

# **Statistics 222, Spatial Statistics.**

## Outline for the day:

1. Exams.
2. fitpoisstoyancovariates.r
3. fithawkesstoyan
4. prototypes

## Prototypes.

Some motivating questions:

- A) What does a *typical* aftershock sequence look like?
- B) How can we tell if a particular sequence is an *outlier*?
- C) How can we group aftershock sequences into *clusters* based on the similarity of their features?

## Global Earthquake Data:

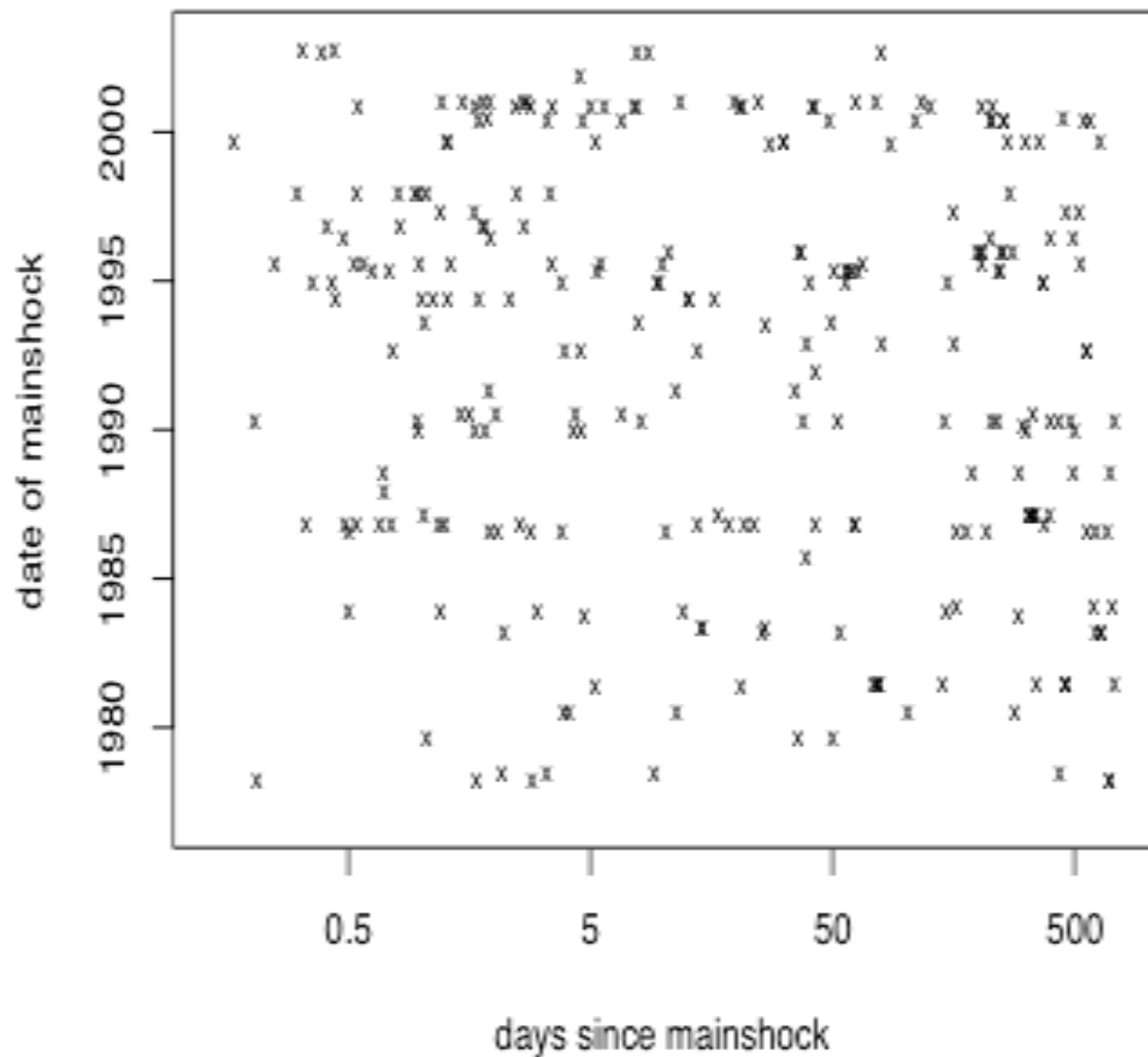
Local e.q. catalogs tend to have problems, esp. missing data.

1977: Harvard (global) catalog created.

Considered the most complete. Errors best understood.

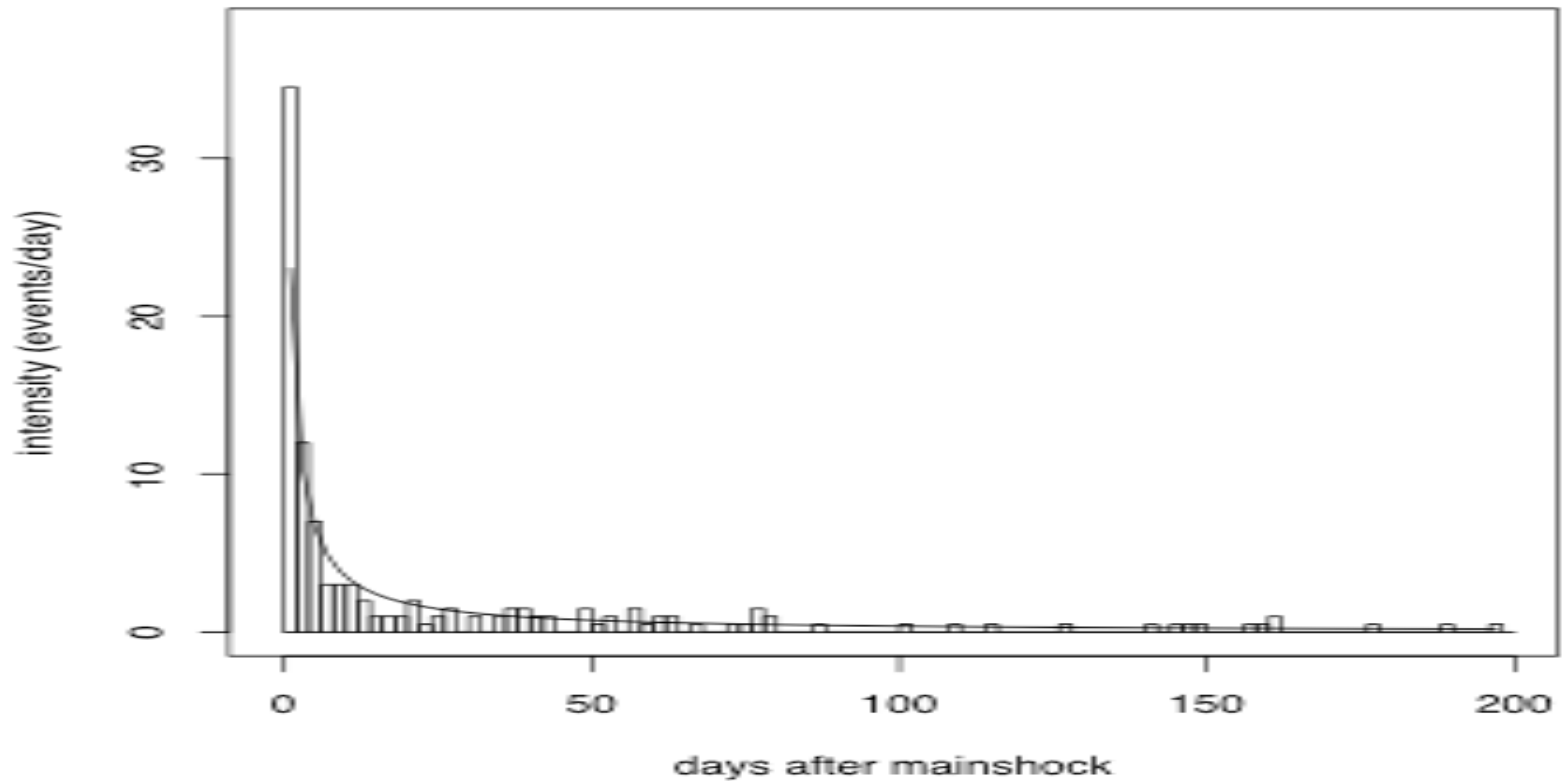
### **A collection of aftershock sequences:**

- Harvard Catalog, 1/1/77 to 3/1/03
- Shallow events only (depth < 70km)
- Mw 7.5 to 8.0
- Aftershocks: Mw > 5.5, within 100km, 0.133 days - 2 yrs.
- No Mw  $\geq$  7.5 within 200km in previous 2 yrs.
- No Mw  $\geq$  8.0 w/in 400km within 4 yrs (Molchan et al., 1997)
  - 49 mainshocks, avg. 5.47 aftershocks, SD = 4.3.



What does the typical aftershock sequence look like?  
e.g. What is typically observed after an eq of  $M_w$  7.5 - 8.0?

Modified Omori:  $K/(t+c)^p$



What does the typical aftershock sequence look like?  
e.g. What is typically observed after an eq of  $M_w$  7.5 - 8.0?

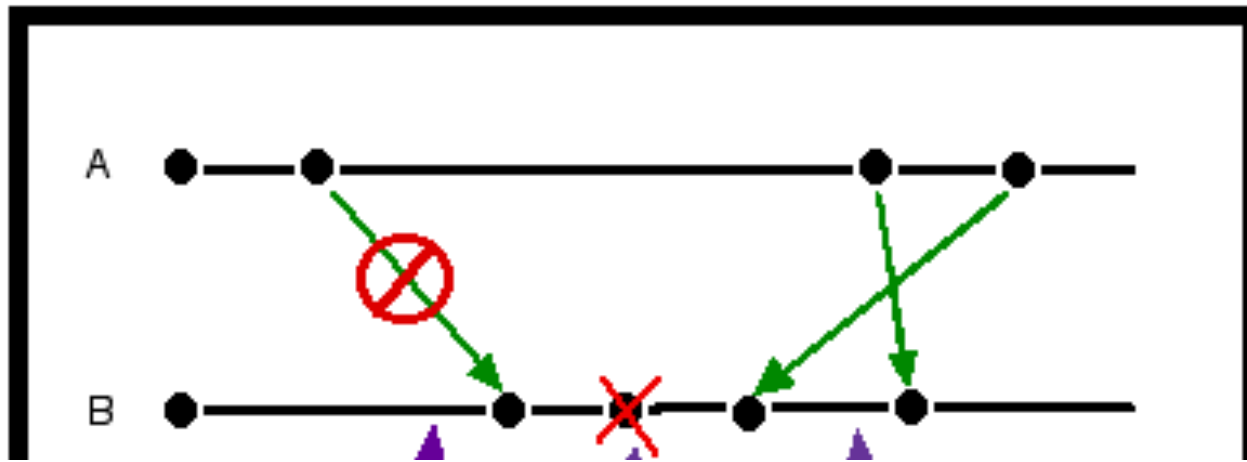
Modified Omori:  $K/(t+c)^p$

May desire a *prototype*:

a point pattern of min. distance to those observed.

Requires distance between point patterns.

# Victor-Purpura (1997) distance

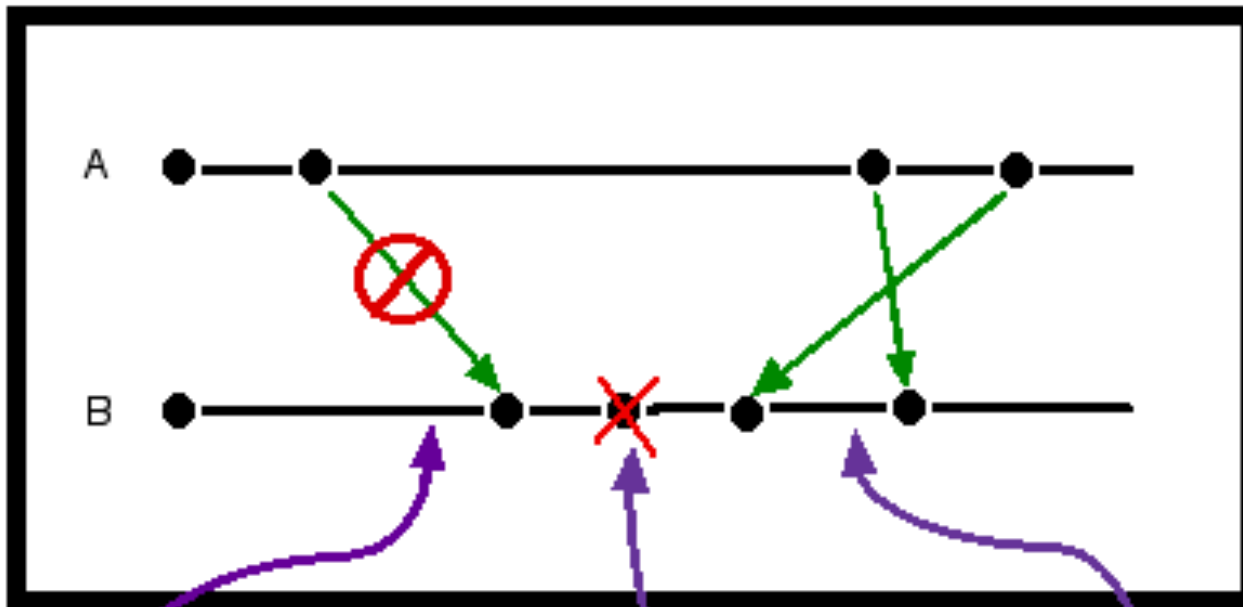


Given two point patterns:

Match each point in A to a point in B and record the horizontal distance moved (penalty  $p_m=1$  per unit moved)

Delete excess points (with penalty  $p_a$ )

# Considerations



Jumps larger than  $2 \cdot P_a$  will not be made

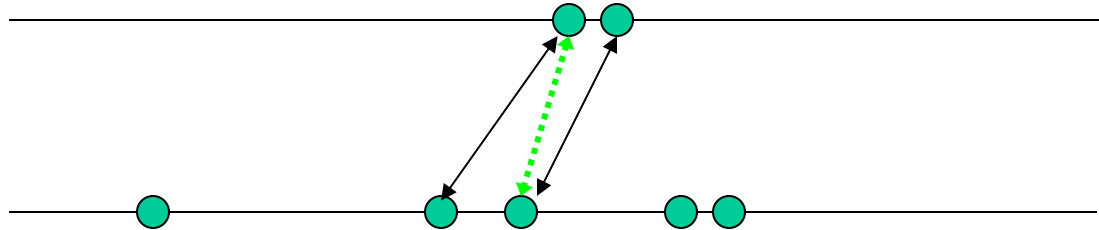
Excess points deleted

Moves that cross must be fixed



## Calculating the distance between two point patterns:

Reduces to which points are kept and which are removed.



- A point  $> 2p_a/p_m$  from its nearest neighbor is automatically removed.
- Mutual nearest neighbors within  $2p_a/p_m$  are automatically kept.

# Prototype Point Pattern

Defined such that the sum of distances from the prototype to all observed point patterns in the data set is minimized.

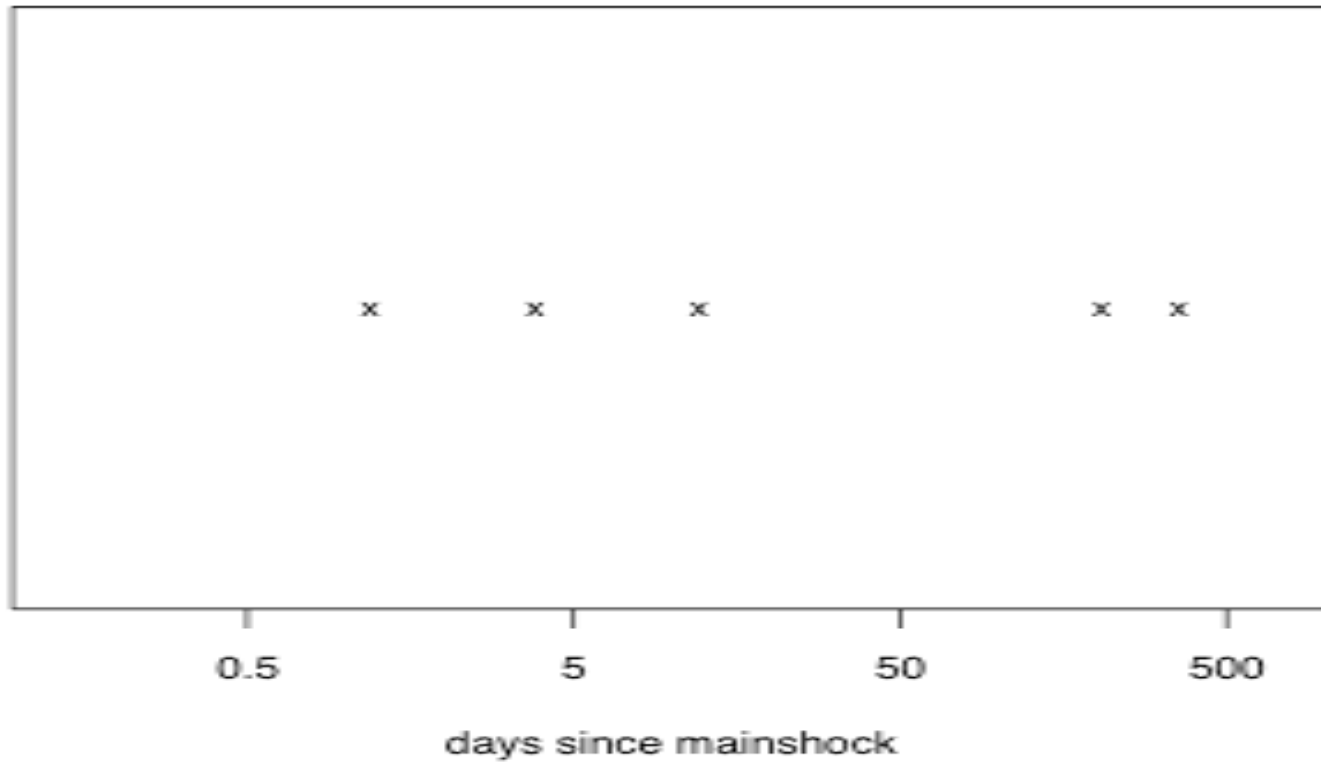
Represents a “typical” observation.

# Some properties of the prototype

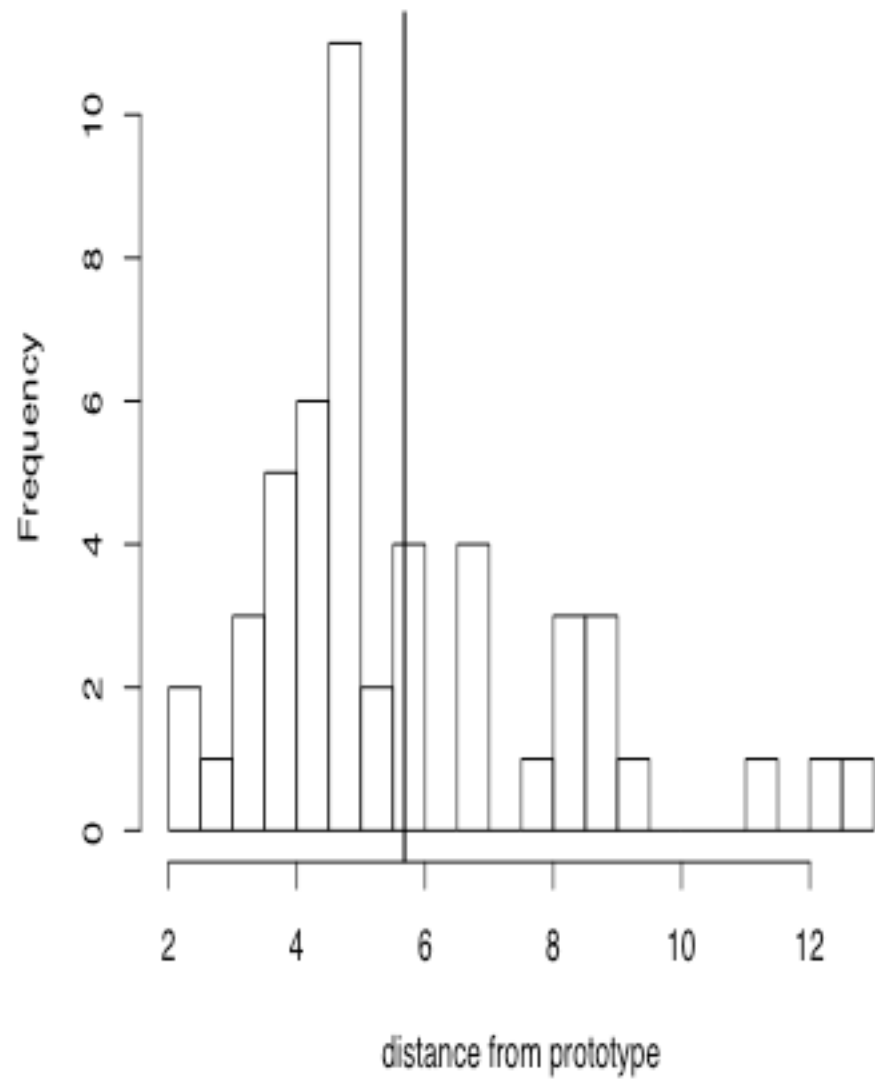
Prototype is not necessarily unique.

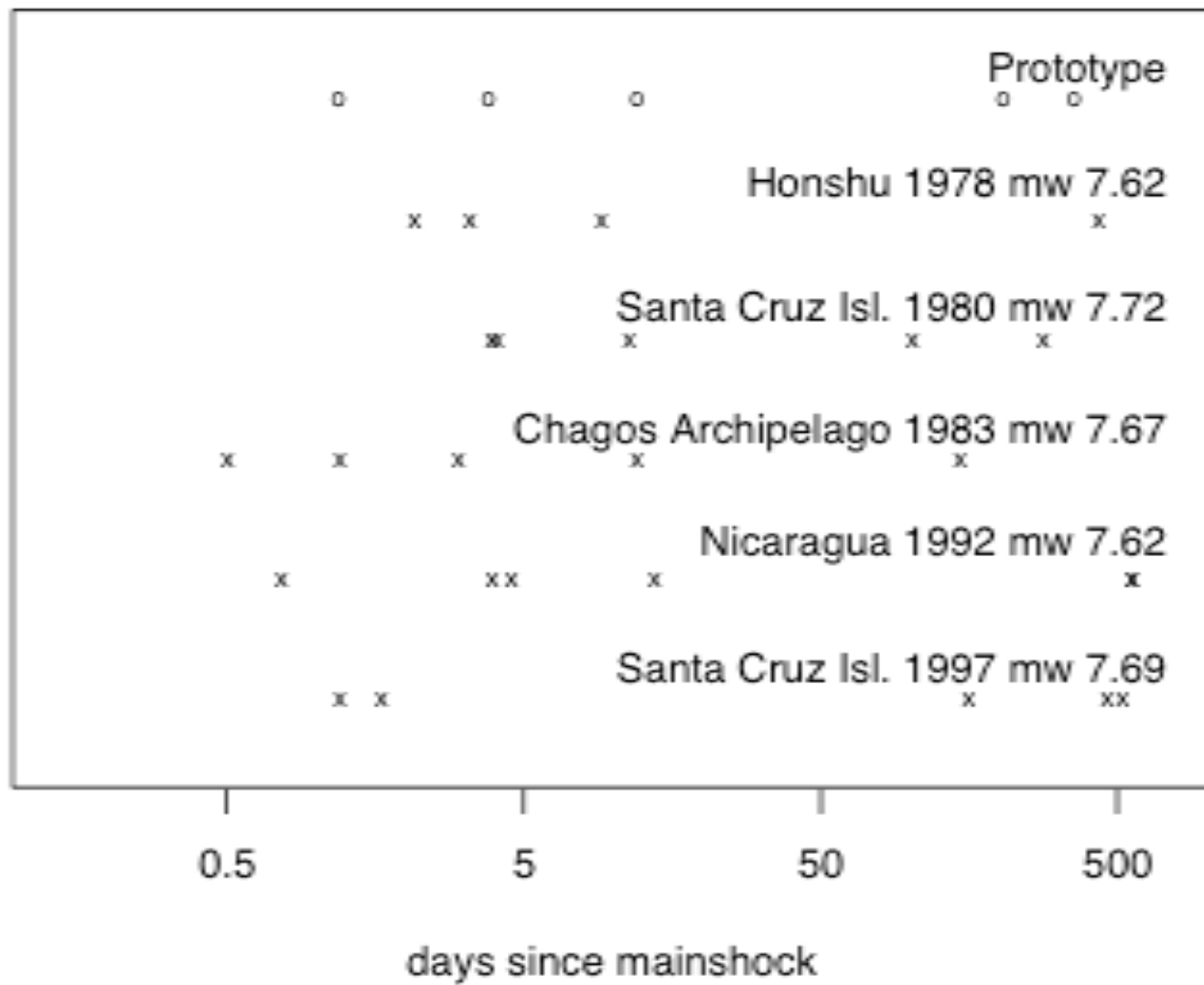
There exists a prototype pattern composed entirely of points in the dataset.

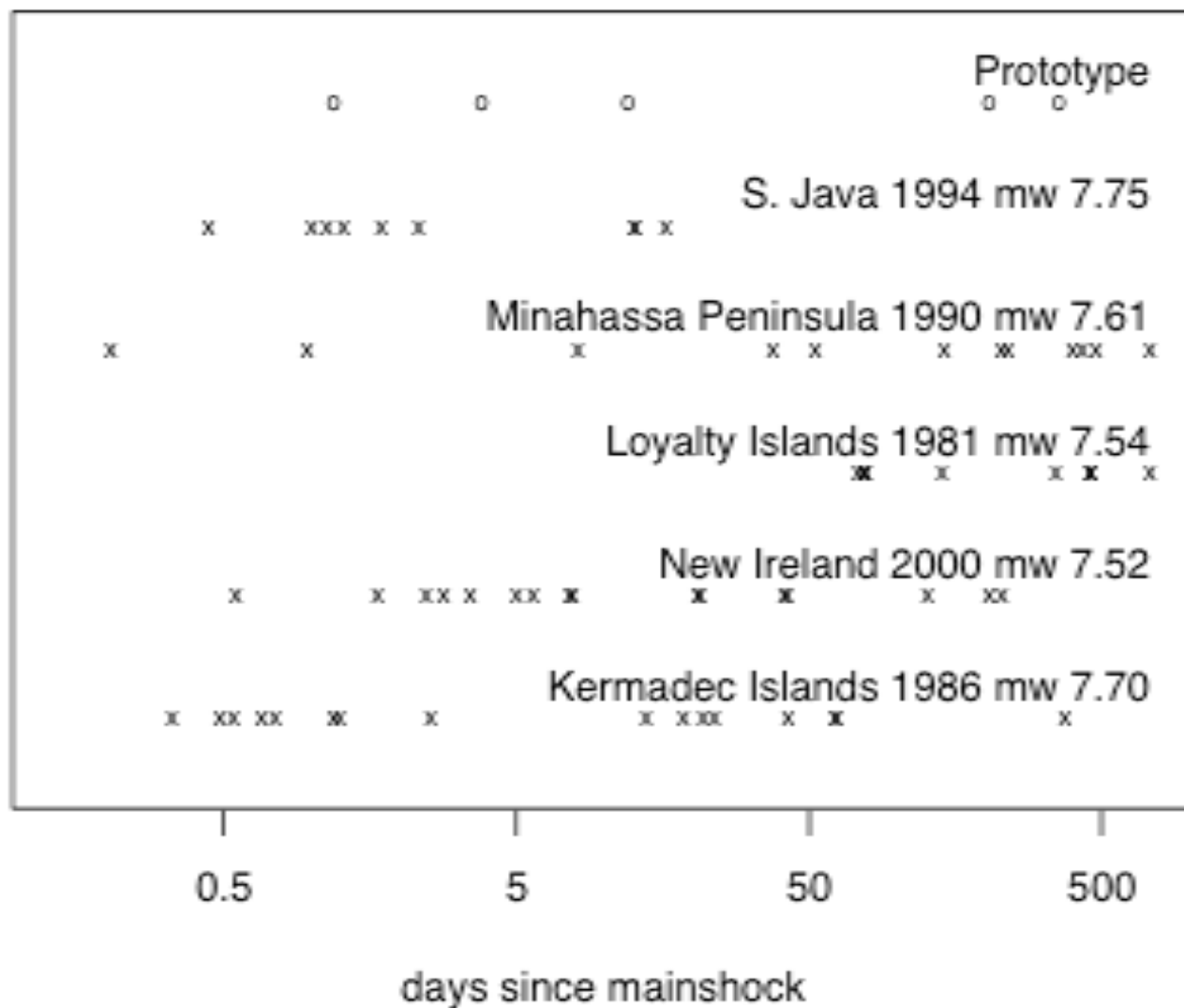
In fact, a prototype can be found such that each point it contains is the median of its associated points in distance calculations.



Uses: Data summary, outlier identification, clustering, ...

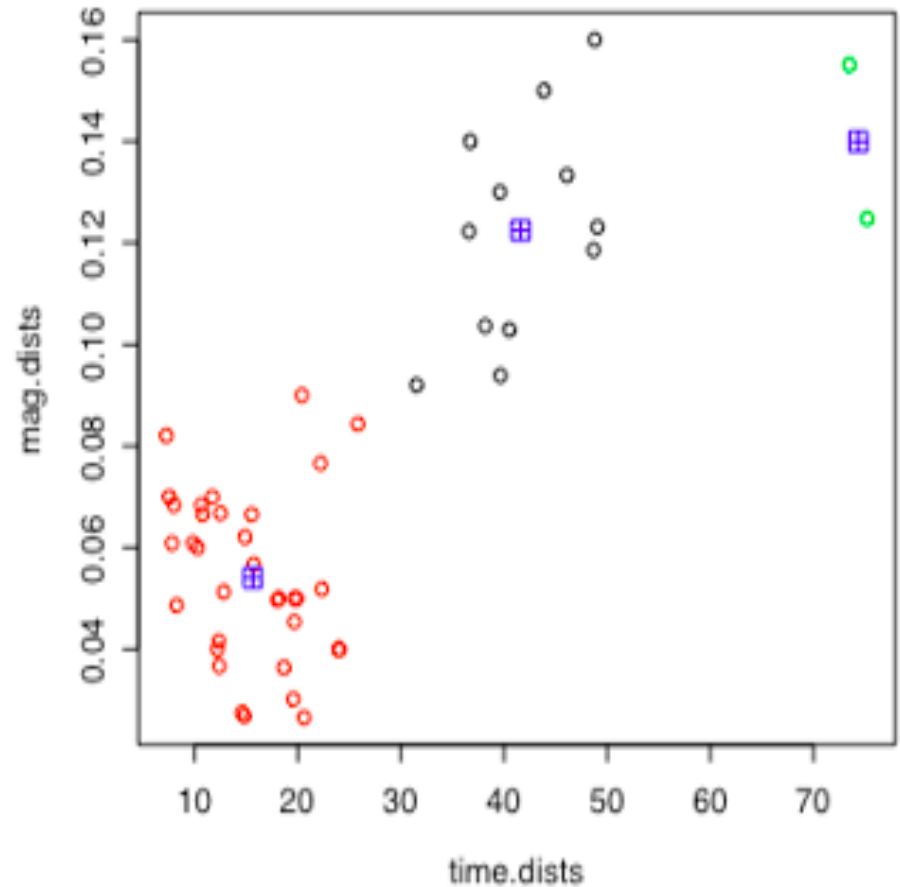






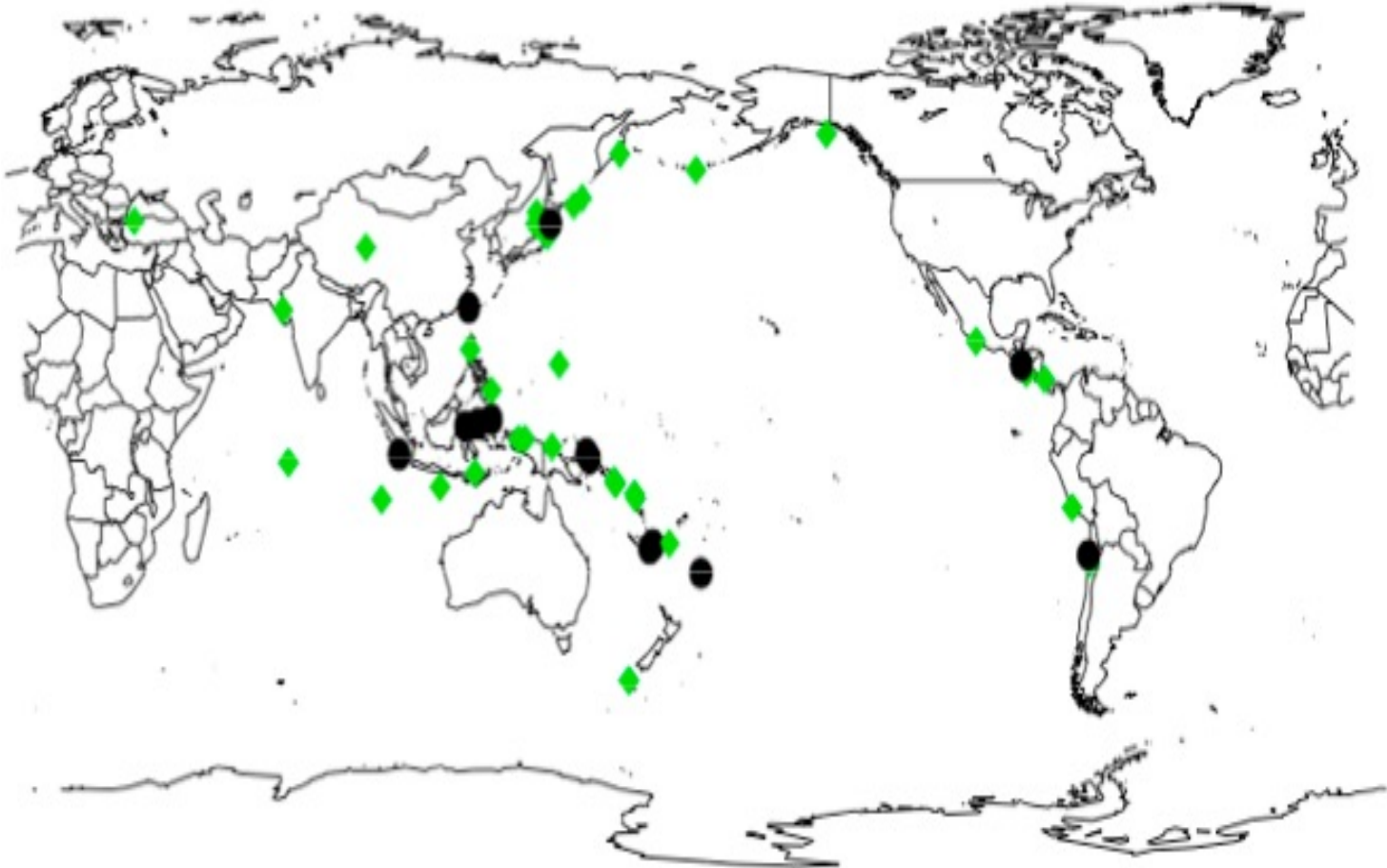
# Clusters of aftershock sequences

Distance of each aftershock sequence to the prototypes for time and magnitude





# Cluster Map

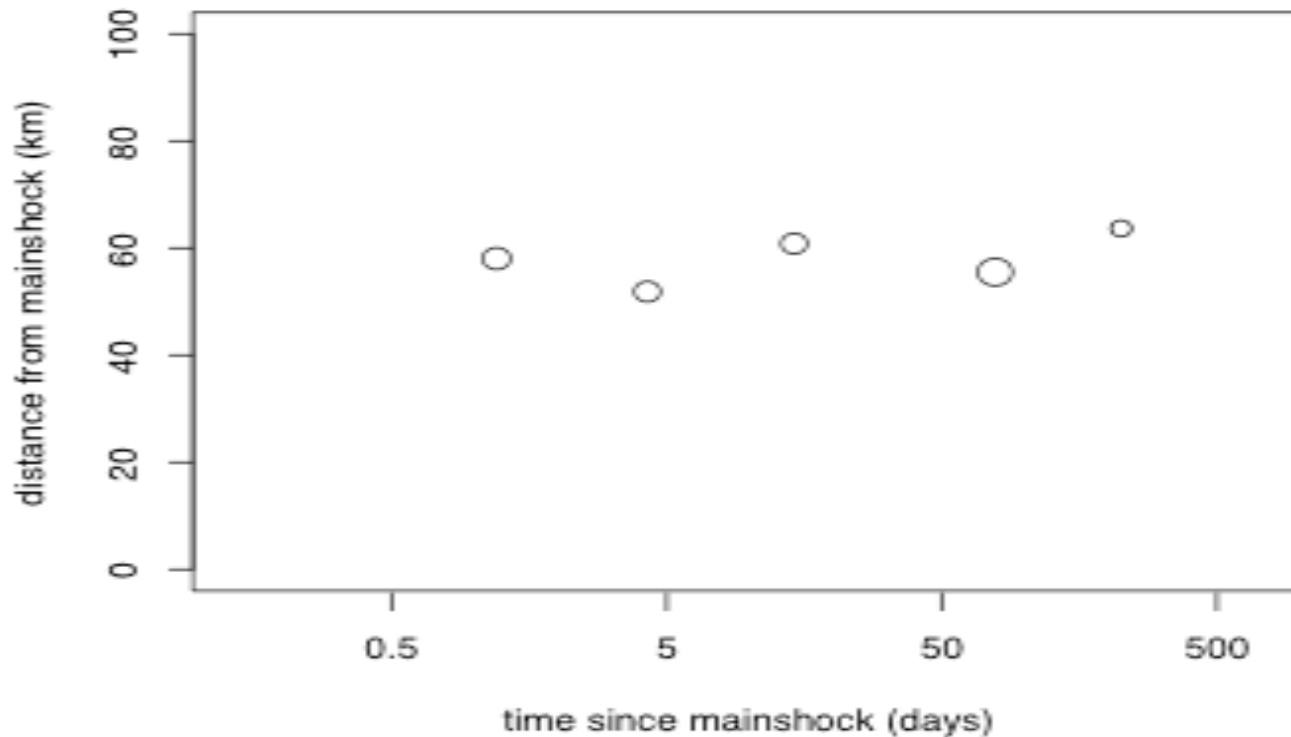


With multidimensional point processes (time,  $m_w$ , location):

No simple sequential pairing.

Mutual nearest neighbors are kept.

There exists a prototype consisting only of points whose coordinates are medians of coordinates of associated pts.



# Wildfires in CA by year

