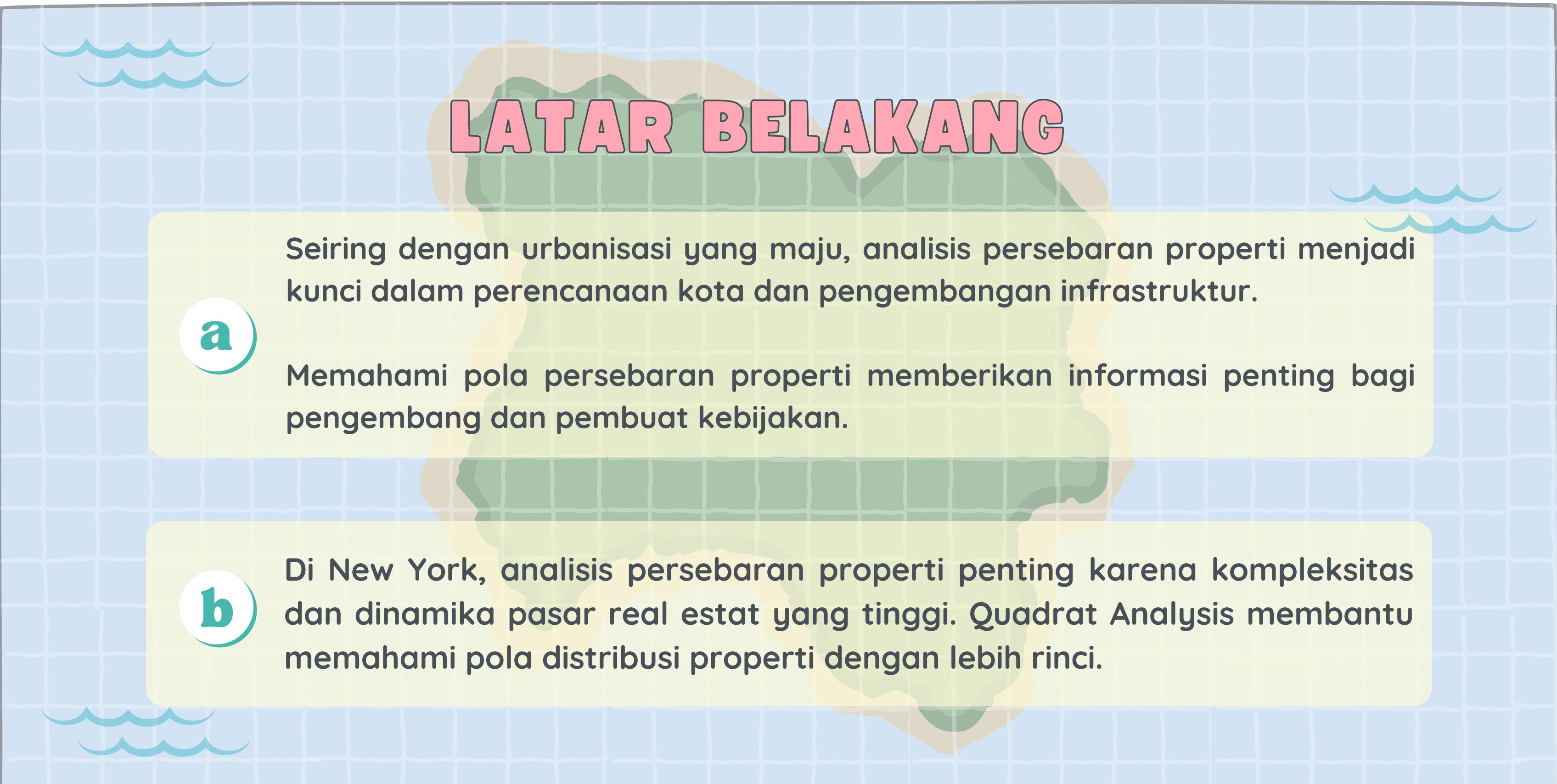


ANALISIS PERSEBARAN PROPERTY DI NEW YORK

**QUADRAT ANALYSIS &
NEAREST NEIGHBOUR ANALYSIS**

ZAHRAH MAHFUZAH



LATAR BELAKANG

Seiring dengan urbanisasi yang maju, analisis persebaran properti menjadi kunci dalam perencanaan kota dan pengembangan infrastruktur.

a

Memahami pola persebaran properti memberikan informasi penting bagi pengembang dan pembuat kebijakan.

b

Di New York, analisis persebaran properti penting karena kompleksitas dan dinamika pasar real estat yang tinggi. Quadrat Analysis membantu memahami pola distribusi properti dengan lebih rinci.

LATAR BELAKANG

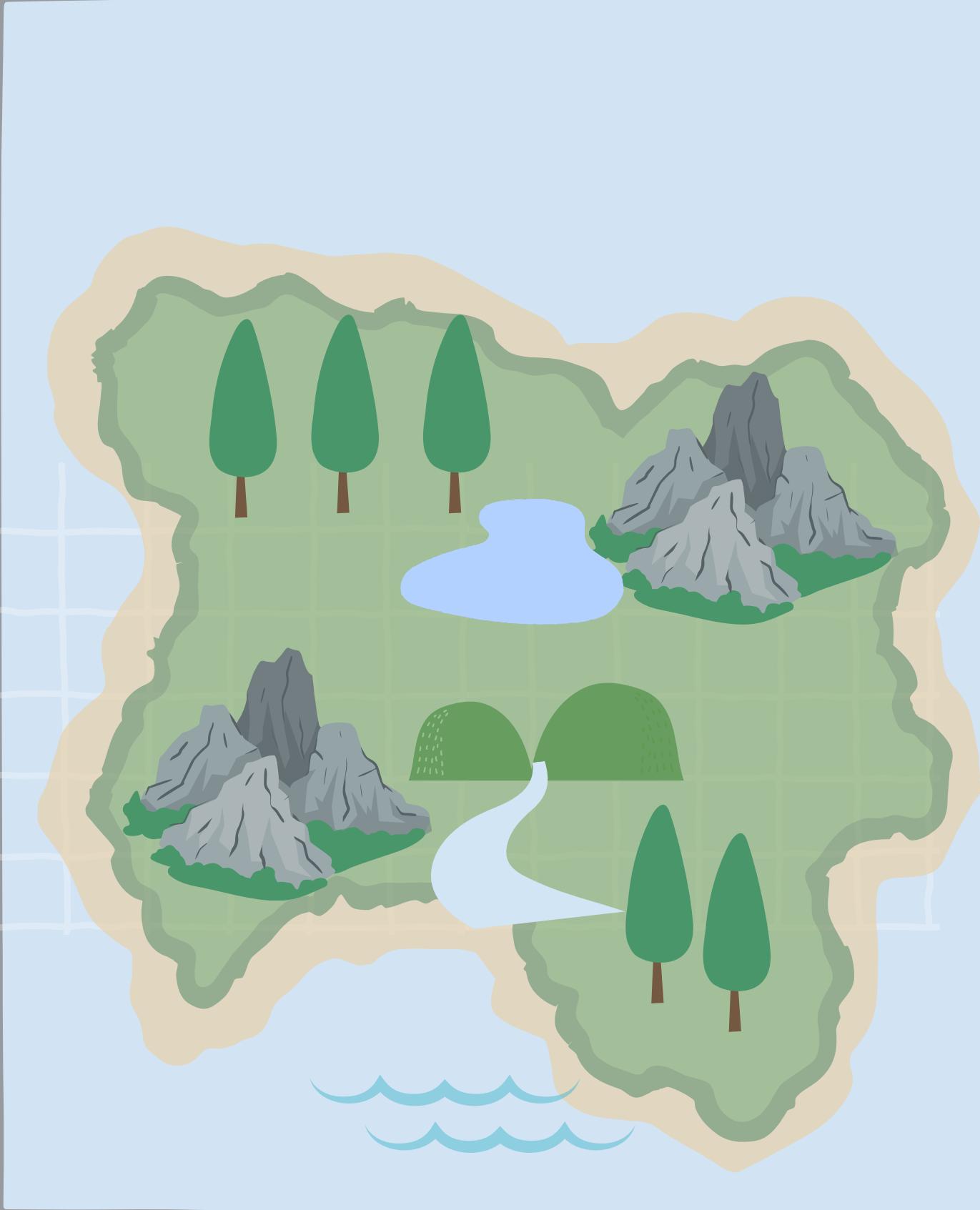
Metode Quadrat Analysis memetakan distribusi titik data dan mengidentifikasi pola distribusi properti berdasarkan kriteria seperti harga dan jenis.

c

Ini berguna untuk analisis pasar, investasi, serta perencanaan kota dan kebijakan publik.

d

Penelitian ini bertujuan untuk menganalisis persebaran properti di New York menggunakan Quadrat Analysis, dengan harapan memberikan wawasan mendalam tentang pasar properti dan mendukung perencanaan serta pengembangan di masa depan.



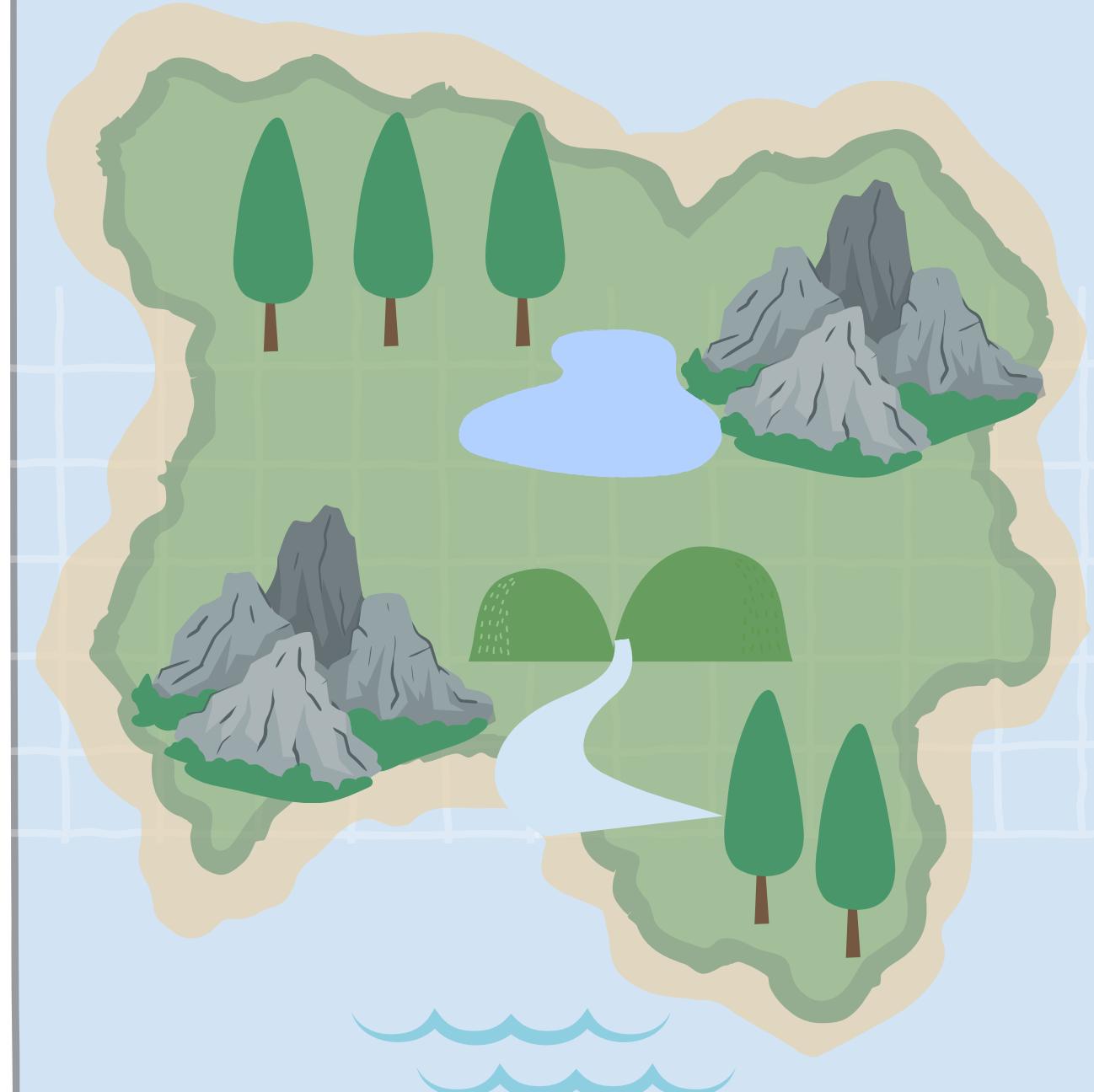
Rumusan Masalah

- Apakah persebaran properti di Kota New York tersebar secara spasial acak/ Complete Spatial Randomness (CSR)?
- Jika tidak, apakah persebaran terjadi secara berkelompok atau teratur (clustered or uniform)?



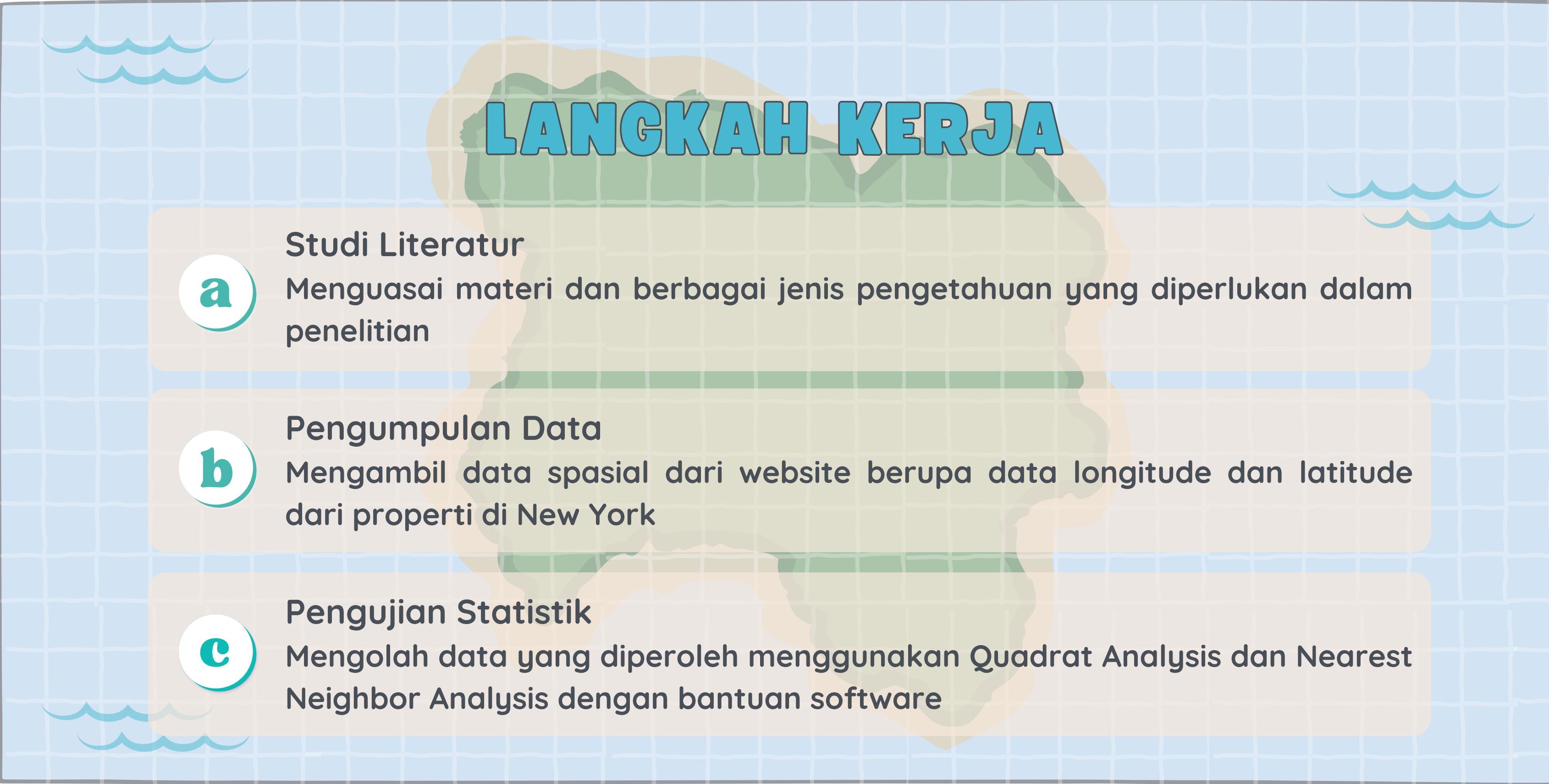
Tujuan Penelitian

- Untuk mengetahui persebaran properti di Kota New York apakah tersebar secara spasial acak (CSR) atau tidak.
- Jika tidak, akan dicari tahu apakah persebaran terjadi secara berkelompok atau teratur.



Manfaat Penelitian

- Memperdalam pemahaman dan keterampilan peneliti dalam menggunakan metode Quadrat Analysis dan Nearest Neighbor Analysis untuk analisis spasial.
- Membantu perencanaan dan pengembangan infrastruktur, seperti transportasi, utilitas, dan fasilitas umum, sehingga lebih sesuai dengan kebutuhan masyarakat.



LANGKAH KERJA

a

Studi Literatur

Menguasai materi dan berbagai jenis pengetahuan yang diperlukan dalam penelitian

b

Pengumpulan Data

Mengambil data spasial dari website berupa data longitude dan latitude dari properti di New York

c

Pengujian Statistik

Mengolah data yang diperoleh menggunakan Quadrat Analysis dan Nearest Neighbor Analysis dengan bantuan software



LANGKAH KERJA

d

Analisis Hasil Analisis Quadrat

- Membuat grafik persebaran
- Melakukan uji hipotesis untuk mengetahui apakah terdapat persebaran secara acak atau tidak mengikuti Complete Spatial Randomness (CSR)
- Jika H_0 ditolak, akan dilakukan uji lanjutan untuk mengetahui apakah data tersebar secara uniform/regular atau clustered menggunakan David-Moore index (ICS)

Analisis Nearest Neighbor

- Membuat plot dan menandai titik koordinat dari data
- Membuat grafik G, F, dan K function untuk melihat apakah data tersebar secara random, uniform/regular, atau clustered

DATA

No	Address	Latitude	Longitude
1	Regis Residence, 2 E 55th St #803, New York, NY 10022, USA	40.76125	-73.97448
2	217 W 57th St, New York, NY 10019, USA	40.76639	-73.98099
3	620 Sinclair Ave, Staten Island, NY 10312, USA	40.54181	-74.19611
4	2 E 55th St, New York, NY 10022, USA	40.76140	-73.97461
5	5 E 64th St, New York, NY 10065, USA	40.76722	-73.96986
:	:	:	:
4801	460 Neptune Ave #14a, Brooklyn, NY 11224, USA	40.57915	-73.97095

Data yang dikumpulkan adalah data titik koordinat latitude dan longitude real estate di New York melalui data sekunder yang diperoleh dari situs [kaggle.com](https://www.kaggle.com) .

PRE-PROCESSING DATA

Data Collection

Data titik koordinat (latitude, longitude) real estate di New York diambil dari situs Kaggle. Data terdiri dari 4801 baris dan 17 kolom, yaitu

- BROKERTITLE: Nama broker
- TYPE: Tipe rumah
- PRICE: Harga rumah
- BEDS: Jumlah kamar tidur
- BATH: Jumlah kamar mandi
- PROPERTYSQFT: Luas persegi dari properti
- ADDRESS: Alamat lengkap rumah
- STATE: Nama state rumah
- MAIN_ADDRESS: Informasi alamat utama
- ADMINISTRATIVE_AREA_LEVEL_2: Informasi tingkat wilayah administrasi
- LOCALITY: Informasi lokalitas
- SUBLOCALITY: Informasi sublokalisasi
- STREET_NAME: Nama jalan
- LONG_NAME: Nama panjang jalan
- FORMATTED_ADDRESS: Alamat yang diformat
- LATITUDE: Koordinat lintang rumah
- LONGITUDE: Koordinat bujur rumah

Penghapusan duplikasi data

Dari 4801 data menjadi 4587 data. Lalu dilakukan penghapusan duplikat data berdasarkan longitude dan latitude serta mengantinya dengan harga properti tertinggi. Hasil akhir data menjadi 4196.

```
# Remove duplicate
data <- distinct(data)

data_cleaned <- data %>%
  group_by(LONGITUDE, LATITUDE) %>%
  filter(PRICE == max(PRICE)) %>% # Memilih baris dengan harga tertinggi
  ungroup()

# Round longitude and latitude to avoid near-duplicate points
data_cleaned <- data_cleaned %>%
  mutate(
    LATITUDE = round(LATITUDE, 5),      # Round to 5 decimal places
    LONGITUDE = round(LONGITUDE, 5)
  ) %>%
  distinct(LATITUDE, LONGITUDE, .keep_all = TRUE)
```

PRE-PROCESSING DATA

Pengeditan kolom

Menghapus awalan yang tidak diperlukan, mengganti nama kolom, dan mengubah tipe data kolom 'BATH' menjadi integer.

```
# Remove prefix "Brokered by " from the "BROKERTITLE" column  
data$BROKERTITLE <- gsub("Brokered by ", "", data$BROKERTITLE)  
# Remove suffix " for sale" from the "TYPE" column  
data$TYPE <- gsub(" for sale$", "", data$TYPE)  
# Convert the "baths" column to integer  
data$BATH <- as.integer(data$BATH)
```

```
# Rename columns and select specific columns  
data <- data %>%  
  rename(  
    SQFT = PROPERTYSQFT,  
  )
```

Handling missing value dan outliers

Menggabungkan entri NA ke sub locality terdekat dan menghapus outlier di Richmond karena kesalahan entri data.

```
# Replace "The Bronx" with "Bronx County"  
data$SUBLOCALITY <- gsub("The Bronx", "Bronx County", data$SUBLOCALITY)  
# Replace "East Bronx" with "Bronx County"  
data$SUBLOCALITY <- gsub("East Bronx", "Bronx County", data$SUBLOCALITY)  
# Replace "Brooklyn Heights" with "Brooklyn"  
data$SUBLOCALITY <- gsub("Brooklyn Heights", "Brooklyn", data$SUBLOCALITY)  
# Replace "Dumbo" with "Brooklyn"  
data$SUBLOCALITY <- gsub("Dumbo", "Brooklyn", data$SUBLOCALITY)  
# Replace "Fort Hamilton" with "Brooklyn"  
data$SUBLOCALITY <- gsub("Fort Hamilton", "Brooklyn", data$SUBLOCALITY)  
# Replace "Jackson Heights" with "Queens"  
data$SUBLOCALITY <- gsub("Jackson Heights", "Queens", data$SUBLOCALITY)  
# Replace "Rego Park" with "Queens"  
data$SUBLOCALITY <- gsub("Rego Park", "Queens", data$SUBLOCALITY)  
# Replace "Snyder Avenue" with "Brooklyn"  
data$SUBLOCALITY <- gsub("Snyder Avenue", "Brooklyn", data$SUBLOCALITY)
```

```
# Correct the outlier  
data$PRICE[data$PRICE == 2147483647] <- 2595000
```

QUADRAT METHOD UNTUK CSR

Menguji Complete Spatial Randomness (CSR) dari persebaran lokasi rumah/properti.

Pembagian area

Area lokasi rumah dibagi menjadi sub area yang sama besar, atau 100 quadrats (10x10).

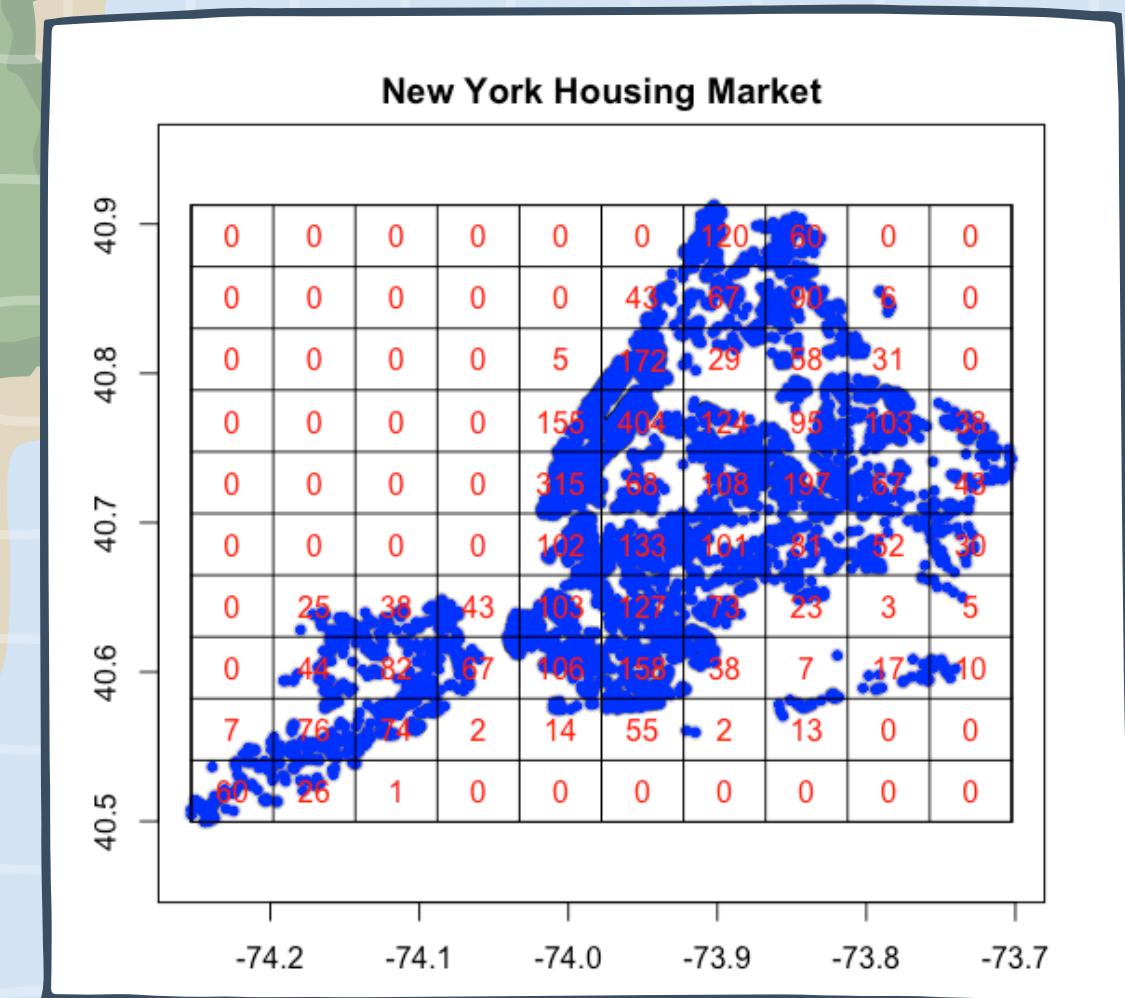
Analisis

Hitung frekuensi dan intensitas titik lokasi pada tiap quadrat serta evaluasi pola persebaran rumah.

```
##QUADRAT METHOD
#Data scaling
x_max<-max(data_cleaned$LONGITUDE)
x_min<-min(data_cleaned$LONGITUDE)
y_max<-max(data_cleaned$LATITUDE)
y_min<-min(data_cleaned$LATITUDE)

#PPP object berdasarkan long and lat
ppp_house<-ppp(data_cleaned$LONGITUDE,data_cleaned$LATITUDE, window = owin(c(x_min,x_max),c(y_min,y_max)))
#Quadrat count
Q <- quadratcount(ppp_house, nx=10, ny=10)
Q

#Plotting PPP, Quadrat, intensity
par(mar = c(2, 2, 2, 2))
plot(ppp_house, axes=T, main="New York Housing Market")
plot(ppp_house, pch=20, cols="blue", main="New York Housing Market", add=TRUE, axes =F)
plot(Q, add=TRUE, col="red")
```



ANALISIS QUADRAT UNTUK DETEKSI POLA SPASIAL

Mengevaluasi distribusi titik dengan memeriksa perubahan densitas di seluruh ruang.

Uji hipotesis

Hipotesis Nol (H_0): Peristiwa menunjukkan keacakan spasial lengkap (CSR)

Hipotesis Alternatif (H_1): Peristiwa secara spasial berkerumun atau tersebar

Hasi

Titik berkerumun di Quadrat Point: 404 (kolom ke-6, baris ke-4) dengan Chi-

Square: $\chi^2 = 10266$, df = 99, p-value < 2.2e-16.

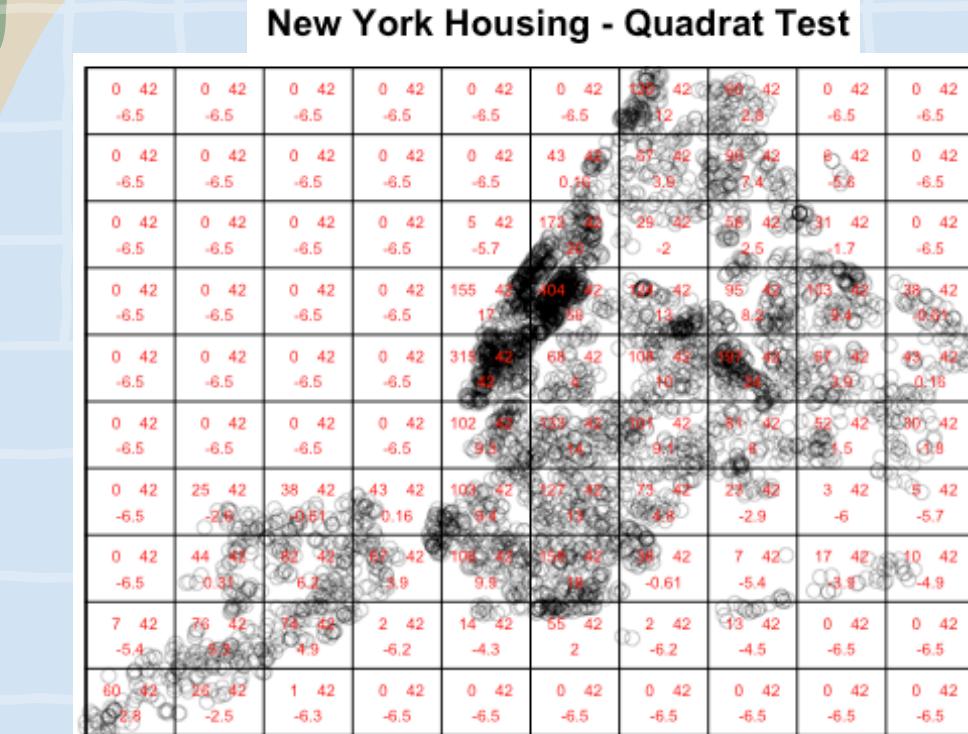
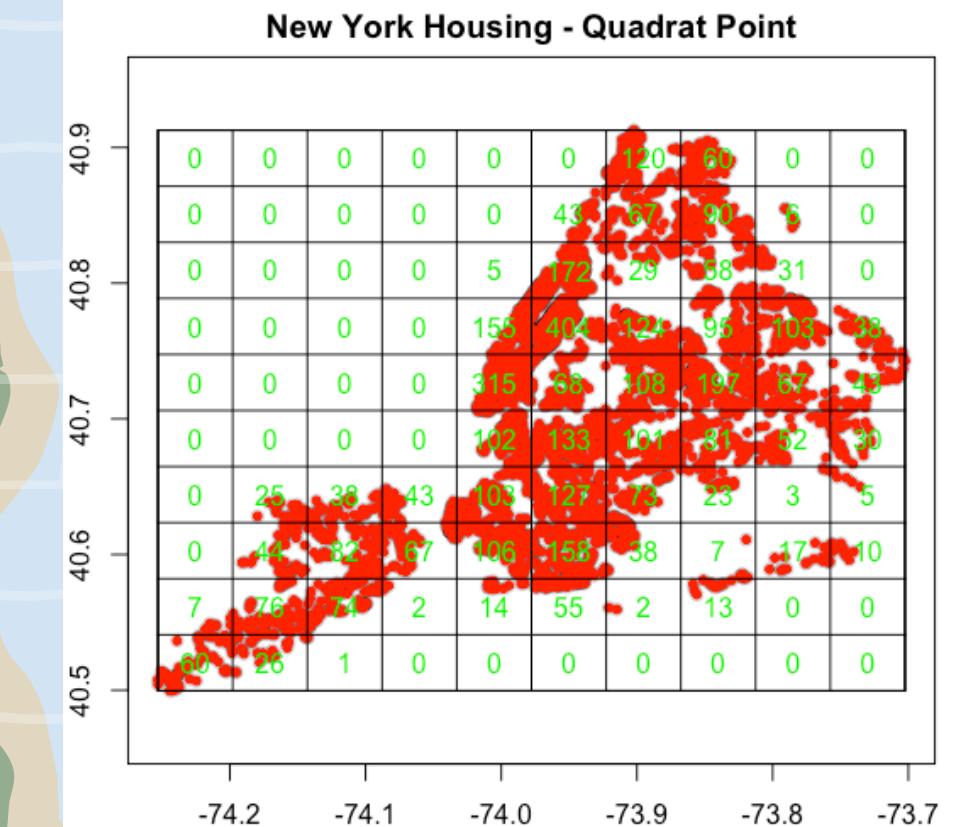
Kesimpulan

H_0 ditolak, maka New York Housing berkerumun dan tidak menunjukkan keacakan spasial lengkap (CSR).

Chi-squared test of CSR using quadrat counts

```
data: ppp_house  
X2 = 10266, df = 99, p-value < 2.2e-16  
alternative hypothesis: two.sided
```

Quadrats: 10 by 10 grid of tiles



UJI LANJUTAN

*Sample
Variance*

4351.049

*Sample
Mean*

41.96

*Relative
Variance (I)*

103.6952

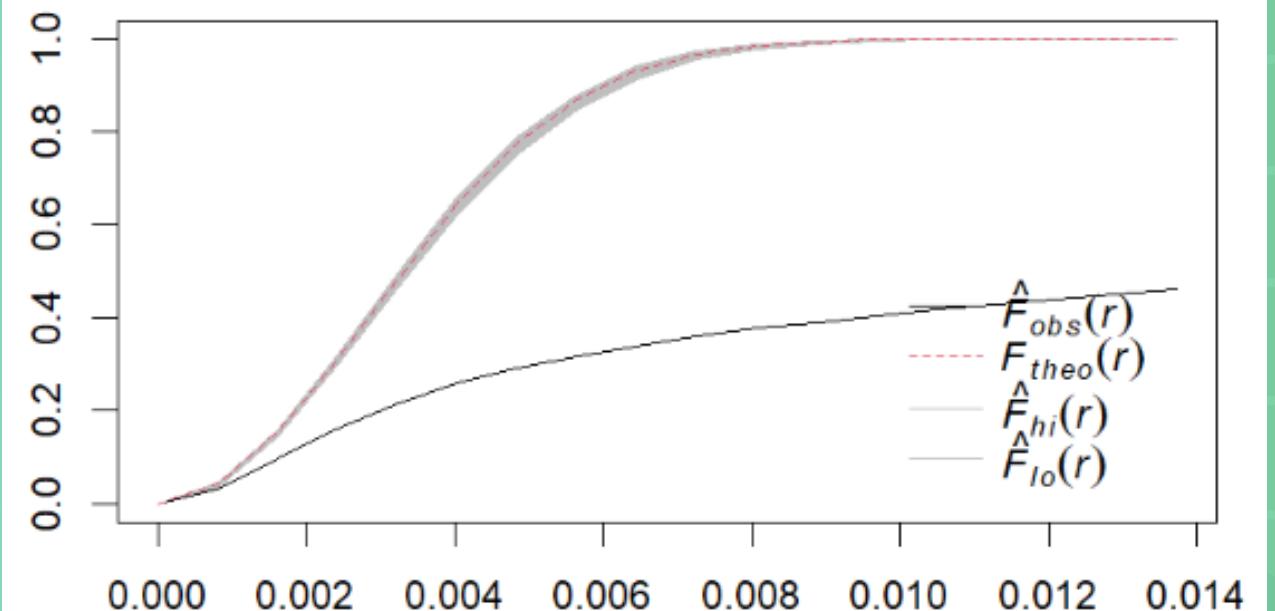
ICS

102.6952

Relative Variance (I) > 1 dan ICS > 0 : Data New York Housing Market berdistribusi clustered.

NEAREST NEIGHBOR ANALYSIS

F Function (Empty Space)



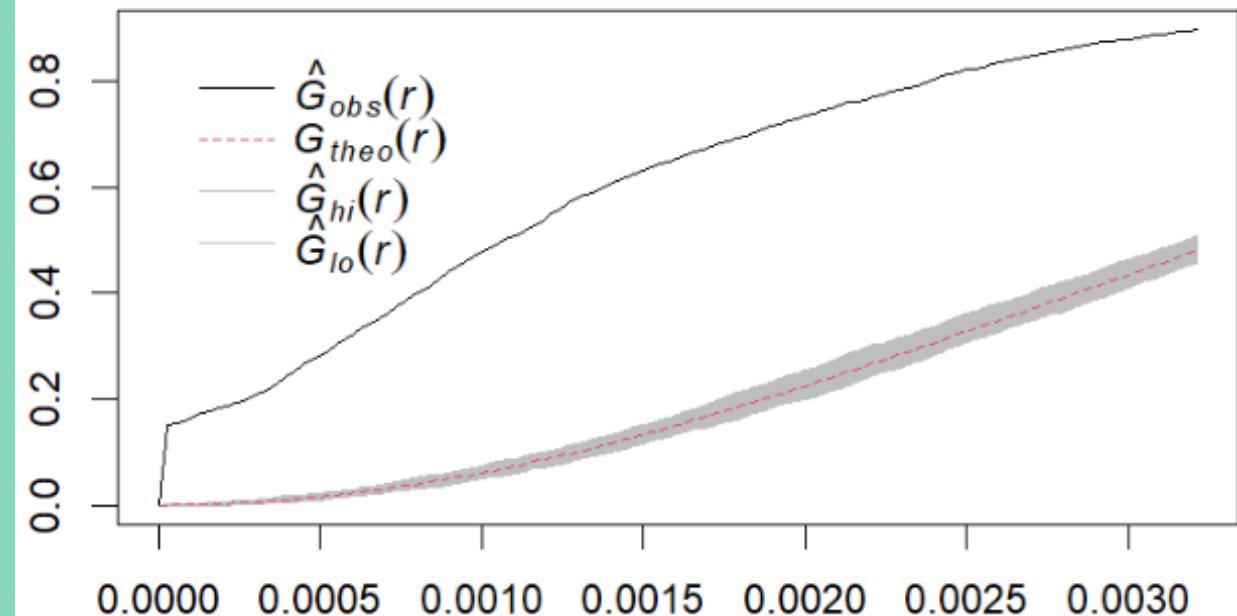
- Garis yang diobservasi (hitam) berada di bawah garis ekspektasi (merah putus-putus).
- Terdapat lebih banyak ruang kosong dalam pola titik tersebut dibandingkan dengan pola acak.

c

F - FUNCTION

NEAREST NEIGHBOR ANALYSIS

G Function (Nearest Neighbor)



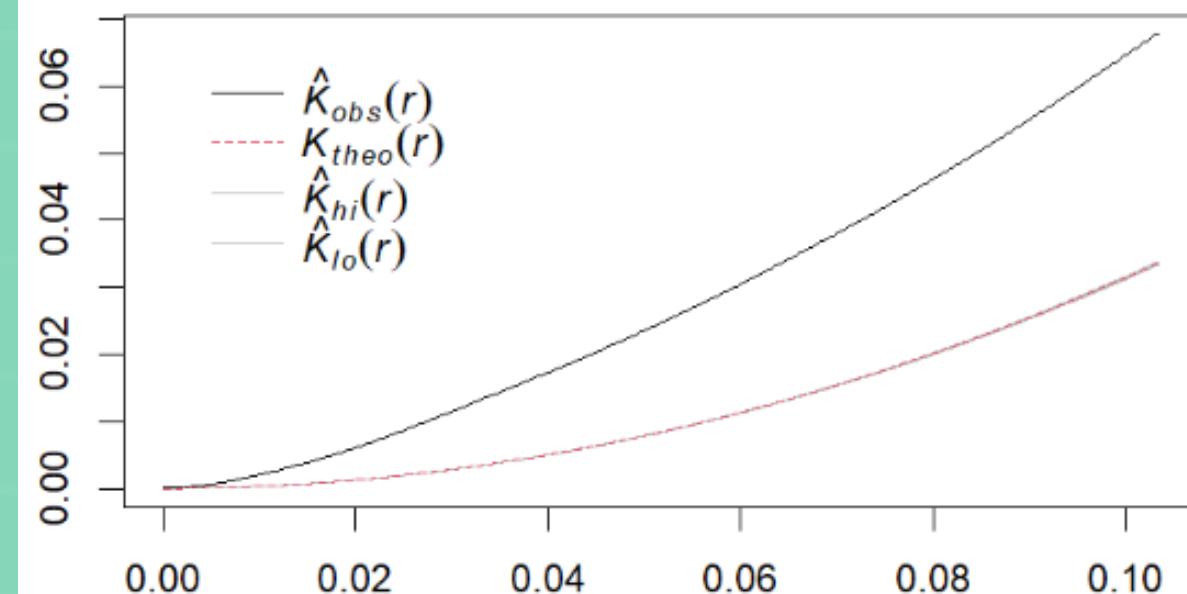
- Garis yang diobservasi (hitam) berada di atas garis ekspektasi (merah putus-putus).
- Ini menandakan bahwa pola titik cenderung berkelompok (clustering)

c

G - FUNCTION

NEAREST NEIGHBOR ANALYSIS

K Function (Ripley's K)



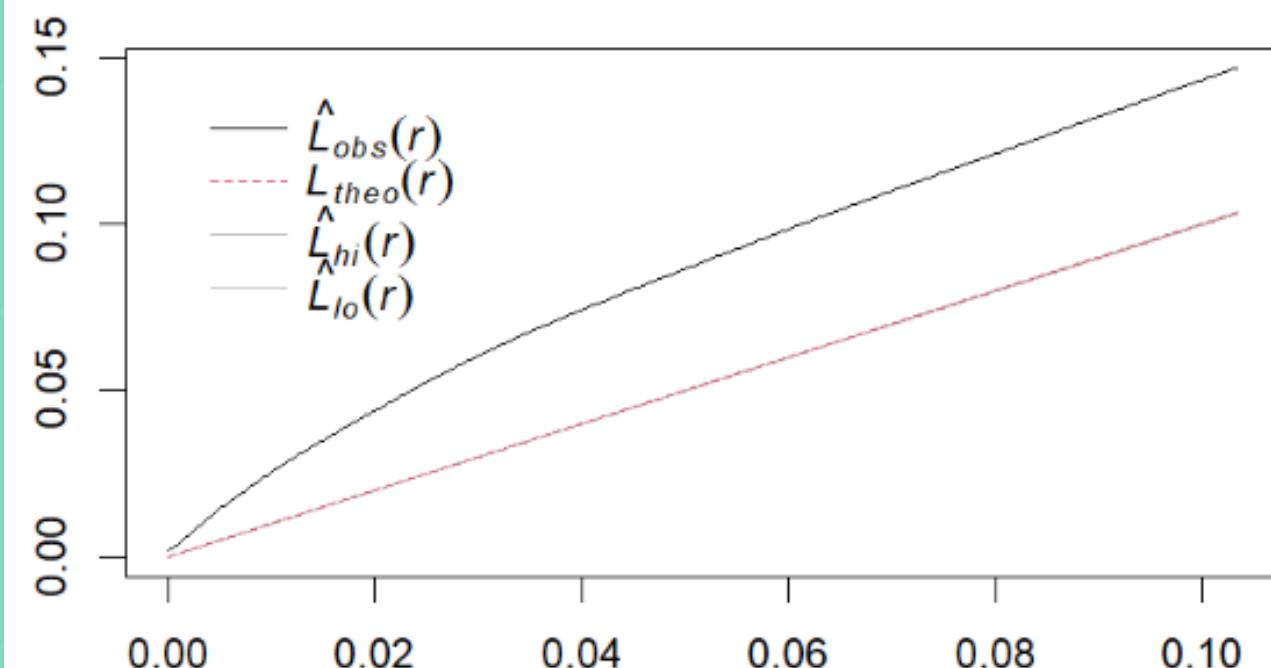
- Garis yang diobservasi (hitam) berada di atas garis ekspektasi (merah putus-putus).
- Ini menandakan bahwa pola titik cenderung berkelompok (clustering)

c

K - FUNCTION

NEAREST NEIGHBOR ANALYSIS

L Function (Modified Ripley's K)



- Garis yang diobservasi (hitam) berada di atas garis ekspektasi (merah putus-putus).
- Ini menandakan bahwa pola titik cenderung berkelompok (clustering)

c

L - FUNCTION

KESIMPULAN

a

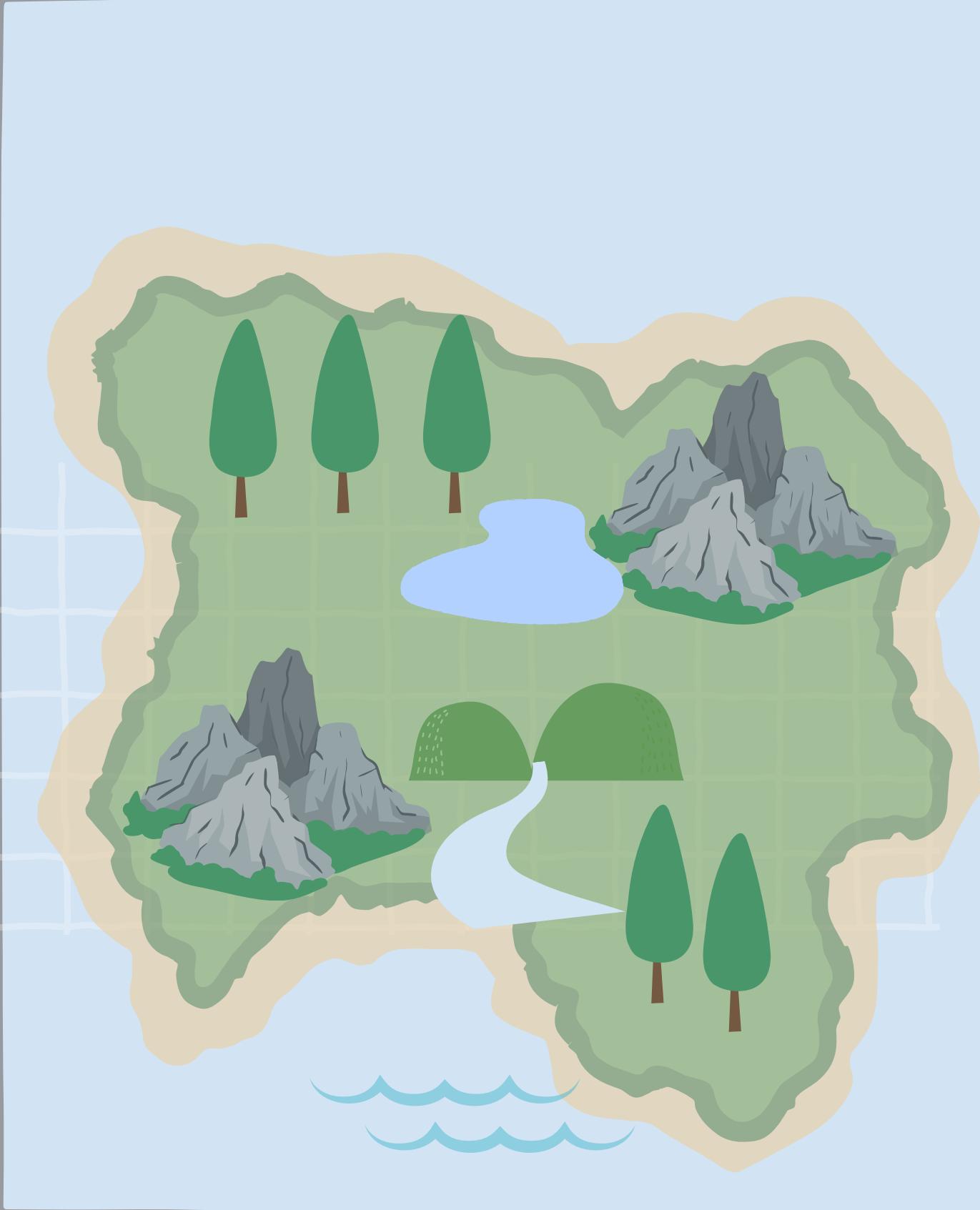
Metode Quadrat Analysis menunjukkan adanya penolakan terhadap hipotesis nol, yang menyatakan bahwa properti di Kota New York tidak menyebar secara acak.

b

Metode Nearest Neighbor Analysis menunjukkan adanya pola clustering yang signifikan dalam penyebaran properti di Kota New York. Hal ini dapat dilihat pada plot fungsi G, F, K, dan L.

c

Dari keseluruhan hasil analisis dapat disimpulkan bahwa penyebaran properti di Kota New York tidak menyebar secara acak dan cenderung mengelompok.



SARAN

- Perlu melakukan uji lanjutan untuk mengetahui karakteristik masing-masing titik lokasi dari persebaran properti di New York.

TCRIMA KASIH