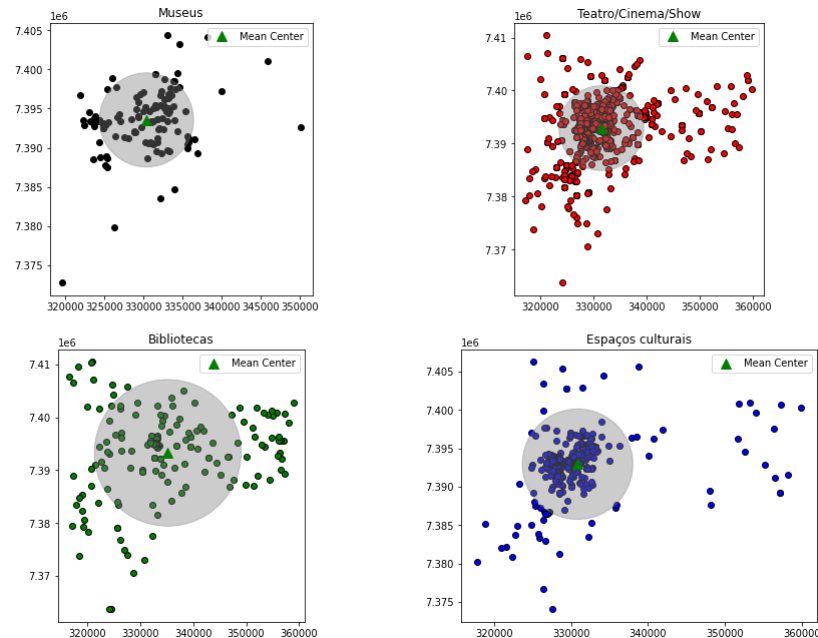


Figura 3-Círculo de distância padrão para cada espaço analisado

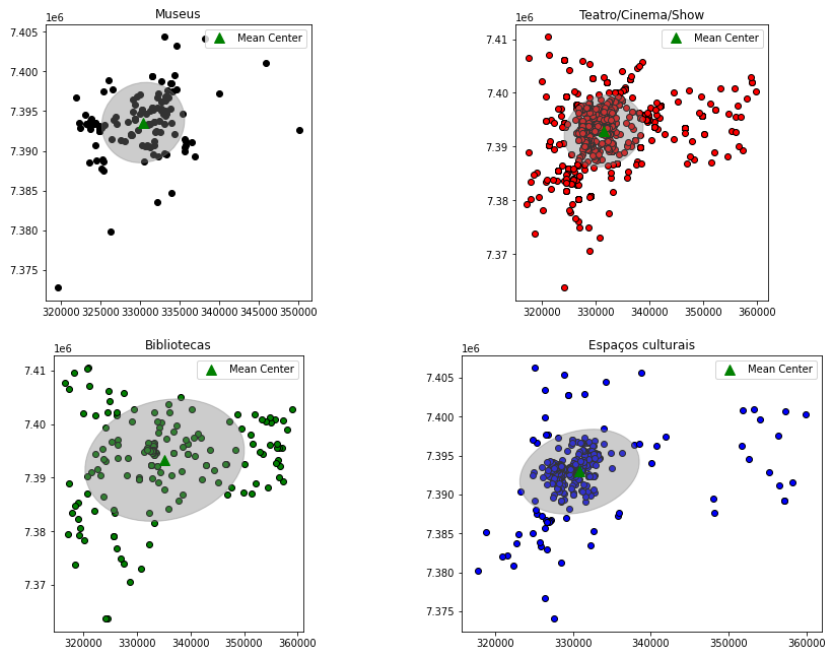
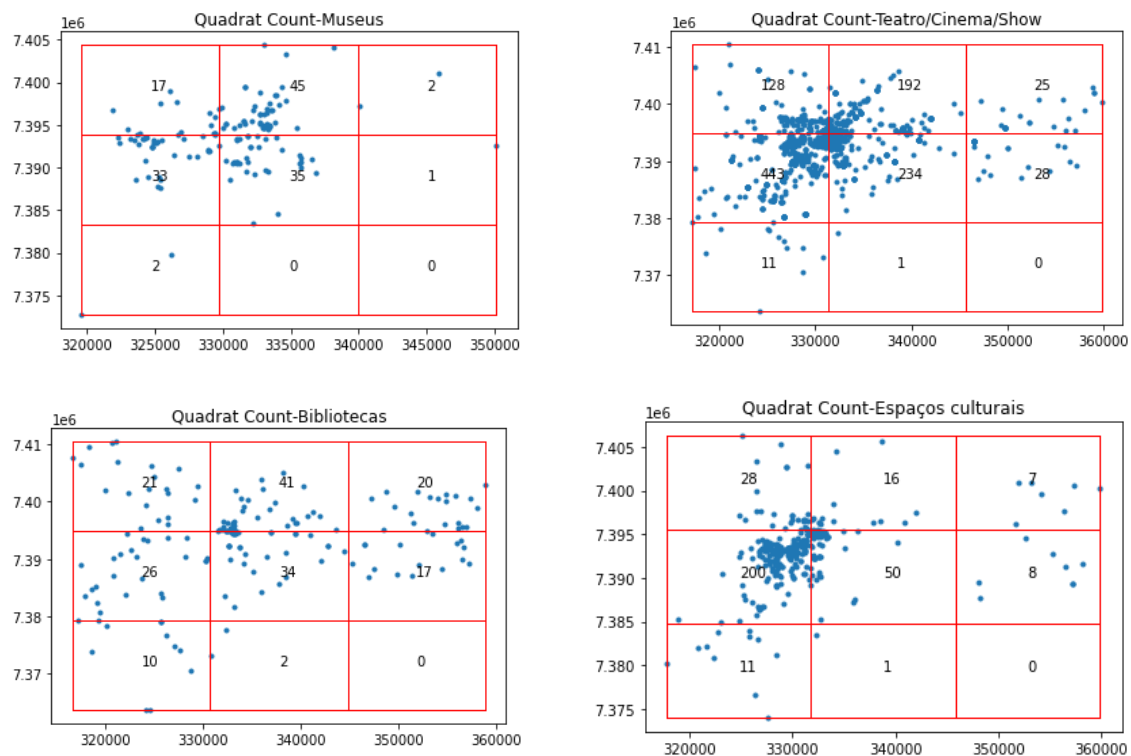


Figura 4-Elipse de distância padrão para cada espaço analisado

Outra forma de analisar a dispersão dos pontos em determinada área é dividi-la em espaços menores e contar quantos pontos ocupam cada espaço, valendo-se, portanto, de uma quadrícula. O resultado da divisão pode ser observado na figura 5.



*Figura 5-Análise de quadrícula para cada equipamento cultural*

Da análise por quadrículas verifica-se a maior dispersão para bibliotecas, comprovando o resultado obtido com a distância padrão.

Para uma análise de densidade de pontos, faz-se o uso de mapas de kernel. Mapas de kernel destacando em cores mais fortes regiões do mapa em que há maior concentração de pontos, facilitando a visualização e identificação de padrões.

Para os equipamentos culturais analisados, a figura 6 ilustra os mapas de kernel gerados.

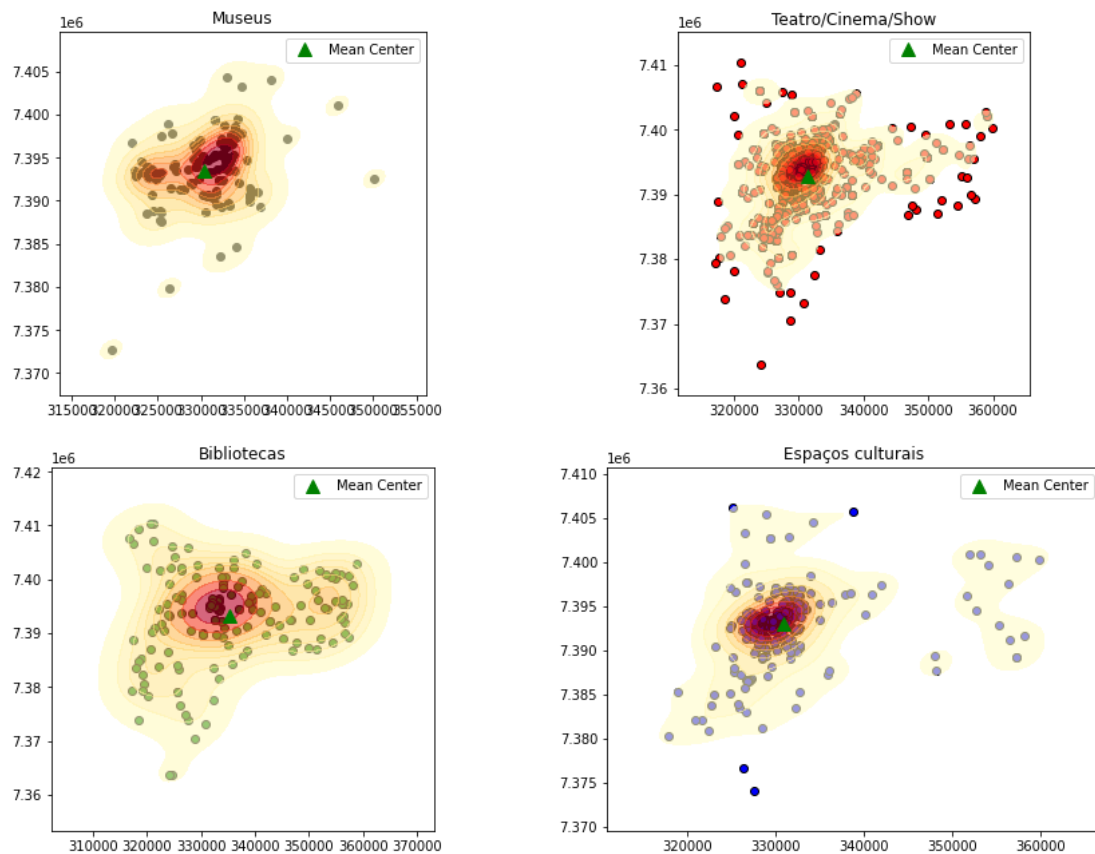


Figura 6-Mapas de kernel para cada espaço analisado

Dos mapas de kernel, visualiza-se concentração de pontos na região central de todos os mapas. Para museus, identifica-se pouca dispersão devido a menor área realçada na imagem. Para teatros, cinemas, shows e espaços culturais, existe grande concentração de pontos na região central, mas também observa-se destaque de cores em regiões mais periféricas, indicando melhor distribuição em relação a museus. Para bibliotecas, apesar da concentração na região central, o mapa de kernel destaca uma área maior do que os outros equipamentos culturais analisados, sendo assim, demonstra a melhor distribuição entre os casos analisados.

Para identificar o espaço que contém os pontos analisados, é possível utilizar uma envoltória convexa (*convex hull*). A envoltória convexa engloba todos os pontos em um único polígono convexo mínimo. A identificação desse polígono é útil para reconhecimento de padrões e análises estatísticas.

As envoltórias convexas identificadas podem ser visualizadas na figura 7.

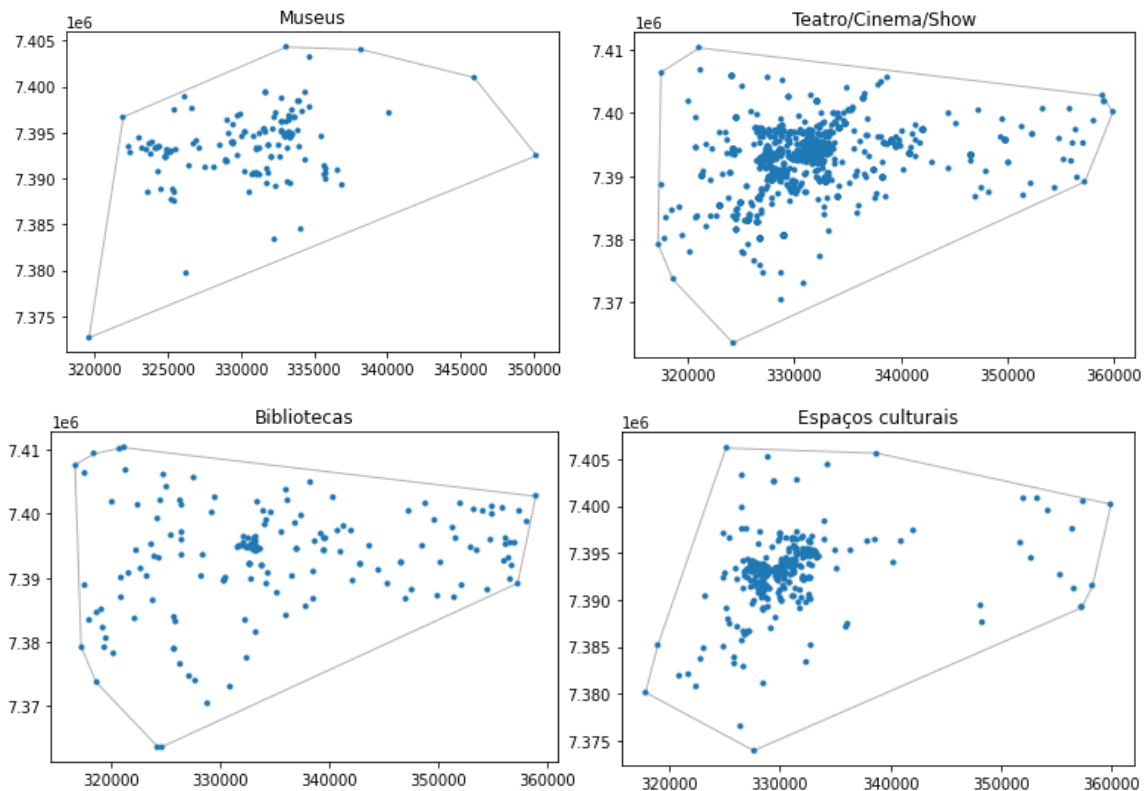


Figura 7-Envoltórias convexas para cada espaço analisado

Em relação a métricas para identificar a dispersão de pontos baseando-se na distância média entre tais, pode-se utilizar da ideia de vizinhos mais próximos entre pontos, *k-nearest neighbours* (K-NN). Este método define a distância média entre  $k$  pontos mais próximos de um determinado ponto e estabelece, também, um valor esperado que depende da quantidade de pontos analisados e do tamanho da área em questão. Assim, valores obtidos abaixo do esperado determinam maior agrupamento de pontos (clusterização), enquanto valores mais próximos ao esperado indicam maior dispersão na área analisada.

Define-se também um índice que serve de padrão para classificar os dados em mais agrupados ou mais dispersos, o *Nearest Neighbour Index* (NNI). O NNI é calculado pela razão entre o valor obtido e o valor esperado, podendo ser classificado segundo o diagrama da figura 8.

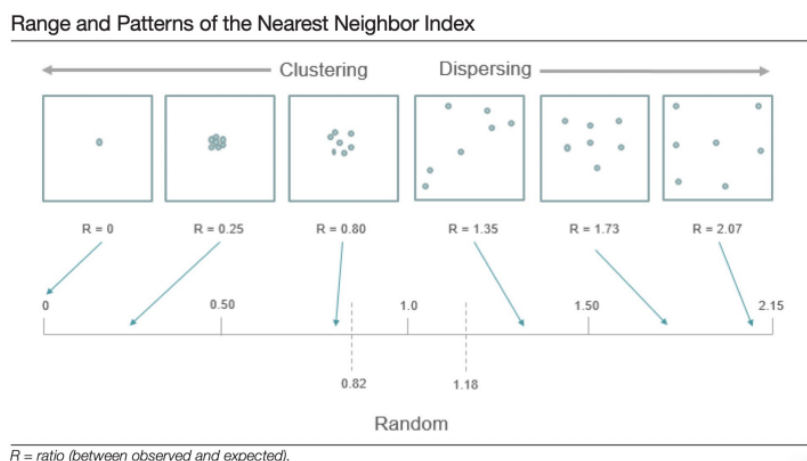


Figura 8-Diagrama de classificação baseando-se no NNI

Os valores obtidos pelo método K-NN são:

Museus:

Valor médio esperado: 1387.2233488719214

Valor médio obtido: 756.6116815506631

Nearest Neighbour Index: 0.545414465641696

Teatro/Cinema/Show:

Valor médio esperado: 694.4585636718741

Valor médio obtido: 227.15902492265963

Nearest Neighbour Index: 0.32710234534596416

Bibliotecas:

Valor médio esperado: 1758.7030848866189

Valor médio obtido: 1179.6136665182646

Nearest Neighbour Index: 0.6707292871976241

Espaços culturais:

Valor médio esperado: 1052.9910458841136

Valor médio obtido: 454.5242077414675

Nearest Neighbour Index: 0.4316505914443359

Logo, todos os equipamentos culturais analisados podem ser considerados clusterizados na cidade de São Paulo, segundo esse método, sendo os teatros, cinemas e shows os mais agrupados e as bibliotecas as mais dispersas.

Além do método dos K-NN, outra métrica muito valorizada na análise espacial de pontos é a função K de Ripley. Neste método, diferentes iterações são realizadas definindo círculos com centros em cada ponto e raios com aumento gradual a cada iteração. Para cada valor de raio verifica-se a quantidade de pontos que cada círculo abrange e cria-se uma métrica ao final de todas as iterações que define o agrupamento ou a dispersão do conjunto. A métrica é dada pela função K.

Assim como no método dos vizinhos mais próximos, a função K de Ripley também define valores esperados caso a distribuição fosse aleatória, além de intervalos de confiança

(envelopes) que englobam este valor esperado. Curvas de K acima do intervalo indicam maior agrupamento dos pontos (clusterização) e curvas abaixo indicam dispersão.

A figura 9 ilustra o resultado da função K de Ripley nos pontos analisados com 250 iterações realizadas.

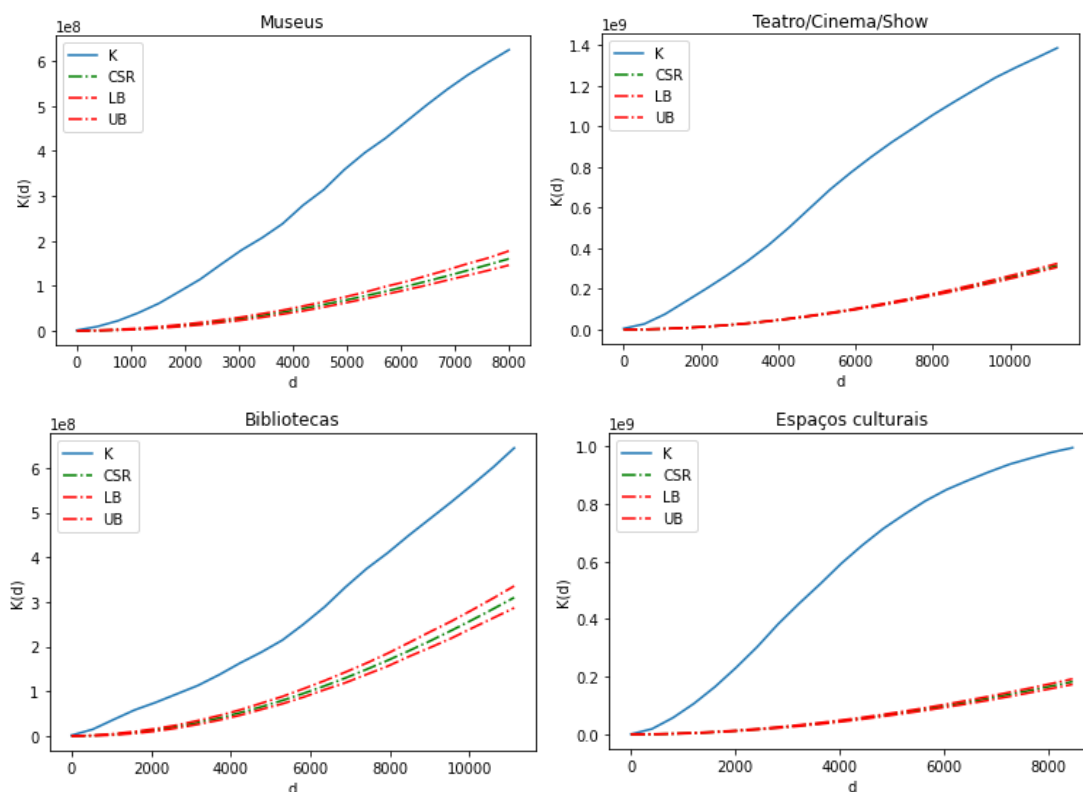


Figura 9-Função K de Ripley aplicada nos pontos analisados

Dos resultados obtidos, pode-se observar que há clusterização para todos os equipamentos culturais, correspondendo com o observado na análise de K-NN. Também se verifica a menor dispersão por parte das bibliotecas, que apresenta a menor discrepância em relação a curva esperada (CSR), e maior agrupamento de teatros, cinemas e show, seguido dos espaços culturais.

De todas as análises realizadas com os equipamentos culturais na cidade de São Paulo, é possível chegar em algumas conclusões. A partir da função K de Ripley e do método dos vizinhos mais próximos (K-NN), verifica-se clusterização dos pontos para todos os 4 tipos analisados, sendo a biblioteca o tipo mais disperso e os teatros, cinemas e shows o tipo mais agrupado.

Em relação a localização do centro médio dos pontos, dos mapas de kernel e da análise de quadrículas, verifica-se que o agrupamento de pontos identificado se localiza mais próximo a região central da cidade de São Paulo para todos os equipamentos. Assim, é possível afirmar que a distribuição de espaços culturais e de entretenimento na cidade de São Paulo é irregular, tendo uma concentração muito grande desses espaços em regiões mais próximas ao centro, enquanto há maior escassez em regiões periféricas e limítrofes da cidade.