Why Location Matters

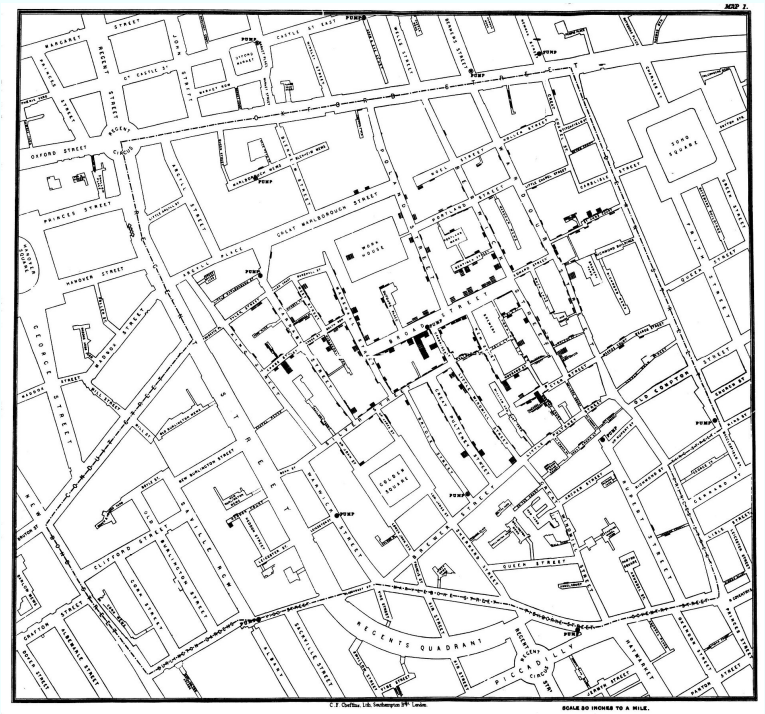
Chris Brunsdon

National Centre for Geocomputation
Maynooth University

contact: christopher.brunsdon@nuim.ie

Historically

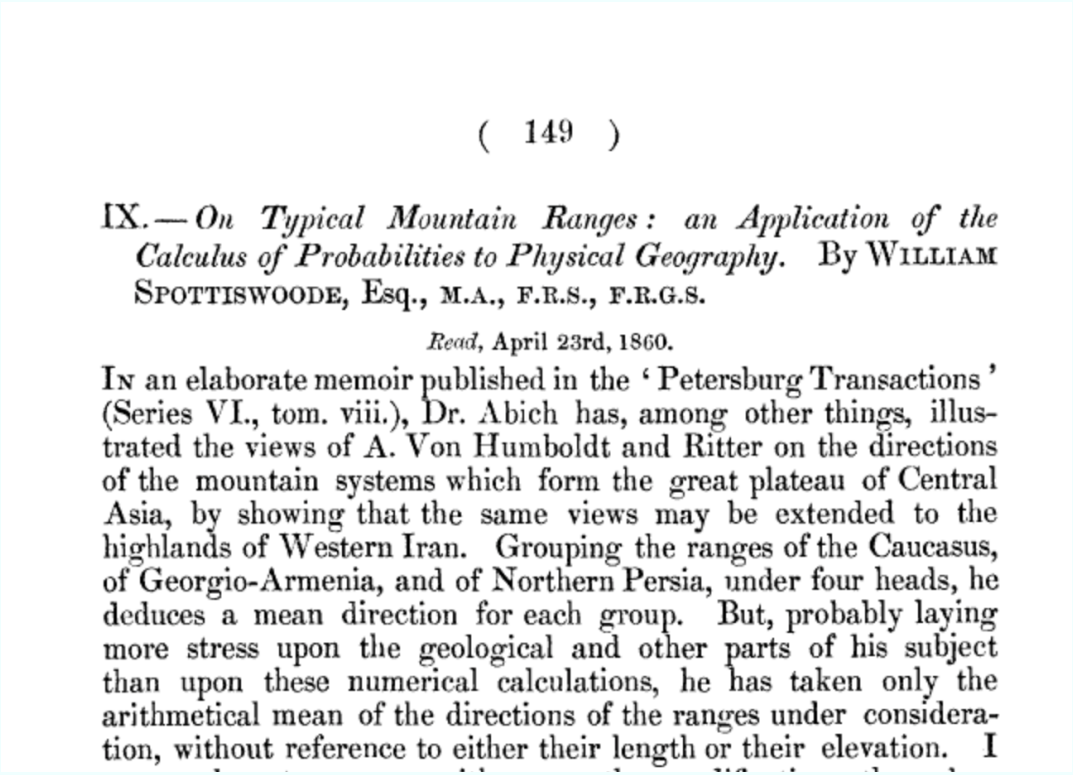
Spatial issues in data thought about at least since 1854



- John Snow's Cholera Map

Historically - Directional Data

Directional data analysis thought about since at least 1861



(149)

IX.—*On Typical Mountain Ranges: an Application of the Calculus of Probabilities to Physical Geography.* By WILLIAM SPOTTISWOODE, Esq., M.A., F.R.S., F.R.G.S.

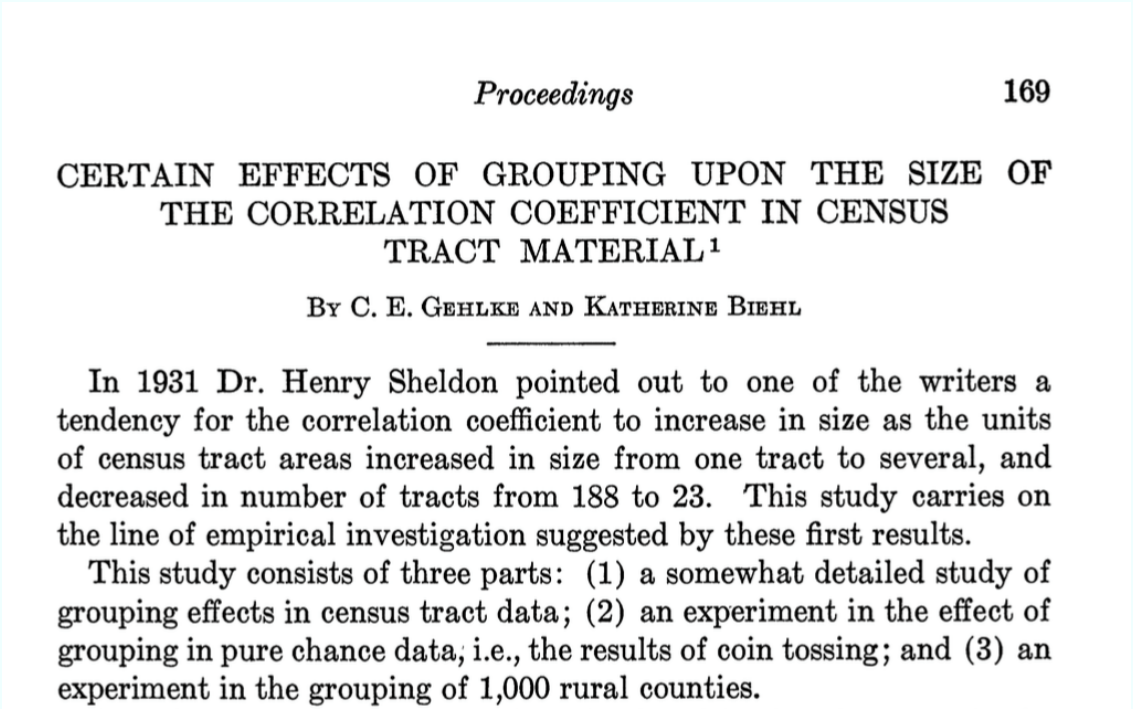
Read, April 23rd, 1860.

IN an elaborate memoir published in the 'Petersburg Transactions' (Series VI., tom. viii.), Dr. Abich has, among other things, illustrated the views of A. Von Humboldt and Ritter on the directions of the mountain systems which form the great plateau of Central Asia, by showing that the same views may be extended to the highlands of Western Iran. Grouping the ranges of the Caucasus, of Georgio-Armenia, and of Northern Persia, under four heads, he deduces a mean direction for each group. But, probably laying more stress upon the geological and other parts of his subject than upon these numerical calculations, he has taken only the arithmetical mean of the directions of the ranges under consideration, without reference to either their length or their elevation. I

William Spottiswoode ,"On Typical Mountain Ranges: An Application of the Calculus of Probabilities to Physical Geography", The Journal of the Royal Geographical Society of London Vol. 31 (1861), pp. 149-154

Modifiable Areal Unit Problem

Thought about since at least 1934



Proceedings 169

CERTAIN EFFECTS OF GROUPING UPON THE SIZE OF THE CORRELATION COEFFICIENT IN CENSUS TRACT MATERIAL[1]

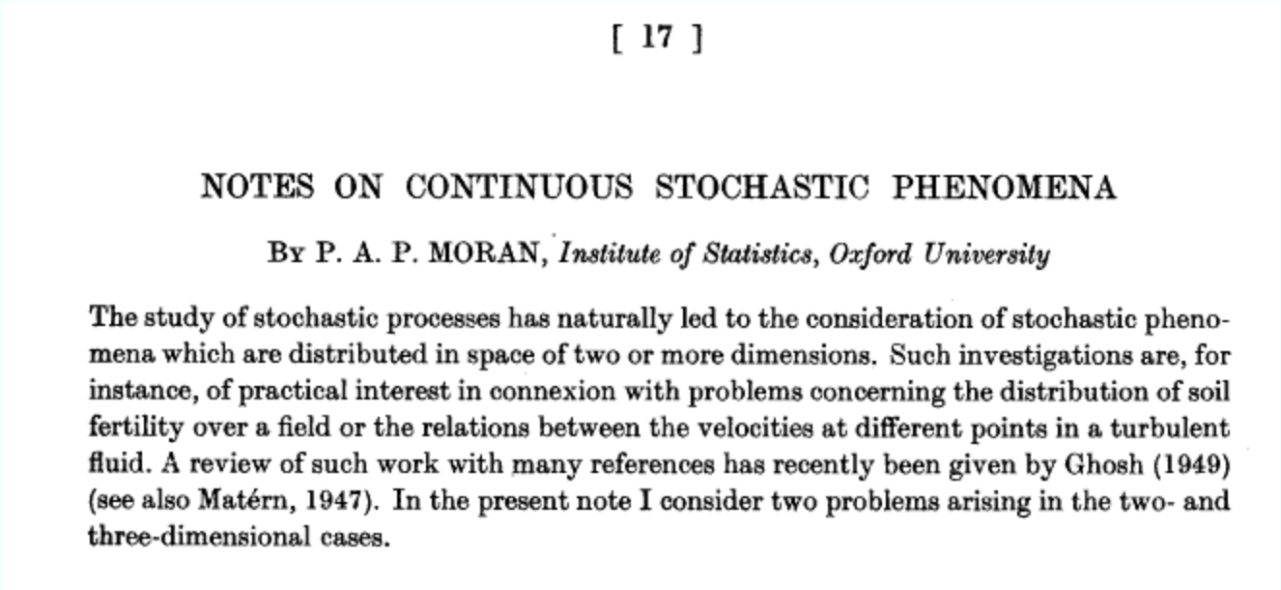
By C. E. Gehlke and Katherine Biehl

In 1931 Dr. Henry Sheldon pointed out to one of the writers a tendency for the correlation coefficient to increase in size as the units of census tract areas increased in size from one tract to several, and decreased in number of tracts from 188 to 23. This study carries on the line of empirical investigation suggested by these first results.

This study consists of three parts: (1) a somewhat detailed study of grouping effects in census tract data; (2) an experiment in the effect of grouping in pure chance data, i.e., the results of coin tossing; and (3) an experiment in the grouping of 1,000 rural counties.

Gehlke, C. E.; Biehl, Katherine (March 1934). "Certain effects of grouping upon the size of the correlation coefficient in census tract material". Journal of the American Statistical Association 29 (185A): 169–170

Spatial autocorrelation in data thought about since at least 1950



[17]

NOTES ON CONTINUOUS STOCHASTIC PHENOMENA

By P. A. P. MORAN, *Institute of Statistics, Oxford University*

The study of stochastic processes has naturally led to the consideration of stochastic phenomena which are distributed in space of two or more dimensions. Such investigations are, for instance, of practical interest in connexion with problems concerning the distribution of soil fertility over a field or the relations between the velocities at different points in a turbulent fluid. A review of such work with many references has recently been given by Ghosh (1949) (see also Matérn, 1947). In the present note I consider two problems arising in the two- and three-dimensional cases.

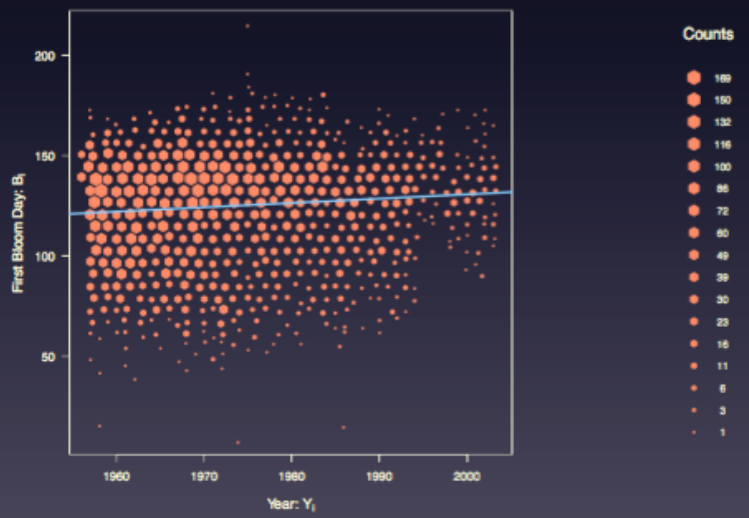
Moran, P. A. P. (1950). "Notes on Continuous Stochastic Phenomena". Biometrika 37 (1): 17–23.

Currently (. . . ish)

Simpson's Paradox

$d_{ij} = S + Ry_{ij} + \epsilon_{ij}$:

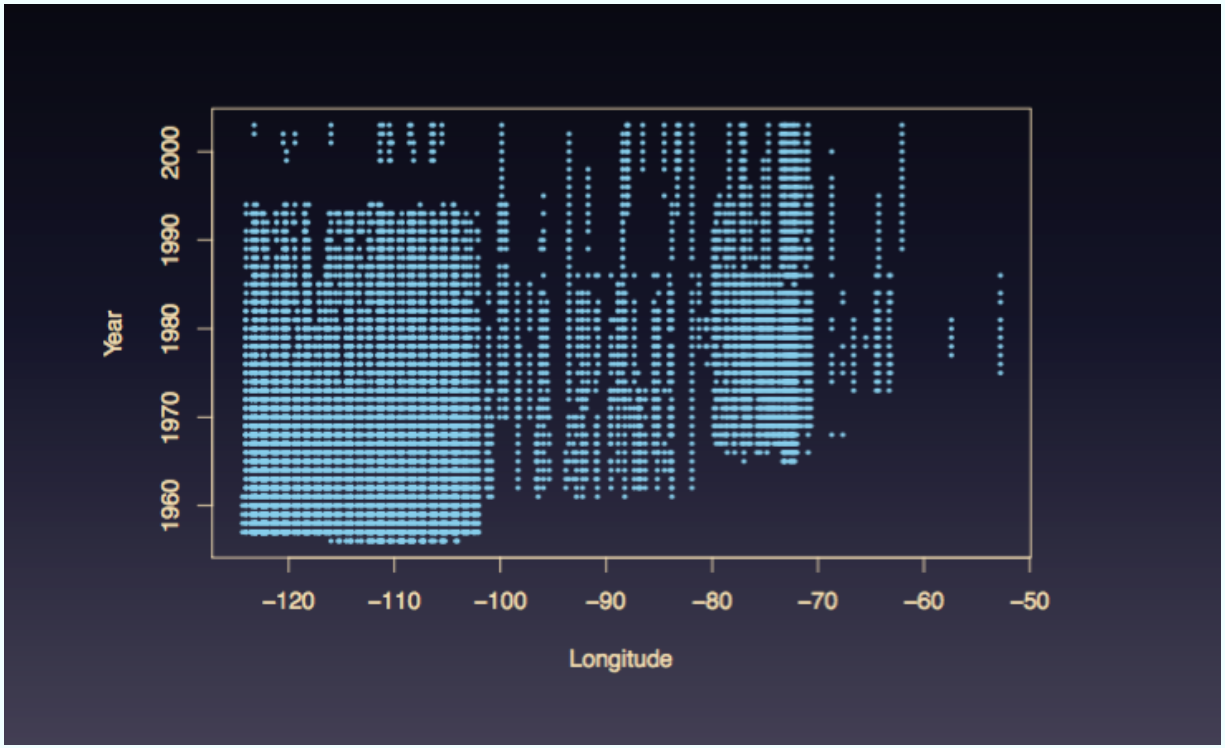
	Estimate	Std. Error	t value	Pr(>|t|)
(Intercept)	-0.0000	0.1861	-0.00	1.0000
year	0.2150	0.0178	12.11	0.0000



Brunsdon, C. and Comber A.(2012) "Assessing the changing flowering date of the common lilac in North America: a random coefficient model approach" GeoInformatica 16(4): 675-690

Currently (. . . ish)

Simpson's Paradox

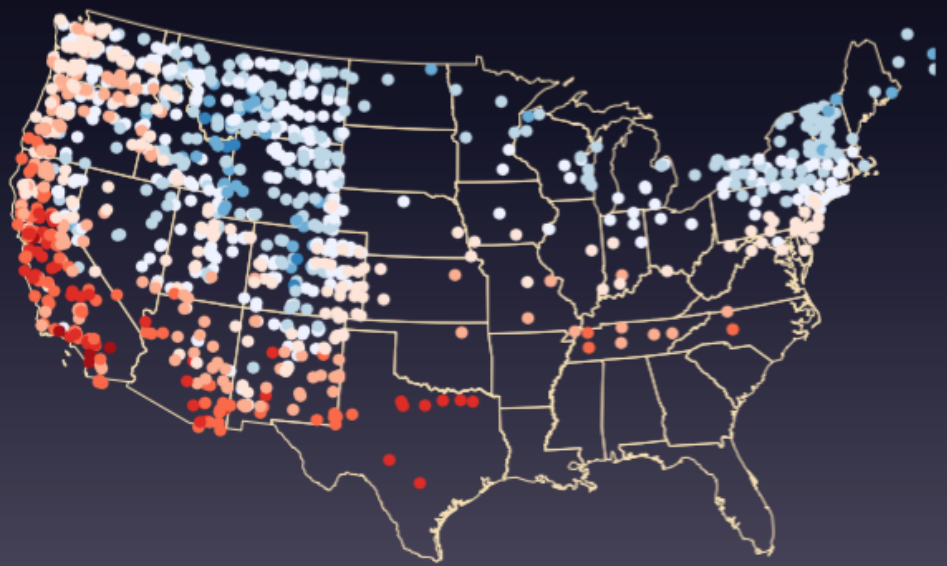


Currently (. . . ish)

Simpson's Paradox

Random Coefficient Model (MLM)

$$d_{ij} = s_i + Ry_{ij} + \epsilon_{ij}$$



	Value	Std.Error	DF	t-value	p-value
(Intercept)	-0.95	0.64	13139.00	-1.50	0.13
year	-0.18	0.01	13139.00	-17.18	0.00

Why be Spatially Aware?

- Atomic fallacy - Errors due to treating elements as separate 'atoms' outside of their spatial context.
- Ecological fallacy - Errors occur in part from spatial aggregation - aggregated values not always representitive.
- Being visual - important for spatial data
- Is everything homogeneous? - relationships between variables could change over space
- Regional Bias in response to surveys
- All of these relevent to surveys...

Introducing the talks

- **17:10-17:30** Using geographically weighted regression to explore spatial variation in survey non-response (Kaisa Lahtinen, University of Liverpool)
- **17:30-17:50** The role of geography in small area estimation (Nikos Tzavidis, University of Southampton)
- **17:50-18:10** Exploring survey results with visualization (Aidan Slingsby, City University London)
- **18:10-18:30** Discussion

Thank you