Explaining the Dynamics of City Size and Related Distributions

Scaling, Rank Size and Rank Clocks

Michael Batty

m.batty@ucl.ac.uk

@jmichaelbatty

http://www.ComplexCity.info/

http://www.casa.ucl.ac.uk/

Outline

- City-Size Distributions: Scaling and the Rank-Size Rule
- Macro-Stability & Micro-Volatility
- The Rank Clock and Other Visual Mnemonics
- Classic Exemplars: US City Populations 1790-2000
- US Metro Area Populations
- Japanese Populations: Cities at Different Scales
- Adding Place to Rank Clocks:
- Animations: Rank Clocks and Rank Space
- Next Steps

<u>The key thesis</u> – we need to visualise dynamic systems where the system appears stable at the macro level but volatile at the micro – there are two nice quotes to start

"I will [tell] the story as I go along of small cities no less than of great. Most of those which were great once are small today; and those which in my own lifetime have grown to greatness, were small enough in the old days"

From **<u>Herodotus – The Histories</u>** –
Quoted in the frontispiece by Jane Jacobs (1969)
<u>The Economy of Cities</u>, Vintage Books, New York

But cities have a remarkable degree of regularity and stability with respect to their size …

"The size distribution of cities in the United States is startlingly well described by a simple power law: the number of cities whose population exceeds P is proportional to 1/P. This simple regularity is puzzling; even more puzzling is the fact that it has apparently remained true for at least the past century."

Paul Krugman, 1996, Confronting the Mystery of Urban Hierarchy, **Journal of the Japanese and International Economies 10**, 399–418.

Why do systems of cities in terms of their sizes appear to conform to such simple representations?

There are many very simple stochastic models that give rise to such laws and all dependent on competition between cities with some cities getting ahead through random change. Simon’s (1954) model is one of the first and simplest: Krugman describes it as ‘nihilistic’.

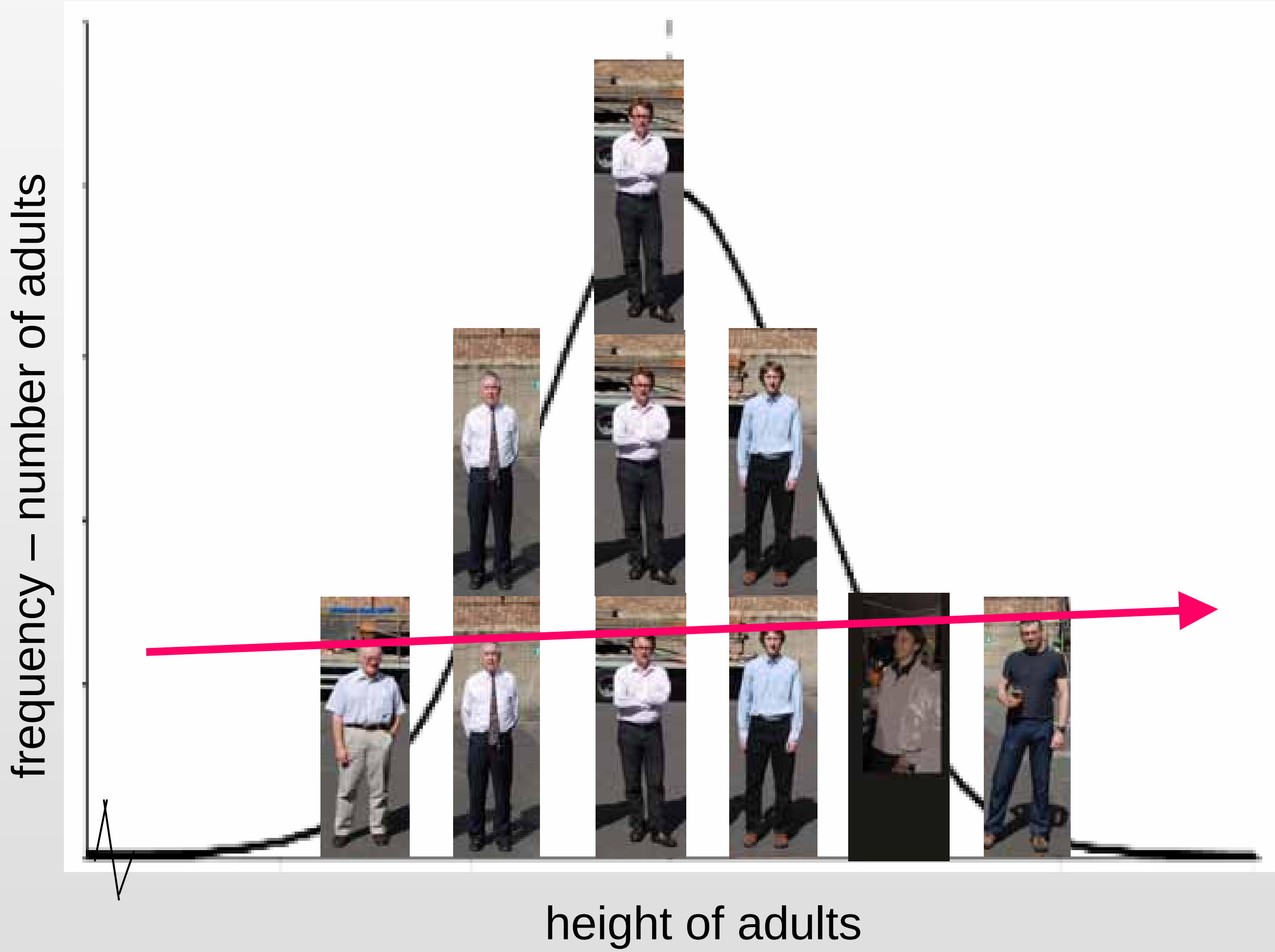
Here we will not dwell on these models but simply on the fact that within the very strong envelop of regularity that these systems appear to exhibit there is enormous volatility, quite enormous and this add to the puzzle even further. Here I will show the puzzle.

Let me begin then by defining size and frequency.

City-Size Distributions: Scaling and the Rank-Size Rule

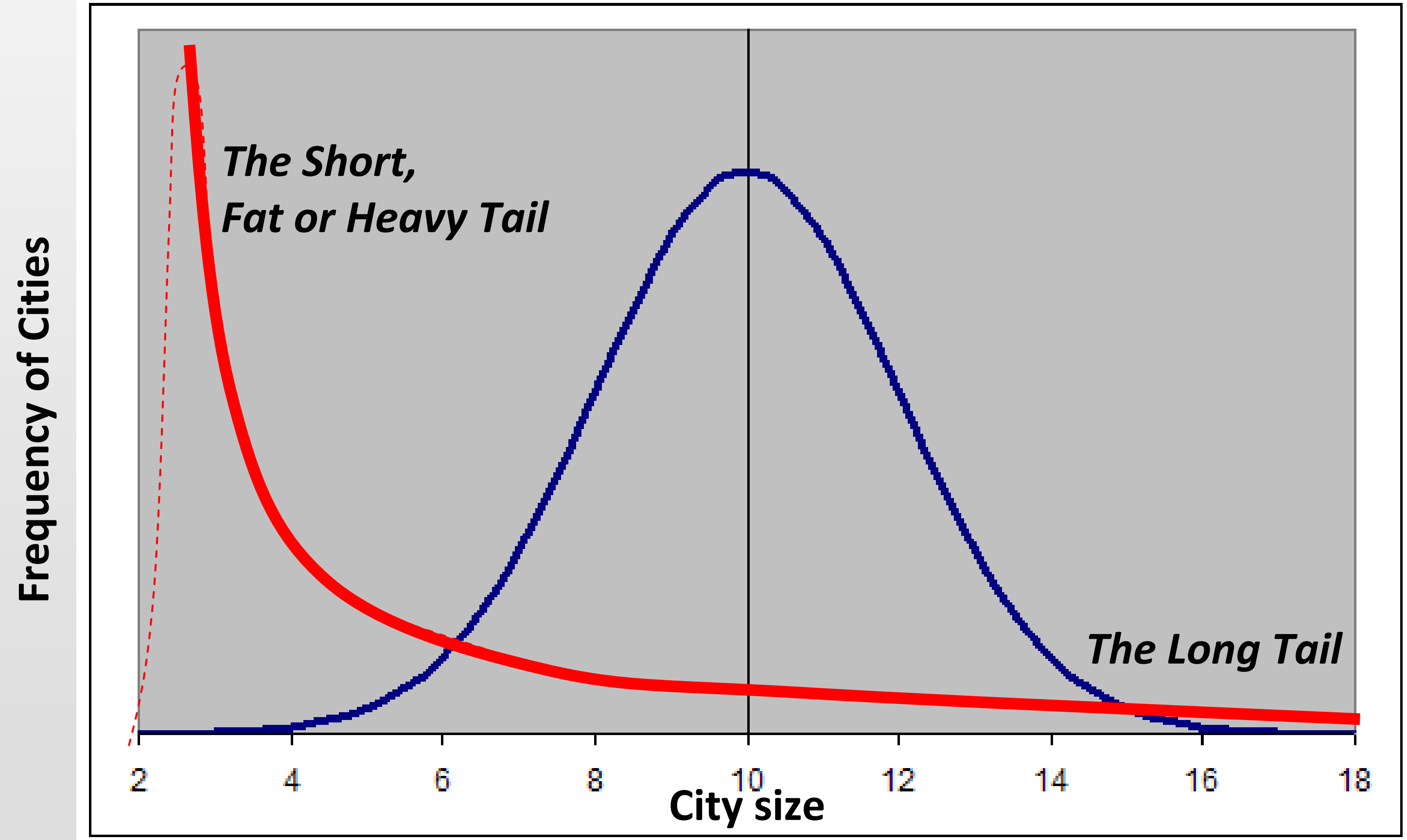
Let us look at a group of people all of different heights

Mike *Rui* *Andy* *Richard* *Duncan* *Phil*

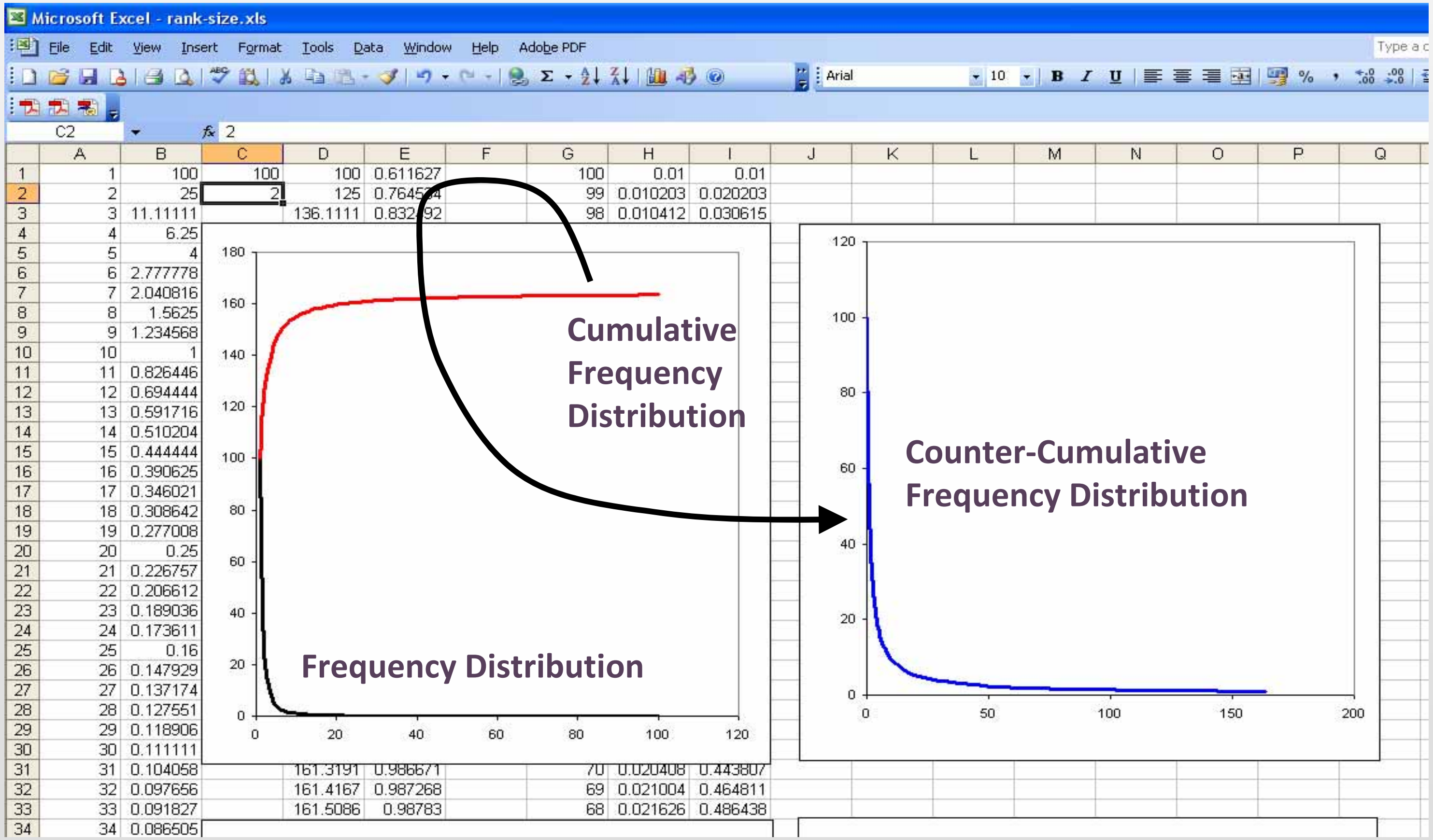


frequency – number of adults
height of adults

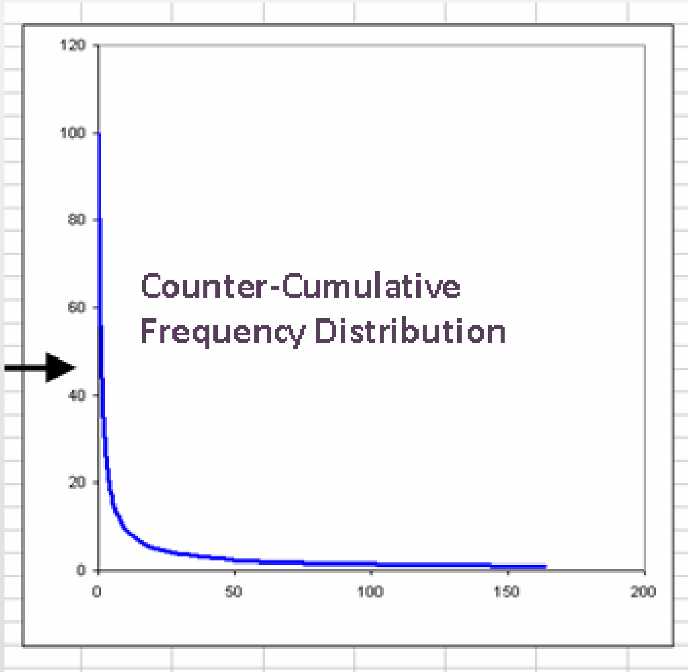
When we look at things like city sizes, we do not find a normal distribution but we find something that is like a power law or at least a lognormal

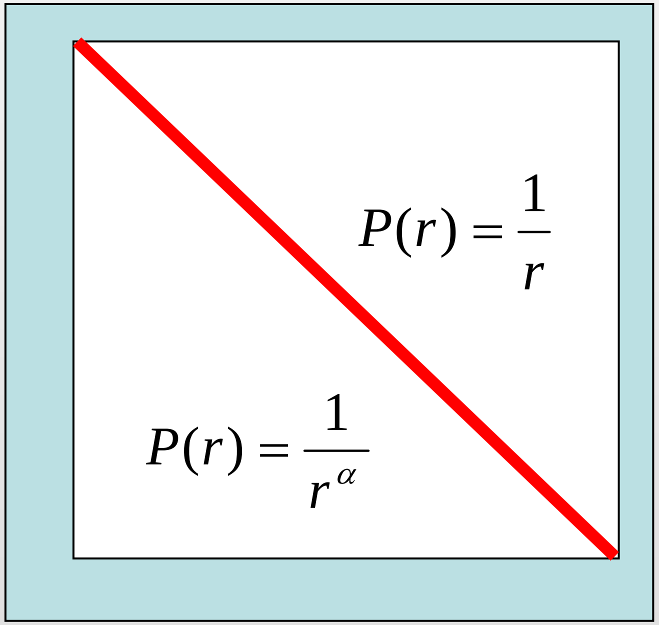


I am going to spare you the algebra although it is easy and this is how we get the rank size distribution from the frequency power



Now if we take logs – i.e. a simple transformation, this power law becomes a straight line in 2-d space and it is this form that we refer to as the rank-size rule





Macro-Stability & Micro-Volatility

Essentially we can illustrate the stability of city size population by showing how this rank size curve changes over time

It remains quite stable for the US from 1790 to 1930 and this is what Zipf, amongst others, discovered in the 1930s. And it prompted the Krugman quote.

It is all in Zipf's famous book. And Paul Krugman in the late 1990s also said that Zipf's Law and Pareto's Law before, are the only real examples of iron laws in the social sciences. Let us see what he meant.

The Rank Clock and Other Visual Mnemonics

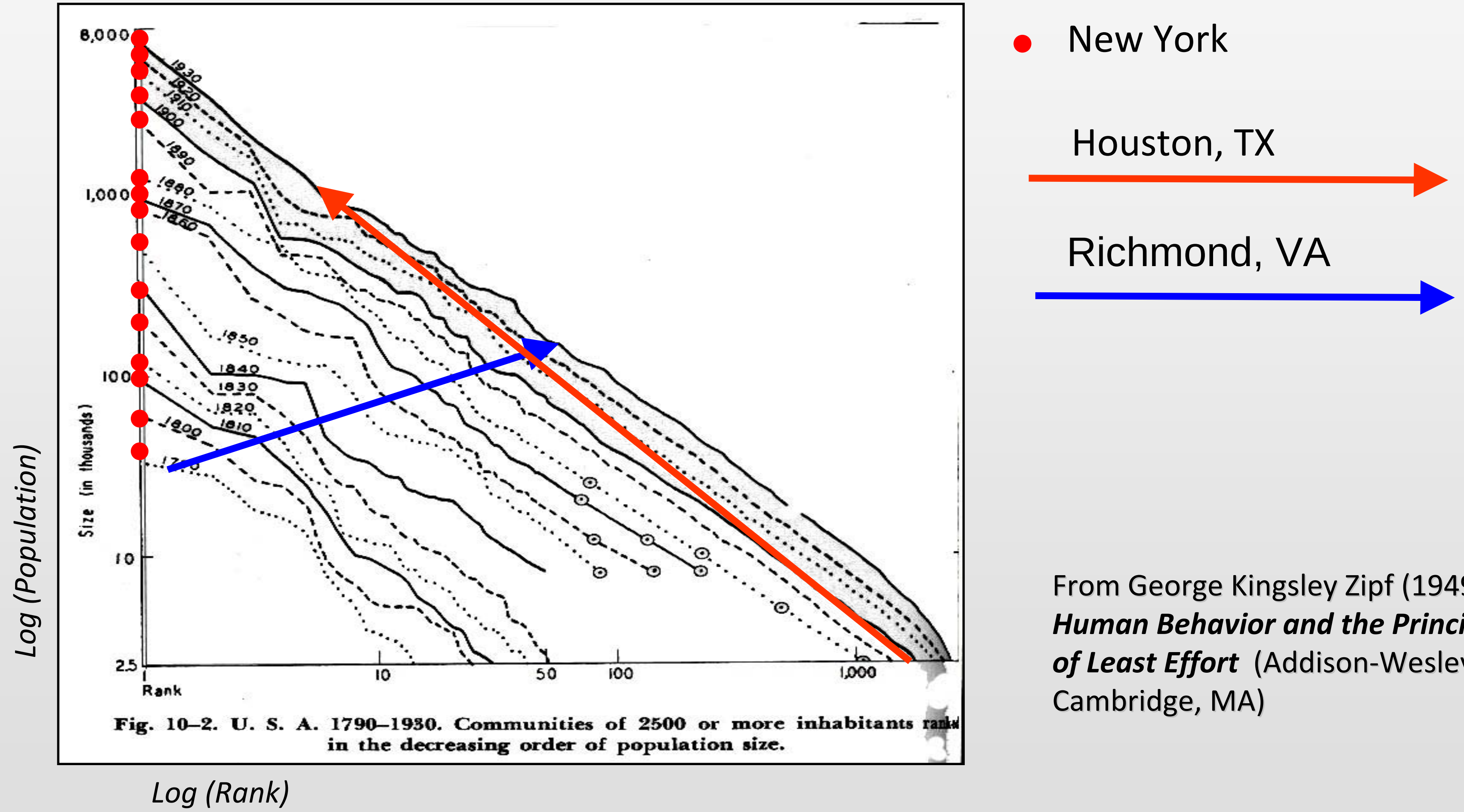


Fig. 10–2. U. S. A. 1790–1930. Communities of 2500 or more inhabitants ranked in the decreasing order of population size.

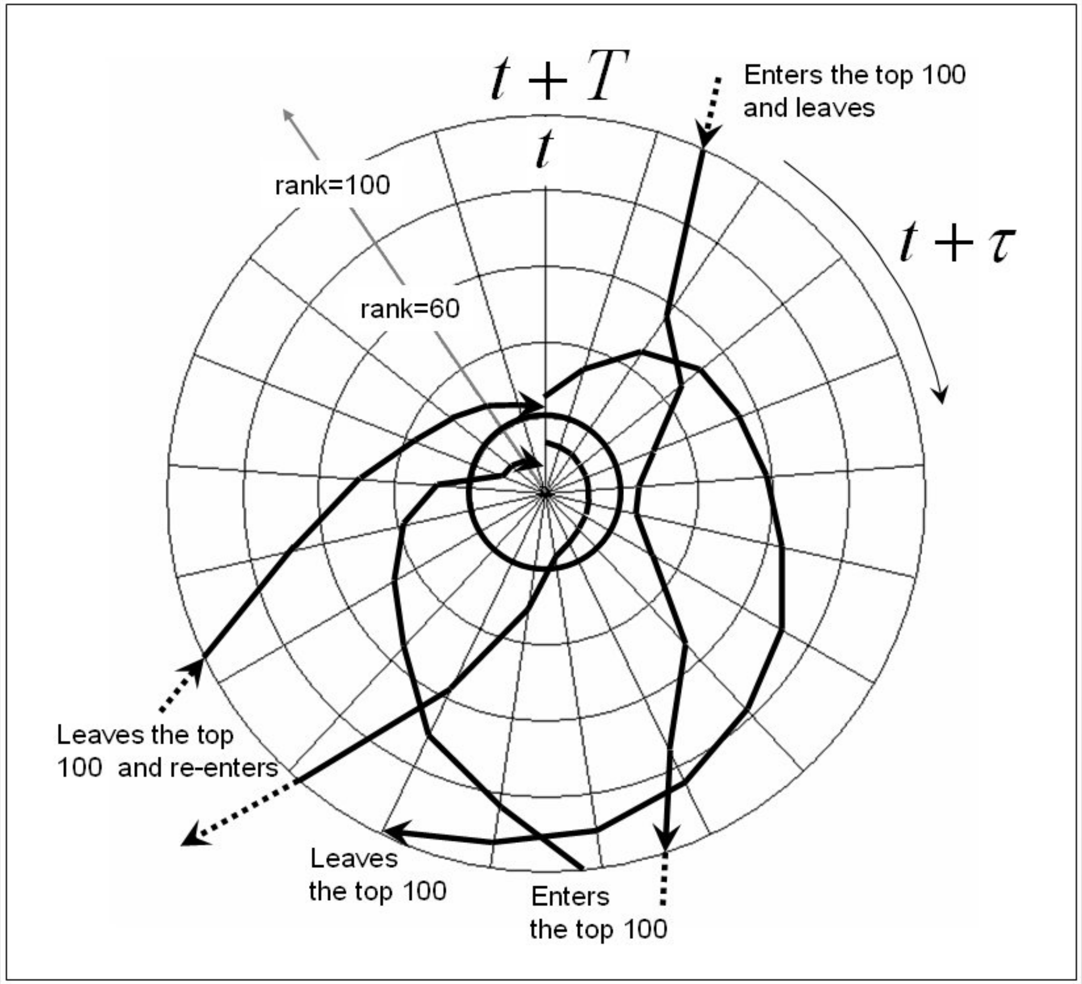
From George Kingsley Zipf (1949) ***Human Behavior and the Principle of Least Effort*** (Addison-Wesley, Cambridge, MA)

The idea of the rank clock is to discard population because rank is its equivalent and just plot changes in rank with respect to time

We can do this in rank space but it is not as effective as in polar coordinate space

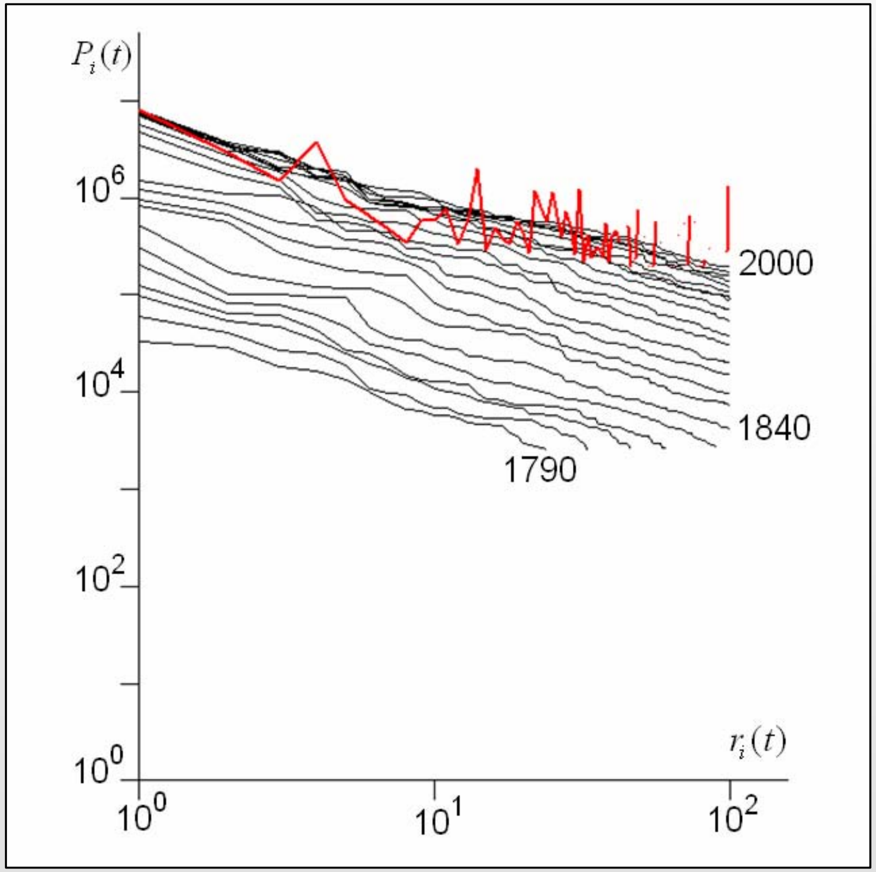
Here is an example of the possible trajectories of cities on the clock and we will stick with our US cities example for a bit longer, showing various possible plots and animations

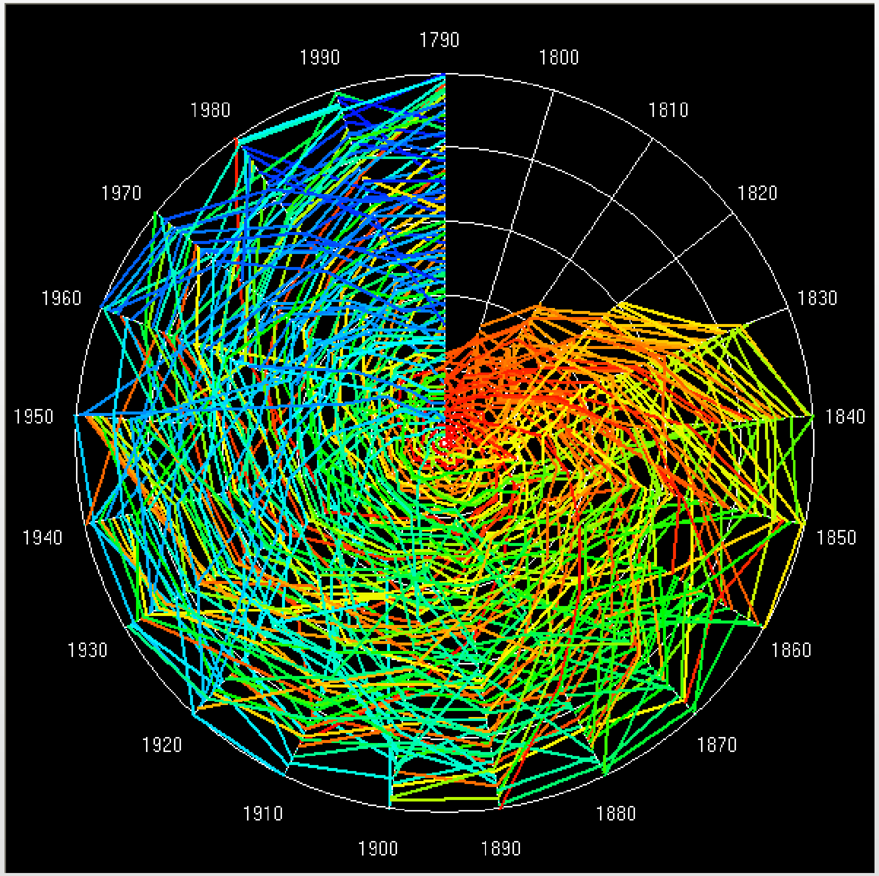
But first how is this clock defined ………



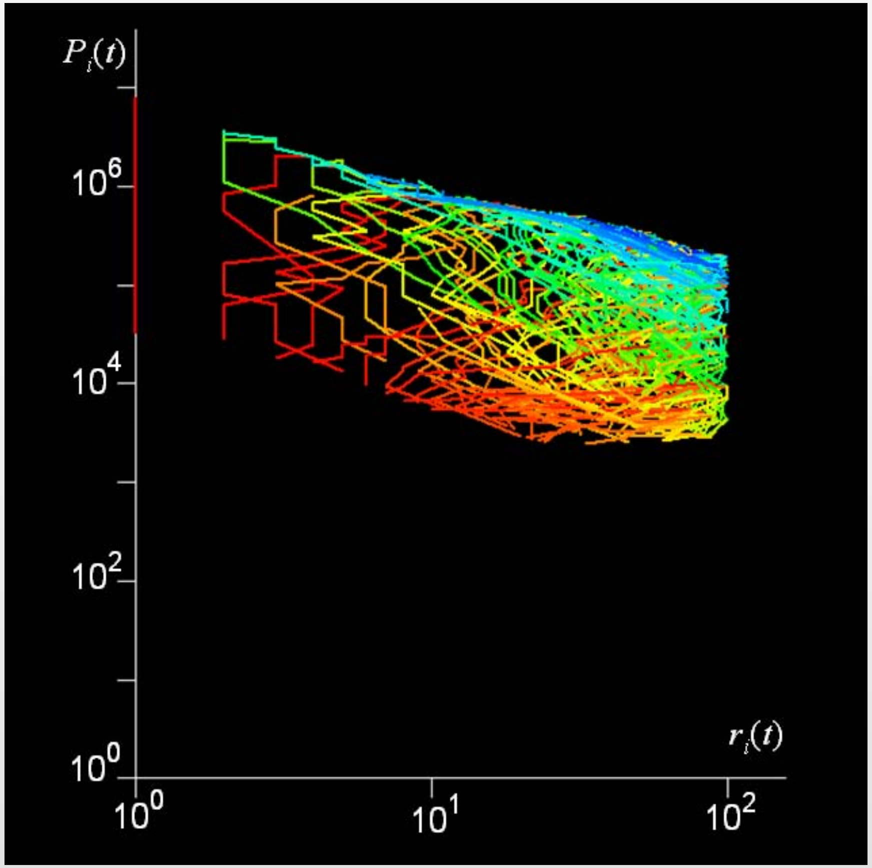
The Idea of a Rank Clock –rank is from number 1 at centre to 100 or whatever maximum is deemed relevant at edge and time goes in years in the usual clockwise direction

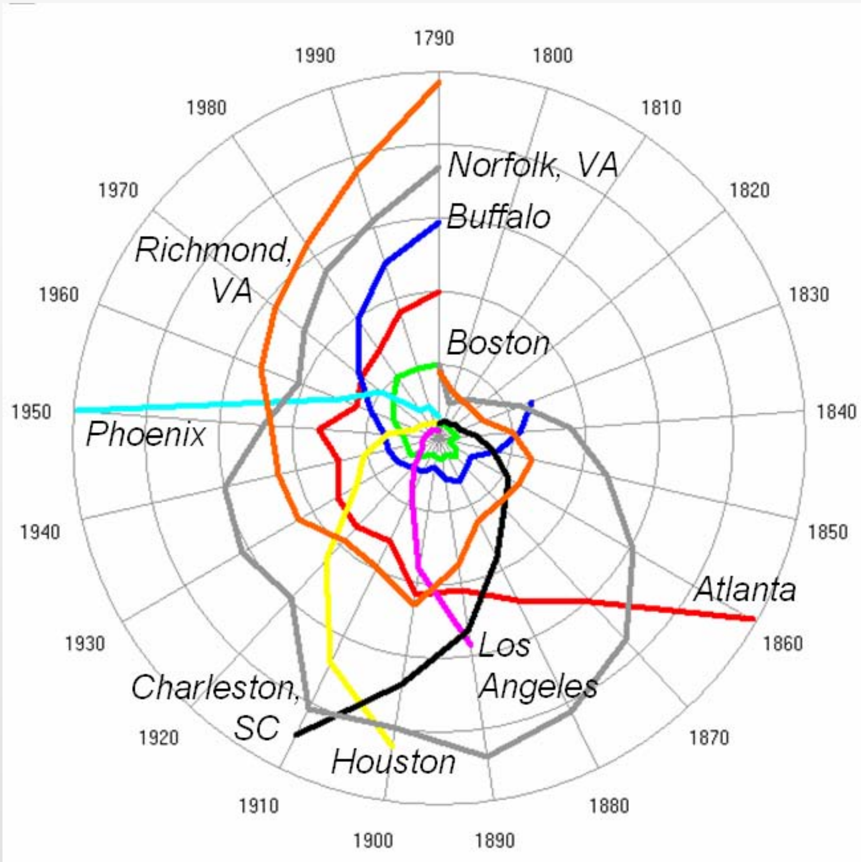
Classic Exemplars: US City Populations 1790-2000



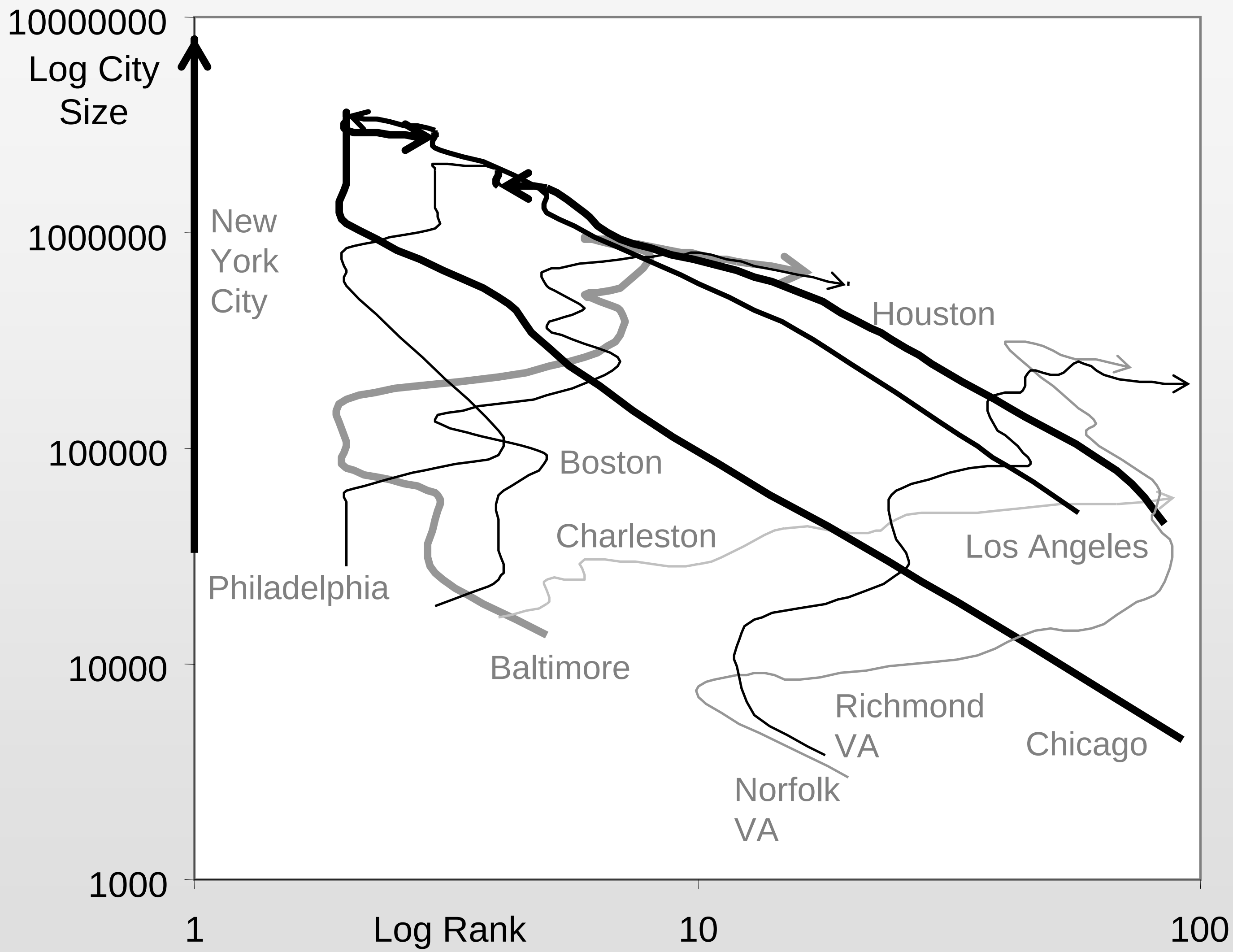


The 'morphology' of the clock should tell us something – i.e. the increase in cities, the volatility of ranks and so on.



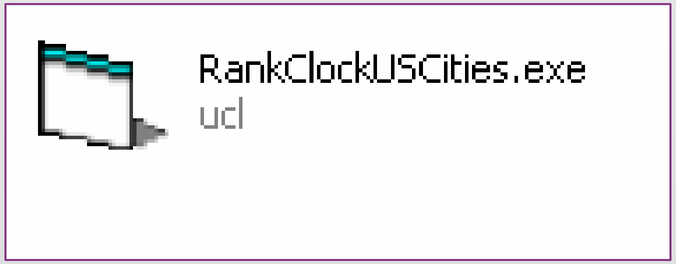


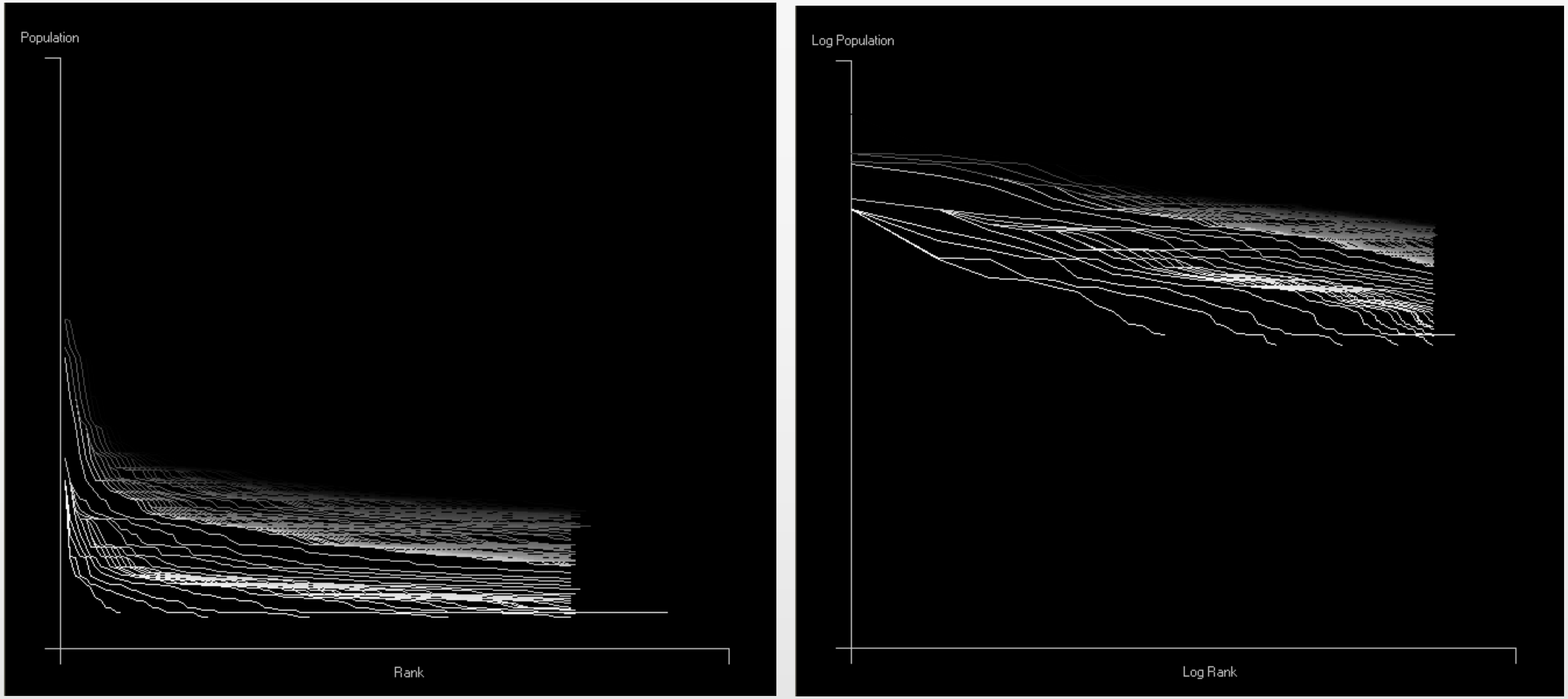
The rudimentary software for this in on our web site at
http://www.casa.ucl.ac.uk/software/rank.asp


10000000
1000000
100000
10000
1000
Log City Size
New York City
Houston
Boston
Charleston
Los Angeles
Philadelphia
Baltimore
Richmond VA
Chicago
Norfolk VA
1
Log Rank
10
100

I will show you an animation of the clock but for a long while I simply plotted the clock and left it at that and measured various properties of the dynamics – but a couple of years ago I decided to animate it and the work that my colleagues Ollie O'Brien and Martin Austwicj are doing comes from the notion that the graphics need to be livened up.

Here is the US city ranks from my desktop software

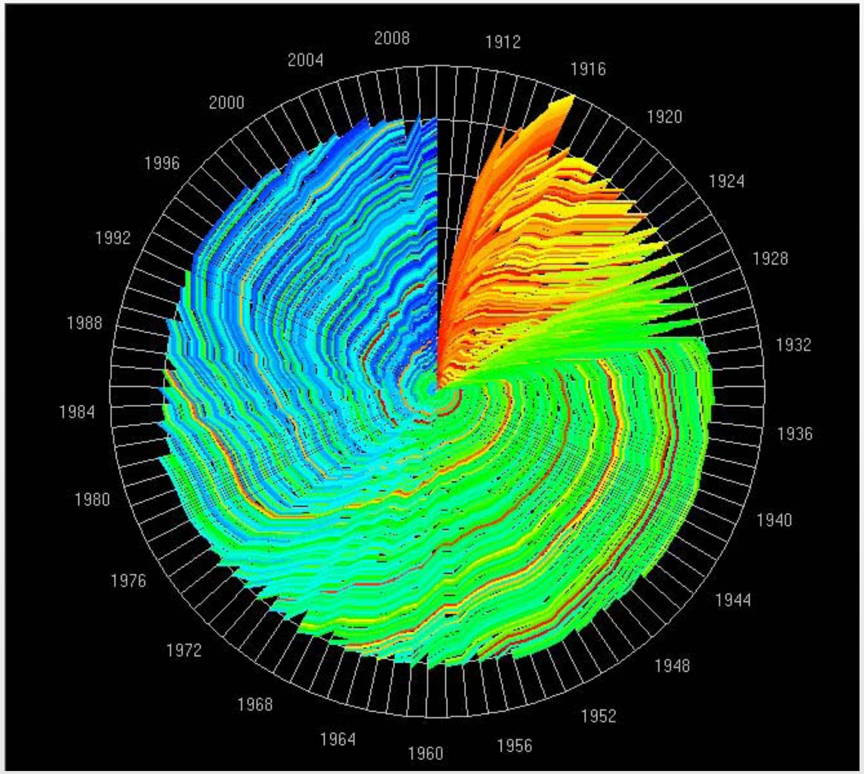


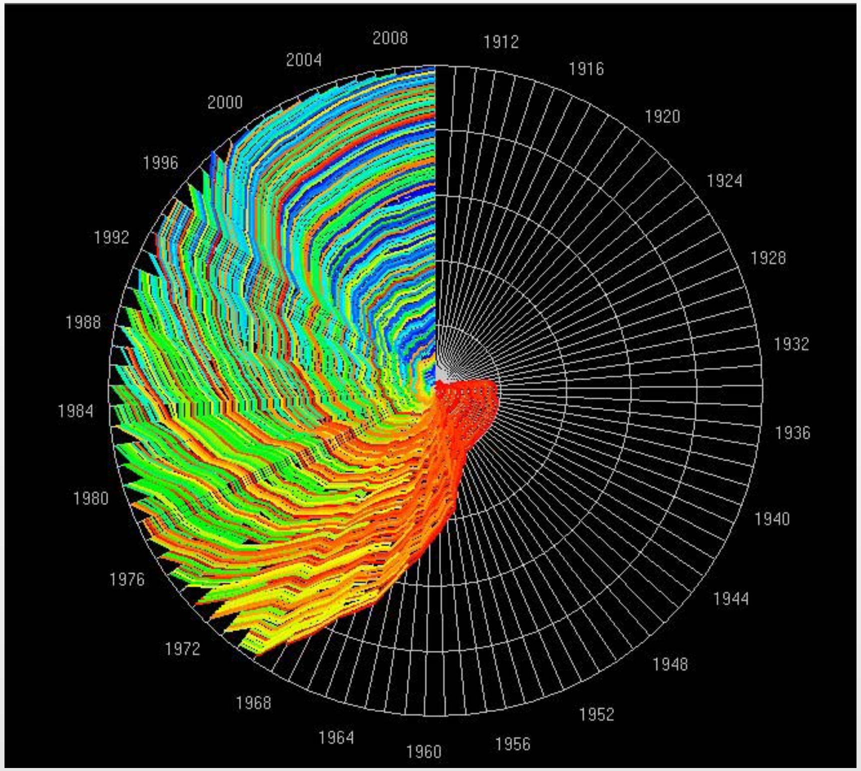


And here is another example of rank size relations for the top 100 high buildings in New York City from 1909 until 2010

power form (left), log form (right) & here is the clock

RankClockNYCitySkyscrapers....
ucl





Rank Clocks of the Top 100 High Buildings in the New York City (a) and the World (b) from 1909 until 2010

There is much more work to do on all this and I am only giving you a taste of this, now back to shape and size

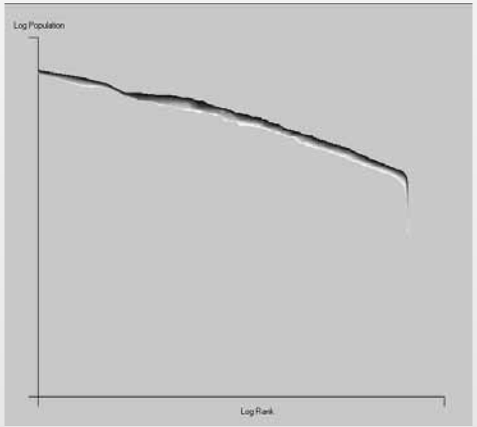
US Metro Area Populations

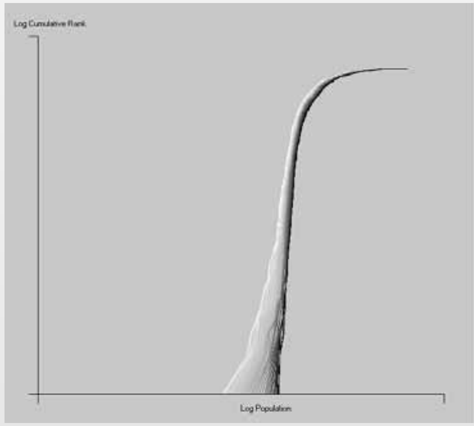
My last example – and we have many now – is from the US Bureau of Economic Analysis on metro areas –SMSAs from which I simply took their 366 regions for which population and income data are available for 37 years from 1969 to 2005.

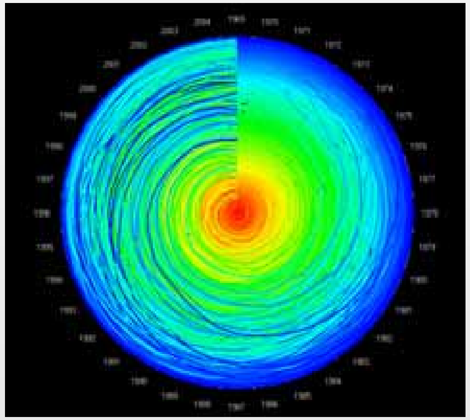
One of the nice things here is that we can plot population ranks etc and income ranks – but we can also look at the ratio – population per capita which is much more volatile than each of the prime variables – let us see

We can easily plot the shifts, spaces, and clocks for these population and income data. These follow very regular scaling laws, at least in their fat tails. Here is a potpourri

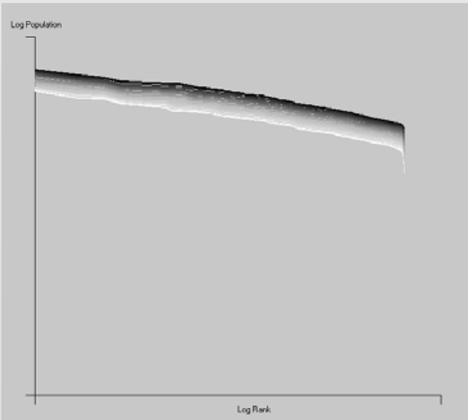
Population

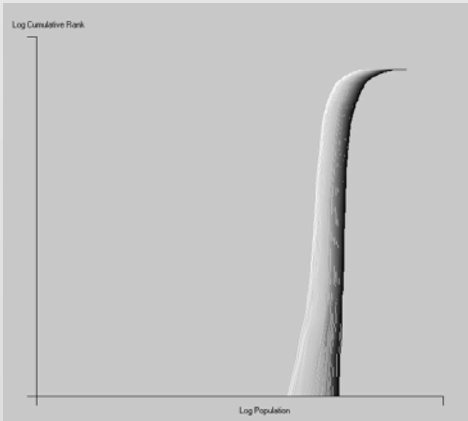


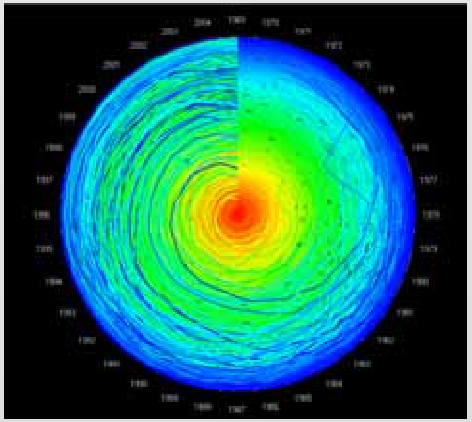




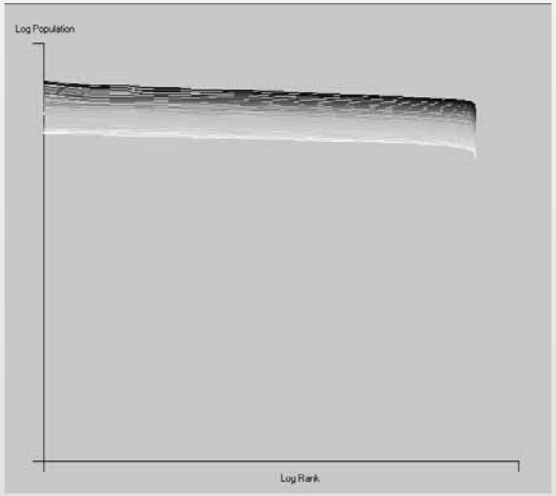
Income

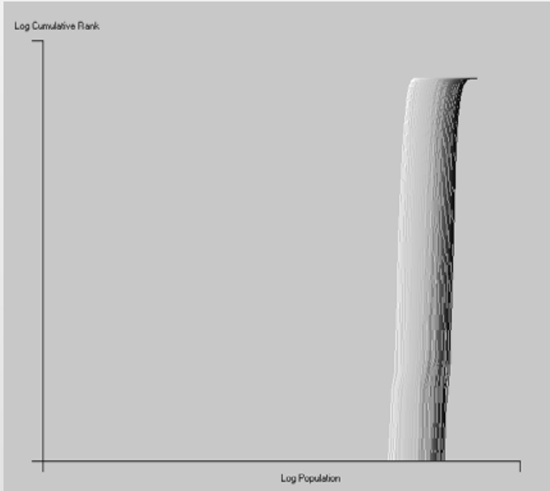


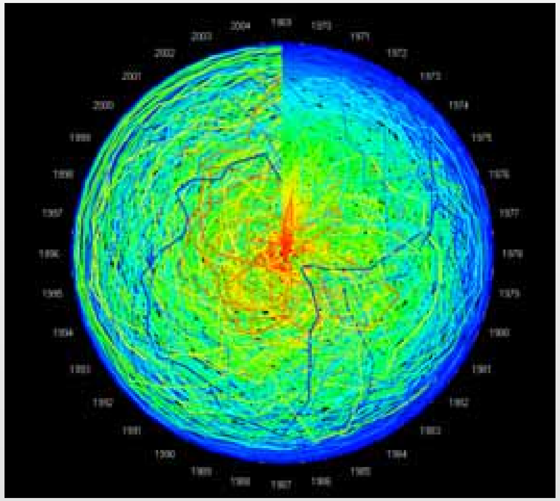




But the real interest is in per capita income/wages – i.e. *Incomes / Population*. How does this rank? And if there are big shifts in rank, this shows divergence of these two variables







As you might expect the rank clock provides a graphic animation of this relative disorder at the micro level

Japanese Populations: Cities at Different Scales

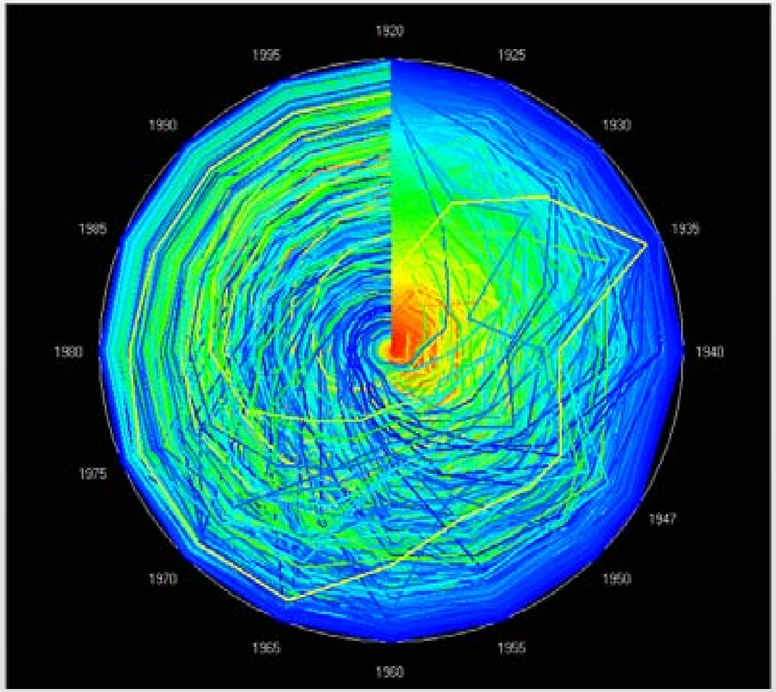
We have a large population data set of 2137 'cities' in Japan that are mutually exclusive subdivisions of the Japanese space and these are not cities in the sense of the US cities we have used.

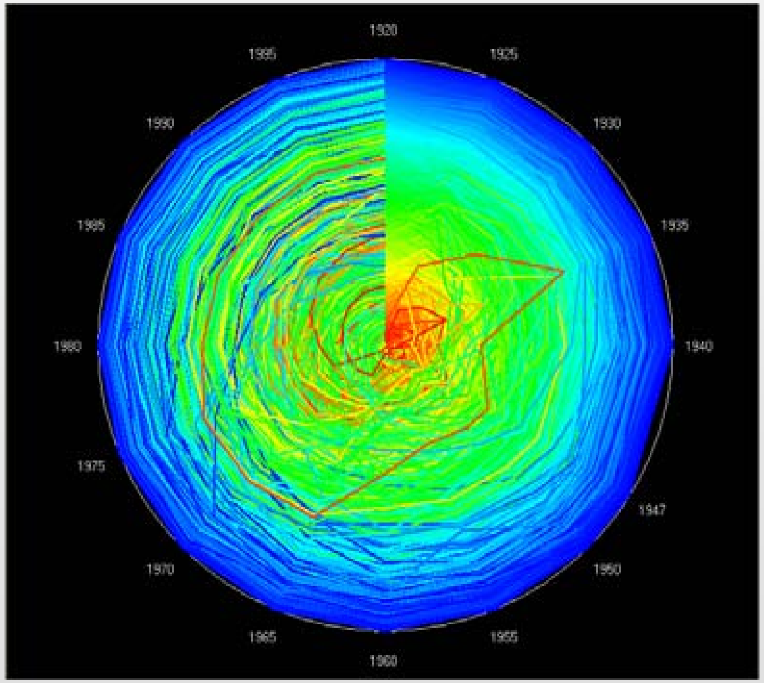
We also have an aggregation into 269 cities at a higher level

We also have other attributes of these cities – with data – such as area so we can compute densities

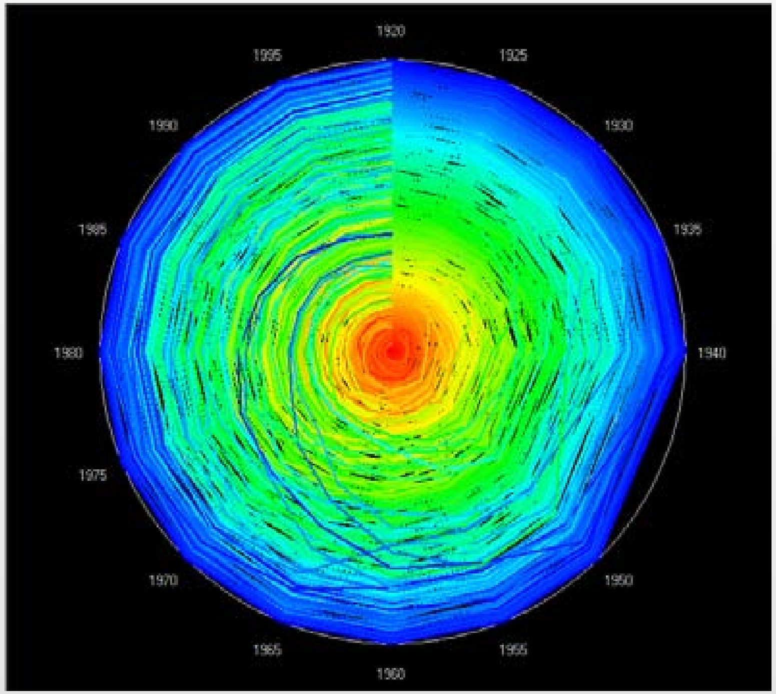
And we have these from 1920 to 2000 in five year periods

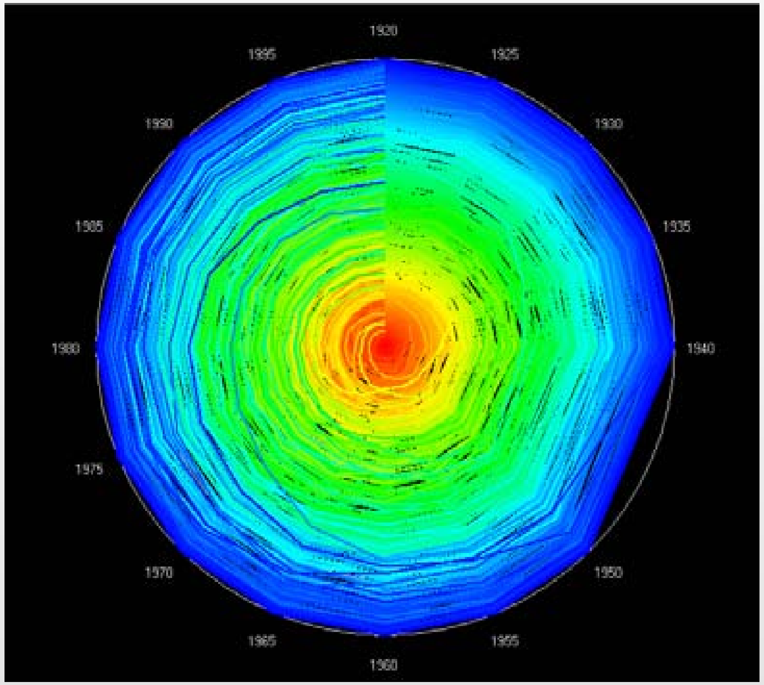
We will look first at the complete set of counts and then densities for 2137 and 269 and then look at Tokyo : first 2137 for counts and densities





then 269 for counts and densities

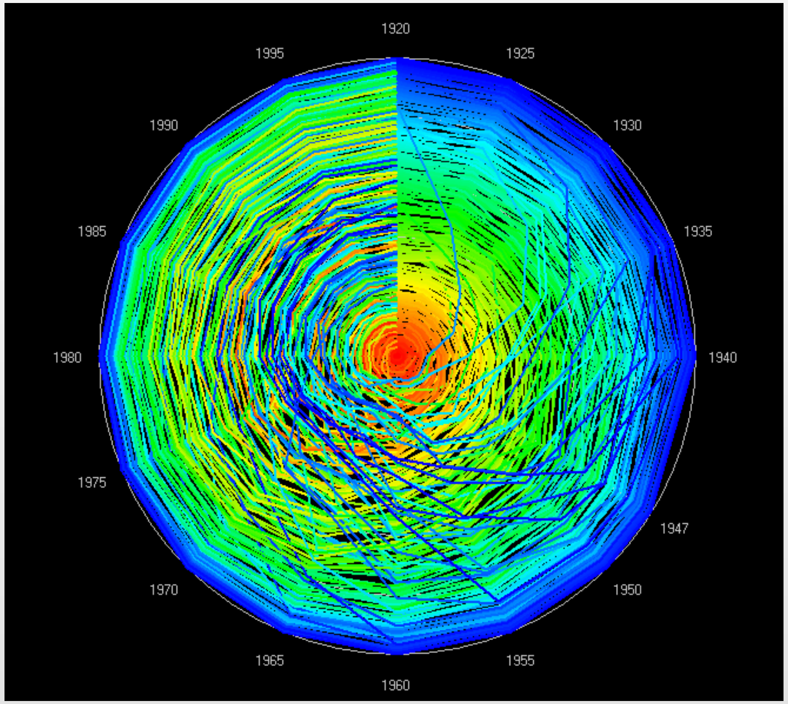


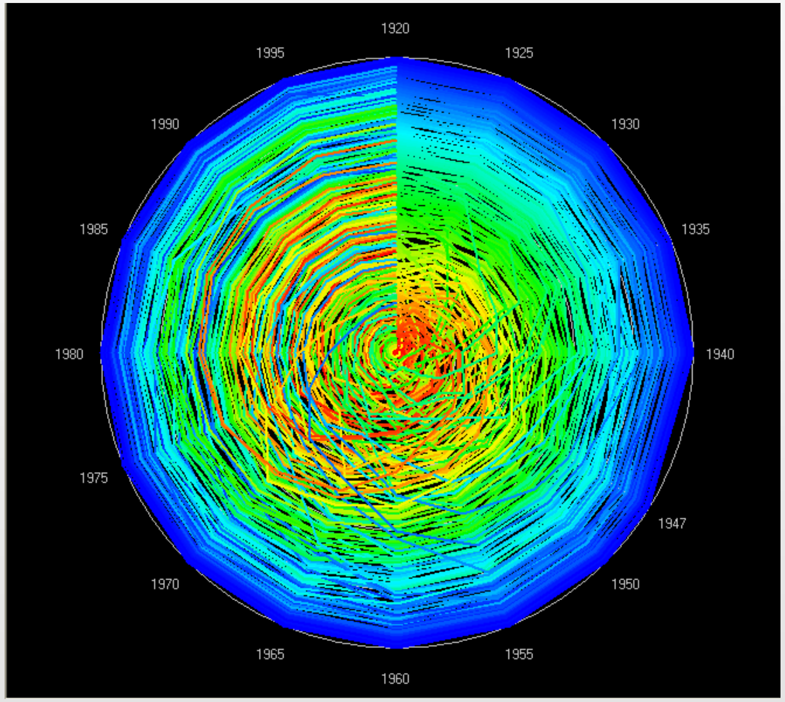


And let us look at the animation of the counts

RankClockJapaneseCities.exe
ucl

then Toyko, counts and densities,





Adding Place to Rank Clocks

Use OpenStreetMap as a base map

Use Google Earth browser plugin for 3D mapping

- Represent each point as an extruded pillar
- Create the pillar size and height based on the extent and average density of data

For simplicity, a uniform distribution of points across a square extent is assumed

Communicate with Google Earth by passing it KML

OpenLayers will readily convert any geometry to KML

Additional Options to Aid Visualisation

Inverting the rank clock

Limiting the size

Colour by:

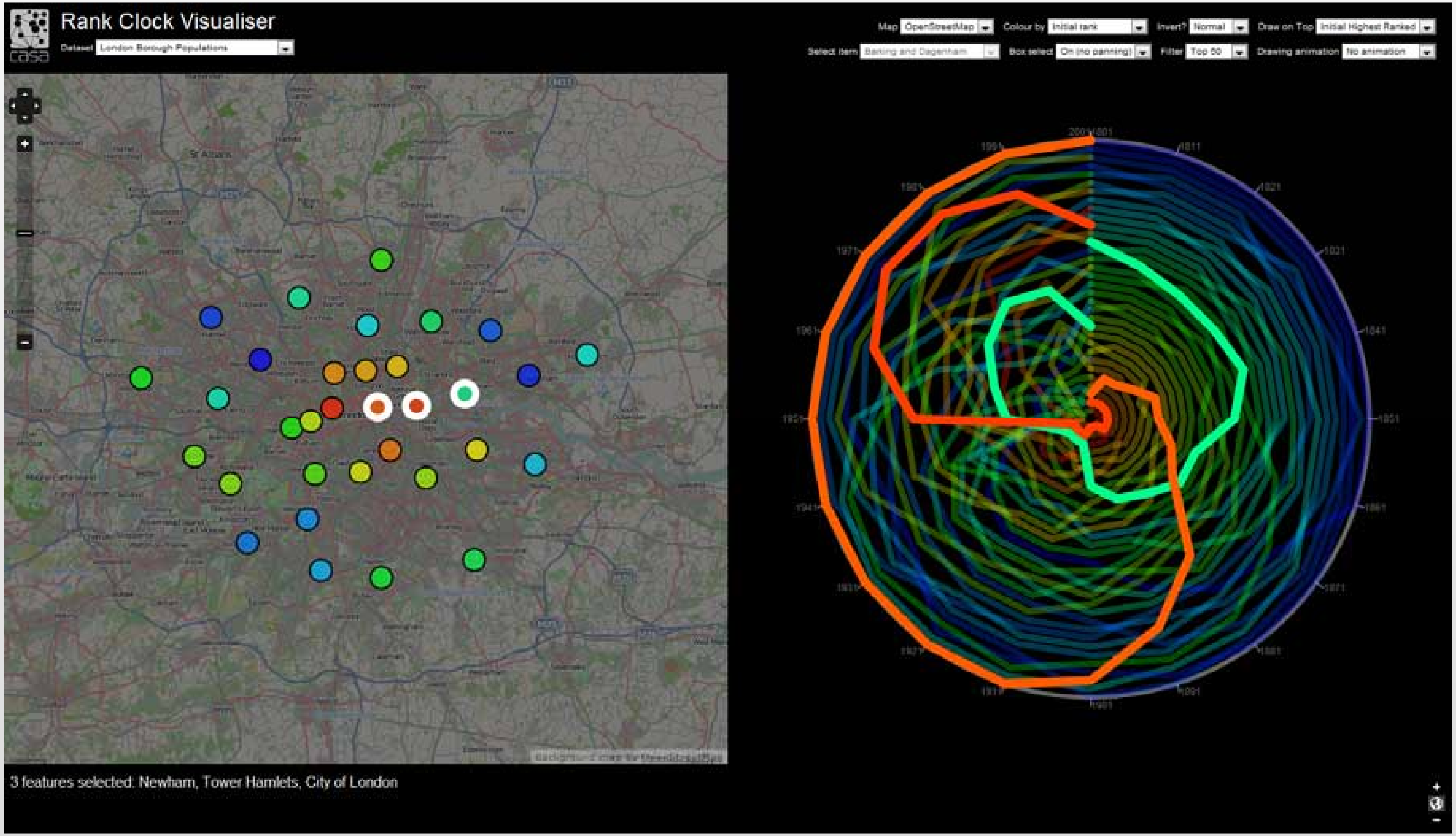
- Initial interval
- Initial rank
- Final rank

Layering the lines (painter's algorithm) by:

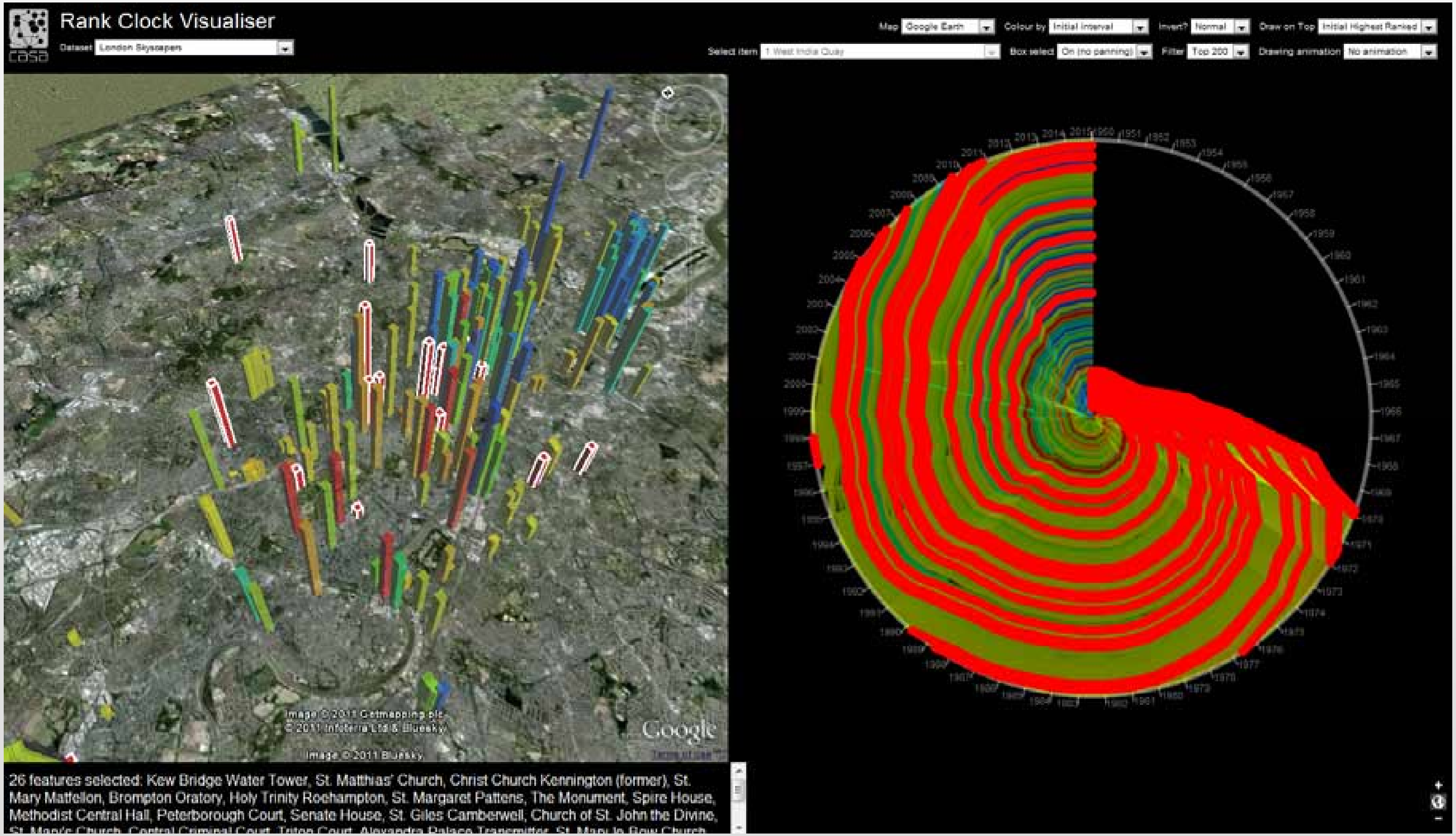
Initial or final highest or lowest ranked

http://casa.oobrien.com/rankclocks/

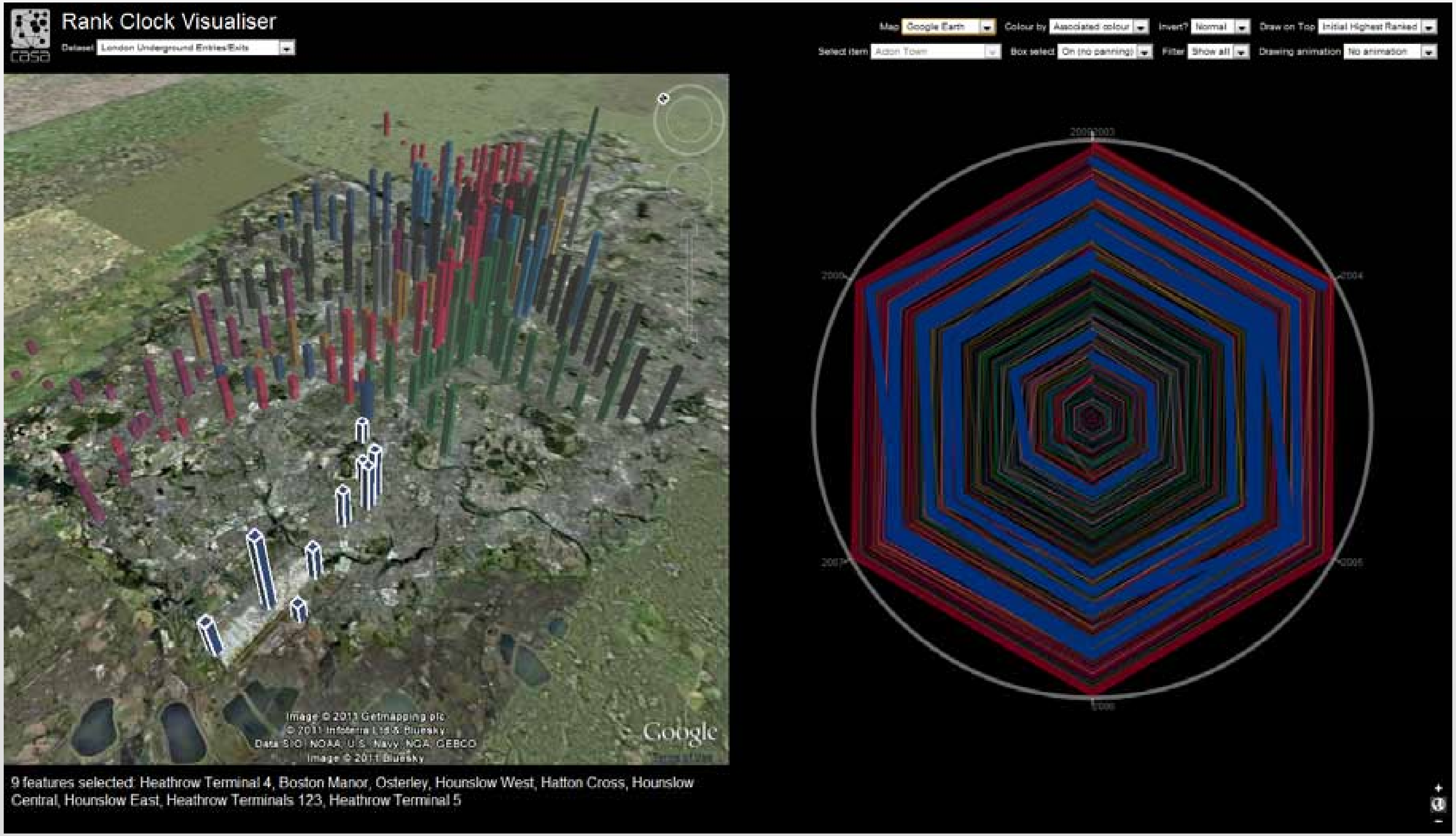
Demo – populations in 33 Greater London Boroughs from 1801



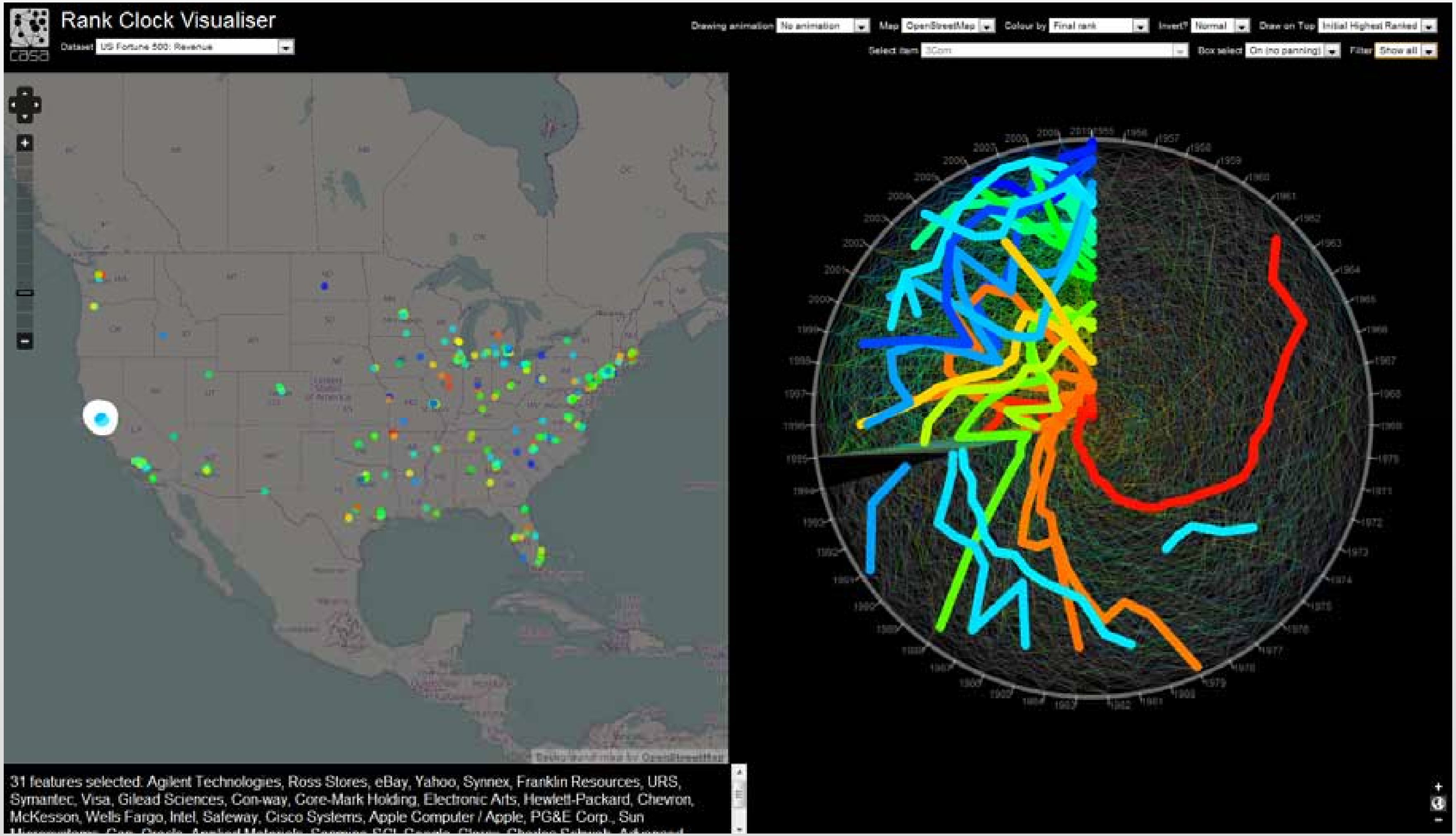
Demo – High buildings in London from 1950 to 2015



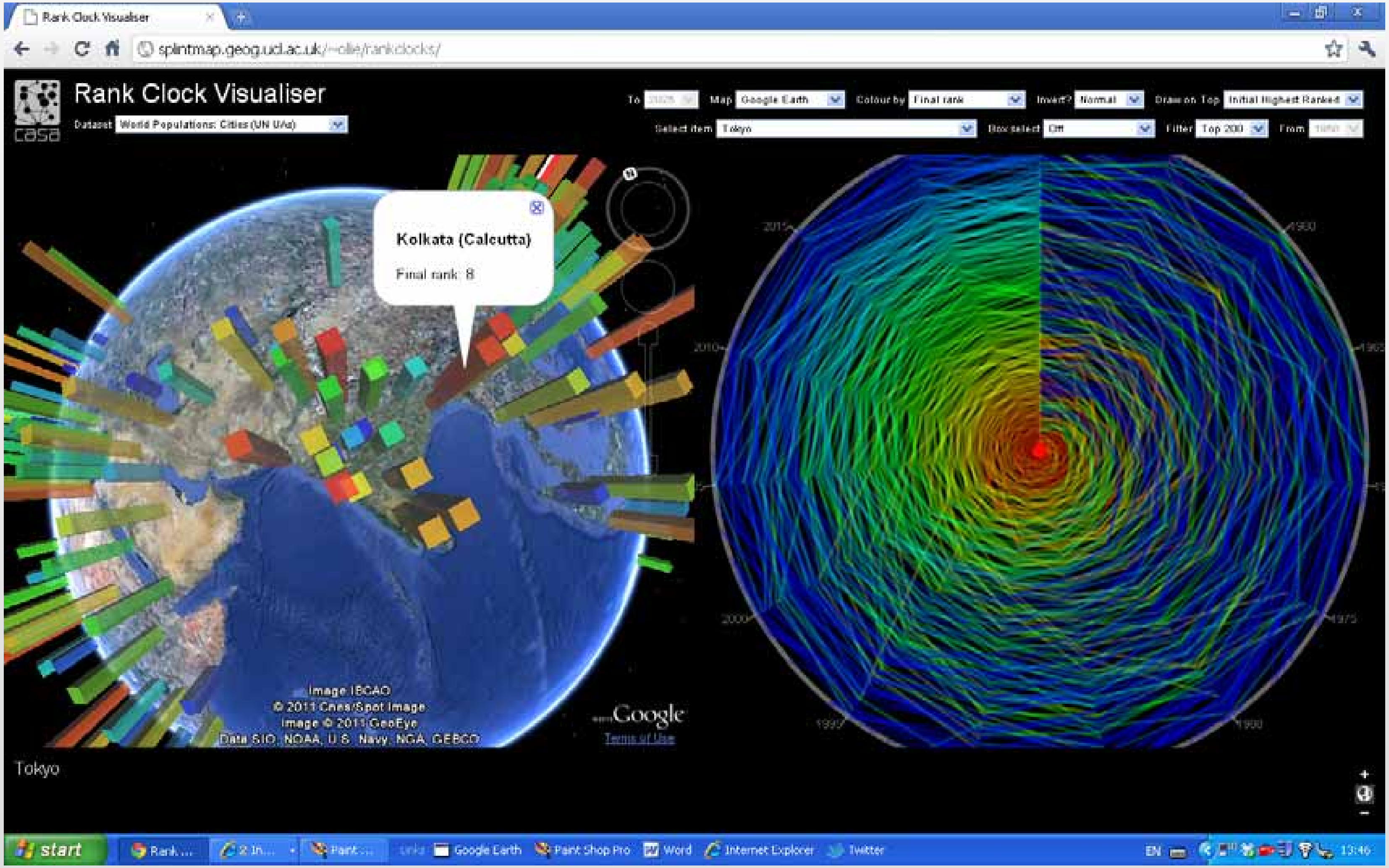
Demo – Tube Exit Volumes from 2003 to 2009

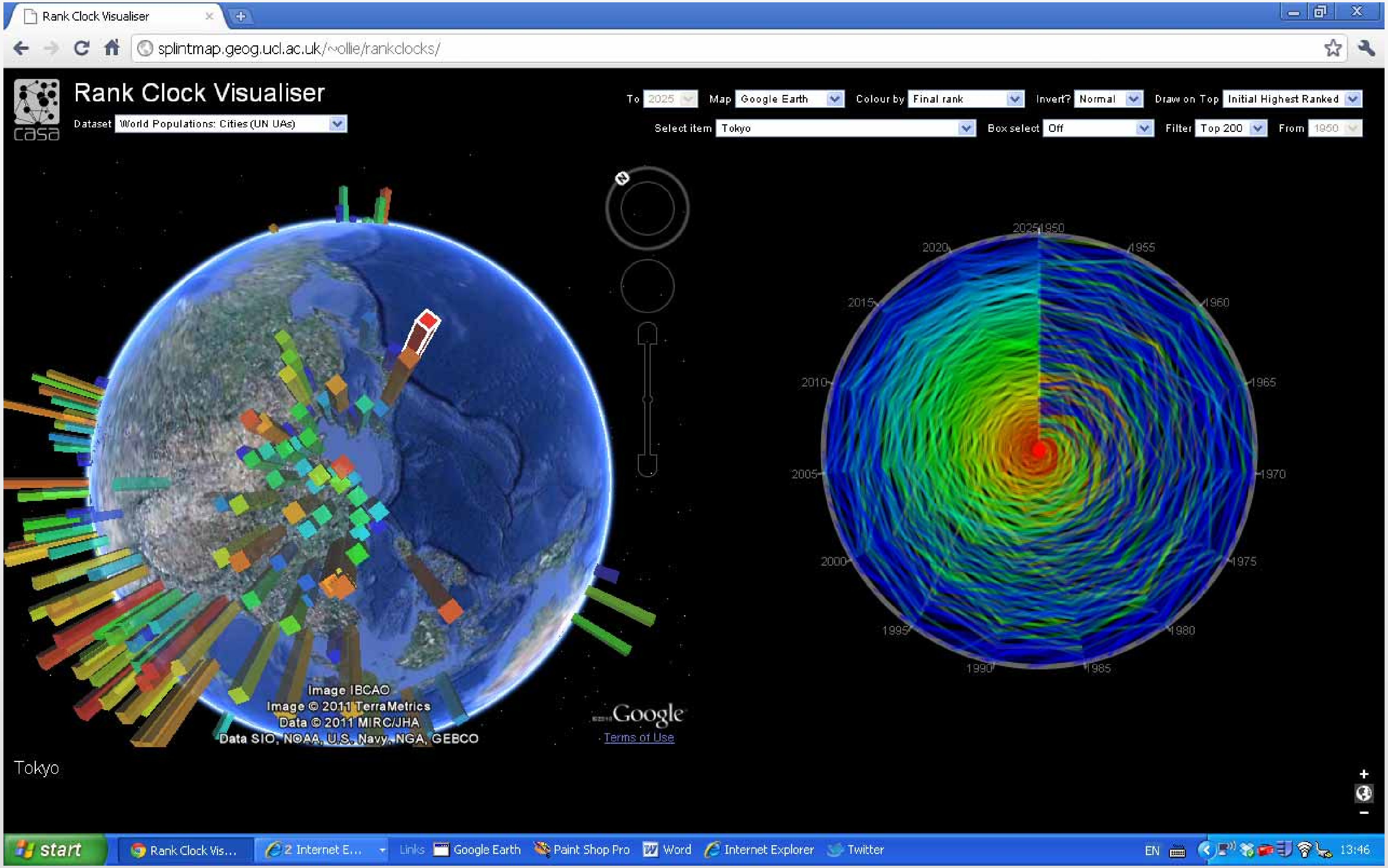


Demo – Fortune 500 from 1955 to 2010 for a sample of fast rising companies



For United Nations World Cities Population (595 cities) from 1950 to 2025



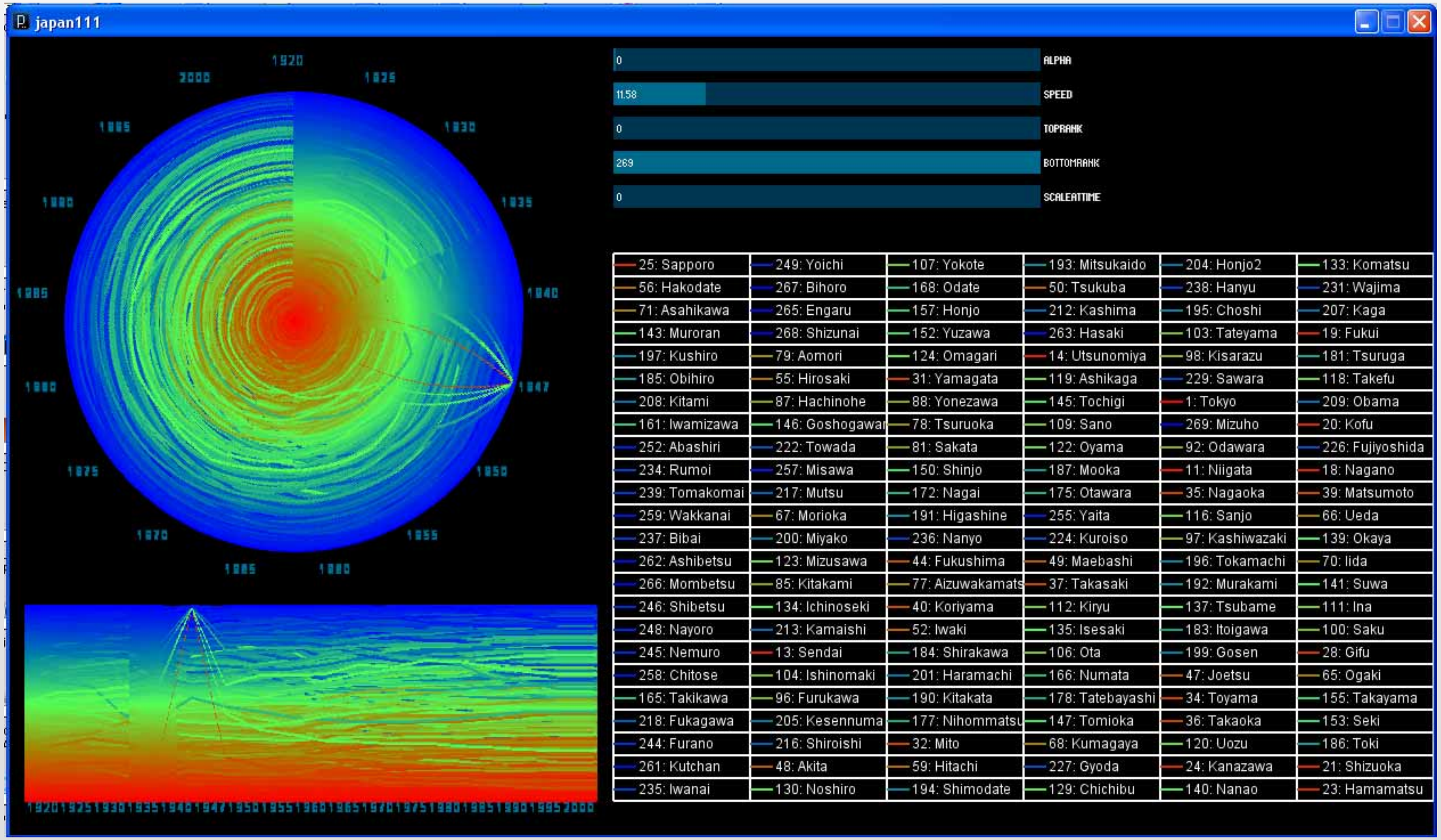


Rank Clock Visualiser
splintmap.geog.ucl.ac.uk/~ollie/rankclocks/
Rank Clock Visualiser
Dataset World Populations: Cities (UN UAs)
To 2025 Map Google Earth Colour by Final rank Invert? Normal Draw on Top Initial Highest Ranked
Select item Tokyo Box select Off Filter Top 200 From 1950
Image IBCAO
Image © 2011 TerraMetrics
Data © 2011 MIRC/JHA
Data SIO, NOAA, U.S. Navy, NGA, GEBCO
Google
Terms of Use
Tokyo
2025 1950
1955
1960
1965
1970
1975
1980
1985
1990
1995
2000
2005
2010
2015
2020
start Rank Clock Vis... 2 Internet E... Links Google Earth Paint Shop Pro Word Internet Explorer Twitter
EN 13:46

Animations: Rank Clocks and Rank Space

- Martin Austwick has developed another set of programs which are superb at animation using the Java based *Processing* language
- The essence of this is to continuously animate the clock and then we can see the deviations
- We can also plot individual trajectories and we can shorten the trajectories to show how each moves relative to one another
- We can do all of this also in rank space but I can't show you this

Here is the typical user interface and I would like to show it to you running ……………… *but* ···►



I don’t have *Processing* on this machine but there is a YouTube clip of what the software does

YouTube

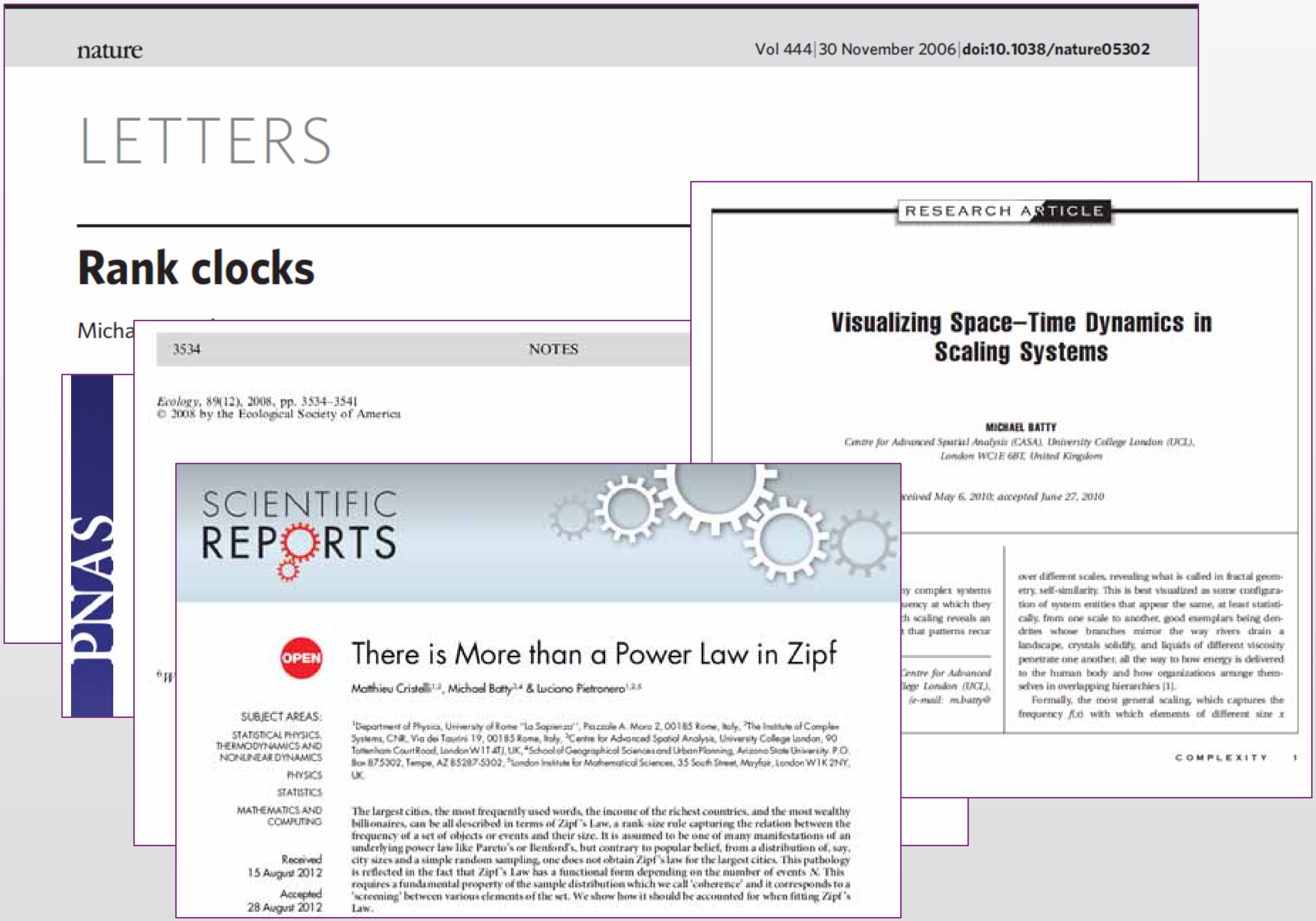
http://sociablephysics.wordpress.com/2011/02/01/rank-clocks-showing-time-as-time/

http://www.youtube.com/watch?feature=player_embedded&v=4V3X8G8Db44

Next Steps

- *Good measures of the volatility*: some of these I have explored but as yet we haven't done anything comprehensive – such as distances measures on the clock – measures of spread such as entropy and so on
- *Extensions to network systems* – trade and migration and traffic flows and their changes over time – in terms of nodal volumes and flows on links
- *Examining different definitions of cities* – and related systems: cities as exhaustive partitions of space versus cities as nodes or points
- *And examining the completeness of cities* and city systems as well as developing lots of ideas about what is being ranked

To finish some papers ………


nature Vol 444 | 30 November 2006 | doi:10.1038/nature05302

LETTERS

Rank clocks

Micha

3534 NOTES

Ecology, 89(12), 2008, pp. 3534–3541
© 2008 by the Ecological Society of America

PNAS

RESEARCH ARTICLE

Visualizing Space–Time Dynamics in Scaling Systems

MICHAEL BATTY
Centre for Advanced Spatial Analysis (CASA), University College London (UCL), London WC1E 6BT, United Kingdom

…ceived May 6, 2010; accepted June 27, 2010

…ny complex systems …uency at which they …h scaling reveals an …t that patterns recur

…Centre for Advanced …llege London (UCL), (e-mail: m.batty@

over different scales, revealing what is called in fractal geometry, self-similarity. This is best visualized as some configuration of system entities that appear the same, at least statistically, from one scale to another, good exemplars being dendrites whose branches mirror the way rivers drain a landscape, crystals solidify, and liquids of different viscosity penetrate one another, all the way to how energy is delivered to the human body and how organizations arrange themselves in overlapping hierarchies [1].

Formally, the most general scaling, which captures the frequency f(x) with which elements of different size x

COMPLEXITY 1

SCIENTIFIC REPORTS

OPEN

There is More than a Power Law in Zipf

Matthieu Cristelli[1,2], Michael Batty[3,4] & Luciano Pietronero[1,2,5]

[1]Department of Physics, University of Rome "La Sapienza", Piazzale A. Moro 2, 00185 Rome, Italy, [2]The Institute of Complex Systems, CNR, Via dei Taurini 19, 00185 Rome, Italy, [3]Centre for Advanced Spatial Analysis, University College London, 90 Tottenham Court Road, London W1T 4TJ, UK, [4]School of Geographical Sciences and Urban Planning, Arizona State University, P.O. Box 875302, Tempe, AZ 85287-5302, [5]London Institute for Mathematical Sciences, 35 South Street, Mayfair, London W1K 2NY, UK.

SUBJECT AREAS:
STATISTICAL PHYSICS, THERMODYNAMICS AND NONLINEAR DYNAMICS
PHYSICS
STATISTICS
MATHEMATICS AND COMPUTING

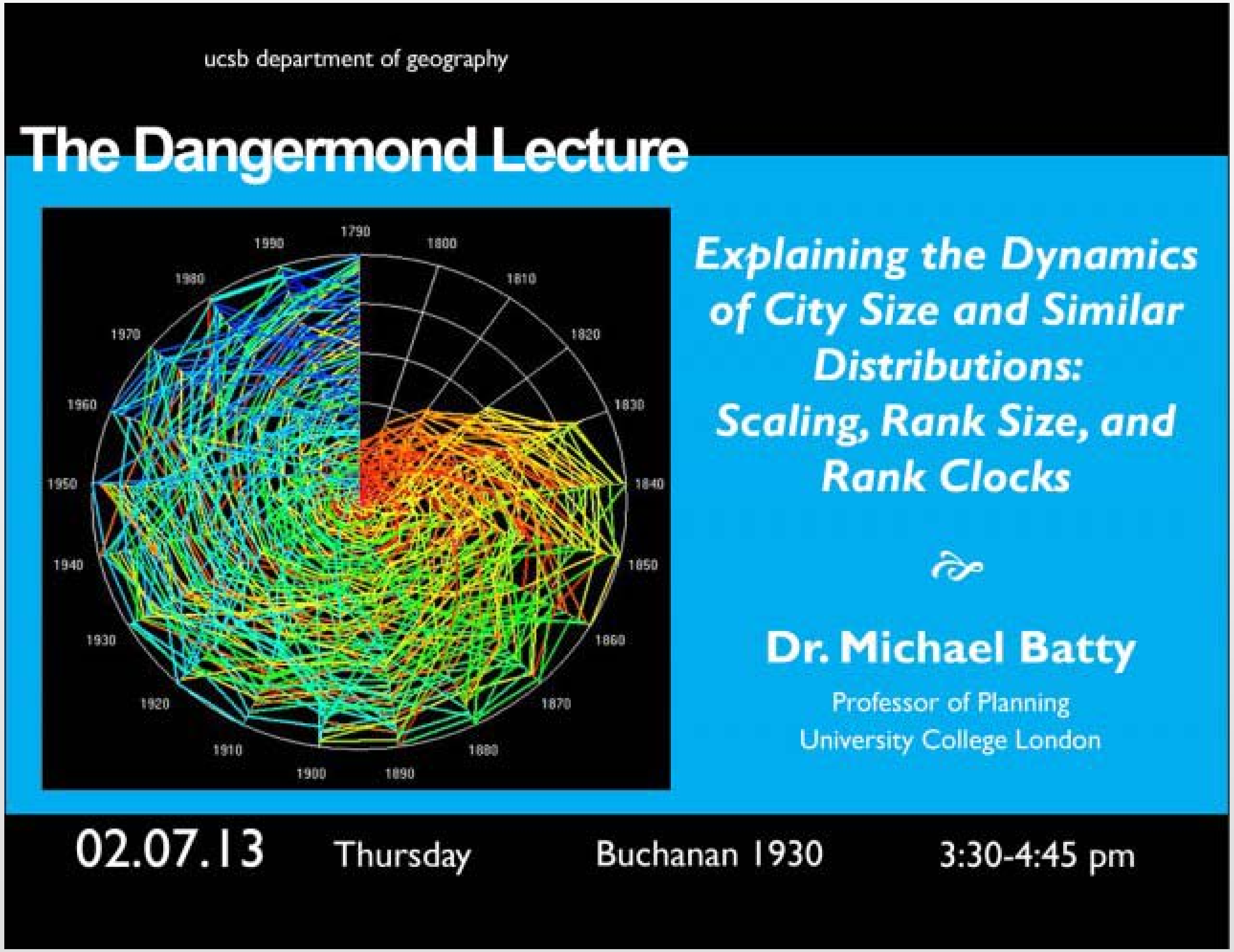
Received 15 August 2012
Accepted 28 August 2012

The largest cities, the most frequently used words, the income of the richest countries, and the most wealthy billionaires, can be all described in terms of Zipf's Law, a rank-size rule capturing the relation between the frequency of a set of objects or events and their size. It is assumed to be one of many manifestations of an underlying power law like Pareto's or Benford's, but contrary to popular belief, from a distribution of, say, city sizes and a simple random sampling, one does not obtain Zipf's law for the largest cities. This pathology is reflected in the fact that Zipf's Law has a functional form depending on the number of events N. This requires a fundamental property of the sample distribution which we call 'coherence' and it corresponds to a 'screening' between various elements of the set. We show how it should be accounted for when fitting Zipf's Law.

Questions?

http://www.casa.ucl.ac.uk/
http://casa.oobrien.com/rankclocks/
http://www.complexcity.info/
http://www.spatialcomplexity.info/

and the PDF of this talk is at


ucsb department of geography
The Dangermond Lecture
1790
1800
1810
1820
1830
1840
1850
1860
1870
1880
1890
1900
1910
1920
1930
1940
1950
1960
1970
1980
1990
Explaining the Dynamics of City Size and Similar Distributions: Scaling, Rank Size, and Rank Clocks
Dr. Michael Batty
Professor of Planning
University College London
02.07.13
Thursday
Buchanan 1930
3:30-4:45 pm