Atomic Spatial Processes

Sean Jewell
Neil Spencer
Alexandre Bouchard-Côté

Department of Statistics
University of British Columbia

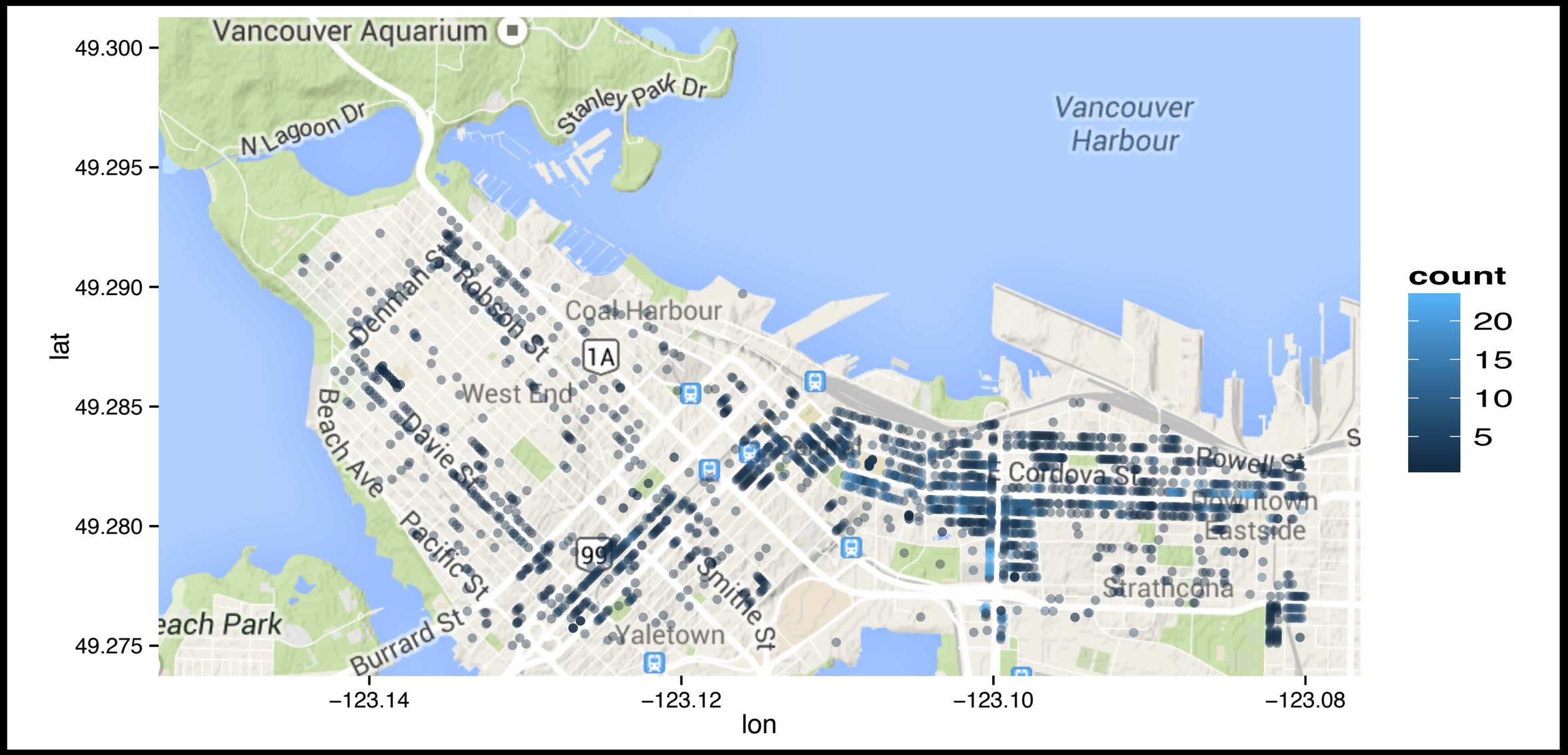
Outline

- Motivation
- Point Processes
- Atomic Spatial Processes
- Tractable Inference Strategy
- Data Analysis Results
- Model Comparisons

Urban Spatial Data

- Often possesses rich and non-standard structure

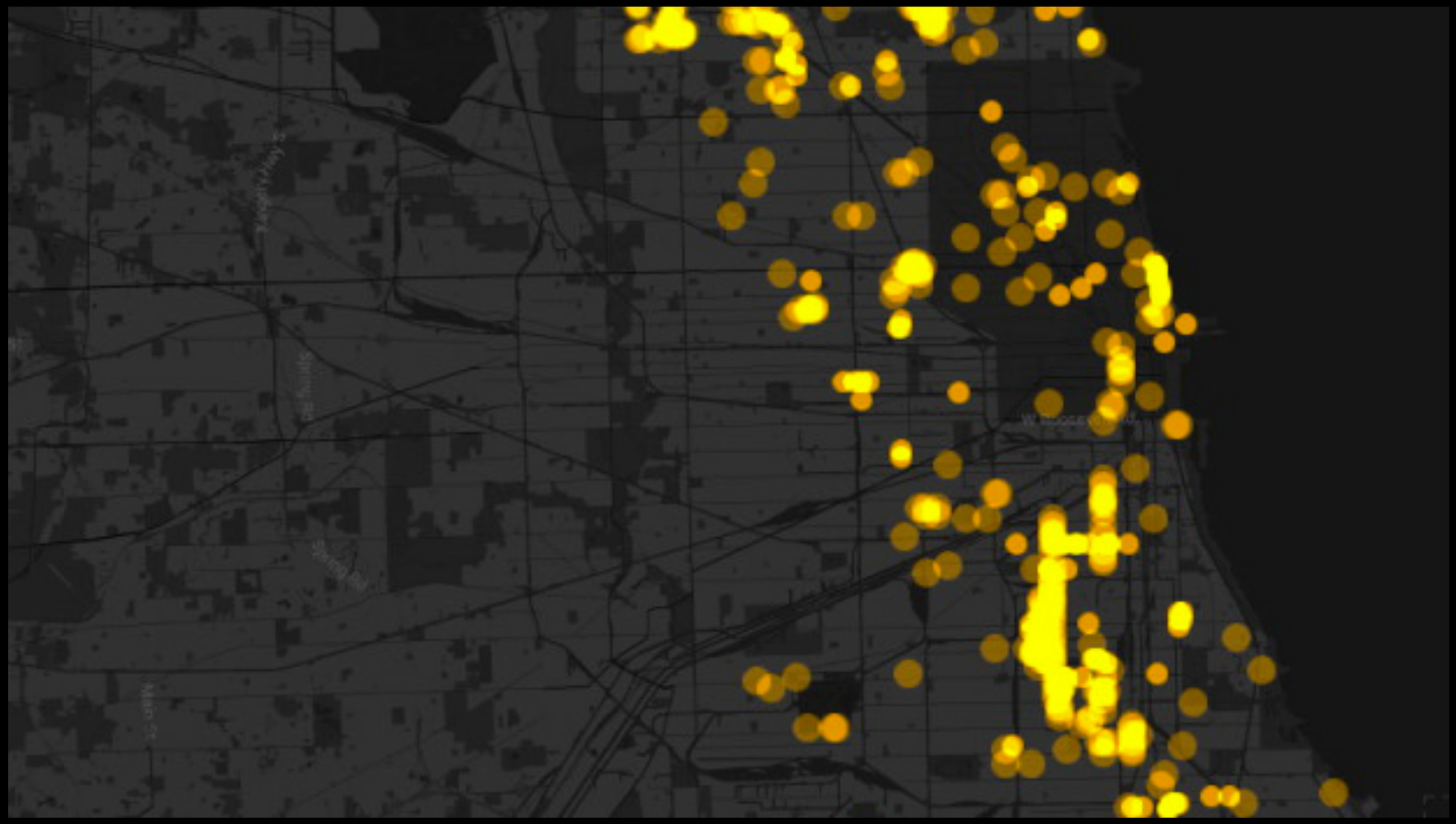
Vancouver Graffiti Locations



Urban Spatial Data

- Often possesses rich and non-standard structure

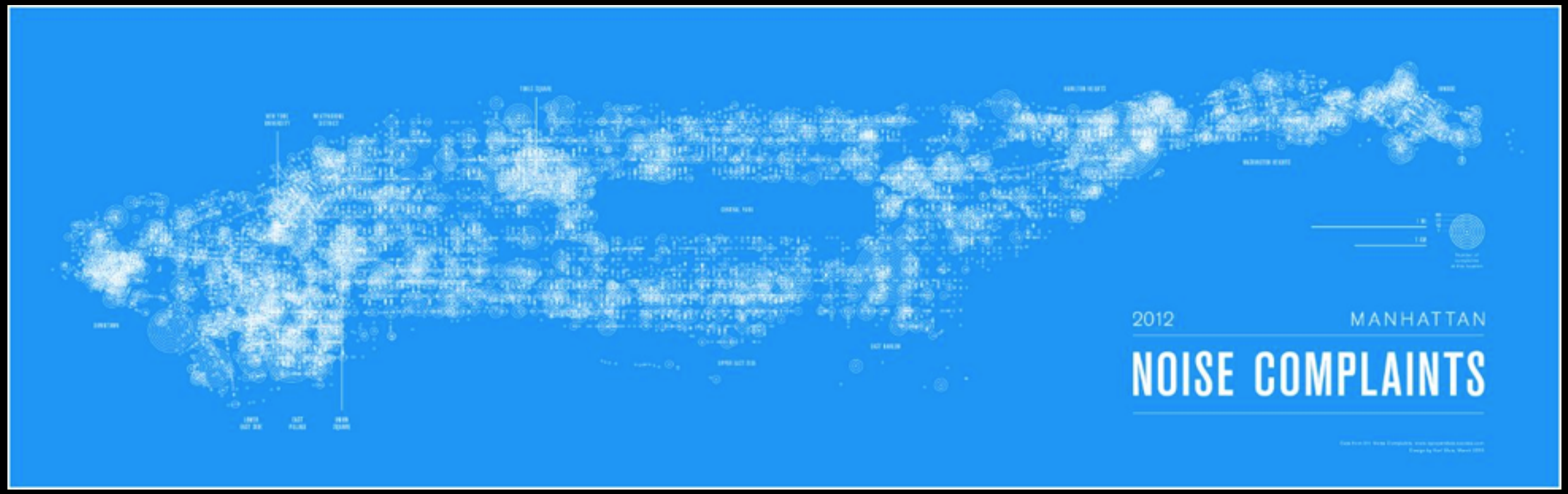
Chicago Pothole Locations



http://chicago.cbslocal.com/2014/04/04/animated-map-tracks-potholes-across-chicago/

Urban Spatial Data

- Often possesses rich and non-standard structure
  Manhattan Noise Complaints

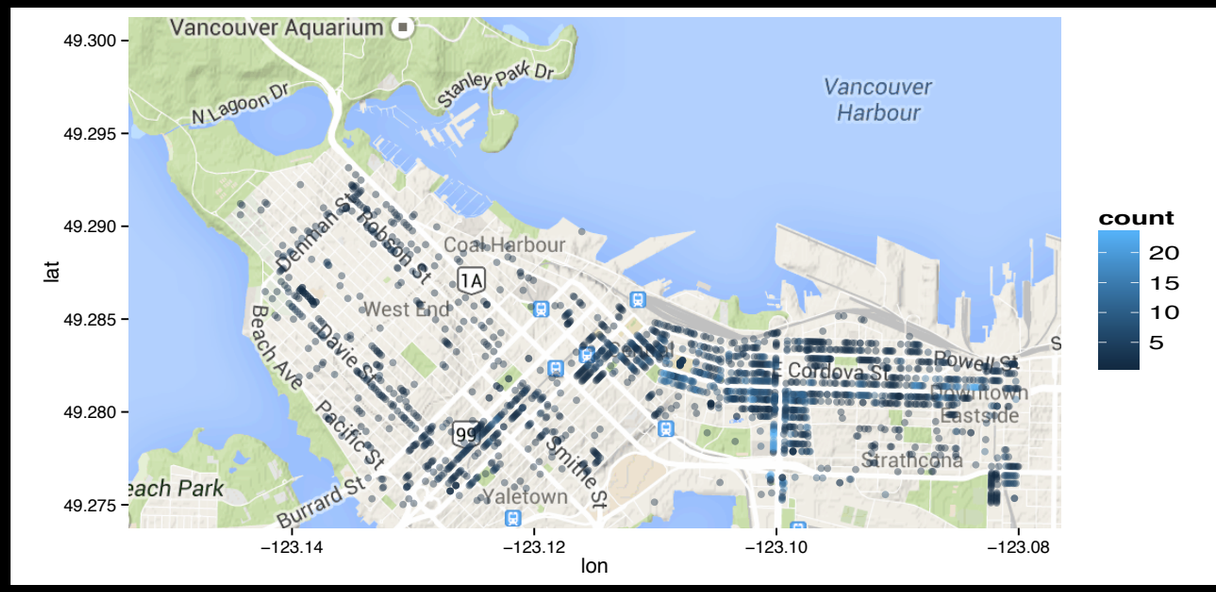


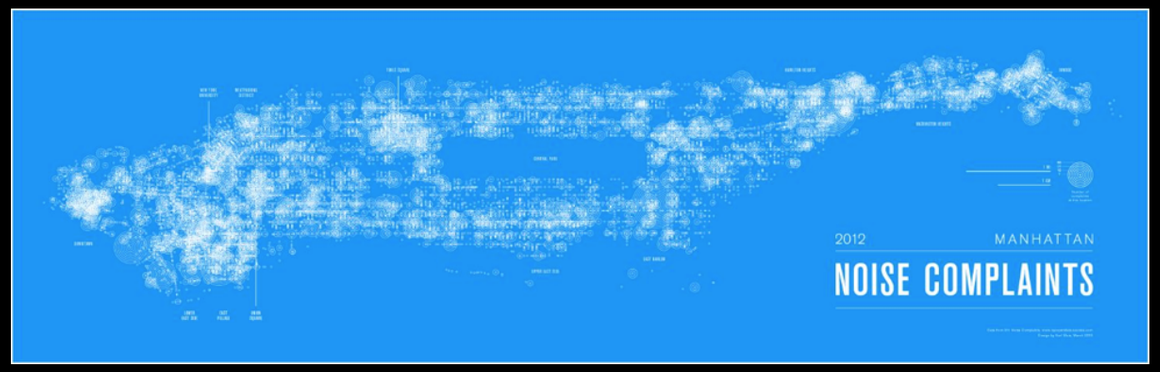
http://www.citylab.com/tech/2013/04/

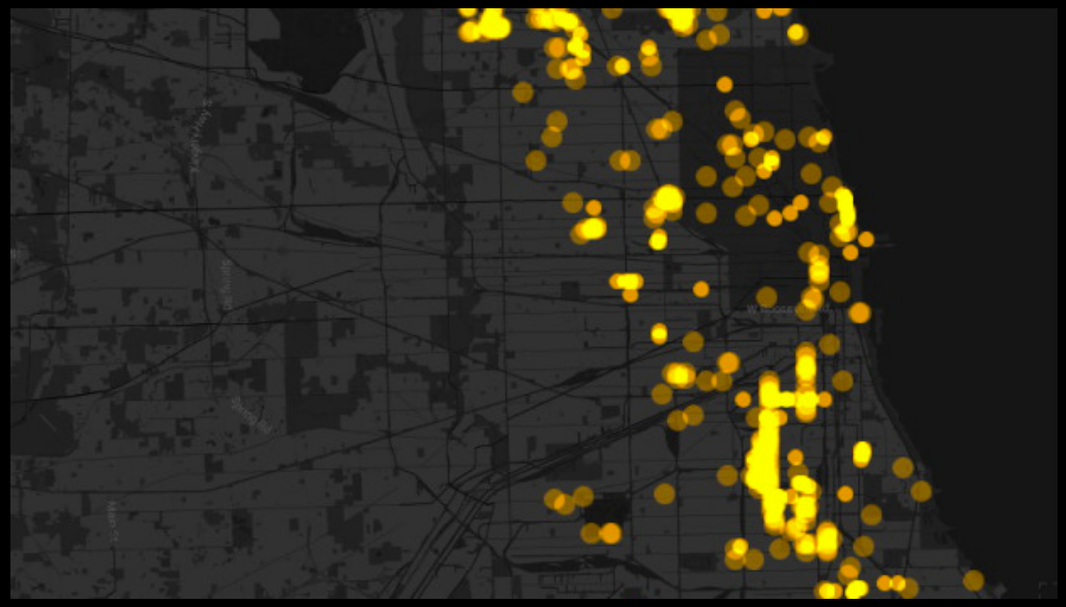
Urban Spatial Data

- Often possesses rich and non-standard structure

- Non-Euclidean
- Non-Continuous
- Inhomogeneous
- Clustered



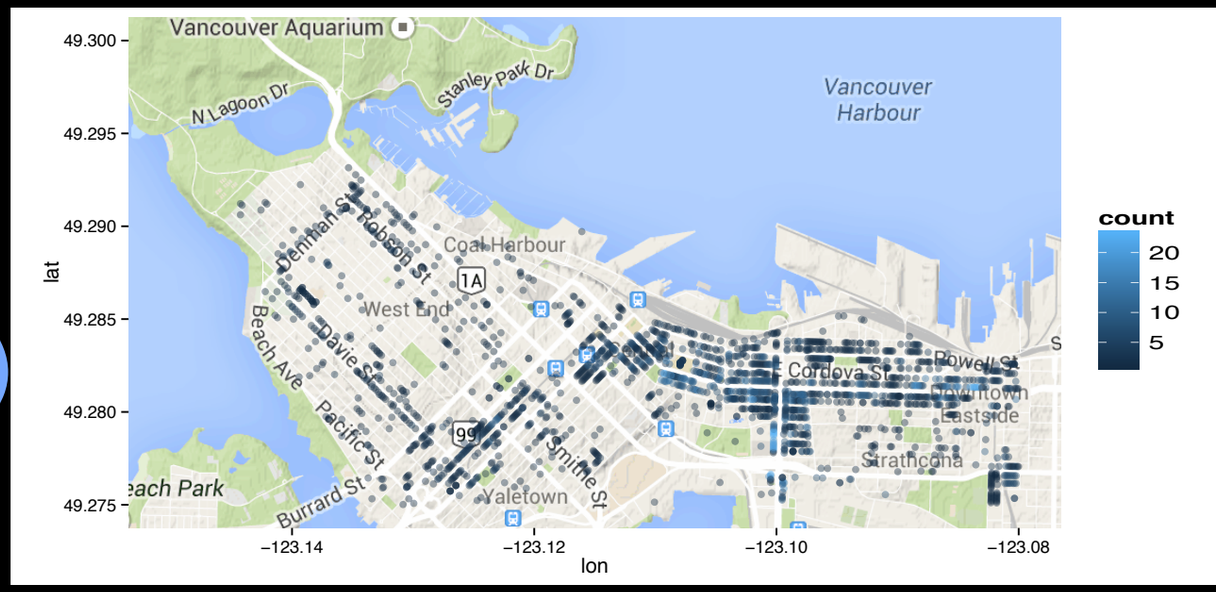


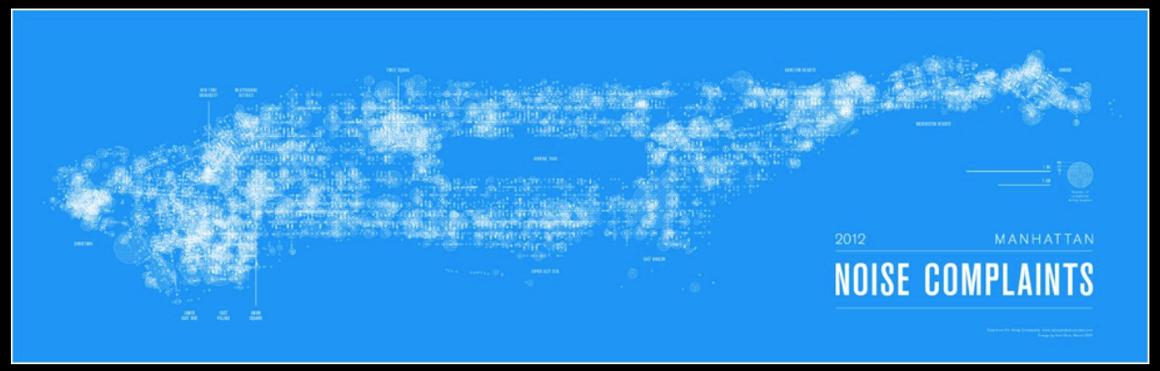


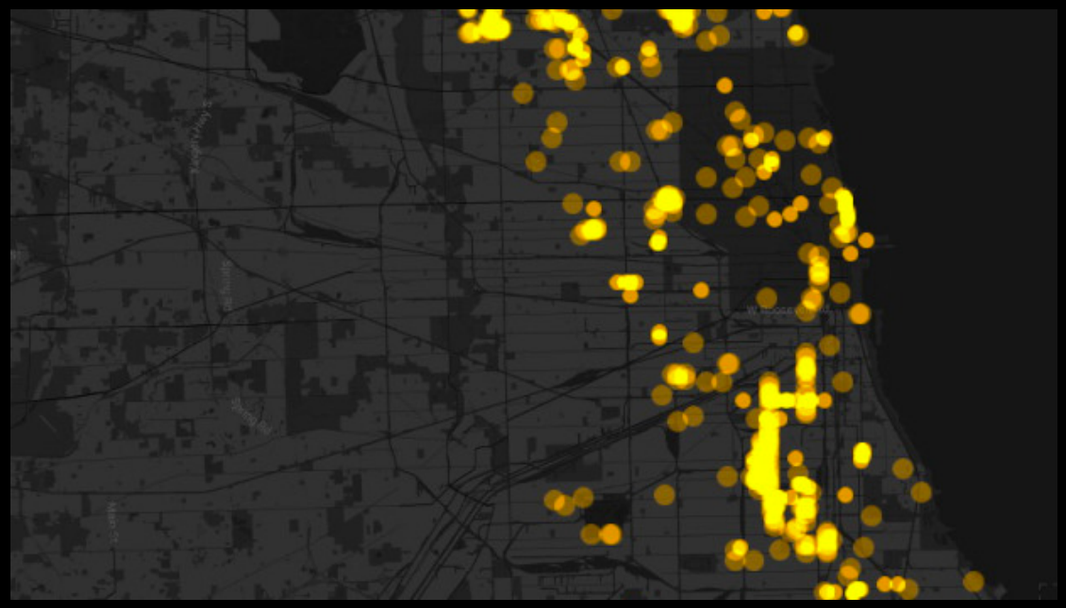
Urban Spatial Data

- Often possesses rich and non-standard structure

- Non-Euclidean
- Non-Continuous (Atomic)
- Inhomogeneous
- Clustered

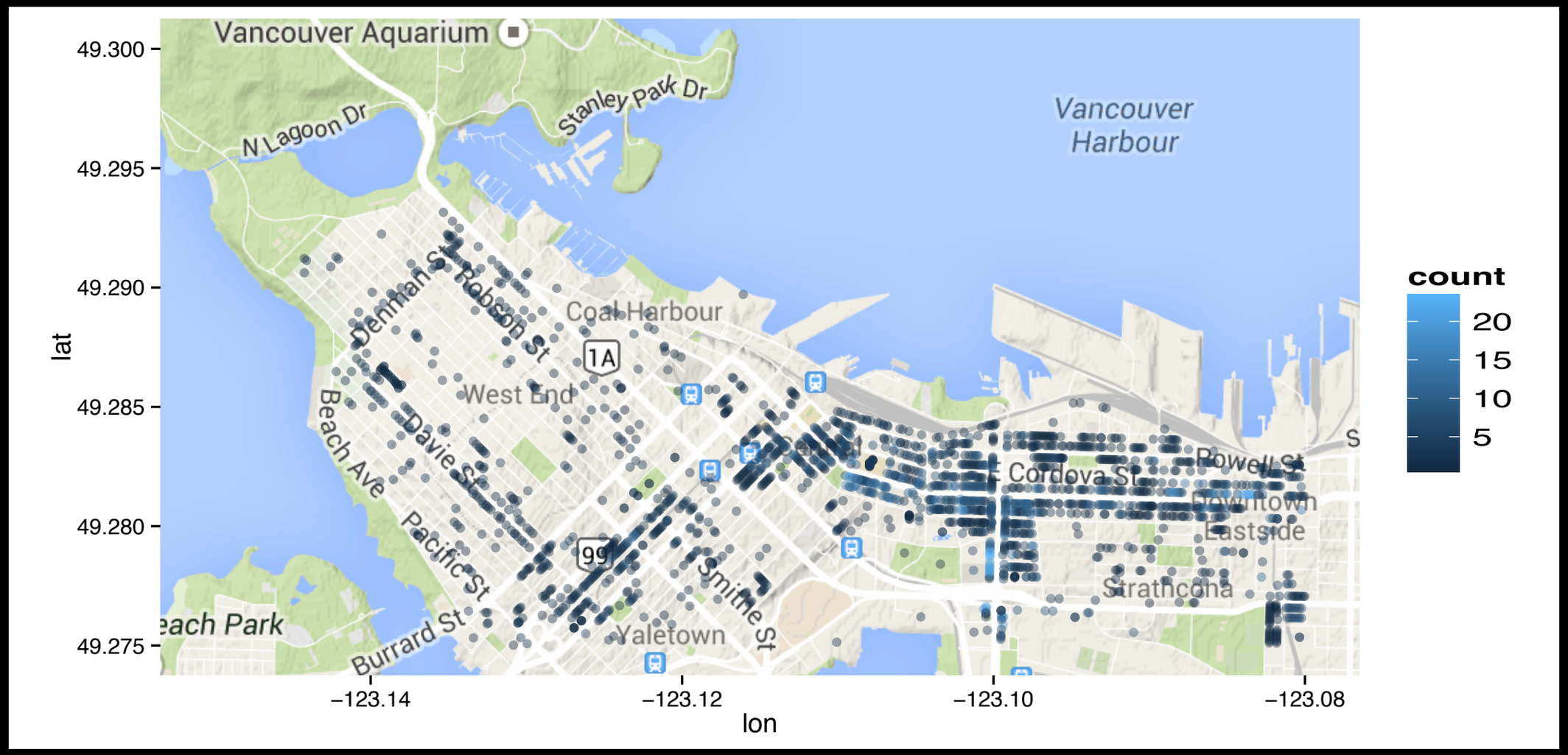






Motivation

Spatial Distribution of Graffiti in Vancouver



7675 pieces of graffiti occurring at 1982 unique locations

Atomic Data

Multiple observations with identical locations

Can arise:

Naturally

Multiple Observations

NeilRules!

Multiple Observations

Multiple Observations

Multiple Observations

3 Tags One Location = Atomicity

Atomic Data

Multiple observations with identical locations

Can arise:

Naturally

During the data collection stages:

Atomic Data

Multiple observations with identical locations

Can arise:

Naturally

During the data collection stages:

- Forced by collection mechanism
- To simplify data collection
- To preserve anonymity

Atomic Data

Multiple observations with identical locations

Can arise:

Naturally

During the data collection stages:

- Forced by collection mechanism
- To simplify data collection
- To preserve anonymity

Other approaches available

Atomic Data

Multiple observations with identical locations

Can arise:

Naturally

Focus Today

During the data collection stages:

- Forced by collection mechanism
- To simplify data collection
- To preserve anonymity

Other approaches available

Atomic Data

Multiple observations with identical locations

Can arise:

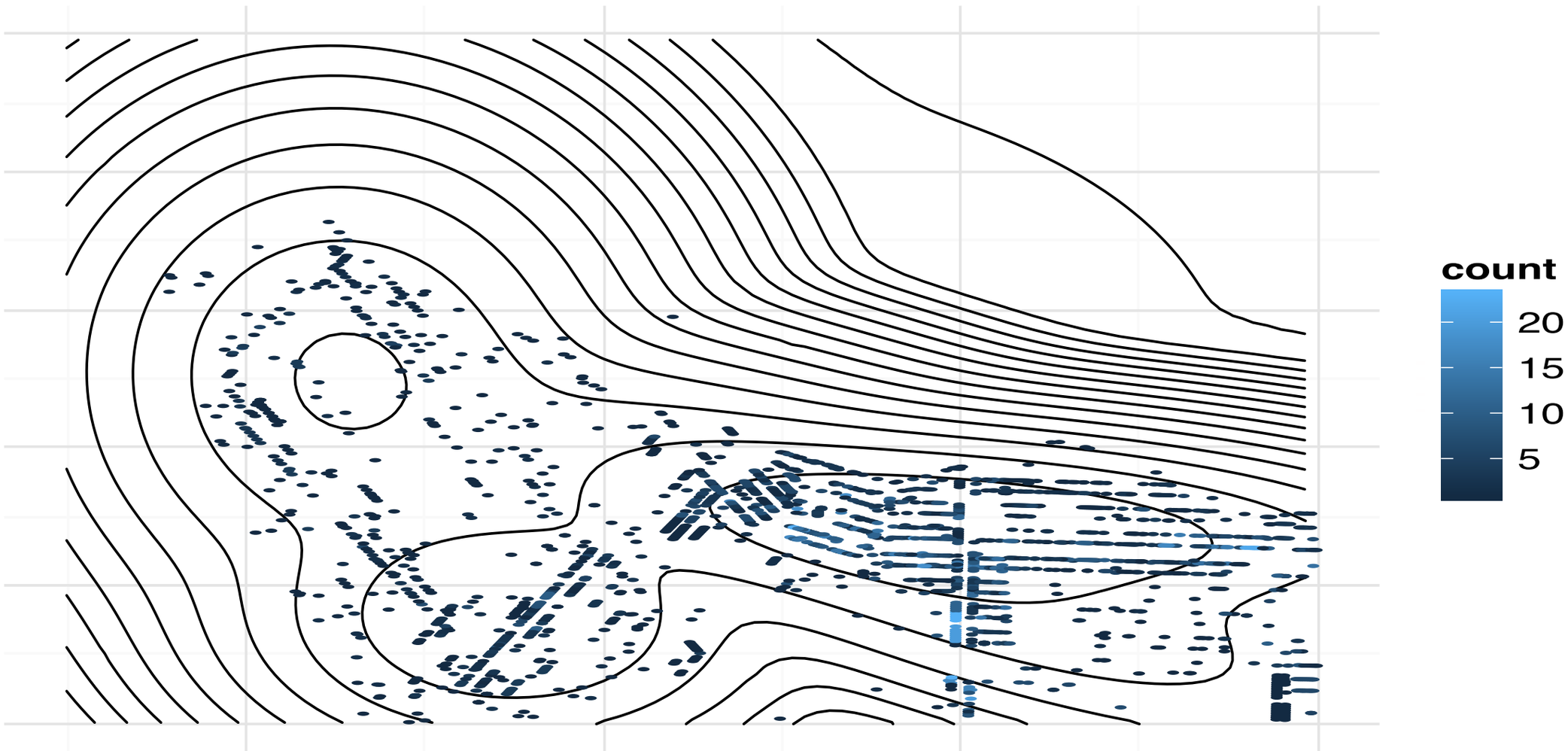
Naturally

Focus Today

- Graffiti
- Potholes
- Parking Tickets
- Police Speeding Checks

Desired Results

Vancouver Posterior Predictive



-Probability masses at graffitied locations

-Probability density at clean locations

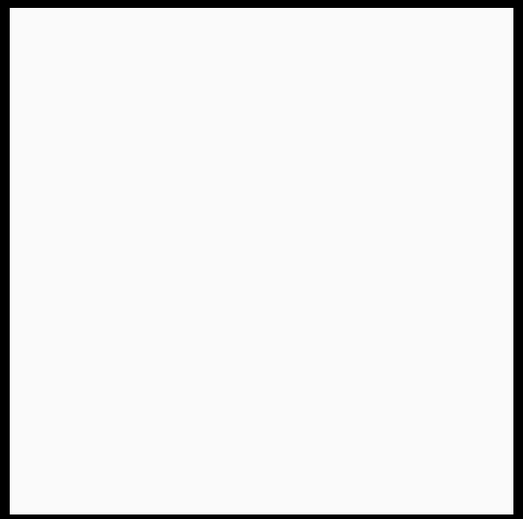
Spatial Modeling Using Point Processes

(Brief Review)

Spatial Modeling Using Point Processes

Spatial Poisson Process

$\Omega = [0, 1] \times [0, 1]$



Spatial Modeling Using Point Processes

Spatial Poisson Process

$\Omega = [0, 1] \times [0, 1]$

$\text{Observed Data} \sim \text{PP}(\mu)$

Observed Data

Spatial Modeling Using Point Processes

Spatial Poisson Process

$\Omega = [0,1] \times [0,1]$

$\text{Observed Data} \sim \text{PP}(\mu)$

$\Rightarrow |\text{Observed Data}| \sim \text{Poi}(\mu(\Omega))$

Observed Data

Spatial Modeling Using Point Processes

Spatial Poisson Process

$\Omega = [0,1] \times [0,1]$

$\text{Observed Data} \sim \text{PP}(\mu)$

$\Rightarrow |\text{Observed Data}| \sim \text{Poi}(\mu(\Omega))$

$\text{Each point} \overset{iid}{\sim} \bar{\mu}$

Observed Data

Spatial Modeling Using Point Processes

Spatial Poisson Process

$\Omega = [0,1] \times [0,1]$

$\text{Observed Data} \sim \text{PP}(\mu)$

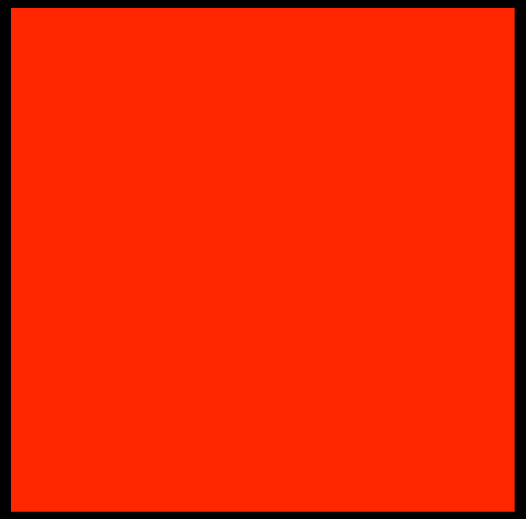
$\Rightarrow |\text{Observed Data}| \sim \text{Poi}(\mu(\Omega))$

$\text{Each point} \overset{iid}{\sim} \bar{\mu}$

Observed Data

Rate Measure: μ

Constant



Spatial Modeling Using Point Processes

Spatial Poisson Process

$\Omega = [0,1] \times [0,1]$

$\text{Observed Data} \sim \text{PP}(\mu)$

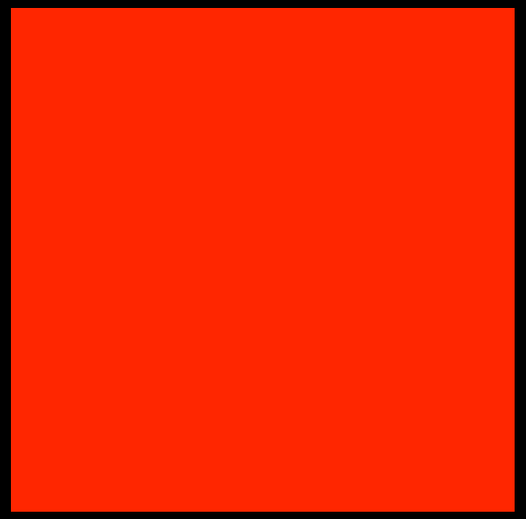
$\Rightarrow |\text{Observed Data}| \sim \text{Poi}(\mu(\Omega))$

$\text{Each point} \overset{iid}{\sim} \bar{\mu}$

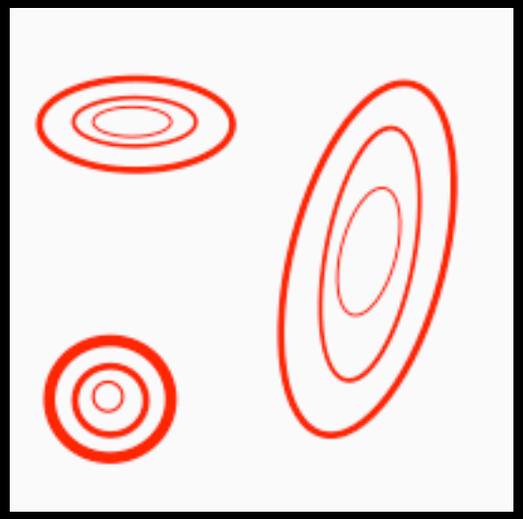
Observed Data

Rate Measure: μ

Constant



Random (e.g. Bayesian Nonparametric)



Spatial Poisson Process for Atomic Data

Colours signify multiple observations

Each point $\overset{iid}{\sim} \bar{\mu}$

$\Rightarrow$ μ possesses discrete points (atoms).

Spatial Poisson Process for Atomic Data

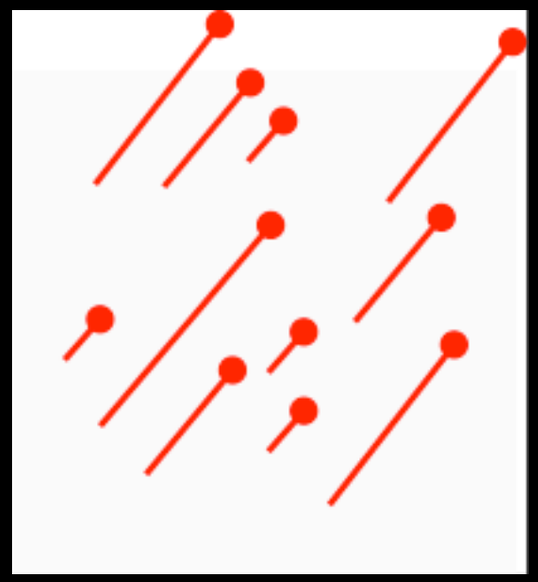
Colours signify multiple observations

Each point $\overset{iid}{\sim} \bar{\mu}$

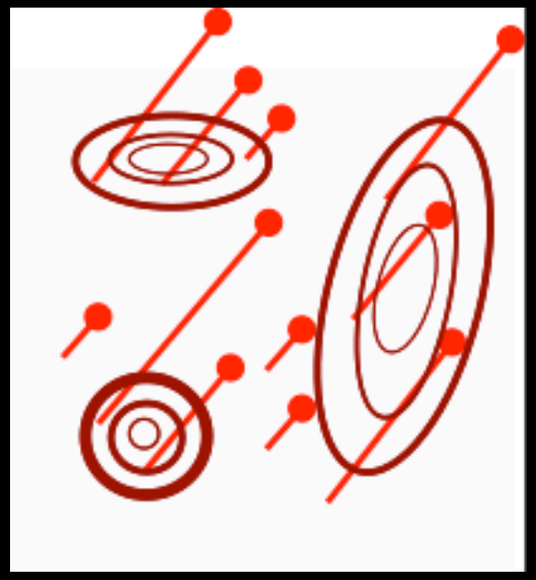
$\Rightarrow$ μ possesses discrete points (atoms).

Rate Measure: μ

Purely Atomic



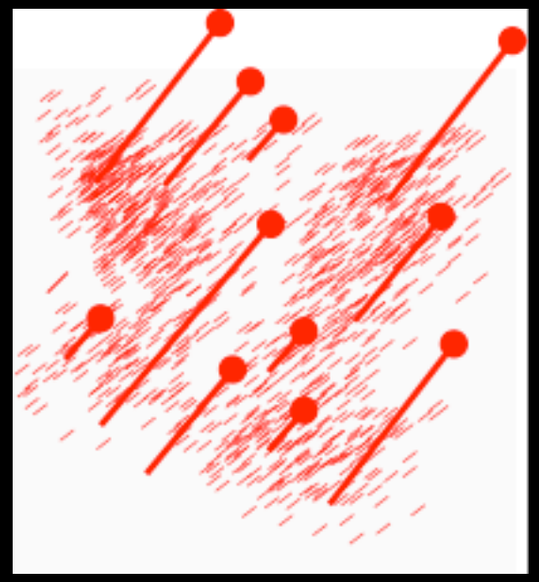
Mixed



Gamma Process

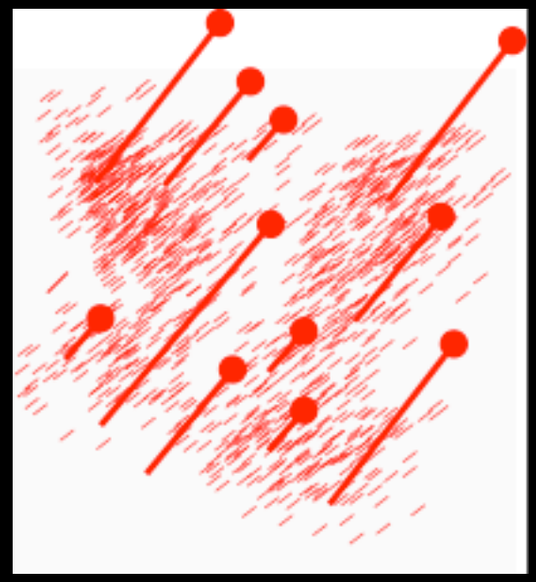
- A Poisson Process
- Augmented Space $\Omega \times \mathbb{R}^+$

(weights drawn from positive reals)



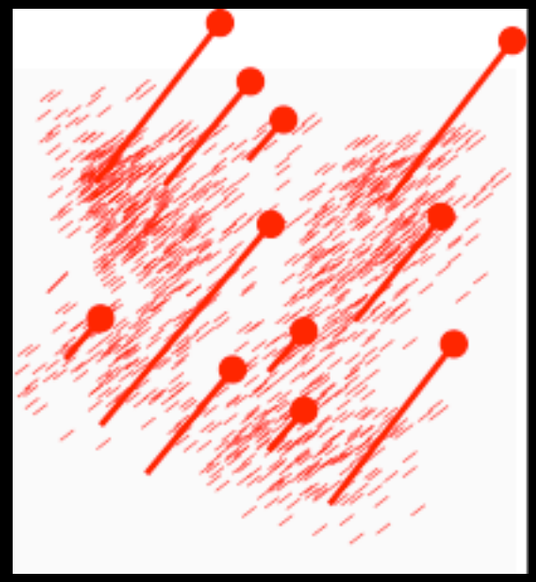
Gamma Process

- A Poisson Process
- Augmented Space $\Omega \times \mathbb{R}^+$
- Infinite number of atoms (with probability one)



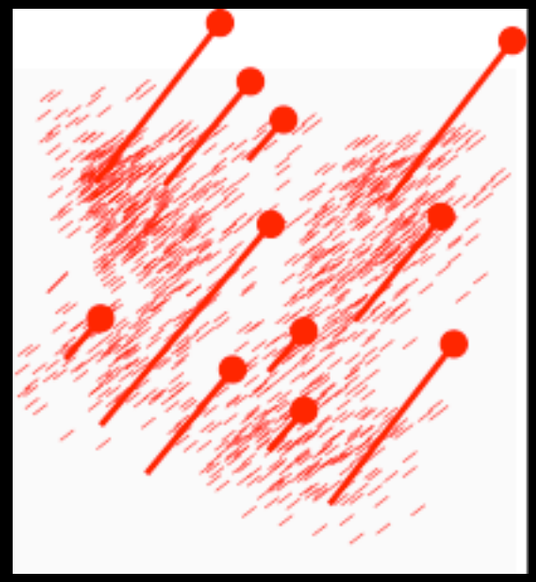
- Expected sum of weights is finite.

Gamma Process



- A Poisson Process
- Augmented Space $\Omega \times \mathbb{R}^+$
- Infinite number of atoms (with probability one)
- Expected sum of weights is finite.
- Three parameters,
  - G_0 controls locations of atoms.
  - c, α_0 parameterize the distribution of atom heights.

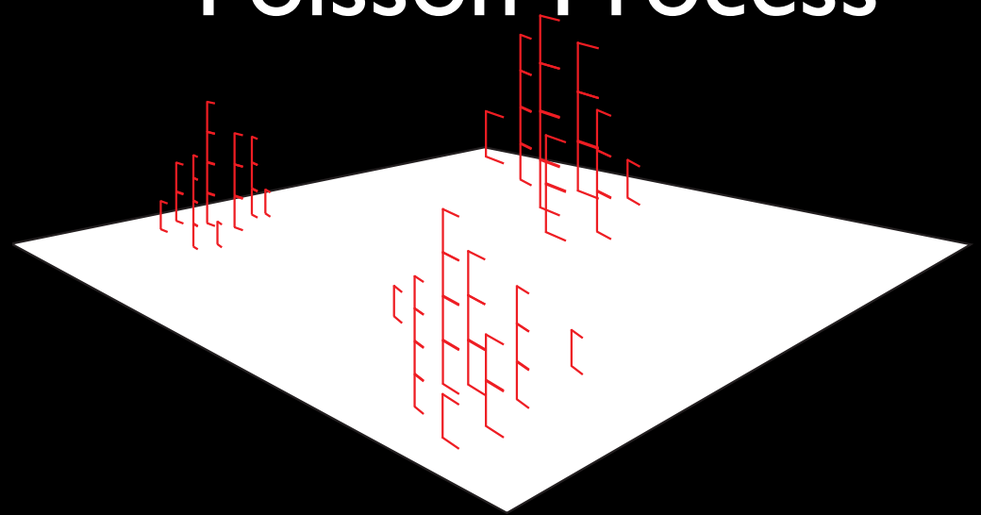
Gamma Process



- A Poisson Process
- Augmented Space $\Omega \times \mathbb{R}^+$
- Infinite number of atoms (with probability one)
- Expected sum of weights is finite.
- Three parameters,
  - G_0 controls locations of atoms.
  - c, α_0 parameterize the distribution of atom heights.
- Normalizes to Dirichlet Process

Atomic Spatial Processes

Poisson Process

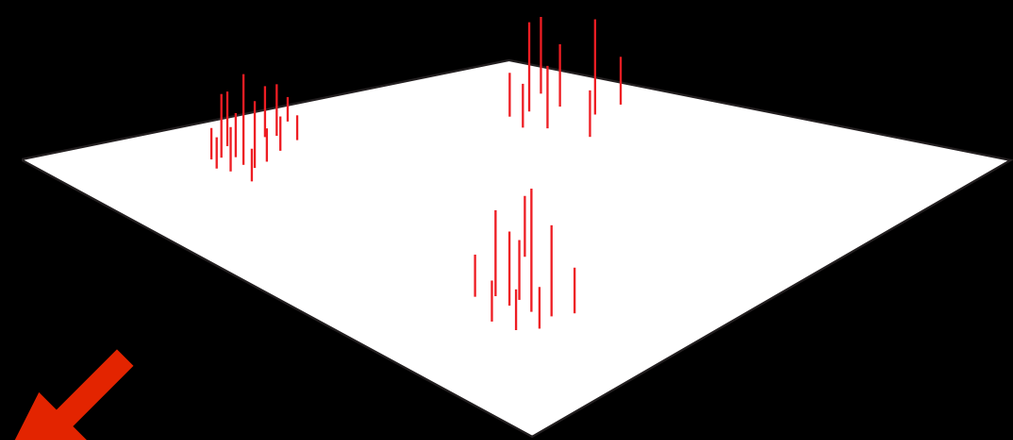


$$\text{Observed Data}|\mu \sim PP(\mu)$$

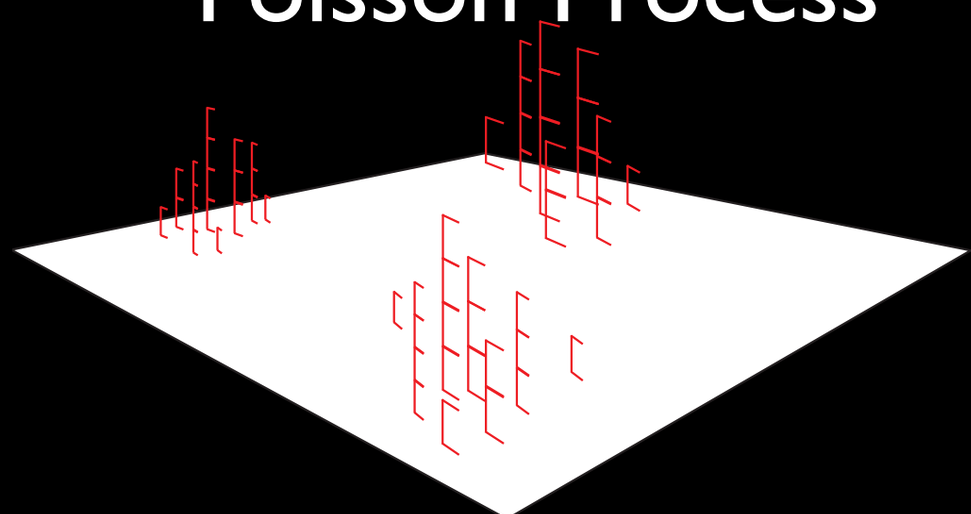
Atomic Spatial Processes

Gamma Process

$$\mu | G_0 \sim \mathrm{GaP}(c^{-1}, \alpha_0, G_0)$$



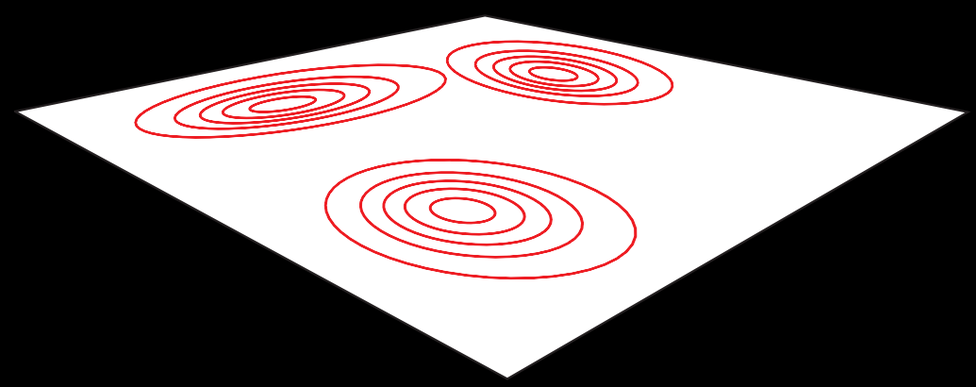
Poisson Process



$$\text{Observed Data} | \mu \sim PP(\mu)$$

Atomic Spatial Processes

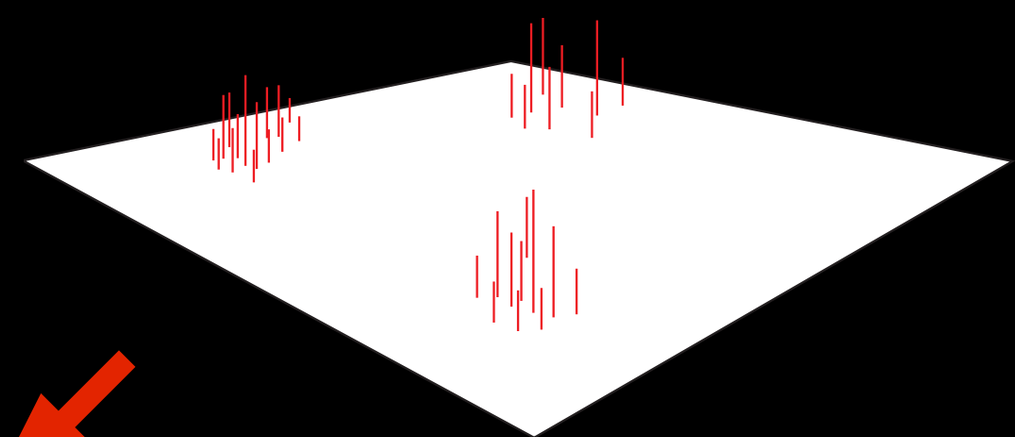
Dirichlet Process Mixture



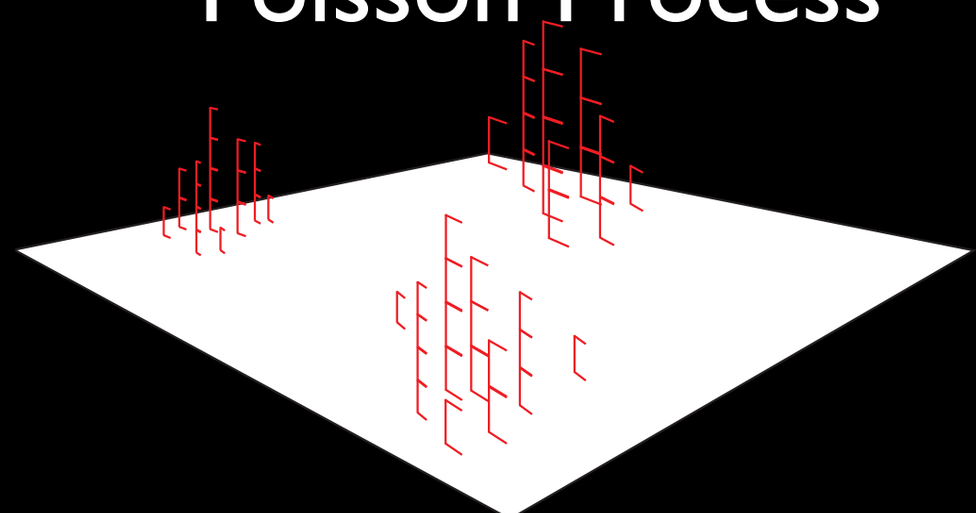
$$G_0 \sim \mathrm{DPM}(\alpha_0, H_0)$$

Gamma Process

$$\mu | G_0 \sim \mathrm{GaP}(c^{-1}, \alpha_0, G_0)$$



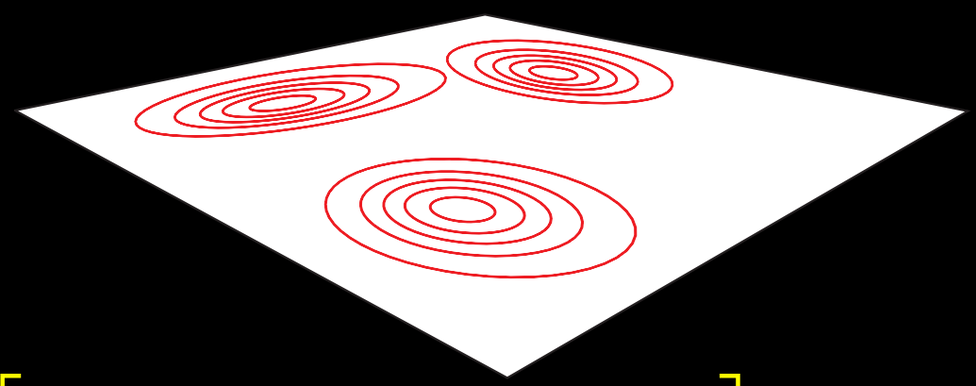
Poisson Process



$$\text{Observed Data} | \mu \sim PP(\mu)$$

Joint Distribution

Dirichlet Process Mixture

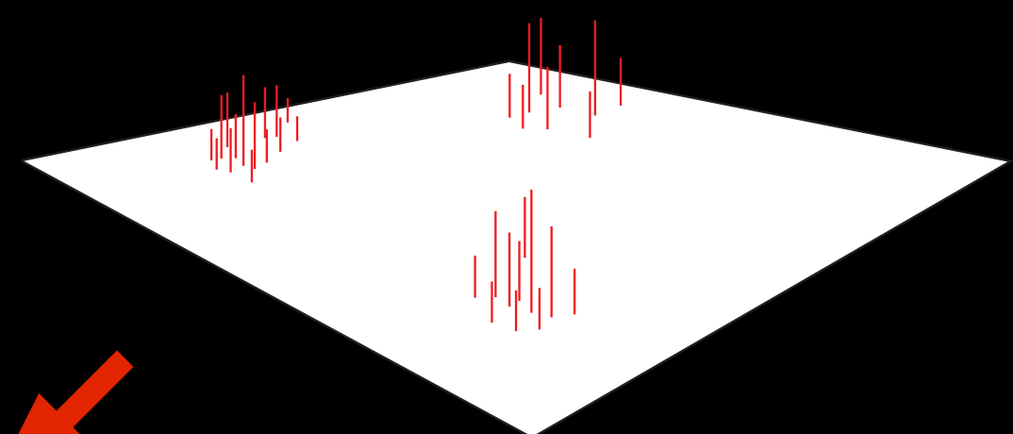


$$\left[\prod_{k=1}^{\infty} \mathrm{Beta}(\mathrm{d}\beta_k^{\pi}|1,\alpha_0^{\pi}) H_0(\mathrm{d}\theta_k)\right]$$

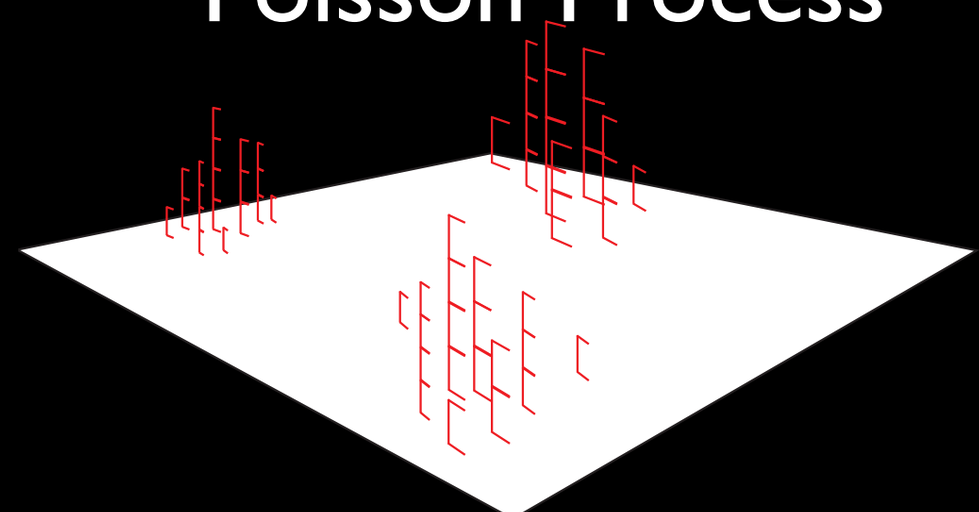
$$\left[\prod_{j=1}^{\infty} \mathrm{Mult}(\mathrm{d}z_j|\pi) L(\mathrm{d}a_j|\theta_{z_j})\right] \times$$

$$\mathrm{Gamma}(\mathrm{d}\gamma|\alpha_0, c^{-1}) \left[\prod_{j=1}^{\infty} \mathrm{Beta}(\mathrm{d}\beta_j^{p}|1,\alpha_0)\right]$$

Gamma Process



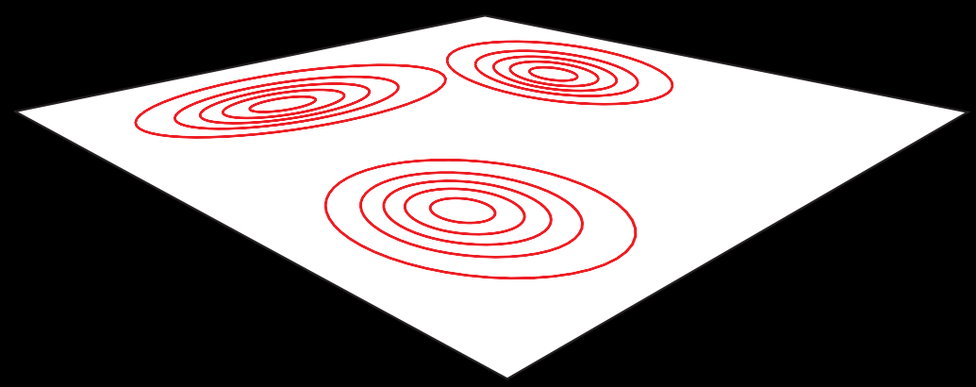
Poisson Process



$$\mathrm{Pois}(n|\gamma) \left[\prod_{i=1}^{n} p_{j(a,x_i)} \delta_{x_i \in a}(\mathrm{d}x)\right]$$

Marginalization

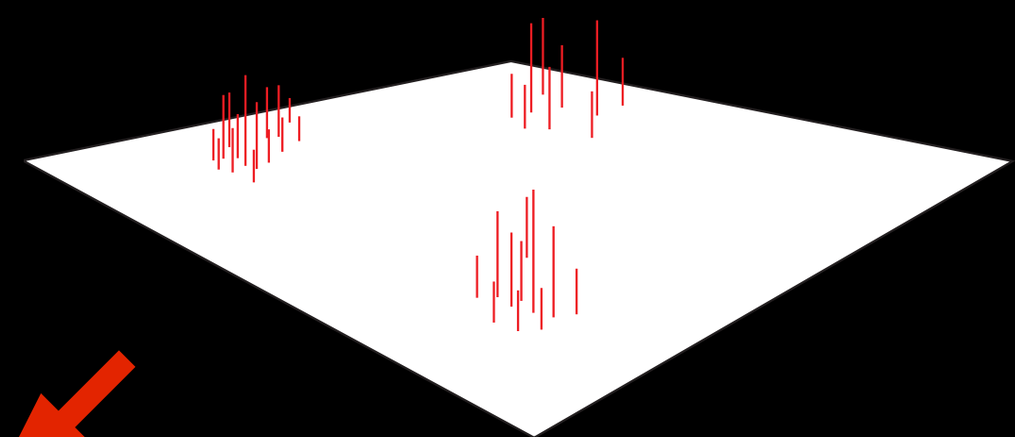
Dirichlet Process Mixture



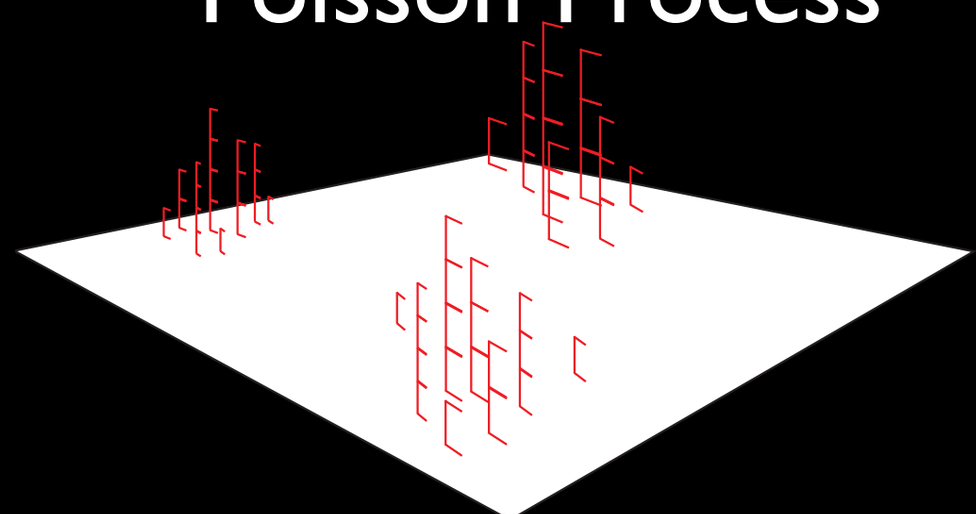
$$\mathrm{CRP}(\rho(\mathrm{d}z)|\alpha_0^\pi)$$

Gamma Process

$$\prod_{B\in\rho(z)} m(\mathrm{d}a_B)$$



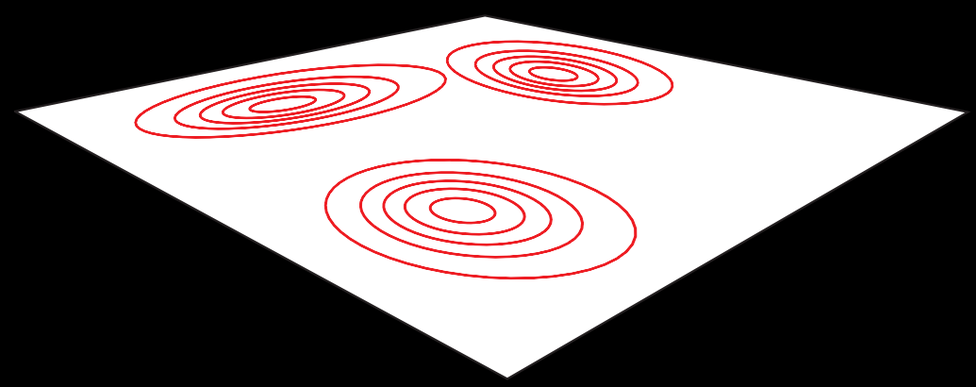
Poisson Process



$$\mathrm{NegBin}\left(n\middle|\frac{1}{c+1},\alpha_0\right)\mathrm{CRP}\left(\rho(\mathrm{d}x)|\alpha_0, N=n\right)$$

Marginalization

Dirichlet Process Mixture

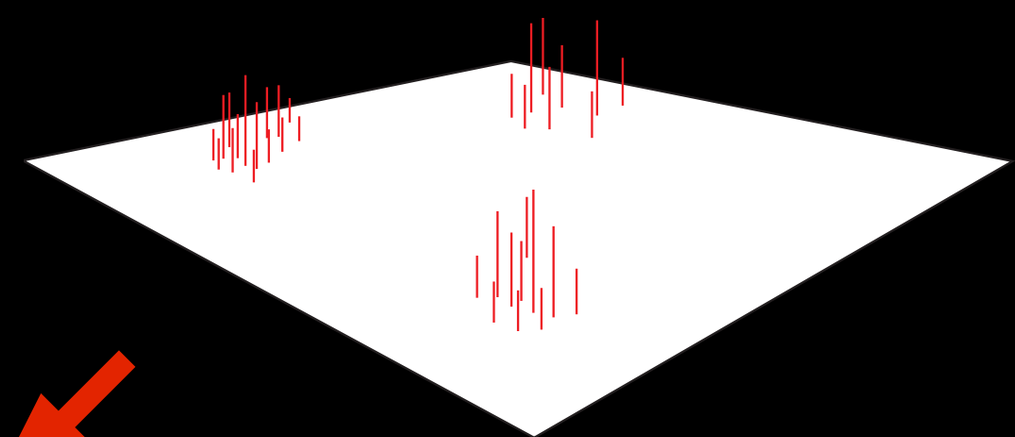


$$\mathrm{CRP}\boxed{(\rho(\mathrm{d}z)|\alpha_0^\pi)}$$

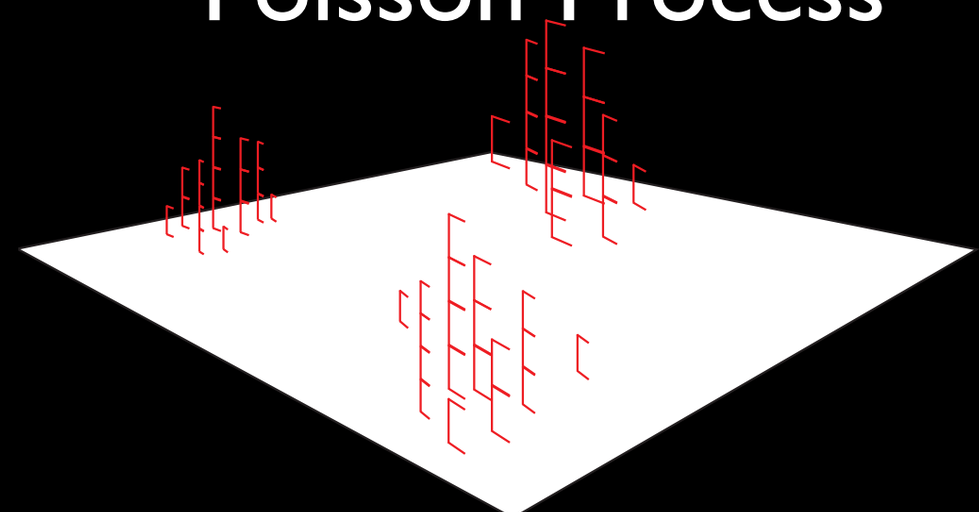
Only Latent Variable

Gamma Process

$$\prod_{B\in\rho(z)} m(\mathrm{d}a_B)$$



Poisson Process



$$\mathrm{NegBin}\left(n\middle|\frac{1}{c+1},\alpha_0\right)\mathrm{CRP}\left(\rho(\mathrm{d}x)|\alpha_0, N=n\right)$$

Data Analysis Results

Vancouver Posterior Predictive

- Non-informative priors on hyperparameters
- 20,000 MCMC iterations
- Each sweep involved:
  - Gibbs step for each atom's cluster membership

Data Analysis Results

Vancouver Posterior Predictive

- Non-informative priors on hyperparameters
- 20,000 MCMC iterations
- Each sweep involved:
  - Gibbs step for each atom's cluster membership

Data Analysis Results

Vancouver Posterior Predictive

- Non-informative priors on hyperparameters
- 20,000 MCMC iterations
- Each sweep involved:
  - Gibbs step for each atom's cluster membership

Data Analysis Results

Vancouver Posterior Predictive

- Non-informative priors on hyperparameters
- 20,000 MCMC iterations
- Each sweep involved:
  - Gibbs step for each atom's cluster membership

Data Analysis Results

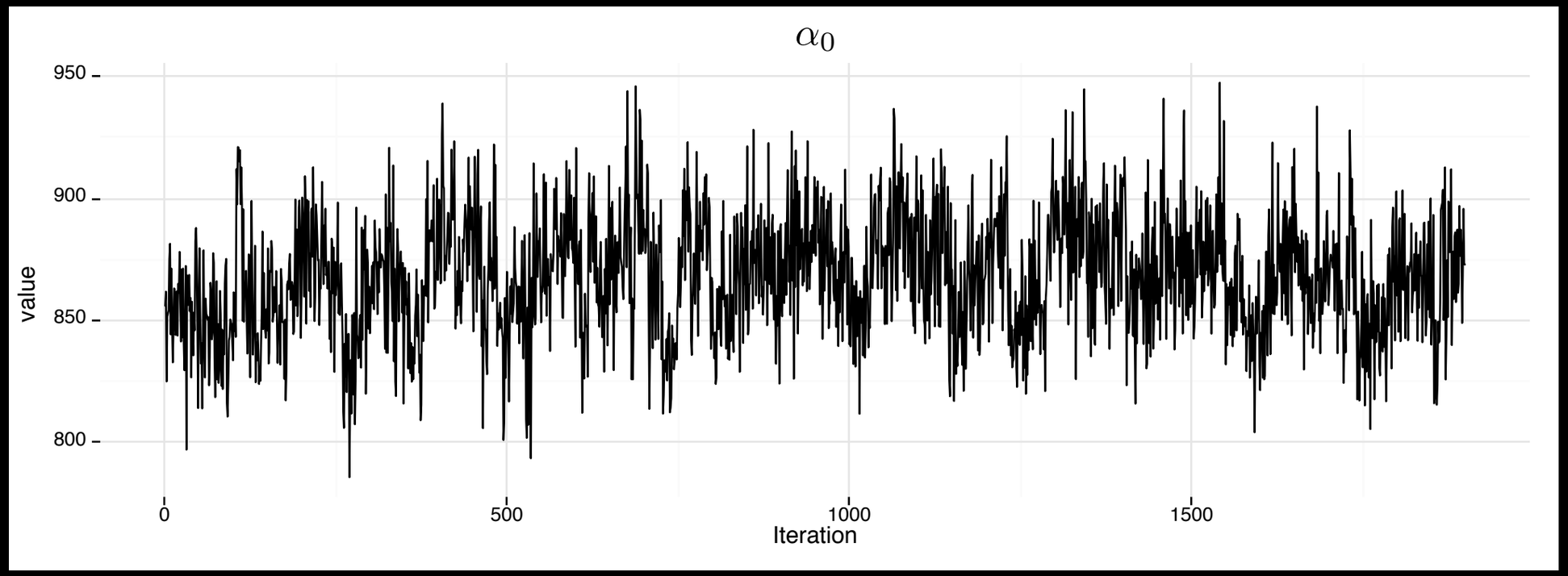
Vancouver Posterior Predictive

- Non-informative priors on hyperparameters
- 20,000 MCMC iterations
- Each sweep involved:
  - Gibbs step for each atom's cluster membership

Data Analysis Results

Vancouver Posterior Predictive

- Non-informative priors on hyperparameters
- 20,000 MCMC iterations
- Each sweep involved:
  - Gibbs step for each atom's cluster membership
  - Metropolis Hastings proposal for hyperparameters



Tractable Inference

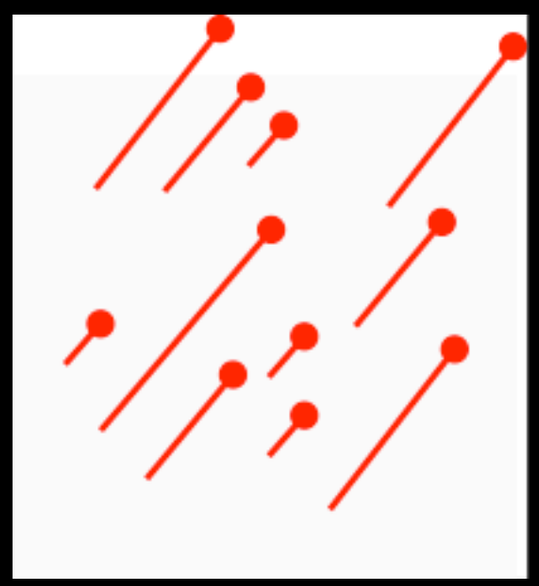
Predictive Posterior Distribution

$$\bar{\mu}'(\mathrm{d}y) = \frac{1}{N+\alpha_0}\sum_{i=1}^{N}\delta_{X_i}(\mathrm{d}y) + \frac{\alpha_0}{N+\alpha_0}\frac{1}{T+\alpha_0^{\pi}}\left(\sum_{i=1}^{K}|B_i|m(\mathrm{d}y|a_{B_i}) + \alpha_0^{\pi}m(\mathrm{d}y)\right).$$

Tractable Inference

Predictive Posterior Distribution

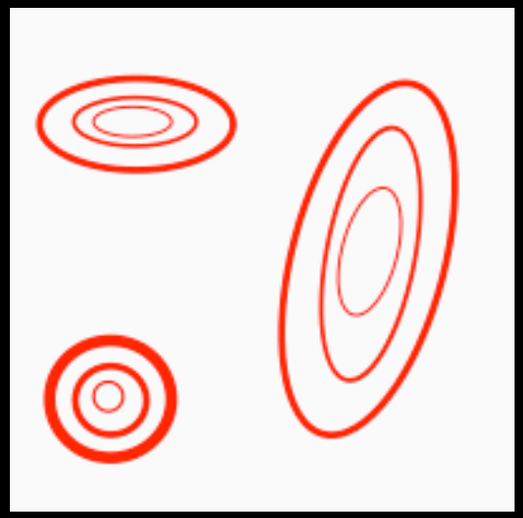
$$\bar{\mu}'(\mathrm{d}y) = \frac{1}{N+\alpha_0}\sum_{i=1}^{N}\delta_{X_i}(\mathrm{d}y) + \frac{\alpha_0}{N+\alpha_0}\frac{1}{T+\alpha_0^{\pi}}\left(\sum_{i=1}^{K}|B_i|m(\mathrm{d}y|a_{B_i}) + \alpha_0^{\pi}m(\mathrm{d}y)\right).$$



Tractable Inference

Predictive Posterior Distribution

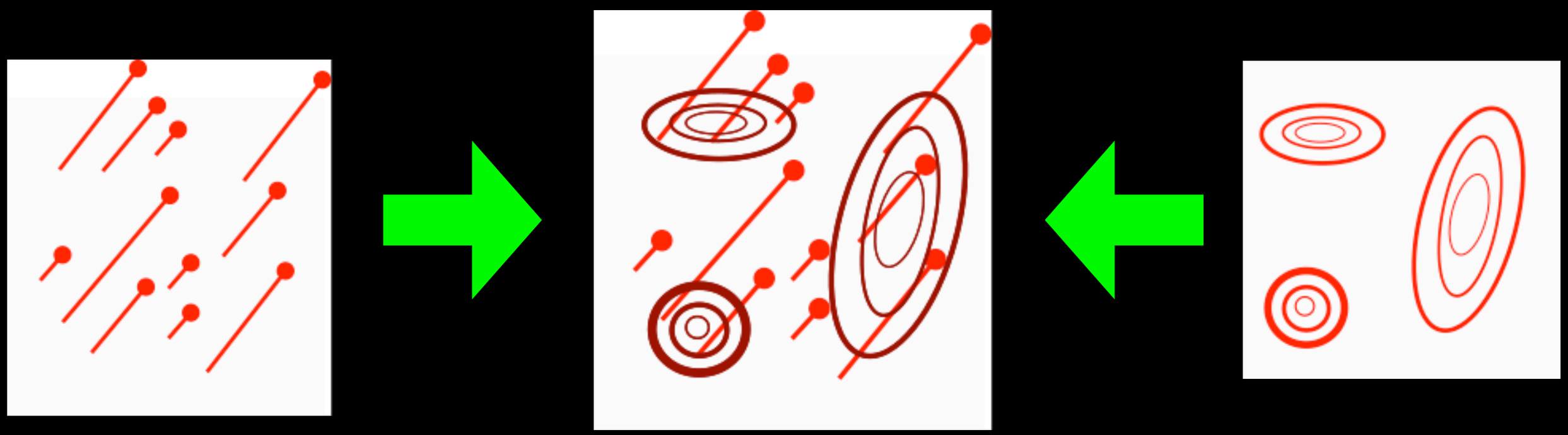
$$\bar{\mu}'(\mathrm{d}y) = \frac{1}{N+\alpha_0}\sum_{i=1}^{N}\delta_{X_i}(\mathrm{d}y) + \frac{\alpha_0}{N+\alpha_0}\frac{1}{T+\alpha_0^{\pi}}\left(\sum_{i=1}^{K}|B_i|m(\mathrm{d}y|a_{B_i}) + \alpha_0^{\pi}m(\mathrm{d}y)\right).$$



Tractable Inference

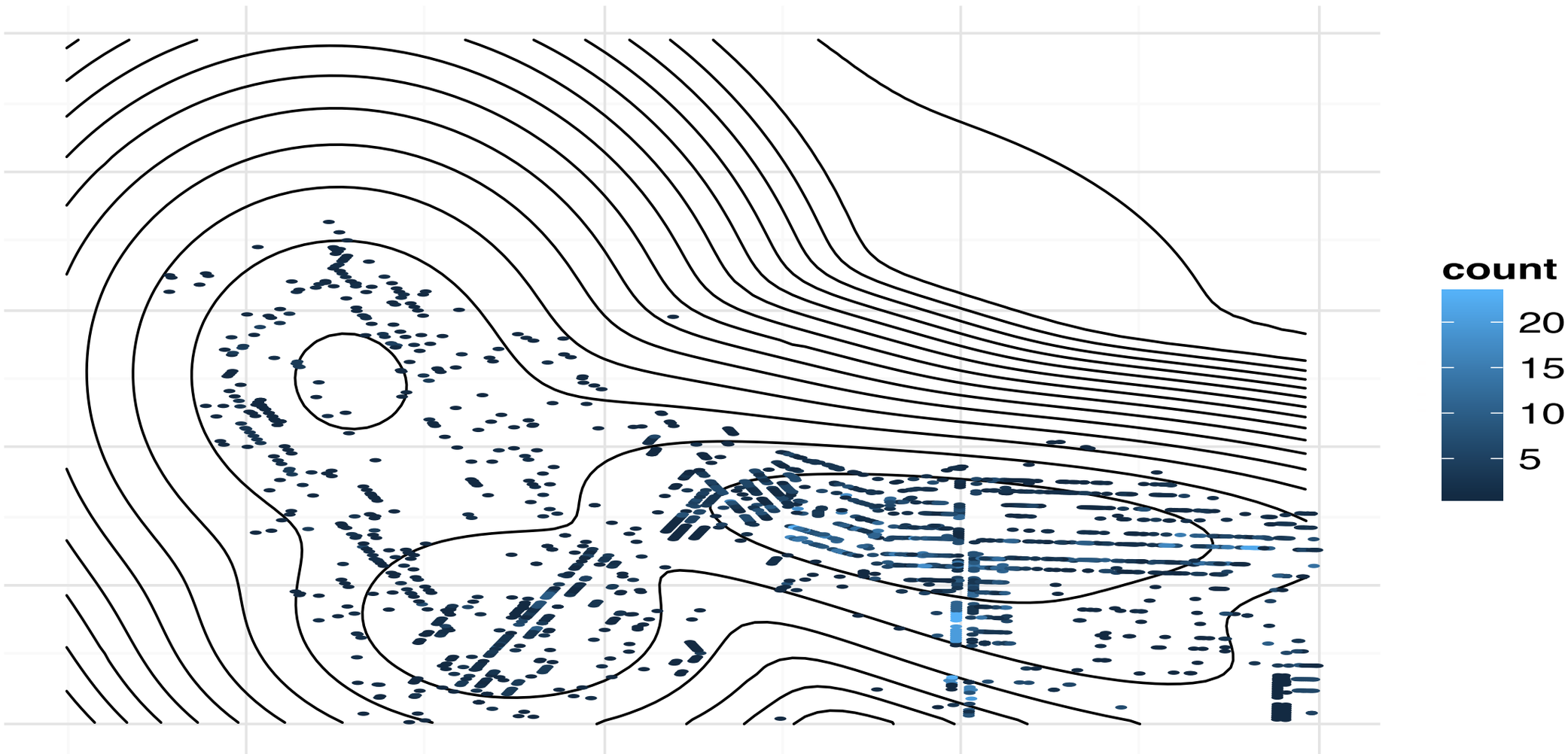
Predictive Posterior Expression

$$\bar{\mu}'(\mathrm{d}y) = \boxed{\frac{1}{N+\alpha_0}\sum_{i=1}^{N}\delta_{X_i}(\mathrm{d}y) + \frac{\alpha_0}{N+\alpha_0}\frac{1}{T+\alpha_0^{\pi}}\left(\sum_{i=1}^{K}|B_i|m(\mathrm{d}y|a_{B_i}) + \alpha_0^{\pi}m(\mathrm{d}y)\right)}$$



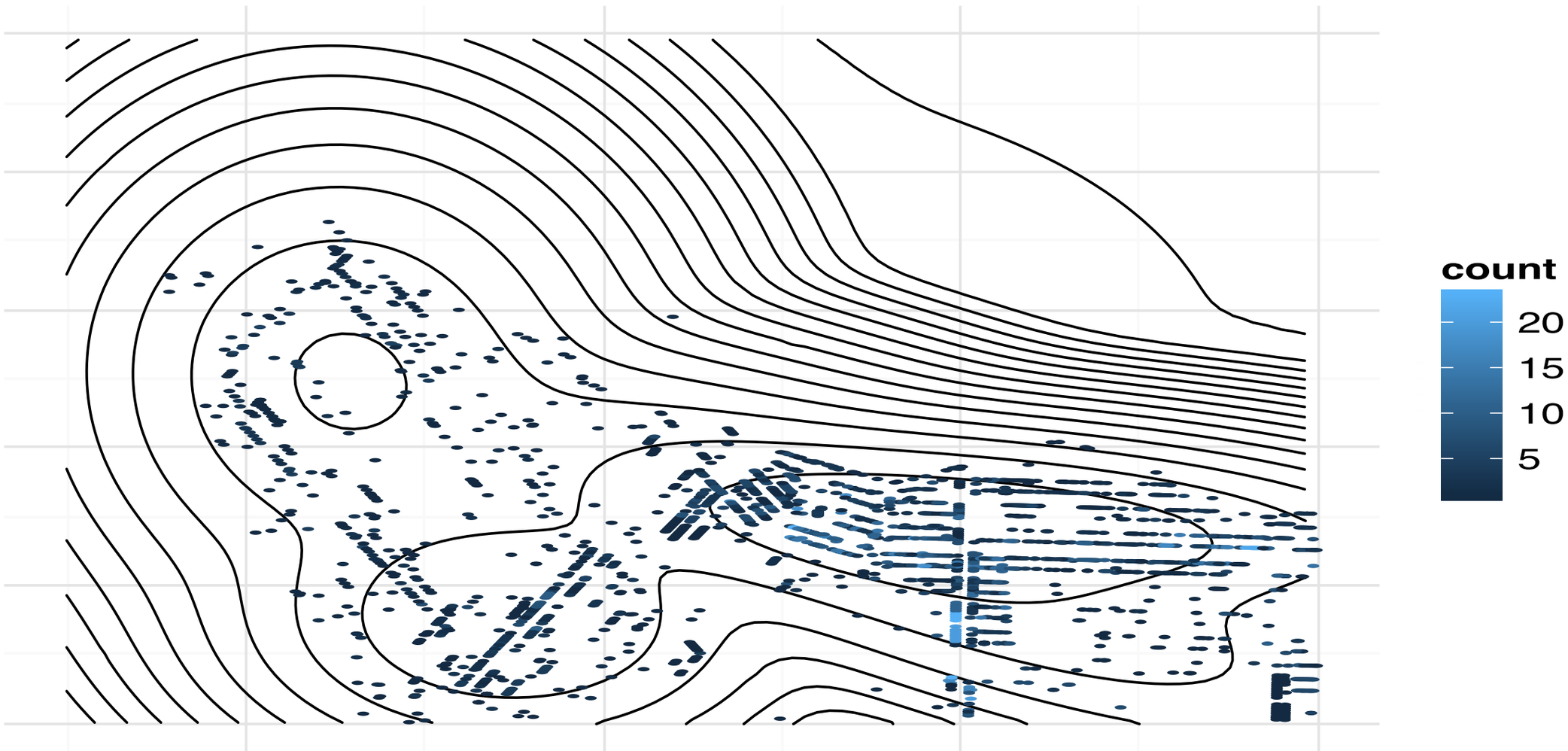
Data Analysis Results

Vancouver Posterior Predictive



Data Analysis Results

Vancouver Posterior Predictive

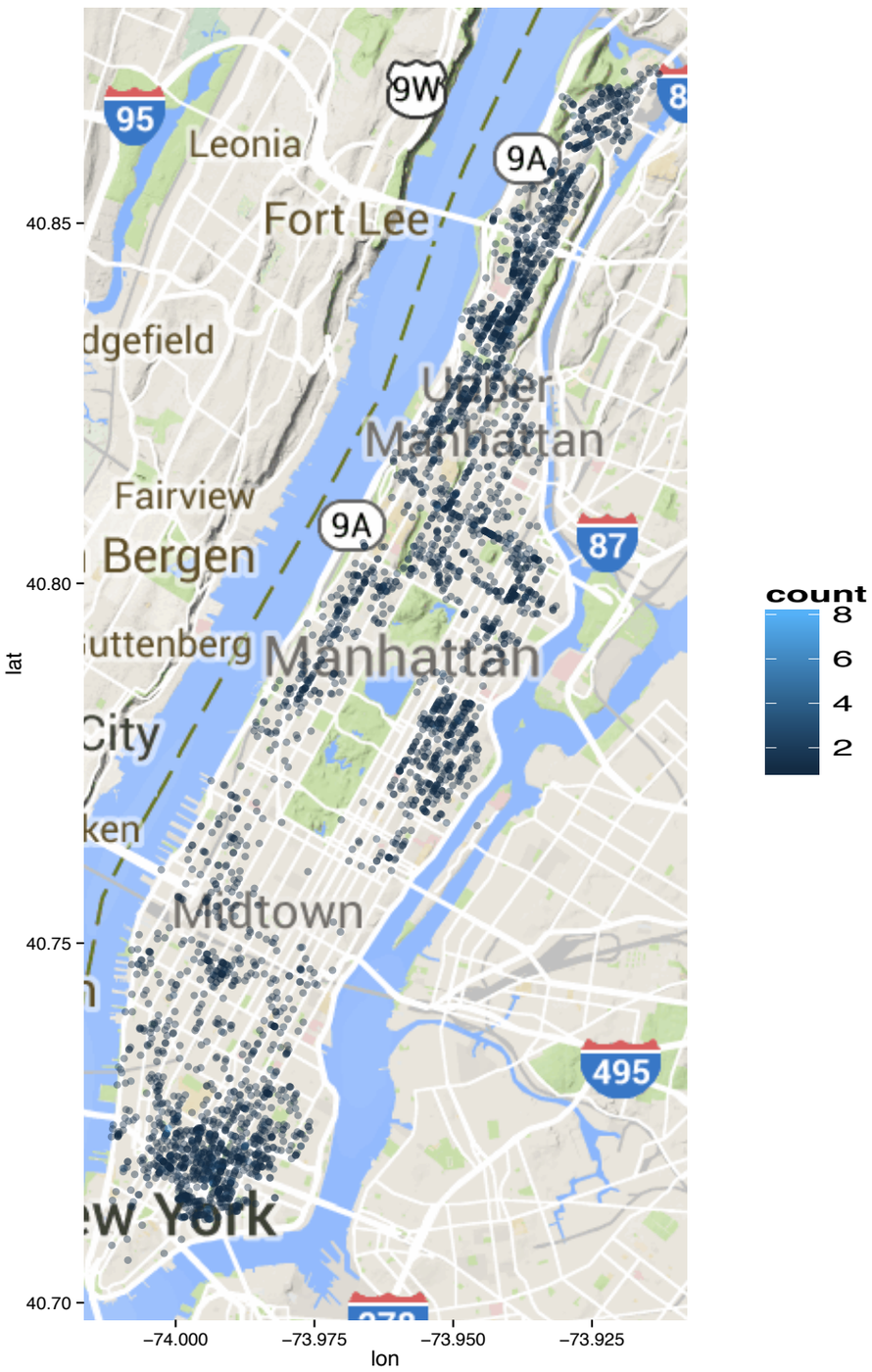


$$\mathbb{P}(\text{new location}) = 0.102$$

Mode number of clusters 4

Another Example

Spatial Distribution of Graffiti in Manhattan

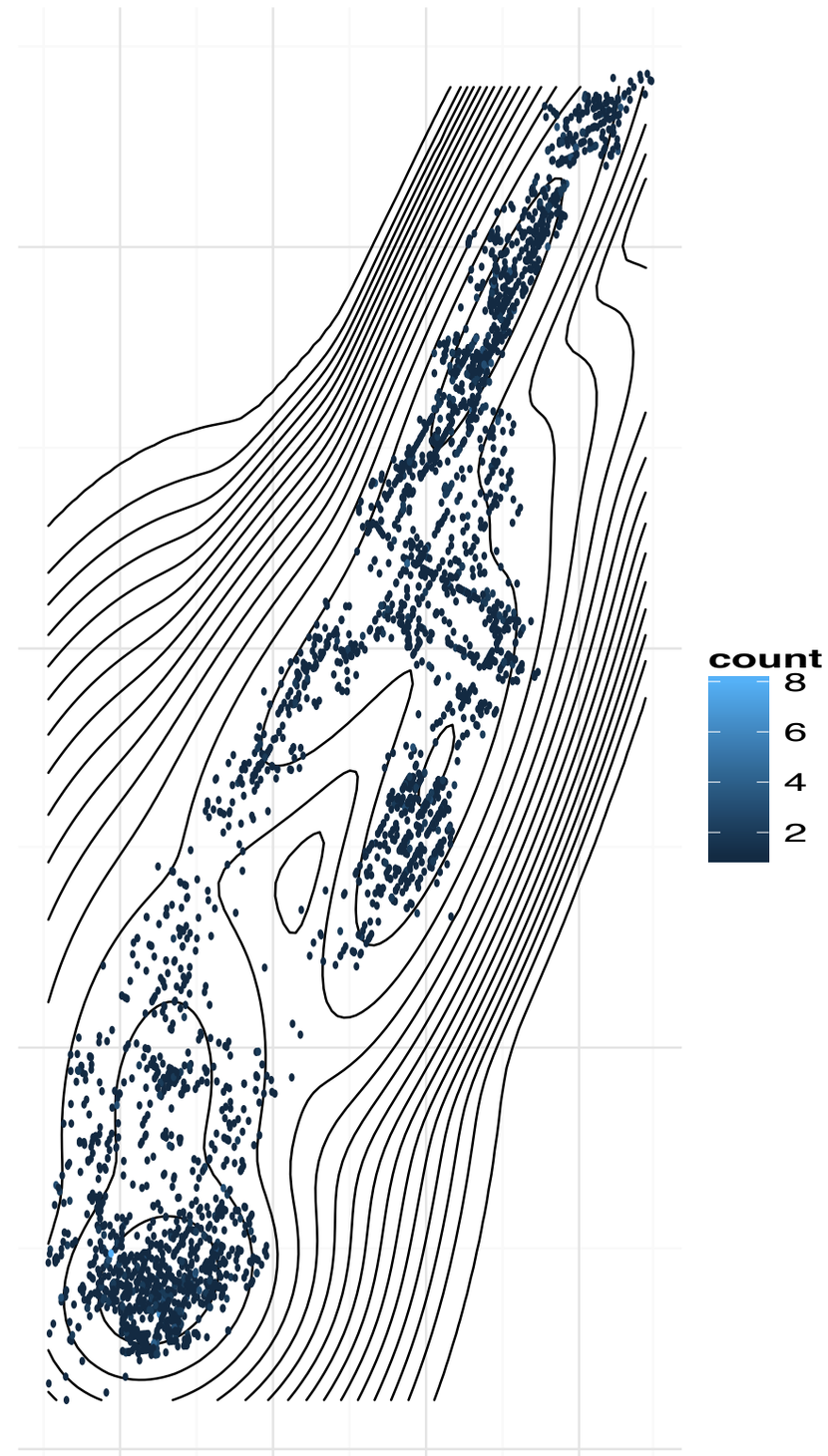


Another Example

Spatial Distribution of Graffiti in Manhattan

$$\mathbb{P}(\text{new location}) = 0.7501$$

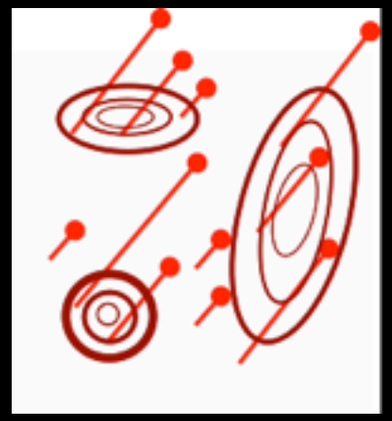
Mode number of clusters 5



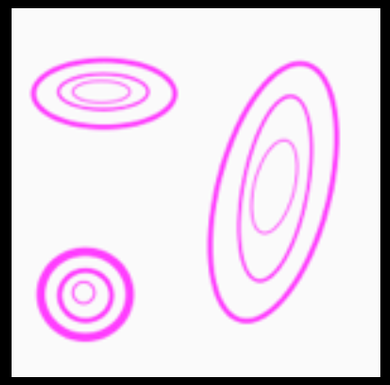
Model Comparisons

Competitors Considered

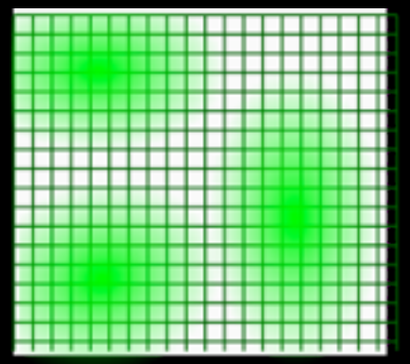
- Atomic Spatial Process



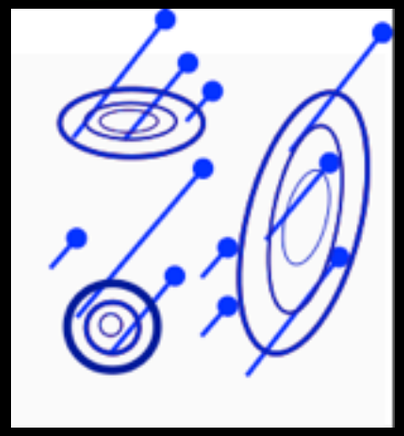
- Dirichlet Process Mixture Model (DPM)



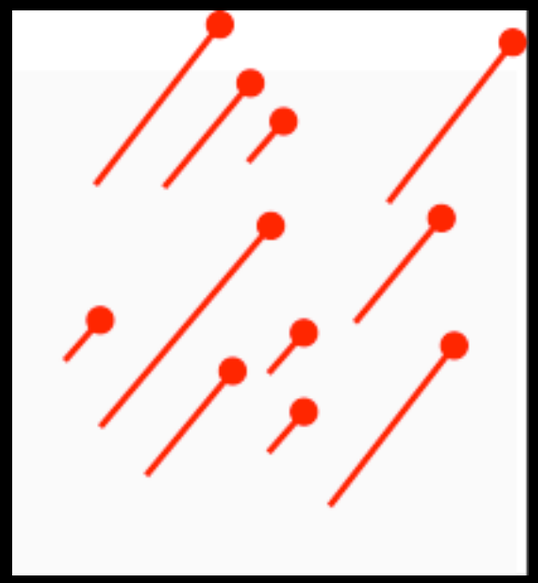
- Empirical Distribution Grid Approach



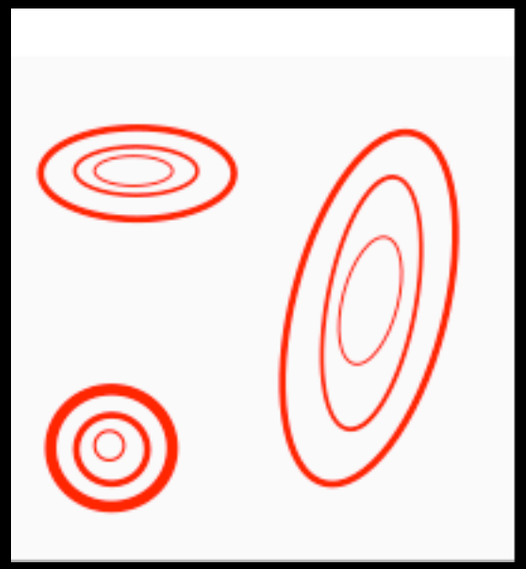
- DPM mixed with Empirical Distribution



Mass Functions versus Densities

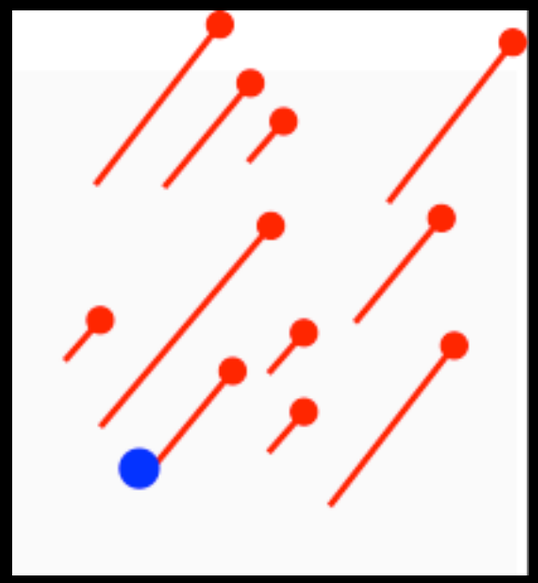


Mass



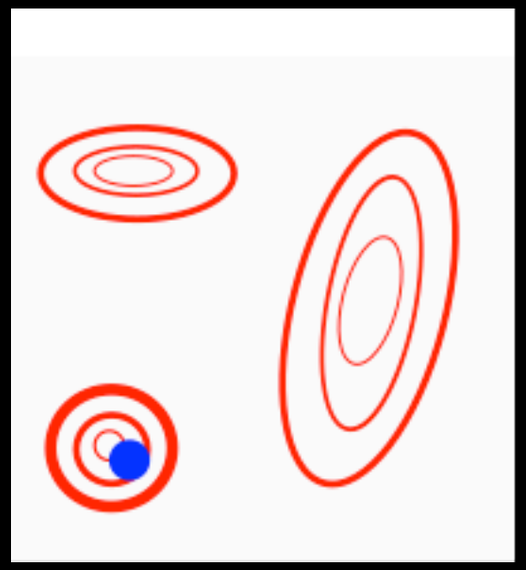
Density

Mass Functions versus Densities



Mass

Unitless (Probabilities)

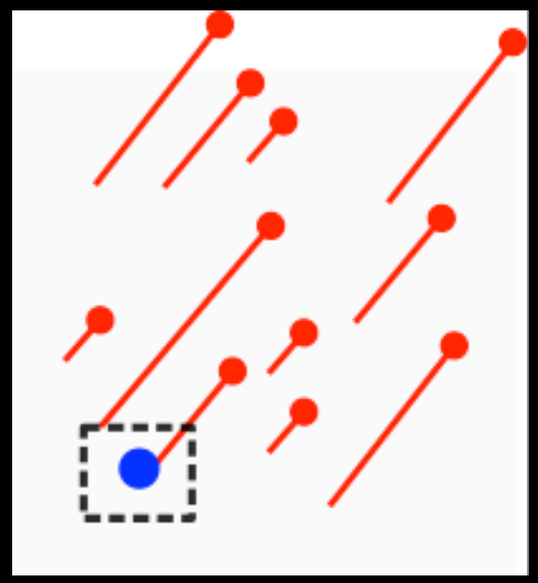


Density

Units: 1/Area (Prob/Area)

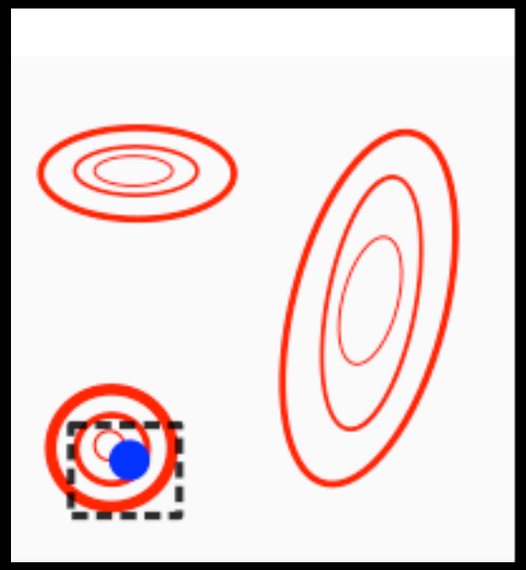
Not directly comparable.

Mass Functions versus Densities



Mass

Unitless (Probabilties)



Density

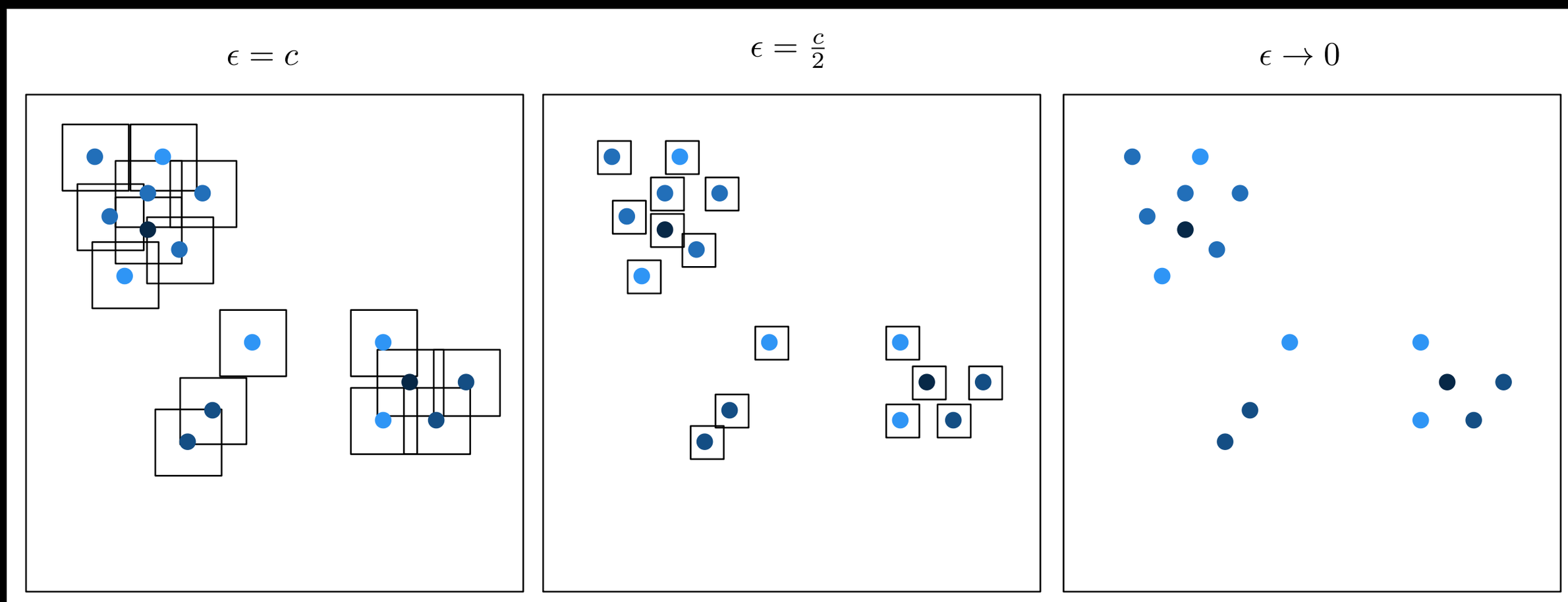
Units: 1/Area (Prob/Area)

Instead:

Compare mass in square neighbourhoods

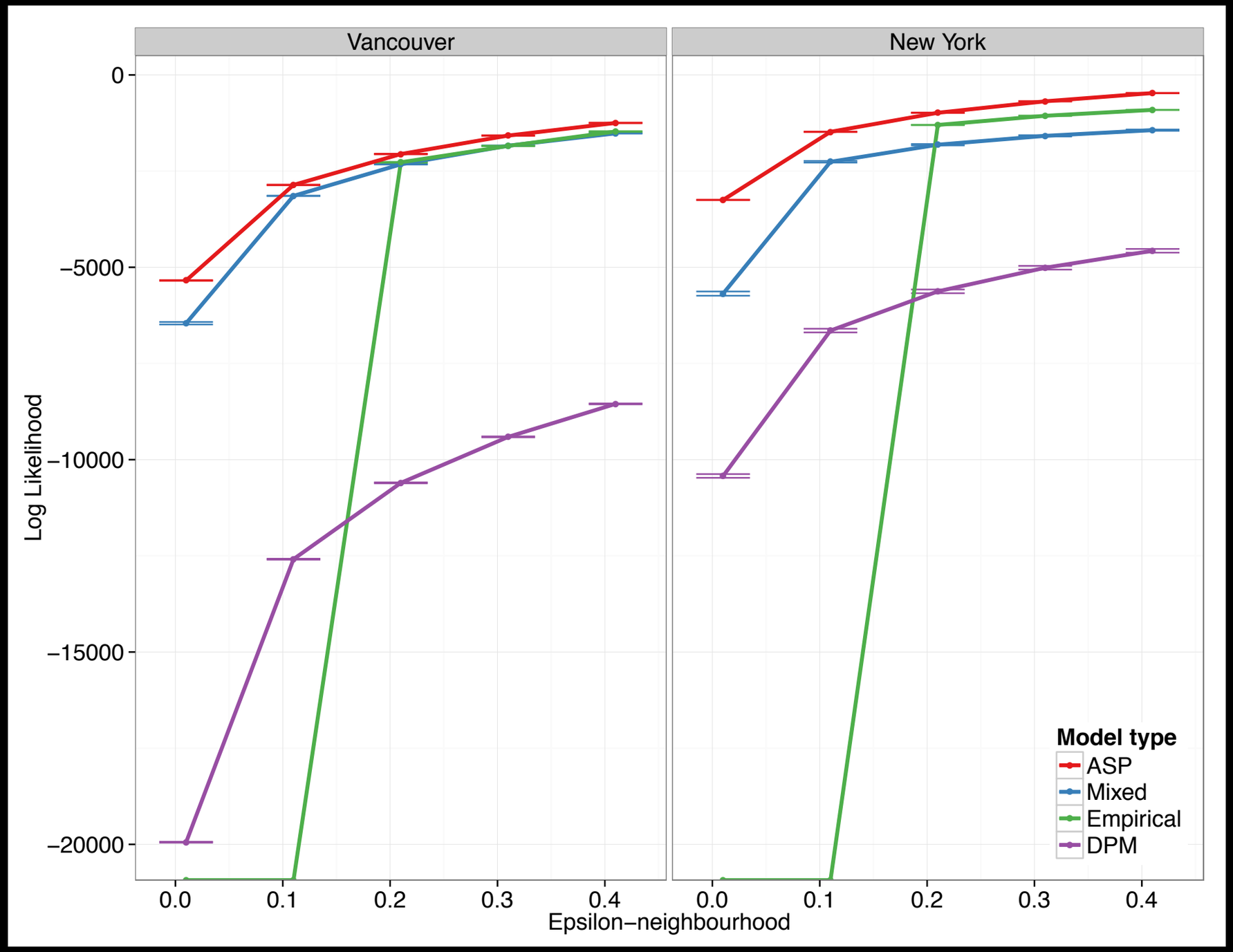
Model Comparisons

Non-standard comparison: predictive mass in square epsilon regions around held-out points.



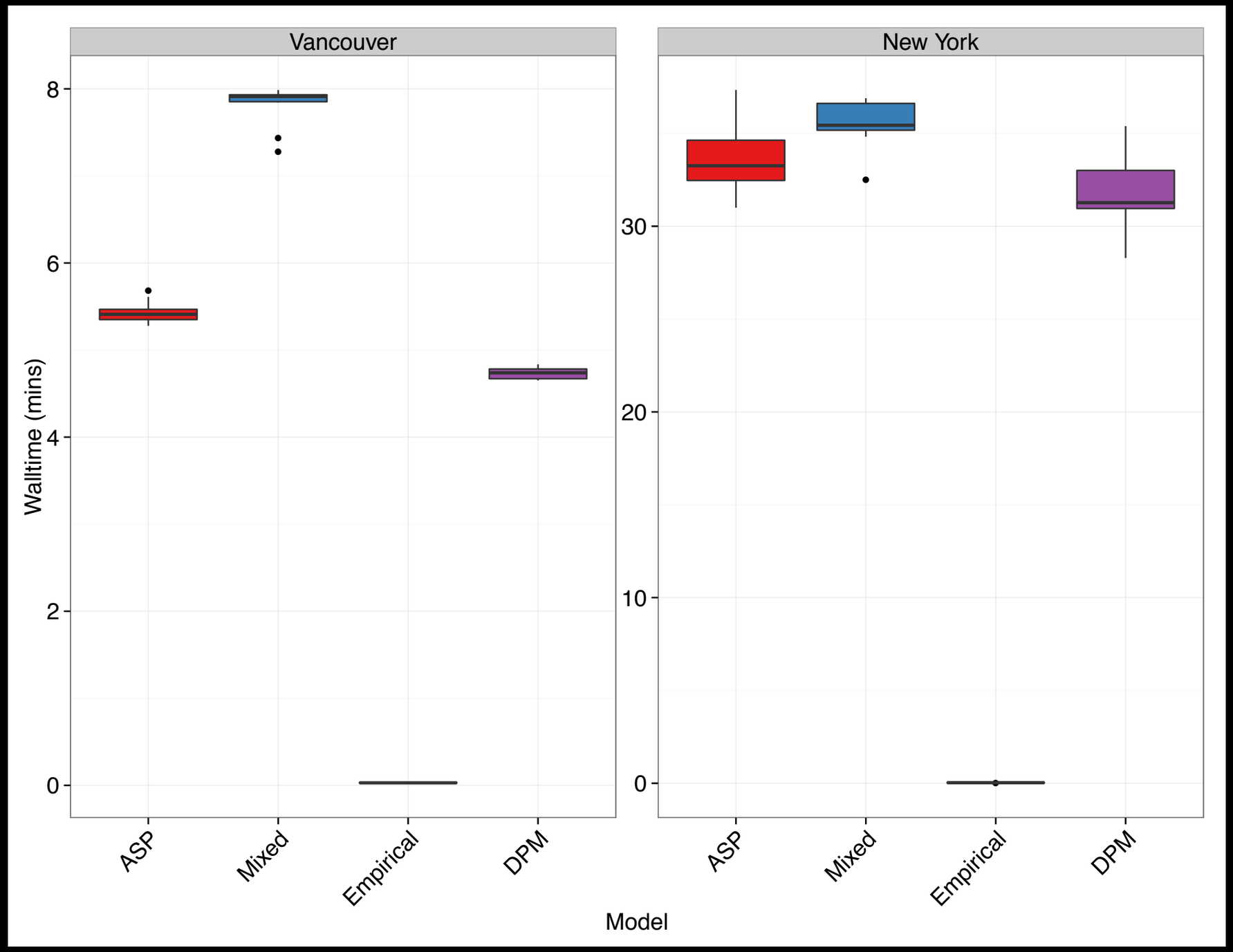
Model Comparisons

Predictive Performance



Model Comparisons

Runtime Performance



Acknowledgements

- Natural Sciences and Engineering Research Council of Canada
- Professor Jim Zidek
- Anonymous Referees
- ICML
- The audience