Progress in Human Geography
2018, Vol. 42(4) 610–621

DOI: 10.1177/0309132517717008
journals.sagepub.com/home/phg
SAGE

Quantitative methods III: Scales of measurement in quantitative human geography

Chris Brunsdon

National Centre for Geocomputation, Maynooth University, Ireland

Abstract

Stevens' scales of measurement are often used in texts outlining statistical approaches for geographers. However, it is sometimes overlooked that these are not universally accepted, and indeed the theory surrounding them is contested. This progress report reviews the key ideas of these scales, and discusses a number of the problems they raise – most notably the fact that certain kinds of data are omitted. The value of an axiomatic approach to measurement scales and appropriate statistical techniques is then considered. The report concludes by considering further areas where these ideas may be developed.

Keywords

data type, measurement, statistical testing

I Introduction: What are measurement scales?

The concept of *scales of measurement* (Stevens, 1946) has been used in the physical and social sciences for decades. Essentially, these are a categorization of the kinds of variable (or *attribute* in GIS vector model terminology) that may occur. Identifying the scales of measurement for a given set of variables is intended to provide guidelines for the appropriate method for data analysis – and possibly visualization tools – that can be applied in a meaningful way. This has been an issue that has long contributed to debates in quantitative human geography – since data relating to people, the economy, migration, travel, social well-being and attitudes to place and many other topics contain examples of measurement on all of these scales. It is hoped that understanding the relationship between scales of measurement and data analysis and visualization techniques may provide some guidance when working with diverse forms of quantitative geographical data. I begin by overviewing and highlighting recent interest in this topic, and then review and critique the underlying ideas, and consider fruitful areas for further research.

II Context and review

Scales of measurement have received much attention since Stevens' paper was first published, and they inform research currently ongoing. Since the theory outlines structure in data types that informs choices on how it may be processed, it takes on new relevance in an era of

Corresponding author:

Chris Brunsdon, National Centre for Geocomputation, Iontas Building, Maynooth University, Kildare, Ireland.
Email: christopher.brunsdon@nuim.ie

Table 1. Stevens' original four scales of measurement.

Scale	Basic Empirical Operations	Mathematical Group Structure	Permissible statistics (invariantive)
NOMINAL	Determination of equality	*Permutation group* $x' = f(x) - f(x)$ means any one-to-one substitution.	Number of cases, mode
ORDINAL	Determination of greater or less	*Isotonic group* $x' = f(x) - f(x)$ means any monotonic increasing function	Median, percentiles
INTERVAL	Determination of equality of intervals or differences	*General linear group* $x' = ax + b$	Mean, standard deviation, rank order correlation
RATIO	Determination of equality of ratios	*Similarity group* $x' = ax$	Coefficient of variation

big data and infographics. For example, this has been considered in the framework of creating a unifying ontology of data visualization (Voigt and Polowinski, 2011), which then leads to consideration of data visualization techniques for city models (Métral et al., 2012, 2014), and possibly city dashboards. Big data processing also draws on these ideas, exemplified by Bimonte, Villanova-Oliver, and Gensel (2012) – who consider measurement scale as one factor in developing methods for automatic aggregation of data on web-based servers. Härtwig, Müller, and Bernard (2014) go on to consider this in the context of spatio-temporal data analysis. It is also highlighted as an issue when considering strategies for mapping American Community Survey-based estimates of statistics, and their associated errors (Francis et al., 2015).

In terms of visualization, Harvey (2017) proposes a novel theoretical framework approach, part of which refers to the scales of measurement paradigm for visualizing big data. Finally, in a visualization context, Lin, Hanink, and Cromley (2017) draw attention to its relevance in the context of isopleth mapping. Consideration of measurement scales in the analysis of social and economic data continues to influence thinking. In a recent PhD thesis Jensen (2014) considers the use of compositional data in econometric modelling, and the approach here is informed in part by measurement scale. Also, Erguven (2014) considers influences of measurement theory on statistical analysis of survey data. Thus, the key idea and Stevens' proposed scales continue to be drawn upon as an aid to understanding many current aspects of quantitative spatial data analysis.

The idea, and the proposed scales, need little introduction to readers who have attended a basic 'quantitative methods for geographers' or similar course, but, following Chrisman (1995), it is helpful to list Stevens' initial specification of these scales (Table 1).

These can be interpreted as an ordered set – the 'Basic empirical operations' are essentially cumulative – so that 'Determination of equality' is possible on each scale, 'Determination of greater or less' is possible on all scales except nominal, and so on. The mathematical group structure is essentially a set of transformations which can be applied to measurements in each scale such that the *determination of equality* basic empirical operation associated with each scale will still hold. Again, these form a sequence since each group transformation is a subset of the previous one. For example, $x' = ax + b$ is a monotonic increasing function, provided $a > 0$. It should be pointed out that the further condition $a > 0$ should be added to the mathematical group structure entries for interval and ratio. Mapping *all* x values onto a

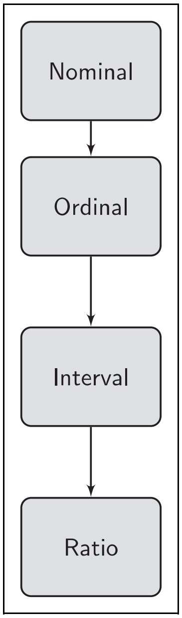

Figure 1. Scales of measurement – as suggested by Stevens.

constant *b* (interval) or zero (ratio), as would be the case if *a* were zero, does not preserve the respective determinations of equality for these scales. Negative values reverse the ordering. I refer to this group of four scales as the NOIR group (from the initial letters of Stevens' original nomenclature). Their interconnection is illustrated in Figure 1.

Around this set of scales of measurement, Stevens develops the idea of *measurement theory*. This argues that the scale of measurement adopted dictates the method of statistical analysis that should be used, in terms of either statistical tests or descriptive statistics. It also sets out a rigorous definition of a *scale of measurement* as a mapping of some quality (abstract or physical) onto a numeric value. Examples include a physical distance to a number of meters or feet, or a response to a certain government policy on to an attitude scale (for a rigorous exposition see Krantz et al., 1971; Suppes et al., 1989; Luce et al., 1990). These ideas persist in many ways at the time of writing. Many statistical textbooks contain some version of Stevens' table as a guide to choosing appropriate statistical methods given certain kinds of data, essentially basing their guide on the NOIR scales. Some software packages use this framework to classify variables. This could involve 'greying out' certain statistical procedures that would be inappropriate for certain variables according to Stevens' dictum. For example, defining variables in SPSS requires variables to be classified into the near-NOIR grouping of 'nominal', 'ordinal' or 'scale'. The last, although confusingly named, is a union of interval and ratio scale data.

Although the NOIR grouping is perhaps the first attempt to classify scales of measurement in this way, and arguably the first to propose this concept in a formal framework, other commentators note that it is by no means complete. The NOIR scales of measurement themselves form an ordinal scale in that the list of group structure transforms as defined above form a cumulative set – and the notion of comparison $x < y$ in this context can be defined in terms of subsets of permissible operations. This self-reference may be seen as aesthetically satisfying, but unfortunately the NOIR set of scales is not exhaustive and some non-NOIR scales do not fit well into this ordered framework. Chrisman (1995) notes that circular data do not find an intuitive place here: 'Angles seem to be ratio, in the sense that there is a zero and an arbitrary unit (degrees, grads or radians). However, angles repeat the cycle. The direction 359° is as far from 0° as 1° is' (p. 274). A similar argument could be applied to time-of-day or time-of-year data.

Others have also identified kinds of measurement that do not fit comfortably into this framework, including Stevens himself (1959), who subsequently argued for a fifth scale (sitting at the same level as interval) for logarithmic measures.

Other attempts to identify scales of measurement also exist. Some of these are simpler than Stevens', and some more complex. Indeed Van den Berg (1991) found that the identification of distinct scales of measurement was an area of notable disagreement among statistical experts. Nelder (1990) identified various *modes* of data, comprising continuous counts, continuous ratios, count ratios and categorical. The last of these are sub-classified into 'nominal', ordered on the basis of an underlying scale, and ordered without an underlying scale. Here there are more than the four levels proposed by Stevens, although arguably these do not fit with his overall idea that level of measurement dictates statistical technique.

Although many sources still use Stevens' initial classification as a basis for structuring statistical analysis (despite the fact that Stevens himself has amended this), the ideas are contested. The contestation falls broadly into two groupings: those who support measurement theory but disagree in some way with Stevens' approach, and those who oppose measurement theory per se.

Below, I review both critiques and offer further suggestions. I proceed by outlining some of the basic ideas in greater detail, then focusing on the critiques, and finally reflecting on these. I consider levels outwith Stevens' original list, and place them within the 'Mathematical Group Structure' framework suggested in that list. In particular, types of spatial data and measurement scales will be considered in this framework as well as the concept of *permissible statistics* applied in this context.

III Detailed view: 'Permissible' statistical operations for the NOIR scales

In addition to outlining scales of measurement Stevens suggested appropriate statistical tests for each of the levels, termed *permissible*. These are tests whose outcomes and interpretation remain unaltered when the mathematical group structure transform is applied to the data. For example, two-sample t-tests are not permissible for nominal level data, even if the nominal levels are denoted numerically. Suppose we had two samples, uniquely labelled numerically, but thought to be nominal, as in Table 2.

Table 2. s_1 and s_2.

s_1	6	7	8	9	10	11
s_2	1	2	3	4	5	

Table 3. Two-sample t-test: s_1 and s_2.

t - Statistic	5.20
p-value (two-tailed)	<0.001

Table 4. $f(s_1)$ and $f(s_2)$.

$f(s_1)$	6	7	8	3	2	1
$f(s_2)$	11	10	9	4	5	

Table 5. Two-sample t-test: $f(s_1)$ and $f(s_2)$.

t - Statistic	-1.82
p-value (two-tailed)	0.10

Applying a t-test to these values has the following result, rejecting a two-tailed test of $H_0 : \mu_1 = \mu_2$ at the 5% significance level (Table 3).

However, if the data really are nominal, the results should be invariant to a transform $x \mapsto f(x)$ provided $f(x)$ is unique for any x value. However, consider the transform

$$f(x) = \begin{cases} x \text{ if } 4 \leq x \leq 8; \\ 12 - x \text{ otherwise} \end{cases}$$

This function 'reverses' the integers between four and eight, leaving the rest unchanged, and meets the uniqueness criterion. The transformed data are shown in Table 4.

And the test performed on the transformed data gives the result in Table 5.

In this case the test fails to reject H_0. Even if the test *outcome* were the same the test statistic would have been altered – the p-values would be altered. This example illustrates Stevens' principle: the transform above *should* leave the outcome invariant at the nominal level, but it actually changes the t-statistic. Therefore t-tests are not meaningful for this level of measurement. In Stevens' terminology, the two-sample t-test is 'not permissible' for nominal level data. Conversely, a transform of the kind $ax + b$ or ax would not change the value of the test statistic or the p value, so two-tailed t-tests *are* permissible for interval or ratio levels.

Similarly, 'permissible statistics' listed in Stevens' table are essentially statistics that are consistent with the appropriate group structure transform, in the sense that if $\{x_1, \ldots, x_n\}$ are a set of measurements on a given level, $f(x)$ is a group structure transform for this level, and $x^* = g(x_1, \cdots, x_n)$ is a permissible statistic then $f(x^*) = g\Big(f(x_1), \cdots, f(x_n)\Big)$. If the measurement scale is somehow 'recalibrated' via $f(x)$, then the recalibrated summary statistic should be the summary statistic of the recalibrated measurements. However, this is not always strictly the case, and Stevens' theoretical overview is 'informal' in some parts. Although standard deviations are permissible for interval scale data, the group structural transform here is $f(x) \mapsto ax + b$, but if the standard deviation of a sample is s, then the standard deviation of the transformed data is as, and not $as + b$.

In addition to the idea of a 'mathematical group structure' and the associated 'basic empirical operations', another characteristic of levels of measurement is that of a *meaningful binary operator*. For example, for nominal level variables x and y the binary operators $x = y$ and $x \neq y$ are meaningful, as they have logical outcomes of *true* or *false*. However, $x \geq y$ is not meaningful, since there is no concept of order in nominal variables. Other levels of measurement will have different meaningful operators.

IV Scales of measurement outwith the basic theory

Since others have identified levels of measurement outwith Stevens' original table, how do these fit in? For example, are there group structure transforms and associated permissible statistics for angular data, or log interval data and if so, what are they?

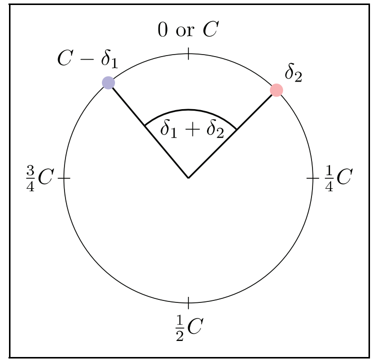

Figure 2. The cyclic scale of measurement.

1 The 'cyclic' scale

The cyclic scale is the generic name by which all scales of the kind given in the example of angles above will be designated. Typical uses in human geography might include analysis of directions of movement (Faggian et al., 2013; Brunsdon and Charlton, 2006), the time of day of crimes (Brunsdon and Corcoran, 2006), or crowd movement dynamics (Wirz et al., 2012). The idea is that they range over a set of numbers from 0 to C but that the distance between 0 and C is zero, and for $\delta_1, \delta_2 \in [0, C/4]$ the difference between $C - \delta_1$ and δ_2 is $\delta_1 + \delta_2$. In other situations the difference is defined by ordinary subtraction. Thus, for angles, $C = 360°$ and the difference between $359°$ and $1°$ is $2°$, x may be thought of as a point of the circumference of a circle (see Figure 2).

The set of group structure transformations take the form $x \mapsto f(x)$ where f maps x onto a revised scale with a possibly different value for C. An offset can be added to the transformed x so that the zero position on the circle is altered. A switch from 'clockwise' to 'anticlockwise' is also possible so that $\frac{3}{4}C$ and $\frac{1}{4}C$ are transposed in the diagram. These operations occur when

Table 6. Average pace and speed for two pairs of running shoes for repeated 10 km runs.

Shoes	Measurement	Individual Runs									
A	Pace Min/Km	6.45	6.43	5.80	5.93	6.08	6.37	6.64	6.30	6.61	6.06
	Speed Km/h	9.30	9.32	10.34	10.12	9.87	9.42	9.04	9.53	9.08	9.90
B	Pace Min/Km	6.46	6.42	6.45	6.47	6.37	6.68	7.00	6.73	6.17	6.36
	Speed Km/h	9.28	9.35	9.30	9.28	9.41	8.98	8.57	8.92	9.72	9.43

switching between degrees and radians, where C changes from 360 to 2π, the zero position moves from 12 o'clock to 3 o'clock, and direction switches from clockwise to anti-clockwise. Stevens-type permissible statistics are outlined in Fisher (1993). Averaging may be exemplified by the *circular mean* $\tilde{x}$ - defined by:

$$\tilde{x} = \tan^{-1}\left(\sum_i \sin(x_i), \sum_i \cos(x_i)\right)$$

for x in radians, where the bivariate form $\tan^{-1}(y,x)$ computes the arctangent, taking into account the correct quadrant given the signs of the argument. For positive x and y, $\tan^{-1}(y,x) = \tan^{-1}(y/x)$. If x is not in radians, it is converted to radians for the calculation of this statistic. The result is converted back to the initial units afterwards. For spread, typical statistics include the *circular standard deviation*

$$\nu = \sqrt{-\ln\left(\left(\frac{1}{n}\sum_i \sin(x_i)\right)^2 + \left(\frac{1}{n}\sum_i \cos(x_i)\right)^2\right)}$$

where again x is in radians, and converted to this scale otherwise.

There is no well-defined ordering for this scale of measurement, due to its cyclic nature, so order-based statistics (such as quantiles) have no meaning. Effectively, in relation to Figure 1, the cyclic scale forms a new branch away from the existing hierarchy at the nominal level, since operations applicable to the nominal scale (such as the mode) may be applied here, as may other operations unique to cyclic measurements. However, operations that are valid for ordinal measurements are not valid for the cyclic case.

Table 7. Two-sample *t*-test for null hypothesis of equal pace.

t - Statistic	−2.1000
p-value (two-tailed)	0.0501

Table 8. Two-sample *t*-test for null hypothesis of equal speed.

t - Statistic	−2.1220
p-value (two-tailed)	0.0480

2 The 'log interval' scale

Suppose I have two pairs of running shoes (denoted A and B here). I take ten 10 km runs in each pair and measure the average pace of each run (in minutes per km – see data in Table 6). I wish to test the hypothesis that the average running pace in both pairs of shoes is the same.

In each run the speed is calculated in km/h (pace divided into 60). If we assume the pace to be measured on a ratio scale (it has a well-defined zero of being stationary!), a t-test may be used to test this hypothesis since, in Stevens' terms, it is permissible. Here insufficient evidence is found to reject this hypothesis, using a two-tailed test, as seen in Table 7.

Now consider the same test, using speed instead of pace (Table 8). This time there is evidence to reject the null hypothesis.

It could be argued that if there is no strong reason to favour one hypothesis over the other; then a valid group structure transform is $f(x) \mapsto 60/x$, which is non-linear, and hence rapidity is measured on an *ordinal* scale. This

Table 9. Two-sample Wilcoxon test for null hypothesis of equal speed.

w - Statistic	75
p-value (two-tailed)	0.063

Table 10. Two-sample logged t-test for null hypothesis of equal speed.

t - Statistic	−2.1120
p-value (two-tailed)	0.0489

being the case, the result of the hypothesis test may be based on a *Wilcoxon test* which compares the *ranks* of the speed (or the pace) for each pair of shoes, which gives the results in Table 9.

Now the p-value implies the null hypothesis fails to be rejected. Furthermore, the p-value itself is notably increased, suggesting the power of the test has diminished.

An alternative approach might be to transform by taking logs (of both pace and speed), noting that $\log(\text{Pace}) = \log(60) - \log(\text{Speed})$. The logged quantities are related by a linear transform, and form an interval scale. In this case a t-test *is* permissible in both cases. For pace, see the result in Table 10.

Applying the same test to logged speed measurements gives t=2.112, which is equivalent to the above (except for a change in sign) with identical p-value due to the use of a two-tailed test.

This suggests the existence of another scale of measurement, whose group structure transform is $\log(x) \mapsto a\log(x) + b$. Given that $x > 0$, this may be stated as $x \mapsto bx^a$ (for a different b). Stevens proposed this in a later paper. For this scale the set of group structure transforms is a subset of those for ordinal scales of measurement but not a superset of those for interval or ratio scale. These form a 'parallel' scale to the interval scale – superseding the ordinal scale but incomparable to interval and ratio scales.

3 The 'absolute' scale

In 'absolute' scale, no transform apart from the identity transform $f(x) \mapsto x$ is permitted. This typically applies to count data. Counts of objects cannot be multiplied by an arbitrary linear transform and maintain their meaning. Like ratio data, they clearly have a well-defined zero. It could arguably also include unobservable attributes, such as probability (which, in order for the laws of probability to hold, must be in the range $[0, 1]$) or correlation – which must lie in the range $[-1, 1]$. This scale can be distinguished from ratio scale by the absence of units of measurement.

4 An augmented set of scales

Putting the extra scales together with those put forward by Stevens prompts re-drawing of the diagram in Figure 1, giving the updated arrangement shown in Figure 3, where arrows with dotted shafts show relationships involving the scales augmenting the original classification. The scales themselves are shaded in a lighter colour. The simple 'ladder' relationship seen earlier is no longer apparent.

V Further issues with measurement theory and NOIR

Criticisms of Stevens' original set of scales include that it is incomplete, and that it is inconsistent and therefore potentially confusing. For example, Spearman's (1904) rank correlation coefficient (ρ) is frequently suggested for ordinal measures (though it is essentially Pearson's coefficient with interval or ratio measurements replaced by their rank, an ordinal measure). Thus, although considering the relationship between the scales of measurement may prove useful, the idea of prescribing and proscribing certain analytical techniques on the basis of measurement scale contains inconsistencies in practice. The paradoxical use of Spearman's coefficient could be circumnavigated by using

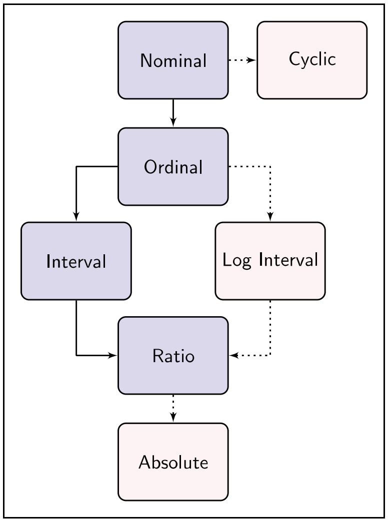

Figure 3. Scales of measurement: An augmentation of Stevens' classification.

the Kendall coefficient of concordance τ instead (Kendall, 1938). This is based on pairs of observations (x_i, y_i) and (x_j, y_j) being either *concordant* such that either $x_i < x_j$ and $y_i < y_j$ or $x_i > x_j$ and $y_i > y_j$ or *discordant* if either $x_i > y_i$ and $y_i < y_j$ or $x_i < x_j$ and $y_i > y_j$. With these definitions, and assuming no ties in the data sets,

$$\tau = \frac{2(\#\text{concordant pairs} - \#\text{discordant pairs})}{n(n-1)}$$

Note this only requires comparison operators.

Despite the contradictory nature of ρ as a descriptor of association for pairs of ordinal measurements, its use still persists, possibly dominating the use of τ. A similar issue exists for the Mann-Whitney U-test, as it is a nonparametric equivalent to the t-test based on the sums of the ranks of observations in two samples. Intended for use with ordinal measurements, it is based on sums of the ranks – an operation not permissible in Stevens' theory. In these objections some confusion arises: although Spearman's correlation coefficient and Mann and Whitney's test both rely on non-permissible operations, they apparently do the jobs that they are intended to do. In the light of this observation, does the concept of levels of measurement scale actually offer any useful contribution?

Rather than questioning its relevance, some consider it to be harmful, and that adhering to its recommendations might lead to misleading or inappropriate analysis. One of the earliest critiques is from Lord (1954), who argued that the appropriateness of data analysis depends on context and the question it is designed to answer, rather than the scale of measurement attributed to the data (Velleman and Wilkinson, 1993). For example, the floor number of an apartment block is an ordinal scale measurement, but if all floors had the same room height, it effectively functions as an interval scale measurement. Regressing this quantity against sound level for some source of noise at ground level (say passing trains) would have meaning given this context.

Guttman (1977: 105) argues similarly: 'Permission is not required in data analysis. What is required is a loss function to be minimized.' Here, 'loss function' encapsulates a penalty for making a wrong decision or a wrong summary statistic in the context of the question being asked. This should be defined differently in different situations. Further, Guttman argues, 'If a mathematician gives or withholds "permission" without reference to a loss function, he [sic] may be accessory to helping the practitioner escape the reality of defining the research problem' (p. 105). Although the idea of a 'loss function' is perhaps in itself reductionist, the last statement highlights a distinction between the mathematician and the practitioner, and perhaps a distinction between the (pure) mathematician and the data analyst. Attempting to base recommendations for appropriate data analysis techniques solely on a set of abstract axioms risks overlooking the issues that the researcher initially wished to address.

Whereas one may disagree that human geography is a science, an important point is made on the role of experience. An axiomatic approach to scales of measurement has more in common with pure mathematics than a science. Velleman and Wilkinson (1993: 7) reinforce this argument: 'Experience has shown that in a wide range of situations that the application of proscribed statistics to data can yield results that are scientifically meaningful, useful in making decisions, and valuable as a basis for further research.' Following these proscriptions could hinder the yielding of such results, suggesting that axiomatic measurement theory should be considered harmful.

Objections to measurement theory certainly cast doubt on the utility of an axiomatic approach. However, knowing something about the characteristics of data types does inform the process. For example, in many contexts the numbers associated with nominal categories are simply proxies for some other attribute, such as political party or gender. Carrying out arithmetical operations on these makes no sense when context is also considered. Similarly, although the average rank of scores in one group could be compared to that for another, there are limits to which this approach is useful. On the other hand, in the apartment floor example, a difference in mean ordinal measures may have further meaning, as each level suggests a constant change in height. Fractions of these also imply differences in height. Thus, considering types of data does inform understanding, although rather than providing a guiding theory to data analysis *of itself*, it provides contextual information.

With this in mind, it is also helpful to consider other suggested classifications of data. These are generally in a less rigorous setting than those of Stevens, but they do shed light on different approaches to analysis. One such list of classifications was suggested by Tukey (1977):

- Names
- *Grades* (ordered labels such as Freshman, Sophomore, Junior, Senior)
- *Ranks* (starting from 1, which may represent either the largest or the smallest)
- *Counted Fractions* (bounded by zero and one; these include percentages, for example)
- *Counts* (non-negative integers)
- *Amounts* (non-negative real numbers)
- *Balances* (unbounded, positive or negative values)

The list is interesting if not axiomatic. It offers greater detail than Stevens' groupings, and brings in the concept of constrained values – *Amounts* may not be negative, *Counted Fractions* must lie between zero and one, and *Counts* must be integers greater than or equal to zero. Such constraints help consideration of the nature of variables, and provide concepts beyond those considered by Stevens' classification. They are also helpful in suggesting appropriate statistical models for data. For example, identifying a quantity as a counted fraction suggests that a binomial distribution may be an appropriate model. Similarly, Poisson models may be appropriate for count data.

VI Further issues

Tukey's suggestions are well established, yet have made less of an impact on textbooks of quantitative geography than Stevens' initial ideas despite the fact that few quantitative human geographers currently adhere to an axiomatic approach. Yet despite the fact that Tukey identifies more richness in variety of data types than Stevens' approach, more recent considerations suggest that even more kinds of data should be considered, many of which are relevant to quantitative human geography.

In many situations, locations in geographical space (modelled as a map projections or a point on the surface of a globe) are in themselves a data type. Like directional data, in general they have no implicit ordering, but mathematical operations such as addition, subtraction and the

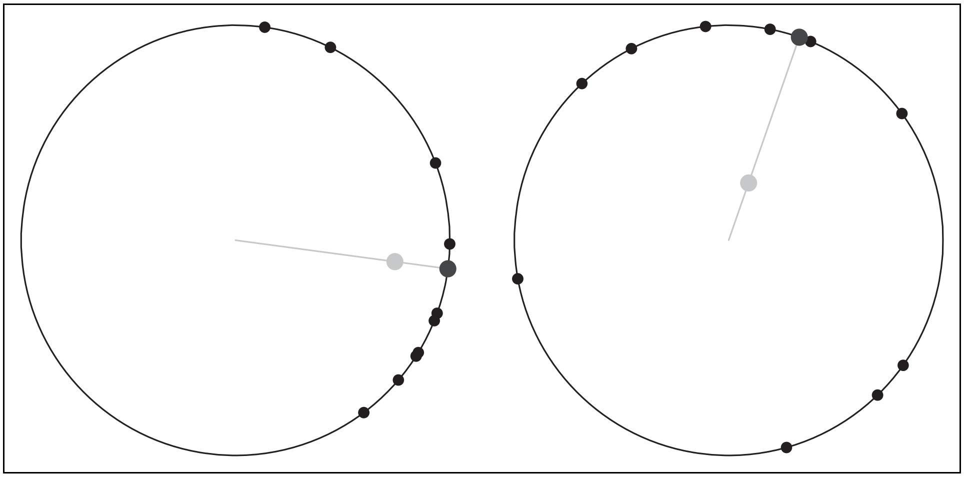

Figure 4. Constrained 2D interpretation of a circular mean.

computation of means (in this case centroids) are meaningful. Referring again to the network diagram in Figure 1, this is another offshoot from ‘nominal’. Further understanding of directional data can be gained by regarding directions as two-dimensional points constrained to lie on a circle of unit radius. The circular mean is the projection of the centroid of a set of directions represented in this way onto the circle of unit radius, as demonstrated in Figure 4.

All scales of measurement are subject to measurement error. However, further uncertainty occurs in the assignment of categories to nominal measures if the categories are unclearly defined. It is difficult, if not impossible, to define the concept of a materially deprived household in a crisp way. In such cases, the idea of *fuzzy sets* is often applied. However, how fuzzy nominal entities fit into this framework, and what statistical operations are meaningful for such data, is not yet clear. Work such as that by Fisher (1999) provides a strong starting-point for this.

Returning to issues related to multi-dimensional data, most basic classifications of scale focus on the unidirectional, where relationships between entities are generally based on ordering. However, there are other kinds of relationship. Geographical regions could be regarded as measurements on a nominal scale. Intuitively there is no order relationship between them but a notion of adjacency, where a pair of regions share a border, could be represented. If this scale is tentatively named ‘nominal positional’, there is more structure in the scale of measure than in ‘pure’ nominal data. Statistics such as Moran’s-*I* (Moran, 1950) identify the degree of linkage between a nominal positional scale and an interval or ratio scale measurement.

VII Closing discussion

Although much of the material here has been considered in previous discussions, some of these ideas take on a new significance in the era of data science, and in particular geographical data science. There may be good arguments that an axiomatic view on measurement types is not the most advantageous way to proceed, yet an understanding of different kinds of data, or (in Stevens’ framework) different measurement scales, and an understanding that such typologies are themselves contested ideas, plays an important role in data science. Even if one does not accept the notion that certain kinds of analysis or statistical summaries may be proscribed purely on the basis of measurement scale, it is informative to reflect on the combination of research question, kind of data, and type of visualization or analysis that is taking place. These ideas are an aid to thought in this area, even if they do not provide a set of concrete rules.

Declaration of conflicting interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/ or publication of this article: I gratefully acknowledge funding from Science Foundation Ireland under the Investigators' Award Programme award number 15/IA/3090.

References

Bimonte S, Villanova-Oliver M and Gensel J (2012) A multidimensional model for correct aggregation of geographic measures. In: Information Resources Management Association (ed.) *Geographic Information Systems: Concepts, Methodologies, Tools, and Applications: Concepts, Methodologies, Tools, and Applications*. Hershey, PA: IGI Global, 377–398.

Brunsdon C and Charlton M (2006) Local trend statistics for directional data: A moving window approach. *Computers, Environment and Urban Systems* 30(2): 130–142. DOI: http://dx.doi.org/10.1016/j.compenvurbsys.2005.08.004.

Brunsdon C and Corcoran J (2006) Using circular statistics to analyse time patterns in crime incidence. *Computers, Environment and Urban Systems* 30(3): 300–319. DOI: http://dx.doi.org/10.1016/j.compenvurbsys.2005.11.001.

Chrisman NR (1995) Beyond Stevens: A revised approach to measurement for geographic information. In: *Twelfth International Symposium on Computer-Aided Cartography, Auto-Carto 12*. Falls Church, VA: ACSM/ASPRS, 271–280.

Erguven M (2014) Influences of measurement theory on statistical analysis and Stevens' scales of measurement. *Journal of Technical Science and Technologies* 2(1): 27–31.

Faggian A, Corcoran J and McCann P (2013) Modelling geographical graduate job search using circular statistics. *Papers in Regional Science* 92(2): 329–343. DOI: 10.1111/pirs.12026.

Fisher NI (1993) *Statistical Analysis of Circular Data*. Cambridge: Cambridge University Press.

Fisher PF (1999) Models of uncertainty in spatial data. In: Longley P, Goodchild M, Maguire D and Rhind D (eds) *Geographical Information Systems: Principles, Techniques, Management and Applications*, *Vol*. 1. New York: John Wiley & Sons, 191–205.

Francis J, Tontisirin N, Anantsuksomsri S, Vink J and Zhong V (2015) Alternative strategies for mapping ACS estimates and error of estimation. In: *Emerging Techniques in Applied Demography*. Dordrecht: Springer, 247–273.

Guttman L (1977) What is not what in statistics. *The Statistician* 26: 81–107.

Harvey F (2017) Improving GIScience visualization: Ideas for a new methodology. In: *Spatial Data Handling in the Big Data Era*. Dordrecht: Springer, 127–137.

Härtwig M, Müller M and Bernard L (2014) A generic web service for ad-hoc statistical spatio-temporal aggregation. *Transactions in GIS* 18(4): 527–543.

Jensen G (2014) *Beyond dichotomy: Dynamics of choice in compositional space*. PhD thesis, Columbia University, USA.

Kendall MG (1938) A new measure of rank correlation. *Biometrika* 30(1/2). JSTOR: 81. DOI: 10.2307/2332226.

Krantz DH, Luce RD, Suppes P and Tversky A (1971) *Foundations of Measurement, Vol. I: Additive and Polynomial Representations*. New York: Academic Press.

Lin J, Hanink DM and Cromley RG (2017) A cartographic modeling approach to isopleth mapping. *International Journal of Geographical Information Science* 31(5): 849–866.

Lord FM (1954) On the statistical treatment of football numbers. *American Psychologist* 9: 264–265.

Luce RD, Krantz DH, Suppes P and Tversky A (1990) *Foundations of Measurement, Vol. III: Representation, Axiomatization, and Invariance*. New York: Academic Press.

Métral C, Ghoula N and Falquet G (2012) An ontology of 3D visualization techniques for enriched 3D city models. In: *Usage, Usability, and Utility of 3D City Models–European Cost Action Tu0801* (02005). Les Ulis, France: EDP Sciences.

Métral C, Ghoula N, Silva V and Falquet G (2014) A repository of information visualization to support the design of 3D virtual city models. In: *Innovations in 3D Geo-Information Sciences*. Dordrecht: Springer, 175–194.

Moran PAP (1950) Notes on continuous stochastic phenomena. *Biometrika* 37(1): 17–23.

Nelder JA (1990) The knowledge needed to computerise the analysis and interpretation of statistical information. In: Fessey MC (ed.) *Expert Systems and Artificial Intelligence: The Need for Information about Data*. London: Library Association, 23–27.

Spearman C (1904) The proof and measurement of association between two things. *The American Journal of Psychology* 15: 72–101.

Stevens SS (1946) On the theory of scales of measurement. *Science* 103: 677–680.

Stevens SS (1959) Measurement, psychophysics and utility. In: Churchman CW and Ratoosh P (eds) *Measurement: Definitions and Theories*. New York: Wiley, 18–63.

Suppes P, Krantz DH, Luce RD and Tversky A (1989) *Foundations of Measurement, Vol. II: Geometrical, Threshold, and Probabilistic Representations*. New York: Academic Press.

Tukey JW (1977) *Exploratory Data Analysis*. Boston, MA: Addison-Wesley.

Van den Berg G (1991) Ideas of statistical experts on the choice of an analysis method. *Journal of Applied Statistics* 18(1): 77–86.

Velleman PF and Wilkinson L (1993) Nominal, ordinal, interval, and ratio typologies are misleading. *The American Statistician* 47(1): 65–72. DOI:10.1080/00031305.1993.10475938.

Voigt M and Polowinski J (2011) Towards a unifying visualization ontology (Technical Report TUD-F11 1-01). Available at: https://books.google.ie/books?id=8gC6XwAACAAJ.

Wirz M, Franke T, Roggen D, Mitleton-Kelly E, Lukowicz P and Tröster G (2012) Inferring crowd conditions from pedestrians: Location traces for real-time crowd monitoring during city-scale mass gatherings. In: *2012 IEEE 21st International WETICE*. Hammamet, Tunisia: IEEE Press.