Ruth Amstutz:
“Biometric Mosaics”

in: Eric Hounshell, Ruth Amstutz (eds.):

cache 03

SHADOW OF THE TREE

intercomverlag

BIOMETRIC MOSAICS Speculative Encounter under a Tree

"So I was in the park just now. The roots of the [plane] tree were sunk in the ground just under my bench. I couldn't remember it was a root any more. The words had vanished and with them the significance of things, their methods of use, and the feeble points of reference which men have traced on their surface. I was sitting, stooping forward, head bowed, alone in front of this black, knotty mass, entirely beastly, which frightened me. Then I had this vision. [...]
[M]y head was empty, or there was just one word in my head, the word 'to be.' Or else I was thinking . . . how can I explain it? I was thinking of belonging, I was telling myself that the sea belonged to the class of green objects, or that the green was a part of the quality of the sea. [...]
And then all of a sudden, there it was, clear as day: existence had suddenly unveiled itself. It had lost the harmless look of an abstract category: it was the very paste of things, this root was kneaded into existence. Or rather the root, the park gates, the bench, the sparse grass, all that had vanished: the diversity of things, their individuality, were only an appearance, a veneer. This veneer had melted, leaving soft, monstrous masses, all in disorder – naked, in a frightful, obscene nakedness."

Jean-Paul Sartre: *Nausea,* New York: New Directions (2013 [1938]), pp. 304–306.

De trop! I shouted from my black turtleneck sweater, noticing at that moment a man in bronze sitting on a steamer trunk in the bushes on the other side of the gravel path, looking at me sympathetically.



Monument to Jan Czekanowski in a park in Szczecin, Poland.
Wikipedia, https://de.wikipedia.org/wiki/Jan_Czekanowski
(16 June 2023).

His suitcase, the man said, introducing himself as Prof. Dr. Jan Czekanowski, anthropologist and statistician, contained 109 skulls as well as notes on a statistical method[1] – the *Differentialdiagnose* – which, he believed, could be expected to transform my experience of exuberant abundance into a world shaped not by questions of the primacy of existence vs. essence but by patterns of relations of objects to one another.

BIOMETRIC MOSAICS Quadrantennetz

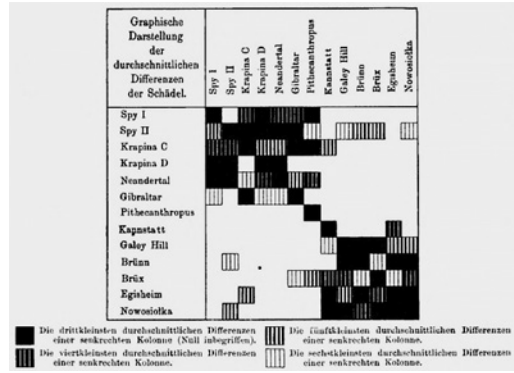
By calculating the extent of dissimilarity of two individuals by means of the average difference of 27 features – from the *Innere Orbitale Gesichtsbreite* to the *Bregma Winkel* to the *Kalottenhöhenindex* and other biometric measurements – of pairs of thirteen fossil hominin skulls, Czekanowski explained, he had developed a criterion for determining whether an individual should be assigned to a group or not. He had visualized the results of his *Differentialdiagnose* in a *Quadrantennetz*, as he called it, consisting of tiles of various shadings representing the average difference between two skulls in relation to the differences to the respective other skulls.

Tabelle II. Durchschnittliche Differenzen.

	Spy I	Spy II	Krapina C	Krapina D	Neander-tal	Gibraltar	Pithec-anthrop.	Kannstatt	Galey Hill	Brünn	Brüx	Egisheim	Nowo-sielka
Spy I	0	5,415	5,402	4,179	5,045	6,583	8,054	16,343	10,310	9,130	9,435	11,389	8,267
Spy II	5,415	0	5,036	2,692	4,248	5,721	10,340	11,717	6,536	5,373	5,887	6,289	5,315
Krapina C	5,402	5,036	0	6,664	6,435	5,032	9,307	9,005	9,850	8,333	6,673	8,100	7,023
Krapina D	4,179	2,692	6,664	0	4,671	9,141	16,525	15,933	13,360	12,680	10,700	12,933	9,141
Neandertal	5,045	4,248	6,435	4,671	0	6,819	10,138	13,150	10,540	9,232	7,301	9,144	7,208
Gibraltar	6,583	5,721	5,032	9,141	6,819	0	11,273	11,793	9,982	9,158	7,708	9,000	9,008
Pithecanthropus	8,054	10,340	9,307	16,525	10,138	11,273	0	15,105	14,582	15,436	10,219	10,486	13,551
Kannstatt	16,343	11,717	9,005	15,933	13,150	11,793	15,105	0	10,473	12,478	8,228	4,761	6,795
Galey Hill	10,310	6,536	9,850	13,360	10,504	9,982	14,582	10,473	0	3,557	5,372	5,550	5,183
Brünn	9,130	5,373	8,333	12,680	9,232	9,158	15,436	12,478	3,557	0	6,289	3,250	3,690
Brüx	9,435	5,887	6,673	10,700	7,301	7,708	10,219	8,228	5,372	6,289	0	5,757	4,586
Egisheim	11,389	6,289	8,100	12,933	9,144	9,000	10,486	4,761	5,550	3,250	5,757	0	4,064
Nowosiotka	8,267	5,315	7,023	9,141	7,208	9,008	13,551	6,795	5,183	3,690	4,586	4,064	0

Table II. Average differences. Jan Czekanowski: "Zur Differentialdiagnose der Neandertalgruppe", in: *Korrespondenzblatt der Deutschen Gesellschaft für Anthropologie, Ethnologie und Urgeschichte* 40 (1909), pp. 44–47, p. 46.

Compilation of the distance values of 13 skulls. The intersection between two skulls (one in the vertical, the other in the horizontal) yields their average difference.



Graphical representation of average differences of skulls. Jan Czekanowski: "Zur Differentialdiagnose der Neandertalgruppe", in: *Korrespondenzblatt der Deutschen Gesellschaft für Anthropologie, Ethnologie und Urgeschichte* 40 (1909), pp. 44–47, p. 46.

To transform the distance values into a *Quadrantennetz,* the values of the average differences are assigned to each square in a matrix with a side length of as many squares as there are objects to be compared. The squares are then filled with different markings indicating different gradations of distance values. These gradations always refer to the extent of similarity between objects within a vertical column, which in turn gives information about the similarity of a certain object (a) (e.g. the Neanderthal skull) with the other skulls. The similarity of object (a) with another object (b) (e.g. the Brüxer skull) is thus determined in relation to the similarity of object (a) to all other objects with which object (a) is compared. In Czekanowski's diagram, the skulls form two groups: The cluster in the upper left corner is generated by those skulls that Czekanowski assigns to the Neanderthal group. The cluster in the bottom right corner is generated by another, unspecified group. Furthermore, the Pithecampus skull proves to be clearly different from the others by its hatchless horizontal column. Note that the gradation of distance values is irregular: It does not start with the smallest differences, but only with the third smallest differences (which, however, also include zero as well as the smallest and second smallest differences) and then proceeds regularly and in smaller steps via the fourth and fifth smallest differences to the sixth smallest difference (each step is represented with differently dense hatchings). The decision of how many and which gradations of difference to represent cannot be derived inductively, i.e. from the data alone. Rather, it is a decision that may highlight or blur existing tendencies of similarity and difference between objects in the data but is based on existing hypotheses about the similarity of objects. Czekanowski assumed that some of the skulls belonged to the Neanderthal group and that the Pithecampus skull belonged to a different group than the other skulls – that of Homo Erectus. This may have contributed to his choosing the gradations in a way that the former appeared in the quadrant network as a distinct cluster and the latter as a distinct object.

By the way, did you know, Czekanowski asked, pointing to the tree spreading its branches above *my* skull, that the leaf of the plane tree is easily confused with that of certain maple species because of its hand-shaped, lobed leaves? Without waiting for my answer, he proceeded to explain that, well before some of his students had set out to use his statistical methods to distinguish racial types, he had expected that

"[...] this method will be called upon to play an important role for *Rassediagnosen*."

Jan Czekanowski: "Zur Differentialdiagnose der Neandertalgruppe", in: *Korrespondenzblatt der Deutschen Gesellschaft für Anthropologie, Ethnologie und Urgeschichte* 40 (1909), p. 44–47, p. 47. Translation by the author.

In particular, he remembered a study by his student Boleslaw Rosinski, who sought to determine the systematic affiliation of Polish smallholders from the western part of Pultusk County via the coefficients of similarity between *Rassekomponenten* – you know, he added: hair color according to Fischer, eye color according to Martin, etc. The *Quadrantennetz*, Czekanowski exulted, had expressed Rosinski's results with extraordinary acuity.
A loud bang, followed by a precipitation of gray-brown, scale-like slabs bursting from the trunk of the plane tree, interrupted his explanations.

Tabelle 2. Ähnlichkeitskoeffizienten der Individuen einer anthropologischen Gruppe.

	13	15	14	16	10	12	11	9	4	1	2	3	7	6	5	8
13	+1·	+·850	+·583	+·583	+·117	−·050	−·467	−·575	−·217	−·70	−·650	−·750	−·033	+·10	−·083	+·083
15	+·850	+1·	+·642	+·733	+·233	+·183	−·133	−·358	−·233	−·717	−·733	−·933	−·317	−·267	−·283	−·167
14	+·583	+·642	+1·	+·617	+·117	+·10	+·067	−·642	−·433	−·967	−·483	−·750	−·383	−·30	−·250	+·067
16	+·583	+·733	+·617	+1·	+·283	+·017	0	−·308	−·050	−·650	−·517	−·708	−·30	−·383	−·908	−·350
10	+·117	+·233	+·117	+·283	+1·	+·883	+·533	+·583	−·267	0	−·40	−·433	−·90	−·833	−·450	−·567
12	−·050	+·183	+·10	+·017	+·883	+1·	+·717	+·575	−·10	+·075	−·267	−·217	−·883	−·833	−·30	−·317
11	−·467	−·133	+·067	0	+·533	+·717	+1·	+·508	+·117	+·167	+·325	0	−·667	−·850	−·550	−·867
9	−·575	−·358	−·642	−·308	+·583	+·575	+·508	+1·	+·242	+·775	+·175	+·375	−·358	−·408	−·092	−·392
4	−·217	−·233	−·433	−·050	−·267	−·10	+·117	+·242	+1·	+·40	+·317	+·333	+·317	+·10	−·267	−·250
1	−·70	−·717	−·967	−·650	0	+·075	+·167	+·775	+·40	+1·	+·60	+·683	+·233	+·117	+·150	−·20
2	−·650	−·733	−·483	−·517	−·40	−·267	+·325	+·175	+·317	+·60	+1·	+·833	+·433	+·183	0	+·067
3	−·750	−·933	−·750	−·708	−·433	−·217	0	+·375	+·333	+·683	+·833	+1·	+·558	+·433	+·450	+·267
7	−·033	−·317	−·383	−·30	−·90	−·883	−·667	−·358	+·317	+·233	+·433	+·558	+1·	+·917	+·433	+·450
6	+·10	−·267	−·30	−·383	−·833	−·833	−·850	−·408	+·10	+·117	+·183	+·433	+·917	+1·	+·683	+·717
5	−·083	−·283	−·250	−·908	−·450	−·30	−·550	−·092	−·267	+·150	0	+·450	+·433	+·683	+1·	+·833
8	+·083	−·167	+·067	−·350	−·567	−·317	−·867	−·392	−·250	−·20	+·067	+·267	+·450	+·717	+·833	+1·

Similarity coefficients of individuals in an anthropological group. Boleslaw Rosinski: "Anthropogenetische Auslese", in: *Anthropologischer Anzeiger* 6/1 (1929), pp. 49–64, p. 51.

Table 2 is laid out such that each individual (1–16) corresponds to both a vertical column and a horizontal row. At the intersection of row and column, the similarity measure of each pairing is indicated. Rosinski arranged the individuals in the table in such a way that the most similar ones are next to each other and, as a result, complexes of strikingly large positive coefficients become visible. To make them even more visually apparent, he enclosed these clusters in frames both in the table and in the *Quadrantennetz* he constructed based on the table.

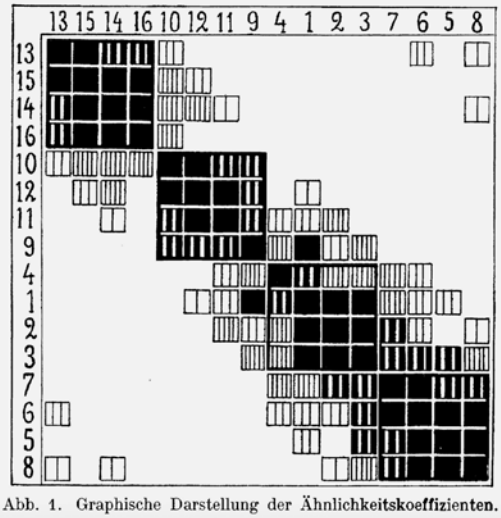


Abb. 1. Graphische Darstellung der Ähnlichkeitskoeffizienten.

Graphical representation of similarity coefficients. Boleslaw Rosinski: "Anthropogenetische Auslese", in: *Anthropologischer Anzeiger* 6/1 (1929), pp. 49–64, p. 51.

On the one hand, the case of Rosinski shows that the columns and rows of the Quadrantennetz can be arranged in any order and that it is the choice of this order that makes the clusters recognizable in the first place. On the other hand, it becomes clear how graphical interventions such as the addition of frames can produce the visual effect of relatively distinct groups. This is particularly evident in the lower right corner of the matrix, where the frames create the impression of two clusters, which would not occur with such clarity from the data alone.

A man crawled out of the bushes. The natural spectacle we had just witnessed, he said, was the bark-shedding of the plane tree. Often the splitting of the bark is accompanied by a loud bang, as we had just witnessed. Dr. Wladimir Iwanowicz introduced himself and added that he had inspected the results of the *Rassediagnosen* that Czekanowski's students carried out using the *Quadrantennetz*. His review showed, he argued, that ¾ of the individuals belonged to a different racial type than the one assigned.
Once again, there was a bang.

"In reality [...],"

Iwanowicz said,

"[Czekanowski's method] raises the similarity of the kind of deviations of the single characteristics of two individuals from a level formed by the complex of arithmetic averages of the single characteristics of the whole group."

Wladimir Iwanowicz: "Die Anwendung der Ähnlichkeitsmethode in der Rassenbestimmung", in: *Anthropologischer Anzeiger* 10/2/3 (1933), p. 225–228, p. 227. Translation by the author.

Meanwhile, the bark continued to crack without pause.

"I believe [...],"

Iwanowicz shouted into the din,

"I do not have to indicate with any greater clarity that the results attained with the aid of Czekanowski's *Ähnlichkeitsmethode* lead directly to absurdity. The essence of the method contains a logical error[!]"

Wladimir Iwanowicz: "Die Anwendung der Ähnlichkeitsmethode in der Rassenbestimmung", in: *Anthropologischer Anzeiger* 10/2/3 (1933), p. 225–228, p. 228. Translation by the author.

Iwanowicz's criticism does not concern the method of the *Differentialdiagnose* by means of the *Quadrantennetz* itself, but the *Ähnlichkeitsmethode* – developed by Czekanowski and used by Rosinski to determine the distance values. The problem Iwanowicz addresses is that the correlation coefficients determined by this method do not represent the absolute magnitude of the deviations of the individual characteristics from their arithmetic mean but only the rank order of these deviations in magnitude. This leads, for example, to individuals who are very similar in their studied characteristics being determined as dissimilar because there are still individuals in the data set that are slightly more similar and these therefore appear closer together in the ranking. The "logical error" in Rosinski's *Rassendiagnose* therefore already occurs in the processing of the data, not in their analysis by means of the *Quadrantennetz*.

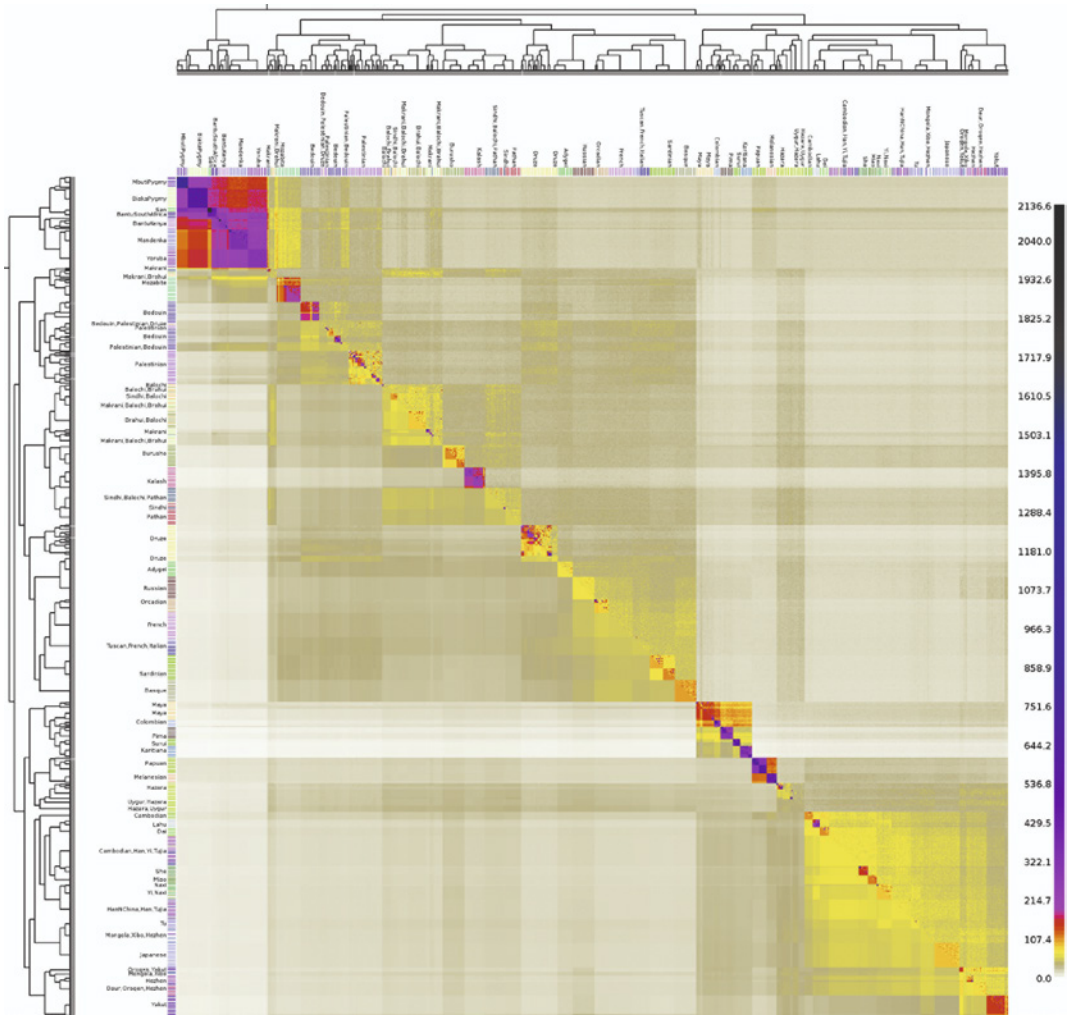
BIOMETRIC MOSAICS Clustered Heat Map

Be that as it may, replied Czekanowski, who in the meantime was standing up to his spats in the bark; in any case, the *Quadrantennetz* bears a striking resemblance to the so-called "clustered heat map," a diagram which, incidentally, is accorded the status of a "postgenomic icon."[2] He wiped a pile of bark from his epaulettes. The postgenomic era, he explained, as if he had just become aware of the historical situatedness of his counterpart, is difficult to define, but its most characteristic feature is probably the accumulation of huge amounts of genotypic and phenotypic data, which have to be organized, visualized, and interpreted.
De trop! I blurted once again.
Czekanowski fanned me with his notebook and reassured me that the clustered heat map modeled on the principle of the *Quadratennetz* made it possible to display these enormous amounts of data in a condensed form in a very small space and thus to uncover patterns in the DNA of any organism – from the plant *Arabidopsis thaliana* (a popular model organism in plant biology and genetics) to rainbow trout and suicidal crickets.[3] Iwanowicz, meanwhile, crawled back through the bark into the bushes and grumbled:

"Perhaps the *Ähnlichkeitsmethode* is useable for other aims [...]. But for the *Rassediagnose* [it is] absolutely not appropriate."

Wladimir Iwanowicz, "Die Anwendung der Ähnlichkeitsmethode in der Rassenbestimmung", in: *Anthropologischer Anzeiger* 10/2/3 (1933), pp. 225–228, p. 227. Translation by the author.

Well, Czekanowski replied, he knew of a clustered heat map called a "coancestry matrix" in which, based on the analysis of 641'000 markers of 938 individuals, their ancestral relationships could be represented and 226 worldwide populations identified.[4] With regard to its purposes, he added, this diagram showed unmistakable parallels to the *Rassediagnose* of Rosinski. Obviously, this clustered heat map did not differ significantly from the *Quadrantennetz* – apart from the coloration and, Czekanowski said with a flick of his wrist, from this simplifying tree appendix.



Whole world HGDP coancestry matrix. Daniel John Lawson, Garrett Hellenthal, Simon Myers and Daniel Falush, "Inference of Population Structure Using Dense Haplotype Data", in: *PLoS Genetics* 8/1 (2012), pp. 1–16. Supplementary Material, Figure S14.

The coancestry matrix represents the pattern of worldwide population structure based on an analysis of the Human Genome Diversity Project data set using a population genetic software called "fineSTRUCTURE." In simple terms, the matrix represents the degree of genetic relatedness based on a set of distance values between two individuals. Like the *Quadrantennetz,* the coancestry matrix compresses these multidimensional relationships onto a two-dimensional space. As the developers of the software point out, the intricate structure of the data (reflected in the complex visual structure of the matrix) indicates that many groups must be connected, not by simple hierarchical tree-like relations but also by various admixture events. However, the program also includes a tree-building algorithm that reduces these complex relations precisely to this simple tree-like structure of relatedness (above and left of the matrix).

By the way, Iwanovicz's voice interrupted him from the bushes, did you know that the plane tree sheds its bark due to rapid growth? This, he explained, without waiting for our answer, leads to the fact that the circumference of the trunk and the branches increase to such an extent that the tree's mantle becomes too tight, so to speak, and bursts at the seams, revealing the underlying mosaic-like structure of the tree's cortex. However, he added emphatically, the vitality of the tree itself does not suffer at any time.

Notes

1 Czekanowski brought the skulls back from his travels to German East Africa and Equatorial Africa between 1905 and 1909. Today, the skulls – believed to have been looted from the graves of high dignitaries of the Hayo on the island of Busira in Lake Victoria – are in the possession of the Prussian Cultural Heritage Foundation.
2 John N. Weinstein: "A Postgenomic Visual Icon", in: *Science* 319 (2008), pp. 1772–1773.
3 Julien Bobe, Jerôme Montfort, Thaovi Nguyen and Alexis Fostier: "Identification of New Participants in the Rainbow Trout (*Oncorhynchus Mykiss*) Oocyte Maturation and ovulation processes using cDNA microarrays", in: *Reproductive Biology and Endocrinology* 4/1 (2006), p. 39; Andrew C. Doxey, Mahmoud W. F. Yaish, Barbara A. Moffatt, Marilyn Griffith and Brendan J. McConkey, "Functional Divergence in the Arabidopsis β-1,3-Glucanase Gene Family Inferred by Phylogenetic Reconstruction of Expression States", in: *Molecular Biology and Evolution* 24/4 (2007), pp. 1045–55; D. G. Biron, F. Ponton, L. Marché, N. Galeotti, L. Renault, E. Demey-Thomas, J. Poncet, S. P. Brown, P. Jouin and F. Thomas, "'Suicide' of Crickets Harbouring Hairworms: A Proteomics Investigation", in: *Insect Molecular Biology* 15/6 (2006), pp. 731–42.
4 Daniel John Lawson, Garrett Hellenthal, Simon Myers and Daniel Falush, "Inference of Population Structure Using Dense Haplotype Data", in: *PLoS Genetics* 8/1 (2012), pp. 1–16.

Further Readings

Arkadiusz Soltysiak, Piotr Jaskulski: "Czekanowski's Diagram. A Method of Multidimensional Clustering", in: Juan A. Barceló, Ivan Briz, Assumpció Vila (eds.): *New Techniques for Old Times, CAA 98. Computer Applications and Quantitative Methods in Archaeology,* Proceedings of the 26th Conference, Barcelona, March 1998, BAR International Series 757 (1999), pp. 175–184.

John N. Weinstein: "A Postgenomic Visual Icon", in: *Science* 319 (2008), pp. 1772–1773.

Leland Wilkinson, Michael Friendly: "The History of the Cluster Heat Map", in: *The American Statistician* 63/2 (2009), pp. 179–184.