https://en.zeugmakongresi.org/ - https://miau.my-x.hu/miau/320/Zeugma%20Program.pdf

13th International ZEUGMA - CONGRESS ON SCIENTIFIC RESEARCH / February 24-26, 2025 / Gaziantep, Turkiye

Building statistical neurons in case of regional development projects

Dániel Váradi (https://orcid.org/0000-0001-9610-8566)
László Pitlik (https://orcid.org/0000-0001-5819-0319)
e-Mails: danielvaradi140@gmail.com, pitlik@my-x.hu
Kodolányi János University and MY-X research team Hungary

Keywords: AI, anti-discrimination, optimization, automation, similarity analysis, group building

Introduction

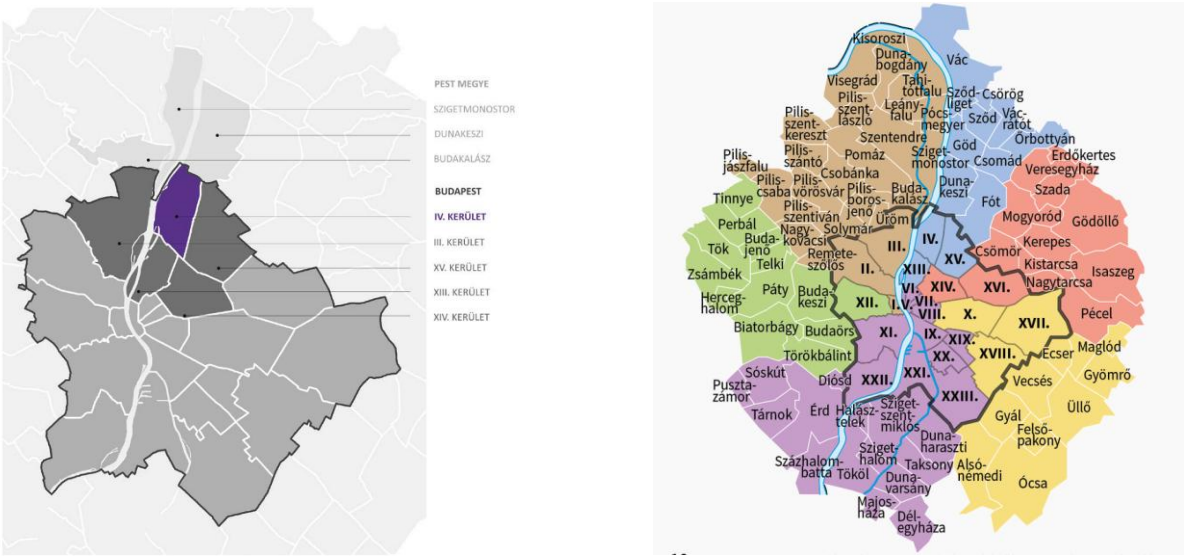
The authors of this paper have been previously working together on topics such as measuring homogeneity of European Countries based on statistical dataset using AI (cf.: https://miau.my-x.hu/miau/302/homogeneity_abs.docx), defining border lines of the Mezőföld region (Hungary) using self designed AI engine models (cf.: https://miau.my-x.hu/miau/311/mezofold/full_region_team_sustainability_analyses.pdf). As a consequence, the latest research topic is about measuring and defining the Cultural Homogeneity of Hungary‘s capital: Budapest’s districts (I-XXIII). This paper was developed, to showcase an unexpected sub-result of the overall project. The term of statistical neurons as such is a new term, but the workflow behind it is already tested for context free time-series, where the cutting points should be derived in a real-time way: c.f. https://miau.my-x.hu/miau2009/index.php3?x=e0&string=szakaszolas

Budapest is currently divided into 23 districts. Our raw dataset was downloaded from the website of the National Spatial Development and Planning Information System (cf.: https://www.oeny.hu/oeny/teir/#/) consisting of 83 attributes in case of each district and between 2011 and 2023. After carrying out necessary steps to clean the raw dataset (cf.: relativizing absolute numbers, cleaning up attributes with missing values etc.) we began with our first experiment to define which district has the most impact on the overall standard deviation, if one of the district’s data-points at a time were deleting from the dataset. To do this, we constructed an OAM (Object-Attribute-Matrix) where the objects are the official statistical indicators for Budapest always deleted a particular district (they are the same in case of all 23 districts) and the attributes are the years (2011-2023). Of course, it is necessary to have a benchmark object, where each district is given. Based on this, a second table was made showing the impacts on standard deviation (concerning parallel each attribute – c.f. impulses for a statistical neuron) in case of the single deleted district for each year. After unifying (replacing the unit of measure of numbers with rankings) the results of this table the output is a barcode-like creation. This result questions the need of optimisation using complicated neural networks in case of such experiments which are based on GEO Statistical datasets. This also means, processing times could be massively reduced – it means the real-time analysis is therefore given. Parallel: the pure statistical analysis delivered high correlation (0.96?) to the complex (optimized) anti-discriminative similarity analysis having the same OAM as input – exactly as before in case of the cutting robots for context-free time-series.

Literature, background, history

The literature will focus on the methodological aspects of the analysis. The observed region (Budapest, Hungary and t agglomeration) is to interpret only as a potential example. Methodologies should always be context-free.

On the other hand, it is necessary to define the used objects (districts = “kerület”): c.f. Figure#1 and neighbourhood-settlements (25+1): c.f. Figure#2):



Figure#1 (left, Gray-scaled): Districts of Budapest (source: https://www.ujpest.hu/galeria/File/foepiteszi_iroda/10_sz_helyzetelemzes_helyzetertekeles.pdf

Figure#2 (right, coloured): Observed settlements (source: https://nat2012.nkp.hu/tankonyv/foldrajz_10/img/3%20fej/10_3_Bp_agglomeracio_szektorok.png?max_width=800 / https://nat2012.nkp.hu/tankonyv/foldrajz_10/lecke_03_010

Big cities (big countries) have a seemingly specific kind of own devilment process (c.f. history of USA – incl. even the recent questions about Greenland and/or Canada). In the background of the different unification processes it is a simple rule: there are regional units being (temporarily) possible to cannibalize and there are regional units which can (temporarily) not be cannibalize... Big systems are like amoebas/rivers: they are tending to directions where the weaker subsystems are weaker than in other directions...

This study tries to make the abstraction weakness more concrete: strong cultural homogeneity can be seen as a kind of magnetic forcefield, and weak cultural homogeneity is a kind of anti-forcefield concerning risk analyses

Databases

Details:

- https://miau.my-x.hu/miau/320/statistical_neurons/disctricts_of_budapest.xlsx
- https://miau.my-x.hu/miau/320/statistical_neurons/agglomeration_of%20budapest.xlsx
- https://miau.my-x.hu/miau/320/statistical_neurons/Budapest%20Districts.xlsx

The raw statistics is part of the TEIR database (https://www.teir.hu/). This database does need specific access rights (e.g. given for Hungarian citizens).

The database can be split into 2 parts: data about the districts, and data about the settlements (where Budapest as such is also a settlement).

The objects (districts, settlements) can have different volumes therefore, each involved attribute of these objects should be relativized (e.g. per capita).

Involved: Attributes and their measurement units for the districts of Budapest (see "BP"):

1. BP: Proportion of females aged 0-14 in the permanent population (percentage)
2. BP: Proportion of males aged 0-14 in the permanent population (percentage)
3. BP: Natural increase or decrease (per mille)
4. BP: Domestic migration balance per 1,000 inhabitants (per mille)
5. BP: Proportion of job seekers registered for more than 180 days (percentage)
6. BP: Proportion of registered job seekers with a maximum of 8 years of primary education (percentage)
7. BP: Proportion of registered job seekers under 25 years old (percentage)
8. BP: Proportion of registered job seekers over 45 years old (percentage)
9. BP: Passenger cars per 1,000 inhabitants (units)
10. BP: Resident population per 100 dwellings (persons)
11. BP: Newly built dwellings during the year per 1,000 dwellings (units)
12. BP: Household gas consumers per 100 dwellings (persons)
13. BP: Electricity supplied to households per capita (kWh)
14. BP: Proportion of district-heated dwellings (percentage)
15. BP: Population per general practitioner and pediatrician (persons)
16. BP: Elderly people receiving daytime care per 100 places (persons)
17. BP: Children enrolled in kindergarten per available place (persons)
18. BP: Children aged 3-5 per kindergarten place (persons)
19. BP: Average number of students in primary schools (persons)
20. BP: Proportion of primary school students commuting from another settlement in daytime education (percentage)
21. BP: Number of full-time primary school students per computer
22. BP: Number of full-time high school students per computer
23. BP: Number of library units borrowed from municipal libraries per 100 inhabitants
24. BP: Registered businesses per 1,000 inhabitants
25. BP: Proportion of registered businesses in industry and construction (percentage)
26. BP: Population over 65 years old per 100 children aged 0-14
27. BP: Total number of registered job seekers
28. BP: Number of registered job seekers per 100 people aged 15-64
29. BP: Housing stock (number of units)
30. BP: Kindergarten places (number of places)
31. BP: Full-time primary school students (number of students)
32. BP: Registered businesses (as of Dec. 31) (number)
33. BP: Registered businesses with 1-9 employees (as of Dec. 31) (number)
34. BP: Registered businesses with 10-19 employees (as of Dec. 31) (number)
35. BP: Registered businesses with 20-49 employees (as of Dec. 31) (number)
36. BP: Registered businesses with 50-249 employees (as of Dec. 31) (number)

37. BP: Registered businesses with 500 or more employees (as of Dec. 31) (number)
38. BP: Registered businesses; Industry (TEÁOR08: B+C+D+E) (as of Dec. 31) (number)
39. BP: Registered businesses; Mining and quarrying (TEÁOR08: B) (as of Dec. 31) (number)
40. BP: Registered businesses; Manufacturing industry (TEÁOR08: C) (as of Dec. 31) (number)
41. BP: Registered businesses; Electricity, gas, steam supply, air conditioning (TEÁOR08: D) (Dec. 31) (number)
42. BP: Registered businesses; Water supply, sewage collection and treatment, waste management, decontamination (TEÁOR08: E) (Dec. 31) (number)
43. BP: Registered businesses; Construction (TEÁOR08: F) (Dec. 31) (number)
44. BP: Registered businesses in services (number)
45. BP: Registered businesses; Trade, motor vehicle repair (TEÁOR08: G) (Dec. 31) (number)
46. BP: Registered businesses; Transport, storage (TEÁOR08: H) (Dec. 31) (number)
47. BP: Registered businesses; Accommodation services, catering (TEÁOR08: I) (Dec. 31) (number)
48. BP: Registered businesses; Information, communication (TEÁOR08: J) (Dec. 31) (number)
49. BP: Registered businesses; Financial, insurance activities (TEÁOR08: K) (Dec. 31) (number)
50. BP: Registered businesses; Real estate activities (TEÁOR08: L) (Dec. 31) (number)
51. BP: Registered businesses; Professional, scientific, technical activities (TEÁOR08: M) (Dec. 31) (number)
52. BP: Registered businesses; Administrative and support service activities (TEÁOR08: N) (Dec. 31) (number)
53. BP: Registered businesses; Public administration, defence, compulsory social security (TEÁOR08: O) (Dec. 31) (number)
54. BP: Registered businesses; Education (TEÁOR08: P) (Dec. 31) (number)
55. BP: Registered businesses; Human health, social care (TEÁOR08: Q) (Dec. 31) (number)
56. BP: Registered businesses; Arts, entertainment, recreation (TEÁOR08: R, GFO14, Dec. 31) (number)
57. BP: Registered businesses; Other services (TEÁOR08: S) (Dec. 31) (number)

Remark: TEÁOR means official catalogue of business activities

Involved: Attributes and their measurement units for the settlements (incl. Budapest):

1. Proportion of population aged 0-14 in the permanent population (percent)
2. Proportion of population aged 65 and over in the permanent population (percent)
3. Natural population growth or decline (per thousand)
4. Internal migration balance per thousand inhabitants (per thousand)
5. Proportion of registered job seekers with at most 8 years of primary education (percent)
6. Proportion of registered job seekers under 25 years old (percent)
7. Proportion of registered job seekers over 45 years old (percent)
8. Proportion of employees in high-prestige employment groups (percent)
9. Passenger cars per 1,000 inhabitants (number)
10. Resident population per 100 dwellings (persons)
11. Number of dwellings built during the year per 1,000 dwellings (number)
12. Proportion of dwellings connected to the public sewerage network (percent)
13. Number of dwellings connected to the sewerage network per 100 dwellings connected to the piped drinking water supply (number)
14. Household gas consumers per 100 dwellings (number)

15. Electricity supplied to households per inhabitant (kWh)
16. Proportion of dwellings with district heating (percent)
17. Members of creative cultural communities per 100 inhabitants (persons)
18. Participants in regular cultural activities per 100 inhabitants (persons)
19. Number of library units borrowed from municipal libraries per 100 inhabitants (number)
20. Number of registered businesses per 1,000 inhabitants (number)
21. Municipal waste collected separately and transported per inhabitant (kg)
22. Proportion of municipal waste transported from the population (percent)
23. Total wastewater discharged into the public sewer per inhabitant (m^3)
24. Proportion of municipal waste collected separately and transported (percent)
25. Population over 65 years old per 100 people aged 0-14 (persons)
26. Total registered job seekers (persons)
27. Registered job seekers per 100 people aged 15-64 (persons)
28. Number of dwellings (number)
29. Number of kindergarten places (persons)
30. Number of general school students in daytime education (persons)
31. Number of general school students commuting from another settlement in daytime education (persons)
32. Number of students in higher education in full-time training (based on training location) (persons)
33. Number of creative cultural communities (number)
34. Number of regular cultural activities (number)
35. Number of museum visitors (persons)
36. Number of registered economic organizations (as of Dec. 31) (number)
37. Registered businesses with 1-9 employees (as of Dec. 31) (number)
38. Registered businesses with 10-19 employees (as of Dec. 31) (number)
39. Registered businesses with 20-49 employees (as of Dec. 31) (number)
40. Registered businesses with 50-249 employees (as of Dec. 31) (number)
41. Registered businesses with 250-499 employees (as of Dec. 31) (number)
42. Registered businesses with 500 or more employees (as of Dec. 31) (number)
43. Registered businesses; Agriculture, forestry, fishing (TEÁOR08: A) (as of Dec. 31) (number)
44. Registered primary producers (as of Dec. 31) (number)
45. Registered businesses; Industry (TEÁOR08: B+C+D+E) (as of Dec. 31) (number)
46. Registered businesses; Mining and quarrying (TEÁOR08: B) (as of Dec. 31) (number)
47. Registered businesses; Manufacturing (TEÁOR08: C) (as of Dec. 31) (number)
48. Registered businesses; Electricity, gas, steam supply, air conditioning (TEÁOR08: D) (as of Dec. 31) (number)
49. Registered businesses; Water supply, sewage collection, waste management, decontamination (TEÁOR08: E) (as of Dec. 31) (number)
50. Registered businesses; Construction (TEÁOR08: F) (as of Dec. 31) (number)
51. Registered businesses in services (number)
52. Registered businesses; Trade, motor vehicle repair (TEÁOR08: G) (as of Dec. 31) (number)
53. Registered businesses; Transport, storage (TEÁOR08: H) (as of Dec. 31) (number)
54. Registered businesses; Accommodation, catering (TEÁOR08: I) (as of Dec. 31) (number)
55. Registered businesses; Information, communication (TEÁOR08: J) (as of Dec. 31) (number)
56. Registered businesses; Financial, insurance activities (TEÁOR08: K) (as of Dec. 31) (number)

57. Registered businesses; Real estate activities (TEÁOR08: L) (as of Dec. 31) (number)
58. Registered businesses; Professional, scientific, technical activities (TEÁOR08: M) (as of Dec. 31) (number)
59. Registered businesses; Administrative and support service activities (TEÁOR08: N) (as of Dec. 31) (number)
60. Registered businesses; Public administration, defence, compulsory social security (TEÁOR08: O) (GFO14, as of Dec. 31) (number)
61. Registered businesses; Education (TEÁOR08: P) (as of Dec. 31) (number)
62. Registered businesses; Human health, social care (TEÁOR08: Q) (as of Dec. 31) (number)
63. Registered businesses; Arts, entertainment, recreation (TEÁOR08: R) (as of Dec. 31) (number)
64. Registered businesses; Other services (TEÁOR08: S) (as of Dec. 31) (number)
65. Registered businesses; Other activities (TEÁOR08: T+U) (as of Dec. 31) (number)
66. Length of municipal bicycle paths and shared pedestrian-bicycle paths (km)
67. Length of unpaved municipal roads and public spaces (km)
68. Length of paved municipal roads and public spaces (km)
69. Proportion of developed municipal roads (percent)
70. Length of state roads (km)
71. Contaminated sites; before environmental assessment (number)
72. Contaminated sites; after environmental assessment (number)
73. Contaminated sites; after technical intervention (number)
74. Total municipal waste transported from the population (tons)
75. Planned daily capacity of public sewage treatment plants (kg O_2)

Methodology

There are two different layers of the methodology:

1. Building of the statistical neurons
2. Antidiscriminative AI-based analyses

Statistical neurons: As it can immediately be seen in the Figures #3 and #4, statistical neurons are a part of antidiscriminative analyses, where the barcode-like patterns (see coloured cells building row-oriented similarity patterns) can be interpreted as mathematical guarantee for avoiding rel. long optimization processes without any relevant impacts for the output. Statistical neurons make real-time-analyses possible. There is no exact threshold (e.g. correlation value), which is a strict border for needing optimization or not. The real-time-needs and the risk management strategies should be evaluated in a parallel way: for which calculation speed which risks (weaking exactness) in the output can be accepted case-specific?!

Antidiscriminative analyses are part of the similarity analyses, where the parameter in the attribute-specific staircase functions can not have the same substitution values in the neighboured positions. This regulation makes namely possible, that objects can finally have the same evaluation (norm) values (see: 1.000.000 in Figure#3&#4). The antidiscriminative analysis means we are searching for forcefield excluding to use abstract terms (like different, poor, rich, etc.) for the given objects. Human terms are products of the human intuition, but these terms (phenomena) are parallel products e.g. of electric/chemical impulses in the human brain. The AI-based term-creation processes are the mechanistic copies of the biochemical processes.

How to build objects: In order to build a statistical neuron first and foremost an Object-Attribute-Matrix need to be built, where the attributes are the years and objects are the districts/settlements and years combined. Using this structure of OAM will allow to implement a standard-deviation-based examination methodology (c.f. Figure #3).

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O
1	District	Examined area	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020	2021	2022	2023
2	1	BP: Proportion of females aged 0-14 in the permanent population (percentage)	5.58	5.8	6.03	6.35	6.53	6.75	6.73	6.79	6.84	6.73	6.54	6.33	6.14
3	1	BP: Proportion of males aged 0-14 in the permanent population (percentage)	5.71	5.78	5.99	6.15	6.34	6.46	6.55	6.61	6.62	6.67	6.47	6.51	6.37
4	1	BP: Natural increase or decrease (per mille)	-6.51	-5.84	-4.02	-2.14	-6.39	-3.45	-5.01	-4.22	-4.42	-6.59	-7.82	-5.71	-5.63
5	1	BP: Domestic migration balance per 1,000 inhabitants (per mille)	2.36	4.34	5.84	9.35	10.26	1.51	1.87	1.08	1.19	-10.81	-4.22	-3.97	-1.16
6	1	BP: Proportion of job seekers registered for more than 180 days (percentage)	51.08	46.13	53.57	54.96	54.85	52.25	56.13	47.29	38.67	45.24	57.49	52.86	47.51
7	1	BP: Proportion of registered job seekers with a maximum of 8 years of primary education (percentage)	17.35	15.19	12.14	9.92	14.35	10.67	8.39	6.98	8	14.88	8.7	7.14	7.69
8	1	BP: Proportion of registered job seekers under 25 years old (percentage)	6.75	6.63	7.14	4.55	2.53	5.06	4.52	3.1	5.33	5.95	5.8	7.62	4.52
9	1	BP: Proportion of registered job seekers over 45 years old (percentage)	44.82	47.51	51.43	57.02	59.49	57.3	61.29	55.04	48.67	43.15	55.07	47.14	52.94
10	1	BP: Passenger cars per 1,000 inhabitants (units)	469.46	450.72	440.95	436.02	452.33	455.62	473.83	477.26	448.01	452.17	464.41	473.11	464.41
11	1	BP: Resident population per 100 dwellings (persons)	147.44	144.86	145.26	146.66	148.06	147.97	147.24	147.59	147.7	143.74	141.22	137.88	138.32
12	1	BP: Newly built dwellings during the year per 1,000 dwellings (units)	0.84	0.59	1.65	0	0.24	1.35	0	0.12	0	5.84	1.52	0	1.31
13	1	BP: Household gas consumers per 100 dwellings (persons)	86.34	84.12	85.89	85.42	84.16	82.77	81.51	80.66	78.56	76.61	76.06	73.18	72.75
14	1	BP: Electricity supplied to households per capita (kWh)	1739.36	1857.85	1855.35	1833.62	1870.58	1904.69	1965.1	2020.6	2025.98	2169.92	2299.31	1983.36	1869.31
15	1	BP: Proportion of district-heated dwellings (percentage)	10	9.81	9.81	9.81	9.81	9.79	9.79	9.79	9.77	9.72	9.74	9.56	9.54
16	1	BP: Population per general practitioner and pediatrician (persons)	1066.43	1068.91	987.68	997.2	1006.88	1049.58	1044.42	966.5	1093.35	1025.71	1009.25	1003.92	1008.46
17	1	BP: Elderly people receiving daytime care per 100 places (persons)	104.55	103.64	105.45	104.55	100.91	103.64	104.55	102.73	101.82	100	105.45	101.82	105.45
18	1	BP: Children enrolled in kindergarten per available place (persons)	0.99	0.97	0.97	1	1	0.95	0.97	0.91	0.89	0.78	0.72	0.78	0.77
19	1	BP: Children aged 3-5 per kindergarten place (persons)	1.04	1.05	1.1	1.13	1.14	1.09	1.09	1.05	1.02	0.94	0.93	0.91	0.91
20	1	BP: Average number of students in primary schools (persons)	242	249.2	269.4	274.8	284.8	287.2	281.6	279	268.6	266.4	261	248	234.4
21	1	BP: Proportion of primary school students commuting from another settlement in daytime education (percentage)	9.5	6.5	8.02	7.21	7.79	7.03	6.96	5.88	8.19	6.01	7.36	7.42	5.72
22	1	BP: Number of full-time primary school students per computer	13.01	13.12	14.64	13.47	13.31	11.49	12.24	9.06	8.5	8.59	8.26	4.81	4.32
23	1	BP: Number of full-time high school students per computer	10.54	10.02	9.48	10.91	11.36	12.31	12.3	10.48	9.59	8.94	7.9	5.3	4.84
24	1	BP: Number of library units borrowed from municipal libraries per 100 inhabitants	229.99	224.48	222.06	210	191.68	186.28	183.68	187.48	190.15	108.03	131.53	180.04	199.58
25	1	BP: Registered businesses per 1,000 inhabitants	352.82	356.68	358.78	356.76	356.79	361.61	373.25	384.54	391.94	388.47	395.38	391.01	377.52
	.	BP: Proportion of registered businesses in industry and construction													

Figure#3: The structure of the OAM in case of a statistical neuron (source: own editing, OAM_1 worksheet, A1-O1312, https://miau.my-x.hu/miau/320/statistical_neurons/disctricts_of_budapest.xlsx)

Role of the standard deviations (see "STDEV"): The use of standard deviations is key for building a statistical neuron for homogeneity analyses. Therefore, as the next step on building the neuron a STDEV impact table is needed, where we can measure the impact of excluding each district/settlement one by one from the whole database per year. To do this, we created a STDEV difference matrix wherefore we used the built in STDEV() function given by Excel, by measuring STDEV values in case of excluding one district or city (c.f. Figure#4).

	Q	R	S	T	U	V	W	X	Y	Z	AA	AB	AC	AD
1	Std.dev differences	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020	2021	2022	2023
2	Alltogether	252.95	255.23	252.36	251.36	255.33	258.79	264.47	266.96	268.78	275.99	283.26	275.15	281.42
3	"23-1	251.93	253.69	250.94	249.99	253.91	257.22	262.84	265.34	266.90	273.84	280.73	274.08	281.25
4	"23-2	252.41	255.05	252.00	251.12	255.12	258.52	264.25	266.77	268.55	275.59	282.91	275.11	280.38
5	"23-3	253.23	255.62	252.83	251.83	255.87	259.36	265.08	267.68	269.43	276.90	283.96	276.09	282.23
6	"23-4	253.15	255.78	252.88	251.91	255.97	259.41	265.45	268.00	269.80	276.85	284.33	276.18	283.12
7	"23-5	250.62	251.61	249.57	248.32	252.38	255.35	260.64	262.80	264.31	270.45	277.25	271.16	280.98
8	"23-6	253.91	255.98	252.88	252.19	255.99	259.37	265.07	267.50	269.37	276.33	283.49	275.39	279.00
9	"23-7	253.47	255.70	253.00	251.81	255.64	258.97	264.68	267.26	268.78	275.61	282.96	275.08	281.94
10	"23-8	250.72	252.47	250.02	249.49	253.87	257.83	263.33	266.14	268.39	276.73	284.25	275.37	282.08
11	"23-9	252.39	254.88	252.06	251.10	255.08	258.41	264.05	266.30	268.09	274.86	282.33	274.37	281.83
12	"23-10	254.13	256.44	253.57	252.59	256.64	260.09	265.70	268.39	270.21	277.59	285.10	276.77	283.42
13	"23-11	253.38	255.83	252.82	251.57	255.62	259.21	265.09	267.43	269.48	276.82	283.95	275.84	282.44
14	"23-12	253.11	255.53	252.52	251.48	255.16	258.54	264.13	266.56	268.49	275.64	283.00	274.50	280.24
15	"23-13	252.99	255.42	252.78	251.93	255.82	259.13	264.78	267.42	268.80	276.24	283.66	275.64	282.40
16	"23-14	253.59	256.05	253.09	252.06	256.10	259.57	265.34	267.82	269.50	276.57	283.88	275.75	282.35
17	"23-15	253.43	255.95	252.96	252.01	255.92	259.70	265.30	267.67	269.58	276.65	283.85	275.71	282.44
18	"23-16	253.28	255.46	252.34	251.41	255.43	258.95	264.75	267.23	269.06	276.20	283.52	275.17	280.37
19	"23-17	252.99	255.00	252.04	251.05	254.94	258.52	264.22	266.62	268.32	276.15	283.28	274.06	279.94
20	"23-18	252.99	255.74	252.65	251.65	255.78	258.90	264.49	267.19	269.40	276.68	283.78	275.79	281.56
21	"23-19	253.89	256.30	253.57	252.58	256.64	260.20	266.00	268.38	269.86	277.42	284.21	276.06	281.75
22	"23-20	253.27	255.88	252.36	251.67	255.58	258.79	264.26	266.52	268.24	275.51	283.01	274.90	281.52
23	"23-21	253.18	255.85	253.04	251.82	255.85	259.42	265.26	267.77	269.49	276.37	283.93	275.82	281.05
24	"23-22	253.30	255.19	252.63	251.86	255.69	259.26	264.84	267.41	269.43	276.81	284.03	275.38	280.31
25	"23-23	252.48	254.93	251.84	249.90	253.66	257.60	263.42	265.97	268.60	276.01	283.57	274.40	280.05

Figure#4: The structure of the OAM in case of a statistical neuron (source: own editing, OAM_1 worksheet, A1-O1312, https://miau.my-x.hu/miau/320/statistical_neurons/disctricts_of_budapest.xlsx)

In terms of the structure of the STDEV difference table, the objects here are the districts/settlements which has been excluded the calculation (e.g. All- 1) and the attributes are the years.

To be able to remove the units behind the STDEV values we used the built in RANK() function of the excel. This is an important step from two perspective, one because in this way we are able to evaluate the input dataset regardless the unit of measurement of the original downloaded dataset and secondly, because this is the input data structure which the COCO Y0 engine requires (see Figure#5 and Figure#6).

For ranking the individual values, we used the "1" ranking logic. This means, that the bugger the STDEV value is, the worst it is in the ranking (c.f. Figure#5). In order to run the antidiscriminative COCO Y0 model, an additional attribute has been added to the ranking dataset, named Y0 with a value of 1.000.000 in case of each object. This is necessary to build the antidiscriminative vector for the AI system (c.f. column AE on Figure#5).

Results

The result should be simply interpretable as follows:

- The higher the evaluation value about the norm (1.000.000) value, the more extreme force fields have the excluded district or settlement.
- It means the higher evaluation values are signs for potential sovereignty of the excluded district or settlement.
- Parallel, the districts/settlement with lower evaluation values (below the norm value) are organic parts of the conglomerates.
- The intrigues to exclude/cut out these organic parts from the entire set of the districts/settlements bring risks for the total system.
- Districts/settlements with high-levelled evaluation values are rather harmless forcefields.

																	0.96	rank-correlation				
Obejcts	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020	2021	2022	2023	Y0	Estimations	<--Ranking		Objects	Conclusions	Ranking-->	Avergae
All	7	9	10	8	9	10	10	10	11	8	8	10	11	1000000	1000031.7	11	Type	All		9	9
"23-1	3	3	3	4	4	2	2	2	2	2	2	3	10	1000000	1000110.7	2	inner	"23-1		2	3
"23-2	5	7	5	7	7	6	8	9	8	5	4	9	7	1000000	1000089.7	5	outer	"23-2		6	7
"23-3	14	13	16	16	18	17	17	19	16	22	19	22	18	1000000	999925.8	19	outer	"23-3	<--min. impact	19	17
"23-4	12	16	18	18	20	19	22	22	22	21	23	23	23	1000000	999893.8	22	outer	"23-4	<--min. impact	22	20
"23-5	1	1	1	1	1	1	1	1	1	1	1	1	8	1000000	1000132.7	1	inner	"23-5		1	2
"23-6	23	21	17	22	21	18	16	17	14	13	10	14	1	1000000	1000039.2	9	inner	"23-6		16	16
"23-7	20	14	20	14	13	13	12	13	10	6	5	8	16	1000000	999988.8	14	inner	"23-7		12	13
"23-8	2	2	2	2	3	4	3	4	6	18	22	12	17	1000000	1000055.7	8	inner	"23-8		7	7
"23-9	4	4	7	6	6	5	5	5	3	3	3	4	15	1000000	1000082.7	6	outer	"23-9		4	5
"23-10	24	24	23	24	23	23	23	24	24	24	24	24	24	1000000	999844.8	24	inner	"23-10		24	24
"23-11	18	17	15	11	12	15	18	16	18	20	18	20	21	1000000	999933.8	17	inner	"23-11		17	17
"23-12	11	12	11	10	8	8	6	7	7	7	6	6	4	1000000	1000073.7	7	outer	"23-12		8	8
"23-13	10	10	14	19	16	14	14	15	12	12	13	15	20	1000000	999968.8	16	inner	"23-13		14	14
"23-14	21	22	22	21	22	21	21	21	20	15	16	17	19	1000000	999894.8	21	inner	"23-14		21	20
"23-15	19	20	19	20	19	22	20	18	21	16	15	16	22	1000000	999905.8	20	outer	"23-15	<--min. impact	20	19
"23-16	16	11	8	9	10	12	13	12	13	11	11	11	6	1000000	1000033.7	10	outer	"23-16		11	11
"23-17	8	6	6	5	5	7	7	8	5	10	9	2	2	1000000	1000102.7	4	outer	"23-17		5	6
"23-18	9	15	13	12	15	11	11	11	15	17	14	18	13	1000000	999978.8	15	outer	"23-18		13	13
"23-19	22	23	24	23	24	24	24	23	23	23	21	21	14	1000000	999854.3	23	inner	"23-19		23	22
"23-20	15	19	9	13	11	9	9	6	4	4	7	7	12	1000000	1000027.7	12	inner	"23-20		10	10
"23-21	13	18	21	15	17	20	19	20	19	14	17	19	9	1000000	999931.8	18	outer	"23-21		18	17
"23-22	17	8	12	17	14	16	15	14	17	19	20	13	5	1000000	999989.8	13	outer	"23-22		15	14
"23-23	6	5	4	3	2	3	4	3	9	9	12	5	3	1000000	1000108.7	3	outer	"23-23		3	5

Figure#5: Statistical neuron in case of districts (source: own editing, OAM_1 worksheet, A1-O1312, https://miau.my-x.hu/miau/320/statistical_neurons/disctricts_of_budapest.xlsx)

																	0.997		ranking-correlation		
St.deviation	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020	2021	2022	2023	Y0	Estimations	<--Ranking	conclusions	Ranking-->	Average	Objects
All	15	14	15	14	14	14	14	14	16	15	15	15	12	1000000	999987.9	14		14	14	All
All - Alsónémendi	2	2	2	2	2	2	2	2	2	1	1	1	17	1000000	1000146.4	2		2	3	All - Alsónémendi
All - Budakalász	3	11	13	12	12	12	12	12	12	13	14	14	5	1000000	1000028.9	11		11	11	All - Budakalász
All - Budakeszi	8	8	8	6	7	7	7	7	6	5	8	7	14	1000000	1000078.4	8		8	8	All - Budakeszi
All - Budaörs	7	5	7	9	9	10	10	10	9	6	5	9	15	1000000	1000064.9	9		9	9	All - Budaörs
All - Csömör	10	3	9	8	4	4	4	4	3	8	7	8	2	1000000	1000103.4	4		4	6	All - Csömör
All - Diósd	18	18	19	21	21	21	21	21	21	21	21	21	3	1000000	999929.4	20		19	19	All - Diósd
All - Dunaharaszti	5	6	4	4	5	6	6	5	7	9	9	11	9	1000000	1000088.9	6		6	7	All - Dunaharaszti
All - Dunakeszi	21	22	22	22	22	22	22	20	20	20	20	19	23	1000000	999897.9	22	min. impact	21	21	All - Dunakeszi
All - Ecser	6	7	5	5	8	8	8	8	8	7	6	5	10	1000000	1000084.4	7		7	7	All - Ecser
All - Érd	12	15	11	11	11	11	11	11	10	11	10	12	13	1000000	1000024.4	12		12	11	All - Érd
All - Fót	11	13	14	13	13	13	13	13	13	14	13	6	6	1000000	1000022.9	13		13	12	All - Fót
All - Gyál	14	20	18	17	17	17	18	17	19	17	17	16	19	1000000	999948.4	17		17	17	All - Gyál
All - Halásztelek	23	21	21	20	18	20	17	22	23	23	23	23	21	1000000	999898.9	21		21	21	All - Halásztelek
All - Kistarcsa	20	17	20	19	20	18	19	18	18	19	19	20	20	1000000	999930.4	19		19	19	All - Kistarcsa
All - Maglód	17	12	12	15	15	15	15	16	15	16	16	18	18	1000000	999975.9	16		16	15	All - Maglód
All - Nagytarcsa	22	24	25	25	25	25	25	25	25	25	25	25	25	1000000	999852.4	25	min. impact	25	25	All - Nagytarcsa
All - Pécel	24	23	23	23	23	23	23	23	22	22	22	22	22	1000000	999879.4	24	min. impact	24	23	All - Pécel
All - Remeteszőlős	26	26	26	26	26	26	26	26	26	26	26	26	26	1000000	999837.4	26	min. impact	26	26	All - Remeteszőlős
All - Solymár	16	4	3	3	3	3	3	3	4	3	3	3	7	1000000	1000117.4	3		3	4	All - Solymár
All - Szigetmonostor	1	1	1	1	1	1	1	1	1	2	2	2	1	1000000	1000156.9	1		1	1	All - Szigetmonostor
All - Szigetszentmiklós	19	19	17	18	19	19	20	19	17	18	18	17	16	1000000	999940.9	18		18	18	All - Szigetszentmiklós
All - Törökbálint	9	9	10	10	10	9	9	9	11	12	11	10	8	1000000	1000049.9	10		10	10	All - Törökbálint
All - Üröm	25	25	24	24	24	24	24	24	24	24	24	24	4	1000000	999880.9	23	min. impact	23	23	All - Üröm
All - Vecsés	4	10	6	7	6	5	5	6	5	4	4	4	11	1000000	1000097.4	5		5	6	All - Vecsés
All - Budapest	13	16	16	16	16	16	16	15	14	10	12	13	24	1000000	999976.9	15		15	15	All - Budapest

Figure#6: Statistical neuron in case of settlements (source: own calculations, LSX = https://miau.my-x.hu/miau/320/statistical_neurons/agglomeration_of%20budapest.xlsx)

Further results: exploring raw-data-errors is also possible through statistic neurons. The trivial given barcode-like pattern may not be disturbed through punctual impacts in the colour-harmony. One of these raw-data-errors was reported to the authority (KSH and/or Lechner Knowledge Centre) and the errors were verified, accepted and improved.

Discussion

The statistical neurons (it means the real-time-calculations) produce results which can not be verified immediately. Therefore, it is so relevant to have the real-time-characteristics. The AI-based optimization processes do need rel. long process durations, but they have internal validation sub-modules too: the function-symmetry in case of staircase function is a relevant tool to detect objects with specific characteristics where an evaluation based on the particular given raw data is not robust enough.

The correlation between the ranked evaluation values based on statistical neurons and based on the optimized results will always be present in a late moment. Therefore, the correlation value can not be used to classify OAMs: whether they are homogeneous enough concerning the coloured row-like (barcode-like) patterns or not.

Conclusions

OAMs as inputs for statistical neurons can be used if the standard deviation for each row is near to zero. It means that the standard deviation of the row-oriented standard deviations must also be minimal!

Figure#5 presents the most unorganic districts with deep green background colours. The whole city (c.f. row of “all”) has a central ranking position (rank=14 from 23+1 objects). Unorganic district are such parts of the city where the aggregated conditions cause higher estimation values (above the norm value of 1.000.000). Without these districts (without their extremities) the situation would be more sustainable. These districts are District I. and district V. and also District II. (the “region” of the “high society”). On the other hand, extremities can be identified in districts where the population is relatively poor: Districts XXIII, XVII. Parallel, the estimation values below the norm value of 1.000.000 highlights the most integrated “regions” of the city with hardly extremities: e.g. Districts III, IV, X, XIV, XV, XIX – they represent a kind of Budapest-lifestyle (not for the tourism but even for the real spirit of the capitol city).

Figure#6 can be interpreted in the same way: Robust extremities can be assumed in the following settlements Szigetmonostor, Alsónémedi, Solymár, Vecsés, Csömör – they are the most sovereign "regions" – with the less city-like lifestyle. Parallel, the less extremities present the following settlements: Remeteszőlős, Nagytarcsa, Pécel, Üröm, Dunakeszi... The population in these "regions" is quasi fully integrated into the capitol-lifestyle.

The two partial analyses point out that [Districts III, IV, X, XIV, XV, XIX] and [Settlements Remeteszőlős, Nagytarcsa, Pécel, Üröm, Dunakeszi] are in a few cases in a neighbourhood-relationship (see Figure#2). This can be interpreted as a kind of organic development.

On the other hand, e.g. [Districts II, XVII] and [Settlements Remeteszőlős, Nagytarcsa, Pécel, Üröm, Dunakeszi] demonstrate a kind of vacuum-process: the population of the city wants a less city-like lifestyle and it migrates to the neighbourhood-settlements.

In case of highlighted inner districts (without any neighbourhood-settlements), we could speak about ghettos/enclaves/... (see District VIII for the time period of 2011-2019 with ranking values 2). District I (castle district) is and was ever a kind of enclave.

Future

The KNUTH-principle (c.f. science/knowledge is when we are capable of transforming, trans-scripting, translating our concepts into source codes) is still not given. It is still necessary to create an algorithm for deriving an index evaluating a particular OAM for being capable of calculating outputs based on the concept of the statistical neurons or not...

References

See URLs in the text stream